Optimisation

of Pixel Modules

for the ATLAS Inner Tracker

at the High-Luminosity LHC

Dissertation by Julien-Christopher Beyer

Optimisation

of Pixel Modules

for the ATLAS Inner Tracker

at the High-Luminosity LHC

Dissertation
an der Fakultät für Physik
der
Ludwig–Maximilians–Universität München

vorgelegt von
Julien-Christopher Beyer
geboren in Braunschweig

München, den 15. Februar 2019

Dissertation

submitted to the faculty of physics of the
Ludwig–Maximilians–Universität München

by Julien-Christopher Beyer

1st Referee: Priv. Doz. Dr. Richard Nisius
2nd Referee: Prof. Dr. Otmar Biebel

Date of submission: 15 February 2019
Date of oral examination: 18 March 2019

Zusammenfassung

Der *Large Hadron Collider* und seine Vorbeschleuniger werden in drei Schritten aufgerüstet, um die *high luminosity* Datennahme zu ermöglichen. In dieser Phase sollen zehnmal so viele Daten gesammelt werden wie zuvor, indem die Intensität auf das Siebenfache des ursprünglich geplanten Wertes gesteigert wird. Im Hinblick auf den Ausbau des Beschleunigers wird ein wesentlich verbesserter Spurdetektor entwickelt, der um 2026 in Betrieb gehen soll. Der neue Detektor greift ausschließlich auf Siliziumdetektoren zurück und verwendet zahlreiche innovative Technologien, um mit den gravierenden Herausforderungen mitzuhalten, welche sich aus der gesteigerten Intensität ergeben. Der Pixeldetektor verwendet einen neuen Auslesechip, der die Pixelgröße auf ein Fünftel reduziert, um einzelne Teilchenbahnen in der durch eine hohe Teilchendichte charakterisierten Region nahe des Kollisionspunktes auseinanderzuhalten. Dünne planare n-in-p Sensoren sind aufgrund ihrer verringerten Verlustleistung und effizienten Ladungssammlung nach Bestrahlung ideal geeignet, um auch nach den massiven Strahlenschäden effektiv zu arbeiten.

Um den Sensoraufbau für die neue Pixelgröße von $50 \times 50\,\mu\text{m}^2$ zu optimieren, werden TCAD Simulationen genutzt. Damit einher gehen Analysen der Ladungssammlungseffizienz, des elektronischen Rauschens sowie der Eigenschaften des elektrischen Feldes jeweils vor und nach Bestrahlung. Darüber hinaus wird der RD53A Prototyp Auslesechip verwendet, um Module basierend auf den vorgeschlagenen dünnen planaren n-in-p Sensoren zu montieren. Die Leistungsfähigkeit unterschiedlicher Sensorvarianten wird basierend auf Daten von Teststrahlmessungen bewertet.

Die Effekte, welche durch eine Lagerzeit bei Raumtemperatur während der Wartungsarbeiten des Pixeldetektors entstehen, werden untersucht. Pixeldetektormodule basierend auf $100-150\,\mu\text{m}$ dünnen Sensoren werden durch Teststrahlmessungen evaluiert. Dabei werden das Ladungssammlungsverhalten sowie die Nachweiseffizienz vor und nach einer mehr als einjährigen Lagerung bei Raumtemperatur verglichen.

ABSTRACT

The Large Hadron Collider and its pre-accelerator complex will be upgraded in three steps to allow for the high luminosity phase. A factor of ten times more data will be collected in this period by facilitating the increased instantaneous luminosity being seven times as large as the original design value. A new inner tracker system is in preparation for the ATLAS detector in view of the high luminosity phase to start operation around 2026. This all silicon tracker relies on various innovative technologies to cope with the severe challenges arising from the increased luminosity. The pixel detector employs a new readout chip to decrease the pixel size to a fifth of the pixel size of the present generation to be able to disentangle all tracks in the high multiplicity environment close to the interaction point. Thanks to their reduced power dissipation and high charge collection efficiency after irradiation, thin planar n-in-p pixel sensors are ideally suited to cope with the expected unprecedented radiation damage.

TCAD simulations are being performed to optimise the sensor layout for the new pixel cell size of $50 \times 50\,\mu m^2$. In this study, charge collection efficiency, electronic noise and electrical field properties are investigated both before and after irradiation. The RD53A prototype readout chip is used to build modules based on the proposed thin planar n-in-p sensors. The performance of different sensor designs is assessed by analysing data from various test-beam campaigns.

The effects of storage time at room temperature for the ITk pixel detector during maintenance periods are reproduced on real modules. Pixel detector modules built with sensors of $100-150\,\mu m$ thickness are characterised with testbeam measurements. The charge collection and hit efficiencies are compared before and after annealing at room temperature up to one year.

TABLE OF CONTENTS

I

Introduction and motivation

In their aim of understanding the most fundamental building blocks of our universe, particle physicists are creating more and more complex experiments. The search begun with comparatively low energetic particles originating from radioactive sources and observing their scattering behaviour off different targets. The development of accelerators in the 1950s and 1960s enabled higher particle energies and thus better resolutions. Particle detectors around the interaction region allowed to record the interaction rates of various collisions of particle beams. By facilitating yet more advanced accelerator and detector techniques, the fundamental particles, classified as leptons, quarks and gauge bosons, were discovered. The Standard Model (SM) of particle physics [1, 2] describes all particles and their interactions via the weak, strong and electromagnetic forces. A particle consistent with the last missing piece of the SM, the Higgs boson, was discovered in 2012 at the Large Hadron Collider.

The upcoming high luminosity phase of the Large Hadron Collider will pose tough challenges to all detector components of the ATLAS experiment. This thesis aims to optimise and test the required sensor technology to instrument the pixel detector of the new Inner Tracker (ITk). The high particle density in the environment close to the interaction point demands smaller pixels, increased radiation tolerance and high rate capability. The thin n-in-p planar sensors described here are perfectly suited for these challenging conditions.

After an introduction to the Large Hadron Collider and the ATLAS experiment in Chapter 2, the working principle of silicon detectors (Chapter 3) and the utilised methods are explained (Chapter 4). Chapter 5 is dedicated to the building block of the ITk - quad modules, describing their assembly and laboratory testing. Two different applications of simulations applied to silicon devices are introduced in Chapter 6, one employs an established TCAD simulation suite while the other is a new implementation of the Hamburg model describing radiation damage and annealing in silicon. The results of TCAD simulations of small pixel cells for the ITk detector are presented in Chapter 7. Chapter 8 discusses simulation and experimental investigations concerning the impact of annealing on irradiated pixel detectors. Lastly, Chapter 9 contains results of pixel modules employing a prototype of the newly designed readout chip for the ITk - the RD53A chip. Results of various sensor designs are compared both before and after irradiation. The main results are summarised in Chapter 10. Natural units $\hbar = c = 1$ are used throughout this thesis.

THE LARGE HADRON COLLIDER AND THE ATLAS EXPERIMENT

The ultimate goal of particle physics is to probe and understand the fundamental elements of nature. During the last decades, particle beams have represented an essential tool for research in high energy physics. The energy of the beam particles determines the resolution with which the target can be examined. Fixed targets have widely been exchanged by collisions with another high energetic particle beam to maximise the available energy in the interaction. Massive particles can be created with the large amounts of available energy within the conservation laws. The short lifetimes of those particles cause instantaneous decays, the remnants of which can be detected by specialised detectors.

The Large Hadron Collider (LHC) [3], located at CERN (European organisation for Nuclear Research) at the Swiss-French border, represents the most energetic man-made particle accelerator. A 26.7 km long tunnel, 45 m to 170 m below ground, houses two-in-one superconducting magnets to accommodate two beam pipes and accelerate the two proton beams in opposite directions. The LHC is designed to operate proton beams with a center-of-mass energy of 14 TeV and a luminosity of $1 \times 10^{34}\,\mathrm{cm^{-2}s^{-1}}$. The proton beams cross each other at four interaction points (IPs). A detector is built around each of them: the two multi-purpose detectors, the ATLAS (A Toroidal LHC Apparatus) [4] and CMS (Compact Muon Solenoid) [5] experiments as well as the more specialised detectors, the ALICE (A Large Ion Collider Experiment) [6] and LHCb (LHC beauty) [7] experiments. All experiments aim for a better understanding of the SM and search for physics beyond.

The following sections will introduce the LHC with its accelerator complex and describe the ATLAS experiment in detail.

2.1 The Large Hadron Collider and its high luminosity upgrade

The LHC is hosted by the huge 26.7 km long tunnel that was built for the LEP (Large Electron Positron) collider from 1983 to 1986. The installation of the LHC machine begun after the LEP was shut down in 2000 and dismantled in the following year. The inauguration of the LHC followed in
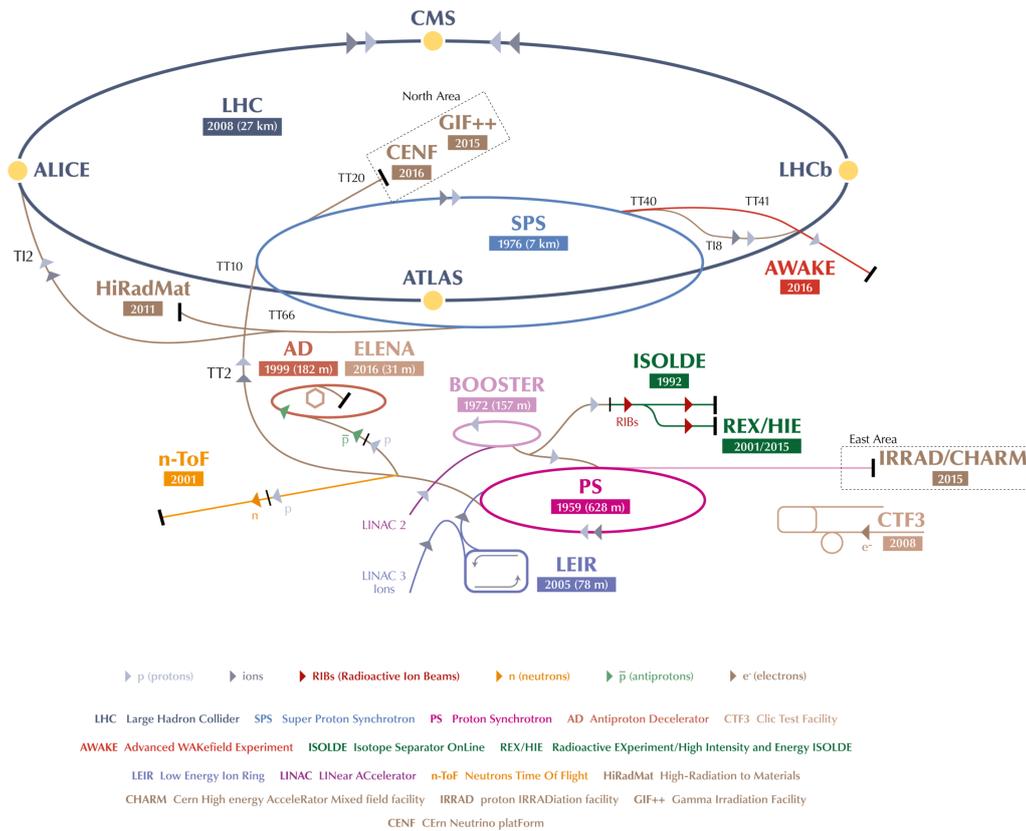
Figure 2.1: The CERN accelerator complex near Geneva. The SPS is the last ring in a chain of accelerators that are used to inject the pre-accelerated beam into the LHC. They also serve various smaller experiments at CERN aiming to investigate specialised aspects of particle physics. The image is taken from Reference [9].

the year 2008, but the first run of the LHC was delayed until 2010 due to technical problems.

The LHC depends on the historically grown accelerator chain at CERN as it is also displayed in Figure 2.1. The presently used acceleration chain starts with the LINAC-2[1] [8] that supplies a proton beam to the Booster. The Booster accelerates the beam in four superimposed rings before it is injected into the Proton Synchrotron (PS). Here, the beam energy is pushed to 25 GeV before the last pre-accelerator, the Super Proton Synchrotron (SPS), increases the beam energy further to 450 GeV. The injection into the LHC is done in buckets of 2.5 ns, each of which is 25 ns apart from the next. Finally, the LHC boosts the proton energy further up to 6.5 TeV, thus, resulting in a proton-proton center-of-mass energy of $\sqrt{s} = 13$ TeV.

It takes about 20 min to accelerate the two proton beams in the LHC to their full energy. The magnetic fields of the superconducting dipole magnets are ramped up to 8.4 T to enable the required bending of the proton beams. After the acceleration, the beams are collided at the four IPs for up to 24 hours. The nominal LHC design foresees 2808 proton bunches with a total energy of 664 MJ. The effectiveness of the collisions is characterised by the instantaneous luminosity which describes

---

[1]LINAC-4 will be used after Run-2 to supply the initial proton beam.

(a) Peak instantaneous luminosity

(b) Integrated luminosity
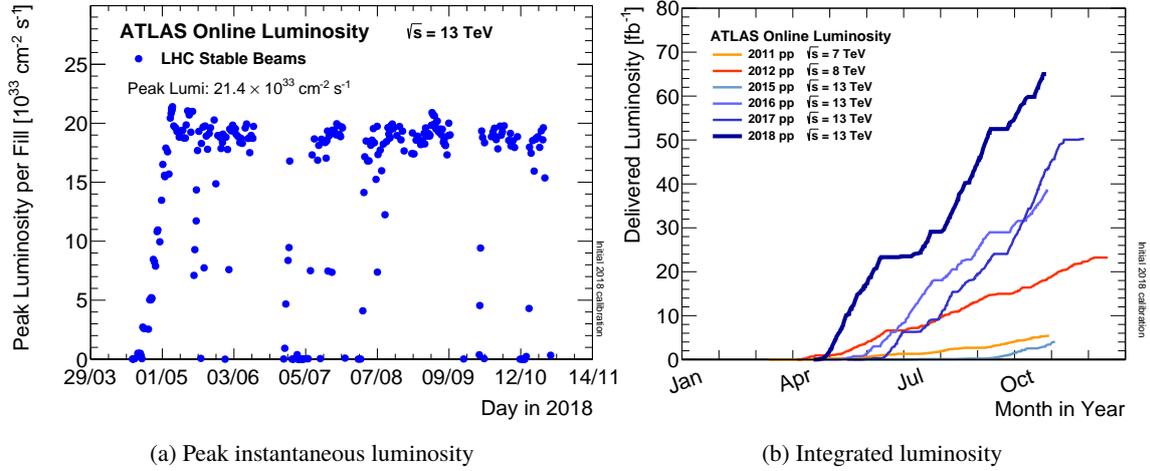
Figure 2.2: Benchmarks of the LHC success. The peak instantaneous luminosity in ATLAS in 2018 is shown in (a). The total integrated luminosity delivered to the ATLAS experiment as a function of time across the whole lifetime up to this date is shown in (b). The images are taken from Reference [10].

the rate of collisions at the interaction point. It is defined as

$$\mathcal{L} = f_{\text{rev}} \cdot N_b \frac{n_1 n_2}{4\pi \sigma_x \sigma_y} \tag{2.1}$$

with f_{rev} the revolution frequency, N_b the number of bunches per beam, $n_{1,2}$ the number of particles per bunch and $\sigma_{x,y}$ the Gaussian transversal beam dimensions in x and y. The instantaneous luminosity in the ATLAS detector in 2018 is shown in Figure 2.2a. The design luminosity is exceeded by more than a factor of two, and is already artificially reduced (*levelled*) to reduce the number of simultaneous interactions in one bunch-crossing (*pile-up*). The integrated luminosity describes the total amount of collisions and is typically expressed in units of the inverse cross-section fb^{-1}. The evolution of the integrated luminosity delivered to the ATLAS experiment is depicted in Figure 2.2b over the whole LHC lifetime up to this date.

The LHC schedule is summarised in Figure 2.3 [11]. The LHC program is organised in periods of data taking, called runs, and extended closure periods, called long shutdown (LS). The LHC started its productive operation in 2011 at a reduced $\sqrt{s} = 7$ TeV given the initial technical issues. It was increased to 8 TeV still in the so called Run-1 that lasted from 2011 to 2013. The period of Run-2 covers operations between 2015 and 2018. The gap in between is known as LS-1. The first shutdown period was used to prepare the machine for the operation at $\sqrt{s} = 13$ TeV that was finally used throughout Run-2. Furthermore, a significant gain in integrated luminosity was achieved during Run-2. The next shutdown phase starting at the end of 2018 is LS-2 and will be focused to already prepare the injectors for the High Luminosity LHC (HL-LHC). The LINAC-2 injector will be replaced by the LINAC-4 during this phase. Before the HL-LHC phase, Run-3 is scheduled from 2021 to 2023 with $\sqrt{s} = 14$ TeV.

All parameters of the initial LHC design as well as the HL-LHC upgrade are summarised in Table 2.1. The final phase of the LHC is supposed to start in 2026 and will increase the instantaneous luminosity

Figure 2.3: Schedule of LHC operations on the road to HL-LHC. The image is adapted from Reference [11].

Table 2.1: Comparison of beam parameters at the nominal LHC and the High Luminosity LHC [12] for the 25 ns bunch spacing.

Name and symbol	[Unit]	LHC	HL-LHC
Beam energy	[TeV]	7	7
Number of bunches N_b		2808	2808
Particles per bunch $n_{1,2}$	[10^{11}]	1.15	2.2
Revolution frequency f_{rev}	[kHz]	11.245	11.245
Normalised emittance ϵ_n	[μm]	3.75	2.5
Beta function at the IP β^*	[m]	0.55	0.15
Crossing angle θ_c	[μrad]	285	590
RMS bunch length σ_z	[cm]	7.55	7.55
RMS beam size at the IP σ^*	[μm]	16.7	13.6
Levelled luminosity $\mathcal{L}_{\text{lev}}$	[10^{34} cm^{-2} s^{-1}]	1	5
Peak luminosity $\mathcal{L}_{\text{peak}}$	[10^{34} cm^{-2} s^{-1}]	1.0	7.18
Collisions per bunch-crossing <μ>	[events]	19	200

to an unprecedented level of 7×10^{34} cm^{-2}s^{-1}. One of the key technologies to enable this huge improvement are the so called crab cavities that are currently being tested. They are made of high-purity niobium superconducting material to generate very high transverse voltages of 3.4 MV. The crab cavities will tilt the proton bunches in each beam to increase their overlap at the interaction point. Consequently, the protons will have to pass through the whole length of the opposite bunch, increasing the probability of collisions. The higher instantaneous luminosity will result in 200 collisions per bunch-crossing. This has to be compared to the 2018 average of 37 collisions per bunch-crossing [10].

2.1.1 Physics motivation

The upgrades to the accelerator and detectors will enhance the discovery potential of the LHC in its high luminosity phase [12]. In general, the increased integrated luminosity will decrease the statistical uncertainties in many analyses of important standard model processes [13]. Even for analysis that are not statistically limited, the enhanced data set will enable a stricter selection

of events resulting in purer data samples and reduced systematic uncertainties. Searches for new physics and rare decays will profit significantly as their limit to claim a discovery depends on the significance that increases with the square root of the number of events. Three exemplary research areas of particular interest are described in the following.

- **Higgs boson:** The discovery of a Higgs boson like particle with a mass of 126 GeV in 2012 [14] was a huge success for the ATLAS and CMS experiments. Further analyses confirmed the spin-0 and even parity properties of the discovered particle [15]. A number of expected decay channels were discovered one after the other such that in addition to the H→ $\gamma\gamma$ discovery channel also the H→ $\tau\tau$, H→WW and H→ZZ were observed [16]. The excellent performance of the LHC recently also allowed for the measurement of the H→b$\bar{\text{b}}$ channel [17]. The measurement of this decay is a mandatory probe of the mass generation for fermions through Yukawa interactions, as predicted in the SM. The particle manifests to be the SM Higgs boson based on all available measurements. With the HL-LHC and ten times more Higgs bosons produced, further parameters will become accessible. Apart from other decay channels of the Higgs boson, the measurement of the Higgs self coupling is a particularly demanding and, at the same time, interesting measurement, that will be possible thanks to the enhanced data set.

- **SM precision measurements:** The precise measurement of SM parameters, such as the mass of the top quark, allows to draw conclusions on many other aspects of the theory. The combination of top quark and Higgs boson mass, for example, has strong implications for the vacuum stability. The large amount of data becoming available with the HL-LHC will reduce the statistical uncertainties and enable new analysis techniques that rely on large data samples.

- **BSM:** Beyond the standard model (BSM) physics denote various theories and extensions to the SM. One of the most prominent ones is supersymmetry (SUSY), which postulates that every particle of the current SM has a supersymmetric partner. The predicted mass range of those partners is within reach of the HL-LHC. A discovery of SUSY could potentially solve the hierarchy problem [18]. Moreover, some SUSY particles are promising dark matter (DM) candidates. Also other models, for example Technicolor or Extra-Dimensions, can explain DM and other phenomena. The detection of those models is feasible as decay cascades with SM particles and missing energy/momentum in the final state are predicted. The increased number of events in the high luminosity phase will either lead to a discovery of a new particle, or significantly extend the constraints on many BSM models [19].

2.2 The ATLAS experiment

The ATLAS experiment is one of the four large experiments at the LHC. A schematic overview of the detector is given in Figure 2.4. The idea behind each of the high energy physics (HEP) experiments is to record the decay products of the particles produced by the proton-proton interaction. The
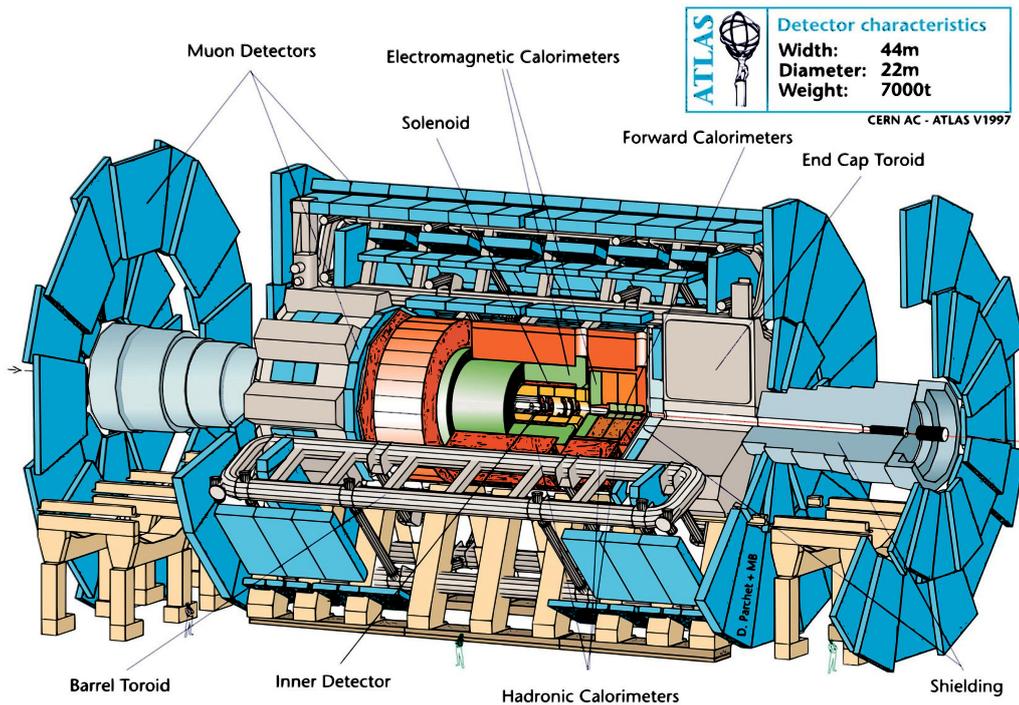
Figure 2.4: Schematic layout of the ATLAS detector. The image is taken from Reference [20].

elementary particles that are created during the interaction like Higgs bosons, top quarks, W and Z bosons and other heavy particles decay immediately after their productions but can be reconstructed from their decay products. The energy, momentum and charge of a particle can be measured, which also yields information about the mass of the particle. A magnetic field is typically used to bend particle trajectories. The charge is derived from the bending direction while the momentum is reconstructed from the curvature. This task is performed by a tracking detector that is placed close to the interaction point. The energy of the particle can be measured by stopping it completely in the calorimeter. The outermost layer is the muon detector that detects the lightly interacting muons. All of those components are present in the ATLAS detector and will be described in the following.

2.2.1   Detector components

Magnets

The design of the ATLAS detector is driven by its magnet system [21]. The central solenoid encloses the tracking system while providing a homogeneous 2 T magnetic field along the beam axis. An additional toroidal magnetic field is present outside the solenoid mainly for muon identification. The inhomogeneous toroidal magnetic field reaches up to 4 T. The design allows for a large calorimeter system since it is not restricted by the size of the central solenoid. Additionally, the large-scale magnetic field provided by the toroid enables muon tracking over a large distance ensuring a long lever arm for precise measurements.

Inner detector

The inner detector (ID) [22, 23] constitutes the innermost sub-detector of ATLAS. It is used to measure the trajectories of charged particles originating from the interaction point. The tracking information is not only used to determine the charge and the momentum of particles, but also to measure the vertex from which each particle is originating. Displaced or secondary vertices likely originate from particles that have a longer lifetime. Prominent examples are hadrons containing bottom (b) quarks. The quark hadronisation process leads to jets which are narrow cones of hadrons. The b-jets can be identified by the tracking system based on their displaced vertex that is typically in the order of 500 µm apart from the primary vertex.

The fundamental concept to measure tracks relies on a very lightweight detector. Otherwise, the particles are severely deflected by multiple scattering and other interactions with detector material. The innermost part of the ID is the pixel detector. It offers the highest granularity to efficiently disentangle tracks in the high multiplicity region close to the IP. A drawing of the present pixel detector is shown in Figure 2.5. The pixel detector consists of four barrel layers and three end-cap disks each side. The innermost barrel layer was inserted as a Phase-0 upgrade in LS-1 and is called insertable b-layer (IBL) [24]. The remaining layers, from the inside to the outside, are called B-Layer, Layer-1 (L1) and Layer-2 (L2) .

The three original layers consist of 1744 pixel detector modules built with FE-I3 [25] readout chips (see Section 3.5) and 250 µm thick planar n-in-n sensors. Each sensor features 328×144 pixel cells measuring $50 \times 400 \, \mu m^2$ each. Therefore, 16 readout chips are interconnected to each sensor in a module. The IBL exploits 448 FE-I4 [26] readout chips (see Section 3.5). The central part of the detector is composed of modules built from 200 µm thick planar n-in-n sensors interconnected to two readout chips. The matrix of pixels per readout chip is increased to 80×336 while the pixel cell size is decreased to $50 \times 250 \, \mu m^2$. The outermost four modules per end are implemented as single chip modules employing 230 µm thick 3D sensors [27, 28]. The collecting electrodes are implemented as columns going through the sensor thickness in 3D sensors resulting in shorter collecting distances decoupled from the sensor thickness and thereby offering intrinsically improved radiation hardness.

The semiconductor tracker (SCT) is a silicon micro-strip detector and is placed around the pixel detector. It is composed of four barrel layers and nine end-caps per side [30]. The strip detector works much like the pixel detector but with much longer electrodes. Each SCT module comprises of two or four single sided 285 µm thick p-in-n sensors. A single sensor has 768 strips and for each module side two of these sensors are wire bonded together. That results in a strip length of 123.2 mm with a pitch of 57−94 µm. A simpler manufacturing and assembly process results in lower cost per area compared to pixel detectors and thus makes strip detectors particularly suited to cover larger surfaces at larger radii.

The outermost part of the ID is the transition radiation tracker (TRT) [31, 32]. As the other tracking detectors, it consists of a barrel part and two end-caps. The detector is built from 4 mm thick gas-filled straw tubes with a length of either 144 cm in the barrel or 37 cm in the end-caps. They contain a gas mixture of 70 % Xenon, 27 % CO_2 and 3 % O_2. Each tube acts as a proportional counter with

Pixel Disk

Pixel Disk

IBL

Pixel B-Layer

Pixel Layer-1

Pixel Layer-2

r = 42.5 mm : IBL Support Tube (IST)

r = 33 mm : IBL Stave

r = 29.3 mm : Inner Positioning Tube (IPT)

r = 24.3 mm : New Beam Pipe

r = 50.5 mm : B-Layer

r = 88.5 mm : Layer-1

r = 122.5 mm : Layer-2
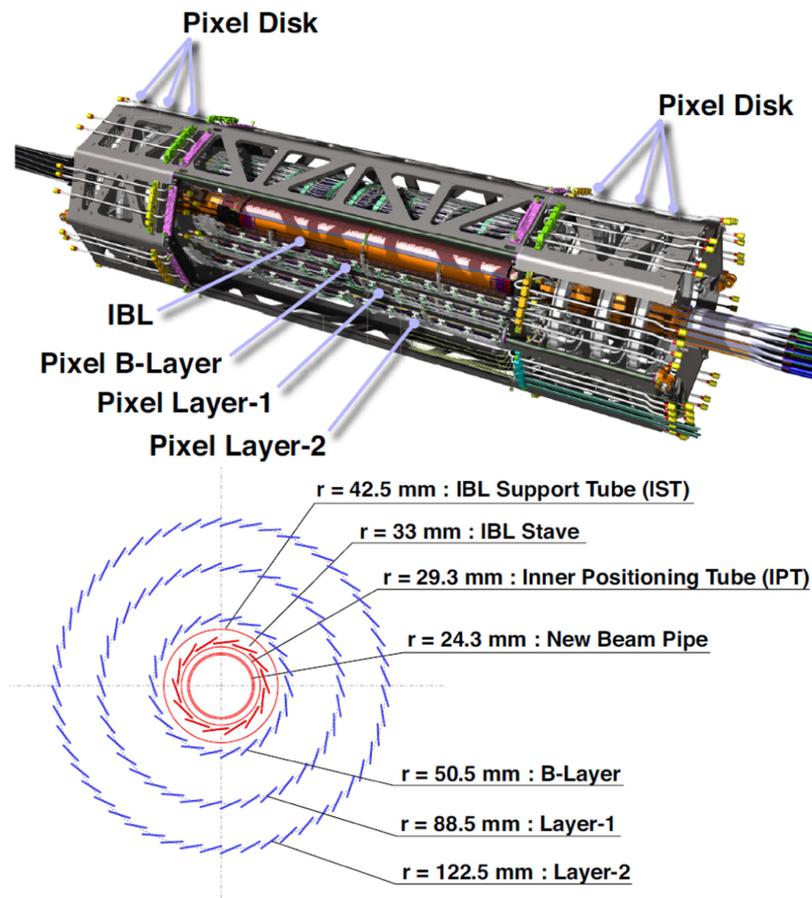
Figure 2.5: Technical drawing of the pixel detector after its Phase-0 upgrade with the IBL. The image is taken from Reference [29].

the wall kept at a voltage of 1530 V and the central wire at ground potential. Radiation foils and fibres between the tubes cause impinging particles to create transition radiation, photons emitted by highly relativistic charged particles as they traverse a material boundary. The effect depends on the ratio of the particles energy and mass and the resulting photons can be used to distinguish pions from electrons.

Calorimeters

The calorimeters, placed outside the ID, measure the energy a particle loses as it passes through the detector. A calorimeter is usually designed to stop particles entirely, forcing them to deposit all of their energy within the detector. The ATLAS calorimeter is split into two specialised parts that account for the properties of different particles.

- **ECAL** The radiation length X_0 describes the mean distance over which the energy of high energetic electrons or photons is reduced by bremsstrahlung to 1/e. As the radiation length is related to the atomic number Z of the material that is penetrated, high-Z materials are used to stop particles in a reasonable distance. For example, the radiation lengths of iron and lead are $X_0(\text{Fe}) = 1.76\,\text{cm}$ and $X_0(\text{Pb}) = 0.56\,\text{cm}$ [33]. In ATLAS, the electromagnetic

calorimeter (ECAL) is used to measure the energy of electrons/positrons and photons. It is built as a sampling calorimeter with passive lead absorbers and liquid argon (LAr) as active material [34]. Its accordion shaped electrodes enable a hermetic coverage and fast signal shaping required by the 25 ns bunch-crossing interval [35]. The barrel part of the calorimeter starts just outside the cryostat of the central solenoid. It is completed with two end-caps in the forward regions.

- **HCAL** The hadronic analogy to the electromagnetic radiation length is the hadronic inter- action length λ_I. It is defined as the average distance between hadronic interactions and is usually significantly larger than the radiation length as for example $\lambda_I(\text{Fe}) \approx 17\,\text{cm} \gg X_0(\text{Fe}) = 1.76\,\text{cm}$ [33]. In ATLAS, the hadronic calorimeter (HCAL) is used to stop and measure the energy of hadrons [36]. The detector technology varies and depends on the environmental requirements in the different regions. The barrel region which is less affected by radiation is built in a sampling approach of iron absorbers and polysteren scintillators as active material. The end-cap is instead located in the high radiation forward region and is therefore equipped with the intrinsically radiation hard LAr active material. Copper and tungsten absorbers are used as passive material in those regions.

Muon system

The last and outermost detector sub-system is built with the purpose of detecting and measuring muons. The muons are mainly electromagnetically interacting and the probability of bremsstrahlung is significantly reduced since muons are much heavier than electrons and positrons. Therefore, they can penetrate even dense materials and escape the detector volume. The muon system, located around the remaining ATLAS sub-detectors, measures the tracks of the escaping muons in the magnetic field of the toroid [37]. This measurement enables the determination of the momentum of the muons based on their curvature.

The barrel part of the muon system is composed of two layers of resistive plate chambers (RPCs) and three layers of monitored drift tubes (MDTs). The MDTs are composed of 30 mm thick drift tubes that work in proportional mode. The end-caps consist of thin gap chambers (TGCs) and cathode strip chambers (CSCs) plus two disks per side of MDTs.

2.2.2 Coordinate system

The detector is described in a common coordinate system throughout ATLAS. The origin is located at the IP in the center of the detector. The azimuth angle ϕ is defined around the beam axis while the z axis points along the beam axis. The x-y plane is perpendicular to the beam direction and is called the transverse plane. The positive x-axis points from the IP to the center of the LHC while the positive y-axis points upwards. An additional η parameter is introduced and defined as

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right) \tag{2.2}$$

with θ the polar angle to the beam line. The value of $\eta = 0$ describes the direction perpendicular to the beam line while $|\eta| = \infty$ is reached for the direction along the beam line. The η parameter
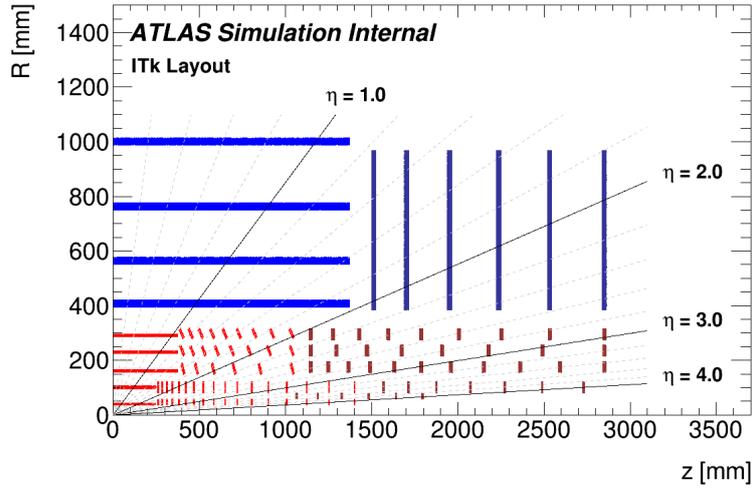
Figure 2.6: Schematic layout of a quarter of the ITk detector. The pixel detector is indicated in red while the strip detector is drawn in blue. Taken from Reference [40].

is particularly interesting in hadron colliders as each η bin contains the same flux of particles originating from the collision at the IP.

The ATLAS detector provides hermetic coverage in ϕ up to $|\eta| = 2.7$. Some sub detectors provide even more coverage, e.g. IBL extends to $|\eta| = 3.0$, the ECAL to $|\eta| = 3.2$ and the HCAL up to $|\eta| = 4.9$.

2.2.3 Inner tracker upgrade

For operation at the HL-LHC, the whole ID of the current ATLAS detector will be replaced by the ITk. The challenges posed by the accelerator upgrade are all driven by the increased instantaneous and integrated luminosity. The increased instantaneous luminosity results in more proton-proton collisions per bunch-crossing, yielding more tracks in the tracking detector that need to be reliably disentangled. The foreseen integrated luminosity at the end of the HL-LHC operations of around $3000-4000\,\text{fb}^{-1}$ will cause severe radiation damage especially in the innermost parts of the pixel detector.

The design of the ITk foresees the usage of an all silicon tracker composed of five layers of pixel detectors [38] and four layers of strip detectors [39]. Several pixel detector rings are installed instead of the typical end-caps to best facilitate the space constraints. The proposed layout can be seen in Figure 2.6. The total silicon surface of the pixel detector will be about $13\,\text{m}^2$ and thus being almost ten times as large as the current detector.

The current level of occupancy per pixel can be maintained despite the higher particle density by reducing the pixel cell size from $50{\times}250\,\mu\text{m}^2$ and $50{\times}400\,\mu\text{m}^2$ to either $50{\times}50\,\mu\text{m}^2$ or $25{\times}100\,\mu\text{m}^2$. Furthermore, the thickness of planar sensors is reduced from $200\,\mu\text{m}$ and $250\,\mu\text{m}$ to $100\,\mu\text{m}$ and $150\,\mu\text{m}$. As a consequence, inclined tracks result in less pixels traversed in thinner sensors such that the required bandwidth for the readout of the data is reduced. The reduced sensor thickness also leads to an increased radiation tolerance and reduced power dissipation on account of the lower

required bias voltage for operation and the reduced leakage current. The sensor technology is chosen according to the requirements at a given radius. Therefore, 3D sensors are chosen as the baseline option to instrument the innermost layer which will be exposed to about 1.4×10^{16} n_{eq}/cm^2[2]. The next layer will be equipped with $100\,\mu m$ thick planar n-in-p sensors mainly developed at the Max-Planck Institut für Physik (MPP) [41]. This second layer will be exposed to a fluence of about $3–4 \times 10^{15}$ n_{eq}/cm^2. The two innermost layers are located in the inner support tube, enabling the exchange of those layers after half the HL-LHC lifetime. The remaining three layers employ $150\,\mu m$ thick planar n-in-p sensors. All modules based on planar sensors will be built from quad chip modules that utilise a single sensor with four readout chips (see Chapter 5). The innermost layer will instead use single and double chip modules.

A dedicated readout chip is developed by the RD53 collaboration [42] (see Section 3.5.2). The chip implements the new pixel cell size and achieves the required readout speed (bandwidth) and radiation tolerance. The readout chip will enable the usage of serial powering chains. The reduction in cables required is one of the largest reductions in material budget compared to the current detector. The ITk design foresees no more than two radiation lengths in any direction up to $|\eta| = 4$. This is to be compared to up to nearly 6 radiation lengths for the present detector. The decreased material budget will result in less scattering in the tracker and a lower fraction of photon conversions, which will improve the energy resolution of the calorimeters.

Expected improvements

The ITk pixel detector upgrade of the ATLAS experiment will be an important component to explore the full physics potential of the proton-proton collisions in the high luminosity phase. It aims not only to maintain the high performance of the current pixel detector despite the much more challenging environment at the HL-LHC, but it will even extend its abilities. The η coverage will be extended from $|\eta| \leq 3.0$ in IBL to $|\eta| \leq 4.0$ for ITk. This will enable the reliable reconstruction of forward jets beyond $|\eta| = 2.5$ which will be essential to distinguish the vector boson fusion (VBF) and gluon fusion Higgs boson production channels. The same ability is crucial to extend the detection of vector boson scattering that is characterised by the presence of high p_T jets in the forward region. Furthermore, the high b-tagging efficiency feasible with the ITk will be essential to reduce the hadronic background in all physics analyses that include b quark final states [43].

---

[2]Fluence expressed in units of 1 MeV neutrons that correspond to the same radiation damage as the original irradiation with arbitrary particles and energies (see Section 3.2).

SILICON PARTICLE DETECTORS

Silicon particle detectors are widely used in HEP applications. Especially, pixel detectors are employed as tracking detectors as they offer an unrivalled combination of excellent spatial and time resolution. In this application, they record the paths of charged particles originating from high energetic particle collisions enabling high precision trajectory reconstruction even in the high track-density environment close to the interaction point. The measurement of the trajectory in a magnetic field reveals the particle momentum. Pixel detectors are also essential for the primary and secondary vertex reconstruction of which the latter is particularly important for b-tagging.

The following sections will provide an overview from the fundamental working principle to manufacturing and application of pixel detectors.

3.1 Working principle of silicon particle detectors

Solid materials can be categorised into crystalline and amorphous structures. While amorphous materials lack a long range ordering, defined shape and pattern, crystalline materials are microscopically highly ordered structures. The lattice extends on a large scale in all directions. As a consequence, the sharply distinct energy levels of electrons (e) in individual atoms influence each other and smear out to energy bands. The highest fully filled energy level in each atom is called valence band. The electrons of the valence band are available for the covalent atom binding with the neighbouring atoms in the lattice. The next higher energy level starts the so called conduction band. Charge carriers in the conduction band have sufficient energy to move independently of the atoms through the lattice and are available for conduction. The conductivity of any material is given by the number of charge carriers in the conduction band. Hence, a material is called conductor if valence and conduction band overlap, and insulator if the two bands are far apart. If the two bands are still separate but closer than $3\,\mathrm{eV}$, the material is defined as a semiconductor. The energy difference between one band and the other is called band-gap energy (E_g).

In semiconductors, electrons can only cross the band-gap if they have enough energy, either from thermal excitation, charged particle interaction or photon absorption. In the thermodynamic equilibrium, the probability to find an electron with a given energy E_e as a function of the temperature

T is

$$f_\mathrm{e}(E_\mathrm{e}) = \frac{1}{1 + \exp\frac{E_\mathrm{e} - E_f}{k_\mathrm{B}T}} \tag{3.1}$$

where E_f is the Fermi energy. The Fermi energy for intrinsic, i.e. defect free and pure, semiconductors is in the center of the band-gap

$$E_f \approx \frac{E_C + E_V}{2}. \tag{3.2}$$

Here, E_V and E_C are the energy limits of valence and conduction band with $E_V - E_C = E_g$. Equation 3.1 shows that at $T = 0\,\mathrm{K}$ all electrons are bound in the valence band with energies up to $E = E_f$. For temperatures $T > 0\,\mathrm{K}$, there is an increasing probability for electrons to be lifted across the band-gap into the conduction band. There is a positively charged vacancy or hole (h) in the valence band for each electron in the conduction band. Therefore, the number of electrons (n) and number of holes (p) available for conduction is equal in intrinsic silicon

$$n = p. \tag{3.3}$$

The intrinsic carrier concentration n_i is constant and given by

$$n_i^2 = n \cdot p. \tag{3.4}$$

The conductivity σ of a semiconductor depends on the carrier concentrations of electrons and holes as well as the mobilities of the respective carrier types ($\mu_\mathrm{e,h}$)

$$\sigma = \mathrm{e}(n\mu_\mathrm{e} + p\mu_\mathrm{h}) \tag{3.5}$$

where e is the elementary charge.

3.1.1 Silicon

Since silicon is the most common semiconductor in modern technology it will be considered from here on, although the general principles mostly work for all semiconductors. The silicon crystal is of diamond type and has four valence electrons each of which is engaged in a covalent binding to an electron of a neighbouring atom. The band-gap energy of silicon at $300\,\mathrm{K}$ is $1.12\,\mathrm{eV}$ [44] leading to an intrinsic carrier concentration of about $10^{10}\,\mathrm{cm}^{-3}$. However, intrinsic silicon is a rather hypothetical construct as even the purest manufactured silicon always comes with some unavoidable impurities and defects. Moreover, artificial impurity atoms called *dopants* are purposefully introduced into the lattice in a controlled way. By intentionally steering and placing certain dopants in the silicon crystal, a junction is created that is the base of the functionality of the pixel sensor. The dopants transform the silicon to either n- or p-type. For n-type material, dopants with one valence electron more from the fifth main group of elements, called donors, are used. The most common donor for silicon is Phosphorous, although also Arsenic and Antimony are adopted. If the Phosphorous atom is integrated into the silicon lattice, the excess valence electron is just weakly bound to the atom. An energy level in the band-gap close to the conduction band is created,

which is usually ionised at room temperature. For p-type material on the other side, elements of the third main group having just three valence electrons are used. The most common dopant of this kind, called acceptor, is Boron. The Boron atom has not enough valence electrons for all covalent bindings to the neighbouring silicon atoms. Therefore, an electron from one of the neighbouring silicon atoms fills the gap at room temperature, such that a hole is released into the valence band. Consequently, the introduction of acceptors creates states in the band-gap close to the valence band. Typical doping concentrations range from some $10^{11}\,\mathrm{cm^{-3}}$ in the bulk of high resistive detector material to $10^{21}\,\mathrm{cm^{-3}}$ in highly doped implants (see e.g. Section 7.1). In most cases, the intrinsic carrier concentration is exceeded by several orders of magnitude.

The new energy levels in the band-gap pull the Fermi energy towards the respective side. Although the introduction of dopants creates free charge carriers available for conduction, the material still stays electrically neutral.

3.1.2 The pn-junction

While n- and p-type materials have interesting functionalities already by themselves, the application of semiconductors relies on the combination of the two. When two pieces of silicon, one n- and one p-type, are brought together, the gradient in carrier concentration causes a diffusion driven exchange of electrons and holes. The majority carriers, electrons for n-type material and holes for p-type material, diffuse into the opposite type material. The carriers start to recombine at the interface growing into the depth of the silicon, the bulk. The previously electrically neutral material charges up as a consequence of the movement of the electrons and holes. This charge causes an electrical field which acts against the diffusion process due to the carrier concentration. An equilibrium is reached as soon as both forces are equal. The electrical potential Φ inside the junction is derived by solving the Poisson equation[1]

$$\triangle\Phi = -\frac{\rho}{\epsilon} \tag{3.6}$$

where ρ is the charge density and $\epsilon = \epsilon_0 \cdot \epsilon_\mathrm{r}$ is the absolute permittivity. The region in which an electrical field is present is called space charge or depletion region. The voltage corresponding to the electrical field in the junction is called built-in voltage V_bi and is typically about $0.7\,\mathrm{V}$. Depending on the resistivity of the material, this is enough to deplete a few micrometer of silicon. The depletion zone can be extended by applying an external bias voltage to the junction. For a reverse bias, the p-side is connected to the negative potential, while the n-side is connected to the positive potential. In this configuration, the depletion zone d_dep is

$$d_\mathrm{dep} = \sqrt{\frac{2\epsilon_0\epsilon_\mathrm{r}(V_\mathrm{b} + V_\mathrm{bi})}{e}\left(\frac{1}{N_\mathrm{D}} + \frac{1}{N_\mathrm{A}}\right)} \tag{3.7}$$

with V_b the external bias voltage, $N_\mathrm{D,A}$ the donor and acceptor concentrations and $\epsilon_{0,\mathrm{r}}$ the permittivity of the vacuum and material. This can typically be simplified since the external bias voltage is usually

---

[1] $\triangle\Phi = \nabla \cdot \nabla\Phi = \mathrm{div}(\mathrm{grad}(\Phi))$.

much larger than the built-in voltage and one doping (N) is dominant in most applications. The simplified formula is given by

$$d_\text{dep} \approx \sqrt{\frac{2\epsilon_0\epsilon_\text{r}V_\text{b}}{eN}}.$$ (3.8)

The voltage required to deplete the entire device thickness is called depletion voltage (V_dep). Then, all free charge carriers are removed from the bulk and newly released free charge carriers are immediately accelerated towards the electrode of opposite polarity by the electrical field. The current caused by the constant thermal emission of electrons and holes is called leakage current. Consequently, the more material is connected to an electrode, the more leakage current is collected. As long as the device is not yet depleted, the leakage current increases with the square-root of the voltage and linearly with the depleted volume as more and more generation centers contribute to the current. In an ideal system, the leakage current would be constant as a function of the voltage as soon as the device is depleted. In reality, higher order effects as the Hurkx trap assisted tunnelling (see Section 6.1.2) cause the leakage current to keep slowly increasing even after full depletion. This continues until the field somewhere in the device is high enough to cause drifting charge carriers to gain sufficient energy to ionise further atoms. This process is known as avalanche and results in the breakdown of the device. In this state, the current increases exponentially with the bias voltage.

3.1.3 Charge generation and signal formation

A depleted pn-junction can be employed to detect high energetic charged particles and photons. Starting point is the depleted silicon without any free charge carriers apart from the small leakage current. Any charged particle passing through the depleted area will loose energy by scattering with the electrons of the silicon. The electrons gain sufficient energy to cross the band-gap and enter the conduction band. Simultaneously, a hole is created in the valence band. The electrons and holes drift apart, each towards the electrode of opposite polarity. The number of generated electron-hole pairs by an impinging particle can be calculated by dividing the energy loss per traveled distance $\frac{\text{d}E}{\text{d}x}$ by the mean energy required to create an electron-hole pair $W = 3.6\,\text{eV}$ [45]. This energy does not correspond to the band-gap energy since silicon is an indirect semiconductor in which the energetic transition from valence to conduction band must be accompanied by a change in momentum. Thus, to lift an electron to the conduction band, a phonon carrying away the remaining momentum has to be created and the required energy has also to be provided.

For relativistic heavy particles, the specific energy loss can be calculated by the Bethe-Bloch equation [2]

$$-\left\langle \frac{dE}{dx} \right\rangle = Kz^2 \frac{Z}{A\beta^2} \left[\frac{1}{2} \ln \frac{2m_\text{e}c^2\beta^2\gamma^2 T_\text{max}}{I^2} - \beta^2 - \frac{\delta(\beta\gamma)}{2} \right]$$ (3.9)

with $K = 0.307\,075\,\text{MeV}\,\text{mol}^{-1}\text{cm}^2$, m_e the mass of the electron, z the charge of the impinging particle in units of the electron charge, Z the atomic number of the absorbing material, A the atomic mass of the absorbing material, β the velocity of the traversing particle as a fraction of the speed of light, γ the Lorentz factor ($\gamma = 1/\sqrt{1-\beta^2}$) and I the mean excitation energy (173 eV).

The parameter T_{max} describes the maximum kinetic energy that can be transferred to an electron by a particle of mass M. The expression $\frac{\delta(\beta\gamma)}{2}$ describes the density-effect correction [2]. Most particles relevant in HEP can be categorised by a single energy loss as the Bethe-Bloch equation exhibits a prominent minimum at $\beta \cdot \gamma = 3 - 4$ with a very slow increase towards higher values. So called minimal ionising particles (MIPs) exhibit an ionisation compatible with the minimum of the Bethe-Bloch equation.

Electrons are not accurately described by the Bethe-Bloch equation since they are very light and dominantly emit bremsstrahlung. Also for high momentum particles, the energy deposition and charge generation is increased through energy losses by bremsstrahlung and hard scattering processes.

There is another deviation from the Bethe-Bloch equation due to the typically employed thin sensors which are subject to large fluctuations of the deposited energy. The effect is caused by δ electrons, which are high-energetic secondary knock-off particles. In thin sensors, they can either be emitted parallel to the surface and be absorbed entirely or they can be emitted perpendicular to the surface and only a fraction of the energy is deposited in the device before the δ electron escapes the device. The energy deposition in thin sensors is described by a Landau distribution with a long tail towards high deposited energies that accounts for high energy depositions by e.g. δ electrons. Since most particles deposit energies below the mean value, the most probable value (MPV) is generally used to describe the typical energy deposition.

Photons can also release charge in the sensitive volume by either photo-effect, Compton-scattering or pair-creation. The most relevant processes, at photon energies below 1 MeV, are the two former ones in which all, or part of the photon energy is transferred to an electron of the silicon. The now high energetic secondary electron releases free charge carriers as previously described.

In any case, the cloud of generated free charge carriers drifts towards the electrodes with a velocity depending on the electrical field E as

$$v_{\mathrm{e,h}}(E) = \mu_{\mathrm{e,h}}E. \tag{3.10}$$

That holds for low and medium electrical fields, but deviates at higher electrical fields as the mobility itself becomes a function of the electrical field and decreases for high field strengths $E > 10^3\,\mathrm{V\,cm^{-1}}$ [46]. A current is induced on the collecting electrodes as soon as charge carriers drift towards them. The current is described by the Shockley-Ramo theorem [47, 48] and is given by

$$I = \mathrm{e}\left(\underline{v} \cdot \underline{E}_{\mathrm{w}}\right) \tag{3.11}$$

with $\underline{E}_{\mathrm{w}}$ the vector of the weighting field and $\underline{v}$ the vector of the carrier velocity [49]. The weighting field is a simple linear function between n- and p-doped electrodes in a 1-dimensional pn-junction, as in the case of a diode, but becomes more complex and relevant in pixelated detectors. In these complex configurations, the weighting field accounts for the geometry of the device. It is derived from the weighting potential Φ_{w} as

$$\underline{E}_{\mathrm{w}} = -\underline{\nabla}\Phi_{\mathrm{w}}. \tag{3.12}$$

Table 3.1: Summary of all doping combinations for pixel sensors.

bulk pixel	n	p
n	+ radiation hard - double-sided process	+ radiation hard + single-sided process
p	+ single-sided process - no under-depleted operation	- double-sided process - no under-depleted operation

The weighting potential has to be derived from the Laplace equation $\triangle \Phi_w = 0$ with special boundary conditions in which the voltage of the respective electrode is $1\,\text{V}$ and all others have $0\,\text{V}$ potential. The charge collected by the electrode can be calculated by integrating the current from the starting time of the carrier movement until the carriers are trapped or arrive at the electrode.

3.1.4 Pixel detectors

In a planar pixel detector, the simple combination of an n- and p-doped piece of silicon is mostly replaced by a bulk material being of one doping type and pixelated electrodes of opposite doping type on the surface. In some cases, a segmented electrode of the same doping type as the bulk is used with the pn-junction located on the backside with a uniform implant of the opposite type. In general, four combinations of dopings are imaginable, summarised in Table 3.1.

The terminology is as x-in-y with x the type of doping of the pixelated electrode and y the bulk material. The technology investigated in this thesis is n-in-p, i.e. an n-doped pixelated surface with p-type bulk material. The junction is located at the collecting electrode making an under-depleted operation after heavy irradiation possible. The single-sided process allows for cost effective production particularly interesting to cover large areas in excess of $1\,\text{m}^2$. The n-in-p technology was developed to replace the n-in-n technology, currently used in the ATLAS pixel detector. Here, the junction is initially located at the backside, requiring a fully depleted sensor to register the signal at the collecting electrode. After irradiation, space charge sign inversion (SCSI) converts the bulk to p-type (see Section 3.2), effectively moving the junction to the front-side. Therefore, n-in-n detectors after irradiation behave like n-in-p detectors and can be operated under-depleted, making both technologies particularly attractive for high radiation environments. The n-in-n technology, however, requires a more expensive double-sided process making it less suited for large area detectors. The p-in-n technology on the other side allows for the cheaper single sided processing but is problematic after irradiation since the SCSI converts the bulk to p-type and moves the junction to the backside which requires full depleted operation. Lastly, the p-in-p technology is a rather theoretical concept since it joins the detrimental properties of the other concepts, high costs due to double-sided process and difficult operation throughout the lifetime because of the location of the junction at the backside.

The main argument for a pixel detector over previously used technologies such as wire chambers or strip detectors is the high spatial resolution. The resolution for a pixel detector with pitch p and a binary readout of just the one pixel with the highest charge for a uniform track distribution

($D(x_t) = p^{-1}$) can be calculated as the average difference of hit position ($x_h = 0$, with the center of the pixel being at 0) and track position (x_t) to be

$$\sigma^2_{\text{position}} = \frac{\int_{-p/2}^{p/2}(x_t - x_h)^2 D(x_t)dx_t}{\int_{-p/2}^{p/2} D(x_t)dx_t} = \frac{\frac{1}{p}\int_{-p/2}^{p/2} x_t^2 dx_t}{\int_{-p/2}^{p/2} \frac{1}{p}dx_t} = \frac{p^2}{12} \tag{3.13}$$

according to [50]. The resolution follows to be

$$\sigma_{\text{position}} = \frac{p}{\sqrt{12}}. \tag{3.14}$$

Consequently, even in the simplest configuration, a resolution of about 0.29 times the pixel pitch is achieved. Since modern pixel detectors usually offer low threshold operation and information about the amount of recorded charge per pixel, clusters consisting of several pixels lead to a much improved resolution by employing charged weighted position reconstruction.

3.2 Radiation damage in silicon detectors

The flux through the silicon sensors in the high particle density environments close to the interaction regions of modern particle detectors causes radiation damage which alters various aspects of the sensors. The damage can be categorised into two categories: bulk damages and surface effects. The details of both categories and the consequences the radiation damage has on detector properties and operation will be discussed in the following part.

3.2.1 Bulk damage

The damage to the silicon bulk is caused by the impinging particles hitting atomic nuclei of the crystal lattice. While low energetic particles tend to create point defects with single atom displacements, high energetic particles typically cause cluster damages spanning $10-200$ nm [51]. The required energy to displace an atom from its original position in the crystal lattice was measured to be around $13-20$ eV [52]. The damage caused by various species of particles at different energies varies and can be compared through the non ionising energy loss (NIEL) hypothesis [53]. A particle-type and energy dependent hardness factor κ is used to scale the radiation damage

$$\Phi_{\text{eq}} = \kappa\Phi_{\text{orig}}, \ [\Phi_{\text{eq}}] = n_{\text{eq}}/\text{cm}^2. \tag{3.15}$$

A particle flux Φ_{orig} of given type and energy is scaled to the required flux Φ_{eq} of neutrons of 1 MeV to cause the same damage. The bulk damages consist of either missing atoms in the lattice (*vacancies*) or of additional misplaced atoms (*interstitials*). Defects are additional energy levels in the band-gap being either electrically neutral or of acceptor or donor type.

The defects cause changes of various detector properties. They act as generation centers regarding the leakage current. As a result, the leakage current I_{leak} is linearly increasing with fluence [54]

$$I_{\text{leak}} = \alpha(t, T) \cdot \Phi_{\text{eq}} \cdot V \tag{3.16}$$

where V is the sensor volume contributing to the leakage current and $\alpha(t, T)$ is a time and temperature dependent scaling factor. The defects responsible for the carrier generation are not stable in time and can anneal at high temperatures or on long time scales. The annealing mechanisms are interstitials being bound to vacancies or free bonds from other elements in the lattice such as oxygen. The mathematical description as well as further details on the annealing of the leakage current is presented in Section 6.2.2.

A second detector property that is affected is the depletion voltage. The electrically active defects act exactly like donor or acceptor atoms and increase the effective doping concentration and thereby the depletion voltage. The depletion voltage V_{dep} of a sensor of thickness d is calculated as a function of the effective doping concentration N_{eff} according to Equation 3.8

$$V_{\text{dep}} = \frac{e \cdot d^2}{2\epsilon_r \epsilon_0} \cdot N_{\text{eff}}. \tag{3.17}$$

The effective doping concentration is calculated as the sum of original dopants and radiation caused defects. The defects introduced by irradiation are mainly of acceptor type and thereby drive the effective doping concentration of the bulk towards p-type. The type inversion occurring if the not-irradiated sensor bulk is n-type is called SCSI. Depending on the initial doping concentration, the SCSI occurs after more or less radiation but, for the usually employed high-resistive material, typically at about $\Phi_{\text{eq}} = 10^{12}\,\text{n}_{\text{eq}}/\text{cm}^2$. This value is low compared to the typical fluences in the LHC environment of several $10^{15}\,\text{n}_{\text{eq}}/\text{cm}^2$ such that SCSI occurs early in the lifetime of the detector. The depletion voltage of detectors built on n-type material decreases up to the point of SCSI and increases from then on. As explained above, defects are not stable and so detector properties change in time even after the irradiation has stopped. The effective doping concentration, according to the Hamburg model [54], is subject to four effects and is calculated as the sum of them

$$N_{\text{eff}} = N_c^{\text{dr}}(t) + N_c^{\text{a}}(t) + N_r^{a,1}(t, T) + N_r^{a,2}(t, T). \tag{3.18}$$

The time dependence of the first two terms is proportional to the fluence rate φ_{eq} while the other two have a direct time dependence and an additional temperature dependence. The description of decaying defects results in the direct time and temperature dependence of the latter two terms. The effects relevant for the depletion voltage are summarised in the following.

- **Dopant removal:** The dopant removal (dr) describes a process within which the initial dopants regardless of type are electrically deactivated. As this contribution depends on the initial doping concentration, it is particularly small for high resistivity material most used in high energy physics applications. Ergo, the initial dopants are usually exhausted at fluences in the order of $10^{13}\,\text{n}_{\text{eq}}/\text{cm}^2$, being at least one order of magnitude smaller than the design fluences of modern vertex detectors. An exponential decay of initial doping atoms is considered in the mathematical description of this effect

$$N_c^{\text{dr}}(t) = N_{\text{eff}}^{0,\text{nr}} + N_{c,0} \cdot \left(1 - e^{-c\varphi_{\text{eq}}t}\right) \tag{3.19}$$

  where $N_{\text{eff}}^{0,\text{nr}}$ is the fraction of not removable (nr) initial dopants and $N_{c,0}$ is the number of removable dopants. The ratio of removable and not removable dopants depends on the kind

of irradiation. Charged particles tend to remove the complete initial doping concentration whereas neutral particles can only remove about 80% or even less. Further investigations showed that this process is also dependent on the oxygen level of the silicon substrate. Given the majority of particles close to the interaction point as it is the case for IBL and B-Layer of the ATLAS Inner Detector being charged particles, 99 % of initial donors are assumed to be removable. The constant c depends on the initial effective doping concentration as $c = 0.109/N_{\text{eff}}$ [54]. The product of dose rate φ_{eq} and a given time period t results in the total dose per time step.

- **Constant damage:** The constant damage describes the introduction of acceptor type defects which are stable in time and cannot anneal. The constant damage depends on the particle type and is typically lower for charged particle irradiation in the case of diffusion oxygenated float zone (DOFZ) silicon [55]. This damage sector can be described by

$$N_{\text{c}}^{\text{a}}(t) = g_{\text{C}}\varphi_{\text{eq}}t \qquad (3.20)$$

with an introduction rate g_{C}. The introduction rate determines the amount of stable acceptors which are generated per received dose $\varphi_{\text{eq}} \cdot t$.

- **Beneficial annealing:** Beneficial annealing is caused by acceptor type defects introduced during irradiation which are not stable in time. These defects start to decay to electrically inactive states at temperatures above $0\,^{\circ}\text{C}$ on a relatively short time scale of a few weeks. This damage type is called beneficial as the decaying acceptors cause the effective doping concentration and thereby the depletion voltage to decrease. The effect is characterised by an introduction rate as well as the temperature dependence of the decay. It can be described by a differential equation with a source term[2]

$$\dot{N}_{\text{r}}^{\text{a},1}(t,T) = -k_{\text{A}}(T) \cdot N_{\text{r}}^{\text{a},1}(t,T) + g_{\text{A}} \cdot \varphi_{\text{eq}} \qquad (3.21)$$

with an introduction rate of reversible (r) acceptors g_{A} and a decay constant k_{A} of currently present reversible acceptors $N_{\text{r}}^{\text{a},1}$. For a constant temperature ($k_{\text{A}}(T) = k_{\text{A}}$), this differential equation can be solved via variation of parameters for the boundary condition $N(t = 0) = N_0^{\text{a},1}$ to be

$$N_{\text{r}}^{\text{a},1}(t) = \frac{g_{\text{A}}\varphi_{\text{eq}}}{k_{\text{A}}}\left(1 - e^{-k_{\text{A}}t}\right) + N_0^{\text{a},1} \cdot e^{-k_{\text{A}}t}. \qquad (3.22)$$

- **Reverse annealing:** The second reversible radiation damage is called reverse annealing. Phenomenologically, reverse annealing describes a process which increases the depletion voltage (after SCSI), which means either an increase in acceptors or a decrease of donors. On a microscopic level, radiation-induced neutral defects in the atomic lattice form acceptor-like energy levels during annealing. The process is slower and occurs at higher temperatures compared to the beneficial annealing. The reverse annealing at room temperature (RT) can

---

[2]The dot above the N indicates the partial derivative: $\partial_t = \frac{\partial}{\partial t}$.

last several months and increases the depletion voltage beyond the value before annealing. The amount of reversible non-active defects N_r^{nd} has to be computed first to describe the interesting quantity of active states. A differential equation of the same form as before describes the generation of non-active defects

$$\dot{N}_r^{nd}(t,T) = -k_Y(T) \cdot N_r^{nd}(t,T) + g_Y \cdot \varphi_{eq} \tag{3.23}$$

with an introduction rate g_Y and a decay constant of $k_Y(T)$ for this kind of defect. The solution analogous to Equation 3.22 is

$$N_r^{nd}(t,T) = \frac{g_Y \varphi_{eq}}{k_Y} \left(1 - e^{-k_Y t}\right) + N_0 \cdot e^{-k_Y t}. \tag{3.24}$$

The relevant number of active defects can be calculated as the number of decaying non-active defects following another differential equation

$$\dot{N}_r^{a,2}(t,T) = k_Y(T) \cdot N_r^{nd}(t,T). \tag{3.25}$$

The solution of this equation is

$$N_r^{a,2}(t,T) = \frac{g_Y \varphi_{eq}}{k_Y} \left(k_Y t + e^{-k_Y t} - 1\right) + N_0^{nd} \left(1 - e^{-k_Y t}\right) \tag{3.26}$$

where N_0^{nd} is the concentration of neutral defects at the beginning of the individual time interval.

For all equations, the temperature dependence is expressed through the variation of the decay constant k_X following an Arrhenius relation

$$k_X(T) = k_X^0 \cdot e^{-E_X^a / k_B T} \tag{3.27}$$

where E_X^a is the activation energy for the specific process. The used values of E_X^a and k_X^0 are shown in Table 3.2 whereas the values of the g_X exhibit a large material and irradiation specific dependence. Hence, the introduction rates have to be fitted to the individual situation.

Table 3.2: Values of E_X^a and k_X^0 used for the simulation of the depletion voltage [54].

k_A^0 [s^{-1}]	k_Y^0 [s^{-1}]	E_A^a [eV]	E_Y^a [eV]
$2.4 \cdot 10^{13}$	$7.4 \cdot 10^{14}$	1.09	1.325

The last sensor property affected by bulk damages is the charge collection efficiency (CCE). Radiation-induced defects can act as recombination centers, similar to the generation centers responsible for the increased leakage current. Those defects can capture free charge carriers (*trapping*) for a time in the order of several µs. Consequently, this charge is lost given the typical integration times of pixel detector electronics of tens of ns. Although, the trapped charge has already contributed to the induced signal on the electrode according to Equation 3.11, the typical weighting fields of small-pitch pixel detectors restrict most of the current generation to the last few µm. Therefore,

charge trapping significantly reduces the CCE after irradiation. The collected charge $Q_{\mathrm{e,h}}$ can be calculated as a function of the effective carrier lifetimes $\tau_{\mathrm{e,h}}^{\mathrm{eff}}$ according to [50]

$$Q_{\mathrm{e,h}}(t) = Q_{\mathrm{e,h}}^0 \cdot \exp\left(-\frac{t}{\tau_{\mathrm{e,h}}^{\mathrm{eff}}}\right). \tag{3.28}$$

The lifetime can be expressed as a function of the irradiation flux as

$$\tau_{\mathrm{e,h}}^{\mathrm{eff}}(\Phi_{\mathrm{eq}}) = \tau_{\mathrm{e,h}}^{\mathrm{eff}}(0) \cdot \left(\frac{1}{1 + \tau_{\mathrm{e,h}}^{\mathrm{eff}}(0)\beta_{\mathrm{e,h}}^{\mathrm{trap}}\Phi_{\mathrm{eq}}}\right) \tag{3.29}$$

with $\beta_{\mathrm{e,h}}^{\mathrm{trap}}$ a trapping related damage parameter.

3.2.2 Surface damage

The second category of radiation damage are surface effects. Ionising radiation such as photons or charged particles release electron-hole pairs in the insulating silicon oxide (SiO_2) layer grown on top of the bulk. Most of the created electron-hole pairs recombine immediately due to the generally low mobility in the silicon oxide. However, the electrons have a sufficient mobility to escape the insulating layer if they do not recombine immediately. Holes, on the contrary, can only be transported via shallow traps in the oxide, and are captured by deep traps forming a fixed positive space charge at the interface between silicon and silicon oxide. The positive charge attracts electrons which accumulate in the silicon close to the surface. This layer of electrons would short all pixel implants in sensors based on n-type pixels. Therefore, a p-spray layer with a low dose p-implant is used here to compensate the accumulating electrons to keep the inter-pixel isolation at sufficient levels. As the oxide charge increases with irradiation, the p-spray layer is partially compensated and the electrical field between p-spray and pixel implant is decreased [56]. Consequently, the breakdown voltage of silicon sensors using a p-spray inter-pixel isolation technique increases significantly after irradiation. On the other hand, a p-spray dose high enough to maintain isolation also after radiation must be implemented.

3.3 Layout and production of silicon sensors

This section describes the production process of typical semiconductor pixel sensors as well as the specific technology implemented for the structures that are investigated within this thesis. This technology was developed by the semiconductor laboratory of the Max-Planck society (HLL) and has been transferred to commercial vendors such as CiS Technology[3]. Also, the different semiconductor productions used for the results in this thesis are presented.

---

[3]https://www.cismst.org

3.3.1   Sensor manufacturing and composition

Sensor manufacturing

The sensor manufacturing starts with clean wafers cut from high resistive float-zone (FZ) p-type silicon ingots. At first, the wafer surface is polished to obtain a homogeneous and reproducible starting point. The wafers are processed in an oven at temperatures beyond 1000 °C to grow a SiO_2 layer on top of the wafer. A photoresist is applied on top of the oxide and is patterned using a photolithography step with ultra-violet (UV) light and either a quartz mask or a laser projector. The illuminated parts of the photoresist are removed in a subsequent step to expose the areas of the oxide which are supposed to be etched away. After etching of most of the oxide in the patterned areas, the implantation of the pixel implants is done with ions of the respective dopant. The patterning of the pixel implants is accomplished by the shielding properties of the oxide that is not etched away as well as the photoresist on it.

Subsequently, layers of nitride (Si_3N_4) and low temperature oxide (LTO) are deposited on the wafers. Afterwards, a via is etched through the stack of thinned oxide, nitride and LTO using another photolithographic step. The wafer is then patterned again using photoresist to form the layer of aluminium previously sputtered on the wafer. The aluminium is left only in the desired places after striping away the photoresist. A layer of benzocyclobutene (BCB) is deposited on the wafer in the next step. Again, photolithography is used to pattern the BCB and etch vias from the surface to the aluminium pads.

A layer of TiW is deposited on the wafer serving as a diffusion barrier for the copper of the last processing step, the under-bump metallisation (UBM). On top of the TiW, a copper layer is sputtered, acting as the seed layer for the UBM subsequently being deposited through electroplating. On each pixel, the UBM is in electrical contact with the aluminium pad through the via in the BCB. The UBM serves as landing pad for the bump bonds placed on the readout chip. The full stack of layers is illustrated in Figure 3.1.

Figure 3.2 shows a three dimensional scan using an optical profiler of the surface of the sensor after the production. The sensor is from the SOI-3 HLL production (see Section 3.3.2 for more details about the different productions) and similar to those of the SOI-4 and SOI-5 runs. The copper UBM pillars stick out of the relatively flat aluminium pad on top of the pixel which causes a pattern in the BCB.

Additional sensor components

Various additional components have to be considered during the design of pixel sensors which are not required for the bare pixel cell but are essential for proper operation. The general layout of sensors with the $50 \times 50\,\mu m^2$ and $25 \times 100\,\mu m^2$ pixel cell geometry is presented in Figure 3.3. A selection of additional components is given in the following.

- **Inter-pixel isolation** An inter-pixel isolation technique has to be used with n-type pixels to ensure proper isolation of the individual pixels, especially after irradiation. Theoretically, a junction is established from the n-type pixel implant to the p-type bulk material surrounding
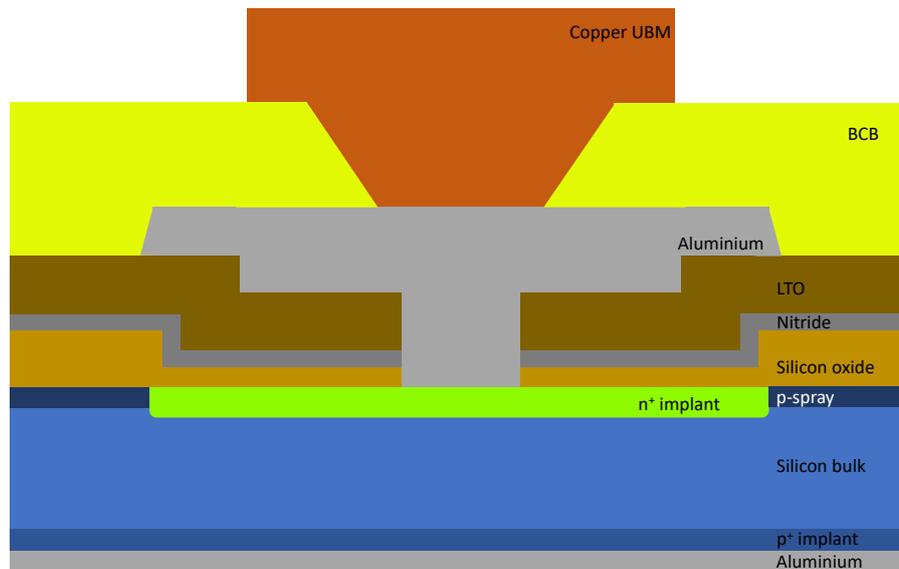
Figure 3.1: Schematic overview of the various layers and elements in a silicon sensor. The individual thicknesses are not to scale.
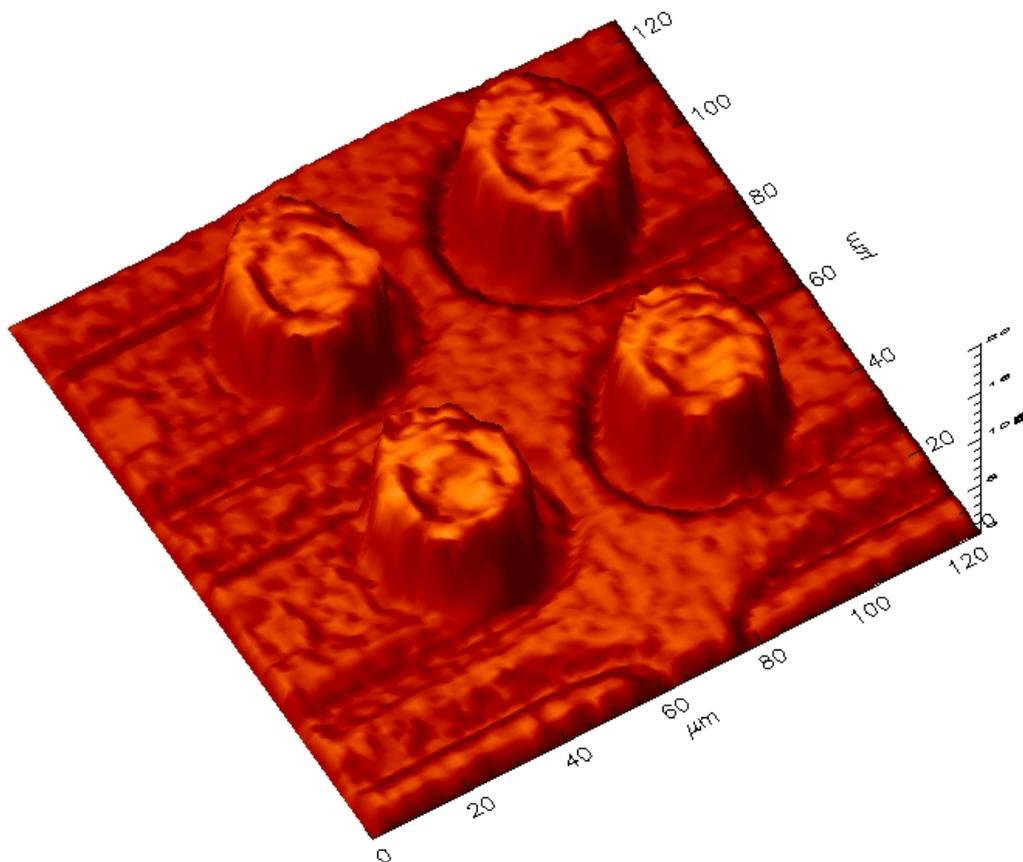

Figure 3.2: 3D scan of a quad sensor of the SOI-3 HLL production zooming onto the border region of four pixels. The copper UBM of each pixel is clearly visible as a pillar. Courtesy of HLL [57].
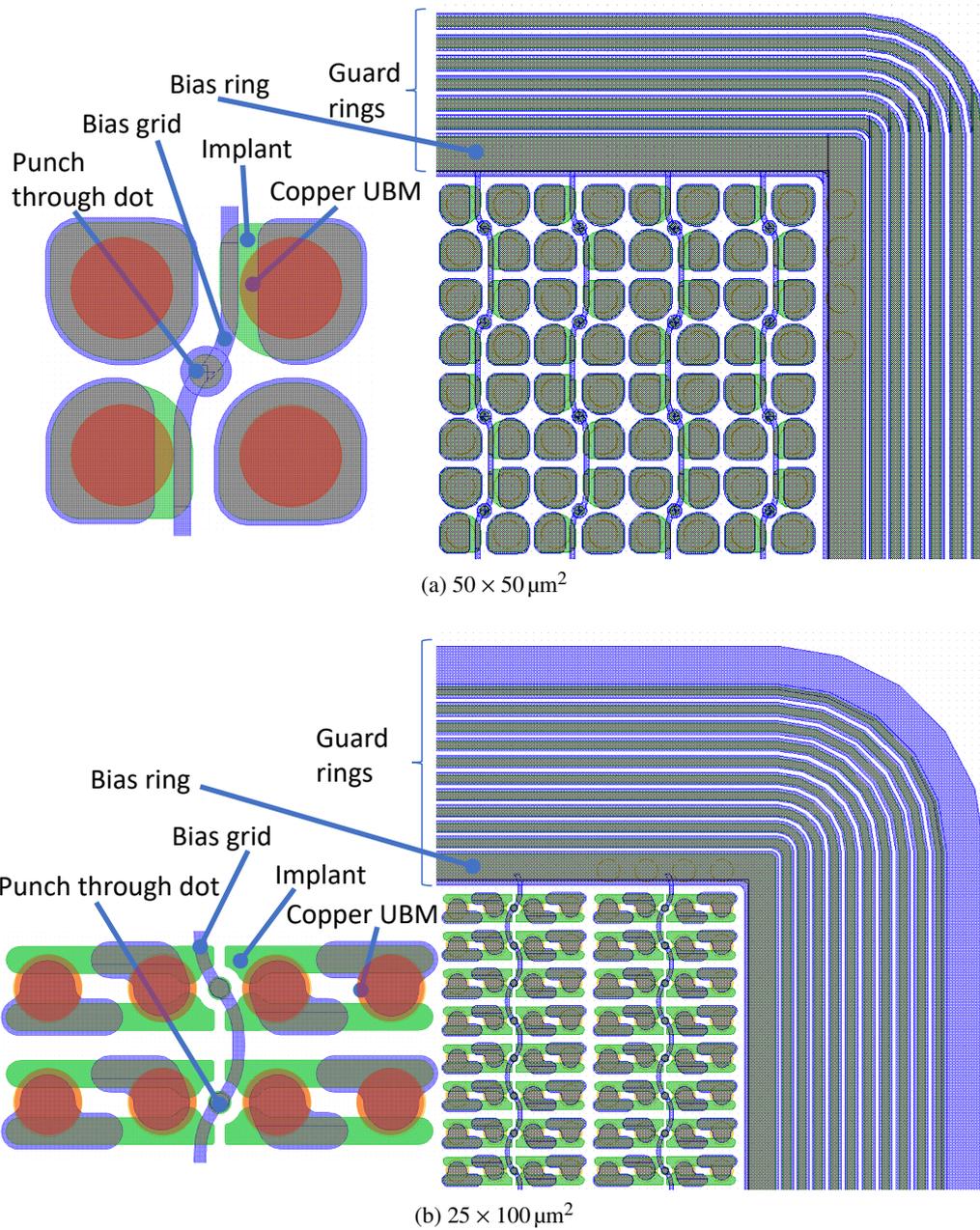
(a) $50 \times 50\,\mu\text{m}^2$

(b) $25 \times 100\,\mu\text{m}^2$

Figure 3.3: Layout of sensors with the $50 \times 50\,\mu\text{m}^2$ (a) and $25 \times 100\,\mu\text{m}^2$ (b) pixel cell geometries. The pixel cell is shown in detail for each geometry as well as an overview of the layout. The guard rings and biasing structures are indicated.

the implant, thus ensuring isolation to neighbouring pixels if no surface effects are considered. However, the production related fixed positive oxide charge at the interface already compensates the very weak p-type doping of the high resistive sensor material. The oxide charge attracts a free electron cloud that shorts the n-implants together. Consequently, the isolation between pixels breaks down. This can be compensated by a homogeneous layer of light p doping called p-spray with a concentration in the order of some $10^{16}\,\mathrm{cm}^{-3}$. The p-spray sufficiently increases the p doping-concentration in the critical area close to the surface, to compensate for the accumulating electrons and ensure inter-pixel isolation. The doping concentration of the p-spray layer has to be chosen high enough to ensure the isolation properties also after irradiation when the oxide charge is increased due to ionising radiation (see Section 3.2).

- **Guard rings** The electrical potential of the sensor backside extends to the frontside through the conductive cutting-edge of the sensor. The guard rings (GRs) are a series of ring implants around the matrix of pixels across which the electrical potential from the cutting-edge is smoothly dropping towards the pixel matrix [58]. Simultaneously, to save space, the innermost GR acts as bias ring (BR). This ensures ground potential next to the pixels at the edges of the matrix. The GRs are required to limit the depleted area towards the edges. The high electrical field within the depletion zone would otherwise cause a huge current if it would extend up to the conductive cutting-edge. The number, the shape and the aluminium overhang of the GRs as well as the distance between the individual GRs is crucial for the breakdown behaviour of the sensor. While numerous GRs ensure a high level of safety and a high breakdown voltage, the inactive area around the pixel matrix grows and reduces both the fraction of active area and compactness of the sensor. Yet, even with numerous GRs, the high voltage of the backside still extends to the front-side through the cutting-edge, such that the maximum bias voltage is limited by the risk of sparks from the cutting-edge to the grounded readout chip only around $30\,\mu\mathrm{m}$ apart. Possibilities to mitigate this problem are discussed in the following.

- **Spark protection** The BCB layer deposited on top of the sensor protects the readout chip from the high voltage potential of the sensor. The advantage of this approach is that the BCB can be applied already during the sensor manufacturing in relatively thick layers of typically around $3\,\mu\mathrm{m}$. With its high dielectric strength, the BCB results in a more reliable protection against sparks with respect to the silicon oxide. However, a remaining problem is that the cutting area and the vertical sides are not covered by BCB. For a complete protection, a layer of parylene can be deposited on the whole structure after module assembly. The high mobility of the gaseous parylene during its application ensures coverage of all surfaces even between sensor and readout chip while the high dielectric strength guarantees high voltage protection beyond $1000\,\mathrm{V}$. The disadvantage of using parylene is that the deposition has to be done in a separate step after module assembly and that all contacts to the module have to be reliably protected not to be insulated as well.

- **Biasing structure** Although the production of semiconductor sensors for HEP applications has improved significantly over the past decades, quality control is essential for process optimisations and the selection of good sensors for module assembly, hence, lowering the costs. A sensor with a simple pixelated surface as described before cannot be tested before interconnection to a readout chip. However, it is important to enable tests on bare-sensor and wafer level already since the hybridisation is the most expensive step in module production. For that purpose, all pixels have to be biased at the same time. On-sensor techniques have to be employed for testing because the connection to several ten thousand pixels at once using mechanical tools, such as probe cards, is impractical. In this approach, all pixels are shorted with a high resistive contact such that low frequency signals such as the voltage sweep of a voltage-current curve are transmitted without problems, while high frequency signals produced by impinging particles are not able to take this path. Two approaches have proven to enable successful testing. In both approaches, a BR on top of the innermost GR provides a potential around the pixel matrix. From there, a bias grid (BG) of aluminium lines distributes the signal through the matrix. The two approaches differ in the way the pixels are connected to the BG. In the case of bias resistors [59], a high ohmic material, typically polysilicon, is used in a curly layout on top of the pixel to achieve a high ohmic connection to it. The disadvantage of this approach is that the resistance scales with the length of the resistor limited by the pixel size. The approach used for all investigated structures in this thesis uses the punch-through effect [60]. Small additional implants are implemented in between pixels (see Figure 3.3a). The resistivity of the connection between the two n^+ implants is given by the potential barrier in the form of the p-spray and the depleted p-type bulk. The height of the potential barrier is given by the distance between the implants, the potential of the backside and, to a smaller extend, by the doping concentration of the bulk. A typical resistance of $O(10)M\Omega$ can be achieved. One disadvantage of both techniques is particularly severe after irradiation when the bias resistor or the punch-through (PT) dot attracts charge carriers generated by impinging particles. Consequently, the signal recorded by the readout chip is reduced and inefficiencies occur. The effect is localised in the area of the biasing structure. Mitigation strategies such as routing the BG mostly over underlying pixel implants have shown improvements in comparison with standard routings in between pixels but may not satisfy the most stringent requirements of ITk. Therefore, sensors without biasing structures are tested taking the risk of no testing capability before interconnection.

- **Intermediate metal** A technique to test sensors without biasing structures before interconnection on wafer level is in development together with HLL. It employs the TiW layer deposited as part of the wafer post-processing. The full-surface coverage on the wafer has to be patterned such that individual structures can be tested. The mask leaves the TiW layer within the pixel matrix and removes it elsewhere. In that way, the thin TiW layer connects all pixels and allows testing of individual structures without any further biasing structure already on wafer level. The required subsequent etching step would have to be done in any case and is selective to the TiW such that no additional risk is introduced into the sensor production process. Still, no

testing after removal of the TiW layer is possible which would usually be done after dicing.

3.3.2 Sensor productions

Several semiconductor pixel sensor productions at CiS and HLL are performed by the MPP. Sensors are designed in an n-in-p technology and are in accordance with previously achieved and discussed results [61–64]. The wafers are manufactured in the previously described technology, any particularities of the individual productions will be mentioned in this section.

CiS-4

The CiS-4 production is carried out by CiS Technology on high resistive ($\approx 20\,\mathrm{k\Omega\,cm}$) FZ p-type 100 mm wafers. This size allows to fit two quad sensors and four single chip sensors compatible with the FE-I4 readout chip. The wafers have a thickness of 525 μm and are thinned to the target thickness of 100 μm and 150 μm. The thinning technology employs an anisotropic wet etching step using potassium hydroxide (KOH) [65]. The etching leaves behind frames around the individual structures that are cut away during dicing. An electroless Ni UBM is used and no BCB protection layer is applied to the sensors during the production. The quad sensors of this production are used for the investigation of quad module assembly in Chapter 5 and no sensor related studies are presented in this thesis. The design is shown in Figure 3.4. The inter-chip region of the quad sensors in y-direction is filled with further pixels that are connected by aluminium lines to pixels within the active matrix of each readout chip. Always two pixels are connected together in that way. The additional pixels without corresponding readout channel are called ganged pixels. The inter-chip region in x-direction is covered by long pixels that extend 400 μm instead of the usual 250 μm. Further details on this production and all sensor related studies can be found in References [61,62].

SOI-3 and SOI-4

The SOI-3 and SOI-4 sensor productions are performed at HLL. As the name suggests, the wafer technology is silicon on insulator (SOI) making them a composite of two permanently connected silicon wafers separated by a thick oxide layer. With a thickness of 400 μm, the support wafers are substantially thicker then the 100 μm and 150 μm thin device wafers. Both wafers are originally 675 μm thick and are thinned to the desired thickness using chemical mechanical polishing (CMP) [66]. After frontside processing, the attached support wafer is removed in a combination of mechanical grinding and chemical etching. Once the wafers are thinned to a safe thickness using fine grinding, the remaining material of the support wafer is removed employing the deep reactive ion etching (DRIE) technique [67]. The wafers have a diameter of 150 mm and are of high resistive ($\approx 18\,\mathrm{k\Omega\,cm}$) p-type FZ silicon. The differences of the two productions are summarised in the following list:

- **SOI-3** The SOI-3 production is focused on FE-I4 compatible quad sensors. It also includes several single chip FE-I4 compatible sensors using different edge designs. The number of GRs is varied from 11 (edge distance, $d_e = 450\,\mu\mathrm{m}$) and 8 ($d_e = 300\,\mu\mathrm{m}$) to 4 ($d_e = 90\,\mu\mathrm{m}$).
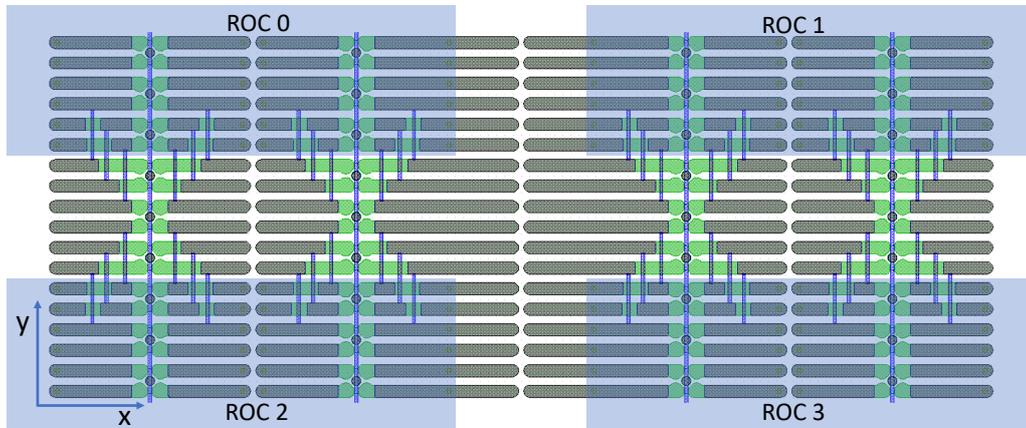
Figure 3.4:   Design of the inter-chip region in the CIS-4 sensor production. Ganged and long pixels are employed to implement the technically required gap between the four readout chips bump bonded to the quad sensor.

The additional RD53A compatible sensors have 8 GRs. The SOI-3 production is the first to test different configurations of the PT potential by not opening the BCB on top of the BR for some of the structures. Consequently, the BR and thus the PT dots stay floating in those structures, even in case the corresponding bump bonds on the readout chip are at ground potential. The post-processing, including BCB deposition, copper UBM and backside thinning, is performed at Fraunhofer IZM. The oxide between support and device wafer is removed in an etching step employing a mask to provide vias between backside aluminium and p$^+$ contact.

- **SOI-4** The structures implemented in the SOI-4 production are similar to those of the SOI-3 production. Again, the majority of the surface is covered by four quad sensors compatible with FE-I4 readout chips. The focus for the remaining structures is shifted to RD53A compatible sensors. Different GR widths and PT potentials are investigated. A new feature of the SOI-4 production is the split design in which the BR potential is decoupled from the potential of the innermost GR. The BR is divided into four straight sections, two of which connect to the BG while the other two receive the potential from the readout chip and distribute it to the underlying implant of the innermost GR. The part of the BR connecting the BG is still located on top of the innermost GR but is isolated from it by the nitride and LTO layers. In the split design, the BG is connected to two straights of the BR square. This part is referred to as BR for the split design. In this way, it is defined by its original purpose - distributing a homogeneous potential amongst the PT dots through the BG.
Differently than for the SOI-3 case, the BCB and UBM deposition is performed at HLL. Another difference is the omission of a mask for the oxide removal during thinning. This is

possible since the etching process is selective to oxide and thus stops as soon as the oxide is removed. The naming scheme of SOI-4 sensors can be found in Table 3.3. A *D* is added in front of the name if the structure is a double chip sensor. Multiple instances of the same type are indicated by trailing letter A,B,C,... .

SOI-5

The SOI-5 production is entirely focused on RD53A compatible sensors, including several variants of single, double and one four chip or quad sensor. The purpose is twofold: firstly completing the research and development phase by investigating various pixel implant widths, pixel metal overhangs and guard ring widths as well as PT potentials. Secondly, preparation for the mass production by means of finalising the production flow and insourcing also the last part of the post-production, the thinning and dicing of the wafers. The naming scheme for SOI-5 sensors can be found in Table 3.3, with a supplement for sensors without PT in Table 3.4. The naming of double chip sensors can be found in Table 3.5. Multiple instances of the same type are indicated by trailing letters.

Table 3.3: Naming scheme of SOI-4 and SOI-5 sensors. The left entry per column is always the employed identifier with the explanation adjacent to the right. Entries in brackets only occur in the SOI-5 production.

Pixel cell		GR width		Grounding		PT	
1	$25 \times 100 \, \mu m^2$	2	$250 \, \mu m$	G	grounded	N	no PT
2	$50 \times 250 \, \mu m^2$	(3	$350 \, \mu m$)	F	floating		
5	$50 \times 50 \, \mu m^2$	4	$450 \, \mu m$	S	split	P	with PT

Table 3.4: Naming scheme of SOI-5 sensors without PT. The left entry per column is always the employed identifier with the explanation adjacent to the right. The resulting ID is inserted between the PT ID and the identification letter.

Implant width		Metal overhang	
2	$20 \, \mu m$		
3	$30 \, \mu m$	2	$2.5 \, \mu m$
4	$38 \, \mu m$	3	$3.5 \, \mu m$
8	$40 \, \mu m$	4	$4 \, \mu m$
M	multiple widths		

Table 3.5: Naming scheme of SOI-5 double chip sensors. The left entry per column is always the employed identifier with the explanation adjacent to the right.

Device type		Pixel cell		PT	
D	double chip	1	$25 \times 100 \, \mu m^2$	N	no PT
		5	$50 \times 50 \, \mu m^2$	P	with PT

As mentioned before, the thinning of SOI-5 wafers is performed at HLL, in contrast to all previous productions. Before thinning, a ring of BCB of a few mm width is applied on the circumference of

<table>
<tr><td>(a)</td><td>(b)</td><td>(c)</td></tr>
</table>
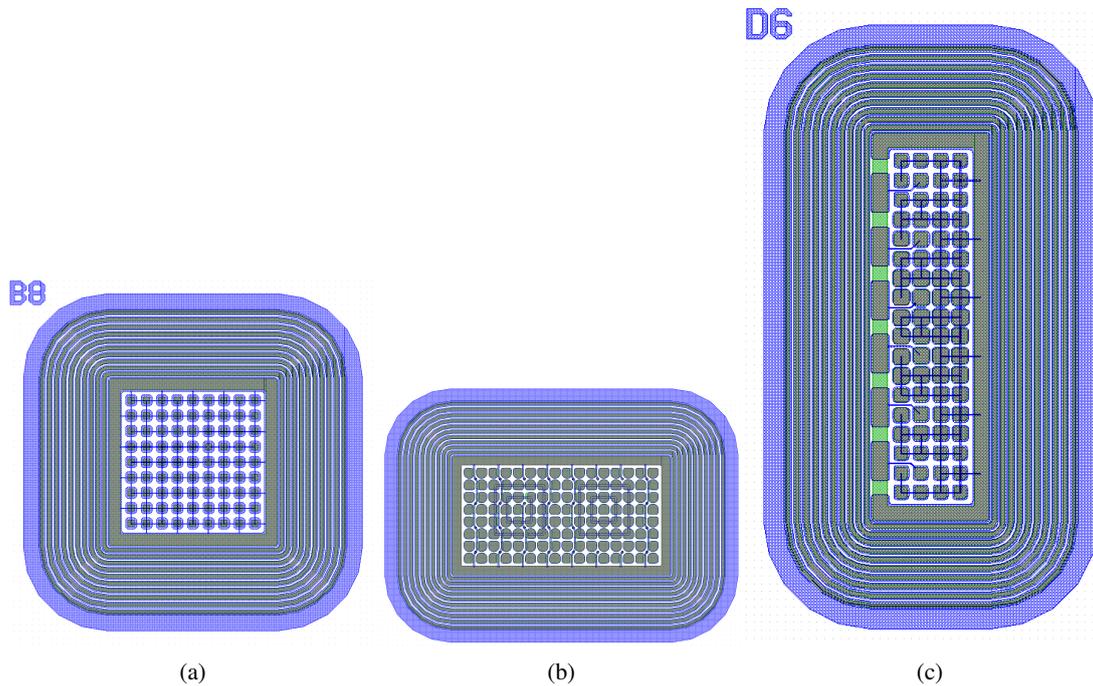
Figure 3.5:   Test structures implemented in the SOI-5 production. The breakdown behaviour of different implant widths and metal overhangs is tested with the structure shown in (a). The structure in (b) offers the possibility to measure the inter-pixel capacitance for the PT and no-PT case. Finally, the structure in (c) can be used in edge TCT measurements to extract the CCE of different implant widths.

the wafer backside. This ring is used to seal the wafer frontside from the backside using a rubber lip. The backside is then exposed to an anisotropic wet etchant to etch away the entire support wafer. The frame at the circumference around the wafer is required for safe handling during the remaining production steps. It is eventually removed during the regular dicing procedure to extract individual structures from the wafer.

The SOI-5 production is also used to enable a comparison of TCAD simulation results (see Chapter 7) to data. Various test structures are implemented for that purpose. The first group consists of test structures to examine the dependence of the breakdown voltage on the pixel implant width. In each structure, 9×9 pixels are contained in a safe GR complex, as exemplarily shown in Figure 3.5a. The implant width is set to be $10\,\mu m$ (A), $20\,\mu m$ (B) or $30\,\mu m$ (C). The metal overhang is varied in the same structures, ranging from $0\,\mu m$ to $8\,\mu m$. In Figure 3.5a, the character defining the implant width is indicated in combination with the metal overhang on each structure. All implants are shorted by thin aluminium lines to provide ground potential to all pixels with a single probe.

Further test structures are designed to enable measurements of the pixel capacitance as function of the implant width. The design is depicted in Figure 3.5b. During the measurement, the central pixel is connected to the first probe, the capacitance is measured against the direct neighbours using a second probe and the remaining pixels need to be grounded with a third probe through the BR.

The last category of test structures is meant to investigate the influence of the implant width on the CCE using edge transient current technique (TCT) (the interested reader is referred to

References [61, 68–71] for further information about edge TCT). One of the structures is shown in Figure 3.5c. Different implant widths are implemented in a single structure since the comparison of collected charge between different structures is delicate due to the difficulty of achieving an absolute charge calibration in TCT measurements. One structure consists of four columns and six groups of three rows each. All pixels in one group feature the same implant width. The width is varied from group to group in the pattern I-II-III-III-II-I such that relative changes across the sensor can be measured and corrected for. For instance, the structure in Figure 3.5c features implant widths ranging from 30 μm to 40 μm, the implants of the first and last group are 30 μm wide, the ones of the second and fifth group are 35 μm wide and the ones of the third and fourth group are 40 μm wide. The central pixel in each group is connected to a larger pad on the innermost GR for easier connection to the periphery using a wire-bond. The combinations of implant widths implemented are summarised in Table 3.6. A letter is assigned to each number depending if the BCB is opened only on the pixels that need to be connected (A and C) or if the full matrix is opened (B and D).

Table 3.6: Naming scheme of the different test structures for measurements using a TCT setup. The implant widths correspond to the widths of the implants of the first/second/third group of pixels, the first at the outside, the third in the center of the test structure and the second in between.

Structure number	Implant widths [μm]
0	28/36/39
1	25/35/40
2	10/15/20
3	15/20/25
4	20/30/40
5	25/30/35
6	30/35/40

3.4 Module concept

The analog signal generated in the sensor as described in Section 3.1 needs to be extracted from the pixel. Also further processing of the analog signal as amplification, discrimination, signal storage and triggered readout through high bandwidth data links are requirements of a pixel detector. For those purposes application specific readout chips have been developed. The actual implementation splits all pixel detectors into two categories, monolithic and hybrid concepts. In the monolithic approach, typically a commercial complementary metal-oxide semiconductor (CMOS) process is used to create chips which incorporate the necessary analog and digital logic circuitry for signal processing. Simultaneously, the chips use the silicon they are built on as sensor volume. Therefore, no hybridisation is required promising potential cost reductions and manufacturing simplifications. Since only hybrid detectors are covered in this thesis, monolithic detectors will not be treated in more detail but the interested reader is referred to References [72–75] for further information on monolithic detectors and their application in high energy physics.

(a) Top view

(b) Side view
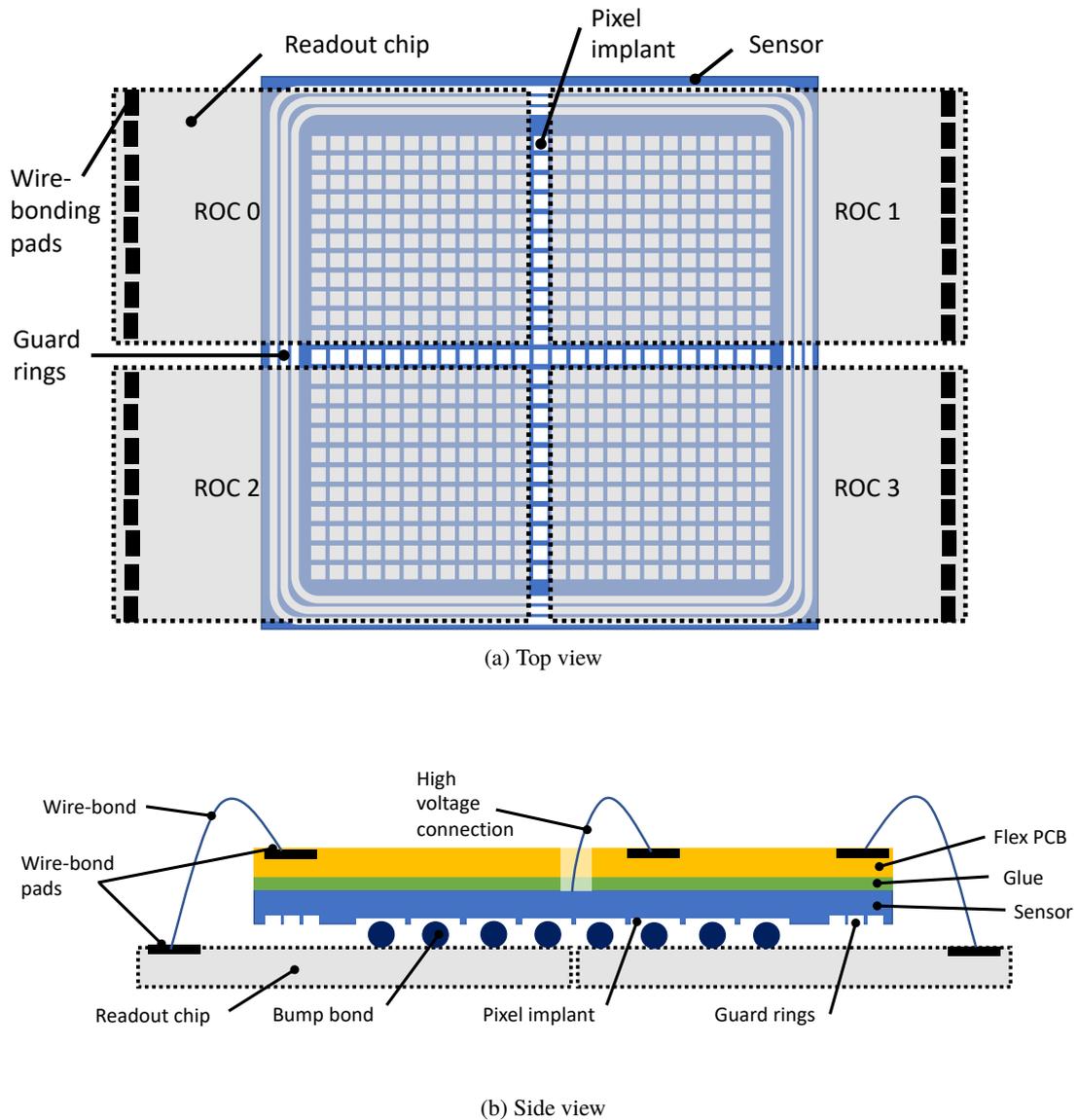
Figure 3.6: The baseline ITk module concept for the quad chip configuration employing thin planar n-in-p sensors. Top view in (a) and cross-section in (b).

Instead, hybrid detectors use separate devices for sensor volume and logic circuitry. The big advantage of this approach is that readout chip and sensor can be optimised individually without the necessity of making compromises regarding bulk material or process technology. Therefore, at present hybrid detectors are the most radiation hard and fastest pixel detectors.

Impinging particles create electrical signals in the sensor which drift to the individual pixels. Each sensor pixel is connected through a bump-bond ball to the readout chip in a process called *flip-chipping*. On the readout chip side, the signal is amplified and further processed as described in Section 3.5.

For the ITk, thin n-in-p planar sensors with a thickness of $100 \, \mu m$ and $150 \, \mu m$ are foreseen to be used in all but the innermost layer. The majority of modules will be built as quad modules (see
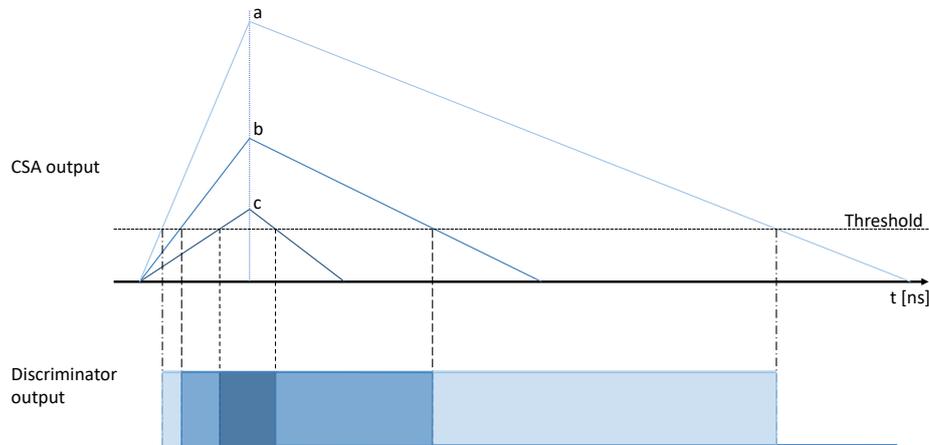
Figure 3.7: Influence of three different feedback current settings on the CSA and discriminator output.

Chapter 5) where four readout chips are attached to a single sensor. For signal routing and voltage distribution, a flexible printed circuit board (PCB), called the flex, is attached to the bare module built from sensor and readout chip. Wire-bonds are used to route the data as well as supply voltages to the readout chip. All signals are bundled in a single connector used to couple the module to the off-detector components. A visualisation of top view and cross-section of the proposed module concept is given in Figure 3.6.

3.5 Read-out chips

First, a general description of the typical concept of readout chips for high energy physics applications is given based on the examples of the FE-I4 and RD53A chips. Afterwards, individual aspects of the two chips are discussed.

The first step in both readout chips is to amplify the typically very small and short signal (about $1\,\mu\text{A}$ for a few ns). A charge sensitive amplifier (CSA) is used which sends out a signal relative to the charge accumulated on the input. The charge on the input has to be cleared to be able to recognise successive signals close in time. For that purpose, a feedback current is used to discharge the input, clearing both actual signal and leakage current. The feedback current can be adjusted on a per-chip and on a per-pixel (only FE-I4) level to adjust the signal duration for a given input charge. The impact of different feedback currents is visualised in Figure 3.7. The time during which the signals stay over the threshold shortens for increasing feedback current from case a to case c.

The output of the CSA is fed into a discriminator to distinguish signal from noise. Again, per-chip and per-pixel trimming registers are available to homogenise the response. The trimming registers change the voltage of the discriminator to which incoming signals are compared to. The threshold $Q_{\text{threshold}}$ is defined to be the charge that is registered in $50\,\%$ of the cases. The probability p_{hit} to

register a hit of charge Q depends on the threshold as well as the average noise level σ_{noise} as

$$p_{\text{hit}}(Q) = 0.5 \cdot \text{erfc}\left(\frac{Q_{\text{threshold}} - Q}{\sqrt{2} \cdot \sigma_{\text{noise}}}\right) \tag{3.30}$$

where erfc() is the complementary error function. The output of the discriminator is directed into the digital circuitry which first of all counts the number of clock cycles (40 MHz) during which the threshold is exceeded. The time over threshold (ToT), together with the time of the first of these clock cycles is stored in the memory of the pixel cell. Further digital processing of data is performed on the chip, as for example a timewalk correction. Timewalk denoted the fact that smaller signals are recorded later, similar to the effect displayed in Figure 3.7 where the registration of the hit shifts towards later times with increasing feedback current.

The trimming or tuning of both threshold and ToT response is necessary since production variations cause different behaviours of pixels of the same chip, and even more differences between chips from different wafers. The global and in-pixel trimming of the readout chips ensures an identical response of different pixels to identical signals. However, the trimming requires a reliable injection mechanism built into each pixel cell. This is done by one (or more) capacitors which are charged with a known voltage defined by an injection register called VCAL. The injected charge Q_{inj} is given by

$$Q_{\text{inj}} = C_{\text{total}} \cdot \left(V_{\text{slope}} \cdot \text{VCAL} + V_{\text{offset}}\right) \tag{3.31}$$

with C_{total} the total capcitance, V_{slope} and V_{offset} being the parameters of a linear fit to the measured VCAL-to-voltage relation. The charge injection in the FE-I4 is done with a pulse relative to a common ground potential whereas the RD53A uses a differential voltage injection between two definable signals. Hence, the injected charge in the case of the RD53A is quoted in terms of ΔVCAL which is the difference between the high and low VCAL parameters.

3.5.1 FE-I4

The FE-I4 chip was developed for the ATLAS pixel detector upgrade IBL. This readout chip is based on the 130 nm IBM CMOS process. Each of the 336×80 pixels has a dimension of $50 \times 250\,\mu\text{m}^2$ resulting in an active area of $16.8 \times 20.0\,\text{mm}^2$ [26]. Particular emphasis is put on the radiation hardness, achieved with encapsulated transistors and avoiding minimal size transistors. The result is radiation hardness at least up to $5 \times 10^{15}\,\text{n}_{\text{eq}}/\text{cm}^2$ [76]. The bandwidth of 160 Mb/s is four times larger with respect to the previous generation [77]. The chip features a low dropout regulator (LDO) which can be operated in constant voltage or constant current (shunt) mode. The chip generates the required operating voltages of 1.2 V (digital) and 1.5 V (analog) from a supplied 2 V voltage or a 2 A current.

One particularity of the FE-I4 readout chip is the HitDiscCnfg register which determines the mapping of ToT code in the output data to the actual duration the threshold is exceeded. The impact of different HitDiscCnfg settings is summarised in Table 3.7. For all measurements in this thesis, a HitDiscCnfg value of 0 is used to resolve also small signals.

Table 3.7: The 4-bit ToT code hex values and the corresponding time above threshold in units of the clock cycles of 25 ns [78]. All values below the dividing line are considered big hits, ToT codes above the dividing line are only read out if they are adjacent to a big hit.

	HitDiscCnfg		
True ToT [25 ns]	0	1	2
below thresh.	F	F	F
1	0	E	E
2	1	0	E
3	2	1	0
4	3	2	1
5	4	3	2
6	5	4	3
7	6	5	4
8	7	6	5
9	8	7	6
10	9	8	7
11	A	9	8
12	B	A	9
13	C	B	A
14	D	C	B
15	D	D	C
≥16	D	D	D

3.5.2 RD53A

A new readout chip is required to cope with the challenges of the HL-LHC. The RD53 collaboration [42] has developed the RD53A readout chip as a prototype to test and benchmark the proposed technologies. The collaboration consists of members of the ATLAS and CMS experiments. Consequently, the developed basic blocks of the RD53A will be used for the final chips of both experiments. Some details of the final chips will be tailored to the requirements of the different experiments, in particular in terms of size and trigger scheme compatibility.

The RD53A prototype chip is developed in the 65 nm CMOS process at Taiwan Semiconductor Manufacturing Company (TSMC) which will be employed also for the final chip. A radiation tolerance of 500 MRad is ensured through intrinsic advantages of the smaller node size compared to previous generations. Stable low threshold operation will be essential to cope with the small signals resulting from the radiation damage on the sensor side. Therefore, a stable threshold of 800 e is foreseen for RD53A.

The cell size is reduced to $50 \times 50\,\mu m^2$ to keep the occupancy at the present levels. The matrix of 400×192 pixels spans a total of $20 \times 9.6\,mm^2$ of active area. There is an end-of-column region at the bottom of the chip housing common logic and communication for the whole chip as well as the main wire-bonding pads (see Figure 3.6a, the bottom of the chip corresponds to the left side, the top

corresponds to the right side of ROC 0). Further wire-bonding pads are placed on top of the chip and can be used to access diagnostic signals. The top row of wire-bonding pads cannot be removed preventing gap-less placement of readout chips in the case of quad modules. Those pads will not be present in the production version of the chip and consequently a gap-less placement will be possible. Also, the matrix will be enlarged to contain 400×384 pixels, resulting in $20 \times 19.2 \, \text{mm}^2$ of active surface.

The RD53A chip includes three analog front end (FE) designs identified as linear, differential and synchronous. They are substantially different ground-up designs and not just small variants of one common design. The linear FE employs a linear amplification stage in front of the discriminator, which compares the signal to a threshold voltage. This design is the most conventional one and is close to the approach of the FE-I4 readout chip. Instead, a differential gain stage is implemented in front of the discriminator which sets a threshold by unbalancing the two branches in the differential FE. The most exotic design is implemented in the synchronous FE which uses a baseline auto-zeroing scheme. A periodic acquisition of a baseline instead of per-pixel threshold trimming is required in this approach [79]. Each FE design is implemented for roughly one third of the columns. In the final chip, a transceiver with a bandwidth of $5.12 \, \text{Gb/s}$ will be available for communication to send the immense amount of recorded data to the periphery [80]. Another difference to previous readout chip generations is the focus on serial powering, which will enable the usage of fewer cables for low voltage supply, thereby, drastically reducing the material budget of the detector. Up to eight modules, each consisting of four readout chips, are connected in a serial power chain, in which the four readout chips per module are connected in parallel. A current of 2 A is expected per chip such that the current per chain will be up to 8 A with a total voltage applied across the chain depending on its length.

EXPERIMENTAL TECHNIQUES

4.1 Basic characterisation of pixel detectors

4.1.1 Laboratory measurements

All laboratory measurements at module level are performed in a climate chamber (Vötsch VT7011) to ensure stable temperature and thus comparable results. The climate chamber is flushed with nitrogen to reduce the humidity. Measurements of not-irradiated devices are performed at 20 °C while irradiated devices are cooled to −50 °C. A Rohde & Schwarz HMP 4040 is used for low voltage (LV) power supply. Not-irradiated devices are connected to a Keithley 6517B for high voltage (HV) supply. An Iseg SHQ 224M is used for irradiated devices due to the much higher leakage currents and the higher power that can be supplied by this instrument in comparison to the former one.

A Cascade Microtech Summit 12000 probe-station is employed for measurements at sensor or wafer level. The thermal chuck system uses an ESPEC ETC-200L chiller. The probe-station is cased with a light-tight electromagnetic interference (EMI) shield. Again, a Keithley 6517B or an Iseg SHQ 224M are available for HV power supply. Moreover, an Agilent E4980A precision LCR (inductance, capacitance, resistance) meter is available for capacitance measurements.

4.1.2 Read-out systems

Both used readout chips need to be calibrated (tuned) before usage to ensure optimal performance. The tuning procedure is described alongside the employed data acquisition (DAQ) systems USBPix and BDAQ.

USBPix

The USBPix system [81] was developed in Bonn and Göttingen in various iterations for the readout of FE-I4 readout chips. The MMC3 board (see Figure 4.1a) of the third generation is used for all measurements within this thesis. It consists of a field programmable gate array (FPGA) for the communication with the readout chip and a micro-controller for communication with a connected

(a) MMC3

(b) BDAQ

(c) FE-I4 SCC

(d) RD53A SCC
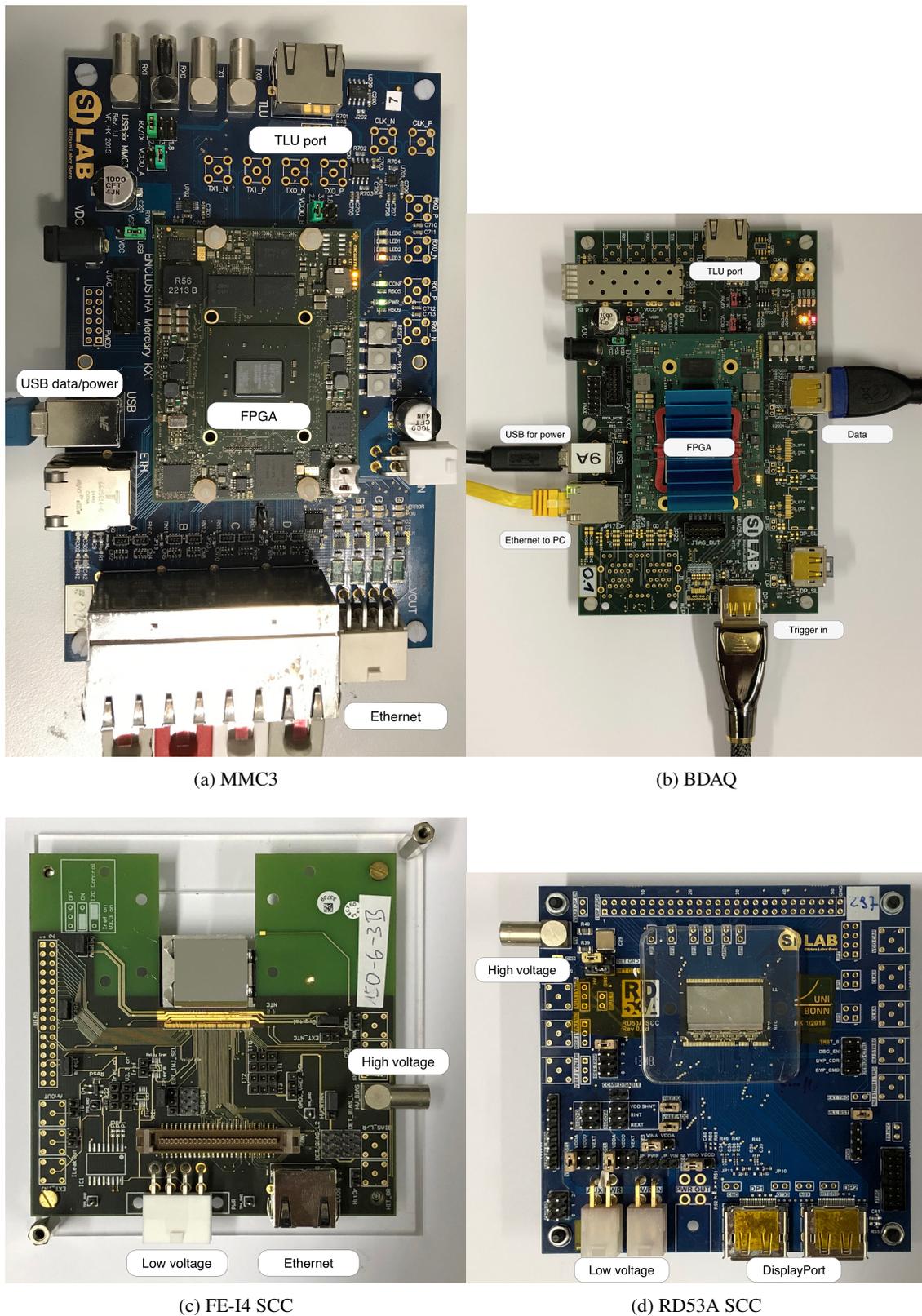
Figure 4.1: Pictures of the MMC3 board for FE-I4 readout (a), the BDAQ board for RD53A readout (b), an FE-I4 module on a single chip card (SSC) (c) and an RD53A module on an SCC (d).

PC. The C++ and ROOT based STcontrol software [82] developed in Bonn and Göttingen is used to initialise and steer all scans and tunings on the MMC3 from a PC connected through USB-3. The USB connection is simultaneously used to provide power to the DAQ board.

A trigger logic unit (TLU) port [83] is available to integrate the MMC3 board into testbeam operation (see Section 4.2.2). Up to eight RJ45 ports allow to connect eight single-chip modules or even more front-ends of multi-chip modules because readout chips can share the same link in this configuration. The following list provides an overview of the most frequently used scans in STcontrol and the typical tuning scheme.

- **Prerequisites** Before the actual tuning starts, a few steps have to be performed. The *Hitdicconf* parameter of the FE-I4 has to be selected according to the needs of the application. The internal injection mechanism has to be calibrated to obtain the VCAL-to-voltage relation (see Section 3.5). Lastly, the global threshold (GDAC) and feedback current (IF) settings have to be chosen according to some initial threshold and ToT scans.

- **Analog scan** This is the most fundamental scan, being employed by most other scans in different configurations and loops. A fixed charge is injected subsequently into each pixel according to a mask pattern. The charge is chosen in such a way that it will exceed most threshold settings. The number of registered hits in 100 injections is given in the occupancy histogram for all pixels. In that way, the response of the readout chip can be tested and a good overview of the readout chip status is obtained. Noisy pixels will be visible as they register more hits than the number of injections, while dead pixels will not report any hits.

- **ToT-verify** In this scan, a specific charge is injected 50 times and the average ToT response per pixel is calculated and visualised in different ways in Figure 4.2a. The top figure shows a two dimensional histogram indicating the ToT response of each pixel. The central figure displays the one dimensional histogram of the ToT response, while the ToT as a function of the channel number is shown on the bottom. This scan is used to verify a homogeneous response to the most likely charge collected by the attached sensor.

- **FDAC tune** The FDAC tune is used to achieve a homogeneous charge response. The FDAC parameter is an in-pixel digital to analog converter (DAC) value which changes the feedback current responsible for the discharge speed of an incoming signal. Lower FDAC parameters correspond to lower feedback currents and thus longer discharge times and higher ToT. The FDAC tune probes all possible FDAC values for all pixels and choses the value resulting in the best fit to the required ToT response. The resulting distribution of FDAC values across the readout chip is shown in Figure 4.2b. The impact of FDAC on the mean ToT is shown in Figure 4.2c.

- **Threshold scan** A threshold scan is composed of a series of analog scans with increasing injected charge. The scan usually starts with an injected charge of zero and ends with a value for which the expected threshold is well exceeded, i.e. for a threshold of 2000 e a well chosen upper limit of the injection is 4000 e. For an average chip, each VCAL step corresponds to
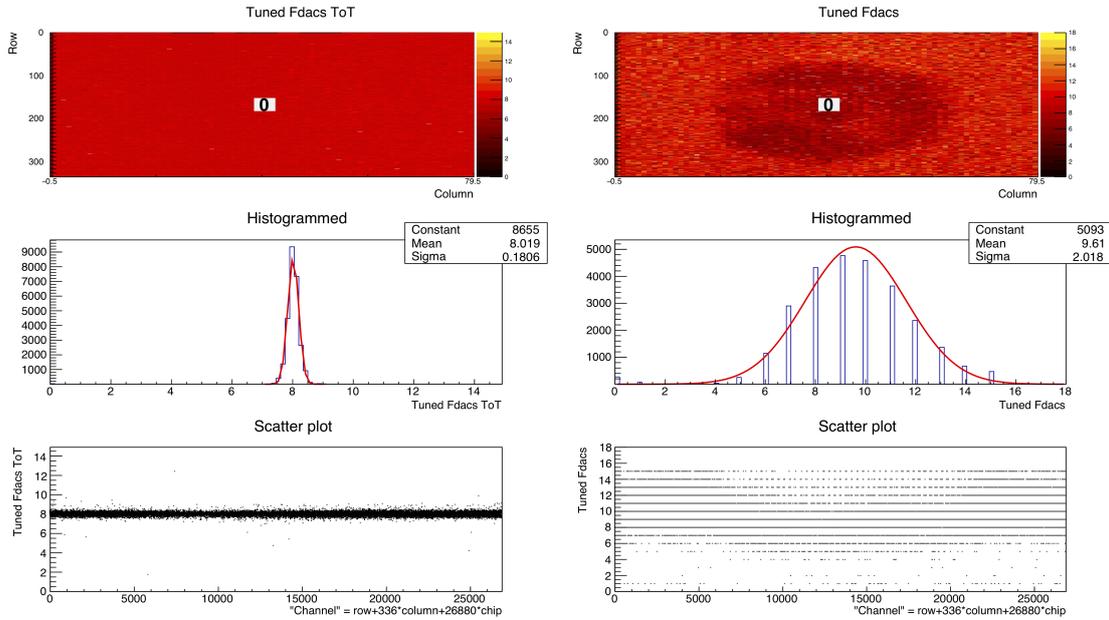
(a) ToT distribution

(b) FDAC distribution
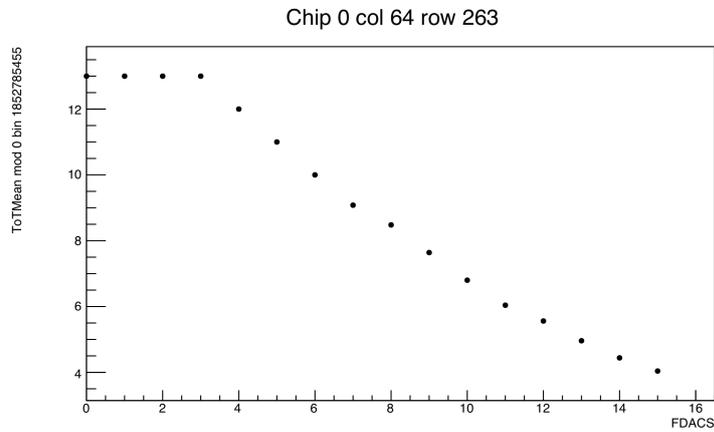
(c) Mean ToT

Figure 4.2: Relevant quantities regarding the ToT tuning mechanism of the FE-I4. The result of a ToT-verify scan with the mean ToT response to 50 injections (a). The distribution of in-pixel FDAC values after tuning (b). The dependence of the mean ToT on the FDAC parameter for one particular pixel (column = 64 / row = 263) (c). The figures are produced by the STcontrol software.

(a) Mean occupancy

(b) Threshold distribution

(c) TDAC distribution
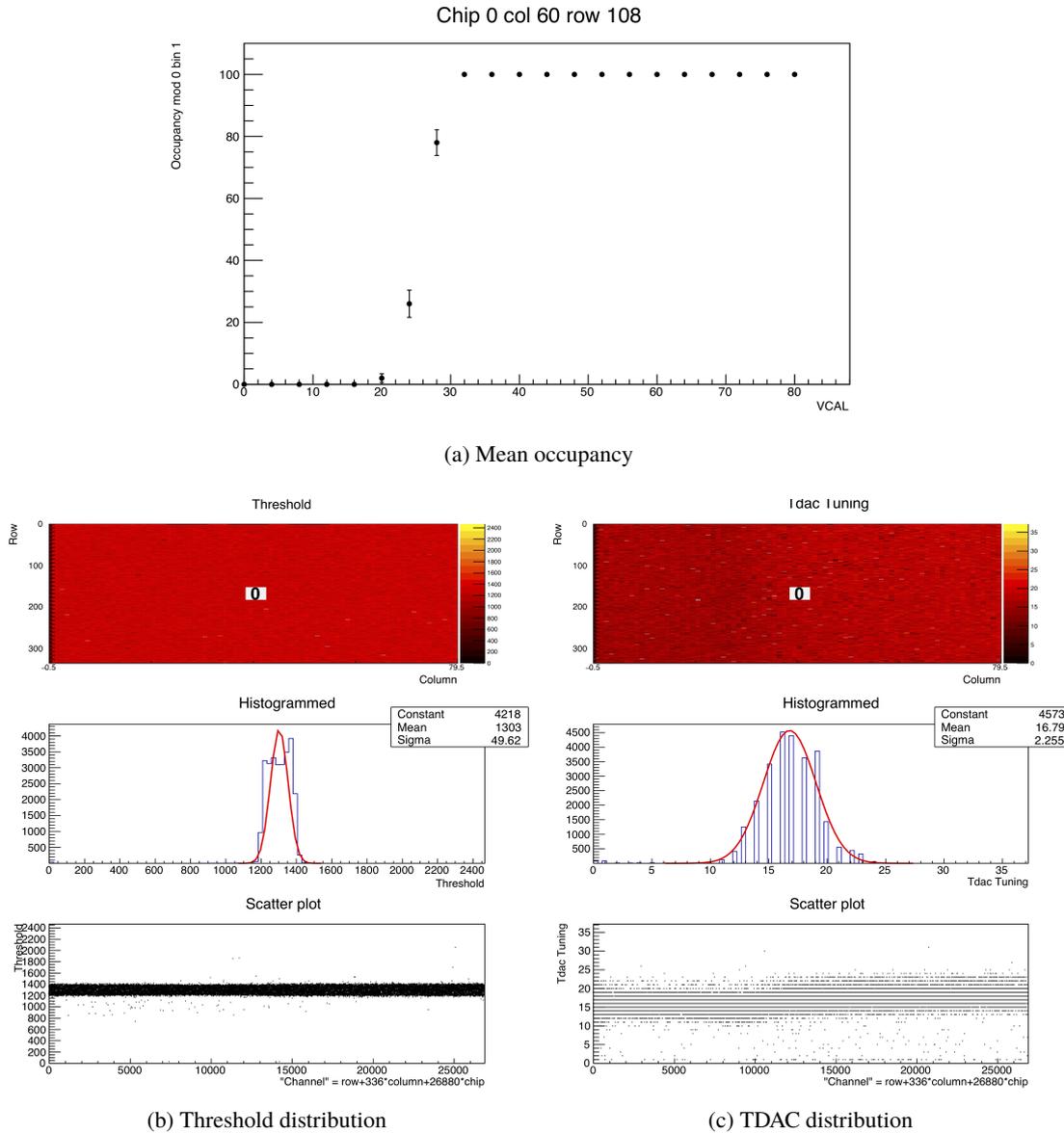
Figure 4.3: Relevant quantities regarding the threshold tuning mechanism of the FE-I4. The mean occupancy of 100 injections as a function of injected charge for one particular pixel (60/108) (a). The distribution of thresholds is shown in (b). The distribution of in-pixel TDAC values after tuning (c). The figures are produced by the STcontrol software.

an injected charge of about 30 e. The mean occupancy is calculated for each charge being injected 100 times per pixel (see Figure 4.3a). The obtained data for a given pixel is fitted to an error function from which the point corresponding to 50 % occupancy is taken as the threshold. The threshold distribution after tuning of all pixels can be seen in Figure 4.3b.

- **TDAC tune** The threshold tuning is performed by this tune. The TDAC parameter is an in-pixel DAC value which sets a voltage based on which the comparator discriminates the input signal. A TDAC tune is composed of a set of threshold scans, performed at varying TDAC values. The TDAC parameter per pixel is calculated after each threshold scan by a
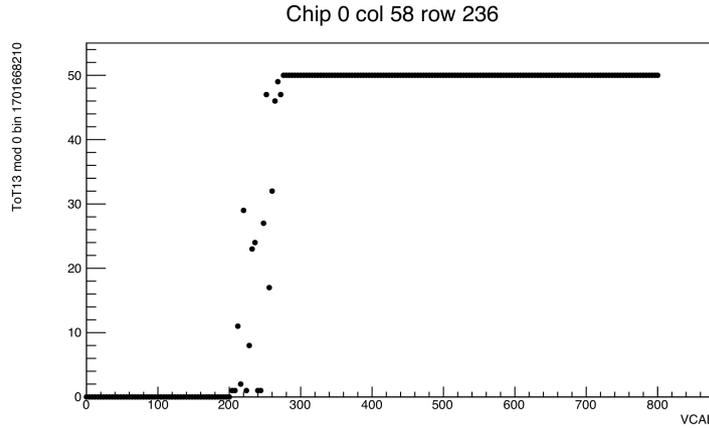
Figure 4.4: Frequency of ToT = 13 in 50 injections as a function of injected charge for a particular pixel (58/236).

nested interval algorithm. The first tested TDAC value for all pixels is 16, representing the center of the available range. Depending on the result of the first threshold scan, the TDAC is decreased if the measured threshold is below the desired threshold and vice versa. The steps and the starting value can be defined by the user. The resulting TDAC distribution is displayed in Figure 4.3c. Given the cross-dependence of ToT and threshold, an iterative tuning procedure of multiple TDAC and FDAC tunes may be required. In any case, after tuning either ToT or threshold, the respective other quantity has to be validated.

- **ToT-to-charge calibration** This calibration can be performed once the module is tuned. The scan yields the occupancy of each ToT per pixel for charges ranging from 0 to 30 ke. For each pixel, the function

$$Q = a + b \cdot \text{ToT} + c \cdot \text{ToT}^2 \tag{4.1}$$

is fitted to the data [84]. Exemplarily, the frequency of ToT = 13 as a function of the injected charge (VCAL value) is presented for one particular pixel (58/236) in Figure 4.4. This scan has to be performed in case the exact charge information is required. In that case the charge per hit has to be calculated with the per-pixel parameters of Formula 4.1 before hit clustering due to the non-linear ToT-to-charge conversion.

BDAQ

The Bonn data acquisition (BDAQ) readout system [85] has been developed in Bonn. It includes both software and hardware components. The BDAQ hardware is shown in Figure 4.1b and is built again on an FPGA module for communication with the RD53A module (see Figure 4.1d). The data and control signals from and to the readout chip are guided through a commercial DisplayPort (DP) cable. The HitOr trigger signal from the chip is connected to the board via a DP cable as well. The communication to the PC running the BDAQ software is realised through an ethernet cable. Power can be provided either by a USB cable (5 V) or a coaxial power cable (5−12 V). The BDAQ board also features a TLU port [83] for full testbeam compatibility.

(a) ToT calibration

(b) Timing

(c) Occupancy

(d) Threshold

(e) TDAC

(f) Noise
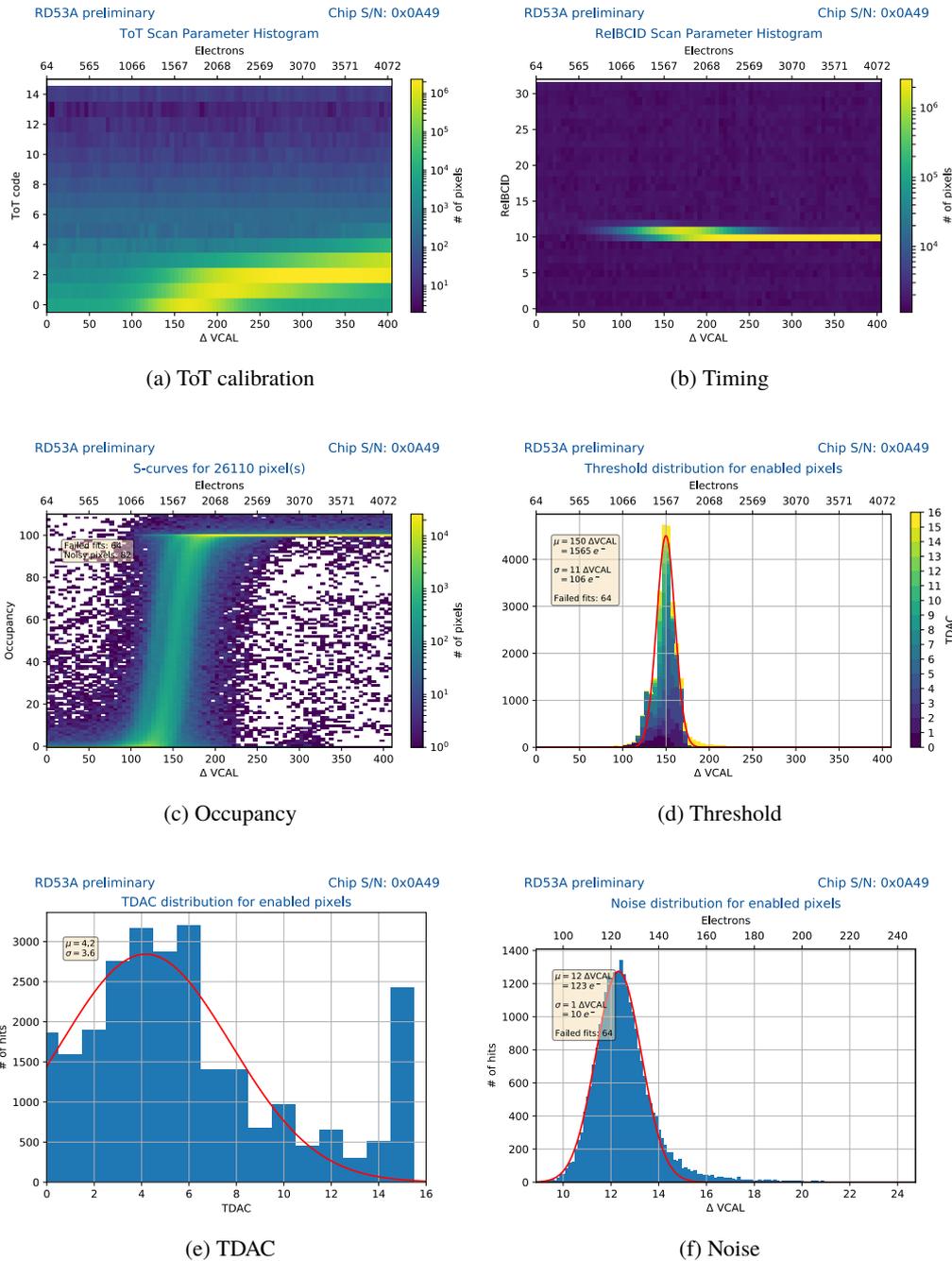
Figure 4.5: Most relevant output for a threshold scan performed with the BDAQ system. The ToT as a function of the injected charge measured in VCAL (a). Time of registration relative to the trigger signal of injections as a function of injected charge (b). Occupancy as a function of injected charge displaying the expected sigmoidal form (c). Threshold distribution extracted from (c) by fitting error functions to the per-pixel measurements (d). The TDAC distribution is indicated by the colour code. The TDAC distribution of all pixels is presented in (e). The per-pixel noise extracted as the width of the error function fitted to the occupancy (f).

The BDAQ system allows for similar scans as the USBPix system apart from the fact that no in-pixel feedback-current trimming is possible in the RD53A. At this point in time, neither ToT-to-charge calibration nor external-trigger scans are possible with BDAQ. A selection of the most frequently used results from a threshold scan are shown in Figure 4.5. The average ToT response is recorded during the threshold scan and displayed in Figure 4.5a.

In the case of RD53A, one VCAL step corresponds to a charge of approximately 10 e. The registration time relative to the trigger signal, denoted by the relative bunch-crossing identification number (BCID), of hits as a function of the injected charge is presented in Figure 4.5b. One relative BCID corresponds to 25 ns, one clock cycle of the LHC beam. Smaller hits are registered later due to the timewalk effect. A superposition of all occupancy measurements in a threshold scan is shown in Figure 4.5c. Fits of error functions to each measurement yields the distribution of thresholds in Figure 4.5d. The bars are colour coded according to the TDAC parameter of the pixels they are composed of. The TDAC distribution resulting from a TDAC tune is shown in Figure 4.5e. The fits of error functions also provide information about the per-pixel noise presented in Figure 4.5f.

4.1.3   Measurements with radioactive sources

The amount of charge a sensor is able to collect is required to judge the performance of sensor designs before and after irradiation. The advantage is that the necessary measurements can be carried out in the laboratory. The three radioactive sources available for measurements are summarised in Table 4.1. The ^{90}Sr source produces β electrons and is used to mimic the realistic charge deposition a MIP would create. The energy spectrum of the ^{90}Sr is continuous up to an energy of 546 keV. The ^{90}Y decay product decays in another β decay with a much shorter half-life and a maximum energy of 2.28 MeV. Consequently, the β electron energy spectra of the two decays are present. Free charge carriers are released along the path of the electron through the silicon. The deposited energy can be calculated from the stopping power of silicon for example for a 546 keV electron [86] and the density of silicon to be

$$1.63 \, \text{MeVcm}^2/\text{g} \cdot 2.33 \, \text{g/cm}^3 = 380 \, \text{eV}/\mu\text{m}. \tag{4.2}$$

The division of this value by the average necessary energy to generate an electron-hole pair in silicon of $W = 3.6$ eV [45] yields 106 electron-hole pairs created per µm. Nonlinear effects in thin sensors such as not contained delta electrons decrease this number to about 80.

There are different uncertainties involved in the measurement of the absolute collected charge. A first important source relates to the calibration of the internal injection mechanism of the readout chip. Also, the only four bit wide ToT register of FE-I4 and RD53A results in difficulties to resolve both the low charge regime caused by charge sharing and the high charge regime due to the Landau distribution of the charge.

The collected charge is calibrated by measurements with either of the two γ sources to reduce the systematic uncertainties. The ^{241}Am source produces mainly photons with an energy of 59.5 keV resulting in a released charge of 16.5 ke while the 22.2 keV photons of the ^{109}Cd source result in a released charge of 6.2 ke. Since the photons are either absorbed entirely or do not interact at all

Table 4.1: Overview of the radioactive sources used for laboratory characterisation of pixel modules [87]. Energies indicated for β decays correspond to the maximum energy of the continuous electron-energy spectrum.

Element	Atomic number	Decay type	Energy [keV]	Branching ratio [%]	Half life
Strontium	90	β^-	546	100	29.8 a
Yttrium	90	β^-	2280	100	64.6 h
Americium	241	γ	59.5	35.9	432.2 a
Cadmium	109	γ	22.2	55.7	461.9 d

with the material, a fixed amount of charge is recorded by the readout chip. The γ source is chosen in such a way that the released charge is closest to the expected charge in the β electron case. Still, the total uncertainty of the recorded charge was found to be about 12 % [62]. For that reason, only relative measurements are used for this thesis, where most of the systematic uncertainties cancel out and the absolute collected charge is not determined.

Apart from the calibration of CCE measurements, the γ sources are also used to inspect the interconnection between readout chip and sensor. The wide opening angles of the γ sources allow to illuminate an entire module at once. Every not responding pixel in the occupancy map is identified to be an open bump-bond connection.

All scans are performed using the internal HitOr mechanism of the chip for triggering the data acquisition. The HitOr is triggered if any of the pixels in the matrix registers a hit. The HitOr signal is transmitted to the DAQ board, where it is discriminated and the generated trigger signal is sent back to the readout chip. The hit corresponding to the incoming trigger is subsequently read out from the readout chip.

4.2 Measurements using particle beams

Measurements using particle beams are the most realistic method to test prototype detectors during the development phase. While some quantities like CCE can be determined in a laboratory using radioactive sources, the hit efficiency for MIPs can only be measured at a testbeam. Furthermore, all kind of position resolved in-pixel information like efficiency, collected charge or cluster information require the usage of beam telescopes for single particle trajectory reconstruction. Those telescopes usually consist of multiple layers of reliable, high-resolution pixel detectors up- and downstream of the device under test (DUT). The information from the telescope is used to reconstruct the particle trajectory to determine the impact point on the DUT with sub-pixel resolution. The methodology of testbeam measurements from the different facilities to the data reconstruction is described in the following.

4.2.1 Beam lines

Since high energetic particle beams are necessary to perform these measurements, only a few facilities worldwide offer testbeams. The two facilities that are used for all measurements within
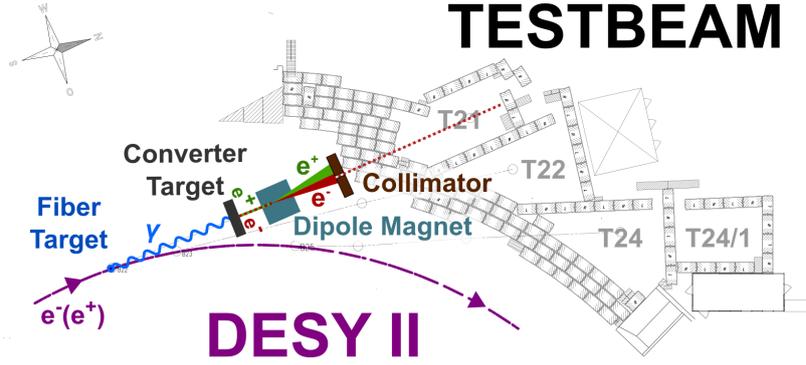
Figure 4.6:   Sketch of the DESY testbeam facility indicating the beam generation through bremsstrahlung photons and pair-creation of electrons and positrons for the three testbeam lines [88].

this thesis, located at the Deutsches Elektronen Synchrotron (DESY) and at CERN, are presented in the following.

DESY

The DESY II synchrotron accelerates electrons and positrons to energies up to 7 GeV. It is used as an injector for the storage rings DORIS and PETRA as well as to feed the testbeam facility. The primary beam is directed on a 25 µm thick carbon-fibre target which generates bremsstrahlung photons. A metal converter target is placed in the path of the photons to initiate electron-positron pair production. The narrow tertiary electron-positron beam is spread out like a fan using a strong magnetic field. A moveable collimator is used to select the particle type and energy [88]. This beam is guided to the experimental areas as displayed in Figure 4.6. Three beam lines are available at DESY out of which two are equipped with EUDET telescopes (see Section 4.2.2). The final beam has a particle energy between 1 and 6 GeV, where the intensity is limited by the $1/E$ dependence of the bremsstrahlung spectrum. Up to 1000 particles per cm^2 are available in a quasi-continuous beam with a 5 % energy spread and an angular divergence of 5 mrad [89].

Typically, a beam energy of 4 GeV is selected being a good compromise of rate and energy. The energy plays an important role as the telescope space-resolution deteriorates for low energies due to multiple scattering. Every passage of a particle through matter causes it to deviate from its original trajectory through Coulomb scattering. The effect increases drastically for low energies and large material budgets $\epsilon_{\mathrm{rad}} = \frac{x}{X_0}$ being the thickness of the traversed medium x in radiation lengths X_0. The root mean square (RMS) of the multiple scattering through small angles at one layer of a particle with atomic number z_p and momentum p is [90]

$$\Theta_{\mathrm{MS}}^{\mathrm{RMS}} = \frac{13.6\,\mathrm{MeV}}{\beta c p} \cdot z_\mathrm{p} \sqrt{\epsilon_{\mathrm{rad}}} \cdot \left[1 + 0.038 \ln\left(\epsilon_{\mathrm{rad}} \frac{z_\mathrm{p}^2}{\beta^2} \right) \right]. \tag{4.3}$$

Only two DUTs are measured at a time at DESY and a light-weight Styrofoam cooling box is used to reduce the material budget as much as possible. The box is shown in Figure 4.7a and was developed in Dortmund [91]. The box is divided into two chambers one of which houses the DUTs while the other contains dry ice blocks to cool the modules. Temperatures below −50 °C are achieved after

(a) (b)
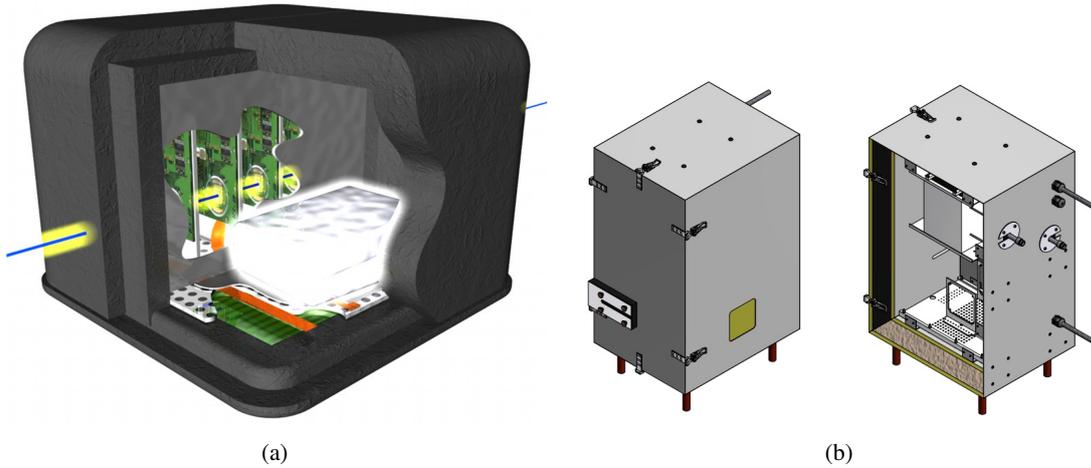
Figure 4.7: Drawings of cooling boxes for testbeam purposes. The box developed by Dortmund [91] employs Styrofoam walls and dry ice cooling. It is used both at CERN and DESY (a). Another type of cooling box is permanently installed at CERN. It was developed at MPP and employs a 46 mm PUR rigid foam isolation layer sandwiched by two 1.5 mm aluminium plates for mechanical stability (b).

the box is filled with dry ice and carefully sealed while it increases to about $-10\,°C$ after several hours. Usually one set of modules can be measured with one set of dry ice, which is exchanged when the box is opened anyway to mount new modules.

CERN

The proton beam at CERN is generated in a linear accelerator and sent to the LHC pre-accelerators PS and further to the SPS. The primary beam with a particle energy of 450 GeV is then directed onto a target at SPS. The secondary beam consists mainly of photons, electrons and hadrons of various energies. The particle type and energy is selected by electromagnetic fields. The beam employed for all measurements in this thesis consists of pions, kaons and protons with an energy of 120 GeV. The SPS machine accelerates several spills in a structure called super cycle (SC). In one SC of about 45 s, various experiments are served by SPS and only one or two spills, depending on the other users, are utilised for the testbeam facility. A typical spill has a duration of about 5 s and contains up to 2×10^8 particles with an energy spread of only 1.5 % [92].

There are two EUDET type telescopes available at CERN placed one after the other in the same beam line. This is possible thanks to the negligible multiple scattering effect at a beam energy of 120 GeV. It also enables the usage of a more solid cooling box shown in Figure 4.7b. The cooling box was developed at the MPP and employs a 46 mm PUR rigid foam isolation layer sandwiched by two 1.5 mm aluminium plates for mechanical stability. The cooling is realised through a JULABO chiller pumping silicon oil through a custom made heat exchanger equipped with a fan for more efficient heat transfer. Temperatures as low as $-50\,°C$ can be achieved and constantly kept. Another copy of the Dortmund cooling box is installed in the second EUDET telescope.

4.2.2   EUDET beam telescopes

Several copies of EUDET type beam telescopes were produced and are used at testbeam facilities around the world. A EUDET type telescope contains six planes of fine segmented MIMOSA-26 pixel detectors [93]. Three layers each are located up- and downstream of the DUTs. The MIMOSA sensors are built in a monolithic active pixel sensor (MAPS) technology. Each of the 576 by 1152 pixels has a size of $18.4 \times 18.4\,\mu m^2$ spanning a total active area of $10.6 \times 21.2\,mm^2$. The sensors are thinned to a thickness of $50\,\mu m$ to reduce the material budget to only $0.3\,\%\,X_0$. The zero-suppressed data are read out in a continuous rolling shutter mode with a maximum possible rate of about $9\,kHz$ set by the integration time of $115.2\,\mu s$. The data of all planes are collected in an aggregator board and are then fead into an FPGA in a National Instruments (NI) crate [94].

The performance of the telescope in terms of tracking efficiency is measured in Reference [95] to be $(99.5 \pm 0.1)\,\%$ at a fake rate of 10^{-4} per pixel at room temperature. This is enabled by the low-threshold operation possible due to the low-noise architecture of the employed MAPS technology. The sensors are water cooled at a temperature of $15\,°C$ inside an aluminum enclosure. The intrinsic resolution for a single pixel is $5.3\,\mu m$. Experimentally, the point resolution is improved to $3.5\,\mu m$ at DESY and $2\,\mu m$ at CERN due to multi-pixel clusters despite the detrimental multiple scattering [95].

The acquisition of data is triggered by two BC408 scintillators each up- and downstream of the telescope. A TLU connected to the scintillators analyses their signals and generates a trigger for all connected devices if a user defined coincidence criterion is satisfied. Connected devices can veto the generation of triggers in the TLU in case they are busy. External devices can be connected via RJ45 or LEMO connectors and can choose between different handshake options [83]. The TLU also provides trigger identifiers to all connected devices to synchronise the data stream. The devices connected to the TLU typically are the telescope itself and the DAQ boards to read out FE-I4 and RD53 readout chips (see Section 4.1.2).

A time reference plane is necessary when using EUDET type telescopes since the readout time of $112\,\mu s$ captures several particles in one frame at the usual particle fluxes. Instead, the DUT, being sensitive for only $400-800\,ns$, is only responsive to the particle which has provoked the trigger. Thus, the reference time is necessary to calculate the number of tracks which should have been registered by the DUT. Usually either a not-irradiated FE-I4 or RD53 module is chosen to give the reference time.

The data taking of all devices is steered by the EUDAQ software [96]. Programs called *producers*, provided by the different devices, connect through TCP/IP to the PC on which the EUDAQ software is executed. The `runControl` interface visualises the connected producers and their status. The data and log collectors serve to write the incoming data streams of the different devices to disk. The incoming data can be controlled by the `OnlineMonitor` program where basic quantities like hit-maps, ToT and relative BCID (LV1) distributions as well as the correlations between all devices can be examined. The data taking is started in `runControl` and organised in runs limited by a certain file size. A small run size reduces the amount of data which is lost in case of corrupted files while larger run sizes lead to faster data taking as the turnaround time from one run to another being about $15\,s$ is avoided. The runs are saved in a raw data file which contains all data streams in

their specific format. The decoding of data streams is handled by user provided converter plugins in EUDAQ and is used not only for monitoring of data but also for the reconstruction.

4.2.3 Testbeam reconstruction

EUTelescope

The testbeam reconstruction framework EUTelescope [96–98] is used for data analysis. The typical workflow is explained in the following. The analysis is organised in jobs called by the modular analysis and reconstruction for linear collider (MARLIN) toolset. Each job is responsible for a specific part in the workflow and is composed of various processors. Each processor reads, modifies and writes data. The input parameters for all processors of each part of the reconstruction are organised in steering templates in the xml data format. The various reconstruction parts in a typical workflow, with the most important input parameters, are presented in the following. Thanks to the versatile architecture of EUTelescope, also different workflows leaving out or adding additional steps can be implemented.

- **Converter** The first step employs the DAQ dependent converter plugins of EUDAQ to convert the raw data into the linear collider input/output data format used by EUTelescope. The occupancy of all pixels is monitored while converting the hit information of each plane. Pixels with occupancy levels above a user-defined limit are flagged and registered in a noisy-pixel database. EUTelescope also analyses the number of flagged pixels as a function of occupancy limit to provide information about the ideal maximum allowed occupancy. This limit should be set in such a way that the shape of the beam profile starts to be recognisable.

- **Clustering** The clustering step converts single hits into clusters. Clusters can be built with various algorithms implemented in EUTelescope. The user can provide limits that define the spatial extension to which hits can be considered to belong to the same cluster. Subsequently, clusters which involve, or are in the vicinity of flagged pixels from the noisy-pixel database, are removed from the cluster collection.

- **Hitmaker** The hitmaker step converts clusters to hit information. For that purpose, different algorithms can be used to determine the hit position from the cluster shape. Usually a ToT weighted center-of-gravity algorithm is used. The hitmaker is also the first step which involves the geometry description provided by the user. Accordingly, the hit coordinates are transformed into a global coordinate system in which correlations between the different planes can be used to obtain a first pre-alignment. This requires a beam with a small divergence as it is usually the case. All rotations around the beam axis and tilts with respect to the beam axis beyond a few degrees have to be correctly described in the geometry file as otherwise no correlations can be obtained.

- **Align** The alignment step aims to produce a final alignment with a precision of a few μm. The pre-alignment from the hitmaker step is applied to all hits to obtain a best guess of the module positions. A pattern recognition algorithm groups together hits on several planes.

The algorithm starts on a user-defined seed plane and searches within an extrapolated cone on the next plane. A certain number of hits per group is required for track candidates to be accepted. Similar tracks are identified based on the overlap of involved hits and only the most probable track is retained. The collection of track candidates is fed to the Millepede II software [99] which performs a least square minimisation of the residuals of the whole system. The algorithm minimises the sum in quadrature of the residuals, which is the distance of an individual hit to the track candidate projected to the plane of the hit. The fitting of track candidates is performed with the deterministic annealing filter (DAF) fitter [100], an extended Kalman fitter [101] optimised for high noise environments. An alternative fitter is available with a general broken line (GBL) algorithm [102] but not used in any reconstruction of this thesis.

- **Fitter** The final step of the reconstruction uses the previously obtained alignment. The hits of the aligned planes are again grouped in track candidates which are afterwards fit with either DAF fitter or GBL. The tracks passing various quality requirements chosen from: $\frac{\chi^2}{\text{ndof}}$, the χ^2 value itself and residual cuts, are written into a ROOT output file. The output file includes the track and hit position per DUT plane such that the final analysis can be performed by an external software.

TBmon II

The reconstructed data needs to be analysed to extract the interesting quantities like the hit- and charge-collection efficiencies. That task is performed by TBmon II [103]. The track and hit positions from EUTelescope are read and some final adjustments to the alignment are performed to correct for rotations and tilts. A central set of information is built up consisting of matched hits, cluster properties and track and hit position information. In this step, hits are matched with tracks and tracks with no corresponding hit are recorded as inefficiency. Subsequently, a user-defined set of analyses is performed on this common data set. One standard analysis determines the hit efficiency and all quantities related to it: the in-pixel hit efficiency, the efficiency as a function of rows, columns and runs, the efficiency of different pre-defined regions of each sensor and other efficiency related quantities. Another standard analysis evaluates the charge collection behaviour of each plane, again with a comparable set of distributions as in the efficiency analysis.

The versatile interface of TBmon II allows for the integration of user-written analyses. A new analysis of in-pixel cluster size is developed in the course of this thesis. For this purpose, the relative position of a track within a pixel cell is connected with the information of the cluster size of the matched hit. The result can be seen for example in Figure 9.12 in Section 9.1.2 being a helpful tool in diagnosing cross-talk effects.

FEI4 analyser

The *FE-I4Analyzer* is a software package developed in the course of the thesis in Reference [62] and can be found at [104]. It was developed for the stand-alone analysis of FE-I4 and RD53A data originating from source scans (raw data input) and testbeams (LCIO input file). The purpose in the

latter case is the analysis of testbeam data without the usage of a telescope. This can be done for long clusters which occur at high inclination angles between detector and impinging particle (see Section 9.1.2). Pixels close by in space and relative BCID are grouped together into clusters. A user defined limit of allowed holes (set to two for all measurements in this thesis) defines up to which number of internal pixels without hits a cluster is still accepted. The single pixel efficiency ϵ_{pix} can be calculated as

$$\epsilon_{\text{pix}} = 1 - \sum_{c=1}^{N} \frac{h_{\text{miss}}^{c}}{w_{x}^{c} - 2} \tag{4.4}$$

with h_{miss}^{c} the number of missing hits inside a cluster c and w_{x}^{c} the cluster width (including holes) in x. The first and last pixels are subtracted from w_{x}^{c}, being always efficient by definition. The per-pixel ToT is calculated in the same way.

FluenceAnalyser

The dedicated program FluenceAnalyser is written in the course of this thesis to analyse TBmon II output files of modules that are irradiated with an inhomogeneous fluence distribution. It is necessary to correlate the efficiency with the received fluence of the different areas of a module as DUTs are never entirely illuminated during a testbeam measurement. Consequently, the average fluence in the illuminated area is different in each measurement. Furthermore, a simple cut in x and y directions would still lead to a different mean fluence than what would be expected in the cut area, assuming a homogeneous particle distribution. This origins from the fact that the hit profile is typically not flat but also follows a Gaussian distribution.

The code calculates the fluence per pixel by integrating the two dimensional fluence density approximated by a Gaussian, as it is shown in Figures 4.8 (a) and (b). For better visibility of the dynamic range, the frequency per fluence is shown in Figures 4.8 (c) and (d). The figures show the fluence profiles that are used for FE-I4 modules in Section 8.2 and RD53A modules in Section 9.2.1. The hits and tracks assigned to each pixel are filled into the corresponding fluence bins. The *TEfficiency* ROOT class is used to evaluate the efficiency per fluence bin. The statistical uncertainties are calculated according to the Clopper-Pearson interval [105]. Additional systematic uncertainties apply according to the estimation discussed in the following section. However, using this technique, the data for a broad fluence band can be derived from a single module, thus, reducing significantly the systematic uncertainties in comparisons between fluences. This is due to the fact that most of the systematic uncertainties are highly correlated between the fluence bins as the systematic uncertainties arise from reconstruction methodology, module temperature, readout chip threshold and further environmental settings which are identical for all fluences in this analysis. Still, despite the small bin width of $1 \times 10^{14}\,\text{n}_{\text{eq}}/\text{cm}^2$, an additional systematic uncertainty of $5 \times 10^{14}\,\text{n}_{\text{eq}}/\text{cm}^2$ on the absolute fluence related to the fluence extraction procedure should be taken into account.
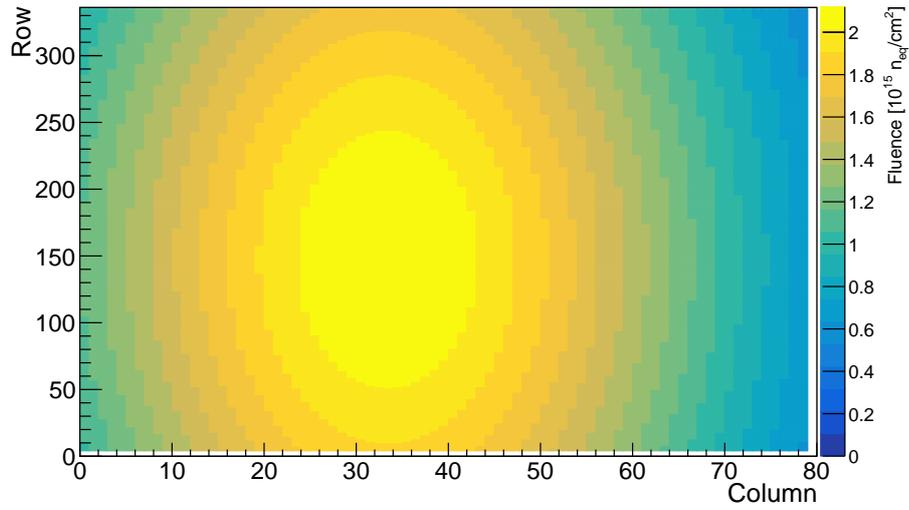
(a) Fluence map FE-I4
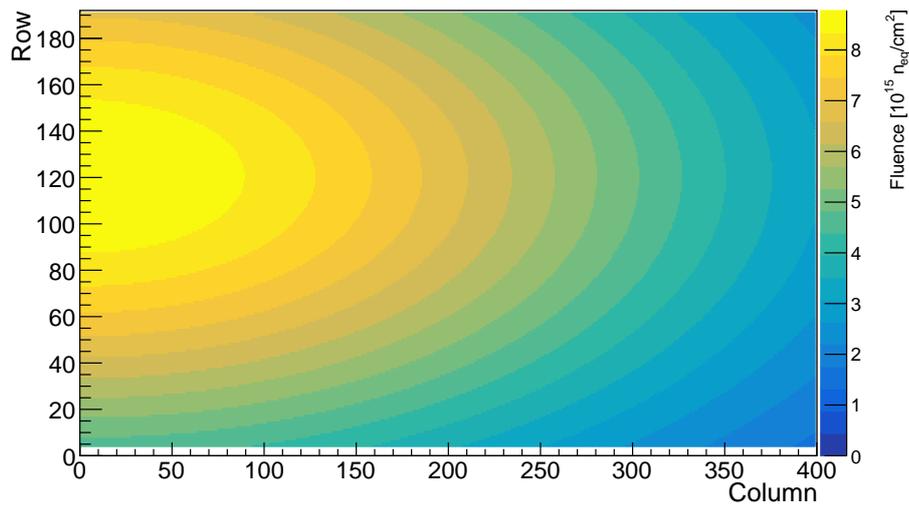
(b) Fluence map RD53A
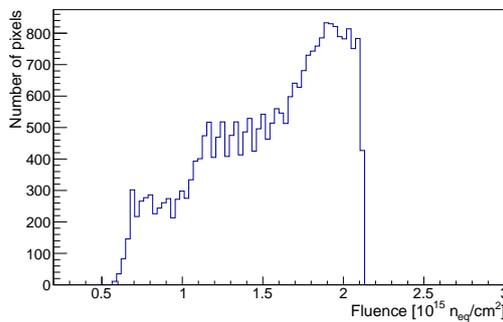
(c) Fluence distribution FE-I4
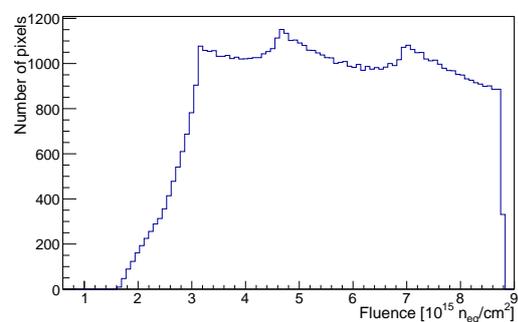
(d) Fluence distribution RD53A

Figure 4.8: Fluence distribution in the investigated FE-I4 and RD53A modules. The origin of the coordinate system in the fluence maps correspond to the bottom left corner of the module. In case of the RD53A, the synchronous FE flavour is on the left side, the linear FE flavour is in the center and the differential FE flavour is on the right side.

Uncertainties on hit efficiency measurements

The uncertainty on the measured hit efficiency is divided into two parts, a statistical and a systematic component. The statistical uncertainty on the hit efficiency is given by

$$\delta\epsilon = \sqrt{\frac{\epsilon \cdot (1 - \epsilon)}{N_{\text{tracks}}}} \tag{4.5}$$

where ϵ is the efficiency and N_{tracks} is the number of tracks used to measure the efficiency. For high efficiencies beyond 90 % and a number of analysed tracks of several 10^5, the statistical uncertainty is typically in the order of 0.01 %.

The systematic uncertainty is estimated to be 0.3 % in References [61, 62]. The systematic uncertainty arises from an uncertainty in the actual readout chip threshold, environmental influences like temperature and humidity, properties of the testbeam (e.g. energy, width, illuminated area), as well as various parameters in the reconstruction which can bias the result. The total uncertainty is computed by the sum in quadrature of the statistical and systematic uncertainties. The systematic uncertainty is usually dominating, although in special cases of small number of tracks or low efficiencies the statistical uncertainty can exceed the systematic uncertainty. The number of tracks can for example be reduced by selecting only subregions of the module to study individual irradiation levels with inhomogeneous irradiations (see Section 8.2). All shown efficiency figures include the total uncertainty according to this procedure unless explicitly quoted.

4.3   Irradiation centers

Four irradiation facilities are used for the irradiations of samples during this thesis. The different facilities are used according to their availability but also depending on their individual advantages and disadvantages, which are shortly summarised in the following.

4.3.1   Karlsruhe Institute of Technology

The irradiation facility at the Karlsruhe Institute of Technology (KIT) uses a compact cyclotron operated by Zyklotron AG. It produces a high intensity, highly focused proton beam with a kinetic energy of 25 MeV [106]. The protons have a hardness factor of $\kappa = 2$. The beam width can be varied from 4−8 mm with a flux of approximately $9 \times 10^{12} \, \text{cm}^{-2} \, \text{s}^{-1}$ at nominal beam current. The beam has to be scanned across the area to irradiate large devices due to the narrow beam width. The final fluence is determined based on the ^{57}Ni content in Ni foils placed in front of the devices which are activated during the irradiation. The devices consist of modules on PCB cards which are protected by a 2 cm thick aluminium screen, and are placed in a cold box during the irradiation. The temperature of $-30\,^\circ$C is reached by flushing cold nitrogen through the box. The nitrogen also results in a low humidity environment that prevents condensation on samples. The cold box is mounted on a movable stage to allow scanning of samples. A homogeneous irradiation profile is guaranteed by moving the turning points out of the sample area.

The comparably low energy of 25 MeV results in a high ionising dose. The ionising dose is irrelevant for sensor bulk properties but is the main damage mechanism to the readout chip. Therefore, module

irradiation at KIT are limited in this study to a fluence of 5×10^{15} n_{eq}/cm^2 corresponding to about 750 Mrad ionising dose.

4.3.2   University of Birmingham

The University of Birmingham (UoB) offers irradiations with a proton beam with a kinetic energy of 28 MeV. The beam is generated in a Medical Physics MC40 cyclotron by Scanditronix similar to the one used at KIT. The 28 MeV beam energy is reduced to about 24 MeV on the sample due to a 0.3 mm thick aluminium absorber to suppress low energetic beam components as well as the energy loss in the entrance window of the irradiation box. The beam spot has a size of about $1 \times 1\,cm^2$ and thus has to be scanned across devices to allow for a homogeneous irradiation. A typical beam current of 400 nA allows for the homogeneous irradiation of several samples to a fluence of 5×10^{15} n_{eq}/cm^2 (hardness factor $\kappa = 2.2$) in 2.5 h [107]. The samples are mounted side by side perpendicular to the beam in a box that is cooled to a temperature of $-27\,°C$ during irradiations. The box allows for a maximum surface of $15 \times 15\,cm^2$ to be irradiated. Again, the exact fluence is determined by spectroscopy of ^{57}Ni in attached Ni foils [108]. As in the case of KIT, the fluence is limited to about 5×10^{15} n_{eq}/cm^2 due to the ionising dose deposited in the readout chip.

4.3.3   CERN-PS

The irradiation facility provided by CERN is called IRRAD and is connected to the PS accelerator complex. The exceptional feature at IRRAD is the available high particle energy of 24 GeV. The proton beam is delivered in bunches of 5×10^{11} protons. There are up to three bunches per about 30 s long PS super-cycle. This leads to a comparably low fluence rate and, consequently, it can take up to two weeks to reach a fluence of 5×10^{15} n_{eq}/cm^2, depending on the availability of the PS accelerator. Thanks to the high proton energy, an ionising dose of only 220 Mrad is expected at a fluence of 5×10^{15} n_{eq}/cm^2. The beam is monitored online using a custom-made beam profile monitor (BPM) [109]. The beam has a full width at half maximum (FWHM) of $1.2 \times 1.2\,cm^2$ to $2 \times 2\,cm^2$ just wide enough to irradiate a whole module. The resulting fluence profile is Gaussian shaped such that all quoted fluences correspond to mean values across the device. The relative fluence profile is extracted from the BPM and calibrated with the spectroscopy of aluminium foils mounted on the samples.

The IRRAD bunker is divided into three areas where devices are irradiated according to their material budget. The lightest samples, like pixel detectors, are irradiated in the first zone while heavier samples like calorimeters are irradiated in the last zone. The modules are arranged in a row on a table inside the irradiation hall. Therefore, there is neither humidity nor temperature control and modules are irradiated at room temperature. The table on which the modules are mounted can be moved with an accuracy of 0.1 mm to align the modules precisely in the beam [110]. The table can also be removed remotely from the beamline which is necessary due to the high activation of samples during the irradiation.

4.3.4 Jožef Stefan Institute

The TRIGA Mark II reactor at Jožef Stefan Institute in Ljubljana is the only irradiation facility, among the four sites used in this study, based on neutrons rather than protons. It is a convection-cooled light-water pool reactor for research purposes with a peak power of 250 kW [111]. Neutrons and γ rays are used for radiation damage studies in high energy physics as well as for many other research fields like material sciences. The reactor has 12 channels (empty fuel rods) for irradiation with a diameter of 15.5 mm each. The reactor has a continuous neutron energy spectrum with a high flux of fast neutrons with energies below 100 keV with a combined hardness factor of $\kappa = 0.9$ [112]. The total neutron flux at peak reactor power is $2 \times 10^{13}\,\mathrm{cm^{-2}\,s^{-1}}$ in the most central irradiation channel. The damage to the silicon lattice occurs as displacements of bulk atoms which requires relatively high energies. Therefore, thermal neutrons and γ rays do not significantly contribute to the bulk damage. The only small contribution to the bulk damage due to low energetic neutrons comes through nuclear reactions with Si atoms and the subsequent damage by radioactive decays. Instead, fast neutrons have the largest contribution to the radiation damage in silicon sensors [113]. No cooling is provided to the sensors due to the limited space in the irradiation channels. The space constraints in the channels also require the modules to be irradiated without the PCB card such that wire-bonding has to be renewed after irradiation.

QUAD MODULES

5.1 Overview

The ITk pixel detector will mostly consist of multi-chip modules. This is necessary to reduce the total amount of modules to a minimum which is beneficial for both production and loading of modules to the mechanical support structures of the detector (staves). The size of a single chip module is driven by production constraints mainly of the readout chip. The small feature size of the readout chip requires a stepping of the photolithography mask across the wafer. The maximum allowed structure size is the one of the mask reticle. The probability to manufacture a faultless structure is decreasing with increasing structure size. This yield determines the cost of each structure which is limiting the size of one readout chip to be even smaller than the reticle. The best compromise is found to be quad modules consisting of a large area sensor equipped with four readout chips. The constraint on the sensor size is given by the cost increase due to the yield that is also in the sensor manufacturing reduced for larger structures and the available space on a circular wafer being optimally used by covering it with smaller structures. Also, single chip modules will be built to instrument the narrow area in the close vicinity of the beam pipe in the innermost barrel layer and the inner part of the ring system as well as to accommodate stronger yield constraints from the 3D-sensors to be used in these regions. The sheer size of the sensor measuring about $4 \times 4\,\mathrm{cm}^2$ requires sophisticated quality control like a measurement of the planarity and bow (Figure 5.1) using a high resolution optical profilometer.

The connection from the module to the off-detector readout system will be substantially different in the production version of the module compared to the prototyping solution. During prototyping, bulky PCB cards are used to house pins with access to many different chip signals for tests and debugging. The production version of this are flex-PCB (henceforth just called flex) cards which have the same footprint as the sensor plus a tail to accommodate an adapter to connect to the stave. All data from the module and all control signals, high and low voltages to the module are passing through this connection.

Since the RD53A readout chip was only available towards the end of this thesis, quad modules built with FE-I4 readout chips are used for all measurements in this chapter.
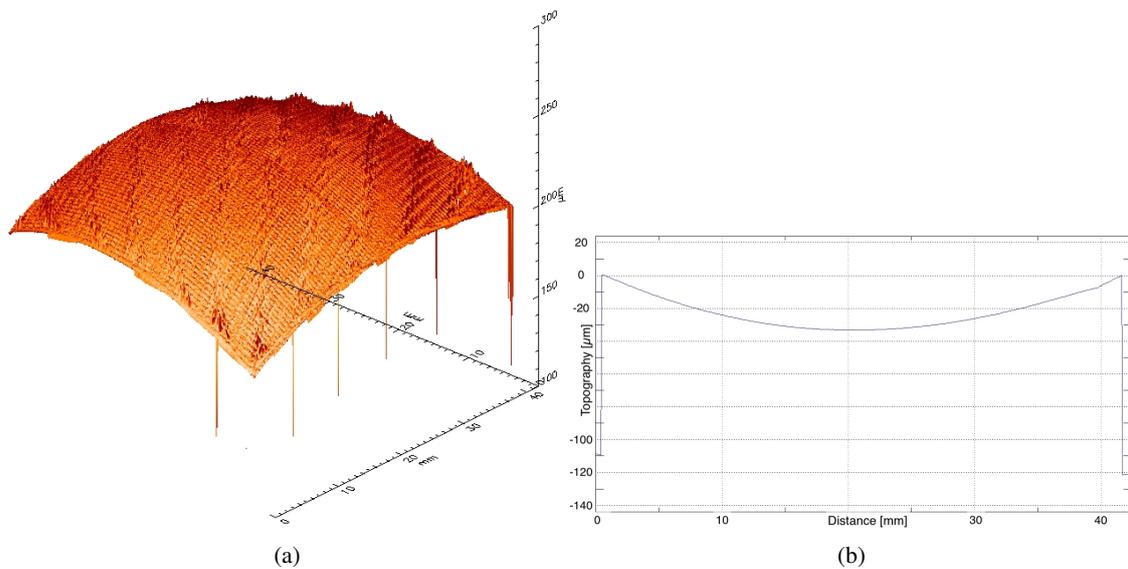
Figure 5.1:   A 3D scan of a quad sensor of the SOI3 production facing down to measure the topography employing a high resolution optical profilometer (a). A profile in the center of one direction along the other direction relative to the height at the edge of the module to extract the bow (b).
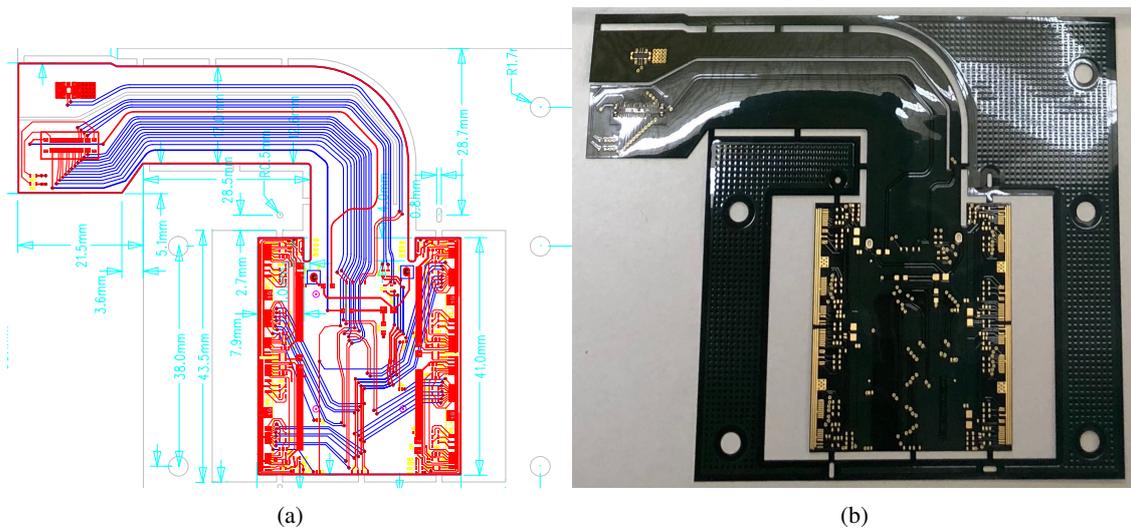
Figure 5.2:   Drawing of the module flex designed by A. Petrukhin from University of Siegen (a). Photo of the module flex in use for the ITk stave demonstrator (b).

5.2   Flex loading

The automated process of loading the passive surface-mount device (SMD) components on the flex is set up in collaboration with the MPP electronic workshop in the course of this thesis. The flex design and a picture are displayed in Figure 5.2.

A photo of the employed fully automatic loading machine is shown in Figure 5.3. After an optical reception test, the flexes are cleaned and dried. Solder is applied through an aluminium mask using
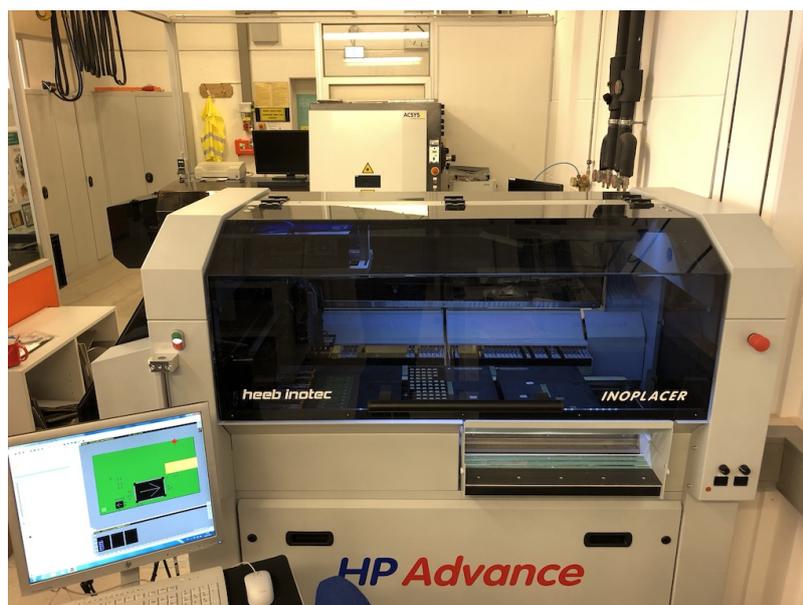
Figure 5.3: Fully automatic *heeb inotec inoplacer HP Advance* SMD loading machine.

a spatula. The prepared flexes are mounted inside the loading machine where the SMD components are placed at the respective positions onto the solder with an automatic procedure. In a subsequent step, the flexes are baked at 200 °C to reflow the solder and to establish the permanent connection between SMD component and flex. As a final step, each flex is cleaned again to remove superfluous solder.

5.3 Module assembly

The bare module has to be attached to the flex during module assembly. The utilised glue has to be as robust against environmental influences as the remaining components. This includes a wide range of temperatures but also a high level of irradiation. Double-sided tapes are used for the assembly of prototype quad modules. This approach is tested in various benchmarks before the tape is used in an actual module. Table 5.1 provides information about the tapes under investigation as well as the outcome of the individual tests.

Table 5.1: Information about the tapes under investigation for quad module assembly as well as the outcome of the individual tests: electrical test, pull test after neutron irradiation and wire-bonding test.

Tape name	Reference	Thickness [µm]	El. test	Neutr. irr. pull test	Wire-bonding
3M VHB-5909	[114]	300	passed	passed	difficult
PPI-7011DS	[115]	85	passed	passed	possible
tesafix®4962	[116]	160	passed	passed	possible

As a first test, the isolation capabilities of the different tapes are investigated. This is important as the tape is the only part in between the sensor backside at HV potential and the flex, housing various contacts at or close to ground potential. Each tape is used to connect two $2 \times 4 \, \text{cm}^2$ stainless steel
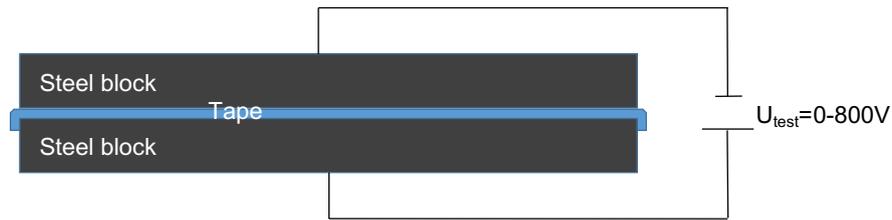
Figure 5.4: Sketch of the test procedure to extract the HV capabilities of the tape candidates.

blocks, as indicated in Figure 5.4. A HV potential is applied between the two blocks by ramping up the voltage from 0 V to 800 V while recording the current. If the current is constant across the whole voltage range, the test is considered successful. All tested tapes have successfully passed this test as they withstand at least 800 V.

The next test concerns the radiation hardness of the tape. Dummy modules are built (see Figure 5.5a) to test the tapes in realistic conditions. The flex is substituted by Kapton foil having similar properties as the real flex material. A bare silicon sensor is used instead of the flip-chipped module. That still represents realistic conditions given the fact that the backside of the sensor being glued to the Kapton foil has the same metallisation as the actual module. The dummies are tested in the setup shown in Figure 5.5b before and after neutron irradiation to fluences of 5×10^{15} n_{eq}/cm^2 and 1×10^{16} n_{eq}/cm^2. In this configuration, the dummy is hung-up inside a climate chamber using a hook attached to the Kapton side of the dummy. A test weight of 100 g is attached to another hook on the silicon side of the sensor. The whole setup inside the climate chamber is thermal cycled at least 5 times from 30 °C to −40 °C and back during a time period of 10 h. The test is considered successful if the dummy is still faultless, including no detachments in any of the corners. All tapes passed this test before and after irradiation without issues. Tests with higher masses (>1 kg) and longer time periods (> 1 week) at room temperature failed because always either of the hooks detached from the dummy. Still, no issues or detachments between flex and sensor backside are observed during those tests.

Further systematic investigations are performed on lap joints. They consist of two pieces of aluminium that are connected in a fraction of the surface through the glue under test. The two laps are inserted into a tensile testing machine that is pulling the laps apart. The tests are performed by CERN EP-DT following closely the ASTM D1002 standard [117]. A Zwick-Roell all-around line tensile testing machine is used. The lap joints are tested before and after irradiation to ionising doses of 2 MGy and 7 MGy. An industrial ^{60}Co source at BGS [118] is used for the gamma irradiation. The results for the tesafix®4962 and the 3M VHB-5909 lap joints are shown in Figure 5.6.

The maximum shear stress indicates the maximum applied shear force in relation to the contact surface before the two laps detach. The strain at which that point is reached is a measure of the flexibility of the glue. The maximum shear stress decreases for the tesafix®4962 by about a factor of four from the state before irradiation to that at 7 MGy. Also the elongation at break is decreased by about a factor of six. On the other hand, the 3M VHB-5909 exhibits even an increasing maximum shear stress. However, also this tape becomes less flexible after irradiation to high fluences. The elongation at break is reduced nearly by a factor of ten. In the ITk detector,

(a)   (b)
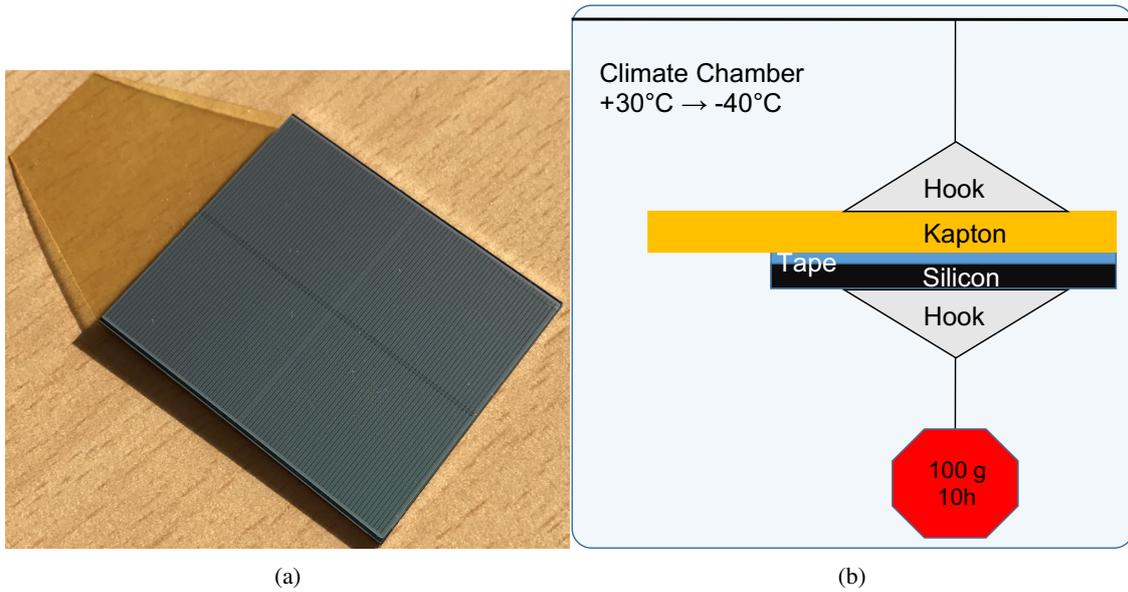
Figure 5.5: Dummy structure to test different double sided tapes (a). Sketch of the test procedure to determine the pull strength of dummy structures before and after irradiation (b). A test weight of 100 g is attached via a hook to the dummy hanging inside a climate chamber.

(a) tesafix®4962   (b) 3M VHB-5909
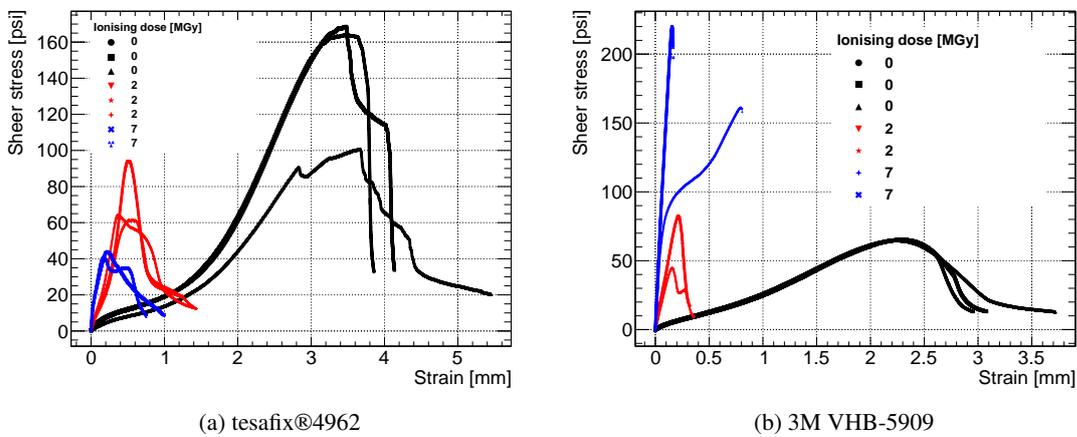
Figure 5.6: Results of apparent lap sheer tests performed by CERN EP-DT. The results of lap joints glued with tesafix®4962 are shown in (a) and those of 3M VHB-5909 in (b). The strain on the x-axis indicates the elongation of the lap joint during the pull test while the y-axis expresses the shear stress on the lap joint.

most stress is induced during handling of modules in the construction and integration phase of the detector. Before irradiation, both glues show sufficiently high maximum shear stresses with a particular advantage compared to the established epoxy glues being the flexibility, manifesting through the high elongation at break.

The last test concerns the wire-bonding of quad modules. The material below the wire-bonding pads has to be stiff enough not to follow the ultra-sonic movement of the wire-bonding wedge head. The underlying tape might cause the flex to vibrate during wire-bonding, given the sponginess of some

of the tapes. More dummies are built (similar to those in Figure 5.5a) using all available tapes to test this issue. Wire-bonding tests are performed and resulted in no issues with tesafix®4962 and PPI-7011DS, with the latter showing just a slightly worse behaviour with respect to the tesafix®4962. The 3M VHB-5909 tape is the worst of the three in this category with several unsuccessful wire-bonds and wire-bond detachments. Therefore, 3M VHB-5909 is not considered a good candidate for the quad module assembly. The final decision to use tesafix®4962 over PPI-7011DS is taken based on the fact that the handling of the tesafix®tape is easier since the PPI-7011DS has to be baked after application at a temperature of 130 °C for 2 h or 150 °C for 1 h.

The accurate positioning and attachment of the bare module with respect to the flex is crucial for a successful assembly since in this module concept the flex is tailored to the sensor. Too large horizontal displacements in Figure 5.2 would result in wire-bonding pads not being accessible. Vertical displacements would at least cause problems during wire-bonding as the wire-bonds would have to be rotated to connect the two displaced rows of pads on the flex and on the chip. Figure 5.7 introduces the tools and workflow used for precise quad module assembly. The following list summarises the necessary steps:

- The vacuum tool (1) is used to pick up the double-sided tape placed on the pick-up pad, the safety foil is removed from the bottom side of the tape once it is attached to the vacuum tool.

- The table is pushed to the right to move the flex on its mount beneath the vacuum tool holding the tape (2).

- After aligning the tape with the flex using a camera facing up and down with an overlaid image, the vacuum tool is lowered onto the flex to apply the tape to the flex.

- Now, the table is pushed to the left where the flex with the tape is stored. The vacuum tool is used to pick up the bare module (3) previously placed on the pick-up pad.

- Once the table is pushed back to the right, the bare module is aligned with the tape on top of the flex.

- Just before attaching the bare module to the flex, the safety foil is removed on the top side of the double-sided tape.

The pick-and-place machine is a very precise tool since x and y can be moved accurately with μm-screws and any rotation can be compensated by turning the vacuum tool with another screw. This allows for reproducible placement within 50 μm. The 3D-printed flex mount (Figure 5.7a) has cut-outs to accommodate the SMD components of the flex to ensure a planar flex surface for the placement of the module. Otherwise, the fragile thin sensor can break when the module is pushed onto the flex.

Figure 5.8a illustrates a schematic view of a designed support structure used to store and move quad modules safely. The module is fixed with screws through the flex, while the adapter card at the bottom left of the picture is connected to the flex to be able to read out the data via four ethernet connectors, supply low voltage via a four-pin Molex connector and HV using a Lemo connector.
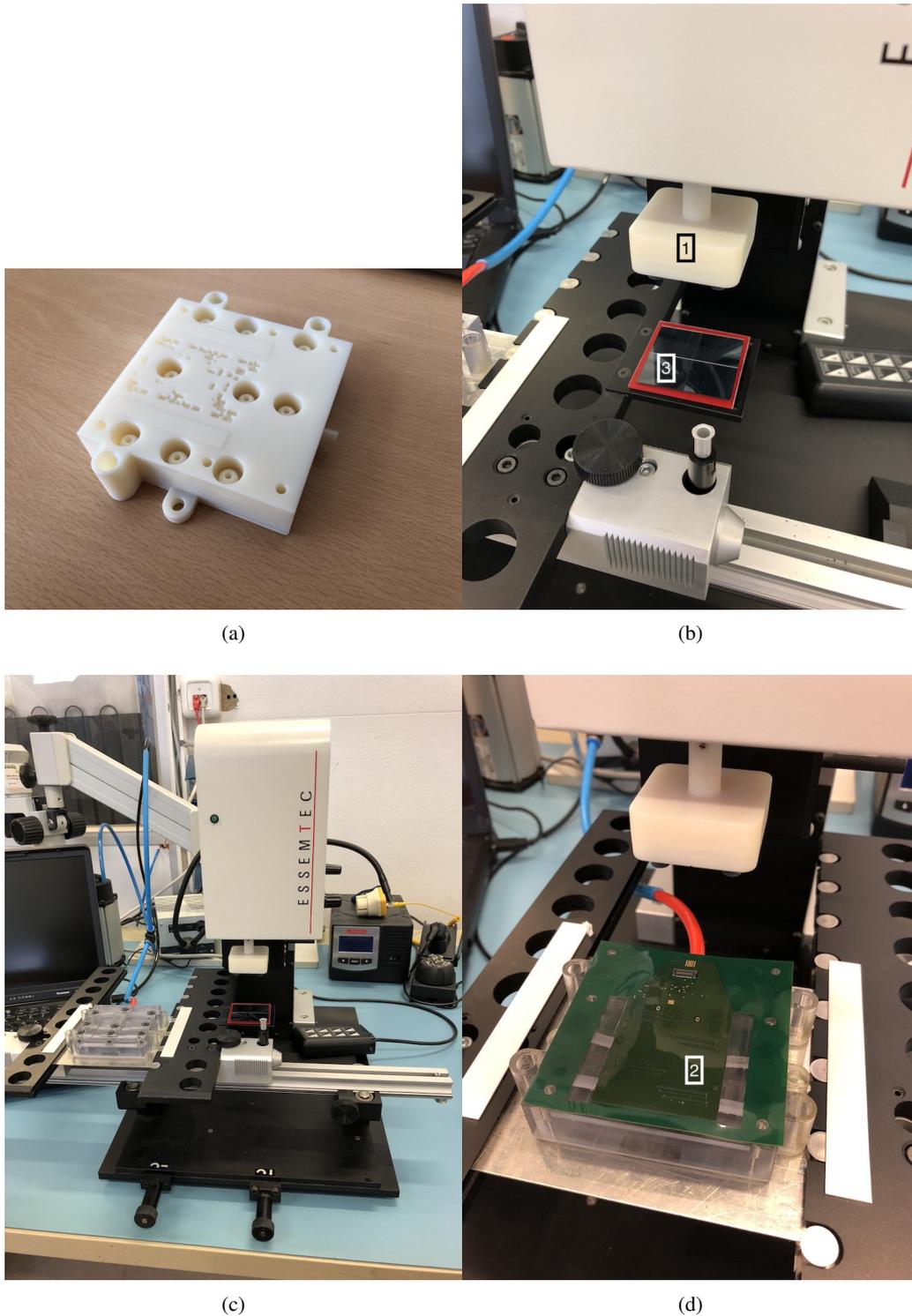
Figure 5.7: Assembly of quad modules using a pick-and-place machine. Close up of the flex mount with visible cut-outs to accommodate the SMD components (a). The bare module (3) is grabbed with the vacuum tool (1) in image (b). Overview of the machine with the bare module located under the vacuum tool and the flex mount on the left (c). The table is pushed to the right to move the flex mount (2) under the vacuum tool to place the tape and the bare module (d).

(a)   (b)
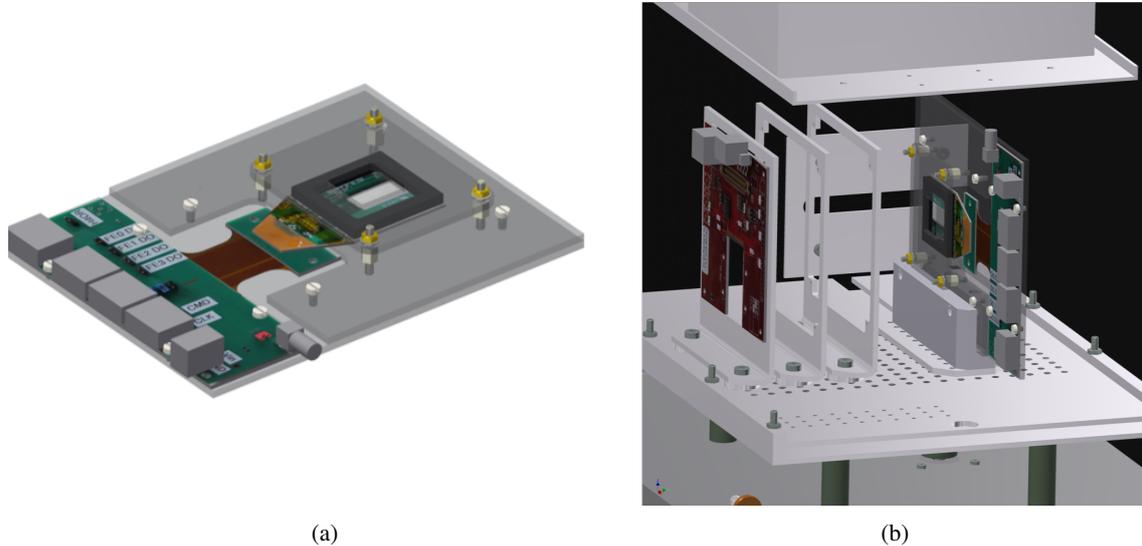
Figure 5.8: Schematic view of the support structure used for quad modules built with flexes (a) and a quad module mounted in the cooling box at the testbeam (b).

Figure 5.8b shows the support structure in a special mount inside the cooling box installed at the SPS testbeam facility at CERN (see Section 4.2). The beam is passing through the box from the left. A reference module is indicated on the left side of the image.

5.4   Quad module measurements

Table 5.2 presents the assembled FE-I4 quad modules that are investigated in the course of this thesis. The module MPP-Q3 was assembled and wire-bonded by Bonn University. The testing of this module revealed that only 3 FEs are working and a source scan indicated many disconnected channels. It is operated in the so called shunt LDO mode, where a constant current of 2 A is supplied to the quad module. Each FE chip internally generates the required digital and analog voltages. All quad modules are operated in shunt LDO mode except for the module MPP-Q5 which is operated in LDO mode, where a constant voltage of 2 V is supplied from which the appropriate digital and analog voltages are generated. In addition, this is the only module of a different bulk type being built in an n-in-n process at CiS.

Quad modules Si-6 and Si-26 are built to be used in a stave demonstrator effort [119]. Consequently, the naming scheme is adopted to the needs of the stave demonstrator effort. The stave demonstrator is used to prove the feasibility of employing serial powering with a realistic set-up in terms of mechanical supports, electrical connections between modules and number of assemblies powered together. After assembly, the modules are characterised and sent to CERN for further tests prior to being loaded on a stave prototype, sharing many key technologies and composition with the staves to be used in the ITk pixel detector.

Table 5.2: Overview of the assembled FE-I4 quad modules and their properties.

Module	MPP-Q3	MPP-Q4	MPP-Q5
Sensor type	n-in-p	n-in-p	n-in-n
Sensor producer	CiS	CiS	CiS
Thickness [μm]	150	150	285
FEs working	3	4	4
Powering	shunt LDO	shunt LDO	LDO
comment	disconnected channels	fully working	tested in beam

Module	MPP-Q6	Si-26	Si-6
Sensor type	n-in-p	n-in-p	n-in-p
Sensor producer	HLL (SOI-3)	HLL (SOI-3)	HLL (SOI-3)
Thickness [μm]	150	100	100
FEs working	2	4	4
Powering	shunt LDO	shunt LDO	shunt LDO
comment	-	for demonstrator	for demonstrator

(a)　　　　　　　　　　　　　(b)
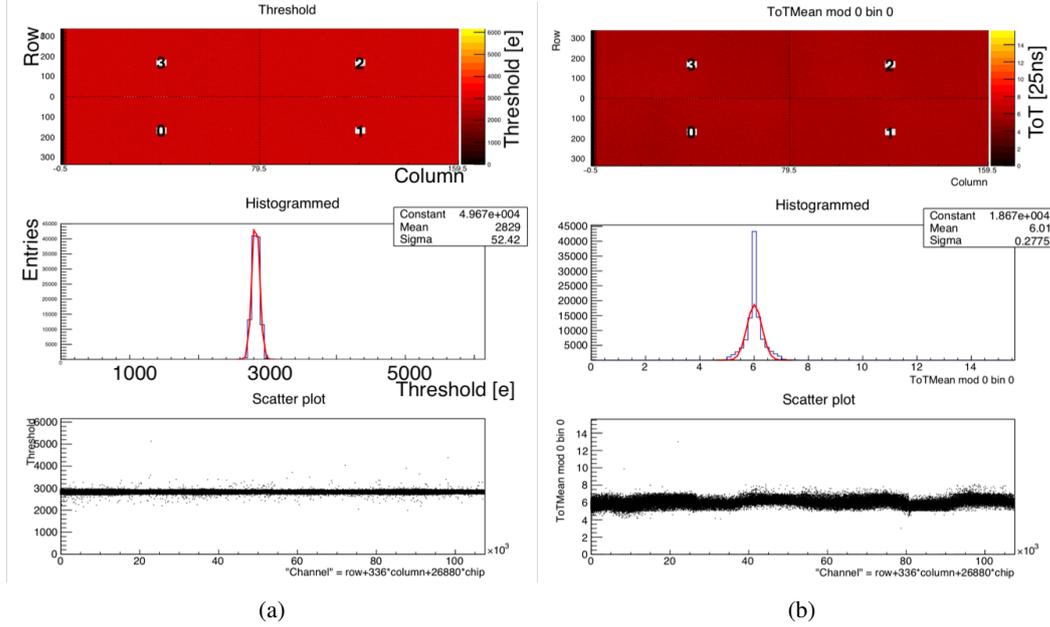
Figure 5.9: Results from the tuning of the module MPP-Q4. The figures are produced by the STcontrol software. The ToT response is tuned to yield a ToT value of 6 for a reference charge of 8000 e while the threshold is tuned to 3000 e. The 2D histogram of the threshold of all channels and the same data is also given as 1D histogram and scatter plot below (a). The threshold is (2829 ± 52) e. The average ToT response of every pixel in a 2D histogram as well as a 1D histogram and a scatter plot of the same data (b). The ToT response is 6.01 ± 0.28.

(a)   (b)
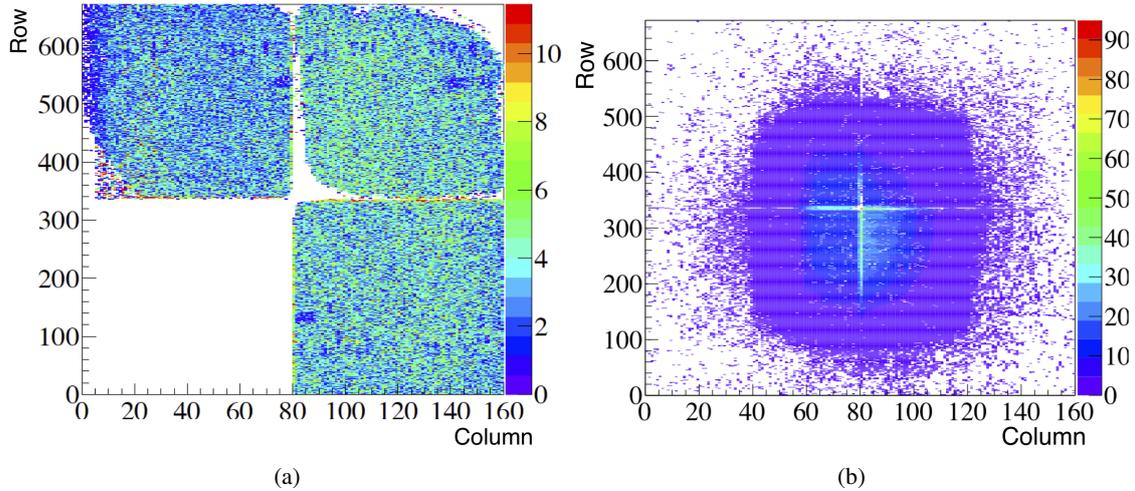
Figure 5.10:   The hit occupancy in a source scan is shown for all 4 FE chips mounted on the quad sensor. The figures are produced by the STcontrol software. The area between FEs has a higher occupancy due to the edge and ganged pixels covering a geometrically larger surface. The measurement of the module MPP-Q3 using the ^{241}Am source and a bias voltage of 50 V is shown in (a) and module MPP-Q4 using the ^{90}Sr source and a bias voltage of 80 V is shown in (b).

All quad modules are tested after assembly by simple analog scans to evaluate the functionality of the different FEs. Subsequently, the threshold and ToT response of each module is tuned using the MMC3 board and STcontrol (see Section 4.1.2). Figure 5.9 illustrates exemplarily the outcome of the tuning for module MPP-Q4. The ToT response is tuned to yield a ToT value of 6 for a reference charge of 8000 e, while the threshold is tuned to 3000 e. The result of the tuning is a mean threshold of (2829 ± 52) e and an average ToT response of 6.01 ± 0.28.

A source scan using the ^{241}Am or ^{109}Cd γ source or the ^{90}Sr β source is taken once the tuning is successfully finished. Figure 5.10a presents the results of a source scan using the ^{241}Am γ source for the module MPP-Q3. Two corners of the top right FE chip show many disconnected channels while the chip on the top left has only a few disconnected channels at its left edge. The FE chip on the bottom right exhibits only a few disconnected channels in the center of the quad indicated through the corner being round instead of square. The most likely cause of the disconnected channels is a problem during the flip-chip process. Instead, Figure 5.10b shows the result of a ^{90}Sr source scan using quad module MPP-Q4 demonstrating a much better interconnection efficiency in the central region, while the periphery can not be thoroughly tested given the low number of events caused by the narrow beam spot of the ^{90}Sr source. However, single events in the corners and edge regions of all FE chips indicate that at least no larger sensor regions are disconnected from the FE chips.

SIMULATION FRAMEWORKS

6.1 Synopsys TCAD simulation

The term *Technology Computer Aided Design* (TCAD) refers to a variation of normal CAD applications with enhanced physics models tailored for semiconductor applications. The CAD offers the possibility to create all kinds of devices in a precise and defined computer environment. The T in TCAD adds the property that the generated structures do have physics characteristics, defined by the material they are made of. The TCAD is employed in the electronic design automation aiming at modelling both semiconductor device manufacturing and operation, called device TCAD. The latter aims at reproducing the expected characteristics of real devices. The TCAD simulation offers the advantage to predict the device performance before it is being manufactured. The profound insight into fundamental physics quantities provided by TCAD goes beyond the observables accessible in actual devices.

Sentaurus TCAD is a commercial simulation suite offered by Synopsys Inc. [120] for the semiconductor device simulation in various applications ranging from solar cells to cutting edge 3D Fin-FET [121] used in logic chips. First, a continuous device is approximated with a mesh of connected vertices. The vertices are the space points on which the finite element simulation is carried out to extract simple observables like the current for a given applied potential, but also complex observables like the electrical field, the number of charge carriers per space point or the occupancy status of various traps.

The following sections explain the workflow with the various tools included in Synopsys TCAD as well as the physics and radiation damage models used in the simulation.

6.1.1 Tools of Synopsys TCAD

Synopsys TCAD comes with a variety of different programs that can all be steered by `swb`, the Sentaurus Workbench. The graphical user interface (GUI) is shown in Figure 6.1. The different programs are presented in the following in the order of the usual workflow.

- **SDE** (Sentaurus Device Editor) is the first step in all simulations. It is used to create the silicon device under study. The device is modelled with a given size in two or three dimen-
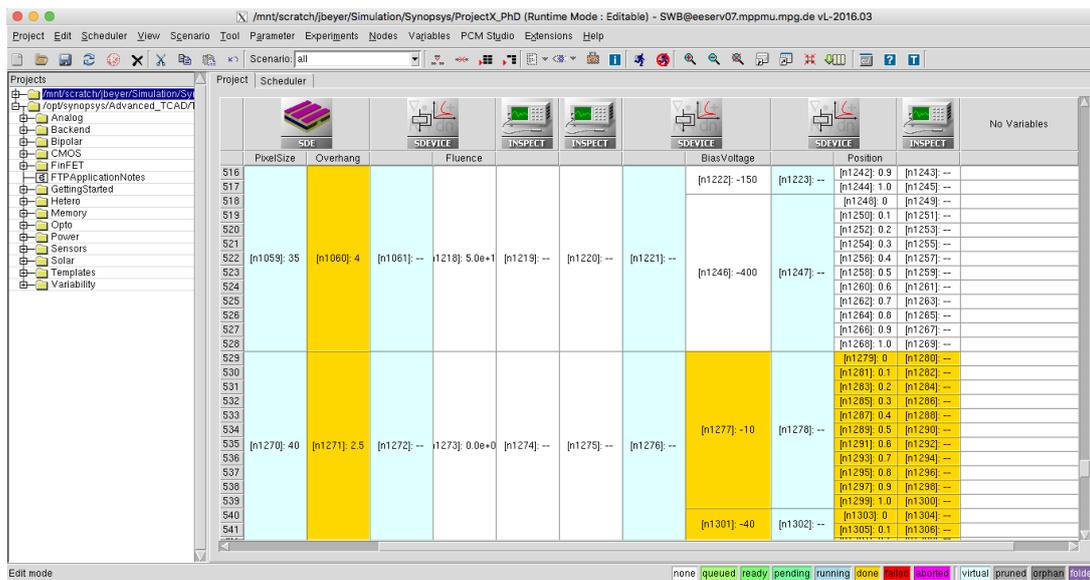
Figure 6.1: Graphical user interface of the Sentaurus Workbench `swb`.

sions. Moreover, the surface structures like a silicon oxide layer, a contact required to apply potentials and measure physics quantities or an aluminium pad, can be created in this tool. Furthermore, the dopings are defined. Since the full sets of information about the manufacturing processes of the foundries are usually not available, analytical profiles aiming to reproduce experimentally measured doping profiles are employed. The continuous structure is modelled with a discrete mesh to be used in the following simulation steps. The meshing tool provided by Synopsys is `sMesh` being integrated into SDE. The mesh configuration has a significant impact on the outcome of the simulation. It has to be detailed enough to catch all physics processes, for example charge carrier multiplication, which can happen on scales as low as $O(10)$nm. On the other hand, it can not involve too many vertices as the simulation-time scales with the amount of vertices. In the worst case, the simulation does not converge due to the excessive number of mesh points. For this reason, different meshing strategies are available which refine the mesh in areas of particular importance, e.g. areas close to the surface, with large changes in the doping concentration or with a high electrical field.

- **sProcess** is an alternative way to generate a structure. In contrast to SDE, where for example an oxide layer is simply defined to have a certain thickness, `sProcess` mimics the physics processes performed by the factory to grow that layer of oxide. In this way, the full process chain is described with all details. This includes for example the duration of the growth in the oven, including the applied temperature and the details of the temperature ramp needed to reach that value, together with the description of the exact composition and flux of the gas flowing through the oven. Also, layers of photoresist can be deposited, patterned or stripped to enable the deposition of different materials or the implantation with dopants. For the implantation, the dopant material and the energy with which the dopant is shot under a certain angle on the structure together with the total dose has to be known. This level of detail

is usually confidential to the foundry. Also, this detailed process simulation is much more time consuming than creating the structure with SDE. Therefore, this way of generating a structure is only used for the secondary implant simulation in Section 7.2, where the implantation details are disclosed by the foundry for the few relevant steps.

- **sDevice** performs the device simulation. It is used after the creation and meshing of the structure in every simulation. It performs all electrical characterisations by providing different possibilities to ramp up static voltages, perform AC measurements or simulate the crossing of charged particles. The steering file for `sDevice` includes a physics section where all models to be used are defined. Additionally, several default settings and physics models valid for most applications are applied in the background. A detailed description of the physics models used for the results obtained in this thesis is given in Section 6.1.2. The steering file also includes details about the numerical models and parameters in use. It also includes a section about the primary goal of the script, e.g. the ramp of the voltage or a transient simulation. For the ramp, a quasi-static approach is used in which the voltage is increased in small steps after which the Poisson equation is solved again for the new situation. To arrive at a typical use case in HEP applications of bias voltages of a few hundred volts, several hundred steps can be necessary. A method to ramp the current instead of the voltage has proven to be successful to avoid convergence problems. In that case, the algorithm tries to increase the current by a fixed amount and calculates the corresponding voltage change. This is particularly helpful in situations where a voltage ramp would result in either very small or very large changes of the current, e.g. close to the breakdown of a device. During the ramping, the complete set of characteristics of the device, like electrical current, charge carrier densities or electrical field, can be saved into a file. Once the target voltage is reached, or a predefined termination condition is met, the next step is executed. For example, this can be the simulation of a charged particle traversing the device and releasing charge carriers along its path. The transient simulation of the device before and after the incidence can be performed as well.

- **sVisual** is the tool of Synopsys TCAD to view the two and three dimensional maps produced e.g. in SDE, `sProcess` and `sDevice`. It offers a variety of tools to inspect the devices by cross-sections, close zooms and a comparison of different structures or settings. As all other Synopsys tools, it can be steered by a GUI but also by the Synopsys programming language. The programming language offers the possibility to automate data analyses to quickly export certain device characteristics to external tools.

- **Inspect** is used in the same way as `sVisual` except that it is focused on the analysis of one dimensional information like the current flowing through a predefined contact as a function of the applied bias voltage. It is also used to extract the capacitance information used in Section 7.3.

6.1.2 Synopsys TCAD physics models

A broad range of physics models is enabled in the program to ensure realistic conditions. All information reported below are taken from the Synopsys TCAD manual [122].

- **Mobility:** The mobility μ of charge carriers is described in a modular approach. The different contributions (μ_1, μ_2, ...) to the mobility for bulk, surface and thin layers are combined according to Matthiessen's rule that handles individual contributions as if they were independent of the others [123]

$$\frac{1}{\mu} = \frac{1}{\mu_1} + \frac{1}{\mu_2} + \dots. \tag{6.1}$$

  The simplest model calculates the carrier mobilities only as a function of the lattice temperature. This approximation, however, does not hold for the doped materials used in this thesis. In doped silicon the carriers scatter off the impurities. The default model used by Synopsys TCAD is the Masetti model [124]. The parameters of the model can be found in the Synopsys TCAD manual.

  Another degradation of mobility occurs at device interfaces crossed by charge carriers. The computation of the necessary field perpendicular to the semiconductor interface is invoked by the `Enormal` statement and performed according to the Lombardi model [125].

  The carrier velocity generally increases with electrical field strength, but it also saturates for very high fields. The reason for this is that the mobility of charge carriers is saturating at high electrical fields. This effect is particularly important for HEP applications where thin sensors and high bias voltages are used, resulting in very high electrical fields in the sensor, and it is described by the Canali model [126] activated by the statement `HighFieldSat` in the physics section. With the high field saturation enabled, the mobility is computed in two steps. The low field mobility (μ_{low}) is determined as in Equation 6.1 and then computed as a model-dependent function of the driving force F_{hfs} [122]

$$\mu = f(\mu_{\text{low}}, F_{\text{hfs}}). \tag{6.2}$$

  The driving force used in all presented simulations is by default the gradient of the quasi-Fermi potential

$$F_{\text{hfs}} = |\nabla \Phi| . \tag{6.3}$$

- **Bandgap energy:** By default, Synopsys TCAD uses a temperature dependent bandgap energy. The bandgap energy is calculated as

$$E_g(T) = E_g(0) - \frac{\alpha T^2}{T + \beta_{\text{g}}} \tag{6.4}$$

  with $E_g(0) = 1.1696\,\text{eV}$ being the bandgap energy at $0\,\text{K}$ and $\alpha = 4.73 \times 10^{-4}\,\text{eV/K}$, $\beta_{\text{g}} = 636\,\text{K}$ according to [127]. However, the bandgap also depends on the doping concentration

by an effect called bandgap narrowing. The model used to describe the bandgap narrowing is called Old Slotboom [128] and is activated by the flag `OldSlotboom`. This model also considers an effect on the electron affinity due to temperature and the bandgap narrowing.

- **Recombination:** The Shockley-Read-Hall (SRH) statistic is used to describe the generation and recombination of charge carriers according to

$$R_{\text{net}}^{\text{SRH}} = \frac{np - n_{\text{i}}^2}{\tau_p (n + n_l) + \tau_n (p + p_l)} \tag{6.5}$$

with

$$n_l = n_{\text{i}} \exp\left(\frac{E_{\text{trap}}}{k_{\text{B}} T}\right), \ p_l = n_{\text{i}} \exp\left(\frac{-E_{\text{trap}}}{k_{\text{B}} T}\right) \tag{6.6}$$

where E_{trap} is the energy-difference between the defect level (l) and the intrinsic level (i). The carrier lifetimes $\tau_{\text{n,p}}$ are modelled as the product of a temperature-dependent $\tau_{\text{T}}(T)$, doping-dependent τ_{dop}, and field-dependent $g_{\text{n,p}}(F)$ factor

$$\tau_{\text{n,p}} = \tau_{\text{dop}} \cdot \frac{\tau_{\text{T}}(T)}{1 + g_{\text{n,p}}(F)}. \tag{6.7}$$

The temperature dependence of the recombination lifetime is calculated according to a power law [129]

$$\tau_{\text{T}}(T) = \tau_0 \left(\frac{T}{300\text{K}}\right)^a \tag{6.8}$$

with τ_0 the temperature independent lifetime and $a = -1.5$ for both electrons and holes. Furthermore, the carrier lifetimes depend on the doping concentration following the Scharfetter relation [130]. Lastly, the carrier lifetime is affected by the field strength. Particularly, near the breakdown, i.e. at the critical electrical field of $E \approx 3 \times 10^5 \, \text{V cm}^{-1}$, defect-assisted tunnelling becomes a significant effect. Defect-assisted tunnelling causes electron-hole pair generation and significantly influences the IV characteristics of reverse biased pn-junctions even before band-to-band tunnelling or avalanche carrier generation comes into effect. The effect is considered in Synopsys by the flag `ElectricField` in combination with a specific field-enhancement model. The `Hurkx` statement invokes the Hurkx model [131] which has proven to be quite successful for silicon particle detectors [132]. Still in the recombination statement, the avalanche process of charge carrier amplification in high electrical fields needs to be activated. The default avalanche model of van Overstraeten and de Man [133] is activated with the flag `Avalanche`.

6.1.3 Radiation damage in TCAD simulation

In Synopsys TCAD, bulk damage is simulated by the introduction of acceptor and donor type traps with certain cross-sections for electrons and holes as well as an introduction rate being the ratio of defect concentration to fluence for each trap. Furthermore, the energy level of each trap characterises significantly the effect on the sensor behaviour. The model is taken from Reference [134] and

contains two acceptor and one donor type traps. A summary of the properties of all traps is given in Table 6.1. This model is chosen amongst the several available models as it is being widely used in the HEP community. Even if the absolute differences between some available radiation damage models for TCAD simulations are significant, the comparison of the simulated relative performance of different pixel geometries are likely much less affected.

Table 6.1: Radiation damage model as presented in Reference [134] for p-type material up to fluences of 7×10^{15} n_{eq}/cm^2. The model is characterised by the energy level relative to the conduction (E_C) and valence (E_V) band, the trapping cross-sections for electrons and holes ($\sigma_{e/h}$) and the introduction rate (η) of each trap.

Type	Energy [eV]	σ_e [10^{-15} cm^{-2}]	σ_h [10^{-14} cm^{-2}]	η [n_{eq}^{-1}cm^{-1}]
Acceptor	E_C-0.42	1	1	1.613
Acceptor	E_C-0.46	7	7	0.9
Donor	E_V+0.36	323	3.23	0.9

Additionally, surface damage is also treated in TCAD simulations. Here, two physics effects are modelled to approximate the reality. The first effect is the introduction of a fixed oxide charge as a homogeneous sheet layer between oxide and silicon. Following the measurements of Reference [135], the oxide charge saturates after high ionising doses $O(10)$MGy at values of $2-3 \times 10^{12}$ cm^{-2}. Due to the spread arising from different production technologies, and given the possibly not completely saturated and already partially annealed oxide charge, the oxide charge in the simulation is fixed to 5×10^{10} cm^{-2} before irradiation [136] and to 1.5×10^{12} cm^{-2} for all irradiations [135]. In addition to the fixed oxide charge, interface traps are placed in accordance with Reference [134]. The properties of the traps can be found in Table 6.2. In this case, the energy level is not just a single level but a Gaussian distribution around the quoted level with a width σ. The trap concentration is given as a fraction of the fixed oxide charge N_{ox}.

Table 6.2: Properties of surface traps according to Reference [134]. Traps are characterised by their energy level as well as the Gaussian width around that energy level (σ). The concentration of traps is given as a percentage of the fixed oxide charge N_{ox}.

Type	Energy [eV]	σ [eV]	Concentration [%N_{ox}]
Acceptor	E_C-0.4	0.07	0.34
Acceptor	E_C-0.6	0.07	0.51
Donor	E_V+0.7	0.07	0.85

6.1.4 Charge collection efficiency

The `HeavyIon` module of Synopsys TCAD is used for simulations of the CCE. Despite the name, it can be tuned also for the general case of any charged particle passing through the material by choosing the amount of released charge carriers per traveled distance as well as the area within which this charge is distributed. All CCE results are obtained by simulating the impact of a MIP releasing 76 electron − hole pairs per µm in a 100 µm thick sensor (the value is higher for thicker

sensors [137]) with a Gaussian width of 1 μm. The current of the device is simulated from 5 ns before the impact to 30 ns after the impact. The collected charge is obtained by integrating the current pulse and subtracting the baseline leakage current extracted from the time before the impact.

6.2 Leakage current simulation software: RadDamSim

Irradiation induced defects in the atomic lattice of semiconductors change basic properties of pixel detectors used for tracking of high energy particles. An irradiation model (Hamburg model [54]) is implemented in a C++ program to predict the leakage current of a silicon sensor operated in a high radiation environment. This information is essential to ensure optimal working conditions. The program is called `RadDamSim` (Radiation damage simulation) and can calculate the mentioned properties of semiconductor sensors for a given radiation and temperature profile. The key motivation for the program is to implement simultaneous irradiation and annealing in arbitrary condition profiles. The employed luminosity-to-fluence conversion factor, also required for many other simulations, can be validated by comparing simulated and measured leakage currents.

6.2.1 Input file

The input file for `RadDamSim` needs to have the following format. It has to include three columns. The first column is filled with the time span covered by each step (in seconds). The second column contains the temperature of the sensor during the time step. The last column contains the (constant) dose rate in a given time step in units of $n_{eq}/(cm^2s)$. The dose rate and time span have to have integer format while the temperature can be of float type. The granularity is important for the accuracy of the simulation and a minimal granularity of four hours has proven to be sufficient. A high granularity will result in long computation times mainly driven by the numerical approach of the leakage current simulation in which every previous time step has to be recomputed when going to the next element of the profile.

The program is applied for the simulation of the leakage current of the present ATLAS pixel detector. In this case, the temperature data is measured using a negative temperature coefficient (NTC) sensor on top of each pixel detector module flex. The flex is attached to the sensor surface by epoxy glue. The readout chips are bump-bonded to the sensor and are the only part in thermal contact to the cooling pipe of the support structure. Therefore, the actual sensor temperature might vary from the recorded temperature depending on the thermal conductivity of the epoxy glue and the environmental temperature. In any case, the temperature difference is expected to be smaller than 2 °C which is used as a systematic uncertainty. The data is recorded by the detector control system (DCS) and retrieved by the DCS data viewer (DDV).

The fluence data is computed based on the official ATLAS luminosity measurements. The luminosity is transformed into a fluence using a conversion factor that is center-of-mass energy and detector layer dependent. This conversion factor is obtained by FLUKA simulations and verified by the comparison of leakage current simulation and data. More information about the input data that is used for the simulation can be found in Section 8.1.

6.2.2 Leakage current

The leakage current is found to be linearly increasing with fluence [54]

$$I_{\text{leak}}(t, T) = \alpha(t, T) \cdot \Phi_{\text{eq}} \cdot V \tag{6.9}$$

where Φ_{eq} is the total accumulated equivalent fluence ($\Phi_{\text{eq}} = \varphi_{\text{eq}} \cdot t$) and V is the detector volume. The physical explanation for this behaviour is that charge-carrier generation centers are created by NIEL processes (see Section 3.2 for more details). Those generation centers anneal with time but unlike in the case of the depletion voltage there is no reverse annealing. The annealing is handled in the α parameter which depends on time and temperature. A numerical approach is chosen in which the continuous irradiation is split up into discrete steps to simulate the behaviour. Each step is followed up in the following steps and treated independently from the following irradiations. The superposition of all individual steps yields the total change in leakage current. The leakage current density G_i for every time step t_k can be calculated from an exponential part

$$G_i^{\text{exp}} = \alpha_I \cdot \sum_{j=1}^{i} \Phi_{\text{eq},j} \cdot e^{-\sum_{k=j}^{i} t_k / \tau_I(T_k)} \tag{6.10}$$

and a logarithmic part

$$G_i^{\text{log}} = \sum_{j=1}^{i} \Phi_{\text{eq},j} \left[\alpha_0^* - \xi \cdot \ln \left(\sum_{k=j}^{i} t_k \Theta(T_k) \right) \right] \tag{6.11}$$

where $\Theta(T_k)$ is the time/temperature scaling function

$$\Theta(T) = e^{-\frac{E_I^*}{k_B} \left(\frac{1}{T} - \frac{1}{T_{\text{ref}}} \right)}. \tag{6.12}$$

There, α_0^*, ξ and E_I^* are constants (determined at $T_{\text{ref}} = 21\,°\text{C}$) and τ_I again exhibits an Arrhenius relation

$$\frac{1}{\tau_I} = k_{0I} \cdot e^{-\frac{E_I}{k_B T}}. \tag{6.13}$$

All constants are summarised in Table 6.3. The summations in G_i^{exp} and G_i^{log} run over all time steps up to the present time i and represents the superposition of individual irradiations in all time steps. Meanwhile, the summations in the logarithm and exponential run from the time j at which each

Table 6.3: Constants used for the simulation of the leakage current [54].

α_I	$[10^{-17}\,\text{A}\,\text{cm}^{-1}]$	1.23 ± 0.06
k_{0I}	$[10^{13}\,\text{s}^{-1}]$	$1.2^{+5.3}_{-1.0}$
E_I	$[\text{eV}]$	1.11 ± 0.05
α_0^*	$[10^{-17}\,\text{A}\,\text{cm}^{-1}]$	7.07
ξ	$[10^{-18}\,\text{A}\,\text{cm}^{-1}]$	3.07 ± 0.18
E_I^*	$[\text{eV}]$	1.30 ± 0.14

irradiation happened up to the present time i. The leakage current at each time-step i can now be calculated as

$$I_{\text{leak}}^i = \left(G_i^{\text{exp}} + G_i^{\text{log}} \right) \cdot V. \tag{6.14}$$

The leakage current can be transformed from the general reference temperature of $21\,°\text{C}$ to any given temperature T according to

$$I_{\text{leak}}(T) = I_{\text{leak}}(T_{\text{ref}}) \cdot \left(\frac{T}{T_{\text{ref}}} \right)^2 e^{-\frac{E_g^*}{2k_B}\left(\frac{1}{T} - \frac{1}{T_{\text{ref}}} \right)} \tag{6.15}$$

where E_g^* is the effective band gap of the (irradiated) semiconductor ($E_g^* = 1.21\,\text{eV}$ used [138]).

RESULTS FROM TCAD SIMULATIONS

This chapter presents the results obtained using the Synopsys TCAD simulation suite introduced in Chapter 6. Part of the results are published in References [139, 140]. The chapter is split into three sections, first introducing the structures that are investigated (Section 7.1), second covering the effort to optimise the production process of sensors (Section 7.2) and third investigating and optimising the performance of small pixel cells (Section 7.3). The investigation of new pixel cell sizes is necessary because the ITk detector requires smaller pixel cells of $50 \times 50\,\mu\mathrm{m}^2$. The current ATLAS pixel detector employs pixel cells of $50 \times 250\,\mu\mathrm{m}^2$ and $50 \times 400\,\mu\mathrm{m}^2$.

7.1 Structures under investigation

Several structures are investigated during the work for this thesis. Their detailed layout is described in this section. The pixel cell size of $50 \times 50\,\mu\mathrm{m}^2$ is common to all of them.

- **Structure A** is used for the simulation of the secondary implant (see Section 7.2). It is a two dimensional structure with one full and two half pixels, generated in SDE. Structure A is illustrated in Figure 7.1 with the x-axis pointing along the pixel and the y-axis pointing in the sensor depth. The doping profiles are imported from a full process simulation with sProcess. The structure consists of a $100\,\mu\mathrm{m}$ thick silicon bulk with a Boron doping concentration of $10^{12}\,\mathrm{cm}^{-3}$, being a typical value in HEP applications. A $220\,\mathrm{nm}$ thick oxide layer is grown on the top surface. The oxide is etched down to $130\,\mathrm{nm}$ in the area of the implant. The area of decreased thickness and thereby the pixel implant has a width of $35\,\mu\mathrm{m}$. This is the area where the primary n^+ implantation is realised with a dose of $1 \times 10^{15}\,\mathrm{cm}^{-2}$ at an energy of $100\,\mathrm{keV}$.

  The secondary n^+ implant uses varying doses and implantation angles at a fixed energy of $350\,\mathrm{keV}$. The implantations are annealed at nearly $1000\,^{\circ}\mathrm{C}$ for three hours (*drive-in*). Within the area of reduced oxide thickness, there is a $10\,\mu\mathrm{m}$ wide gap to establish an electrical contact from the top to the implant using an electrode. The electrode extends $2.5\,\mu\mathrm{m}$ beyond the area of reduced oxide thickness and mimics the aluminium of the actual process (see Section 3.3). The p-spray layer is a homogeneous p-doping layer on the surface of the sensor. A dose of
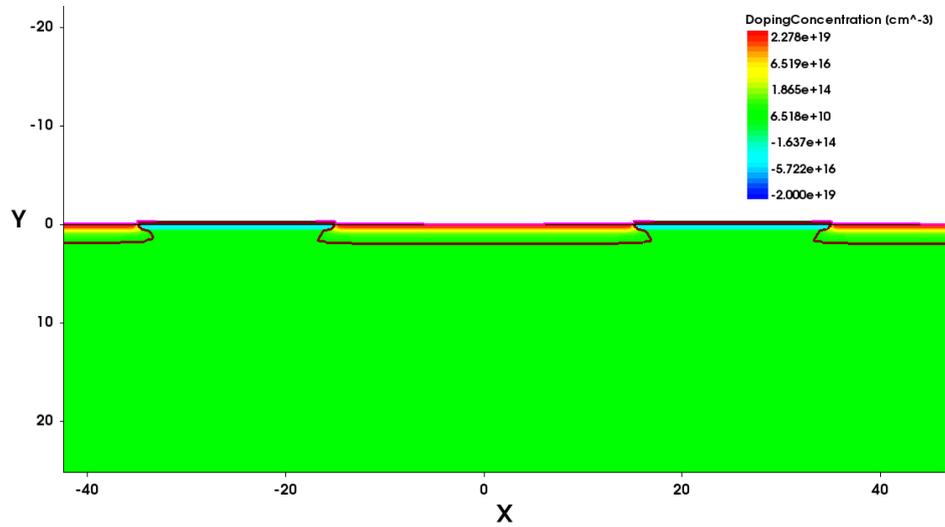
(a)
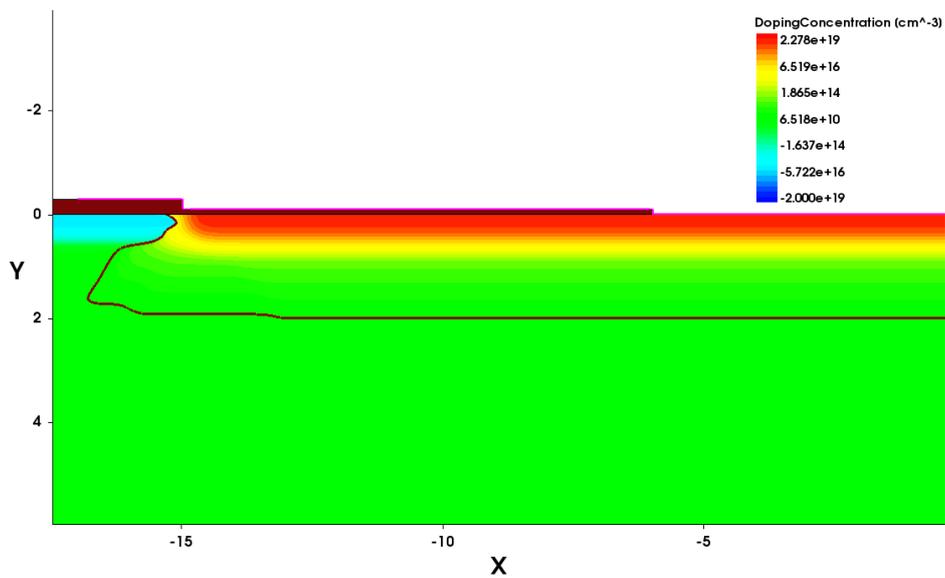
(b)

Figure 7.1: Illustrations from `sVisual` of the geometrical setup and doping concentration for structure A. The parameters for the shown secondary implant are a dose of $4 \times 10^{12}\,\mathrm{cm}^{-2}$ and an angle of $55°$. The wine-red line beneath the pixel implant displays the position of the junction. An overview of the structure with its one full and two half pixels in two dimensions is given in (a). The extension of the secondary implant below the p-spray is indicated in (b).
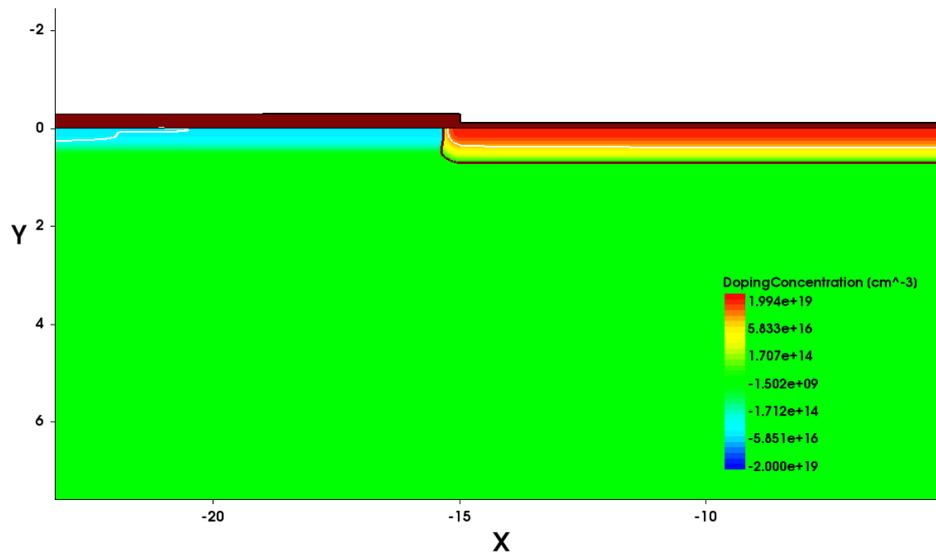
(a)

Figure 7.2: Illustrations from `sVisual` of the geometrical setup and doping concentration for structure B. The structure is generated by SDE and uses analytical doping profiles. Similar to structure A, one full and two half pixels are used in a two dimensional simulation. The analytical doping profile of p-spray and pixel implant is shown in detail. The wine-red line beneath the pixel implant displays the position of the junction.

$1 \times 10^{12} \, \text{cm}^{-2}$ is used for the p-spray implantation. The backside of the device consists of a p^+ implant with a high concentration of $2 \times 10^{19} \, \text{cm}^{-3}$ and another electrical contact over the full surface.

- **Structure B** is used for the simulation of the influence of the aluminium overhang on the breakdown voltage. It is generated by SDE and employs analytical doping profiles for all dopings. Figure 7.2 illustrates the composition of structure B. As in structure A, the structure consists of one full and two half pixels in two dimensions with a $100 \, \mu\text{m}$ thick bulk. The bulk has a Boron doping concentration of $10^{12} \, \text{cm}^{-3}$. The oxide has a thickness of $300 \, \text{nm}$ while in the reduced area of the pixel implant it only has $100 \, \text{nm}$. The contact opening in the oxide is $12 \, \mu\text{m}$ wide and the pixel implant is $30 \, \mu\text{m}$ wide.

  The analytical doping profile of the pixel implant has a peak concentration of $2 \times 10^{19} \, \text{cm}^{-3}$ at $y = 0$ and is extending to $y = 0.7 \, \mu\text{m}$ into the bulk. This means that at a depth of $0.7 \, \mu\text{m}$ the Gaussian doping profile has fallen off to a concentration equivalent to the bulk doping. The p-spray has a peak concentration of $6 \times 10^{16} \, \text{cm}^{-3}$ extending $0.5 \, \mu\text{m}$ into the bulk. The p-type backside implant has the same profile as the front side pixel implant but covers the whole backside. An electrode covers each implant and is used to contact the device.

- **Structure C** is the most sophisticated structure. It includes four quarter pixels in a three dimensional model which is necessary for the accurate description of charge collection and capacitance by the simulation. Structure C is illustrated in Figure 7.3. It also has the most accurate description of the passivation and isolation layers on top of the bulk material,

mimicking the real production process. On top of the p-type silicon bulk (doping concentration of 10^{12} cm^{-3}), a 220 nm thick oxide layer is deposited, which is reduced to 130 nm in the area of the pixel implant. The oxide charge at the interface of oxide and silicon bulk is a sheet layer with a concentration of 5×10^{10} cm^{-2} (1.5×10^{12} cm^{-2}) before (after) irradiation [135, 136]. A 30 nm thick nitride layer is added on top of the oxide, followed by a 150 nm thick LTO layer. All isolation layers have a $5 \times 5 \, \mu m^2$ large hole in the center of each pixel implant for the electrical contact to the pixel.

An electrode covers each pixel to simulate the aluminium pad in the real process. For the simulation of the IV characteristics in Figure 7.11 and the extracted breakdown voltages in Figure 7.12, the square electrode constantly overhangs the square implant by 2 μm to disentangle the effects of implant width and metal overhang. For the remaining investigations, an overhang of 6 μm is chosen for an implant width of 10 μm, 4 μm overhang for 20 μm and 30 μm implant width and 2.5 μm for 40 μm implant width. All dopings are defined as analytical profiles. The p-spray layer has a peak concentration of 6×10^{16} cm^{-3} extending 0.6 μm into the bulk. The doping profile of the n$^+$ pixel implant has a peak concentration of 2.5×10^{19} cm^{-3} extending 0.8 μm into the bulk. The p$^+$ implant covering the entire backside differs from the top n$^+$ implant with a peak concentration of 1.5×10^{19} cm^{-3} extending to a depth of 1.2 μm.

- **Structure D** is used for the investigation of p-spray properties. It is shown in Figure 7.4. Structure D is a copy of structure C, except that it only includes one quarter pixel instead of four. This is sufficient since here only the interplay of pixel implant and p-spray is investigated.

7.2   Sensor production process optimisation

This section contains the results of different process optimisation studies. The main purpose of these studies is to increase the breakdown voltage, in particular before irradiation. As the depletion voltages of the thin sensors under investigation are usually very low before irradiation, in the order of a few 10 V, the typically achieved breakdown voltages between 100−200 V are not critical in the first place. However, the increase of breakdown voltage is desired as it leaves a larger margin for breakdown voltage reductions during the module assembly steps after the sensor fabrication is completed.

The mechanism of the breakdown needs to be understood to improve the breakdown behaviour of a sensor. As explained in Section 3.1, the breakdown of a sensor occurs if the electrical field within the sensor exceeds a critical value. The electrical field depends on the gradient of the doping concentration and the potential difference between anode and cathode. Thus, the critical region is the transition between n$^+$ pixel implant and p-spray layer, as it contains the largest doping concentration gradient (see Figure 7.3b).

However, the simulation considers an infinite pixel matrix by applying periodic boundary conditions to the simulated symmetric objects. Therefore, edge effects of actual sensors are not reproduced although they represent the dominating breakdown mechanism for not-irradiated structures with
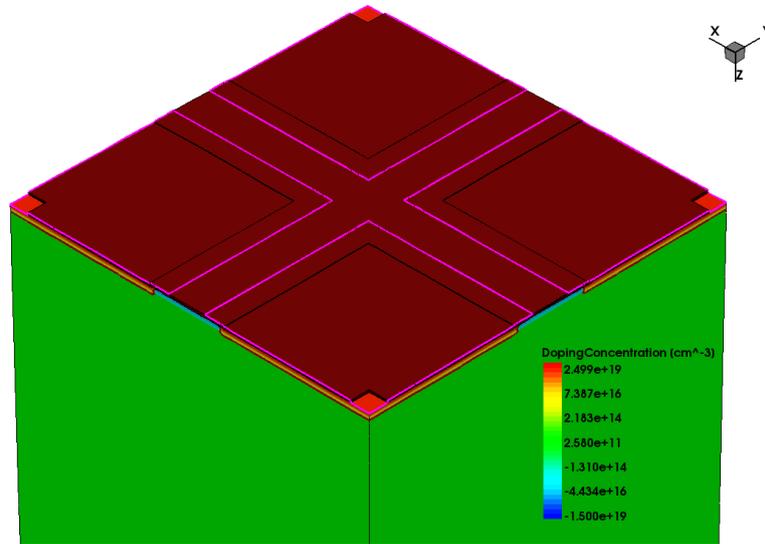
(a)
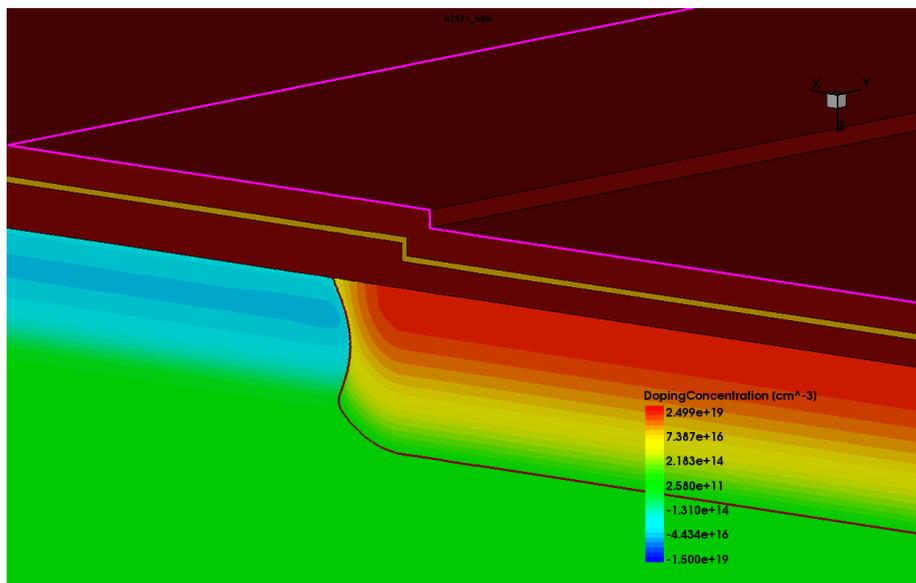
(b)

Figure 7.3: Illustrations from `sVisual` of the geometrical setup and doping concentration for structure C. The structure is generated by SDE and uses analytical doping profiles. Four quarter pixels are simulated in three dimensions (a) and the detailed composition of the passivation layers is displayed in (b). The wine-red line beneath the pixel implant displays the position of the junction.
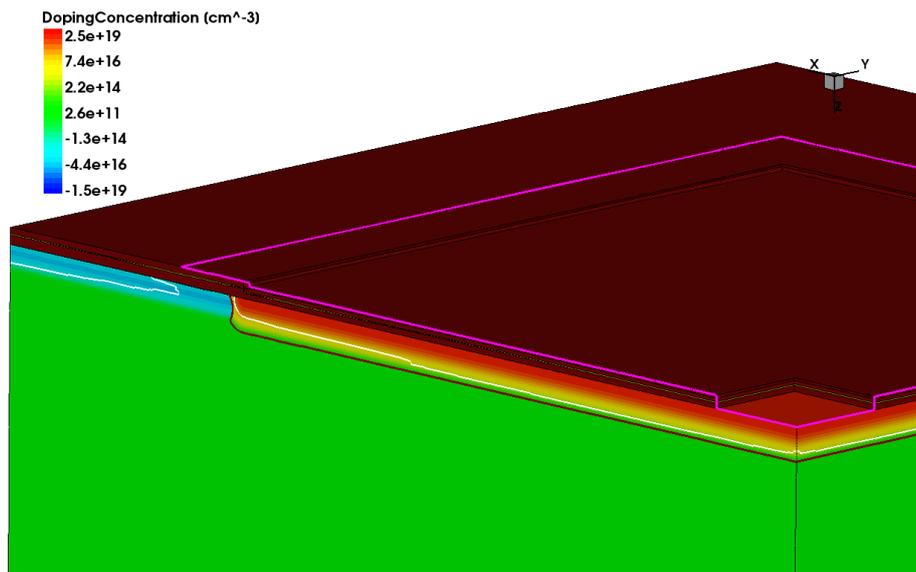
Figure 7.4: Illustrations from `sVisual` of the geometrical setup and doping concentration for structure D. The structure is generated by SDE and uses analytical doping profiles. It is identical to structure C but uses only one quarter pixel. The wine-red line beneath the pixel implant displays the position of the junction.

narrow GR structures as they are required for the $100\,\mu m$ thin sensors of the ITk. The reason to still optimise the breakdown behaviour of the pixel cells themselves is twofold. First, the GR structure can be optimised separately to circumvent this bottleneck. The pixel matrix itself then has to be working in the best possible way. Secondly, after irradiation, real sensors typically do not break anymore at the GR structure but again in the pixel matrix. Consequently, the optimisation of the breakdown behaviour of the pixel cell itself is important to ensure reliable operation of the devices at high bias voltages after irradiation. Nonetheless, the results obtained from the simulations are not directly comparable to actual sensors, but have to be compared to special test structures with generous GR structures like they are implemented in the ongoing SOI-5 sensor production.

7.2.1   Secondary implant

The first approach to increase the breakdown voltage is to smooth out the gradient of the doping concentration by adding a second implant with modified parameters to the main n^+ pixel implant. These investigations are carried out on structure A. The first implantation is shot with an energy of $100\,keV$, a dose of $1 \times 10^{15}\,cm^{-2}$ and at an angle of $7°$ with respect to the direction perpendicular to the sensor surface, the standard orientation used by HLL. This prevents dopants to slip through channels in the crystal lattice known as channeling. Instead, the secondary implant with doses of $1 - 8 \times 10^{12}\,cm^{-2}$ is shot under a shallow angle between $10°$ and $70°$ with a higher energy of $350\,keV$. Consequently, the anticipated smaller gradient of the doping concentration in the p-spray and pixel implant interface region is achieved.

As an example, the doping concentration is shown in Figure 7.5 for the standard process (black squares) and for the modified process based on a secondary implant with a dose of $4 \times 10^{12}\,cm^{-2}$ at an angle of $55°$ (blue triangles). The profiles are recorded in a depth of $200\,nm$ and $600\,nm$. In a

(a) 200 nm depth

(b) 600 nm depth
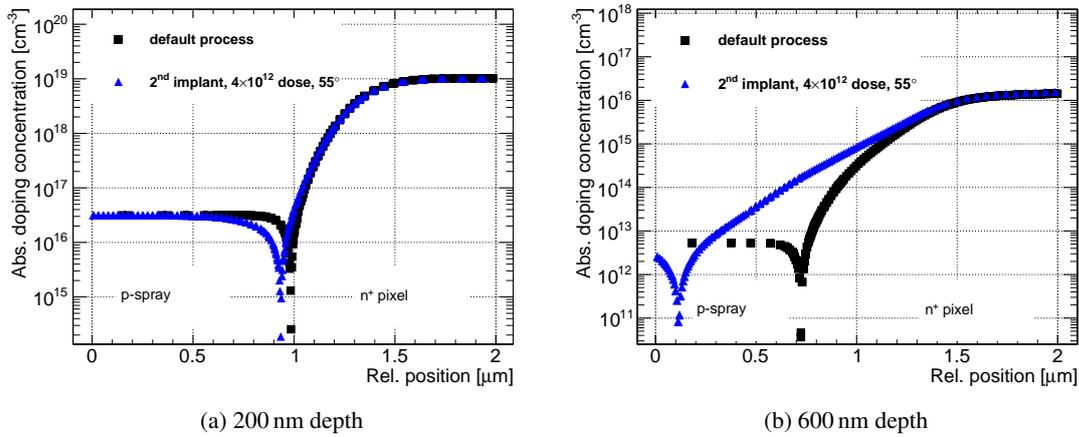
Figure 7.5: Horizontal profile of the absolute doping concentration of default implantation vs default implantation plus second implantation performed at an angle of $55°$ and a dose of $4 \times 10^{12}\,\mathrm{cm^{-2}}$ after drive-in. The horizontal profile is recorded at the border from p-spray region (at $0\,\mu m$ relative position) to n^+ implant (at $2\,\mu m$ relative position) in a depth of 200 nm (a) and 600 nm (b).
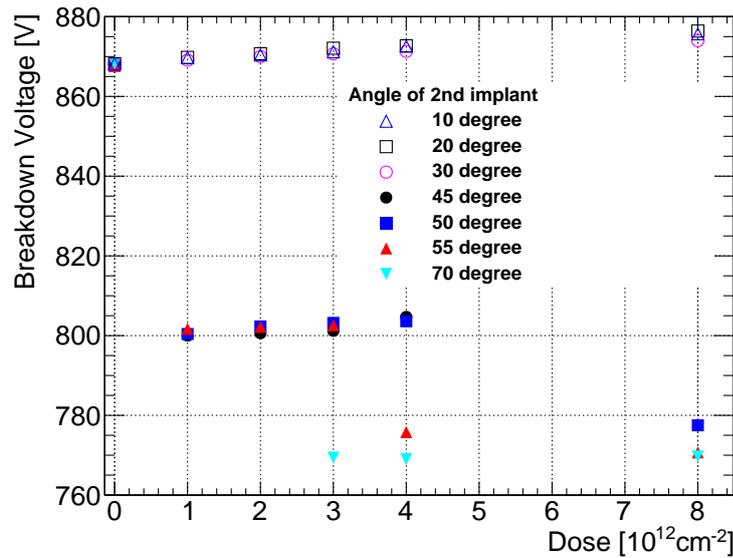
Figure 7.6: Breakdown voltage as a function of the implanted secondary dose for different implantation angles. All points consistently overlap for zero dose of the secondary implant.

depth of 200 nm, the modified process exhibits a smaller gradient in the doping profile compared to the default process. This difference gets even larger for the horizontal profile in a depth of 600 nm. The effect is also more relevant for larger implantation angles and higher doses as the secondary implant extends further under the edge of the oxide layer. After a full process simulation of the device, the breakdown voltage is extracted by ramping up the voltage until a predefined current limit (100 nA) is reached. The limit is chosen to be several orders of magnitude higher than the current before breakdown, hence being a reliable breakdown criterion. The breakdown voltage as a function of the implanted secondary dose for different implantation angles is presented in Figure 7.6. A small

improvement in breakdown voltage from 868 V in the default process without secondary implant to about 875 V is observed for small implantation angles of 10°, 20° and 30° and the highest dose of $8 \times 10^{12}\,\mathrm{cm^{-2}}$. However, for larger implantation angles the breakdown voltage even decreases with respect to the default process. This behaviour is further investigated since the original assumption is an improvement for larger implantation angles. The electrical field is chosen as observable to evaluate each structure.

Figure 7.7 shows the electrical field at a bias voltage of 720 V for the default process (a) and a secondary dose of $4 \times 10^{12}\,\mathrm{cm^{-2}}$ implanted under an angle of 55° (b). This bias voltage is approximately 50−100 V lower than the breakdown voltage. The peak of the electrical field is actually reduced for the modified process. This is still the case when the device with the modified process starts to break around 780 V. Figure 7.7(c,d) shows the hole-current density for the modified and the default process. The device with the modified process exhibits a huge current flowing from the backside of the device through a channel in between the pixel implants to the top side. From there, the carriers are accelerated through the high field towards the pixel implant. The secondary implants with implantation angles larger than 30° apparently create a second junction in vertical direction from the pixel implant to the p-spray, since in the case of the modified process the implant extends under the p-spray layer. Despite achieving lower electrical fields in the default junction region using the modified process, the breakdown occurs earlier due to the second junction and the extension of the pixel implant under the p-spray layer.

Since the possible improvements demonstrated by the small angle secondary implantations are rather small, it was not tried to include this modified process in the production runs for n-in-p pixel sensors.

7.2.2   P-spray concentration

Another way of increasing the breakdown voltage is simulated by decreasing the peak concentration of the p-spray implant. For this, the p-spray depth is varied as a second variable. The depth is defined as the point at which the Gaussian p-spray concentration profile reaches the nominal bulk doping concentration. This investigation requires a three dimensional simulation to accurately describe the potential of the p-spray. Therefore, structure D is employed. Figure 7.8 presents the breakdown voltage as a function of the p-spray peak concentration for various p-spray depths. The breakdown voltage clearly increases for decreasing p-spray peak concentrations. The increase in breakdown voltage over the investigated dose range from $1-6 \times 10^{16}\,\mathrm{cm^{-3}}$ is about 450−550 V. Also shallower p-spray implants are beneficial with respect to the breakdown voltage. The difference in breakdown voltage between $0.6\,\mu\mathrm{m}$ and $1.2\,\mu\mathrm{m}$ is about 150−250 V.

Some wafers of the SOI-4 production have a reduced p-spray dose to verify the simulated increase of breakdown voltage with decreasing p-spray dose. For three wafers, the implanted dose is reduced from $1.4 \times 10^{12}\,\mathrm{cm^{-2}}$ to $1 \times 10^{12}\,\mathrm{cm^{-2}}$ which should result in an equally lower, although unknown, peak concentration of the p-spray. Beyond that, the low dose wafers received an additional annealing step in which the p-spray doping is driven into the bulk. This causes a significant shift of the p-spray peak position deeper into the bulk, while simultaneously further decreasing the peak height.

(a)

(b)

(c)

(d)
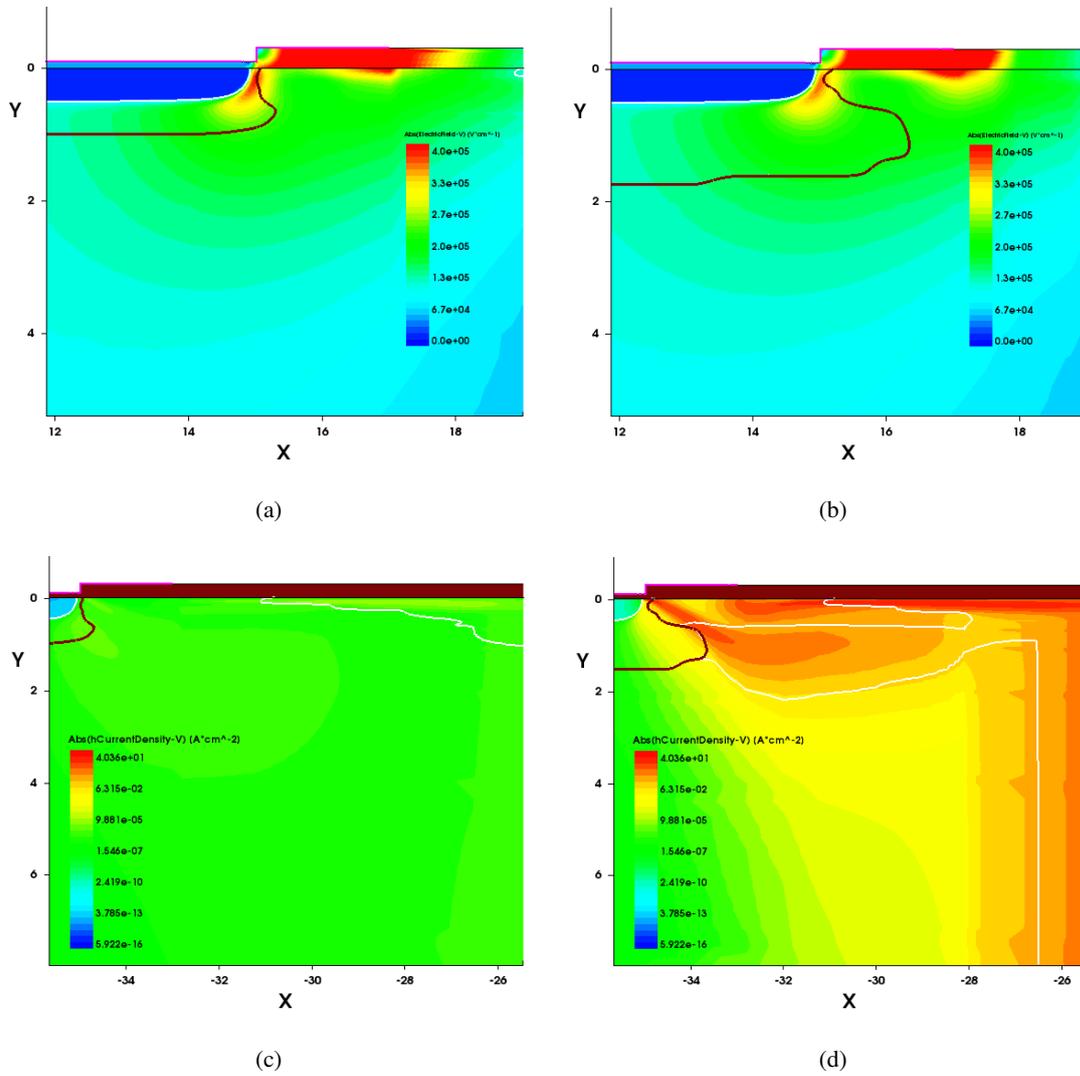
Figure 7.7: Absolute value of the electrical field for a bias voltage of 720 V. The default process without secondary implant is shown in (a), while the modified process with a secondary implant dose of 4×10^{12} cm^{-2} implanted under an angle of 55° is shown in (b). At a higher voltage of 780 V, the hole-current density in the case of the default process is shown in (c) and does not exhibit breakdown behaviour while the sensor employing the secondary dose breaks as it can be seen in (d).
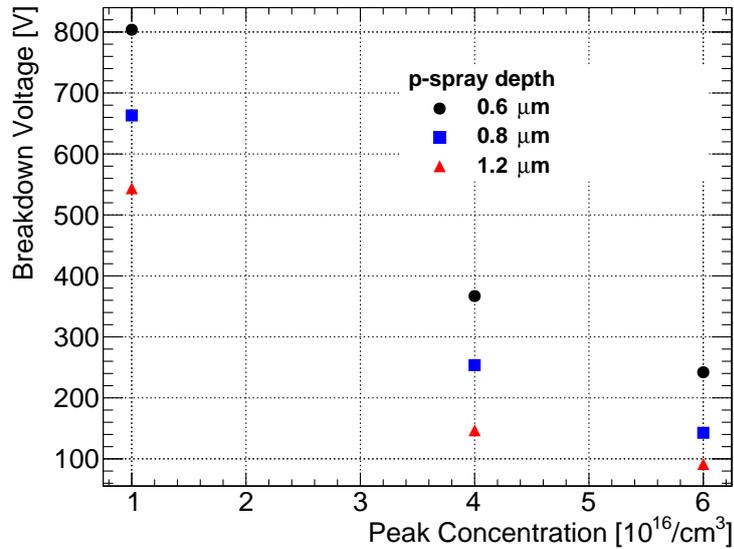
Figure 7.8: Simulated breakdown voltage as a function of the p-spray peak concentration for different p-spray depths.
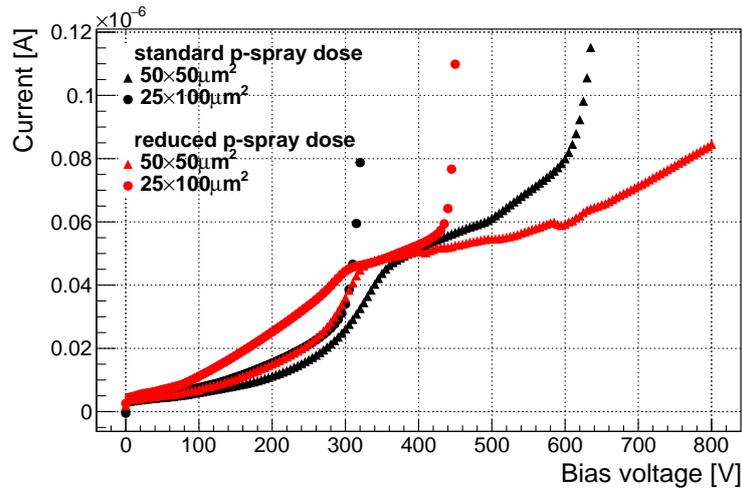
Figure 7.9: Measured IV curves of sensors of the SOI-4 production. Red colour indicates sensors from wafers produced with the reduced p-spray dose while black corresponds to the standard dose. Triangles show $50 \times 50\,\mu m^2$ pixels while circles indicate $25 \times 100\,\mu m^2$ pixels. Vertical bars indicate the RMS of ten consecutive current measurements per voltage and are smaller than the symbol size.

Two identical sensors of each pixel cell geometry ($50 \times 50\,\mu m^2$ and $25 \times 100\,\mu m^2$) are electrically characterised in an IV measurement for one wafer with the default dose and one wafer with the reduced dose. The result is shown in Figure 7.9. The modified process causes an increase of breakdown voltage of about 350 V for both structures. Since the additional annealing step causes an unknown decrease of peak height, the exact values are difficult to compare, but the qualitative effect and the order of magnitude is reproduced by the simulation.
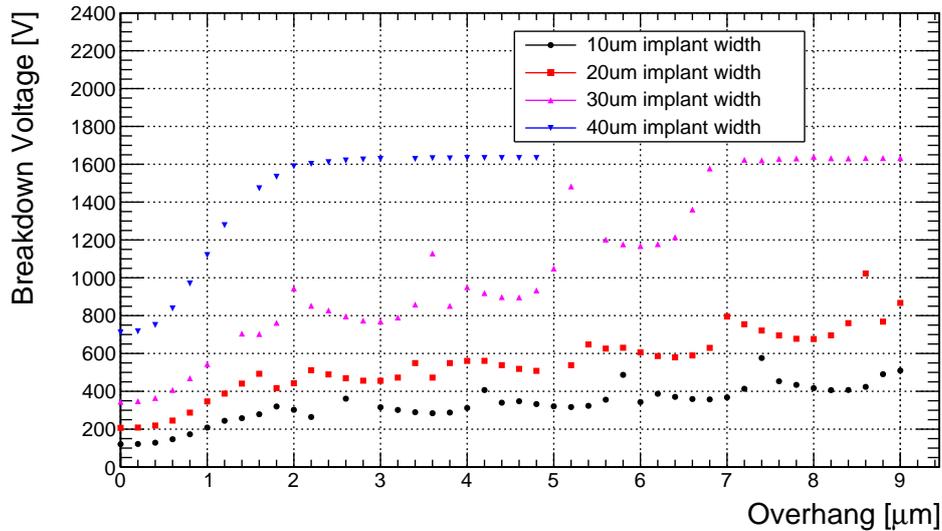
Figure 7.10: Simulated breakdown voltage as a function of the aluminium overhang for different implant widths. The overhang for the 40 μm wide implant has to be smaller than 5 μm given the pixel pitch of 50 μm.

7.2.3 Metal overhang

Lastly, the effect of the metal overhang is investigated. The metal overhang is the part of the aluminium pad on top of the pixel implant which extends beyond the implant. In the simulation, an overhang of 0 μm corresponds to an aluminium pad exactly as large as the opening for the pixel implant itself. In practice, however, the lateral diffusion of the n^+ pixel implantation during the drive-in would extend the implantation beyond the boundaries of the aluminium.

The ground potential of the readout chip is connected through the bump-bond to the aluminium pad. If the aluminium pad is larger than the implant below, the ground potential of the pad influences the properties in the adjacent p-spray layer. In this case, the potential within the p-spray layer is locally forced to be closer to the ground potential of the pad. This happens primarily in the critical region where n^+ pixel implant meets p-spray layer. Due to the decreased potential difference between p-spray layer and n^+ implant, the electrical field is decreased at the same bias voltage and, consequently, the breakdown voltage is increased.

Figure 7.10 shows the breakdown voltage as a function of the aluminium overhang for different implant widths. For all implant widths, the breakdown voltage strongly increases with increasing metal overhang. The overhang for the 40 μm wide implant has to be smaller than 5 μm given the pixel pitch of 50 μm. Some points are not computed as the simulator was not able to find a converging solution of the Poisson equation for those points.

When considering the largest implant width of 40 μm, the breakdown voltage increases from about 700 V for no metal overhang to more than 1600 V for a metal overhang larger than 2 μm. All other implant widths exhibit peaks every 1.5−2 μm, which are due to numerical issues in the computation of the avalanche processes. The numerical problems are related to the mesh with which the structure is approximated.

Besides the beneficial influence of larger metal overhangs on the breakdown voltage, they also have a detrimental effect. The inter-pixel capacitance is dominated by the capacitance between the n^+ implants, but very large metal overhangs with small gaps in between can have a significant influence. As demonstrated in Section 7.3.2 below, the capacitance of the electrode to the backside can be neglected after depletion. Furthermore, there is a minimal distance between aluminium pads defined by the manufacturer which can not be violated. Therefore, when looking for the optimal metal overhang, not only the positive influence on the breakdown voltage but also negative effects as an increased inter-pixel capacitance have to be considered. Especially the plateau that is clearly visible for the 40 µm wide implant in Figure 7.10 indicates that increasing the overhang to the maximum does not necessarily result in an increased breakdown voltage. As a consequence, the suggested optimal metal overhang is 2 µm for an implant width of 40 µm, close to the usually employed implant width.

7.3   Simulation of RD53A compatible pixel cells

This section covers the layout optimisation of small pixel cells compatible with the RD53A readout chip. The variable to be optimised is the width of the n^+ pixel implant. It is evaluated based on three observables which are the breakdown voltage, the capacitance of the implant and the charge collection properties before and after irradiation. All simulations in this section are carried out on structure C that implements four quarter pixels in three dimensions.

7.3.1   Breakdown voltage

The pixel itself is grounded while the sensor backside is ramped to a negative potential to determine the breakdown voltage. The resulting IV curves for different implant widths and a constant metal overhang of 2 µm are displayed in Figure 7.11. After depletion, the leakage current keeps slowly increasing as an effect of the Hurkx trap-assisted tunneling. Once the electrical field strength exceeds the limit for avalanche processes, the leakage current exponentially increases and the sensor goes into breakdown.

The breakdown voltage for a given implant width is determined by a fixed current limit of 0.1 nA. For better visibility, the extracted breakdown voltage as a function of implant width is summarised in Figure 7.12 for two sensor thicknesses of 100 µm and 150 µm. For all implant widths, the breakdown voltage of the thicker sensor is higher. The reason is that the voltage difference between front- and backside drops off across a larger distance resulting in lower electrical field strengths. In the case of the thicker sensor, the distance from the p-spray layer to the backside is increased and, thus, the potential of the p-spray layer is closer to the potential of the n^+ pixel implant creating less tension between the two at a given backside potential. Consequently, the breakdown voltage increases with increasing thickness. Also, the breakdown voltage increases exponentially with the implant width. This is due to the fact that the larger implants shield the p-spray layer from the backside potential more efficiently, keeping the p-spray potential closer to the n^+ pixel implant potential. Again, with a smaller potential difference between the p-spray layer and the n^+ pixel implant at a given bias
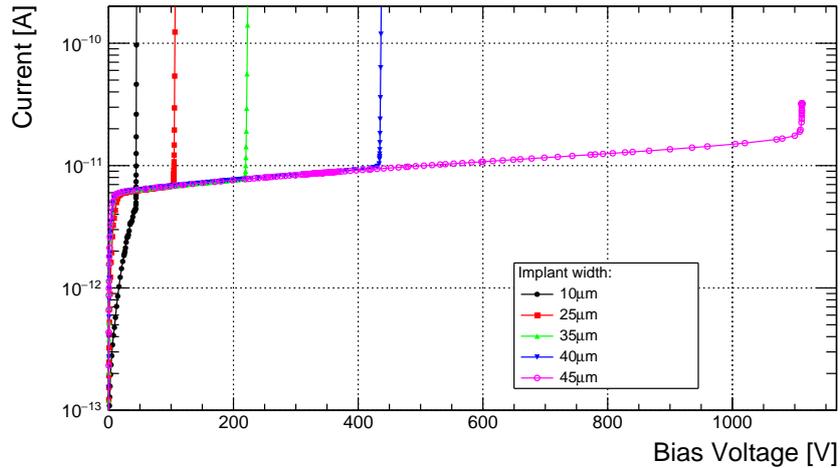
Figure 7.11: Simulated IV curves for different implant widths. The sensor thickness is 100 μm and the metal overhang is 2 μm.
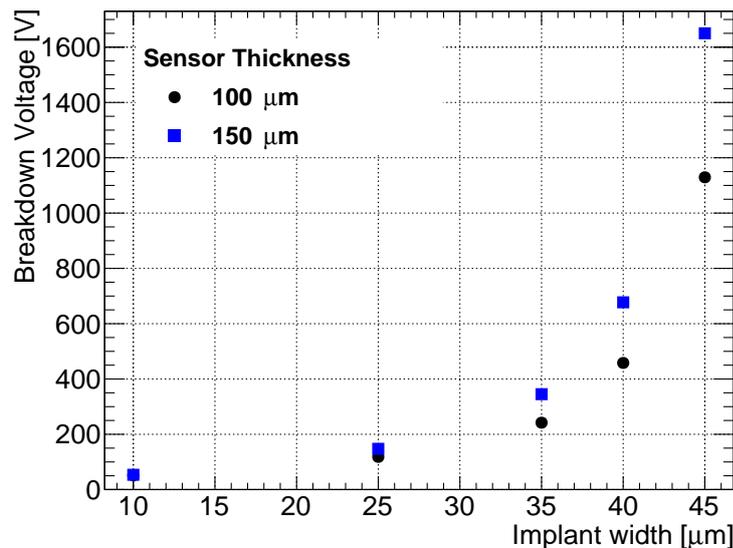
Figure 7.12: Breakdown voltage as a function of implant width for sensor thicknesses of 100 μm and 150 μm.

voltage, the breakdown voltage is increased with increasing n^+ pixel implant width. The negative impact of large implants is discussed next.

7.3.2 Capacitance

The input capacitance seen by the pre-amplifier of the readout chip is the main source of the electronic noise as typical leakage currents are low and the shaping times are short. Operation of the ITk pixel detector at very low thresholds of 500−700 e will be essential to maintain the required hit efficiency after the high fluences expected at the HL-LHC. Consequently, the noise has to be as low as possible. Therefore, the capacitance of the implant is one important parameter in the

(a)   (b)
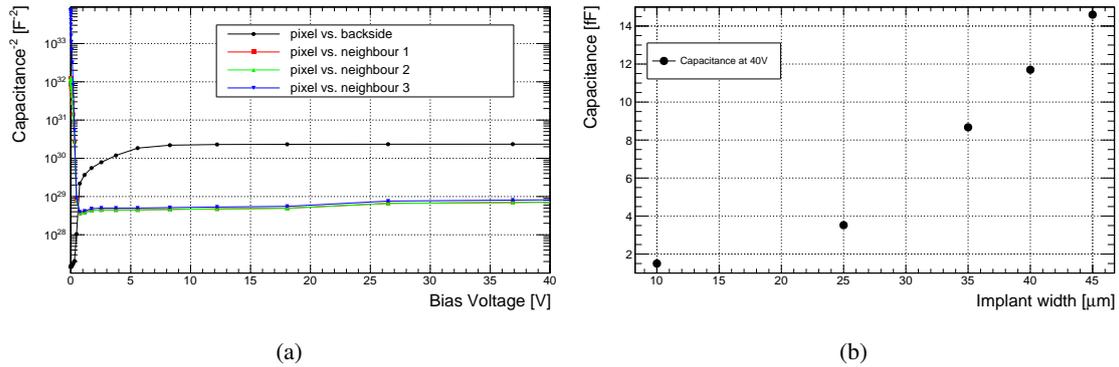
Figure 7.13: Investigation of the pixel implant capacitance for a sensor with a thickness of $100\,\mu m$. Inverse capacitance squared of the different capacitance contributions as functions of the bias voltage for a quarter pixel with an implant width of $40\,\mu m$ at a frequency of $10\,kHz$ (a). The contributions of neighbours 2 and 3 are identical such that their points overlap. The total capacitance of the quarter pixel is shown as a function of the implant width for a bias voltage of $40\,V$ (b).

overall optimisation of the implant width. The bias voltage of the backside is ramped to negative potential while applying an AC signal with a frequency of $10\,kHz$ to the n^+ implant to determine the capacitance in the simulation. The influence of the AC signal on the neighbouring pixels as well as on the backside itself is monitored while raising the bias voltage. From this, the different contributions to the capacitance can be extracted as shown in Figure 7.13a.

The capacitance of the quarter pixel to the direct neighbours is equal and is shown in overlapping red squares and green pyramids (neighbours 1 and 2), the capacitance to the diagonal neighbour is a bit smaller and is shown in blue triangles (neighbour 3) while the contribution from the backside is shown in black circles. Once the sensor is depleted, the contribution originating from the inter-pixel capacitance outweighs the contribution from the backside by far. The sum of all contributions at a fixed bias voltage of $40\,V$ is shown for different implant widths for the quarter pixel in Figure 7.13b. The capacitance is increasing with increasing implant width. For example, an implant width of $35\,\mu m$ results in a total pixel capacitance of $36\,fF$, thus, being well below the $50\,fF$ limit up to which the minimal operable threshold of the RD53A readout chip is guaranteed. Instead, the $40\,\mu m$ wide implant already has a total pixel capacitance of about $48\,fF$, marking the operational limit.

7.3.3   Charge collection efficiency

The influence of the pixel implant width on the CCE is investigated as another important benchmark. The CCE is calculated as the fraction of the charge that is being collected by the largest implant ($40\,\mu m$) at $40\,V$ before irradiation (further methodology of CCE measurements can be found in Section 6.1.4). Fluences of 0, 1, 3 and $5 \times 10^{15}\,n_{eq}/cm^2$ are investigated. Bulk and surface traps are implemented according to the Perugia model [134]. The collected charge is evaluated for a fixed impact position ($10\,\mu m/10\,\mu m \equiv d_{rel} = 40\,\%$, with 0/0 being the center of the four pixel region) while varying the bias voltage. From this, a voltage before and after saturation of collected charge is chosen. The spatial dependence of the collected charge is studied at the previously determined

(a) not-irradiated

(b) 1×10^{15} n$_{\text{eq}}$/cm^2

(c) 3×10^{15} n$_{\text{eq}}$/cm^2
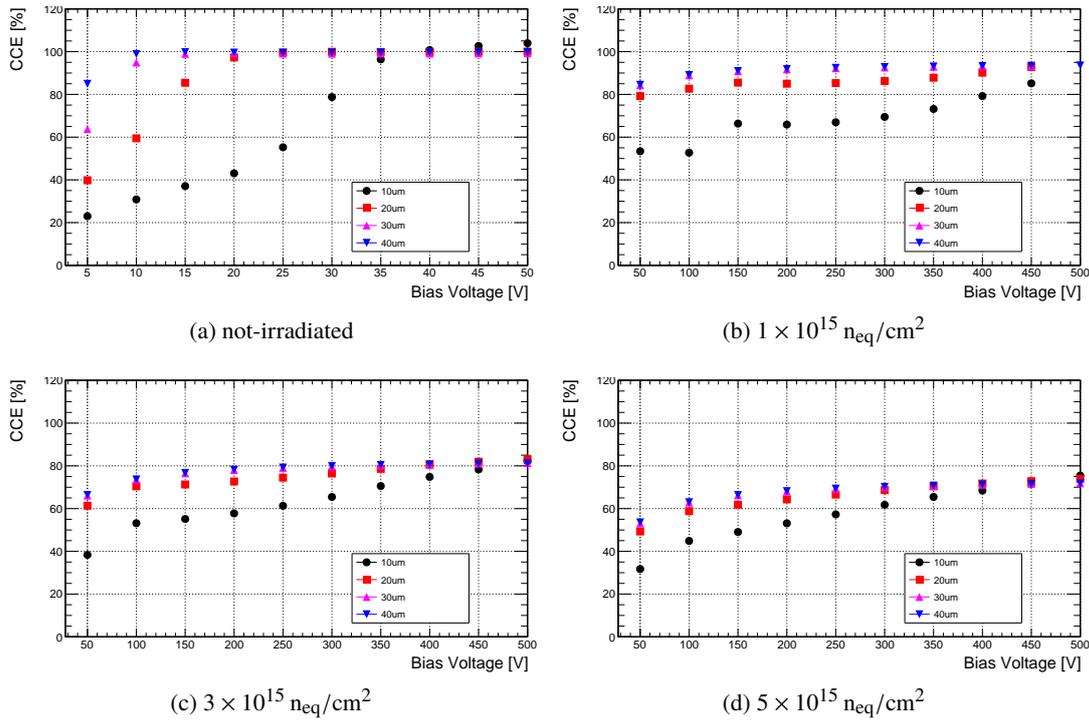
(d) 5×10^{15} n$_{\text{eq}}$/cm^2

Figure 7.14: Simulated charge collection efficiency (CCE) for different implant widths as a function of the bias voltage for fluences of 0 (a), 1 (b), 3 (c) and 5×10^{15} n$_{\text{eq}}$/cm^2 (d). The impact position corresponds to $d_{\text{rel}} = 40\,\%$.

voltages by changing the penetration point across the diagonal of the collecting pixel. The position on the diagonal is characterised by the relative position d_{rel} as a percentage of the distance from the four-pixel center ($d_{\text{rel}} = 0\,\%$) to the center of the pixel under investigation ($d_{\text{rel}} = 100\,\%$).

Influence of the bias voltage

The results of the CCE simulation at a fixed position are shown in Figure 7.14. Before irradiation, the only difference between different implant widths is the bias voltage at which the full charge is collected, as shown in Figure 7.14a. Once the sensor is depleted, all deposited charge is collected in all cases. Noteworthy, not much difference can be detected between implant widths of 30−40 µm while there is a significant difference to the 20−10 µm wide implants. After irradiation, across all fluences, the picture is still very coherent. No significant difference is observed between the 30 µm and 40 µm wide implants. In both cases, the depletion occurs between 200 − 300 V for all irradiations. The 20 µm wide implant shows a slightly worse CCE. Before depletion, around 10 % less charge is collected and 300 − 400 V are necessary to deplete the sensor and collect as much charge as the larger implants. Again, a serious CCE reduction is found in the 10 µm wide implant. Up to 40 − 50 % less charge is collected before depletion which only occurs beyond 450 − 500 V.

Influence of the impact position

From the simulation at a fixed position before irradiation, 10 V and 40 V are chosen to represent the situation before and after depletion. At those voltages, the CCE is evaluated as a function of the relative impact position across the diagonal of one pixel implant. Figure 7.15a reveals a big difference in CCE at 10 V. The implants with widths of 30 µm and 40 µm perform nearly equally good across the whole diagonal of the implant. Meanwhile, the 20 µm wide implant is only able to collect about 60 % of the charge of the larger implants across the diagonal. Furthermore, only about 40 % of the charge is collected in the case of the 10 µm wide implant. Consistently amongst the different implant widths, the collected charge decreases towards the four-pixel center ($d_{rel} = 0 \%$) where charge sharing is largest. The collected charge is reduced when injecting into the center of the pixel under investigation located in the corner of the structure at $d_{rel} = 100 \%$ as the mesh boundary is reached and part of the deposited charge is lost across the boundary.

The CCE improves drastically for all implant widths when increasing the bias voltage to 40 V (Figure 7.15b). In the core region of the implant between $d_{rel} = 50 \%$ and $d_{rel} = 80 \%$ all implant widths are able to collect the full charge. The only difference occurs in the charge sharing region $d_{rel} \leq 30 \%$ where the larger implants of 30 µm and 40 µm are able to collect about 10 % more charge than the 10 µm wide implant. As in the fixed position simulations, the smaller implant appears to be able to collect the same charge, although at the price of a higher bias voltage required to reach the full efficiency.

The results of simulations of irradiated structures are shown in Figures 7.15(c)-(h) for fluences of 1, 3 and $5 \times 10^{15} \, n_{eq}/cm^2$. Two voltages are chosen per fluence to represent the situation at low and high bias voltages. The low voltages are meant to reveal particular problems of the implant widths while the high bias voltages are meant to show the optimal performance of each structure.

The results are fairly consistent across the various fluences. At low bias voltages, all implant widths are equally efficient in the core region of the implant between $d_{rel} = 70 \%$ and 90 % for all irradiated structures. However, significant differences occur towards the charge sharing region in the center of the four pixels. While the 40 µm wide implant is still able to collect 80 % of the charge at $d_{rel} = 20 \%$ at 50 V and a fluence of $1 \times 10^{15} \, n_{eq}/cm^2$, the 10 µm wide implant can only collect about 30 % of the charge. Consequently, charge sharing is enhanced for smaller implants. This is beneficial for the resolution before irradiation, but after irradiation it can make the difference between detecting a particle or not. Therefore, localised inefficiencies around the implant have to be expected for the small implants after irradiation at low bias voltages.

The CCE of all implant widths decreases with increasing irradiation. While about 90 % of charge can be collected in the core region at 200 V and a fluence of $1 \times 10^{15} \, n_{eq}/cm^2$, only 70 % of the charge is collected in the core region at 400 V and a fluence of $5 \times 10^{15} \, n_{eq}/cm^2$. Comparing different bias voltages, the increase in CCE sums up to a value of about 5−10 % points across all irradiations, implant widths and impact positions between low and high bias voltages.

In general, the charge collection efficiency increases from low to high bias voltages. The qualitative picture does not depend on the fluence: the charge sharing region between $d_{rel} = 10 \%$ and 50 % is the critical point where the larger implants can collect significantly more charge at all voltages.
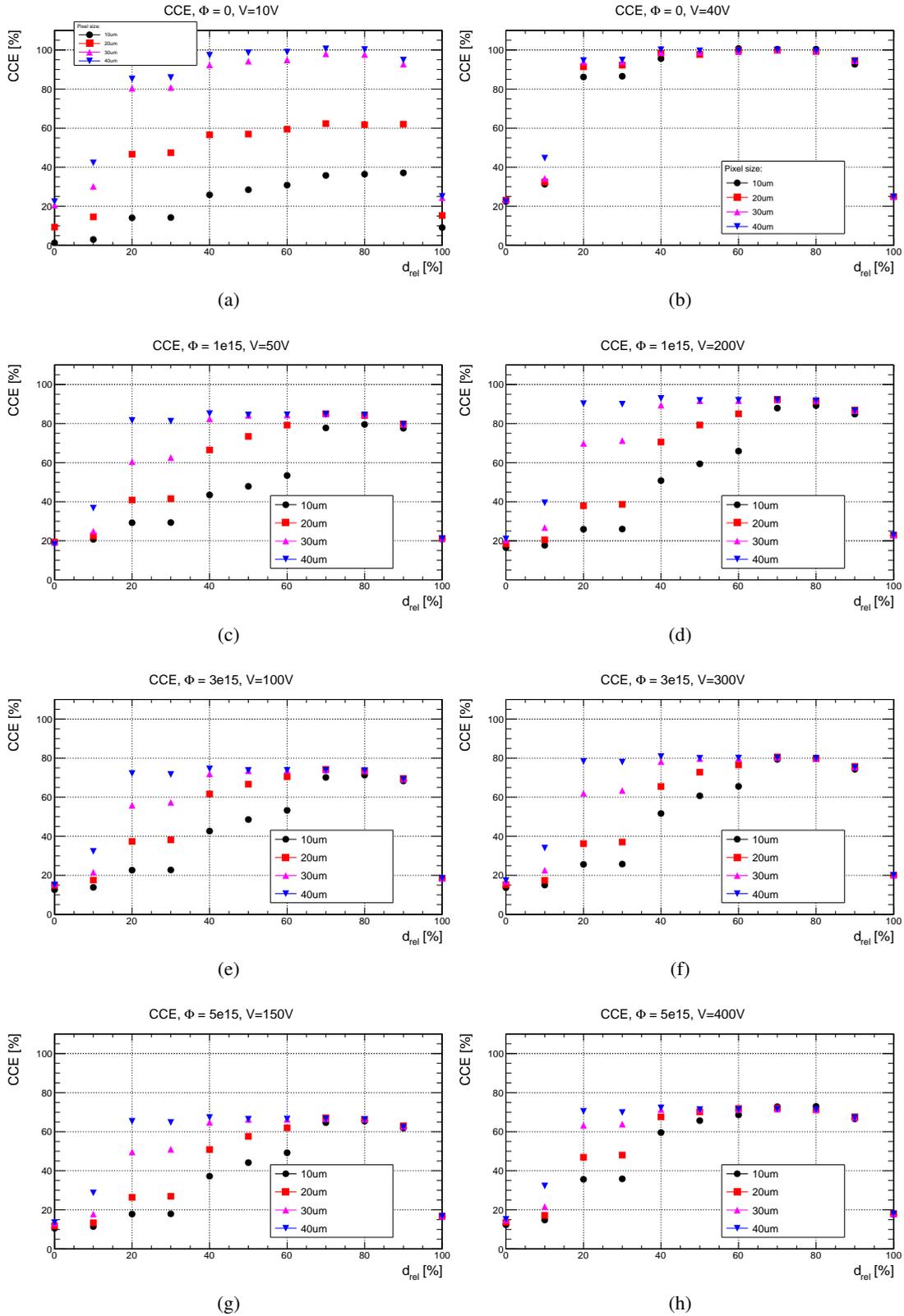
Figure 7.15: Simulated charge collection efficiency (CCE) as a function of the impact position across the diagonal of the pixel under investigation. Two bias voltages are investigated per fluence. The fluences range from 0, 1, 3 to 5×10^{15} n_{eq}/cm^2 from top to bottom.

Conclusion on the pixel implant optimisation

The power dissipation of irradiated sensors is a significant heat source in the detector. Therefore, reducing the power dissipation allows to reduce the cooling capabilities and, thereby, the material budget. The leakage current typically rises significantly with increasing bias voltage for irradiated sensors. Consequently, requiring a higher bias voltage to achieve the same CCE severely increases the power dissipation. Hence, the smallest investigated implant widths of $10\,\mu\text{m}$ and $20\,\mu\text{m}$ can be excluded as they require higher bias voltages to perform competitively. There is no straightforward choice in the comparison of the remaining $30\,\mu\text{m}$ and $40\,\mu\text{m}$ wide implants. While the small differences in CCE is not a large advantage for the $40\,\mu\text{m}$ wide implant, the capacitance (and thus noise) is significantly smaller for the $30\,\mu\text{m}$ wide implant. As a result, the signal-to-noise ratio (SNR) is smaller for the larger implant. On the other hand, a clear improvement in terms of breakdown voltage is found for the $40\,\mu\text{m}$ wide implant. In conclusion, the optimal pixel implant width is found to be between $30\,\mu\text{m}$ and $40\,\mu\text{m}$ considering the breakdown voltage, the capacitance and the CCE before and after irradiation.

ANNEALING STUDIES

8.1 Results for the ATLAS pixel detector

The leakage current increases linearly with fluence as explained in Sections 3.2 and 6.2.2. This section presents a comparison of the evolution of the measured leakage current of pixel modules of the ATLAS pixel detector to the simulated leakage current according to the Hamburg model [54]. The work is done in cooperation with the University of New Mexico [141] that is responsible for the measurements of pixel leakage currents in ATLAS. The simulation part employs the RadDamSim code developed in the course of this thesis (see Section 6.2).

This analysis focuses on the barrel part of the pixel detector from which a representative selection of modules in $r - \phi$ direction around the beam pipe and in z direction along the beam pipe is used. The three barrel layers are located at radii of $r = 50.5 \, \text{mm}$ (B-Layer), $r = 88.5 \, \text{mm}$ (Layer-1) and $r = 122.5 \, \text{mm}$ (Layer-2) while they all extend up to $z = \pm 380 \, \text{cm}$, where $z = 0$ is defined as the nominal interaction point.

The comparison of simulated and measured leakage current enables a verification of the luminosity-to-fluence conversion factor for the present ATLAS pixel detector, derived with FLUKA simulations. Thereby, it does not only improve the understanding of the detector and the underlying simulations but also yields important input to other simulations that require an accurately determined fluence. The particle transport code FLUKA [142, 143] provides insight into the complex radiation fields of the ATLAS detector [144]. In this approach, the high-energy hadron cascades in the material of the individual detector components are modelled being the dominant cause of displacement defects in the silicon bulk. The underlying events of inelastic proton-proton scattering are generated by `Pythia8` [145] based on the MSTW2008LO parton distribution functions [146] and the A2 tune [147]. The simulations yield predictions of the fluence per luminosity in units of $n_{eq}/(\text{cm}^2 \text{fb}^{-1})$. The resulting conversion factors are summarised in Table 8.1 for each layer of the detector at the respective center-of-mass energy of the LHC.

The luminosity is measured by ATLAS employing dedicated bunch-by-bunch detectors [149]. The luminosity data is recorded approximately once per minute for every luminosity block. The input data for the simulation is completed with the temperature of the detector that is measured with NTCs

Table 8.1: Conversion factors from luminosity to fluence based on FLUKA simulations for the barrel layers of the ATLAS detector at a given center-of-mass collision energy [148]. All values are given in units of $10^{11} \, n_{eq}/(cm^2 fb^{-1})$.

Energy [TeV]	B-Layer	Layer-1	Layer-2
7	21.26	8.78	5.43
8	22.50	9.20	5.71
13	28.95	12.35	7.84

on the surface of each module flex being directly glued to the silicon sensor. The temperature is recorded at a similar frequency as the luminosity and is combined offline to a profile according to the description in Section 6.2.1.

The leakage current data is recorded in two different ways. The per-module high voltage patch panel subsystem (HVPP4) serves as a fan-out and distributes the HV generated by Iseg power supplies [150] to the individual modules. The HVPP4 system monitors the leakage current per module based on a modified commercial current monitoring board. The second data source is the current reading from the Iseg power supplies themselves being used throughout Run-2 when the granularity was reduced to only two detector modules per power supply. The number of modules per power supply was six or seven in Run-1, and for this reason the single module resolved data of the HVPP4 system is used in this period, even if less accurate. More information about the measurement of the leakage current utilising the HVPP4 system can be found in Reference [151].

Quality requirements are applied to the leakage current data to improve the reliability. Data are only accepted if the nominal bias voltage is applied and the stable beam flag is set for the proton beams in the LHC. The measured leakage current is transformed to a reference temperature of $0\,°C$, employing Equation 6.15, by using the recorded temperature per module. The total systematic uncertainty of the final current is estimated in [151] to be $12.5\,\%$ for current measurements of the HVPP4 system and $5.4\,\%$ for readings of the power supplies. This includes uncertainties arising from the luminosity determination, temperature measurement, current measurement and the granularity of the measurements.

Figure 8.1 shows the measured and simulated leakage current for all layers from February 2011 to November 2017. The top panels show the average module temperature as well as the applied bias voltage. Some important dates indicated in the main panel point to shutdown periods during which beneficial annealing reduces the leakage currents. The B-Layer exhibits the steepest increase in bias voltage being closest to the interaction point. The simulated leakage current is compared to the data after a constant scaling factor is applied to the luminosity-to-fluence conversion factor calculated by FLUKA. The scaling factor is obtained by a fit to the data and it is different for each layer. The extracted scaling factors are summarised in Table 8.2. The uncertainty is obtained from the fit to the data.

There is a good overall qualitative description of the data by the Hamburg model simulation after the scaling factors are applied. In particular, the annealing periods are described well and confirm both the temperature readings and the assumptions about the annealing process made by the Hamburg
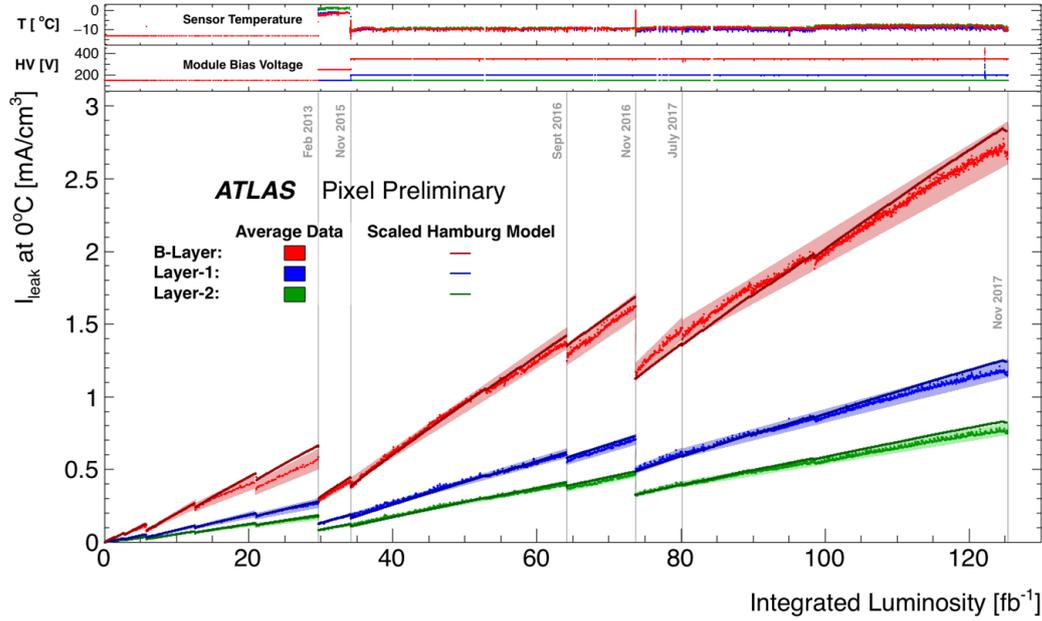
Figure 8.1: Measured and simulated leakage current of the ATLAS pixel detector as a function of the integrated luminosity. Measurements on each layer are averaged over a representative sample of modules in z and r-ϕ. The leakage current is normalised to the sensor volume and to a temperature of $0\,°C$. The simulated leakage current is scaled to match the measured leakage current. The indicated uncertainties include statistical and systematic uncertainties of the measured leakage current indicated as colour band. This figure is published in Reference [152].

Table 8.2: Scaling factors applied to the conversion factors presented in Table 8.1.

B-Layer	Layer-1	Layer-2
1.27 ± 0.02	1.27 ± 0.02	1.34 ± 0.02

model.

The extracted scaling factors are used as an accurate input for further simulations of other detector properties. One example is the simulation of the depletion voltage of the pixel sensors that has inherently more degrees of freedom and, thus, does not allow to simultaneously determine the fluence.

8.2 Investigations of annealed irradiated pixel modules

The fluence a pixel detector module has received affects various aspects of its behaviour. The different aspects of the radiation damage were studied extensively in the past (see for example References [60–63]). In addition to the fluence, the temperature at which the modules are operated and stored influences their properties. The temperature of the detector is usually well below $0\,°C$ during regular operation. These low temperatures prevent the sensors to undergo reverse annealing

Table 8.3: Overview of the investigated modules and their properties.

Module	W7-450-1	W3-450-1	W4-300-2
Sensor production	HLL (SOI-3)	HLL (SOI-3)	HLL (SOI-3)
Thickness [μm]	150	100	100
Irradiation facility	CERN PS	CERN PS	Karlsruhe
Fluence [10^{15} n_{eq}/cm^2]	2	2	5
Total annealing time [days]	336	336	480

which would increase the depletion voltage and result in a decreased performance [54]. However, the cooling for the detector might not be available during maintenance phases such that longer periods of RT exposure can occur. This happened during the LS-1 of the present ATLAS pixel detector and it is foreseen to happen again during the planned exchange of the two innermost pixel detector layers of the ITk at about half the HL-LHC luminosity. The available data on the impact of long-term annealing on the hit efficiency is sparse. The hit efficiency is examined here rather then the traditional quantities of depletion voltage and charge-collection efficiency as is is the most relevant figure of merit for the detector operation.

Modules built with FE-I4 readout chips are used due to the unavailability of the RD53A readout chip during the initial phase of this thesis and the long time period needed during which the devices are measured. A summary of the four modules included in this investigation is presented in Table 8.3. First, the functionality of each module is verified before irradiation using laboratory and testbeam measurements. Then, the modules are irradiated to different fluences at different facilities after which they are stored in a freezer (typical temperature $-18\,°C$) until they can be measured in the testbeam. This first measurement analyses the performance before annealing. Subsequent to the first testbeam after irradiation, the modules are stored in different lockers at the testbeam facilities of CERN and DESY. The periods at RT are tracked in time but not in temperature. The actual temperature of the module depends on the temperature of the testbeam hall which is affected by the ambient temperature. A temperature between $15\,°C$ (during winter) and $25\,°C$ (during summer) with a mean around $20\,°C$ is estimated. The modules are remeasured during various testbeam campaigns. The modules are operated in a cold box at temperatures well below $0\,°C$ during the measurements but stored in the locker at RT again between measurements. The hit efficiency as a function of bias voltage is measured each time to identify any change.

The performance of modules W7-450-1 and W3-450-1 is summarised in Figure 8.2. It shows the efficiency as a function of the bias voltage at various times of annealing although not all bias voltages are tested at each measurement. The modules are irradiated to a fluence of 2×10^{15} n_{eq}/cm^2 at CERN PS. The Gaussian beam profile of the irradiation leads to an inhomogeneous irradiation profile. Without a proper selection of pixels with similar fluence each testbeam would correspond to a different effective fluence since a different part of the module is illuminated and thus probed during each testbeam. Therefore, the analysis code FluenceAnalyser is developed (see Section 4.2.3) to compute the mean fluence per pixel and connect this value to the per-pixel efficiency. The underlying fluence distribution is also shown in Section 4.2.3. Using this procedure, the efficiency can be
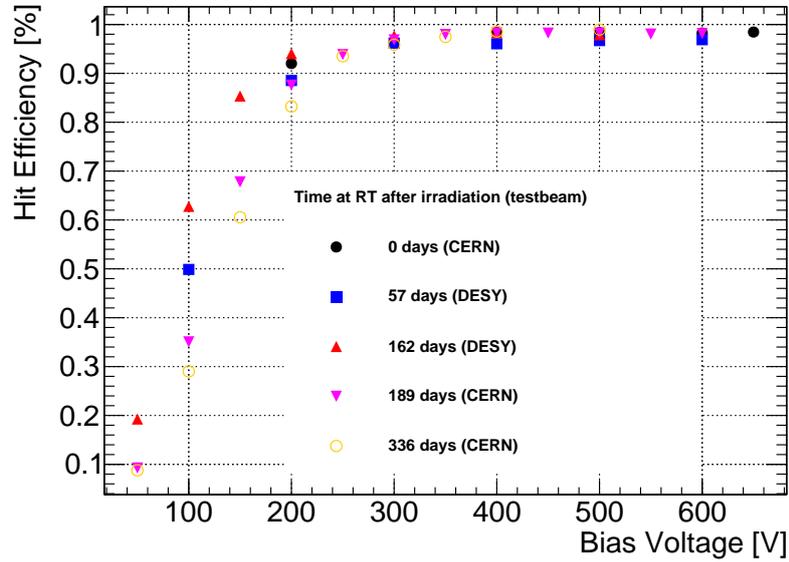
(a) W7-450-1, $d = 150\,\mu m$
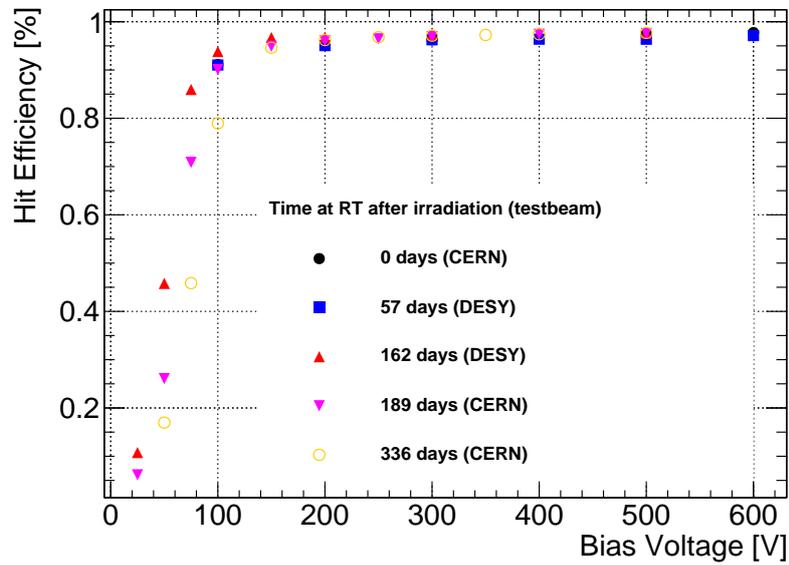
(b) W3-450-1, $d = 100\,\mu m$

Figure 8.2: Measured hit efficiency of modules W7-450-1 (a) and W3-450-1 (b) before and after annealing. Both modules are irradiated to a fluence of $2 \times 10^{15}\,n_{eq}/cm^2$ at CERN PS and annealed up to 336 days at RT. Not all bias voltages are tested at each measurement. Vertical bars indicate systematic and statistical uncertainties but are smaller than the symbol size.
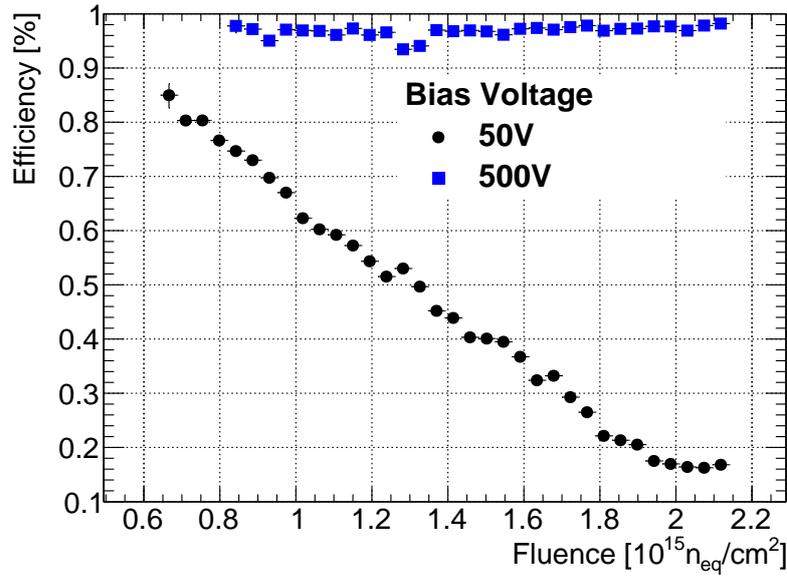
Figure 8.3:   Measured hit efficiency of the module W3-450-1 ($d = 100\,\mu m$) after 336 days of annealing as a function of the fluence at bias voltages of 50 V and 500 V. The uncertainties only include statistical uncertainties according to the Clopper-Pearson intervall.

extracted as a function of the fluence. The same fluence bin of $(2.00 \pm 0.04) \times 10^{15}\,n_{eq}/cm^2$ is selected for all measurements in Figure 8.2 to ensure the comparability between different testbeams. However, an additional systematic uncertainty of $5 \times 10^{14}\,n_{eq}/cm^2$ on the absolute fluence should be taken into account, related to the fluence extraction procedure.

Focusing on a low voltage of 100 V, the performance is more sensitive to any kind of effect such that the previously described effects of beneficial and reverse annealing are most apparent here. The efficiency increases slightly in the course of the early annealing phase from 0 days and 57 days (black dots and blue squares) to the measurement after 162 days of annealing (red pyramids). The necessary bias voltage to obtain full hit efficiency is reduced by about 50 V corresponding to the phase of beneficial annealing. During the two measurements at annealing times of 189 and 336 days (magenta triangles and cyan open circles), the efficiency decreases again. The necessary bias voltage to obtain full hit efficiency is increased again by about $50-100$ V. This phase corresponds to reverse annealing.

Despite the significant changes at low bias voltages, both modules achieve the full hit efficiency around 97 % at a bias voltage of 300 V for the 150 μm thick sensor of W7-450-1 and of 200 V for the 100 μm thick sensor of W3-450-1 at all annealing times up to 336 days. Therefore, no effect on the operation of these sensors is expected during annealing times at RT up to one year.

The FluenceAnalyser also allows for further investigations. Figure 8.3 shows the hit efficiency as a function of the fluence for W3-450-1 after 336 days of annealing. The results at the lowest bias voltage of 50 V are compared to those obtained at 500 V. The full hit efficiency is maintained in the entire fluence range at a bias voltage of 500 V while the hit efficiency exhibits a strong dependence on the fluence for the low bias voltage of 50 V.
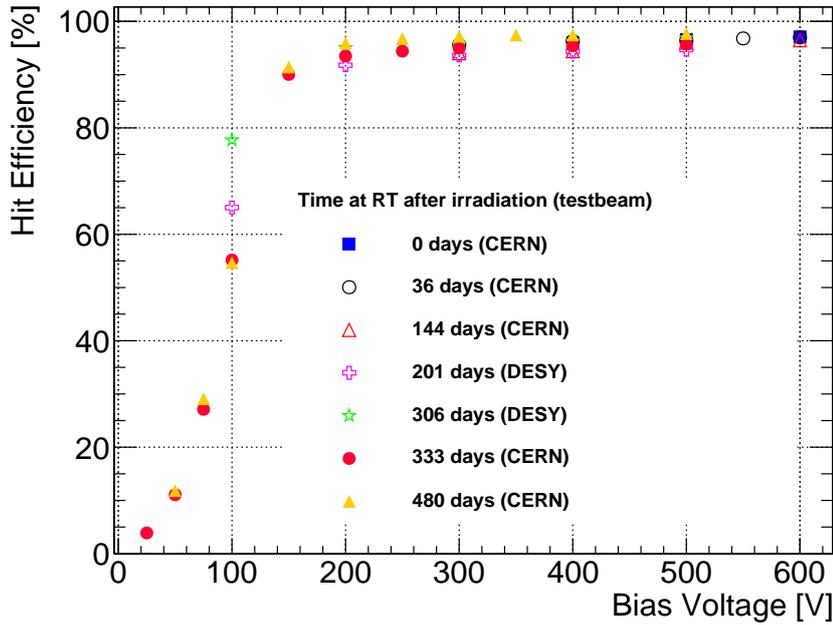
Figure 8.4: Measured hit efficiency of the module W4-300-2 ($d = 100\,\mu\text{m}$) before and after annealing. The module is irradiated to a fluence of $5 \times 10^{15}\,n_{eq}/\text{cm}^2$ at KIT and annealed up to 480 days at RT. Vertical bars indicate systematic and statistical uncertainties but are smaller than the symbol size.

The last module under investigation is W4-300-2. It is irradiated homogeneously to a fluence of $5 \times 10^{15}\,n_{eq}/\text{cm}^2$ at KIT. The hit efficiency as a function of the bias voltage at various annealing times is shown in Figure 8.4. The evolution of the efficiency is tracked up to 480 days of annealing. Unfortunately, no measurements at the lower voltages are available for the early annealing times. However, the variations at 100 V again indicate the effect of beneficial annealing as an improvement of the efficiency in the period from 201 to 306 days. Furthermore, the effect of reverse annealing manifests as a decrease of efficiency in the period from 306 to 333 days. Surprisingly, the time span of just 27 days seems very short compared to the previous beneficial annealing happening between 201 days and 306 days of annealing. Still, the behaviour of the module is not traced continuously such that no information about the magnitude and dynamics of the respective annealing phases are present. The data points at 480 days of annealing are very close to the points after 333 days indicating that most volatile defects contributing to the depletion behaviour of the sensor have already decayed.

Despite being the best indication for the annealing processes, the behaviour at low bias voltages is irrelevant for the operation of the detector in the ATLAS experiment. Instead, the optimal hit efficiency has to be guaranteed and the detector will always be operated beyond its full depletion voltage, if possible. The variation in the bias voltage needed to achieve full hit efficiency is again around 50 V. A bias voltage of 300 V is sufficient to achieve full hit efficiency at all annealing times up to 480 days. Consequently, again no effect on the operation of these sensors is expected during annealing times at RT up to about 16 months.

CHARACTERISATION OF RD53A MODULES

This chapter presents results of the first modules built with RD53A readout chips (see Section 3.5.2) and sensors from the SOI-3 and SOI-4 productions (see Section 3.3.2). All measurements are performed using the BDAQ system (see Section 4.1.2).

9.1 Characterisation before irradiation

Several not-irradiated RD53A modules are evaluated to investigate different effects which impact their performance and therefore influence the design of the ITk planar sensors. If not explicitly stated, presented results origin from measurements employing the linear FE flavour of the RD53A read-out chip. In addition, a reduced set of measurements is also conducted using the differential FE flavour yielding qualitatively similar results.

9.1.1 Laboratory measurements

A first characterisation of the modules is performed in the laboratory. After module assembly and wire-bonding, an IV curve is taken and each module is tuned. The IV curves in Figure 9.1 are measured with configured modules and at a constant ambient temperature of 20 °C within a climate chamber. Depending on the sensor, the breakdown voltages before irradiation range between 80-350 V and are well above their full depletion voltage of around 15-30 V.

Section 4.1 describes the setup that is used to perform source scan measurements, which marks the final step in the initial module characterisation. Figure 9.2 demonstrates the CCE as a function of the applied bias voltage. The ToT spectra are recorded at different bias voltages while the module is exposed to β electrons from the ^{90}Sr source. A convoluted Landau- and Gauss-function is fitted to the ToT distribution and the MPV as well as its uncertainty are extracted. The CCE is defined as the MPV at a given voltage divided by the maximum observed MPV in the investigated voltage range. Since this procedure does not use a ToT-to-charge conversion, it relies on a linear relation of the two that is generally given by the RD53A. The dependence of the CCE on the sensor thickness is displayed for one sensor each with 100 µm and 150 µm thickness. As expected, the thinner sensor of 100 µm thickness reaches full CCE already at lower bias voltages compared to the 150 µm thick

Figure 9.1: The current-voltage characteristics of various RD53A modules, see Table 3.3 for the naming convention. Open (closed) symbols correspond to modules with $100\,\mu m$ ($150\,\mu m$) thick sensors. All modules are configured and kept at a constant ambient temperature of $20\,°C$ within a climate chamber. Vertical bars include statistical uncertainties only and are smaller than the symbol size.
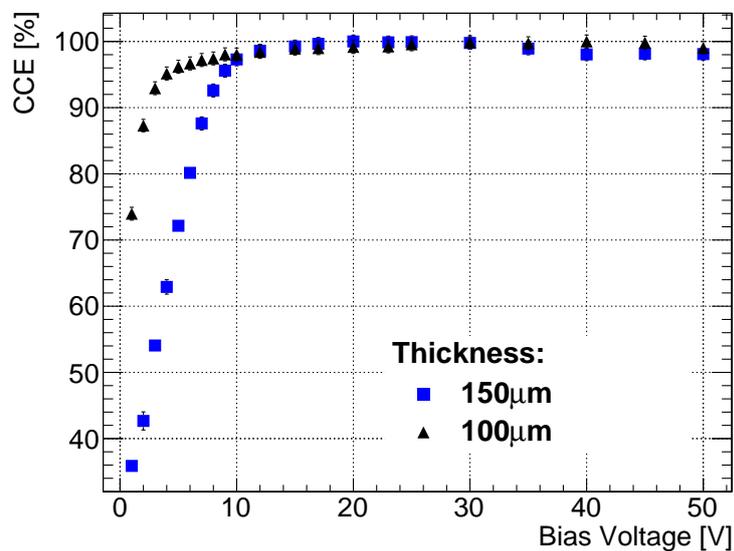
Figure 9.2: Charge collection efficiency as a function of the applied bias voltage for RD53A modules with sensor thicknesses of $100\,\mu m$ ($1400\,e$ readout chip threshold) and $150\,\mu m$ ($1500\,e$ readout chip threshold). Vertical bars correspond to the uncertainty in MPV derived from the fit to the ToT spectrum.

sensor. That indicates a lower depletion voltage for the thinner sensor compared to the $150\,\mu m$ thick sensor.

(a) Grounded BR

(b) Floating BR
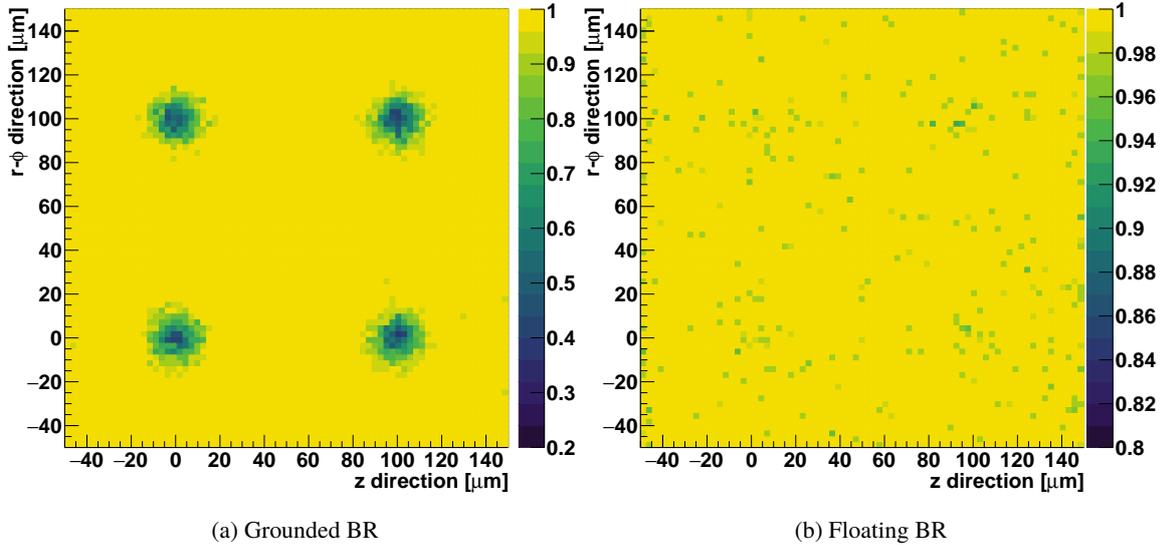
Figure 9.3: In-pixel hit efficiency for 4×4 pixels with $50 \times 50\,\mu m^2$ pixel cell geometry, the PT dots are located at $(x,y) = (0,0)$, $(0,100)$, $(100,0)$ and $(100,100)$. A bias voltage of 50 V is used to deplete the 150 μm thick sensor. The results for both figures are obtained with the same module tuned to a threshold of 2000 e. While the BR, and thereby the PT dots, is grounded in (a) (overall efficiency: 98.5 %), it is floating in (b) (overall efficiency: 99.9 %).

9.1.2 Testbeam measurements

After the initial laboratory characterisation, the modules are examined with particle beams (see Section 4.2). For the modules tested before irradiation, all data are recorded at the SPS testbeam facility at CERN.

Effect of the punch-through structure on the hit efficiency

Biasing structures were subject to many studies, documented in References [61, 62, 153, 154]. Despite this, it is not yet decided if and if yes in which form a PT mechanism will be implemented in the final design of ITk sensors. Hence, different options are implemented in the SOI-3 and SOI-4 sensor productions (see Section 3.3.2). Since the layout of the PT was already optimised in the previously mentioned studies, this investigation focuses on another aspect: the electrical potential of the PT. The effect of the PT dot on the charge collection efficiency of a module depends on the difference in the electrical potential of PT dot and pixel implant. The lower the potential of the PT dot, the less it competes in the charge collection with the pixel implant. In this configuration, the performance of a sensor with PT is indistinguishable from a sensor without PT.

One way of changing the PT potential is to leave the BG floating, which connects all PT dots and is connected to the BR. Then, the PT system will be grounded only through a reverse punch-through effect from the pixel implant. Consequently, the resulting potential of the PT system will be close to ground, but still significantly lower than the potential of the grounded pixel implants. Figure 9.3 shows the in-pixel hit efficiency for a group of 4×4 pixels for the same module, once with the PT

(a) Grounded GR   (b) Floating GR
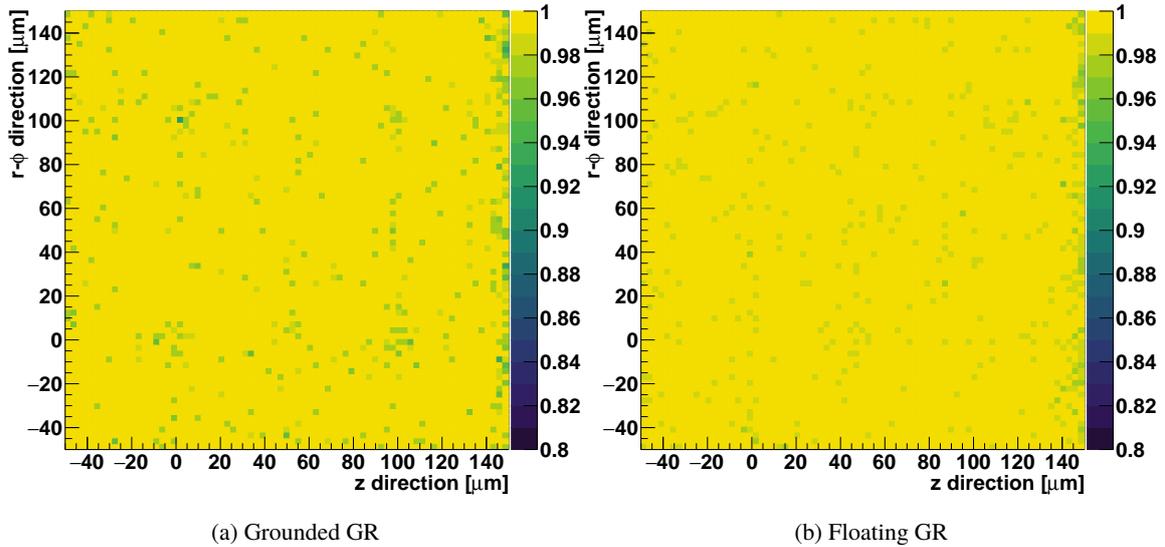
Figure 9.4:   In-pixel hit efficiency for 4×4 pixels with $50 \times 50\,\mu m^2$ pixel cell geometry, the PT dots are located at (x,y) = (0,0), (0,100), (100,0) and (100,100). A bias voltage of 50 V is used to deplete the $100\,\mu m$ thick sensor. The results are obtained with the same module tuned to a threshold of 1500 e. The sensor has a split design which decouples the potential of the PT dots from the potential of the GR. For the left figure, the GR is grounded while for the right figure the GR is left floating. The overall efficiency is 99.9 % in both cases.

dots connected to the grounded BR (Figure 9.3a), once the BR is left floating (Figure 9.3b). This is possible because a jumper is implemented on the module PCB to select the BR potential. In that way, the results differ only in the BR potential settings, while the tuning and all other conditions are identical for the measurements. The module is tuned to a threshold of 2000 e. The efficiency degradation in the area affected by the PT dot is severe in the case of a grounded BR. Therefore, the overall efficiency decreases by 1.4 % compared to the case of leaving the BR floating (98.5 % (grounded) to 99.9 % (floating)).

Unfortunately, there are also some detrimental effects expected for a floating BR potential. Since the BR is usually connected to the innermost GR, leaving the BR floating means leaving this GR floating. With a floating GR, the large current being created in the edge region of the sensor especially after irradiation cannot be absorbed by the GR before entering the pixel matrix. Hence, the outermost pixels have to compensate this current and are consequently potentially more noisy. However, this effect is absent in not-irradiated modules.

Furthermore, a new design (henceforth called split, see Section 3.3.2) implemented in the SOI-4 production allows to set the GR and BR potentials independently. In this design, the BR is always left floating, while the GR can either be grounded or also left floating, depending on the jumper setting on the module PCB. A module with a sensor featuring the split design is measured with grounded and floating GR potential to validate the isolation of BR and GR. In case the two are perfectly isolated, the grounding of the GR does not have an effect on the BR and consequently also not on the PT dots. Figure 9.4 presents the comparison of the hit efficiency results obtained with
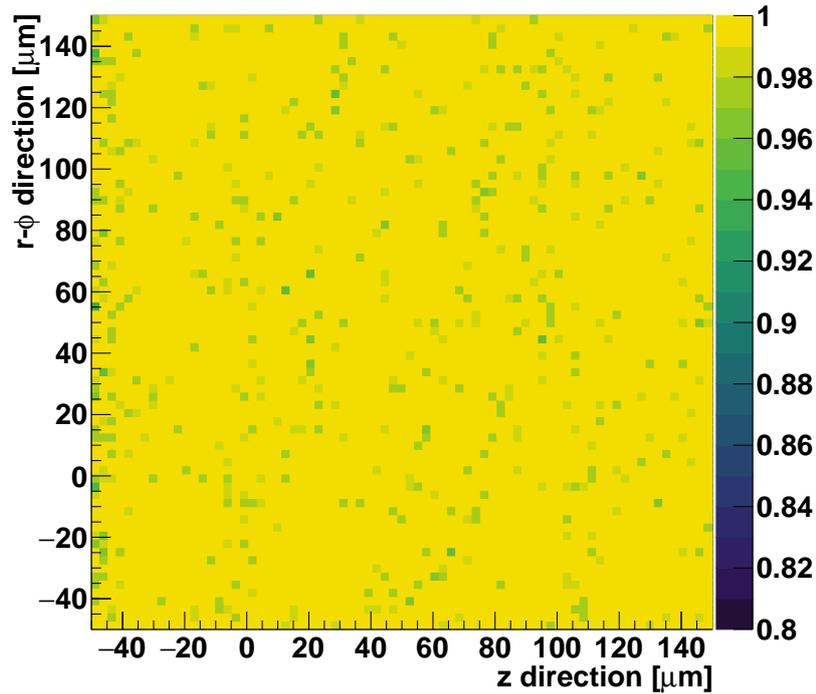
Figure 9.5:   In-pixel hit efficiency for 4×4 pixels with $50 \times 50\,\mu m^2$ pixel cell geometry with no biasing structure implemented. A bias voltage of $50\,V$ is used to deplete the $100\,\mu m$ thick sensor. The module is tuned to a threshold of $1000\,e$. The overall efficiency is $99.8\,\%$.

both GR potentials for a module built with a sensor of the split design and tuned to a threshold of $1500\,e$. The overall efficiency with the split design is not affected by the change of potential being $99.9\,\%$ in both cases. On the other hand, a tiny but visible inefficiency can be spotted at the four PT dots for the grounded GR case, located at $(x,y) = (0,0)$, $(0,100)$, $(100,0)$ and $(100,100)$ (Figure 9.4a). The size of this effect is not relevant for the operation of the ITk detector before irradiation.

The last option under consideration for the final ITk layout is a sensor without any PT structures. The in-pixel hit efficiency map of a module employing a $100\,\mu m$ thick sensor without PT is presented in Figure 9.5. The module is tuned to a threshold of $1000\,e$. The overall efficiency is $99.8\,\%$. There are no local inefficiencies in the in-pixel hit efficiency distribution before irradiation. While the design without biasing structure is quite safe in terms of hit efficiency, it does not allow to perform IV characterisations of the bare sensors, other than at wafer level, in case a temporary or intermediate metal technique (see Section 3.3) is applied.

Impact of the bias voltage on the hit efficiency

Figure 9.6 illustrates the influence of the bias voltage on the hit efficiency. The module under investigation is built with a $100\,\mu m$ thick sensor of the SOI-3 production. The bias voltage is increased in steps, and for each of them the hit efficiency is evaluated. Since the measurements are all performed in the same configuration (e.g. tuning, temperature, beam properties, ...), the systematic uncertainties of the individual measurements are strongly correlated. Statistical uncertainties are
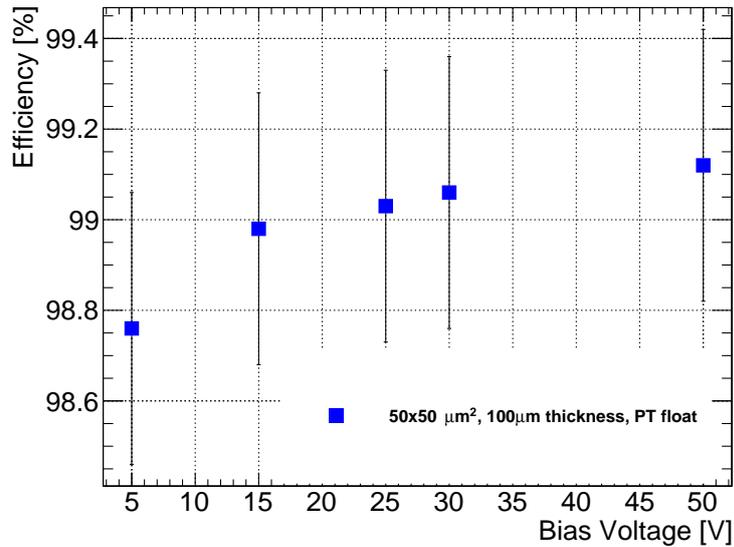
Figure 9.6: Hit efficiency as a function of the applied bias voltage for a module built with a $100 \, \mu m$ thick sensor employing the $50 \times 50 \, \mu m^2$ pixel cell geometry. The PT is floating during the measurement. The vertical bars correspond to the statistical and systematic uncertainty added in quadrature.

one to two orders of magnitude smaller than the systematic uncertainties. A small but significant increase of the hit efficiency of about 0.4 % is observed when increasing the bias voltage from 5 V to 50 V. Nonetheless, including the information from the CCE measurements in Section 9.1.1, the sensor is depleted below 10 V and the hit efficiency is saturated between 30 V and 50 V for a sensor thickness of $100 \, \mu m$.

Hit efficiency as a function of the threshold

The readout chip threshold is investigated and is found to have a large impact on the module performance. Figure 9.7 displays the efficiencies of different modules at various thresholds for the linear and differential FE flavours. A bias voltage of 50 V is used for all measurements. The difference in overall efficiency is about 2.5 % over 1500 e of difference in threshold. Consequently, in the relevant threshold regime between 600 e and 2100 e, a loss in efficiency of 1 % is expected per 600 e increased threshold. Therefore, low threshold operation will be essential to maintain a high hit efficiency especially towards the end of lifetime of the ITk when the collected charge is reduced due to radiation damage. Additionally noteworthy is the measurement of a sensor employing the $25 \times 100 \, \mu m^2$ pixel geometry without PT employing the differential FE flavour (black triangles). The efficiency decreases as expected from a threshold of 580 e to 1200 e before it increases again for a threshold of 2000 e. The cause is unknown but might be related to an artefact of the reconstruction or an influence of the reference module that is more noisy during the measurement at a threshold of 1200 e.
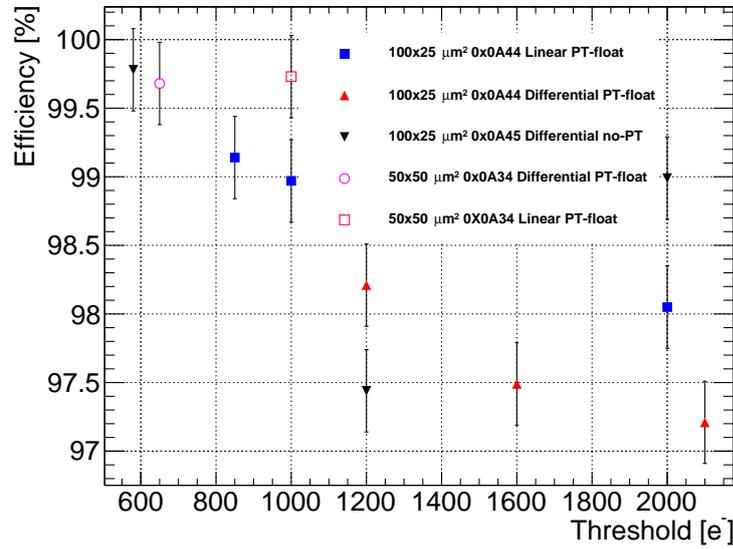
Figure 9.7: Hit efficiency as a function of the threshold for different modules. The measurements are performed at a bias voltage of 50 V. The vertical bars correspond to the statistical and systematic uncertainty added in quadrature.
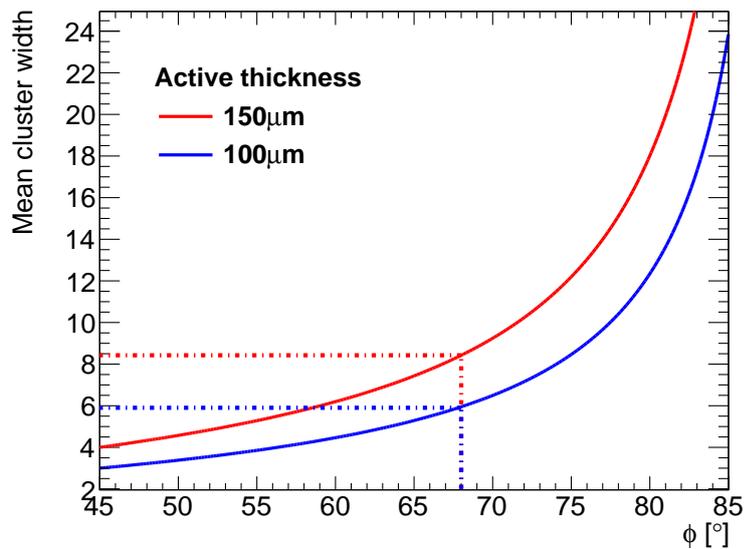
Figure 9.8: Expected mean cluster width as a function of the inclination angle Φ for a pixel cell size of $50 \times 50\,\mu m^2$.

Measurements at inclined incidence

Measurements at perpendicular incidence as presented in the previous sections correspond to $\eta = 0$ in the detector. They need to be complemented by measurements with inclined tracks. The wide η coverage of the ITk results in a sizeable inclination for most of the tracks coming from the interaction point. Therefore, measurements at the geometrically highest possible inclination for L1 of the ITk of 68° (about $\eta = 1.6$) are performed with one module each with sensors of 100 μm and 150 μm thickness. At this inclination, the particle tracks are passing through the sensor at a shallow angle to the surface, resulting in long clusters. Information about the electrical field as a function of the sensor depth can be deduced in addition to the hit efficiency determination. The expected mean cluster width can be calculated as a function of the inclination angle ϕ according to

$$\overline{\omega} = \frac{d_a \cdot \tan \phi}{p} + 1, \tag{9.1}$$

where d_a is the active thickness of the sensor and p is the pitch of the pixel in the traversed direction. Figure 9.8 displays the expected mean cluster width as a function of the inclination angle for the two thicknesses under investigation. Accordingly, a mean cluster width of 5.95 (8.43) is expected for 100 μm (150 μm) thick sensors. However, the formula is only an approximation for the assumption that no threshold is applied to the signal. Simulations of the actual situation indicate that the cluster width can be reduced from the original 5.95 to 5.75 (5.55) in the case of a 100 μm thick sensor and a threshold of 1000 e (2000 e). The threshold inhibits the readout of the partially hit pixels on the edge of the cluster if the deposited charge in the respective pixel is lower than the threshold.

During the analysis of the data, the most likely cluster width is found to be 5 (7) for the 100 μm (150 μm) thick sensor. The deviation from the expected values is likely caused by a combination of the previously discussed threshold effect and a slightly smaller than expected inclination. Also, Reference [61] has found indications that the actual active sensor thickness of these sensors is reduced by around 10 μm leading to a reduced expected cluster width. Consequently, the actual most likely cluster widths of 5 and 7 are used for the analysis of the data.

The analysis of the recorded data can be done by studying cluster properties without the tracking information of the telescope. As explained in Section 4.2, the efficiency as well as the ToT can be extracted for each pixel position within the cluster. Figure 9.9a shows the efficiency as a function of the pixel number, whereas Figure 9.9b indicates the corresponding ToT. In both figures, pixel 0 corresponds to the first pixel of a cluster, being located at the entry point of the particle on the surface of the sensor while the last pixel (4 or 6) corresponds to the exit point of the particle on the backside of the sensor. A maximum of 2 holes are allowed in a cluster to be considered for the analysis. The sensors of both modules are biased with a voltage of 50 V. The module with a 100 μm thick sensor is tuned to a threshold of 800 e, while the module with a 150 μm thick sensor is tuned to a threshold 1000 e.

By definition, the efficiency of the first and last pixel in the cluster has to be 100 % as those determine the begin and end of the cluster. Generally, the ToT is lower in the first and last pixel, as those pixels include the entry/exit point of the impinging particle such that only a part of this pixel cell

(a) Efficiency

(b) ToT
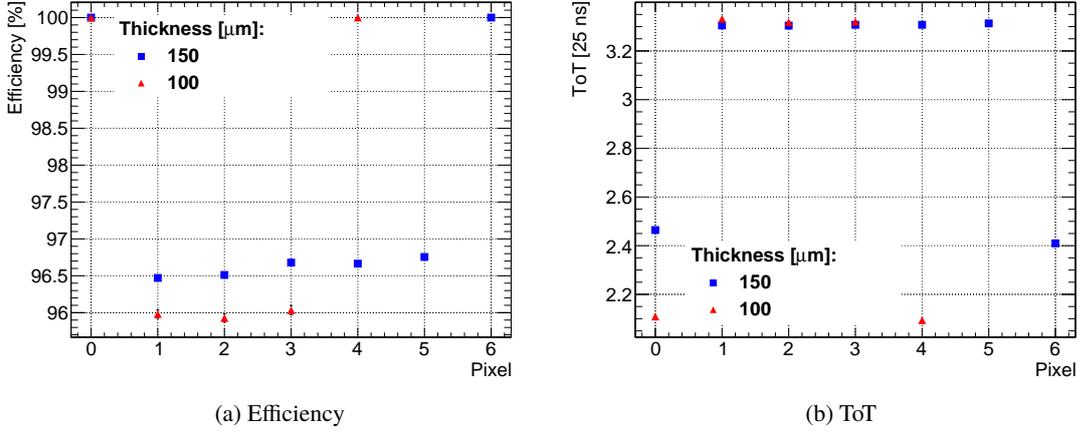
Figure 9.9: Efficiency (a) and ToT (b) as a function of the pixel number in the cluster. Pixel 0 corresponds to the first pixel of a cluster, being located at the entry point of the particle on the front surface of the sensor. The module with a $100\,\mu m$ thick sensor is tuned to a threshold of $800\,e$, while the module with a $150\,\mu m$ thick sensor is tuned to a threshold $1000\,e$. The vertical bars correspond to the statistical uncertainty (a) and to the RMS of the ToT (b).

is traversed by the particle. With less traveled distance within a pixel, there is also less charge deposited on average. The efficiency between the first and last pixel is flat to within about $\pm 0.15\,\%$. Similarly, the ToT is flat to within ± 0.025.

The flat distribution of collected charge points to an electrical field throughout the sensor depth which is high enough to accelerate all deposited charge carriers towards the electrodes. This indicates that both sensors are depleted at the used bias voltage of $50\,V$. This is in agreement with the results from the CCE measurements in the laboratory as well as the testbeam measurements of the hit efficiency as a function of bias voltage.

The lower efficiency with respect to the perpendicular incidence at high η is due to the shorter silicon path traversed by the charged tracks in this configuration (approximately $54\,\mu m$ for the $50 \times 50\,\mu m^2$ pixel geometry). The efficiency of about $96\,\%$ per pixel means for this layer in the real detector that the probability to miss a particle completely at this incidence angle is less than $0.04^5 = 1 \times 10^{-7}$ ($0.04^3 = 6 \times 10^{-5}$) for the $150\,\mu m$ ($100\,\mu m$) thick sensor.

Properties of different pixel cell geometries

The RD53A chip allows for two different pixel geometries, as explained in Section 3.3. The natural pixel geometry is $50 \times 50\,\mu m^2$ fitting exactly to the pattern of the RD53A chip. A $25 \times 100\,\mu m^2$ pixel size can also be interfaced to the readout chip pixel matrix, by displacing the UBM pad positions with respect to the pixel implants (see Figure 3.3b). While all previous results are based on measurements of sensors with the standard geometry of $50 \times 50\,\mu m^2$, this section focusses on the $25 \times 100\,\mu m^2$ geometry and the differences to the standard geometry.

The main argument for considering the $25 \times 100\,\mu m^2$ geometry is to improve the resolution in the r-Φ direction by halving the pixel pitch in that direction compared to the default $50 \times 50\,\mu m^2$ geometry.

(a) $25 \times 100\,\mu m^2$

(b) $25 \times 100\,\mu m^2$
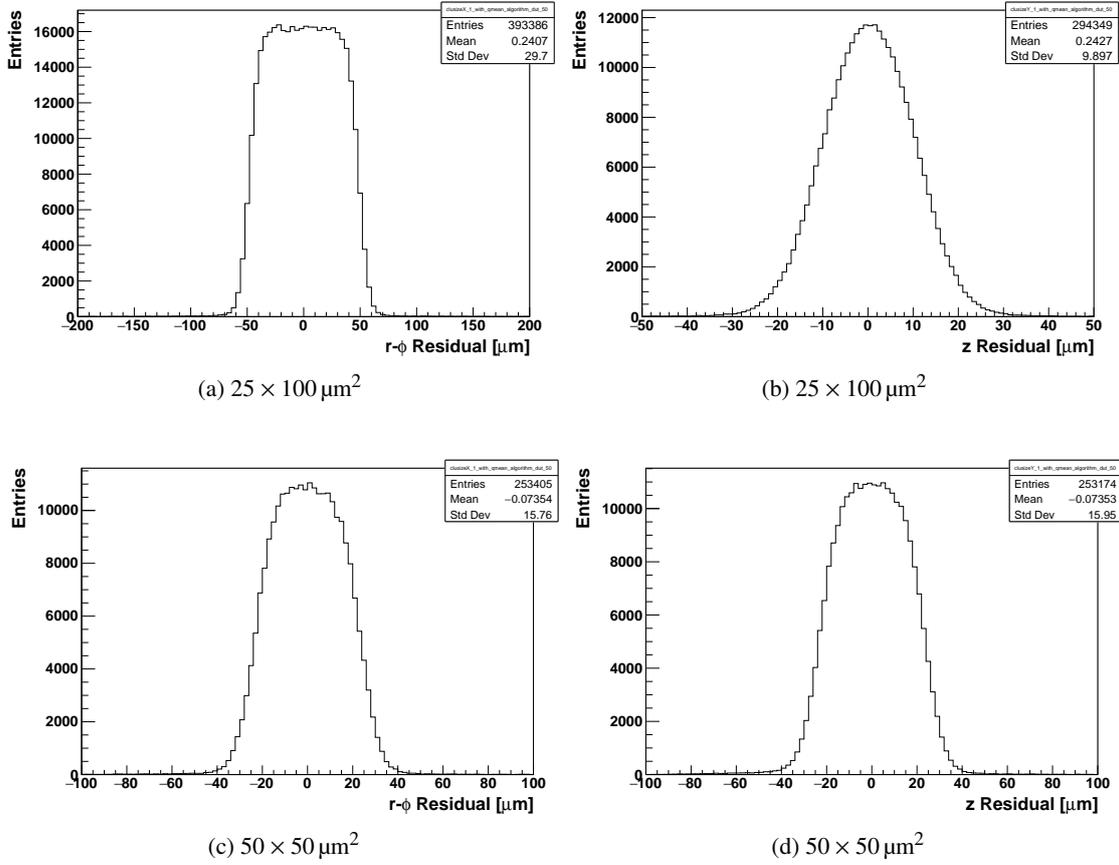
(c) $50 \times 50\,\mu m^2$

(d) $50 \times 50\,\mu m^2$

Figure 9.10: Residuals of hits to fitted tracks using only hits with cluster size one. Both sensors do not have any biasing structures and employ $25 \times 100\,\mu m^2$ (top) and $50 \times 50\,\mu m^2$ (bottom) large pixel cells. Residuals in r-ϕ direction are shown on the left while residuals in z direction are shown on the right. The sensor with $25 \times 100\,\mu m^2$ pixels has a thickness of $150\,\mu m$ and the readout chip is tuned to a threshold of $1500\,e$. The sensor with $50 \times 50\,\mu m^2$ pixels has a thickness of $100\,\mu m$ and the readout chip is tuned to a threshold $1000\,e$.

Many different physics analyses and BSM searches require precise measurements of the transverse momentum (p_T) of charged particles. Therefore, the precise measurement in r-Φ direction is prioritised compared to the z direction. Consequently, an improved transverse resolution, at the cost of a loss in the longitudinal resolution, will be beneficial for the global physics performance [80].

Figure 9.10 presents the residuals of hits to fitted tracks (see Section 4.2) on modules with pixels of the two possible geometries. Only one hit clusters are accepted for the analysis and the hit positions correspond to the center of the hit pixel. The extracted width (RMS) of the residual distribution corresponds to the resolution achieved by the device itself folded with the resolution of the beam telescope and other experimental uncertainties (e.g. multiple scattering). Using the RMS is preferred over other quantities like a Gaussian width, since it also catches non-Gaussian tails mostly caused by δ-electrons. Table 9.1 summarises the pitch and the corresponding theoretical resolution according to Equation 3.14, that assumes a homogeneous track distribution over the pixel area. It also includes the measured RMS and the value of the combined telescope and experimental

Table 9.1: The expected resolution for a binary readout of a pixel with given pitch in comparison to the RMS of the one hit cluster residual distribution and the derived pointing/experimental resolution. No biasing structures are implemented in any of the structures. The sensor with $25 \times 100 \, \mu m^2$ pixels has a thickness of $150 \, \mu m$ and the readout chip is tuned to $1500 \, e$. The sensor with $50 \times 50 \, \mu m^2$ pixels has a thickness of $100 \, \mu m$ and the readout chip is tuned to $1000 \, e$.

Direction	Pitch [µm]	Theo. resolution [µm]	RMS [µm]	Pointing resolution [µm]
z	25	7.2	9.9	6.8
r-ϕ	100	28.9	29.7	6.8
z	50	14.4	16.0	6.7
r-ϕ	50	14.4	15.8	6.5

uncertainty. This is evaluated by subtracting in quadrature the digital resolution from the measured RMS. These calculations are based on the assumption that the calculated digital resolution is valid for the DUT.

The pointing resolution extracted from the measured RMS for the $25 \times 100 \, \mu m^2$ pixels based on the theoretical prediction is compatible with the values obtained for the module employing $50 \times 50 \, \mu m^2$ pixels. The important value is the resolution in the r-ϕ direction for $25 \times 100 \, \mu m^2$ pixels. The RMS for the $25 \, \mu m$ side is $9.9 \, \mu m$ being exactly what is expected considering the theoretical and experimental resolution, the latter extracted from the module using $50 \times 50 \, \mu m^2$ pixels. Therefore, the measured resolution of the $25 \, \mu m$ side is consistent with the predicted value being half that of the $50 \times 50 \, \mu m^2$ case.

As in Section 9.1.2, the influence of the BR potential is investigated for sensors employing the $25 \times 100 \, \mu m^2$ geometry. Two modules with $150 \, \mu m$ thick sensors are investigated. Figure 9.11 reveals a similar effect of the BR potential on the hit efficiency as for the $50 \times 50 \, \mu m^2$ geometry. Figure 9.11a indicates a strong localised inefficiency at the positions of the PT dots when the BR is grounded. The overall efficiency is reduced to $99.0 \, \%$. In comparison, there are no local inefficiencies in Figure 9.11b in the case of floating BR. The overall efficiency is $99.8 \, \%$. No significant difference between the geometries can be identified with respect to the hit efficiency and its sensitivity to the BR potential. In both cases, the overall hit efficiency is significantly reduced by a grounded BR due to local inefficiencies while it is completely restored by leaving the BR floating.

Lastly, the hit efficiency and the cluster size properties for modules without PT are investigated and the results are presented in Figure 9.12. Results of two modules employing sensors of the $25 \times 100 \, \mu m^2$ and $50 \times 50 \, \mu m^2$ geometries are studied. The sensors of both modules are $150 \, \mu m$ thick and are biased at $50 \, V$. While the threshold is tuned to $1500 \, e$ for the module using the $25 \times 100 \, \mu m^2$ geometry, it is tuned to a threshold of $800 \, e$ for the other module. The overall efficiency is not significantly different between the two modules being $99.7 \, \%$ for the module with the $50 \times 50 \, \mu m^2$ geometry and $99.8 \, \%$ for the module with the $25 \times 100 \, \mu m^2$ geometry. Local inefficiencies cannot be detected in either case. In contrast, the in-pixel cluster size differs significantly between the two.

The peak values are extracted from the 2D histogram in the following. In the $50 \times 50 \, \mu m^2$ case, the average cluster size in the center of the pixel is in the range of 1.0-1.1 while it increases to around 1.3
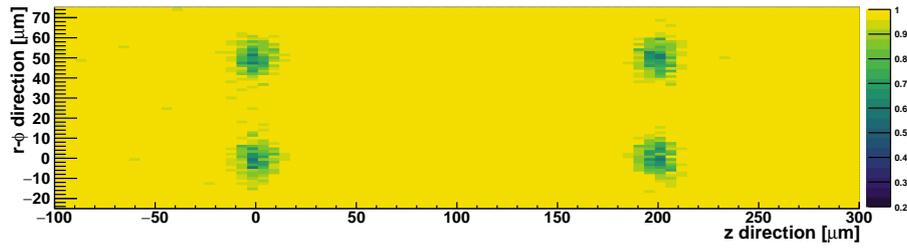
(a) Grounded BR
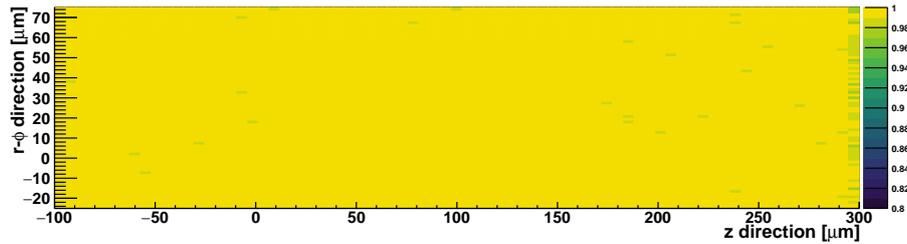
(b) Floating BR

Figure 9.11: In-pixel hit efficiency for a group of 4×4 pixels with $25 \times 100\,\mu m^2$ pixel cell geometry for two modules employing $150\,\mu m$ thick sensors. A bias voltage of $50\,V$ is used in both measurements. One module featuring a grounded PT system is tuned to a threshold of $1000\,e$ and reaches an overall efficiency of $99.0\,\%$ (a). Instead, the other module has the PT system floating, is tuned to a threshold of $1200\,e$ and reaches an overall efficiency of $99.8\,\%$ (b).

in the region between two pixels. The highest average cluster size is observed in between four pixels where it reaches 1.5. For the $25 \times 100\,\mu m^2$ geometry, the average cluster size is about 1.3 in the pixel center, 1.5-1.6 between two pixels and approximately 1.8 between four pixels. This corresponds to an increase of 20-30 % when comparing the $25 \times 100\,\mu m^2$ geometry to the $50 \times 50\,\mu m^2$ geometry. The different thresholds of the two modules can have an additional impact on the cluster size, such that the cluster size of the module using the $25 \times 100\,\mu m^2$ pixels could be even higher. However, additional data, that could not be analysed due to independent issues, indicates no further increase of the average cluster size at a threshold of $1000\,e$.

Another question regarding the usage of the $25 \times 100\,\mu m^2$ geometry is whether the layout of aluminium pads from one cell overlaying neighbouring implants causes cross-talk between the pixels. For better visibility, the projections of Figure 9.12b are shown in Figures 9.13 (a,b). Similarly, Figures 9.13 (c,d) show the projections of Figure 9.12d. The measurements show that the cluster size is different between gaps with and without bump-bond[1] in both geometries. Averaging the bins of the gap region in the z projection leads to an increase of cluster size of about 2.6 % in the case of the $25 \times 100\,\mu m^2$ geometry while it is 2.1 % in the case of the $50 \times 50\,\mu m^2$ geometry when comparing gaps with and without bump-bond. It can be concluded, that a small but acceptable increase in cross-talk is found for the $25 \times 100\,\mu m^2$ geometry when comparing with the $50 \times 50\,\mu m^2$ geometry.

---

[1]In the design of the sensor, there is an overlapping bump-bond or not in the gap between pixels, compare Figure 3.3b.
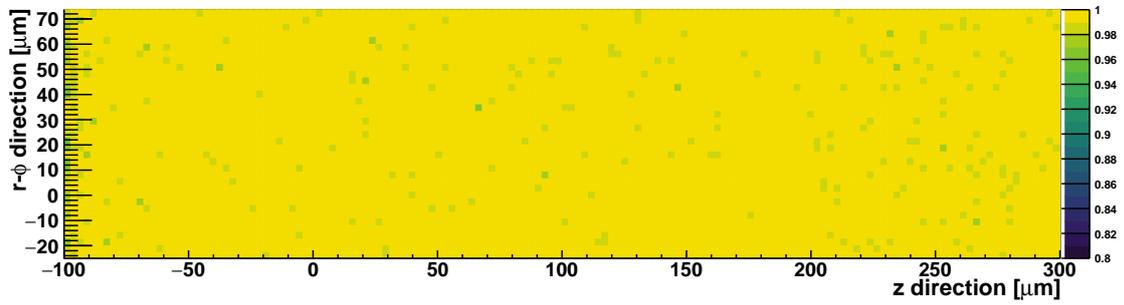
(a) Hit efficiency $25 \times 100\,\mu m^2$
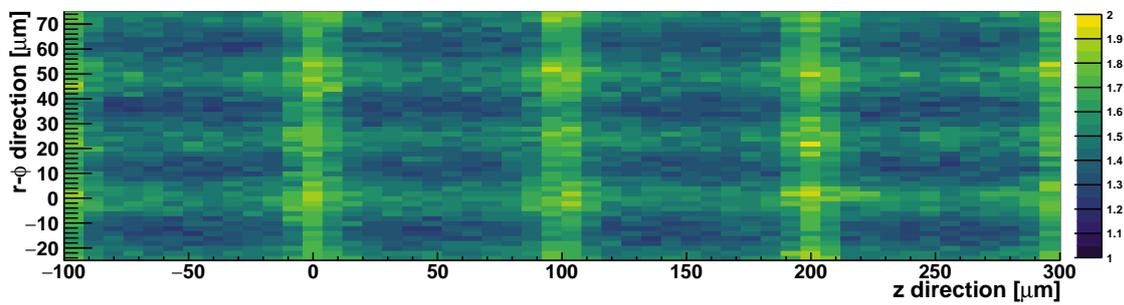
(b) Cluster size $25 \times 100\,\mu m^2$
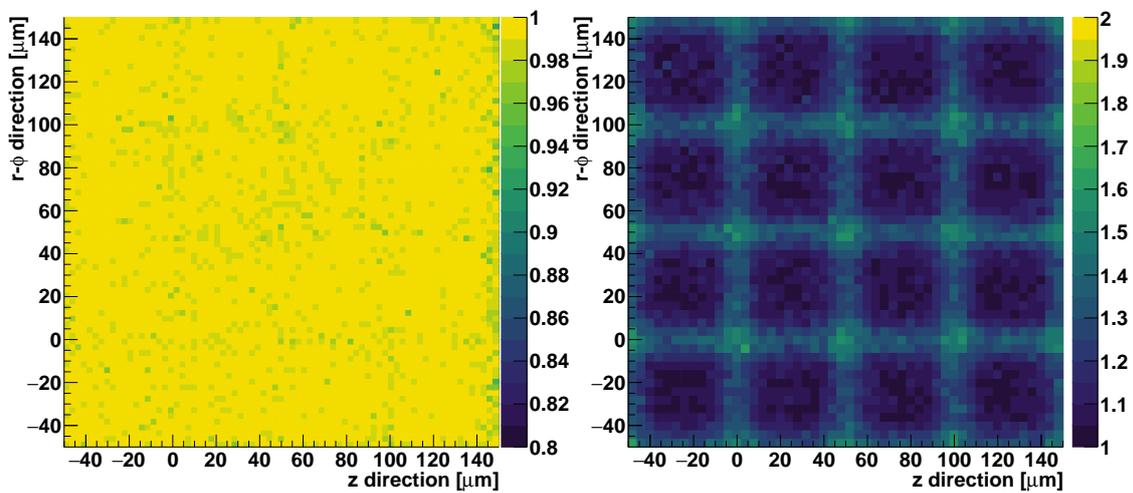
(c) Hit efficiency $50 \times 50\,\mu m^2$

(d) Cluster size $50 \times 50\,\mu m^2$

Figure 9.12: In-pixel hit efficiency for 4×4 pixels of two modules with $150\,\mu m$ thick sensors without PT system biased at $50\,V$. The sensors are segmented in pixel cells of $25 \times 100\,\mu m^2$ (a) and $50 \times 50\,\mu m^2$ (c). In-pixel cluster size maps for 4×4 pixels for the same modules are shown in (b) and (d). The module employing a sensor of the $25 \times 100\,\mu m^2$ geometry is tuned to a threshold of $1500\,e$, while the $50 \times 50\,\mu m^2$ module is tuned to $800\,e$.

(a) r-ϕ projection $25 \times 100\,\mu m^2$

(b) z projection $25 \times 100\,\mu m^2$

(c) r-ϕ projection $50 \times 50\,\mu m^2$

(d) z projection $50 \times 50\,\mu m^2$
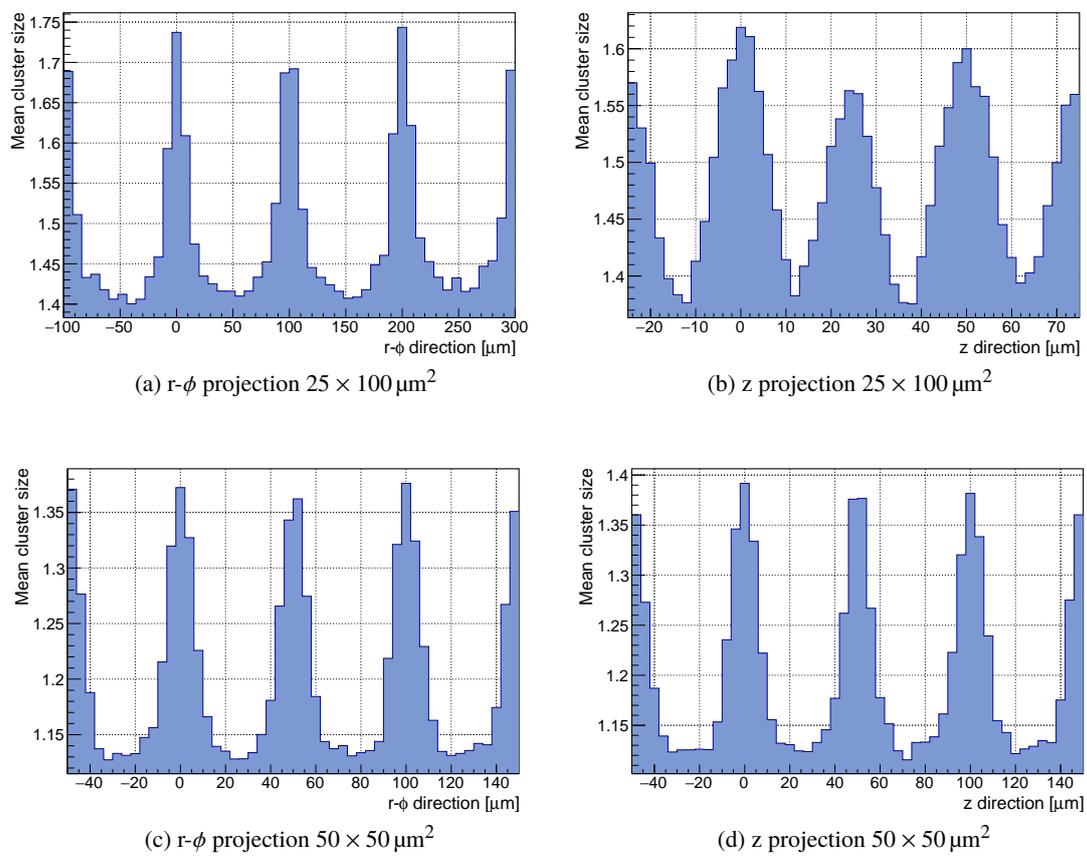
Figure 9.13: Projections in r-ϕ and z direction of the in-pixel cluster size maps of Figure 9.12. Figures (a,b) show the $25 \times 100\,\mu m^2$ geometry while (c,d) show the $50 \times 50\,\mu m^2$ geometry.

(a) Grounded GR

(b) Floating GR
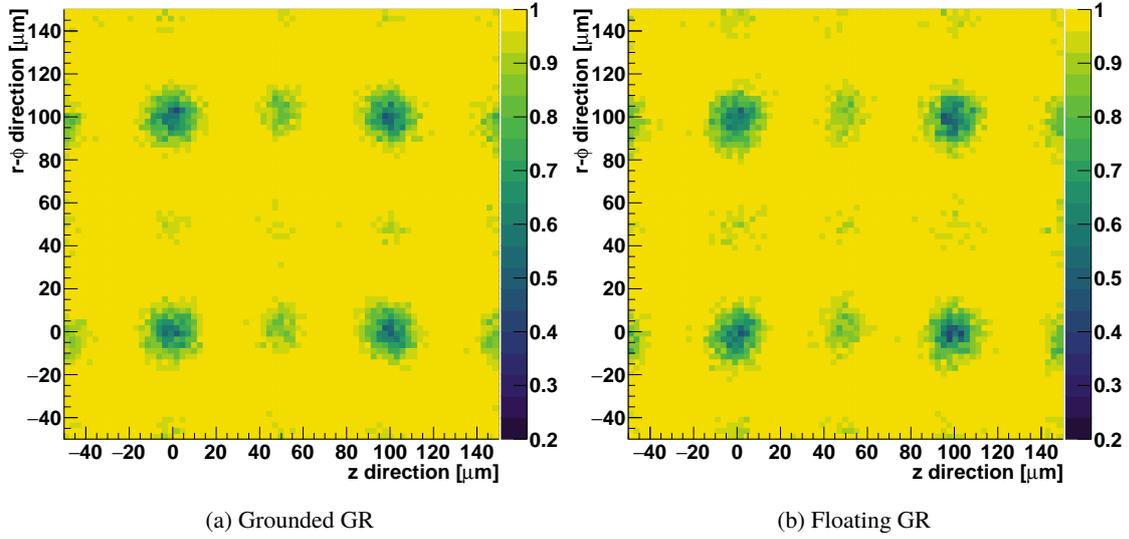
Figure 9.14: In-pixel hit efficiency for 4×4 pixels with $50 \times 50\,\mu m^2$ pixel cell geometry on a $100\,\mu m$ thick sensor using the split design. A bias voltage of $500\,V$ is used. The readout chip is tuned to a threshold of $1200\,e$. The module is irradiated to a fluence of $5 \times 10^{15}\,n_{eq}/cm^2$ at CERN PS. The sensor exhibits an efficiency of $98.0\,\%$ in case the GR is grounded (a) while the efficiency is $97.8\,\%$ when the GR is left floating (b).

9.2 Characterisation after irradiation

9.2.1 Testbeam measurements

All modules are irradiated at either CERN PS employing the Gaussian distributed particle beam or at KIT using a uniform irradiation by implementing a scanning method (see Section 4.3). Quoted doses for PS irradiations correspond to the mean fluence. Again, all measurements are performed with the linear FE of the RD53A readout chip if not explicitly stated differently.

Effect of the punch-through structure on the hit efficiency

The first investigation focuses on a module irradiated at CERN PS to a fluence of $5 \times 10^{15}\,n_{eq}/cm^2$. The $100\,\mu m$ thick sensor incorporates the split BR design, implementing PT dots that are at a floating potential after interconnection to the readout chip. The GR can be either grounded or left floating but it should have no influence on the potential of the PT dots.

Figure 9.14 compares the in-pixel hit efficiency of the module in both configurations. No performance reduction is observed when grounding the GR. In particular, the local inefficiency caused by the PT grid does not significantly change. Noteworthy, the floating PT located at $0\,\mu m$ and $100\,\mu m$ in both directions is causing a significant local performance reduction which can be seen when comparing the effect due to charge sharing only (e.g. at $50\,\mu m/50\,\mu m$) to the one resulting from the combination of charge sharing and PT dots together (e.g. at $100\,\mu m/100\,\mu m$).

The efficiency as a function of the bias voltage is shown in Figure 9.15. In this analysis, only
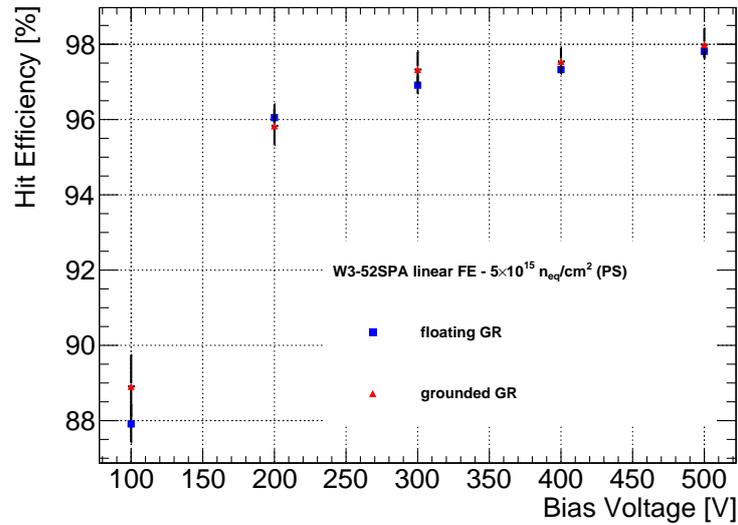
Figure 9.15:   Hit efficiency as a function of the bias voltage of a module irradiated to a fluence of $5 \times 10^{15}\,n_{eq}/cm^2$ at CERN PS. The sensor is $100\,\mu m$ thick and employs the split design. Different GR potentials are compared. The vertical bars correspond to the statistical uncertainty according to Clopper-Pearson.

the pixels for which the fluence is estimated to be in the range of $(5.00 \pm 0.05) \times 10^{15}\,n_{eq}/cm^2$ are considered using the FluenceAnalyser code (see Section 4.2.3). Figure 9.15 shows that high efficiencies around 96 % are already obtained at 200 V but it also shows that 300 V are necessary to achieve the ITk requirement of at least 97 %. Even higher bias voltages of 400 V are necessary to saturate the hit efficiency at around 98 %.

The next module under investigation incorporates a classical layout with a grounded PT system. The sensor of the module is again $100\,\mu m$ thick. It is irradiated to a fluence of $5 \times 10^{15}\,n_{eq}/cm^2$ at KIT. The readout chip is tuned to a threshold of 1000 e. The in-pixel hit efficiency as well as the overall hit efficiency as a function of the bias voltage are shown in Figure 9.16. The resolution of the in-pixel hit efficiency is worse compared to all other results since the measurement is performed at DESY with 5 GeV electrons. The low energetic electrons are subject to more severe multiple scattering, thus deteriorating the resolution. Accordingly, the localised efficiency loss caused by the PT dot is smeared out across a wider area. The contribution from the highly efficient pixel core reduces the relative magnitude of the efficiency loss at the PT dot, while the absolute inefficiency is reflected in the overall hit efficiency. The module reaches an overall hit efficiency of 96.9 % at a bias voltage of 500 V. Although this value is compatible within the uncertainties with the ITk criterion of 97 %, the sensor performs significantly worse than the sensor employing the split design. Again, an operational voltage of 300 V seems realistic given the efficiency of already 96.5 %.

The efficiency reduction due to the PT is assessed by comparing it to a sensor without PT. For this purpose, a module built with a $100\,\mu m$ thick sensor without PT system is irradiated to a fluence of $3 \times 10^{15}\,n_{eq}/cm^2$ at KIT. The readout chip is tuned to a threshold of 1000 e during all measurements. The in-pixel hit efficiency is shown in Figure 9.17a (using a different colour scale

(a) Hit efficiency map

(b) Efficiency vs. bias voltage
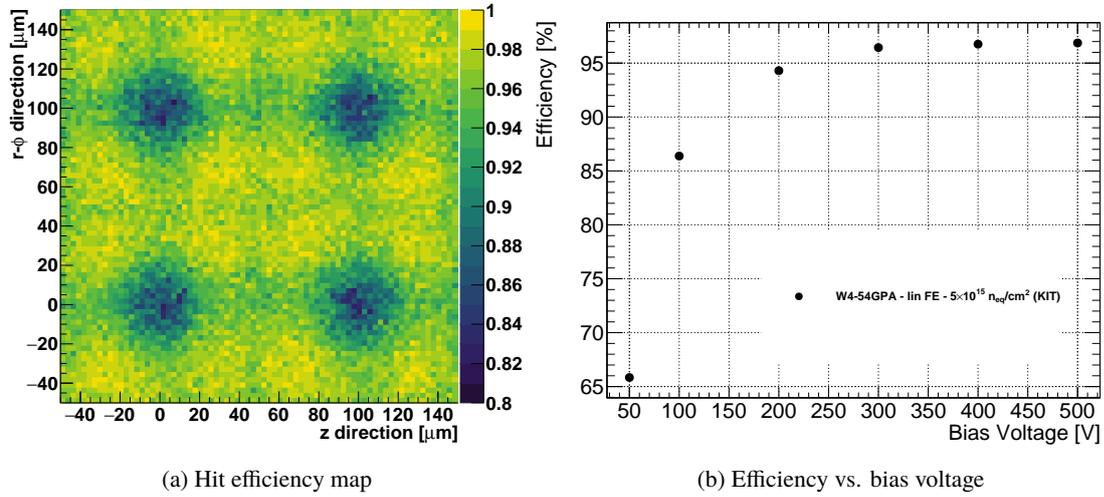
Figure 9.16: The in-pixel hit efficiency for 4×4 pixels is shown in (a) and the hit efficiency as a function of the bias voltage in (b). The module employs a $100\,\mu m$ thick sensor with grounded PT with $50 \times 50\,\mu m^2$ pixels. The module is irradiated to a fluence of $5 \times 10^{15}\,n_{eq}/cm^2$ at KIT and the readout chip is tuned to a threshold of $1000\,e$. The vertical bars correspond to the statistical and systematic uncertainty added in quadrature, and they are smaller than the symbol size.

(a) Hit efficiency map

(b) Efficiency vs. bias voltage
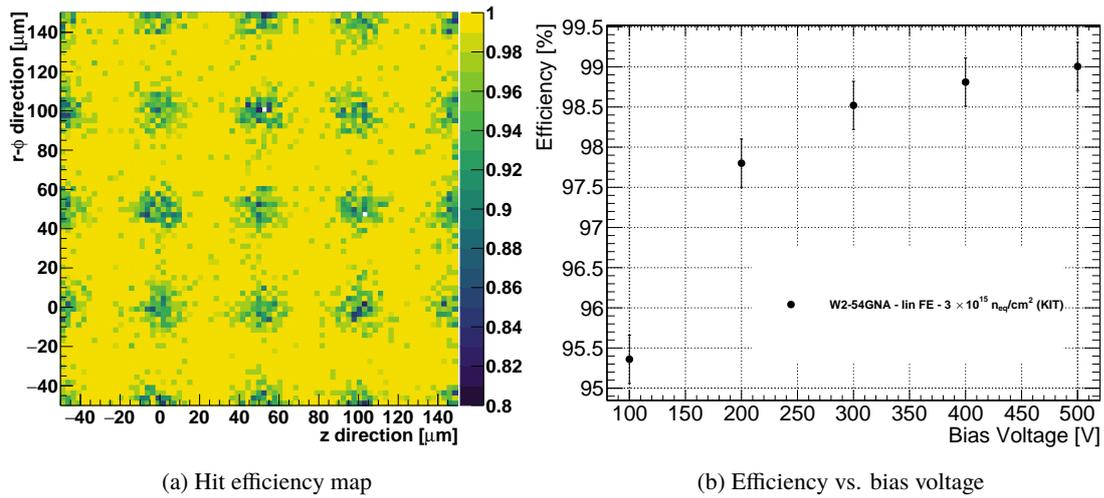
Figure 9.17: The in-pixel hit efficiency for 4×4 pixels is shown in (a) and the hit efficiency as a function of the bias voltage in (b). The module employs a $100\,\mu m$ thick sensor without PT with $50 \times 50\,\mu m^2$ pixels. The module is irradiated to a fluence of $3 \times 10^{15}\,n_{eq}/cm^2$ at KIT and the readout chip is tuned to a threshold of $1000\,e$. The vertical bars correspond to the statistical and systematic uncertainty added in quadrature.
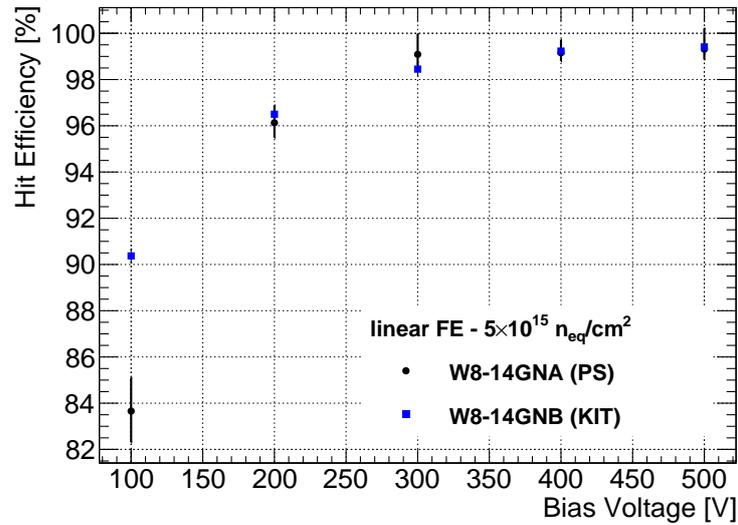
Figure 9.18: Hit efficiency as a function of the bias voltage for two modules employing 150 μm thick sensors with a pixel cell size of $25 \times 100\,\mu\mathrm{m}^2$ without a PT system. The readout chips are tuned to a threshold of 1200 e. Both modules are irradiated to a fluence of $5 \times 10^{15}\,\mathrm{n_{eq}/cm^2}$ but one is irradiated at CERN PS and the other at KIT. The vertical bars correspond to the statistical uncertainty according to Clopper-Pearson.

than the counterparts of modules with PT, also measured at CERN). The local inefficiency in the four pixel region is due to charge-sharing effects and causes an efficiency reduction to about $80-90\,\%$. This is to be compared to the $40-50\,\%$ localised efficiency reduction caused by the floating PT in Figure 9.14 for the sensor with the split design. Furthermore, Figure 9.17b shows the excellent efficiency of the module as a function of the bias voltage. Even a low bias voltage of 100 V leads to a hit efficiency of 95.4 %, the ITk criterion is fulfilled at just 200 V with an efficiency of 97.8 %. The saturation is again reached at around 500 V, but with a significantly higher hit efficiency of 99 %.

Properties of different pixel cell geometries

The performance of the alternative pixel cell geometry of $25 \times 100\,\mu\mathrm{m}^2$ also needs to be evaluated after irradiation. Two modules are used for this investigation. Both modules have a 150 μm thick sensor without PT system. They are irradiated to a fluence of $5 \times 10^{15}\,\mathrm{n_{eq}/cm^2}$, one at KIT, the other one at CERN PS. Both modules are tuned to a threshold of 1200 e for the measurements. The results are compared in Figure 9.18. The module irradiated at KIT shows an excellent performance with a hit efficiency of 96.5 % close to the ITk criterion already at 200 V. The efficiency saturates around 400 V with a value of 99.4 %. The module irradiated at CERN PS is indistinguishable from the module irradiated at KIT at bias voltages larger than 100 V. The two modules only differ significantly at 100 V even considering the additional systematic uncertainties of 0.3 % that are not accounted for in the figure. In this comparison, for the module irradiated at CERN PS, only pixels for which the fluence is estimated to be in the range of $(5.00 \pm 0.05) \times 10^{15}\,\mathrm{n_{eq}/cm^2}$ are considered. In addition, the spatial resolved in-pixel hit efficiency is shown in Figure 9.19. The module irradiated
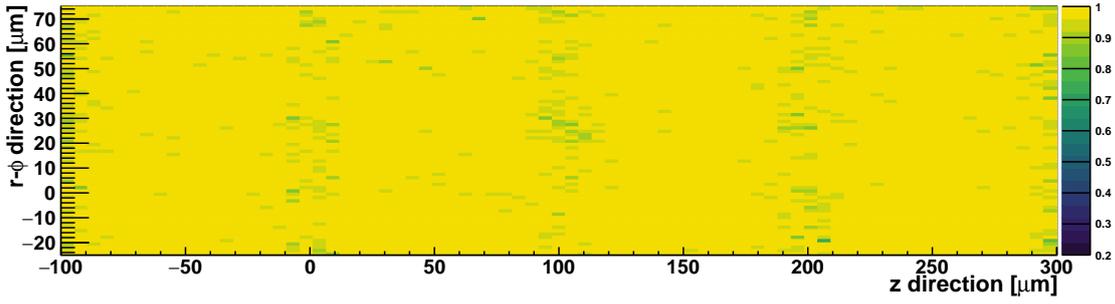
(a) CERN PS
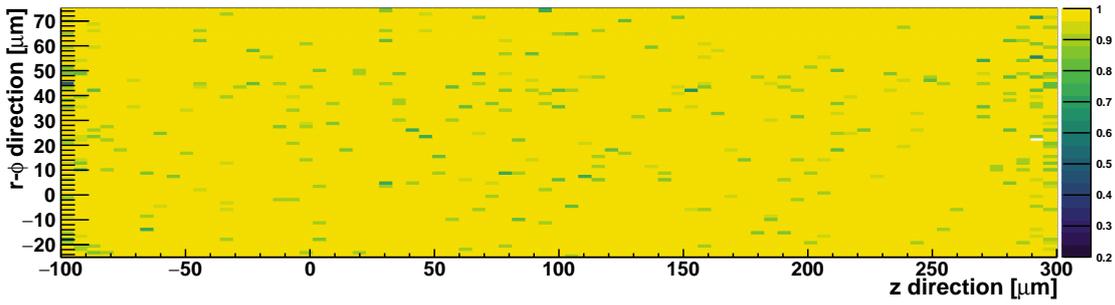
(b) KIT

Figure 9.19: In-pixel hit efficiency for 4×4 pixels with $25\times100\,\mu m^2$ pixel cell geometry. Two modules using $150\,\mu m$ thick sensors without PT system are compared. A bias voltage of 500 V is used and the readout chip is tuned to a threshold of 1200 e in both cases. Both modules are irradiated to a fluence of $5\times10^{15}\,n_{eq}/cm^2$ but in different facilities: (a) at CERN PS and (b) at KIT.

at CERN PS exhibits a more pronounced local inefficiency in the four pixel regions. On the other hand, the localised effect is comparably small ($<8\,\%$ efficiency loss in four pixel region) in both cases and a continuously high efficiency is obtained.

Secondly, the in-pixel cluster size is compared between modules built with sensors of each pixel cell geometry to investigate the possible cross-talk effect discussed in Section 9.1.2 after irradiation. Again, sensors without PT are used for the comparison. An increased cluster size in gaps with or without bump-bond would indicate significant cross-talk between direct neighbours. Also, a continuously high cluster size around two would be an indication of cross-talk. Figure 9.20 shows the in-pixel cluster size for one module per geometry. Both measurements are performed at 500 V. The module with the $25\times100\,\mu m^2$ pixel cell geometry has a $150\,\mu m$ thick sensor and is irradiated to a fluence of $5\times10^{15}\,n_{eq}/cm^2$ at CERN PS. The readout chip is tuned to a threshold of 1200 e. The module with the $50\times50\,\mu m^2$ pixel cell geometry has a $100\,\mu m$ thick sensor and is irradiated to a fluence of $3\times10^{15}\,n_{eq}/cm^2$ at KIT. The readout chip is tuned to a threshold of 1100 e. Neither of the modules exhibit a high overall cluster size being mostly in the range $1-1.5$. There is also no sign of increased cluster size and thus cross-talk between gaps with or without bump-bond.

The comparison with the cluster size before irradiation shows a reduced probability of larger clusters. The effect is caused by the reduced amount of available charge due to trapping and the consequent increased probability for a hit in the cluster to be below threshold. Additionally, the mobility of the charge carriers is reduced at high electrical field strengths that are resulting from the required high
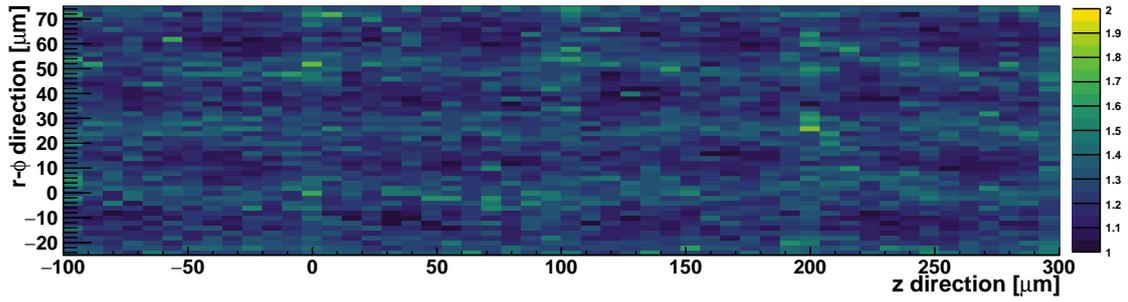
(a) $25 \times 100\,\mu\text{m}^2$
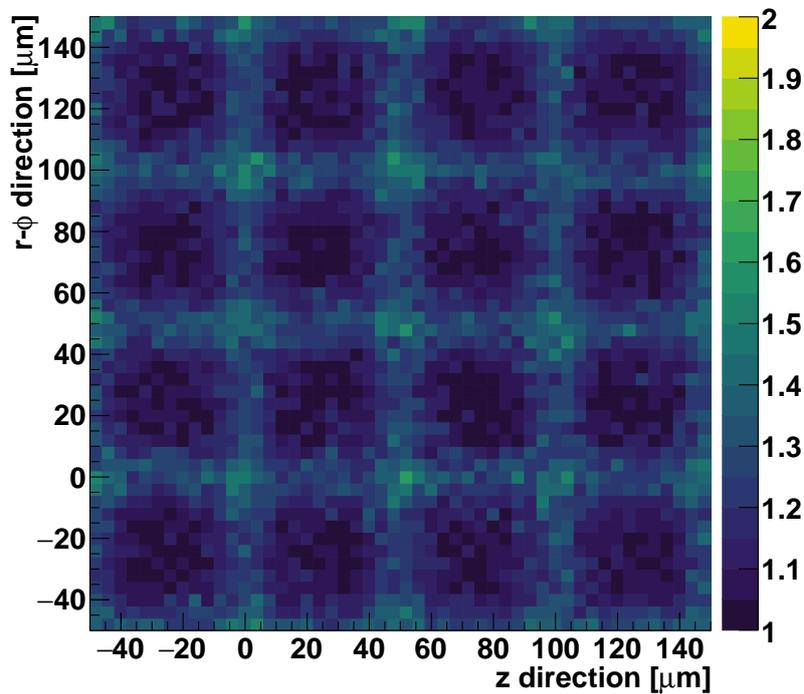
(b) $50 \times 50\,\mu\text{m}^2$

Figure 9.20: In-pixel cluster size for 4×4 pixels. Both measurements are performed at a bias voltage of 500 V. One module has a sensor using the $25 \times 100\,\mu\text{m}^2$ pixel cell geometry with a thickness of 150 µm. It is irradiated to a fluence of 5×10^{15} n_{eq}/cm^2 at CERN PS and the readout chip is tuned to a threshold of 1200 e (a). The other module has a sensor with a thickness of 100 µm using $50 \times 50\,\mu\text{m}^2$ pixels. It is irradiated to a fluence of 3×10^{15} n_{eq}/cm^2 at KIT and the readout chip is tuned to a threshold of 1100 e (b). Neither of the sensors has a PT system.

bias voltages to deplete the irradiated sensors.

Fluence dependence of the hit efficiency

The FluenceAnalyser allows to extract the hit efficiency in different fluence regions over the module in case of inhomogeneous irradiations (see Section 4.2.3). It is used to derive further information about the change of performance as a function of the fluence. The module assembled with the 100 µm

thick sensor W3-52SPA of the SOI-4 production is irradiated to a mean fluence of 5×10^{15} n_{eq}/cm^2 at CERN PS. The fluence profile is measured by the activation of an aluminium foil that is attached to each module during irradiation. The foil is cut into several pieces such that the activity of the different pieces reveal the fluence profile. The fitted two dimensional Gaussian is read by the FluenceAnalyser and the per pixel fluence is calculated. The underlying fluence distribution is displayed in Figure 4.8. The origin of the coordinate system corresponds to the bottom left corner of the module. Figure 4.8 shows that the irradiation beam was misaligned with respect to the module in such a way that the highest fluence is received by the synchronous FE flavour.

The data is acquired at a threshold of 1200 e at CERN. First, the efficiency as a function of the applied bias voltage is extracted for different fluences and is shown in Figure 9.21. Each fluence bin in Figure 9.21 is 1×10^{14} n_{eq}/cm^2 wide. As an additional visualisation, the hit efficiency as a function of the fluence for different bias voltages is shown in Figure 9.22.

Figure 9.21 reveals differences at low bias voltages of $100-200$ V while the saturation values for all fluences are very similar. A more throughout analysis of the fluence dependence is possible in Figure 9.22. The efficiency exhibits an inverse proportionality to the fluence at 100 V. Here, the efficiency is about 95 % at 3×10^{15} n_{eq}/cm^2 and decreases to 60 % at 7.5×10^{15} n_{eq}/cm^2. The fluence dependence of the efficiency strongly decreases with increasing bias voltage. While a clear trend can still be identified at 200 V, the efficiency at the highest bias voltage of 500 V does not depend on the fluence anymore. Therefore, the ITk pixel detector will reliably maintain a sufficient hit efficiency throughout its lifetime. This is also the case for Layer 1, which will experience the highest fluence among the layers instrumented with planar sensors.

The inhomogeneous fluence profile also enables an analysis beyond the intended mean fluence of the irradiation. In this case, the peak fluence nearly reaches a value of 8×10^{15} n_{eq}/cm^2 where the sensor still exhibits an excellent performance with possible operational voltages still in the range of $300-400$ V.
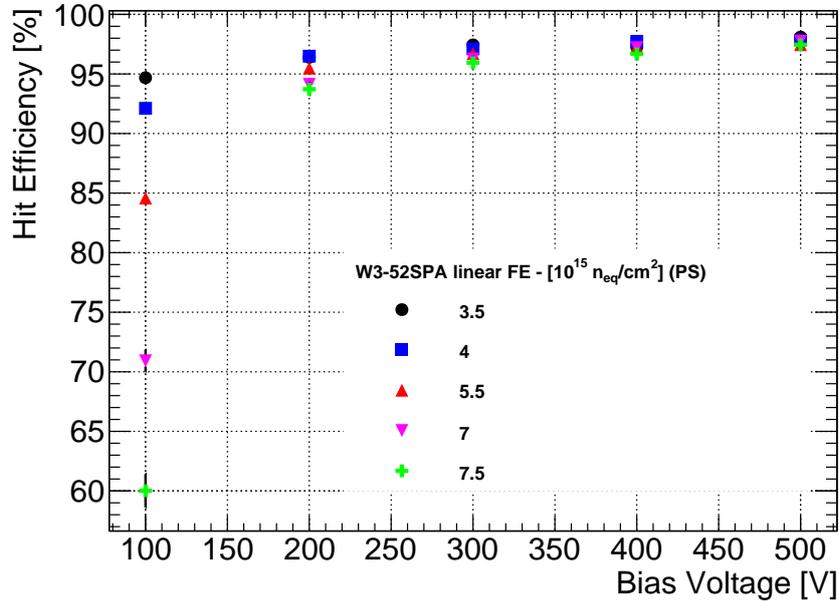
(a) Full scale
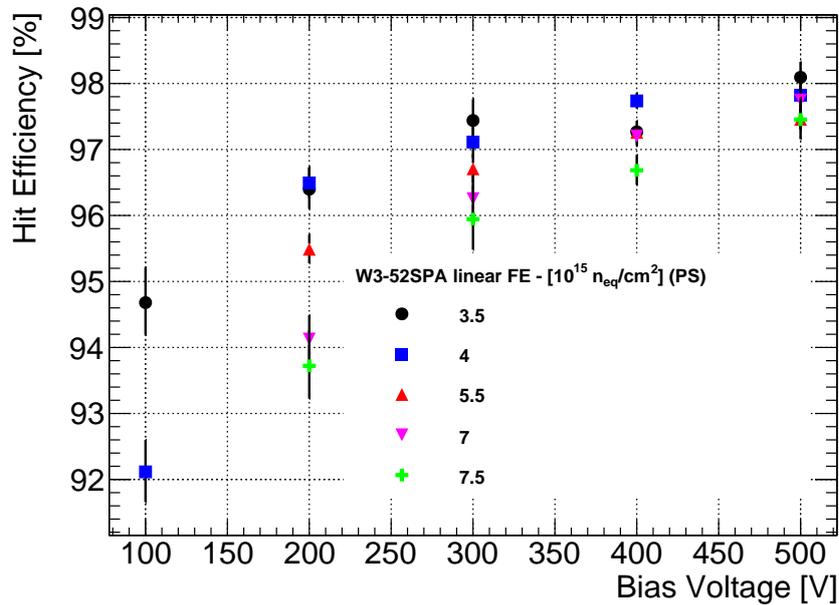
(b) Zoomed in

Figure 9.21: Hit efficiency as a function of the bias voltage for different regions of equal fluence on the same module. The module is tuned to a threshold of 1200 e and employs a 100 μm thick sensor. The vertical bars include the statistical uncertainties according to Clopper-Pearson.
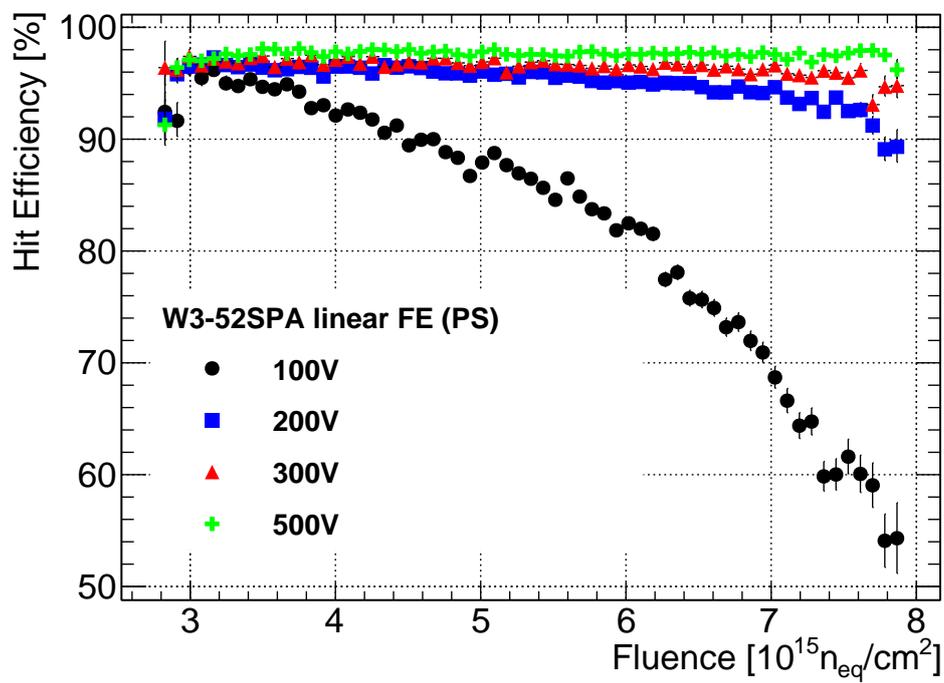
Figure 9.22: Hit efficiency as a function of the fluence for different bias voltages using the data set of Figure 9.21. The module is tuned to a threshold of 1200 e and employs a 100 μm thick sensor. The vertical bars include the statistical uncertainties according to Clopper-Pearson.

CONCLUSION AND OUTLOOK

The tracking detector of the ATLAS experiment enables precise measurements of the trajectories of charged particles. The ITk upgrade is in preparation to maintain and improve the tracking performance despite the increased demand posed by the high luminosity upgrade of the LHC. Namely, high radiation damage and increased track density will be most challenging for the detectors. Many new different technologies were developed during the last years to address the requirements coming from this harsh environment. Among those, the thin planar n-in-p sensors, presented in this thesis, have proven to be a reliable detector component, with excellent charge collection and hit efficiency both before and after irradiation.

As the ITk pixel detector will be built mostly from lightweight serially powered quad modules, the prototyping of devices employing flexible PCBs has started. Multiple double sided tapes were qualified for usage in the high radiation environment of the HL-LHC. The gained experience in the loading of SMD components to the flexes as well as in the overall assembly process of pixel modules with thin sensors will be valuable for the upcoming production.

TCAD simulations have proven to be an important tool assisting the design of pixel detectors. The Synopsis TCAD suite has been employed in this study and, after the first validation phase, various geometries were implemented with the goal of design optimisation. While the introduction of a secondary n$^+$ implant at a high inclination does not significantly improve the breakdown voltage, it has been verified that a decreased p-spray dose and an increased aluminium overhang both increase the breakdown voltage. A first production employing the lower p-spray concentration resulted in significant improvements of the breakdown voltage. The effect of different aluminium overhangs will be evaluated on dedicated test-structures in the SOI-5 production at the semiconductor laboratory of the Max-Planck society. The implant width for the new pixel size of $50 \times 50\,\mu\text{m}^2$ was optimised in a full three dimensional simulation of four quarter pixels. Here, the breakdown voltage was found to be strongly increasing for larger implant sizes. The shielding effect of the implants themselves regarding the HV potential of the backside electrode was found to be responsible for lower p-spray potentials at the same bias voltage, thereby decreasing the potential difference to the n$^+$ pixel implant and improving the breakdown behaviour. On the contrary, larger pixel implants were found in an AC simulation to increase the capacitance of the pixels, dominated by the inter-pixel contribution rather than the bulk capacitance. However, implant widths up to $40\,\mu\text{m}$ are determined to fulfil

the criterion of the RD53 chip, which is a total per pixel input capacitance of less than 50 fF. Nonetheless, the increased capacitance will cause higher electronic noise and thus increase the minimal analog threshold achievable for stable operation of the readout chip. The usage of the `HeavyIon` package enabled the investigation of the charge collection characteristics of the different implant widths. The comparison of different bias voltages showed that smaller implants need higher bias voltages to reach a plateau in CCE, although the value in the plateau is not affected by the implant width. Further investigations at various impact positions within the pixel cell revealed that the collection efficiency of small implants is particularly affected in the inter-pixel region. Also, the effect of saturation of collected charge at larger voltages in smaller implants was confirmed. The ability of larger implants to collect the same amount of charge at lower bias voltages enables a lower operational voltage which decreases the power dissipation significantly. The recommended implant width is between 35 μm and 40 μm, with the upper limit set by the increasing electronic noise and the lower limit defined by the high bias voltages required for efficient charge collection in conjunction with the decreasing breakdown voltage.

The detector performance is not only affected by the previously considered radiation damage itself, but is also subject to an ageing process that depends on both irradiation and thermal history. The effect of this annealing was measured in three different modules irradiated to fluences of 2 and 5×10^{15} n_{eq}/cm^2. Their performance was monitored with testbeam measurements several times during an annealing time of 336 and 480 days at room temperature. While changes are observed in the turn-on behaviour at 100 V, no significant dependence of the possible operational voltage on the annealing time was found. The variations at low voltages are within the expectation based on the Hamburg model. Nonetheless, the operational voltage required for reliable high hit efficiency is not affected and remains at 200 V (100 μm thickness and 2×10^{15} n_{eq}/cm^2 fluence) and 300 V for either 150 μm thickness or 5×10^{15} n_{eq}/cm^2 fluence. Therefore, even if the operation of the cooling system has to be interrupted during the exchange of the two innermost layers of the ITk detector after half of the HL-LHC phase, no major consequences are to be expected for the remaining layers. No performance degradation is expected up to a storage time of one year at room temperature. Therefore, no internal barrier is required between the two innermost and the remaining layers which enables a reduced material budget and relaxed constraints during the exchange of the inner two layers.

Finally, the comprehensive investigation of the very first modules using the RD53A readout chip was presented. Modules built with various sensor designs were tested and assessed based on laboratory and testbeam measurements. Initially, the performance was evaluated before irradiation. The employed thin sensors reach their full charge collection efficiency already at 5 V (10 V) for sensors with a thickness of 100 μm (150 μm). Moreover, the saturation of the hit efficiency is reached at just 10−20 V more. Even before irradiation, the PT dot was found to cause significant localised inefficiencies, deteriorating the overall hit efficiency to (98.5 ± 0.3) %. This efficiency loss is avoided by leaving the potential of the PT dots floating, yielding (99.9 ± 0.3) %. The new split design allows to ground the important innermost GR, while the PT dots are independently floating. Furthermore, the impact of the threshold on the hit efficiency was investigated and a strong dependence of about 1 % per 600 e was found in the investigated threshold range from

600 e to 2100 e. Consequently, low threshold operation will be essential to maintain a high hit efficiency especially after irradiation. The alternative cell geometry of $25 \times 100 \, \mu m^2$ was also evaluated. The resolution improves as expected by a factor of two to about 8 μm in the r-ϕ direction, while no evidence of increased cross-talk in testbeam data was found. Lastly, the performance of irradiated modules was assessed. The 100 μm thick sensors, to be used in the second layer of the ITk, revealed that the standard design with a grounded PT can barely meet the ITk criterion of 97 % yielding (96.9 ± 0.3) % hit efficiency at 500 V and a fluence of $5 \times 10^{15} \, n_{eq}/cm^2$. However, leaving the GR and thereby the PT dots floating is not an option after irradiation since the large current from the edge region causes severe noise in the pixels close to the edges. The split design, instead, enables to ground the GR and still leave the PT floating, thereby increasing the efficiency to (97.8 ± 0.3) %, again at 500 V bias voltage and a fluence of $5 \times 10^{15} \, n_{eq}/cm^2$. The efficiency is further improved to (99.0 ± 0.3) % in modules built with sensors without PT dots. Although this module was only irradiated to $3 \times 10^{15} \, n_{eq}/cm^2$, based on the localised inefficiencies of the split sensor it is concluded that this would also be possible at higher fluences. This is also supported by a special analysis that utilises the inhomogeneous irradiation profile at CERN. Hits and tracks were histogrammed according to the fluence to which the intercepted pixels were exposed. That offers a measurement of the efficiency as a function of the fluence for the various bias voltages. While there is a reduction of hit efficiency at low bias voltages of 100 V, almost no dependence of the hit efficiency on the actual fluence is observed at the highest bias voltage of 500 V. A fluence range $3 - 8 \times 10^{15} \, n_{eq}/cm^2$ was investigated with this analysis employing a single module that was irradiated once to an average fluence of $5 \times 10^{15} \, n_{eq}/cm^2$. In conclusion, thin planar n-in-p sensors are ideally suited to instrument the ITk pixel detector given the observed excellent performance across a wide fluence range up to $8 \times 10^{15} \, n_{eq}/cm^2$.

ACKNOWLEDGEMENTS

GLOSSARY

ALICE   A large ion collider experiment

ATLAS   A toroidal LHC apparatus

BCB   Benzocyclobutene

BCID   Bunch-crossing identification number

BDAQ   Bonn data acquisition system

BG   Bias grid

BPM   Beam profile monitor

BR   Bias ring

BSM   Beyond standard model

CCE   Charge collection efficiency

CERN   Centre européenne pour la recherche nucléaire

CMOS   Complementary metal-oxide semiconductor

CMP   Chemical mechanical polishing

CMS   Compact muon solenoid

CSA   Charge sensitive amplifier

CSC   Cathode strip chamber

DAC   Digital-to-analog converter

DAF   Deterministic annealing filter

DAQ   Data acquisition

DCS   Detector control system

DDV   DCS data viewer

DESY   Deutsches Elektronen Synchrotron

DM   Dark matter

DOFZ   Diffusion oxygenated float zone

DP   Display port

DRIE   Deep reactive ion etching

DUT   Device under test

ECAL   Electromagnetic calorimeter

EMI   Electromagnetic interference

FE   Front end

Fin-FET   Fin field effect transistor

FPGA   Field programmable gate array

FWHM   Full width half maximum

FZ   Float-zone

GBL   General broken line

GR   Guard ring

GUI   Graphical user interface

HCAL   Hadronic calorimeter

HEP   High energy physics

HL-LHC   High luminosity LHC

HLL   Halbleiter Labor der Max-Planck Gesellschaft (MPG Semiconductor Lab)

HV   High voltage

IBL   Insertable B-Layer

ID   Inner detector

IP   Interaction point

ITk   Inner tracker

IZM   Fraunhofer Institute for Reliability and Microintegration

KIT   Karlsruhe Institute of Technology

KOH   Potassium hydroxide

LAr   Liquid argon

LCIO   Linear collider input/output data format

LDO   Low dropout regulator

LEP   Large electron positron collider

LHC   Large hadron collider

LHCb   LHC beauty experiment

LS Long shutdown

LTO Low temperature oxide

LV Low voltage

MAPS Monolithic active pixel sensor

MARLIN Modular analysis and reconstruction for linear collider

MDT Monitored drift tube

MIP Minimal ionising particle

MPP Max-Planck Institut für Physik

MPV Most probable value

NI National Instruments

NIEL Non ionising energy loss

NTC Negative temperature coefficient

PCB Printed circuit board

PS Proton synchrotron

PT Punch through

RD53A RD53A read-out chip

RMS Root mean square

RPC Resistive plate chamber

RT Room temperature

SC Super cycle

SCC Single chip card

SCSI Space charge sign inversion

SCT Semiconductor tracker

SM Standard model

SMD Surface-mount device

SNR Signal-to-noise ratio

SOI Silicon on insulator

SPS Super proton synchrotron

SUSY Supersymmetry

TCT Transient current technique

TGC Thin gap chamber

TLU Trigger logic unit

ToT Time over threshold

TRT Transition radiation tracker

TSMC Taiwan semiconductor manufacturing company

UBM Under-bump metallisation

UoB University of Birmingham

UV Ultra-violet radiation

VBF Vector boson fusion

LIST OF FIGURES

139

LIST OF TABLES

BIBLIOGRAPHY

[1] S. Glashow, *Partial-symmetries of weak interactions*, Nucl. Phys. **22** (1961) 579

[2] Particle Data Group, *Review of Particle Physics*, Chin. Phys. C **40** (2016) 100001, URL http://stacks.iop.org/1674-1137/40/i=10/a=100001

[3] O. S. Bruning et al., *LHC Design Report Vol.1: The LHC Main Ring* (2004)

[4] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003, URL http://stacks.iop.org/1748-0221/3/i=08/a=S08003

[5] CMS Collaboration, *The CMS experiment at the CERN LHC*, JINST **3** (2008) S08004, URL http://stacks.iop.org/1748-0221/3/i=08/a=S08004

[6] ALICE Collaboration, *The ALICE experiment at the CERN LHC*, JINST **3** (2008) S08002, URL http://stacks.iop.org/1748-0221/3/i=08/a=S08002

[7] LHCb Collaboration, *The LHCb Detector at the LHC*, JINST **3** (2008) S08005, URL http://stacks.iop.org/1748-0221/3/i=08/a=S08005

[8] L. Arnaudon et al., *Linac4 technical design report*, Technical report

[9] E. Mobs, *The CERN accelerator complex. Complexe des accélérateurs du CERN*

[10] ATLAS Collaboration, *ATLAS public results* (2018), URL https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2

[11] HL-LHC Project, *LHC project schedule* (2018), URL https://project-hl-lhc-industry.web.cern.ch/content/project-schedule

[12] G. Apollinari et al., *High-Luminosity Large Hadron Collider (HL-LHC): Technical Design Report V. 0.1*, CERN Yellow Reports: Monographs (2017), URL https://cds.cern.ch/record/2284929

[13] ATLAS Collaboration, *Physics at a High-Luminosity LHC with ATLAS*, Technical Report ATL-PHYS-PUB-2013-007, CERN, Geneva

[14] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012) 1

[15] ATLAS Collaboration, *Evidence for the spin-0 nature of the Higgs boson using ATLAS data*, Phys. Lett. B **726** (2013) 120

[16] ATLAS Collaboration, *Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s}$ = 7 and 8 TeV*, J. High Energy Phys. (2016)

[17] ATLAS Collaboration, *Observation of $H \rightarrow b\bar{b}$ decays and VH production with the ATLAS detector*, Phys. Lett. B **786** (2018) 59, arXiv:1808.08238

[18] S. P. Martin, *A Supersymmetry primer* (1997), arXiv:hep-ph/9709356

[19] H. Haberand and G. Kane, *The search for supersymmetry: probing physics beyond the standard model*, Phys. Rep. **117** (1985) 75

[20] CERN AC, *Layout of ATLAS. Dessin representant le detecteur ATLAS*

[21] ATLAS Collaboration, *ATLAS magnet system: Technical Design Report, 1*, CERN, Geneva (1997), URL https://cds.cern.ch/record/338080

[22] ATLAS Collaboration, *The ATLAS Inner Detector commissioning and calibration*, Eur. Phys. J. C **70** (2010) 787

[23] ATLAS Collaboration, *ATLAS inner detector: Technical Design Report, 1*, CERN, Geneva (1997), URL https://cds.cern.ch/record/331063

[24] M. Capeans et al., *ATLAS Insertable B-Layer Technical Design Report*, Technical Report CERN-LHCC-2010-013

[25] I. Peric et al., *The FEI3 readout chip for the ATLAS pixel detector*, Nucl. Instr. Meth. A **565** (2006) 178

[26] M. Garcia-Sciveres et al., *The FE-I4 pixel readout integrated circuit*, Nucl. Instr. Meth. A **636** (2011) S155

[27] E. Vianello et al., *Optimization of double-side 3D detector technology for first productions at FBK*, IEEE Nucl. Sci. Symp. Conf. Rec. (2011) 523

[28] G. Pellegrini et al., *First double-sided 3-D detectors fabricated at CNM-IMB*, Nucl. Instr. Meth. A **592** (2008) 38

[29] M. Backhaus et al., *The upgraded Pixel Detector of the ATLAS Experiment for Run 2 at the Large Hadron Collider*, Nucl. Instr. Meth. A **831** (2016) 65

[30] Y. Unno, *ATLAS silicon microstrip detector system (SCT)*, Nucl. Instr. Meth. A **511** (2003) 58

[31] E. Abat et al., *The ATLAS Transition Radiation Tracker (TRT) proportional drift tube: Design and performance*, JINST **3** (2008) P02013

[32] JD Degenhardt, *ATLAS Transition Radiation Tracker (TRT) Electronics Operation Experience at High Rates*, Technical report

[33] M. Thomson, *Modern Particle Physics*, Cambridge University Press, Cambridge (2013), ISBN: 978-1-107-03426-6

[34] ATLAS Collaboration, *ATLAS liquid-argon calorimeter: Technical Design Report*, CERN, Geneva (1996), URL https://cds.cern.ch/record/331061

[35] G. Sauvage, *The ATLAS electromagnetic calorimeter: construction, commissioning and selected test beam results*, Nucl. Phys. B - Proceedings Supplements **150** (2006) 98

[36] ATLAS Collaboration, *ATLAS calorimeter performance: Technical Design Report*, CERN, Geneva (1996), URL https://cds.cern.ch/record/331059

[37] S. Palestini et al., *The muon spectrometer of the ATLAS experiment*, Nucl. Phys. B - Proceedings Supplements **125** (2003) 337

[38] ATLAS Collaboration, *Technical Design Report for the ATLAS ITk Pixel Detector*, Technical Report ATL-COM-ITK-2017-040, CERN, Geneva

[39] ATLAS Collaboration, *Technical Design Report for the ATLAS Inner Tracker Strip Detector*, Technical report

[40] M. Elsing and D. Giugni, *Final Report of the second ITk Pixel Layout Task Force*, Technical Report ATL-COM-ITK-2018-053, CERN, Geneva

[41] A. Macchiolo et al., *Development of thin pixel sensors and a novel interconnection technology for the SLHC*, Nucl. Instr. Meth. A **591** (2008) 229

[42] RD-53 Collaboration, URL http://rd53.web.cern.ch/RD53/

[43] ATLAS Collaboration, *ATLAS Phase-II Upgrade Scoping Document*, Technical Report CERN-LHCC-2015-020, CERN, Geneva

[44] G. Lutz et al., *Semiconductor radiation detectors*, Springer (1999), ISBN: 978-35-407-1678-5

[45] F. Scholze, H. Rabus, and G. Ulm, *Mean energy required to produce an electron-hole pair in silicon for photons of energies between 50 and 1500 eV*, J. Appl. Phys. **84** (1998) 2926

[46] C. Canali et al., *Electron and hole drift velocity measurements in silicon and their empirical relation to electric field and temperature*, IEEE Trans. Electron Devices **22** (1975) 1045

[47] W. Shockley, *Currents to conductors induced by a moving point charge*, J. Appl. Phys. **9** (1938) 635

[48] S. Ramo, *Currents induced by electron motion*, Proceedings of the IRE **27** (1939) 584

[49] W. Riegler, *Electric fields, weighting fields, signals and charge diffusion in detectors including resistive materials*, JINST **11** (2016) P11002

[50] L. Rossi et al., *Pixel detectors: From fundamentals to applications*, Springer Science & Business Media (2006), ISBN: 978-3-540-28332-4

[51] JR Srour et al., *Review of displacement damage effects in silicon devices*, IEEE Trans. Nucl. Sci. **50** (2003) 653

[52] H. Pernegger, *High mobility diamonds and particle detectors*, Phys. Status Solidi A **203** (2006) 3299

[53] A. Van Ginneken, *Nonionizing energy deposition in silicon for radiation damage studies*, Technical Report FERMILAB-FN-522 (1989), URL http://cds.cern.ch/record/203592

[54] M. Moll, *Radiation damage in silicon particle detectors: Microscopic defects and macroscopic properties*, Ph.D. thesis, Hamburg U. (1999), URL http://www-library.desy.de/cgi-bin/showprep.pl?desy-thesis99-040

[55] G. Lindström et al., *Radiation hard silicon detectors-developments by the RD48 (ROSE) collaboration*, Nucl. Instr. Meth. A **466** (2001) 308

[56] H.F.-W. Sadrozinski et al., *Total dose dependence of oxide charge, interstrip capacitance and breakdown behavior of sLHC prototype silicon strip detectors and test structures of the SMART collaboration*, Nucl. Instr. Meth. A **579** (2007) 769

[57] L. Andricek, *private communication*

[58] A. Bischoff et al., *Breakdown protection and long-term stabilisation for Si-detectors*, Nucl. Instr. Meth. A **326** (1993) 27

[59] Y. Unno et al., *Development of n-in-p silicon planar pixel sensors and flip-chip modules for very high radiation environments*, Nucl. Instr. Meth. A **650** (2011) 129

[60] T. Rohe, *Planung, Bau und Test des Sensor-Bausteins für einen hybriden Silizium-Pixel-Detektor zum Einsatz unter den extremen Strahlenbelastungen am LHC*, Ph.D. thesis, Ludwig Maximilians Universität München (2000), URL https://edoc.ub.uni-muenchen.de/434/

[61] N. Savic, *Development of pixel detectors for the inner tracker upgrade of the ATLAS experiment*, Ph.D. thesis, Ludwig Maximilians Universität München (2018), URL https://publications.mppmu.mpg.de/2017/MPP-2017-390/FullText.pdf

[62] S. Terzo, *Development of radiation hard pixel modules employing planar n-in-p planar silicon sensors with active edges for the ATLAS detector at HL-LHC*, Ph.D. thesis, Technische Universität München (2015), URL https://publications.mppmu.mpg.de/2015/MPP-2015-291/FullText.pdf

[63] P. Weigell, *Investigation of Properties of Novel Silicon Pixel Assemblies Employing Thin n-in-p Sensors and 3D-Integration*, Ph.D. thesis, Technische Universität München (2013), URL https://publications.mppmu.mpg.de/2013/MPP-2013-5/FullText.pdf

[64] M. Beimforde, *Development of thin sensors and a novel interconnection technology for the upgrade of the atlas pixel system*, Ph.D. thesis, Technische Universität München (2010), URL https://publications.mppmu.mpg.de/2010/MPP-2010-115/FullText.pdf

[65] K. Bean, *Anisotropic etching of silicon*, IEEE Trans. Electron Devices **25** (1978) 1185

[66] A. Nenandic and R. W. Pasco, *Method of chemical-mechanical polishing an electronic component substrate and polishing slurry therefor*, IEEE Trans. Electron Devices **25** (1992) 1185

[67] F. Lärmer and A. Schilp, *Verfahren zum anisotropen Ätzen von Silicium*, Deutsches Patent DE **42** (1994) 045

[68] G. Kramberger et al., *Investigation of irradiated silicon detectors by edge-TCT*, IEEE Trans. Nucl. Sci. **57** (2010) 2294

[69] I. Mandic et al., *Edge-TCT measurements with the laser beam directed parallel to the strips*, JINST **10** (2015) P08004

[70] I. Mandic et al., *TCT measurements of irradiated strip detectors with a focused laser beam*, JINST **8** (2013) P04016

[71] Particulars, *Advanced measurement systems*, URL http://www.particulars.si

[72] R. Turchetta et al., *A monolithic active pixel sensor for charged particle tracking and imaging using standard VLSI CMOS technology*, Nucl. Instr. Meth. A **458** (2001) 677

[73] I. Peric, *A novel monolithic pixelated particle detector implemented in high-voltage CMOS technology*, Nucl. Instr. Meth. A **582** (2007) 876

[74] N. Wermes, *Pixel detectors for particle physics and imaging applications*, Nucl. Instr. Meth. A **512** (2003) 277

[75] I. Peric et al., *High-voltage pixel detectors in commercial CMOS technologies for ATLAS, CLIC and Mu3e experiments*, Nucl. Instr. Meth. A **731** (2013) 131

[76] M. Barbero et al., *FE-I4 ATLAS pixel chip design*, PoS (2009) 027

[77] M. Karagounis, *Development of the ATLAS FE-I4 pixel readout IC for b-layer Upgrade and Super-LHC*, Technical report (2008)

[78] FE-I4 Collaboration, *FE-I4B Integrated Circuit Guide*, Technical Report 03 (2012)

[79] RD53 Collaboration, *The RD53A Integrated Circuit*, Technical Report 3.42 (2018)

[80] ATLAS Collaboration, *ATLAS pixel detector: Technical Design Report*, Technical Report CERN-LHCC-98-013, Geneva (1998), URL http://cds.cern.ch/record/381263

[81] M. Backhaus et al., *Development of a versatile and modular test system for ATLAS hybrid pixel detectors*, Nucl. Instr. Meth. A **650** (2011) 37

[82] J. Große-Knetter, *Vertex measurement at a hadron collider. The ATLAS pixel detector*, Habilitation, Universität Bonn (2008), URL http://www.iaea.org/inis/collection/NCLCollectionStore/_Public/39/088/39088841.pdf

[83] D. Cussans, *Description of the JRA1 Trigger Logic Unit (TLU), v0.2c*, EUDET-Memo-2009-4 (2009)

[84] J. Rieger, *Comparison of Thin n- and p-type Bulk Silicon Sensors*, Master thesis, Georg-August-Universität Göttingen (2012)

[85] BDAQ Software and Hardware, URL https://gitlab.cern.ch/silab/bdaq53

[86] S. Seltzer, *Stopping-Powers and Range Tables for Electrons, Protons, and Helium Ions, NIST Standard Reference Database 124* (1993), URL http://www.nist.gov/pml/data/star/index.cfm

[87] Live Chart of Nuclides, URL https://www-nds.iaea.org/relnsd/vcharthtml/VChartHTML.html

[88] *DESY Test Beam - Deutsches Elektron Synchrotron DESY*, URL testbeam.desy.de/

[89] T. Behnke et al., *Test beams at DESY*, EUDET-Memo-2007-11 (2007)

[90] Particle Data Group, *Review of particle physics*, Chin. Phys. C **38** (2014) 090001

[91] G. Troska, *Development and operation of a testbeam setup for qualification studies of ATLAS pixel sensors*, Ph.D. thesis, Technische Universität Dortmund (2012), URL https://eldorado.tu-dortmund.de/handle/2003/29351

[92] *CERN Testbeam - secondary beam areas of the PS and SPS machines*, URL http://sba.web.cern.ch/sba/

[93] J. Baudot et al., *First test results of MIMOSA-26, a fast CMOS sensor with integrated zero suppression and digitized output*, IEEE Nucl. Sci. Symp. Conf. Rec. (2009) 1169

[94] I. Rubinskiy et al., *An EUDET/AIDA pixel beam telescope for detector development*, Phys. Procedia **37** (2012) 923

[95] H. Jansen et al., *Performance of the EUDET-type beam telescopes*, EPJ Tech. Instrum. **3** (2016) 7

[96] H. Perrey, *EUDAQ and EUTelescope Software Frameworks for Testbeam Data Acquisition and Analysis*, PoS (TIPP2014) (2014) 353

[97] T. Bisanz et al., *EUTelescope 1.0: Reconstruction Software for the AIDA Testbeam Telescope*, Technical Report AIDA-NOTE-2015-009, URL http://cds.cern.ch/record/2000969

[98] A Generic Pixel Telescope Data Analysis Framework, URL http://eutelescope.web.cern.ch

[99] V. Blobel, *Software alignment for tracking detectors*, Nucl. Instr. Meth. A **566** (2006) 5

[100] S. Fleischmann, *Track Reconstruction in the ATLAS Experiment: The deterministic annealing filter*, Ph.D. thesis, Universität Wuppertal

[101] R. Kalman et al., *A new approach to linear filtering and prediction problems*, J. Basic Eng. **82** (1960) 35

[102] C. Kleinwort, *General Broken Lines as advanced track fitting method*, Nucl. Instr. Meth. A **673** (2012) 107

[103] M. George, *Testbeam Measurements with Pixel Sensors for the ATLAS Insertable b-Layer Project*, Ph.D. thesis, Georg-August-Universität Göttingen (2014), URL https://ediss.uni-goettingen.de/handle/11858/00-1735-0000-0022-5F07-A

[104] S. Terzo, *FE-I4 Analyzer*, URL https://github.com/terzo/fei4Analyzer

[105] C. J. Clopper and E. S. Pearson, *The use of Confidence or Fiducial Limits Illustrated in the Case of the Binomial*, Biometrika **26** (1934) 404

[106] A. Dierlamm, *Untersuchung zur Strahlungshärte von Siliziumsensoren*, Ph.D. thesis (2003), URL https://ekp-invenio.physik.uni-karlsruhe.de/record/45039

[107] P. Dervan et al., *Upgrade to the Birmingham Irradiation Facility*, Nucl. Instr. Meth. A **796** (2015) 80

[108] P. Allport et al., *Recent results and experience with the Birmingham and MC40 irradiation facility*, JINST **12** (2017) C03075

[109] B. Gkotse et al., *The Beam Profile Monitoring System for the IRRAD Proton Facility at the CERN PS East Area*

[110] F. Ravotti et al., *A New High-Intensity Proton Irradiation Facility at the CERN PS East Area*, PoS (2014) 354

[111] L. Snoj et al., *Computational analysis of irradiation facilities at the JSI TRIGA reactor*, Int. J. Rad. Appl. Instrum. A **10** (2012) 483

[112] L. Snoj et al., *Computational analysis of irradiation facilities at the JSI TRIGA reactor*, Appl. Radiat. Isot. **70** (2012) 483

[113] I. Mandic et al., *Bulk Damage in DMILL npn Bipolar Transistors Caused by Thermal Neutrons Versus Protons and Fast Neutrons*, IEEE Trans. Nucl. Sci. **51** (2004) 1752

[114] 3M Industrial Adhesives and Tapes Division, *VHB-5909*, URL https://multimedia.3m.com/mws/media/946719O/3m-vhb-thin-foam-tape-5906-5907-5908-5909.pdf

[115] PPI Adhesive Products Limited, *7011-DS*, URL http://www.polydecoupe.com/admin/general/upload/dt-gamme-electrotechnique/24-01-2013_7011-_6.pdf

[116] Tesa SE, *tesafix 4962*, URL https://www.tesa.com/de-ch/industrie/tesafix-4962.html

[117] ASTM International, *Test Method for Apparent Shear Strength of Single-Lap-Joint Adhesively Bonded Metal Specimens by Tension Loading (Metal-to-Metal)* URL https://doi.org/10.1520/d1002-10

[118] BGS, *BGS Beta-Gamma-Service GmbH and Co. KG* (2018), URL https://en.bgs.eu

[119] S. Kühn, *Results of larger structures prototyping for the Phase-II upgrade of the pixel detector of the ATLAS experiment*, URL https://indico.cern.ch/event/669866/contributions/3235377/attachments/1770453/2877537/ProtoypingPixel-Pixel18-SK.pdf

[120] Synopsys, Inc., URL https://www.synopsys.com

[121] C. Hu et al., *FinFET-a self-aligned double-gate MOSFET scalable to 20 nm*, IEEE Trans. Electron Devices **47** (2000) 2320

[122] Synopsys Inc., *Sentaurus^{TM} Device User Guide* (Version L-2016.03, March 2016)

[123] Y. Takeda and T. P. Pearsall, *Failure of Matthiessen's rule in the calculation of carrier mobility and alloy scattering effects in Ga0.47In0.53As*, Electron. Lett. **17** (1981) 573

[124] G. Masetti, M. Severi, and S. Solmi, *Modeling of carrier mobility against carrier concentration in arsenic-, phosphorus-, and boron-doped silicon*, IEEE Trans. Electron Devices **30** (1983) 764

[125] C. Lombardi et al., *A physically based mobility model for numerical simulation of nonplanar devices*, IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst. **7** (1988) 1164

[126] C. Canali et al., *Electron and hole drift velocity measurements in silicon and their empirical relation to electric field and temperature*, IEEE Trans. Electron Devices **22** (1975) 1045

[127] W. Bludau, A. Onton, and W. Heinke, *Temperature dependence of the band gap of silicon*, J. Appl. Phys. **45** (1974) 1846

[128] D. Klaassen, J. Slotboom, and H. de Graaff, *Unified apparent bandgap narrowing in n- and p-type silicon*, Solid-State Electron. **35** (1992) 125

[129] M. Tyagi and R. V. Overstraeten, *Minority carrier recombination in heavily-doped silicon*, Solid-State Electron. **26** (1983) 577

[130] J. Fossum et al., *Carrier recombination and lifetime in highly doped silicon*, Solid-State Electron. **26** (1983) 569

[131] G. A. M. Hurkx, D. B. M. Klaassen, and M. P. G. Knuvers, *A new recombination model for device simulation including tunneling*, IEEE Trans. Electron Devices **39** (1992) 331

[132] R. Eber, *Investigations of new Sensor Designs and Development of an effective Radiation Damage Model for the Simulation of highly irradiated Silicon Particle Detectors*, Ph.D. thesis (2013), URL `https://ekp-invenio.physik.uni-karlsruhe.de/record/48328`

[133] R. V. Overstraeten and H. D. Man, *Measurement of the ionization rates in diffused silicon p-n junctions*, Solid-State Electron. **13** (1970) 583

[134] F. Moscatelli et al., *Effects of Interface Donor Trap States on Isolation Properties of Detectors Operating at High-Luminosity LHC*, IEEE Trans. Nucl. Sci. **64** (2017) 2259

[135] J. Zhang et al., *Investigation of X-ray induced radiation damage at the Si-SiO 2 interface of silicon sensors for the European XFEL*, JINST **7** (2012) C12012, URL `http://stacks.iop.org/1748-0221/7/i=12/a=C12012`

[136] R. H. Richter, *private communication*

[137] H. Bichsel, *Straggling in thin silicon detectors*, Rev. Mod. Phys. **60** (1988) 663

[138] A. Chilingarov, *Temperature dependence of the current generated in Si bulk*, JINST **8** (2013) P10003, URL `http://stacks.iop.org/1748-0221/8/i=10/a=P10003`

[139] J.-C. Beyer et al., *TCAD simulations of pixel sensors for the ATLAS ITk upgrade and performance of annealed planar pixel modules*, JINST **13** (2018) C11001, URL `http://stacks.iop.org/1748-0221/13/i=11/a=C11001`

[140] J.-C. Beyer et al., *Characterisation of novel thin n-in-p planar pixel modules for the ATLAS Inner Tracker upgrade*, JINST **13** (2018) C01009, URL `http://stacks.iop.org/1748-0221/13/i=01/a=C01009`

[141] A. Grummer, *private communication* (University of New Mexico)

[142] G. Battistoni et al., *The FLUKA code: Description and benchmarking*, AIP Conf. Proc. **896** (2007) 31

[143] A. Ferrari et al., *FLUKA: A multi-particle transport code (Program version 2005)* (2005)

[144] S. Baranov et al., *Estimation of Radiation Background, Impact on Detectors, Activation and Shielding Optimization in ATLAS*, Technical report, CERN, Geneva

[145] T. Sjöstrand et al., *An Introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191** (2015) 159, `arXiv:1410.3012`

[146] A. D. Martin et al., *Parton distributions for the LHC*, Eur. Phys. J. **C63** (2009) 189, `arXiv:0901.0002`

[147] ATLAS Collaboration, *Summary of ATLAS Pythia 8 tunes* (2012), ATL-PHYS-PUB-2012-003

[148] ATLAS Collaboration, *Radiation Background Simulation for Run-2*, URL `https://twiki.cern.ch/twiki/bin/viewauth/Atlas/RadiationBackgroundSimulationsRun2`

[149] ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s}$ = 8 TeV using the ATLAS detector at the LHC*, Eur. Phys. J. **C76** (2016) 653, arXiv:1608.03953

[150] Iseg Spezialelektronik GmbH, *Rossendorf, Germany*, URL www.iseg-hv.com

[151] I. Gorelov, *A leakage current-based measurement of the radiation damage in the ATLAS Pixel Detector*, JINST **10** (2015) C04024, URL http://stacks.iop.org/1748-0221/10/i=04/a=C04024

[152] A. Grummer et al., *Measurements of ATLAS Pixel Detector leakage current data compared to Hamburg Model predictions*, Technical Report ATL-INDET-INT-2018-011, CERN, Geneva

[153] A. Macchiolo et al., *Optimization of thin n-in-p planar pixel modules for the ATLAS upgrade at HL-LHC*, JINST **12** (2017) C01024

[154] N. Savic et al., *Thin n-in-p planar pixel modules for the ATLAS upgrade at HL-LHC*, Nucl. Instr. Meth. A **A845** (2017) 154