



UNIVERSITY OF
LIVERPOOL



UNIVERSITÉ
DE GENÈVE

Development of CMOS Sensors for High-Luminosity ATLAS Detectors

Thesis submitted to the University of Liverpool in accordance
with the requirements for the degree of Doctor in Philosophy
by

Lingxin Meng

5th September 2018

Physics Department
University of Liverpool



Supervisors: Professor Dr. Joost Vossebeld, University of Liverpool
Professor Dr. Gianluigi Casse, University of Liverpool
Professor Dr. Giuseppe Iacobucci, Université de Genève
Dr. Mathieu Benoit, Université de Genève

Abstract

In 2024 the Large Hadron Collider will undergo a High Luminosity upgrade in order to better accumulate data and improve statistics. The existing inner detector of the ATLAS experiment cannot cope with this kind of luminosities. In addition, central silicon trackers will meet their end of lifetime and be severely damaged by radiation. For an accurate description of the detectors a radiation damage pixel digitiser was developed for the ATLAS simulation framework. The digitiser is discussed in detail and a comparison of simulation with data is shown.

For the High Luminosity upgrade, new silicon sensors are currently under development to completely replace the present ATLAS Inner Detector. One of the candidates for silicon pixel sensors is the high voltage complementary metal-oxide-semiconductor (HV-CMOS) sensor. Compared with traditional silicon pixel sensors, it has integrated electronics and can be fabricated using industrially standard CMOS processes. The main advantage is therefore the low cost due to inexpensive glue-bonding to the readout-chip and the commercial availability, which plays an important role in replacing a large area of detectors. In the scope of this work Technology Computer Aided Design (TCAD) simulations on two different technologies, ams 350 nm and LFoundry 150 nm, are carried out, each with a different focus. Measurements of an HV-CMOS demonstrator sensor are also presented here.

Acknowledgements

First, I would like to express my sincere gratitude to Prof. Dr. Joost Vossebeld, Prof. Dr. Gianluigi Casse, Prof. Dr. Giuseppe Iacobucci and Dr. Mathieu Benoit for giving me the chance to perform the work for this thesis, throughout which I gained invaluable experience. Without their supervision and guidance this work would not have been possible. Dr. Eva Vilella enriched me with valuable ideas and explanations on the H35DEMO and Sam Powell helped me with the readout system. Many thanks to Dr. Benjamin Nachman for discussions on the digitiser. I owe thanks to everybody who proofread this thesis and provided useful feedback and I appreciate the effort of everybody who helped me to complete this thesis in anyway.

Special thanks go to Daniel, for elating me with HV-CMOS sensors, making this opportunity possible and convincing me to seize it; and Bane for indispensable IT help desk services.

My heartfelt gratitude goes to my fellow PhD students, office colleagues, lunch companions and people from the CERN ski club who contributed to the fun side of my PhD.

Many thanks to my wonderful friends for their listening, encouragement and motivation, for showing me a life outside a PhD and sharing their adventures with me.

Last but not least, I would like to thank my family for their eternal understanding and support.

Contents

1	Introduction	1
2	The ATLAS Experiment at the LHC	5
2.1	The Large Hadron Collider	5
2.2	The ATLAS Detector	8
2.2.1	Inner Detector	10
2.2.1.1	Pixel Detector	11
2.2.1.2	Semiconductor Tracker	12
2.2.1.3	Transition Radiation Tracker	12
2.2.2	Calorimeters	12
2.2.2.1	LAr Calorimeters	13
2.2.2.2	Tile Calorimeter	14
2.2.3	Muon Spectrometer	15
2.3	ITk — HL-ATLAS Inner Detector Upgrade	16
2.4	Physics at the ATLAS Experiment	19
2.4.1	The Standard Model	19
2.4.2	Beyond the Standard Model	22
2.4.3	Physics at the High Luminosity LHC	22
3	Particle Interactions in Silicon	25
3.1	Energy Loss of Particles in Matter	25
3.1.1	Ionising Energy Loss	27
3.1.2	Radiative Energy Loss	30
3.1.3	Non-ionising Energy Loss	31
3.2	Silicon Particle Detector	33
3.2.1	Physical Properties of Silicon	33

CONTENTS

3.2.2	Impurities and Doping	39
3.2.3	P-N Junction	42
3.2.4	Current-Voltage Characteristics	44
3.2.5	Capacitance-Voltage Characteristics	46
3.2.6	Hall Effect	46
3.2.7	Metal-Oxide-Semiconductor	47
3.2.8	Semiconductor Fabrication Process	48
3.2.9	Charge Collection	50
3.2.9.1	Shockley-Ramo Theorem	52
3.3	Radiation Damage in Silicon Sensors	54
3.3.1	Surface Damage	55
3.3.2	Bulk Damage	55
3.3.2.1	Effective Doping Concentration	56
3.3.2.2	Leakage Current	58
3.3.2.3	Charge Collection Efficiency	59
3.3.2.4	Donor and Acceptor Removal	59
3.3.2.5	Annealing	60
4	Radiation Damage Pixel Digitiser	63
4.1	Development of the Digitiser	65
4.1.1	Electric Field	66
4.1.2	Charge Position	68
4.1.3	Time to Electrode	72
4.1.4	Ramo Potential	73
4.1.5	Hall Effect	73
4.1.6	Radiation Damage	75
4.1.7	Chunk Correction	76
4.1.8	Diffusion and Mobility	76
4.1.9	Time over Threshold	78
4.2	Test Beam at SLAC	79
4.2.1	Track Reconstruction	81
4.2.2	Test Beam Simulation	89
4.2.3	Comparison between Data and Simulation	91

CONTENTS

5	High Voltage CMOS Sensors	99
5.1	The ams 350 nm Demonstrator	101
5.2	TCAD Simulation	101
5.3	Simulations of H35DEMO	107
5.3.1	Electric Field	110
5.3.2	Current-Voltage Characteristics	111
5.3.3	Depletion Region	111
5.3.4	Capacitance-Voltage Characteristics	113
5.3.5	Radiation Damage	115
5.3.6	Charge Collection	117
5.3.7	Edge TCT Simulation	121
5.4	Simulations of Devices in the LFoundry 150 nm Technology . . .	131
5.4.1	Doping Profiles	131
5.4.2	Capacitance	137
5.4.3	Size of the PSUB	140
5.4.4	I-V Curve	144
5.4.5	Other Cross-sections	144
5.5	Measurements of H35DEMO	148
5.5.1	I-V Measurements	148
5.5.2	Comparison with Simulation	153
5.5.2.1	I-V Characteristics	153
5.5.2.2	Depletion Depth	158
5.5.3	Test Beam Measurements of the Analog Matrices	159
5.5.3.1	Samples	162
5.5.3.2	Results	163
5.5.4	Laboratory Tests of the NMOS Matrix	165
5.5.4.1	The Measurement Setup	167
5.5.4.2	Measurements with Pulse Injection	169
5.5.4.3	The Scintillator Setup and Source Measurements	174
6	Summary and Outlook	183
	Appendix A Radiation Damage Pixel Digitiser	187
A.1	Time-over-Threshold Encoding	187
A.2	Test Beam Plots	188

CONTENTS

Appendix B TCAD Comparisons	203
B.1 TCAD Models	203
B.2 Normal vs. Reduced Physics	203
B.3 Periodic Boundary Condition	204
B.4 Alternative Cross-section	204
B.5 Radiation Models	213
B.6 Focus Length of the Laser	217
Appendix C Measurements	219
C.1 I-V Curves of Test Beam Samples	219
C.2 Measurements of the NMOS Matrix	222
List of Figures	i
List of Acronyms	vii
Bibliography	xiii

Chapter 1

Introduction

Curiosity killed the cat¹. But it is curiosity that keeps the scientific progress of humankind going. We look into the sky and see the entire history of the universe in front of our eyes. We question its origin, its destiny and what it is made of. The search for the smallest building blocks of matter is an eternal ongoing quest. The ancient Greeks named these atoms, literally unsplittable. Yet today, we split what we call the unsplittable in large amounts to satisfy our high energy demands. To be precise, it is the nucleus of an atom of certain elements we split to harvest the energy while the rest of the atom, the electron, is what delivers this energy to us. Today, more subatomic particles have been discovered. What we now call the elementary particles are described by the Standard Model of particle physics. They do not only make up nucleons, components of the nucleus, but also more exotic composites (Section 2.4).

All the discoveries of tiny structures could not be made without an adequate probe. As things get smaller the probes needed to resolve them also decrease in size: from wavelength of visible light to (de Broglie) wavelength in the range of the weak interaction, from photons of a few electron volt (eV) to particles of several trillion eV.

After the gold foil experiment to proof the structure of the atom, in a speech before the Royal Society in 1927 [1], Ernest Rutherford expressed his wish for his fellow scientists to find a way to accelerate particles beyond their natural decay energies. Following up this request, already in 1930 the Cockcroft-Walton accelerator was born [2], the required high voltage originates from a

¹No animal was harmed during the process of this thesis.

voltage multiplier. The Van de Graaff accelerator followed in 1931 [3], based on the principles of the Van de Graaff generator. Apart from accelerators using static electric fields, in 1928 and 1930, Wideröe and Lawrence invented the linear accelerator and the cyclotron, respectively [4, 5]. Around 1944 and 1945, Vladimir Veksler and Edwin McMillan invented the synchrotron independently [6, 7] which enables the building of large-scale storage rings. Today, the Large Hadron Collider (LHC) (Section 2.1) remains the largest particle storage ring and accelerator on earth.

Parallel to the development of accelerators, particle detectors also evolved. The first detection of ionising radiation dates back to 1717, when Johann Heinrich Schulze discovered the darkening of silver nitrate in sunlight [8]. Based on this principle, photographic plates were developed, which were used by Victor Franz Hess around 1910 in the discovery of cosmic rays where the plates were sent into the higher atmosphere a balloon [9]. This led to a Nobel Prize in 1936 [10]. They also verified the existence of x-rays by Röntgen in 1895 [11], led to the discovery of radioactivity by Becquerel in 1896 [12] and, about 50 years later, to the discovery of pions in cosmic rays [13]. All were awarded Nobel Prizes.

In the early days of particle physics, cloud chambers, invented by Charles Wilson around 1911 [14], played a major role in discoveries of new particles like the positron, the muon and the kaon [15, 16, 17]. Also bubble chambers, invented by Donald Glaser [18], bestowed us with discovery of the neutral currents shown in Fig. 1.1 [19] which resulted in the discoveries of the $W^\pm$ and Z^0 bosons [20, 21, 22, 23]. However, these particle detectors have a major drawback which is the readout. It is necessary to conserve the event by photographs which are then evaluated by human eye. This method is slow and unreliable compared to today's standard.

The detector that revolutionised it all was the multi-wire proportional chamber (MWPC), invented by Georges Charpak in 1968 [24] who earned a Nobel Prize in 1992. The MWPC works similar to the Geiger-Müller counter, but has a grid of wires to resolve the particle's position. The most remarkable thing is however, that each wire can be read out electronically. Compared to bubble chambers, this allows the data to be recorded a thousand times faster, stored electronically and evaluated by computers in an instant.

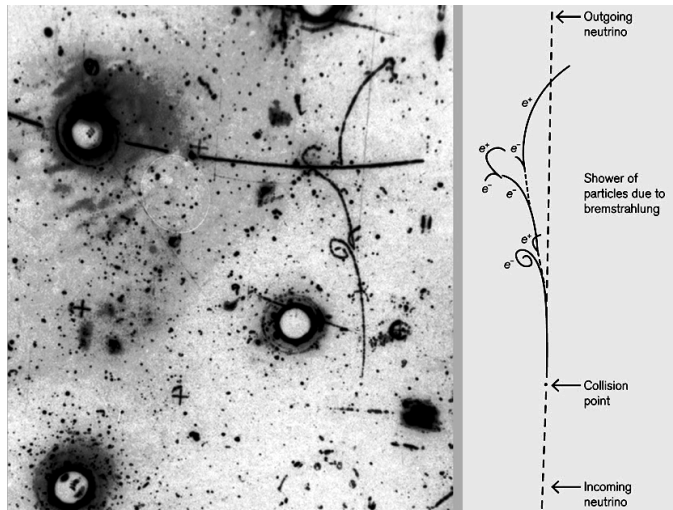


Figure 1.1: Neutral current observation in the Gargamelle bubble chamber [27].

Around the same time, semiconductor detectors were being developed. Compared to the widely used MWPCs, the semiconductor devices had a position resolution that was better by a magnitude [25, 26].

The ATLAS detector at the LHC displays many different types of subdetectors and will be described in Section 2.2. The principles how particles interact with matter are explained in Section 3.1. This work mainly focusses on silicon HV-CMOS pixel sensors, thus some relevant properties of silicon and radiation damage in silicon are addressed in Section 3.2 and Section 3.3, respectively. The ensuing chapter is dedicated to the development of a radiation damage implementation in a planar pixel sensor simulation and the comparison with data. Chapter 5 contains TCAD simulation works on two different HV-CMOS detector technologies plus measurements done on one of them. Conclusions will be drawn in the last chapter.

Chapter 2

The ATLAS Experiment at the LHC

The Conseil Européen pour la Recherche Nucléaire (European Organisation for Nuclear Research) (CERN), located in Geneva, Switzerland, was founded in 1954. CERN's first operational accelerator was the Synchrocyclotron, built in 1957, that provided 33 years of service. The Proton Synchrotron (PS) started up in 1959 and the Super Proton Synchrotron (SPS) in 1976. Both are still in use today. In 1989 the Large Electron-Positron Collider (LEP) started operation. After eleven successful years it was shut down in 2000 to make place in the same tunnel for the LHC, a proton-proton collider inaugurated in 2008. After initial low energy tests, when the data-taking began in 2010 the LHC replaced the Tevatron as the world's most powerful particle accelerator.

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is located at CERN, between Lake Geneva and the Jura mountains. It is a proton-proton collider with a 26.659 km [28] circumference situated on average 90 m below ground. Fig. 2.1 shows the whole accelerator complex and the experiments at the LHC: A Toroidal LHC Apparatus (ATLAS), Compact Muon Solenoid (CMS), A Large Ion Collider Experiment (ALICE) and Large Hadron Collider Beauty (LHCb).

Protons (and ions) have to travel several stages in the accelerator complex shown in Fig. 2.3. Hydrogen from an ordinary gas bottle is released into

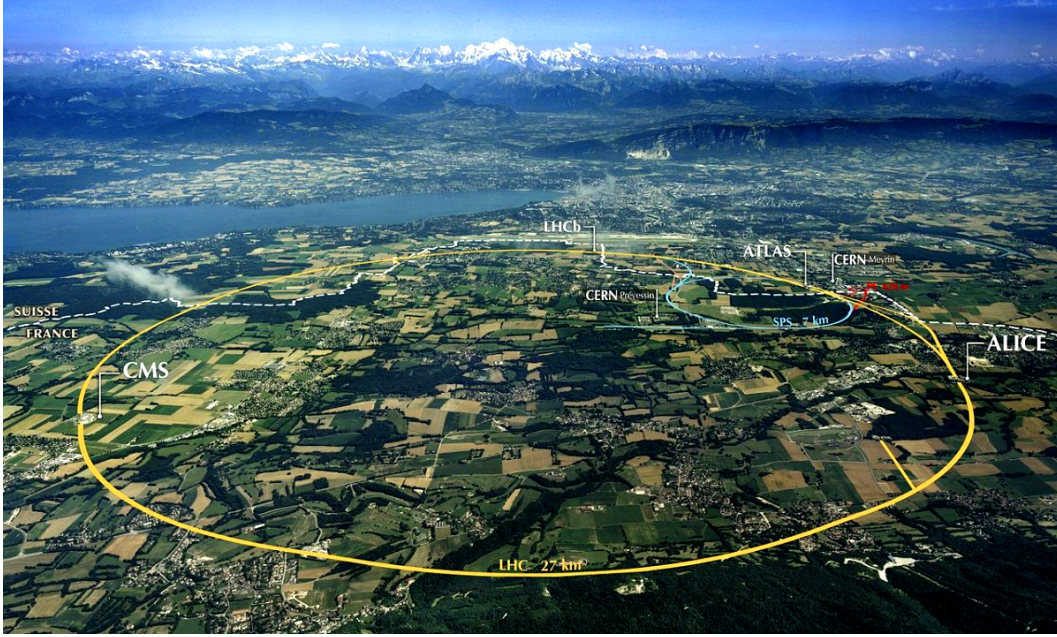


Figure 2.1: CERN aerial view ©Maximilien Brice (CERN)

a duoplasmatron where it is ionised and stripped off its electrons. The proton plasma is extracted and packed into bunches in a radio frequency (RF) quadrupole and then accelerated in RF cavities up to 50 MeV in the linear accelerator Linac 2. The protons continue to the Booster, which then delivers them with an energy of 1.6 GeV to the PS, which in turn feeds the 26 GeV protons into the SPS. Finally, the protons, having 450 GeV now, are guided as two beams into the LHC. It takes four minutes and twenty seconds to fill each of the LHC rings [29] and a total of about 45 minutes to bring the protons to their final energy. Most of the time is spent in transitions from SPS to the LHC and the ramping up of the magnets to high energy. The design energy per beam of 7 TeV is the upper limit given by the superconducting magnets. Since the LHC is a storage ring, the proton beams have to be constantly bent. The magnetic field required to hold the 7 TeV proton beam on its orbit of about 27 km is given by the Lorentz force which acts as the centripetal force:

$$B \approx \frac{E}{cRq} \quad (2.1)$$

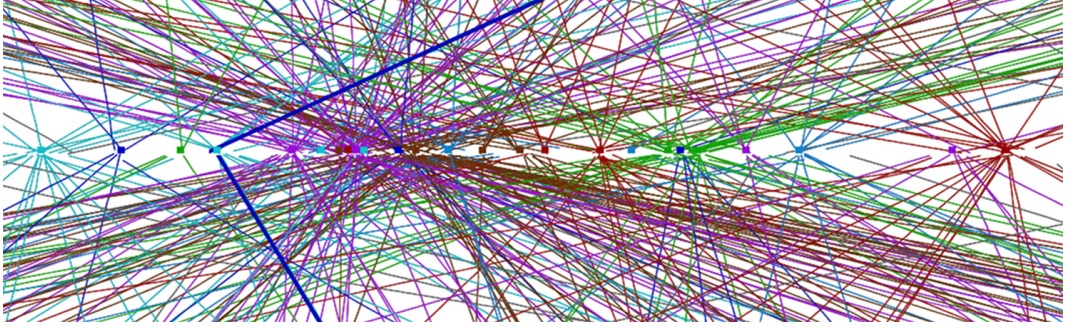


Figure 2.2: Reconstructed $Z \rightarrow \mu\mu$ event with 25 vertices within 5 cm [30].

with the beam energy E , the speed of light c , the radius of the LHC R and the proton charge q , and yields approximately 5.5 T. However, because the LHC is not fully circular, thus the bending radius in the curved sections is only 2804 m and the required field is higher. Apart from the 1232 dipole bending magnets, the rest of the ring is straight, consisting of 858 focusing quadrupoles and an even larger number of smaller correction magnets. The maximum magnetic field is 8.33 T, created by almost 12 kA current flowing through niobium-titanium superconductors, which are cooled down to 1.9 K by 5000 tons of superfluid helium.

The acceleration of the protons occurs at a single location in the LHC via eight RF cavities supplying a total of 16 MeV energy per proton bunch per turn. The RF cavities operate with 400 MHz, which is a multiple of the beam revolution frequency of $f_{\text{rev}} = 11.2$ kHz. Therefore, it is theoretically possible to have up to $\frac{400 \cdot 10^6}{11.2 \cdot 10^3}$ proton bunches in one ring. However, this would exceed what the accelerator and detector systems can handle, which is $n_b = 2808$ bunches per beam in a 25 ns bunch spacing. As a result, the two antiparallel beams cross each other's path in each of the four experiments with a bunch crossing (BC) rate of 40 MHz. Each proton packet has a transverse dimension of minimum $\sigma = 16.7 \mu\text{m}$, a longitudinal dimension of $\sigma_z = 7.55$ cm, and contains about $N = 1.15 \cdot 10^{11}$ protons. At each bunch crossing the probability of proton-proton collisions is given by the geometric cross-section σ_c . The number of collisions per BC is called pile-up, illustrated in Fig. 2.2. The total

number of collisions per second is given by $\mathcal{L} \cdot \sigma_c$ where

$$\mathcal{L} = \frac{N^2 n_b f_{\text{rev}}}{4\pi \epsilon_n \beta^*} \gamma R \quad (2.2)$$

is the instantaneous luminosity of the order of $10^{34} \text{ cm}^{-2} \text{ s}^{-1}$, taken into account the normalised beam emittance ϵ_n , the beta function related to the transverse size of the beam at collision point as β^* , the relativistic factor (proton beam energy in unit of proton rest mass) $\gamma = E_{\text{beam}}/m_p$, and a geometrical reduction factor $R \approx 1$ which corrects for the collision angle. The (time) integrated luminosity $\int \mathcal{L} dt$ is a measure for the amount of data taken, given in inverse barns $b^{-1} = 10^{24} \text{ cm}^{-2}$. The lifetime of one proton fill is about 10 hours of data-taking. Protons are “lost” mainly in collision processes during bunch crossings and with gas molecules. Collimation and Coulomb scattering also worsen the beam quality. Therefore, the beam has to be dumped after a certain time. While the LHC is working with one fill of protons, the pre-accelerators deliver protons to many other experiments.

The LHC is designed to have 14 TeV centre of mass energy $\sqrt{s}$, and an integrated luminosity of 300 fb^{-1} by the end of 2023. After the initial operation experience and the incident detailed in [31], it was operated at 7 TeV in 2010–2011, and at 8 TeV in 2012. After reparation of the interconnections between magnets during Long Shutdown 1 (LS1), the LHC has been operating at 13 TeV from 2015 until present. In LS3 the LHC is going to be upgraded to the High Luminosity (HL) LHC to accumulate larger amount of data more quickly. In order to accommodate the much higher luminosity and therefore pile-up, detectors have to be upgraded too.

2.2 The ATLAS Detector

A Toroidal LHC Apparatus (ATLAS) is one of two general-purpose detectors at the LHC. Fig. 2.5 is a computer generated cut-away view showing the ATLAS detector and its subdetectors. It is cylinder-shaped with a diameter of 25 m and a length of 44 m. The subdetector systems are arranged in an onion-like configuration. The origin of the coordinate system [34] is the nominal collision point. The xy-plane is perpendicular to the beam direction

CHAPTER 2. THE ATLAS EXPERIMENT AT THE LHC

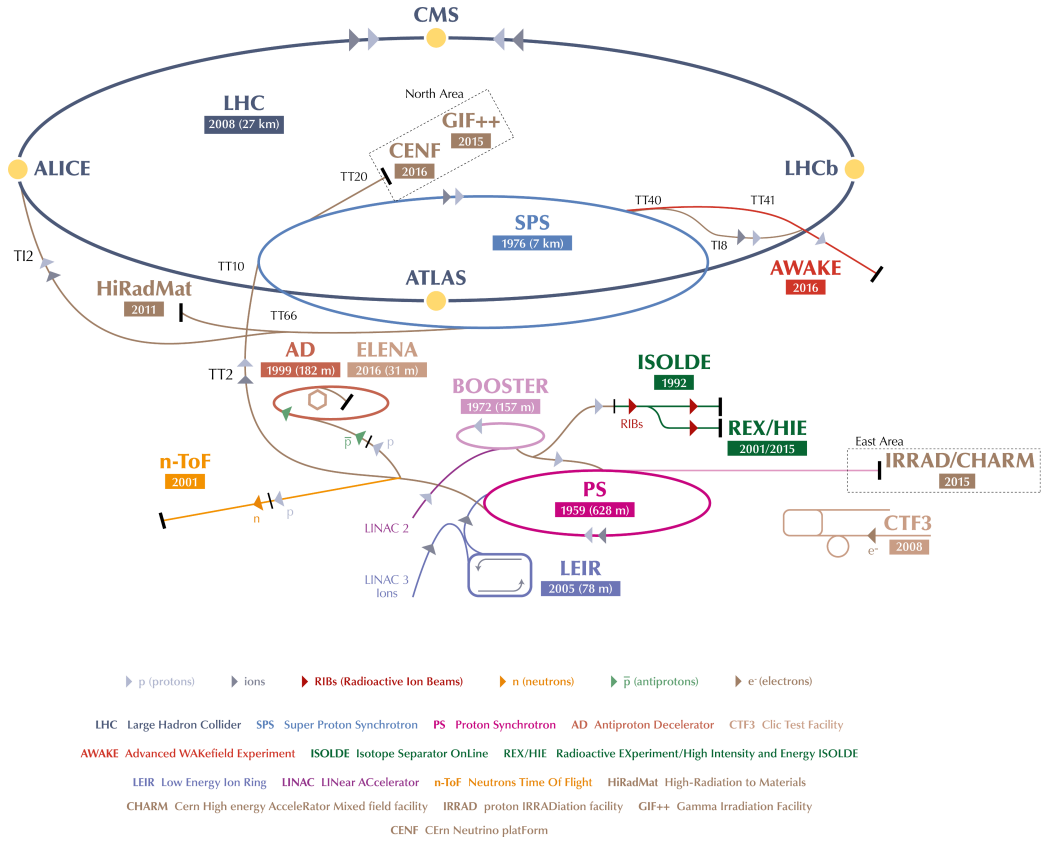


Figure 2.3: CERN accelerator complex [32].

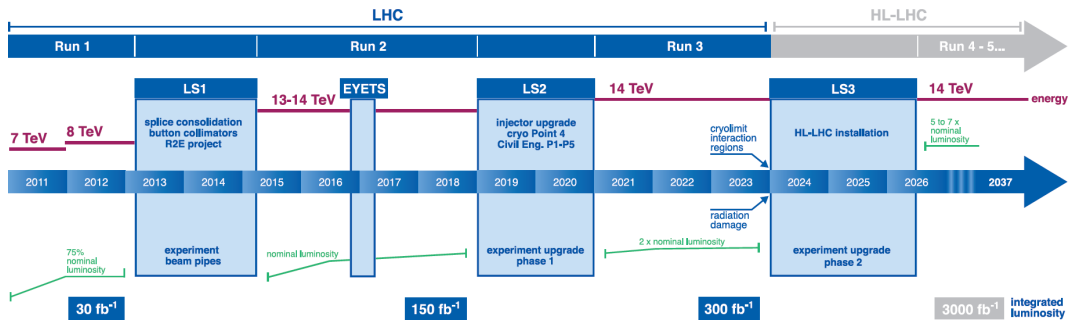


Figure 2.4: LHC/HL-LHC schedule [33].

which is the z-axis. The positive x-direction points towards the centre of the LHC ring, the positive y-direction points upwards. The positive z-axis thus results from a right-handed coordinate system. The azimuthal angle ϕ lies in the xy-plane while the polar angle θ is measured from the positive z-direction

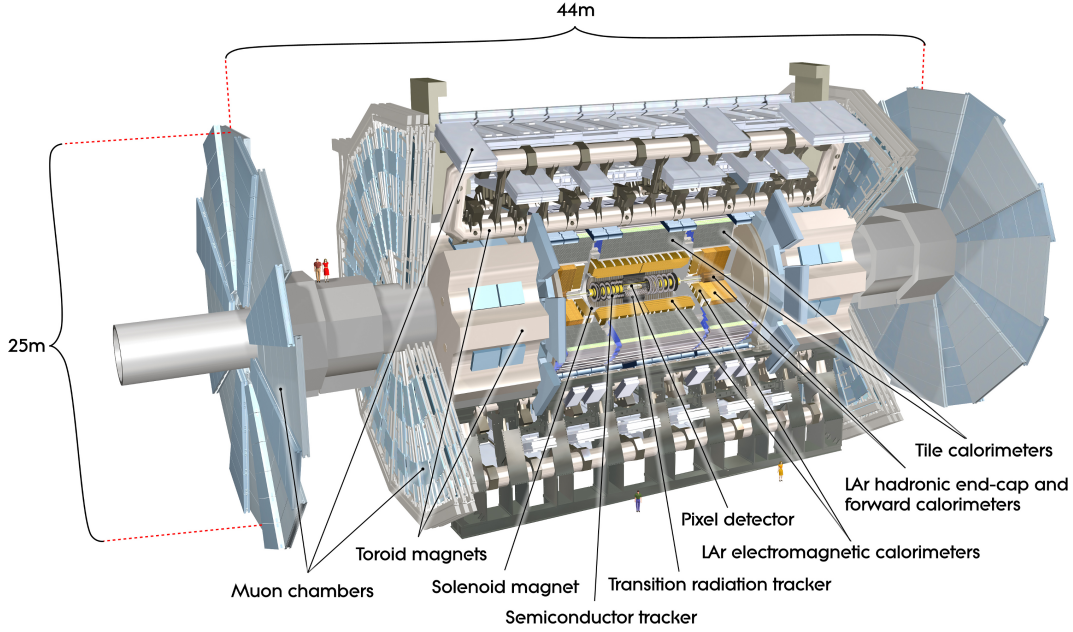


Figure 2.5: Cut-away view of the ATLAS detector [35].

and lies in the rz -plane. The pseudorapidity is defined as $\eta = -\ln \tan(\theta/2)$ and the transverse momentum p_T and energy E_T are defined in the xy -plane.

2.2.1 Inner Detector

The current Inner Detector (ID), shown in Fig. 2.6, sits around the beam axis in the very centre of the ATLAS detector and extends to $|\eta| = 2.5$ [34]. It is made of four cylindrical barrel layers at radius 3.3–12 cm and six end-cap disks of silicon pixel detectors, four barrel layers at radius 30–51 cm and 18 end-cap disks of silicon strip detectors, also known as the Semiconductor Tracker (SCT), and the Transition Radiation Tracker (TRT) made of straw tubes at radius 55–108 cm. The inner layers of the ID are just a few centimetres away from the primary interaction point where the particle track density is highest. This requires an excellent spatial and time resolution which is why silicon detectors are used here.

The ID is surrounded by a superconducting solenoid magnet with a field of 2 T to bend the tracks of charged particles in the xy -plane, perpendicular to the beam axis (z).

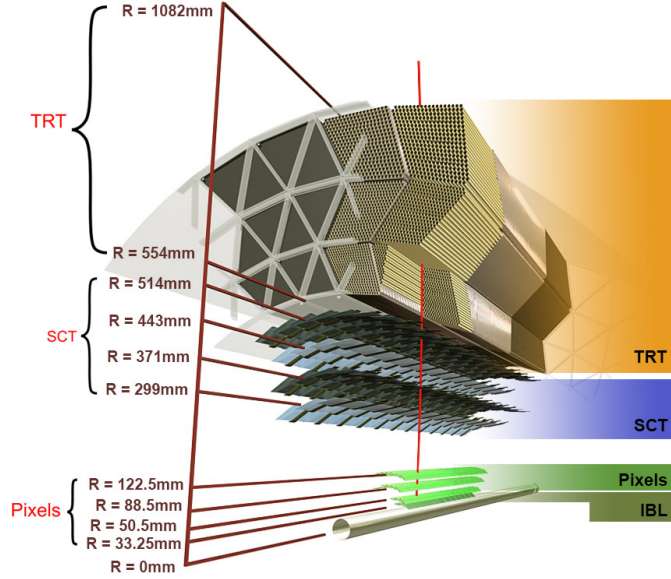


Figure 2.6: Computer generated graphic of the ATLAS inner detector [36].

2.2.1.1 Pixel Detector

The pixel detectors are mainly hybrid planar sensors — diodes made from high purity float zone (FZ) silicon DC coupled onto front-end (FE) readout chips using tiny metal bumps (bump bonding). The sensors of the three original pixel layers are $250\text{ }\mu\text{m}$ thick and 90 % of the pixels have an area of $400 \times 50\text{ }\mu\text{m}^2$, predetermined by the readout chip FE-I3, while the rest are $600 \times 50\text{ }\mu\text{m}^2$. The intrinsic spatial resolution is $10\text{ }\mu\text{m}$ in the $r\phi$ -direction and $115\text{ }\mu\text{m}$ in z -direction for the $400\text{ }\mu\text{m}$ pixels [34]. Enhanced by charge sharing between pixels in one hit cluster, the intrinsic spatial resolution can be better than $\text{pitch}/\sqrt{12}$. Sensors are made in oxygenated n -type wafers with n^+ doping (n^+ -in- n). The reverse bias voltage is 150 V before and up to 600 V after irradiation to maintain a stable charge collection. The Insertable B-Layer (IBL) is the innermost layer, added in 2014 during the LS1 between Run 1 and Run 2. The planar sensors of the IBL use an active-edge design to maximise the area usable for charge collection. Beside 60 % planar sensors there are also n^+ -in- p 3D sensors at high η region. In 3D sensors the n^+ implant are columns perpendicular to the sensor surface. Because the distance between the columns can be smaller than the sensor thickness, the charge drift distance can be signifi-

cantly reduced compared to planar sensors. The IBL sensors have a thickness of $200\text{ }\mu\text{m}$ (planar) and $230\text{ }\mu\text{m}$ (3D), and a pixel size of $250 \times 50\text{ }\mu\text{m}^2$ matching the readout chip FE-I4 [37, 38] with 80 columns and 336 rows.

2.2.1.2 Semiconductor Tracker

The SCT strip detectors use $(285 \pm 15)\text{ }\mu\text{m}$ thick p-in-n sensors. In the barrel, two rectangular sensors each with 6 cm edge length are stitched together to make strips that are $12\text{ cm} \times 80\text{ }\mu\text{m}$. The end-cap modules are trapezoidal with radial strips. In both cases an identical sensor is glued back-to-back to the first one and rotated by 40 mrad to allow 2-dimensional hit coordinates module. The bias voltage is 150 V before and 250–350 V after irradiation, whereas currently the SCT is operated at 150 V [39].

2.2.1.3 Transition Radiation Tracker

The Transition Radiation Tracker (TRT) consists of drift tubes (straw tubes). These are 4 mm in diameter and made of aluminium-coated Kapton tubes with gold-plated tungsten wires in the middle. The potential difference between the wire and the tube is 1.5 kV. The 1.5 m long barrel tubes are parallel to the beam axis and are read out at both ends, while the end-cap tubes are 0.4 m long, radial and perpendicular to the beam axis. The tubes are filled with a gas mixture consisting of 70 % xenon. The space between the straws is filled with polymer fibres in the barrel and foils in the end-caps so that high energetic particles can create transition radiation when they cross the boundary of different materials.

2.2.2 Calorimeters

The ATLAS calorimeters are placed outside the solenoid. Fig. 2.7 shows the two calorimeter systems: the sampling Liquid Argon (LAr) calorimeter which comprises the forward calorimeter (FCAL), the electromagnetic barrel and end-cap calorimeters and the hadron end-cap, and the Tile Hadronic Calorimeter (TileCal). Apart from muons and neutrinos, all particles should deposit all of their energy into calorimeters due to the energy loss mechanisms described in Section 3.1. Electromagnetic calorimeters exploit ionising and

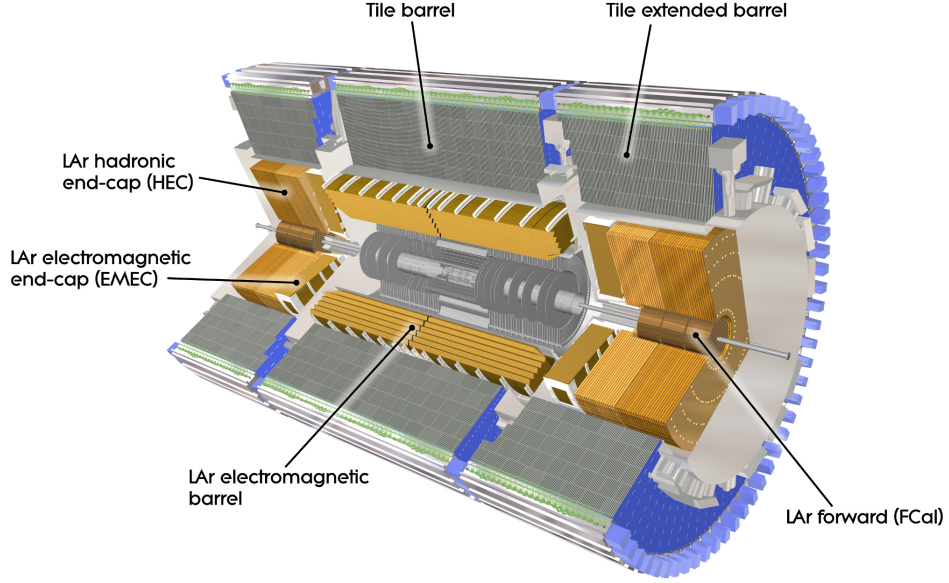


Figure 2.7: Cut-away view of the ATLAS calorimeters [32].

radiative energy deposition (cf. Section 3.1.1 and Section 3.1.2) where the stopping power of the material strongly depends on Z (ionising loss) or even Z^2 (radiative loss). Therefore, materials with high Z , e.g. lead, are used as absorbers like in the case of the LAr calorimeter. Hadron calorimeters rely on the interaction of the incident particle with the nuclei of the atoms. Therefore, materials with a high density, like copper, are used for the hadron end-cap of the LAr calorimeter, and iron is used for TileCal.

2.2.2.1 LAr Calorimeters

The barrel and the electromagnetic end-caps of the LAr calorimeter are made of lead as the absorber and liquid argon as the active material. They are arranged in an accordion geometry with alternating layers of lead absorber and readout boards made of copper-etched insulating layers. Between each two lead layers are two honeycomb spacers serving as ionisation chambers and supporting the readout boards in the middle. The barrel is 6.4 m long and covers $|\eta|$ up to 1.475, it has a thickness of 53 cm which corresponds to 22 radiation lengths X_0 [40].

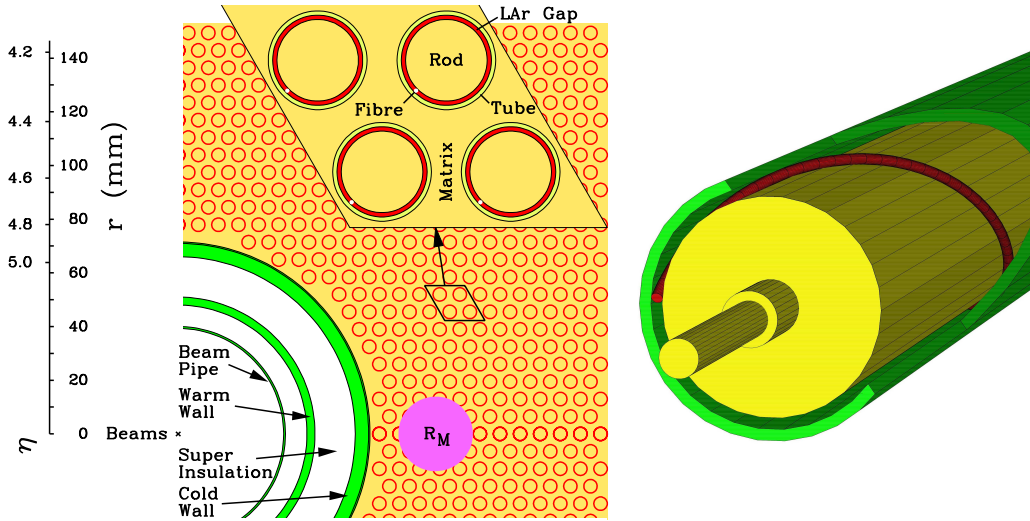


Figure 2.8: A section of the ATLAS FCAL (left) and a cut-away view of the assembly of an electrode with tube, fibre and rod (right) [40].

The hadronic end-cap has a similar layout as the electromagnetic end-cap but with a flat geometry and copper absorbers. It has a radius of approximately 2 m and a thickness of 10 nuclear interaction lengths.

The FCAL, which covers the $|\eta|$ regions from 3.2 to 4.9, has a very different geometry. It consists of copper and tungsten blocks with holes containing a tube held at ground. In its centre a rod is held at high voltage and acts as the collection electrode. Optical quartz fibres winding around the rod hold them in place. The space in between is filled with liquid argon. The tube and the rod are made of the same material as the absorber block.

The advantage of using LAr are the linear response to deposited energies and the radiation tolerance. As the argon has to be cooled to 90 K the whole LAr calorimeter are inside an aluminium cryostat.

2.2.2.2 Tile Calorimeter

Because hadronic end-caps already exist in the LAr calorimeter, the tile hadron calorimeter only consists of a central barrel and two barrel extensions. All barrels are made of ϕ -wedges with staggered layers made of steel-scintillator stacks. Each scintillator tile is read out from both sides through wavelength-

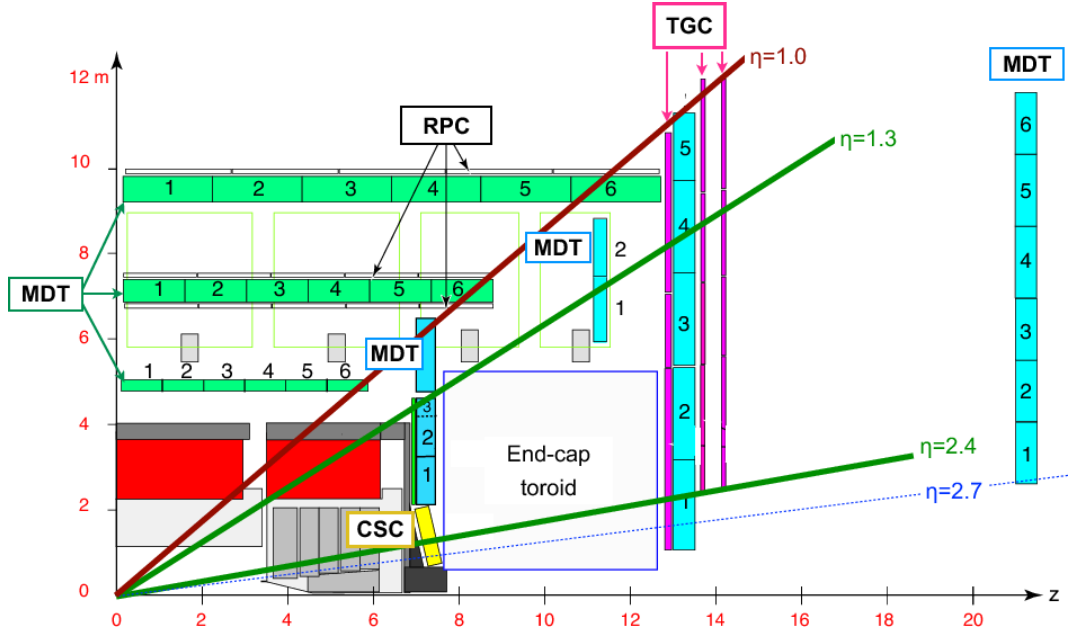


Figure 2.9: Cross-section schematic of the detectors in the muon spectrometer [41].

shifting fibres by photomultiplier tubes (PMTs). The TileCal covers an $|\eta|$ region of < 1.7 and has a thickness of 7.4 nuclear interaction lengths.

2.2.3 Muon Spectrometer

The muon spectrometer is the outermost layer in the ATLAS detector. All known particles are detected and stopped with the calorimeters, except for muons and neutrinos for their small interaction cross-section. The ATLAS muon spectrometer consists of many different detector systems as shown in Fig. 2.9: Monitored Drift Tubes (MDTs), Thin Gap Chambers (TGCs), Cathode Strip Chambers (CSCs) and Resistive Plate Chambers (RPCs).

The toroid magnet (not shown above) consists of a barrel with 20 m outer diameter made of eight coils of 25 m length, and end-caps with 11 m outer diameter and 5 m length. The superconducting toroid creates a magnetic field of 4 T which creates curved muon tracks for the momentum measurement. The MDTs are made of aluminium tubes with 3 cm diameter and 0.9–6.2 m length. For large $|\eta|$ however, CSCs with higher granularity are required to handle the high particle flux. These are MWPCs with anode wire spacing of

2.54 mm and segmented cathode strips with 5.08 mm spacing at a 90° stereo angle with respect to the wires. The wires and strips are both read out to give a two-dimensional position of the particle track. The MDT and the CSC make up the muon precision tracking system which measures the particle track in the bending direction with excellent spatial resolution of $50\text{ }\mu\text{m}$. But because of the long drift times of a few hundred nanoseconds [42], muon triggers are provided by the trigger chambers: RPCs for the barrel region and TGCs for the end-cap region. They also add a second coordinate measurement of the particle tracks in the magnetic field direction with an accuracy of up to ten centimetres. The RPC, operated in avalanche mode, is made of high-resistive bakelite plates with $\sim 10^{10}\text{ }\Omega\text{cm}$. The TGC is a MWPC where the distance between the wire and the readout plane is smaller than the distance between the wires.

2.3 ITk — HL-ATLAS Inner Detector Upgrade

The LHC aims to achieve over 4000 fb^{-1} integrated luminosity over its full lifetime [43]. For this purpose there will be a Phase-II upgrade during LS3. The total luminosity of the HL-LHC has to be accumulated over ten years with an instantaneous luminosity of $7.5 \cdot 10^{34}\text{ cm}^{-2}\text{s}^{-1}$, which corresponds to a pile-up of ~ 200 collisions per bunch crossing every 25 ns in the ATLAS detector. Earlier studies [44] considered an integrated luminosity of 3000 fb^{-1} . Fig. 2.10 shows the expected 1 MeV neutron equivalent fluence of the HL-LHC for the Inner Tracker (ITk) region of the ATLAS detector according to these studies. The fluence is normalised to 3000 fb^{-1} of 14 TeV events.

The current ID was designed for an operational lifetime of ten years at a maximum instantaneous luminosity of $10^{34}\text{ cm}^{-2}\text{s}^{-1}$ and 23 pile-up per 25 ns bunch crossing. The pixel detector can withstand 10^{15} 1 MeV neutron equivalent fluence (n_{eq}) per cm^2 (Section 3.1.3), which is equivalent to 400 fb^{-1} . As Fig. 2.10 shows the pixel system of the ITk will be subjected to fluences of up to $1.4 \cdot 10^{16}\text{ }n_{\text{eq}}/\text{cm}^2$ and a total ionising dose (TID) of up to 7.7 MGy in the inner most layer. The simulation agrees with data obtained with the current ID within 50 % [44]. The current ID would not survive this much harsher radiation environment. As for the occupancy, front-end electronics of the silicon

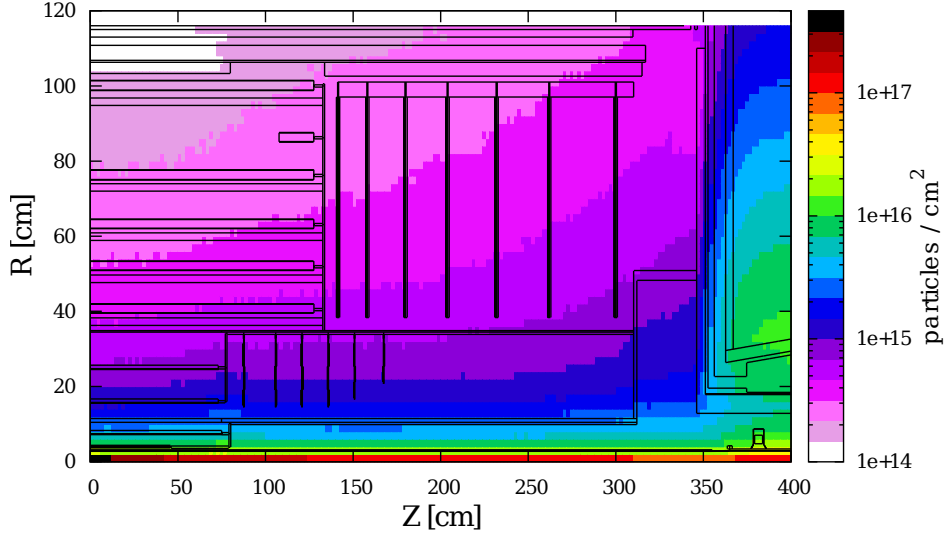


Figure 2.10: FLUKA simulation of NIEL fluence of the ITk region [44]

detectors cannot accommodate such a large number of pile-up events and at such a high rate the TRT would be 100 % occupied.

Apart from the necessary higher radiation tolerance also better performances in resolution and occupancy are required of the new ITk in order to cope with the higher particle rate and dense tracks. A minimum number of 11 hits per track within the ITk helps to reduce the percentage of fake reconstructed tracks at high rates. The best momentum resolution can be obtained by maximising the lever arm, placing the innermost layer at the shortest possible radius and the outermost layer at the largest possible radius with consideration of the mechanical constraints. In addition, a precise hit position is required which can be achieved by minimising the pixel and strip size. E.g. the innermost pixels should have a size of $50 \times 50 \mu\text{m}^2$ or $25 \times 150 \mu\text{m}^2$ and the strips should have a much shorter length of 23.8 mm for the innermost strip layer. With this fine granularity and thinner detectors, dense tracks can be more easily separated and detector occupancy is reduced. Also, a smaller material budget can be achieved using thinner sensors which would reduce multiple scattering and the production of secondary particles. This results in a better precision, fewer tracks and lower occupancy. While the current ID incorporates more than $1.2 X_0$ for $|\eta| > 1$, the ITk should contribute less than $0.7 X_0$ to the total material budget up to $|\eta| = 2.7$.

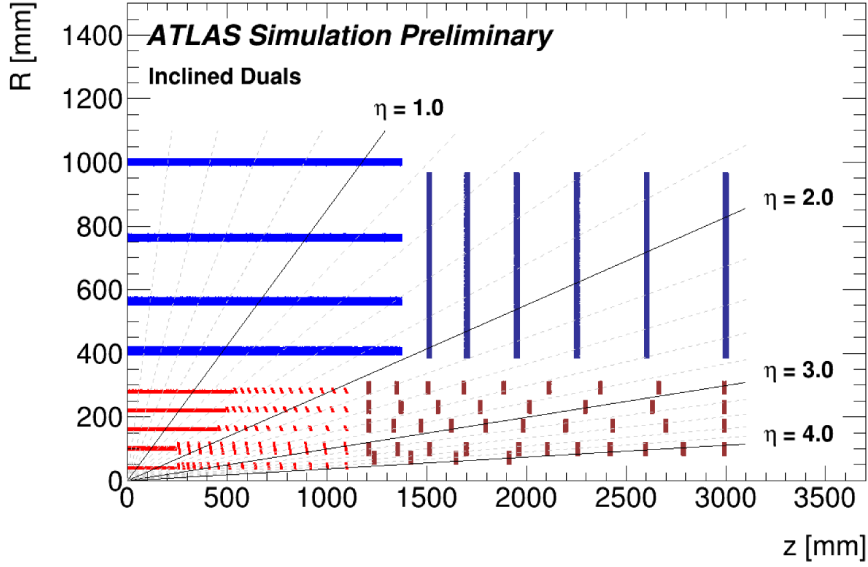


Figure 2.11: Inclined layout of the ITk [43]

To meet these demanding requirements, a full-silicon ITk is foreseen. Since the ATLAS Phase-II Letter of Intent [44] the layout for the ITk has evolved in order to meet the optimal configuration. Design iterations have converged on an inclined layout which is referred to as Inclined Duals (Fig. 2.11) in the technical design report (TDR) of the ITk pixel detector [43]. The design extends the coverage up to $|\eta| = 4$. The “inclined” refers to the tilted sensors in the high $|\eta|$ region of the pixel barrel. This way the required area of pixel detectors is reduced compared to an extended layout where the innermost barrel layers exceed the length of the outer layers. Consequently less service material is required. The material a forward particle has to traverse is minimised and thus the hit cluster length, while the number of hits in a track is maximised. This results in a better spatial resolution. In the current Inclined Duals layout the number of hits per track is > 13 for $|\eta| < 2.7$ and > 9 for higher $|\eta|$.

The sensors considered for the ITk in both the ATLAS Phase-II Letter of Intent [44] and in the ITk pixel TDR are 3D and high resistivity planar sensors, which have been extensively researched and tested. Both are proven to be very radiation tolerant whereas the planar sensor is more cost-efficient than the 3D sensor. However, in recent years a rapidly increasing R&D effort has been

put into high voltage complementary metal-oxide-semiconductor (HV-CMOS) sensors with the number of collaborating institutes growing. The technologies in considerations are ams AG (ams) 180 nm and 350 nm [45], LFoundry (LF) 150 nm [46] and TowerJazz 180 nm [47]. In order to synergise the effort a choice on one of the technologies has yet to be made later in 2018 or in 2019.

2.4 Physics at the ATLAS Experiment

Physics at ATLAS can be divided into two categories: precision measurements of the Standard Model (SM), and searches for particles within and beyond the SM. The Standard Model of particle physics, briefly summarised in Section 2.4.1, describes the physics, which we know quite well. But it still has many free parameters and countless open questions, of which some are probed at ATLAS and described in Section 2.4.2.

2.4.1 The Standard Model

There are four fundamental forces in physics: the gravitational, electromagnetic, weak and strong force. At the small scale of the particle world the gravitational interaction is negligible. The Standard Model describes the electroweak interaction, which combines the electromagnetic and the weak interaction, and the strong interaction, also known as the Quantum Chromodynamics (QCD).

Fig. 2.12 pictures the fundamental building blocks of matter in the Standard Model. For each force there are mediators and participants. The mediators are the vector bosons not shown in Fig. 2.12, which are the photon (γ) for the electromagnetic force, the $W^\pm$ and Z^0 bosons for the weak force, and the gluons (G) for the strong force. The participants are elementary particles: leptons and quarks. Each come in three families or generations of particles. The electrically charged leptons are the electron e , the muon μ and the tau τ , with their corresponding chargeless neutrinos $\nu_{e,\mu,\tau}$. The quarks charged with $\frac{2}{3}e$ are up u , charm c and top t , with $-\frac{1}{3}e$ are down d , strange s and bottom b . All leptons and quarks participate in the weak interaction. Charged leptons and quarks participate in the electromagnetic interaction and all quarks par-

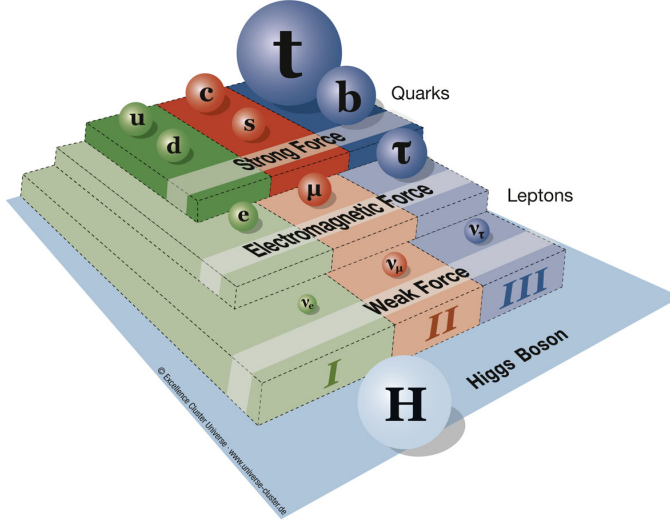


Figure 2.12: Quarks and leptons of the Standard Model grouped by the type of interactions and families [48].

ticipate in the strong interaction. The masses of the charged particles and the Higgs are indicated with differently sized spheres in Fig. 2.12. The charged elementary particles acquire their masses by interaction with the Higgs field with its scalar boson called the Higgs boson and have an antiparticle counterpart.

Precision measurements of the Standard Model are crucial. Not only are SM processes the main background to any new physics processes, but precision measurements are also a way to search for these. Thus an accurate description and modelling help in searches for new physics. Fig. 2.13 shows the measured SM production cross-section for different processes compared with theoretical calculations. Until recently the picture of the Standard Model was still missing the Higgs boson. The discovery of the Higgs boson was announced on 4th July 2012 as a “Higgs-like” particle by the ATLAS and the CMS experiment. The results are published in [49] and [50]. Today, more data are collected and evaluated to confirm the properties of e.g. its mass, spin 0 and parity 1 and coupling to other SM particles to ensure it is indeed the SM Higgs boson.

Even though the SM has deemed itself very successful by being consistent with experiment and fulfilling theoretical predictions, there are still missing pieces. The many free parameters in the SM like the fermion masses, the coupling constants, the Higgs mass and its self-coupling, and the coefficients

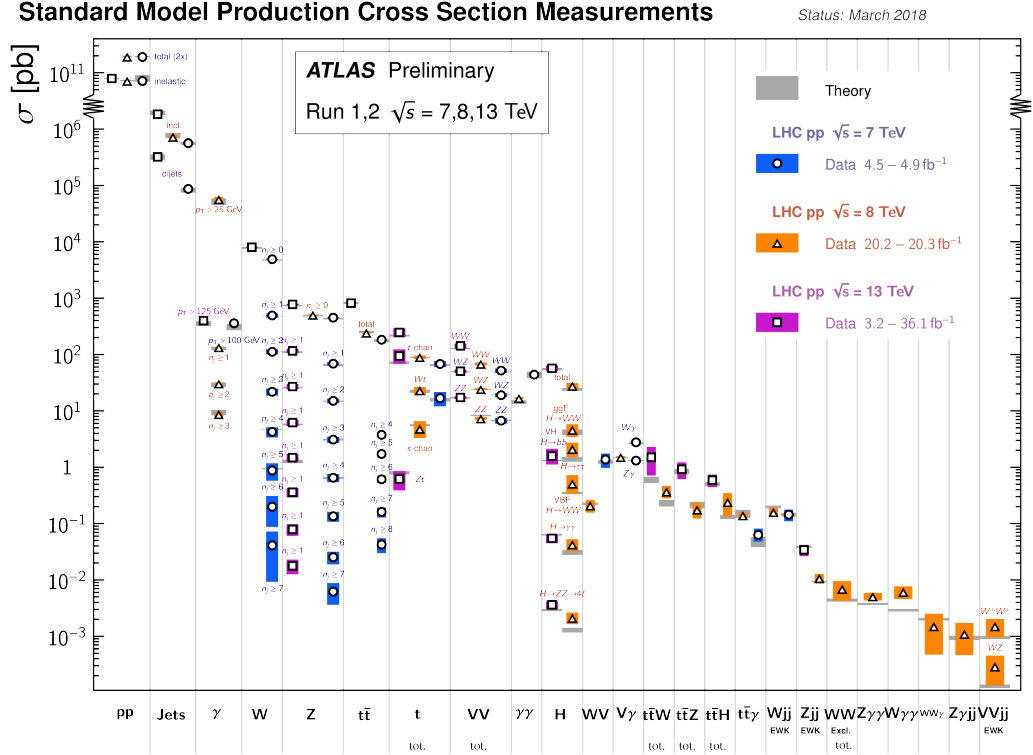


Figure 2.13: ATLAS SM production cross-sections measured and predicted [51].

of the Cabibbo-Kobayashi-Maskawa (CKM) matrix, are not predicted by the Standard Model but have to be determined experimentally. Therefore, various processes and their production and decay cross-sections are being measured. Originally the Standard Model (SM) assumed a zero mass for neutrinos. Observations show that neutrinos oscillate in flavours, which would not be possible if they had a zero mass. Therefore, these have since been included.

The Standard Model also leaves many open questions. It does not address gravity at all. What is the origin of the CP violation? Why are there three families of leptons and quarks, and why do they have increasing mass? What about the other 85% of matter in the universe we have not observed yet? It is evident that there is more physics beyond the Standard Model.

2.4.2 Beyond the Standard Model

As the most powerful particle accelerator in the world, the LHC is a good place to look for physics beyond the Standard Model. At ATLAS the search can be roughly divided into two sections: search for supersymmetry (SUSY) and search for exotics. However, the boundary is diffuse and many aspects overlap.

The search for SUSY is highly motivated by the elegant theories addressing most of the SM problems like unification, Dark Matter (DM) and hierarchy. In the SM only the electromagnetic and weak interactions can be combined into the electroweak theory. In the Minimal Supersymmetric SM (MSSM) the running couplings of the electromagnetic, weak and strong forces unify at very high energy scale of about $5 \cdot 10^{16}$ GeV. SUSY also includes the lightest supersymmetric particle as a stable particle and therefore a suitable candidate for DM. In the SM the Higgs mass has divergent corrections but it is protected in SUSY.

All other Beyond SM (BSM) physics are generally referred to as exotics. The search for heavy gauge bosons is motivated by many SM extension models which predict such particles. An excess in events of di-photon invariant mass could be a hint of extra dimensions, therefore di-photon final states are being studied. The di-jet final state is well-described by perturbative QCD, therefore a thorough study of such data might be a source of discovering new physics. Signatures of black holes could manifest themselves in data of multi-jet final state events and would play a key role in quantum gravity. Another important final state includes missing transverse energy (MET), which could be the signature of a DM or SUSY particle.

Even though nothing beyond the SM is discovered yet, measurements at the LHC so far allowed to exclude the existence of new physics in the searched energy ranges [52].

2.4.3 Physics at the High Luminosity LHC

With the data the LHC delivered so far, the Higgs boson was discovered and the SM predictions were confirmed. Even though the evidence for any BSM particles and interactions is still missing, it does not prove that they do

not exist because evidently, the SM cannot explain e.g. Dark Matter. The data so far put constraints on a wide range of models and parameters.

The increase of the centre-of-mass energy from 7 or 8 TeV to 13 TeV after LS1 and eventually to the nominal 14 TeV opens a window to search for particles with higher mass. The High Luminosity dataset will surpass the 300 fb^{-1} of the LHC dataset within a few months. This benefits searches for physics that are harder to distinguish from the SM or for very rare phenomena by building up high statistics. For example, the average production cross-section times branching ratio of the SM Higgs boson is measured to agree within 15–20 % with expectations. This could be reduced to a few percent with the HL-LHC [53]. Observations of rare decay processes like $H \rightarrow Z\gamma$ or $H \rightarrow \mu^+\mu^-$ are facilitated by building up statistics. The Higgs pair production, which is sensitive to Higgs self-coupling, is not within reach of the LHC due to small cross-section and large background. This requires large luminosity and will be within reach of the HL-LHC. In BSM physics, processes with low cross-sections, like the production of neutralino and chargino, the supersymmetric partners of the Z^0 and $W^\pm$ bosons could be discovered with a 5σ significance [54]. Apart from proving SUSY, this would also provide a candidate for DM in the lightest SUSY particle which is stable.

In addition, some theoretical predictions of the SM lack precision because there are no higher-order calculations or there are uncertainties in the knowledge of parton distribution functions. The precision measurements that can be done with higher luminosity help in refining these predictions.

Chapter 3

Particle Interactions in Silicon

Tracking detectors are an essential part of any particle detector. There are numerous requirements they have to fulfill, such as high rate capability, excellent time and spatial resolution, and short interaction and radiation lengths. In this chapter, the fundamental concepts of energy loss of particles in matter are explained in Section 3.1; how particle detectors work is described by means of silicon detectors in Section 3.2, including some aspects of which the understanding is relevant for later chapters; finally, the damage caused by particles in silicon is discussed in Section 3.3.

3.1 Energy Loss of Particles in Matter

A particle entering a medium interacts with it by exchanging energy. The different particles can be roughly divided into three categories: photons, charged particles and neutral particles.

Photons interact with matter mainly through three processes: photoelectric effect, Compton scattering and pair production. The cross-section for each process depends on the photon energy as can be seen in Fig. 3.1. Photons can produce secondary electrons through these interactions which again can produce photons. This is called the electromagnetic cascade or shower and plays an important role in electromagnetic calorimetry.

Charged particles can be detected via Coulomb interaction with the detector material through which the incident particle transfers some energy to the electrons of the detector material via excitation or ionisation. This will be

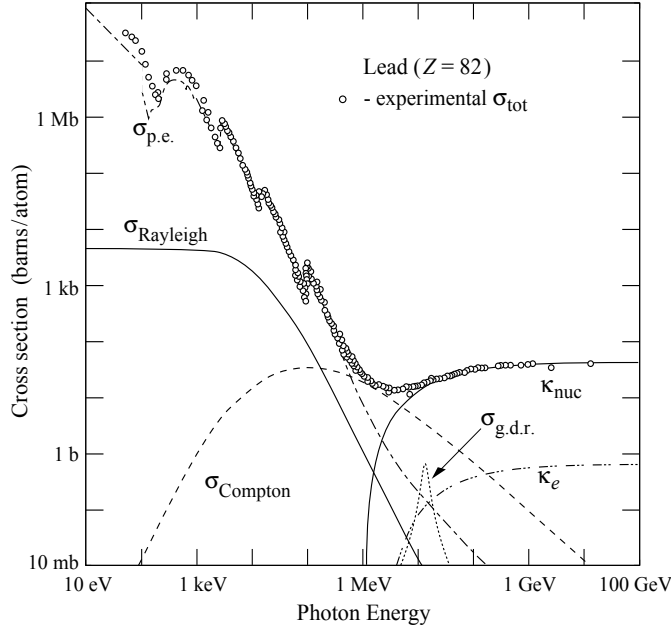


Figure 3.1: Photon cross-section in lead vs. photon energy [55]. κ_{nuc} and κ_e are pair production in nuclear field and electron field, respectively. $\sigma_{\text{g.d.r.}}$ is the cross-section of giant dipole resonance.

explained in Section 3.1.1. At higher energies radiative effects dominate for light particles such as electrons (ref. Section 3.1.2).

Neutral particles like neutrons or neutrinos carry no electric charge and therefore cannot participate in Coulomb interaction. Neutrinos only weakly interact with material and are thus extremely difficult to detect. They either exchange Z bosons with electrons which subsequently emit Cherenkov radiation, or transform into their charged lepton partners via charged current. To observe neutrinos very large dedicated neutrino detection facilities are required. In ATLAS, neutrinos contribute to the missing energy of an event along with any other undetectable particles. Neutrons however, and other strongly interacting charged particles too, participate in nuclear interactions with the material they pass through and lose energy due to non-ionising energy loss (NIEL). This is described in Section 3.1.3 with emphasis on the mechanism of the induced damage.

3.1.1 Ionising Energy Loss

A charged particle travelling through matter loses energy mainly via ionising collisions with atomic electrons. For incident particles with energy in the region $0.1 \lesssim \beta\gamma \lesssim 1000$ and for target materials with intermediate Z , the Bethe-Bloch formula describes the energy loss per unit length:

$$\left\langle -\frac{dE}{dx} \right\rangle_B = \frac{Z\rho N_A}{A} \cdot \frac{4\pi}{m_e} \cdot \left(\frac{z\alpha\hbar}{\beta} \right)^2 \left[\frac{1}{2} \ln \frac{2m_e c^2 \beta^2 \gamma^2 W_{\max}}{I^2} - \beta^2 - \frac{\delta(\beta\gamma)}{2} \right], \quad (3.1)$$

where

Z	is the atomic number of the medium,
ρ	the density of the medium in g/cm ³ ,
N_A	Avogadro's constant in mol ⁻¹ ,
A	the relative atomic mass of the material in g/mol,
m_e	the mass of the electron in MeV/c ² ,
z	the charge multiplicity of the incident particle,
α	$= e^2/4\pi c\epsilon_0\hbar$, the electromagnetic fine-structure constant,
$\hbar$	the reduced Planck constant in MeVs,
c	the speed of light,
β	$= v/c$, the relativistic speed of the projectile,
γ	$= 1/\sqrt{1-\beta^2}$, its Lorentz factor,
$W_{\max}$	$= \frac{2m_e c^2 \beta^2 \gamma^2}{1+2\gamma m_e/M+(m_e/M)^2}$, maximum energy transfer in eV,
M	the mass of the incident particle,
I	the mean ionisation energy of the medium in eV, and
$\frac{\delta(\beta\gamma)}{2}$	$\rightarrow \ln(\hbar\omega_p) + \ln \beta\gamma - 1/2$ the density correction with plasma frequency ω_p .

The unit of the linear stopping power is MeV/cm. The unit of the mass stopping power $\langle -dE/\rho dx \rangle$ is MeV cm²/g. The energy deposition is proportional to $z^2(Z/A)$ and depends only weakly on the material since $Z/A \approx 1/2$ for most light to medium-weight materials. At maximum energy transfer, a dependence on the mass of the projectile is given by $W_{\max}$. However, at low energies below about 100 GeV the approximation $W_{\max} \approx 2m_e c^2 \beta^2 \gamma^2$ is valid for $2\gamma m_e/M \ll 1$ and the mean energy loss becomes independent of the mass of the incident particle.

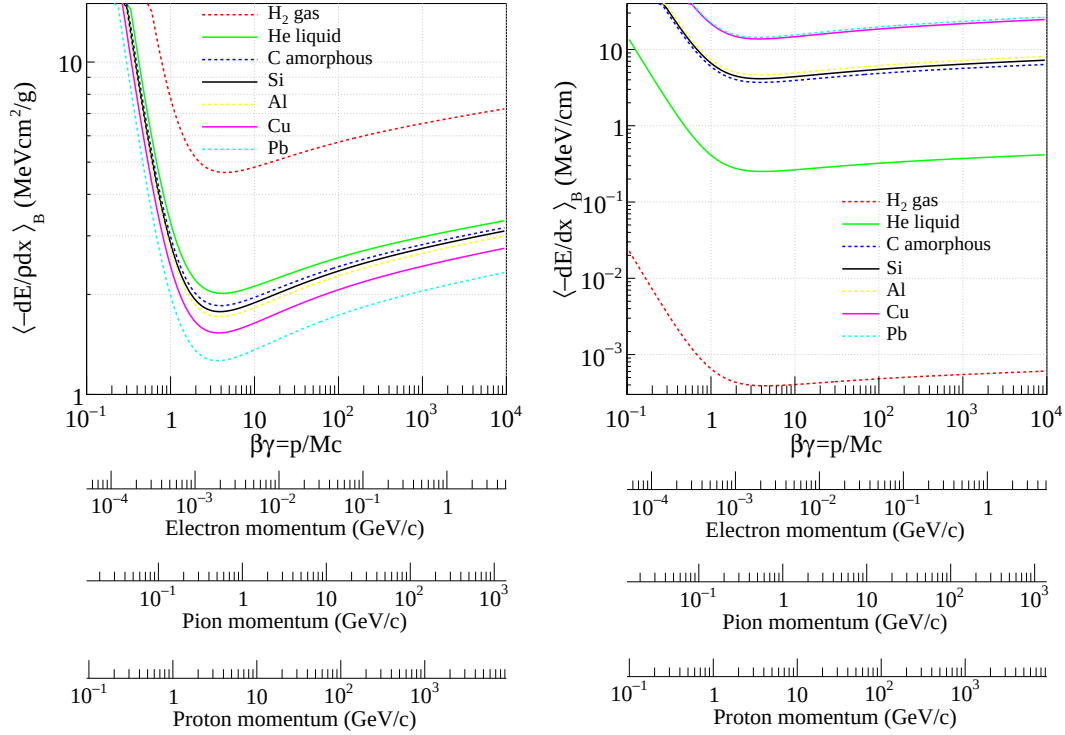


Figure 3.2: Mass (left) and linear (right) stopping power for charged particles as a function of $\beta\gamma$ according to the Bethe-Bloch formula. Shown are some common materials used in particle detectors plus the very light hydrogen and very heavy lead. Energy scales for particles, commonly used in test beams, are added. Data used for the graphs are taken from [56].

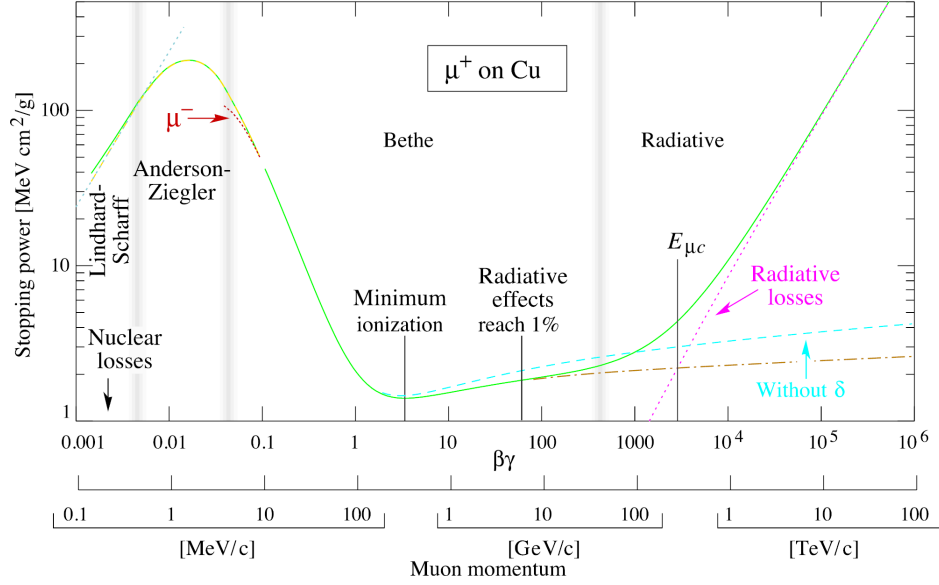


Figure 3.3: Stopping power of μ^+ in copper in full energy range [57].

The stopping power in different materials is shown in Fig. 3.2 as a function of $\beta\gamma$. It is roughly proportional to $1/\beta^2$ for low energies and proportional to $\ln(\beta\gamma)$ for high energies. The minimum stopping power is at $\beta\gamma \approx 3$ and particles with that energy are thus called minimum ionising particles (MIPs). The last term of the Bethe formula, the density correction, is only important at higher energies when the transverse electric field of the particle expands proportionally to γ . This term takes into account that the medium is polarised by the passing particle and thus its electric field is screened from electrons with large impact parameters. At energies above the critical energy radiative losses dominate.

In the low-energy region, i.e. below $\beta\gamma \approx 0.3$, the energy loss mechanism is more complicated. Shell corrections need to be taken into account for projectile velocities v smaller than the orbital velocity of bound e^- : $v_0 = Z\alpha c$. The assumption of stationary shell electrons is no longer valid which adds the term $-C/Z$ into brackets of the Bethe formula. For slow heavy ions, e^- -capture is present which results in a lower effective charge of the incident particle. The scattering cross-section has to be reconsidered with higher orders because only the first order (Born approximation) is used for the Bethe formula. Barkas effect becomes distinctive which yields a smaller stopping power for negatively

charged particles than for positively charged particles. This is due to polarisation of the target material which depends on the charge polarity of the incident particle. For the region $0.01 \lesssim \beta \lesssim 0.05$ there is no fully reliable theory but a phenomenological fit by Andersen and Ziegler exists.

The mechanism of ionisation occurs on a microscopic level. A charged particle electromagnetically scatters off both electrons and nuclei of the lattice atoms. However, the energy lost in interactions with the nuclei can be neglected whilst the energy lost in collisions with electrons dominates due to the significant difference in their masses. The probability of a collision is given by the collision cross-section and the energy transfer per collision is stochastic. This results in a statistical macroscopic picture in which energy loss is naturally stochastic and fluctuates. Large energy transfers up to $W_{\max}$ to single electrons produce δ -electrons perpendicular to the path of the projectile. They make up the long tail in the energy loss distribution. Therefore, the mean energy loss described by the Bethe formula differs from the most probable energy loss which corresponds to the peak in the straggling function [58].

3.1.2 Radiative Energy Loss

Charged particles moving in a medium also experience bremsstrahlung. However, as $\langle -\frac{dE}{dx} \rangle_{\text{rad}} \propto \frac{Z^2 e^4 E_0}{Am^2}$, this depends strongly on the mass, and radiative loss is thus only noticeable for light particles. For fast electrons, the term for the collisional loss is logarithmic and proportional to $Z \ln E_0$, similar to the Bethe formula. Their total energy loss is the sum of collisional and radiative loss:

$$\left\langle -\frac{dE}{dx} \right\rangle_{\text{tot}} = \left\langle -\frac{dE}{dx} \right\rangle_{\text{coll}} + \left\langle -\frac{dE}{dx} \right\rangle_{\text{rad}} . \quad (3.2)$$

The characteristic length of a material by which an electron has lost all but $1/e$ of its initial energy is called the radiation length X_0 . Fig. 3.4 shows the fraction of ionising and radiative losses of electrons and positrons in lead. The energy at which the two losses are equal is called critical energy E_c . For electrons with more than 100 MeV the energy loss by bremsstrahlung dominates and can be approximated by $\langle -\frac{dE}{dx} \rangle_{\text{rad}} \approx \frac{E}{X_0}$.

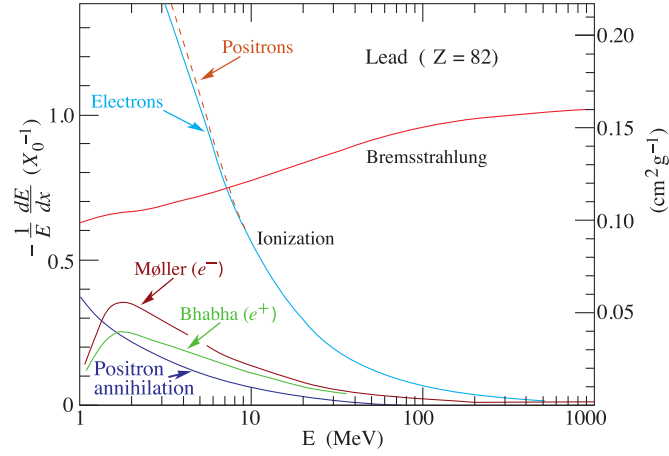


Figure 3.4: Contributions of electron energy loss in lead with $X_0(\text{Pb}) = 6.37 \text{ g/cm}^2$ [55].

3.1.3 Non-ionising Energy Loss

The non-ionising energy loss (NIEL) (in MeV/cm) can be described by

$$\text{NIEL}(E) = \frac{N_A}{A} \int_{\theta_{min}}^{\pi} \left(\frac{d\sigma(\theta, E)}{d\Omega} T(\theta, E) L(T(\theta, E)) \right) d\Omega \quad (3.3)$$

where

- E is the initial energy of the incident particle,
- N_A Avogadro's constant,
- A the relative atomic mass of the medium,
- θ collision angle of the incident particle with the atom,
- $\frac{d\sigma}{d\Omega}$ the differential cross-section of atomic displacements,
- T average recoil energy of target atoms, and
- L the Lindhard partition factor.

The Lindhard partition factor [59] describes the fraction of energy loss due to NIEL. A very detailed calculation can be found in [60]. NIEL occurs due to Coulomb interaction of the incident particle with lattice atoms, mainly resulting in displacement of a primary knock-on atom. All massive particles and photons (gammas) are involved in NIEL, whereas the contribution of photons is indirect through creation of secondary electrons in the detector which then contribute to NIEL [61]. The NIEL of protons and neutrons (and other par-

icles) is due to elastic and inelastic interactions [62, 63]. They are similar for both protons and neutrons. However, at low energies of < 20 MeV the Coulomb interaction has a dominant effect on the protons (ref. Fig. 3.5) and the cross-section is well-described by the Rutherford cross-section. There are also elastic nuclear reactions that play almost no role in the total displacement cross-section but increase the recoil energy [62]. The inelastic reactions cover excitations and fragmentation of the nucleus where it is broken by the incident particle. Inelastic reactions leave a range of fragments that contribute to the recoil and thus displacement. As NIEL depends on the type of the incident particle, it is helpful for analyses to have a unit that facilitates comparisons. A much used source of neutrons for NIEL studies are neutrons from nuclear fission reactors. These neutrons approximate well the displacement damage of 1 MeV neutrons. Therefore, the total displacement damage energy is normalised to the equivalent of fluences of 1 MeV neutrons: $1 \text{ MeV } n_{\text{eq}}/\text{cm}^2$. Fig. 3.5 shows the calculated displacement damage cross-section D , which is equivalent to 1 MeV neutrons, for neutrons as well as other particles. The same curves can also be experimentally determined by measuring the change in resistivity through radiation induced defects (ref. Section 3.3.2). The much larger cross-section of protons below ~ 20 MeV comes from the Coulomb contribution. The neutron cross-section that exists well below the threshold energy of 175 keV (Section 3.3.2) is due to neutron capture with subsequent gamma emission which leads to a recoil energy of ~ 1 keV [64].

As shown in [64], other particle fluences and energies can be converted to $1 \text{ MeV } n_{\text{eq}}$ fluence using

$$\Phi_{\text{eq}} = \kappa \Phi = \kappa \int \phi(E) \text{d}E \quad (3.4)$$

by integrating the differential spectral fluence $\phi(E)$ of that particle and scale it by the hardness factor

$$\kappa = \frac{\int D(E) \phi(E) \text{d}E}{D(E_n = 1 \text{ MeV}) \cdot \int \phi(E) \text{d}E} \quad (3.5)$$

where $D(E)$ is the displacement damage cross-section.

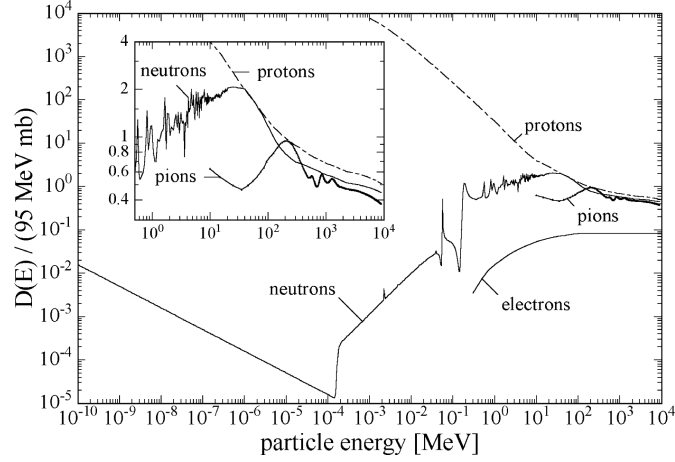


Figure 3.5: Displacement damage functions, normalised for 1 MeV neutrons, for different particles. The data for the graphs are obtained through calculations (see [64] and references therein).

3.2 Silicon Particle Detector

Different kinds of detectors are described in Section 2.2. This section focusses on solid state detectors and why silicon is widely used.

3.2.1 Physical Properties of Silicon

The difference between amorphous and crystalline solids is the presence of a symmetric, periodically repeating elementary atomic grid for the latter, whereas the former lacks this kind of ordered structure. This grid is called the Bravais lattice shown in Fig. 3.6 (a), with primitive unit cells and lattice vectors $\mathbf{a}_i$. One kind of primitive unit cells with the full symmetry of the Bravais lattice is called Wigner-Seitz cells. Fig. 3.6 (b) shows how to construct such a cell in two dimensions. The distances from one atom to its eight neighbours are cut in half. Then the smallest area around the atom enclosed by the cuts is the Wigner-Seitz cell.

For a crystalline solid the Miller indices (hkl) indicate the orientation of the Bravais lattice. In such a lattice the atoms are bound in different ways, from Van der Waals bond to ionic bond, covalent bond and metallic bond, with increasing strength. Furthermore, solids can be grouped according to their electrical properties into conductors, semiconductors and insulators. The

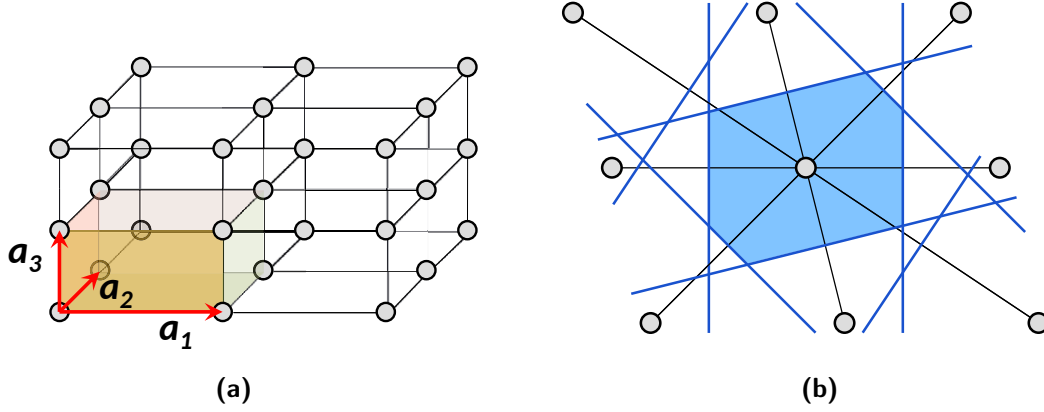


Figure 3.6: Three dimensional Bravais lattice and a possible choice of primitive lattice vectors (a). Two dimensional Wigner-Seitz cell (b). [65]

deciding factor for this property is the structure of their electron energy bands. To understand the energy bands, it is worthwhile to make a small excursion into crystallography. The Bravais lattice can be described as

$$\mathbf{R} = n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2 + n_3 \mathbf{a}_3 \quad (3.6)$$

with three-dimensional Bravais lattice $\mathbf{R}$, primitive lattice vectors $\mathbf{a}_i$ as shown in Fig. 3.6, and integers n_i , so that every point in the lattice can be described by Eq. (3.6). In the free electron model a plane wave with wave vector $\mathbf{k}$ can be described as

$$\Psi_{\mathbf{k}}(\mathbf{r}) = \Psi_0 e^{i\mathbf{k}\mathbf{r}} \quad (3.7)$$

For all vectors $\mathbf{k}$ with the periodicity of the Bravais lattice it must prevail

$$\Psi_{\mathbf{k}}(\mathbf{r}) = \Psi_{\mathbf{k}}(\mathbf{r} + \mathbf{R}) \quad (3.8)$$

which leads to

$$e^{i\mathbf{k}\mathbf{R}} = 1 \quad (3.9)$$

This set of wave vectors $\mathbf{k}$ that satisfies the plane wave with the periodicity of a Bravais lattice is called the reciprocal lattice

$$\mathbf{G} = h\mathbf{b}_1 + k\mathbf{b}_2 + l\mathbf{b}_3 \quad (3.10)$$

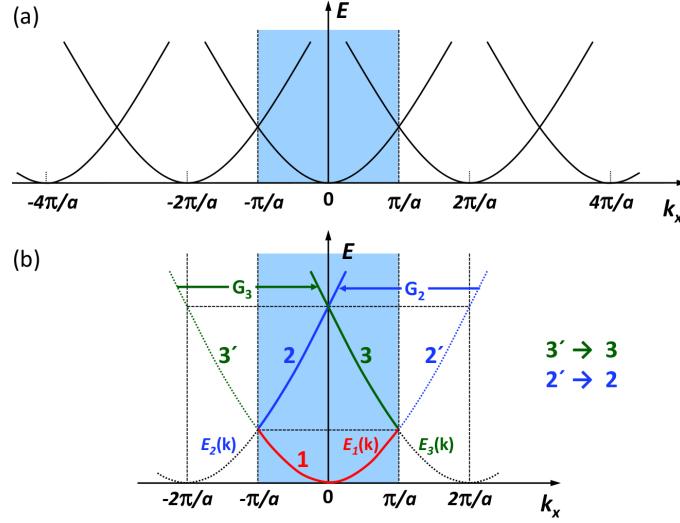


Figure 3.7: Normal (top) and reduced (bottom) zone schemes [65].

with integers h, k and l . The primitive cell of the reciprocal lattice is equivalent to the Wigner-Seitz cell and is called the first Brillouin zone. For the first Brillouin zone in one dimension it is $G = 2\pi/a$ with lattice constant a , and $k \in [-G/2, G/2]$.

Using the free electron gas model the Schrödinger equation of the plane wave can be solved

$$\hat{H}\Psi_k(\mathbf{r}) = E_k\Psi_k(\mathbf{r}) = \frac{\hbar^2}{2m}k^2\Psi_k(\mathbf{r}) \quad (3.11)$$

so that the dispersion relation E_k versus k lies on a parabola for every Brillouin zone as shown in Fig. 3.7 (a). Because of the symmetry the reduced zone scheme as shown in Fig. 3.7 (b) can be used.

While the free electron gas model describes the transport properties quite well, it does not explain the different electrical properties. Therefore, the nearly-free electron (NFE) model should be considered, which is based on the free electrons model but regards the periodical potential of the positive lattice atom cores as small perturbations. The Laue condition for the border of the first Brillouin zone in one dimension is $2\mathbf{k} \cdot \mathbf{G} = G^2$ where $k = \pm G/2 = \pm\pi/a$.

The plane wave at the border of the first Brillouin zone is

$$\Psi_1(x) = e^{iGx/2} = e^{i\pi x/a} \quad (3.12)$$

$$\Psi_2(x) = e^{-iGx/2} = e^{-i\pi x/a} \quad , \quad (3.13)$$

which is one plane wave moving in x direction and one Bragg reflected wave moving in $-x$ direction. Both waves interfere constructively and linear superposition yields standing waves

$$\Psi_+(x) \propto (e^{iGx/2} + e^{-iGx/2}) \propto \cos(\pi x/a) \quad (3.14)$$

$$\Psi_-(x) \propto (e^{iGx/2} - e^{-iGx/2}) \propto \sin(\pi x/a) \quad (3.15)$$

with the probability density $|\Psi_i|^2$ shown in Fig. 3.8. The potential of the ion cores in the lattice is negative and thus attractive to electrons. While $|\Psi_+|^2$ has maximum probability density at the atomic cores and $|\Psi_-|^2$ has minimum, it means that compared to the potential of a free electron which has a homogeneous potential everywhere, the potential of Ψ_+ is lower and Ψ_- is higher. The band gap is the difference between these potentials. In addition, for a standing wave the propagation velocity $\mathbf{v}_k = \frac{1}{\hbar} \frac{\partial E(\mathbf{k})}{\partial \mathbf{k}}$ must be 0. The function $E(\mathbf{k})$ must have a zero slope at the border of the first Brillouin zone (Fig. 3.9). The resulting gap between the minimum of the conduction band and the maximum of the valence band is called the band gap.

For the electronic properties the energy bands are filled up with electrons according to the Pauli principle. At 0K the lowermost empty band in an insulator or semiconductor is called the conduction band and the uppermost fully filled band is called the valence band, as can be seen in Fig. 3.10. They are separated by the band gap where the Fermi energy level is situated in the middle. If an electric field is applied, there is no free state where the electrons can move to. For metals, the uppermost fully filled band is also called the valence band, but the conduction band is only half filled up to the Fermi level as shown in Fig. 3.10. By applying an electric field the electrons in the conduction band can change states by gaining a momentum. Semi-metals have an overlap of the conduction band with the valence band so that the Fermi level is situated in both bands. However, neither of them is fully filled or

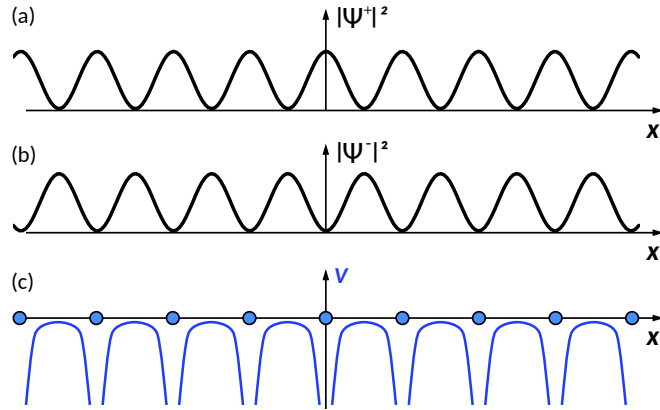


Figure 3.8: (a) and (b): Probability densities of the nearly-free electron model as a function of the position in the lattice. (c) potential of the lattice atoms [65].

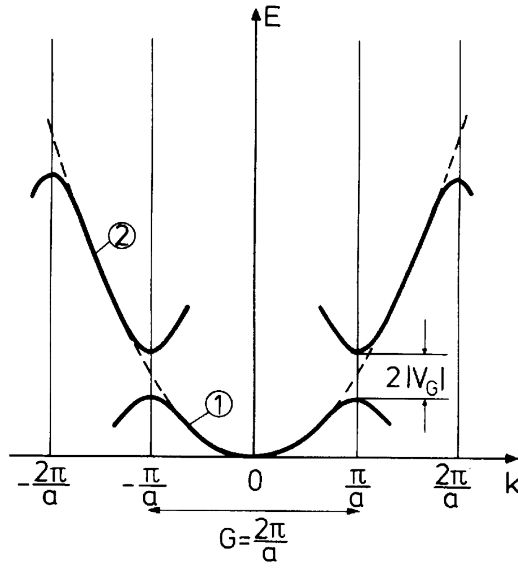


Figure 3.9: Energy band structure [66]. The dashed lines indicate the dispersion relation of the free electrons, the solid lines that of band electrons. $2|V_g|$ is the size of the band gap.

completely empty. The equivalent of the Fermi energy for temperatures larger than 0 K is called the chemical potential μ .

The difference between a semiconductor and an insulator lies only in the size of the band gap. At room temperature the semiconductor is conductive due to

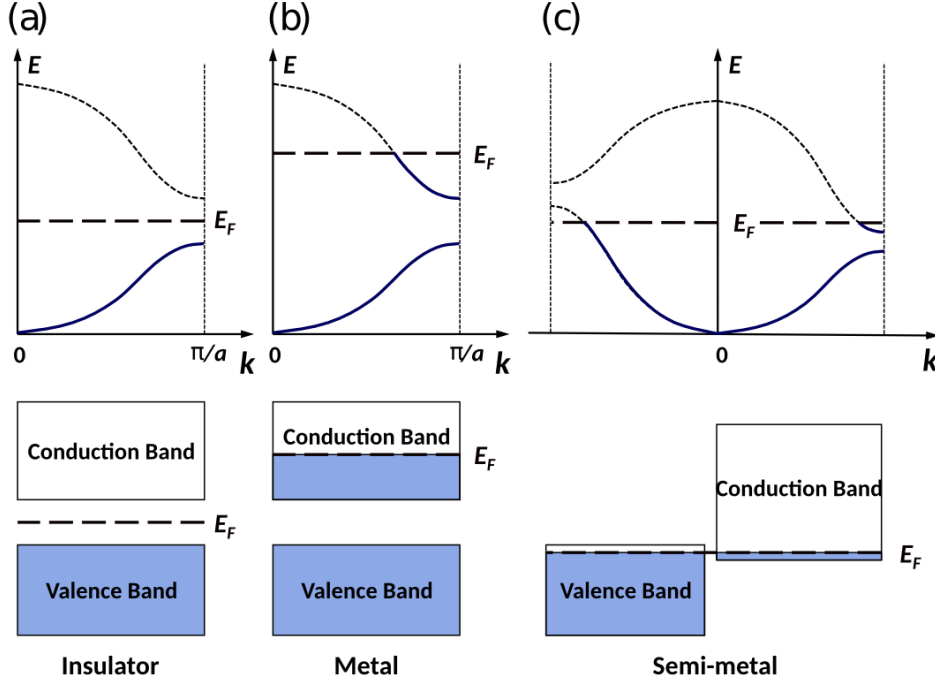


Figure 3.10: Energy band structures of insulator (a), metal (b) and semi-metal (c) [65].

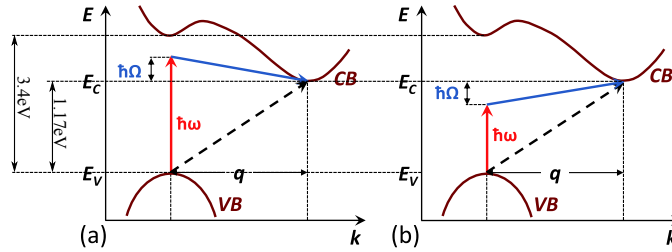


Figure 3.11: Indirect optical transition from a lower band to an upper band via phonon interaction to bridge a momentum gap of q through phonon emission (a) and phonon absorption (b) [65]. Ω is the phonon frequency and ω is the electron frequency.

thermal excitation of some electrons from the valence band to the conduction band. If the maximum of the valence band is situated right below the minimum of the conduction band at the same k , it is called a direct semiconductor. Otherwise, it is called an indirect semiconductor, in contrast to Fig. 3.10. The band gap is still the difference between the minimum of the conduction band and the maximum of the valence band. For a direct semiconductor

the excitation of a valence band electron into the conduction band requires only the energy of the band gap. For an indirect semiconductor the energy required to promote an electron from the maximum of the valence band to the conduction band at the same $\mathbf{k}$ is the difference between these energy bands at this $\mathbf{k}$. Interband transitions through photoexcitation can only promote an electron to a higher energy band at the same $\mathbf{k}$ due to the small photon momenta, while thermal excitation through phonons, which are quasiparticles of the crystal oscillation with larger momenta compared to photons, adds a momentum transfer as shown in Fig. 3.11.

The size of the band gap becomes smaller with increasing temperature due to two factors: the thermal expansion of the crystal increases the lattice constant a , and the temperature dependence of the phonon distribution and thus the band structure. Both effects result in a decrease of the band gap [65].

Within the crystal the transport properties of a charge carrier is described with the effective mass tensor $m_{i,j}^*$ with

$$\left(\frac{1}{m^*}\right)_{i,j} = \frac{1}{\hbar^2} \frac{\partial^2 E_n(\mathbf{k})}{\partial k_i \partial k_j} \quad (3.16)$$

where n is the band index, and mobility $\mu = \frac{e\tau}{m^*}$, where τ is the mean free time. The mobility also depends on external fields and doping, thus has to be described empirically like in Section 4.1.5.

Silicon is a crystalline solid with diamond lattice structure, as illustrated in Fig. 3.12, and a covalent bond between every two atoms. It is an indirect semiconductor with a minimum band gap of 1.17 eV at 0 K (ref. Table 3.1). To promote one electron from the valence band to the conduction band without phonon participation a minimum energy of 3.4 eV energy is required as shown in Fig. 3.11 [65], while through phonon excitation much less energy is required since phonon energies are typically 0.01–0.03 eV [67].

3.2.2 Impurities and Doping

In reality a crystalline solid has neither a perfect lattice nor is it completely pure. The structural defects include point defects, e.g. vacancies due to a missing atom in the lattice, or interstitials when an extra atom is present in the

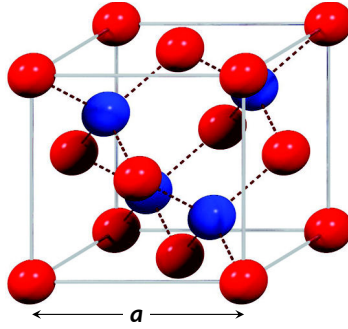


Figure 3.12: Diamond lattice with lattice constant a . The red atoms make up a face-centred cubic structure while the blue atoms are in the centre of the tetraedra consisting of their closest neighbours [65].

lattice, or Frenkel defects, associated with a pair of vacancies and interstitials; line defects, e.g. dislocations; planar defects, e.g. grain boundaries, stacking faults; and bulk defects which are more macroscopic like pores or cracks. The most common and relevant for detector material are point defects. Impurities occur through presence of other elements which bond into the lattice instead of the elementary atom. The lowest impurity concentration that can be reached nowadays is in the order of 10^{12} cm^{-3} . Both point defects and impurities create additional energy levels within the band gap. For silicon as a detector material, electrically active impurities are intentionally induced while defects come to exist through usage in highly radiating environments. They are often detrimental and therefore to be avoided.

Electrically active impurities can be divided into two categories. If one silicon atom, which is in the IVth main group in the periodic system, is replaced with an element from the Vth main group, e.g. phosphorus, even though the impurity is neutral, the bonding in the lattice results in one unbound electron e^- which can propagate freely in the lattice. Hence such impurities are called donors and the silicon is n-doped. On the other hand, if the replacement atom is from the IIIrd main group, e.g. boron, there would be a lack of an electron, and therefore a hole (e^+ or h^+) can propagate in the lattice. In this case the impurities are called acceptors and the silicon is p-doped. The free moving electrons and holes are called mobile charge carriers. Fig. 3.13 shows both kinds of doping and the energy levels, donor levels E_D and acceptor levels E_A they create. By doping the silicon, the chemical potential is moved just below

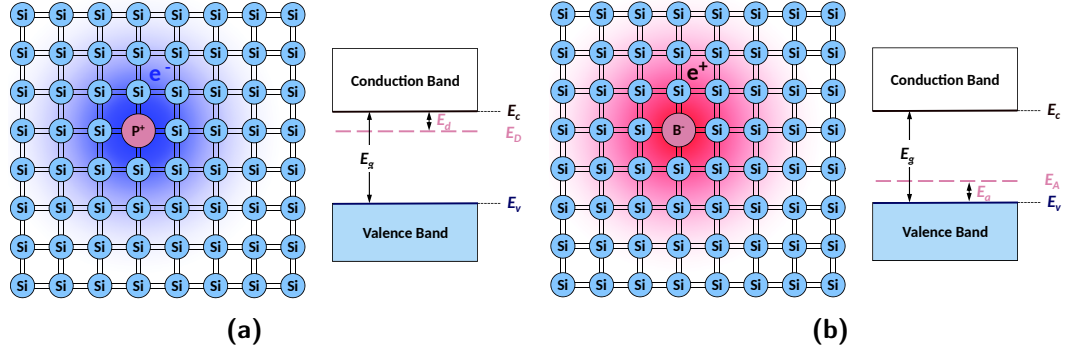


Figure 3.13: Donor (phosphorous) (a) and acceptor (boron) (b) as dopants in a silicon lattice with donor level E_D and acceptor level E_A [65].

the donor level for the n-doped silicon and above the acceptor level for the p-doped silicon.

The intrinsic charge carrier density of a semiconductor is

$$n_i = \sqrt{n_c p_v} = 2 \left(\frac{k_B T}{2\pi \hbar^2} \right)^{3/2} (m_e^* m_h^*)^{3/4} e^{-E_g/2k_B T} \quad (3.17)$$

where n_c and p_v are the electron density in conduction band and hole density in valence band, respectively, $m_{e,h}^*$ is the effective mass of an electron or hole, E_g is the band gap energy, T is the absolute temperature. At room temperature, the intrinsic charge carrier density for silicon is $1.45 \cdot 10^{10} \text{ cm}^{-3}$ (ref. Table 3.1). The average ionisation energy required to create an electron-hole (e-h) pair in silicon is 3.63 eV [68]. A MIP deposits on average about 400 eV/ μm in silicon as can be calculated using Eq. (3.1) and seen in Fig. 3.2 right. Assume a $1 \times 1 \text{ cm}^2$ silicon sensor of 300 μm thickness. Intrinsic (non-doped) silicon corresponding to this volume would have $4.35 \cdot 10^8$ free charge carriers. A MIP would typically create $3.3 \cdot 10^4$ e-h pairs in a piece of silicon of that thickness. Thus there are 10^4 times more free charge carriers than induced by the signal. Some possibilities exist to remedy this problem: operate at very low temperature as it is done for germanium detectors; use a material with a larger band gap such as diamond which results in far fewer e-h pairs created, but is a very costly material; or deplete the free charge carriers via a p-n junction and a reverse bias.

Crystal	Gap	$E_g(0\text{ K})$ (eV)	$E_g(300\text{ K})$ (eV)	Intr. charge carrier density (cm^{-3})
Germanium	indirect	0.74	0.66	2.4×10^{13}
Silicon	indirect	1.17	1.12	1.45×10^{10}
Diamond	indirect	5.48	5.47	-

Table 3.1: Band gap of germanium, silicon and diamond [68].

3.2.3 P-N Junction

On their own, p- and n-doped silicon have their individual energy levels with the chemical potentials at different places in the band gap as described in Section 3.2. By joining them together into a single structure, the chemical potentials have to be at the same level on both p- and n-doped sides so that the band structure is altered at the p-n interface as shown in Fig. 3.14. Two forces are active to create two opposite currents. First, due to the different concentrations of electrons and holes in each region, mobile charge carriers diffuse: electrons diffuse from n-region into p-region where they recombine with holes. This diffusion current is therefore also called recombination current. Electrons in n-type silicon and holes in p-type silicon are the majority charge carriers, as opposed to minority charge carriers which are electrons in p-silicon and holes in n-silicon. The diffusion creates a potential difference called diffusion voltage or built-in voltage V_{bi} . It causes a drift current of minority charge carriers in the opposite direction. Since the minority carriers in the p-region have to be generated through thermal excitation, the drift current is also called generation current. In equilibrium, the region around the p-n junction is depleted of mobile charge carriers which is called the depletion region. Since it is therefore no longer neutral, it is also called the space charge region.

Through the diffusion and drift currents the Einstein relation which describes the diffusion constant of electrons and holes can be derived as

$$D_{e,h} = \frac{k_B T}{e} \mu_{e,h} \quad , \quad (3.18)$$

with electron and hole mobilities $\mu_{e,h}$ respectively.

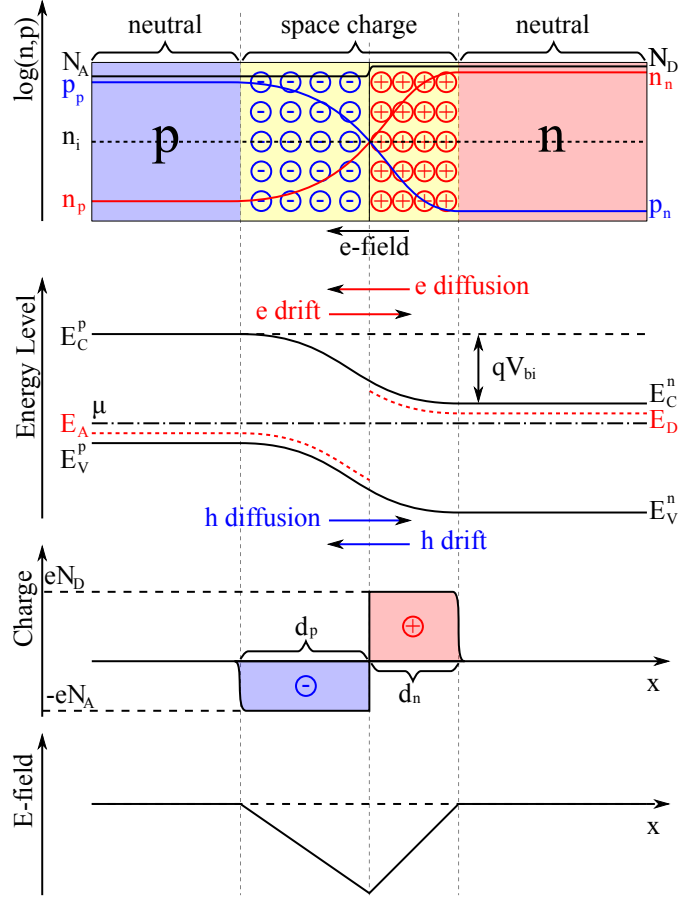


Figure 3.14: Charge carrier density, energy levels, charge density and electric field at a p-n junction.

The width of the depletion region is described by the following relations:

$$d_n = \sqrt{\frac{2\epsilon(V_{bi} - V_b)}{e} \frac{N_A/N_D}{N_A + N_D}} \quad \text{and} \quad (3.19)$$

$$d_p = \sqrt{\frac{2\epsilon(V_{bi} - V_b)}{e} \frac{N_D/N_A}{N_A + N_D}} \quad , \quad (3.20)$$

where $d_{n,p}$ is the depletion width in the n- and p-doped silicon, respectively, ϵ is the dielectric constant of silicon, $N_{A,D}$ is the acceptor and donor concentration,

$$V_{bi} = \frac{k_B T}{e} \ln\left(\frac{N_A N_D}{n_i^2}\right) \quad , \quad (3.21)$$

is the built-in voltage, and V_b the bias voltage which has a positive value for forward bias and a negative value for reverse bias. Eq. (3.19) and Eq. (3.20) show that if an external voltage is applied to the diode, the band structure is changed and therefore the width of the depletion region. In addition, the smaller the majority carrier concentration, the larger the depletion depth at the same bias voltage. For this reason, high resistivity silicon requires smaller bias voltages to reach the same depletion width.

3.2.4 Current-Voltage Characteristics

The current-voltage (I-V) characteristics of an ideal p-n junction are described by the Shockley equation [69]

$$I = I_n + I_p = I_0(e^{qV_b/k_B T} - 1) \quad (3.22)$$

with saturation current I_0 . For a reverse-biased diode the exponential term in the bracket disappears. The electron diffusion current on the p-side and hole diffusion current on the n-side are

$$\begin{aligned} I_n &= \frac{qD_n n_{p0}}{L_n} \left(\exp\left(\frac{qV_b}{k_B T}\right) - 1 \right) \\ I_p &= \frac{qD_p p_{n0}}{L_p} \left(\exp\left(\frac{qV_b}{k_B T}\right) - 1 \right) \quad , \end{aligned} \quad (3.23)$$

respectively, with the electron and hole diffusion coefficient $D_{n,p}$ from Eq. (3.18), diffusion length $L_{n,p} = \sqrt{D_{n,p}\tau_{n,p}}$, charge carrier lifetime with the temperature dependence of $\tau_{n,p} \propto T^{-1/2}$ [70]. For an n^+ -in-p sensor in equilibrium, the concentration of electrons on the p-side, n_{p0} , is much larger than the concentration of holes on the n-side, p_{n0} . Therefore the saturation current can be approximated as

$$I_0 \approx I_n = \frac{qD_n n_{p0}}{L_n} \approx q \sqrt{\frac{D_n}{\tau_n}} \frac{n_i^2}{N_A} \propto T^{\frac{1}{2}} \cdot T^3 \exp\left(-\frac{E_g}{k_B T}\right) \quad (3.24)$$

with $n_{p0} = n_i^2/p_{p0}$, $p_{p0} \approx N_A$, which only depends on the acceptor concentration, and using the temperature dependencies above and in Eq. (3.17).

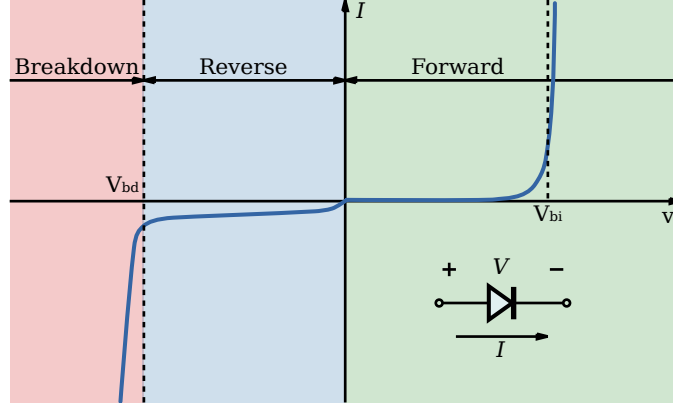


Figure 3.15: Diode I-V characteristics. The plot is strongly asymmetric. While the built-in voltage V_{bi} for silicon is around 0.7 V, the breakdown usually occurs at some 10–100 V, depending on the size and doping [71].

Without any external voltages the net current in equilibrium is zero. If a voltage is applied across the diode generating an electric field, the drift current is changed. As a result also the diffusion current. Therefore, the net current is not zero but is as shown in Fig. 3.15. For the forward bias, the external potential has to overcome the built-in voltage, which is ~ 0.7 V for silicon. Then the energy levels of the p-n junction are even, the depletion region is collapsed and the current that flows through the junction scales exponentially with the bias voltage. If the diode is reverse-biased the depletion zone expands with decreasing voltage through the full width of the silicon bulk. Through thermal excitation e-h pairs are generated in the depletion zone and result in a small current which is called leakage current or reverse current. Due to its thermal origin, the leakage current can be reduced by lowering the temperature. With decreasing bias the current saturates quickly until the external field is stronger than 10^5 V/cm. Then an electron would gain sufficient energy to create secondary e-h pairs and cause an avalanche breakdown. The leakage current of a detector contributes as integrated charge to equivalent noise charge (ENC) $\propto \sqrt{It_{\text{int}}}$ with t_{int} the integration time of the signal.

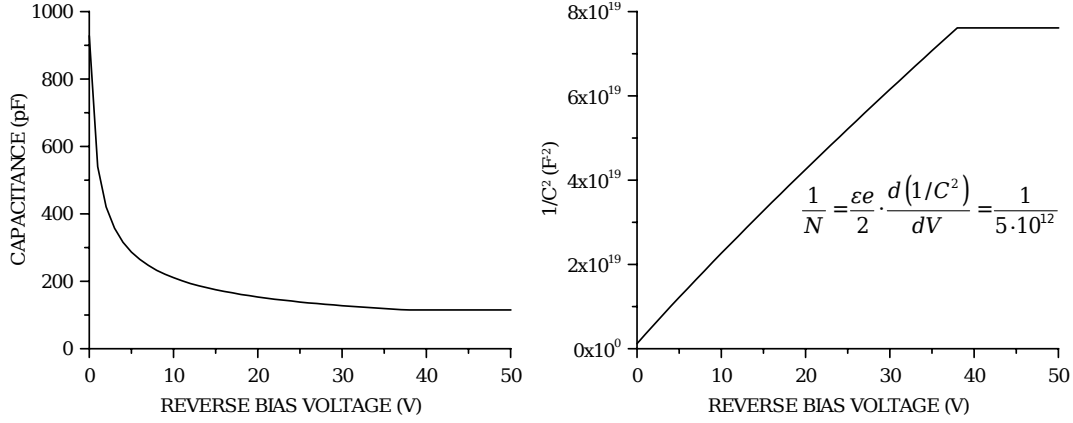


Figure 3.16: Diode C-V characteristics. Left is capacitance C plotted against the absolute reverse bias voltage, right is a different representation of $1/C^2$ vs. voltage [72].

3.2.5 Capacitance-Voltage Characteristics

Because of the depletion region, a diode is also a capacitor. The capacitance of a flat p-n junction is analogous to one of a parallel plate capacitor

$$C = A \cdot \frac{\varepsilon}{d} = A \sqrt{\frac{\varepsilon e N}{2(V_{bi} - V_b)}} \quad (3.25)$$

It depends on the interface area A of the p-n junction and its width d . With increasing reverse bias the width d of the depletion region becomes larger and the capacitance decreases as shown in Fig. 3.16. Because of the relation to the bias voltage it is also often plotted as $1/C^2$ vs. V_b . The sensor capacitance is important because it directly contributes to noise as $ENC \propto C$. Therefore it should be kept as low as possible.

3.2.6 Hall Effect

The ID of the ATLAS detector is placed inside a magnetic field which bends the tracks of the outgoing particles. The magnetic field also affects the free charge carriers within the detector medium. The Hall effect is the generation of a potential difference V_H perpendicular to an applied electric field in the presence of an external magnetic field B that is perpendicular to

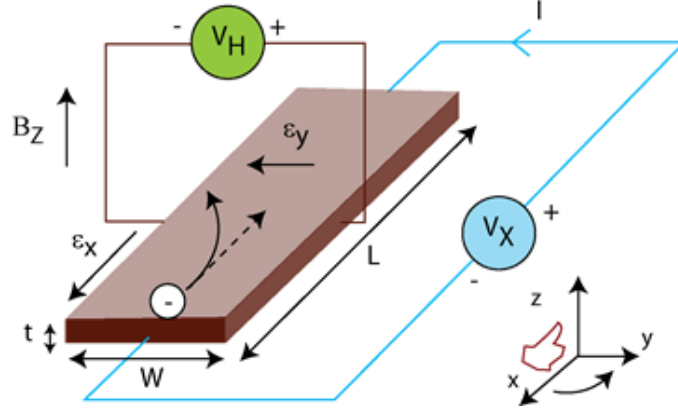


Figure 3.17: Hall effect where the electron is deflected to the left [76].

both as shown in Fig. 3.17. The reason is the contribution of the Lorentz force $\mathbf{F}_L$ in $\mathbf{F} = \mathbf{F}_E + \mathbf{F}_L = -e(\mathbf{E} + \mathbf{v} \times \mathbf{B})$.

The Lorentz force on the electrons works in negative y-direction. This results in the electric field component $E_y = \frac{eB\tau}{m} E_x$ and thus the deflection angle of the electrons from their original track is

$$\tan \theta_H = \frac{E_y}{E_x} = \frac{eB\tau}{m} = \sigma_0 R_H B \quad (3.26)$$

with Hall angle θ_H , also known as Lorentz angle, mean scattering time τ , the electric conductivity $\sigma_0 = \frac{ne^2\tau}{m}$ and Hall constant $R_H = -\frac{1}{ne}$.

For silicon the Hall constant is more complex, see Section 4.1.5. The above expression can also be written as

$$\tan \theta_H = \mu_H B = r \mu_d B \quad (3.27)$$

with Hall mobility μ_H , Hall factor r and drift mobility μ_d . Other works e.g. [73] and [74] have measured numerical values for these parameters. In [75] the Lorentz angle was modelled for the ATLAS silicon detectors.

3.2.7 Metal-Oxide-Semiconductor

Thanks to the semiconducting property, the easy manipulation thereof through doping, and the inexpensive fabrication (cf. Section 3.2.8), silicon de-

vices are widely used. One of the prevalent semiconductor devices in integrated electronics are metal-oxide-semiconductor field-effect transistors (MOSFETs).

A MOSFET consists of a doped semiconductor substrate with so-called source and drain, which are regions that have a different doping than the substrate. A gate is between these highly doped regions and consists of an oxide and metal layer on top of the substrate. Depending on the source and drain doping, there are p-MOS and n-MOS transistors. Fig. 3.18 shows an n-MOS transistor. By applying a positive voltage on the gate, the minority charge carriers in the bulk, the electrons, would wander to the Si-SiO₂ interface and recombine with the holes. When the gate-source voltage V_{GS} is equal to the threshold voltage V_{TH} , all the holes are filled with electrons. For $V_{GS} > V_{TH}$ the electrons can move freely and create a conductive channel below the gate so that current can flow from source to drain. The I-V characteristics of a MOSFET are shown in Fig. 3.19. For small drain voltages the current increases approximately linearly and reaches a saturation for larger drain voltages. A MOSFET can be used as an amplifier by connecting a resistor in series with the drain output and using a fixed source-drain voltage, e.g. 800 Ω and 10 V. Then the drain-source voltage is 8 V if a current of 10 mA passes through the resistor. This way a load line can be drawn and the amplification for each value of $V_{GS} - V_{TH}$ can be read along the line as shown in Fig. 3.19.

MOSFETs as amplifiers and fast switching components require very little current to operate but can deliver large currents. They have long replaced vacuum tubes. The smallest possible gate size that can be reached and is mass-produced today is 7 nm [77] which is already not far from the quantum mechanic limit of 2–3 nm [78]. Due to the small feature size, MOSFETs are the core of integrated circuits on computer chips and readout chips of particle detectors.

3.2.8 Semiconductor Fabrication Process

Monocrystalline semiconductor wafers, e.g. silicon, can be produced using float zone (FZ) silicon or Czochralski (CZ) silicon. While FZ silicon is high-purity detector grade, the CZ silicon is industrial grade, much cheaper to produce and used in integrated circuits.

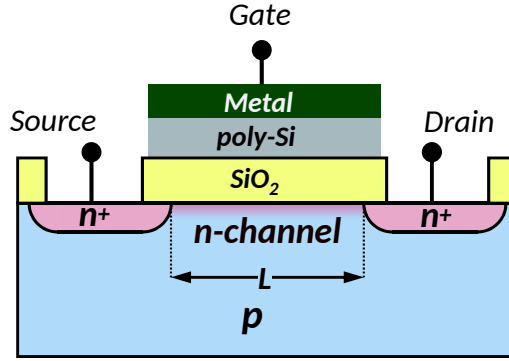


Figure 3.18: 2D schematic of a MOSFET [65].

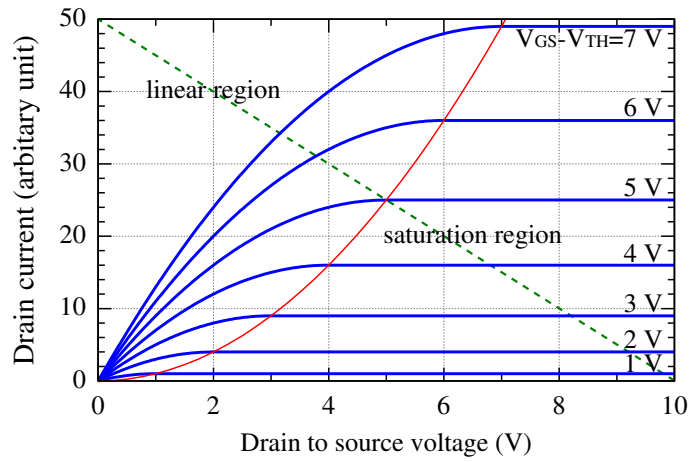


Figure 3.19: I-V characteristics of a MOSFET [79]. The red solid line separates the linear region saturation region of the transductance, and the load line is represented by the green dashed line.

The FZ silicon is produced using a prepared polycrystalline silicon rod surrounded by a ring-shaped heater. The heater is passed from one end to the other, melting the rod only in one location at a time so that the impurities wander into the melt and move along with it, hence float zone.

The CZ process is named after its inventor. A small seed of single crystal silicon is dipped into a pot of silicon melt and slowly pulled up so a single crystal is grown onto the seed.

After growing the ingot that becomes up to 30 cm in diameter it is sliced into single wafer layers of some hundreds μm which are then polished. The structural patterns are created using different masks.

P-N junctions are usually introduced into a doped wafer by ion implantation. Ions of different elements are accelerated and bombard onto the wafer in different concentrations, energies and angles. Making use of the Bragg curve the ions are stopped precisely in the desired depths. To create a pattern usually a thick oxide layer is deposited using a positive mask. The implanted ions are essentially intentionally introduced impurities and point defects which have to be cured and electrically activated, i.e. integrated into the lattice, by Rapid Thermal Annealing at high temperatures, e.g. boron only becomes active at above 950 °C [80].

Surface structures can be created using photolithography by the means of a positive or negative photoresist layer. A positive photoresist becomes soluble through UV light and hardens without UV while a negative photoresist acts the opposite. The lift-off method deposits first a pattern of photoresist, then the material is deposited on top of it as shown in Fig. 3.20 (a). When the photoresist is removed (lifted-off), the material on top of the photoresist will also be removed while the material in the gaps will remain attached to the underground. Another way is deposition and etching where the desired material is directly deposited onto the surface of the silicon. A patterned photoresist is applied on top of it. During etching the photoresist protects the content beneath while the deposit in the gap is etched away. The lift-off method is less neat and clean so usually it is only applied when the effects of etching onto the material surface are undesirable.

3.2.9 Charge Collection

Energy deposition of an incident particle results in creation of pairs of opposite charge carriers. In this section the concept of charge collection is explained.

The term “charge collection” is actually misleading because it would suggest that the signal current is only formed once the moving charge q reaches the electrode. However, this is not the case. In fact it has been observed that the collection electrode starts to bear current before any charge reaches it [81]. The signal current starts to form as soon as the charge begins to move. Therefore, the signal formed on the electrode is the induced charge. In an ideal world

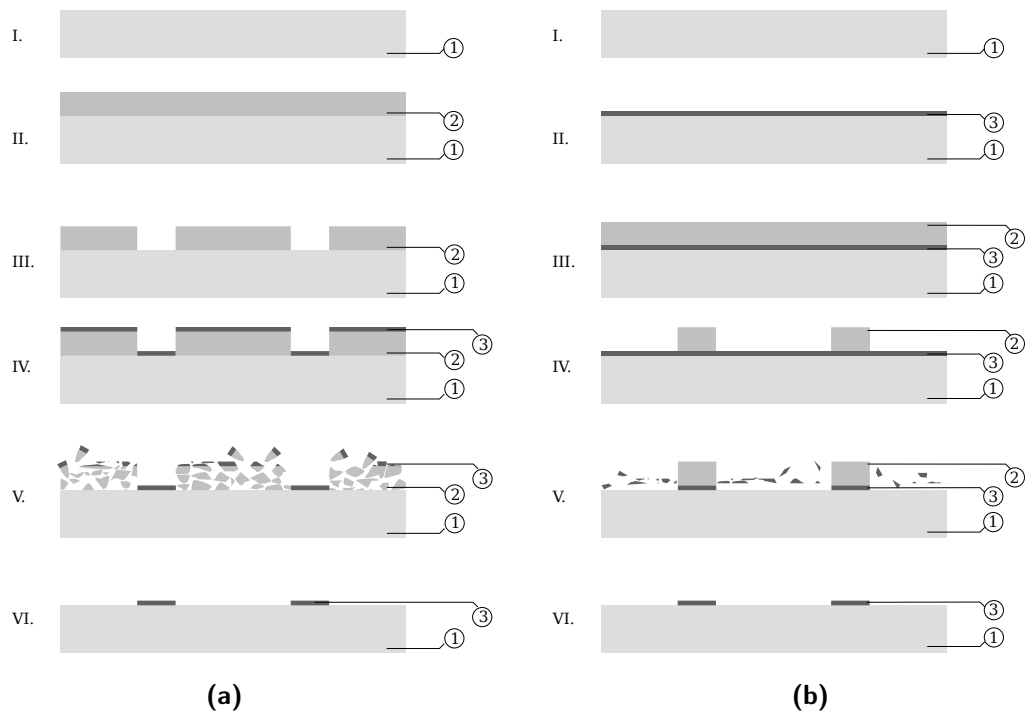


Figure 3.20: Lift-off (a) and deposit and etch (b) photolithography. ① is the silicon substrate, ② is the photoresist and ③ is e.g. the deposit.

there would be one infinite electrode in vacuum. According to Gauss' law

$$Q = \oint_S \varepsilon \mathbf{E} \cdot d\mathbf{S} = -q \quad , \quad (3.28)$$

the induced charge Q on the electrode can be obtained by integrating the electric field $\mathbf{E}$ over the surface S enclosing the electrode, whereas ε is the dielectric constant of the medium. In this ideal case the induced charge is just the opposite of the moving charge q . In real life however, there are more than one single electrode in a sensor and this method requires the calculation of the electric field at every point in space of the moving charge.

3.2.9.1 Shockley-Ramo Theorem

Shockley and Ramo independently found a simpler approach [82, 81, 83, 72] where only one field calculation is required.

First assume that induction due to magnetic effects is negligible and the electric field propagates instantaneously. Consider a system of collection electrodes on certain voltages and a charge q as shown in Fig. 3.21. Due to the superposition principle the system can be disassembled into a component of charged electrodes and space charge between them but without the moving charge, and a component of grounded electrodes without space charge between them but with the moving charge. It is then evident that if Eq. (3.28) is broken into these components, only the second component would contribute to the induced charge.

ϕ is the potential between the electrodes with the charge q . Consider a second case where one electrode is raised to the unit potential, the others are left grounded and the charge is removed. The second potential is ϕ' . It is

$$\nabla^2 \phi = 0 \quad \text{and} \quad \nabla^2 \phi' = 0 \quad . \quad (3.29)$$

By playing this trick Green's theorem can be applied

$$\int_V (\phi' \nabla^2 \phi - \phi \nabla^2 \phi') dv = \int_S \left(\phi' \frac{\partial \phi}{\partial n} - \phi \frac{\partial \phi'}{\partial n} \right) ds \quad , \quad (3.30)$$

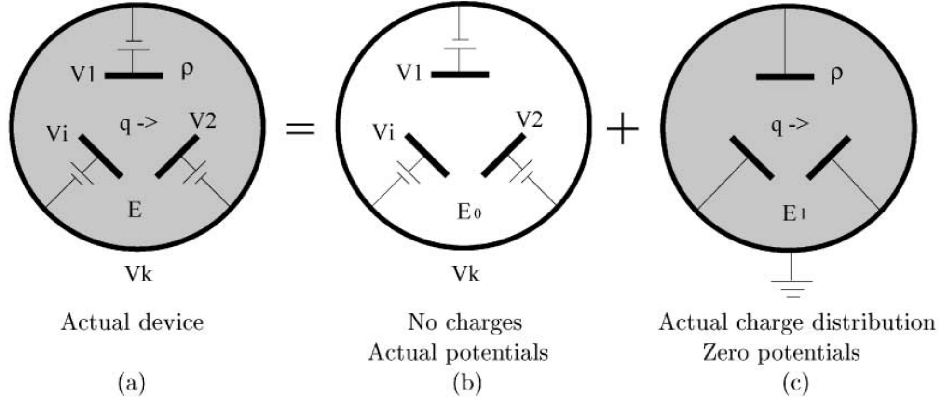


Figure 3.21: Linear superposition principle: the actual device (a) with a moving charge q , electrodes on potentials V_i and a space charge ρ can be linearly composed of its components displayed in (b) and (c) [83].

where n notes the surface normal. The left-hand side evaluates to zero and the right-hand side can be split into three parts:

1. the integral over the surfaces of all electrodes except for the one with the unit potential. In this case it is $\phi = \phi' = 0$.
2. the integral over the electrode that is considered. In this case it is $\phi = 0$ and $\phi' = 1$ for this electrode and the right-hand side reduces to $\int_S \frac{\partial \phi}{\partial n} ds = 4\pi Q$.
3. the integral over a tiny spherical surface of the charge q : $\int_S (\phi'_q \frac{\partial \phi}{\partial n} - \phi_q \frac{\partial \phi'}{\partial n}) ds$ whereas $\int_S \frac{\partial \phi'}{\partial n} ds$ evaluates to zero because in the second case the charge q was removed. This part yields $\int_S \phi'_q \frac{\partial \phi}{\partial n} ds = 4\pi q \phi'_q$.

In total it is

$$0 = \int_{\text{electrode}} \frac{\partial \phi}{\partial n} ds + \int_{\text{sphere}} \phi'_q \frac{\partial \phi}{\partial n} ds = 4\pi Q + 4\pi q \phi'_q \quad , \quad (3.31)$$

which leads to the induced charge on the electrode

$$Q = -q \phi'_q \quad (3.32)$$

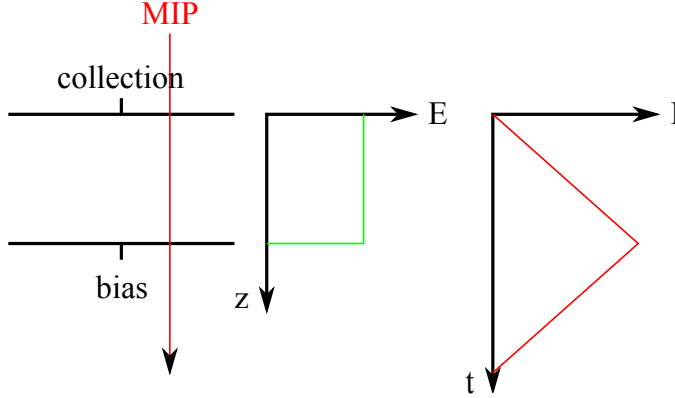


Figure 3.22: Induced current by a MIP on the collection electrode according to Shockley-Ramo theorem, assuming uniform electric field and weighting field.

and the induced current

$$I = \frac{dQ}{dt} = -q \frac{d\phi'_q}{dt} = -q \frac{\partial \phi'_q}{\partial \mathbf{x}} \frac{d\mathbf{x}}{dt} = q\mathbf{v}\mathbf{E}_q \quad , \quad (3.33)$$

with $\mathbf{v}$ as the velocity of the moving particle. ϕ'_q is called the weighting potential and $\mathbf{E}_q$ the weighting field.

In a silicon sensor, a charge carrier's velocity depends on the mobility as discussed in Section 4.1.5. Assuming a uniform electric field and weighting field as indicated in Fig. 3.22, then the velocity is constant throughout the depth of a planar sensor. The induced current on the collection electrode increases linearly as the MIP is passing through the sensor and starts to decrease as the first charge carriers reached the collection electrode.

3.3 Radiation Damage in Silicon Sensors

As described in previous sections, a heavy charged particle loses energy primarily due to ionisation, but a fraction also goes into NIEL. However, the damage caused by the former is reversible and exists in a much shorter time scale than the latter. Ionising radiation affects the silicon surface while NIEL damages persist in the bulk.

3.3.1 Surface Damage

An ionising particle creates e-h pairs in silicon. If no electric field is applied, the electrons and holes recombine immediately. Otherwise the created electrons and holes are collected by the electrodes applying the bias voltage. On the surface of silicon however, silicon oxide (SiO_2) is present, e.g. as a passivation layer, gate oxide or shallow trench isolation for embedded electronics. In SiO_2 the mobility of electrons and especially holes is much lower: $\sim 20 \text{ cm}^2/\text{Vs}$ and $\sim 10^{-9} \text{ cm}^2/\text{Vs}$ at 300 K, whereas the electron mobility is temperature-independent and the hole mobility depends exponentially on the temperature [84]. Thus they do not recombine immediately. The electrons diffuse out of the oxide. The holes become trapped by defects. In addition, trapping centres activated by ionising radiation at the Si- SiO_2 interface can trap electrons or holes. Hence, a positive charge accumulates in the oxide. One example of this effect in action can be found in [85].

3.3.2 Bulk Damage

As shown in Section 3.1.3 the bulk damage is induced by non-ionising radiation. Despite the success of the NIEL scaling hypothesis which essentially states that the damage is proportional to the NIEL, more recent studies like [86] and [87] show that not all non-ionising energy deposition goes into dislocations of lattice atoms. Depending on the type of particles and their energies, some fraction of NIEL is channelled into phonon excitations.

Bulk damage comprises two types of displacement: point defects and defect clusters [88]. A point defect is e.g. a (multiple) vacancy, interstitial or a Frenkel pair with a threshold recoil energy of the primary knock-on atom (PKA) of $\sim 25 \text{ eV}$. The threshold energy for the incident neutron or electron to induce a single defect is 175 eV and 260 keV, respectively. A defect cluster is a dense agglomeration of defects and it is formed at the end of a track of a primary knock-on atom with a recoil energy $> 5 \text{ keV}$. The threshold energy for the neutron or electron in this case is much larger: $\sim 35 \text{ keV}$ and $\sim 8 \text{ MeV}$, respectively. On average, a 1 MeV neutron can create a 50 keV primary knock-on atom [88]. The generated clusters and their dimensions are shown in simula-

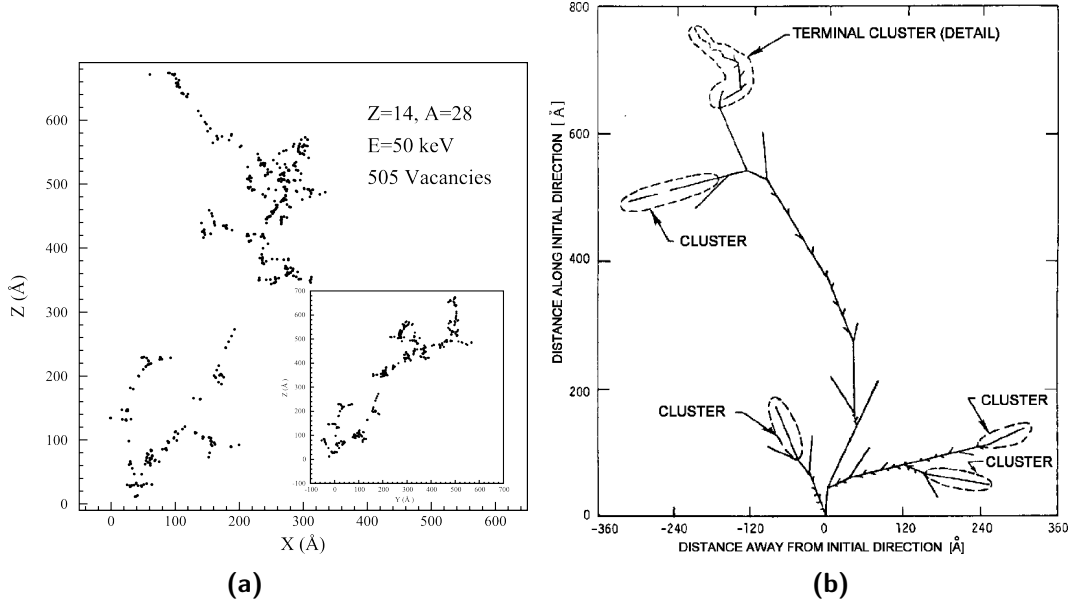


Figure 3.23: Simulation of defects generated by a 50 keV primary knock-on atom (a) [86] and (b) [64].

tion in Fig. 3.23. If occupied, the depletion region caused by a cluster is much larger than the cluster itself [89].

The defects in the bulk create additional donor and acceptor levels in the band gap as shown in Fig. 3.24. Shallow levels are near the edges of conduction or valance band and are ionised at room temperature, in contrast to deep levels. This has three consequences: the change of effective doping concentration, an increased leakage current and a decrease of charge collection efficiency.

3.3.2.1 Effective Doping Concentration

The effective doping concentration is the difference between the donor concentration N_D and acceptor concentration N_A :

$$N_{\text{eff}} = N_D - N_A \quad . \quad (3.34)$$

The change in effective doping concentration ΔN_{eff} is the difference between initial concentration $N_{\text{eff}0}$ and concentration $N_{\text{eff}}(\Phi_{\text{eq}}, t(T_a))$ after fluence Φ_{eq} and annealing (ref. Section 3.3.2.5) time t at temperature T_a . It is often parametrised containing three components: short term annealing N_a , stable

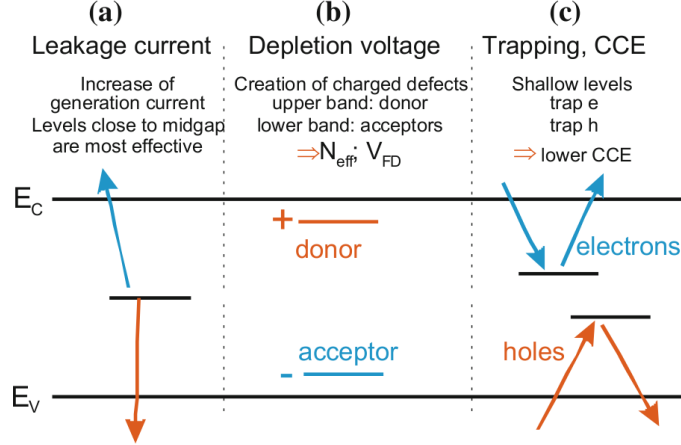


Figure 3.24: Energy level of defects [90]. (a) Mid-gap levels generate leakage current. (b) additional donor and acceptor levels affect the depletion behaviour. (c) Deeper levels act as trapping centres and decrease the charge collection efficiency.

damage N_c and long term annealing N_y :

$$\Delta N_{\text{eff}} = N_a(\Phi_{\text{eq}}, t(T_a)) + N_c(\Phi_{\text{eq}}, t(T_a)) + N_y(\Phi_{\text{eq}}, t(T_a)) \quad . \quad (3.35)$$

At low fluences up to $10^{12} - 10^{13} \text{ n}_{\text{eq}}/\text{cm}^2$ the resistivity of both n- and p-type silicon actually increases [91] due to donor and acceptor removal described in Section 3.3.2.4. For n-type silicon the change of N_{eff} with increasing fluence leads to space charge sign inversion, also known as type inversion as displayed in Fig. 3.25. A severe result of this effect is the requirement of higher and higher bias voltage to reach the full depletion depth d :

$$V_{\text{dep}} = \frac{q_0}{2\epsilon\epsilon_0} |N_{\text{eff}}| d^2 \quad [88]. \quad (3.36)$$

Silicon processed using the diffusion oxygenated float zone (DOFZ) technology proved to offer more resistance against change in N_{eff} due to proton and pion irradiations, although it is less effective against neutron irradiation. This “proton-neutron puzzle” has not been fully understood yet. The works [64, 88] and literature cited there report a wealth of studies.

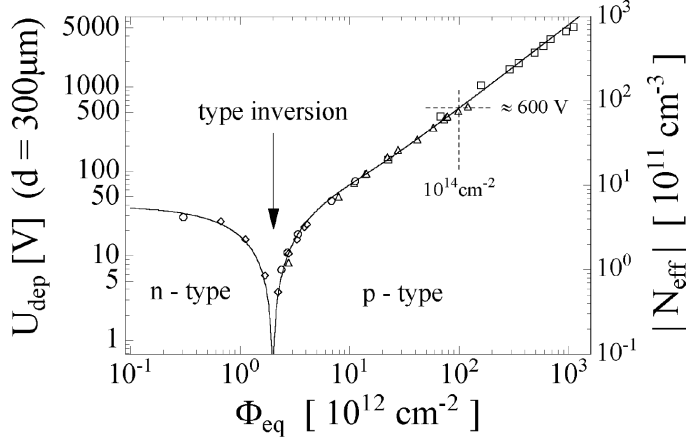


Figure 3.25: Space charge sign inversion (also known as type inversion) of n-silicon at high fluences [64].

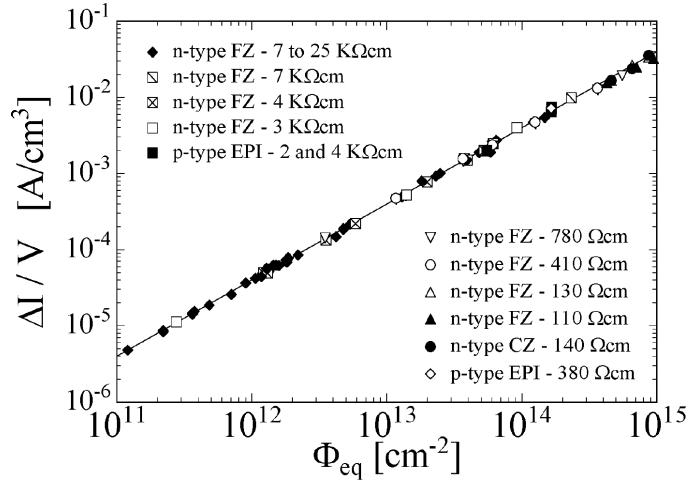


Figure 3.26: Increase of leakage current with fluence [64].

3.3.2.2 Leakage Current

Due to the additional energy levels in the band gap which act as generation or trap-assisted tunneling centres, the leakage current increases with the fluence as shown in Fig. 3.26. It can be reduced with beneficial annealing (ref. Section 3.3.2.5). In detector operation, the leakage current can be reduced by reducing the operation temperature down to cryogenic temperatures in order to reduce the thermal excitation.

3.3.2.3 Charge Collection Efficiency

Loss in charge collection efficiency (CCE) originates from trapping. Free charge carriers can be captured by radiation induced deep defects levels. The release time of the trapped charge might exceed the signal processing time and thus some portion of the charge is “lost”. This results in a small decrease in signal and hence in CCE. Another process that deteriorates the charge collection is that the additional radiation-induced defects can act as scattering centres and thus decrease the mobility of free charge carriers.

There are also effects that actually enhance the CCE. Donor and acceptor compensation by radiation-induced defects causes removal of free charge carriers. Similarly, if the detector is operated at cryogenic temperatures, trapped charges are frozen out due to lack of thermal energy to re-emit them. This is called the Lazarus effect. Both the dopant compensation and the Lazarus effect result in a decrease of free charge carriers and smaller space charge, meaning lower depletion voltage is needed. When all the free carriers are trapped and frozen out, the bulk is fully depleted and the CCE improves.

3.3.2.4 Donor and Acceptor Removal

At small fluences up to 10^{12} – 10^{13} n_{eq} one can observe the effect of donor and acceptor removal [91, 88, 92] where both n- and p-type silicon increase in resistivity which results in an increased depletion depth and reduced depletion voltage. For n-type silicon donor removal dominates and for p-type silicon acceptor removal prevails. Several different models exist trying to explain the mechanism of dopant removal which are based on the assumption that the doping atoms, phosphorous and boron, are removed from their dopant position and become electrically inactive [93, 91]. The reason is that silicon self-interstitial and -vacancies are unstable thus would either self-annihilate or interact with impurities [93]. The rate of the increase in resistivity and thus the dopant removal rate is different for n- and p-types, as shown in Fig. 3.27 as a function of material resistivity ρ with

$$N_{eff} = -\frac{1}{q\mu_{e,h}\rho} \quad , \quad (3.37)$$

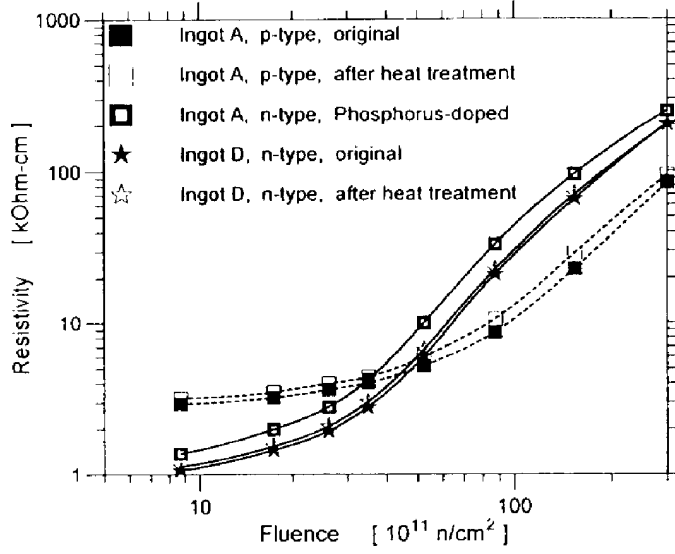


Figure 3.27: Resistivity as a function fluence [91]. Ingot A n-type is the same p-type wafer as ingot A p-type but with a neutron transmutation doping to convert silicon atoms into isotopes which subsequently decay into phosphorous atoms.

where N_{eff} is the net impurity concentration, $\mu_{e,h}$ is the electron and hole mobility, and q is the electron or hole charge, $-e$ and e , respectively.

The dopant removal can be described by

$$N_{P,B}(\Phi) = N_{P0,B0} e^{-c_{D,A}\Phi} \quad . \quad (3.38)$$

$N_{P,B}$ is the net phosphorous and boron concentration after irradiation, respectively, the subscript 0 indicates the initial doping concentration before irradiation. $c_D = 2.41 \cdot 10^{-13} \text{ cm}^2$ and $c_A = 1.98 \cdot 10^{-13} \text{ cm}^2$ are the donor and acceptor removal rate, respectively, according to [91]. However, as both dopants are in a silicon wafer with unknown concentration, normally only the change in their relative difference N_{eff} can be measured.

3.3.2.5 Annealing

Not all displacement damage caused by NIEL is permanent. With time and temperature some of them can be repaired, which is called annealing. As described in Eq. (3.35), there are short term and long term annealing (N_a and

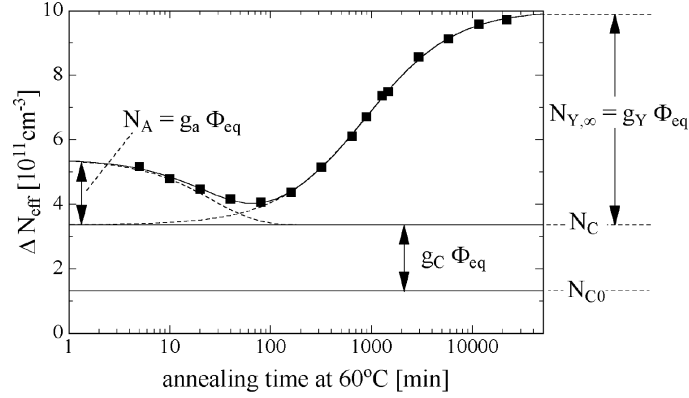


Figure 3.28: Change in N_{eff} as a function of annealing time. The sample is an n-type silicon with $25 \text{ k}\Omega\text{cm}$ irradiated with a fluence of $1.4 \cdot 10^{13} \text{ cm}^{-2}$ [64].

N_y). The constant damage that cannot be repaired is called stable damage N_c . The composition of the three components can be seen in Fig. 3.28. Annealing occurs due to reordering of defects. Initially produced defects are not stable and would wander through the lattice to recombine, forming more stable or complex defects, or they could be dissolved by phonons.

Annealing on a short time scale is observed to be beneficial. The effective doping concentration N_{eff} increases so that for type-inverted detectors the negative N_{eff} becomes less negative and for non-type-inverted detectors the positive N_{eff} becomes more positive. On a long time scale however, more acceptors build up and N_{eff} would decrease. That's why it is also called beneficial and reverse annealing, respectively.

Chapter 4

Radiation Damage Pixel Digitiser

The correct modelling of the interaction of particles in the detector requires thorough consideration of the properties of the materials the particles interact with. However, the current Athena framework — the software framework that manages almost all production workflows in ATLAS: event generation, simulation, reconstruction and derivation production, and is also used online in the High Level Trigger — does not account for radiation damage effects in silicon detectors. As can be seen from Section 2.3 and Section 3.3.2 these effects are not negligible. In the simulation framework these can be accounted for in the digitisation step that converts charge depositions by particles into digital signals.

In a real detector system the digitisation refers to the conversion of the analog signal from the sensor to a digital signal that is created in the FE readout chip. The signal current of a pixel (ref. Section 3.2.9.1) can be digitised using a sampling **adc!** (**adc!**) to save the shape of the signal at discrete points, or, if the signal shape is known, the time it exceeds a certain threshold, time-over-threshold (ToT), can be saved, which is the case for the IBL, using the FE-I4 readout chip. The threshold is important to distinguish a real signal from noise. In the simulation, the ToT can be calculated using the threshold and the charge-to-ToT conversion (ref. Section 4.1.9). However, the impact of radiation damage in silicon is also manifested in the shape of the electric field and the trapping probability of charge carriers. Therefore, the implementation of this digitiser has to consider the whole process from charge deposition to the output of the digital signal.

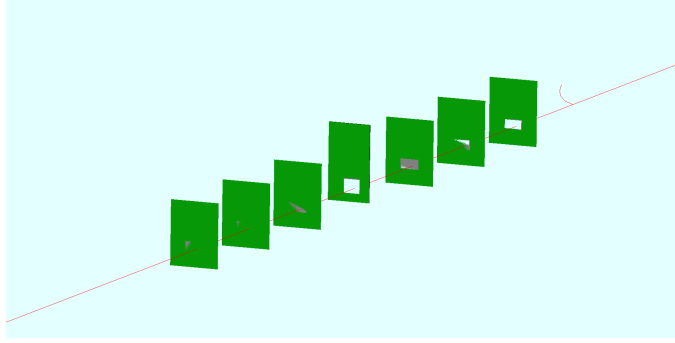


Figure 4.1: AllPix event display of a particle passing (red) passing through six telescope planes and a DUT (middle plane).

Before implementing a radiation damage pixel digitiser into a complex working framework, it is necessary to develop and test it in smaller frameworks and validate it with data. The AllPix simulation framework [94, 95] is a small and less complex framework ideal for this purpose. Using also the EUTelescope framework for reconstruction and pyEudetAnalysis for analysis, the radiation damage pixel digitiser was developed and validated.

AllPix is a simulation framework based on Geant4 [96, 97]. It provides a popular platform for simulations of e.g. test beam setups. It includes a range of different digitisers, a collection of different pixel detector geometries and offers different macro templates. Each detector can have several unique identifiers (ID), one identifier for each instance in the simulation so that individual sensors of a setup of the same type have the same physical properties but can have e.g. different digitisers. In the simulation macro the user can define the detector arrangement and non-detector materials, the beam properties, the number of events and control visualisations. The output formats include telescope files with one text file for each event, ROOT format and slcio format.

EUTelescope is a reconstruction framework developed for European Detector R&D (EUDET)-type beam telescopes [26]. The reconstruction steps are converting, clustering, hit making, aligning and fitting, where the converting step is optional if the data are already in the correct format and noisy pixels do not have to be dealt with. Each step needs an input and an output file where the output file feeds into the next step. Each step requires a corresponding steering file, in which different data processors can be used. The main config-

uration file sets the parameters for all the steps which also includes a gear file describing the setup and sensor geometries. A job submission module is used to run each step.

pyEudetAnalysis is a small set of standalone analysis codes written in Python to process the EUTelescope output. It also filters noisy pixels, builds clusters and does alignment. In order to analyse the DUT using this framework, the reconstruction in EUTelescope only uses data from the telescope planes.

4.1 Development of the Digitiser

The radiation damage pixel digitiser described here is developed for Insertable B-Layer (IBL)-like planar silicon pixel modules (ref. Section 2.2.1). A digitiser for 3D sensors is also being developed but is not the scope of this work.

To limit the computing power required to carry out calculations of the physical properties of the silicon detector and the propagation of each individual charge carrier, the digitiser makes use of many lookup maps and bundles charge carriers into chunks. The digitiser takes as input Technology Computer Aided Design (TCAD) maps in 1D or 3D for the electric field and weighting potential (ref. Shockley-Ramo theorem in Section 3.2.9.1). Two further lookup tables are usually calculated: the position of a charge depending on the location of creation and the time travelled (Section 4.1.2), and the time the charge needs to reach the electrode (Section 4.1.3). The total energy deposition is split into a discrete number of subcharges or charge chunks. Each charge chunk is drifted through the detector under consideration of the Hall effect and whether a Gaussian diffusion is applied to the chunk. The time a particle needs to reach the collecting electrode is looked up and compared to the randomly generated drift time, which is the lifetime of a charge chunk before it gets trapped. The induced charge is calculated according to Eq. (3.32) for the pixel at the coordinates of the subcharge and two neighbouring pixels to each side.

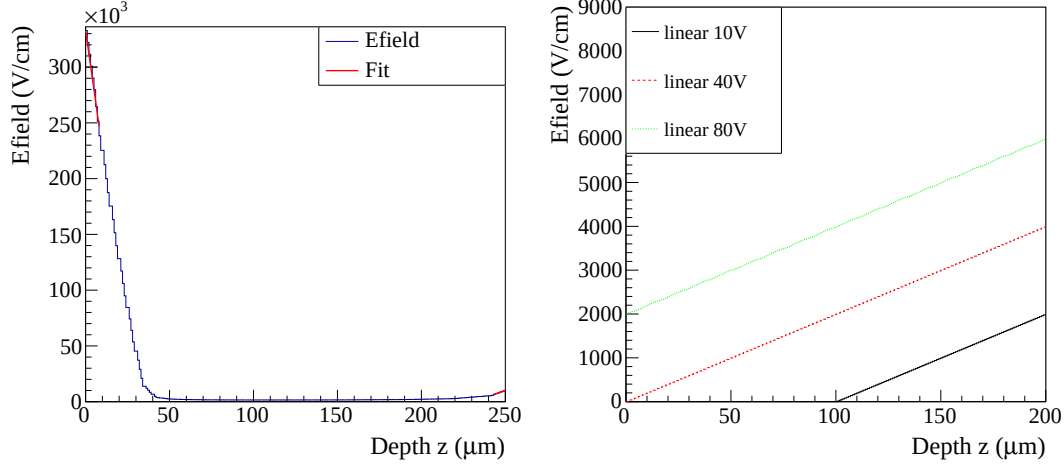


Figure 4.2: Left: An example of an electric field with linear fits at top and back sides since the actual TCAD map is shorter. Right: 1D linear electric field maps generated in the digitiser for a 200 μm thick sensor biased at different voltages.

4.1.1 Electric Field

The 3-dimensional electric field maps are created using a Technology Computer Aided Design (TCAD) model (ref. Section 5.2) of the sensor, which outputs the field as a text file that is converted into a ROOT file as input for the digitiser. In the absence of a 3D electric field map the digitiser looks for a 1D map. Due to implants on the top and back side of the sensor the electric field maps might not contain the extreme fields at both ends when exported from TCAD. Thus the field map is actually shorter than the sensor thickness. In ROOT, the missing values for the electric field are replaced by zero by default, thus causing infinite loops when running the simulation. To prevent the non-physical values of zero time-to-electrode at zero field values, the top and back side of the sensor with missing electric field values are determined using a linear extrapolation, which is a simplification of the real and more complicated field topology.

In case no electric field map is provided, an approximated 1D map can be generated where the user can choose between either a uniform field or a linear field. The uniform field strength can be changed in the source code and the depletion depth is $d = L\sqrt{V_{\text{bias}}/V_{\text{depl}}}$, where L is the sensor thickness. V_{depl}

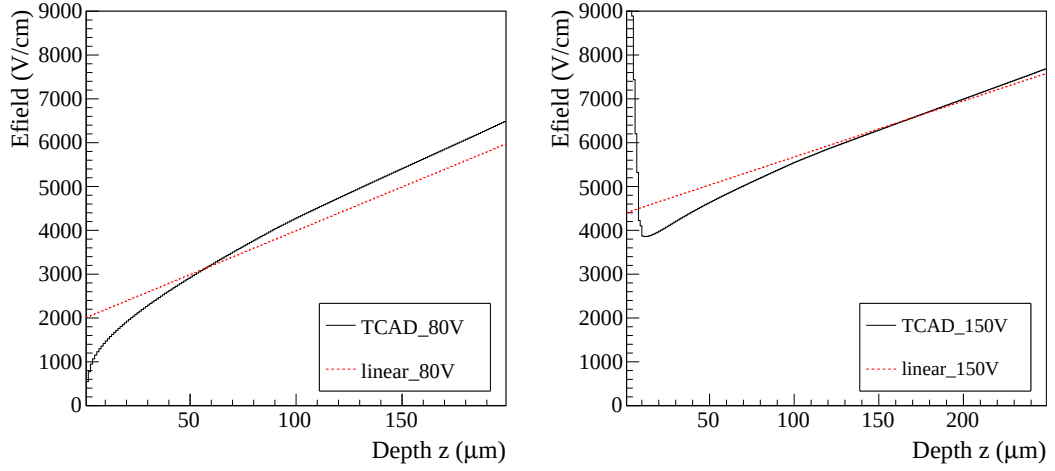


Figure 4.3: Comparison of 1D electric field maps generated in the digitiser (linear) and using TCAD. Left: for a 200 μm thick sensor biased at 80 V. Right: for a 250 μm thick sensor biased at 150 V. The collection electrode and thus the n^+ doping is located at $z = 0$. The steep decline near $z = 0$ in the right plot is due to the implant.

and V_{bias} are the absolute depletion and bias voltage, respectively. The linear field for position $z \in [0, L]$ based on [72] is

$$\begin{aligned}
 E &= \frac{2V_{\text{bias}}}{d} \cdot \left(1 - \frac{z}{d}\right) && \text{for } V_{\text{bias}} < V_{\text{depl}} \text{ and} \\
 E &= \frac{2V_{\text{depl}}}{L} \cdot \left(1 - \frac{z}{L}\right) + \frac{V_{\text{bias}} - V_{\text{depl}}}{L} && \text{for } V_{\text{bias}} \geq V_{\text{depl}},
 \end{aligned} \tag{4.1}$$

The built-in voltage of a p-n junction in detector grade silicon is typically ~ 0.5 V and therefore negligible. The collection electrode (top/front side) is at $z = 0$ and the bias electrode (back side) is at $z = L$. The depletion voltage of an unirradiated sensor is around 30–40 V [98, 99] at which the sensor is fully depleted, which, of course, depends on its thickness. 40 V is used in the digitiser for 200–250 μm thick sensors. Fig. 4.2 shows the electric field at 10 V, 40 V and 80 V. The detector is not fully depleted when the applied bias voltage is smaller than the depletion voltage. The linear approximation is comparable with TCAD simulations for non-irradiated sensors as shown in Fig. 4.3.

4.1.2 Charge Position

Based on the electric field two other lookup tables are created if they are not provided, describing the charge carrier's position and time needed to reach the electrode, respectively. The former is a 2D map of the charge carrier's location depending on the initial position in z and on the time it travelled. The location is calculated separately for electrons and holes since they have different mobilities and travel in opposite directions. The time a charge carrier travelled is the partial sum of $\Delta t = \Delta z / \mu_{e,h} E$, where Δz equals the bin width in z , and the charge carrier mobility $\mu_{e,h}$ depends on the electric field E , which in turn depends on the position. Due to the discrete nature of the electric field map the calculation does not yield a value for every bin of the 2D map. The lookup tables with missing values are shown in Fig. 4.4. If the missing values are assumed to be zero, this results in inaccuracies and errors in the simulation. This effect is present at lower bias voltages — below the depletion voltage. While for a fully depleted sensor this effect does not have a large impact on the electron distance map, since the missing values would be mainly near zero. For the partially depleted sensor however, the missing values would be in the medium range. Therefore, a 2D fit is applied to electrons and holes separately to fill in the missing values in the map as shown in Fig. 4.5. The fit has the form

$$f(x, y) = \sum_{n=0}^1 p_{5n}(x + p_{5n+1})^n + p_{5n+2}(x + p_{5n+3})^n(y + p_{5n+4}) \quad (4.2)$$

with free parameters p_i and the possibility of extending the sum. It only fits in the z region where the electric field is nonzero. The fit is good for the range of bias voltages where using a generated electric field map is reasonable as can be seen in Fig. 4.6. For a partially depleted sensor the electron will never travel to $z = 0$ and reach the collection electrode. However, as long as it is drifting in the depleted area, it induces charge, until it reaches the zero-field region and recombines.

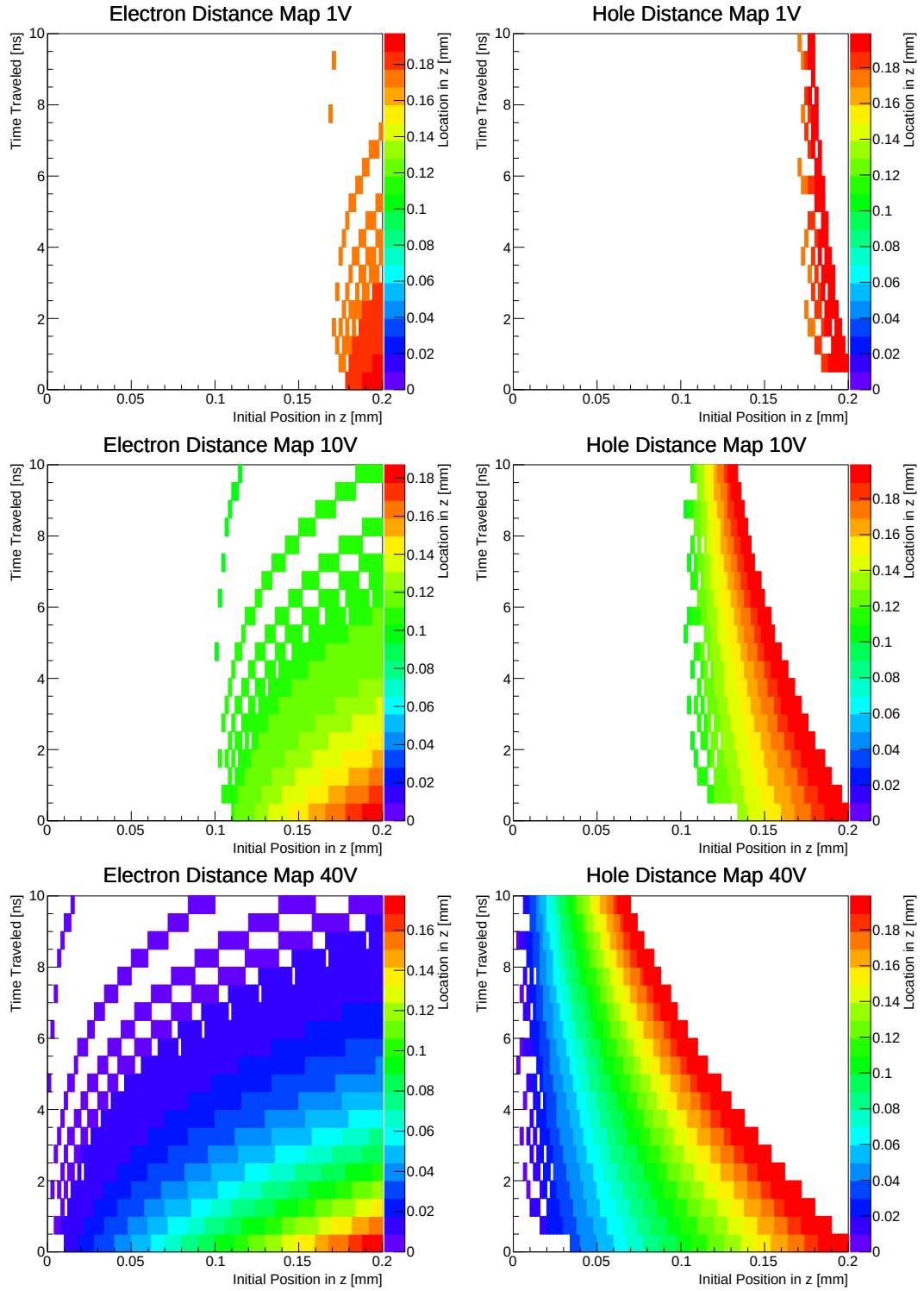


Figure 4.4: Distance maps for electrons (left) and holes (right) for bias voltage of 1 V (top), 10 V (middle) and 40 V (bottom).

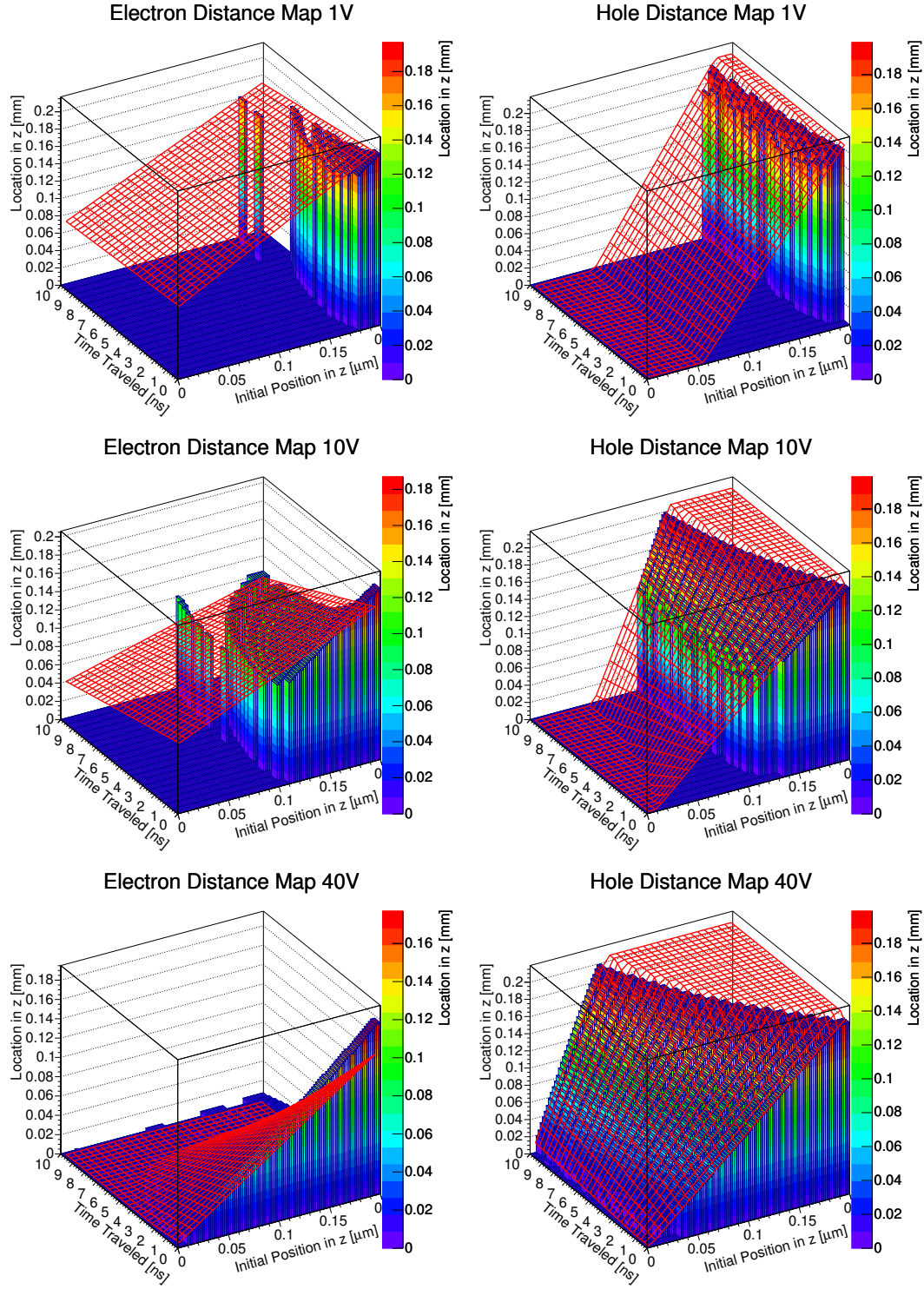


Figure 4.5: 2D fit of distance maps for electrons (left) and holes (right) for bias voltage of 1 V (top), 10 V (middle) and 40 V (bottom).

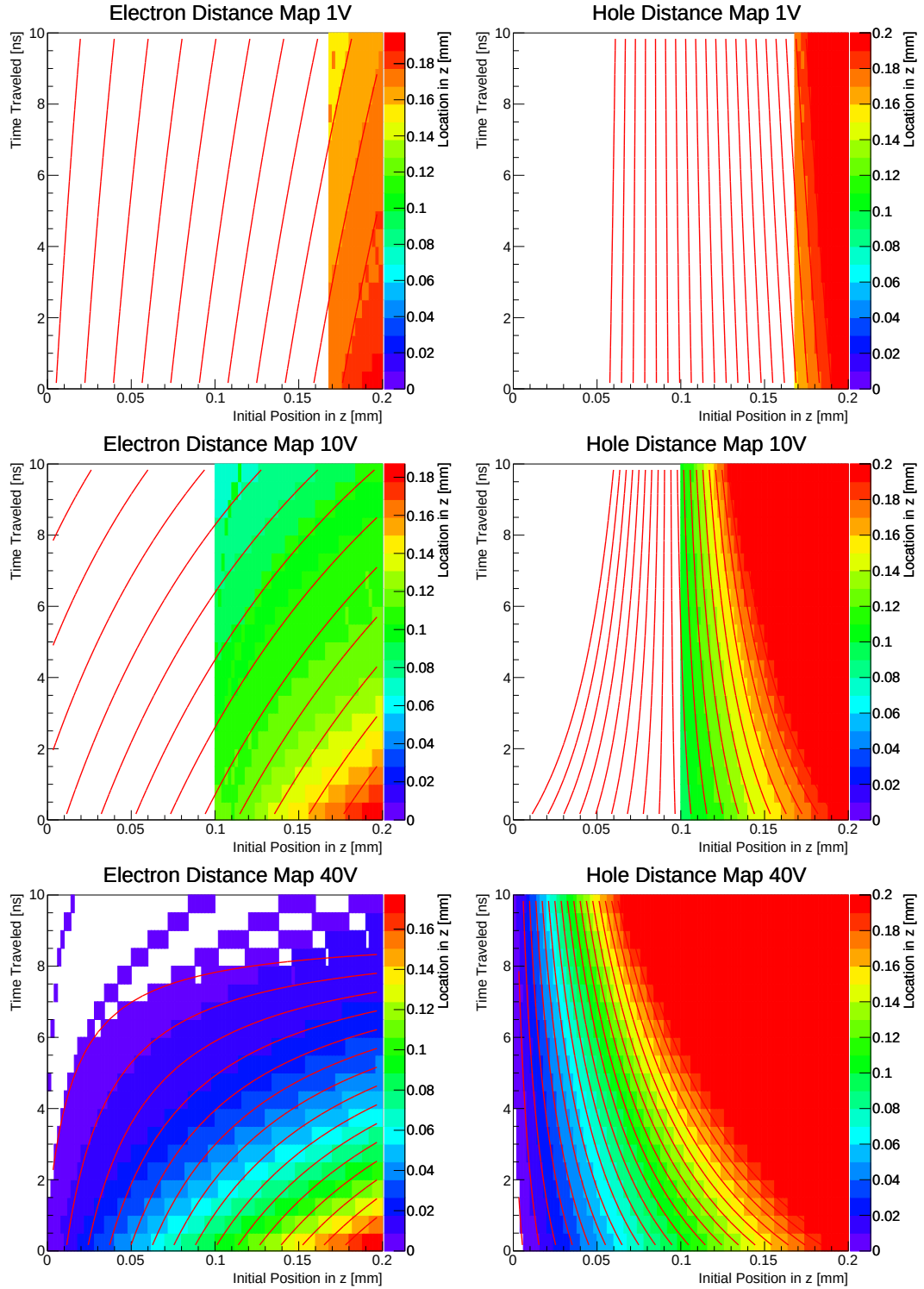


Figure 4.6: Distance maps after 2D fit with filled missing values for electrons (left) and holes (right) for bias voltage of 1 V (top), 10 V (middle) and 40 V (bottom).

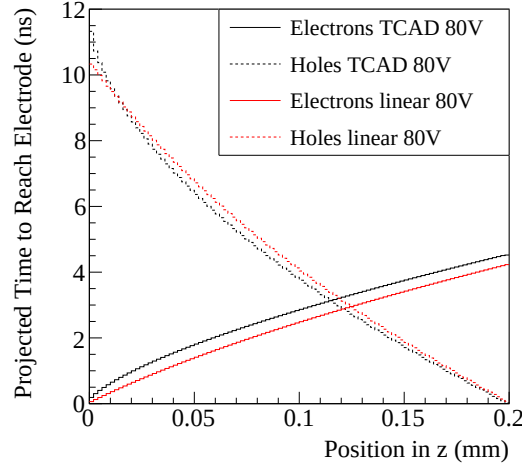


Figure 4.7: Time-to-electrode maps for electrons and holes for a sensor biased at 80 V based on electric field maps generated by TCAD and the digitiser (linear field).

4.1.3 Time to Electrode

The time to electrode is the time a charge carrier needs to travel from its initial position to the collection electrode. It is calculated as the total sum of Δt . A comparison of time maps generated using a linear electric field and a TCAD field is shown in Fig. 4.7. While the generation of time maps is straight forward for a fully depleted sensor, it is more challenging for a partially depleted sensor as the electric field does not span across the whole sensor depth. In regions without electric field there is also no drift mobility. Thus charges generated in zero-field would recombine and never reach any electrode. In regions with finite electric field, the holes moving towards the back side electrode can be treated normally while the electrons will never reach their collection electrode on the top side of the sensor. Because they still drift through the depletion region, they induce charge until they reach the zero-field region. Therefore, the time map for electrons of a partially depleted sensor is the time to reach z with $E(z) = 0$.

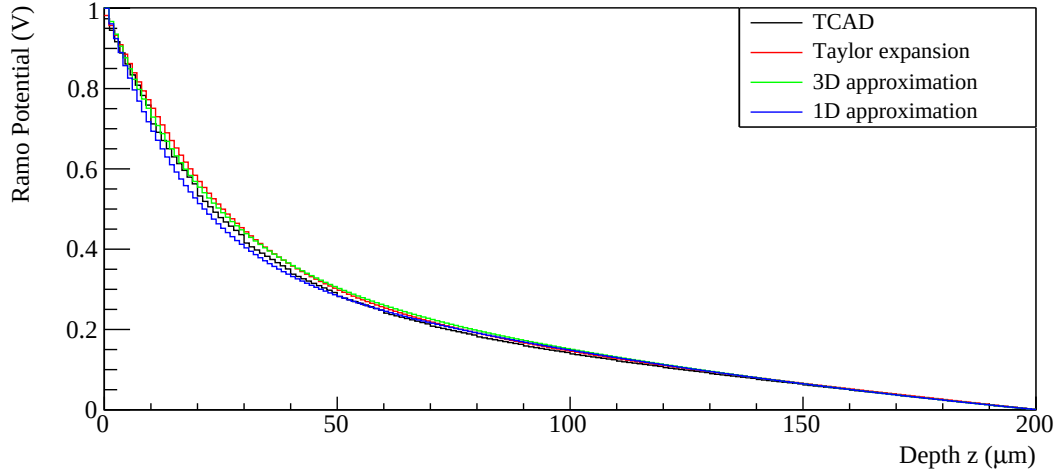


Figure 4.8: Comparison of Ramo maps generated using TCAD and three defaults in the digitiser.

4.1.4 Ramo Potential

The Ramo map follows the same procedure as all other maps. In absence of an input map the user has three choices with increasing complexity and computing time between

- an approximate z -dependent map as a sum of exponential functions,
- an approximate xy -dependent map with compensating z -dependence, and
- a solution to Poisson's equation for a simplified geometry with Ramo terms in series.

Fig. 4.8 shows the difference between the different defaults and the TCAD generated Ramo map. The calculation is explained in [100].

4.1.5 Hall Effect

In the presence of a magnetic field B , the deflection by the Hall angle or Lorentz angle θ_L must be considered (Section 3.2.6). The value of this deflection is looked up in yet another table. In the digitiser this is implemented

	Electrons	Holes
r	$1.13 + 0.0008(T - 273)$	$0.72 - 0.0005(T - 273)$
v_m (cm/s)	$1.53 \cdot 10^9 \cdot T^{-0.87}$	$1.62 \cdot 10^8 \cdot T^{-0.52}$
E_c (V/cm)	$1.01 \cdot T^{1.55}$	$1.24 \cdot T^{1.68}$
β	$2.57 \cdot 10^{-2} \cdot T^{0.66}$	$0.46 \cdot T^{0.17}$

Table 4.1: Numerical values for Hall factor in Eq. (4.3) from [75] and parameters in drift velocity Eq. (4.4) from [101]. T is the absolute temperature.

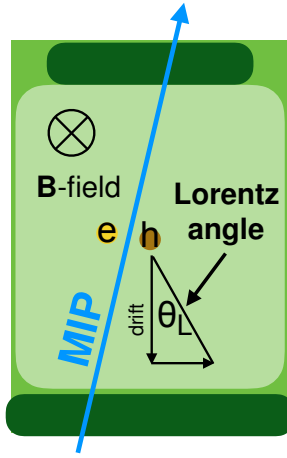


Figure 4.9: Lorentz angle [102].

according to [75]. The magnitude of the deflection depending on the Lorentz angle is given in Eq. (3.27) as

$$\tan \theta_L = \mu_H B = r_{e,h} \mu_{d,e,h} B \quad (4.3)$$

with the Hall mobility μ_H as a product of the Hall factor $r_{e,h}$ and drift mobility $\mu_{d,e,h}$, which is described in [101] by the phenomenological equation

$$\mu_d = \frac{v_m / E_c}{(1 + (E / E_c)^\beta)^{1/\beta}} \quad (4.4)$$

The Hall factor r , the velocity v_m , the electric field E_c , and β have different values for electrons and holes and depend on the temperature T as given in Table 4.1.

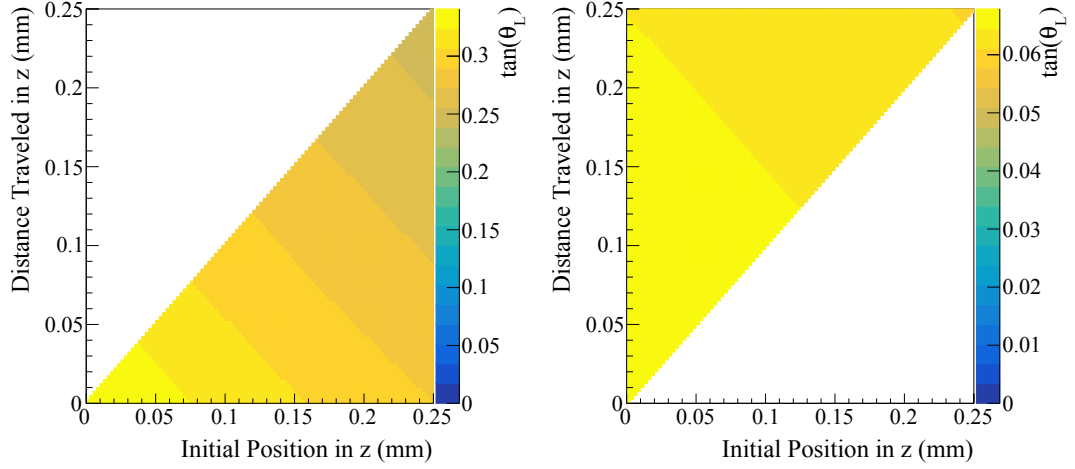


Figure 4.10: Lorentz maps for electrons (left) and holes (right) in a sensor biased at -60 V at the presence of a 2 T electric field.

4.1.6 Radiation Damage

The effects of radiation damage in silicon are included in the electric field map generated using TCAD and in charge trapping implemented in the code which depends on the fluence Φ . Fig. 4.11 shows the shape of the electric field which changes with increasing fluence. Due to the n^+ implant on the top and the p^+ implant on the back side, the double-peak becomes more pronounced with increasing fluence which is known as the double-junction effect [103].

The trapping lifetime τ and trapping probability p for electrons and holes are calculated using

$$\tau_{e,h} = \frac{1}{\beta_{e,h}\Phi} \quad \text{and} \quad (4.5)$$

$$p_{\text{trap}_{e,h}} = e^{-t/\tau_{e,h}} \quad . \quad (4.6)$$

Note that the damage parameter β here is different from the parameter β used in the Hall effect in the previous section. For an irradiated sensor, the drift time is obtained using Eq. (4.6) with a random trapping probability p_{trap} as

$$t_{e,h} = \tau_{e,h} \cdot |\log p_{\text{trap}}| \quad , \quad p_{\text{trap}} \in (0, 1] \quad . \quad (4.7)$$

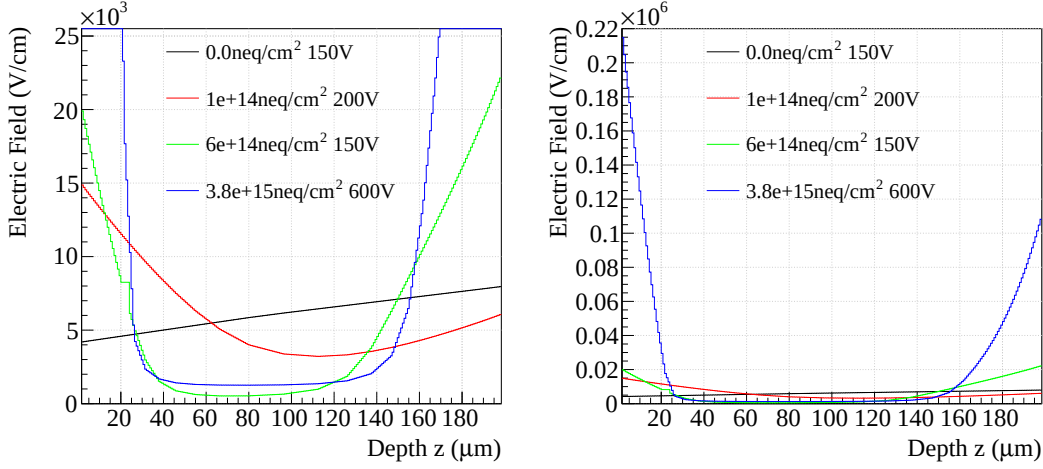


Figure 4.11: TCAD generated 1D electric field at different fluences. The left are the right are the same plot at different scales.

4.1.7 Chunk Correction

A MIP creates ~ 80 e-h pairs per μm silicon travelled as the most probable value (MPV) in the Landau distribution. The computing time can be drastically reduced if the calculations are not executed for every individual elementary charge. Therefore, the total created charge is divided into sub-charges. The number of subcharges (charge chunks) can be defined by the user in the code. It has been observed that the fluctuation in collected charge is larger with smaller numbers of chunks as can be seen in Fig. 4.12. In order not to increase the precision and thus the computing time, an unsmearing has to be applied which is described in [104].

4.1.8 Diffusion and Mobility

The diffusion radius is $r = \sqrt{Dt}$ where the diffusion coefficient D is calculated according to Einstein relation $D = \mu k_B T / q$. The mobility $\mu(E)$ of the subcharge is determined by looking up the given position in the electric field map and evaluating Eq. (4.4).

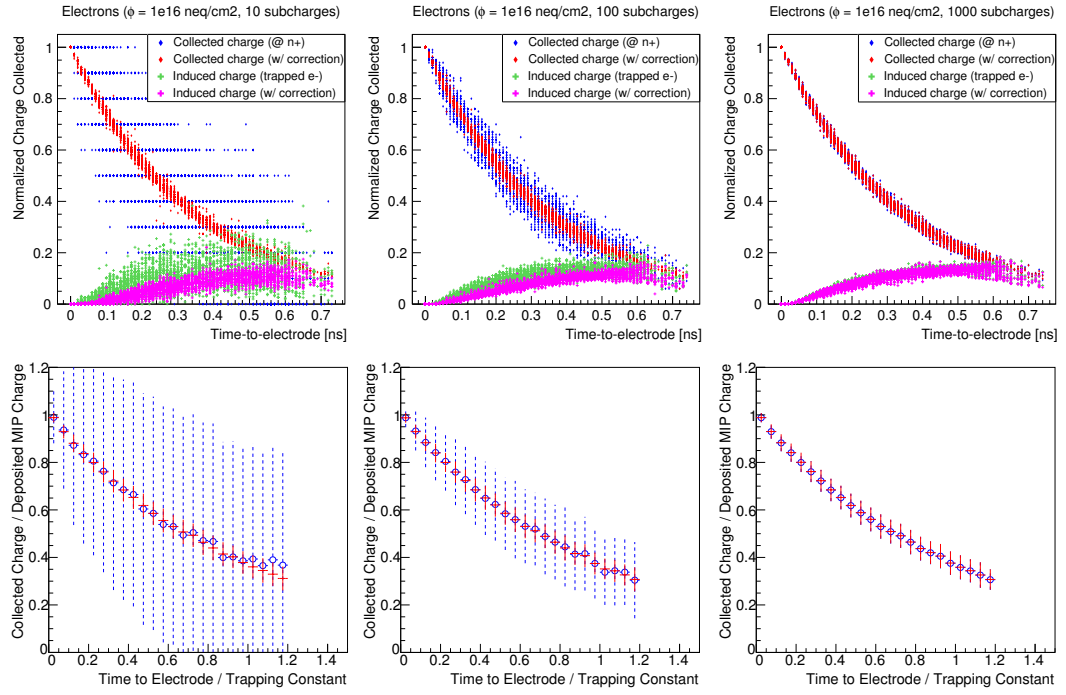


Figure 4.12: Top: Induced and collected charge change with the numbers of subcharges. Bottom: Average and error (blue) of collected charge. The average of the collected charge is the same for charge chunks and individual elementary charges (red). Shown are 10 (left), 100 (middle) and 1000 (right) subcharges [105].

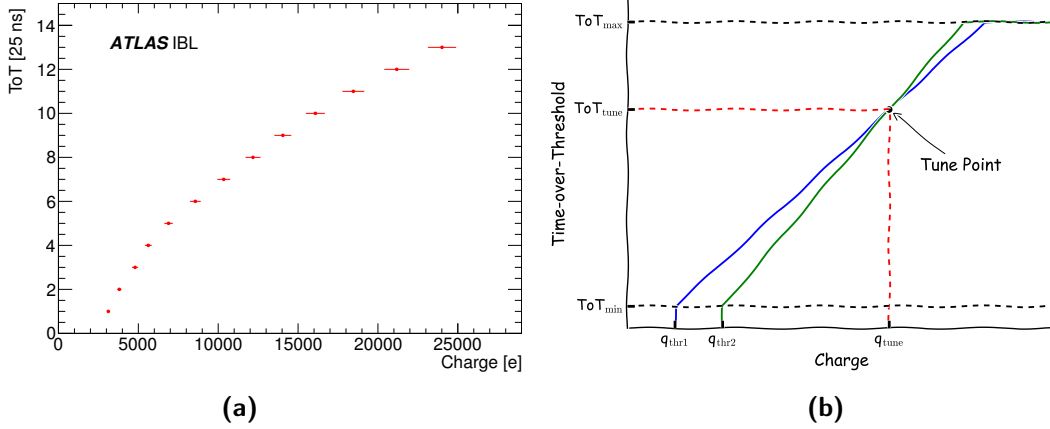


Figure 4.13: (a) ToT-charge calibration for an IBL sensor module with threshold at 3000 electrons and a tuning of 10 ToT at 16000 electrons [106]. (b) Linear ToT-charge conversion used in the digitiser. Compared to (a) the ToT is underestimated for charge below the tuning point and overestimated for charge above the tuning point.

4.1.9 Time over Threshold

The time-over-threshold (ToT) is the time the charge signal is higher than the threshold. The relation between ToT and the deposited charge is shown in Fig. 4.13 (a). This function is unique for each threshold and each tuning point. For simplicity a linear relation is assumed in the digitiser. As the ToT is 4-bit, any charge larger than the maximum possible ToT acquires the value of the maximum ToT which is therefore also the overflow bin. Charge below the threshold charge yields $\text{ToT} = 0$. The conversion follows Eq. (4.8) which is visualised in Fig. 4.13 (b). The result value is then rounded to an integer accordingly.

$$\text{ToT}(q) = \begin{cases} \frac{(\text{ToT}_{\text{tune}} - \text{ToT}_{\text{min}})}{q_{\text{tune}} - q_{\text{thr}}} \cdot (q - q_{\text{thr}}) + \text{ToT}_{\text{min}} & q \geq q_{\text{thr}} \\ 0 & \text{otherwise} \end{cases} \quad (4.8)$$

4.2 Test Beam at SLAC

For the validation of the radiation damage pixel digitiser a test beam campaign took place in May and October 2016 at Stanford Linear Accelerator Center (SLAC), using the EUDET [26] pixel telescope Caladium (Fig. 4.14 top) from Carleton University, read out by a high speed input/output (HSIO)-II and Reconfigurable Cluster Element (RCE) system [107]. The device under tests (DUTs) are all IBL-type (Section 2.2.1) n^+ -in- n planar sensors: an unirradiated double-module sensor 93-04-03 characterised in [108], LUB1 (301061-01-06) and LUB2 (30106102-03) of $250\text{ }\mu\text{m}$ thickness with slim edge design [109], both irradiated at Jožef Stefan Institute (JSI), Ljubljana, Slovenia, with reactor neutrons up to $4 \cdot 10^{15}\text{ n}_{\text{eq}}/\text{cm}^2$ according to [99]. SLAC provides a pulsed electron beam at 5 Hz with 11 GeV energy. Data were taken for all three DUTs at different bias voltages, temperatures and with angles of 0° and 73° (LUB1 only has data at 73°). While the unirradiated double module was kept at room temperature of 20°C , the irradiated sensors were cooled with dry ice at a temperature fluctuating between -50 and -25°C due to sublimation. Each run has about 10000 events. The positions of telescope planes [110] and the DUT is shown in Fig. 4.14 bottom and the same values are used for simulations.

Runs 176–266 are with the unirradiated double module as the DUT, biased at -60 V . In runs 176–193 the DUT is perpendicular to the beam. Most data are taken with a threshold of 3000 electrons and a tuning point of 16000 electrons at 10 ToT, two runs (190 and 191) with a threshold at 1000 electrons, and run 193 with a threshold of 2000 electrons and a tuning point of 11000 electrons at 8 ToT. The tuning is chosen to ensure that the most probable value (MPV) of the Landau distribution of the deposited charge is in the middle of the available ToT range.

Runs 294, 299 and 300, as well as 682–715 are recorded with LUB2, mostly with a bias voltage at 800 V , threshold at 1500 electrons and tuning of 8000 electrons at 10 ToT. The threshold and tuning of the irradiated sensor are reduced by a factor of two to account for loss in charge collection efficiency. As shown in [111], n^+ -in- n IBL planar sensors exposed to comparable level of radiation have an MPV of about 8000 electrons at 800 V bias voltage.

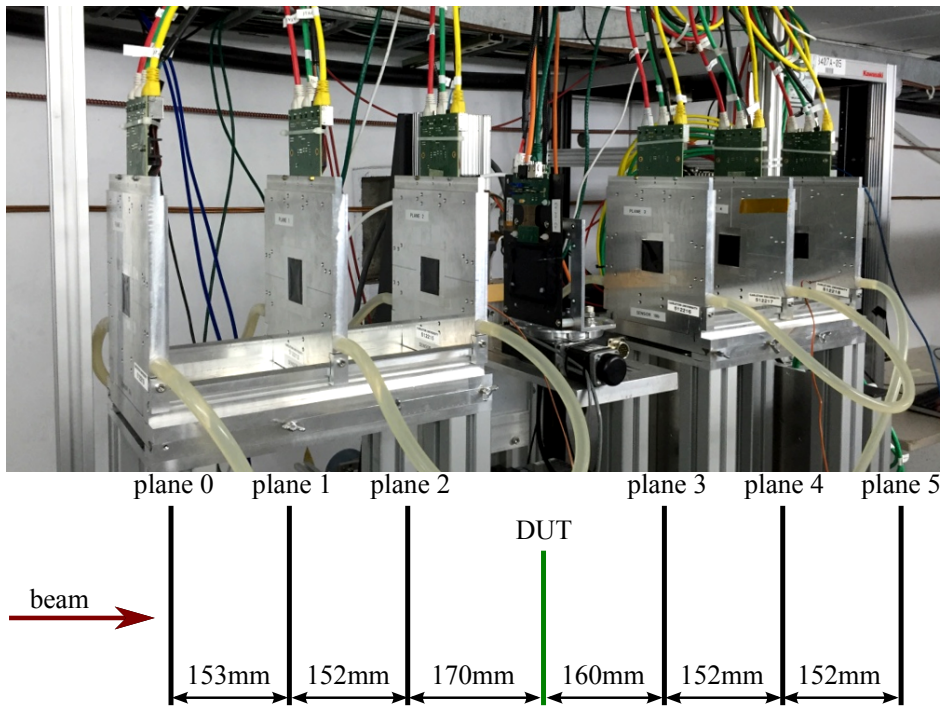


Figure 4.14: Top: Test beam at SLAC using EUDET telescope with DUT [112]. Bottom: Positions of telescope planes and DUT.

4.2.1 Track Reconstruction

Most data at 0° could be reconstructed using the EU Telescope framework. The data were taken in the `.raw` format. The first step in the reconstruction is to convert these data to `.slcio` format which is needed in the EU Telescope framework. In the same step noisy pixels are removed if they have a higher firing rate than 2%. Then the hits are grouped into clusters, which are used to find a correlation between the telescope hits and the DUT hits. A pre-alignment is done and the cluster positions in the local sensor frame are translated into the global telescope frame using the geometry of the setup. Finally, the pre-alignment is fitted using a deterministic annealing filter (DAF), taking multiple scattering into account, and another full alignment is done using only the telescope hits. This alignment is fitted again and the data is written out.

The total multiplicity, which is the total number of hits, per event is shown in Fig. 4.15. The top plot shows the event multiplicity of the first telescope plane and the bottom plot shows the DUT before removing noisy pixels. While the unirradiated sensor detects the same multiplicity as the telescope planes, the irradiated sensor has a much higher multiplicity. Comparing Fig. 4.16 and Fig. 4.17 it is obvious that the higher multiplicity comes from the higher sensor noise despite cooling. The absence of event multiplicity for run 294 is due to a very high particle rate of 350 particles per bunch which is off the scale. The dip of multiplicity around run 685 is due to a position scan where the beam spot was only partially on the sensor. The decline towards run 714 is a result of a decrease of the bias voltage from 1100 V to 1000 V, 800 V and 600 V for runs 711, 713 and 714, respectively. The multiplicity of run numbers greater than 700 is generally much larger than detected by the telescope planes and they also seem to be noisier compared with run 294–300 as shown in Fig. 4.17. The total cluster size (Fig. 4.18 top) mostly peaks at small cluster sizes of 1–2 pixels for the non-irradiated sample and for LUB2 at low event multiplicity, and at 1–4 pixels for a hit rate of around 40 hits per event registered by the telescope.

The timing distribution (Fig. 4.18 middle) shows the bunch crossing ID (BCID) in a multiple of 25 ns, at which a signal is registered within a given time window opened by the trigger. This timing with respect to the trigger

should be consistent for the same detector as is the case for the unirradiated double module. For LUB2 this distribution still shows a clear peak but there is also a constant noise background as noise is randomly distributed.

The ToT distribution peaks at 14 which is the largest ToT value even though there are four bits assigned to it. This is due to encodings (mappings) and configurations which use some of the values to store extra information as described in Appendix A.1. Other reasons for the peak at ToT 14 are the low threshold and the tuning. The low threshold of 1500 electrons results in a lot of noise. This is shown in Fig. 4.19 as a constant distribution of $\text{ToT} = 14$ across all timing distributions. While the collected charge is only about half as much in irradiated sensors compared to unirradiated ones, the tuning of the MPV of 8000 electrons [111] at a rather high ToT of 10 could lead to an overflow of the largest ToT.

Fig. 4.20 shows a side-by-side comparison of cluster size, timing and ToT between run 300 and 715. The data were taken using the same sensor but at different times.

To get rid of the noise shown in Fig. 4.17, different data selections, applied before passing the reconstructed data to pyEudetAnalysis are considered. Normally it would be logical to set a higher threshold of 3000–4000 electrons to avoid noise like in [111]. Since it has been set quite low, low ToT hits could be cut away to mimic a higher threshold. However, while this approach reduces the number of entries, it does not improve the shape of the row and column projections of the hitmap as shown in Fig. 4.21. Since the charge-ToT conversion depends on the threshold as shown in Fig. 4.13 cutting away low ToTs does not reflect the correct threshold because the tuning would have to be changed accordingly. A second approach is to make a selection based on timing and thus the timing value, since a real particle hit in the detector has a fixed delay in time with respect to the trigger while the noise is random in time. For runs 294–300 it can be clearly seen that the hit timing occupies two bins while for the later runs the timing is mainly in one single bin as can be seen in Fig. 4.20. In the latter case it has to be taken into account that the timing bin before and after the one with the maximum entries is likely to contain the fastest pixel within a cluster or a slow pixel, respectively, that is far away from the centre of a cluster where the incident particle and the shared charge would

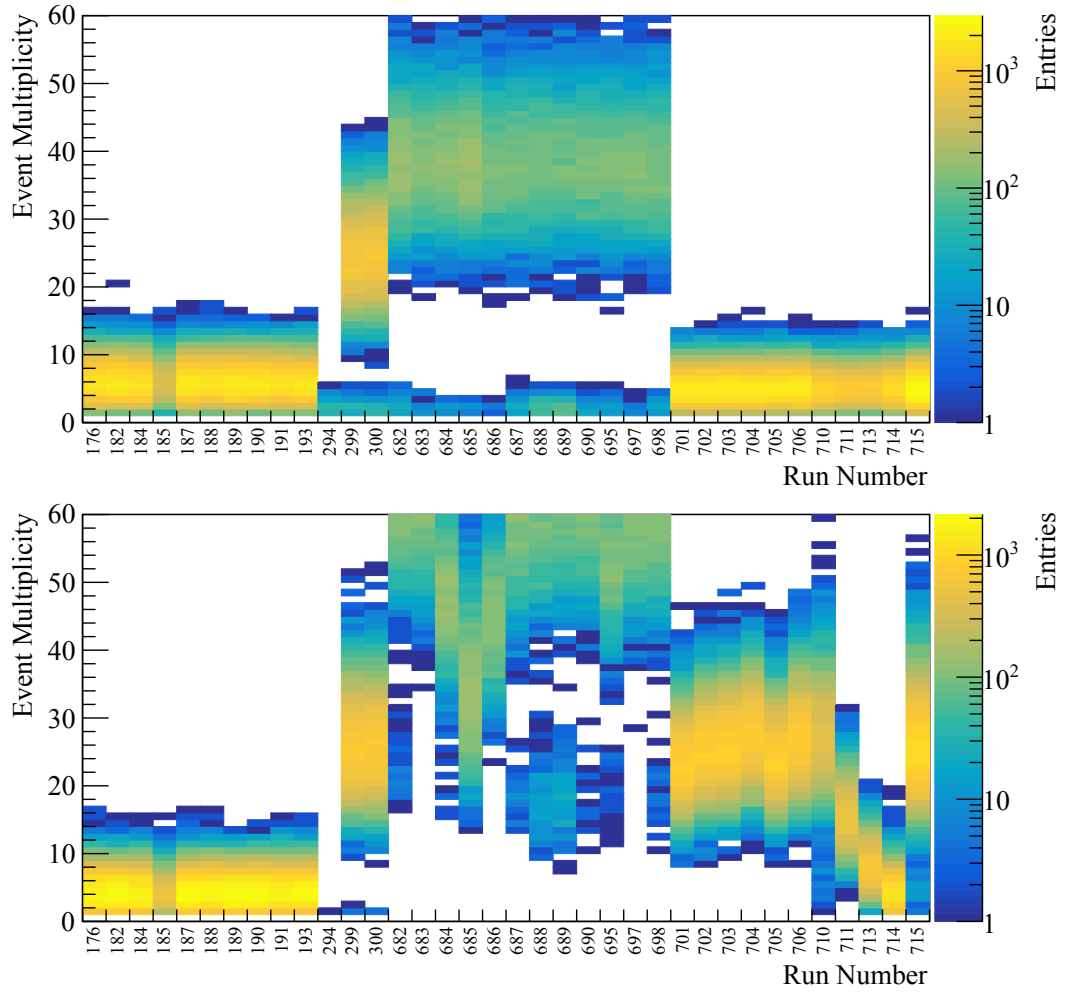


Figure 4.15: Total event multiplicity (number of hit clusters per event) for all reconstructed runs for the first telescope plane (top) and the DUT (bottom). For runs up to 266 the DUT is the unirradiated double module, the rest is LUB2.

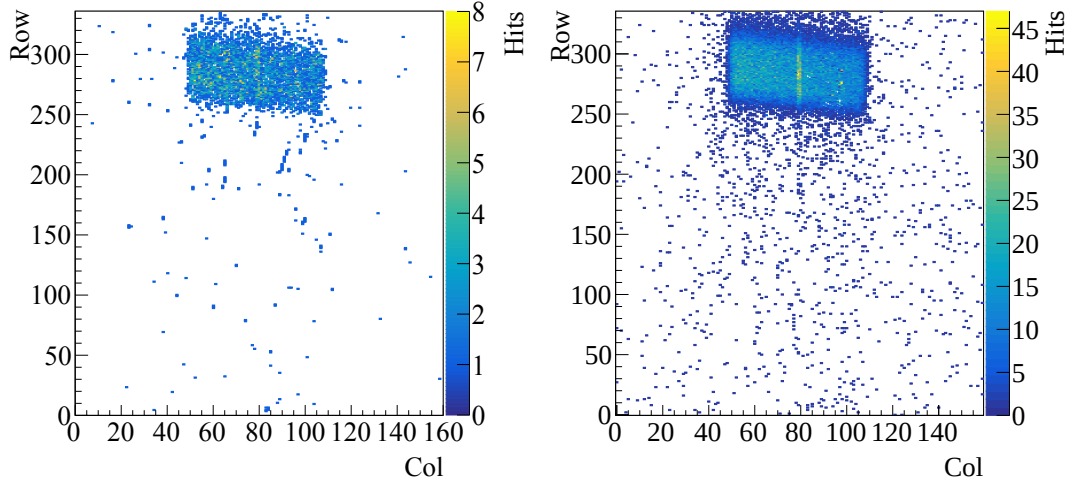


Figure 4.16: Hitmaps of the unirradiated double module from online monitoring (left) and reconstruction (right) of the same run. In the reconstructed hitmap it looks like the two halves of the double module overlap in the middle. This is due to ganged pixels in this region. The reconstruction was done using EUTelescope framework and its standard FE-I4 double module library. The monitoring hitmaps have too little data to give a significant comparison.

need more time to travel. Despite the fact that for most of the runs, the hit frequency of these fast and slow pixels is in the same order as the noise, as shown in the distributions, they are kept in the data, together with the timing bin with the maximum of the distribution. Thus hits in only three timing bins are kept while all other timing bins are cut away to remove the evenly distributed random noise and released trapped charge (Section 3.3.2). Compared to the ToT cut, this approach is more effective as can be seen in Fig. 4.21, which also shows the difference in keeping only the maximum timing bin. Finally, a cut on the region of interest (RoI) is examined in addition to the timing cut in order to keep only the hits in the obvious beam spot. For this, the content of each pixel and its eight neighbours is averaged over the run to ensure it is not an individual noisy pixel but the area is within the beam spot. If the content of one of the nine pixels is too high this would suggest a noisy pixel which is then skipped. If this average is above a certain threshold determined for the run then the examined pixel does not have random noise hits but probably lies within the beam spot. However, this method also removes hits that are within the beam spot, which causes difficulties in the analysis, therefore only

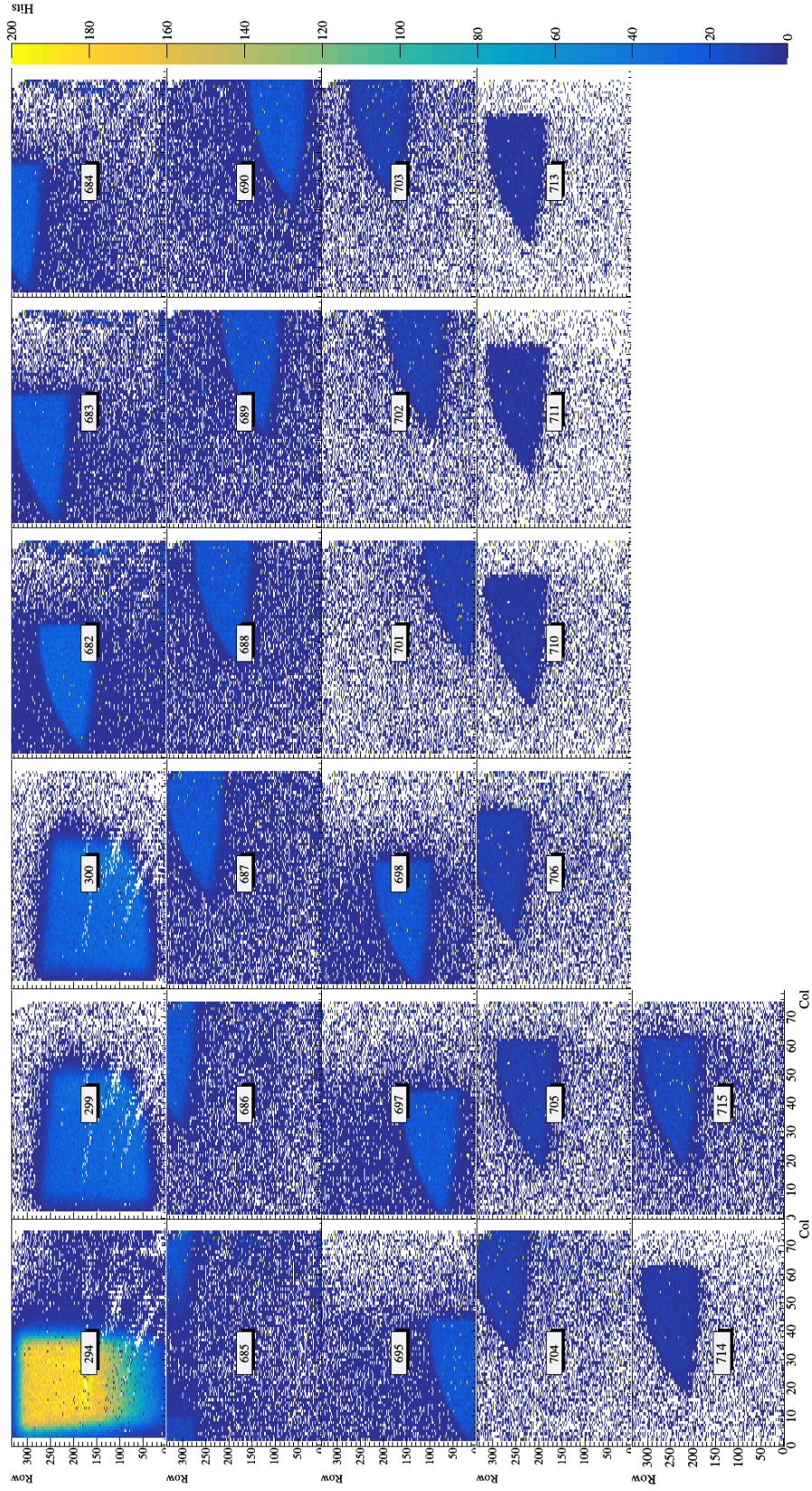


Figure 4.17: These are hitmaps from all reconstructible runs with LUB2 at 0° angle.

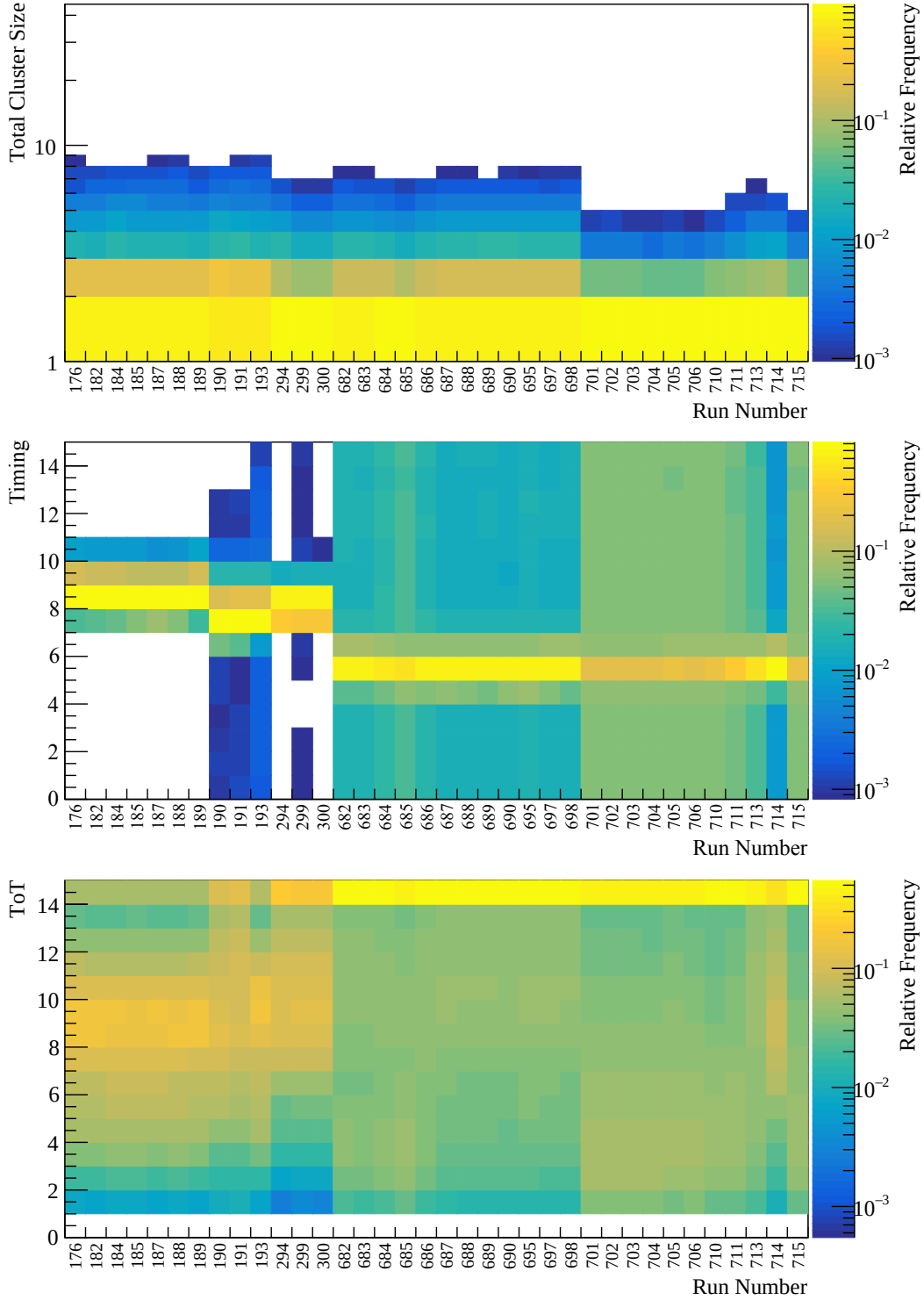


Figure 4.18: Overview of total cluster size (top), timing (middle) and ToT (bottom) distribution of hits for all reconstructed runs.

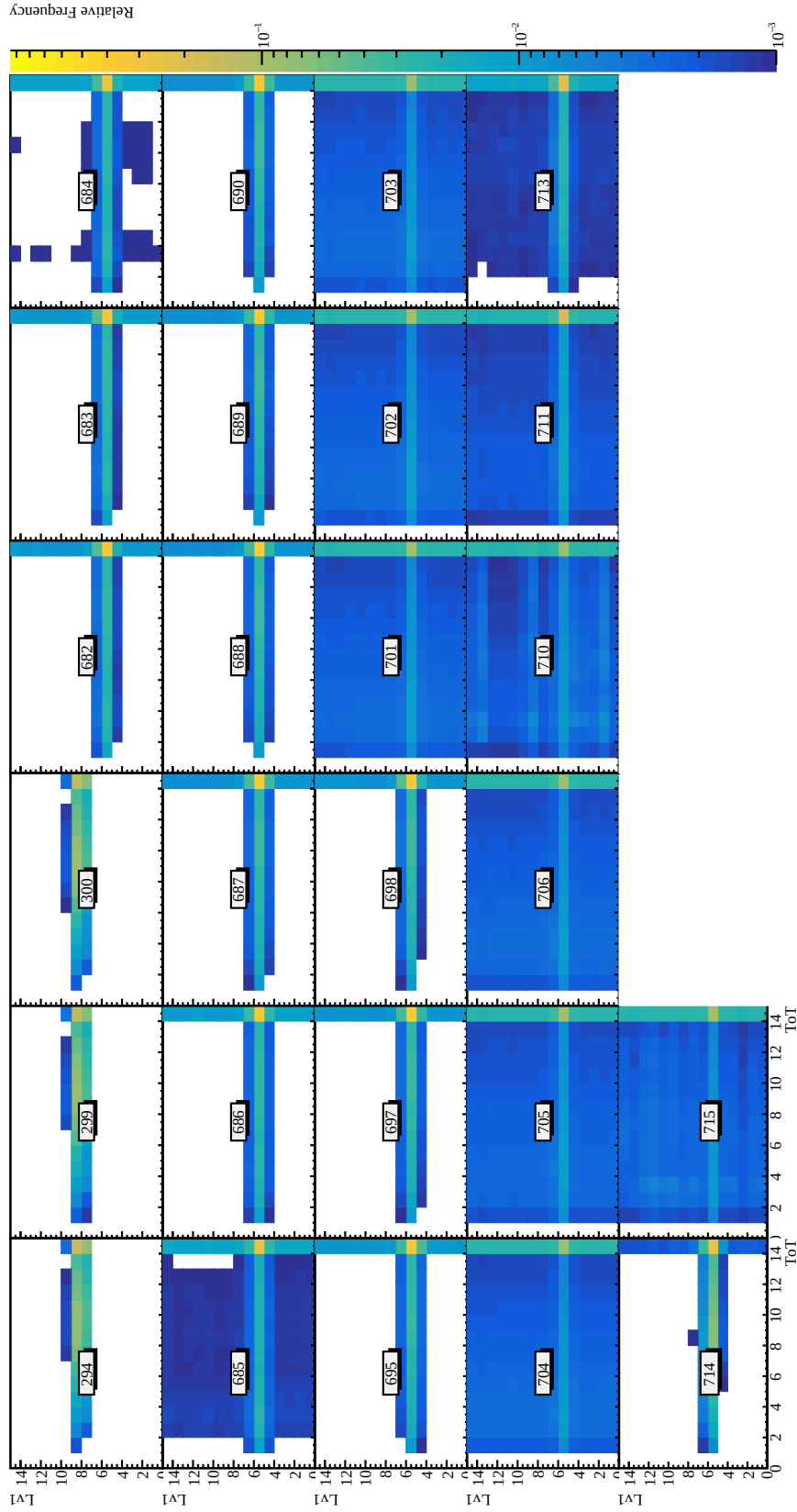


Figure 4.19: Overview of timing vs. ToT of all runs.

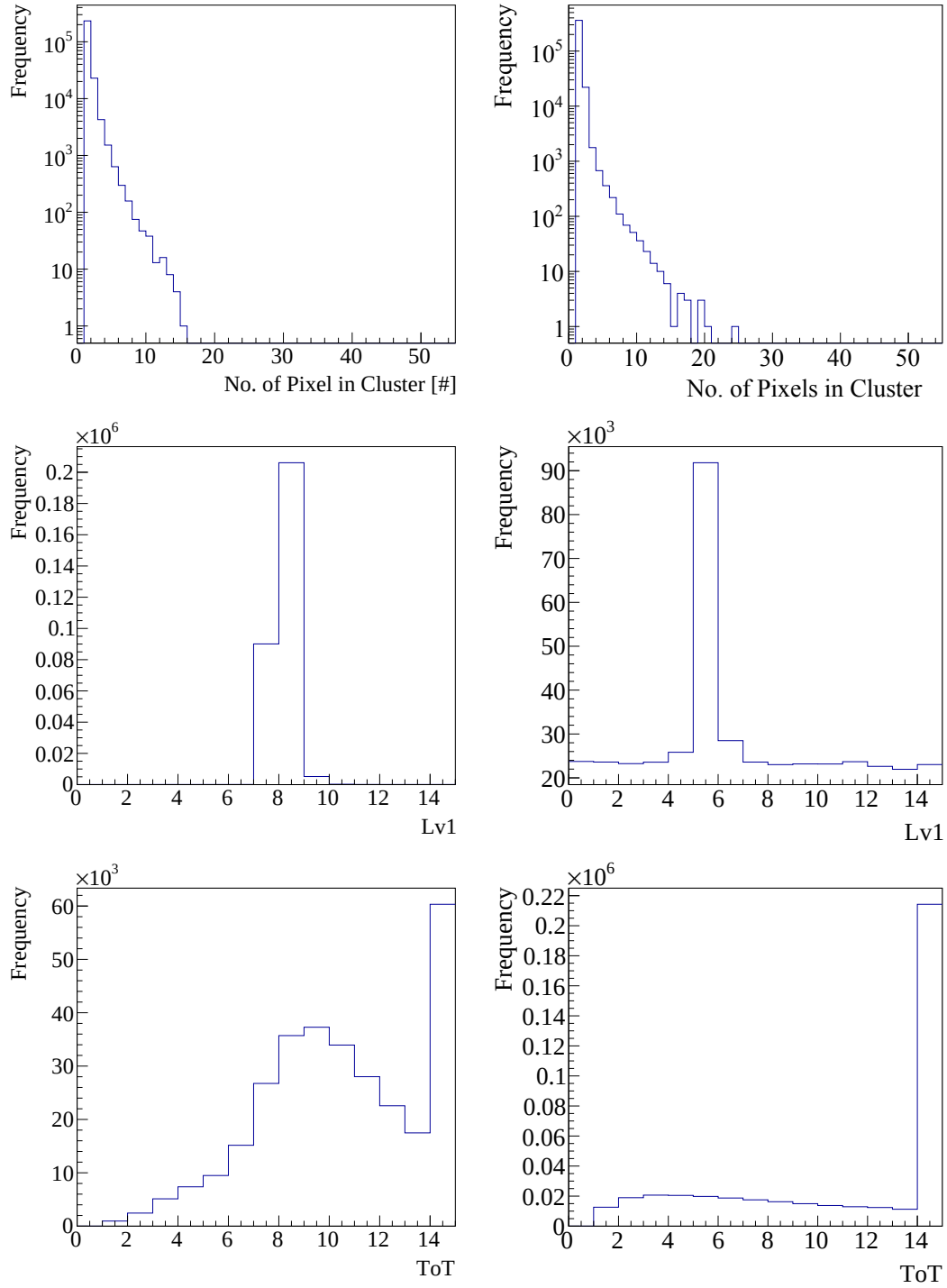


Figure 4.20: Cluster size, timing and ToT for run 300 (left) and run 715 (right). For unknown reasons earlier runs (e.g. 300) are less noisy than the later runs (e.g. 715) which is shown in the smaller cluster size and the noiseless timing distribution.

the three-bin timing cut is applied. In pyEudetAnalysis the cluster weighting method EtaCorrection is used since this gives the correct agreement of track and cluster position as shown in [113].

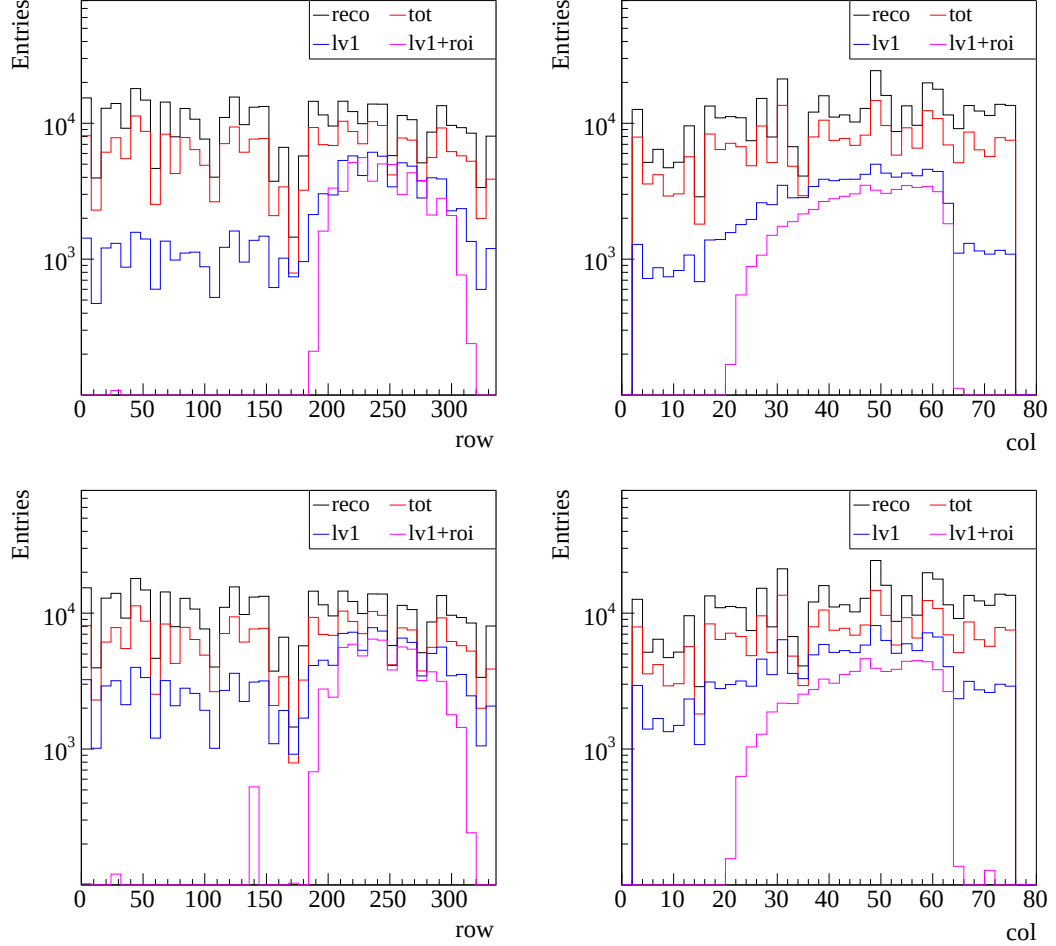


Figure 4.21: Hitmap projections in row and column of reconstructed runs with different cuts. The ToT cut here selects all hits with a ToT higher than 10. The timing cut in the top plots is on the maximum timing bin, while the timing cut in the bottom plots cuts on three timing bins, consisting of the maximum timing bin, the bin before and the bin after it.

4.2.2 Test Beam Simulation

The simulation is run using one of the most recent versions of AllPix [114]. The geometric setup of telescope and DUT planes and the beam type and

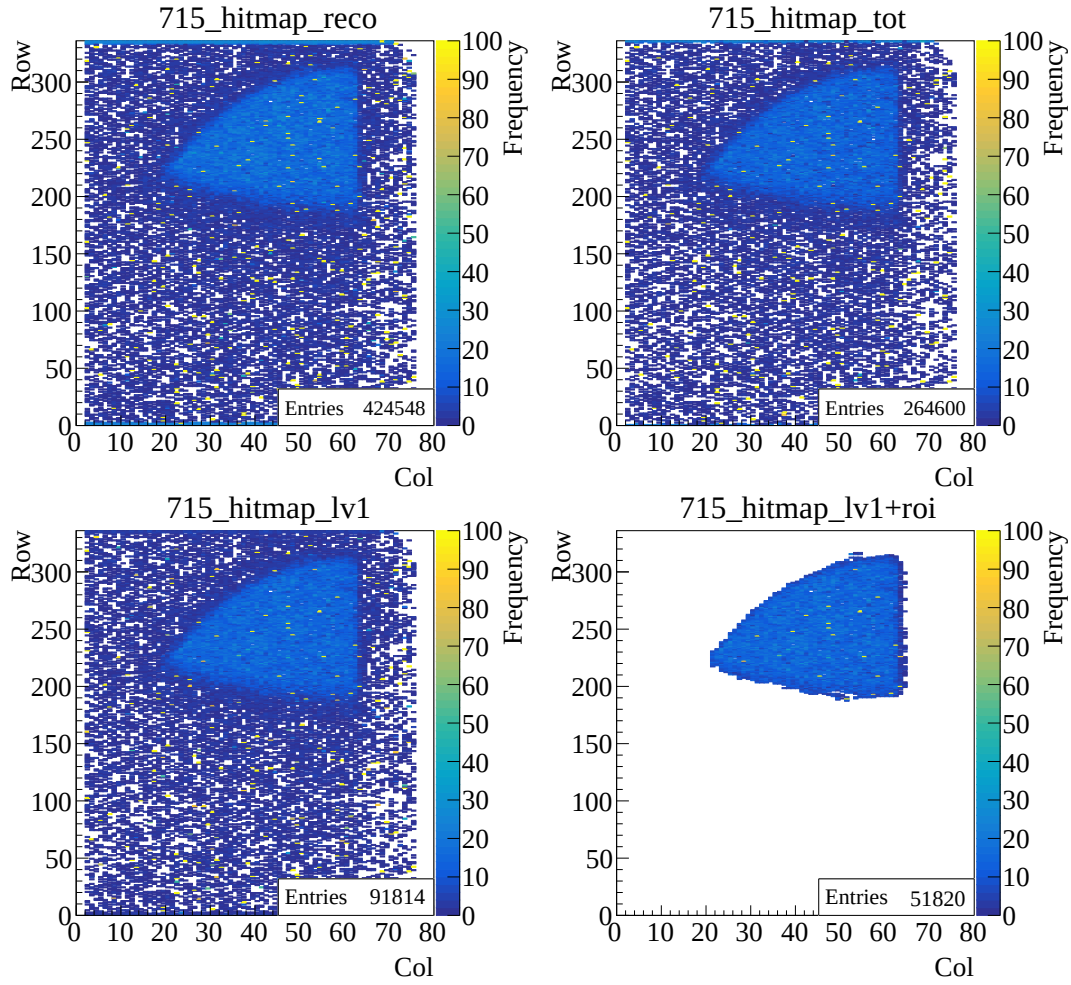


Figure 4.22: Hitmap after reconstruction and cuts.

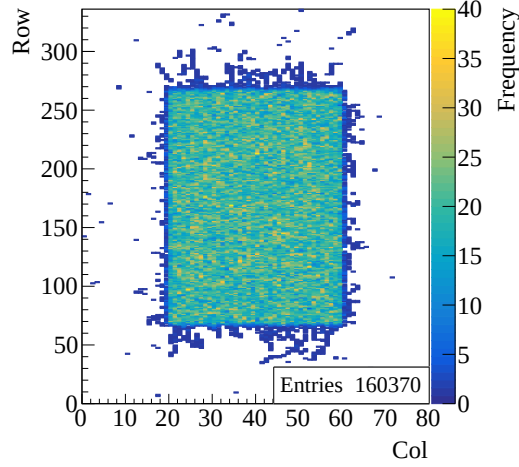


Figure 4.23: Hitmap from the simulation.

energy are configured to reproduce the test beam conditions. However, the beam multiplicity cannot be reproduced because of a lack of timing options in the simulation. Therefore, a single particle per event is used. The TCAD maps used are of $3.8 \cdot 10^{15} \text{ n}_{\text{eq}}/\text{cm}^2$ and $6 \cdot 10^{15} \text{ n}_{\text{eq}}/\text{cm}^2$ with 600 V, 800 V and 1000 V bias voltage. The threshold is 1500 electrons and a tuning of 8000 electrons at 10 ToT is used. For the unirradiated DUT a linear electric field map of 60 V is generated by the digitiser. The threshold is set to 1000, 2000 and 3000 electrons as in the test beam. The tuning is 16000 electrons at 10 ToT while for the 2000 electron threshold simulation a tuning of 11000 electrons at 8 ToT is used. The Ramo map is generated and the same for all simulations. In total, 100000 events per configuration are simulated with a beam spot spread over $10 \times 10 \text{ mm}^2$ as shown in Fig. 4.23.

4.2.3 Comparison between Data and Simulation

The intercept between the hit position on a plane and the position of the reconstructed track through that plane is called a residual. The residuals for run 715, with LUB2 biased at 1100 V, are shown in Fig. 4.24 top and for the simulation with $3.8 \cdot 10^{15} \text{ n}_{\text{eq}}/\text{cm}^2$ and 1000 V bias shown at the bottom. The fit is made with two error functions. In the ideal case this distribution is

rectangular, centred at 0 and have the width of the pixel. This tendency can be seen in the simulation.

The location of the reconstructed track through a pixel on the DUT plane is shown in Fig. 4.25 for different cluster sizes. It is a superposition of all reconstructed runs because each run has a low number of events which would not suffice to show a significant trend. And since the sensor was moved from run to run, it is not possible to merge the datasets before reconstruction. In general the distribution is somewhat as expected: the tracks go mostly through the centre of a pixel in one-pixel clusters, but through the edges in two pixel clusters and through the corners in clusters made of more than two pixel hits. It can be seen that for one-pixel clusters the track position distribution also extends to the upper and lower edge which are missing in the distribution of two-pixel clusters. A comparison can be made with Fig. 4.26 which shows the track position distribution of the simulation. The discrepancy is probably due to the wider residual distributions. The residual distributions of the unirradiated sensor are narrower. However, the statistics are too low to show any trend in track positions within a pixel.

For the following comparisons only irradiated data from runs 294 and 300 are considered because of the higher noise in the later data, which are taken 3 months later. In addition to the noise, Fig. A.7 shows that the discrepancy in ToT and cluster size distributions between them is fairly large.

Fig. 4.27 and Fig. 4.28 show the distribution of the cluster size (CS), ToT for all cluster sizes and for individual cluster sizes. All plots are normalised by the total number of events in each data set. Fig. 4.27 compares run 300 (LUB2, irradiated up to $4 \cdot 10^{15} \text{ n}_{\text{eq}}/\text{cm}^2$, biased at 800 V, with a threshold at 1500 electrons and a tuning of 8000 electrons at 10 ToT) with simulations of $3.8 \cdot 10^{15}$ and $6 \cdot 10^{15} \text{ n}_{\text{eq}}/\text{cm}^2$ fluences at the same bias voltage and with the same threshold and tuning. In addition it is also compared with the simulation of an unirradiated sensor biased at 150 V with a threshold at 3000 electrons and tuning of 16000 electrons at 10 ToT. As trapping would occur in radiation-damaged silicon sensor, it is expected that the cluster size and the ToT become smaller with increasing radiation. The data has some fluctuations depending on the multiplicity of the beam and the temperature. In general, larger particle multiplicity or higher temperature cause slightly larger cluster size and larger

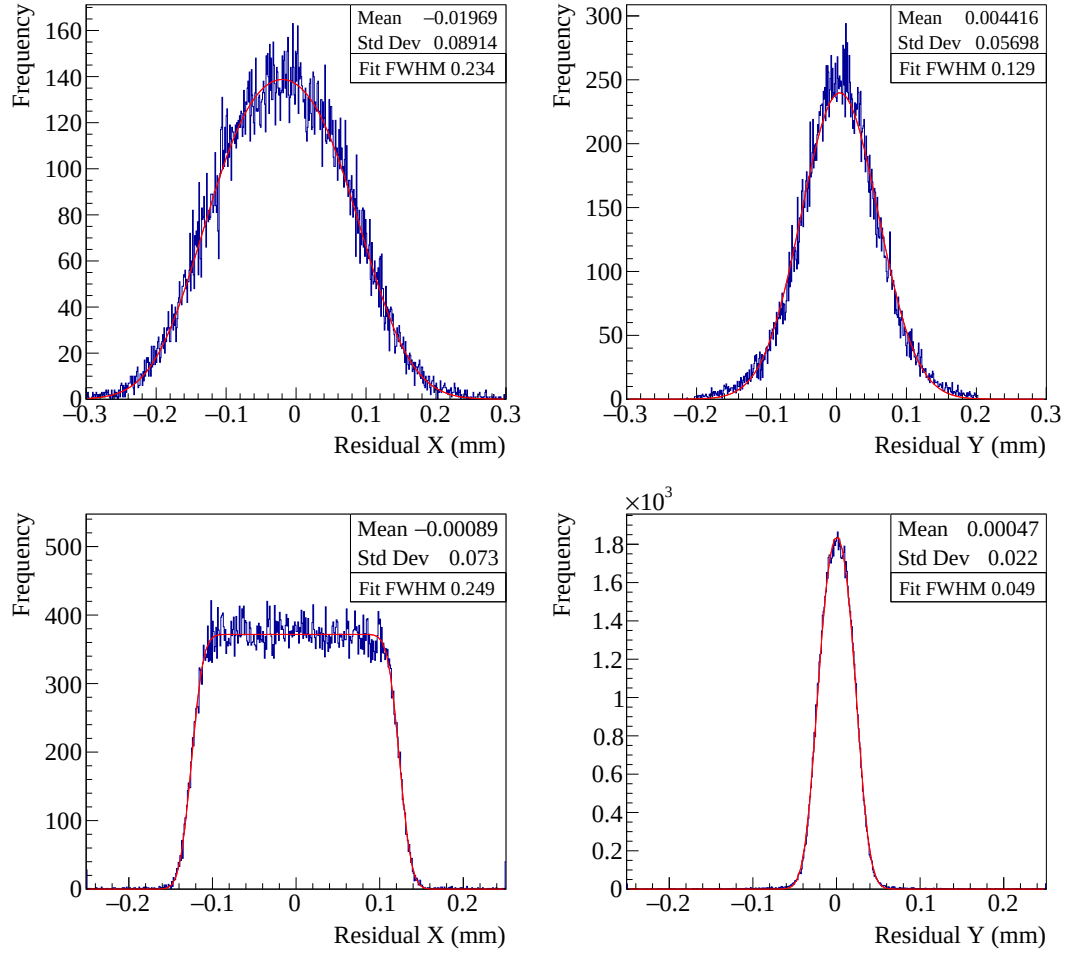


Figure 4.24: Residual distributions of run 715 (top) and of the simulation (bottom) are shown in blue. The fit using two error functions is shown in red. The FWHM is obtained from the fit.

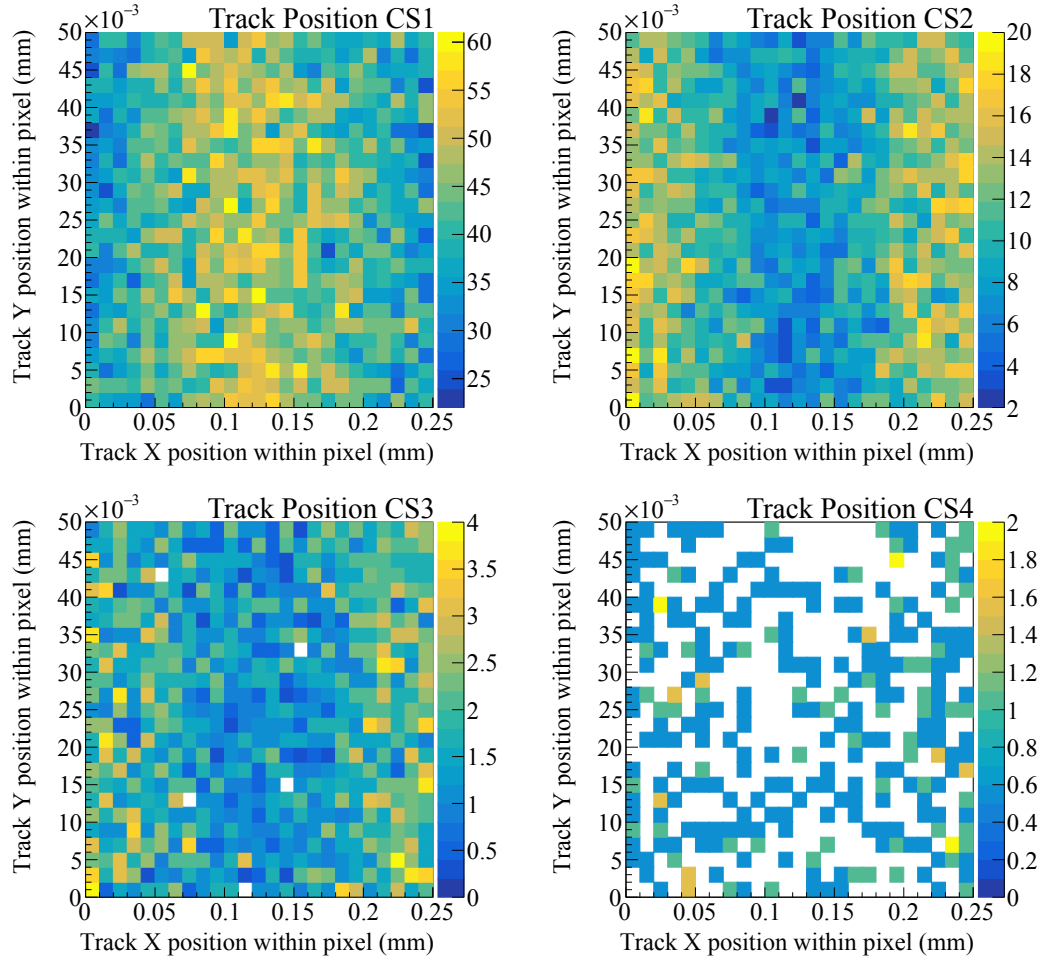


Figure 4.25: Distribution of track positions within a pixel for different cluster sizes (CS) from test beam data, sum of all reconstructed runs.

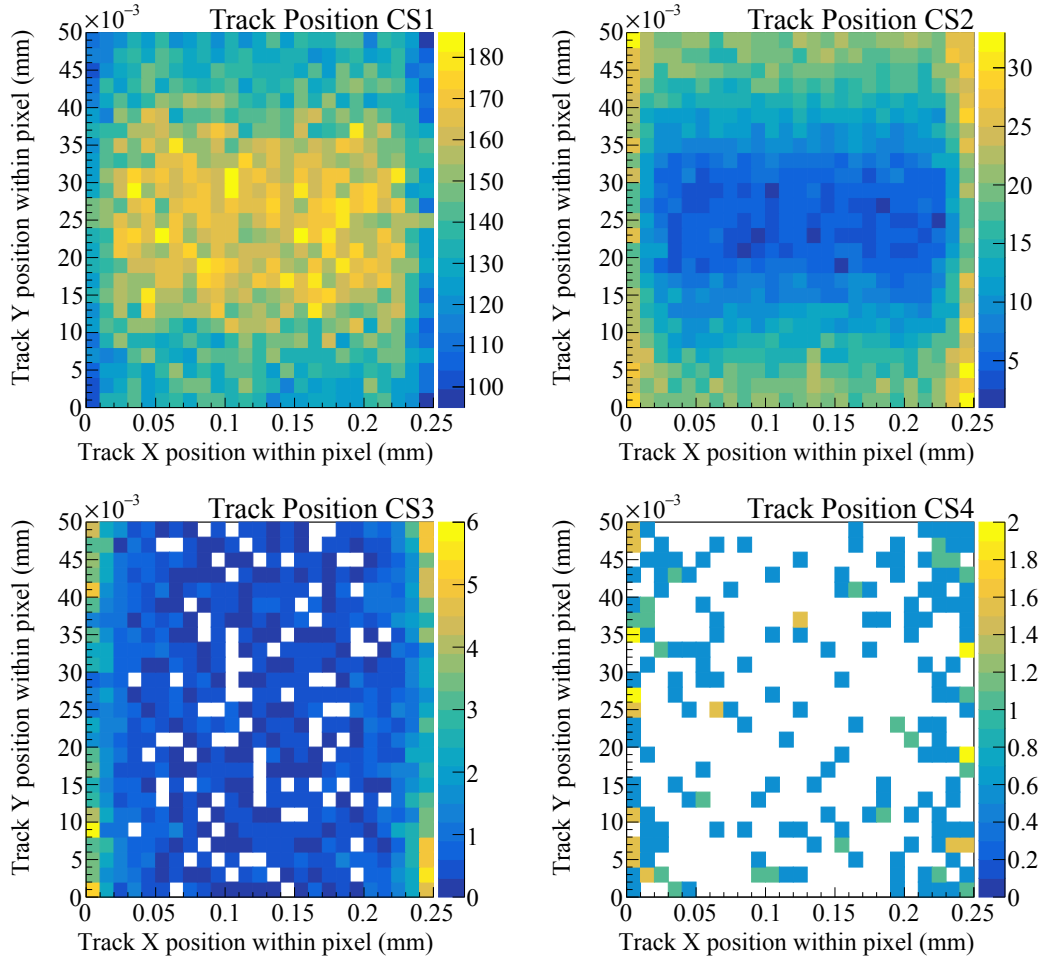


Figure 4.26: Distribution of track positions within a pixel for different cluster sizes (CS) from simulation.

ToT. Also some fluctuation is observed between the different positions of the beam spot and between earlier runs (294–300) and later runs (682–715) (see Appendix A.2). The cluster size distribution of the data is very close to the simulated distribution. For CS1 it agrees well with the $3.8 \cdot 10^{15} \text{ n}_{\text{eq}}/\text{cm}^2$ simulation, but is larger for CS2. The measured ToT distribution lies between the unirradiated and the irradiated simulation with the tendency towards the unirradiated simulation with larger ToT for CS1. This raises the question if and how much the sensor has annealed during years of storage. Although it is unlikely because it was kept in a freezer, this can be found out by measuring the leakage current of the sensor.

For the comparison of the unirradiated case, simulations are run with different thresholds and tunings as used in the data. This is shown in Fig. 4.28. The discrepancy is larger than expected: the simulation in general has larger cluster sizes and ToTs than the data. Similar discrepancy has shown in the irradiated case for the ToT distribution. The simulation might be imprecise with regard to: the generated linear electric field used for a bias voltage of 60 V; the depletion voltage that is likely different in the real sensor; and the oversimplified way of the charge-ToT conversion. All these factors might contribute to the rather large difference at such a low bias voltage. The generated linear electric field at about 150 V, matches the simulation using TCAD input map very well (Fig. A.11). The effect of the different depletion voltage can be seen in Fig. A.13. The trend of how threshold and tuning affects the ToT distribution of e.g. cluster size 1 is comparable in data and simulation. However, one can see that the charge-ToT conversion in the simulation underestimates the ToT for lower charge and overestimates ToT for higher charges as stated in Section 4.1.9.

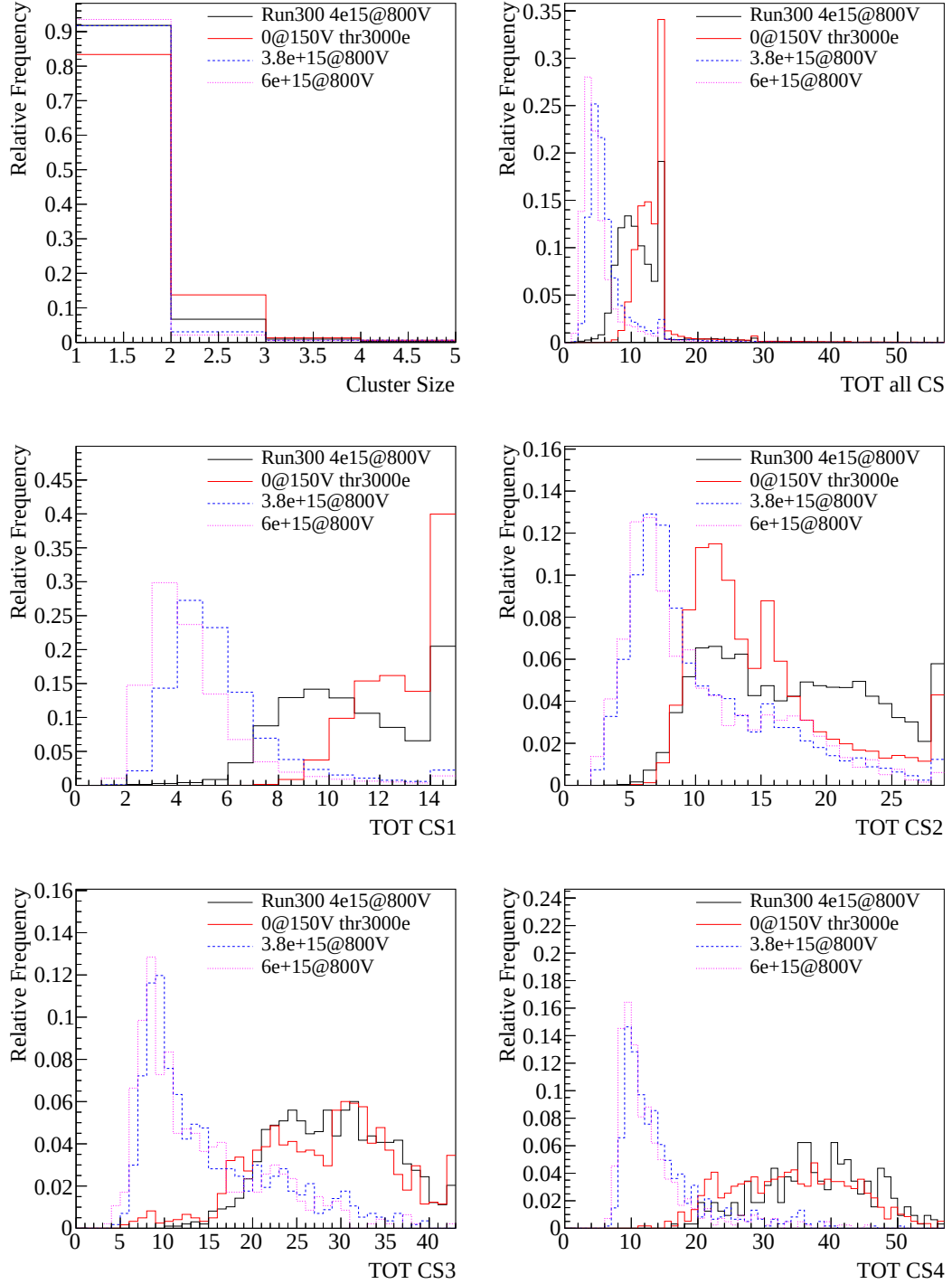


Figure 4.27: Comparison of irradiated test beam data with simulation. The data set is chosen to have a low noise and compared with several simulations of 0, $3.8 \cdot 10^{15}$ and $6 \cdot 10^{15}$ n_{eq}/cm^2 fluences. Shown are cluster size (CS) distribution and ToT distributions.

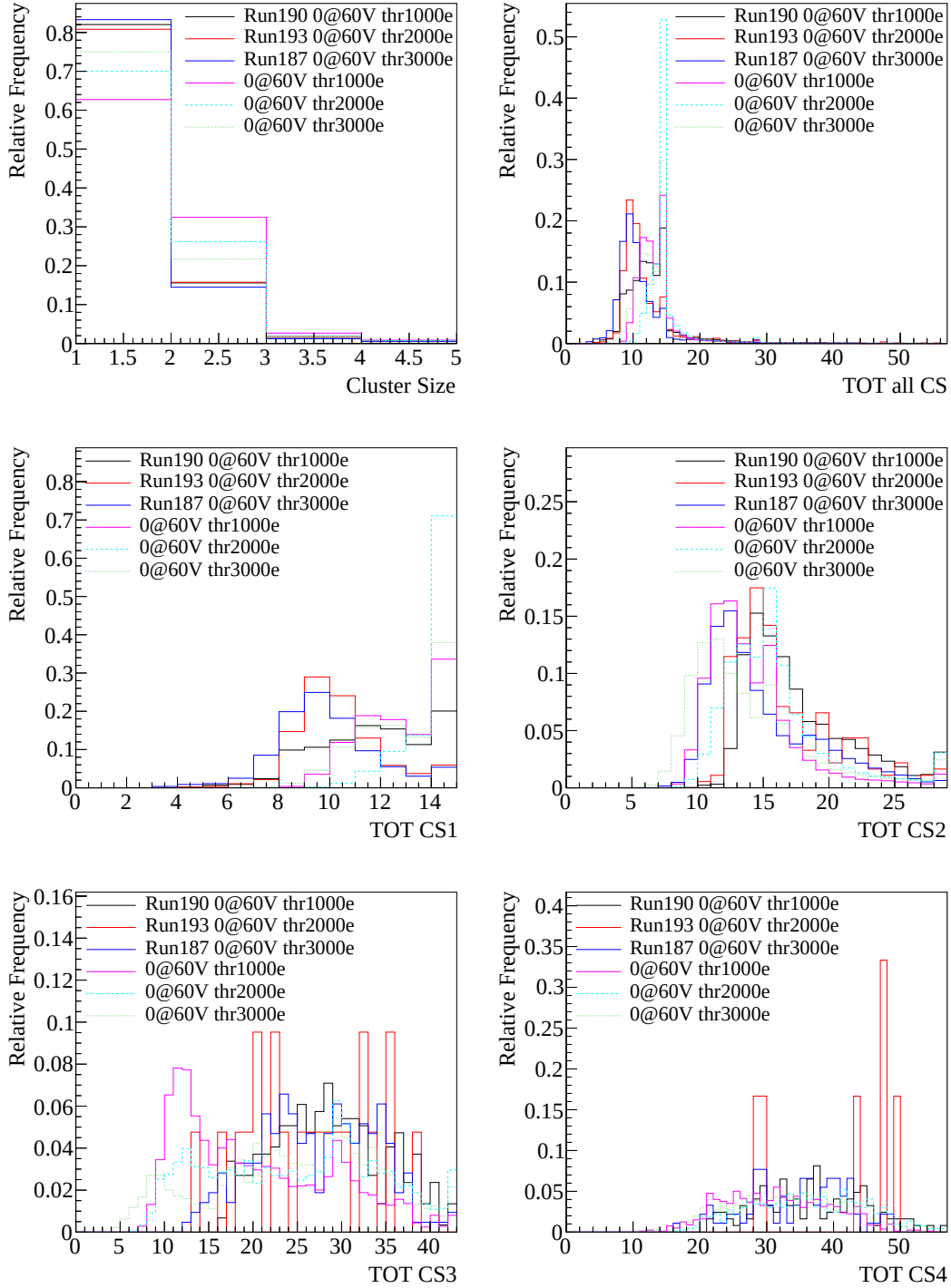


Figure 4.28: Comparison of unirradiated data at different thresholds and tunings with corresponding simulations. The tuning for threshold of 1000 and 3000 electrons is 10 ToT at 16000 electrons and for 2000 electrons threshold is 8 ToT at 11000 electrons.

Chapter 5

High Voltage CMOS Sensors

Complementary metal-oxide-semiconductor (CMOS) pixel sensors are fabricated using industrial CMOS processes on inexpensive Czochralski (CZ) silicon substrates. Compared to the high purity float zone (FZ) silicon substrate of conventional planar pixel sensors, CMOS has usually a low resistivity of $\sim 10\text{--}1000\,\Omega\text{cm}$, some foundries offer $2\text{--}3\,\text{k}\Omega\text{cm}$ too. The charge collection of a CMOS sensor is based on diffusion, like in the TowerJazz 180 nm sensor which was selected for the ALICE tracker [115], or the MIMOSA-26 sensor [116] used in EUDET telescopes.

HV-CMOS sensors are typically reverse biased at a voltage larger than 60 V. In a p-substrate, a deep n-well is not only the collection electrode but also shields the electronics from the applied high voltage (HV). The embedded CMOS electronics, consisting of an NMOS and a PMOS, are placed inside multiple nested wells as can be seen in Fig. 5.1 left. The shallow depletion depth of some tens of μm enables a fast drift-based charge collection of $\sim 200\,\text{ps}$ [117]. This shielding is needed because the operation voltage of a CMOS is only a few volt. Due to the on-sensor signal amplification, it is possible to capacitively couple the sensor onto a readout chip by glueing (Fig. 5.1 right) which replaces the complicated and expensive bump bonding as shown in Fig. 5.2 and described in section 2.2.1. They can also be thinned down to $50\,\mu\text{m}$ because large depletion depths are not necessary. Apart from the capacitively coupled pixel devices (CCPDs) there are also monolithic active pixel sensors (MAPSs) which supersede the need for the separate readout chip by embedding all the readout circuits onto the same substrate as the sensor.

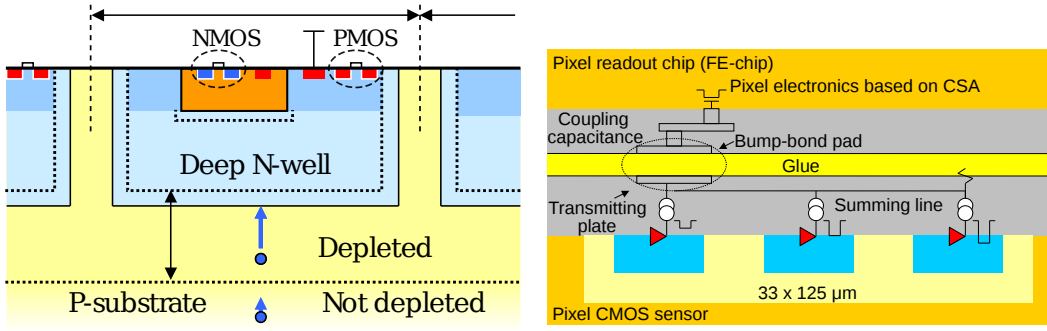


Figure 5.1: Left: schematic of a HV-CMOS pixel [120]. A CMOS consists of an NMOS and a PMOS. The NMOS is embedded in a shallow p-well inside the deep n-well while the PMOS is located inside the deep n-well in the p-substrate. Right: glue bonding of FE chip and a CMOS pixel sensor [121].

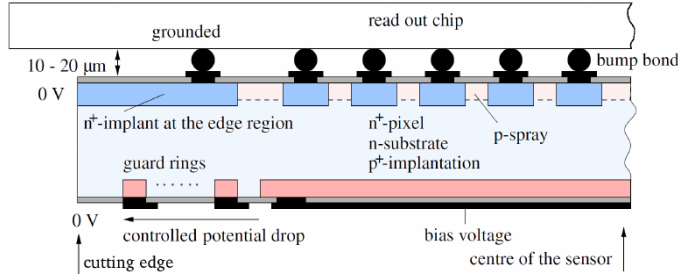


Figure 5.2: Cross-section of an ATLAS pixel module [109].

The commercial availability of HV-CMOS processes, the usage of CZ wafer and the simpler glue-bonding make HV-CMOS sensors more cost-efficient than the conventional hybrid pixel modules. The radiation tolerance has been tested to perform comparably with the unirradiated sensor up to a fluence of $7 \cdot 10^{15} n_{eq}/cm^2$ [118] which would fulfill the radiation tolerance requirements of the outer pixel layers. Together with the lower voltage requirement compared to the traditional planar pixel sensors, they make excellent candidates to be used for large areas as needed for the ITk. They are also the sensor choice for the tracker of the Mu3e experiment [119].

5.1 The ams 350 nm Demonstrator

The HV-CMOS demonstrator in ams 350 nm technology (H35DEMO), designed by Institut de Fisica d'Altes Energies (IFAE), Karlsruher Institut für Technologie (KIT) and University of Liverpool (UoL), is a large area demonstrator sensor fabricated in an engineering run at the end of 2015, using p-substrates of four resistivities: 20, 80, 200 and 1000 Ωcm . The 350 nm technology refers to the width of the gate explained in Section 3.2.7. The die size of the sensor is $18650\text{ }\mu\text{m} \times 25520\text{ }\mu\text{m}$ including the scribe lines outside which are some test structures by the foundry [122]. The actual sensor dimensions without the scribe lines are indicated in Fig. 5.3. The sensor is divided into four matrices: standalone NMOS, analog 1 and 2, and standalone CMOS matrix. The standalone matrices have NMOS or CMOS as amplifiers, respectively. The pixels in the NMOS matrix have two different comparators: without and with time-walk compensation. Time-walk of a signal refers to difference in time a signal crosses the threshold depending on its amplitude. The analog matrices have three different kinds of pixels each, and the CMOS matrix has only one pixel flavour. The top part of Fig. 5.4 shows the cross-section of most of the pixels, which is referred to as the alternative cross-section, whereas the first and last 100 columns of the analog matrix 1, the first 100 columns of the analog matrix 2, and the bottom test structure have the bottom cross-section with extra DNTUBs [123], referred to as the default cross-section. This is because most simulations are done on the default cross-section. DNTUB stands for deep n-wells whereas e.g. SPTUB stands for shallow p-well. The doping concentration increases from depth to top with the highest concentrations being the n^+ and p^+ implants. The pixel size is $50\text{ }\mu\text{m} \times 250\text{ }\mu\text{m}$, matching the pixel size of the FE-I4 readout chip for the analog matrices.

5.2 TCAD Simulation

TCAD is a process and device simulation framework for semiconductor devices, usually transistors. It is capable of 1D, 2D and 3D simulations, using Graphical Design Station II (.gds) files as input masks and finite element method (FEM) to solve physics equations. For the content of this thesis

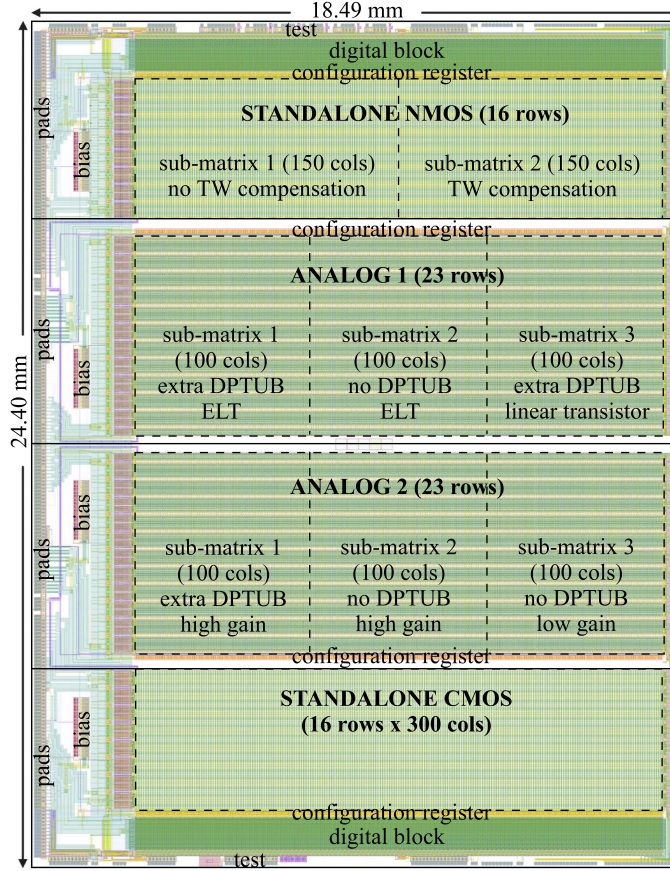


Figure 5.3: Sensor layout of H35DEMO.

Synopsis[®] Sentaurus[™] TCAD is used. Fig. 5.5 shows the schematic of the simulation framework. The core comprises the structure simulation for which different tools can be used, and device simulation. The project tree can contain parameters with different values in each branch. Each new parameter value results in a new node and thus a new simulation.

The device geometry can be created using the script-based process simulation or the graphical user interface (GUI)-based flow simulation. Both simulate the fabrication process which can vary from being as detailed as reproducing every fabrication step like gas mixture, temperature rampings ion implantation and photolithography, or can be simplified to just defining the doping profile. The time needed to run the process simulation strongly depends on its complexity. Even though multithreading is supported, not all steps are suitable for multithreading, e.g. temperature ramps which occur frequently. The device

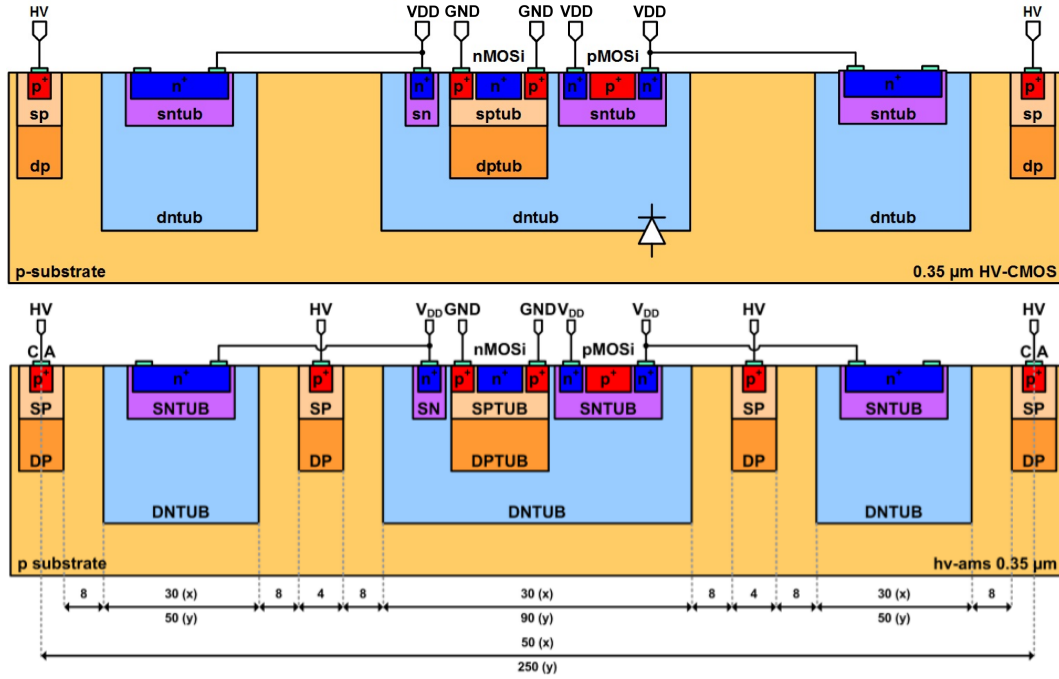


Figure 5.4: H35DEMO pixel cross-sections [124]. The upper and lower case names refer to the same thing. VDD is the electrical connection of the n-implants and is at 1.8 V. GND is the electrical connection of the p-implants which is held at ground. The p-n junction is at the interface between the DNTUB and the p-substrate. The bottom cross-section is used for only a small part of the pixels as described in the text. However, as all TCAD simulations were based on it, it is referred to as the default cross-section. The top cross-section without the extra DPTUBs is therefore referred to as the alternative cross-section.

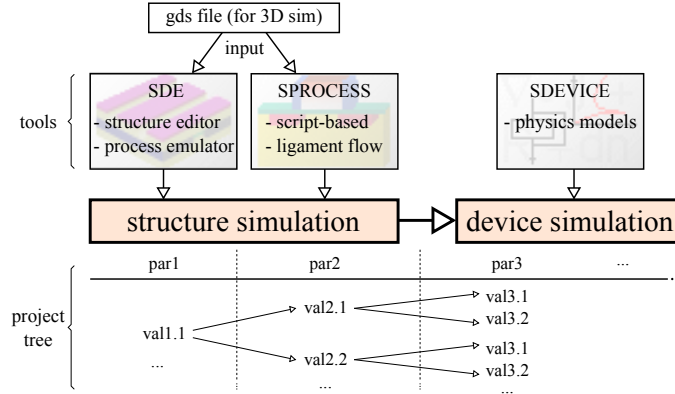


Figure 5.5: Schematic of the Sentaurus™ TCAD framework.

geometry can also be achieved using structure editor or its process emulator. The structure editor offers a sparse GUI where the user can define all features in a coordinate system and apply the desired material or doping profile like in a CAD programme. The process emulator is similar to the process simulation. The simulation time using structure editor is drastically reduced compared to process simulations. Usually, the fabrication process is kept secret by the foundry. So, for the sake of a short simulation time, structure editor, the first tool in Fig. 5.5, is used in this thesis.

The finite element method (FEM) works by dividing a domain into many subdomains where the problems or differential equations of the original domain can be reduced to local approximations and are therefore easier to solve. A good mesh is essential to optimise a simulation. Sentaurus™ Mesh is integrated in the structure tools and offers many possibilities to work with meshes, e.g. transformations, intelligent creation using adaptive meshing or local refinements. For example, in regions with rapidly changing doping concentration it can automatically create a finer mesh, whereas in regions with constant doping the segments can be larger so the overall number of elements is reduced and thus the simulation time. To generate the mesh, the Delaunay triangulation is used [125]. This method creates triangular elements in a way that all three points of a triangle lie on a circle so that no point of any other element lies within this circumcircle. This ensures a high quality of triangular mesh elements due to their maximised internal angles. Fig. 5.6 shows a mesh at a surface area with small segments at doping type boundaries.

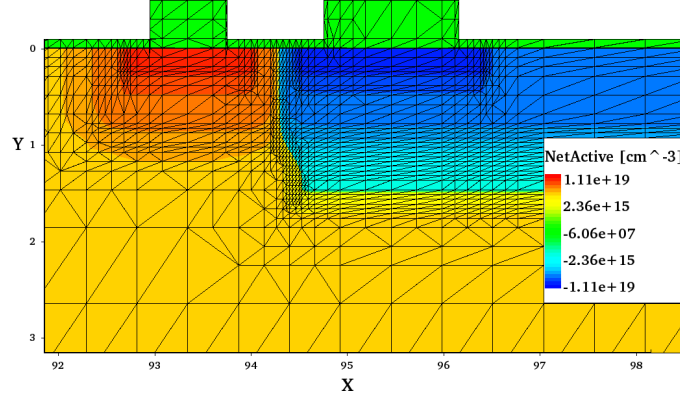


Figure 5.6: An example of a TCAD mesh. Shown is a region of a 2D simulation. NetActive denotes the active doping whereas blue is p-doping and red is n-doping. The higher green regions are aluminum contacts while the flatter green regions is the oxide layer.

The device simulation is executed using sdevice as shown in Fig. 5.5. Here, the Poisson equation

$$\nabla \cdot \varepsilon \nabla \phi = -q(p - n + N_D - N_A) - \rho_{\text{trap}} \quad , \quad (5.1)$$

where

ε	is the electrical permittivity,
ϕ	the electrostatic potential,
q	the elementary charge,
n and p	the electron and hole density, respectively,
N_D and N_A	the ionised donor and acceptor concentrations, respectively,
ρ_{trap}	the density contributed by traps and fixed charges,

is solved for the electrostatic potential ϕ using the densities given by physics models. SentaurusTM TCAD provides many different physics models that are all described in the manual [126]. Those used in this work are show in Table B.1. The doping dependence of carrier mobility and recombination always applies whereas other models are sometimes omitted when the simulation is more complicated or time-consuming, e.g. if radiation damage, MIP passage or capacitance are simulated. This will be explicitly stated and a comparison with the default physics set can be found in Appendix B. All simulations shown were run at the default temperature of 300 K. Due to the nature of vol-

ume densities which are impossible to implement for 2D simulations, an area factor is defined which represents the width of the 2D slice. In this work the area factor is always defined to be 1 μm .

The default carrier transport model is drift-diffusion which is described by the continuity equations

$$\begin{aligned}\nabla \cdot \mathbf{J}_n &= q\left(R_{\text{net}} + \frac{\partial n}{\partial t}\right) \quad \text{and} \\ \nabla \cdot \mathbf{J}_p &= -q\left(R_{\text{net}} + \frac{\partial p}{\partial t}\right) \quad ,\end{aligned}\tag{5.2}$$

where R_{net} is the net recombination rate and $\mathbf{J}_{n,p}$ the electron and hole current density. The current densities are

$$\begin{aligned}\mathbf{J}_n &= -nq\mu_n \nabla \Phi_n \quad \text{and} \\ \mathbf{J}_p &= -pq\mu_p \nabla \Phi_p \quad ,\end{aligned}\tag{5.3}$$

with $\mu_{n,p}$ the electron and hole mobility and $\Phi_{n,p}$ the electron and hole quasi-Fermi potential. These are valid when the Einstein relation holds (ref. Section 3.2.3).

The net recombination rate R_{net} depends on the input models. In Table B.1 a list of used recombination models is shown. The recombination through deep defect levels in the band gap is usually Shockley-Read-Hall (SRH) recombination. Here it is chosen to use the doping dependent model plus a tunneling model which considers field enhancement which reduces the recombination lifetimes in regions where the electric field exceeds $\sim 3 \cdot 10^5 \text{ V/cm}$. This model is recommended to be included since the I-V characteristics of reverse biased p-n junctions is extremely sensitive to trap-assisted tunneling which sets in before band-to-band tunneling and avalanche breakdown. Adding the Auger, Band2Band and Avalanche models enables simulation of the breakdown behaviour of the device.

The electron and hole mobilities depend mainly on the lattice temperature which affects the phonon scattering that is included in the simulation by default. They also depend on the doping concentration thus the DopingDependence model is applied. Additional models take account of the carrier-carrier scattering, the saturation of carrier velocity at high electric fields and the

Resistivity (Ωcm)	20	80	200	1000
Concentration (cm^{-3})	6.725e14	1.667e14	6.650e13	1.328e13

Table 5.1: Substrate doping concentration of corresponding resistivities, calculated and checked using [128] and [129].

mobility degradation at e.g. semiconductor-insulator interface due to the perpendicular electric field. All the different mobility models add up according to the Matthiessen rule

$$\frac{1}{\mu} = \sum_i \frac{1}{\mu_i} \quad . \quad (5.4)$$

The effective intrinsic density is the density of states in the band gap and at band edges. The width of the band gap in semiconductors can shrink at high doping concentrations larger than $\sim 10^{17} \text{ cm}^{-3}$ and can be described by the Slotboom model [127]. Finally, the van Dort quantisation model is applied, which is robust, fast and proven and therefore the preferred choice [126]. Quantum effects become important at the scale of a MOSFET. This adds a potential-like correction to the classic electron and hole state densities.

5.3 Simulations of H35DEMO

Before the submission of the sensor, extensive TCAD simulations are made on the model with the cross-section as shown in Fig. 5.4 bottom and some variations thereof, shown in Fig. 5.7. The wells are very deep, e.g. the DNTUB is about $6 \mu\text{m}$, and it is not known whether they are implemented using ion implantation or through epitaxial growth. For this reason a simplified structure simulation is used.

While the substrate resistivities are given and can be converted to the corresponding doping concentrations in [128] and [129], the concentrations of the implants can only be estimated. Both values are shown in the Table 5.1 and Table 5.2, respectively. The doping concentrations of the implants are always the sum of the concentration in the substrate and the estimated concentrations of the given structures. Sprocess is used to prepare the mesh and sdevice to run the device simulation using the physics models listed in Table B.1.

Feature	DNTUB	DPTUB	SNTUB	SPTUB	n ⁺	p ⁺
Concentration (cm ⁻³)	1e17	1e17	1e18	1e18	1e19	1e19

Table 5.2: Doping concentration of implants based on educated guess.

All simulations are in 2D using the cross-section along the long side of the pixel of 250 μm as shown in Fig. 5.4 and assuming a thickness of 100 μm . The simulation includes the HV contacts which are shared with neighbouring pixels. Thus the simulated structure is $270 \times 100 \mu\text{m}^2$ in the x- and y-direction with an area factor of 1 μm , which is the width in the z-direction. The floating n⁺ and p⁺ implants for the n- and p-MOS in the central tub are removed for convergence reasons. On top of the substrate is a thin layer of oxide of 0.1 μm thickness wherever there is no electric contact. The contacts are defined as a thin aluminium layer of 0.5 μm . However, in the device simulation these are treated as ideal conductors.

In the process, it is possible for the space between DNTUBs to be filled with p-wells which would have the inverse mask compared to the DNTUB [130]. These would form an insulating layer between DNTUBs underneath the oxide to compensate for a conducting electron channel which can form due to radiation-induced oxide charge, as described in Section 3.3.1. The option of having an additional p-spray or p-wells layer is also explored in the simulation and is shown in Fig. 5.7 middle. The p-spray layer has the same doping concentration as DPTUB and the same thickness as the DNTUB.

A back-side-processed sensor was also simulated. The back side processing spans across the whole sensor width and includes an SPTUB, a p⁺ implant and an aluminium layer of 0.5 μm as the HV contact. The contacts on the top are repurposed as floating guard rings in this scenario.

Many different ways of modelling radiation damage in TCAD exist, e.g. as proposed by the RD48 (ROSE) and RD50 collaborations [87, 131] and others [132, 133, 134, 135]. All of the models have parameters tuned in such a way that they fit the specific data. For the H35DEMO the behaviour after radiation damage had to be predicted and there were no data to compare, therefore a recent model, [136], was applied. The model used for p-type silicon is a modified version of the model by University of Perugia (“Perugia model”),

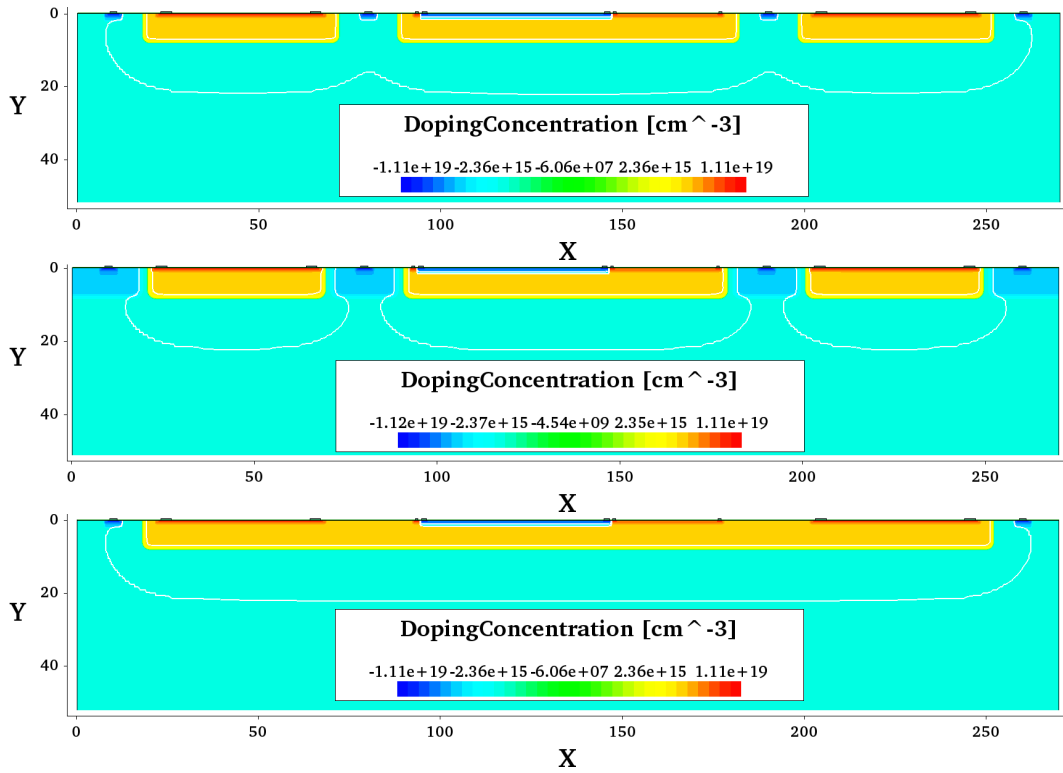


Figure 5.7: H35DEMO doping profile in the default layout (top), with additional p-spray (middle) and with continuous DNTUB (bottom). In the z-scale donors a positive value and acceptors negative. The white lines indicate the edge of the depletion regions.

Type	Energy (eV)	Defect	σ_e (cm ²)	σ_h (cm ²)	η (cm ⁻¹)
Acc.	$E_C - 0.42$	VV	9.5×10^{-15}	9.5×10^{-14}	1.613
Acc.	$E_C - 0.46$	VVV	5.0×10^{-15}	5.0×10^{-14}	0.9
Don.	$E_V + 0.36$	C_iO_i	3.23×10^{-13}	3.23×10^{-14}	0.9

Table 5.3: Radiation damage model for p-type silicon (modified Perugia model) from [136]. VV and VVV are double and triple vacancies, respectively. C_iO_i is carbon-oxygen interstitial.

which is an effective model that consists of two acceptor levels and one donor level as shown in Table 5.3. The introduction rate of defects is η and the defect concentration is $c = \eta\phi$. For the considered fluences the concentration of interface defects in the silicon-oxide interface is assumed to be saturated with a concentration of $\sim 10^{12} \text{ cm}^{-2}$ [137]. In these simulations the acceptor removal effect was attempted but at the end not considered as its parameters depend even more on the technology and the data.

Finally, the behaviour of a particle, typically a MIP passing through the device is simulated and the charge collection of the diode is studied.

5.3.1 Electric Field

The absolute electric field strength in a pixel reverse-biased at -120 V is displayed in Fig. 5.8 and Fig. 5.9, for the standard layout and the p-spray layout, respectively, both biased from the top. Fig. 5.10 shows the standard layout biased through the back side contact. All four resistivities of 20, 80, 200 and $1000 \Omega\text{cm}$ are represented in (a) to (d), respectively. The p-spray layer has a higher doping than the substrate and thus results in a stronger electric field at the junction to DNTUB. For low resistivities the electric field is stronger near the p-n junction than for higher resistivities. However, the field strength decreases faster resulting in only a shallow high-field region. For the higher resistivities the field strength is more uniform and reaches deeper into the substrate. The black lines qualitatively represent the electric field lines. They show a huge difference between the top- and back-side-biased sensors being more uniform for the latter case. The effect of the p-spray and the back side biasing is more prominent for higher resistivities and can be seen in Fig. 5.11.

1D plots are produced by extracting the desired property along the centre of the pixel at $x = 135 \mu\text{m}$, unless otherwise stated.

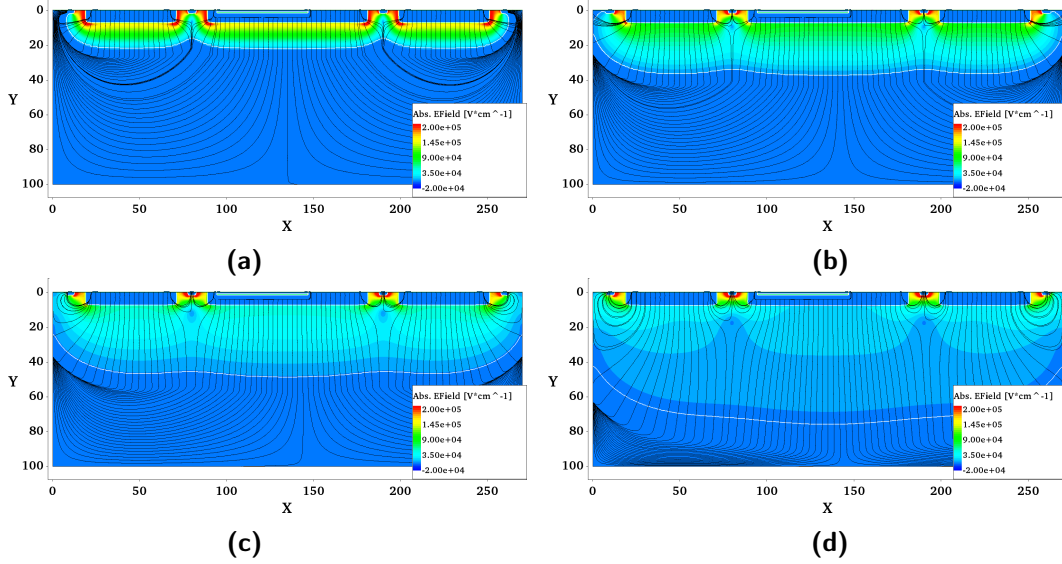


Figure 5.8: Absolute electric field strength of H35DEMO for resistivities 20 (a), 80 (b), 200 (c) and $1000 \Omega\text{cm}$ (d), biased from the top at -120 V for the standard layout.

5.3.2 Current-Voltage Characteristics

Fig. 5.12 shows the leakage current (top) for the default, p-spray and back-side-biased layouts and their breakdown voltages (bottom). It is noticeable that the layout with a larger electric field extension also has a larger leakage current. For the back-side-biased sensor no breakdown is observed up to -400 V . For the $20 \Omega\text{cm}$ substrate the breakdown occurs at around -450 V , for $80 \Omega\text{cm}$ at about -600 V , while no breakdown up to -600 V is observed for higher resistivities. All simulated breakdown voltages are far above the expected operation voltages of up to -120 V .

5.3.3 Depletion Region

In Fig. 5.8–Fig. 5.10 the white lines indicate the boundaries of depletion or space charge regions. For the back-side-biased sensor the depletion region

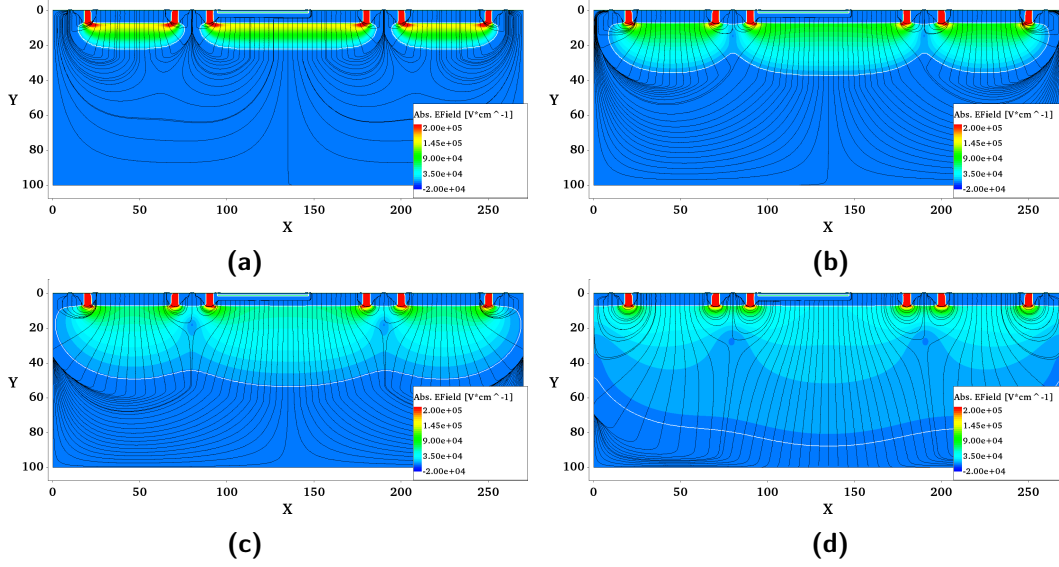


Figure 5.9: Absolute electric field strength of H35DEMO for resistivities 20 (a), 80 (b), 200 (c) and 1000 Ωcm (d), biased from the top at -120 V for the standard layout with additional p-spray layer.

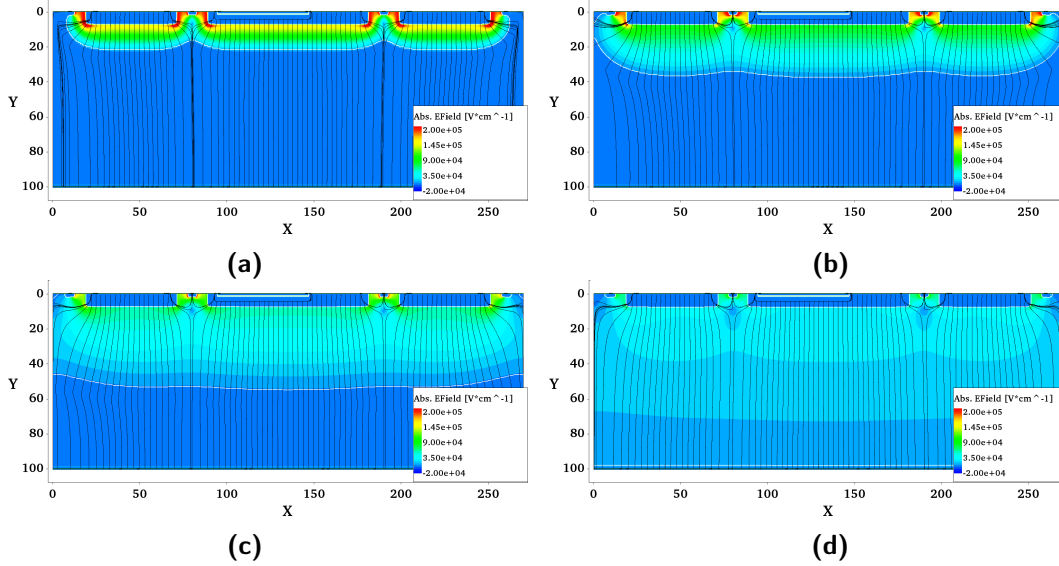


Figure 5.10: Absolute electric field strength of H35DEMO for resistivities 20 (a), 80 (b), 200 (c) and 1000 Ωcm (d), biased from the back at -120 V for the standard layout with back side process.

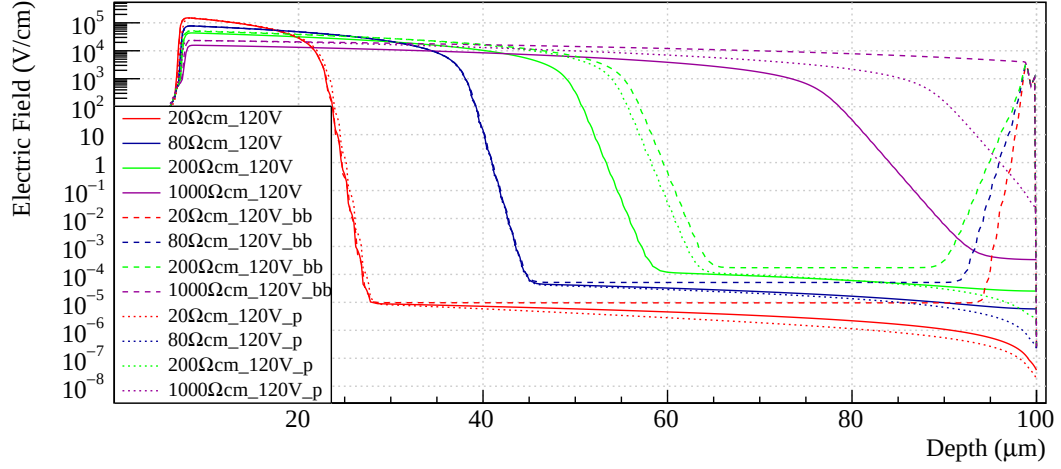


Figure 5.11: Absolute electric field as a function of the depth into the sensor. The label “_p” in the legend denotes the geometry with the extra p-spray and “bb” the back-side-biased sensor.

is substantially larger at the same resistivity and the same bias voltage. This is even clearer to see in Fig. 5.13 which shows the depletion depth extracted using space charge. The boundary of the depletion region is defined as the point at which the space charge density is halfway between the values in the depletion region and in the undepleted substrate as shown in Fig. 5.15. The fill factor is the ratio of depleted area to the total sensor area and is shown in Fig. 5.14.

5.3.4 Capacitance-Voltage Characteristics

The design with three separated DNTUBs (Fig. 5.4 top) was introduced to reduce the sensor capacitance compared to a design with a continuous DNTUB (Fig. 5.4 bottom). Fig. 5.16 displays the total DNTUB capacitance which shows that the opposite is the case for all resistivities. However, the difference is not severe. The sudden declines in the curves are owing to the fact that the depletion region reaches the lateral edge of the simulated pixel as shown in Fig. 5.17.

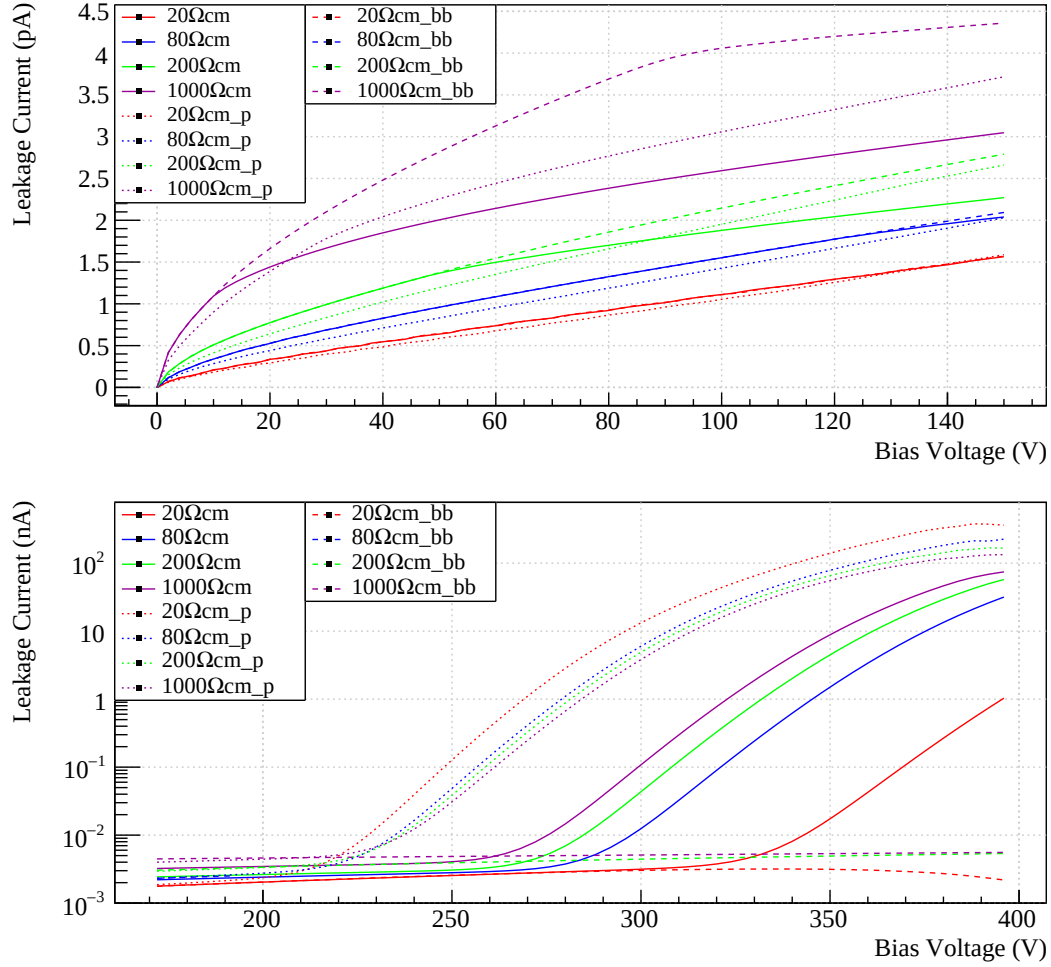


Figure 5.12: Simulated H35DEMO leakage current vs. bias voltage.

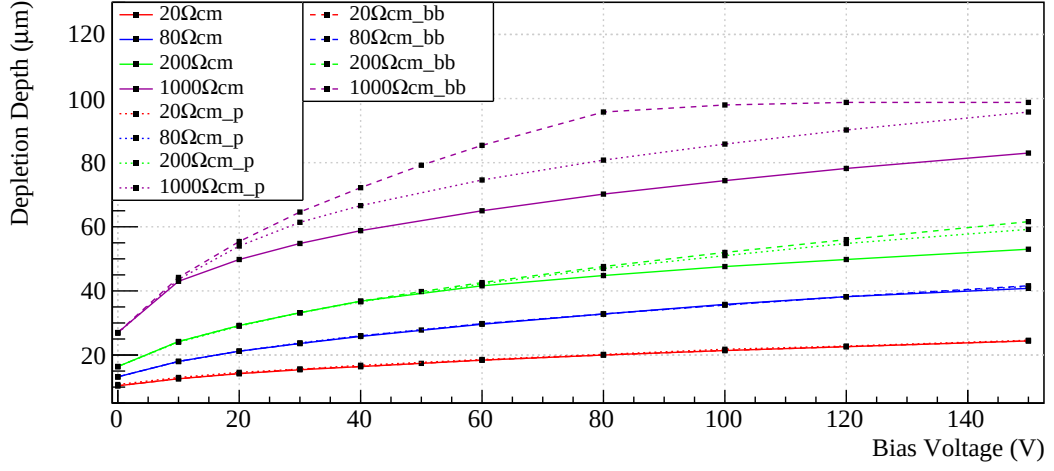


Figure 5.13: Simulated depletion depth for all layouts and resistivities.

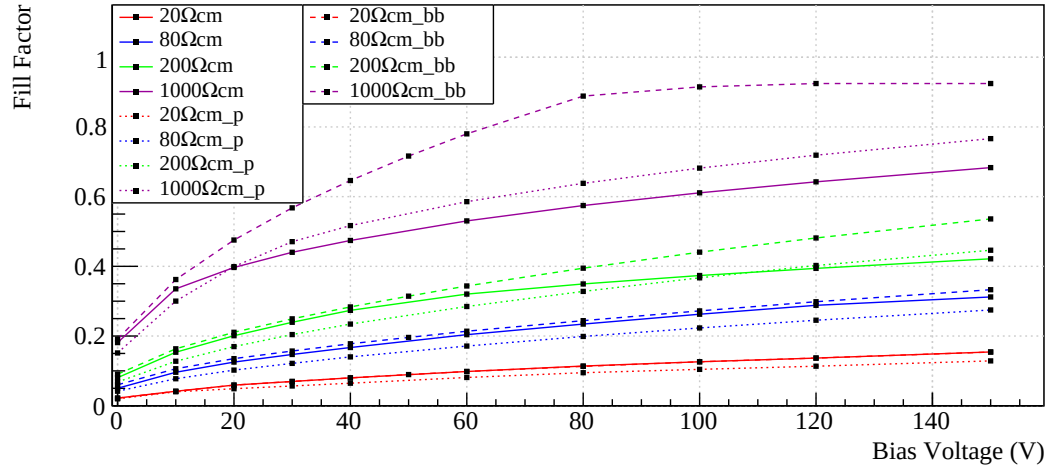


Figure 5.14: Simulated fill factor of the depleted area of the H35DEMO as a function of the bias voltage.

5.3.5 Radiation Damage

The simulation of radiation damage is performed for the top-side-biased standard layout. The obtained electric field after different fluences is shown in Fig. 5.18. The four plots are for substrate resistivities of 20, 80, 200 and 1000 Ωcm substrate biased at -120 V . Each plot shows the electric field along $x = 135\text{ }\mu\text{m}$ for fluences of $1 \cdot 10^{14}$, $1 \cdot 10^{15}$ and $1 \cdot 10^{16}\text{ n}_{\text{eq}}/\text{cm}^2$. In the 20 Ωcm

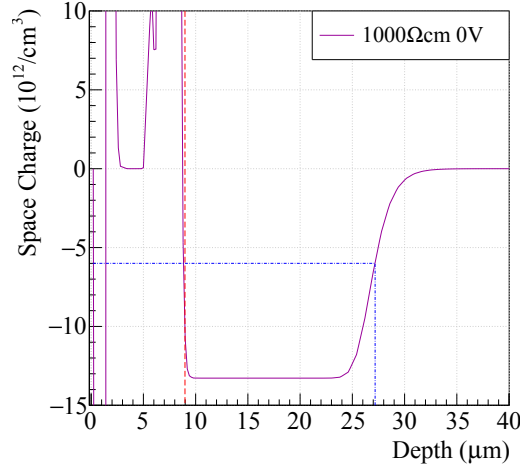


Figure 5.15: Space charge at $x = 135 \mu\text{m}$ as a function of sensor depth for an unbiased pixel on a $1000 \Omega\text{cm}$ substrate. The red dashed line shows position of the p-n junction at about $9 \mu\text{m}$ and the horizontal blue dot-dashed line shows the space charge threshold below which it is considered as depletion region whereas the vertical line leads to the corresponding depth.

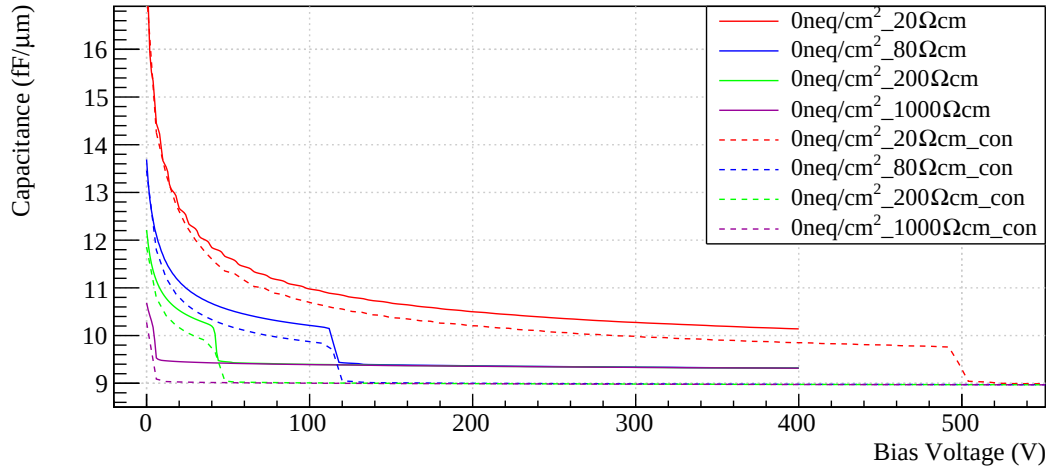


Figure 5.16: Sensor capacitance. Design with three separated DNTUBs and one continuous “con” DNTUB are compared.

substrate the high-field region (on a logarithmic scale) does not change with increasing fluence, whereas in the low-field region the strength increases with radiation. This might be a result of the radiation adding defects and resulting in higher effective doping concentrations. This can be confirmed by looking

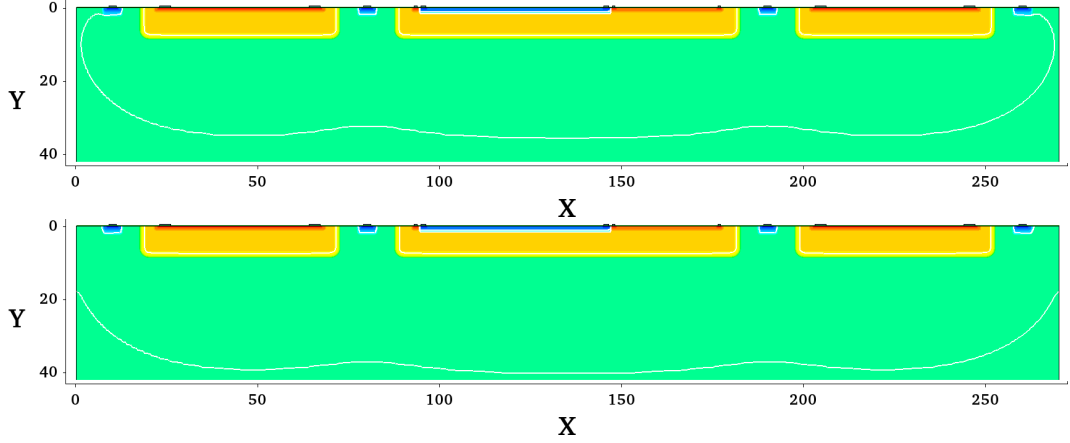


Figure 5.17: Depletion region (white line) at -40 V (top) and -60 V for a $200\ \Omega\text{cm}$ substrate. While at -40 V the depleted area is entirely within the cross-section, at -60 V the depleted area extends beyond it and thus causes a sudden drop in the capacitance.

at the higher resistive substrates. The field strength in the high-field region increases with radiation but the field becomes shallower and more similar to the behaviour of the electric field in the lower resistivity substrates. Furthermore, it can be seen that the leakage current increases with radiation, as expected.

5.3.6 Charge Collection

To simulate the charge collection behaviour of the H35DEMO a minimum ionising particle is sent perpendicular to the surface through the sensor at different positions as shown in Fig. 5.21: in the middle of the pixel and the central DNTUB, and between two DNTUBs. It passes through the sensor at $t = 0.1\text{ ns}$ and creates 80 e-h pairs per μm travelled. The time and spatial distribution of generation rate of the MIP is Gaussian with a standard deviation of 2 ps (default) and 0.025 nm (user-defined), respectively. Fig. 5.22 shows the signal current and the accumulated charge for the top-side-biased and the back-side-biased sensors. The dip at the start of the signal in the plot with the central MIP is due to the fact that the MIP also deposits charge in the PTUB. As can be seen in Fig. 5.24 and Fig. 5.25, the holes within the PTUB are collected by the GND electrodes before the electrons reach VDD

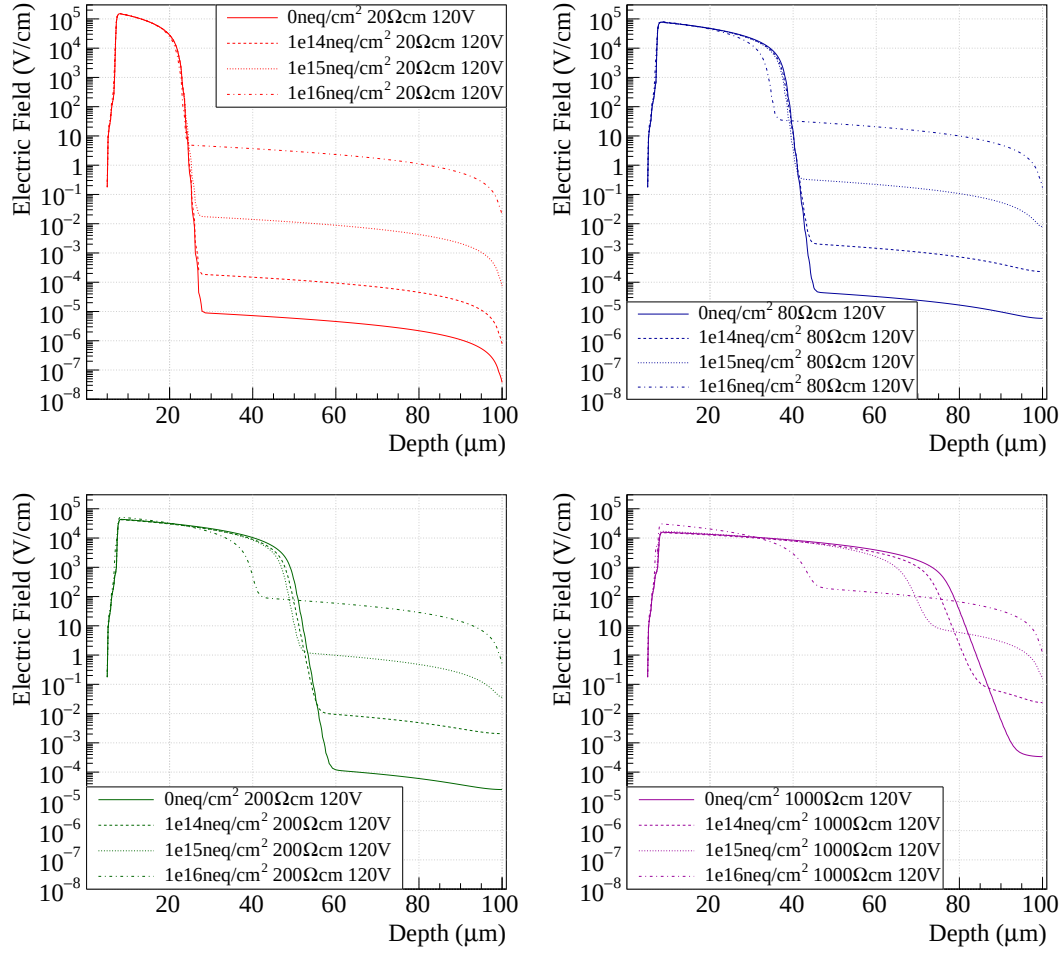


Figure 5.18: Absolute electric field for along $x = 135 \mu\text{m}$ at fluences of $1 \cdot 10^{14}$, $1 \cdot 10^{15}$ and $1 \cdot 10^{16} \text{ neq/cm}^2$ for resistivities 20 (top left), 80 (top right), 200 (bottom left) and 1000 Ωcm (bottom right).

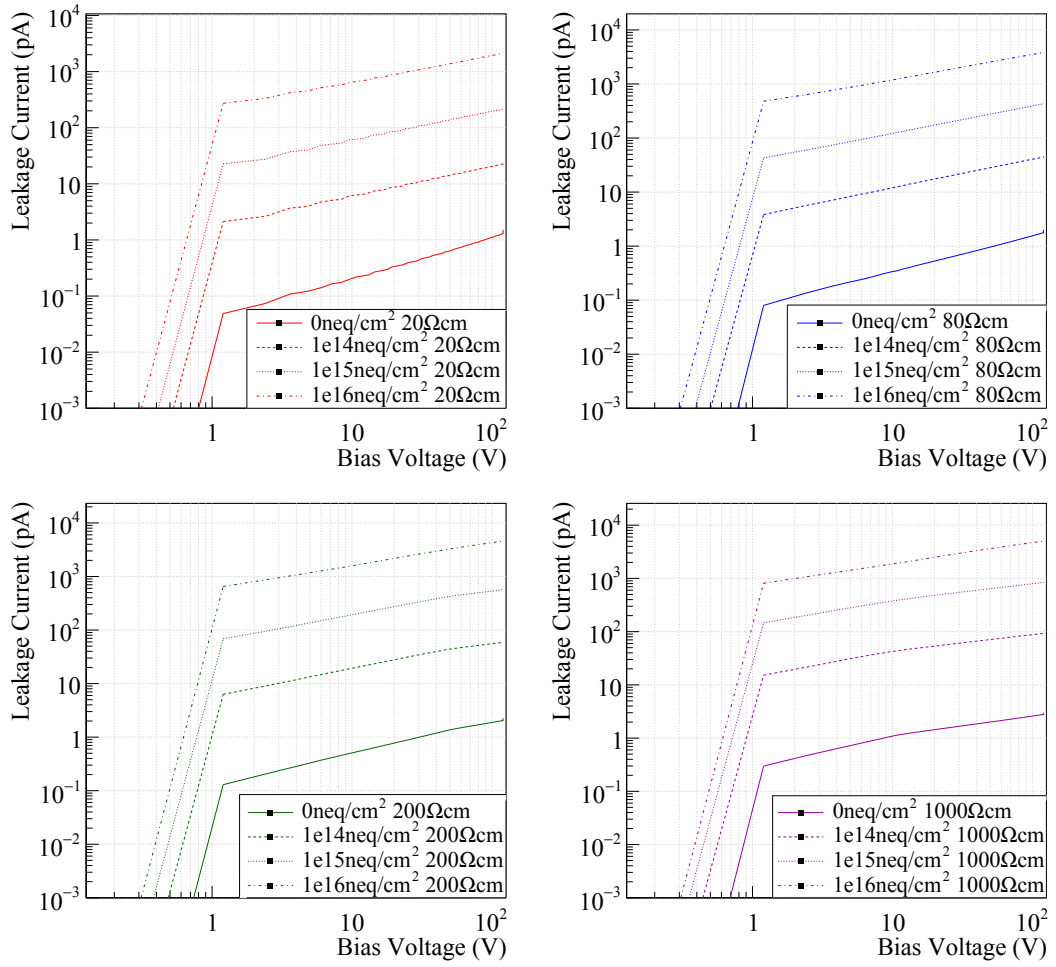


Figure 5.19: Leakage current at fluences of $1 \cdot 10^{14}$, $1 \cdot 10^{15}$ and $1 \cdot 10^{16} \text{ neq}/\text{cm}^2$ for resistivities 20 (top left), 80 (top right), 200 (bottom left) and 1000 Ωcm (bottom right).

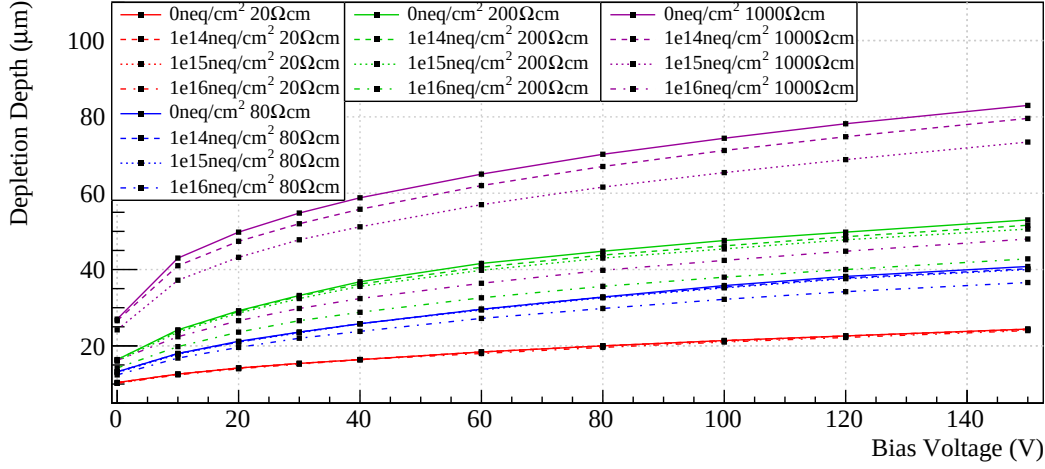


Figure 5.20: Simulated H35DEMO depletion depth vs. bias voltage for the standard layout for different fluences.

(ref. Fig. 5.4). A particle passing through the SNTUB does not create such signal as shown in Fig. 5.25.

The signal has contributions from drift and diffusion. A more detailed broken-down simulation can be found in [138]. While the drift signal is almost immediate, has a high amplitude and thus contributes to a fast rise in the collected charge, the diffusion signal is slower and peaks only after 0.1 ns. The drift signal gains amplitude with increasing resistivity and thus depletion depth, while the diffusion signal behaves the opposite due to a decrease in undepleted depth with increasing resistivity. The accumulated charge saturates, meaning all charge is collected, only after several μs as can be seen in Fig. 5.22 bottom. The particle passing between two DNTUBs has a much higher initial drift and diffusion signal. This is probably due to the much stronger electric field between the DPTUB and the DNTUBs and due to the charge being collected by a larger length of p-n junctions. However, for low resistivities such as $20\ \Omega\text{cm}$, or for small bias voltages, the depleted area between two DNTUBs is smaller which results in a smaller drift signal. An extreme example is the $20\ \Omega\text{cm}$ biased at $-20\ \text{V}$ as shown in Fig. 5.26 where no drift signal can be observed. The top left picture in Fig. 5.22 illustrates the reason: the particle does not pass through any depleted region and thus the generated charge.

Compared to the back-side-biased sensor where the depletion zone reaches deeper into the bulk and the electric field is stronger, the drift contribution for the central particle improves drastically for the MIP passing through the centre of the DNTUB. Both the drift and diffusion components have a higher amplitude in the signal. The $1000\,\Omega\text{cm}$ sensor is fully depleted when back-side-biased at $-120\,\text{V}$ and thus the charge collection happens via drift alone. It only takes a few ns for all charge to be collected. For the MIP passing between DNTUBs the signal is weaker because some of the generated charge escape through the guard ring as shown in Fig. 5.23. In addition, the electric field between the DPTUB of the guard ring and the DNTUBs are weaker with increasing resistivity as can be seen in Fig. 5.10.

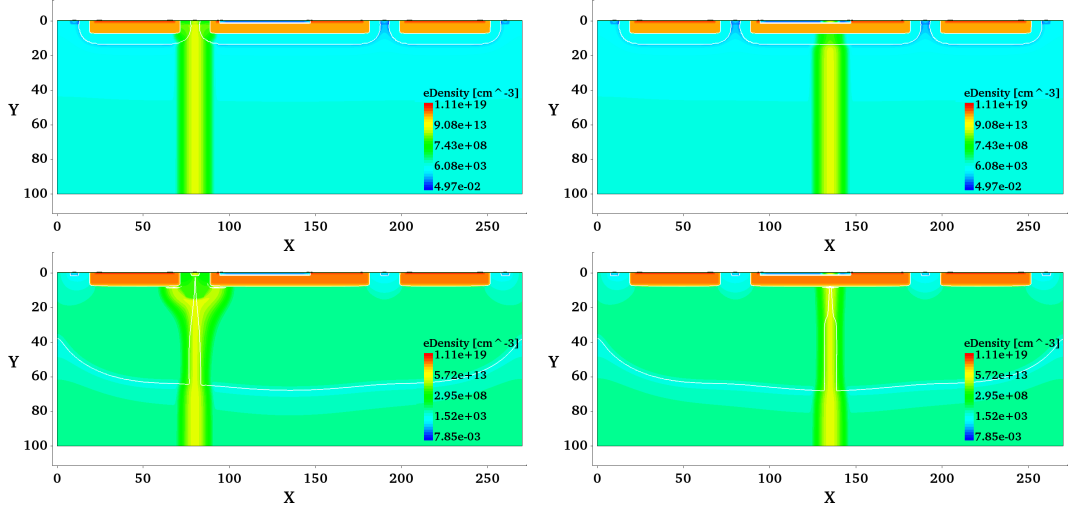


Figure 5.21: MIPs in between two DNTUBs (left) and in the middle of the pixel (right) in a $20\,\Omega\text{cm}$ sensor biased at $-20\,\text{V}$ (top) and a $1000\,\Omega\text{cm}$ sensor biased at $-80\,\text{V}$ (bottom).

5.3.7 Edge TCT Simulation

Transient Current Technique (TCT) measurements are a way to simulate particles passing through a detector by using a laser to ionise the material. TCAD simulations were done and compared to the measurements. The objective is to compare the shape of the depletion region with the measurement which is shown in Fig. 5.28.

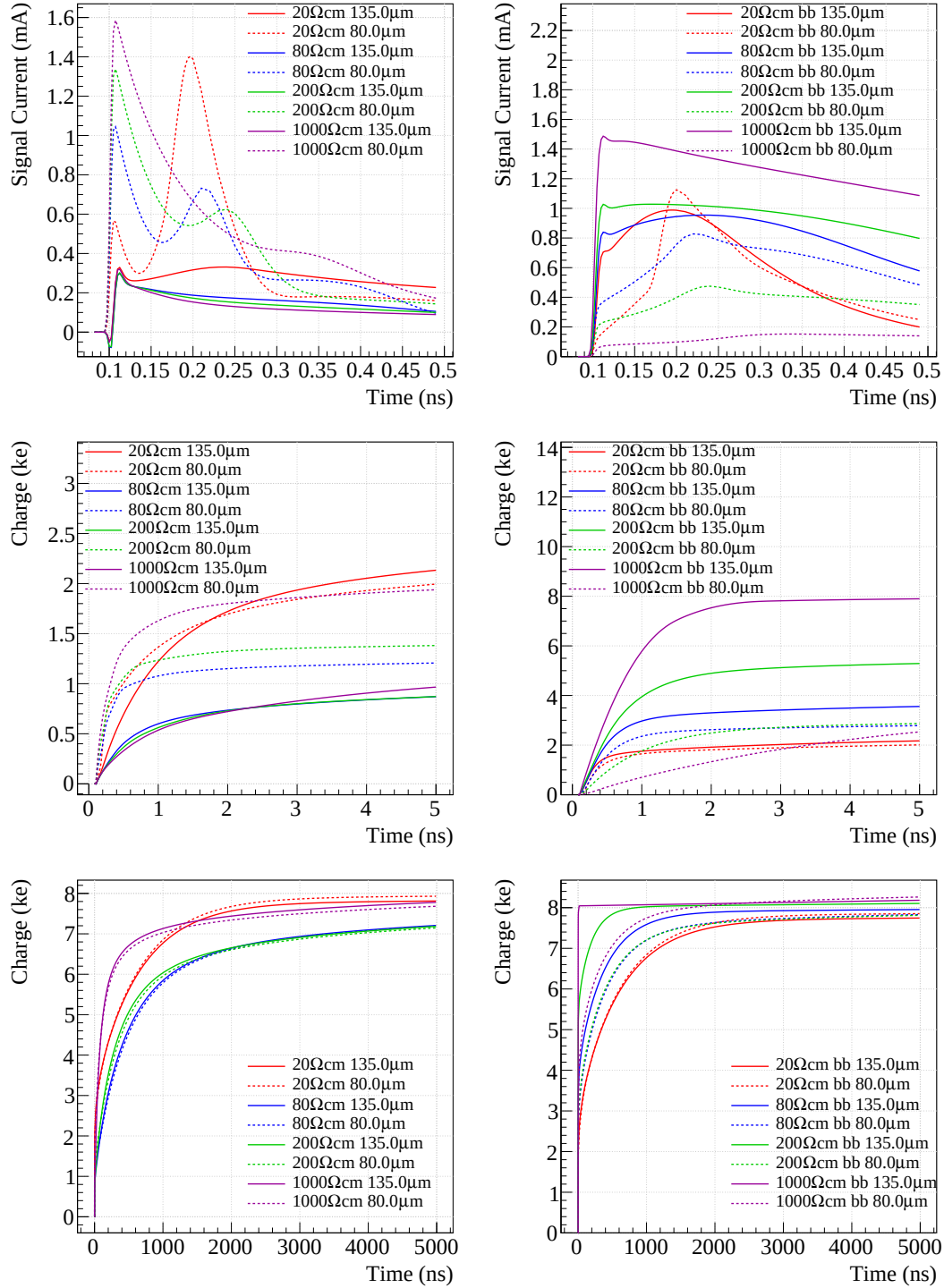


Figure 5.22: Signal (top) and accumulated charge (middle and bottom) of a MIP entering the pixel at different positions for top-side-biased (left) and back-side-biased sensor (right). The MIP enters the device at 0.1 ns.

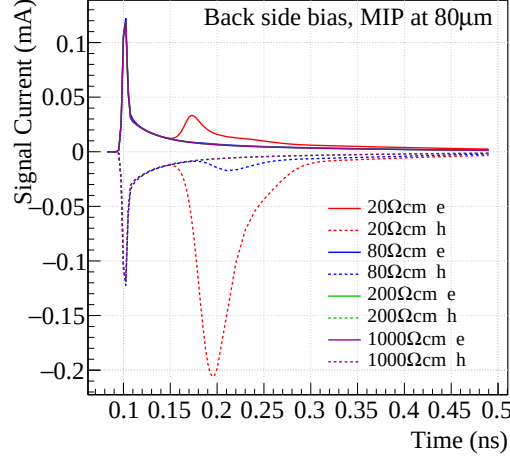


Figure 5.23: Electron (solid lines) and hole (dashed lines) current through the guard ring.

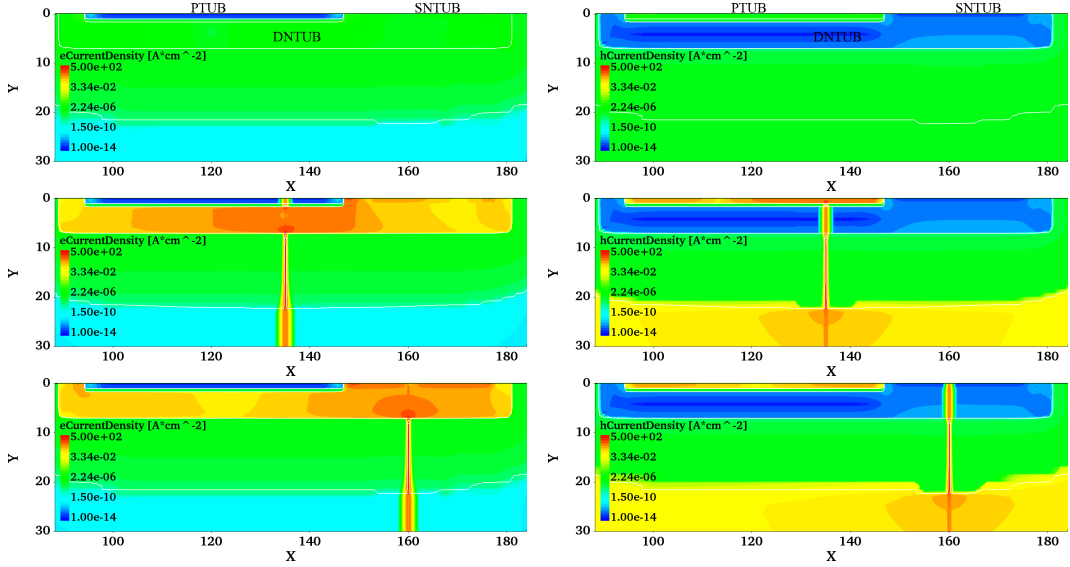


Figure 5.24: Electron (left) and hole (right) current densities for without MIP (top), MIP passing through the central nested PTUB (middle) and SNTUB (bottom) at 110 ps in a $20\ \Omega\text{cm}$ sensor biased at $-120\ \text{V}$ (bottom).

The TCT measurements use the dedicated 3×3 pixel test structure on the lower edge of the sensor. In this test structure the central pixel has an analog output. The eight surrounding pixels are shorted together and have a common output. A laser with 1060 nm wavelength is shot at the edge of the

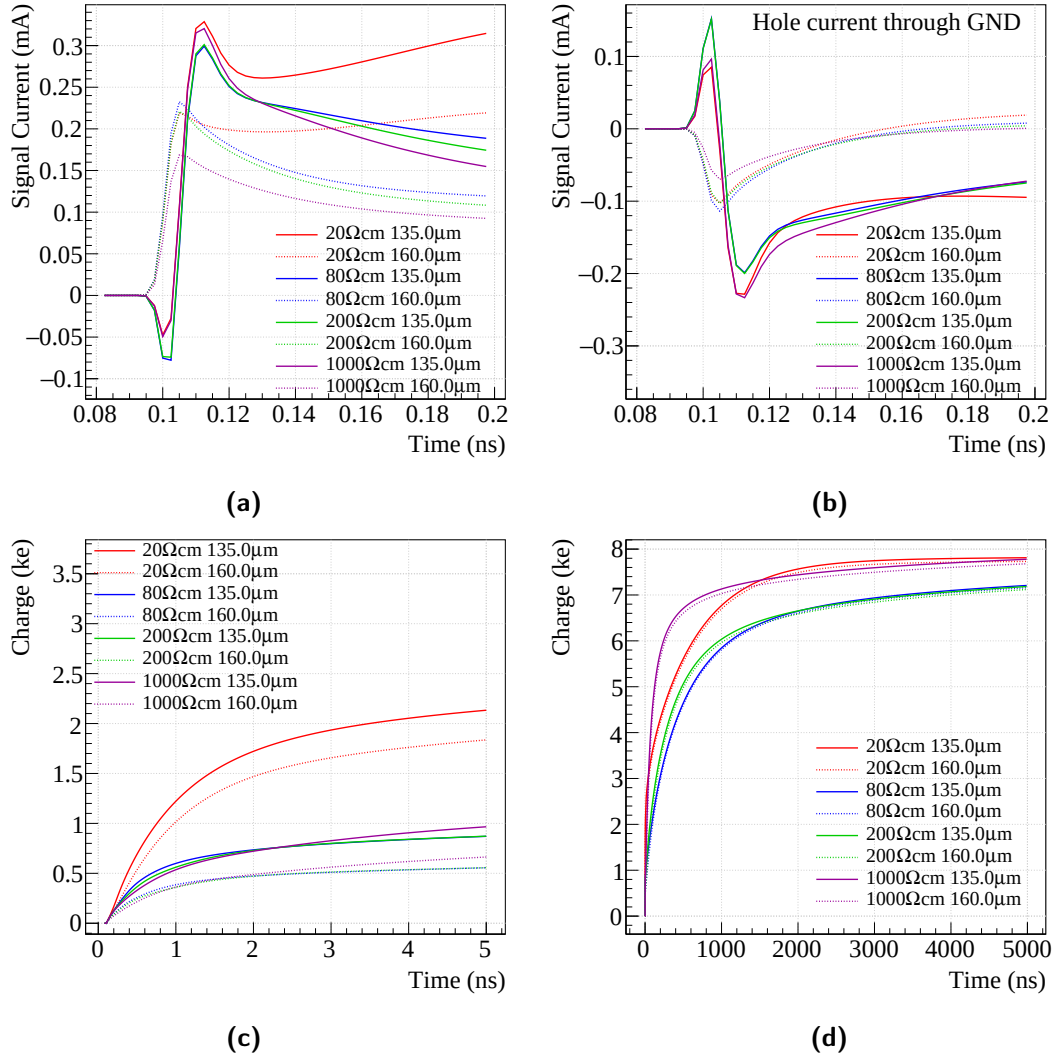


Figure 5.25: Current signal (a) and integrated charge (c) and (d) of MIP going through the nested PTUB and hole current of the PTUB (b) for a top-side-biased sensor at -120 V. The MIP enters the device at 0.1 ns.

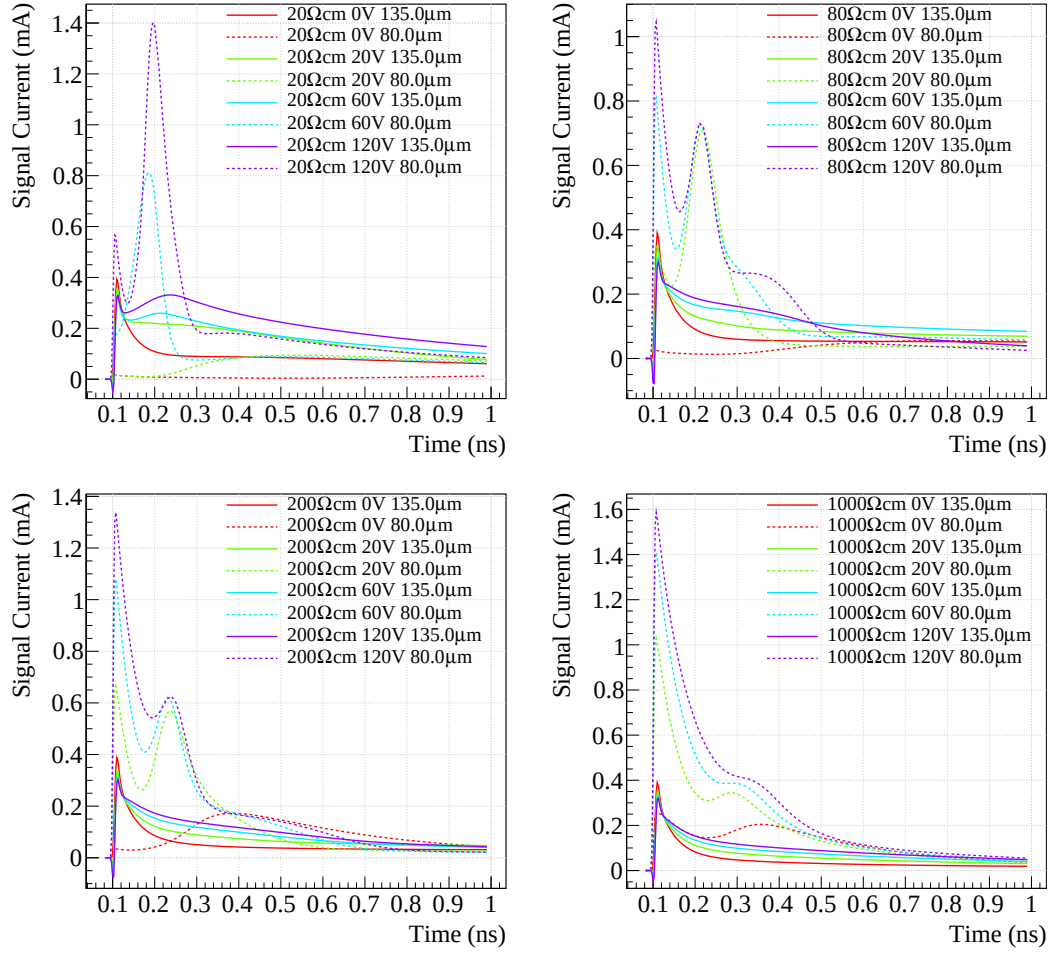


Figure 5.26: MIP signal at different bias voltages for a top-side-biased sensor with different resistivities.

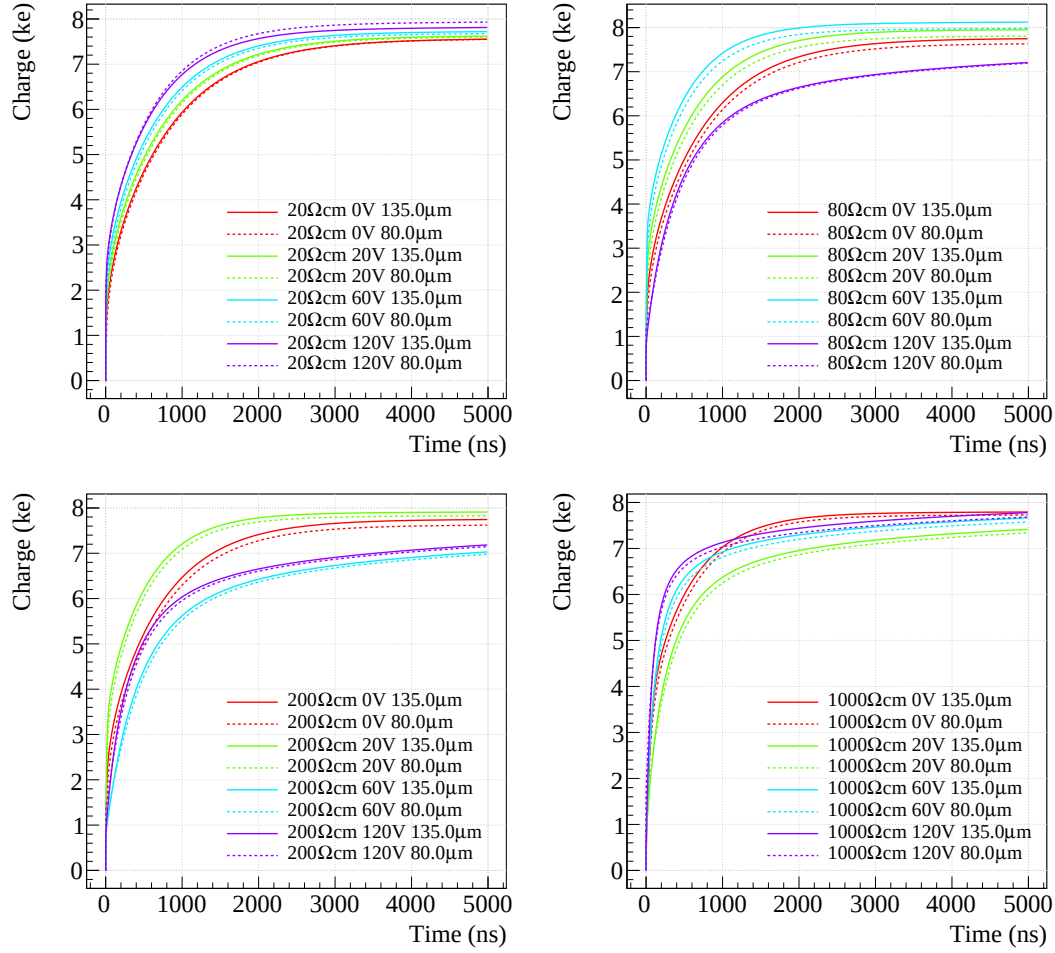


Figure 5.27: MIP charge at different bias voltages for a top-side-biased sensor with different resistivities.

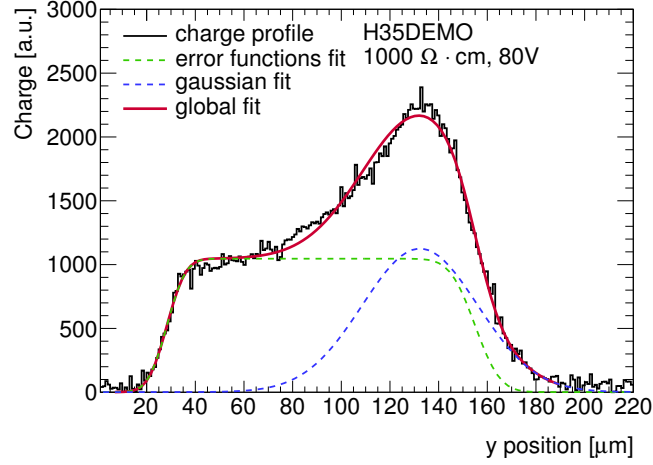


Figure 5.28: Depletion depth of the test structure of an unirradiated $1000\ \Omega\text{cm}$ sensor measured using eTCT. The top side of the sensor is at about $25\ \mu\text{m}$ [141].

sensor in edge-TCT (eTCT). The laser is focussed inside the silicon with a focus diameter of $5\text{--}11\ \mu\text{m}$ [139, 140] and a focus length of at least the width of the test structure of $150\ \mu\text{m}$ (ref. Fig. B.14, Fig. B.15 and [141]).

The $1000\ \Omega\text{cm}$ sensor is unirradiated and has a $760\ \mu\text{m}$ thickness. The test structure is biased at $-80\ \text{V}$ from the top. The measurement is a scan through the depth of the sensor and the collected charge of the central pixel is obtained by integrating the signal current over $\sim 10\ \text{ns}$ for each step. In contrast to expectation, the deeper regions of the depletion zone collect more charge than the region immediately below the p-n junction where the electric field is the strongest. This is only observed for the high resistivity while for all other resistivities the charge collection is more constant throughout the depletion zone.

In the simulation the short edge of the 3×3 matrix is used, with a total width of $500\ \mu\text{m}$ to ensure that there is enough inactive material on both sides of the structure. The bulk thickness is $300\ \mu\text{m}$. TCAD offers the option of optical generation; however, the simulation time would exceed that for a MIP simulation by far. For a full eTCT scan this plays an significant role. Even using MIPs each step of the scan can take up to 2 days to finish. For the MIP simulation, a horizontal Gaussian particle track with $\sigma = 4\ \mu\text{m}$ is used,

comparable with the focus diameter of the laser. The MIP enters the sensor at different depths between 5–150 μm in 5 μm steps as shown in Fig. 5.30, and at 0.1 ns in time. The sensor is biased at -100 V and the electric field is shown in Fig. 5.29. The physics list used for the device simulation is the reduced version to save simulation time (cf. Appendix B.2). The current on the collection electrode is recorded for 300 ns to cover the integration time of the measurements and have some buffer. The signal shapes of the first few nanoseconds are shown in Fig. 5.31 for 20 and 1000 Ωcm .

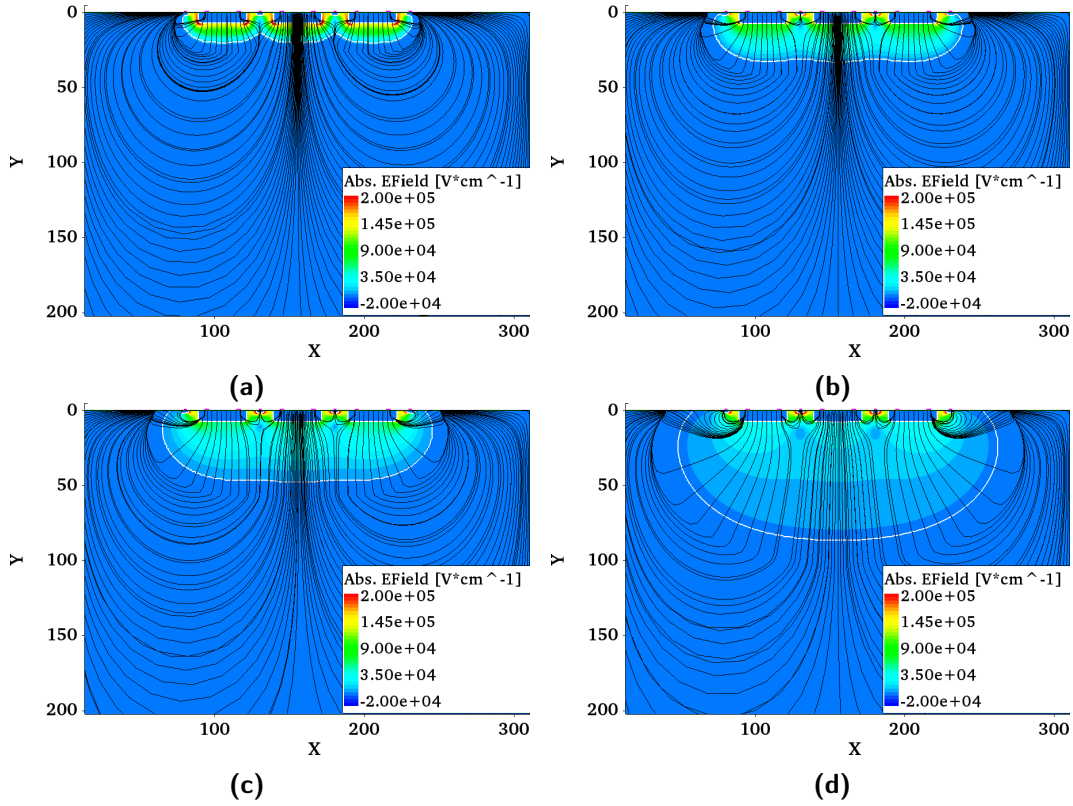


Figure 5.29: Electric field strength and streamlines of short side of the 3×3 test structure biased at -100 V for 20 (a), 80 (b), 200 (c) and 1000 Ωcm (d) resistivities.

In the 20 Ωcm sensor the charge collection happens almost immediately via drift in the DNTUBs and near the p-n junction. At 5 μm depth the MIP travels through the DNTUB and most charge is collected immediately. A small peak at around 0.2 ns can be seen, which comes from the charge deposited outside the DNTUB. This is more pronounced at 10 μm which is just below the

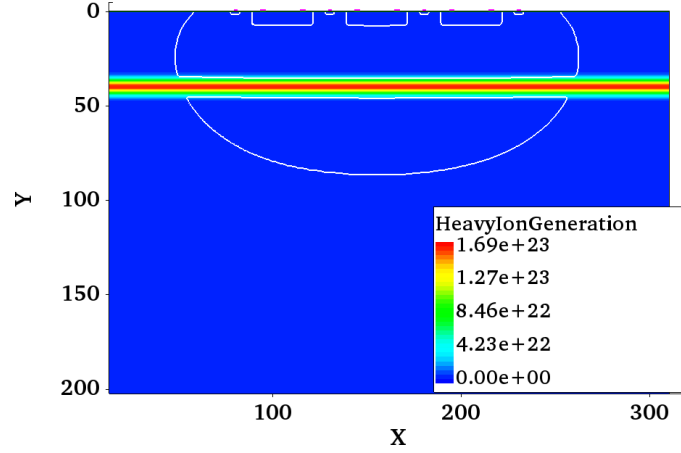


Figure 5.30: MIP passing through the test structure at $y = 45 \mu\text{m}$ for a $1000 \Omega\text{cm}$ substrate and biased at -100 V .

DNTUB. The immediate peak is smaller than at $5 \mu\text{m}$. For the $1000 \Omega\text{cm}$ substrate the MIP at $5 \mu\text{m}$ produces the same signal as for the $20 \Omega\text{cm}$ substrate. The signal becomes slower the deeper the MIP is in the depleted region due to a larger drift distance. In the deeper regions of the much larger depletion zone charge sharing between neighbouring pixels becomes visible through the negative component in the signal, which always corresponds to a peak in the signal of the neighbouring pixel. The integrated charge over 10 ns is shown in Fig. 5.32. For the $20 \Omega\text{cm}$ sensor the charge collection behaves as expected and the profile is the same across the depth for the central and the periphery pixels, whereas the two periphery pixels collect more than twice the charge of the central pixel because the particle deposits charge across the full width of the bulk. The lateral charge is collected by the periphery pixels. For the $1000 \Omega\text{cm}$ substrate the observed “double peak” of the central pixel can be reproduced. The total collected charge by all pixels corresponds to the shape of the depletion zone or the electric field. The further down into the sensor the MIP charge is deposited, the narrower the depletion region and the weaker the electric field; thus the collected charge decreases. This is also true for the lower resistivities. At medium depth into the depletion region, a horizontal particle track creates e-h pairs along the track and thus creates a non-depleted zone of free charge carriers in the middle of a depletion zone and perpendicular to the electric field. The field becomes slightly distorted as shown in Fig. 5.33.

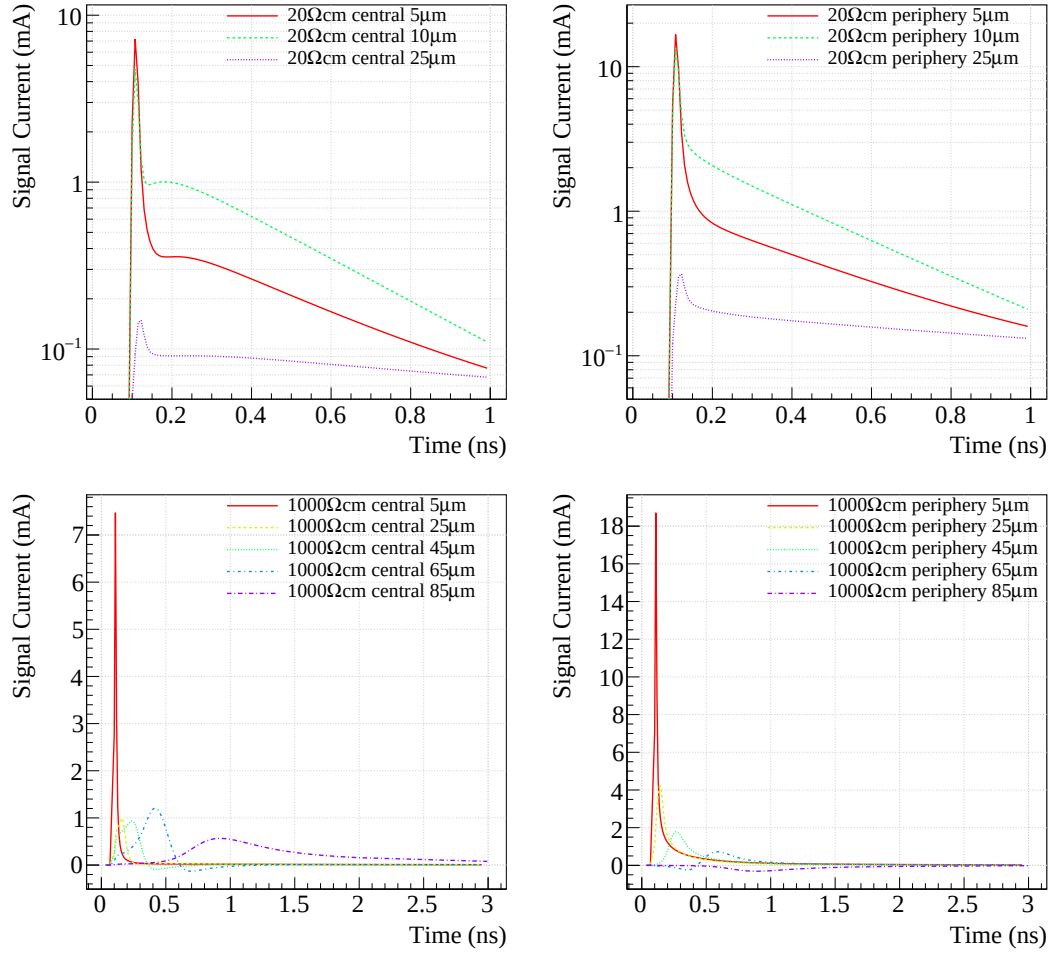


Figure 5.31: Simulated eTCT signals for 20 (top) and 1000 Ωcm (bottom) substrates at different depths for the central pixel (left) and periphery pixels (right). The MIP enters the sensor at 0.1 ns.

This decreases the mobility of the charge carriers. Integrating the signal by a longer time increases the charge collected in this area as shown in Fig. 5.34. It is apparent that the periphery pixels collect more charge due to the outspread of the electric field. At the lower edge, indicated by white lines, the depletion region is not being ripped apart by the particle track and thus the charge collection should behave normally. However, the collected charge clearly exceeds the normal case. This might be explained by the fact, that at the edge of a depletion region the electric field is still reasonably strong but very spread out. Therefore the charge carriers are funnelled into the central pixel, which can also be seen in the charge collected by the periphery pixels. Further simulations were run with a MIP track with a length equals the maximum depletion width (“contained MIP”) to avoid such edge effects in the simulation; a MIP track that creates half of the e-h pairs on the sides of the track than in the centre (“segmented MIP”) to emulate the length of the focussed laser beam; and with a higher maximum charge carrier lifetime of the order of 1 ms for electrons and 0.1 ms for holes (ref. e.g. [142]) instead of the default values [126] of 10 and 3 μ s, respectively, and default physics list, which was run for all scenarios. All simulations return the same charge collection profile. It is obvious that the collected charge within the first 300 ns is less than the total deposited charge since the saturation is not yet reached.

5.4 Simulations of Devices in the LFoundry 150 nm Technology

Simulations were performed for a multi-project wafer (MPW) submission in LF 150 nm technology. The simulated cross-section is shown in Fig. 5.35.

5.4.1 Doping Profiles

For the LF simulation a 2D TCAD process deck of a transistor on a standard 10 Ω cm substrate was provided by the foundry. However, in order to adapt this from the size of a transistor of about 1 μ m to the size of a pixel of 50 μ m, a process simulation would require a large amount of time, i.e. > 40

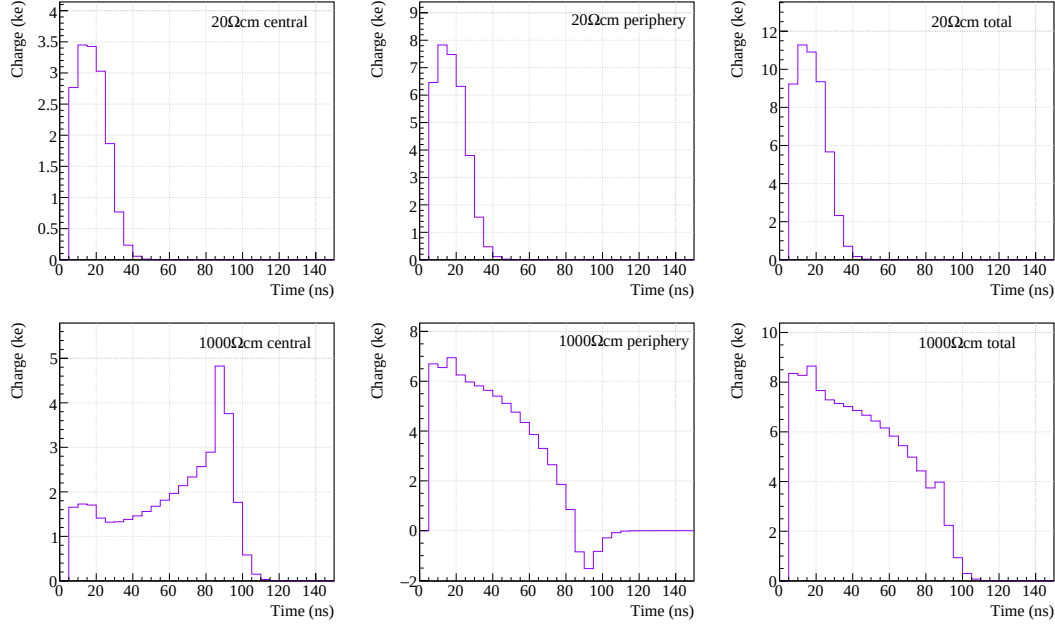


Figure 5.32: Simulated eTCT scan of 20 (top) and 1000 Ωcm sensor (bottom). Shown are the collected charge of the central pixel (left), the periphery pixels (middle) and the total collected charge (right), integrated over 10 ns.

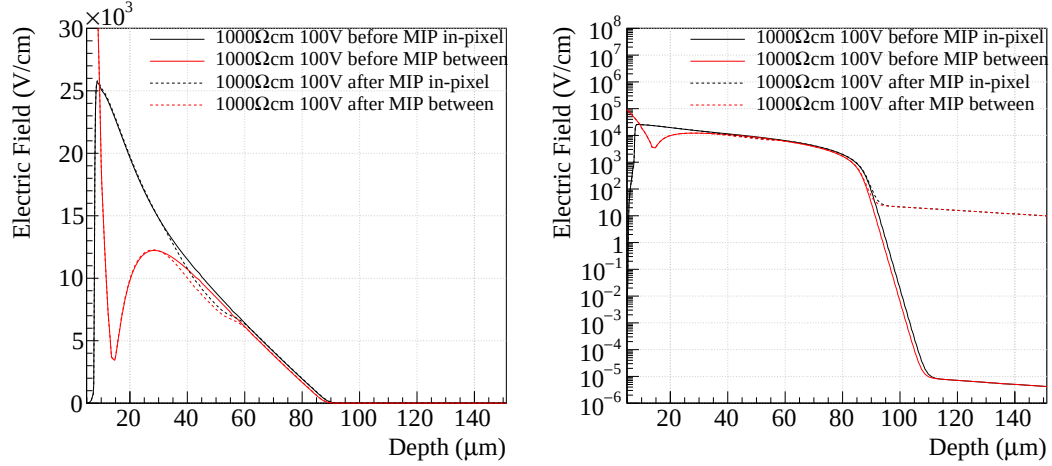


Figure 5.33: Electric field in the middle of a pixel (black) and between two pixels (red) before (solid) and after (dashed) a particle passed through at 50 μm . The two plots show the linear (left) and logarithmic (right) scale.

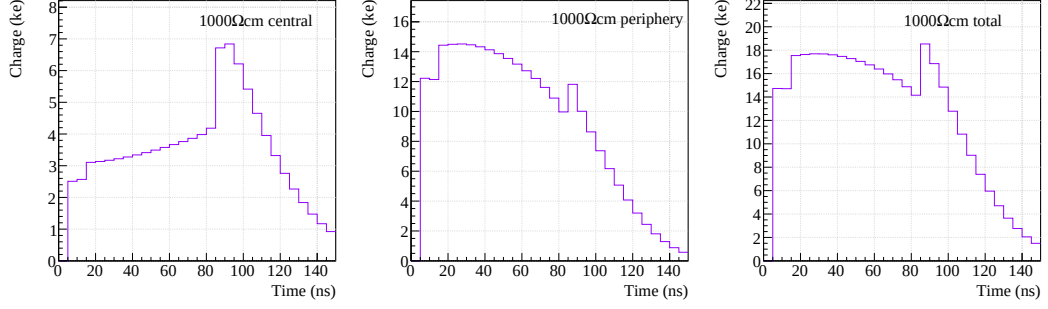


Figure 5.34: Simulated eTCT scan of 1000 Ωcm sensor with 300 ns integration time of the signal.

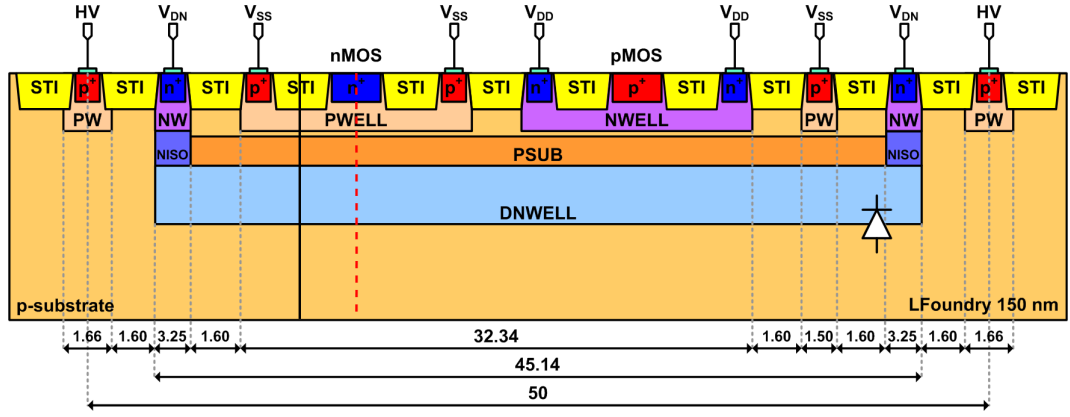


Figure 5.35: Preliminary pixel cross-sections for an LF MPW submission in 2016 [123]. Most work done is based this layout which is referred to as preliminary cross-section while the final layout has mainly different dimensions and is referred to as the MPW layout.

CPU hours. Therefore the doping profiles of the desired features were extracted from the tiny process simulation and fed into a new structure simulation.

The doping profiles and the implant depths of the same wells differ depending on whether there is a shallow trench isolation (STI) at the surface to separate the implants. Ion implantation happens after STI creation. Therefore, the combined stopping power of the different materials results in a different doping profile. As shown in Fig. 5.36, under an STI the implant depth is shallower but the concentration is higher. Assuming most part of the pixel is not covered by STI, the profile shown as a red dashed line is used. To extract the doping profiles, a small process simulation is run with the features to be

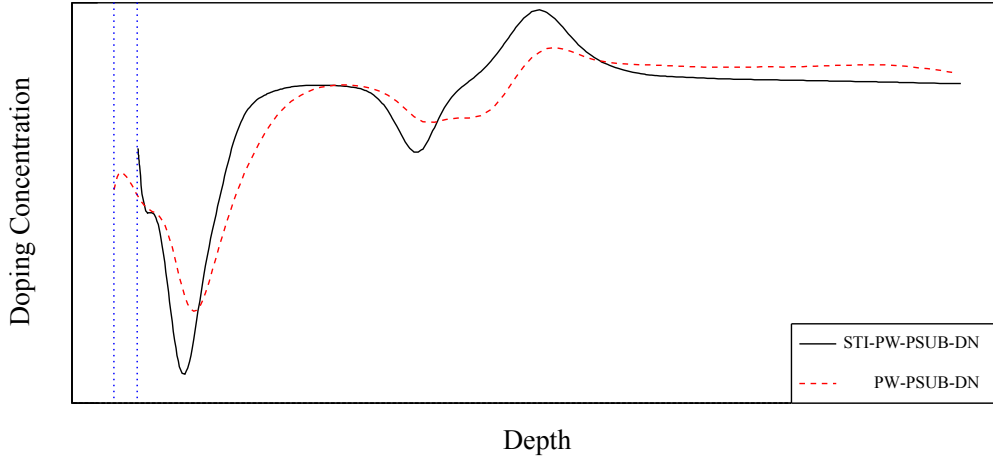


Figure 5.36: Comparison of doping profiles extracted from process simulation. Black solid line represents the doping profile of a region under an STI and red dashed line is the profile in regions without STI for the same implants PW-PSUB-DN, corresponding to the cut lines in Fig. 5.35 without the n^+ doping. No scales are shown due to confidentiality.

extracted, e.g. p-well, p-substrate and deep n-well (PW-PSUB-DN). This is done for all occurring combinations of features. The doping profile for each layer is created using a superposition of Gaussian functions which is the same across the sensor. Under the assumption that the same implantation process is applied for different substrate resistivities, the doping levels of the features are added together along with the substrate doping, as was done in the H35DEMO simulation. Thus, the same profiles can be used to simulate other substrate resistivities. The profiles do not always agree very well with the original process simulation but are still in the acceptable range compared to the difference shown in Fig. 5.36. A comparison of the extracted and reproduced profiles can be seen in Fig. 5.37. They can be further improved by using fit functions.

Several different cross-sections were simulated during the design phase of the pixel. Fig. 5.35 shows the preliminary cross-section considered for an MPW submission in 2016. The DNWELL is biased at 1.75 V through VDN, NW and NISO which are all connected. The PSUB is biased at 0 V through VSS and PW. The NWELL that is biased through VDD is at 1.8 V. As for

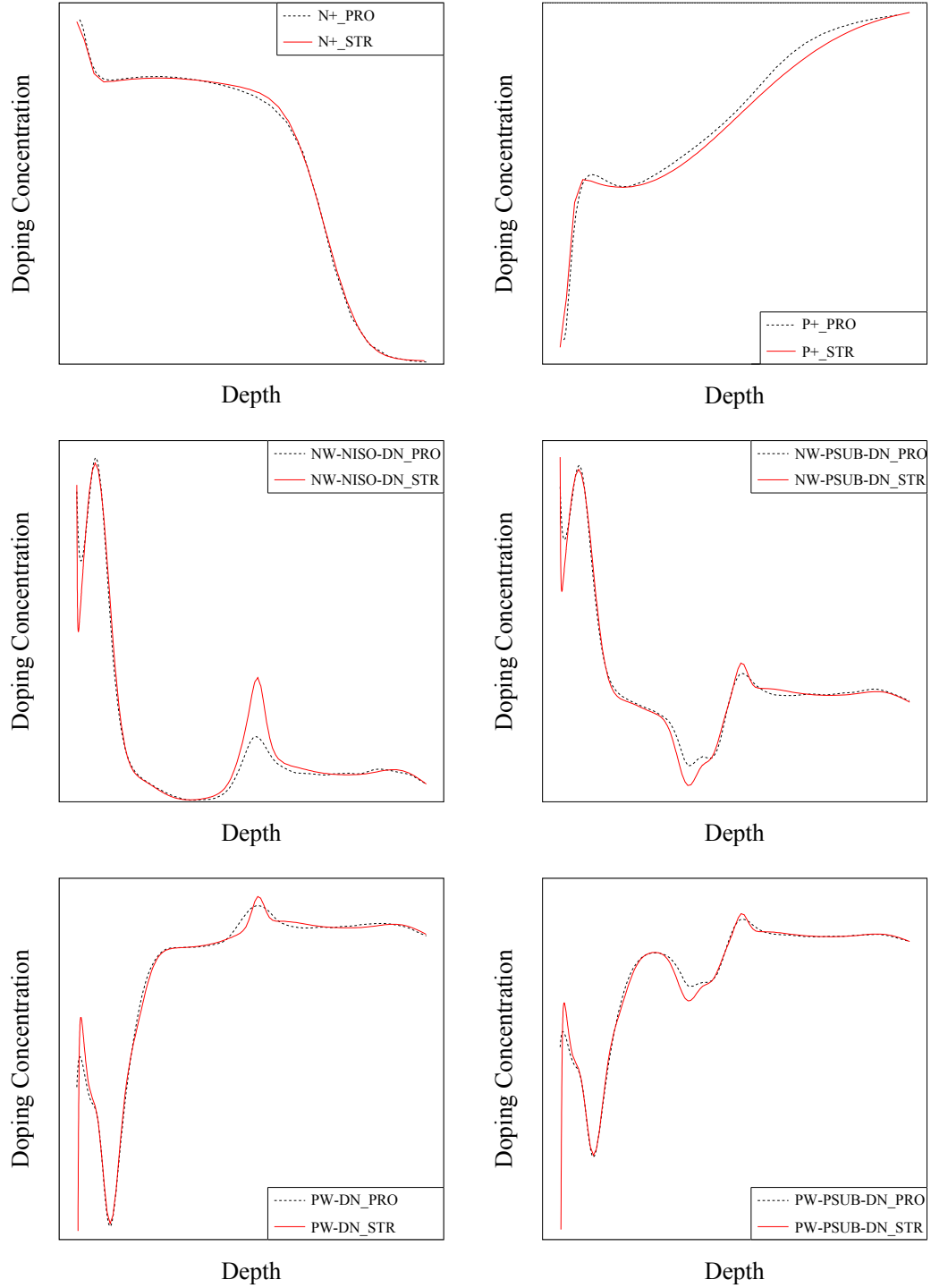


Figure 5.37: Comparison of doping profiles extracted from process simulation with structure simulation. Regions are indicated in the legend. Black dashed lines represent doping profile in the process simulation (PRO) and red solid lines are the reproduced doping profiles used for the structure simulation (STR). No scales are shown due to confidentiality.

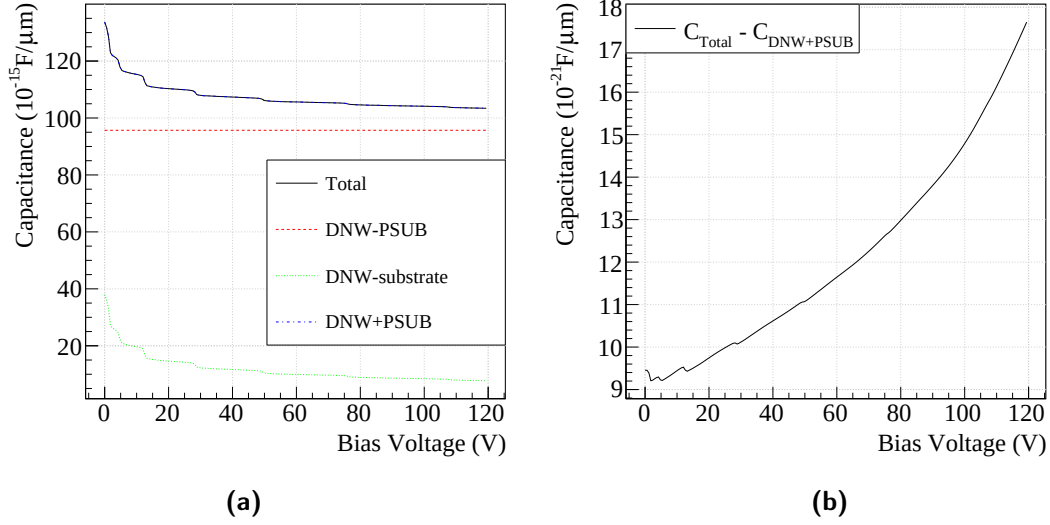


Figure 5.38: Contributions to total sensor capacitance. On the top plot, the graph labeled with “DNW+PSUB” is the sum of the contributions of capacitance of DNW-PSUB and DNW-substrate, which overlaps with the total capacitance, labeled with “Total”, extracted from the simulation. The difference of these two graphs is shown in the right plot.

the H35DEMO simulation the n^+ and p^+ implants of the n- and p-MOS are omitted.

The primary objective of this simulation was to determine the sensor capacitance which is needed as a design input. The 2D capacitance of the preliminary cross-section is shown in Fig. 5.38 (a). Each structure in the sensor contributes to the total sensor capacitance. However, the main contributions come from DNW-PSUB and DNW-substrate while other contributions are of the order of $10^{-6} \text{ fF}/\mu\text{m}$, as shown in Fig. 5.38 (b), and therefore negligible. The DNW acts as a collection electrode but also shields the structures within from the high voltage applied to the substrate. The depletion region between PSUB and DNW is constant as VSS stays zero and therefore the capacitance between DNW and PSUB is constant (ref. Eq. (3.25)) for any value of the substrate bias. The capacitance between DNW and the substrate decreases with increasing bias voltage since the depletion width follows Eq. (3.20) and is proportional to $\sqrt{V_{\text{bias}}}$.

Due to different issues the fabricated sensor did not turn out as intended, as shown in Fig. 5.39. The gap between NW and PW, where the ground and

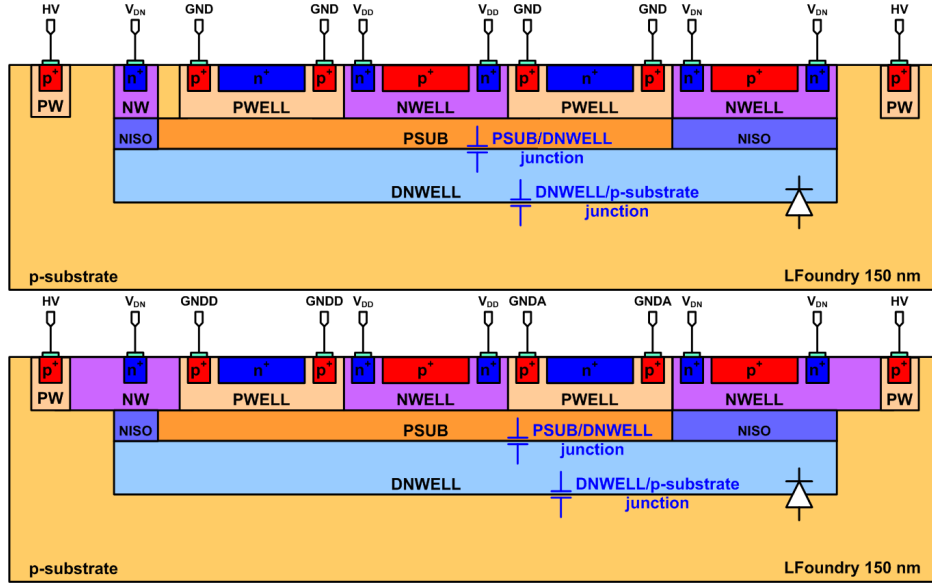


Figure 5.39: Cross-sections of the intended (top) and manufactured (bottom) pixel of the 2016 submission.

the HV are applied, respectively, is filled with extra n-implant and thus no gap is present. Therefore the sensor has a very low breakdown voltage at around -24 V and no further measurements could be conducted.

5.4.2 Capacitance

All simulations in this work were performed in 2D due to the large size of pixels compared to regular transistors and to save time and computing resource. In order to scale the capacitance of a 2D simulation to a 3D device, more considerations have to be made than to simply multiply by the width of the sensor. The capacitance comprises area and perimeter components as illustrated in Fig. 5.40. While the area capacitance is fixed by the process of the technology, i.e. the depth of the implants which influences the depletion region at p-n junctions (ref. Eq. (3.25)), the perimeter capacitance is given by the design, i.e. by design rules and by the layout. In this simulation the capacitance might be overestimated, which depends on the full pixel design, since the doping profiles do not consider STI regions where the implants are

shallower and the distance between PSUB and DNWELL is larger, as shown in Fig. 5.36.

In order to calculate the contributions of area and perimeter capacitance, two simulations with different geometries are used. Simulation 1 uses the normal geometry; simulation 2 uses a geometry with twice the length of the PSUB and accordingly scaled length of the DNWELL. In 2D, the perimeter capacitance c_p is the same for both geometries, while the area capacitance c_a varies with the length of the structures. The total capacitance c_{tot} is simply the sum of those two. For the capacitance between PSUB and DNWELL, the area capacitance of geometry 2 should logically be twice as much as of geometry 1:

$$\begin{aligned} c_{tot1} &= c_{a1} + c_p \quad \text{and} \\ c_{tot2} &= c_{a2} + c_p = 2 \cdot c_{a1} + c_p \quad . \end{aligned} \tag{5.5}$$

Therefore:

$$\begin{aligned} c_{a1} &= c_{tot2} - c_{tot1} \quad \text{or} \\ c_p &= 2 \cdot c_{tot1} - c_{tot2} \quad . \end{aligned} \tag{5.6}$$

Fig. 5.40 shows a simplified pixel layout seen from the top and the side. To scale up the area capacitance of the PSUB, the capacitance per width (area factor) from the simulation is multiplied by the PSUB width. The periphery capacitance calculated in Eq. (5.6), which includes two sidewalls per width, is multiplied by the width and by the length of the structure. Finally, some contributions from the corners might be considered as shown in Fig. 5.40. For the DNW-PSUB capacitance the area of a corner is assumed to be $(3\mu)^2\text{m}$. The same calculation is applied for the DNW-substrate capacitance. This is not very precise yet not substantial because PSUB makes the most significant contribution to the total capacitance.

The simulation is repeated in 10 V steps from 0 to -120 V . The results are compared with the Cadence simulation using the process design kit (PDK) v1.2.0 for the LF 150 nm process. The comparison is shown in Fig. 5.41. In the left plot the TCAD and the Cadence simulations have the same feature size: the PSUB is $(38.64\mu\text{m})^2$ and the DNW is $(45.14\mu\text{m})^2$. In the right plot the

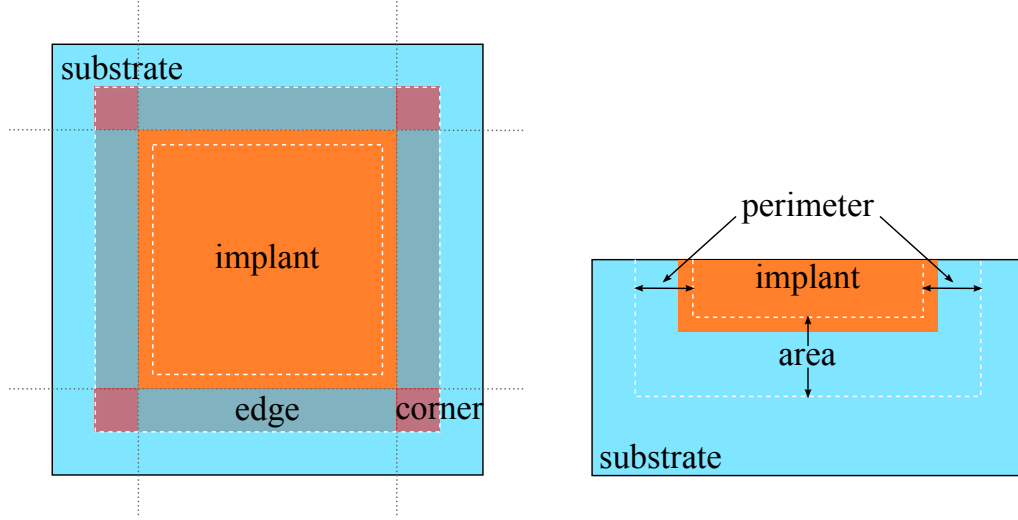


Figure 5.40: Schematic overview (left) and sideview (right) of a simple diode to demonstrate area and perimeter (edge and corner) components of the capacitance. The depletion region lies between the dashed lines.

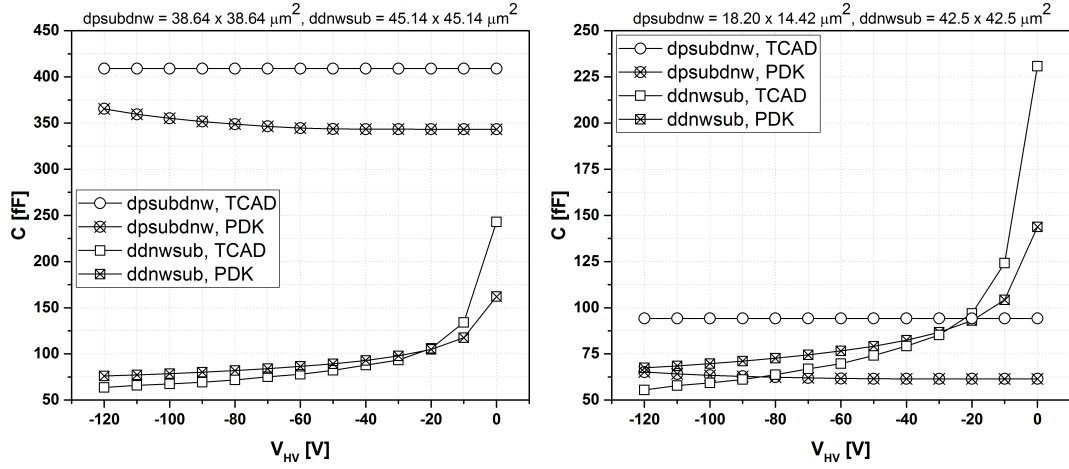


Figure 5.41: Comparison of simulated capacitance [123].

results of the TCAD simulation are extrapolated to different structure sizes with a rectangular PSUB of $18.2 \times 14.42 \mu\text{m}^2$ in a square DNW of $(42.5 \mu\text{m})^2$. The agreement is within 50–90 fF for both cases.

The first iteration of the estimations was not very precise. The corner estimation for PSUB-DNW capacitance was overestimated because in the simulation the width of the depletion region is approximately $0.5 \mu\text{m}$ wide at the

PSUB-NISO interface and $1\text{ }\mu\text{m}$ at the PW-NW interface (ref. Fig. 5.42), independent of the bias voltage which is shielded by the DNW. For the DNW-substrate capacitance the same assumption was used for PSUB, twice the area capacitance for the larger geometry. This is an overestimation because the length of DNW in geometry 2 is less than twice due to the NISO implants. The corner contribution with $3\text{ }\mu\text{m}$ edge length is underestimated since here the depletion width depends on the bias voltage and is approximately $6\text{ }\mu\text{m}$ for -60 V and $8\text{ }\mu\text{m}$ for -120 V . Therefore, another iteration of the calculation was made considering these criteria. For the PSUB-DNW interface a depletion width of $0.5\text{ }\mu\text{m}$ is assumed. For the DNW-substrate interface the depletion width is approximated using the two width at -60 V and -120 V . Following Eq. (3.20) the depletion width can be calculated as

$$d = 0.623 \cdot \sqrt{|V_{\text{bias}}|} + 1.18\text{ }\mu\text{m} \quad , \quad (5.7)$$

where $1.18\text{ }\mu\text{m}$ is the intrinsic depletion width at 0 V . The depletion region lies mostly in the p-substrate, but also partially within the DNW implant which is $\sim 1\text{ }\mu\text{m}$, therefore the offset in the calculation of $1.18\text{ }\mu\text{m}$, which is the intrinsic depletion width at 0 V , is neglected, because this is already considered in the perimeter calculation.

Fig. 5.43 shows that these considerations made only small improvements to the scaled PSUB-DNW capacitance compared to the discrepancy to the Cadence simulation because only the corner contribution was recalculated. For the components of the DNW-substrate capacitance however, it results in a total capacitance for which the trend is more comparable with the Cadence simulation shown in Fig. 5.41. It has a lower value at 0 V and a higher value towards -120 V than the old estimation.

5.4.3 Size of the PSUB

The 2016 MPW layout in Fig. 5.35 bottom has a full-width PSUB of $37.3\text{ }\mu\text{m}$. For further submissions based on this technology the question arises whether it is necessary to have a PSUB with a maximum extension inside a DNW or whether its contribution to the capacitance can be reduced by minimising its size. The PSUB does not have to cover the whole DNW. Sim-

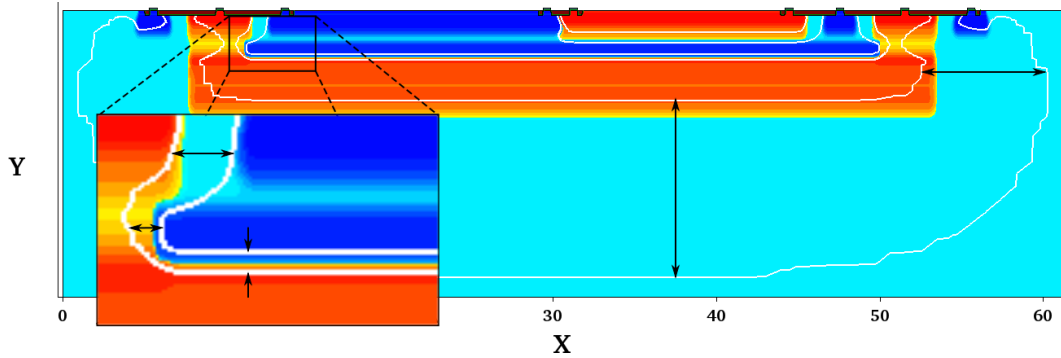


Figure 5.42: Cross-section displaying the depleted regions (white lines) at -120 V bias. The vertical scale is removed due to confidentiality.

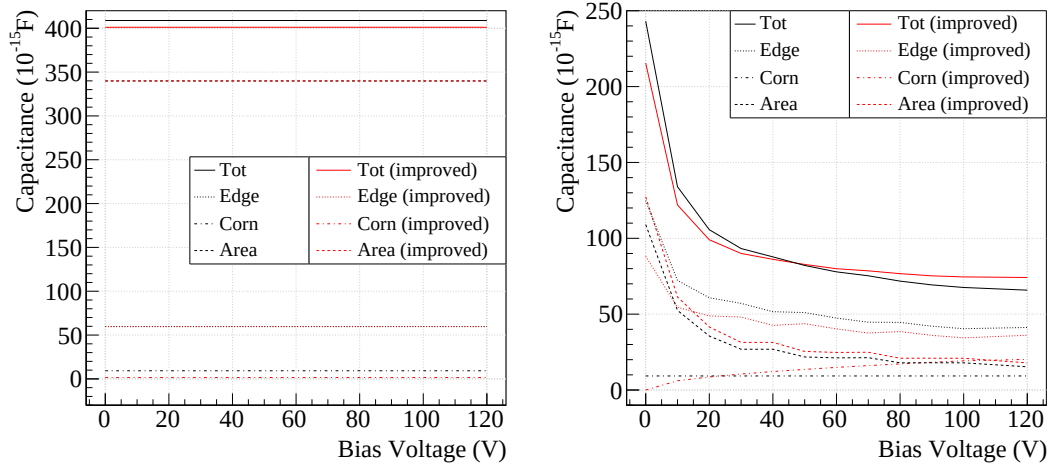


Figure 5.43: Contribution of area, edge and corner capacitance of PSUB-DNW (left) and DNW-substrate (right) capacitance. Black lines are from old calculations and red lines represent a better estimation. For the PSUB-DNW capacitance only the edge component was recalculated.

ulations were done to study the impact of the size of the PSUB layer on the capacitance. The minimum width for a PSUB layer has to be the same as the NW which is $14.7 \mu\text{m}$. However, as the PSUB has to be biased through the PW and it should provide a good isolation of the NW from the DNW, it has to be larger than the NW and have some overlap with the PWs as shown in Fig. 5.44. The width of the PSUB is studied with an overhang of 0, 0.5, 1 and $1.5 \mu\text{m}$ on each side.

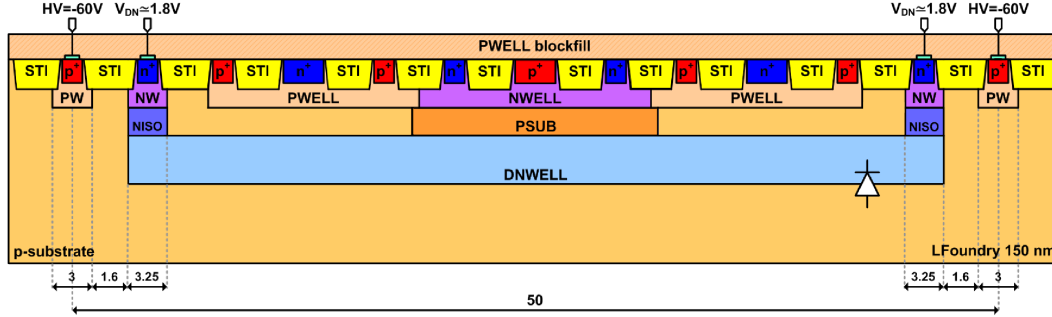


Figure 5.44: LF 2016 MPW layout with a small PSUB.

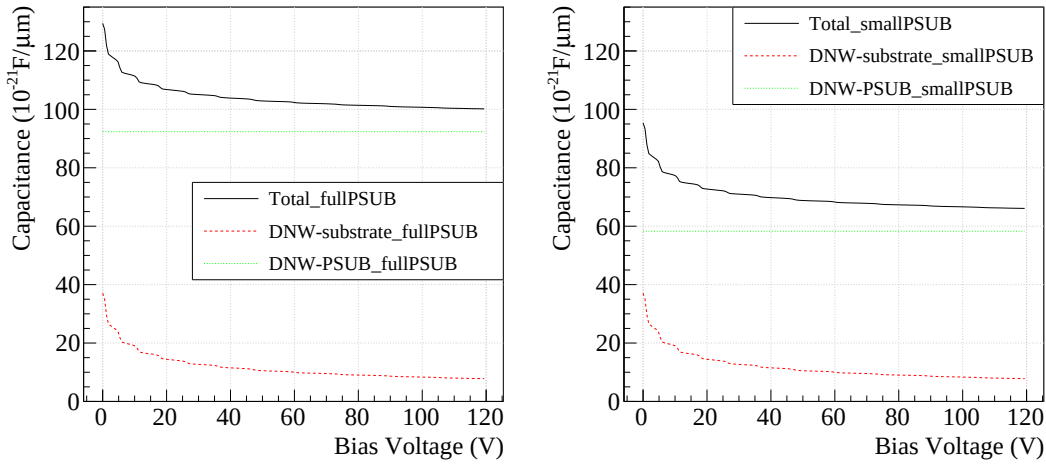


Figure 5.45: Contributions to sensor capacitance using full PSUB of $37.3 \mu\text{m}$ and reduced PSUB of $16.7 \mu\text{m}$ with $1 \mu\text{m}$ overhang on each side.

Fig. 5.45 shows the difference in capacitances and Fig. 5.46 shows the electron and hole current densities in the critical regions. In the absence of an overlap of the PSUB with the PW, the depletion regions at the junctions PW-DNW and NW-PSUB, indicated by the white boundary lines, are merged. A considerable amount of electron current flows between NW and DNW although a dense hole current flows between PW and PSUB as well. The overlap of PW and PSUB creates a separation between these depletion regions and reduces the electron current between the NW and DNW. Even though $0.5 \mu\text{m}$ is already sufficient, the high hole current density can be substantially reduced using a larger overlap of $1 \mu\text{m}$.

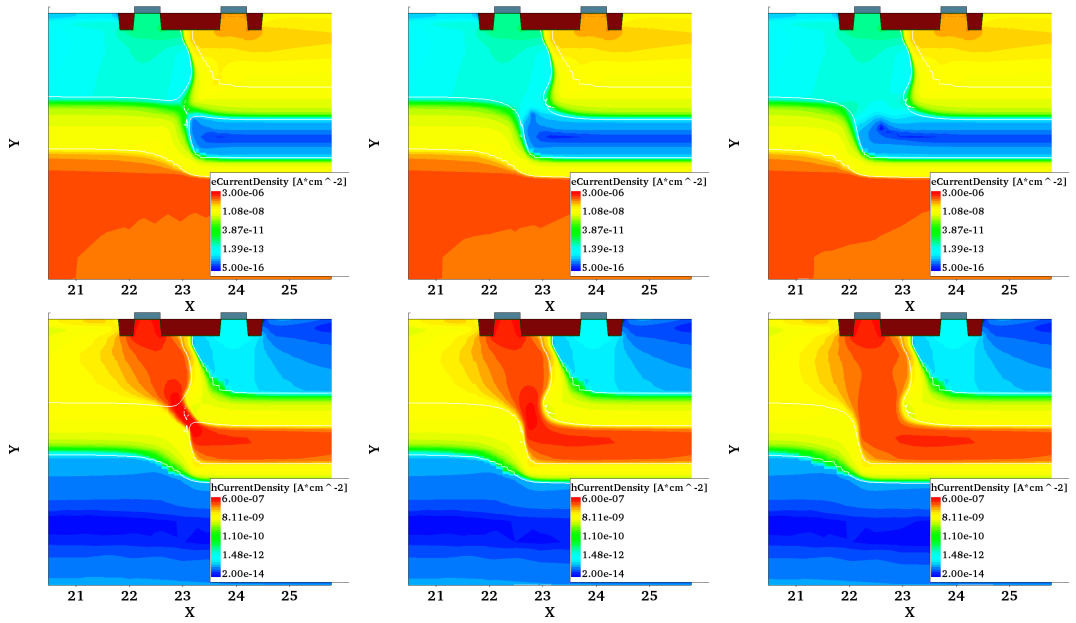


Figure 5.46: Electron (top) and hole (bottom) current densities of PSUB with 0 (left), 0.5 (middle) and 1 μm (right) of overhang. The white line indicates the boundary of the depletion region. The sensor has a resistivity of 10 Ωcm and is biased at $-120 V$. Vertical scale removed due to confidentiality.

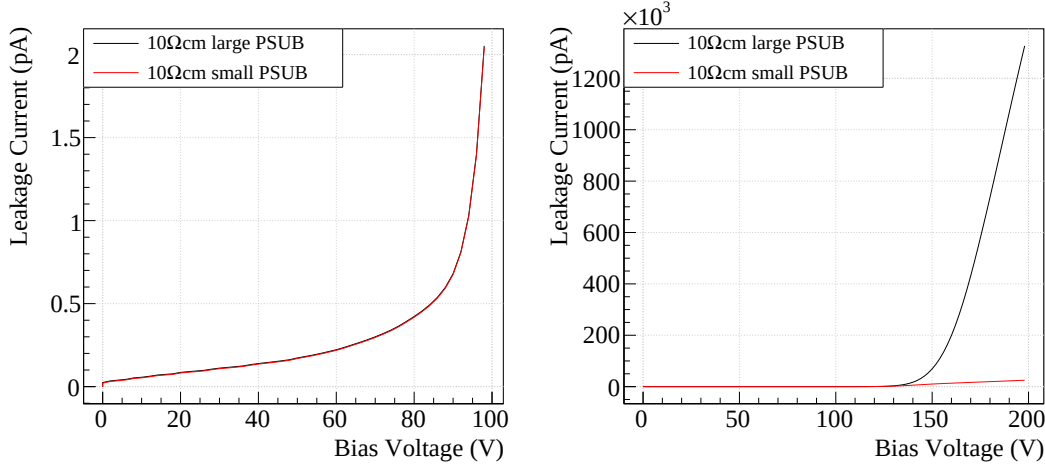


Figure 5.47: I-V curves (left) and breakdown voltages (right) of the 2016 MPW cross-section for a full and a small PSUB layer.

5.4.4 I-V Curve

An I-V plot is shown for the 2016 MPW cross-section in Fig. 5.47. The leakage current starts to increase at around -100 V but a breakdown only happens at around -150 V for the full PSUB and no breakdown is observed up to -200 V for the small PSUB. The white lines in Fig. 5.48 and Fig. 5.49 indicate the boundary of the depletion area. In case of the full PSUB the depleted area separates the regions NW and DNW at the NISO layer at a bias voltage earlier than -400 V, creating a punch-through. This is not yet the case for a layout with a small PSUB at the same bias voltage but the tendency is evident. This is a result of the weak doping of the NISO layer as previously shown in Fig. 5.37. However, it is not relevant for the expected operation voltages which will not go beyond -120 V.

5.4.5 Other Cross-sections

Another design was considered for a submission in 2017. Fig. 5.50 shows that two NW and two PW are present where the PWs carry analog and digital grounds, GNDA and GNDD, respectively. However, as the PWs are in the same p-type substrate, the grounds are shorted. Fig. 5.50 (b) shows a possible solution by separating the PW with a NW guard ring. Both the gap and the

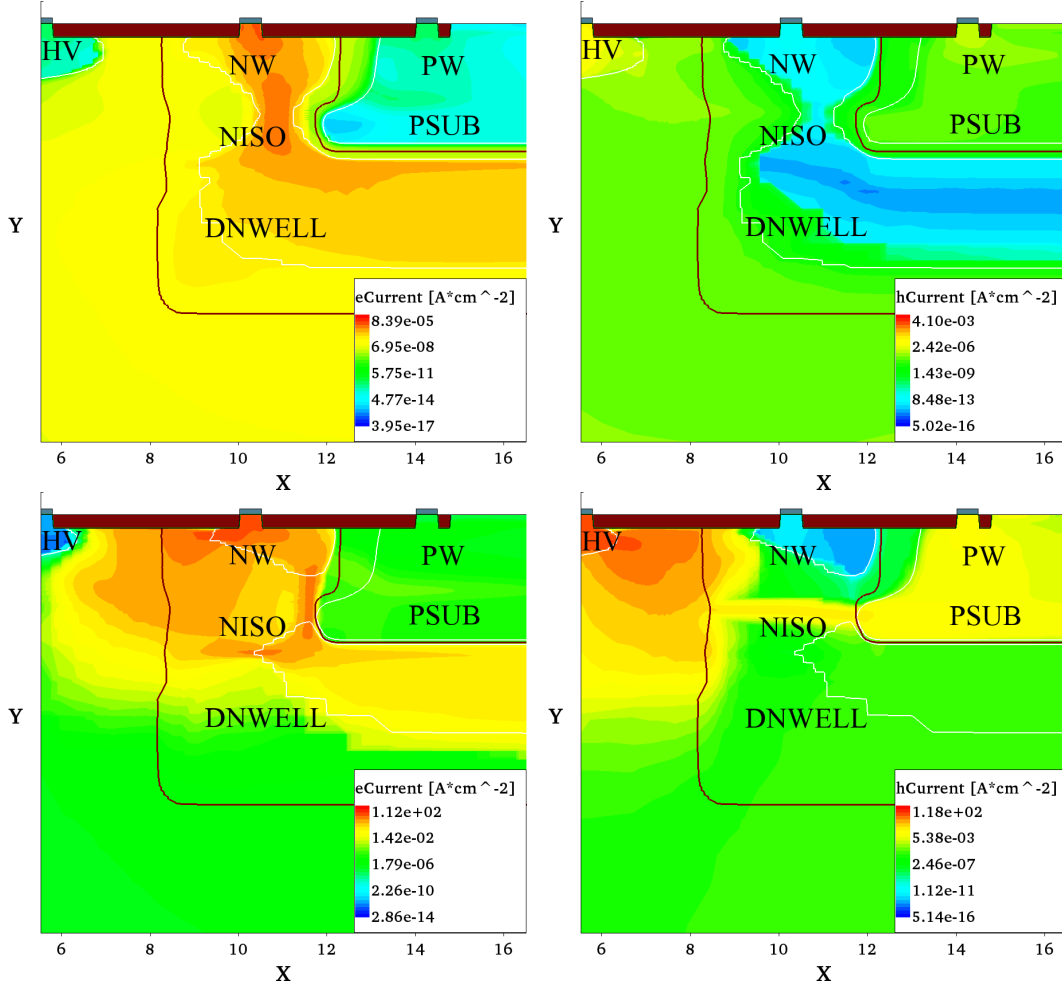


Figure 5.48: Electron (left) and hole (right) current densities of LFoundry 2016 MPW cross-section with full PSUB at -120 V (top) and -400 V (bottom) before and after breakdown, respectively. It can be clearly seen that the depletion region reached through the region between NW and NISO layers and separated them. The vertical scale is removed due to confidentiality.

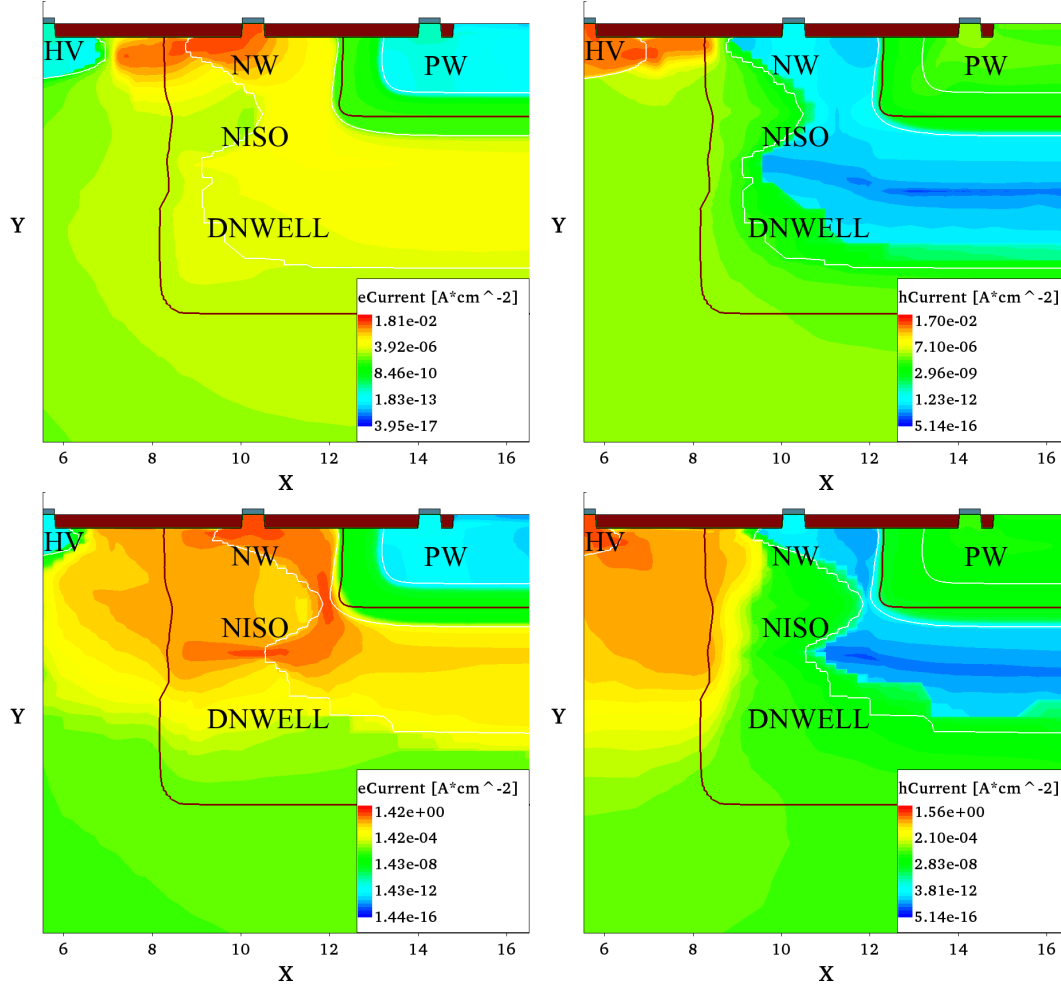


Figure 5.49: Electron (left) and hole (right) current densities of LFoundry 2016 MPW cross-section with small PSUB at -120 V (top) and -400 V (bottom) before and after breakdown, respectively. There is not yet a punch-through as in case of the full PSUB. The vertical scale is removed due to confidentiality.



(b)

147

It is not possible to measure the resistance between these two contacts in the simulation. However, it is assumed that they are shorted through the substrate and thus the resistance scales with the distance. The current through GNDA and GNDD for each cross-section is plotted against the sensor bias voltage as shown in Fig. 5.51, while the low voltages are fixed. The leakage current increases steeply at about -120 V , following the increasing leakage current of the p-n junction (ref. Fig. 5.47). Increasing the gap width and adding the extra NW both increase the current through the electrodes with bias voltage. In Fig. 5.52 a small fast signal of 1 V and 10 ns duration is injected into GNDA to simulate an electrical noise in an unbiased sensor. While there is no voltage cross-talk in GNDD for all layouts, there is current cross-talk for all gap widths in geometry 1. This decreases with increasing gap width between GNDA and GNDD. While for $1.5\text{ }\mu\text{m}$ the current is $15\text{ }\mu\text{A}$, for $5.5\text{ }\mu\text{m}$ it already reduces to only $2\text{ }\mu\text{A}$. However, geometry 2 with a separating NW does not have any current cross-talk at GNDD and is therefore more effective, despite the fact that this design is not permitted by the design rules. The additional NW adds another p-n junction with the PSUB and therefore the capacitance is higher as shown in Fig. 5.53. The electron and hole currents for different gap sizes are shown in Fig. 5.54.

The cross-section that was chosen for the MPW submission in September 2017 is shown in Fig. 5.55. To avoid a short circuit between GNDD and GNDA it was decided to have one common ground.

5.5 Measurements of H35DEMO

5.5.1 I-V Measurements

I-V measurements of the H35DEMO were done using a probe station in the clean room at Université de Genève. The setup is shown in Fig. 5.56. Three probe needles are connected to a Keithley voltage source where a compliance current limit was set to 200 nA . Two needles with the same ground are connected to VDDA (L64) which biases the DNTUB, and GNDA (L69) which biases the DPTUB. They correspond to VDD and GND in Fig. 5.4, respectively. The other needle is connected to HV. For the standalone matrices there

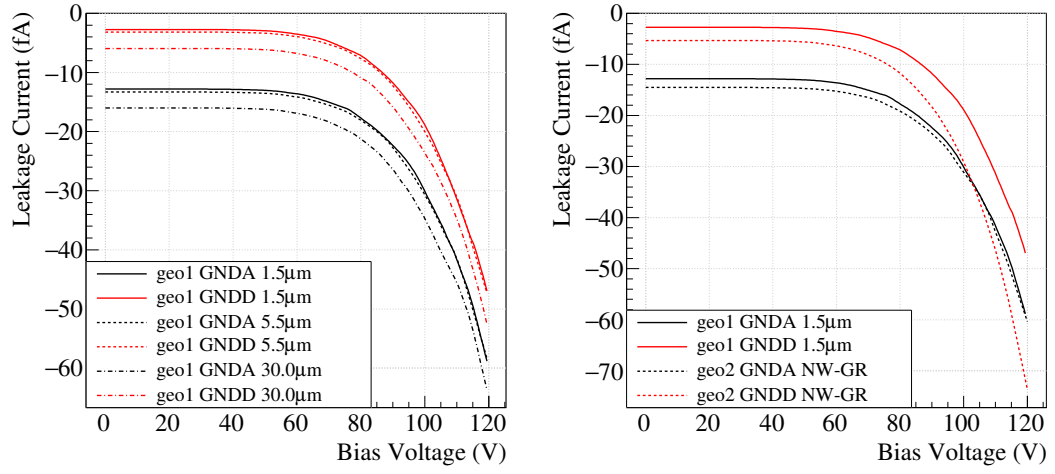


Figure 5.51: Currents through GNDA (black) and GNDD (red) vs bias voltage. The left plot shows geometry 1 with different gap widths. The right plot compares standard gap width of geometry 1 with geometry 2.

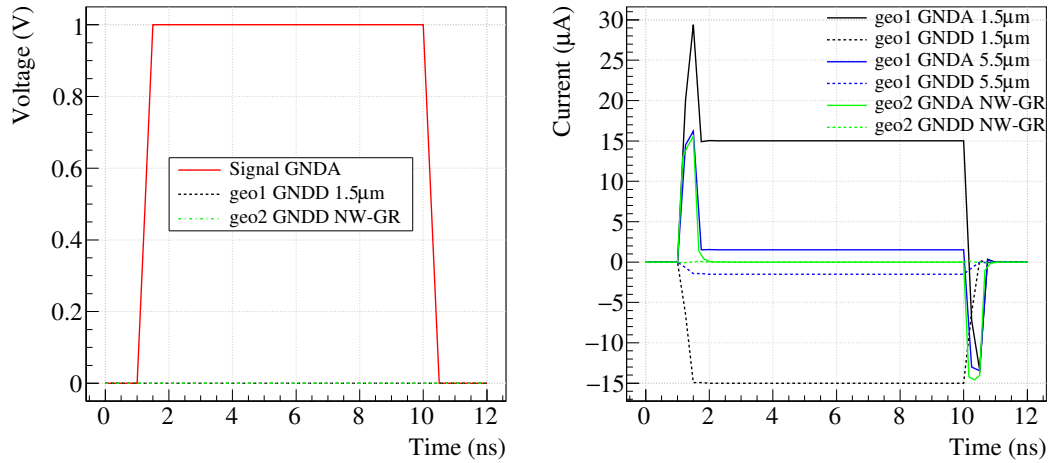


Figure 5.52: Response to a 10 ns signal of 1 V injected into GNDA. The left plot shows the voltage response in GNDD for standard cross-section 1 and 2. The right plot shows the current in GNDA (signal injection) and GNDD (response) for different gap width of geometry 1 and for geometry 2.

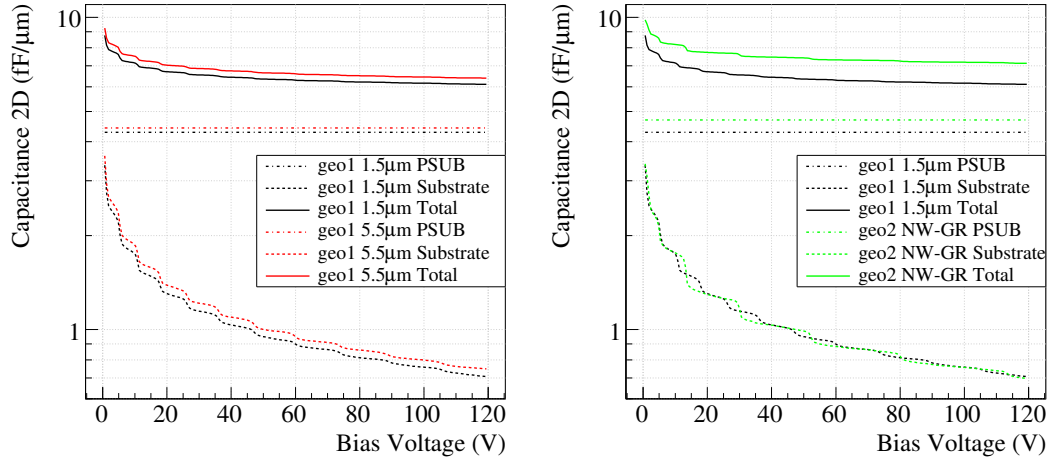


Figure 5.53: Comparison of capacitance for cross-section 1 with different gap width between the PWs (left) and comparison between layout 1 and 2 (right).

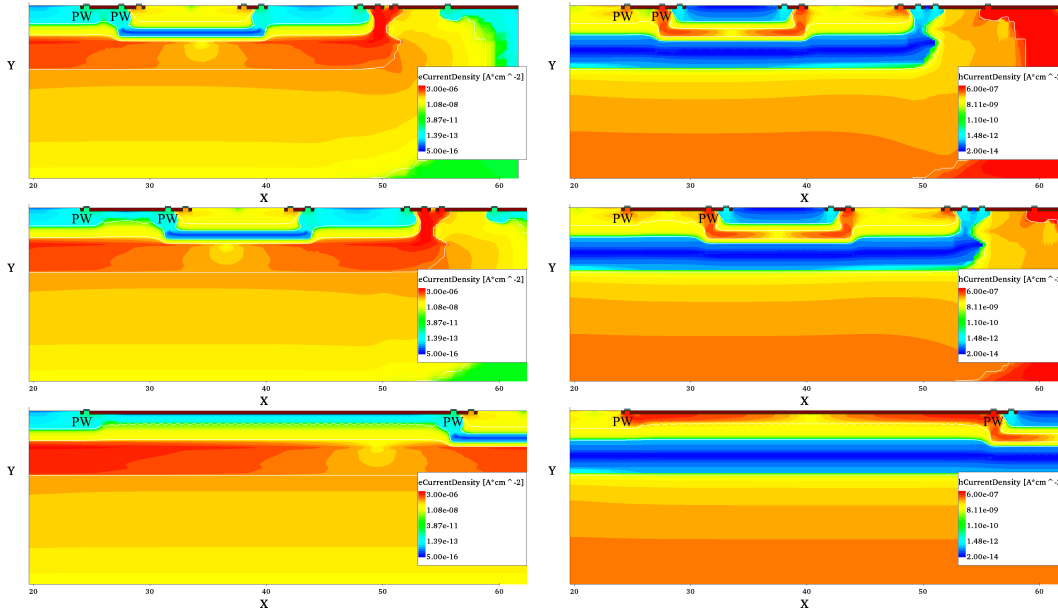


Figure 5.54: Electron (left) and hole (right) current densities for the region between two PWs at different distances. The default distance is 1.5 μm (top). Distances of 5.5 μm and 30 μm are shown in the middle and bottom row. The vertical scale is removed due to confidentiality.

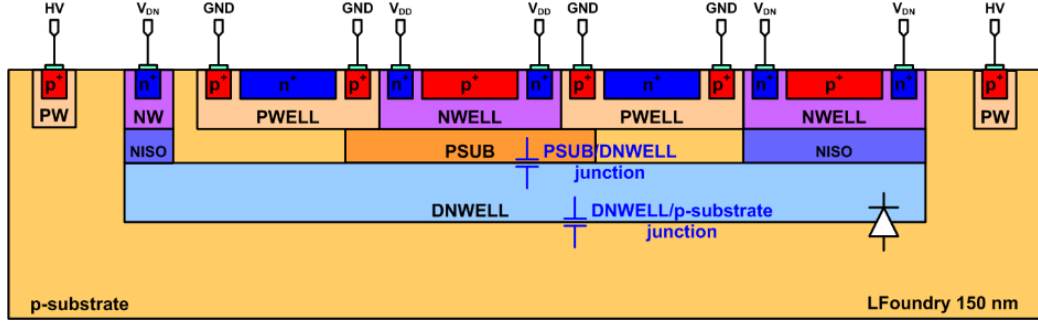


Figure 5.55: Submitted cross-section for RD50 MPW.

are three different HV pads to apply HV through the chip ring, the periphery ring and the pixel ring (L70, L71 and L72 for the NMOS), which are bias for the entire sensor/substrate, the individual matrix and individual pixels. The analog matrices can only be biased through the pixel ring. The H35DEMO has a test structure at the bottom for TCT measurements, as shown in Fig. 5.57. The pixels in this matrix do not have any circuits inside. The bias pads for this 3×3 matrix are TS_D0–5, whereas TS_D0 is the analog input, TS_D1 the analog output, TS_D2 is VDDA, TS_D3 is PW, TS_D4 is GNDA and TS_D5 is substrate HV. PW and GNDA are shorted whereas PW has protective diodes and GNDA has not [123].

Fig. 5.58 and Fig. 5.59 show measured the I-V curves. The Keithley measures the current through HV which is negative. That means the current through ground is complementary and thus positive. Therefore, the measured current is inverted to display the positive leakage current through VDD or GND. This is consistent with the simulations. Plots in Fig. 5.58 show the differences in the I-V characteristics where the bias voltages are the absolute values and the current is inverted. In plot (a), connecting VDDA shifts the total measured leakage current. Further measurements are all taken without VDDA but with only HV and GNDA, where DNTUB is biased through the DPTUB in a forward bias of the p-n junction created by DNTUB and DP-TUB. Plot (b) shows that there is no difference in the leakage current between the different HV pads used as expected since all bias connections are shorted through the bulk. In (c) the leakage current of different matrices is compared. The NMOS matrix has the smallest leakage current. One has to bear in mind

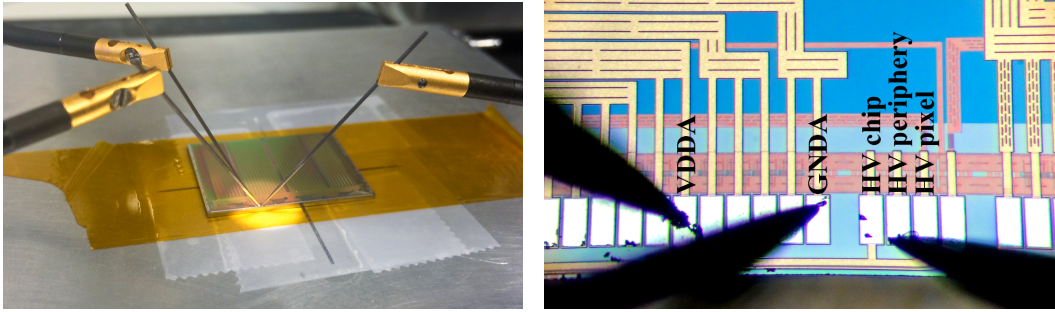


Figure 5.56: H35DEMO I-V measurement setup using a probe station (left) and the corresponding probing pads (right). Photos by [145], modified.

that the number of pixels of 16×300 are the same for the NMOS and CMOS matrix while each of the analog matrices has 23×300 pixels and therefore expected higher leakage current. (d) shows that the matrices on wafer 19, which is the $1000 \Omega\text{cm}$ substrate, all have a very early breakdown. This might be caused by the “Rise-and-Flatten” effect [143, 144] which was not understood at the time of the measurement. The comparison to other resistivities can be seen in Fig. 5.59, measured using different sensors at Université de Genève (UniGe) and UoL. Shown in (e) are the I-V curves for sensors of the same resistivity but at different positions on the wafer. Sensor 1 and 22 are compared. Finally, the I-V curve of the bottom test structure, shown in Fig. 5.57, is also measured. However, the sample was specifically taken from the edge of a wafer which means it is not a full sensor. The I-V measurement was made to verify its functionality for TCT measurements. The leakage current is measured between each power pad and HV and the results is shown in plot (f). Some curves are rather jumpy which was probably due to the sensor being an edge piece, but the overall trend is clear. It is worth noting that the breakdown voltage is about 10–20 V larger than for other matrices on the same resistivity wafer. The I-V curves of the NMOS matrix for sensors of different resistivities are shown in Fig. 5.59. The left and right plots correspond to measurements done at UniGe and UoL, respectively. For the measurement in Liverpool the bias is slightly different. VDDA is connected to ground instead of GND while the chip ring pad (L70) is used to bias the whole substrate. Both results agree well even in the case of the $1000 \Omega\text{cm}$ wafer which has a very early breakdown.

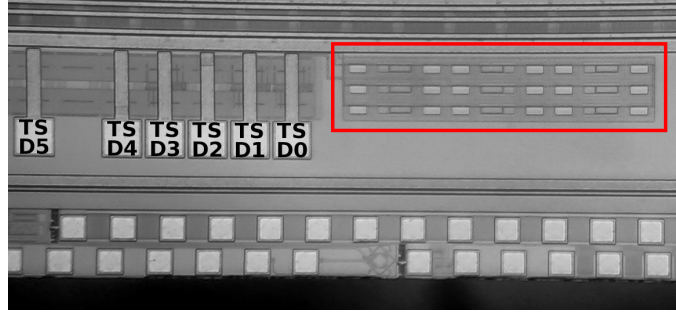


Figure 5.57: H35DEMO test structure with bias pads at the bottom of the sensor.

5.5.2 Comparison with Simulation

5.5.2.1 I-V Characteristics

For this comparison a simulation with periodic boundary conditions is used. The difference is shown in Fig. B.5.

Fig. 5.60 left shows the ratio between simulated and measured current, and right shows the temperature-corrected ratio. The temperature in the simulation was defined as 300 K while the temperature in the clean room T_R varies with day time. Each I-V measurement also recorded the temperature of the probe station, whereas the temperature of the sensor itself was not measured directly. The temperature distribution is shown in Fig. 5.61. An average of $T_R = (20.76 \pm 0.68)^\circ\text{C}$ is used. The leakage current from the simulation is multiplied by the width of DNTUB of $30\text{ }\mu\text{m}$ and by the total number of pixels in the nMOS matrix (16×300). However, the effect of the perimeter coming from the depth of the DNTUB is not fully taken into account in this calculation.

Bludau *et al.* [146] and O'Donnell and Chen [147] described the temperature dependence of the band gap of silicon. Bludau *et al.* use a simple approach of

$$E_g(T) = A + BT + CT^2 \quad (5.8)$$

with the fit variables given in Table 5.4. O'Donnell and Chen use

$$E_g(T) = E_g(0\text{ K}) - S\hbar\omega(\coth(\hbar\omega/2k_B T) - 1) \quad (5.9)$$

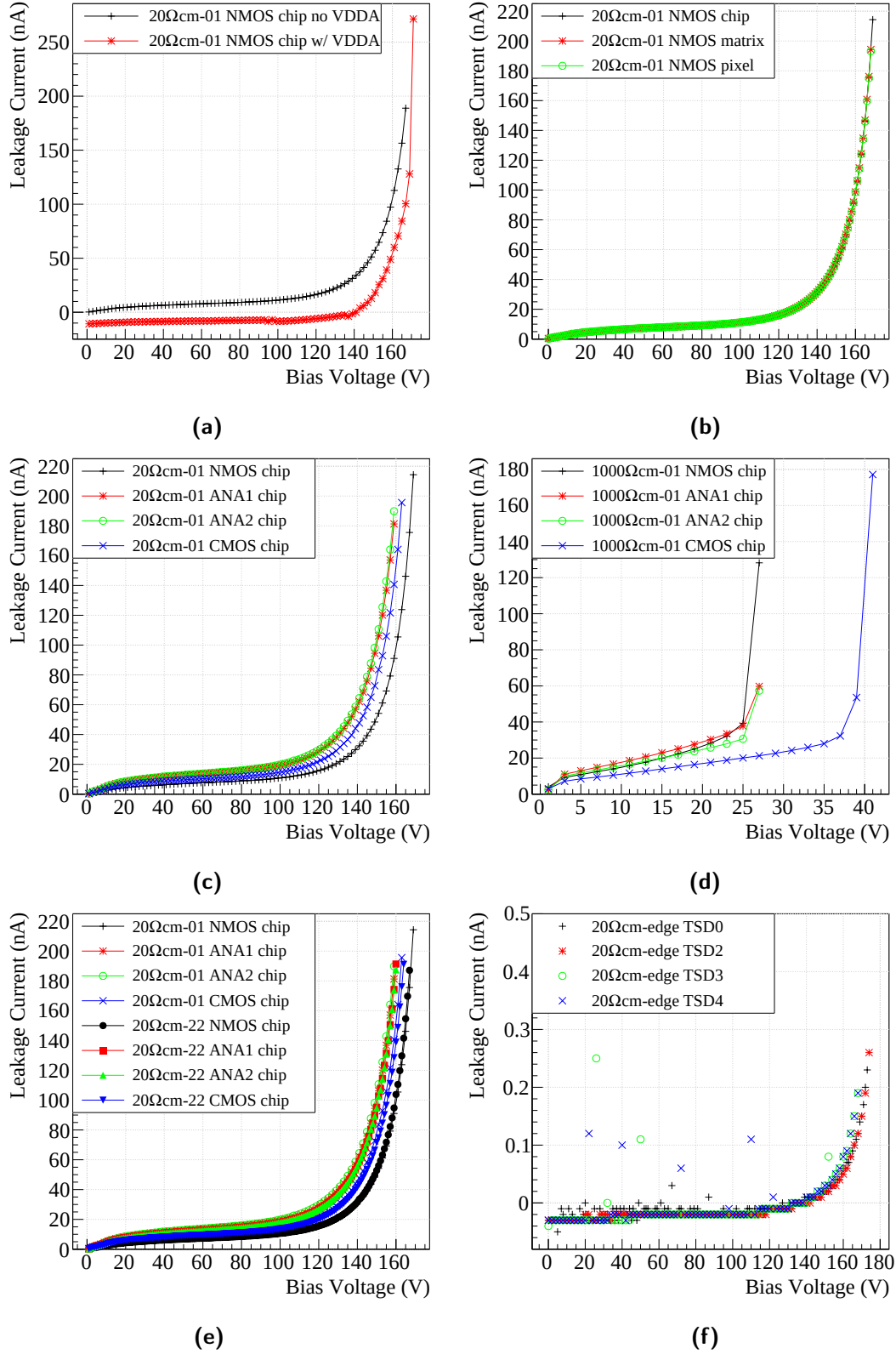


Figure 5.58: I-V measurements of different H35DEMO sensors. Except for (a), all measurements are done without VDDA connection.

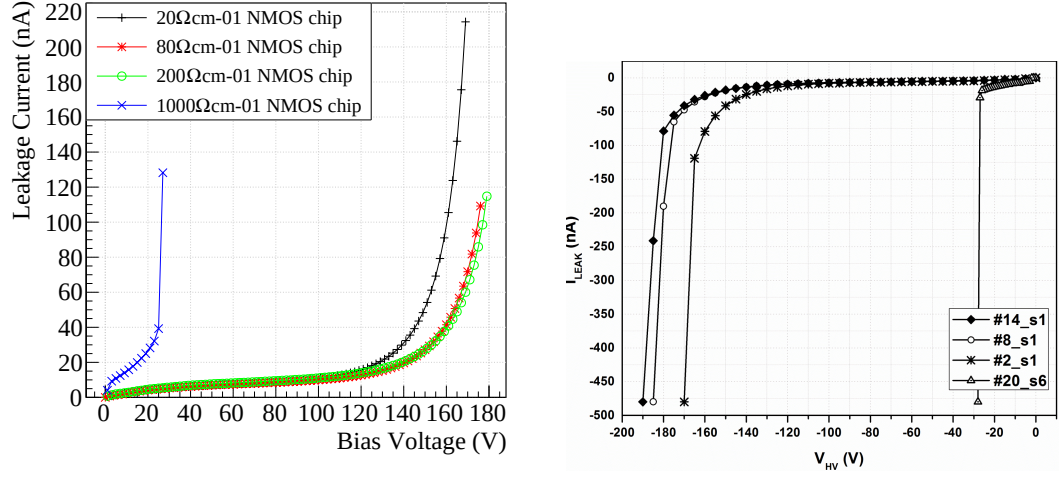


Figure 5.59: I-V measurement of the NMOS matrix on sensors of each resistivity. The left plot shows the sensors measured at UniGe, using absolute bias voltage and leakage current through VDN. The right plot shows the sensors measured at UoL [123], with the current through HV plotted against the actual bias voltage. #2, #8, #14 and #20 correspond to wafers with 20, 80, 200 and 1000 Ωcm resistivity.

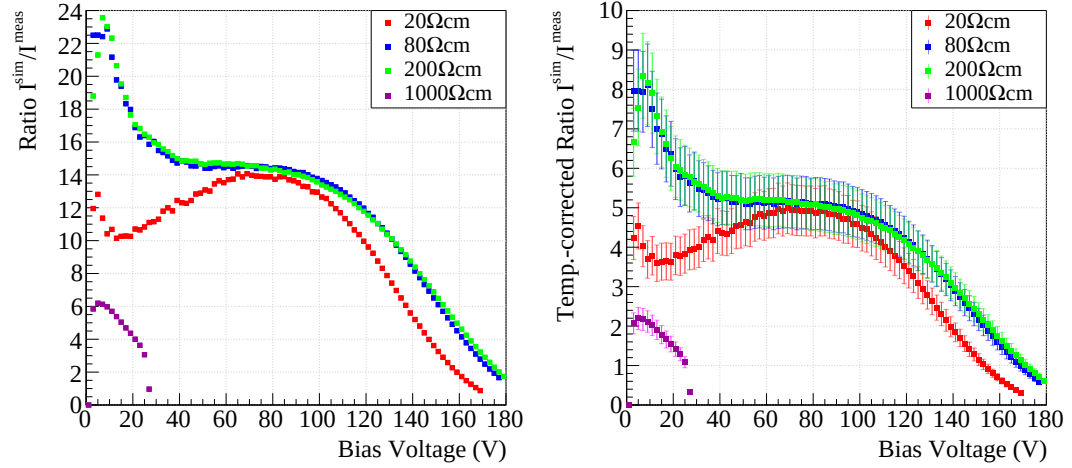


Figure 5.60: Ratio of simulated and measured leakage current (left) and the same plot corrected by the temperature difference and thus width of the band gap (right).

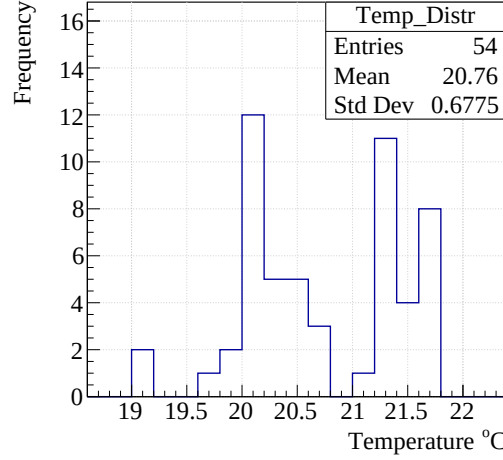


Figure 5.61: Distribution of temperature recorded during the I-V measurements.

with the same parameters valid for the whole range of $0 < T \leq 300$ K: $E_g(0 \text{ K}) = 1.17 \text{ eV}$, $S = 1.49$, and $\hbar\omega$ the average phonon energy of 25.5 meV . This fit with different parameters also describes the temperature dependence of band gaps of other semiconductor materials.

Range (K)	A (eV)	B (10^{-5} eV/K)	C (10^{-7} eV/K^2)
$0 < T \leq 190$	1.170	1.059	-6.05
$150 \leq T \leq 300$	1.1785	-9.025	-3.05

Table 5.4: Values of fit parameters used by Bludau *et al.* in Eq. (5.8).

Both approaches are shown in Fig. 5.62. Bludau *et al.* are able to fit data with a precision of 0.2 meV . However, the data have an uncertainty of 1 meV due to uncertainties in the energy of the exciton and the transversal optical phonon. Because the paper only shows the fit and does not have a plot comparing with data, the work by O'Donnell and Chen was also taken into account. The difference between the fits is shown in Fig. 5.62 right. They agree within 1.1 meV which is within the quoted total uncertainty of 1.2 meV of the fit by Bludau *et al.*. Despite being more recent, Fig. 1 (c) in [147] suggests that the discrepancies between the O'Donnell and Chen fit and some data ranges would be better described by Bludau *et al.*.

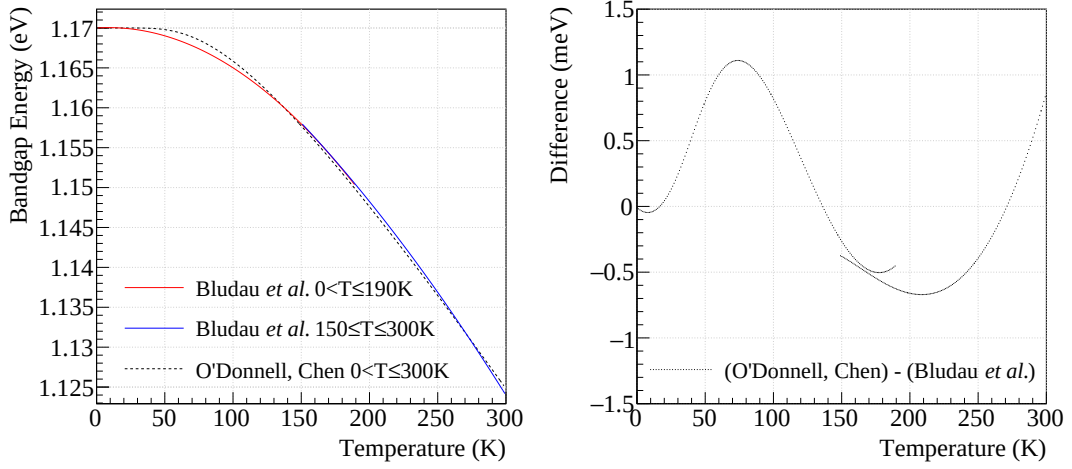


Figure 5.62: Left: Temperature dependence of bandgap energy of silicon fitted by Bludau *et al.* [146] and O'Donnell and Chen [147]. Right: difference of the two fits.

In the Shockley equation (Eq. (3.31)), $\exp(qV_b/k_B T)$ tends to zero for low bias voltages as in this case and thus the term in the bracket becomes independent from temperature. The ratio r of the simulated current can then be described as

$$\begin{aligned}
 r = \frac{I^{\text{sim}}}{I^{\text{meas}}} &= \frac{I_0^{\text{sim}}}{I_0^{\text{meas}}} = \frac{n_{\text{p0}}^{\text{sim}} T_{300}^{3+1/2} \exp\left(\frac{-E_g(300 \text{ K})}{k_B T_{300}}\right)}{n_{\text{p0}}^{\text{meas}} T_R^{3+1/2} \exp\left(\frac{-E_g(T_R)}{k_B T_R}\right)} \\
 &= \frac{n_{\text{p0}}^{\text{sim}}}{n_{\text{p0}}^{\text{meas}}} \left(\frac{T_{300}}{T_R}\right)^{3+1/2} \exp\left(-\frac{1}{k_B} \left(\frac{E_g(300 \text{ K})}{300 \text{ K}} - \frac{E_g(T_R)}{T_R}\right)\right) .
 \end{aligned} \tag{5.10}$$

$E_g(300 \text{ K})$ and $E_g(T_R)$ can be taken from the Bludau *et al.* fit as $(1.1240 \pm 0.0012) \text{ eV}$ and $(1.1256 \pm 0.0012) \text{ eV}$, respectively. The uncertainty due to room temperature variations and the uncertainty from the fit are independent measures, thus the temperature-dependent uncertainty is calculated as

the quadratic addition of these contributions

$$\Delta r_T = \sqrt{\left(\frac{\partial r}{\partial T_R} \Delta T_R\right)^2 + \left(\frac{\partial r}{\partial E_g(300\text{ K})} \Delta E_g(T)\right)^2 + \left(\frac{\partial r}{\partial E_g(T_R)} \Delta E_g(T)\right)^2}. \quad (5.11)$$

The temperature-dependent ratio yields $r_T = I^{\text{sim}}/I^{\text{meas}} = 2.82 \pm 0.36$.

The relationship $n_{p0} = n_i^2/N_A$ for Eq. (3.24) indicates that the leakage current does not depend on the doping concentration of the DNTUB but only on that of the substrate. The doping concentrations of the substrate used for the simulation are given in Table 5.1. However, there are small discrepancies between different resistivity calculators of about up to 3 %. The silicon wafers delivered by the foundry also did not have the exact resistivity. Instead of 20, 80, 200 and 1000 Ωcm they are given as 20, 80–100, 200–300 and 1000–2000 Ωcm . When this is taken into account, the total uncertainty increases to that given in Table 5.5.

Resistivity (Ωcm)	20	80	200	1000
Δr_{tot}	0.37	2.35	1.97	1.49

Table 5.5: Total uncertainty on the ratio of simulated and measured leakage current for different resistivities.

Despite accounting for these obvious uncertainties, the simulation still predicts a leakage current that is around five times the measured. In the simulation the current depends on the physics models applied. Fig. B.2 shows that omitting the models necessary for breakdown analysis the leakage current becomes lower which is most significant for lower resistivities by up to a factor of two. In addition, in TCAD an ideal pixel is simulated without any impurities except for the one type of dopant, while in reality the concentrations of impurities are not known, the sensor is much larger and has more complex structures like metal layers and embedded electronics.

5.5.2.2 Depletion Depth

The depletion depth is compared with the eTCT measurements and with calculations in Fig. 5.63. The data are from eTCT measurements on the bottom test structure obtained by E. Cavallaro from IFAE using top bias, as

described in [148]. They are compared with the default simulated pixel biased from the top. As can be seen, while the low resistive substrates agree well with the simulation, for the high resistivity of $1000\ \Omega\text{cm}$, the discrepancy between data and simulation is quite large. Efforts to obtain the depletion region using other properties like electric field or hole concentration yielded the same results. Variation of the resistivity up to $2000\ \Omega\text{cm}$, whose value is given as the technical uncertainty of fabrication, showed a slight improvement, but the depletion depth still differs substantially from the measured value.

The simulation is compared with the calculation which is done using Eq. (3.20). Because $N_D \gg N_A$ at the p-n junction, Eq. (3.19) is negligible. The values for N_D and N_A are the same as used in the simulation for the substrate and the DNTUB given in Table 5.1 and Table 5.2. The value used for the permittivity ε is $(11.7 \pm 0.2) \cdot \varepsilon_0$ [149], whose uncertainty of ± 0.2 is negligible and whose value can be taken to be constant because the doping of the substrate is low enough [150]. The intrinsic charge carrier density n_i at 300 K is $1.45 \cdot 10^{10}\ \text{cm}^{-3}$ [72]. Since the calculation assumes an ideal diode, it is compared with a back-side-biased pixel of $100\ \mu\text{m}$ thickness. In the calculation the position of the junction at approximately $8\ \mu\text{m}$ ($9\ \mu\text{m}$ for resistivities of $1000\ \Omega\text{cm}$ or higher), taken from the junction line yielded in TCAD shown in Fig. 5.64, is taken into account. The results from the TCAD simulation of a back-side-biased pixel agree well with the calculation. For the $1000\ \Omega\text{cm}$ resistivity the sensor thickness is not considered in the calculation whereas in the simulation the sensor of $100\ \mu\text{m}$ is fully depleted from $-80\ \text{V}$ on. The depletion depth does not reach $100\ \mu\text{m}$ because of the back side p-implants of about $1\ \mu\text{m}$.

It is noticeable that the depletion depth for the back-side-biased sensor is much larger and has a similar depth as obtained from eTCT measurements. Therefore these values are compared and shown in the middle plot in Fig. 5.63.

5.5.3 Test Beam Measurements of the Analog Matrices

So far four test beam campaigns have taken place with the H35DEMO sensor. These took place at SPS at CERN in October 2016 and October 2017, and at Fermi National Accelerator Laboratory (FNAL) spring 2017 and spring 2018. The combined results of the first three campaigns can be found in [144]. In the 2016 SPS test beam only two samples were tested on a prototype Control

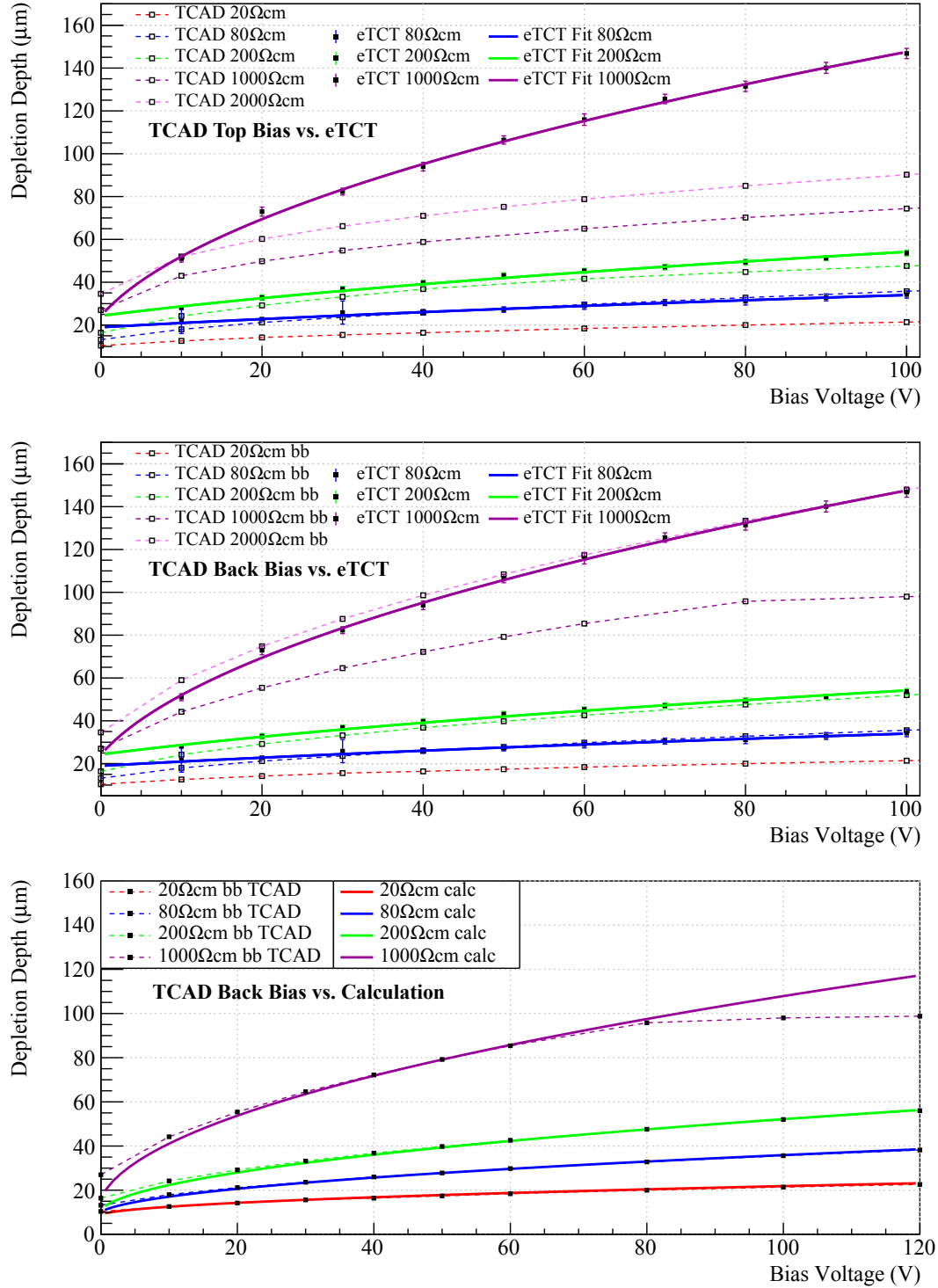


Figure 5.63: Comparison between measured depletion depth, top-side-biased simulation (top), and back-side-biased simulation (middle). Comparison between simulation and calculation (bottom). Data are obtained by E. Cavallaro (IFAE) and back-side-biased simulation has a sensor thickness of 100 μm .

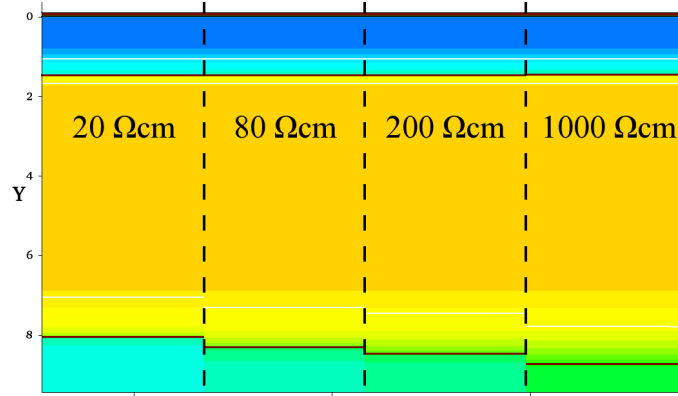


Figure 5.64: Junction line positions for different resistivities at around $y = 8 \mu\text{m}$.

and Readout Inner-Tracking Board (CaRIBou) system. This section focusses on the 2017 FNAL campaign which is the first one with a large number of samples.

In 2017 it was decided to have a test beam at the Fermilab Test Beam Facility (FTBF) in Batavia, Illinois, United States. It took place in the MT6.2 area. The MTest beamline delivers 120 GeV protons with ~ 4 spills per minute and $1 \cdot 10^6 - 3 \cdot 10^6$ particles per spill. Data were taken with the FE-I4 telescope [151] shipped from Geneva. The FE-I4 telescope has six planes, three in each arm of which the middle plane is rotated around the beam axis by 90° as can be seen in Fig. 5.66. The horizontal position of each plane can be adjusted within each arm. The angle with respect to the y-axis can also be changed by using a different base for the plane. The telescope planes are made of spare planar IBL FE-I4 modules as described in Section 2.2.1. Each module is wirebonded onto a flexible printed circuit board (PCB) which provides electrical connections like high voltage, low voltages, Ethernet and temperature reading. The hitbus signal of the first and last planes together provide a coincident trigger for the telescope. The DUT box between the two arms has a base plate which has threaded holes to fix the DUT and can be cooled by a chiller. The box itself is made of polystyrene for thermal insulation. During cooled operation the box is flushed with dry air. The temperature and humidity inside the box is constantly monitored. The DUT plate can be moved in x- and y-positions using two stages. There is a Wiener Mpod mini crate providing up to 24



Figure 5.65: Aerial view of the Fermilab Test Beam Facility [152].

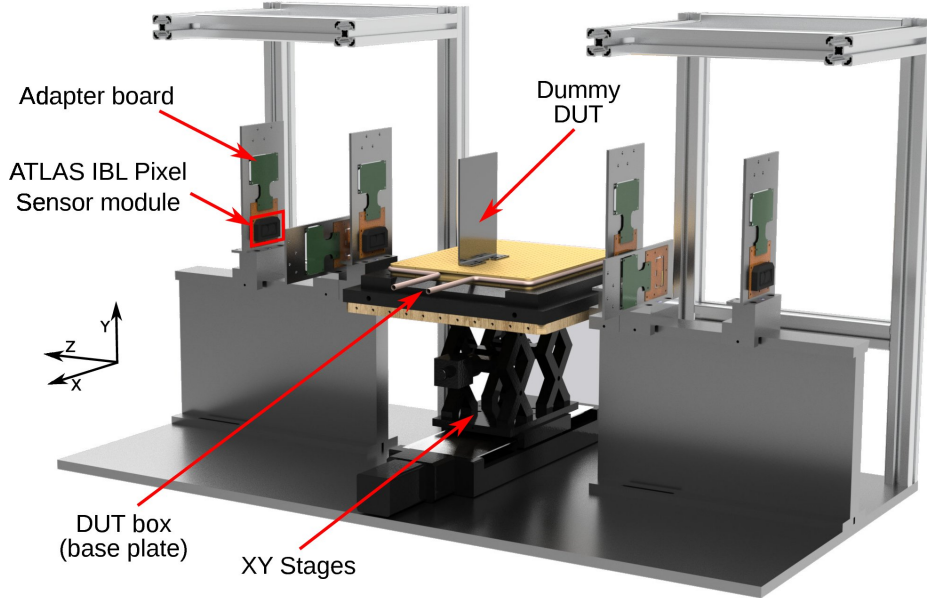


Figure 5.66: Computer generated image of the FE-I4 telescope [151].

low voltages and 8 high voltages. DUTs with an FE-I4 chip can be read out together with the telescope planes by the data acquisition (DAQ) system consisting of a HSIO board and the RCE system [107].

5.5.3.1 Samples

The goal of the test beam was to measure the two capacitively coupled analog matrices, adding more data points to the two prototypes in the previous test beam. The samples used for the test beam were prepared in the clean

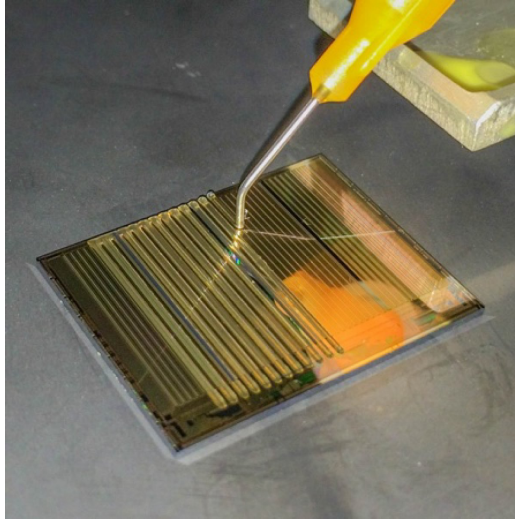


Figure 5.67: Dispensing Araldite[®] onto a H35DEMO sensor using a syringe with the flip-chip machine [144].

room at UniGe. The I-V characteristics of the analog matrices were measured using an automatic probe station by Cascade Microtech [153] and can be found in Appendix C.1. The recorded I-V curve shows the current through the GNDA connection. Apart from assemblies of standard thickness of about $760\text{ }\mu\text{m}$ in each resistivity, also some sensors of $200\text{ }\Omega\text{cm}$ thinned to $100\text{ }\mu\text{m}$ and and of $1000\text{ }\Omega\text{cm}$ thinned to $100\text{ }\mu\text{m}$ with back side process were tested. Each sensor that went into the test beam was tested, cleaned, and glued with Araldite[®] onto an FE-I4 readout sensor using an ACC μ RA100 flip-chip machine by Smart Equipment Technology (SET) [154] which has a precision of $1.5\text{ }\mu\text{m}$ and a parallelism of $1\text{ }\mu\text{rad}$ or less than 100 nm across the assembly. The thickness of the glue layer was $4\text{--}5\text{ }\mu\text{m}$. The samples were shipped to Fermilab and wirebonded onto the chip board (Fig. 5.68) of the CaRIBou system [155] on site.

5.5.3.2 Results

Measurements with a range of threshold and bias values were performed. The data are then reconstructed using the Proteus reconstruction framework [156, 157] which is based on the reconstruction framework Judith [158]. As

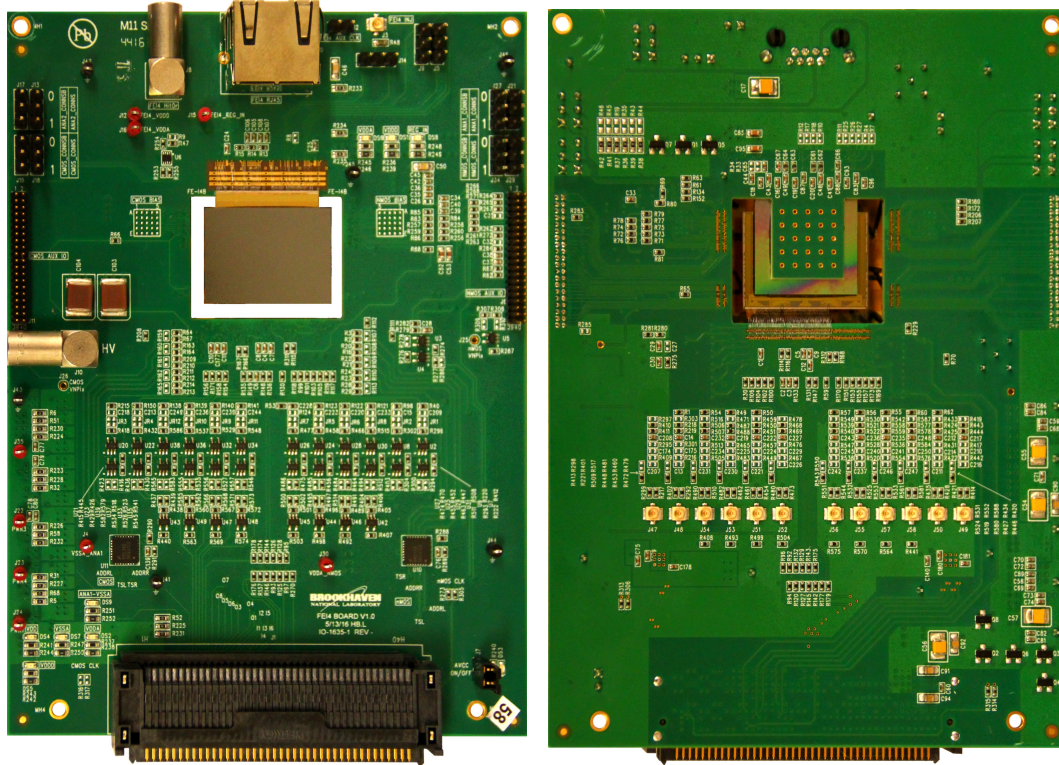


Figure 5.68: CaRIBou chip board with a H35DEMO assembly. The FE-I4 readout ASIC is glued and wirebonded onto the top side (left) of the PCB while the first three matrices of the H35DEMO are wirebonded onto the pads on the back side (right) of the PCB.

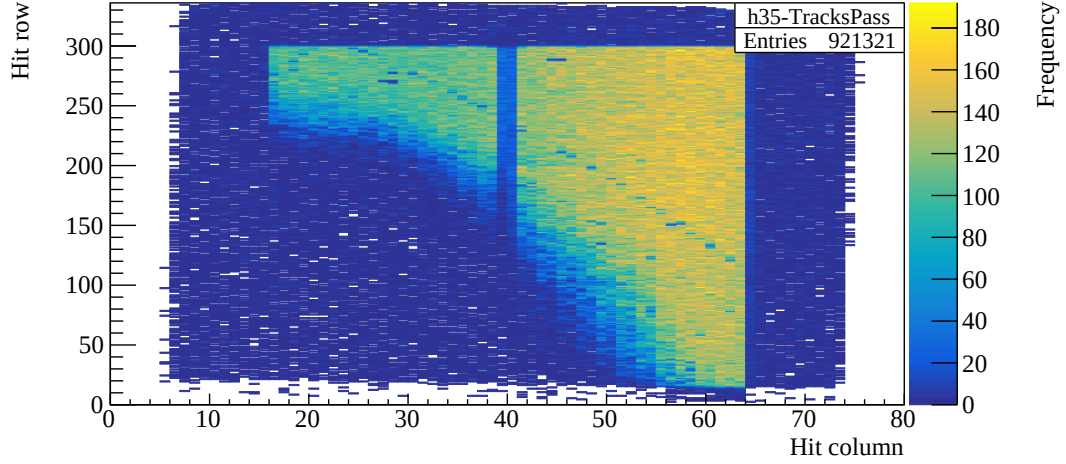


Figure 5.69: All successfully reconstructed tracks of the sample TBP-03 (thinned to $100\text{ }\mu\text{m}$ and back-side-processed) at room temperature biased at -100 V with a threshold at 1000 electrons. Analog matrix 1 can be seen from column number 16–38 and analog matrix 2 is in columns 41–63. The gradient is due to a miscalibration of the flip chip machine causing an uneven capacitive coupling.

Fig. 5.69 shows there is a gradient in the parallelism between the sensor and the readout chip which was due to loss of calibration of the machine.

In the autumn of 2017 new samples were measured at the SPS test beam and the results are published in [144]. The spatial resolution is mainly given by the pixel size. The performance of the alternative cross-section is superior to the default one due to a smaller input capacitance in the absence of the biasing DPTUBs between the DNTUBs which can be confirmed with TCAD simulation shown in Appendix B.4. The overall detection efficiency is excellent and superior to 99% for all resistivities with sufficiently high bias voltage and with a threshold of 2000 e. Especially remarkable is the high-gain and no-DPTUB matrix. With new calibration of the flip-chip machine a good uniformity of the glue layer can be achieved across the large-area sensor.

5.5.4 Laboratory Tests of the NMOS Matrix

In order to do measurements of the NMOS matrix it is essential to understand how the readout works.

The NMOS matrix is a standalone matrix that consists of 16 rows and 300 columns. Column 1–150 are normal pixels and column 151–300 have a time walk (TW) compensated comparator which compensates for the time difference when a signal crosses the threshold depending on its amplitude [122]. These two halves are referred to as the left and right matrix, respectively. In these measurements only the first 150 columns or the left matrix is considered, which is referred to here as “NMOS” matrix, if not explicitly stated otherwise. Only the pixels in the first column of this matrix have an analog output that can be connected to the outside world.

The digital readout is designed in a column-drain architecture. The analog output of each pixel is connected to a readout cell (ROC) in the periphery which contains a discriminator, amongst other processing electronics, to check whether there is any analog signal above threshold coming from this pixel. In case of a signal, the time stamp of the leading edge of the signal and address of this pixel are stored separately each in an 8-bit memory. The pixel with the largest address has the highest readout priority so the columns are read out from top to bottom. Every 40 ROCs (2.5 columns) are connected to an end-of-column (EOC) cell making 60 EOCs per half-matrix. This was done due to layout constraints. The 60 EOC cells are connected in series. Each cell has two 8-bit shift registers in parallel for address and time stamp, respectively. Each readout clock (RoCk) cycle these data from the first EOC are sent to the control unit and the data in the rest of the EOCs move one EOC further towards the control unit.

The control unit generates clocks and time stamps, handles ROCs and EOCs, and controls the data stream. The clock structure is sketched in Fig. 5.70. There is an external fast clock (FastCk) which dictates the data transmission to and from the control unit and some logic blocks. From this external FastCk the control unit generates the RoCk which is 8 times slower. The time stamp is generated every 8 FastCk cycles (= 1 RoCk cycle). For the readout a sequence of control signals have to take place over 60 RoCk cycles: Ld, PullDn, Rd and ParEn. The length corresponds to the shifting speed of the 60 EOCs into the control unit. The control sequence starts with Ld asserted which validates the pixel with the highest readout priority. Then PullDn and Rd are asserted during which the address and time stamp are

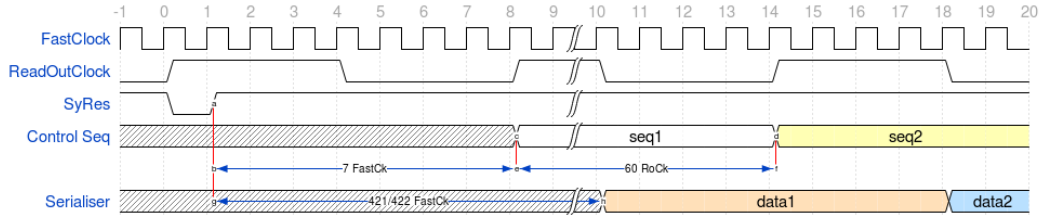


Figure 5.70: Clocks of the standalone readout.

prepared to be stored. And finally, these are stored in the corresponding EOC when ParEn is asserted. With each RoCk cycle the data is sent to the next (lower) EOC so the data in the first EOC is passed to the control unit to be stored in a register and sent to the serialiser. The serialiser sends out the data with a speed of 1 bit/FastCk where address and time stamp are as transmitted in parallel. I.e. it takes 8 FastCk cycles (= RoCk cycle) to transmit one hit. The pixels are zero-suppressed, i.e. only pixels with a hit are read out. If there are more than one pixel hits in one EOC then the pixel with the highest address is read out first, the others are read out in the following cycles with corresponding priorities.

To make sure that everything works synchronously a synchronous reset (SyRes) signal has to be sent once in a while. This signal is active high. It resets the control sequence and resets the time stamp to 0. The control sequence restarts 7 FastCks after SyRes is deasserted. If the time stamp reaches 255 between two SyRes signals, it restarts from 0 again. In order to ensure that all time stamp values are generated and that the readout sequence is complete, the SyRes signal has to be sent every $256 \cdot 60 \cdot 8$ FastCk cycles. The FastCk can be operated at moderate frequencies < 50 MHz and also at high frequencies up to 320 MHz which corresponds to 40 MHz RoCk and time stamp, which is exactly the bunch crossing frequency of the LHC. The serialiser starts to output data 421 or 422 FastCk cycles after deasserting the first SyRes at moderate or high frequency, respectively, and it is unaffected by further SyRes signals.

5.5.4.1 The Measurement Setup

The measurement setup for the H35DEMO consists of a Xilinx Virtex Field-Programmable Gate Array (FPGA) board connected to a motherboard which

provides several probing points, inputs of HV and low voltages, and an interface to a small carrier board with a wirebonded H35DEMO sensor. The FPGA can be programmed in such a way that either address or time stamp information can be output into a scope.

A Domino Ring Sampler (DRS)4 evaluation board [159] by Paul Scherrer Institute (PSI) is connected to the computer to record the measurements with the associated software. The DRS4 has a sampling frequency of 0.7–5 gigasample per second (GSPS) with 1024 sampling points. It needs voltage and timing calibration which are applied to the measurements. While the voltage calibration reduces the root mean square (RMS) noise from 8 mV to 0.35 mV and is independent from the sampling rate, the timing calibration should be redetermined whenever the sampling rate is changed. However, the deviation from the nominal bin width at 5 GSPSs is only ± 100 ps which is negligible for our measurements. The DRS4 can record measurements in binary and Extensible Markup Language (XML) format. While the binary data are very small, the XML data are much larger but human-readable and therefore easy to parse. For measurements using pulse injection, the transistor-transistor logic (TTL) output of the pulse generator is connected to one input channel of the DRS4 as a trigger while the signal to be recorded is connected to another channel. The analog in- and output characteristics can be taken from the datasheet in [159] and are roughly $V_{\text{out}} = 0.9V_{\text{in}}$. The waveforms shown here are the output voltages V_{out} of the DRS4.

The software used to control the pixels is a Java applet, which sets the digital-to-analog converter (DAC) values, pixel thresholds and writes out the address or time stamp into a text file. The DAC values programme the bias block which gives the transistors the ideal operating voltages. The default DAC values used are shown in Table C.1. There are 17 6-bit on-chip DACs. Each 7-bit DAC register segment has therefore one spare bit that can be used for additional configurations, for example enabling pulse injection and monitor. The in-pixel comparator baseline is set to 1.5 V and the second stage CMOS comparator has a baseline at 2.2 V [123]; therefore, the thresholds NTHPIX and NTH are set to 1.6 V and 2.5 V, respectively.

Using the spare DACs the analog part of a pixel (before the comparators) can be configured for testing by injecting an external signal from a pulse gen-

erator. The sensor is programmed in a way that only one pixel of a specific row and column is activated at a given time.

5.5.4.2 Measurements with Pulse Injection

In order to have some first measurements of analog signals of the first 16 pixels, measurements with pulse injection were carried out. The response signals were recorded and their amplitude, rise time, fall time, and the delay between the input and output signals are measured.

The original default value for DAC 13 was 38. However, it is found that not all analog channels respond well to it as shown in Fig. 5.71. Therefore, without any systematic analysis the value 2 is found to work better for most of the 16 pixels. An effort of a systematic study to find the best setting of all DAC values for all pixels in the matrix is currently ongoing but not part of this thesis. The analog part of the readout can be tested using pulse injection and analog output. When the analog part is tested, the pulse is injected directly onto the collection electrode on the substrate as shown in Fig. 5.72 [123], thus applying bias voltage affects the signals as shown in Fig. 5.73. It reduces the noise but also increase the amplitude. However, even with the better DAC value, some pixels, especially pixel 1 and 4, do not have any, or only very flat signals.

For each measurement with pulse injection 2000 events are recorded with the DRS4. Each waveform is saved in the XML file which is then parsed. The components of a signal, i.e. the baseline, rising edge and falling edge, are fitted individually because it would be more complex to use the sum of three functions. The baseline of a signal is fitted using only the data points on both ends of the waveform but discarding the first five points.

The baseline was measured by recording an empty channel while using another channel as trigger which is connected to the pulse generator output. The baseline can be described by a very flat Gaussian distribution as shown in Fig. 5.74 and is suspected to be cross-talk from the trigger channel. A Gaussian fit works well for a good baseline but reacts very drastically to noise in the fitting region, and would therefore be even less suited as a fit when a signal is present. In contrast, a linear fit would be robust against noise by its nature and applicable as a baseline fit. As shown in the right plot of Fig. 5.74, all

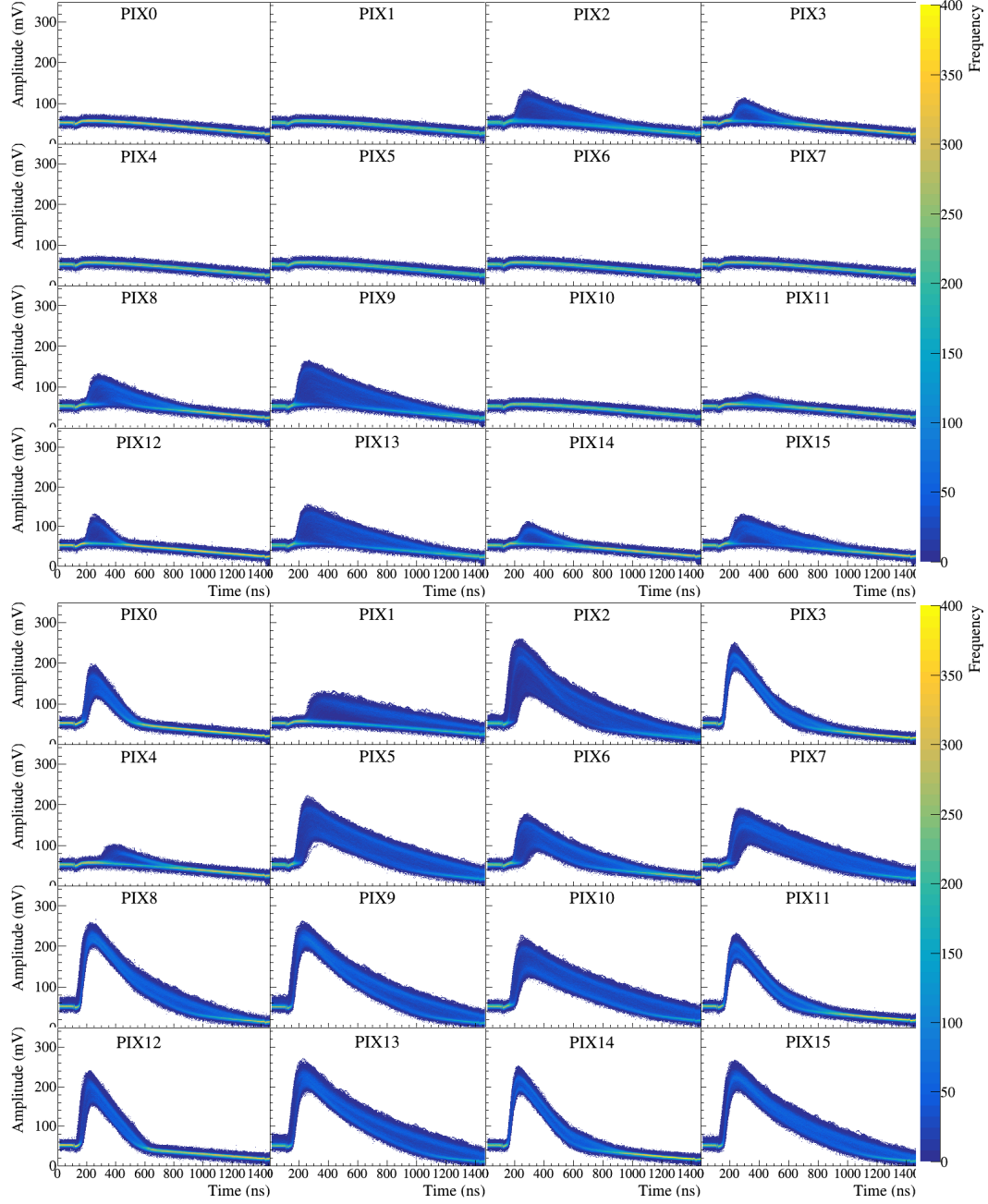


Figure 5.71: Analog signals for DAC13 values of 38 (top) and 2 (bottom). Measurements done in the dark with sensor biased at 71 V.

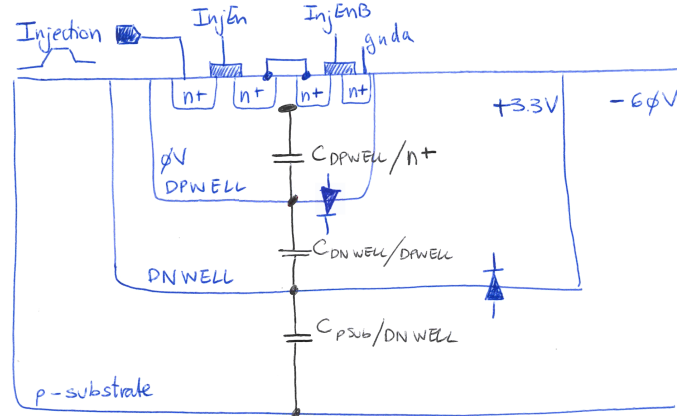


Figure 5.72: Sketch of pulse injection schematic [123].

recorded baselines are akin on average and the non-linearity is small compared to typical signals. Therefore, the baseline of the waveform is approximated by a linear fit and subtracted from the signal. The remaining signal is fitted using an error function for the rising edge and a polynomial 7th order for the falling edge to compensate for the Gaussian baseline and for the noise as shown in Fig. 5.75. The rise and fall times are calculated as the time from 10 % to 90 % of the amplitude that is taken from the fit. The amplitude distribution is shown in Fig. 5.76. For well-working channels with narrow distributions, the amplitude is between 150 and 200 mV. Pixels 1 and 4 have mostly no signal, therefore the first narrow peak comes from fitting the baseline and the smaller second peak comes from some events with low amplitudes. The rise and fall time distributions are shown in Fig. 5.77 top and bottom, respectively. The rise time is about 36 ns while the fall time has a very wide distribution due to noise, and is generally larger than 500 ns. But values as small as about 260 ns also occur. The sensor was biased at -120 V for these measurements. The effect of different bias voltages can be seen in Fig. 5.78.

The delay between the generated pulse and the response is measured separately so that both the trigger and the analog signal can be in the same time window. Fig. 5.79 shows the signals and their timing distribution. The discrepancy in the signal amplitude with respect to the amplitude measurement is due to changes in the firmware and software. The signal timing is determined by approximating the rising edge with a linear function and extrapolated to

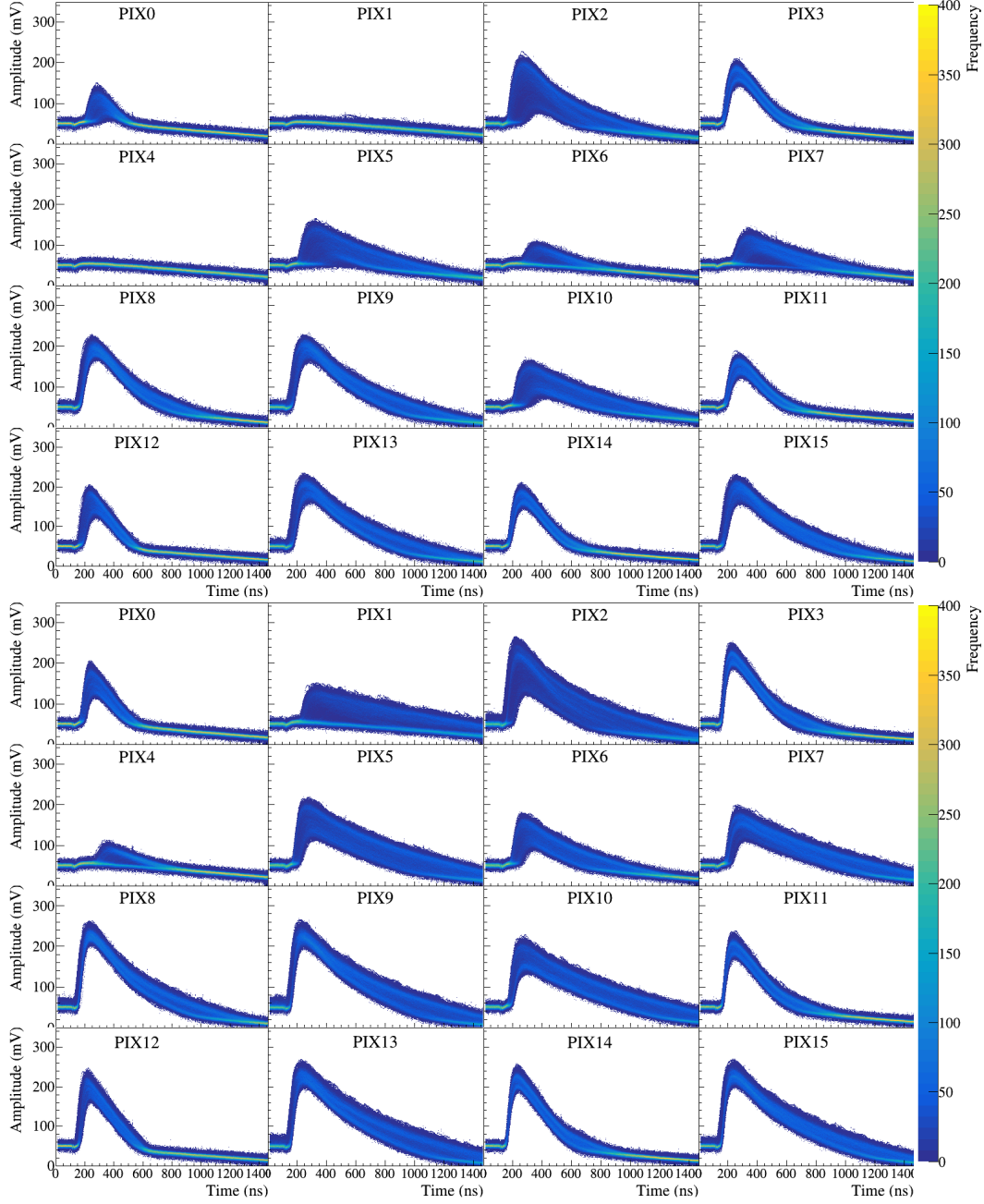


Figure 5.73: Analog response to pulse injection for bias voltages of 0 V (top) and -160 V (bottom). Measurements done in the dark.

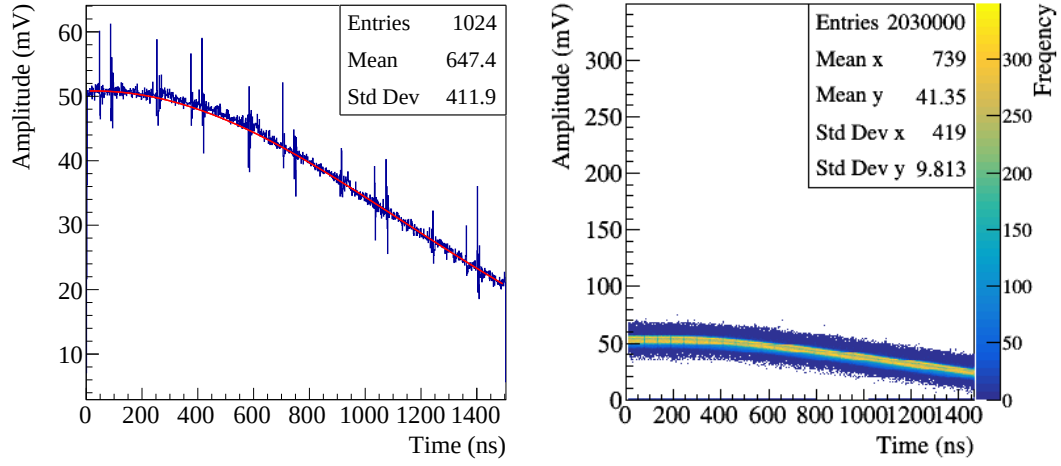


Figure 5.74: Baseline of the measurement taken with the TTL output of the pulse generator connected to the DRS4 as trigger and no signal connected. It can be fitted using a Gaussian function (left) but it is not robust against noise.

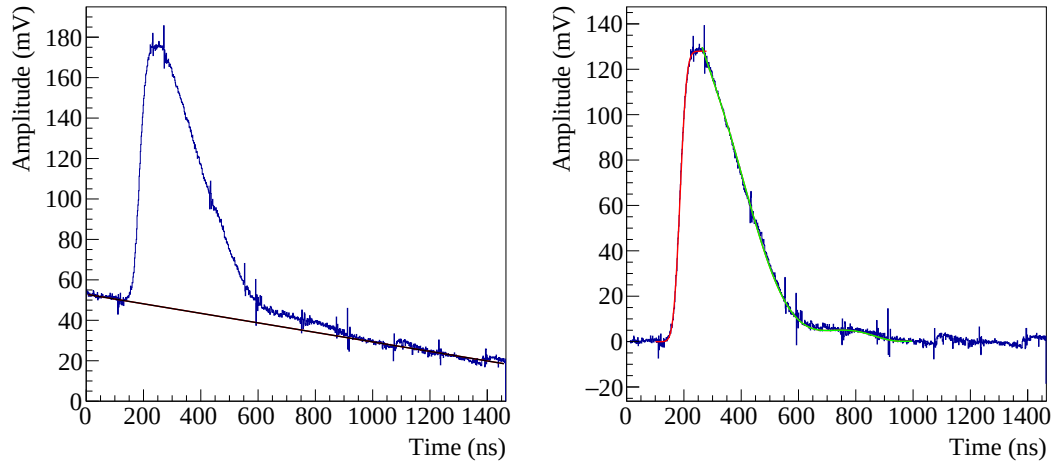


Figure 5.75: Analog signal recorded using DRS4. The baseline is fitted by a linear function (black) and subtracted. The remaining signal (right) is fitted using an error function for the rising edge (red) and a polynomial for the falling edge (green).

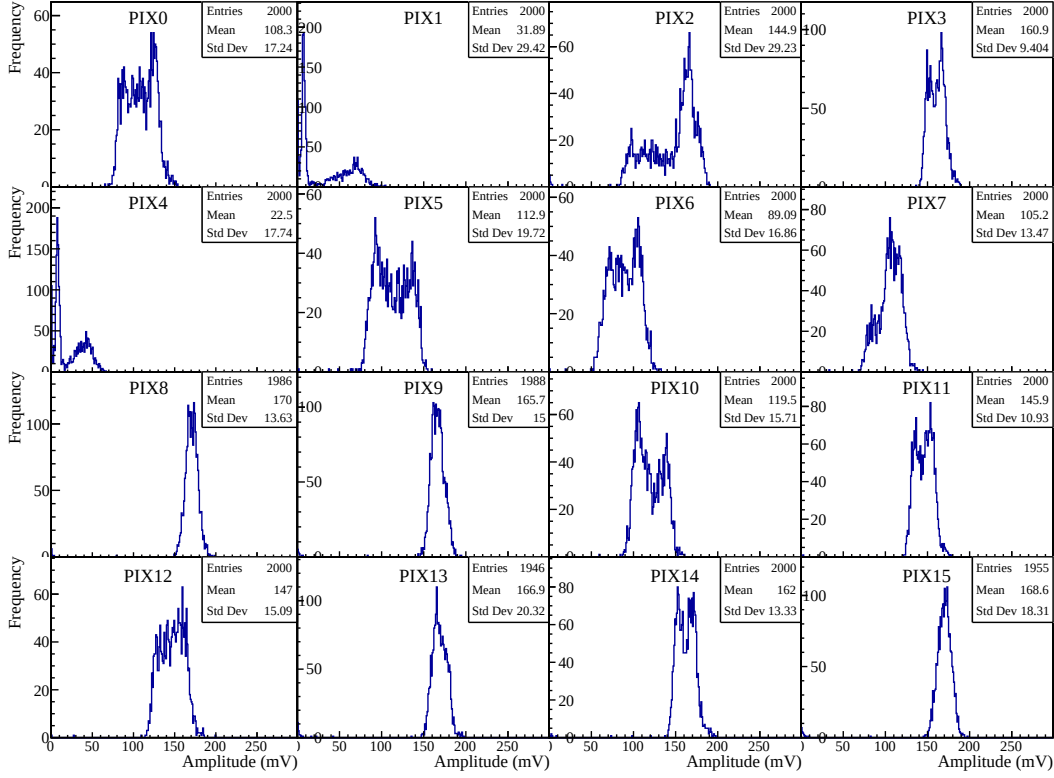


Figure 5.76: Distribution of analog signal amplitude for all 16 pixels with an analog output. The signals are response to pulse injection.

the point it crosses the baseline which is corrected to zero. The trigger signal has a very steep rise of maximum two bins (< 3 ns) therefore no fit function is needed to determine the timing. The delay due to the difference in cable lengths of the trigger and the signal, which is about 20 cm, is about 1 ns and therefore negligible. The distributions for each pixel with an analog output are shown in Fig. 5.80. One might notice that signal timing distribution has a standard deviation of about 70 ns. This would be problematic for operation within the ATLAS detector since events from 3 consecutive bunch crossings cannot be distinguished. However, as the H35DEMO is still a prototype and the final technology for the ITk is not decided yet, this can yet be improved.

5.5.4.3 The Scintillator Setup and Source Measurements

The scintillator setup is an add-on to the existing setup and should work as the trigger for source measurements. The scintillator and light guide are

CHAPTER 5. HIGH VOLTAGE CMOS SENSORS

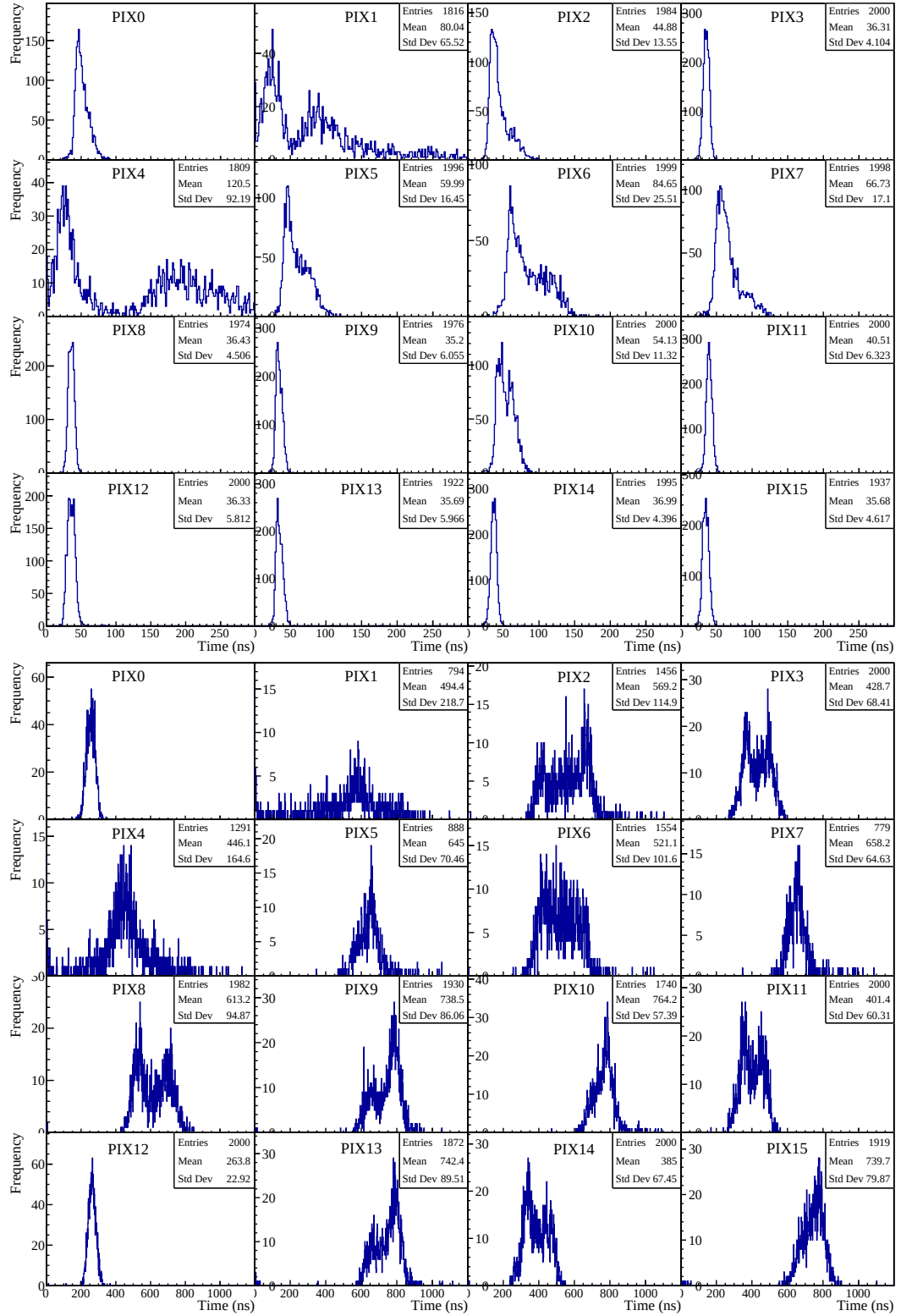


Figure 5.77: Distribution of analog signal rise (top) and fall (bottom) times calculated as time from 10 % to 90 % of the amplitude.

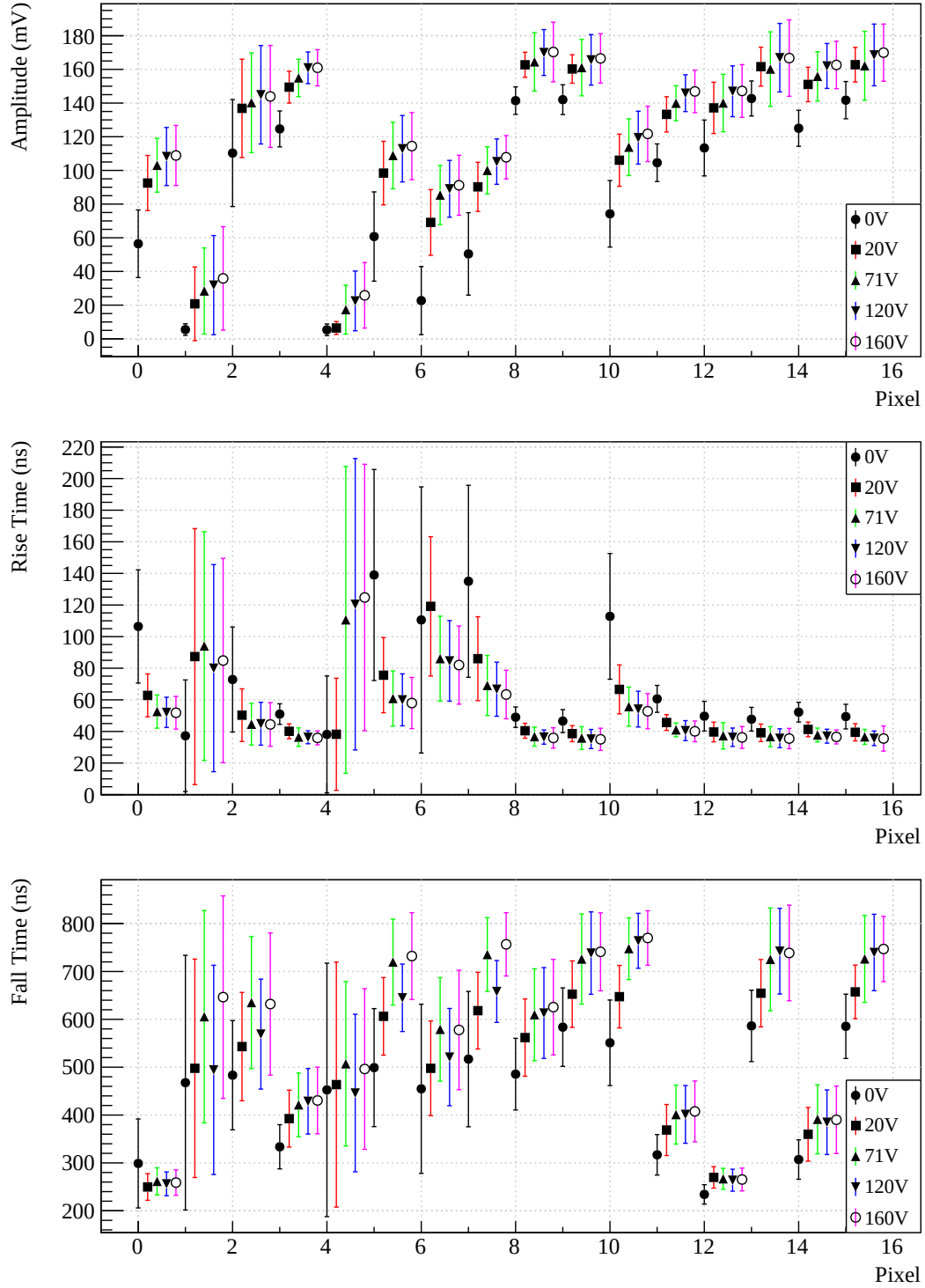


Figure 5.78: Distributions of the amplitude, rise and fall time of the analog signals for different bias voltages. The signals are obtained using pulse injection. The data points for each voltage is shifted slightly within the same pixel bin for a better legibility.

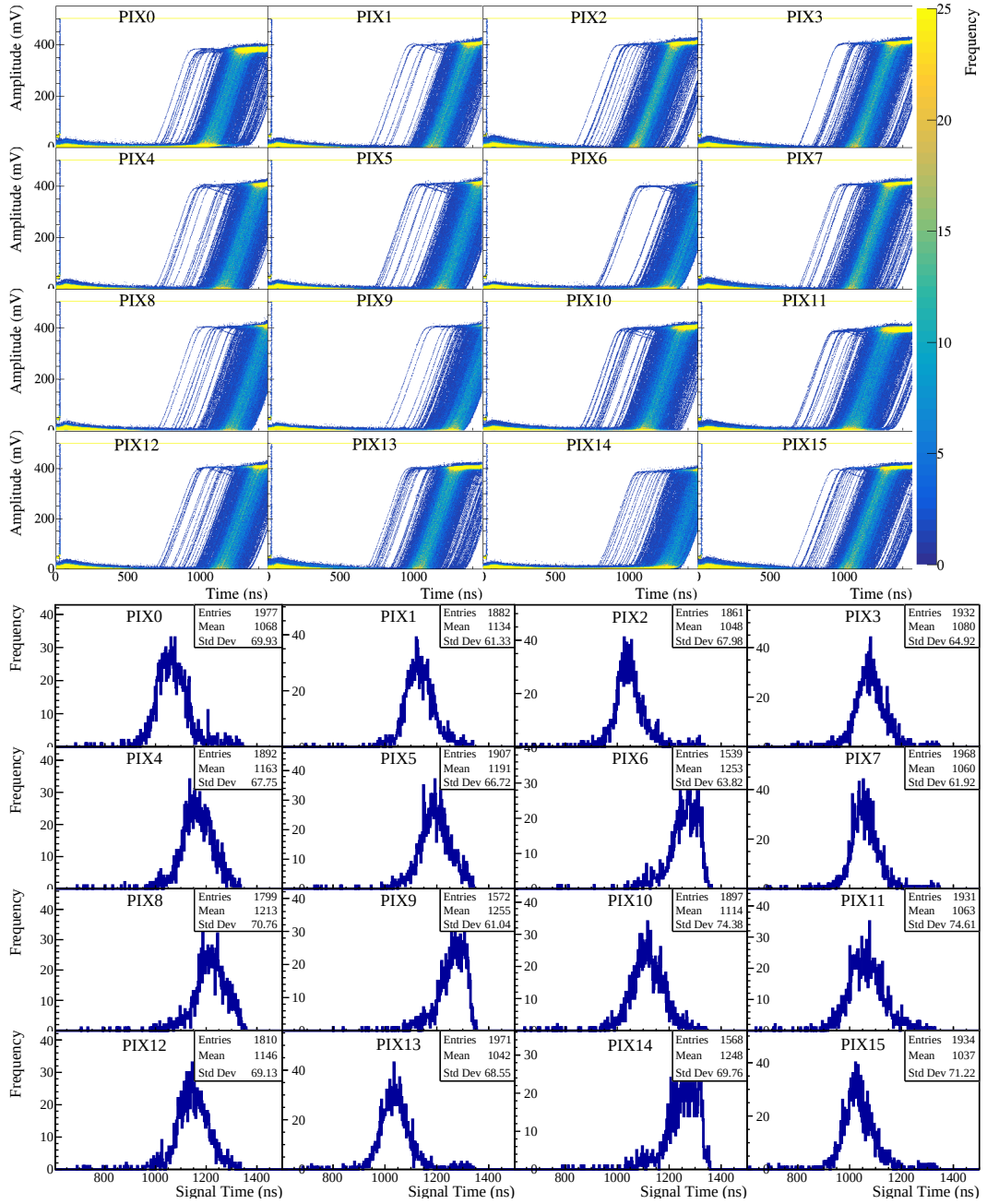


Figure 5.79: Analog signals (top) and their timing distribution (bottom).

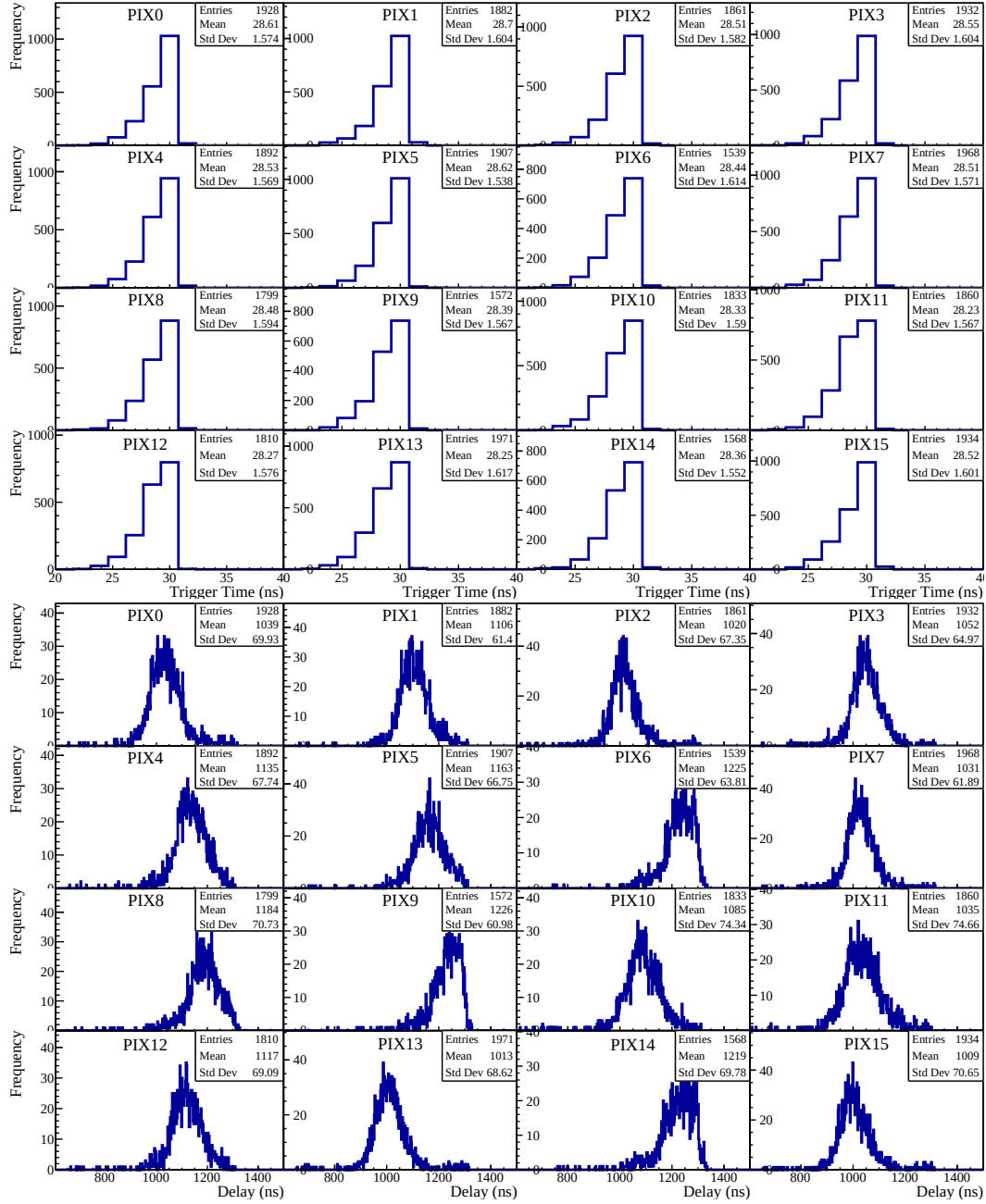


Figure 5.80: Distribution of trigger timing (top) and delay of the response (bottom).

recovered spare parts of which no precise specifications are known. As it works well with the PMT, the assumption can be made that it emits in the ultra violet (UV) range. The PMT is a Hamamatsu H11901.

For source measurements ^{90}Sr is used which undergoes β^- decay to ^{90}Y by emitting an electron and an electron anti-neutrino. ^{90}Y also undergoes a β^- decay to become ^{90}Zr which is stable. ^{90}Sr has a half-life time of 28.79 years and a decay energy of 0.546 MeV which is distributed among the decay products while ^{90}Y has a half-life time of 64 hours and a decay energy of 2.28 MeV.

To minimise the noise of particles other than from the source the scintillator is placed below the sensor. To ensure a reliable trigger signal there should not be too much material between the sensor and the scintillator. Therefore, the latter is placed snugly between the carrier board and the motherboard as shown in Fig. 5.81. The PMT is connected according to the datasheet [160] with a potentiometer for adjusting the gain, as shown in Fig. 5.82. The potentiometer is adjusted so that the gain is maximal. Using this setting the noise appears below 180 mV so the signal with an amplitude of about 360 mV can be well separated.

Fig. 5.83 shows source scans of the NMOS matrix with the sensor unbiased (top) and at -120 V . In the unbiased case the sensor response has an abnormal pattern which disappears with bias voltage. However, in the biased sensor there seems to be some fine structures. E.g. there is a clear edge between row 0 and 1, and between 8 and 9. The middle and the bottom rows generally see more hits, as does every fifth column. This is very likely an indication of malfunctioning of the firmware or software. These plots are work in progress and further improvements on the readout are being made.

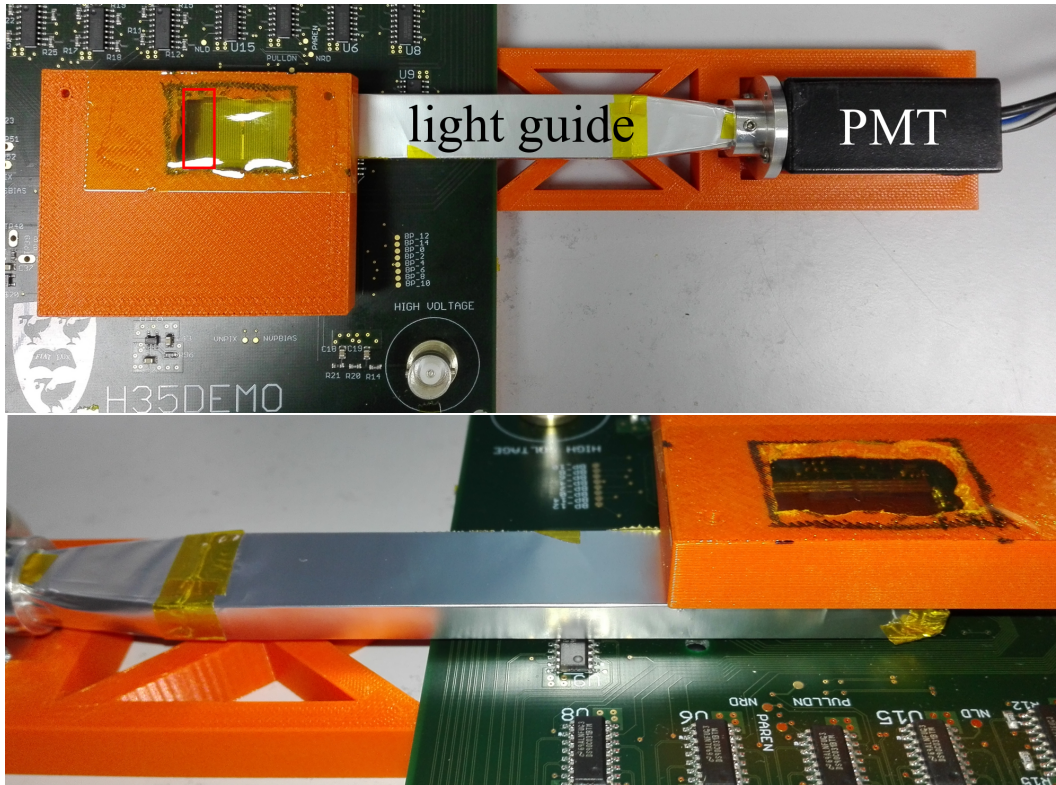


Figure 5.81: Scintillator trigger setup consisting of a scintillator and light guide wrapped in aluminium foil attached to a Hamamatsu PMT fixed with a 3D printed holder (orange). The light guide fits just between the two PCBs.

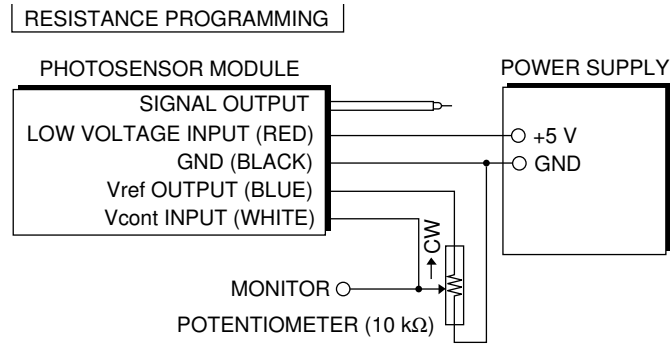


Figure 5.82: PMT amplification adjustment circuit [160]

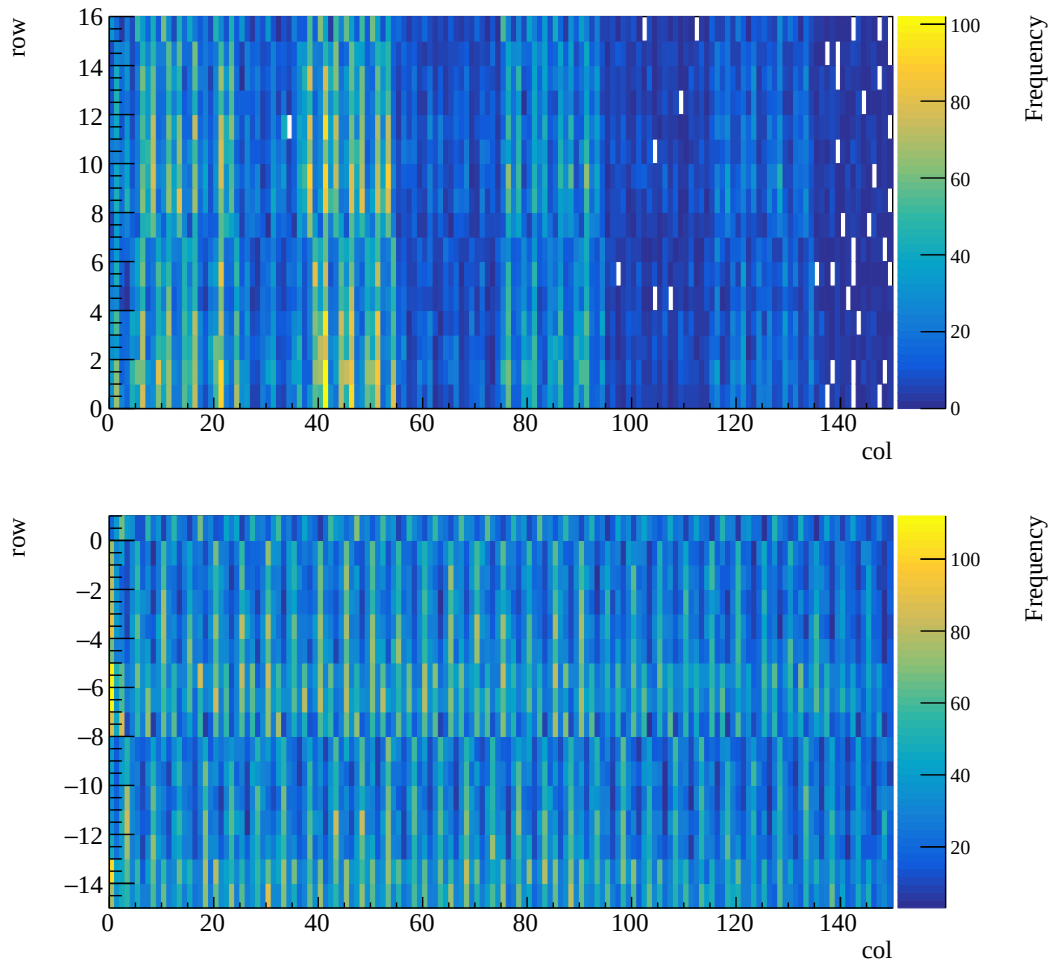


Figure 5.83: Source scan over the NMOS matrix at 0 V (top) and -120 V (bottom). The response of the pixels are not uniform.

Chapter 6

Summary and Outlook

In this thesis two main topics were discussed: the ATLAS radiation damage pixel digitiser as well as simulations and measurements for prototype HV-CMOS sensors.

The work done on the radiation damage pixel digitiser is based on previous work by other people. It has been further developed and refined as shown in this thesis. For example, different approaches to generate the electric field and Ramo field were added and tested, Lorentz angle for magnetic field and chunk correction are introduced, and calculations of parameters are improved, like the diffusion coefficient and ToT conversion. Test beam data were taken at SLAC using an EUDET telescope. In comparison with the simulation the tendency is as expected. For the irradiated sensor, while the cluster size distribution matches the simulation, the ToT does not very well. This might be an indication that the sensor has been annealed, which could be verified with I-V measurements of the sensor. It is more likely to be due to the unprecise ToT conversion. For the unirradiated sensor, there are some discrepancies between data and simulation in the cluster size and ToT distributions. Again, the latter might be due to the ToT conversion. Therefore, no precise conclusions can be drawn. The tendency of the ToT being higher than the data is similar in both the irradiated and unirradiated cases. The ToT conversion as a function of charge is not linear, as assumed in the digitiser, and thus is worthwhile to be updated using data from sensor tunings. Validation in the same way can be easily done with well-known sensors. Further efforts are made to study radiation damage on the TCAD level [161] since TCAD maps are important

inputs to the digitiser. The digitiser is currently being implemented in Athena and validated with data from the ATLAS detector [162, 100].

A TCAD structure simulation of the H35DEMO was created using estimated doping concentrations. Extensive simulations were performed to determine the leakage current, depletion region, electric field, radiation damage and charge collection. Due to the nature of 2D simulations the simulated leakage current was about four times more than measured while the simulated depletion depth of the top-side-biased sensor agrees well for low resistivities but is half that measured at -100 V for a resistivity of $1000\ \Omega\text{cm}$. The electric field is stronger and more uniform for the back-side-processed and -biased sensor. The “double-peak” in the charge collection profile as a function of depth can be reproduced using simulation.

Measurements on the H35DEMO in test beam campaigns showed good results in the analog matrices with a detection efficiency of better than 99 % for all resistivities with sufficiently high bias voltage and a threshold of 2000 electrons. The alternative layout without biasing DPTUBs between the DNTUBs and the high-gain pixels are found to perform best. The NMOS matrix was measured in the laboratory. The measurements with pulse injection on the analog part of the sensor showed irregular responses of the 16 pixels with analog output. These are due to DAC values which are yet to be tweaked. The obtained amplitude at -120 V is between 150 and 200 mV, the signal rise time is about 35 ns and the fall time larger than 250 ns. The bias voltage of the sensor influences these values. The delay time between the injection and the response is 1–1.2 μs . Source measurements of the whole matrix work and pixel addresses can be received. However, they are not correct and the firmware or software has still to be improved.

Different layouts of the LF 150 nm technology were simulated with structure simulation using approximated doping profiles extracted from process simulation. The sensor capacitance can be accurately described and extrapolated to 3D. For the extrapolation assumptions for the perimeter contribution and approximations of depletion regions were made. However, more exact values could be used. The extrapolation yields about 80 fF and 400 fF for the DNW-substrate and DNW-PSUB capacitance, respectively, at a bias voltage of -120 V for a $(38.64\ \mu\text{m})^2$ PSUB and a $(45.14\ \mu\text{m})^2$ DNW in a $10\ \Omega\text{cm}$ sub-

strate. These two components are the main contributions to the total sensor capacitance. The former is the smaller part and decreases with increasing bias voltage, whereas the latter is the main contribution which is constant. Therefore, reducing the size of the PSUB is essential while ensuring enough overlap with PW to allow for the bias. Smaller PSUB also benefits the breakdown characteristic which cannot be observed up to -200 V while for the full-width PSUB the breakdown happens at about -150 V . The sensors were fabricated while this thesis was being completed.

Appendix A

Radiation Damage Pixel Digitiser

A.1 Time-over-Threshold Encoding

The true sensor ToT is encoded to the FE-I4 ToT which is different than the true ToT value. The mapping also depends on the value of the Hit Discriminator Configuration (HitDiscCnfg) [163]. In the ATLAS readout the FE-I4 ToT is decoded back with a scheme to represent a ToT similar to the true ToT by using the Readout Driver (ROD). The exact mapping can be taken from Fig. A.1. The same scheme is embedded in the RCE software [164] used for the test beam. The idea behind HitDiscCnfg is to catch small hits which arrive 1 or 2 BCs delayed due to time walk, corresponding to HitDiscCnfg = 1 or 2, respectively. If a hit is below the analog threshold, the FE-I4 ToT is set to 15 meaning no hit; if it exceeds the analog threshold it is considered a small hit with FE-I4 ToT value 14; and if it exceeds both the analog and the digital threshold, a hit is considered a bit hit and the FE-I4 ToT corresponds to the real ToT minus 1. Thus the FE-I4 ToT value ranges from 0 to 13. For the default HitDiscCnfg = 0 as used in the test beam, no small hits are stored and thus the ToT values range from 1 to 14 with 14 as the overflow.

Real ToT	Output					
	HitDiscCnfg=0		HitDiscCnfg=1		HitDiscCnfg=2	
	FE-I4	ROD	FE-I4	ROD	FE-I4	ROD
Below analog thr.	15	0	15	0	15	0
1	0	1	14	1	14	1
2	1	2	0	2	14	1
3	2	3	1	3	0	3
4	3	4	2	4	1	4
5	4	5	3	5	2	5
6	5	6	4	6	3	6
7	6	7	5	7	4	7
8	7	8	6	8	5	8
9	8	9	7	9	6	9
10	9	10	8	10	7	10
11	10	11	9	11	8	11
12	11	12	10	12	9	12
13	12	13	11	13	10	13
14	13	14	12	14	11	14
15	13	14	13	15	12	15
>15	13	14	13	15	13	2

Figure A.1: ToT encoding at different levels for values of HitDiscCnfg [165].

A.2 Test Beam Plots

The properties of the test beam data is studied for different factors.

Fig. A.2 and Fig. A.3 shows the multiplicity dependence of cluster size and ToT. The larger the multiplicity, the larger the cluster size gets. This is consistent for runs 294, 299 and 300, where in the latter two runs the DUT multiplicity agrees with the telescope multiplicity, as well as for later runs, where the runs larger than 700 seem to be more noisy and has a larger multiplicity than in the telescope. The dependence on the temperature is shown in Fig. A.4 for consecutive datasets. The difference between the highest and lowest measured temperature is only 5 K which is already enough to make a noticeable trend. Both the cluster size and ToT decreases with increasing temperature. With increasing bias voltage the cluster size and ToT both tend to move to the larger values as shown in Fig. A.5 In Fig. A.6 the consecutive datasets are taken by sweeping the beam spot across the sensor as shown in Fig. 4.17. Compared with Fig. A.4 the fluctuation with temperature is dominated by the

position-dependent fluctuations. Finally, in Fig. A.7 the difference between the unirradiated sensor, and the earlier and later runs of the irradiated sensor is illustrated.

Fig. A.8 and Fig. A.9 show the data of the unirradiated sensor at room temperature. The variations of both the cluster size and ToT are very small for the same conditions. The change observed with different threshold and tuning is as expected.

The simulation plots in Fig. A.10 show that for the same bias voltage, the cluster size decreases with increasing radiation fluence due to trapping. The ToT distribution of the same fluence increases with increasing bias voltage as described in Section 3.3.2 but decreases for higher fluence.

Fig. A.11 compares the effect of the linear electric field generated by the digitiser and the electric field from a TCAD map (ref. Fig. 4.3).

In Fig. A.12 the effect of different bias voltages is shown and in Fig. A.13 different depletion voltages can be seen.

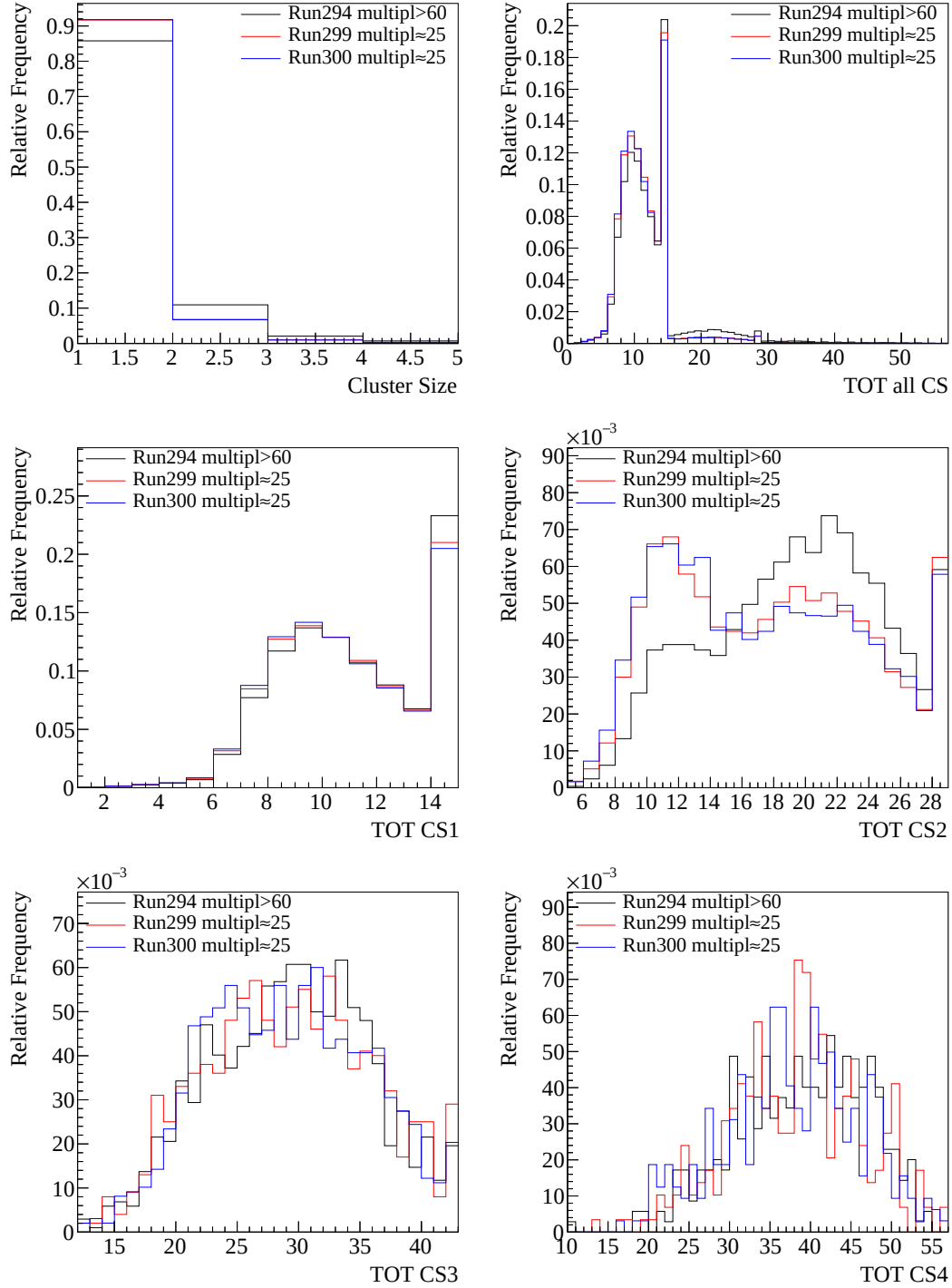


Figure A.2: Variations on the cluster size and ToT distributions depending on the event multiplicity for earlier runs.

APPENDIX A. RADIATION DAMAGE PIXEL DIGITISER

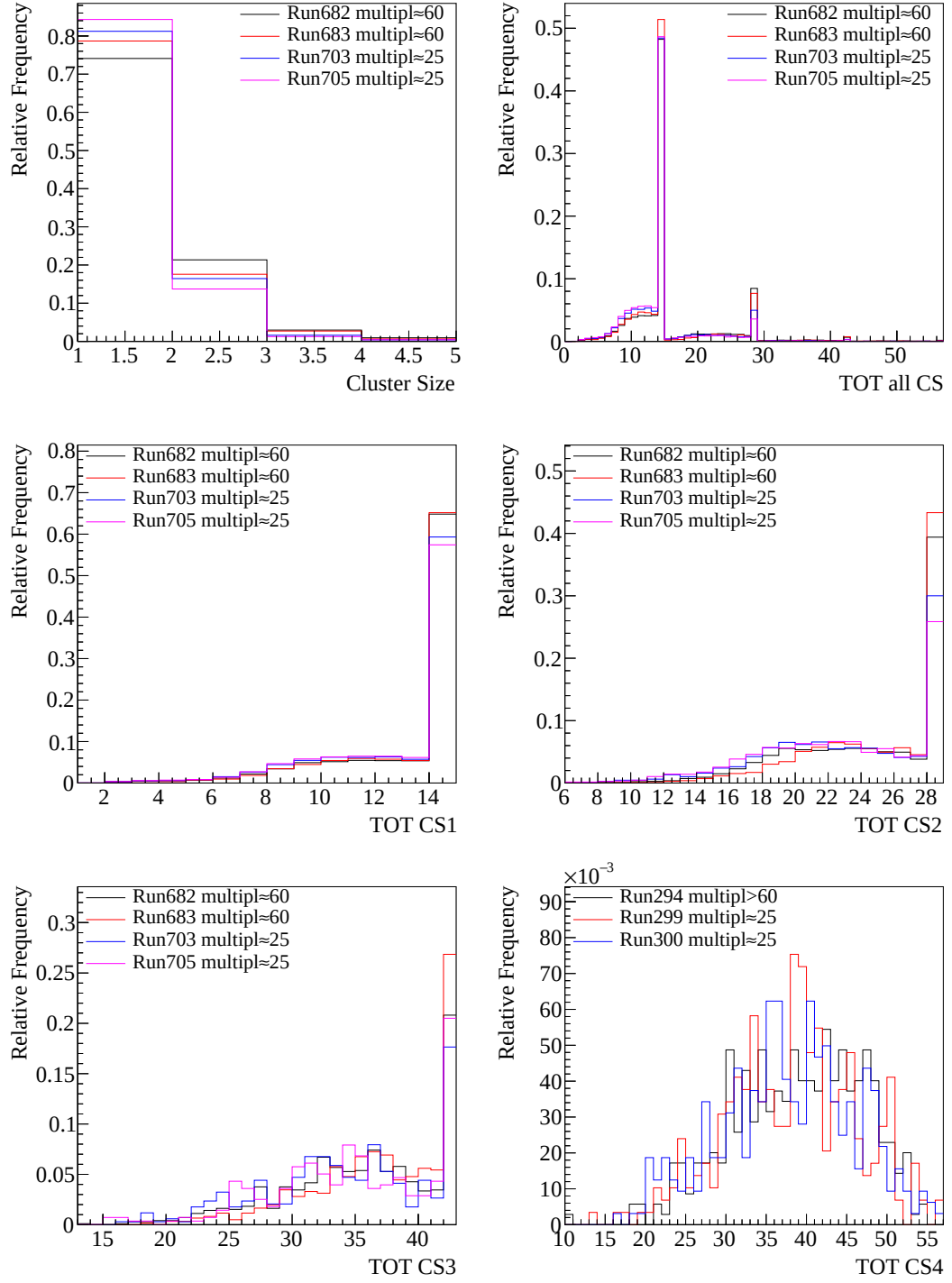


Figure A.3: Variations on the cluster size and ToT distributions depending on the event multiplicity for later runs.

APPENDIX A. RADIATION DAMAGE PIXEL DIGITISER

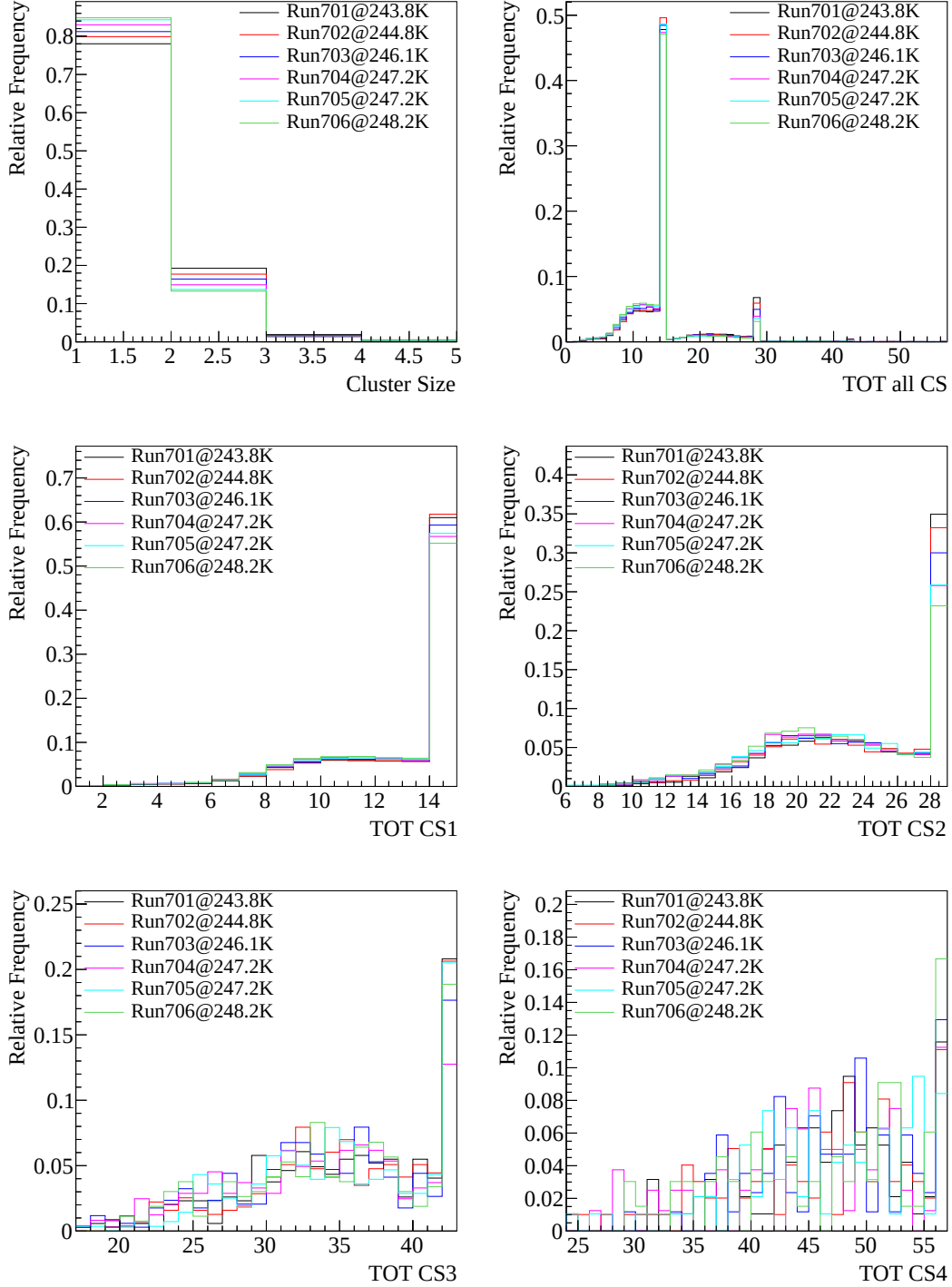


Figure A.4: Variations on the cluster size and ToT distributions depending on the temperature.

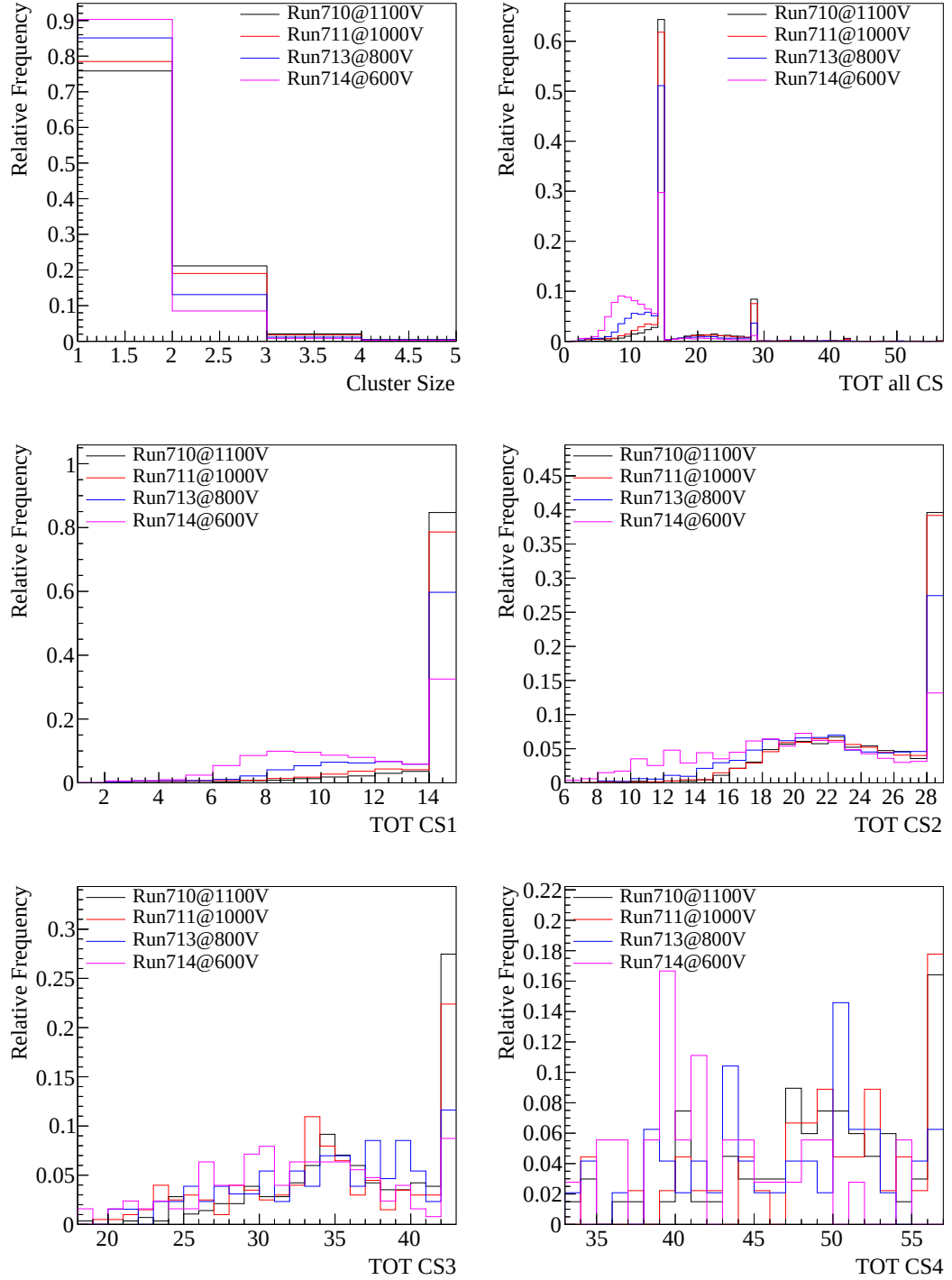


Figure A.5: Variations on the cluster size and ToT distributions depending on the sensor bias voltage.

APPENDIX A. RADIATION DAMAGE PIXEL DIGITISER

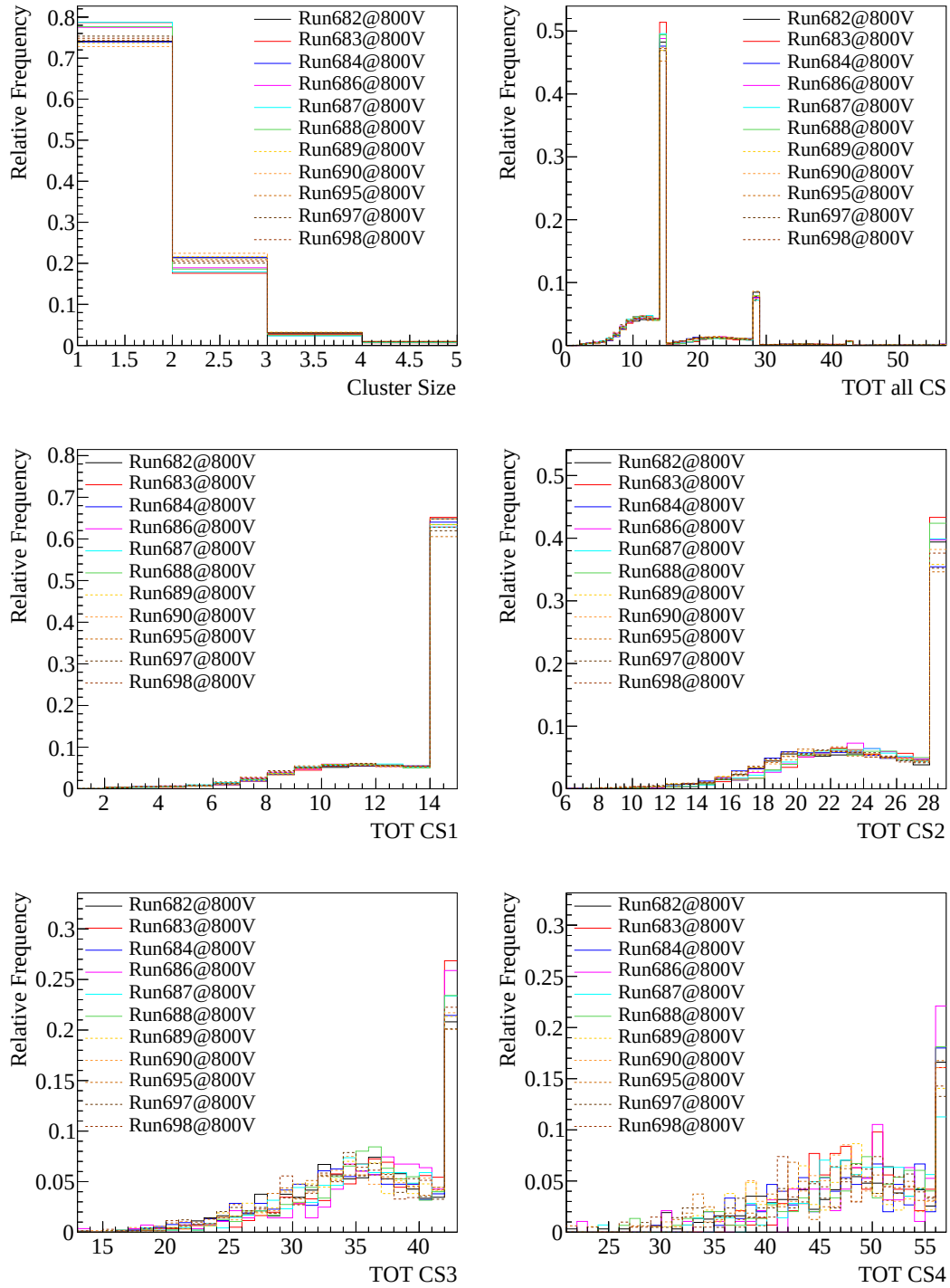


Figure A.6: Variations on the cluster size and ToT distributions depending on the position of the beam spot.

APPENDIX A. RADIATION DAMAGE PIXEL DIGITISER

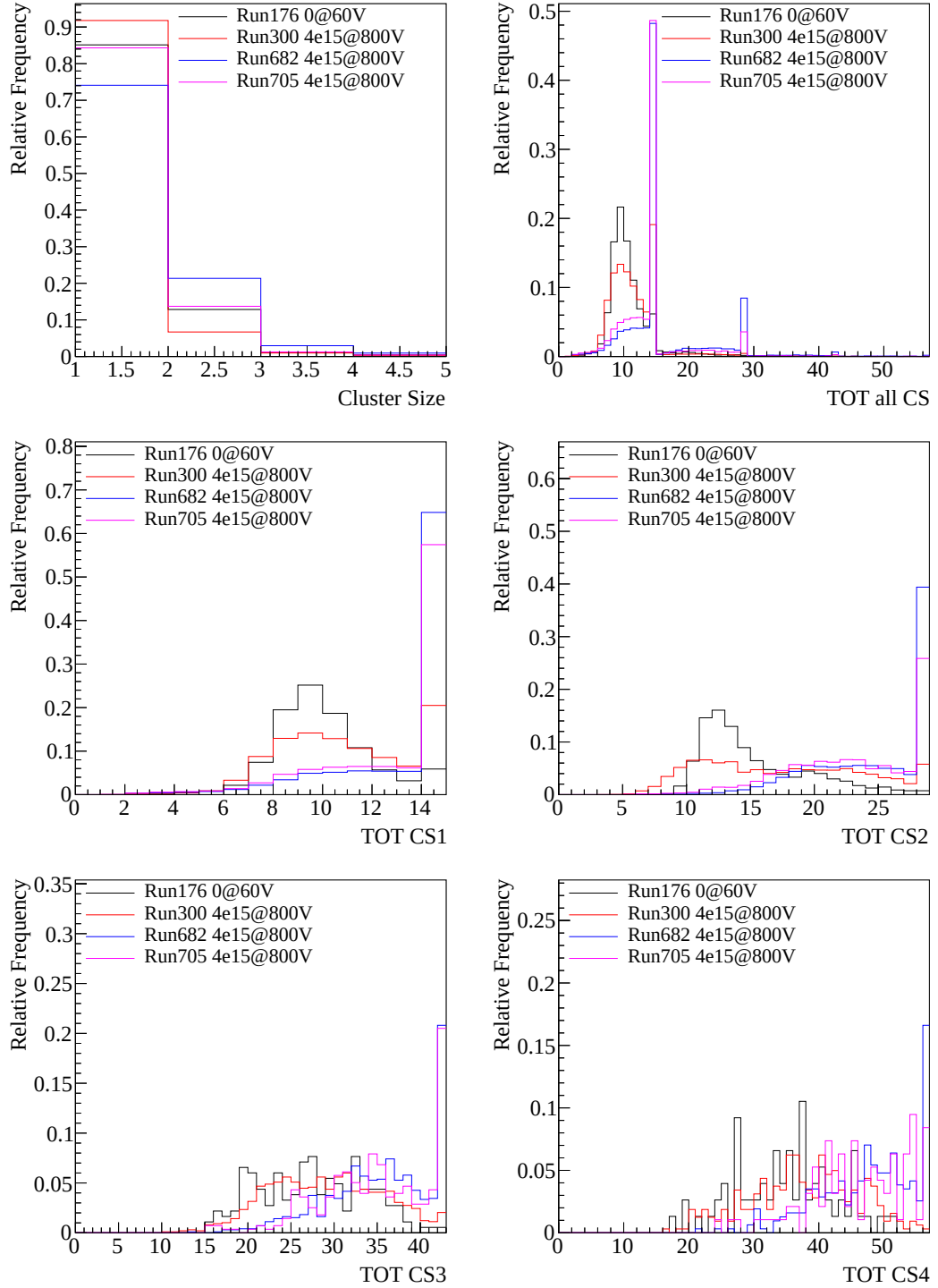


Figure A.7: Comparison of the cluster size and ToT distributions between different batches of datasets.

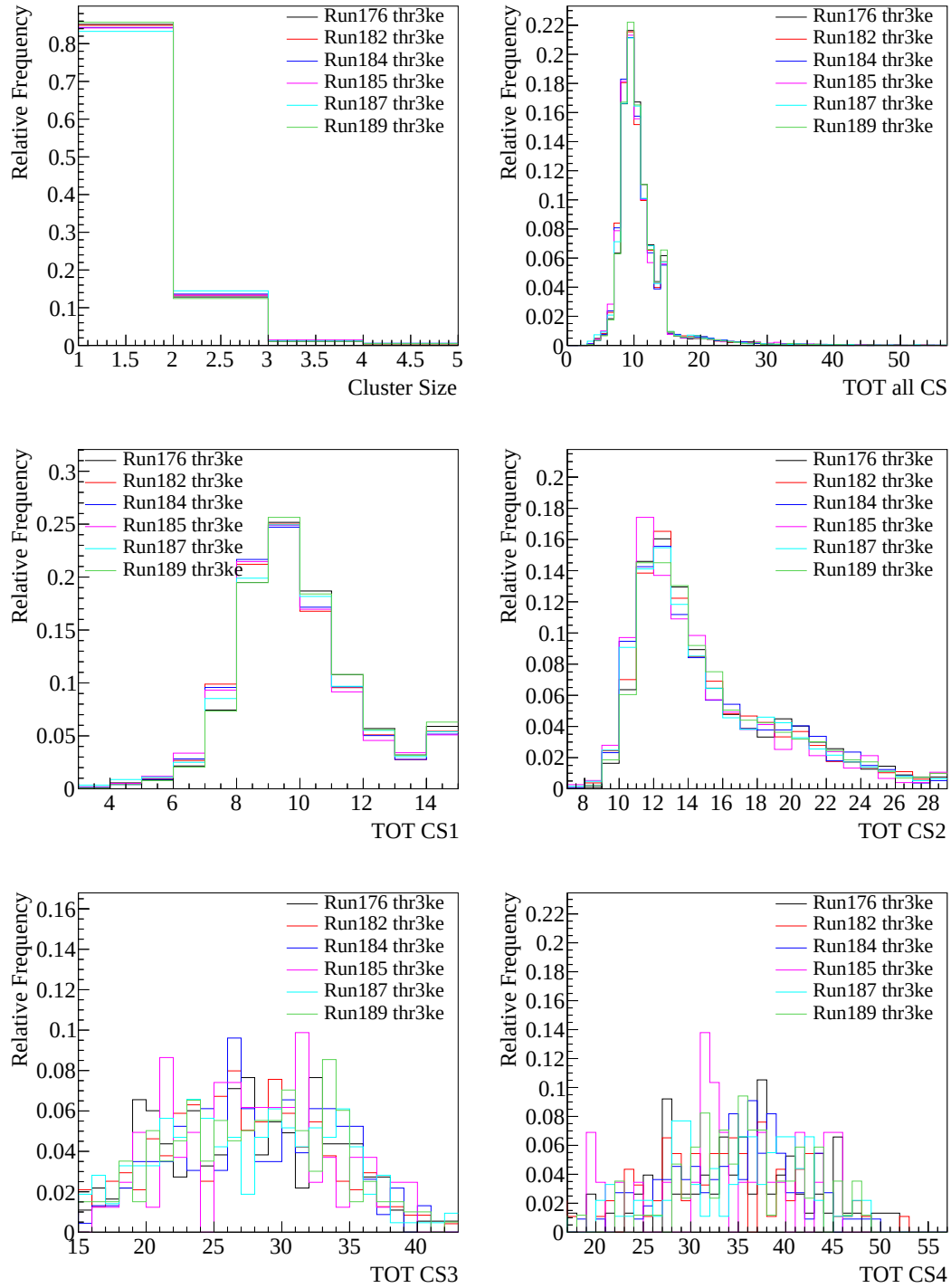


Figure A.8: Distribution of cluster size and ToT of unirradiated sensor. Data taken under the same condition but at different beam position.

APPENDIX A. RADIATION DAMAGE PIXEL DIGITISER

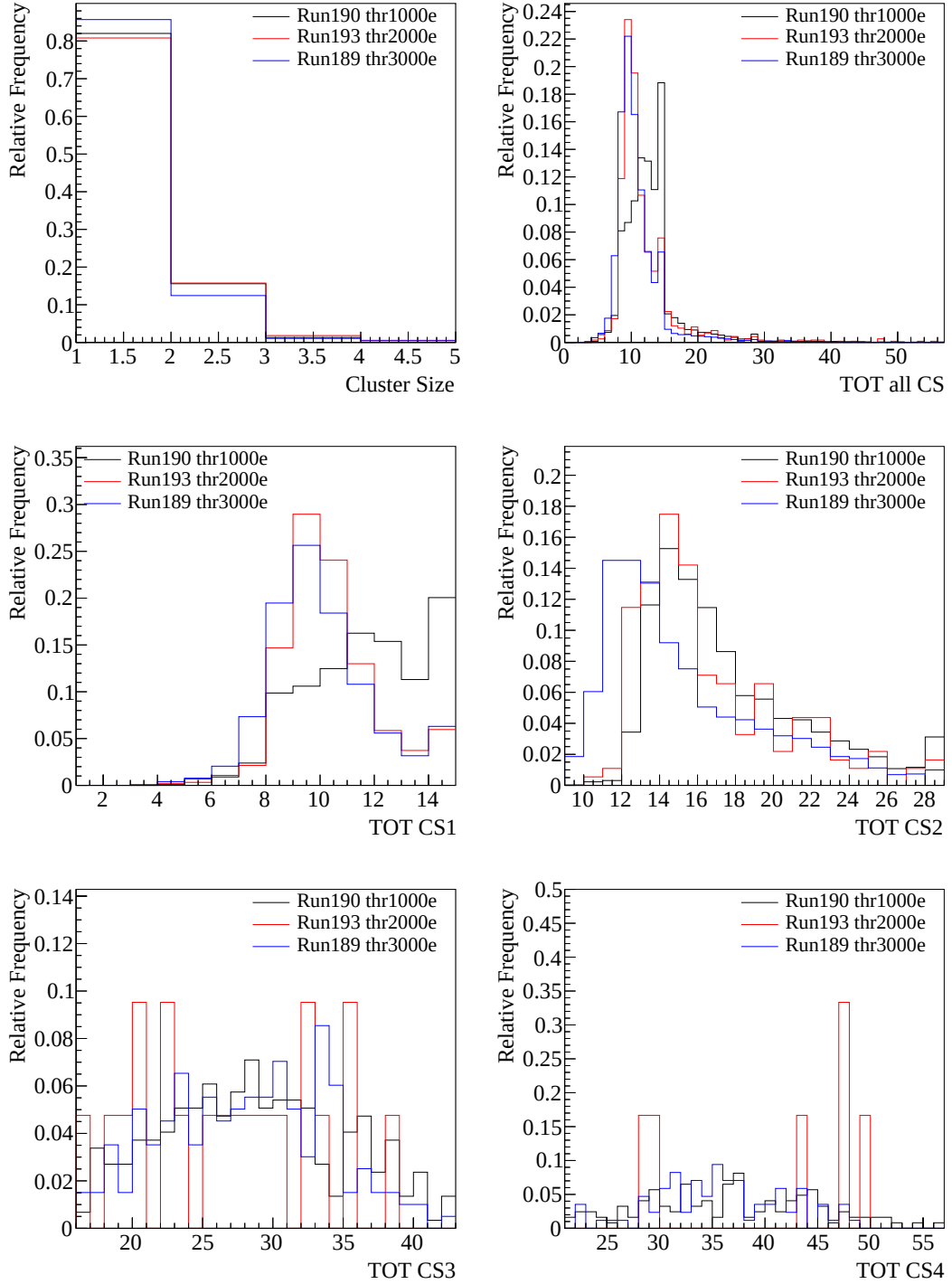


Figure A.9: Comparison of threshold and tuning of the unirradiated sensor. The tuning of 1000 and 3000 electrons threshold is 10 ToT at 16000 electrons and of 2000 electrons threshold is 8 ToT at 11000 electrons.

APPENDIX A. RADIATION DAMAGE PIXEL DIGITISER

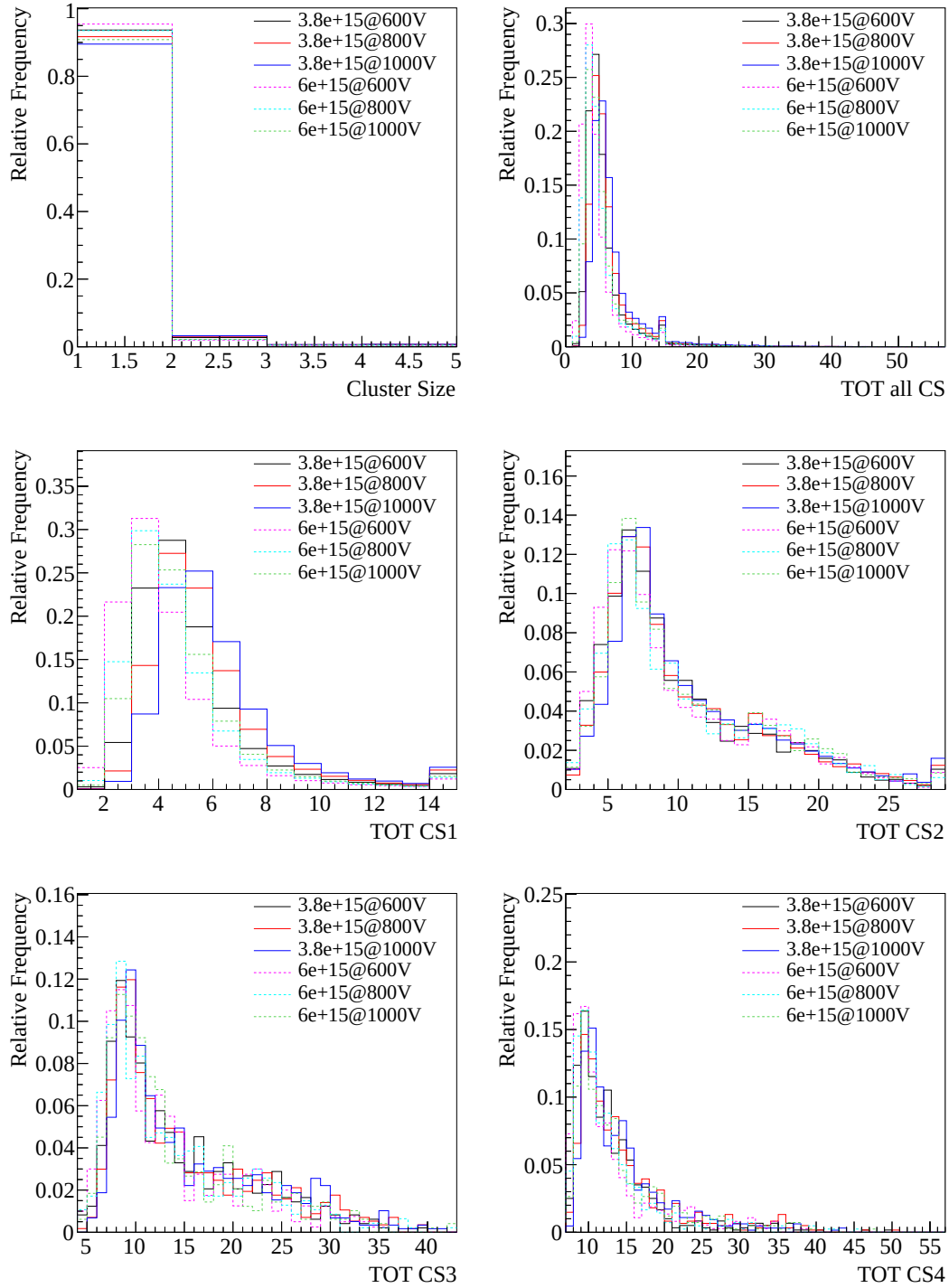


Figure A.10: Simulation of irradiated sensors at the same temperature, threshold and tuning.

APPENDIX A. RADIATION DAMAGE PIXEL DIGITISER

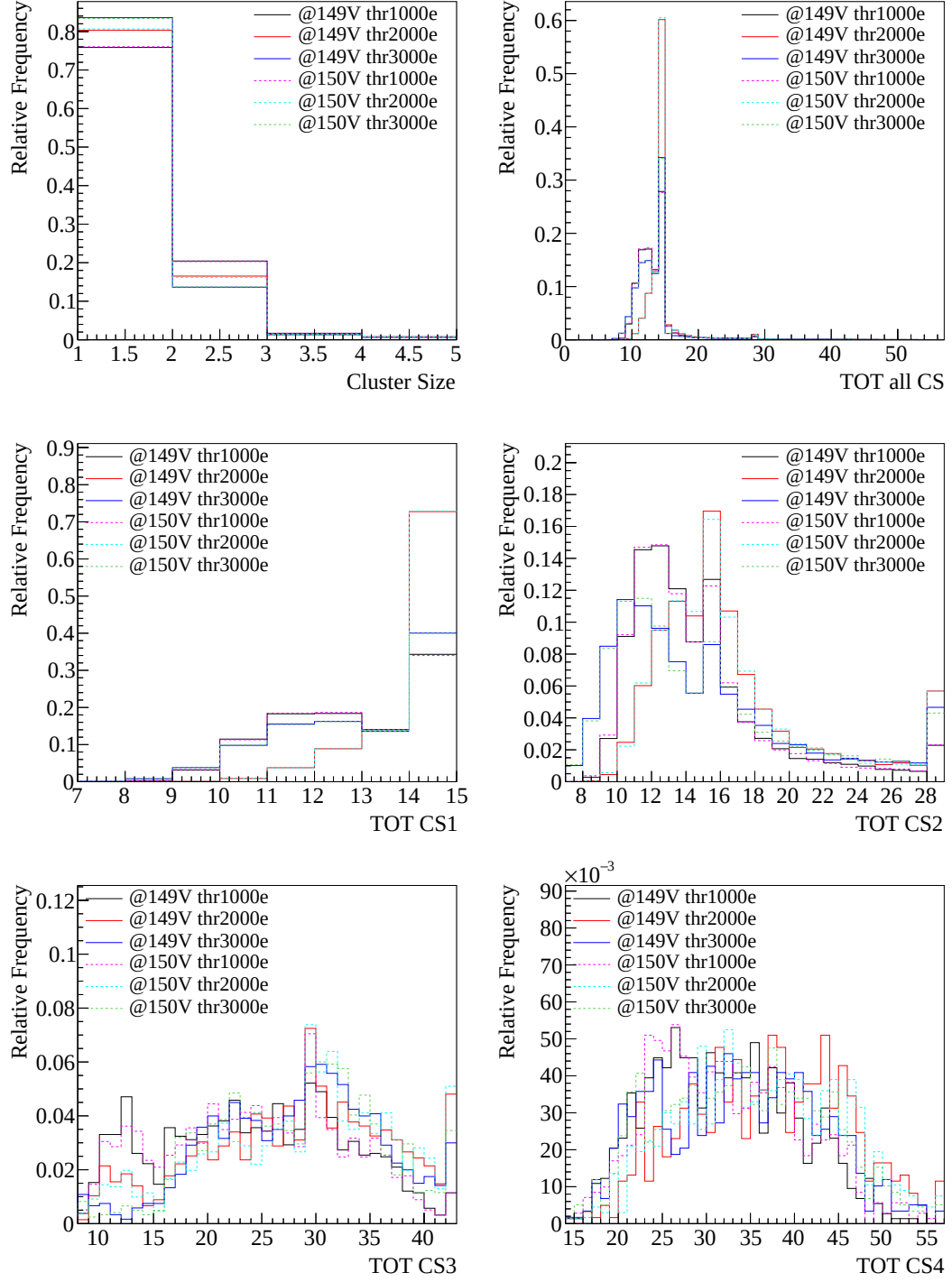


Figure A.11: Comparison of effects between generated linear electric field and simulated using TCAD.

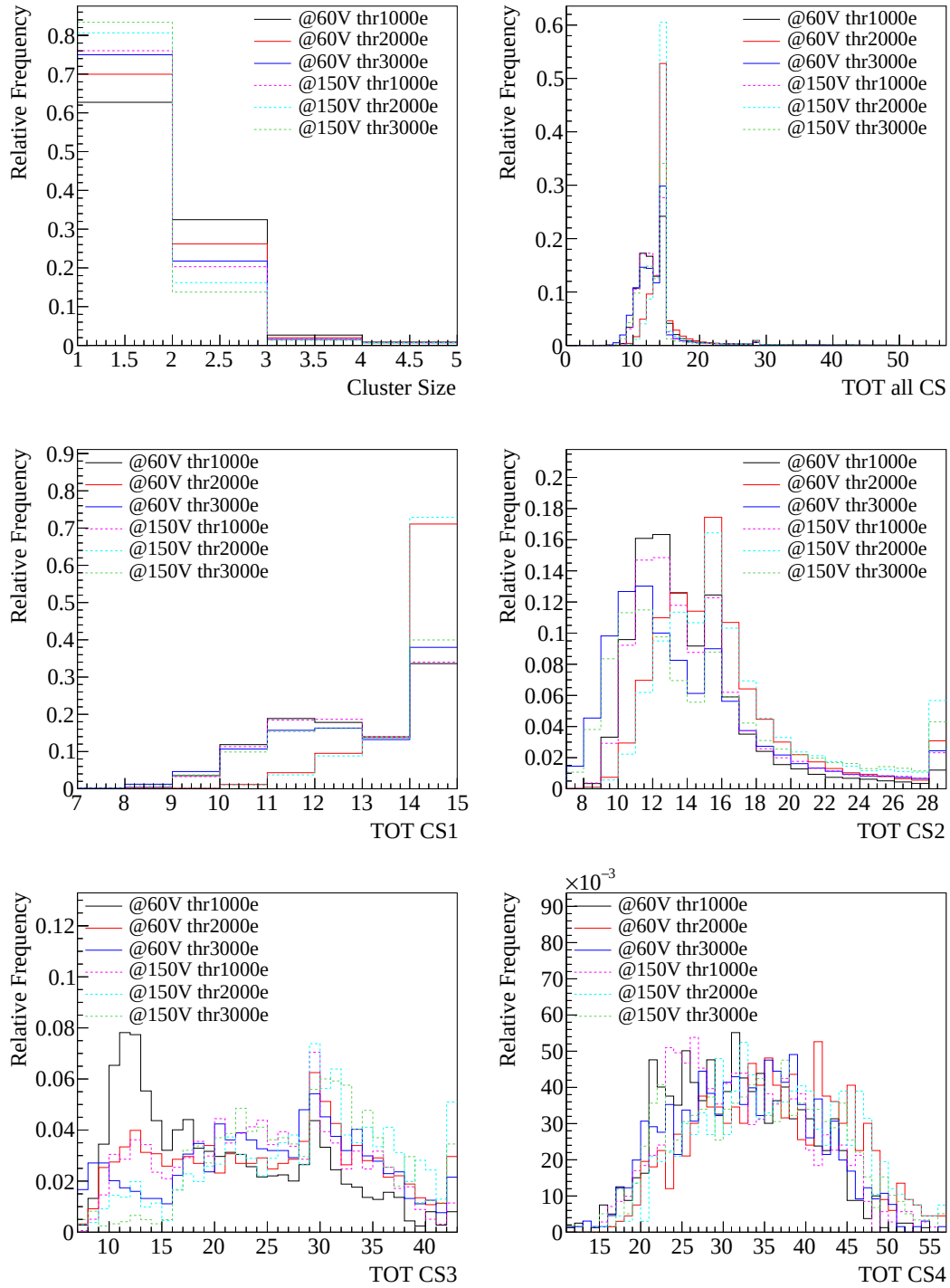


Figure A.12: Comparison of simulated unirradiated sensors at different bias voltages.

APPENDIX A. RADIATION DAMAGE PIXEL DIGITISER

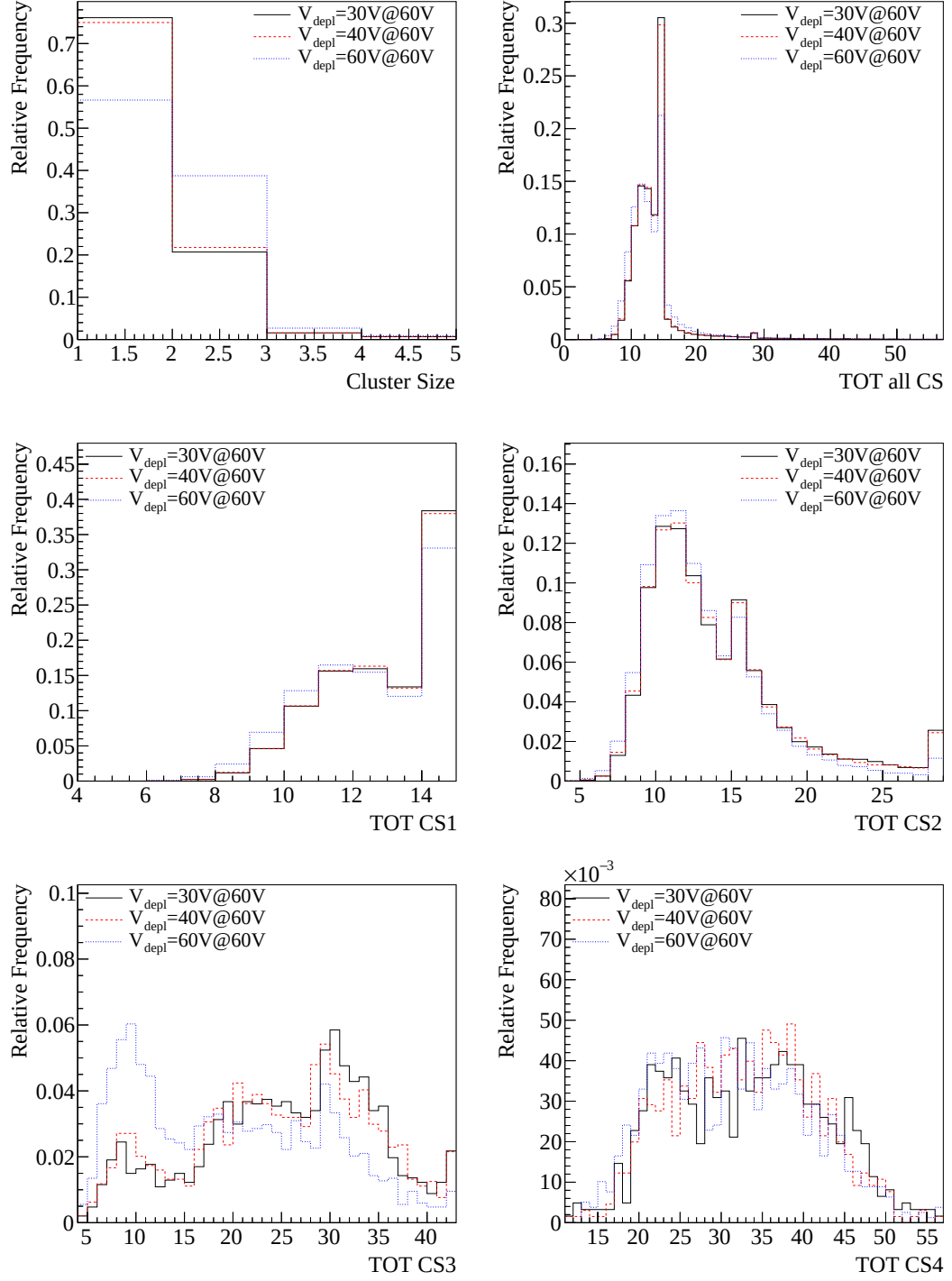


Figure A.13: Comparison of simulated unirradiated sensors with different depletion voltages.

Appendix B

TCAD Comparisons

B.1 TCAD Models

Physics	Model	Sub-model
Mobility	DopingDependence	
	Enormal	
	HighFieldSaturation	
	CarrierCarrierScattering	
Recombination	SRH	DopingDependence
		Tunneling (Hurkx)
	Auger	WithGeneration
	Band2Band	Hurkx
	Avalanche	UniBo
EffectiveIntrinsicDensity		Eparallel
QCvanDort	BandGapNarrowing	Slotboom

Table B.1: Physics models used in TCAD simulations.

B.2 Normal vs. Reduced Physics

In section 5.2 it was mentioned that the default list of physics as shown in Table B.1 are used, unless other physics are specified, e.g. radiation damage, MIP or capacitance (AC analysis). For these latter simulation using minimum physics significantly reduces the simulation time. The reduced physics list

APPENDIX B. TCAD COMPARISONS

consists only of doping dependent models plus the effective intrinsic density model. The impact is shown in Fig. B.1, Fig. B.2 and Fig. B.3. The electric field is slightly weaker, the leakage current is less, but the depletion depth is unchanged. In charge collection the signal shape is affected and thus the accumulated charge.

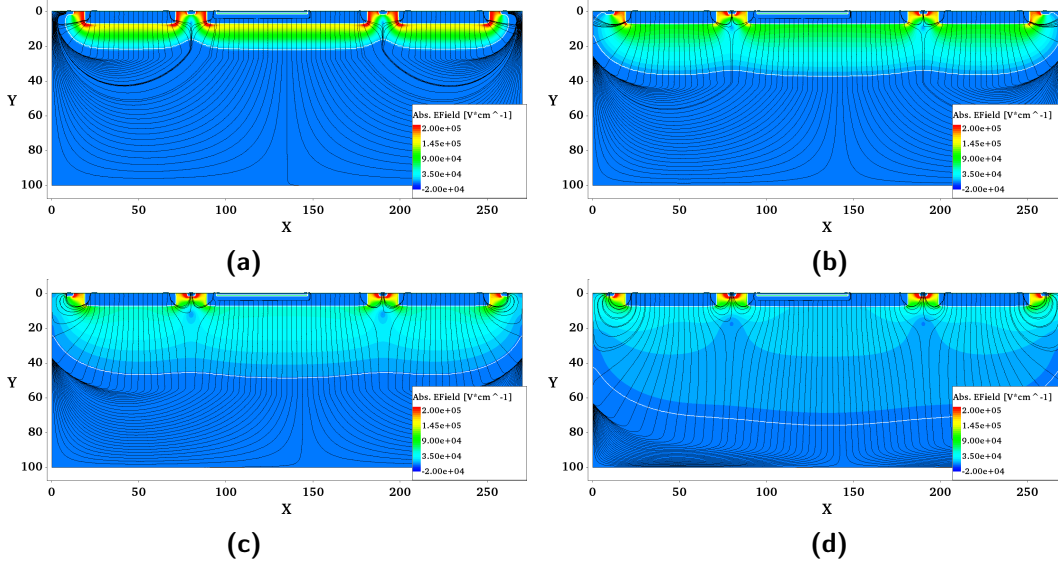


Figure B.1: Absolute electric field strength of H35DEMO for resistivities 20 (a), 80 (b), 200 (c) and 1000 Ωcm (d), biased from the top at -120 V with reduced physics.

B.3 Periodic Boundary Condition

Periodic boundary condition is not used for most of the simulation but is applied in the comparison of the measured leakage current. The difference in the electric field, I-V, depletion depth and capacitance is shown in Fig. B.4, Fig. B.5 and Fig. B.6.

B.4 Alternative Cross-section

The alternative cross-section used for most of the pixels of the H35DEMO without the DPTUBs used for biasing between the DNTUBs. Comparisons of

APPENDIX B. TCAD COMPARISONS

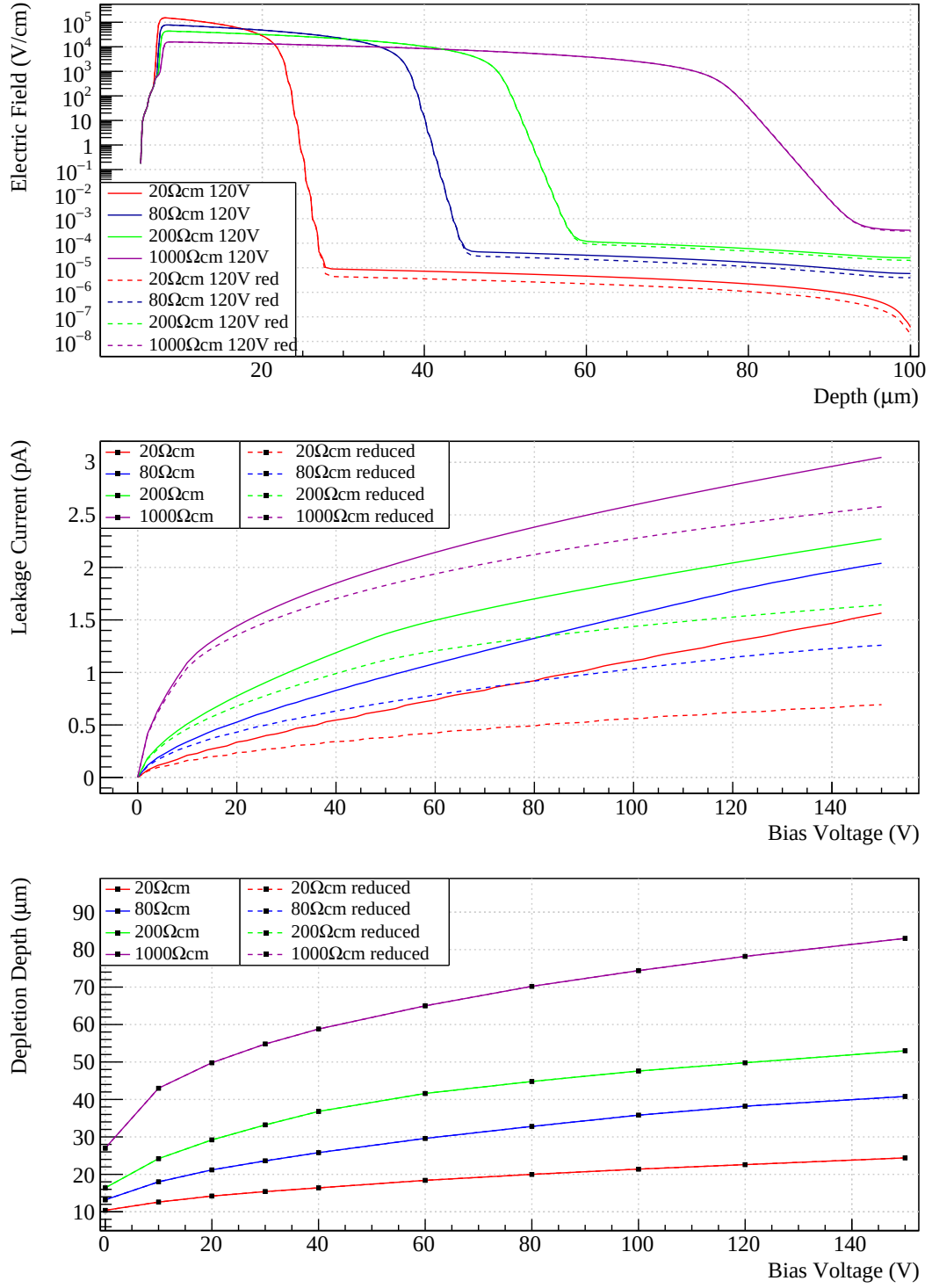


Figure B.2: Comparison of electric field, leakage current and depletion depth using reduced set of physics models.

APPENDIX B. TCAD COMPARISONS

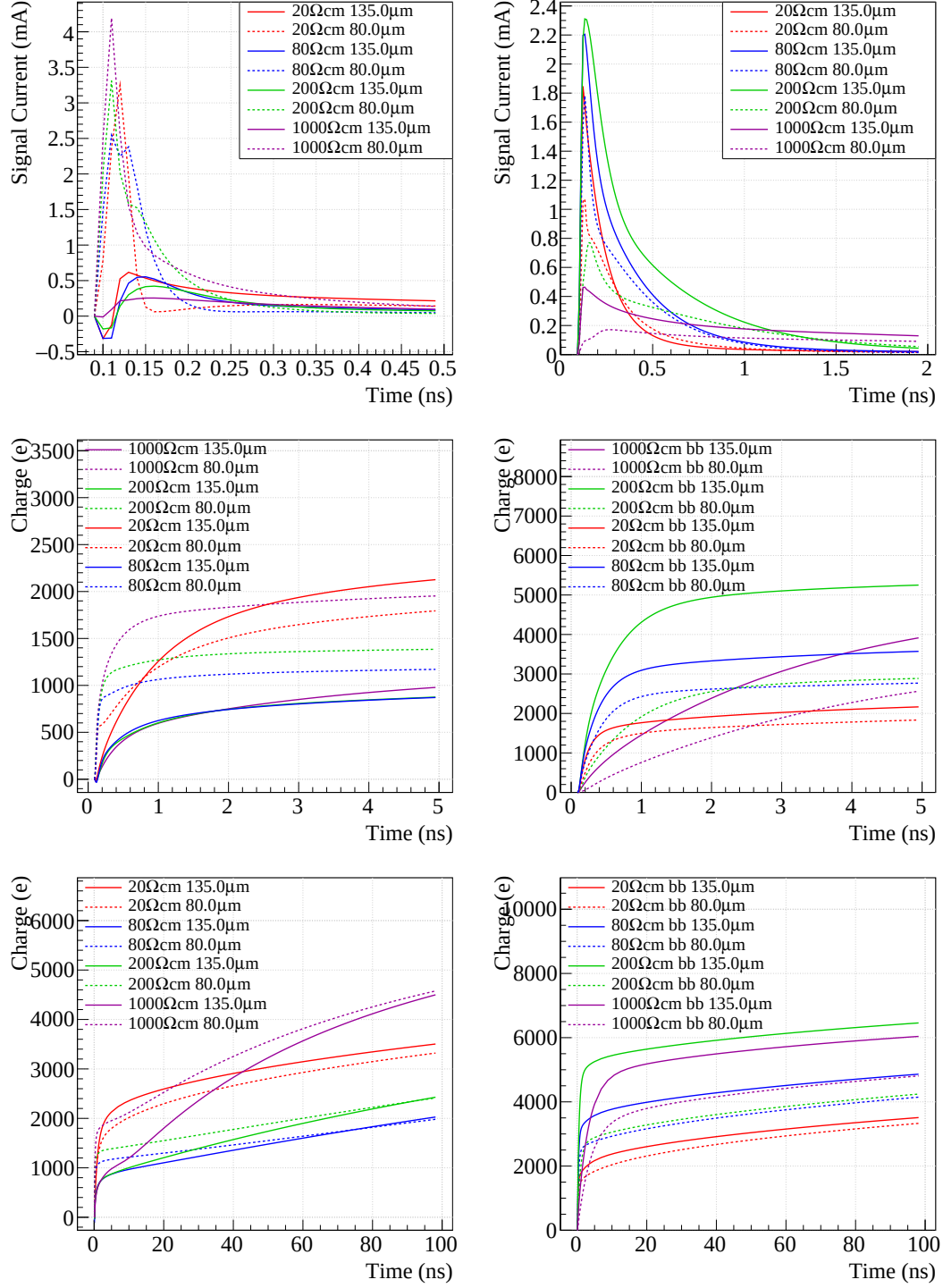


Figure B.3: Signal (top) and accumulated charge (middle and bottom) of a MIP entering the pixel at different positions for top-side-biased (left) and back-side-biased sensor (right) simulated with reduced physics.

APPENDIX B. TCAD COMPARISONS

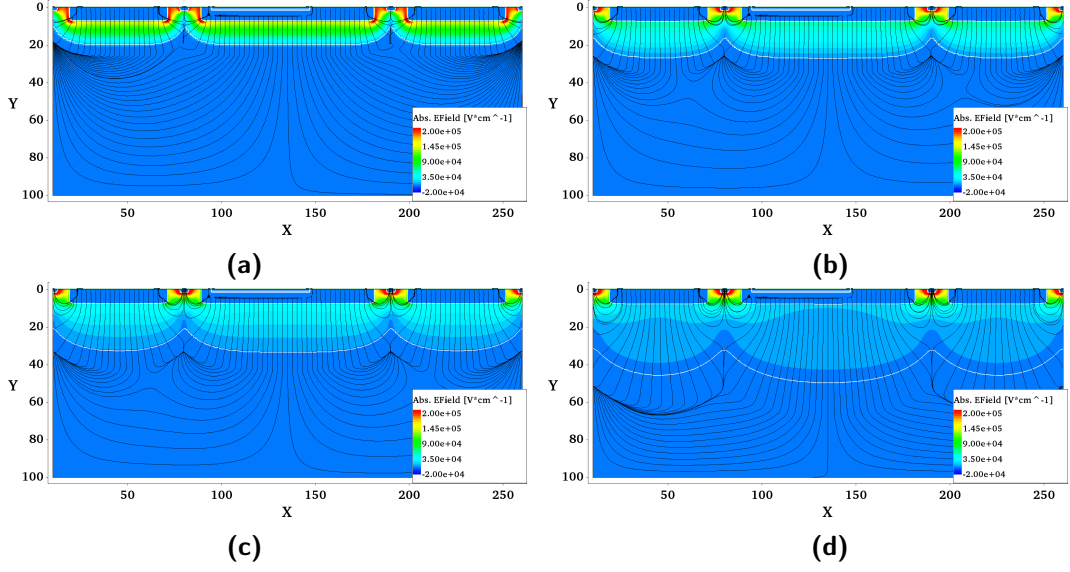


Figure B.4: Absolute electric field strength of H35DEMO for resistivities 20 (a), 80 (b), 200 (c) and 1000 Ωcm (d), biased from the top at -120 V with periodic boundary condition.

the electric field, leakage current, depletion region and capacitance are shown in Fig. B.7, Fig. B.8 and Fig. B.9. Test beam data show that this layout has a better efficiency due to smaller input capacitance which comes from the absence of two DPTUBs between the DNTUBs as shown in Fig. B.9 bottom. When the depletion regions of the separate DNTUBs grow and merge into one depletion region the extra DPTUBs contribute a constant amount to the capacitance. In their absence also the fill factor increases which benefits the charge collection. The increase in depletion region does not come primarily from the region occupied by the DPTUBs. This contribution is tiny since the space charge there is partially below the threshold and will be therefore accounted for depletion region. The main contribution comes from the larger shape further down in the substrate as shown in Fig. B.10 where the red line indicates the depletion region of the default cross-section with extra DPTUBs and the white line indicates the alternative cross-section without the extra DPTUBs.

APPENDIX B. TCAD COMPARISONS

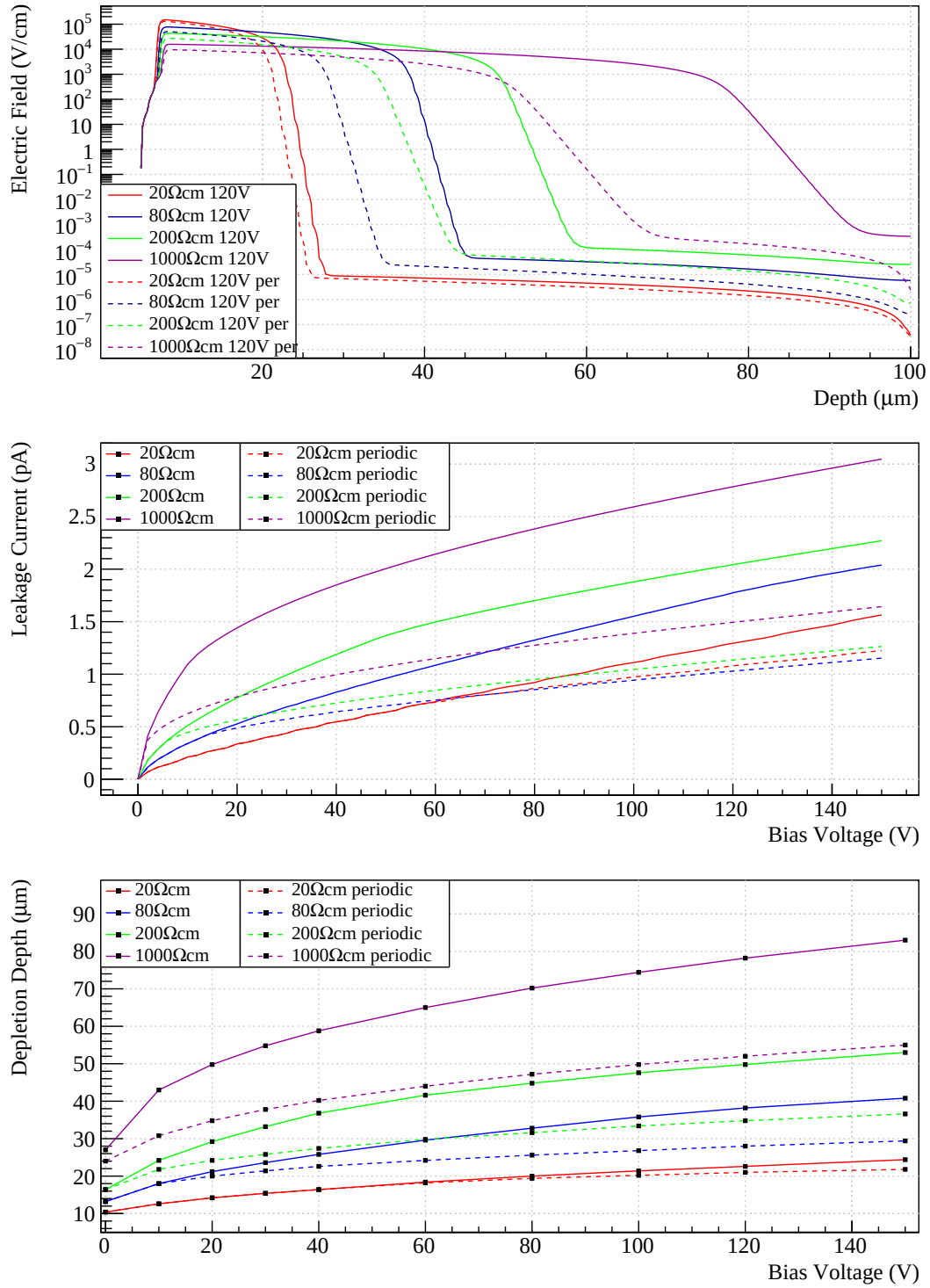


Figure B.5: Electric field, leakage current and depletion depth with periodic boundary condition.

APPENDIX B. TCAD COMPARISONS

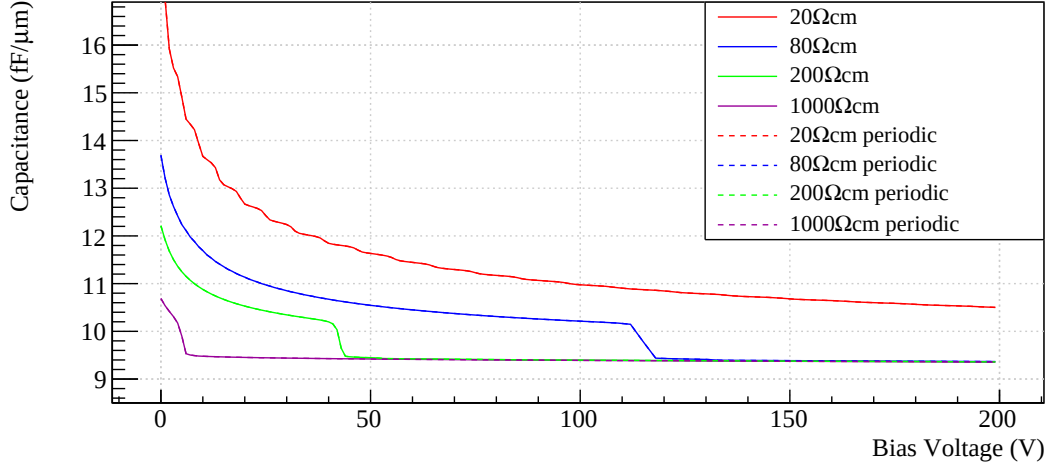


Figure B.6: As expected, applying periodic boundary condition does not change the total capacitance of the pixel.

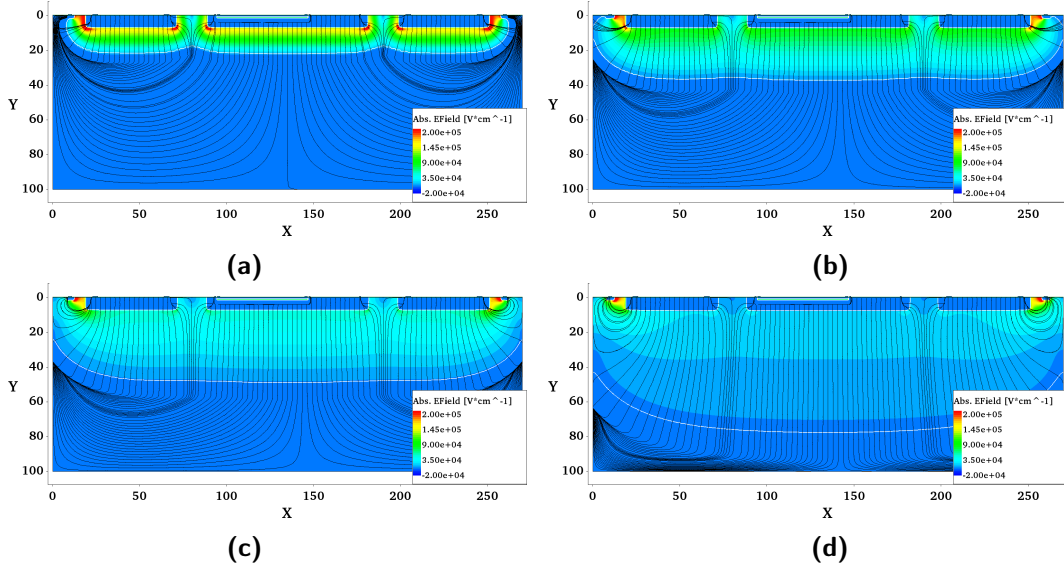


Figure B.7: Absolute electric field strength of H35DEMO for resistivities 20 (a), 80 (b), 200 (c) and 1000 Ωcm (d), biased from the top at -120 V for the alternative cross-section without biasing DPTUBs.

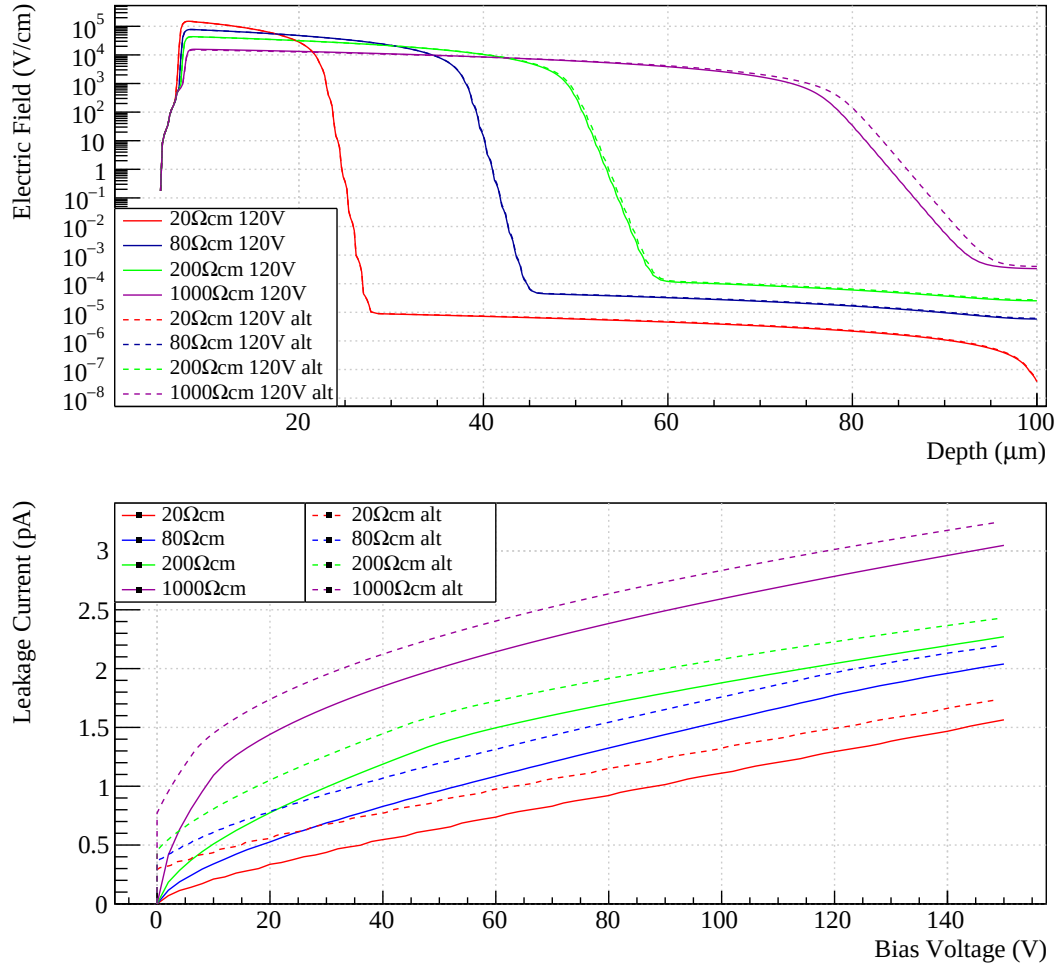


Figure B.8: Electric field strength and I-V characteristics of the alternative cross-section compared with the default one.

APPENDIX B. TCAD COMPARISONS

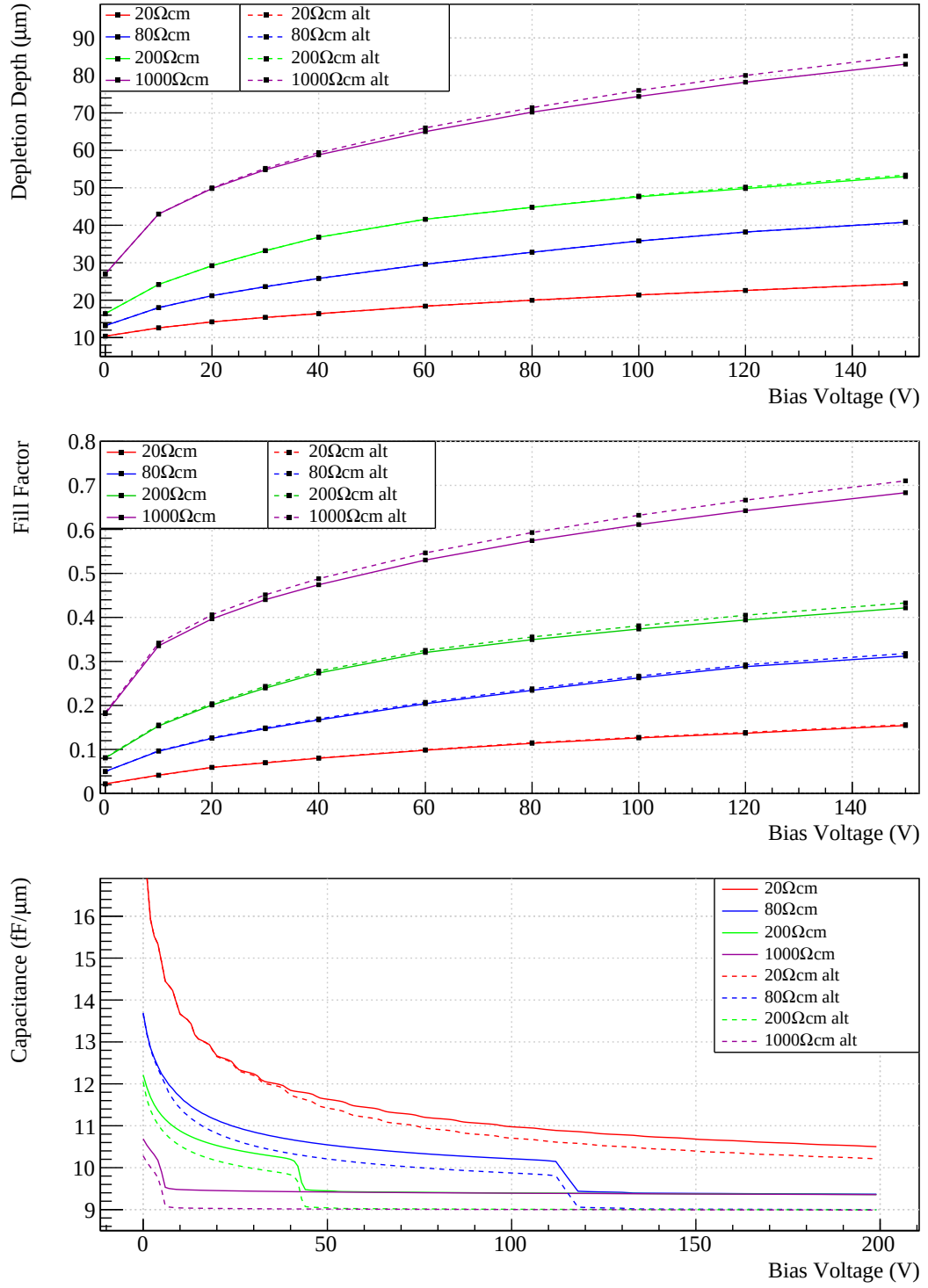


Figure B.9: Depletion depth and capacitance of the alternative cross-section compared with the default cross-section.

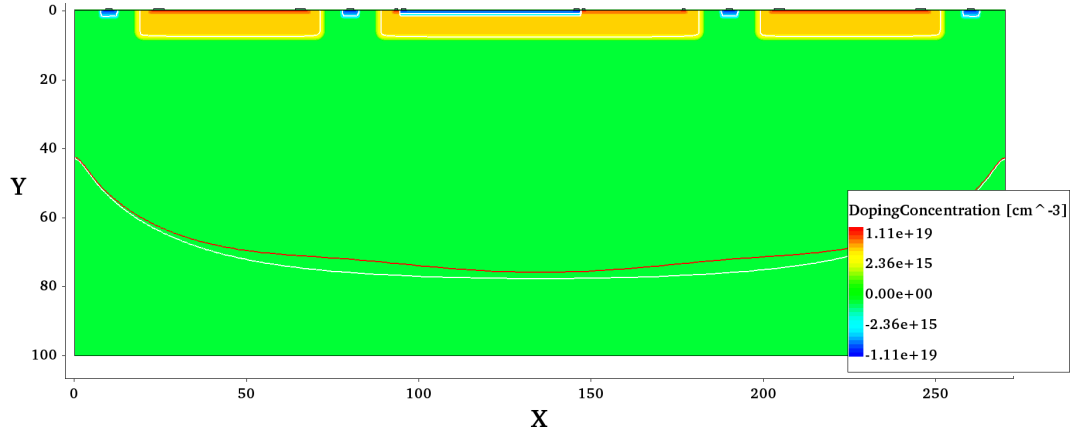


Figure B.10: Comparison of the depletion region of the default (red line) and alternative (white line) cross-sections.

B.5 Radiation Models

Different radiation models exist for TCAD simulations. While the Perugia model is shown in Table 5.3, the according RD50 model is shown in Table B.2 [166]. The difference in the used radiation damage model is shown below in Fig. B.11, Fig. B.12 and Fig. B.13.

Type	Energy (eV)	σ_e (cm ²)	σ_h (cm ²)	η (cm ⁻¹)
Acc.	$E_C - 0.42$	2.0×10^{-15}	2.0×10^{-14}	1.613
Acc.	$E_C - 0.46$	5.0×10^{-15}	5.0×10^{-14}	0.9
Acc.	$E_C - 0.10$	2.0×10^{-15}	2.5×10^{-15}	100
Don.	$E_V + 0.36$	2.5×10^{-14}	2.5×10^{-15}	1.9

Table B.2: Radiation damage model for p-type silicon “RD50” model from [166].

APPENDIX B. TCAD COMPARISONS

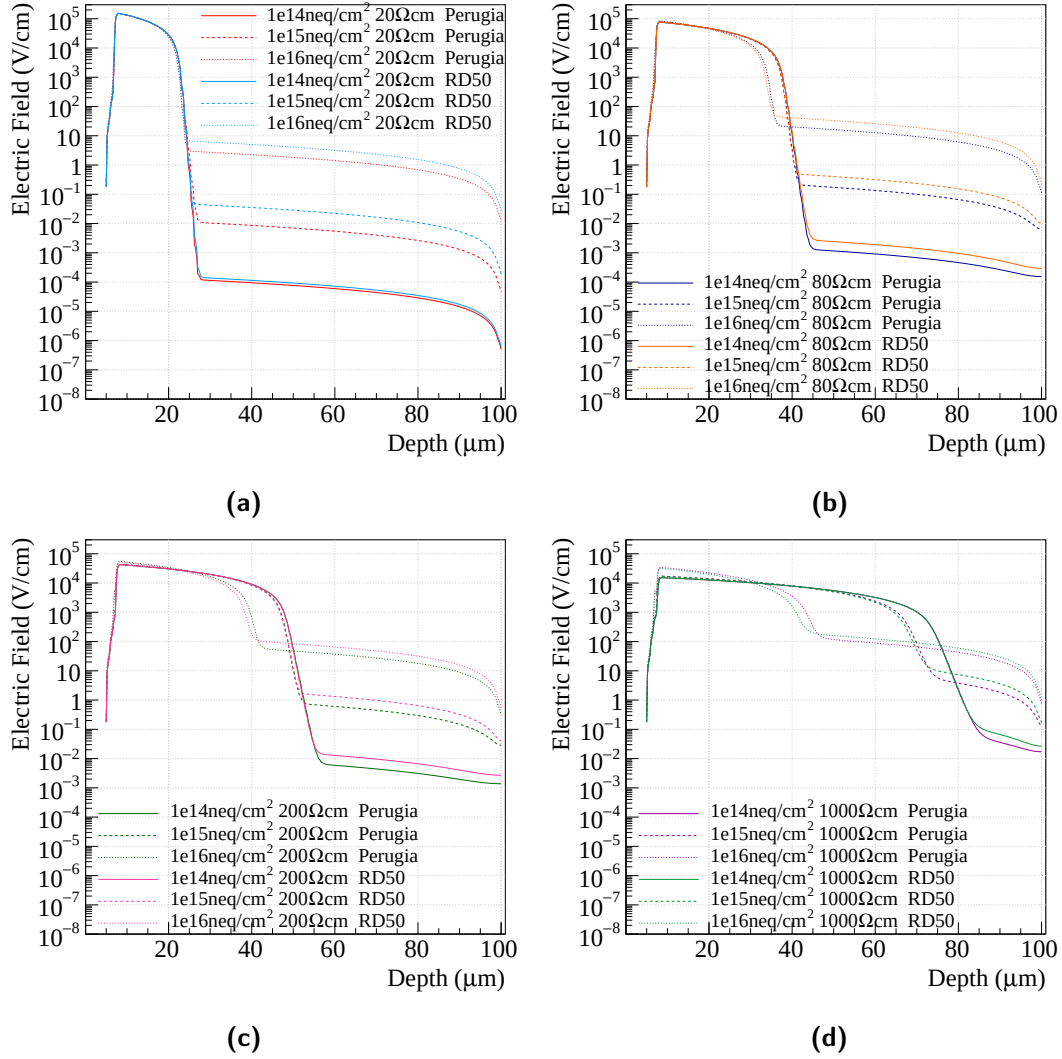


Figure B.11: Comparison of electric field simulated using Perugia and RD50 model at -120 V for fluences of $1 \cdot 10^{14}$, $1 \cdot 10^{15}$ and $1 \cdot 10^{16}$ $\text{n}_{\text{eq}}/\text{cm}^2$ for all resistivities.

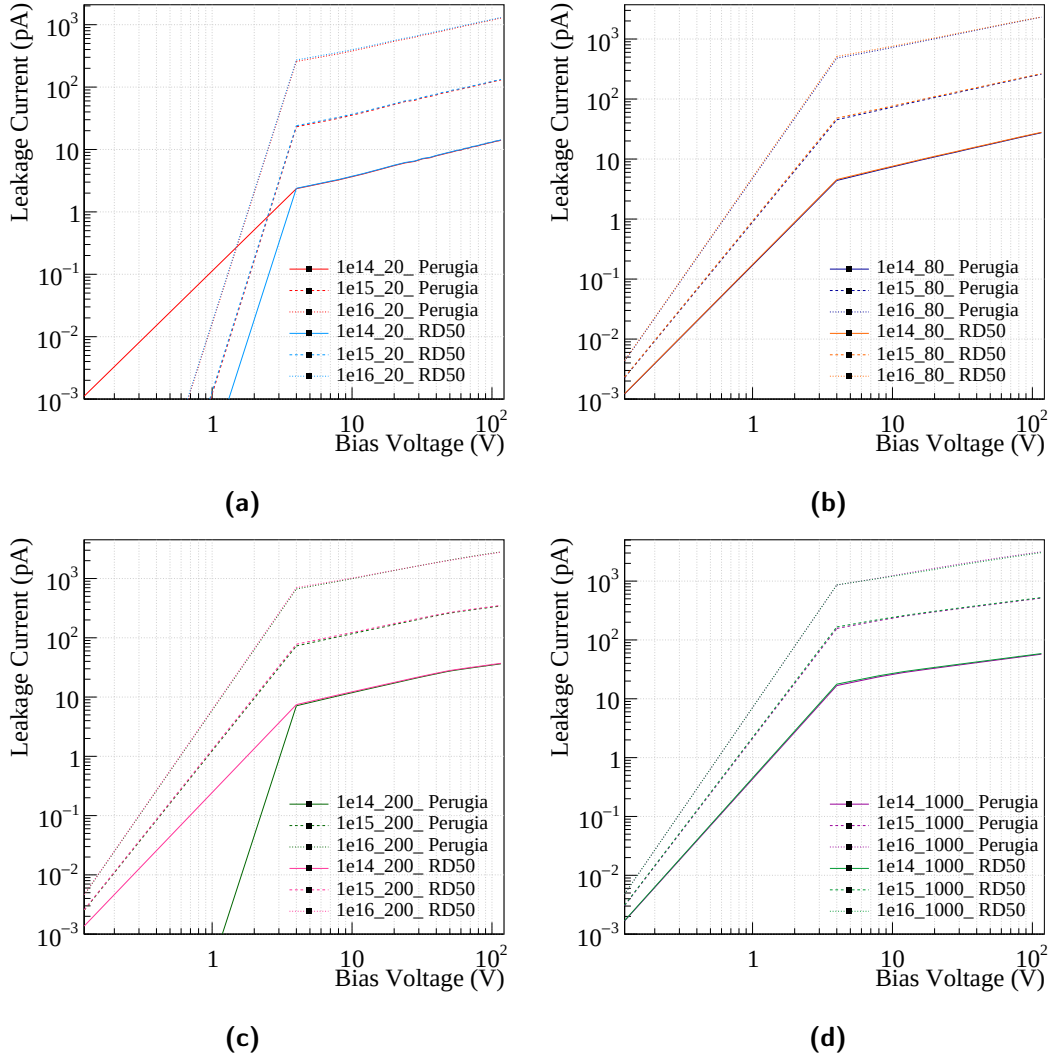


Figure B.12: Comparison of electric field simulated using Perugia and RD50 model for fluences of $1 \cdot 10^{14}$, $1 \cdot 10^{15}$ and $1 \cdot 10^{16} \text{ n}_{\text{eq}}/\text{cm}^2$ for all resistivities. The legend is labeled as “fluence_resistivity_model” with resistivity in Ωcm .

APPENDIX B. TCAD COMPARISONS

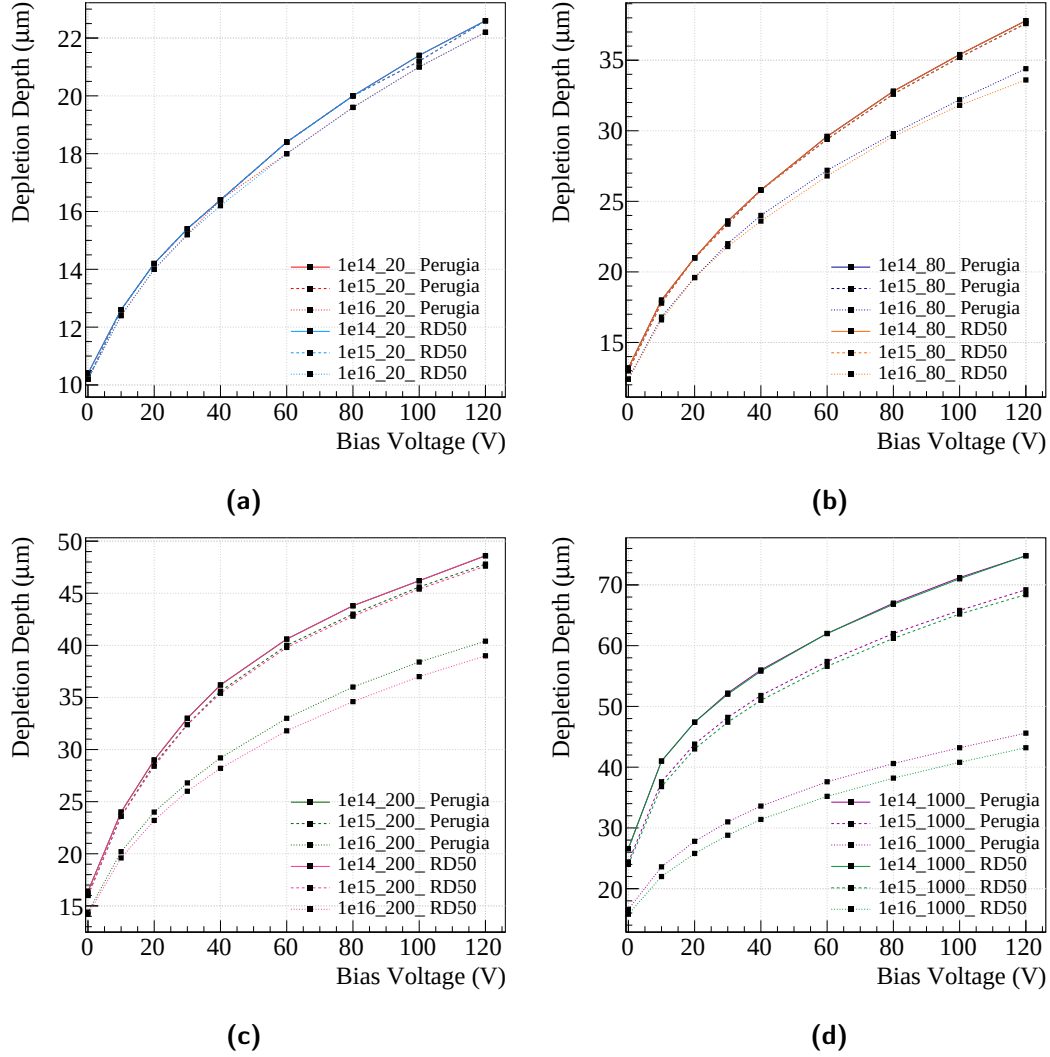


Figure B.13: Comparison of depletion depth simulated using Perugia and RD50 model for fluences of $1 \cdot 10^{14}$, $1 \cdot 10^{15}$ and $1 \cdot 10^{16} \text{ n}_{\text{eq}}/\text{cm}^2$ for all resistivities. The legend is labeled as “fluence_resistivity_model” with resistivity in Ωcm .

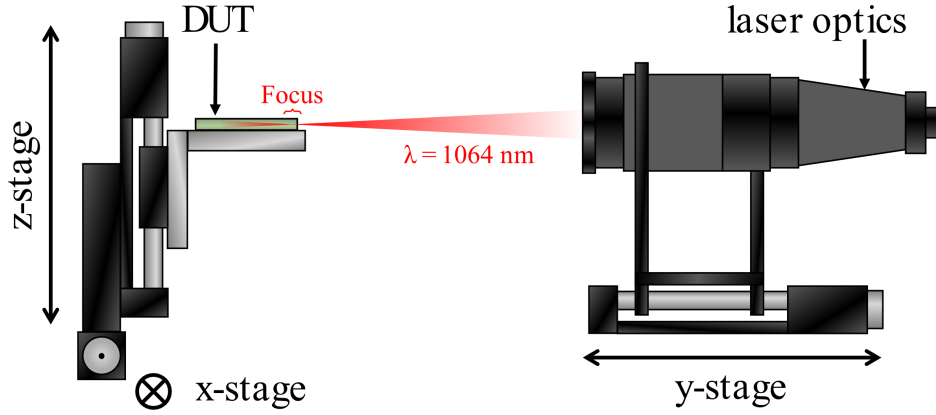


Figure B.14: Schematic of the TCT setup [167].

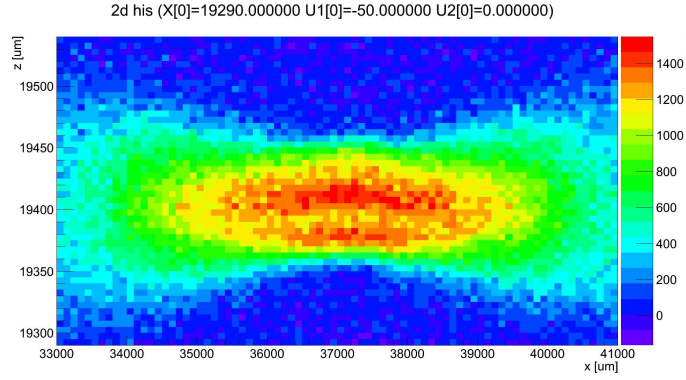


Figure B.15: TCT focus scan along the horizontal axis. The z-axis is the depth of the sensor bulk, the colour scale of the plot is the collected charge in arbitrary units. The centre of the image shows clear structures of the sensor, indicating the length of the horizontal focus of about 1 mm [167].

B.6 Focus Length of the Laser

Appendix C

Measurements

C.1 I-V Curves of Test Beam Samples

The sensors are produced on differently numbered wafers. Wafer 1–6 are $20\ \Omega\text{cm}$, wafer 7–12 are $80\text{--}100\ \Omega\text{cm}$, wafer 13–18 are $200\text{--}300\ \Omega\text{cm}$ and wafer 19–24 are $1000\text{--}2000\ \Omega\text{cm}$. On each wafer the sensors on the tape are numbered according to a schematic e.g. W01-01 which corresponds to the first sensor on wafer 1.

I-V characteristics of most samples used in the Fermilab test beam measured using the automatic probe station. Here the compliance limit was set much higher than with the manual probe station: $10\ \mu\text{A}$ for the $1000\ \Omega\text{cm}$ back-side-processed sensors (Fig. C.1 top) and $5\ \mu\text{A}$ for the regular ones (Fig. C.1 centre and bottom and Fig. C.2). One of the originally four thinned sensors with processed backside had passivation on the pads and therefore could not be used. The logarithmic plot of the three remaining sensors indicates a plateau towards higher voltages which could have been a hint to the then not understood rise-and-flatten (RAF) effect. From the regular sensors, the first analog matrix of W07-07 and the second analog matrix of W13-05 are broken. Therefore the measurements are not shown here.

APPENDIX C. MEASUREMENTS

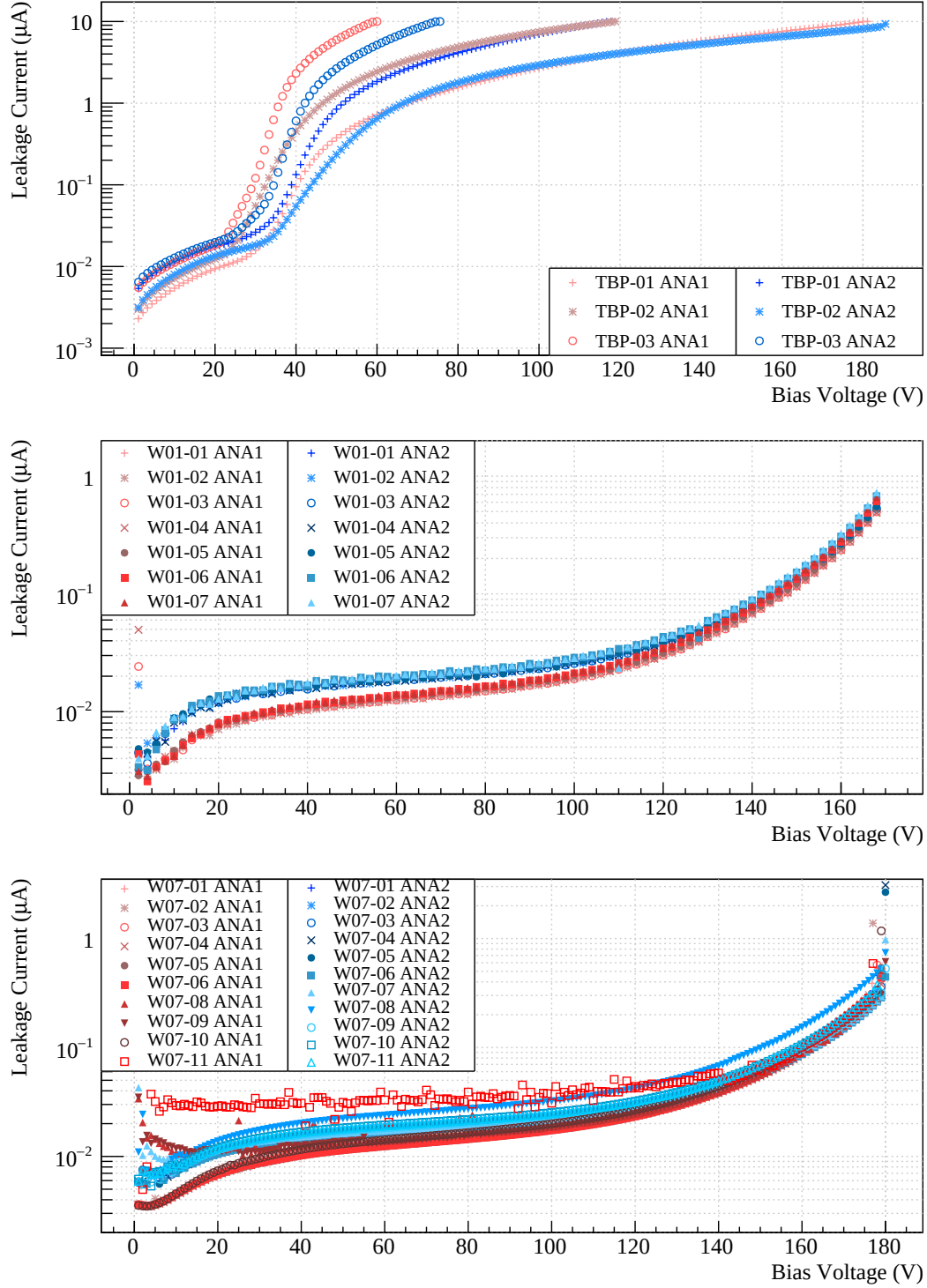


Figure C.1: I-V measurements of H35DEMO sensors for the Fermilab test beam. Top: thinned H35DEMO sensors with 1000 Ωcm and a processed back side to apply bias voltage. Middle and bottom: standard 20 and 80 Ωcm sensors.

APPENDIX C. MEASUREMENTS

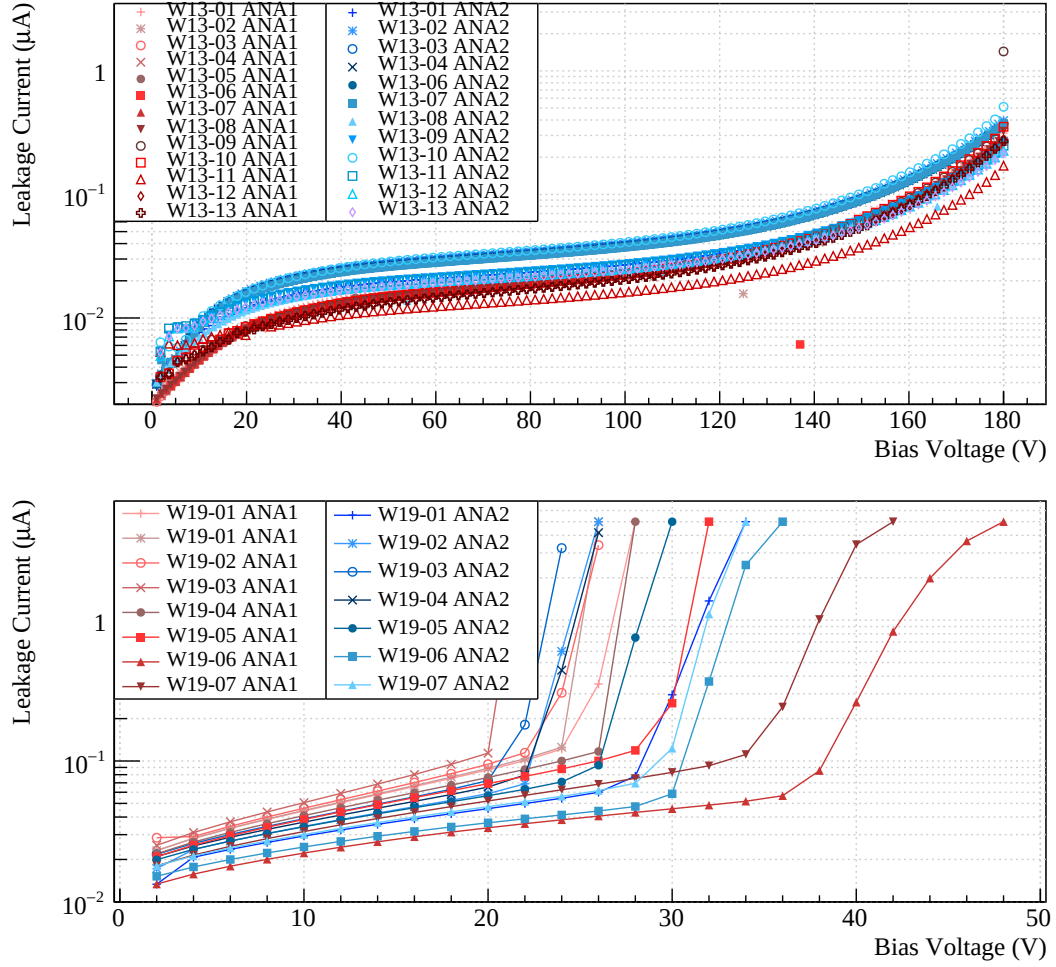


Figure C.2: I-V measurements of H35DEMO sensors for the Fermilab test beam. 200 (top) and 1000 Ωcm (bottom).

C.2 Measurements of the NMOS Matrix

Default DAC values used for the NMOS measurements in Table C.1.

DAC #	Signal Name	nMOS	Default Value
DAC 0	VNHBdig/VPHBdig	VNHBdig/VPHBdig	40
DAC 1	VPDelDigital	VPDel	30
DAC 2	VNDelDigital	VNDel	30
DAC 3	VPTrimDigital	VPTrim	0
DAC 4	VNCompDigital	VNComp	10
DAC 5	VBLResDigital	BLR	32
DAC 6	VBLRes	BLRpix	60
DAC 7	VPBiasRes/VNBiasRes	NC/VNBiasPix	52
DAC 8	VNFB	VNFBPix	15
DAC 9	VPTrim	VPTrimPix	0
DAC 10	VNTWDown	VNTwDownPix	30
DAC 11	VNTW	VNTwPix	38
DAC 12	VNLogic	VNLogicPix	30
DAC 13	VPLoadAmp	VPLoadPix	2
DAC 14	VNSF	VNSFPix	5
DAC 15	VPamp/VNAmp	NC/VNPix	19
DAC 16	VPAB/VNHB	VPABPix/NC	15

Table C.1: Table of default DAC values used for measurements of the NMOS matrix.

List of Figures

1.1	Neutral current observation in the Gargamelle bubble chamber.	3
2.1	CERN aerial view	6
2.2	A bunch crossing with 25 reconstructed pile-up events.	7
2.3	CERN accelerator complex	9
2.4	LHC/HL-LHC schedule.	9
2.5	The ATLAS detector.	10
2.6	The ATLAS inner detector.	11
2.7	ATLAS calorimeters.	13
2.8	Section of the FCAL and assembly of an electrode.	14
2.9	Schematic of the detectors in the muon spectrometer.	15
2.10	FLUKA simulation of NIEL fluence of the ITk region.	17
2.11	Inclined layout of the ITk.	18
2.12	Standard Model - quarks and leptons	20
2.13	ATLAS SM production cross-sections.	21
3.1	Photon cross-section in lead vs. photon energy.	26
3.2	Stopping power in Bethe range.	28
3.3	Stopping power of μ^+ in copper in full energy range.	29
3.4	Contributions of electron energy loss.	31
3.5	Displacement damage functions.	33
3.6	Bravais lattice and Wigner-Seitz cell.	34
3.7	Brillouin zones.	35
3.8	Probability densities of the nearly-free electron model.	37
3.9	Energy band structure.	37
3.10	Insulator, metal, semi-metal	38

LIST OF FIGURES

3.11 Indirect optical transition via phonon.	38
3.12 Diamond lattice.	40
3.13 Donor and acceptor	41
3.14 P-N junction.	43
3.15 Diode I-V characteristics.	45
3.16 Diode C-V characteristics.	46
3.17 Hall effect.	47
3.18 MOSFET schematic in 2D.	49
3.19 IV characteristics of a MOSFET.	49
3.20 Lift-off and deposit and etch photolithography.	51
3.21 Linear superposition for Ramo setup.	53
3.22 Charge collection according to Shockley-Ramo theorem.	54
3.23 Defects generated by a 50 keV primary knock-on atom	56
3.24 Energy levels of defects.	57
3.25 Type inversion.	58
3.26 Increase of leakage current with fluence.	58
3.27 Resistivity vs. fluence.	60
3.28 Change in N_{eff} as a function of annealing time.	61
4.1 AllPix event display	64
4.2 Electric field with fits and generated 1D linear fields.	66
4.3 1D electric field maps.	67
4.4 Distance maps for electrons and holes.	69
4.5 Distance fits for electrons and holes.	70
4.6 Distance maps after fit for electrons and holes.	71
4.7 Time maps for electrons and holes.	72
4.8 Ramo maps comparison.	73
4.9 Lorentz angle.	74
4.10 Lorentz maps for electrons and holes.	75
4.11 1D electric field at different fluences.	76
4.12 Distribution of collected charge depending on the number of subcharges.	77
4.13 ToT-charge plot	78
4.14 EUDET telescope Caladium and DUT with dimensions.	80
4.15 Event multiplicity for first telescope plane and the DUT.	83

LIST OF FIGURES

4.16	Hitmaps of the unirradiated double module.	84
4.17	Hitmaps of LUB2.	85
4.18	Total cluster size, timing and ToT.	86
4.19	Timing vs. ToT.	87
4.20	Cluster size, timing and ToT for run 300 and run 715.	88
4.21	Noise cuts on reconstructed runs.	89
4.22	Hitmap after reconstruction and cuts.	90
4.23	Hitmap simulation.	91
4.24	Residual distribution of run 715 and simulation.	93
4.25	Track position within a pixel, data.	94
4.26	Track position within a pixel, simulation.	95
4.27	Comparison of irradiated data with simulation.	97
4.28	Comparison of unirradiated data with simulation.	98
5.1	Schematic of a HV-CMOS pixel and glue bonding.	100
5.2	Cross-section of an ATLAS pixel module.	100
5.3	Sensor layout of H35DEMO	102
5.4	H35DEMO pixel cross-sections.	103
5.5	Schematic of the TCAD framework.	104
5.6	TCAD mesh	105
5.7	Simulated H35DEMO doping profiles.	109
5.8	Abs. electric field of H35DEMO standard layout.	111
5.9	Abs. electric field of H35DEMO with p-spray.	112
5.10	Abs. electric field of H35DEMO back-side-biased.	112
5.11	Abs. electric field of standard layout, with p-spray and back-side-biased.	113
5.12	Simulated H35DEMO leakage current vs. bias voltage.	114
5.13	Simulated depletion depth for all layouts and resistivities.	115
5.14	Simulated H35DEMO fill factor of the depleted area vs. bias voltage.	115
5.15	Space charge density at $x = 135 \mu\text{m}$	116
5.16	H35DEMO sensor capacitance.	116
5.17	Depletion region extends beyond the cross-section.	117
5.18	Absolute electric field for different fluences.	118
5.19	Leakage current for different fluences.	119

LIST OF FIGURES

5.20	Simulated H35DEMO depletion depth vs. bias voltage for different fluences.	120
5.21	MIP positions.	121
5.22	Signal and charge of a MIP at different positions.	122
5.23	Current through the guard ring.	123
5.24	MIP positions.	123
5.25	MIP signal through PTUB and SNTUB.	124
5.26	MIP signal at different bias voltages.	125
5.27	MIP charge at different bias voltages.	126
5.28	Depletion depth measured using eTCT.	127
5.29	Electric field of the eTCT simulation at -100 V.	128
5.30	MIP passing through the test structure.	129
5.31	Simulated eTCT signals for 20 and 1000 Ωcm substrates at different depths.	130
5.32	Simulated eTCT scan.	132
5.33	Electric field before and after a particle passed through.	132
5.34	Simulated eTCT scan of 1000 Ωcm sensor with longer integration time.	133
5.35	Pixel cross-section for LF MPW submission.	133
5.36	Comparison of doping profiles	134
5.37	Comparison of doping profiles of LF simulation.	135
5.38	Contributions to total sensor capacitance.	136
5.39	Cross-sections of the intended and manufactured pixel of the 2016 submission.	137
5.40	Schematic of area and perimeter components.	139
5.41	Comparison of simulated capacitance.	139
5.42	Cross-section displaying the depleted regions.	141
5.43	Contribution of area, edge and corner capacitance.	141
5.44	LF 2016 MPW layout with a small PSUB.	142
5.45	Sensor capacitance with full PSUB and small PSUB.	142
5.46	Different size of PSUB overhang.	143
5.47	I-V curve of LFoundry 2016 MPW cross-section.	144
5.48	Current densities at -120 V and -400 V for full PSUB.	145
5.49	Current densities at -120 V and -400 V	146

LIST OF FIGURES

5.50	LF cross-sections for submission in 2017.	147
5.51	Current through GNDA and GNDD.	149
5.52	Response to signal injected through GNDA.	149
5.53	Comparison of capacitance for cross-section 1 and 2.	150
5.54	PW gap	150
5.55	Submitted cross-section for RD50 MPW.	151
5.56	H35DEMO I-V measurement probe station setup and probing pads.	152
5.57	H35DEMO test structure and bias pads.	153
5.58	I-V measurements of H35DEMO sensors.	154
5.59	I-V measurement of all resistivities at UniGe and UoL.	155
5.61	Temperature distribution during I-V measurements.	156
5.62	Temperature dependence of bandgap energy of silicon.	157
5.63	Depletion depth comparison with eTCT.	160
5.64	Junction lines for different resistivities.	161
5.65	Aerial view of the Fermilab Test Beam Facility.	162
5.66	FE-I4 telescope.	162
5.67	Dispensing glue onto H35DEMO.	163
5.68	CaRIBou chip board with a H35DEMO assembly.	164
5.69	Reconstructed tracks.	165
5.70	Clocks of the standalone readout.	167
5.71	Analog signals for DAC13 values of 2 and 38.	170
5.72	Sketch of pulse injection schematic.	171
5.73	Analog response to pulse injection for different bias voltages. . .	172
5.74	Baseline of the measurement.	173
5.75	Analog signal with fits.	173
5.76	Distribution of analog signal amplitude.	174
5.77	Distribution of analog signal rise and fall times.	175
5.78	Distributions of analog output.	176
5.79	Distribution of analog signal for delay measurement.	177
5.80	Distribution of trigger timing and delay.	178
5.81	Scintillator setup	180
5.82	PMT amplification adjustment circuit.	181
5.83	Source scan over the NMOS matrix.	181

LIST OF FIGURES

A.1	ToT encoding	188
A.2	Distributions: multiplicity.	190
A.3	Distributions: multiplicity.	191
A.4	CS and ToT distributions depending on the temperature.	192
A.5	CS and ToT distributions depending on bias voltage.	193
A.6	CS and ToT distributions depending on position.	194
A.7	CS and ToT distributions of run 300 and 705.	195
A.8	Distributions of the unirradiated sensor.	196
A.9	Comparison of threshold and tuning (unirradiated sensor).	197
A.10	Simulation of irradiated sensor.	198
A.11	Effects of linear and TCAD electric field maps.	199
A.12	Simulated unirradiated sensors at different bias voltages.	200
A.13	Simulated unirradiated sensors with different depletion voltage.	201
B.1	Abs. electric field of H35DEMO reduced physics.	204
B.2	Comparison of properties using reduced physics models.	205
B.3	Signal and charge of a MIP with reduced physics.	206
B.4	Abs. electric field of H35DEMO periodic boundary condition.	207
B.5	Electric field, leakage current and depletion depth with periodic boundary condition.	208
B.6	Capacitance with periodic boundary condition.	209
B.7	Abs. electric field of H35DEMO alternative cross-section.	209
B.8	Electric field and I-V characteristics of the alternative cross- section.	210
B.9	Depletion depth and capacitance of the alternative cross-section.	211
B.11	Perugia vs. RD50 model, electric field.	214
B.12	Perugia vs. RD50 model, leakage current.	215
B.13	Perugia vs. RD50 model, depletion depth.	216
B.14	Schematic of the TCT setup.	217
B.15	TCT focus scan.	217
C.1	I-V of Fermilab test beam samples.	220
C.2	I-V of Fermilab test beam samples.	221

List of Acronyms

LHCb Large Hadron Collider Beauty

ALICE A Large Ion Collider Experiment

ATLAS A Toroidal LHC Apparatus

ams ams AG

BC bunch crossing

BCID bunch crossing ID

CaRIBou Control and Readout Inner-Tracking Board

CCE charge collection efficiency

CCPD capacitively coupled pixel device

CERN Conseil Européen pour la Recherche Nucléaire (European Organisation for Nuclear Research)

CKM Cabibbo-Kobayashi-Maskawa

CMOS complementary metal-oxide-semiconductor

CMS Compact Muon Solenoid

CSC Cathode Strip Chamber

C-V capacitance-voltage

CS cluster size

DAC digital-to-analog converter

DAF deterministic annealing filter

DAQ data acquisition

DOFZ diffusion oxygenated float zone

DM Dark Matter

DRS Domino Ring Sampler

DUT device under test

ENC equivalent noise charge

EOC end-of-column

EUDET European Detector R&D

FastCk fast clock

FCAL forward calorimeter

FEM finite element method

FE front-end

FE-I4 Front End chip IBM version 4

FPGA Field-Programmable Gate Array

FWHM full width at half maximum

FZ float zone

CZ Czochralski

FNAL Fermi National Accelerator Laboratory

FTBF Fermilab Test Beam Facility

.gds Graphical Design Station II

GUI graphical user interface

HL High Luminosity

HV-CMOS high voltage complementary metal-oxide-semiconductor

H35DEMO HV-CMOS demonstrator in ams AG (ams) 350 nm technology

HSIO high speed input/output

IBL Insertable B-Layer

ID Inner Detector

ITk Inner Tracker

IFAE Institut de Fisica d'Altes Energies

I-V current-voltage
JSI Jožef Stefan Institute
KIT Karlsruher Institut für Technologie
LAr Liquid Argon
LEP Large Electron-Positron Collider
LF LFoundry
LHC Large Hadron Collider
LS Long Shutdown
MAPS monolithic active pixel sensor
MET missing transverse energy
MDT Monitored Drift Tube
MIP minimum ionising particle
MOSFET metal-oxide-semiconductor field-effect transistor
MPV most probable value
MWPC multi-wire proportional chamber
MPW multi-project wafer
MSSM Minimal Supersymmetric Standard Model (SM)
NIEL non-ionising energy loss
NFE nearly-free electron
PCB printed circuit board
PDK process design kit
PMT photomultiplier tube
PS Proton Synchrotron
PSI Paul Scherrer Institute
QCD Quantum Chromodynamics
RAF rise-and-flatten
RCE Reconfigurable Cluster Element

RF radio frequency
ROD Readout Driver
RoI region of interest
ROC readout cell
RoCk readout clock
RPC Resistive Plate Chamber
SRH Shockley-Read-Hall
SM Standard Model
BSM Beyond SM
SCT Semiconductor Tracker
SyRes synchronous reset
SLAC Stanford Linear Accelerator Center
SPS Super Proton Synchrotron
STI shallow trench isolation
SUSY supersymmetry
TCAD Technology Computer Aided Design
TCT Transient Current Technique
eTCT edge-Transient Current Technique (TCT)
TDR technical design report
TGC Thin Gap Chamber
TID total ionising dose
TW time walk
TileCal Tile Hadronic Calorimeter
ToT time-over-threshold
TRT Transition Radiation Tracker
TTL transistor-transistor logic
UniGe Université de Genève

UoL University of Liverpool

XML Extensible Markup Language

Bibliography

- [1] E Rutherford. Address of the President, Sir Ernest Rutherford, O. M., at the Anniversary Meeting, November 30, 1927. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 117(777):300–316, 1928.
- [2] JD Cockcroft and ETS Walton. Experiments with high velocity positive ions. *Proc. R. Soc. Lond. A*, 129(811):477–489, 1930.
- [3] RJ Van de Graaff. A 1,500,000 volt electrostatic generator. *Phys. Rev*, 38:1919, 1931.
- [4] R Wideröe. Über ein neues Prinzip zur Herstellung hoher Spannungen. *Archiv für Elektrotechnik*, 21(4):387–406, 1928.
- [5] EO Lawrence and MS Livingston. The production of high speed light ions without the use of high voltages. *Physical Review*, 40(1):19, 1932.
- [6] VI Veksler. A new method of acceleration of relativistic particles. *J. Phys.*, 9:153–158, 1945.
- [7] EM McMillan. The synchrotron – a proposed high energy particle accelerator. *Physical Review*, 68(5-6):143, 1945.
- [8] JH Schulze. Scotophorus pro phosphoro inventus: seu experimentum curiosum de effectu radiorum solarium. *Bibliotheca Novissima observationum et recensionum. Ed. J. Chr. Franck, Sectio V, Nr. VII. Halae Magdeburgicae*, pages 234–240, 1719.
https://opus.bibliothek.uni-augsburg.de/opus4/frontdoor/deliver/index/docId/3001/file/Schulze_Scotophorus.pdf Retrieved 9 Apr. 2018.

-
- [9] VF Hess. Über den Ursprung der durchdringenden Strahlung. *Z. Phys.*, 14:610, 1913.
- [10] Nobelprize.org. The Nobel Prize in Physics 1936, Aug. 2017. http://www.nobelprize.org/nobel_prizes/physics/laureates/1936.
- [11] WC Röntgen. On a new kind of rays. *Science*, 3(59):227–231, 1896.
- [12] H Becquerel. On the rays emitted by phosphorescence. *Compt. Rend. Hebd. Seances Acad. Sci.*, 122:420–421, 1896.
- [13] CMG Lattes, H Muirhead, GPS Occhialini, and CF Powell. Processes involving charged mesons. *Nature*, 159(4047):694, 1947.
- [14] CTR Wilson et al. On a method of making visible the paths of ionising particles through a gas. *Proc. R. Soc. Lond. A*, 85(578):285–288, 1911.
- [15] CD Anderson. The positive electron. *Physical Review*, 43(6):491, 1933.
- [16] SH Neddermeyer and CD Anderson. Note on the nature of cosmic-ray particles. *Physical Review*, 51(10):884, 1937.
- [17] GD Rochester and CC Butler. Evidence for the existence of new unstable elementary particles. *Nature*, 160(4077):855, 1947.
- [18] DA Glaser. Some Effects of Ionizing Radiation on the Formation of Bubbles in Liquids. *Phys. Rev.*, 87:665–665, Aug 1952.
- [19] FJ Hasert et al. Observation of neutrino-like interactions without muon or electron in the gargamelle neutrino experiment. *Physics Letters B*, 46(1):138 – 140, 1973.
- [20] G Arnison, A Astbury, B Aubert, C Bacci, G Bauer, A Bezaguet, R Böck, TJV Bowcock, M Calvetti, T Carroll, et al. Experimental observation of isolated large transverse energy electrons with associated missing energy at $\sqrt{s}=540$ GeV. *Physics Letters B*, 122(1):103–116, 1983.
- [21] M Banner, R Battiston, P Bloch, F Bonaudi, K Borer, M Borghini, J-C Chollet, AG Clark, C Conta, P Darriulat, et al. Observation of single isolated electrons of high transverse momentum in events with missing

-
- transverse energy at the CERN $\bar{p}p$ collider. *Physics Letters B*, 122(5-6):476–485, 1983.
- [22] G Arnison et al. Experimental Observation of Lepton Pairs of Invariant Mass Around 95 GeV/c² at the CERN SPS Collider. *Phys. Lett.*, B126:398–410, 1983.
- [23] P Bagnaia et al. Evidence for $Z^0 \rightarrow e^+ e^-$ at the CERN $\bar{p}p$ Collider. *Phys. Lett.*, B129:130–140, 1983.
- [24] G Charpak, R Bouclier, T Bressani, J Favier, and C Zupancic. The Use of Multiwire Proportional Counters to Select and Localize Charged Particles. *Nucl. Instrum. Meth.*, 62:262–268, 1968.
- [25] VI Davkov, I Gregor, D Haas, SV Mouraviev, VV Myalkovskiy, L Naumann, VD Peshekhonov, C Rembser, IA Rufanov, NA Russakovich, P Senger, SY Smirnov, and VO Tikhomirov. Spatial resolution of thin-walled high-pressure drift tubes. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 634(1):5 – 7, 2011.
- [26] I Rubinskiy. An EUDET/AIDA Pixel Beam Telescope for Detector Development. *Physics Procedia*, 37(Supplement C):923 – 931, 2012. Proceedings of the 2nd International Conference on Technology and Instrumentation in Particle Physics (TIPP 2011).
- [27] symmetry magazine. Weak neutral currents | symmetry magazine. <https://www.symmetrymagazine.org/article/august-2009/weak-neutral-current>.
- [28] LHC machine outreach, Oct. 2017. <http://lhc-machine-outreach.web.cern.ch>.
- [29] Taking a closer look at LHC, April 2018. <https://www.lhc-closer.es>.
- [30] LRF Castillo. The large hadron collider. In *The Search and Discovery of the Higgs Boson*, 2053-2571, pages 3–1 to 3–9. Morgan & Claypool Publishers, 2015.

-
- [31] M Bajko et al. Report of the Task Force on the Incident of 19th September 2008 at the LHC. Technical Report LHC-PROJECT-Report-1168. CERN-LHC-PROJECT-Report-1168, CERN, Geneva, Mar 2009.
- [32] CERN document server, Oct. 2017. <http://cds.cern.ch>.
- [33] The HL-LHC project, Oct. 2017. <https://hilumilhcds.web.cern.ch/about/hl-lhc-project>.
- [34] The ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *Journal of Instrumentation*, 3(08):S08003, 2008.
- [35] ATLAS experiment at CERN, Oct. 2017. <http://atlas.cern>.
- [36] K Potamianos. The upgraded Pixel detector and the commissioning of the Inner Detector tracking of the ATLAS experiment for Run-2 at the Large Hadron Collider. *PoS*, EPS-HEP2015:261, 2015.
- [37] D Ferrere, ATLAS Collaboration, et al. The upgrade of the ATLAS Inner Detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 718:30–38, 2013.
- [38] M Garcia-Sciveres, D Arutinov, M Barbero, R Beccherle, S Dube, D Elledge, J Fleury, D Fougeron, F Gensolen, D Gnani, et al. The FE-I4 pixel readout integrated circuit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 636(1):S155–S159, 2011.
- [39] ATLAS Run Query - Run 357355. <http://sroe.home.cern.ch/sroe/cgi-bin/avgmodule.py?run=357355>.
- [40] H Wilkens and the ATLAS LArg Collaboration. The atlas liquid argon calorimeter: An overview. *Journal of Physics: Conference Series*, 160(1):012043, 2009.
- [41] G Aad et al. Performance of the ATLAS muon trigger in pp collisions at $\sqrt{s} = 8$ TeV. *Eur. Phys. J.*, C75:120, 2015.

-
- [42] E Diehl. Calibration and performance of the precision chambers of the atlas muon spectrometer. *Physics Procedia*, 37:543 – 548, 2012. Proceedings of the 2nd International Conference on Technology and Instrumentation in Particle Physics (TIPP 2011).
- [43] ATLAS Collaboration. Technical Design Report for the ATLAS ITk Pixel Detector. Technical Report ATL-COM-ITK-2018-019, CERN, Geneva, Mar 2018.
- [44] The ATLAS Collaboration. Letter of Intent for the Phase-II Upgrade of the ATLAS Experiment. Technical Report CERN-LHCC-2012-022. LHCC-I-023, CERN, Geneva, Dec 2012. Draft version for comments.
- [45] ams AG. <https://www.ams.com>.
- [46] LFoundry. <https://www.lfoundry.com>.
- [47] TowerJazz. <https://www.towerjazz.com>.
- [48] Excellence Cluster Universe. <http://www.universe-cluster.de>.
- [49] ATLAS Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Physics Letters B*, 716(1):1–29, 2012.
- [50] CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Physics Letters B*, 716(1):30–61, 2012.
- [51] ATLAS Experiment. Standard Model Public Results. <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/StandardModelPublicResults>.
- [52] ATLAS Experiment. Public Results. <https://twiki.cern.ch/twiki/bin/view/AtlasPublic>.
- [53] M Mangano. The Physics Landscape of the High Luminosity LHC. *Adv. Ser. Direct. High Energy Phys.*, 24:19–30, 2015.

-
- [54] B Schmidt. The High-Luminosity upgrade of the LHC: Physics and Technology Challenges for the Accelerator and the Experiments. In *Journal of Physics: Conference Series*, volume 706, page 022002. IOP Publishing, 2016.
 - [55] C Patrignani et al. Review of Particle Physics. *Chin. Phys.*, C40(10):100001, 2016.
 - [56] Particle Data Group. Atomic and Nuclear Properties of Materials, 2017. <http://pdg.lbl.gov/2017/AtomicNuclearProperties/index.html>.
 - [57] J Beringer et al. 2013 Review of Particle Physics. *Phys. Rev.*, D86(010001), 2012. and 2013 partial update for the 2014 edition.
 - [58] H Bichsel. Straggling in thin silicon detectors. *Rev. Mod. Phys.*, 60:663–699, Jul 1988.
 - [59] SR Messenger, EA Burke, GP Summers, MA Xapsos, RJ Walters, EM Jackson, and BD Weaver. Nonionizing energy loss (NIEL) for heavy ions. *IEEE Transactions on Nuclear Science*, 46(6):1595–1602, 1999.
 - [60] J Lindhard, V Nielsen, M Scharff, and PV Thomsen. Integral equations governing radiation effects. (notes on atomic collisions, iii). *Kgl. Danske Videnskab., Selskab. Mat. Fys. Medd.*, Vol: 33: No. 10, Jan 1963.
 - [61] GP Summers, EA Burke, P Shapiro, SR Messenger, and RJ Walters. Damage correlations in semiconductors exposed to gamma, electron and proton radiations. *IEEE Transactions on Nuclear Science*, 40(6):1372–1379, 1993.
 - [62] EA Burke. Energy dependence of proton-induced displacement damage in silicon. *IEEE Transactions on Nuclear Science*, 33(6):1276–1281, Dec 1986.
 - [63] A Akkerman, J Barak, MB Chadwick, J Levinson, M Murat, and Y Lifshitz. Updated NIEL calculations for estimating the damage induced by particles and γ -rays in Si and GaAs. *Radiation Physics and Chemistry*, 62(4):301–310, 2001.

-
- [64] M Moll. *Radiation damage in silicon particle detectors: Microscopic defects and macroscopic properties*. PhD thesis, Hamburg U., 1999.
 - [65] R Gross and A Marx. *Festkörperphysik*. Oldenbourg Wissenschaftsverlag, 2012.
 - [66] H Ibach and H Lüth. *Festkörperphysik: Einführung in die Grundlagen*. Springer-Lehrbuch. Springer Berlin Heidelberg, 2009.
 - [67] C Kittel. *Introduction to Solid State Physics*. Wiley, 2004.
 - [68] G Lutz. *Semiconductor Radiation Detectors: Device Physics*. Springer Berlin Heidelberg, 2007.
 - [69] SM Sze and KK Ng. *Physics of Semiconductor Devices*. Wiley, 2006.
 - [70] A Chilingarov. Temperature dependence of the current generated in si bulk. *Journal of instrumentation*, 8(10):P10003, 2013.
 - [71] Creative Commons. Diode IV Curves. <https://commons.wikimedia.org/wiki/File:Diode-IV-Curve.svg>.
 - [72] Spieler, H. *Semiconductor Detector Systems*. Oxford University Press, October 20, 2005 edition, 2005.
 - [73] EH Putley and WH Mitchell. The Electrical Conductivity and Hall Effect of Silicon. *Proceedings of the Physical Society*, 72(2):193, 1958.
 - [74] M Wenzel, G Irmer, J Monecke, and W Siegel. Determination of the effective Hall factor in p-type semiconductors. *Semiconductor Science and Technology*, 13(5):505, 1998.
 - [75] T Lari. Lorentz angle variation with electric field for ATLAS silicon detectors. Technical Report ATL-INDET-2001-004, CERN, Geneva, Apr 2001.
 - [76] Hall effect, June 2018. <http://nptel.ac.in/courses/115103030/26>.
 - [77] TSMC Kicks Off Volume Production of 7nm Chips. <https://www.anandtech.com/show/12677/tsmc-kicks-off-volume-production-of-7nm-chips>.

-
- [78] M Mitchell Waldrop. The chips are down for moore’s law. *Nature News*, 530(7589):144, 2016.
- [79] Creative Commons. Diode IV Curves, Oct. 2017. https://commons.wikimedia.org/wiki/File:IvsV_mosfet.svg.
- [80] R Baron, GA Shifrin, OJ Marsh, and James W Mayer. Electrical behavior of group iii and v implanted dopants in silicon. *Journal of Applied Physics*, 40(9):3702–3719, 1969.
- [81] S Ramo. Currents induced by electron motion. *Proc. Ire.*, 27:584–585, 1939.
- [82] W Shockley. Currents to conductors induced by a moving point charge. *Journal of Applied Physics*, 9(10):635–636, 1938.
- [83] Z He. Review of the shockley-ramo theorem and its application in semiconductor gamma-ray detectors. *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 463(1-2):250–267, 5 2001.
- [84] VA Gritsenko, RM Ivanov, and Yu N Morokov. Electronic structure of amorphous SiO₂: Experiment and numerical simulation. *JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS C/C OF ZHURNAL EKSPERIMENTAL’NOI I TEORETICHESKOI FIZIKI*, 81:1208–1212, 1995.
- [85] K Dette. Total Ionising Dose effects in the FE-I4 front-end chip of the ATLAS Pixel IBL detector. *Journal of Instrumentation*, 11(11):C11028, 2016.
- [86] M Huhtinen. Simulation of non-ionising energy loss and defect formation in silicon. *Nucl. Instrum. Meth.*, A491:194–215, 2002.
- [87] G Lindström, M Ahmed, S Albergo, P Allport, D Anderson, L Andricek, MM Angarano, V Augelli, N Bacchetta, P Bartalini, et al. Developments for radiation hard silicon detectors by defect engineering – results by the CERN RD48 (ROSE) Collaboration. *Nuclear Instruments and Methods*

-
- in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 465(1):60–69, 2001.
- [88] G Lindström. Radiation damage in silicon detectors. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 512(1):30 – 43, 2003. Proceedings of the 9th European Symposium on Semiconductor Detectors: New Developments on Radiation Detectors.
- [89] RM Fleming, CH Seager, DV Lang, PJ Cooper, E Bielejec, and JM Campbell. Effects of clustering on the properties of defects in neutron irradiated silicon. *Journal of Applied Physics*, 102(4):043711, 2007.
- [90] F Hartmann. *Evolution of Silicon Sensor Technology in Particle Physics*. Springer Tracts in Modern Physics. Springer International Publishing, 2017.
- [91] R Wunstorf, WM Bugg, J Walter, FW Garber, and D Larson. Investigations of donor and acceptor removal and long term annealing in silicon with different boron/phosphorus ratios. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 377(2-3):228–233, 1996.
- [92] G Kramberger. Initial acceptor removal in p-type silicon. 10th Trento Workshop, 2015.
- [93] I Tsveybak, W Bugg, JA Harvey, and J Walter. Fast neutron-induced changes in net impurity concentration of high-resistivity silicon. *IEEE transactions on nuclear science*, 39(6):1720–1729, 1992.
- [94] Benoit, M and Idarraga, J and Arfaoui, S. AllPix Twiki at CERN, Oct. 2017. <https://twiki.cern.ch/twiki/bin/view/Main/AllPix>.
- [95] Benoit, M and Idarraga, J and Arfaoui, S. AllPix repository at Github, Oct. 2017. <https://github.com/ALLPix>.
- [96] Geant4 Collaboration. Geant4: a toolkit for the simulation of the passage of particles through matter, Oct. 2017. <http://geant4.cern.ch>.

-
- [97] S Agostinelli, J Allison, K Amako, J Apostolakis, H Araujo, P Arce, M Asai, D Axen, S Banerjee, G 2 Barrand, et al. GEANT4 – a simulation toolkit. *Nuclear instruments and methods in physics research section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250–303, 2003.
- [98] S Grinstein. Overview of silicon pixel sensor development for the ATLAS insertable B-layer (IBL). In *Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 2011 IEEE*, pages 1986–1991. IEEE, 2011.
- [99] J Jentsch. Characterization of Planar n^+ -in-n ATLAS FE-I4 Single Chip Assemblies in Laboratory and Testbeam Measurements. Master’s thesis, TU Dortmund, 2011.
- [100] M Benoit, M Bomben, R Carney, G Giugliarelli, T Lari, L Meng, BP Nachman, L Rossini, V Wallangen, JC Beyer, A Ducourthial, A Macchiolo, J Llorente Merino, P Miyagawa, I Dawson, N Dann, and C Da Via. Modeling Radiation Damage Effects for Pixel Sensors in the ATLAS Detector. Technical Report ATL-COM-INDET-2017-011, CERN, Geneva, June 2018.
- [101] C Jacoboni, C Canali, G Ottaviani, and A Alberigi Quaranta. A review of some charge transport properties of silicon. *Solid-State Electronics*, 20(2):77 – 89, 1977.
- [102] B Nachman. Modeling Radiation Damage to Pixel Sensors in the ATLAS Detector. In *Meeting of the APS Division of Particles and Fields (DPF 2017) Batavia, Illinois, USA, July 31-August 4, 2017*, 2017.
- [103] M Swartz, V Chiochia, Y Allkofer, D Bortoletto, L Cremaldi, S Cucciarelli, A Dorokhov, C Hörmann, D Kim, M Konecki, et al. Observation, modeling, and temperature dependence of doubly peaked electric fields in irradiated silicon pixel sensors. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 565(1):212–220, 2006.

-
- [104] M Benoit, C Bertsche, M Bomben, R Carney, M Garcia-Sciveres, L Meng, BP Nachman, V Wallangen, T Lari, and L Rossini. A model for pixel digitization including radiation damage effects. Technical Report ATL-COM-INDET-2016-068, CERN, Geneva, Sep 2016.
- [105] M Benoit, M Bomben, R Carney, G Giugliarelli, T Lari, L Meng, BP Nachman, L Rossini, V Wallangen, JC Beyer, A Ducourthial, A Macchiolo, J Llorente Merino, P Miyagawa, I Dawson, N Dann, and C Da Via. Modeling Radiation Damage Effects for Pixel Sensors in the ATLAS Detector. Technical Report ATL-COM-INDET-2017-011, CERN, Geneva, Mar 2017.
- [106] B Abbott, J Albert, F Alberti, M Alex, G Alimonti, S Alkire, P Allport, S Altenheiner, L Ancu, E Anderssen, et al. Production and integration of the atlas insertable b-layer. *arXiv preprint arXiv:1803.00844*, 2018.
- [107] RCE Development Lab. <https://twiki.cern.ch/twiki/bin/view/Atlas/RCEDevelopmentLab>.
- [108] SlacTestBeamDec2014. <https://twiki.cern.ch/twiki/bin/view/Sandbox/SlacTestBeamDec2014>.
- [109] S Altenheiner, C Goessling, J Jentzsch, R Klingenberg, T Lapsien, D Muenstermann, A Rummler, G Troska, and T Wittig. Planar slim-edge pixel sensors for the ATLAS upgrades. *Journal of Instrumentation*, 7(02):C02051, 2012.
- [110] B Nachman. private communications.
- [111] C Goessling, R Klingenberg, D Muenstermann, A Rummler, G Troska, and T Wittig. Planar n^+ -in-n silicon pixel sensors for the ATLAS IBL upgrade. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 650(1):198–201, 2011.
- [112] RadDamageTBSpring2016. <https://twiki.cern.ch/twiki/bin/view/Atlas/RadDamageTBSpring2016>.

-
- [113] B Nachman et al. FEI4 Digitizer with Radiation Effects. Talk at ITk Charge Collection Workshop, Oct 2015. <https://indico.cern.ch/event/441959/contributions/1100542/>.
- [114] RadDamage-dev at AllPix repository, Jan. 2018. <https://github.com/ALLPix/allpix/tree/RadDamage-dev>.
- [115] TowerJazz Was Selected as Silicon Provider for the Upgrade of the Inner Tracking System of the ALICE Experiment at CERN. <http://ir.towerjazz.com/phoenix.zhtml?c=79678&p=irol-newsArticle&ID=1878512>.
- [116] J Baudot et al. First test results of mimosa-26, a fast cmos sensor with integrated zero suppression and digitized output. In *2009 IEEE Nuclear Science Symposium Conference Record (NSS/MIC)*, pages 1169–1173, Oct 2009.
- [117] I Perić. Active pixel sensors in high-voltage CMOS technologies for ATLAS. *Journal of Instrumentation*, 7(08):C08002, 2012.
- [118] B Ristić. Measurements on HV-CMOS active sensors after irradiation to HL-LHC fluences. *Journal of Instrumentation*, 10(04):C04007, 2015.
- [119] N Berger, H Augustin, S Bachmann, M Kiehn, I Perić, AK Perrevoort, R Philipp, A Schöning, K Stumpf, D Wiedner, B Windelband, and M Zimmermann. A tracker for the mu3e experiment based on high-voltage monolithic active pixel sensors. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 732:61 – 65, 2013. Vienna Conference on Instrumentation 2013.
- [120] F Bompard, S Feigl, M Garcia-Sciveres, L Meng, D Muenstermann, P Pangaud, I Perić, and A Rozanov. Radiation-hard active pixel detectors for tracking of charged particles based on HV-CMOS technology. In *Proc. Int. Image Sensors Workshop*, pages 273–276, 2013.
- [121] I Perić et al. High-voltage pixel detectors in commercial CMOS technologies for ATLAS, CLIC and Mu3e experiments. *Nuclear Instruments and*

Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 731(Supplement C):131 – 136, 2013. PIXEL 2012.

- [122] E Vilella, R Casanova, and I Perić. H35DEMO Documentation. Technical report, IFAE, KIT, University of Liverpool, Jan 2016.
- [123] E Vilella. private communications.
- [124] E Vilella, M Benoit, R Casanova, G Casse, D Ferrere, G Iacobucci, I Perić, and J Vossebeld. Prototyping of an HV-CMOS demonstrator for the High Luminosity-LHC upgrade. *Journal of Instrumentation*, 11(01):C01012, 2016.
- [125] B Delaunay. Sur la sphère vide. *Izv. Akad. Nauk SSSR, Otdelenie Matematicheskii i Estestvennyka Nauk*, 7(793-800):1–2, 1934.
- [126] Sentaurus. TCAD Sentaurus Documentation. Technical report, synopsis, 2015.
- [127] JW Slotboom and HC De Graaff. Measurements of bandgap narrowing in si bipolar transistors. *Solid-State Electronics*, 19(10):857–862, 1976.
- [128] BYU Cleanroom. <http://cleanroom.byu.edu/ResistivityCal> Retrieved Jan. 2018.
- [129] Solecon Laboratories - Resistivity and Concentration Calculators. <http://www.solecon.com/sra/rho2ccal.htm> Retrieved Jan. 2018.
- [130] M Benoit. private communications.
- [131] M Moll. Rd50 status report 2009/2010-radiation hard semiconductor devices for very high luminosity colliders. Technical report, RD50, 2012.
- [132] M Petasecca, F Moscatelli, D Passeri, and GU Pignatell. Numerical simulation of radiation damage effects in p-type and n-type FZ silicon detectors. *IEEE transactions on nuclear science*, 53(5):2971–2976, 2006.

-
- [133] D Pennicard, G Pellegrini, C Fleta, R Bates, V O'Shea, C Parkes, and N Tartoni. Simulations of radiation-damaged 3D detectors for the Super-LHC. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 592(1-2):16–25, 2008.
- [134] A Bhardwaj, A Messineo, K Lalwani, K Ranjan, M Printz, R Ranjeet, R Eber, T Eichhorn, and THT Peltola. Development of Radiation Damage Models for Irradiated Silicon Sensors Using TCAD Tools. Technical Report CMS-CR-2014-120, CERN, Geneva, Jun 2014.
- [135] D Passeri, F Moscatelli, A Morozzi, and G Bilei. Modeling of radiation damage effects in silicon detectors at high fluences HL-LHC with Sentaurus TCAD. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 824:443–445, 2016.
- [136] D Pennicard. *3D Detectors for Synchrotron Applications*. PhD thesis, Glasgow U., 2009.
- [137] F Moscatelli, D Passeri, A Morozzi, R Mendicino, G Dalla Betta, and GM Bilei. Combined bulk and surface radiation damage effects at very high fluences in silicon detectors: Measurements and tcad simulations. *IEEE Transactions on Nuclear Science*, 63(5):2716–2723, Oct 2016.
- [138] M Buckland. TCAD Simulation of CCPDv3 and C3PD for CLIC. Talk at AIDA CMOS simulation Workshop, CPPM Marseille, May 2016. <https://indico.cern.ch/event/497449/contributions/2152829>.
- [139] G Kramberger, V Cindro, I Mandic, M Mikuz, M Milovanovic, M Zavrtnik, and K Zagar. Investigation of irradiated silicon detectors by edge-TCT. *IEEE transactions on Nuclear Science*, 57(4):2294–2302, 2010.
- [140] Particulars, Advanced measurement systems, Ltd. <http://www.particulars.si/products.php>.
- [141] J Anders et al. Charge collection characterisation with the Transient Current Technique of the ams H35DEMO CMOS detector after proton irradiation, 2018.

-
- [142] JG Fossum and DS Lee. A physical model for the dependence of carrier lifetime on doping density in nondegenerate silicon. *Solid-State Electronics*, 25(8):741–747, 1982.
- [143] DMS Sultan, R Mendicino, M Boscardin, S Ronchin, N Zorzi, and G-F Dalla Betta. Characterization of the first double-sided 3D radiation sensors fabricated at FBK on 6-inch silicon wafers. *Journal of Instrumentation*, 10(12):C12009, 2015.
- [144] M Benoit, S Braccini, R Casanova, E Cavallaro, H Chen, K Chen, FA Di Bello, D Ferrere, D Frizzell, T Golling, et al. Test beam measurement of ams H35 HV-CMOS capacitively coupled pixel sensor prototypes with high-resistivity substrate. *arXiv preprint arXiv:1712.08338*, 2017.
- [145] M Vicente. private communications.
- [146] W Bludau, A Onton, and W Heinke. Temperature dependence of the band gap of silicon. *Journal of Applied Physics*, 45(4):1846–1848, 1974.
- [147] KP O’Donnell and X Chen. Temperature dependence of semiconductor band gaps. *Applied physics letters*, 58(25):2924–2926, 1991.
- [148] E Cavallaro, R Casanova, F Förster, S Grinstein, J Lange, G Kramberger, I Mandić, C Puigdengoles, and S Terzo. Studies of irradiated AMS H35 CMOS detectors for the ATLAS tracker upgrade. *Journal of Instrumentation*, 12(01):C01074, 2017.
- [149] WC Dunlap Jr and RL Watters. Direct measurement of the dielectric constants of silicon and germanium. *Physical Review*, 92(6):1396, 1953.
- [150] M Heather Andrews, Alan H Marshak, and Ritu Shrivastava. The effect of position-dependent dielectric constant on the electric field and charge density in ap-n junction. *Journal of Applied Physics*, 52(11):6783–6787, 1981.
- [151] M Benoit, J Bilbao De Mendizabal, FA Di Bello, D Ferrere, T Golling, S Gonzalez-Sevilla, G Iacobucci, M Kocian, D Muenstermann, B Ristic, et al. The FE-I4 telescope for particle tracking in testbeam experiments. *Journal of Instrumentation*, 11(07):P07003, 2016.

-
- [152] FTBF. Fermilab Test Beam Facility. <http://ftbf.fnal.gov>.
- [153] Cascade Microtech, Inc. Cascade Microtech, Inc. <https://www.cascademicrotech.com>.
- [154] SET Smart Equipment Technology. High Accuracy Bonder. <http://www.set-sas.fr/en/pp644993-421389-ACC%C2%B5RA100.html>.
- [155] H Liu et al. Development of a modular test system for the silicon sensor R&D of the ATLAS Upgrade. *Journal of Instrumentation*, 12(01):P01008, 2017.
- [156] M Kiehn et al. Proteus pixel telescope reconstruction. <https://gitlab.cern.ch/unige-fei4tel/proteus>.
- [157] Kiehn, M. Introduction to Proteus. <https://indico.cern.ch/event/661107>.
- [158] G McGoldrick, M Červ, and A Gorišek. Synchronized analysis of test-beam data with the Judith software. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 765:140–145, 2014.
- [159] S Ritt. DRS4 datasheet. Technical report, Paul Scherrer Institut, 2018. <https://www.psi.ch/drs/evaluation-board>.
- [160] Hamamatsu. Photosensor Modules H11900/H11901 Series. Technical report, Hamamatsu, 2016. https://www.hamamatsu.com/resources/pdf/etd/H11900_H11901_TPM01076E.pdf Retrieved Jan. 2018.
- [161] M Bomben. TCAD Simulations of Silicon Sensors at HL-LHC Conditions. 13th Trento Workshop - MPP, München, Feb. 2018. <https://indico.cern.ch/event/666427/contributions/2881825/>.
- [162] L Rossini. Modeling Radiation Damage to Pixel Sensors in the ATLAS Detector. 13th Trento Workshop - MPP, München, Feb. 2018. <https://indico.cern.ch/event/666427/contributions/2881831/>.
- [163] FE-I4 Collaboration. The FE-I4A Integrated Circuit Guide. Technical Report ATL-COM-INDET-2015-073, CERN, Geneva, August 2011.

-
- [164] Pixel RCE. <https://gitlab.cern.ch/rce/pixelrce> Retrieved 7 March 2018.
- [165] H Oide, T Ince, K Lantzsch, K Potamianos, L Jeanty, M Bindi, M Backhaus, J Grosse-Knetter, I Ibragimov, and D Ferrere. Timing adjustment of the IBL and Pixel Detectors for the LHC Run 2. Technical Report ATL-COM-INDET-2015-073, CERN, Geneva, Oct 2015.
- [166] M Benoit, A Lounis, and N Dinu. Simulation of charge multiplication and trap-assisted tunneling in irradiated planar pixel sensors. In *Nuclear Science Symposium Conference Record (NSS/MIC), 2010 IEEE*, pages 612–616. IEEE, 2010.
- [167] M Franks. E-TCT characterisation of irradiated backside biased H35DEMO pixel demonstrators. 13th Trento Workshop - MPP, München, Feb. 2018. <https://indico.cern.ch/event/666427/contributions/2885024/>.