UNIVERSITY OF BERGEN

Department of Physics and Technology

MASTER'S THESIS

Characterisation of a Digital Sampling Calorimeter Prototype for Proton Computed Tomography With Electron Beams

Author: Emilie Haugland Solheim

Supervisor: Professor Dieter Röhrich

Co-supervisors: Ph.D. Pierluigi Piersimoni

Ph.D. Ganesh Tambave

Ph.D. candidate Viljar Nilsen Eikeland

November 2020

Abstract

Proton CT is a novel imaging modality designed to improve dose planning and treatment monitoring in proton therapy. The precision of the beam of charged particles, like protons, requires accurate diagnostics, to avoid unnecessary irradiation of healthy tissue. CT measurements using photons can be used to determine the range and stopping power of the particles, but with an uncertainty of 2-3% [1]. It is, therefore, preferable to use proton CT, since it provides a more accurate representation of the range of the beam in the patient.

Using technology like the ALPIDE (ALICE Pixel Detector) sensor, developed at CERN for high energy physics, a proton CT system is now under development. The ALPIDE is a Monolithic CMOS Active Pixel Sensor with a pixel matrix of 1024 × 512 sensitive pixels, where each pixel measures 29.24 μm × 26.88 μm; giving better resolution and more accurate determination of particle paths. The ALPIDE can also retain a high data rate, resulting in a short scan time.

In this thesis, the focus has been on the characterisation of a digital sampling calorimeter prototype. The properties were investigated based on a test beam experiment at DESY using the Electromagnetic Pixel Calorimeter, EPICAL-2, prototype. EPICAL-2 is composed of 48 ALPIDE sensors, consisting of 24 layers, with two ALPIDE sensors and 3 mm of tungsten absorber per layer. This thesis conducts a systematic study of the sensor performance and the shower profiles at various energies for single and multiple electromagnetic showers in one readout frame. The preliminary results of the data analysis show that the prototype performs as a calorimeter should, i.e. the response scales linearly with the energy of the incoming electron.

Acknowledgement

First of all, I would like to thank my supervisors, Professor Dieter Röhrich, Ph.D. Pierluigi Piersimoni, Ph.D. Ganesh Tambave and Ph.D. candidate Viljar Nilsen Eikeland for all their excellent guidance throughout this project. Thank you, Dieter, for providing me with a truly interesting project, and for your valuable insight along the way. I would also like to thank you for your great lectures in radiation physics, giving me the motivation to pursue this project. I would like to send a thank you to Pierluigi and Ganesh for their input and ideas concerning the project. Thank you, Viljar, for all of the incredibly helpful conversations, and for always having an answer when I was stuck on a problem.

I would also like to express my appreciation to the EPICAL-2 analysis team. A thank you to Naomi Van Der Kolk, Nigel Watson, Thomas Peitzmann, Fabian Pliquett, Qasim Waheed Malik, Robert Ross Bosley, Marcus Johannes Rossewij, Hiroki Yokoyama and Rene Georges Ernst Barthel for everything you have taught me during the test beam experiment and the weekly analysis meetings. I would also like to thank you for all the helpful discussions and for input regarding the analysis—an additional thank you to Qasim Waheed Malik and Hiroki Yokoyama for sharing your scripts for inspiration.

Thank you to all my friends and fellow students at the Department of Physics and Technology, for making these years some of the best years of my life. Thank you, Lena Marie Setterdahl, Minh Chi To, Ingrid Marie Stuen and Andreas Havsgård Handeland for your friendship, motivation, love and support. A special thanks to Ingrid Marie Stuen for our great teamwork and to Andreas Havsgård Handeland for all the helpful discussions and for all your comments and suggestions. Lastly, I want to thank my family, especially my parents and my sister, for all your love and for always encouraging and supporting me.

Emilie Haugland Solheim
Bergen, November 2020

Contents

List of Figures

List of Tables

List of abbreviations

ALICE A Large Ion Collider Experiment

ALPIDE ALICE Pixel Detector

C++ A programming language

CERN Conseil Européen pour la Recherche Nucléaire

CMOS Complementary Metal-Oxide-Semiconductor

CT Computed Tomography

DESY Deutsches Elektronen-Synchrotron

DNA Deoxyribonucleic Acid

DTC Digital Tracking Calorimeter

EPICAL Electromagnetic Pixel Calorimeter

FoCal Forward Calorimeter

ITS Inner Tracking System

LET Linear Energy Transfer

LHC Large Hadron Collider

MAPS Monolithic Active Pixel Sensor

MCS Multiple Coulomb Scattering

MIP Minimum Ionising Particles

MRI Magnetic Resonance Imaging

NMOS N-type Metal-Oxide-Semiconductor

pCT Proton Computed Tomography

PET Positron Emission Tomography

PMOS P-type Metal-Oxide-Semiconductor

RBE Relative Biological Effectiveness

ROOT Object-oriented data analysis framework developed by CERN

RSP Relative Stopping Power

SOBP Spread-Out Bragg Peak

TB Test Bench

WEPL Water Equivalent Path Length

Chapter 1

Introduction

Radiotherapy is a focused treatment course often used on cancer patients, most commonly delivered using external sources emitting x-rays or other high-energy photons. Radiotherapy aims to primarily deliver ionising radiation to the tumour and limit the unwanted irradiation of healthy tissue. A disadvantage of conventional radiotherapy is that the photons have an exponentially decreasing dose delivery as the radiation traverses the patient, leading to a dose delivery to the surrounding healthy tissue [1].

In the last decades, there has, however, been an increase in radiotherapy using charged particles, such as protons and heavier ions. The physical properties of charged particles traversing a medium provide methods for dose delivery at a desired depth in the patient. The accuracy in dose delivery reduces the unwanted irradiation of healthy tissue, resulting in a reduced risk of long-term side effects [1].

In Norway, the first cancer patient will be treated with proton therapy in 2024, at Radiumhospitalet in Oslo. The planned proton centre in Bergen will also open the same year [2]. Currently, measurements using x-ray CT are used in dose planning to determine the range and stopping power of the particle. This technique introduces uncertainties that can be reduced by using a proton CT scanner. In this context, the Bergen pCT collaboration has been granted funding to design and build a pCT prototype.

1.1 About the Thesis

The primary objective of this thesis is to characterise a digital sampling calorimeter prototype, which has a similar design to the one that will be used in the Bergen proton CT prototype. The aim is here to investigate the performance of the ALPIDE chip when used as an electromagnetic calorimeter. The characterisation of this prototype, referred to as the

EPICAL-2 prototype, is conducted based on experiments carried out at the DESY II Test Beam Facility. This prototype consists of 24 layers, where each layer contains two ALPIDE chips and 3 mm of tungsten absorber. The Digital Tracking Calorimeter, which will be used for tracking and energy measurements in the proton CT prototype, will, on the other hand, consist of 2 tracking layers and 41 calorimeter layers.

The general aim of this study is to understand the performance of such a prototype since it is the first time a stack of 24 layers consisting of absorbers and ALPIDE chips has been built and tested. Previously this kind of setup has only been studied digitally, using simulations. This prototype is, therefore, a great test bench to study the overall performance of the ALPIDE when packed in such a dense calorimeter.

This study is performed using electron beams since the showers have a similar hit density as can be expected for proton CT. The prototype could also have been tested using a proton beam with aluminium as an absorber instead of tungsten. This was, however, not a possibility due to the COVID-19 situation.

Time was spent collecting data during the beam test at DESY. This beam test resulted in a substantial amount of data where different parameters and setups were tested. A considerable amount of time went into analysing the collected data, and results were discussed in weekly meetings with the EPICAL-2 analysis group who conducted the experiment. The software used for analyses was written in the programming language C++ using analysis tools provided by ROOT.

1.2 Thesis Outline

Chapter 2: Radiation and Its Medical Applications This chapter describes the medical applications of radiation and provides a basic description of how charged particles and photons interact with matter. Electromagnetic showers and the biological effects are covered, as well as the concept of radiation therapy. Lastly, it will present the concept of proton CT and the Bergen pCT Collaboration.

Chapter 3: Semiconductors and Silicon Pixel Detectors This chapter presents the fundamental principles of semiconductors and silicon pixel detectors, explaining how charged particles can produce a signal. The concepts and design of silicon pixel detectors and Monolithic Active Pixel Sensors are covered, as well as the principal variables influencing the size of clustered pixel hits.

Chapter 4: ALPIDE Chip This chapter aims to describe the essentials of the ALPIDE detector. The ALPIDE operations will be described focusing on the front-end, priority en-

coders, triggering and strobes. Finally, the applications of the detector in calorimeters are presented, concentrating on FoCal and proton CT.

Chapter 5: Characterisation of the EPICAL-2 Prototype In this chapter the experiment conducted at the DESY II Test Beam Facility in February 2020 is described, as well as the result of the data analysis. Data from a similar experiment carried out in November 2019 were also used for analysis.

Chapter 6: Discussion and Conclusion This chapter reviews and discuss the results presented in the thesis.

1.3 Citation Principles

In this thesis, citations listed before the paragraph's ending punctuation refer to all the statements of that paragraph, in cases where there is only one citation per paragraph. Citations registered at the end of a sentence within a paragraph refers to the statement itself.

Chapter 2

Radiation and Its Medical Applications

This chapter outlines the medical applications of radiation and provides a basic description of how charged particles and photons interact with matter—then explaining electromagnetic showers before presenting the biological effects of radiation. Thereafter, the concept of radiation therapy will be disclosed, followed a discussion of the advantages of particle therapy compared to conventional radiotherapy using photons. Finally, proton computed tomography and the Bergen pCT Collaboration will be described.

2.1 Radiation

Radiation is a general term used to describe a process that carries energy through space in the form of electromagnetic waves or as streams of energetic particles [3]. There are several ways to characterise radiation based on how it interacts with matter. We distinguish between non-ionising and ionising radiation, where ionising radiation can be further characterised as either directly or indirectly ionising [4].

Ionisation describes the process where a neutral atom gains a positive or negative charge, and ionising radiation can cause this process by stripping electrons from atoms as the beam traverses the medium. The atom that loses its electron ends up as a positive ion, and subsequently, the stripped electron might further combine with a neutral atom, forming a negative ion [4].

Provided that they have an amount of kinetic energy that is adequate to produce ionisation by collisions as they penetrate matter, charged particles such as α-particles, protons and electrons are referred to as directly ionising radiation. The incident particle loses its energy in a large number of small steps along the ionisation track within the medium, and occasionally the ejected electron acquires enough energy to form its own, secondary ionisation track.

These electrons are known as δ rays. Alternatively, if the energy loss of the incident particle is used to raise the electrons to a higher energy level, but is not large enough to eject it from the atom, the process is called excitation [4].

Uncharged particles, such as photons and neutrons, are referred to as indirectly ionising radiation since they do not ionise themselves, but instead, release directly ionising particles from the matter that they interact with [4].

2.2 Interaction of Photons With Matter

Photons are one of the fundamental particles, which are particles without a known substructure. Photons are force carriers in a force field, and according to the theory of quantum electrodynamics (QED) classified as messenger particles. Being force carriers of the electromagnetic force, they are quanta of the electromagnetic field [4].

Electromagnetic radiation is characterised by electric and magnetic fields that are oscillating perpendicular to each other, and perpendicular to the direction of the energy propagation. Photons have no electric charge or mass and their energy is proportional to the frequency of oscillation. The amount of energy carried by a photon, E_γ, is given by

$$E_\gamma = h\nu \tag{2.1}$$

where h is the Planck's constant and ν is the frequency of oscillation [4].

The utilisation of higher energy photons like x-rays and γ-rays is common in both conventional radiotherapy and medical imaging [4]. In the interest of understanding the medical applications of photons, knowledge of the fundamental interactions between photons and matter is essential.

The interaction of photons with matter (in our case, x-rays and γ-rays) is drastically different from that of charged particles, in large part due to the photon's lack of electric charge. The many inelastic collisions with atomic electrons, characteristic of the interaction between charged particles and matter, are impossible for photons. The main interactions of x-rays and γ-rays with matter are instead [5]:

1. Photoelectric Effect
2. Compton scattering (including Thomson and Rayleigh Scattering)
3. Pair Production.

Photons also interact through other mechanisms, such as nuclear dissociation reactions, but these are significantly less common, and thus not mentioned in detail in this thesis [5].

X-rays and γ-rays are considerably more penetrating than charged particles when passing through matter. This characteristic is a result of the three processes mentioned earlier having a much smaller cross-section compared to the cross-section of the inelastic collisions between charged particles and electrons [5].

Furthermore, the energy of a photon beam does not degrade as it traverses a medium; the beam only attenuates in intensity. This feature is also a consequence of these three interactions. Photons are removed from the beam entirely, either by scattering or by absorption, meaning that the photons passing straight through will not have undergone any interactions at all, retaining their original energy [5].

When a mono-energetic photon beam travels through matter, there is a reduction in the total number of photons by the number of photons that interacts with the medium. The following expression gives the intensity of the photon beam after having travelled a distance x in the medium:

$$I(x) = I_0 \, \mathrm{e}^{-\mu x} \tag{2.2}$$

In this equation, I_0 is the beam's intensity before it enters the medium, and μ is the absorption coefficient of the medium. The absorption coefficient is characteristic of the absorbing material and is dependent on the density of the medium. It is also a quantity which is directly related to the total cross-section of the photon interactions [5].

The following approximation gives the total cross-section, σ_{tot}, of a photon interaction:

$$\sigma_{\mathrm{tot}} \approx \sigma_{\mathrm{pe}} + \sigma_{\mathrm{c}} + \sigma_{\mathrm{pp}} \tag{2.3}$$

where σ_{pe}, σ_{c} and σ_{pp} are the cross-sections of photoelectric effect, Compton scattering and pair production, respectively [4][6].

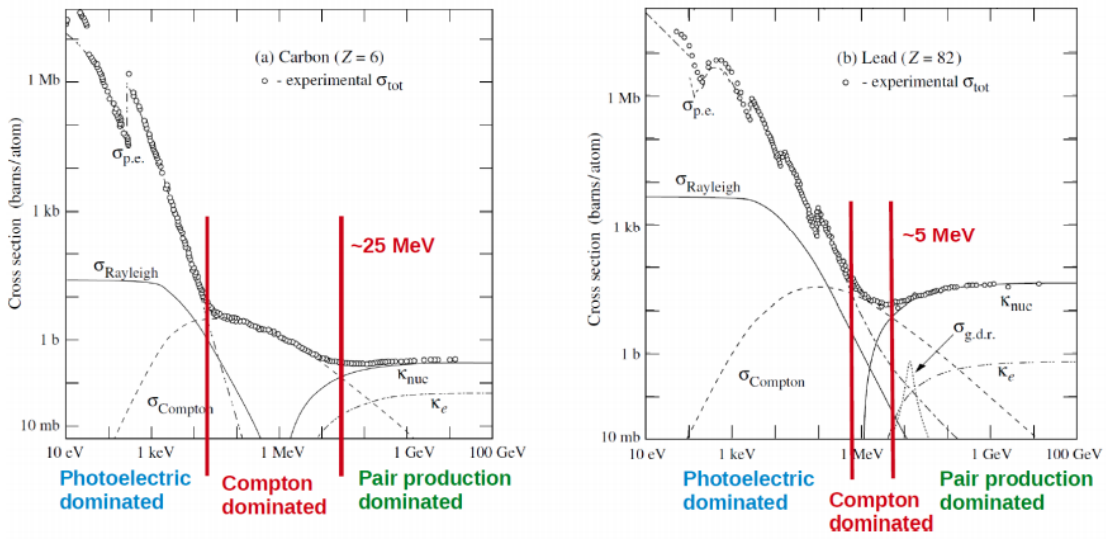


Figure 2.1: Total cross-section (circles) and partial cross-sections (lines) of photons in carbon (left) and lead (right). The partial cross-sections correspond to the following interactions: photoelectric effect ($\sigma_{\text{p.e.}}$), Compton scattering (σ_{Compton}), coherent scattering (σ_{Rayleigh}), pair production in the field of the nucleus (κ_{nuc}) and the field of the electrons (κ_{e}), and lastly the giant dipole resonance ($\sigma_{\text{g.d.r}}$), a photonuclear process [7][8]. The figure is reprinted from [7], 2519-8041 – ©CERN, 2018. Published by CERN under the Creative Common Attribution CC BY 4.0 Licence. https://doi.org/10.23730/CYRSP-2018-005.47

Figure 2.1 illustrates how the cross-section of the different interactions change with the photon energy, and which process is the dominating factor for energy loss at different energies.

The main interactions of photons, as well as their cross-sections, will be addressed in the following sections.

2.2.1 Photoelectric Effect

The photoelectric effect, illustrated in Figure 2.2, is the process of an atom absorbing a photon and subsequently emitting one of its orbital electrons [9]. In this phenomenon, the atom first absorbs the entire energy of the photon ($h\nu$) before, essentially, transferring all of the energy to the atomic electron. The photoelectron, which is the electron ejected from the atom, has kinetic energy, E_{k}, equal to

$$E_{\text{k}} = h\nu - E_{\text{B}} \tag{2.4}$$

where E_{B} is the electron's binding energy. The photoelectric effect can occur with electrons in the K, L, M or N shells [4].

When the electron gets ejected, it creates a vacancy in the shell and leaves the atom in an excited state. An outer orbital electron can fill the vacant position by emitting a characteristic x-ray. The emission of Auger electrons is also a possibility. The Auger effect occurs when an electron in a higher shell gets ejected after absorbing the energy generated by the outer electron filling the vacancy [4].

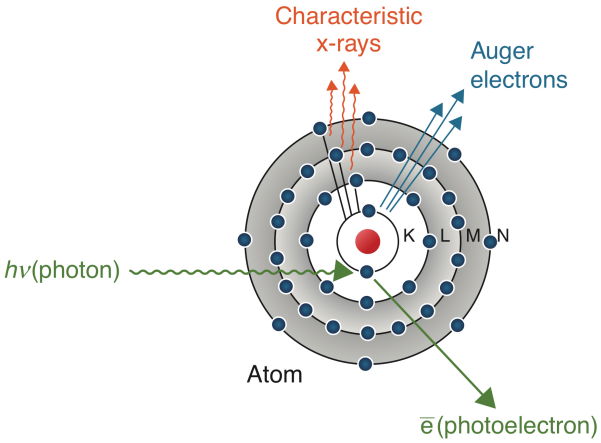


Figure 2.2: Diagram illustrating the photoelectric effect. The K-shell electron gets ejected from its orbit after absorbing the incident photon. An outer orbital electron can fill the vacancy created by the ejected electron, by emitting a characteristic x-ray. Emitting an electron (called an Auger electron), instead of a photon, can also return the excited atom to a lower energy state. The figure is taken from [4].

For photons with energy less than 100 keV, the photoelectric effect is the dominant mode of interaction. The following expression gives an approximation of the energy dependence of the cross-section:

$$\sigma_{\text{pe}} \approx \text{Const} \frac{Z^n}{E_\gamma^{3.5}} \tag{2.5}$$

Here, E_γ is the energy of the photon and Z is the charge of the nucleus. n varies between 4 and 5, depending on the energy of the photon [8].

As illustrated in Figure 2.1, the cross-section for the photoelectric effect decreases sharply with increasing photon energy. However, when the photon energy exceeds the threshold energy corresponding to the binding energy of a lower electron orbital, there is a sudden increase in the cross-section. These leaps are also visible in Figure 2.1 [8].

2.2.2 Compton Scattering

Compton scattering is the phenomenon where a photon scatters off an atomic electron, as shown in Figure 2.3. In this process, the electron's binding energy is negligible compared to the energy of the incident photon. The photon, therefore, interacts with the electron as if

it was a "free" electron. The electron gets emitted at an angle θ to the path of the incident photon, after receiving some of the photon's energy. The photon scatters at an angle ϕ from its original path, with reduced energy [4].

One can analyse Compton scattering by treating it as a collision between two particles. Further, one can obtain the following expressions by using the conservation laws for momentum and energy:

$$E'_\gamma = \frac{E_\gamma}{1+\varepsilon(1-\cos\phi)} = \frac{h\nu_0}{1+\varepsilon(1-\cos\phi)} \tag{2.6}$$

$$E_\mathrm{k} = E_\gamma - E'_\gamma = h\nu_0 \frac{\varepsilon(1-\cos\phi)}{1+\varepsilon(1-\cos\phi)} \tag{2.7}$$

In these equations E_γ and E'_γ are the energies of the incident photon and the scattered photon, respectively. E_k is the kinetic energy of the electron, h is Planck's constant and ν_0 is the frequency of the incident photon. Furthermore, $\varepsilon = h\nu_0/m_\mathrm{e}c^2$, where $m_\mathrm{e}c^2 = 0.511$ MeV is the rest energy of the electron [4][5].

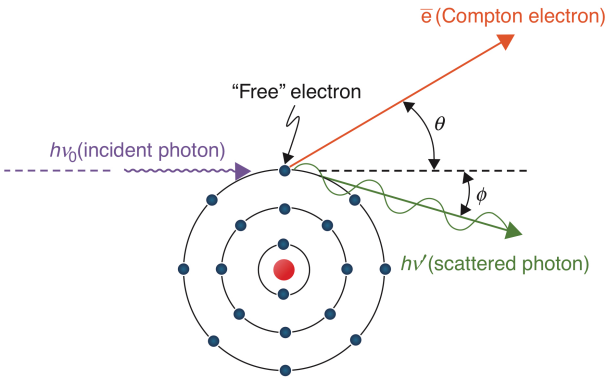


Figure 2.3: Illustration showing the process of Compton scattering, where a photon scatters off an atomic electron. The electron gets emitted at an angle θ to the path of the incident photon, and the photon scatters at an angle ϕ from its original path. [4].

The scattered photon might undergo several interactions after being scattered off the atomic electron, before finally being absorbed by the medium. The Compton electron may, alternatively, cause further excitation and ionisation events as it traverses the medium.

A special case of Compton scattering is when the photon makes a direct hit with the electron. After the interaction, the electron will move along the path of the incident photon with an angle $\theta = 0°$, and the scattered photon will move in the opposite direction, $\phi = 180°$. This event leaves the scattered photon with a minimum energy, $E'_{\gamma,\mathrm{min}}$, and transfers a maximum amount of energy, $E_{\mathrm{k,max}}$, to the electron. One can determine these energies by solving Equa-

tions 2.6 and 2.7 for $\phi = 180°$ in [4]:

$$E'_{\gamma,\min} = h\nu_0 \frac{1}{1+2\varepsilon} \tag{2.8}$$

$$E_{\mathrm{k,max}} = h\nu_0 \frac{2\varepsilon}{1+2\varepsilon} \tag{2.9}$$

Equation 2.9 gives the maximum recoil energy a Compton electron can get according to kinematics, also known as the *Compton edge*. The plot in Figure 2.4 displays the energy distribution for various energies of incoming photons [5].

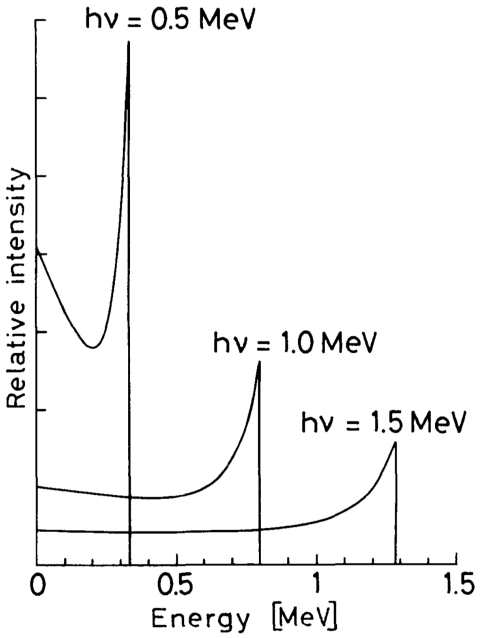


Figure 2.4: Plot showing the energy distribution of Compton recoil electrons. At maximum recoil energy, there is a sharp decline known as the *Compton edge*. The figure is retrieved from [5].

Another special case is when the photon only grazes the electron, resulting in the electron being emitted at $\theta = 90°$ to the incident photon path, while the scattered photon continues forward with $\phi = 0°$. From Equations 2.6 and 2.7, with $\phi = 0°$, the energies become $E'_{\gamma,\min} = h\nu_0$ and $E_{\mathrm{k,max}} = 0$ [4].

The cross-section for Compton scattering is known as the Klein-Nishina formula:

$$\frac{d\sigma}{d\Omega} = \frac{r_\mathrm{e}^2}{2} \frac{1}{[1+\varepsilon(1-\cos\phi)]^2} \left(1+\cos^2\phi + \frac{\varepsilon^2(1-\cos\phi)^2}{1+\varepsilon(1-\cos\phi)}\right) \tag{2.10}$$

where $d\sigma/d\Omega$ is a differential cross-section, $d\Omega$ an infinitesimal solid angle element and r_e

the classical radius of the electron. The total probability for a Compton scattering to occur, per electron, is given by integrating Equation 2.10 over $d\Omega$ [5].

$$\sigma_c = 2\pi r_e^2 \left\{ \frac{1+\varepsilon}{\varepsilon^2} \left[\frac{2(1+\varepsilon)}{1+2\varepsilon} - \frac{1}{\varepsilon}\ln(1+2\varepsilon) \right] + \frac{1}{2\varepsilon}\ln(1+2\varepsilon) - \frac{1+3\varepsilon}{(1+2\varepsilon)^2} \right\} \tag{2.11}$$

Thomson and Rayleigh Scattering

The classical phenomena that are known as Thomson and Rayleigh scattering closely are related to the process of Compton scattering. Thomson scattering is the process of free electrons scattering photons within the classical limit, while Rayleigh scattering describes when an atom as a whole scatters the photons. The latter process is also known as coherent scattering, stemming from the fact that all the electrons of the atom participate coherently [5].

Common for both phenomena is the fact that there is no net energy transfer to the medium. By only altering the direction of the photon, these processes neither ionise nor excite the atoms. For most purposes, Thomson and Rayleigh can be neglected for photons with relatively high energies, as there is minimal scattering occurring at these energies [5].

2.2.3 Pair Production

Pair production describes the process of a photon producing one electron and one positron within the field of an atomic nucleus, as shown in Figure 2.5. This process can only occur if the photon has an energy higher than 1.02 MeV. Interacting with the electromagnetic field of a nucleus, the photon loses all its energy and establishes a pair consisting of an electron, e^-, and its antiparticle, e^+, also known as a positron. A minimum photon energy of 1.02 MeV is required for this process, because both the electron and the positron has a rest energy of 0.511 MeV. The particles share any excess energy as kinetic energy, and they tend to get ejected in a path forward relative to the incoming photon [4].

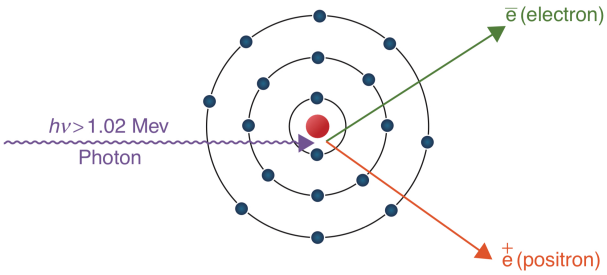


Figure 2.5: Diagram displaying the process of pair production. The photon produces an electron-positron pair within the field of an atomic nucleus. The figure is taken from [4].

If atomic electrons do not screen the nuclear charge, (photons with low energy must be rel-

atively proximate to the nucleus for pair production to be probable; thus the photon views only a 'bare' nucleus), ε will be within the following limits

$$1 \ll \varepsilon < \frac{1}{\alpha Z^{1/3}} \tag{2.12}$$

the following expression then gives the cross-section of pair production

$$\sigma_{\mathrm{pp}} = 4\alpha r_{\mathrm{e}}^2 Z^2 \left(\frac{7}{9} \ln(2\varepsilon) - \frac{109}{54} \right) \tag{2.13}$$

If the nuclear charge, on the other hand, is completely screened,

$$\varepsilon \gg \frac{1}{\alpha Z^{1/3}} \tag{2.14}$$

which makes the cross-section

$$\sigma_{\mathrm{pp}} = 4\alpha r_{\mathrm{e}}^2 Z^2 \left(\frac{7}{9} \ln\left(\frac{183}{Z^{1/3}} \right) - \frac{1}{54} \right) \tag{2.15}$$

The variables in these equations are described earlier in this section.

In these equations, α is the fine-structure constant given by

$$\alpha = \frac{e^2}{4\pi\varepsilon_0 \hbar c} \approx \frac{1}{137} \tag{2.16}$$

where e is the elementary charge, ε_0 is the permittivity of free space, $\hbar = h/2\pi$ is the reduced Planck constant, and c the speed of light in vacuum [5][10].

2.3 Interaction of Charged Particles With Matter

As mentioned in the previous section, photons interact with the medium through processes such as photoelectric effect, Compton scattering and pair production. Charged particles, on the other hand, mainly interact through ionisation and excitation. Charged particles may also interact by radiative collisions, which are collisions between charged particles where a portion of their kinetic energy is radiated away as one or more photons. Bremsstrahlung is an example of a radiative collision, and it is a process more likely for electrons and positrons, than for more massive charged particles [4].

The electric fields of the travelling charged particles interact with the fields of the nuclei and orbital electrons along the charged particles' path. Atoms in the medium are ionised and excited through collisions between the particle and the orbital electrons, while collisions

between the particle and the nucleus give rise to bremsstrahlung and other radiative energy loss. There is also a possibility of the particle scattering without notable energy loss. In addition to the interactions governed by the Coulomb force, charged heavy particles can also produce radioactive nuclides by causing nuclear reactions [4].

Hence, the following interactions are the leading cause of the energy loss and deflection of charged particles traversing a medium:

1. Inelastic collisions with atomic electrons; leading to ionisation and excitation of atoms and is the primary cause of the charged particle's energy loss
2. Elastic collisions with nuclei, also known as Coulomb scattering or angular deflection
3. Elastic collisions with atomic electrons
4. Inelastic collisions with nuclei (nuclear reactions); e.g. absorption of the charged particle, leading to the production of secondary particles
5. Bremsstrahlung (for high energy particles); caused by the deceleration of charged particles
6. Cherenkov radiation; arising from charged particles moving faster than the speed of light in a particular medium
7. Transition radiation; a weak effect that resembles the Cherenkov radiation and is caused by the charged particle polarising the medium

These interactions may occur numerous times per unit path length within the medium which cumulatively causes the observed effects. The three last reactions are, however, of minimal consequence compared to the processes of atomic collision [5][8].

2.3.1 Energy Loss by Ionisation for Heavy Particles

When a massive charged particle passes through a medium, the inelastic collisions caused by the Coulomb fields are almost entirely responsible for its energy loss. The energy transfer per collision is usually minimal compared to the total kinetic energy of the particle. Nevertheless, there is also a large number of collisions per unit path length within the typically dense matter, resulting in substantial cumulative energy loss [5].

The average energy loss per unit length, also known as stopping power, is described by the Bethe-Bloch formula:

$$-\frac{dE}{dx} = 2\pi N_{\mathrm{A}} r_{\mathrm{e}}^2 m_{\mathrm{e}} c^2 \rho \frac{Z}{A} \frac{z^2}{\beta^2} \left[\ln\left(\frac{2 m_{\mathrm{e}} c^2 \gamma^2 \beta^2 W_{\max}}{I^2} \right) - 2\beta^2 - \delta - 2\frac{C}{Z} \right] \tag{2.17}$$

with the following parameters

$$2\pi N_{\mathrm{A}} r_{\mathrm{e}}^2 m_{\mathrm{e}} c^2 = 0.1535 \text{ MeVcm}^2/\text{g}$$

N_{A}	Avogadro's number = $6.022 \cdot 10^{23}$ mol^{-1}
r_{e}	classical radius of the electron = $2.817 \cdot 10^{-13}$ cm
$m_{\mathrm{e}}c^2$	rest energy of the electron = 0.511 MeV
ρ	density of the absorbing material
Z	atomic number of the absorbing material
A	atomic weight of the absorbing material
z	charge of the incident particle, in units of the elementary charge e
β	velocity of the incident particle in units of the speed of light, v/c
γ	Lorentz factor, $1/\sqrt{1-\beta^2}$
W_{max}	maximum transfer of energy in a single collision
I	mean excitation potential
δ	density correction
C	shell correction

A head-on collision produces the maximum energy transfer possible in a single collision, and kinematics gives the following expression for the maximum energy transfer of an incoming particle with mass M:

$$W_{\mathrm{max}} = \frac{2m_{\mathrm{e}}c^2\beta^2\gamma^2}{1 + 2\frac{m_{\mathrm{e}}}{M}\sqrt{1+\beta^2\gamma^2} + \left(\frac{m_{\mathrm{e}}}{M}\right)^2} \tag{2.18}$$

Further, if $M \gg m_{\mathrm{e}}$, then

$$W_{\mathrm{max}} \simeq 2m_{\mathrm{e}}c^2\beta^2\gamma^2 \tag{2.19}$$

The mean excitation potential, I, is essentially the average orbital frequency of the bound electrons, $\bar{\nu}$, times Planck's constant, $h\bar{\nu}$ [5].

The correction for density effects, δ, is important at high energies. When the energy, and thereby the velocity, of the incoming particle, is high, its electric field will flatten out and elongate due to relativistic length contractions. This electric field tends to polarise atoms along its path. The polarisation will shield electrons far from the incoming particles from the full effect of the particles' electric field. Collisions with these shielded electrons will contribute less to the total energy loss than what the uncorrected Bethe-Bloch formula would have predicted. The density effect increases for materials with higher densities due to greater polarisation [5].

The correction for shell effect, C, is, however, more important at lower energies. This effect occurs when the velocity of the incoming particle is close to the orbital velocity of bound electrons in the absorbing material. The assumption that the electrons are stationary com-

pared to the velocity of the incoming particles is no longer valid, causing the Bethe-Bloch formula to break down without this correction [5].

There is also a relativistic correction, $\ln(1-v^2/c^2)-v^2/c^2$, since the particles will interact less at higher velocities.

It is also worth noting that the average energy loss is proportional to z^2/β^2, which means that the energy loss of the particle is mostly dependent on the charge carried by the particle, and its velocity.

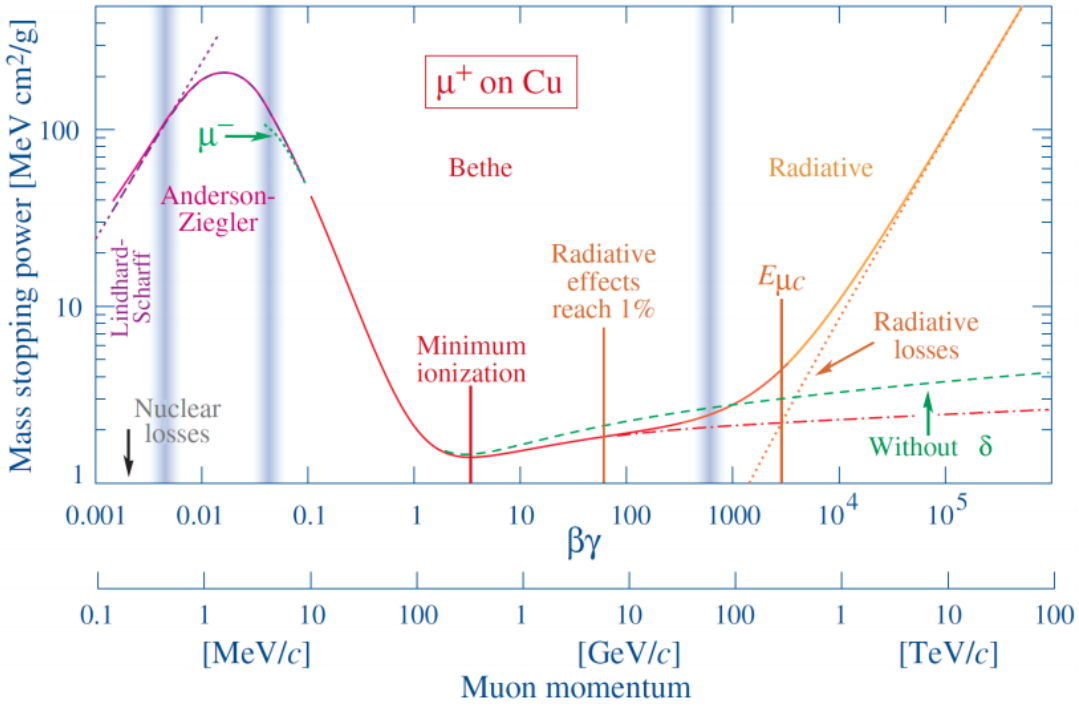


Figure 2.6: The mass stopping power plotted against the momentum for positively charged muons in copper. The solid line gives the total stopping power; the dotted line is the radiative stopping power; whereas the dash-dotted line is the electronic stopping power, according to Bethe. The figure is reprinted from [7], 2519-8041 – ©CERN, 2018. Published by CERN under the Creative Common Attribution CC BY 4.0 Licence. https://doi.org/10.23730/CYRSP-2018-005.47

By studying the plot of the Bethe-Bloch formula, in Figure 2.6, one can see that the energy loss per unit length is most significant at lower velocities. The energy loss decreases with increasing velocity until it reaches a point of minimum ionisation. A particle with average energy loss close to this minimum is called a minimum ionising particle, or MIP. After that, there is a relativistic rise, where the energy loss increases due to the relativistic elongation of the transverse electric field, caused by Lorentz-contractions of the Coulomb forces. The graph finally flattens out in the density effect plateau caused by polarisation and shielding.

2.3.2 Energy Loss of Electrons and Positrons

Electrons and positrons mainly lose their energy through inelastic collisions and bremsstrahlung. While the primary mechanism outlined for collision loss for massive particles still is valid for electrons and positrons, the Bethe-Bloch formula for collision energy needs modification. Since the projectile and target now have identical mass, the assumption that the projectile only transfers a minimal amount of energy per interaction is not valid anymore. The projectile and target are now identical seen from a quantum-mechanically perspective, so there is no way to distinguish the target from the projectile. Consequently, the calculations must consider their indistinguishability [5].

Taking these considerations into account, several terms must be adjusted in the formula. For instance, the maximum transfer of energy permitted becomes $W_{\max} = E_{\mathrm{k,e}}/2$, where $E_{\mathrm{k,e}}$ is the kinetic energy of the incoming electron or positron. Thus, the Bethe-Bloch formula becomes

$$-\frac{dE}{dx} = 2\pi N_{\mathrm{A}} r_{\mathrm{e}}^2 m_{\mathrm{e}} c^2 \rho \frac{Z}{A}\frac{1}{\beta^2}\left[\ln\left(\frac{\tau^2(\tau+2)}{2\left(\frac{I}{m_{\mathrm{e}}c^2}\right)^2}\right) - F(\tau) - \delta - 2\frac{C}{Z}\right] \tag{2.20}$$

In this equation, τ is the projectile particle's kinetic energy in units of $m_{\mathrm{e}}c^2$, and $F(\tau)$ is

$$\begin{aligned} F(\tau) &= 1 - \beta^2 + \frac{\frac{\tau^2}{8} - (2r+1)\ln 2}{(\tau+1)^2} && \text{for } \mathrm{e}^- \\ F(\tau) &= 2\ln 2 - \frac{\beta^2}{12}\left(23 + \frac{14}{\tau+2} + \frac{10}{(\tau+2)^2} + \frac{4}{(\tau+2)^3}\right) && \text{for } \mathrm{e}^+ \end{aligned} \tag{2.21}$$

The remaining parameters are as defined in Section 2.3.1 [5].

Positrons behave precisely like electrons, except at lower energies. After coming to rest, positrons will annihilate since there are always electrons present. The process of annihilation, where antiparticles collide, produces a pair of back-to-back gamma rays, each with an energy equivalent to the rest energy of the electron. If the particles collide before coming to rest, this energy will be higher [3][8].

The small mass of electrons and positrons results in an additional mechanism for energy loss. The electric field of nuclei in the medium causes the particle trajectories to bend around the nuclei; the particles accelerate and radiate energy by releasing photons as part of the bremsstrahlung process. At relativistic velocities, the contribution from bremsstrahlung exceeds the contribution from collisions and it becomes the dominating cause of energy loss

for electrons and positrons [5].

Hence, the total energy loss for electrons and positrons is composed of the energy loss caused by collisions and the loss due to radiation:

$$\left(\frac{dE}{dx}\right)_{\text{tot}} = \left(\frac{dE}{dx}\right)_{\text{coll}} + \left(\frac{dE}{dx}\right)_{\text{rad}} \tag{2.22}$$

Figure 2.7 shows how the primary mechanism for energy loss contributes to the total energy loss for different energies.

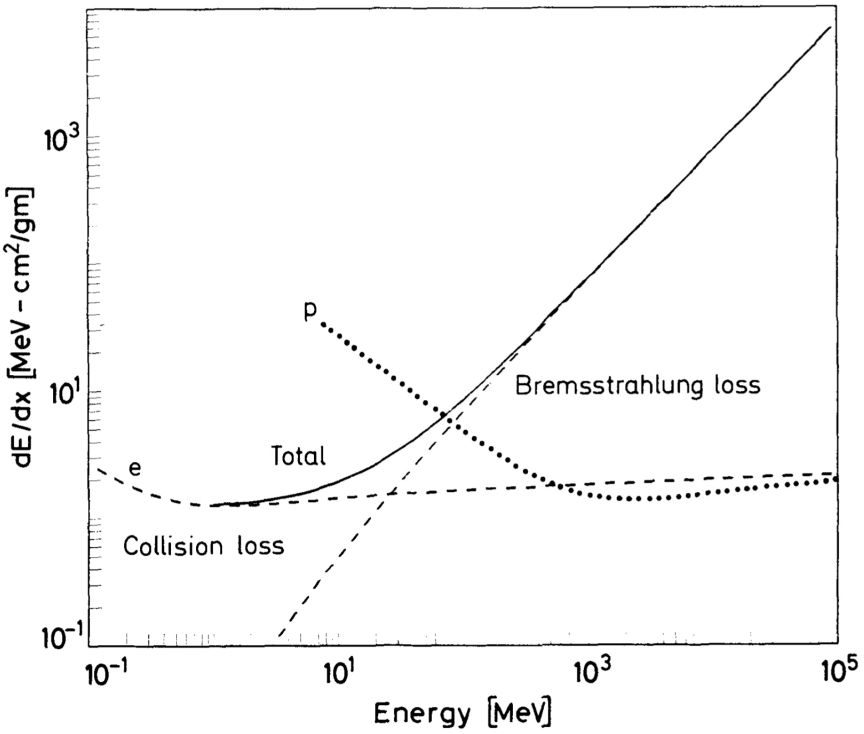


Figure 2.7: Plot illustrating how the energy loss from radiation and collision contributes to the total loss for electrons in copper. The plot also depicts the energy loss for protons for comparison. The figure is retrieved from [5].

2.3.3 Multiple Coulomb Scattering

As previously discussed, when charged particles traverse a medium, the electric field of atomic nuclei will result in small elastic collisions or deflection of the particles, commonly referred to as Coulomb scattering. While a simple scattering usually gives a negligible angular deflection, the sum of multiple scatterings can result in a large deflection from the original path of the particles [5].

The Rutherford formula governs an individual Coulomb scattering, but if the average num-

ber of independent scatterings is large enough, and the individual scattering events have a small or negligible energy loss, one can treat the problem statistically. The phenomenon of multiple repeated collisions is called multiple Coulomb scattering (MCS) [5].

A particle that enters a medium of thickness x perpendicularly and undergoing multiple scattering will come out a particular distance y from the original trajectory. The particles course will then continue with an angle θ_{MCS} from the original perpendicular path, as illustrated in Figure 2.8 [9].

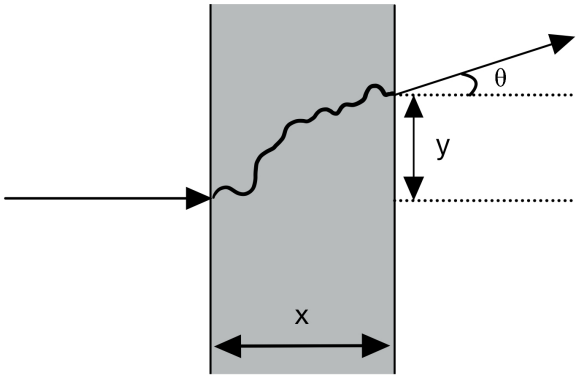


Figure 2.8: Illustration outlining a particle undergoing multiple Coulomb scattering after entering a material with thickness x. The particle emerges at position y at an angle θ with respect to its original path. The sketch is based on a figure from [9].

For small scattering angles, a normal distribution distributed around an average scattering angle of $\theta_{\text{MCS}} = 0$ describes the random variable θ_{MCS}. Larger scattering angles, resulting from elastic collisions with nuclei, occur, nevertheless, more frequently than expected from a Gaussian distribution [10].

The following formula gives the root mean square of the projected scattering-angle distribution

$$\theta_{\text{rms}}^{\text{proj.}} = \sqrt{\langle\theta^2\rangle} = \frac{13.6\ \text{MeV}}{\beta c p}\, z\sqrt{\frac{x}{X_0}}\left[1 + 0.038 \ln\left(\frac{x}{X_0}\right)\right] \tag{2.23}$$

where βc is the velocity, p [MeV/c] the momentum, z is the charge of the scattered particle and x/X_0 is the thickness of the medium per radiation length [10].

2.3.4 Range and Straggling

The range of a particle is the distance the particle will travel before losing all of its energy, a quantity that is dependent on particle type, its energy, as well as the type of medium it pen-

etrates. By measuring the ratio between transmitted and incident particles after sending a beam of particles of desired energy through various thicknesses of a specific material, one can determine the range experimentally. A *range number-distance* curve is a curve displaying this ratio as a function of absorber thickness and is exemplified in Figure 2.9 [5].

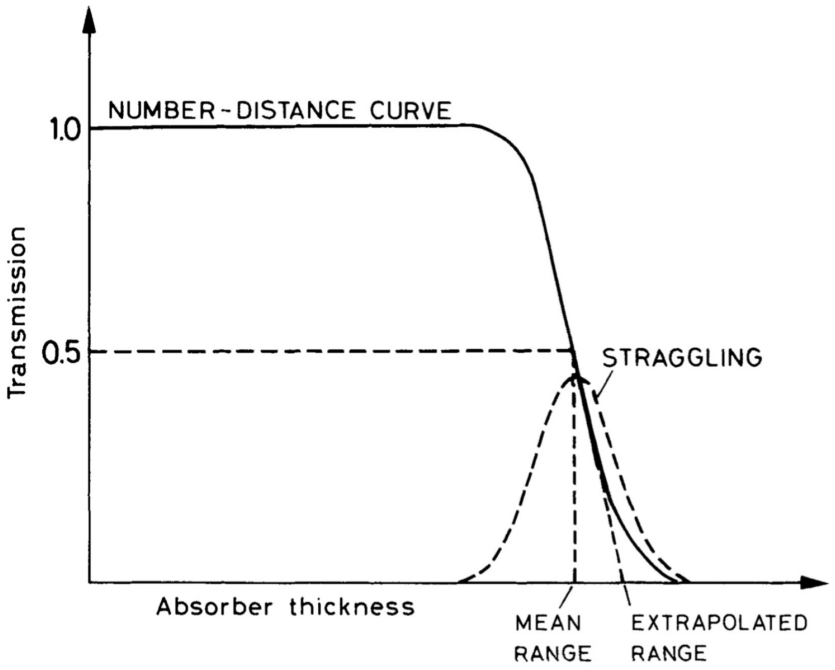


Figure 2.9: A curve showing the ratio of transmitted to incident particles as a function of absorber thickness, also known as a *range number-distance* curve. The range distribution due to range straggling is approximately Gaussian in form. The figure is retrieved from [5].

As illustrated in Figure 2.9, virtually all particles manage to penetrate small absorber thicknesses. The ratio of transmitted to incident particles drops, on the other hand, as one approaches the range. The drop is, however, not as sharp as expected for well-defined quantities, instead it decreases over a particular interval in absorber thickness. This effect is a consequence of the statistical nature of the energy loss. Two identical particles with equal initial energy will rarely undergo the same amount of collisions and will therefore have different energy loss [5].

Consequently, by measuring several identical particles, the ranges will be statistically distributed, centred around a mean value. This is a phenomenon called range straggling, and the distribution is approximately Gaussian in form. The mean value, corresponding to the midpoint of the declining slope in Figure 2.9, is known as the *mean range*. The mean range is the thickness where about half of the particles have been absorbed. Finding the point where all particles are absorbed can be done by using the curve's tangent at the midpoint to extrapolate to the zero-level. This quantity is known as the *practical* or *extrapolated range*

[5].

From a theoretical standpoint, one can determine the mean range of a particle of energy E by integrating dE/dx from initial energy E to zero, hereby assuming that ionisation is the only cause of energy loss [5][9].

$$R(E) = \int_E^0 \left(\frac{dE'}{dx}\right)^{-1} dE' \tag{2.24}$$

This calculation only gives an approximation to the travelled path length, seeing that it neglects effects such as multiple Coulomb scattering. Hence, the length of the zigzag path caused by scattering, as seen in Figure 2.8, will normally be larger than the range defined as a straight-line path. The approximation is, however, relatively good for massive charged particles, since the effect of multiple scattering is mostly negligible in these cases [5].

Electrons, on the other hand, are more susceptible to multiple Coulomb scattering, causing their range to differ quite a bit from the approximated path length obtained by Equation 2.24. There is also more fluctuation in the energy loss for electrons than for massive particles. The energy loss fluctuation is a consequence of both the emission of bremsstrahlung and the permission of much larger energy transfer per collision for electrons. Both cases make it possible for a few photons or collisions to absorb a significant portion of the electron's energy, resulting in a higher degree of range straggling [5].

2.3.5 Bragg Peak

For massive charged particles, the rate of the energy loss by ionisation, per travelled distance, is slowly varying and approximately constant over a widespread kinematic range. In Figure 2.6, this corresponds to the momentum region between about 3-1000 GeV/c. Below the point of minimum ionisation, at $\beta\gamma \approx 3GeV/c$, there is a rapid increase in the rate of energy loss. In other words, the majority of energy loss occurs in the range below the point of minimum ionisation [9].

Hence, a curve showing the energy loss distribution as a function of penetration depth will seem approximately constant until the momentum is below this point, after which a peak of energy deposition will follow. This curve is called the *Bragg Curve*, illustrated in Figure 2.10, and the peak of energy deposition is known as the Bragg Peak [9].

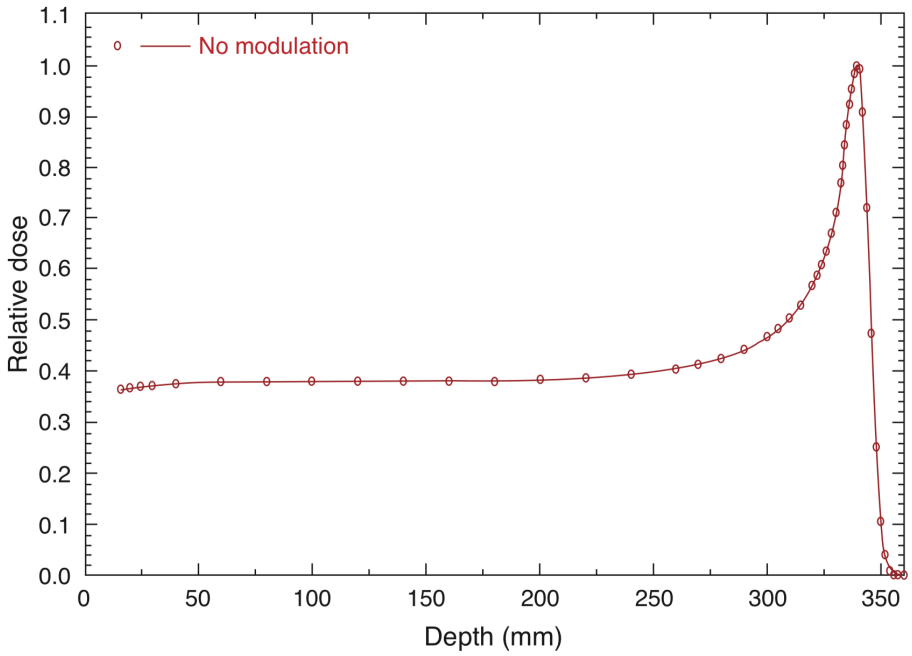


Figure 2.10: Bragg Curve for an unmodulated 250-MeV proton beam illustrating the relative dose versus depth. The sharp energy deposition is known as the Bragg peak. The figure is from [4].

The Bragg Peak can be a remarkably beneficial feature in both medical applications and nuclear physics. On one side, the Bragg Peak gives the opportunity of precisely depositing energy at the desired depth, a characteristic utilised in cancer therapy as an alternative to conventional treatment with x-rays. On the other side, one can also use the feature to calibrate the energy of a particle beam by directing the beam to a stopping material of known thickness, composition and peak position [9].

A substantial part of the technique of particle radiotherapy depends on skilful manipulation of the Bragg Peak. An example is spreading it out over a larger area, a method used to spare as much healthy tissue as possible, as well as covering the target volume. The Spread-out Bragg Peak (SOBP), as shown in Figure 2.11, is obtained by adding weighted and shifted Bragg Peaks [11]. It is generated by using a mono-energetic beam with enough range and energy to reach the distal end of the target volume, followed by adding beams of decreasing intensity and energy to cover the proximal part [4].

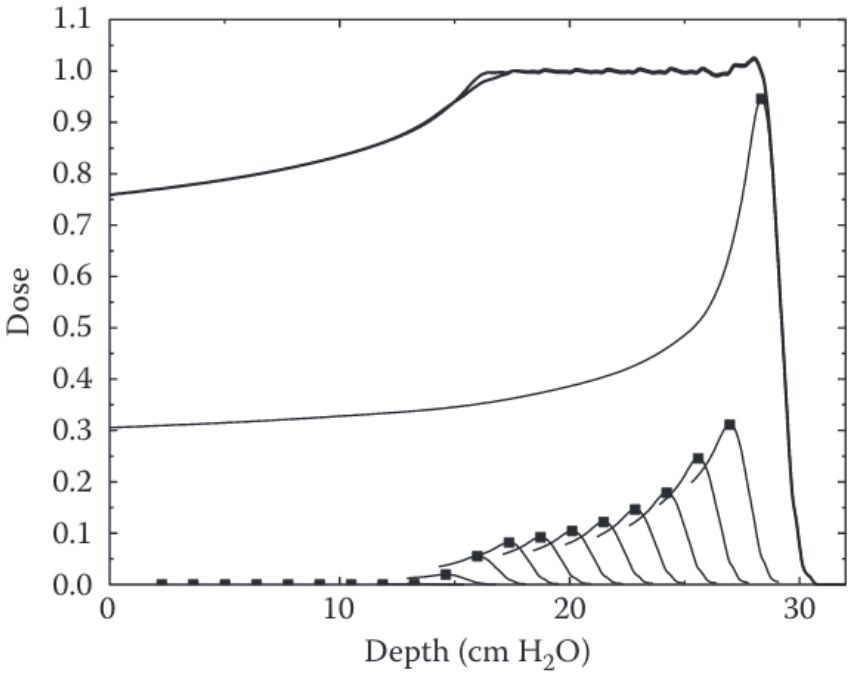


Figure 2.11: Spread-out Bragg Peak (SOBP) gained from adding appropriately weighted and shifted Bragg peaks. The figure is taken from [11].

2.4 Electromagnetic Showers

Initiated by an energetic photon, electron or positron, an electromagnetic shower is a chain reaction into a substance where multiple photons, electrons and positrons are formed or released. The chain of interaction processes is triggered when the energetic particle enters a dense target material [9]. Energetic electrons and positrons mainly lose energy through the bremsstrahlung mechanism, as mentioned in Section 2.3.2. At the same time, Figure 2.1 shows that the interactions that dominate for high-energy photons are the process of pair production, described in Section 2.2.3.

Figure 2.12 shows an example of this process, where an electron loses energy through bremsstrahlung, producing a photon. This photon undergoes pair production producing an electron-positron-pair, while the electron keeps loosing energy by bremsstrahlung. These processes continue, where the photons continue creating pairs of electrons and positrons, and the electrons and positron continuously loose energy by creating photons. The number of particles increases gradually, resulting in an electromagnetic shower.

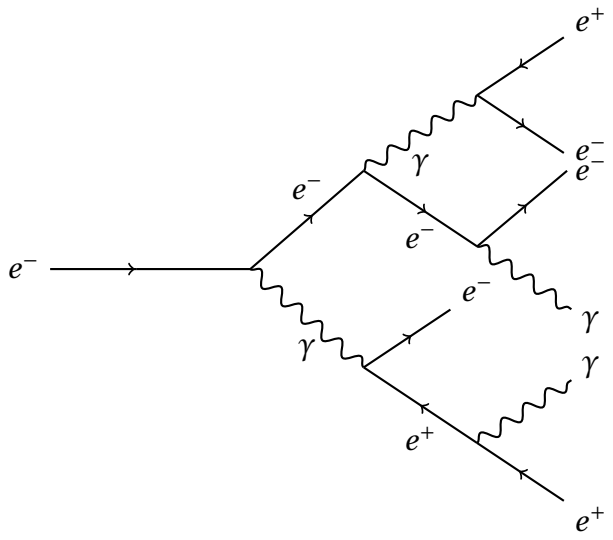


Figure 2.12: Diagram showing the development of an electromagnetic shower. The diagram is a modification of a figure from [9].

The mean thickness of a material (measured in g/cm^2) where an electron's or positron's energy has been reduced by a fraction of 1/e of its initial energy by bremsstrahlung alone, is known as a radiation length. After covering a distance of about one radiation length, a photon will proceed to create an electron-positron-pair. The shower will continue to develop until the secondary photons no longer have a sufficient amount of energy for further pair production. The longitudinal depth of the shower will, therefore, depend on the energy of the initial particle. The transverse size of the shower is, however, given by the following empirical formula:

$$R_{\mathrm{M}} = \frac{\sqrt{4\pi\alpha}\, m_{\mathrm{e}} c^2}{E_{\mathrm{c}}} X_0 \simeq \frac{21\mathrm{MeV}}{E_{\mathrm{c}}} X_0 \tag{2.25}$$

In this equation the parameter R_{M} is called the Molière radius, E_{c} is the critical energy in the material and X_0 is the radiation length. The critical energy in a material is the energy of an electron or a positron where the rate of energy loss by bremsstrahlung equals the rate of energy loss by ionisation [7][9].

It is also possible to estimate the longitudinal size of the shower by using a simplified model. By letting t be a measure of depth in units of radiation length, in short $t = x/X_0$, and by assuming that half the energy of the electron is lost to a single photon after every travelled radiation length, in other words X_0 measures the steps of each bremsstrahlung and pair production event. At a depth t, the number of particles in the shower, $N(t)$, will be:

$$N(t) = 2^t \tag{2.26}$$

and the numbers of photons, positrons and electrons are approximately the same:

$$N_\gamma \simeq N_{e^+} \simeq N_{e^-} \tag{2.27}$$

Furthermore, the average energy per particle at depth t, $E(t)$, is given by

$$E(t) = \frac{E_0}{N(t)} = E_0 \cdot 2^{-t} \tag{2.28}$$

where E_0 is the initial energy of the particle [9].

The shower continues to grow until the electrons or positrons have less energy than the critical energy E_c for bremsstrahlung. By inverting Equation 2.28 and using $E(t) \leq E_c$, the maximum penetration depth, t_{max}, is given by

$$t_{max} \simeq 1.4 \ln\left(\frac{E_0}{E_c}\right) \tag{2.29}$$

Moreover, the maximum number of particles, n_{max}, is given by

$$n_{max} = 2^{t_{max}} = \frac{E_0}{E_c} \tag{2.30}$$

The total integrated track length is also proportional to the initial energy E_0 [9].

2.5 Biological Effects

Radiotherapy aims to treat cancer by killing or sterilising the cancer cells and simultaneously minimising complications. By studying the radiobiology, combining elementary principles from biology and physics, one can investigate how living organisms and biological tissue react to ionising radiation.

When a biological system is irradiated, the radiation transfers energy to the system. On a molecular level, the transfer of energy to the biological system can be divided into two main categories [12]:

1. Direct effect, where the interactions between the atoms of the biological system and the radiation give rise to excitation or ionisation of the atoms.
2. Indirect effect, containing all the chemical and physical interactions with the radicals produced by the radiation.

The DNA is the main target for the ionising radiation, to avoid metastasis of the cancer cells. The ionising radiation can damage or break the DNA by single and double-strand breaks, where double-strand breaks are more severe than single-strand breaks. There has to be enough energy deposited in the DNA to result in a large enough number of double-strand breaks to destroy the cell [6].

A particles' *stopping power* is the average rate of energy loss per unit path length of the particle, $-dE/dx$. The linear stopping power of the particle, also known as the *linear energy transfer*, or LET, is closely correlated to the dose deposition and biological effect of radiation in the medium [4].

Furthermore, the *relative biological effectiveness*, or RBE, of a type of radiation is the ratio between the dose of 250-kVp x-rays providing a specific biological effect and the dose of a given radiation type producing the same effect. The specified biological effect can be tissue damage, mutation, cell killing or any other form of biological result [4].

LET is the most important factor of the RBE, although it also depends on the dose fractionation, the radiation quality and type, in addition to the biological endpoint. An increasing LET is associated with a greater RBE. Charged particles generally have a higher LET than x-rays, leading to an RBE greater than or equal to 1.0. The LET, and thereby the RBE, of charged particles, increase as the particles decelerate, approaching the end of their range. Hence, the RBE is at its highest in the Bragg Peak region [4].

Figure 2.13 shows the RBE plotted versus LET, where each point represents a cell experiment, and the colours indicate the different ions used, ranging from protons to heavier ions. As

the value of RBE increases, less dose is needed to achieve a particular effect. For points with higher LET and lower RBE, there will, however, be delivered a higher dose than needed, resulting in less effective treatment.

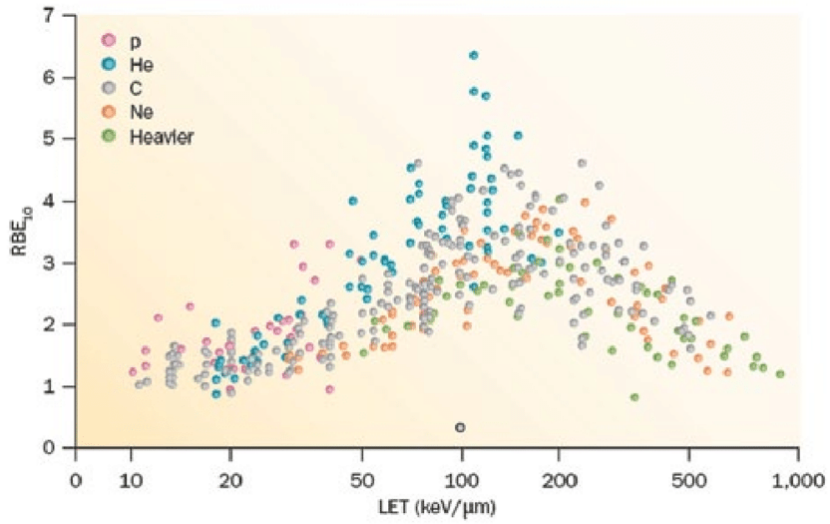


Figure 2.13: Relative biological effectiveness (RBE) versus linear energy transfer (LET) for in vitro cell lines, where the LET values are given in keV/μm in water. The different colours indicate the different ions used as radiation, ranging from protons to heavy ions. The figure is retrieved from [13].

Clinically, one has adopted a single rounded-off value of RBE. For protons, most treatment facilities use an RBE value of 1.1 for protons relative to megavoltage x-rays or cobalt-60. An adaptation implemented to bring the clinical response of photon and proton beams to a rough agreement [4].

2.5.1 Absorbed Dose

Absorbed dose is a quantity defined as the absorbed energy per unit mass in a medium:

$$\text{radiation absorbed dose} = \frac{dE}{dm} \tag{2.31}$$

where dE is the average energy deposited by the ionising radiation in a material of mass dm, with uniformly distributed mass. The unit for absorbed dose in the SI system is called *gray* [Gy] and corresponds to joule per kilogram, 1 Gy = 1 J/kg [6].

Unlike exposure, defined as the quantity of x-rays or γ-radiation that produces ions carrying a charge of one coulomb in air, per kilogram of air, the absorbed dose can describe the quantity for all types of ionising radiation [6].

2.6 Radiotherapy

Radiotherapy is a focused course of treatment that is frequently used for cancer patients, regularly in combination with other courses of treatments, like chemotherapy and surgery. Radiation therapy utilises radioactive sources, both internal and external sources. The most common is, however, the use of external sources emitting x-rays and other high energetic photons [1].

The molecules in the patient interact with the high energy particles by ionisation, producing charged molecules known as free radicals. These free radicals can damage the cancer cells' DNA structure. The damage inflicted by the radiation, by ionising molecules and damaging DNA, can also cause late effects and even give the patient new cases of cancer [1].

Thus, radiotherapy aims to limit the unwanted irradiation of healthy tissue, and primarily deliver ionising radiation to the tumour. The radiation dose delivered by high energetic x-rays has an exponential decay through the patient, meaning that the radiation gradually deposits energy in the tissue. Hence, the patient is irradiated from various angles using multiple beams, creating a so-called low dose bath, limiting the dose to healthy tissue whilst targeting the tumour with a high dose [1].

Although conventional radiotherapy employs photons, most commonly x-rays, to deliver radiation, there has, in the last decades, been an increase in the use of charged particles such as protons and heavier ions. As discussed in previous sections and shown in Figure 2.14, charged particles deposit energy differently than photons. As the velocity of the particles traversing the patient decreases, the particles have more time to ionise the molecules. Consequently, there is a limited area where almost all the particles stop, giving a high ionisation frequency and hence a high dose deposition. This limited area is known as the Bragg Peak, reviewed in Section 2.3.5. As can be seen in Figure 2.14, there is almost no dose deposition after this point [1][14].

Henceforth, the physical characteristics of charged particles traversing a medium provide methods of delivering high dose at a desired depth in the patient. The accuracy of the dose delivery will, therefore, limit the irradiation of healthy tissue resulting in a reduced risk of long-term side effects caused by undesired dose to healthy tissue [1].

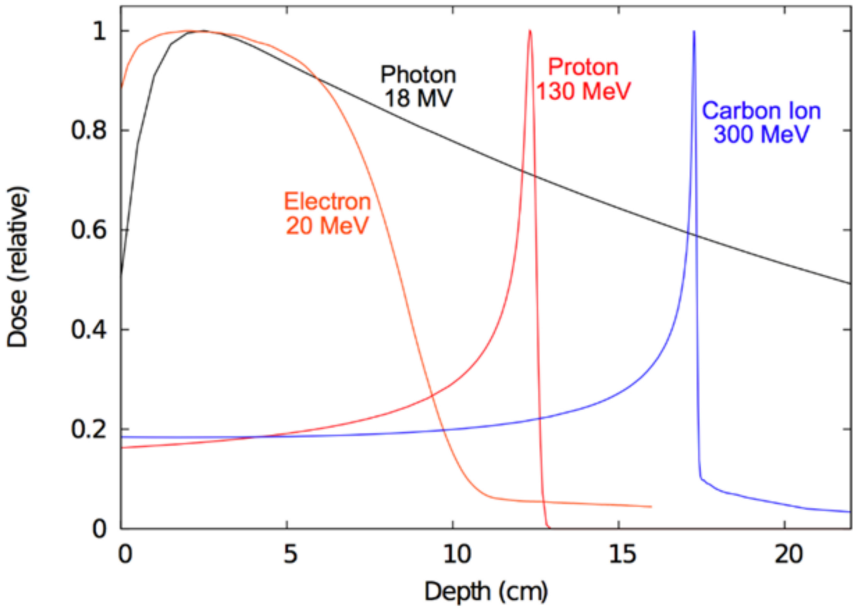


Figure 2.14: Distribution of dose versus depth in water for several clinical radiation beams. The figure is from [15].

2.6.1 Photon Therapy

Conventional radiotherapy utilises a linear accelerator to deliver a high energy beam of x-rays to the site of the patient's tumour to immobilise the cancer cells. This treatment method is also known as external beam radiation, since an external source produces the beams. External beam radiation can be delivered using various techniques which aim to improve the precision of dose delivery to the tumour [16].

Image-guided radiotherapy makes use of frequent imaging to give highly accurate and precise irradiation. One can create and inspect images displaying the tumour site both before and during treatment [16].

Intensity-modulated radiation therapy is a radiation technique that employs radiation beams controlled by a computer together with three-dimensional computed tomography images of the cancer site and surrounding areas. By modulating the intensity of the radiation beams, this technology delivers targeted doses to the tumour site designed to match the tumour shape [16].

Volumetric Modulated Arc Therapy is a treatment procedure that significantly reduces the treatment time by the use of single or multiple beams sweeping around the patient. Supported by three-dimensional imaging modalities, aiding in radiation delivery and precision, it allows visualisation of the cancer site at the time of treatment [16].

2.6.2 Particle Therapy

Proton therapy is a method of targeted radiotherapy that utilises the energy deposition of protons. Due to the characteristics of dose depositions for charged particles, protons minimise the radiation exposure of neighbouring healthy organs and tissue and deliver most of their energy to the tumour site. This precise therapy treatment, therefore, reduces the risk of late effects as a consequence of the treatment itself [16].

Accelerators producing proton beams for medical use are also conventional in the fields of nuclear and atomic research. Cyclotrons and synchrotrons are accelerators that can produce protons, with the beams typically having energies between 200 MeV and 250 MeV [17][18].

Carbon ions exhibit the same characteristic dose distribution as protons, where the majority of the energy is deposited in the area where the Bragg Peak is located. Unlike for protons, there is an energy deposition in the distal tissue due to nuclear fragmentation, as can be observed from Figure 2.14. However, carbon exhibits a higher LET than both photons and protons, leading to a higher RBE. As a result of its higher LET compared to other radiation methods, in addition to its characteristic energy distribution with respect to depth, carbon ion radiation therapy provides a promising treatment option for reducing irradiation of organs at risk, as well as providing higher dose deposition to the targets [19].

2.7 Imaging Modalities

Medical imaging plays a significant role in both the diagnosis and assessment of the response of treatment with radiotherapy. Computed tomography, CT, is often considered to be the gold standard with regards to imaging for treatment planning, particularly concerning geometric accuracy. The modality utilises the transmission of x-rays through the body to produce a medical image, resulting in excellent bone contrast and overall good soft-tissue contrast. Furthermore, it can achieve a small pixel size, resulting in high resolution, in addition to reasonably quick and comfortable image acquisition [17][20][21].

Magnetic resonance imaging, or MRI, makes use of radio waves and a strong magnetic field to generate detailed diagnostic images, and provides important additional information, especially with regards to the target description. Due to the acquisition method and principles behind MR, the bones appear as black areas since they do not exhibit any signal. Small differences in soft tissue composition are, on the other hand, easy to observe [17][21].

Positron emission tomography, or PET, is a modality based on biologically active, positron-emitting radiotracers that get injected into the patient. Hence, the functional PET images provide complementary information to the structural images obtained from CT and MR,

due to their reflection of the biochemical function of organs and tissue. This is illustrated in Figure 2.15 [17][21].

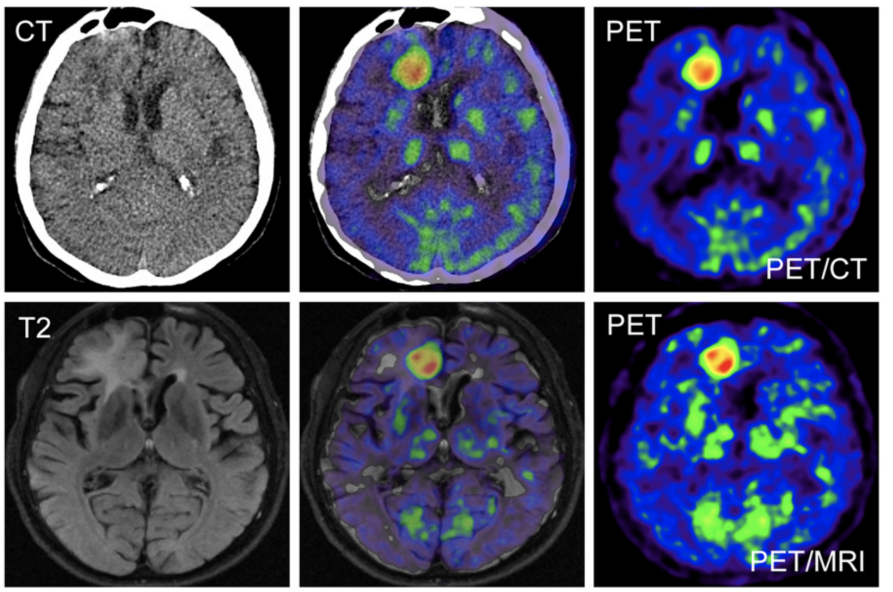


Figure 2.15: (Top) Non-contrast-enhanced low-dose CT scan (left), ^{11}C-methionine PET scan (right) and the corresponding fusion image (centre). (Bottom) MRI scan producing a T2-weighted FLAIR image (left), PET scan (right) and the fusion image (centre). The figure is retrieved from [22].

In the following section, computed tomography using photons will be discussed before undertaking the concept of proton computed tomography. CT using x-rays will be reviewed in more detail than the other modalities, seeing that this technique is closely related to proton CT, and regarding proton therapy, a modality that proton CT aims to exceed.

2.7.1 Computed Tomography (CT)

X-rays are widely used in medical imaging, for instance, in planar x-ray radiography and computed tomography, CT. A different number of x-rays are absorbed as they pass through the body, depending on the density of the medium they traverse. Materials of high density, e.g. bone and metal, will appear white, while media of lower densities will appear darker; fat and muscles will appear as shades of grey, while the air in the lungs looks black [21].

The fundamental principle behind computed tomography is that the two-dimensional structure of an object, in this case the position-dependent attenuation coefficient of x-rays, can be reconstructed from a series of one-dimensional projections obtained from different angles, followed by appropriate image reconstruction [21]. Three-dimensional reconstruction of the imaged object can be obtained by combining several of the two-dimensional pictures

generated by the CT. Gaps separate these two-dimensional images, so to generate a volume of data, these gaps are filled using interpolation [23].

A system consisting of semiconductor detectors is located on the opposite side of the x-ray tube, and together they acquire a one-dimensional projection of the patient, as shown in Figure 2.16. The detectors and the x-ray source are rotated a full revolution around the patient, where data is obtained nearly continuously. Two collimators are used in front of the source. The first collimator limits the beam to an angular width of 45-60°, while the second, which is placed perpendicular to the first, limits the beam to the desired slice thickness, typically 1-5 mm in the patient's head-to-foot-direction [21].

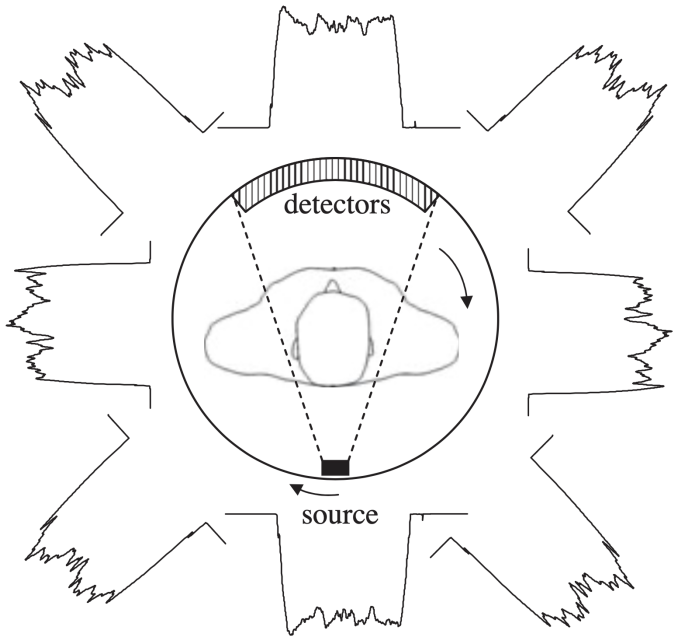


Figure 2.16: This figure illustrates the physical principle of computed tomography, recording a series of one-dimensional projection by synchronous rotation of multiple detectors and the x-ray tube. The figure is from [21].

An image generated form a CT scan in itself does not display a map of spatially-dependent tissue attenuation coefficients, but instead, it exhibits a map of the tissue CT numbers defined by:

$$\mathrm{CT}_{\mathrm{tissue}} = 1000\,\frac{\mu_{\mathrm{tissue}} - \mu_{\mathrm{H_2O}}}{\mu_{\mathrm{H_2O}}} \tag{2.32}$$

where $\mathrm{CT}_{\mathrm{tissue}}$ is the CT number for a given tissue expressed in Hounsfield units (HU), and μ_{tissue} and $\mu_{\mathrm{H_2O}}$ are the linear attenuation coefficients of the tissue and water in each pixel, respectively [21].

2.7.2 Proton CT

The precision of radiotherapy using charged particles, such as protons, makes extensive demands for accurate diagnostics, considering the damage the beam can cause if the Bragg peak misses the target and instead deposits a large dose in healthy tissue. Further, other factors may affect the positioning of the beam, such as the patient losing weight, or the tumour shrinking since the approval of the dose plan, or the target may be moving due to the patients breathing. Although these variations only change the level of dose in photon therapy by a few per cent, these are significant challenges for modern particle therapy. [1].

A suitable map displaying the stopping power of the particle in the patient is, therefore, needed to calculate the exact position of the Bragg peak in the patient. The stopping power, and hence the range of the particle, can be determined by using tomographic images obtained from a CT scan, but only within an uncertainty of 2-3%. The accuracy of the map displaying the three-dimensional stopping power in the patient will drastically increase by using the same particle for the CT scan as for the treatment. It is, therefore, preferable to use a proton CT, pCT, due to the demands of particle therapy for an accurate representation of the beam's range in the patient [1].

Proton CT has been acknowledged for its high potential in reducing uncertainties relating to treatment planning in proton therapy. As mentioned above, the strength of the pCT lies in the direct reconstruction of a 3D map of the target's relative stopping power, RSP. Relative stopping power is the stopping power of a particular material relative to that of water. The human body has a complex geometry; thus, an accurate model of the RSP values for each of the various materials the particles traverse before reaching the tumour is needed to predict the range of the beam correctly [24].

The concept behind the modern pCT scanners is based on tracking the history of every single proton. This tracking is done by measuring the position and direction before and after the imaged object and registering the particle's range or residual energy after crossing the object. A typical pCT system must, therefore, consist of both an energy or range detector and thin tracking detectors [24].

As a result of Multiple Coulomb scattering, protons do not leave straight tracks across the target, which is affecting the spatial resolution of an image retrieved from pCT. Hence, each single proton trajectory is reconstructed by applying several methods of trajectory estimation using the most likely path formalism. The measured particle information can then be processed through sophisticated algorithms for image reconstruction, and thereby directly generate a 3D map of the object's RSP values [24].

In the past 20 years, many pCT systems have been proposed and developed with both exper-

imental and Monte Carlo simulated setups, delivering encouraging results concerning both spatial resolution and RSP accuracy. Recently, it was proven that the RSP accuracy for pCT is well below 1%, meaning that it has a better performance than the modern techniques used today, such as dual-energy CT scans. It has also been shown that there is a lower dose deposition in the patient and a reduced noise level when using pCT compared to conventional x-ray CT. Hence, pCT scans can be performed frequently, e.g. just before delivering each treatment fraction, facilitating adaptive and personalised therapy [24].

2.7.3 The Bergen pCT Collaboration

Established at the University of Bergen amongst numerous institutions across the world, the Bergen pCT collaboration aims to design and build a pCT prototype that overcomes the crucial limitations of the currently existing prototypes. The most distinctive feature of this Bergen pCT design is the utilisation of a digital tracking calorimeter, DTC; a layer-by-layer pixel detector consisting of pixelated silicon sensors [24].

A calorimeter with these features has been built and tested using particle beams, and despite some shortcomings in the form of dead or unstable pixels, it is showing great performance. The Bergen pCT collaboration has benchmarked the DTC prototype for proof-of-concept for pCT purposes, through participation in the ALICE collaboration at CERN. Here the idea is to use the same technology for both tracking and measurement of residual energy, where the aim is to ensure stable operation in a clinical environment and to simplify the assembly of the pCT system [24]. The setup of this system is illustrated in Figure 2.17.

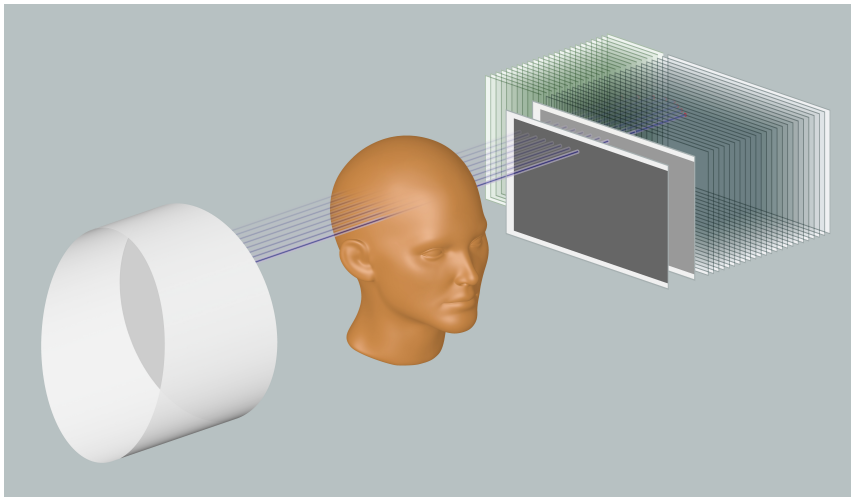


Figure 2.17: The setup of the Bergen pCT system. The beam of protons traverses through the patient before reaching the digital tracking calorimeter placed behind the patient.

The digital tracking calorimeter will operate as both the tracking system and as a detector

for range or energy measurements, and it is designed as a structure of multiple detector/absorber sandwich layers. The detector will track the protons crossing the imaged object and assign a water equivalent path length, WEPL, or energy difference to each of the particles. This data will then be applied in CT and image reconstruction [24].

The tracking system applied in most of the currently available pCT scanners are composed of two tracking layers in front of the imaged object, and two more layers behind it. In the Bergen pCT scanner, on the other hand, the front tracker is excluded. Information about the direction and position of the impinging protons will thus be gathered from the monitoring system's beam optical response. This design will reduce the final cost, the complexity of the setup, as well as the physical impact on the treatment room. It will also allow for a higher particle rate since the pairing of particle hits measured in the front with the measurement on the rear tracker can be avoided [24].

Chapter 3

Semiconductors and Silicon Pixel Detectors

This chapter outlines the fundamental principles of semiconductors and silicon pixel detectors. It will explain how charged particles can produce a signal, as well as the concepts and design of silicon pixel detectors and Monolithic Active Pixel Sensors. Lastly, it will cover the principal variables influencing the size of clustered pixel hits.

3.1 Semiconductors

Atomic electrons in a solid-state matter can populate several permitted energy levels, where one can consider closely spaced energy values as energy bands. For materials having semiconducting or insulating electrical properties, the electrons bound to lattice sites is considered to be a part of the *valence band*. In contrast, mobile electrons migrating through the lattice atoms belong to the *conduction band* [9].

At a temperature of absolute zero ($T = 0$ K), where there is an absence of thermal excitation, the conduction band is empty. The valence band is, on the other hand, full, which results in no electrical conduction. Under the influence of an external electric field, the electrons in the conduction band are free to move, unlike the electrons in the valence band [9].

Semiconductors and insulators differ by the size of the energy gap between the conduction and valence bands. Materials with semiconducting properties have an energy gap that is narrow enough to allow electrons to enter the conduction band as a consequence of normal thermal excitation. The size of the energy gap is, consequently, essential in determining a material's electrical properties [9].

Silicon is one of the most frequently used semiconductor materials for particle detectors and it has an energy gap of 1.16 eV at a temperature of absolute zero. It is abundant and relatively inexpensive and has a sufficiently large energy gap at room temperature compared to its thermal energy. Silicon, therefore, only requires cooling to decrease radiation damage and for applications demanding very low noise [9].

Silicon has an atomic arrangement with four electrons in the outermost shell, forming four covalent bonds with neighbouring silicon atoms in a cubic crystal structure. Covalent bonds are a chemical bond between atoms consisting of shared pairs of electrons.

The *band model of solids* states that the discrete electron energy levels of individual atoms or ions in a crystal are merged, creating energy bands. Each band can, according to the Pauli exclusion principle, only hold a limited number of electrons. Hence, low energy bands might be filled with electrons, while bands of higher energy may be empty - especially at low temperatures [10].

Between the bottom of the conduction band, V_C, and the top of the valence band, V_V, there is, a gap often referred to as the *energy gap* or the *forbidden band*. This gap has a width of $E_g = V_C - V_V$ [10].

Electrons can enter the conduction band at temperatures higher than absolute zero, as a result of thermal excitation, leaving an empty spot in the valance band. These empty spots are generally referred to as *holes* [9]. Under the impact of an external electric field, the electron starts to migrate. The same applies to the hole, which represents a net positive charge, but the hole moves in the opposite direction relative to the electron's motion [25].

There is, at any given temperature, an equilibrium in the number of free electrons and holes, and the following expression gives the probability per unit time for thermal generation of an electron-hole pair:

$$p(T) = CT^{3/2}\exp\left(-\frac{E_g}{2kT}\right) \tag{3.1}$$

In this equation C is a proportionality constant which is characteristic of the material, T is the temperature of the semiconductor in Kelvin, E_g is the bandgap energy at absolute zero ($T = 0$ K), and k is Boltzmann's constant [9][25].

As seen in Equation 3.1, the thermal excitation probability is significantly dependent on the ratio between the bandgap energy and the absolute temperature. The electron-hole pairs created by thermal excitation will ultimately recombine in the absence of an external electric field. If this is the case, an equilibrium will be established, where the electron-hole pair concentration observed at any given time is proportional to the formation rate. This concentration is, according to Equation 3.1, strongly dependent on the temperature, decreasing considerably with the cooling of the material [25].

When primary radiation or secondary particles pass through a semiconductor detector, pairs of electrons and holes are generated along the path. These electron-hole pairs are the fundamental information carriers, and when applying an electric field, their motion generate the primary electrical signal from the detector [25].

Following their creation, both the holes and the electrons diffuse away from their point of origin, due to random thermal motion. If the electron-hole pairs were all produced at the same location, the distribution of charges would spread over time. A Gaussian function with a standard deviation of σ can be used to approximate the cross-section of this distribution. The standard deviation can be parametrised as follows:

$$\sigma = \sqrt{2Dt} \tag{3.2}$$

Here, D is the diffusion coefficient that depends on the temperature and the mobility of the electrons, and t is the time elapsed [9][25].

The following relationship can be used to estimate the values of D:

$$D = \mu \frac{kT}{e} \tag{3.3}$$

where the mobility of the charge carriers are given by μ, k is the Boltzmann constant, T is the absolute temperature, and e is the elementary charge [25].

If one applies an external electric field to the semiconductor, the holes and electrons will start to *drift* with an average velocity given as follows:

$$v_{\mathrm{e}} = \mu_{\mathrm{e}}(E,T) \cdot E \tag{3.4}$$

$$v_{\mathrm{h}} = \mu_{\mathrm{h}}(E,T) \cdot E \tag{3.5}$$

μ_{e} and μ_{h} are the mobility functions for electrons and holes, respectively and depend on both the temperature, T, and the applied electric field, E. For electric fields up to 1 kV/cm, the mobility functions are approximately constant. For higher electric fields, the mobility functions decrease with an inverse proportionality to the field, $\mu \sim 1/E$. Hence, the drift velocities grow linearly with E before displaying a tendency to saturate to a constant value as higher fractions of energy are lost to collisions with atoms in the lattice [9][25].

The mobilities of electrons and holes have a functional dependence of temperature which is approximately given by

$$\mu \sim T^m \tag{3.6}$$

where $m \approx -2$ [9].

As mentioned in the previous section, the charge carriers are affected by diffusion in addition to their drift. Without this effect, the charge carriers would follow the electric field lines connecting their collection and origin point precisely. A Gaussian distribution can characterise the spread in arrival position introduced by the diffusion effect. For a drift distance of length x, the width of the distribution is as follows:

$$\sigma = \sqrt{\frac{2kTx}{eE}} \tag{3.7}$$

The broadening of the charge distribution limits the precision to which measurements of position can be performed using the location of charge collection at the electrodes. Diffusion also spreads out the collection time of the charges. Although the effect can be considered negligible for small volumes, the outcomes can become significant for detectors of large volume, or measurements requiring high spatial or temporal precision [25].

3.2 P-N Junction and the Depletion Zone

All the holes in the valence band and the electrons in the conduction band in pure semiconductors are, in the absence of ionising radiation, a result of thermal excitation. Consequently, the number of holes in the valence band is precisely equal to the number of electrons in the conduction band, since each electron must leave a hole behind as it bridges the energy gap. Materials with these properties are referred to as natural semiconductors. Although one can describe their characteristics theoretically, it is virtually impossible to accomplish in practice. Natural materials tend to have microscopic levels of residual impurities, which is the case even for semiconductors with the highest practical purities possible, such as silicon [25].

The term crystal impurities refer to the case where other atoms are positioned in the crystal lattice, and the presence of these impurities results in additional energy levels within the bandgap. These supplementary energy levels can trap electrons, and they also have the ability to speed up the recombination of holes and electrons, slow the process down or even block it altogether [9].

In a standard crystalline structure, silicon forms four covalent bonds with neighbouring silicon atoms and is therefore referred to as a tetravalent atom. One can dope a silicon crystal by adding, for instance, pentavalent atoms such as phosphorus, creating an *n-type* semiconductor. When the phosphorus is added to the crystal structure, a parentless electron is located at an energy level slightly under the level of the conduction band. This electron can easily reach the conduction band with the help of thermal excitation, where the difference

in energy between the impurity level and the conduction band, for silicon, is as small as 0.01 eV. By dopant concentration of only one per million, the donor density is approximately 10^{17} cm^{-3}, potentially resulting in various orders of magnitude more free electrons than the pure semiconductor's typical thermal concentration [9].

The concentration of donor electrons, N_D, dominates the concentration of free electrons, n, after all the weakly bound donor electrons have been excited. Hence, $n \simeq N_D$. Additionally, there is a reduction in the number of free holes, p, caused by the abundance of electrons in such a manner that, in thermal equilibrium

$$n \cdot p = n_i \cdot p_i \tag{3.8}$$

where n_i and p_i are the intrinsic concentration of electrons and holes in the semiconductor, without the dopant. The total number of charge carriers is much greater than the number of inherent charge carriers. Consequently, the electrical conductivity in an n-type doped semiconductor is almost exclusively due to electrons [9].

Silicon crystals can also be doped with trivalent impurities like boron. For each boron atom in the silicon lattice, there is one covalent bond left unsaturated. This arrangement is called a *p-type* semiconductor, and the configuration creates an empty energy level positioned right above the valence band. This vacant energy level, often referred to as an acceptor, will be filled with electrons from the crystal, leaving holes in the electronic structure of the crystal. The concentration of holes, p, will be approximately equal to the concentration of acceptors, N_A. Equation 3.8 implies that the product of the concentration of free holes and electrons is equal to the product of the intrinsic concentrations. Hence, for p-type semiconductors, the majority of charge carriers are holes [9].

A p-n junction is created by combining a p-doped and an n-doped semiconducting material, which in practice is done by doping two parts of a single semiconductor differently. Electrons are abundant in the n-doped side of the semiconductor which migrates to fill the excess of holes in the p-doped region. This process leads to a net electric charge building up in the proximity of the junction, resulting in an electric field that opposes additional charge diffusion from one region to the other. There is a relative vacancy of free charges in this zone under these conditions. This physical zone is therefore referred to as a *depletion* zone, where any free hole or electron is swept away by the zone's electric field [9].

The properties of the depletion zone are especially attractive for the purpose of radiation detection. When a semiconductor is irradiated with ionising radiation, electron-hole pairs will be liberated in the material. The charge carriers will then drift under the impact of an electric field. The width of the depletion region is, however, usually small, and depends on the concentration of dopants. The width of the depletion region can be estimated as follows;

assuming a charge density, $\rho(x)$, approximated by a constant within the depletion zone and a distance, x, from the contact point between the n-doped and p-doped semiconductors, then

$$\rho(x) = \begin{cases} +eN_D, & \text{for } 0 < x < x_n \\ -eN_A, & \text{for } -x_p < x < 0 \end{cases} \tag{3.9}$$

where N_D and N_A are the concentrations of donor and acceptor impurities. Furthermore, x_n and x_p are the extensions of the depletion region on the n-side and the p-side, respectively. By conservation of total charge, it follows that:

$$N_A x_p = N_D x_n \tag{3.10}$$

Although the depth of the depletion zone is relatively small, the device can become a diode by applying an external potential of appropriate size across the p-n junction, making it one of the most useful detectors. By implementing a forward bias, meaning that the positive terminal is attached to the p-type side of the semiconductor, the current will flow across the device. A reverse bias, where the positive terminal is connected to the n-side, will, on the other hand, pull the electrons away from the junction. Likewise, the holes on the p-doped side will pull away from the negative terminal. Consequently leading to a widening of the depletion zone. The applied voltage, V_{bias}, will arise over the depletion region due to the much higher resistivity of the depletion zone compared to the p-side and n-side alone. There will in practice, however, be a small *leakage* current that flows across the p-n junction when a reversed bias is applied, and device-manufacturers are attempting to minimise this effect [9].

In the p-n junction of a semiconductor, the depletion zone represents a volume where the electron-hole pairs, released by the interaction of radiation with the semiconducting material, can easily be collected at the terminals. The electron-hole pair's low average ionisation energy of about 3 eV yields advantages such as a better signal to noise ratio for weaker signals and lower uncertainty in energy resulting from a smaller energy fraction per carrier [9].

3.3 Pixel Detectors

In image processing, the pixel, which is short for "picture element", describes the smallest element in a given device or process. Consequently, a pixel detector is a device capable of detecting an image, and the granularity of this image is given by the size of the pixel [26].

Figure 3.1 displays a diagram describing a pixel cell of a hybrid pixel detector. The detector is referred to as hybrid, since the sensor and electronics are produced separately and then

paired. The signal is produced by an ionising particle passing through the sensor. The particle generates charges that, under the action of an electric field, moves in the depletion zone. The moving charges are then detected as signals. These signals are amplified, and the pixels hit by a particle are identified and stored by the electronics [26].

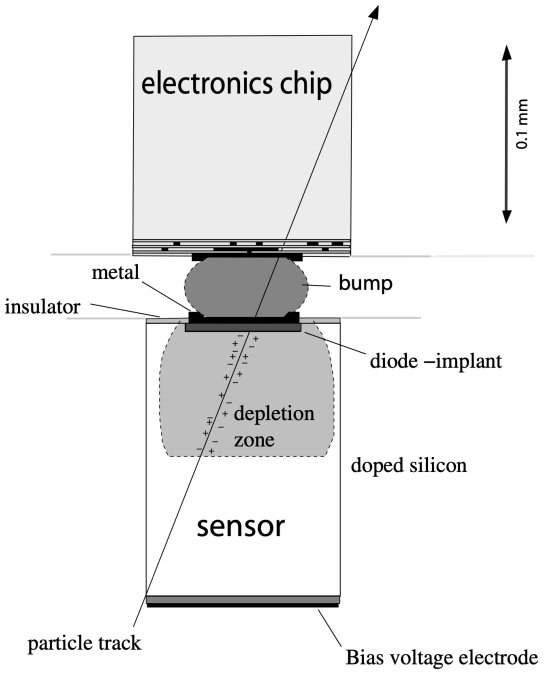


Figure 3.1: A schematic outline of a pixel cell, which is the basic module of a hybrid pixel detector. The figure is retrieved from [26].

In particle physics, the development of pixel detectors has principally been set off by two particular requirements, which in most applications have to be met simultaneously:

1. The possibility of examining short-lived particles
2. The capacity to cope with the increasing energy and interaction rates of modern particle accelerators

A detector with high granularity is needed to satisfy the two requirements listed above, where the detector should be able to detect multiple tracks with high temporal and spatial resolution [26].

Monolithic Active Pixel Sensors

Detectors using monolithic active pixel sensors, also known as MAPS, are produced using CMOS, Complementary Metal-Oxide-Semiconductor, technology. Here, the epitaxial layer,

which is the active volume of the detector, has a thickness of a few to 20 μm. The epitaxial layer is composed of silicon which is slightly p-doped and grown on a highly p++ doped substrate. The complementary wells of n- and p-type are embedded on top of the epitaxial layer. The electrons released by ionising particles entering the detector are collected by a diode which is established by the junction of an n-well and the p-type epitaxial layer [27]. A sketch indicating the operation principle of a monolithic active pixel sensor is shown in Figure 3.2.

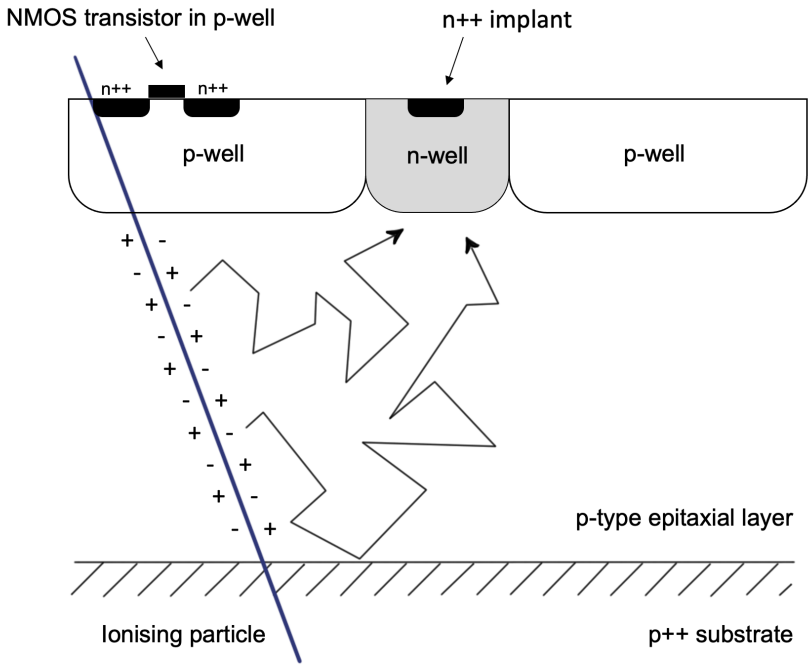


Figure 3.2: Sketch displaying the principle of how monolithic active pixel sensors operates. The sketch is based on a figure from [27].

The charge carriers primarily reach the collection diodes by thermal diffusion, since most of the epitaxial layer is unaffected by the electric field. Although the charge released in the highly doped substrate is mostly lost as a result of fast recombination of carriers, some fraction of these carriers can diffuse to the epitaxial layer and thereby contribute to the total charge collection. There is a difference in magnitude of three orders between the doping levels of the lightly doped epitaxial layer and the p++ substrate and p-well. This difference gives rise to a potential barrier at the boundaries, acting as mirrors for excess electrons [27].

In the MAPS detector, each pixel has electronics for signal processing. These electronics are based on NMOS, N-type Metal-Oxide-Semiconductor, transistors. The transistors are installed in a p-well, as seen in Figure 1.4, resulting in shielding of the source and drain junction from the epitaxial layer. The placement of the NMOS transistors is crucial, since these n+ drains and sources would otherwise serve as collection electrodes, preventing the n-well

from collecting all the charges that generate the signal [27][28].

MAPS detectors possess excellent tracking abilities, due to their high detection efficiency exceeding 99% as well as their excellent single point resolution. In addition to this, MAPS detectors are less sensitive to radiation damage and can be thinned down to 50 μm which minimises the number of multiple scattering events. Furthermore, these detectors have low power dissipation and fast readout due to their readout architecture [27].

Clusters of Pixel Hits

The charge created by a particle hitting the detector will lead to a cluster of pixel hits, due to charge diffusion [29], and the average size of this cluster has a significant effect on the detector's resolution and on the amount of data that is being transferred from the chip. The distribution of cluster sizes can change based on the operating environment, the settings of different chip parameters and the momentum of the measured particles [30].

By lowering the threshold settings, the charge sharing between adjacent pixels can reach an adequate level making these neighbouring pixels fire as well, resulting in larger clusters. A reverse substrate bias may also be applied, leading to a significant altering of the size of the depletion zone, where higher values of negative bias give rise to a larger depletion region. It is easier for the electrons to reach the depletion zone if a larger part of the epitaxial layer is depleted, resulting in less diffusion and less firing pixels for each crossing particle [30].

On average, there are also less firing pixels when the tracks pass through the centre of the pixel as opposed to tracks passing at the pixel's corners, since the charge's path length to adjacent pixels is longer if the particle passes through at the centre rather than at the corners. Additionally, the majority of charges are generated near, or inside, the depletion zone for particles crossing the centre. Electrons within the depletion zone drift towards the collection diodes and are, thus, less prone to the collection at a different pixel than at the pixel where it originated. On the other hand, if the particle passes through at the border between two pixels, the charge generated has to first reach a depletion zone by diffusion, and this region can be in one of the adjacent pixels [30].

Chapter 4

ALPIDE Chip

This chapter describes the essentials of the ALPIDE detector. Firstly the ALICE ITS upgrade will be presented, before introducing the ALPIDE chip. Then the ALPIDE operations will be described focusing on the front-end, priority encoders, as well as triggering and strobes. Ultimately, the applications of the detector in calorimeters are presented, concentrating on FoCal and proton CT.

ALICE, an acronym for *A Large Ion Collider Experiment*, is one of the experiments at the Large Hadron Collider, also known as the LHC, at the European Council for Nuclear Research, CERN. During the second Long Shutdown, LS2, in 2019/2020 the Inner Tracking System, ITS, for ALICE will be upgraded. The ITS Upgrade, ITS2, is based on the ALPIDE chip, developed for this upgrade by the ALICE Collaboration. The ITS2 will hence be the first large-scale silicon tracker operating at a collider that is based on the CMOS Monolithic Active Pixel Sensor technology [31].

As seen in Figure 4.1, the layout of the ITS2 contains seven cylindrical layers which are arranged in two sub-systems: the Inner Barrel and the Outer Barrel. The Inner Barrel is composed of three layers that are 27 cm long with radii of 2.3, 3.1 and 3.9 cm, respectively. The Outer Barrel consists of two Middle Layers with a length of 84 cm located at 24 and 30 cm from the line of interaction, in addition to two Outer Layers with a length of 148 cm at 41 and 48 cm, respectively. The layers have 12.5 G pixels and cover a total surface area of ~10 m^2 [31].

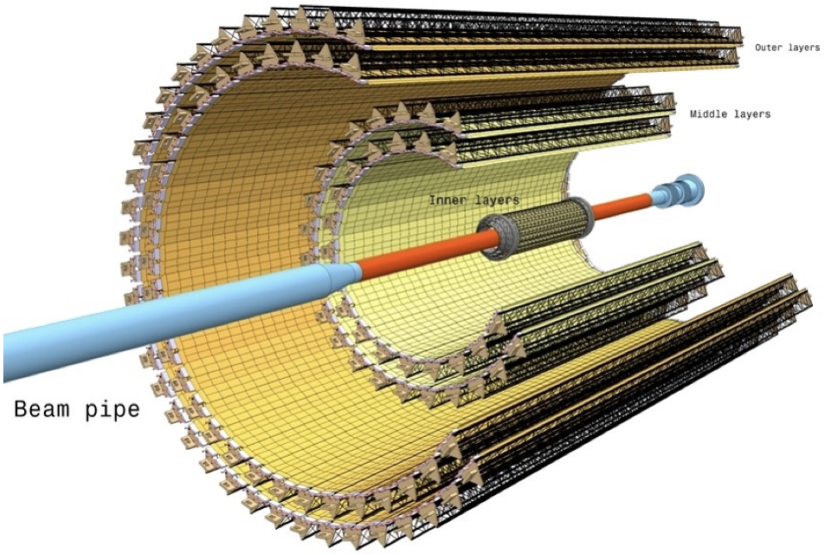


Figure 4.1: An overview of the ITS2 layout, with two Outer layers, two Middle layers and three Inner layers near the beam pipe. The figure is retrieved from [31].

The properties of the ITS2 design will result in an improvement of the tracking performance of the ALICE detector, due to the spatial resolution of the ALPIDE chip combined with the low material budget of the three innermost layers, which is as low as 0.35% of a radiation length. The readout electronics will also be capable of recording events at a typical rate of 50 kHz and a few 100 kHz for minimum bias Pb-Pb and pp collisions, respectively [31].

4.1 The ALPIDE Architecture

The ALPIDE, short for ALICE Pixel Detector, is a CMOS monolithic active pixel sensor where the readout and sensor matrix is integrated on a single chip. For the past eight years, an intensive effort within research and development was carried out by the ALICE collaboration. An unprecedented performance was reached in terms of spatial resolution, readout speed, material budget and signal-to-noise ratio. There is, however, still much potential to further improve MAPS for detectors in high-energy physics by fully utilising the fast progress made for this type of technology in the field of imaging for consumer applications [32].

The ALPIDE chip is produced in the 180 nm CMOS technology by TowerJazz [31][33], and Figure 4.2 shows a schematic cross-section of a MAPS pixel for this technology. Generally, the n-well of a PMOS transistor will compete for collection of signal charge with the n-well collection electrodes. This technology offers a deep p-well in the p-type substrate to avoid this. The deep p-well shields the n-well from the substrate, or epitaxial layer, preventing it from collecting signal charge, and thereby allowing the full use of CMOS circuitry in the pixel

area without hindrance from parasitic charge collection by the n-wells. This process also enables the use of a high-resistive epitaxial layer on a p-substrate, increasing radiation tolerance. Furthermore, the signal-to-noise ratio can be improved by increasing the depletion region around the collection diode, which is done by applying a moderate negative voltage to the substrate [28][31].

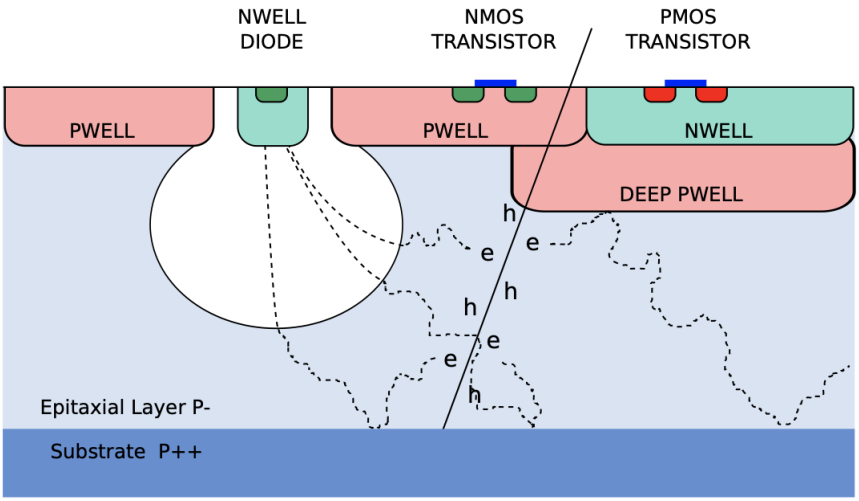


Figure 4.2: Diagram displaying the cross-section of a MAPS pixel by TowerJazz with 180 nm CMOS technology and the deep p-well feature. The figure is from [33].

The ALPIDE chip measures 30 mm × 15 mm, containing a pixel matrix of 1024 × 512 sensitive pixels, where each pixel measures 29.24 μm × 26.88 μm. The sensor also has an integration time of approximately 2 μs and a maximum readout speed of 1.2 Gbit/s. Moreover, it has an in-matrix address encoder with asynchronous sparsified readout, as well as in-pixel discriminators [31][34].

4.2 ALPIDE Operation

The architecture of the ALPIDE, displayed in Figure 4.3, implements a pixel-level sparsified readout, meaning that each pixel cell of the architecture contains a comparator followed by one (or more) latches. The latch is set to one every time a pixel has an analogue signal greater than the threshold of the comparator. All pixels that are set to one are read out asynchronously according to their priority in the chain. Their position within the chip determines the priority of the pixels. Two neighbouring columns of pixels are connected to the same priority encoder, and the addresses of the triggered pixels are stored at the column end in a regional buffer [33].

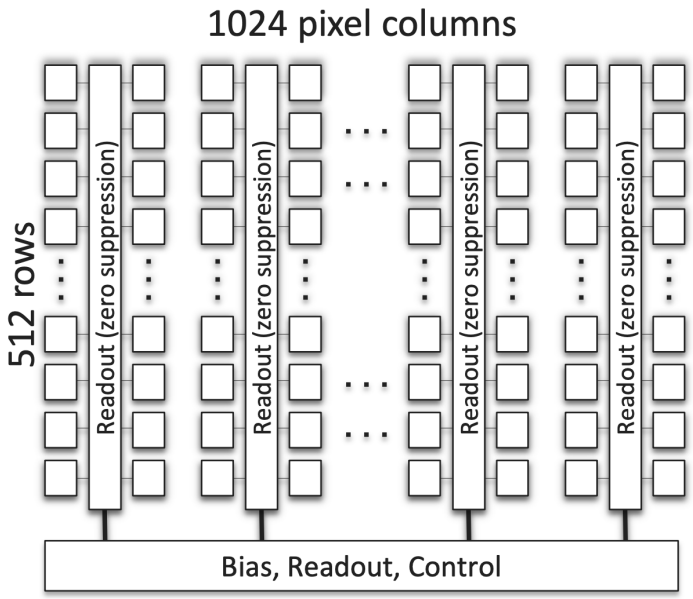


Figure 1.1: General architecture of the ALPIDE chip.

Figure 4.3: The ALPIDE chip's general architecture. The figure is from [34].

4.2.1 Front-End

As seen in Figure 4.4, each pixel of the ALPIDE chip holds a sensing diode, a front-end shaping an amplifying stage, a discriminator and a digital section. The digital section contains a pixel masking register, pulsing logic and three Multi-Event-Buffers, used for hit storage [34].

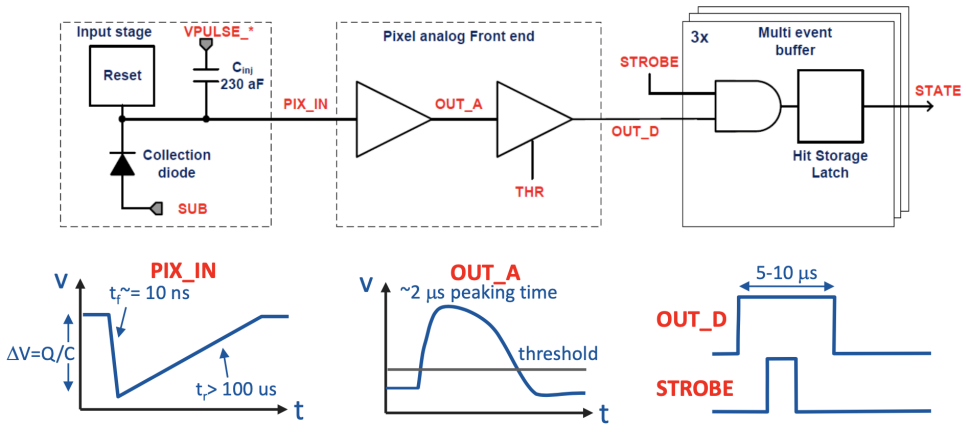


Figure 4.4: Block diagram displaying a pixel cell on the ALPIDE chip. The figure is retrieved from [34].

The discriminator and front-end are continuously active, featuring a non-linear response. The typical duration of the discriminator pulse is 10 μs, while the front-end output has a peaking time in the order of 2 μs. The discriminator and front-end serve as an analogue delay line, which allows the chip to operate in trigger mode in cases where the delay of the incoming trigger is comparable with the front-end peaking time [34].

Figure 4.5 displays a comprehensive scheme of the front-end circuit of a pixel. The diode named D1 is the sensor p-n junction, and diode D0 continuously resets the input node, where VRESETD establishes the reset voltage of pix_in, which is the charge collecting node. When a particle hits the pixel, the potential at the pixel input pix_in will be lowered. This lowering in potential will cause the current source M0, and the source follower created by M1 to force the source node to follow this voltage excursion. It will then dump the capacitance of the source node and charge associated with the change in voltage onto pix_out, which is the analogue output node [34].

ITHR, VCASN and IDB define the charge threshold of the pixel, where the effective charge threshold is decreased by increasing VCASN and increased by increasing ITHR or IDB [34]. Minimum, maximum and nominal values for these DAC specifications are listed in Table 4.1.

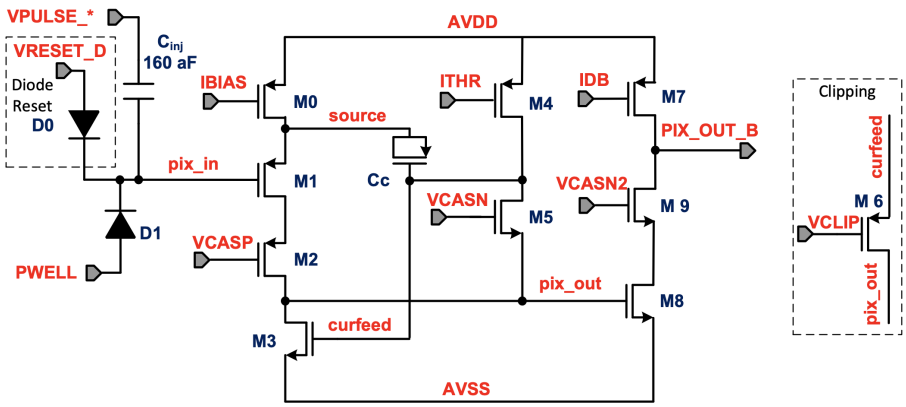


Figure 4.5: Scheme of the ALPIDE front-end. The figure is taken from [34].

Table 4.1: Overview of some of the DAC specifications [34]. The nominal settings are given in DAC units, where 1 DAC unit is equivalent to 10 e^- [35].

	Minimum	**Maximum**	**Nominal Setting**	**Nominal Value**
ITHR	0 nA	80 nA	51 DAC	0.5 nA
IDB	0 nA	80 nA	64 DAC	10 nA
VCASN	0 V	1.8 V	57 DAC	0.4 V

4.2.2 Priority Encoders

The ALPIDE is divided into 32 regions, where each region holds 16 double columns that are read out by 16 Priority Encoder circuits. The Priority Encoders occupy the regions in the middle of each double-column and define the pixel indexing in the readout data. Located in the chip periphery are corresponding 32 region readout units. Each submatrix is read out by one of the readout modules. The 32 submatrices are read out in parallel, while the 16 double columns in each region are read out sequentially [34].

During a transfer cycle, a hit pixel is selected before its address is generated and transmitted to the periphery. Afterwards, the memory element in the pixel is reset. Then the next pixel in the double-column with a valid hit will then have its address generated. The transfer cycle is repeated until all of the hit pixels have had their addresses transmitted to the periphery, and the memory elements have been reset. Hence, the data transferred to the periphery from the matrix is zero-suppressed [34].

4.2.3 ALPIDE Triggering and Strobes

The same threshold level is applied to all pixels in the matrix. Global strobe signals control the latching of the discriminated hits in the storage registers, and the duration of the strobe pulses are programmable. Then, if a strobe pulse is applied to the selected pixel cell at the same time as the output from the front-end is above the set threshold, a hit will be latched into one of the three memory cells of the pixel. The storage of the hit information in the pixel event buffers is controlled by three distinct strobe signals that are generated at the periphery and applied to all pixels. An external command, a trigger, can trigger the generation of the internal strobe signals, or it can be initiated by an internal sequencer [34].

The storage of the front-end discriminators' internal state within each pixel, and the transmission of this off-chip is the core functionality of the ALPIDE. In this context, a frame is a collection of pixel states at a specific time. The generation and transmission of a frame follow the reception of a trigger [34].

The ALPIDE can be operated in two different readout modes; triggered mode and continuous mode. The triggered mode is designed to sample over a relatively short time interval, where the coincidence between the assertion of a strobe signal and a discriminated signal determines the latching of a pixel hit. Thus, all the pixels that fire during the strobe are latched as hits. An external trigger source typically controls the start of the strobing intervals. Continuous mode, on the other hand, provides the readout of pixel hits sampled during periodically repeated strobing intervals. Here, the framing intervals are typically longer compared to the framing intervals used in triggered mode, and the gap between framing intervals are kept as

short as possible [34].

4.3 Applications of the ALPIDE Chip in Calorimeters

In addition to the ALICE ITS Upgrade, the ALPIDE chip will be used for various other applications like FoCal, and the digital tracking calorimeter employed in the Bergen pCT prototype. The next sections will describe how an electromagnetic calorimeter operates, and it will also cover the applications of the ALPIDE chip mentioned above.

4.3.1 Electromagnetic Calorimeters

Detectors in electromagnetic calorimeters utilise the properties of electromagnetic showers. A calorimeter is a type of detector that measures energy by total absorption of particles. The particles can deposit their energy in different forms, and the signal generated in the detector, being light or electric charge, is proportional to the total energy deposited in the detector volume [9].

The proportionality constant(s) are determined as part of the calibration of the detector. The deposited energy, E, is, typically, proportional to the number of interactions, n, that occur in the detector volume:

$$E \propto n \tag{4.1}$$

As a consequence, the width of the deposited energy, σ_{E}, is given as in a Poisson process:

$$\sigma_{\mathrm{E}} \propto \sqrt{n} \tag{4.2}$$

Combining Equations 4.1 and 4.2 gives the following relation:

$$\frac{\sigma_{\mathrm{E}}}{E} \sim \frac{\sqrt{n}}{n} = \frac{1}{\sqrt{n}} \sim \frac{1}{\sqrt{E}} \tag{4.3}$$

Thus, according to Equation 4.3, the calorimeter's energy resolution improves with increasing energy, which is feasible seeing that the number of interactions is higher for larger energy deposits [9].

When it comes to the design and technology of electromagnetic calorimeters, there are two broad categories: homogeneous or sampling. For homogeneous calorimeters, there is a high-density material filling the whole detector volume, that works as both a sensing and

an absorbing medium. The sampling calorimeters are, on the other hand, constructed using alternating layers of active and absorber material [9].

The resolution of a sampling calorimeter is usually worse than for a homogeneous calorimeter. The difference in resolution is a result of variations in the number of particles reaching the active regions, and because the sampling calorimeter only detects part of the energy deposited. Nevertheless, the sampling calorimeter is typically less expensive than its homogeneous peer; in addition to the possibility of it achieving a very compact design [9].

Furthermore, the energy resolution, which at first order is given by Equation 4.3, is more accurately represented by the following expression

$$\frac{\sigma_{\mathrm{E}}}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c \tag{4.4}$$

where the symbol $\oplus$ implies summation in quadrature, and the constants a, b and c take into account the energy fluctuations, electronics noise and shower leakage (where the detector volume does not contain the entire shower), respectively [9].

4.3.2 FoCal

FoCal, a forward calorimeter, is considered in connection with the upgrade of the ALICE detector and would be used to measure isolated photons [29][36]. As part of the FoCal research and development program, a full prototype of a MAPS calorimeter has been designed and constructed. This prototype consisted of 24 layers of MIMOSA chips and absorber plates made of 3.3 mm thick tungsten. It was constructed to validate the concept of a high-granularity calorimeter using CMOS pixel sensors with digital readout. As the MIMOSA chip is not adequate for application in a full detector, due to its readout speed of the rolling shutter, the ALPIDE chip is currently under review for this application. Here, the readout is hit-driven, and it has a speed between 400 Mb/s and 1.2 Gb/s [36].

A small silicon-tungsten (SiW) calorimeter, called EPICAL-2, has been designed and constructed. The calorimeter has 24 layers, each containing two ALPIDE chips and 3 mm tungsten absorber, hence resulting in a sensitive area of 30×30 mm^2. The design of the EPICAL-2 prototype is seen in Figure 4.6. The construction of this prototype allows for testing of the performance of the ALPIDE chip when used in a digital calorimeter application, and consequently provide input to the final design parameters of the FoCal [36].

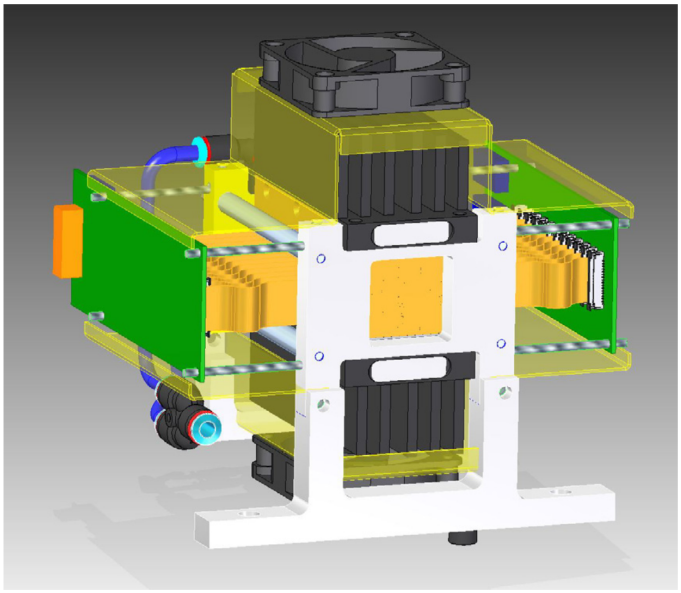


Figure 4.6: Design of the EPICAL-2 prototype. The figure is retrieved from [36].

Compact and thin cabling is used in the prototype to maintain a small Molière radius. In each layer, a flex cable of multi-layered aluminium-polyimide is connected to the chips through aluminium-polyimide chip cables. The plates of tungsten absorber support the chips, as can be seen in Figure 4.7. The ALPIDE chips, cabling and glue, have a combined thickness of less than 0.5 mm [36].

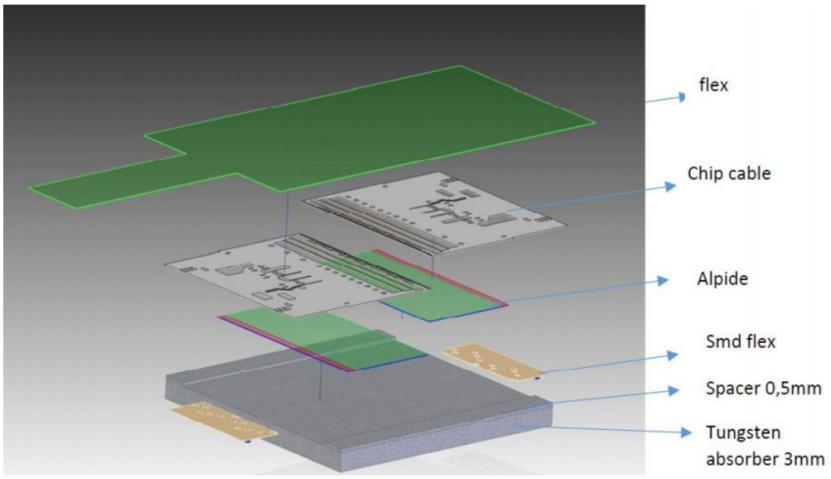


Figure 4.7: Design of a single layer of the EPICAL-2 prototype. The figure is from [36].

4.3.3 Proton CT

With a pixel size of 29.24 μm × 26.88 μm, the high granularity ALPIDE sensors are capable of tracking multiple particles simultaneously. The pixel matrix is inactive as long as there are no hits, since the readout is hit-driven, providing efficient particle detection at a high rate. In addition, the ALPIDE has a detection efficiency of 99.99% and a track position resolution of 5 μm for minimum ionising charged particles, making the ALPIDE chip an ideal candidate for the digital tracking calorimeter [24].

Figure 4.8 displays the general structure of the Bergen pCT scanner, with the tracking layers facing the incoming particles, and 41 layers of aluminium absorbers, excluding the rear trackers, to fully contain the range of a proton beam of 230 MeV. The sensor layers are stacked so that the transition card of alternate layers appear on the opposite side of the main stack. This arrangement is to make room for the readout electronics [24].

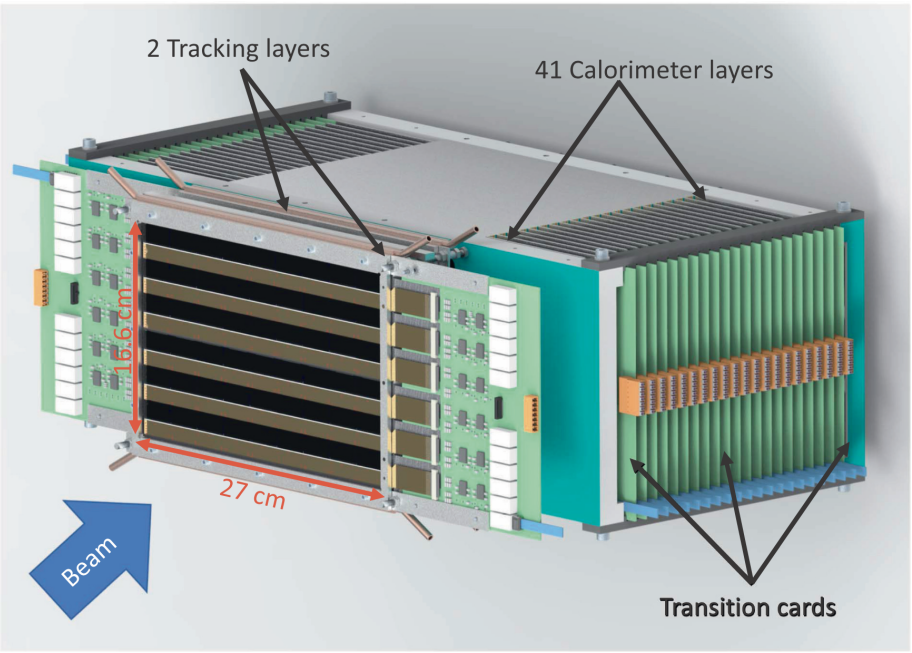


Figure 4.8: A schematic representation of the general structure of the Bergen pCT prototype. The support where the ALPIDE sensors will be mounted in the rear trackers is not displayed to exhibit the arrangement of the sensors in the layer. The figure is retrieved from [24].

Each sensitive layer is based on the same design and building blocks, whether they are used in the calorimeter or as a tracker. The ALPIDE chips are mounted on a flex cable, where a string is a set of nine mounted ALPIDEs. Furthermore, when three strings of this type are glued on an aluminium carrier with dimensions 100 mm × 290 mm × 1 mm, it is referred to as a slab. There are two types of slabs; a top slab and a bottom slab, which combined establish a half layer, as depicted in Figure 4.9A. As seen in the layer structure's side-view in

Figure 4.9B, the two half layers are stacked so that the ALPIDEs are facing each other with a 2mm air-gap between the aluminium carrier boards, assured by an aluminium spacer [24].

As depicted in Figure 4.9B, in the direction of the proton beam, one full layer of the calorimeter consists of: a 1 mm aluminium carrier, a half layer that is not directly facing the proton beam, a 2 mm air-gap ensured by aluminium spacers of 2 mm, another half layer which is facing the proton beam, a 1 mm thick aluminium carrier on which the sensors are mounted, and a 1.5 mm thick aluminium absorber plate. The thickness of the ALPIDE chips used in the calorimeter is 100 μm, and the absorber layer has a thickness of 3.5 mm. The calorimeter will, in total, contain 41 layers of this type [24].

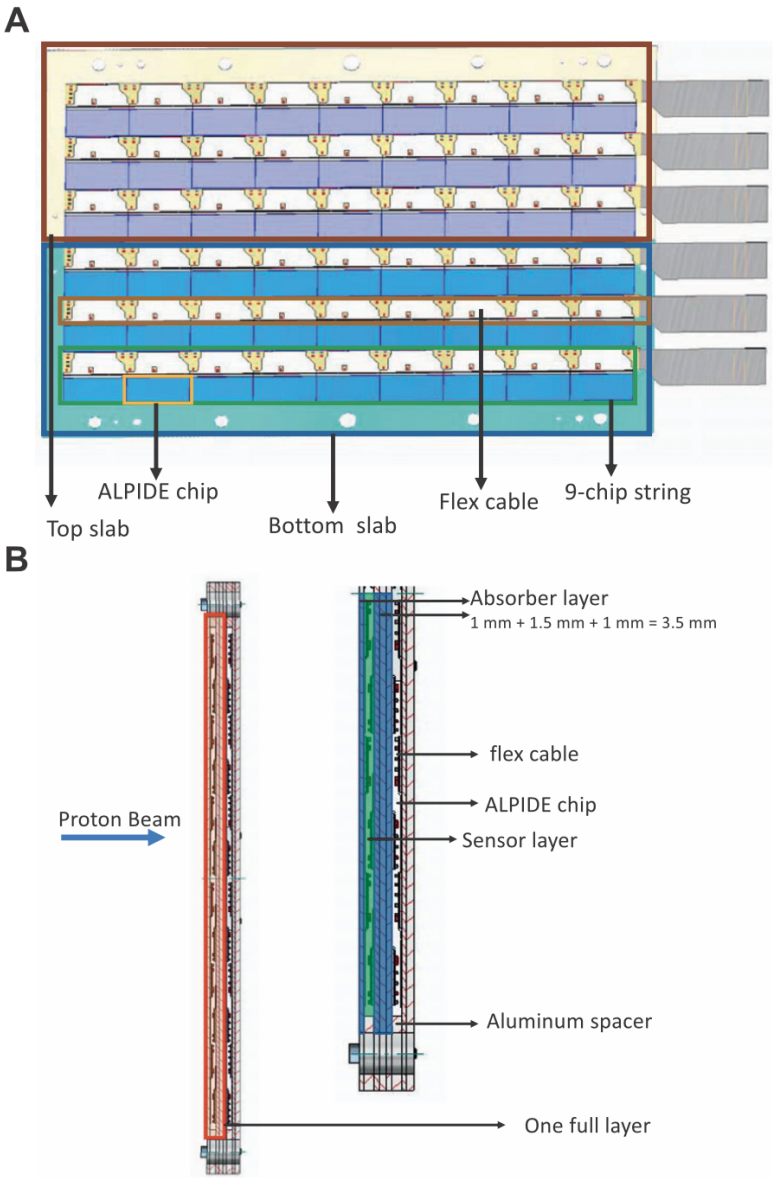


Figure 4.9: (A) A top slab and a bottom slab, together forming a half layer. Each slab consists of three 9-chip strings glued to an aluminium carrier. (B) Side-view of two calorimeter layers (left), and half layer in more detail (right). The figure is from [24].

Similar to the calorimeter layer, in the proton beam direction, one full tracker layer consists of: a 0.2 mm thick carbon-epoxy sandwich sheet with a half layer of sensors not directly facing the proton beam, a 2 mm air-gap, another half layer where the sensors are facing the beam, and another carbon-epoxy sandwich sheet. The half layer has a thickness of 0.2 mm, and ALPIDE chips with a thickness of 50 μm are used to minimise the non-sensor material in the tracking layers. The second layer is placed at a distance of 50 mm from the first layer and is built in the same way [24].

Chapter 5

Characterisation of the EPICAL-2 Prototype

The mechanical design and structure of the EPICAL-2 prototype were both created by Utrecht University. The design of flex cables and ALPIDEs was provided by the University of Bergen together with LTU in Ukraine. Each of the layers in the EPICAL-2, consisting of ALPIDE chips, absorbers and cabling, was constructed at LTU. The readout units are from the ALICE ITS. The setup and data acquisition at DESY were performed by a team from Utrecht University, Birmingham University, University of Bergen, University of Oslo and Johann Wolfgang Goethe University in Frankfurt, where I, personally, took part in the data taking and data analysis, together with this team.

The following chapter describes the experiment conducted at the DESY II Test Beam Facility in February 2020, as well as the results of the data analysis. Data from a similar experiment carried out in November 2019 was also used for analysis.

5.1 Experimental Setup

As mentioned in Section 4.3.2, the construction of a small silicon-tungsten calorimeter prototype allows for testing of the ALPIDE chip's performance when used in a digital calorimeter. This prototype, initially referred to as the mTower, has now been given the name EPICAL-2, an acronym for Electromagnetic Pixel Calorimeter. EPICAL-1 refers to the MIMOSA prototype also mentioned in Section 4.3.2, while EPICAL-2 is the one constructed using the ALPIDE chip.

The purpose of the experiment is thus to test the performance of the EPICAL-2 prototype and consequently provide input to the FoCal's final design parameters with regards to electron-

ics, cabling and readout. The prototype consists of 24 layers, where each layer contains two ALPIDE chips and 3 mm of tungsten absorber. Since each ALPIDE chip has the dimensions 30×15 mm^2, the sensitive area will be 30×30 mm^2.

The stack of 24 layers is kept stable in position by a housing of aluminium, providing options for water and air cooling. Water cooling was used during the test beam measurements to keep the prototype at a stable temperature during operation [37].

Each layer of the EPICAL-2 prototype has a thickness of 3.5 mm and consists of a 3 mm tungsten absorber plate and two 0.5 mm thick tungsten spacers, in addition to the ALPIDE chips, chip cables and flex cables with a total thickness of less than 0.5 mm, as discussed in Section 4.3.2 and seen in Figure 4.7. The EPICAL-2 was placed on a remotely controllable stage.

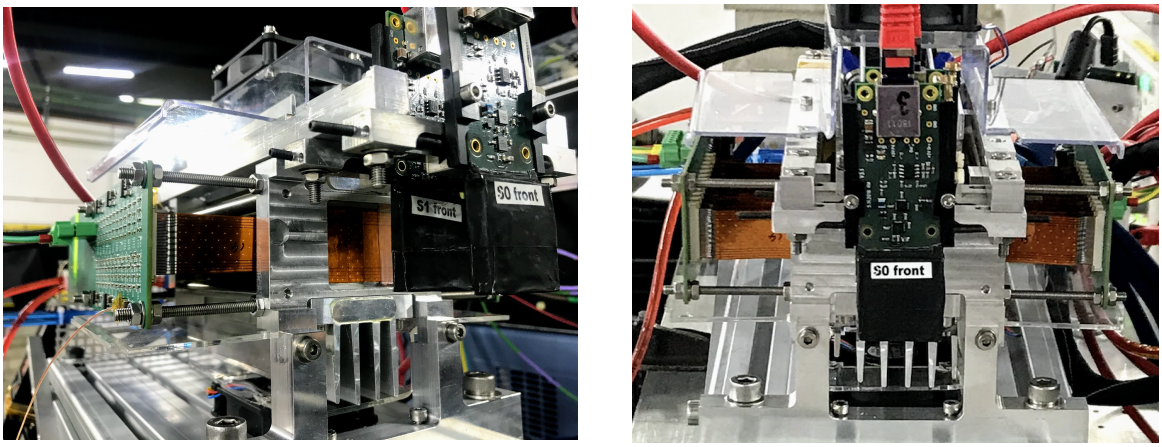
The data acquisition is performed using an adapted version of the python code developed for the ALICE ITS Upgrade. To allow for readout of more than nine ALPIDE data-streams and to enable external triggering, adjustments were also made to the firmware of the readout unit RUv2. The 48 ALPIDE chips are read out by two readout units version 2, using two readout units of version 0 to emulate a common readout unit [37].

The EPICAL-2 is connected to a printed circuit board via 12 Kapton flex cables using ZIF connectors. The printed circuit board sections the 12 modules into three groups, each having two local regulators and a FireFly link connecting the eight 1.2 Gbps high-speed data lines and the shared control and clock signal to the RUv2. The RUv2 is connected to the RUv0 through an optic link. The RUv0 acts as a common readout unit and is connected to the PC by a USB3 connection [37].

The EPICAL-2 is divided into two equal sections, or test benches, TB1 and TB2, where each test bench has 24 ALPIDE chips and one set of readout units; RUv0 and RUv2. When the python script executes the readout procedure for a certain number of cycles, the computer tells both of the RUv0 units to send triggers to the RUv2 units. When the trigger signal arrives at the RUv2, the event readout is initiated. The total number of events is given by the number of cycles times the number of triggers. A BUSY mechanism was included to assure that both of the test benches accepted the triggers synchronously so that the triggers were only accepted if both of the RUv2 units could handle them.

A trigger signalling an electron passing through the EPICAL-2 initialised the readout of an event. The trigger signal was provided by the two scintillation tiles placed in front of the prototype, as seen in Figure 5.1. The scintillator tiles have a thickness of 3 mm and an area of 30×30 mm^2 to match the prototype's sensitive area. The readout of the tiles is performed by a low noise silicon photo-multiplier mounted on a printed circuit board. The tiles are coated with a reflective adhesive foil, before the tiles and the printed circuit board are further

covered in black light-tight tape. This is done to increase the light yield and the uniformity of the response [37].



(a) Side-view of the prototype.

(b) Front-view of the prototype.

Figure 5.1: Pictures of the experimental setup, where (a) displays the EPICAL-2 from the side and (b) from the front.

5.1.1 DESY II Test Beam Facility

Deutsches Elektronen-Synchrotron, DESY, operates a test beam facility that offers three independent beam-lines attached to the DESY II synchrotron. The beam-lines offer beams of electrons or positrons with selectable momenta between 1 and 6 GeV/c. This Test Beam Facility has an infrastructure that is essential for the testing and development of detectors used in particle and nuclear physics. It is also one of the few facilities worldwide offering the users access to beams of multi-GeV [38].

Rather than using a direct extraction of the primary beam in DESY II, the test beams at the DESY II Test Beam Facility are generated by double conversion. A fibre target located in the DESY II beam orbit produces bremsstrahlung photons. These photons then hit a secondary target producing pairs of electrons and positrons. A dipole magnet then selects particles with a certain charge and momentum by changing the strength and polarity of its magnetic field accordingly. The three independent beam-lines of the facility, TB21, TB22 and TB24, are named based on the location of the primary targets positioned behind the quadrupoles QF21, QF22 and QF24, respectively. Figure 5.2 displays a schematic view of the beam production [38].

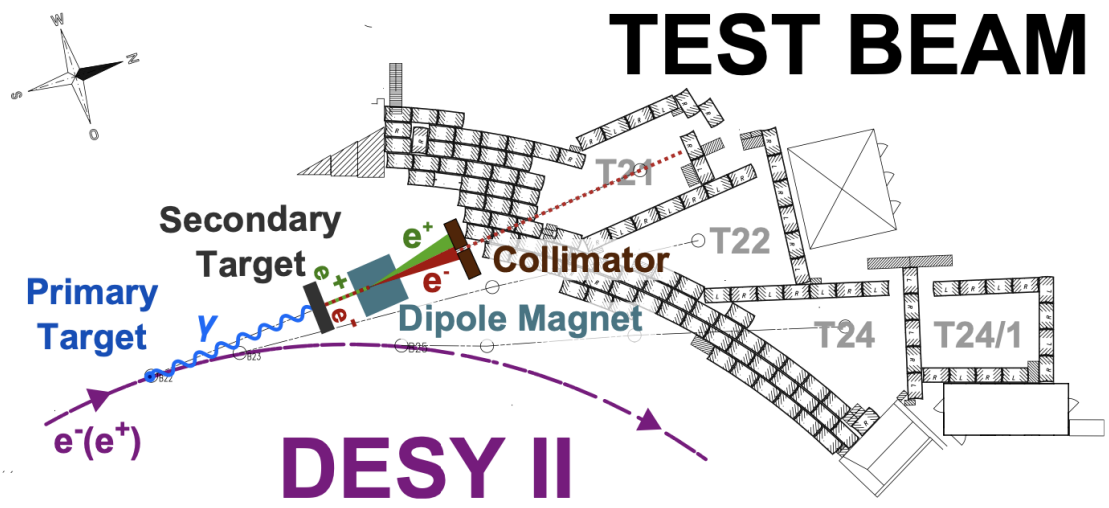


Figure 5.2: Diagram showing a schematic view of how the test beam is generated at the DESY II Test Beam Facility, using beam-line TB21 as an example. In the primary target, bremsstrahlung photons are generated before travelling through the tunnel and hitting the secondary target. Electron-positron-pairs are generated at the secondary target before a dipole magnet selects particles according to charge and momentum. Further collimation can be applied to the particle beam before it enters the test beam area. The figure is taken from [38].

5.1.2 Test Beam Experimental Setup and Data Acquisition

The experimental setup described above was installed in the test beam area, and beam-line TB22 was used. The pixel threshold settings of the ALPIDE chips were also adjusted. The default setting used the same settings for all the ALPIDE chips, with VCASN = 50 and ITHR = 51. The optimised settings, on the other hand, had adjusted settings for 9 ALPIDE chips. The layers, chip IDs and adjusted settings for VCASN and ITHR are listed in Table 5.1.

Table 5.1: Overview of the ALPIDE chips with adjusted VCASN and ITHR settings.

Layer	Chip ID	VCASN [DAC]	ITHR [DAC]
0	20	48	50
2	23	48	50
4	19	48	50
6	17	52	44
8	14	52	44
10	13	52	44
12	11	52	44
14	8	46	60
14	9	46	50

Two collimators can be applied to the beam in the beam-line, where the primary collimator

is located behind the magnet defining the momentum of the beam. A secondary collimator is placed right after the end of the beam pipe and is used to collimate the beam further [38]. In Table 5.2, the primary collimator is referred to as the internal collimator, while the secondary collimator is called the lead collimator.

Data was taken using different beam energies, collimator settings, ALPIDE threshold settings, stage positions, temperatures and strobe lengths. Data was taken at different angles of the EPICAL-2 with respect to the beam axis, as well as in the reverse direction. There was also data collected where positrons were utilised. The ALPIDE threshold settings were either set to default or optimised, as described above, and the temperature was regulated using water-cooling. Pedestal runs were also conducted, which are data runs taken without a beam. These runs are used for noise monitoring, where pixels that fired in more than 90% or 98% of the readout frames are characterised as noisy pixels and masked before the data acquisition.

The data used for analysis from this test beam experiment is listed in Table 5.2 and 5.3. Table 5.2 gives an overview of the beam data, while Table 5.3 lists information about the pedestal runs. The data runs listed in both tables all have optimised pixel threshold settings, and the beam runs all had the same stage position, $(0.3, -448.4)$. Furthermore, Chip ID 27 in Layer 21 was also excluded for most of the data runs as it caused the readout to fail.

An overview of the lane numbers and chip IDs corresponding to the different layers of the EPICAL-2 prototype can be found in Table A.1.

The first six runs listed in Table 5.2 are the runs used in the energy studies, with energies ranging from 1 GeV to 5.8 GeV. The runs all have the same collimator settings, water temperature and strobe length. The next three runs all have an energy of 3 GeV, the same collimator settings and strobe length, but here the temperatures are set to 20°C, 25°C and 30°C, respectively. The last two runs also have an energy of 3 GeV, the same collimator settings and water temperature, but the strobe lengths are 2 μs and 20 μs respectively. The run shared between these three sets is Run 1335, with an energy of 3 GeV, a temperature of 20°C and a strobe length of 2 μs. This run is the basis for most of the data analysis.

Table 5.2: Run condition table for electron beam runs at the DESY Test Beam in February 2020. The table gives an overview of beam energy, collimator settings, water temperature, number of events, beam rate and strobe length for the respective run numbers. Chip ID 27 is excluded in all of the listed data runs. The beam rate values marked by a star is based on the average beam rate for similar runs, as the beam rate was not available for these particular runs.

Run Number	**Beam Energy** [**GeV**]	**Lead Collimator** [$\mathbf{mm^2}$]	**Internal Collimator** [$\mathbf{mm^2}$]	**Water Temperature** [**°C**]	**Number of Recorded Events**	**Beam Rate** [**kHz**]	**Strobe Length** [$\boldsymbol{\mu}$**s**]
1336	**1**	12×12	14×14	20	594000	4.8	2
1337	**2**	12×12	14×14	20	594000	4.5*	2
1335	**3**	12×12	14×14	20	297000	7.5	2
1338	**4**	12×12	14×14	20	297000	2.7*	2
1339	**5**	12×12	14×14	20	747000	1.5*	2
1346	**5.8**	12×12	14×14	20	617749	0.12 – 0.25	2
1335	3	12×12	14×14	**20**	297000	7.5	2
1305	3	12×12	14×14	**25**	285003	5.9*	2
1358	3	12×12	14×14	**30**	522000	5.0 – 8.7	2
1335	3	12×12	14×14	20	297000	7.5	**2**
1324	3	12×12	14×14	20	278837	7	**20**

The pedestal runs listed in Table 5.3 are sorted based on water temperature and strobe length. Run 1332, 1303 and 1356 have temperatures of 20°C, 25°C and 30°C, respectively, while Run 1330 has a strobe length of 20 μs.

Table 5.3: Run condition table for pedestal runs at the DESY Test Beam in February 2020. The table gives an overview of water temperature, number of events and strobe length for the respective run numbers. Chip ID 27 is excluded in all of the listed data runs.

Run Number	**Water Temperature** [**°C**]	**Number of Recorded Events**	**Strobe Length** [$\boldsymbol{\mu}$**s**]
1246	20.1	46000	2
1332	20	96000	2
1303	25	46000	2
1356	30	92000	2
1330	20	8000	20

A similar setup to the one described above was used to gather data in November, but this setup only had 12 layers. The data runs analysed in this thesis are listed in Table 5.4. Default threshold settings were applied for all the data runs listed below, with VCASN = 50 and ITHR = 51. The stage position (188.2, 737.2) was also identical for all the runs. These beam runs were then used to study the strobe length dependence of the number of hits per event.

Table 5.4: Run condition table for electron beam runs at the DESY Test Beam in November 2019. The table gives an overview of beam energy, collimator settings, number of events and strobe length for the respective run numbers.

Run Number	**Beam Energy** [**GeV**]	**Lead Collimator** [$\mathbf{mm^2}$]	**Internal Collimator** [$\mathbf{mm^2}$]	**Number of Recorded Events**	**Strobe Length** [$\boldsymbol{\mu}$**s**]
423	5	6×6	14×14	104201	**0.1**
422	5	6×6	14×14	104400	**0.2**
421	5	6×6	14×14	74999	**0.5**
415	5	6×6	14×14	105000	**1**
414	5	6×6	14×14	104476	**2**
413	5	6×6	14×14	100881	**5**
301	5	6×6	14×14	247732	**10**
412	5	6×6	14×14	99277	**20**

The data used in the analysis is not corrected for internal misalignment, static misalignment or divergence of the beam, hence assuming that the beam hits perpendicular to the first layer. Internal alignment refers to the layer-by-layer alignment of the chips, which can be performed using cosmic data. In contrast, the static alignment refers to the alignment of the calorimeter with respect to the beam axis.

5.2 Data Analysis

In this section, the algorithms used in the data analysis are described. Starting with the clustering algorithm, and the definition of a cluster, before the method used for event selection is explained. The data analysis software was written in C++ using ROOT, a data analysis framework developed by CERN.

5.2.1 Clustering Algorithm

Particles traversing the sensitive volume of the ALPIDE will register hits in the sensor if their energy is above a set threshold. A cluster is defined as a collection of adjacent pixels that

register hits within the same time frame.

A clustering algorithm was used to identify the clusters within a readout frame. For each readout frame, this algorithm loops over the pixel matrix with 1024 columns and 512 rows until it locates a hit. After a hit is registered, the algorithm locates the nearest neighbour of that hit. If the nearest neighbour is adjacent to the hit, it will be counted as part of the cluster, increasing the cluster size by one. However, if the nearest neighbour is not an adjacent pixel, the pixel size is stored, and it moves on to estimate the size of the next cluster. The same applies if there are no more hits in the selected readout frame. This procedure is repeated for all the 48 ALPIDE chips in the 24-layered prototype.

5.2.2 Event Selection

The event selection is primarily based on the number of tracks present in the first layer, Layer 0, of the EPICAL-2. The identification and counting of clusters are based on the clustering algorithm described above. An event is here defined as a one readout frame from the entire stack of sensors.

The event selection algorithm runs through all the data and stores the event numbers in vectors based on the number of tracks present in the first layer. In addition to the event selection based on the number of tracks in the first layer, an event selection for zero tracks in the first layer was also included. This event selection was added, because certain events had shower development, despite no tracks being generated in the first layer. Hence, a particle can hit the first layer without generating a track, as can be seen in Figure 5.42 and Figure 5.43, where there is shower development in the later layers for this event selection. An event selection merely based on the number of tracks in the first layer will, therefore, exclude several showers. Lastly, the analyses were also performed using all the events of the data runs, and these are displayed in the graphs labelled "No Selection" in the various figures.

Position of Incoming Particles

In addition to event selection based on the number of tracks in Layer 0, cuts were made concerning the position of incoming particles. Here all the hits outside a 15×15 mm^2 area around the centre of the layer were disregarded, thus, only keeping hits with column coordinate values within the interval $[256, 768]$ and row coordinates less than or equal to 256. This filter was applied to ensure shower containment by selecting events based on the initial position of the shower. This event selection is indicated by the red square in Figure 5.3.

5.3 Results

In the following section, the results of the data analysis are presented. It covers the beam profile, as well as cluster studies, noise and shower studies.

The uncertainty calculations are based on the standard deviation of the quantities, and by assuming that the measurements are a Poisson process, the variance, σ^2, will equal the mean value, $\bar{x}$. The values are thus presented as $\bar{x} \pm \sqrt{\bar{x}}$, where the uncertainty is displayed using error-bars.

5.3.1 Beam profile

Figures 5.3, 5.4 and 5.5 all display the spatial distribution of hits for Run 1335 with an energy of 3 GeV, a temperature of 20°C and strobe length of 2 μs, as listed in Table 5.2. Figure 5.3 shows the spatial distribution for the first layer integrated over all the readout frames, while in Figure 5.4, the spatial distribution is integrated over all the layers as well. Moreover, Figure 5.5 depicts the integrated spatial distribution for all layers of the EPICAL-2, but displaying the distribution of each layer separately.

The coordinate system used to display the spatial distributions is based on Table A.1 in Appendix A, where the first two columns of the table indicate the lane numbers and chip IDs used for the upper half of the coordinate system, having positive row values. The other two table columns contain the lane numbers and chip IDs used in the lower half, with negative row values.

As seen in all three figures, there appear to be a feature or hot spot present. In Appendix C, spatial distributions for the different layers integrated over all the events of a run are shown for the beam runs listed in Table C.1. Here it seems that this feature is present for different energies, temperatures and strobe lengths. The hot spot appears to be in approximately the same position even when the stage position is changed, thus, causing the beam axis to be located in a different part of the prototype, as seen in Figure C.6. When the EPICAL-2 is fixed at an angle of 10° with respect to the beam axis, the position of the feature is altered slightly in the x-direction, as shown in Figure C.7.

The spot is also present in several layers and is most prominent where there are a maximum of recorded hits. The feature can be seen to increase with an increasing number of hits, which is also shown in the reversed run depicted in Figure C.8. For the reversed run the prototype was placed with the back-end facing the beam. In the regular runs, the feature is most noticeable in the first layers, especially around the shower maximum. At the same time, it is most prominent in the rearmost layers for the reversed run.

The feature observed in the beam profiles appears to be independent of the beam energy, layer number, strobe length and beam position. The position of the feature also changes slightly when the beam approaches from a different angle. Since the structure is present in all layers, the chance that the mechanical setup could be the cause, for instance by a lump of glue pressing on a sensor, is low. The feature is, thus, probably caused by an external factor, such as the trigger scintillators, which would explain why the spot's position changes according to the angle of the incoming beam.

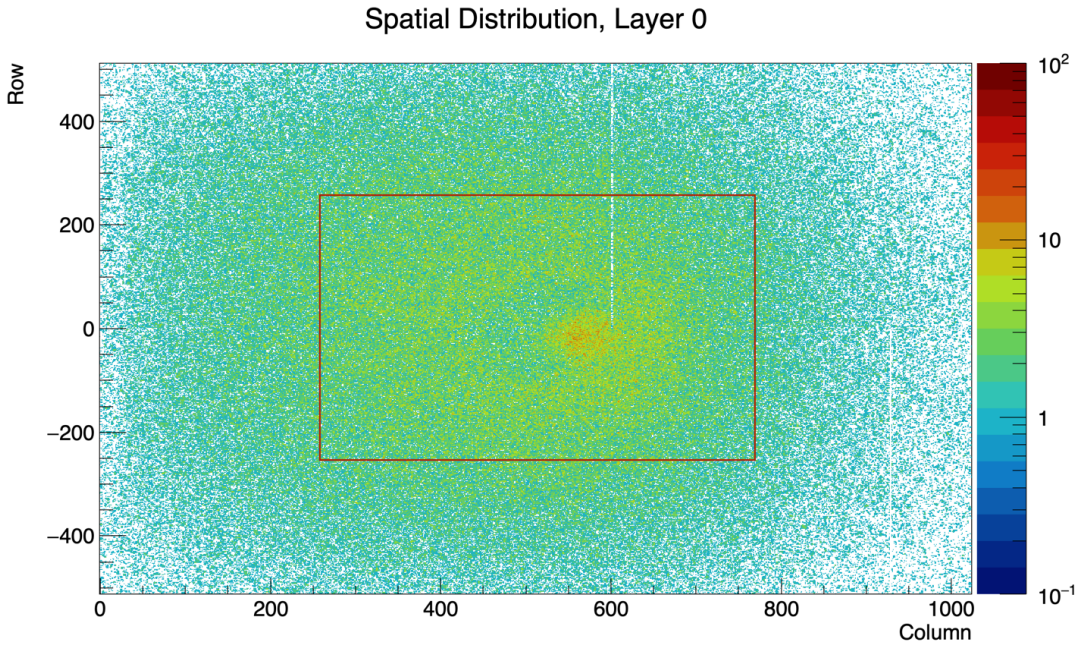


Figure 5.3: An integrated spatial distribution for the first layer, Layer 0. Plotted with a logarithmic scale on the z-axis. The red square indicates the event selection that only includes the hits within a 15 × 15 mm^2 area around the centre of the layer. **Run 1335**: energy of 3 GeV, temperature of 20°C and strobe length of 2 μs.

A pixel is referred to as a dead pixel if it does not give a response when hit by an incoming particle or when directly injected with a test signal. Furthermore, malfunctioning pixels are prone to cause malfunction of whole double-columns of a chip, as a consequence of the setup of the internal readout electronics. The occurrence of dead double-columns is the reason why there is a white line from row number 0 to row number 512 for a double-column nearby column number 600 in Figure 5.3. The double-column appears to be functional for a slight few of the total amount of events, since a few hits can be observed, but they are very sparse compared to the rest of the chip.

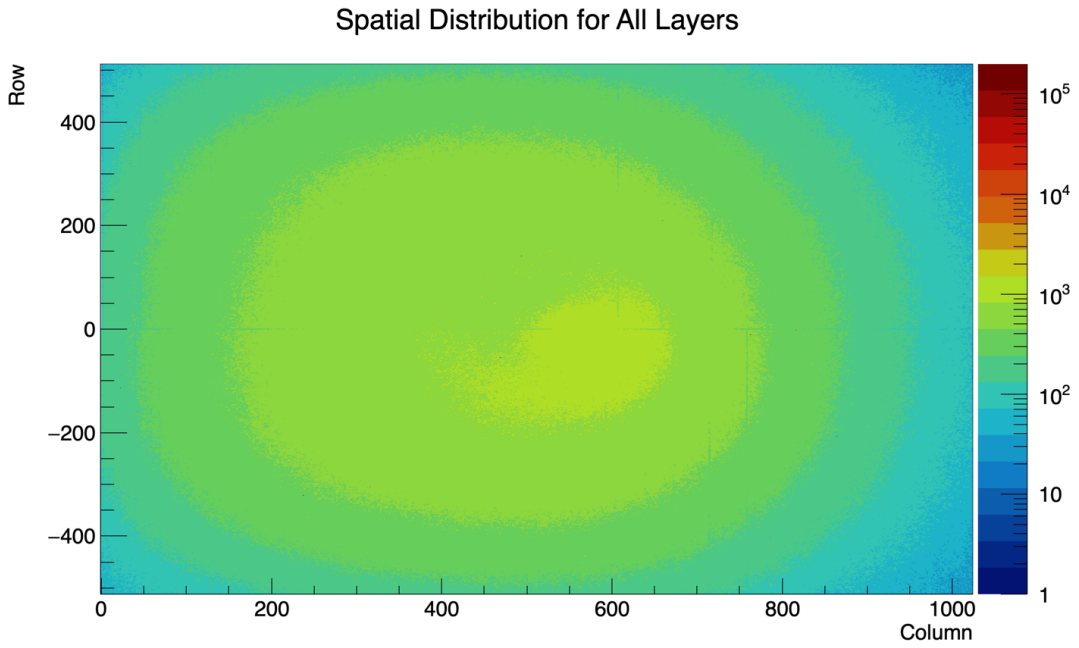


Figure 5.4: An integrated spatial distribution for all layers. Plotted with a logarithmic scale on the z-axis. **Run 1335**: energy of 3 GeV, temperature of 20°C and strobe length of 2 μs.

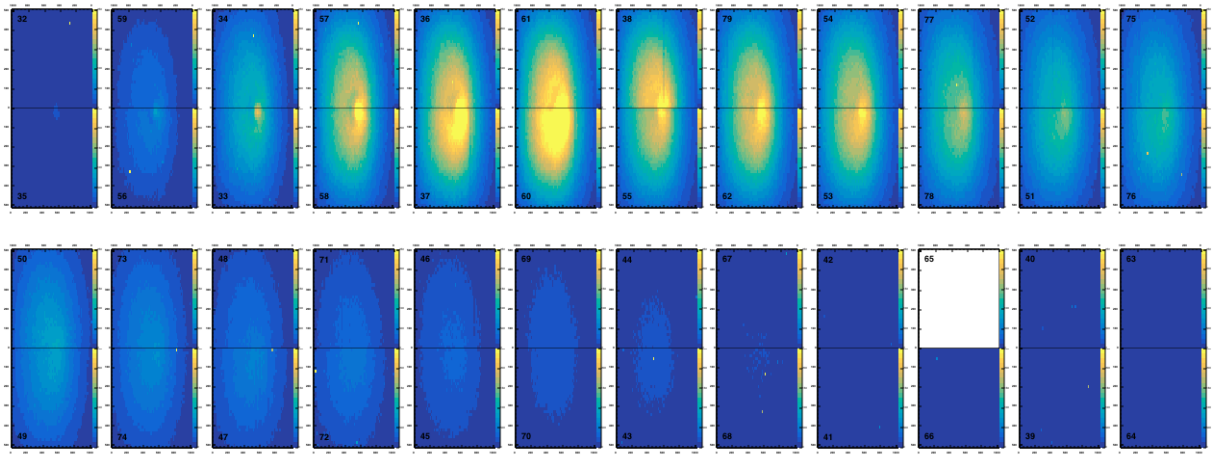
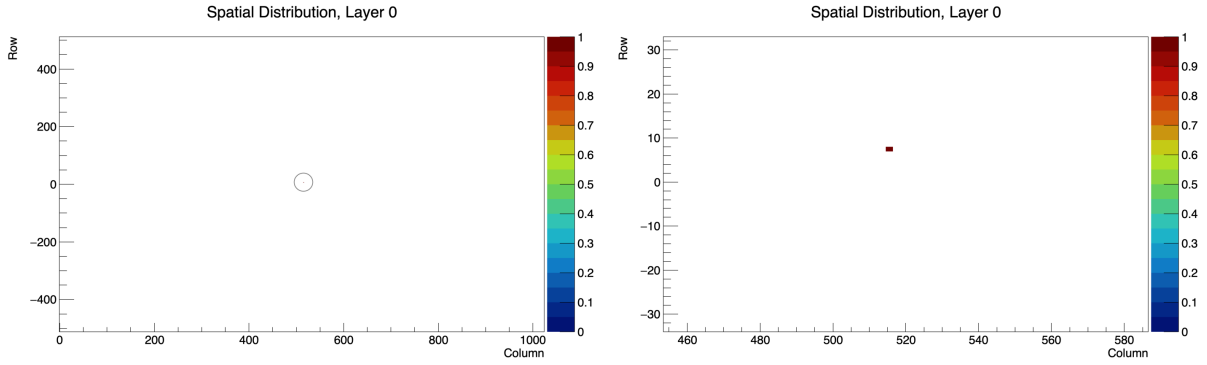


Figure 5.5: Integrated spatial distributions, displaying the distribution for each layer separately. The first 12 layers are located in the upper part of the plot, while the 12 last layers are in the bottom. Chip ID 27 in Layer 21 has no registered hits as it was excluded from the data taking. **Run 1335**: energy of 3 GeV, temperature of 20°C and strobe length of 2 μs.

The spatial distribution for Event 31147 in Run 1335, a readout frame with a single track, is depicted in Figures 5.6 and 5.7. Figure 5.6 shows the position of the single cluster in the first layer, while Figure 5.7 displays the spatial distribution for the shower, integrated over all layers.



(a) Circled hit position.

(b) Zoomed in on hit position.

Figure 5.6: Spatial distribution in Layer 0 for Event 31147. **Run 1335**: energy of 3 GeV, temperature of 20°C and strobe length of 2 μs.

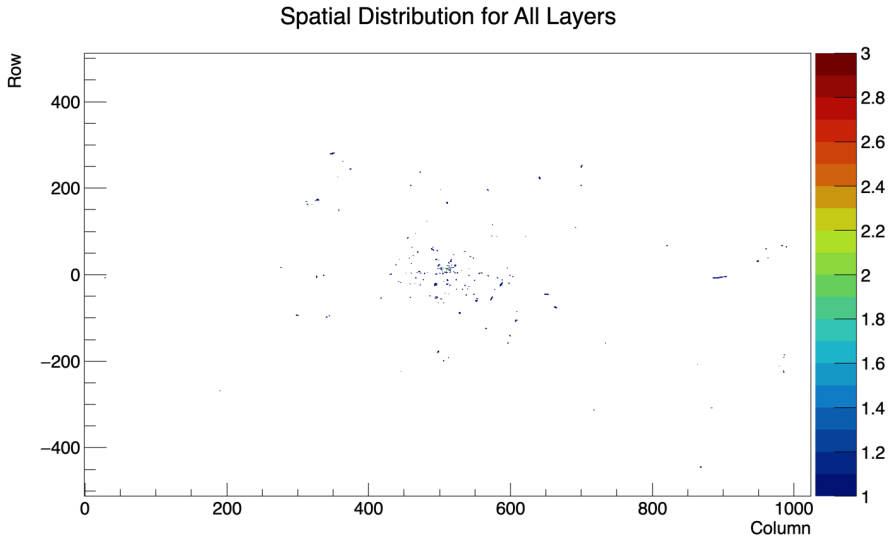


Figure 5.7: Spatial distribution integrated over all layers for Event 31147. **Run 1335**: energy of 3 GeV, temperature of 20°C and strobe length of 2 μs.

The spatial distribution for Event 4168 in Run 1335 is displayed in Figures 5.8 and 5.9. This readout frame has three clusters in the first layer, resulting in the development of three showers, as shown in Figure 5.9, where the spatial distribution is integrated over all layers.

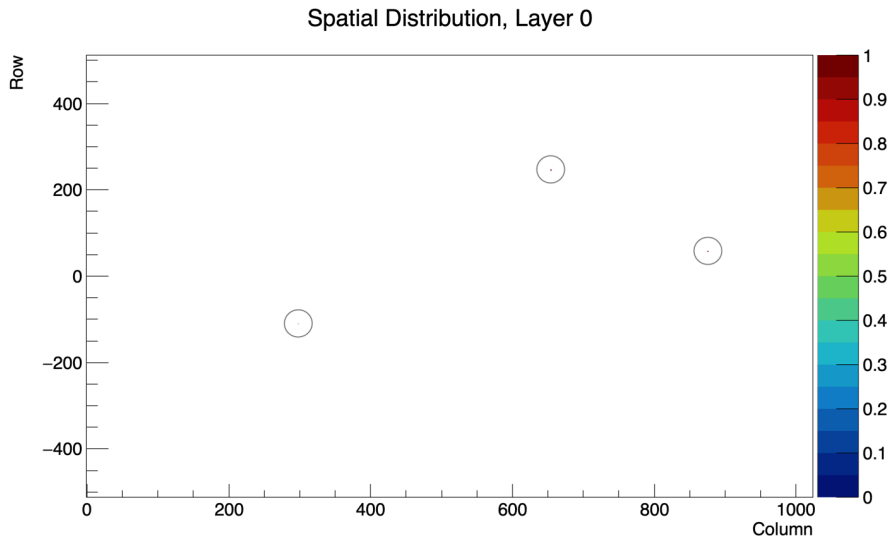


Figure 5.8: Spatial distribution in Layer 0 for Event 4168, where the hit positions are circled. **Run 1335**: energy of 3 GeV, temperature of 20°C and strobe length of 2 μs.

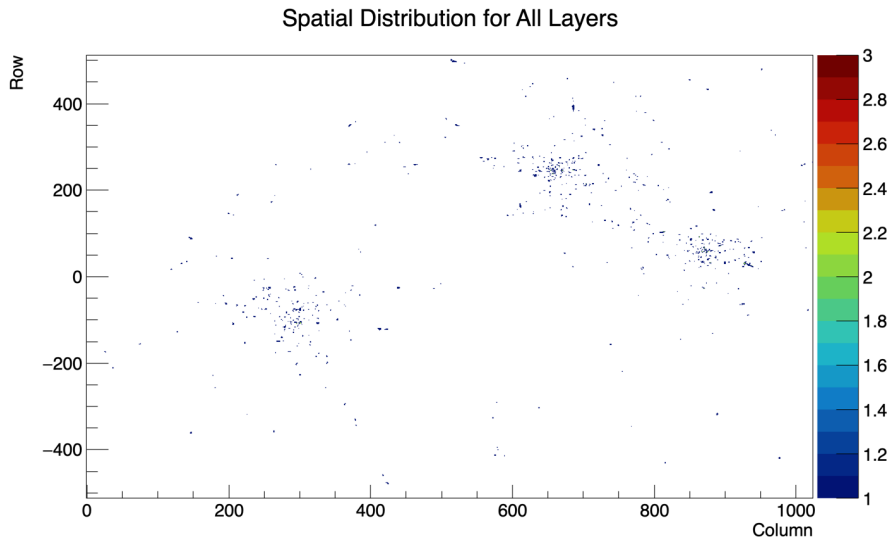


Figure 5.9: Spatial distribution integrated over all layers for Event 4168. **Run 1335**: energy of 3 GeV, temperature of 20°C and strobe length of 2 μs.

Moreover, Figures 5.10 and 5.11 show the spatial distribution for Event 34495. This event has five clusters in the first layer, as seen in Figure 5.10, where the hit positions are circled. However, the spatial distribution for all the layers, shown in Figure 5.11, displays a shower not indicated by a cluster in the first layer.

The showers developing behind the three clusters located between column number 600 and 700, and around row number 200, are not easily separated in Figure 5.11. However, the size of the shower in this area indicates that it is a result of multiple tracks. The two showers developing behind clusters in the first layer, located in the lower half of the coordinate system

close to row zero, are, on the other hand, easier to distinguish. Finally, the shower marked with a red circle, which has not generated a cluster in the first layer, shows that a particle can hit the first layer without generating a cluster, as mentioned in Section 5.2.2.

The last shower may also be caused by a photon, more specifically bremsstrahlung originating from the scintillator. If this is the case, there might be some percentage of showers with unknown energy, since the energy can be extrapolated by adding up the showers.

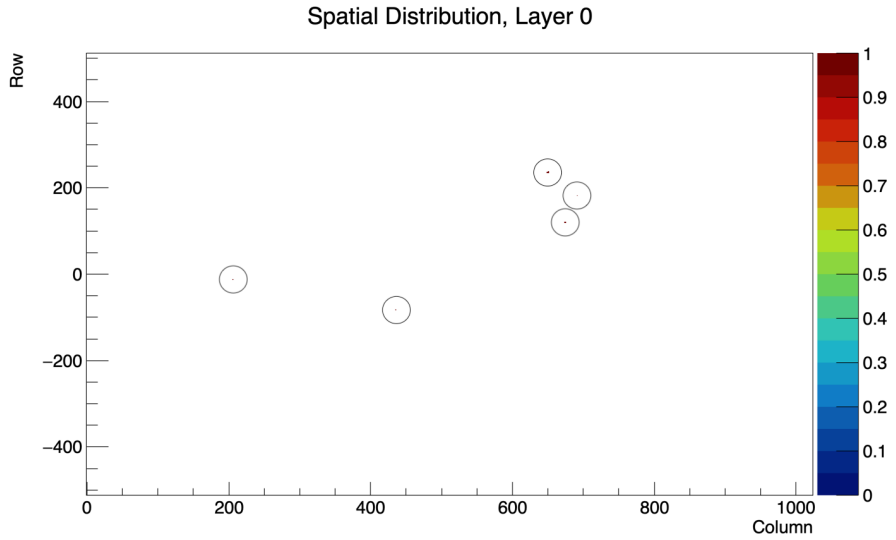


Figure 5.10: Spatial distribution in Layer 0 for Event 34495, where the hit positions are circled. **Run 1335**: energy of 3 GeV, temperature of 20°C and strobe length of 2 μs.

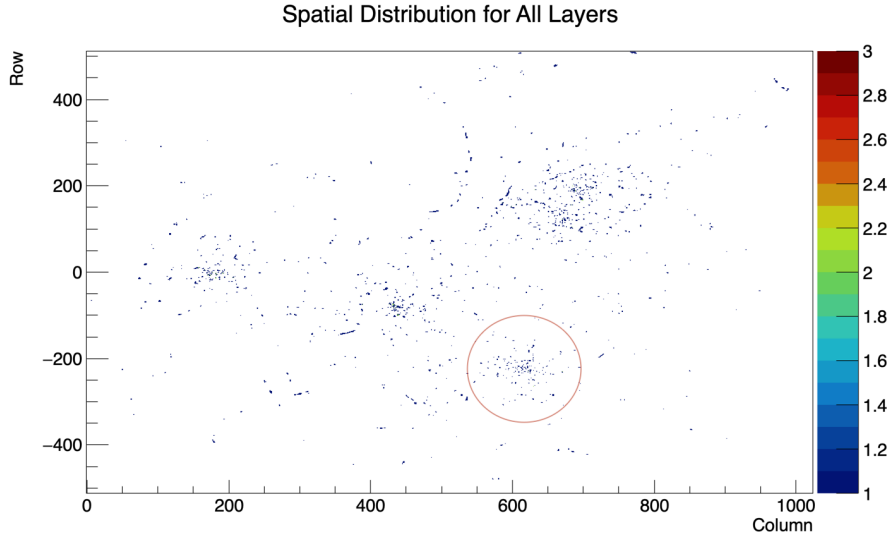


Figure 5.11: Spatial distribution integrated over all layers for Event 34495. **Run 1335**: energy of 3 GeV, temperature of 20°C and strobe length of 2 μs.

The plot in Figure 5.12 displays the number of entries for each column number in Layer 0 for Run 1335, but with a noisy pixel located in Lane 32, with the coordinate (273, 445) filtered out.

This plot displays the starting point of the showers in the front of the EPICAL-2 prototype. The distribution is roughly Gaussian, with a peak close to the centre of the layer. However, the distribution has a larger number of entries on the left side, also indicated by the spatial distribution in Figure 5.3.

As seen in the spatial distributions, such as the one for Layer 0 depicted in Figure 5.3, there is a feature present. There is a dip approximately at the centre of the chip, around column number 512. This dip results in two peaks on each side of the centre. The largest peak is located around column number 600, as indicated by the red points displayed in the spatial distribution of the first layer. There is also a crescent-shaped area with fewer recorded hits located around the centre of the layer, slightly to the left of the area with the most hits. The dead double-column present in Figure 5.3, located in the vicinity of column number 600 can also be seen in this distribution.

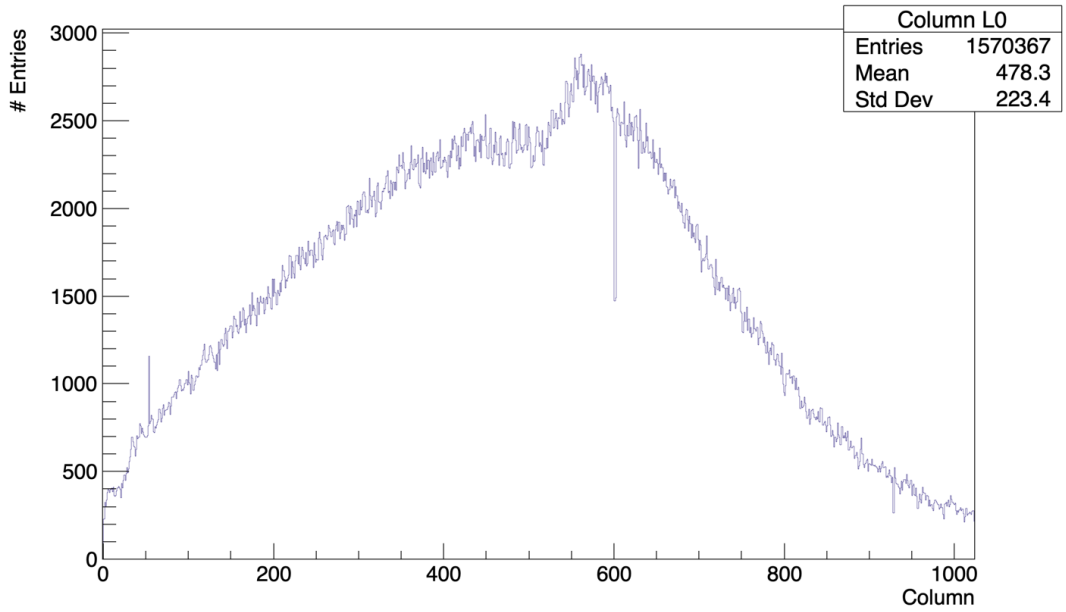


Figure 5.12: Distribution showing the number of entries for each column number in Layer 0. Noisy pixel (273,445) is filtered out. **Run 1335**: energy of 3 GeV, temperature of 20°C and strobe length of 2 μs.

Figure 5.13 displays the number of entries for the different column numbers, where the entries are integrated over all events and all layers. The shape of the distribution resembles a Gaussian distribution. Here the dip around the central column is not as noticeable as in Layer 0. There is, however, still a slight peak around column number 600, as indicated by the area containing the most hits in Figure 5.4, also located around column 600.

The plot in Figure 5.13 also shows that there are fewer hits on the edge of the sensor, around column 1023. This edge effect might be caused by the n-well guard ring that surrounds the

pixel matrix of the ALPIDE. This guard ring is n-doped, which means that it might compete with the n-doped collection electrodes—thus leading to part of the charge being collected by the n-well ring, and not by the collection electrodes that are located at the edges of the pixel matrix [39].

The distribution in Figure 5.13 also shows that there are many hits outside the central part of the layer, providing an argument for using the event selection disregarding events where the hit in Layer 0 is outside the 15×15 mm^2 area around the centre.

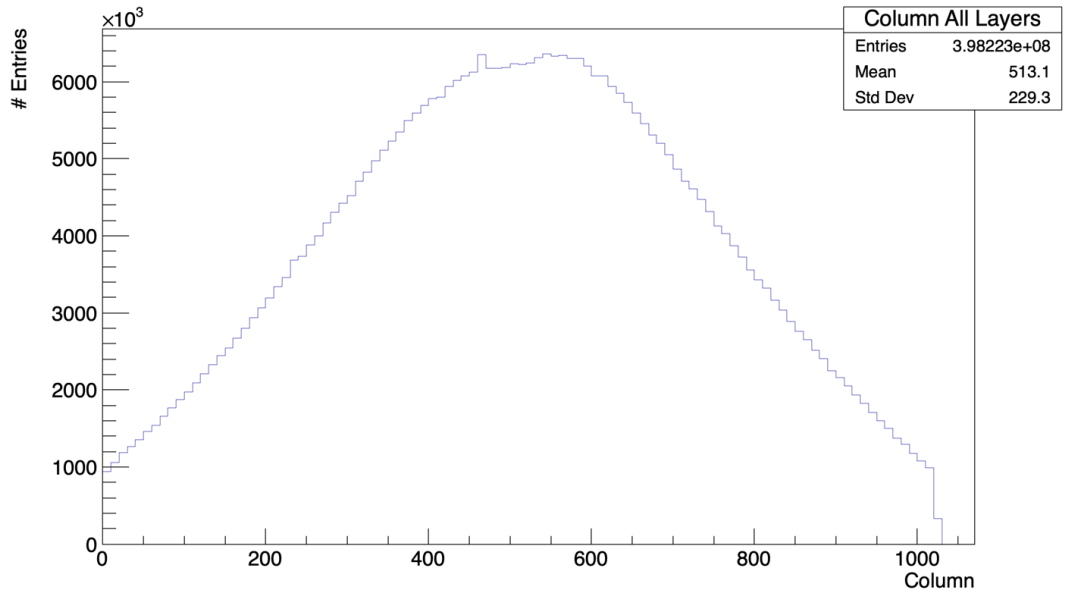


Figure 5.13: Distribution showing the number of entries for each column number integrated over all layers. **Run 1335**: energy of 3 GeV, temperature of 20°C and strobe length of 2 μs.

The number of entries for the different number of hits integrated over all events and layers are shown in Figure 5.14. The first peak, located at about 800 number of hits, indicates the number of events with a single track. The second peak, at around 1600 number of hits, indicates the number of events with two tracks occurring within a readout frame. The third peak is not as prominent as the previous two and is located at about 2400 number of hits, while the barely noticeable fourth peak is at approximately 3200 number of hits. These last minor peaks represent three and four tracks going through the prototype within a readout frame, respectively. The interpretation of this plot is based on the fact that the peaks are equally spaced, indicating that each peak represents a given number of incoming particles.

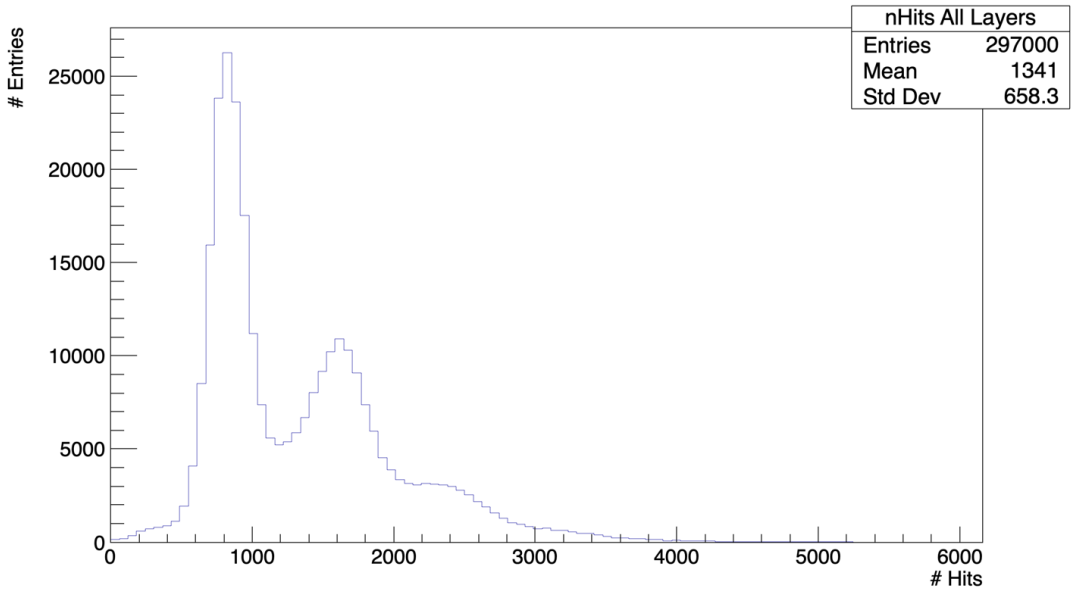


Figure 5.14: Distribution showing the number of entries for the different number of hits, integrated over all layers. **Run 1335**: energy of 3 GeV, temperature of 20°C and strobe length of 2 μs.

5.3.2 Cluster Studies

Cluster Size

By using the clustering algorithm described above, the cluster size distributions for the different layers are depicted in Figure 5.15 for Run 1335. The widest distributions displayed in shades of blue are the distributions for the layers in and around the shower maximum. The narrower distributions in shades of red are for the rearmost layers. Underneath the red distribution is the cluster size distribution for Layer 0, plotted in dark blue.

Figure B.1 in Appendix B displays more detailed cluster size distributions for selected layers, where the plot is zoomed in on the first ten bins.

Figure 5.16 displays the same distributions, but for Layer 0, Layers 5-7, Layer 18 and Layer 22, normalised against Layer 2. Although the shapes of the distributions appear different for the different layers based on Figure 5.15, they look quite similar in Figure 5.16, when the distributions are normalised against Layer 2. Seeing that electrons are MIP particles, one would expect the distributions to have approximately the same shape in all layers, as indicated in Figure 5.16. Layer 0 deviates from the other distributions, both in the number of clusters and the cluster size, as expected since the incoming particle has yet to develop a shower.

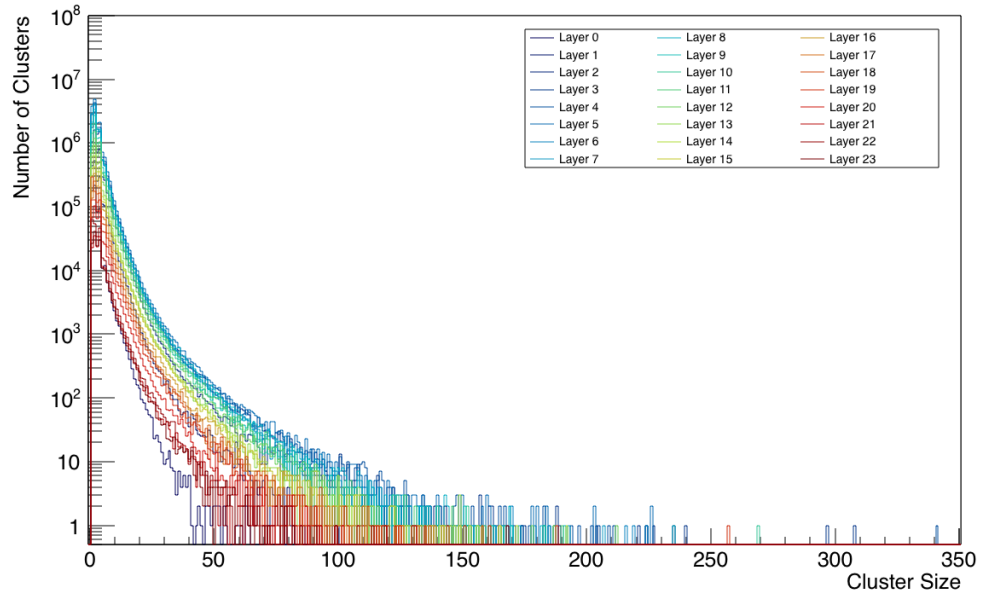


Figure 5.15: Cluster size distribution for the different layers, showing the number of clusters with different cluster sizes. Plotted with a logarithmic scale on the y-axis. **Run 1335**: energy of 3 GeV, temperature of 20°C and strobe length of 2 μs.

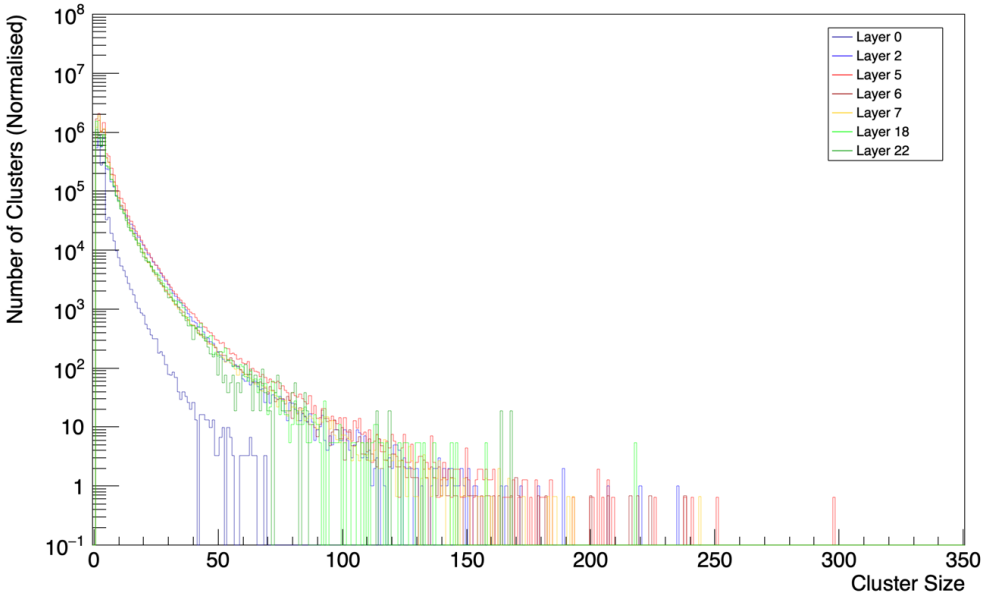


Figure 5.16: Cluster size distribution for Layer 0, Layers 5-7, Layer 18 and Layer 22, normalised against the cluster size distribution of Layer 2, plotted with a logarithmic scale on the y-axis. **Run 1335**: energy of 3 GeV, temperature of 20°C and strobe length of 2 μs.

The broadening of the distribution might be caused by the angle of the shower as it develops. Large angles can cause very large clusters, as seen in Figure 5.17. This plot displays the spatial distribution for events containing a cluster with a minimum cluster size of 200 pixels.

If overlapping clusters result in larger cluster sizes, this might occur in the cluster size range between four and eight pixels in Figure 5.15. However, by looking at the normalised distributions in Figure 5.16, the difference between Layer 2 and the layers in shower maximum for the cluster sizes in this range is minimal. Hence, leading to the assumption that the effect of overlapping clusters is small.

By looking at Figures 5.18 through 5.22, the larger clusters appear to have a track-like nature. These figures display the spatial distributions for clusters in Layer 0 with sizes in the ranges 10-20, 20-30, 30-40, 40-50 and 50-60, respectively. To make it more clear only the clusters within these ranges were included. Layer 0 was also chosen as there are fewer large clusters in this layer compared to later layers. Later layers were often too crowded to study the nature of the clusters, especially for the smaller cluster sizes.

The question is then, why is this effect seen in Layer 0? The incoming particle has yet to scatter on anything; hence there should be no track-like structures. This effect could be explained by backscattering from the tungsten absorber behind the first layer, or it could be caused by δ-electrons.

The spatial distributions shown below indicate that the larger cluster sizes are tracks caused by effects such as Compton scattering or particles moving in the epitaxial layer of the sensor, and not an effect of overlapping clusters.

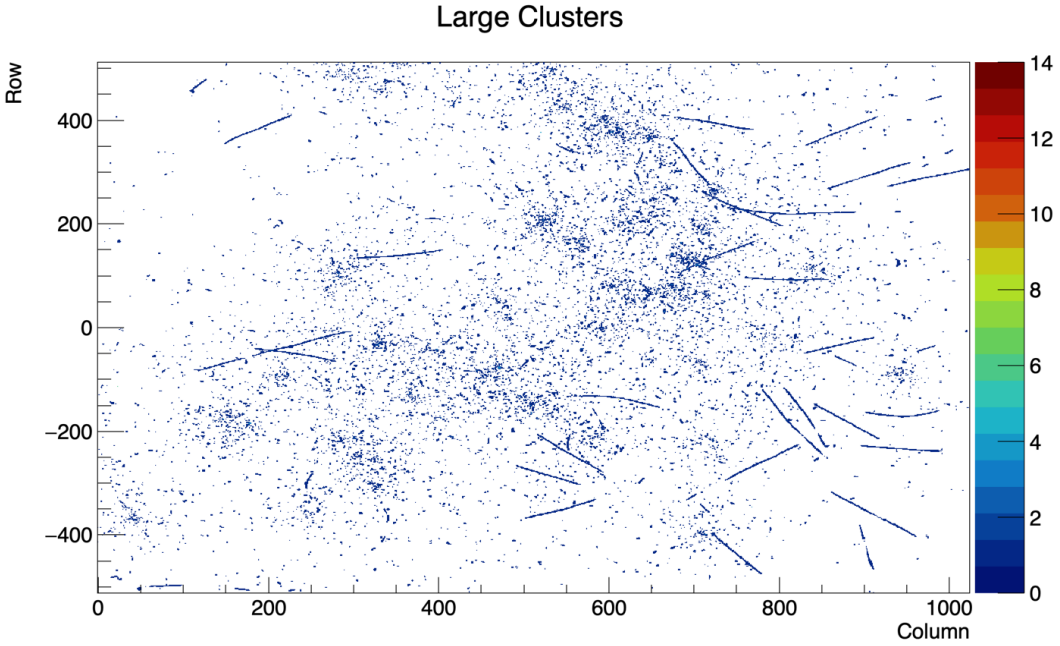


Figure 5.17: Spatial distribution integrated over all layers and all the events containing clusters with a cluster size larger than 200 pixels. **Run 1335**: energy of 3 GeV, temperature of 20°C and strobe length of 2 μs.

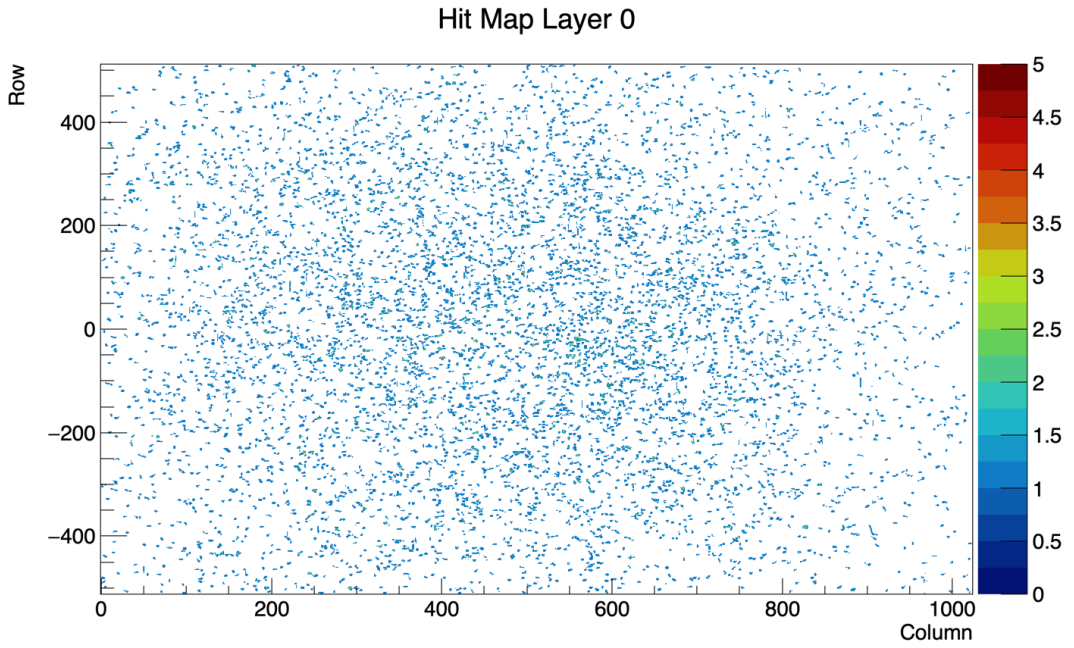


Figure 5.18: Spatial distribution for Layer 0, displaying only clusters with a cluster size between 10 and 20 pixels. **Run 1335**: energy of 3 GeV, temperature of 20°C and strobe length of 2 μs.

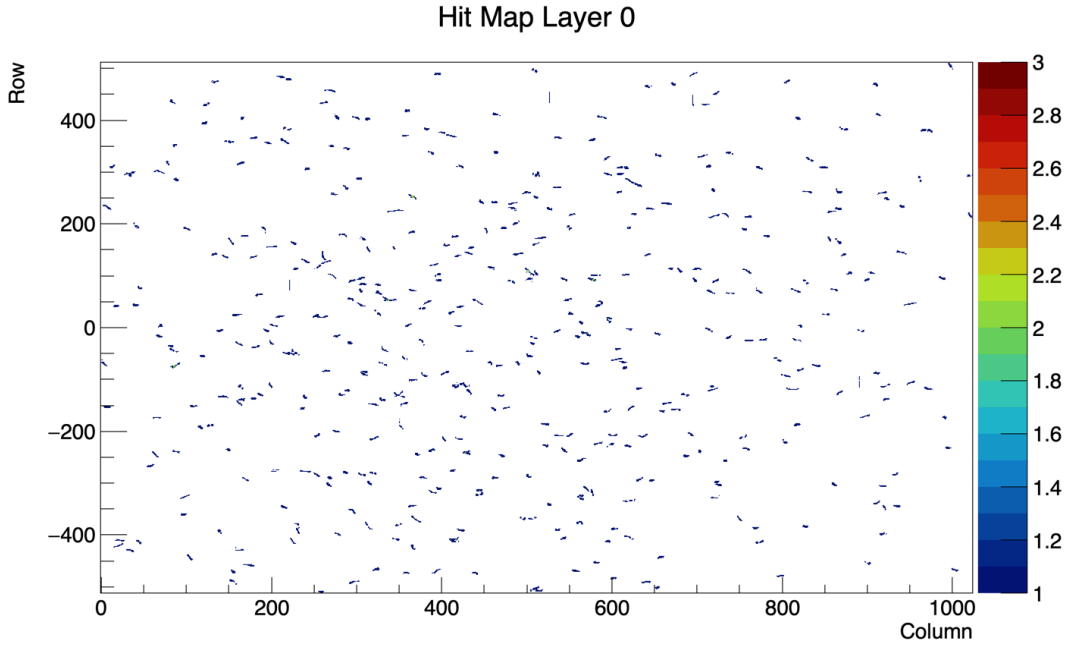


Figure 5.19: Spatial distribution for Layer 0, displaying only clusters with a cluster size between 20 and 30 pixels. **Run 1335**: energy of 3 GeV, temperature of 20°C and strobe length of 2 μs.

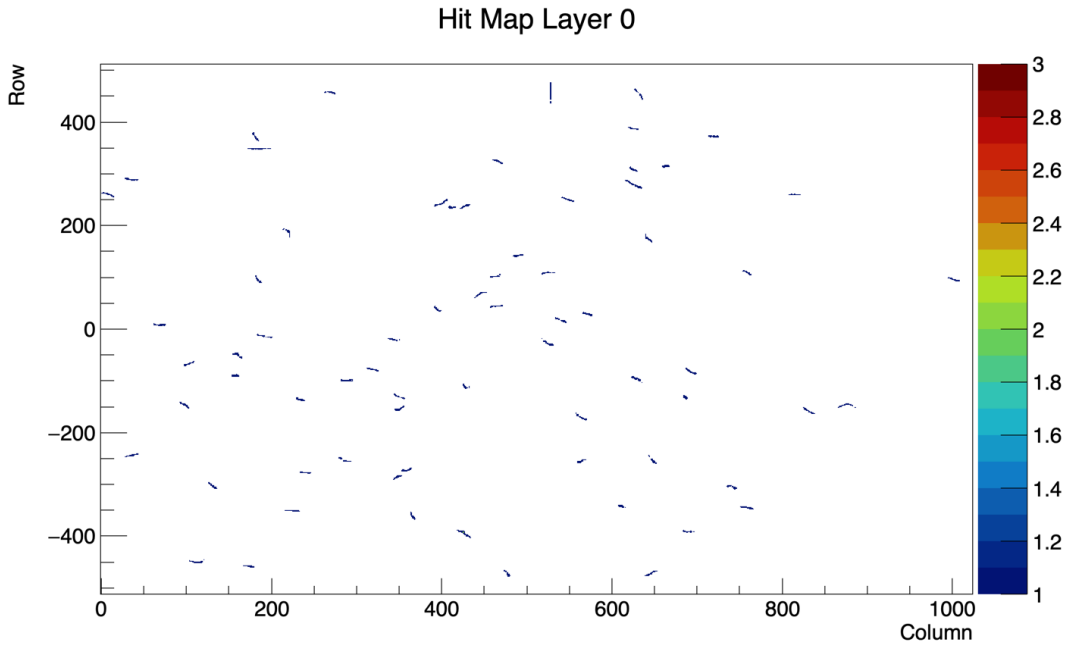


Figure 5.20: Spatial distribution for Layer 0, displaying only clusters with a cluster size between 30 and 40 pixels. **Run 1335**: energy of 3 GeV, temperature of 20°C and strobe length of 2 μs.

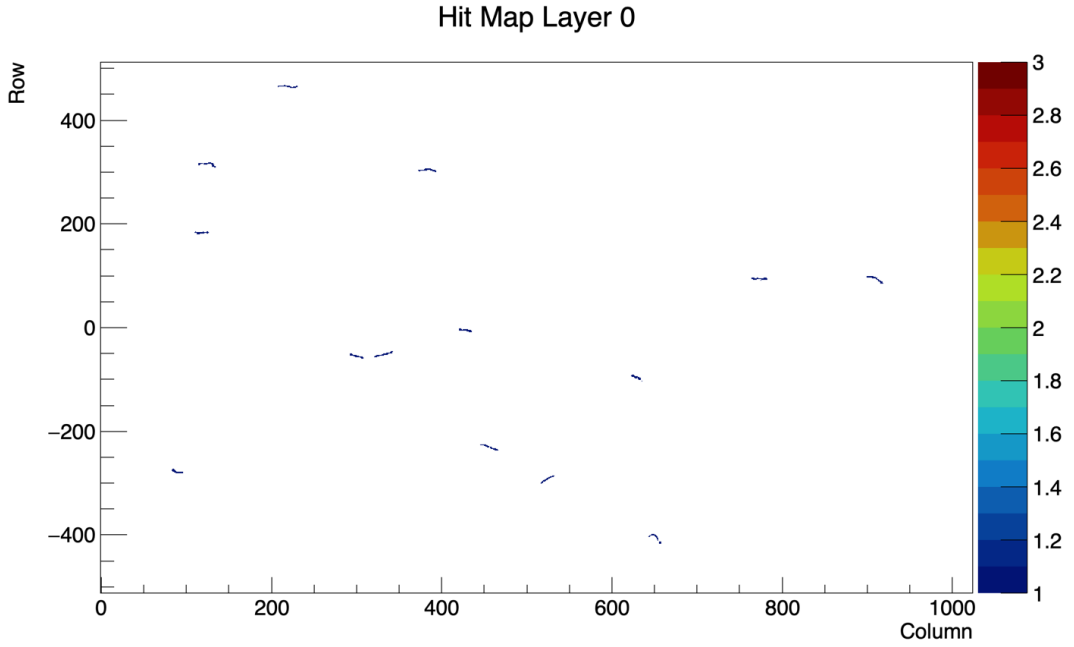


Figure 5.21: Spatial distribution for Layer 0, displaying only clusters with a cluster size between 40 and 50 pixels. **Run 1335**: energy of 3 GeV, temperature of 20°C and strobe length of 2 μs.

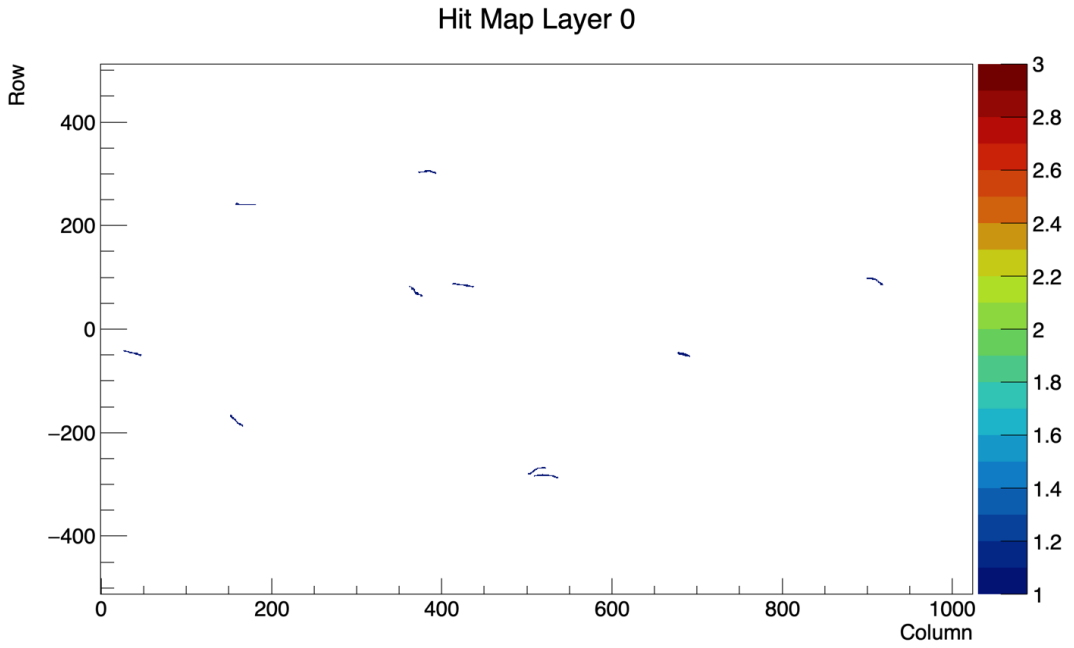


Figure 5.22: Spatial distribution for Layer 0, displaying only clusters with a cluster size between 50 and 60 pixels. **Run 1335**: energy of 3 GeV, temperature of 20°C and strobe length of 2 μs.

Figures 5.23 and 5.24 show the spatial distribution of clusters with a cluster size between 50 and 60 pixels for Layer 5 and Layer 22, respectively. The ratio between a layer's number of clusters with a size between 50-60 pixels and its total number of clusters was calculated. This was done to study if the effect of track-like structures is more prominent in shower maximum and as the shower develops, or if it is equally distributed over all the different layers. The ratio calculation was performed for Layer 0, Layer 5 and Layer 22, and the numbers are presented in Table 5.5.

Table 5.5: The number of clusters in a layer with cluster size between 50 and 60 pixels, the total number of clusters in a layer and the ratio between them, given for Layer 0, Layer 5 and Layer 22.

Layer	**Number of Clusters With Size 50-60**	**Total Number of Clusters**	**Ratio**
0	10	620056	$1.6 \cdot 10^{-5}$
5	1113	14478774	$7.7 \cdot 10^{-5}$
22	24	263180	$9.1 \cdot 10^{-5}$

As seen in Table 5.5, the ratio between the number of large clusters and the total number of clusters increases as the shower develops, where the ratio is largest for Layer 22. This is also the case for the other cluster size ranges, indicating that these clusters are a result of particles from the shower approaching the layer at an angle.

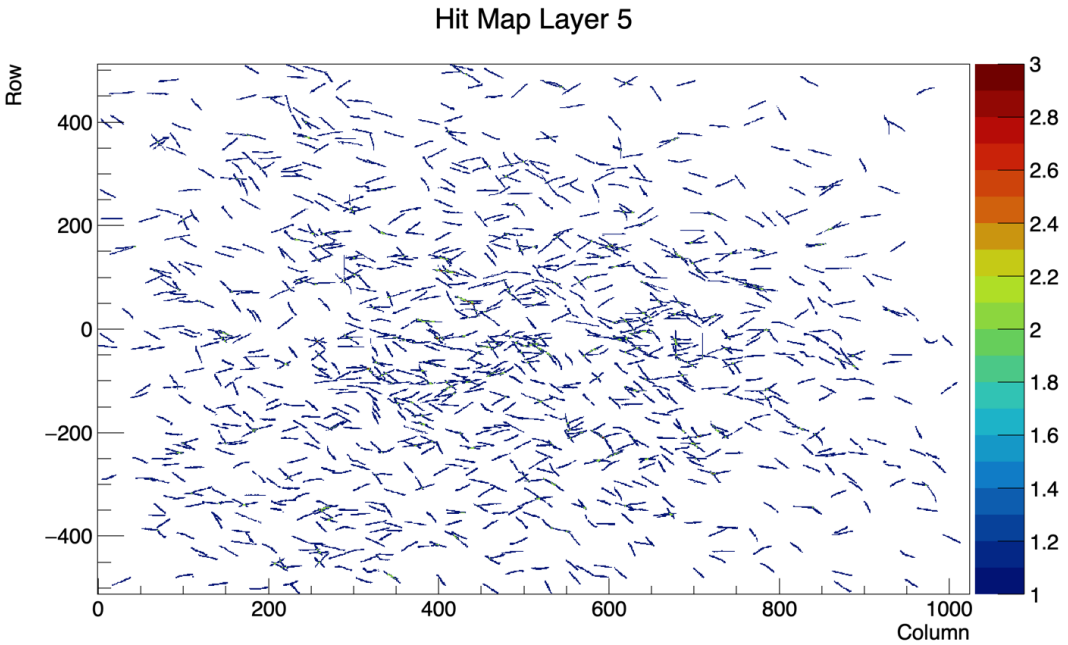


Figure 5.23: Spatial distribution for Layer 5, displaying only clusters with a cluster size between 50 and 60 pixels. **Run 1335**: energy of 3 GeV, temperature of 20°C and strobe length of 2 μs.

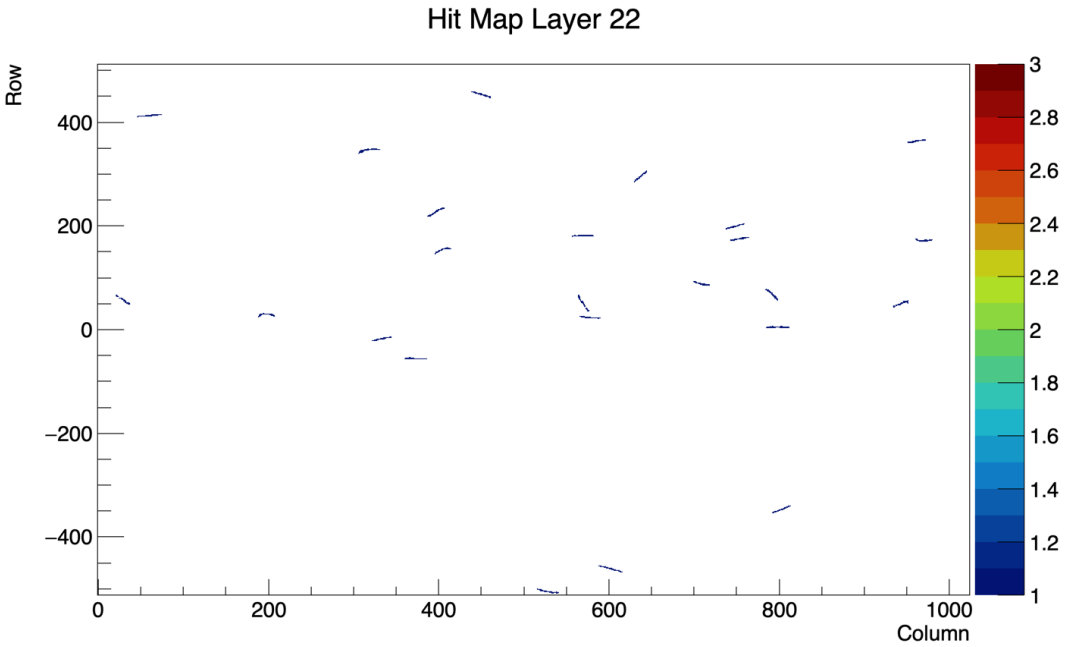


Figure 5.24: Spatial distribution for Layer 22, displaying only clusters with a cluster size between 50 and 60 pixels. **Run 1335**: energy of 3 GeV, temperature of 20°C and strobe length of 2 μs.

Figure 5.25 displays the average cluster size for the different layers in the EPICAL-2. The average cluster size is shown for different event selections, where the yellow graph is for a single

track within a 2 μs readout frame. The green, light blue, blue and purple graphs display the relationship for two, three, four and five tracks within a readout frame, respectively. Lastly, the red displays the average cluster size for no tracks in the first layer and burgundy is based on no event selection.

As seen from the figure, there is a fluctuation in the average cluster size with respect to the layers. The average size starts at about 2.6 in the first layer before it increases to about 3.6. Subsequently, it fluctuates around 3.8 with a minimum average size of 2.8 and a maximum of 4.4. The average cluster size is larger than one would expect for a MIP particle. If this is an effect of overlapping clusters, one should probably observe a peak in average cluster size in shower maximum, in this case, Layer 5. However, the size increases as one moves deeper into the calorimeter, indicating that it might, instead, be caused by the angle of the shower.

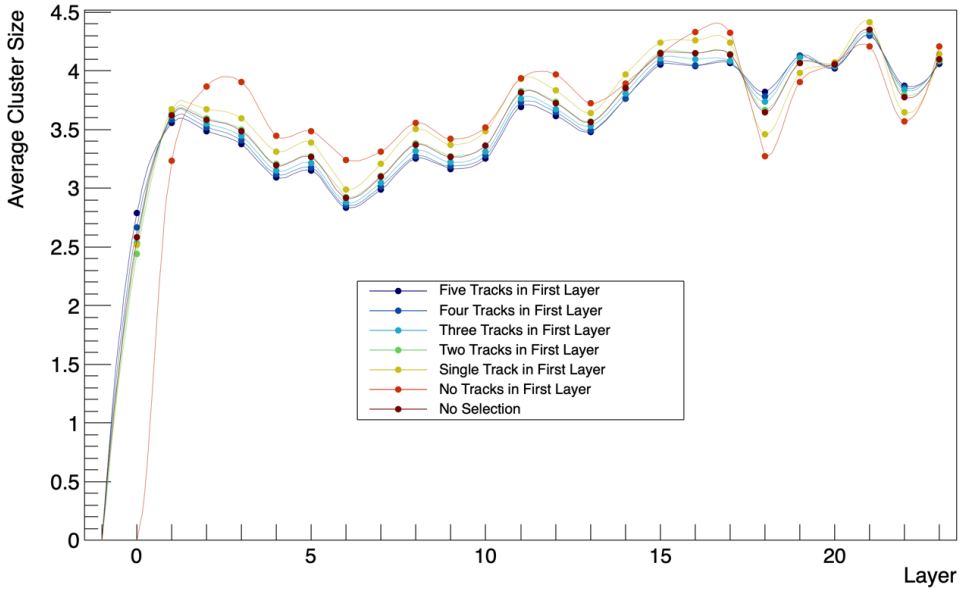


Figure 5.25: Average cluster size versus layer for different event selections. **Run 1335**: energy of 3 GeV, temperature of 20°C and strobe length of 2 μs.

Each of the individual ALPIDE chips have a different threshold, which could explain why there is a fluctuation in average cluster size for the different layers. As mentioned earlier, the runs used in this analysis all have optimised settings, meaning that nine of the ALPIDE chips had adjusted settings as described in Table 5.1. This adjustment was made trying to minimise the difference in threshold values between the different chips. Hence, the fluctuation would likely be more significant without the adjusted settings.

Energy Dependence of the Cluster Size

Figure 5.26 depicts the average cluster size versus layer for different beam energies ranging from 1 GeV to 5.8 GeV. Each graph displays the average cluster size for a single track within a 2 μs readout frame for the various beam energies. As seen from the figure, the cluster size does not seem to depend on the energy. The average cluster size fluctuates, but there is no systematic trend.

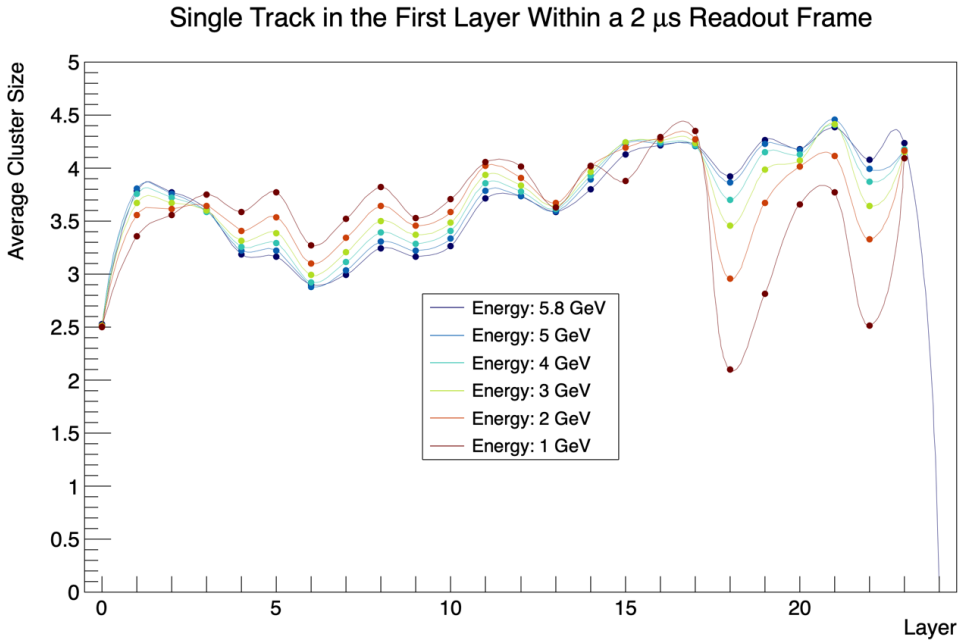


Figure 5.26: Average cluster size versus layer for a single track within a readout frame and different energies. All the runs used here has a temperature of 20°C and a strobe length of 2 μs. The runs are listed in Table 5.2.

Temperature Dependence of the Cluster Size

The average cluster size versus layer is displayed for different temperatures in Figure 5.27. Each graph shows the average cluster size for a single track within a 2 μs readout frame for temperatures of 20°C, 25°C and 30°C. From this plot, one can observe that the average cluster size increases slightly with temperature for almost every layer, except for Layer 21, where the average size for 20°C is larger than for 25°C. In Layer 13, the average size is approximately the same for all the temperatures. One can thus observe a small systematic effect, seeing that the trend is similar for all the temperatures.

The increase in average cluster size is probably caused by a higher diffusion in the epitaxial layer for higher temperatures. As a result, the electrons generated by the incoming particle can spread to several collection diodes.

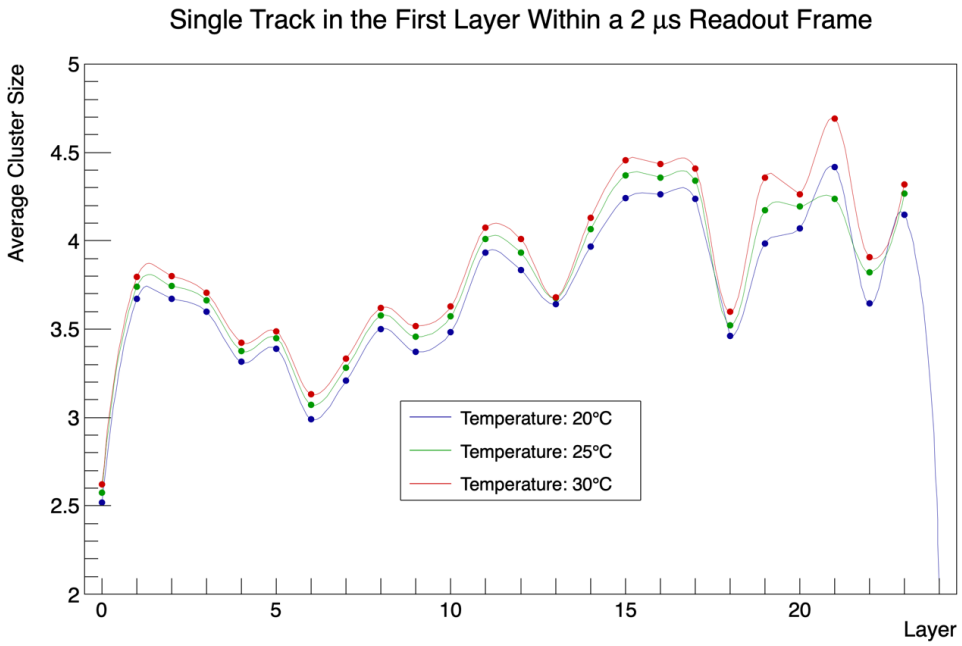


Figure 5.27: Average cluster size versus layer for a single track within a readout frame and different temperatures. All the runs used here have an energy of 3 GeV and a strobe length of 2 μs. The runs are listed in Table 5.2.

Figure 5.28 displays the average cluster size versus temperature, where the cluster size is averaged over all layers. The average cluster size increases from 3.690 ± 0.006 for 20°C to 3.770 ± 0.006 for 25°C and, finally, 3.850 ± 0.004 for 30°C. The uncertainty is here given by the standard error of the mean.

The data points are fitted to a polynomial of first degree using the ROOT-function `TH1::Fit`. The fit function calculates the χ^2 between the data and the first-degree polynomial, to find the fit that gives the lowest χ^2-values. This procedure results in the following linear function:

$$f(x) = (0.02 \pm 0.3)x + (3 \pm 7) \tag{5.1}$$

with $\chi^2 = 8.5 \cdot 10^{-10}$.

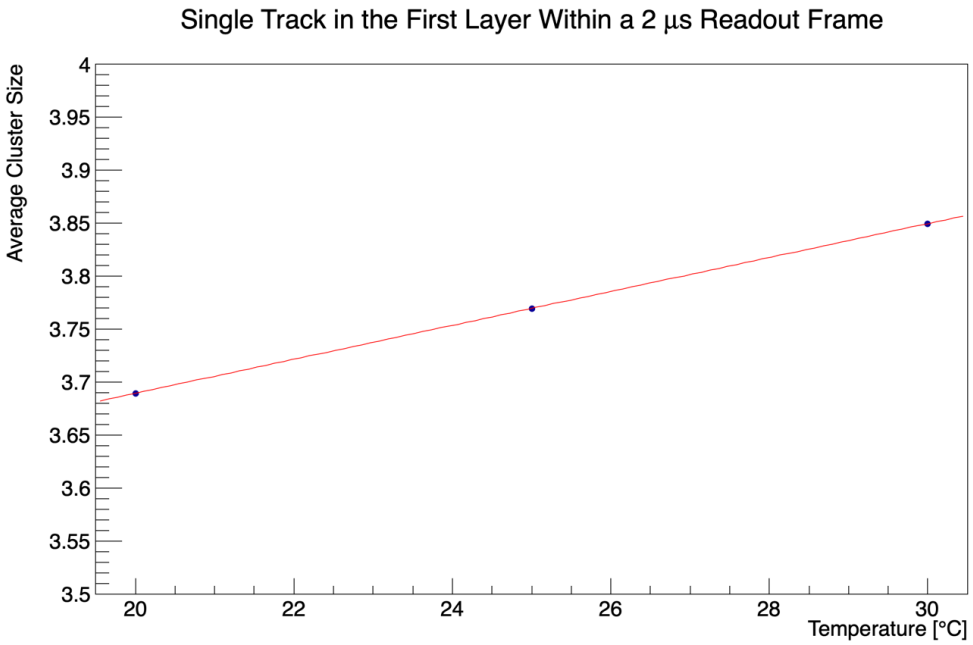


Figure 5.28: Average cluster size, averaged over all layers, versus temperature for a single track within a readout frame. The data points are fitted with the linear function described in Equation 5.1. All the runs used here have an energy of 3 GeV and a strobe length of 2 μs. The runs are listed in Table 5.2.

Although the temperature is not a critical parameter, one can use this increase in average cluster size to argue the importance of maintaining a stable temperature during the data acquisition, in order to use the cluster size as a measure of energy deposition. The importance does not lie in what value the temperature measures, but that it is kept relatively constant. This effect can also be an argument for a water-cooling system for the pCT to preserve a stable temperature during a pCT scan. It should, however, be noted that a temperature change of a few degrees does not lead to a significant difference in cluster size, as seen in Figure 5.28.

Strobe Length Dependence of the Cluster Size

Figure 5.29 displays the average cluster size for a single track within a readout frame versus layer for strobe lengths of 2 μs and 20 μs. The graphs are in principle expected to be similar, as the only effect that should affect the cluster size is that there is a bit more collection of noise for the longer strobe length of 20 μs. Hence, the average size might decrease slightly, as observed in the plot. This effect is most noticeable in the later layers as there are fewer hits in these layers, resulting in a more considerable contribution from the noise to the average cluster size.

By counting the noisy pixels as a cluster, the average cluster size will decrease, since the noisy pixels are single isolated pixels. The problem with MIP particles is that they can create clus-

ters of cluster size one, depending on their impact position. Thus, a filter removing clusters of size one might also filter out a lot of data, not just noise. Hence, the noise has to be removed statistically or before the data taking.

As mentioned in a previous section, noisy pixels are masked based on pedestal runs if the pixels are active in more than 90% or 98% of the readout frames. However, despite masking being implemented prior to data acquisition, there are still noisy pixels present in the data.

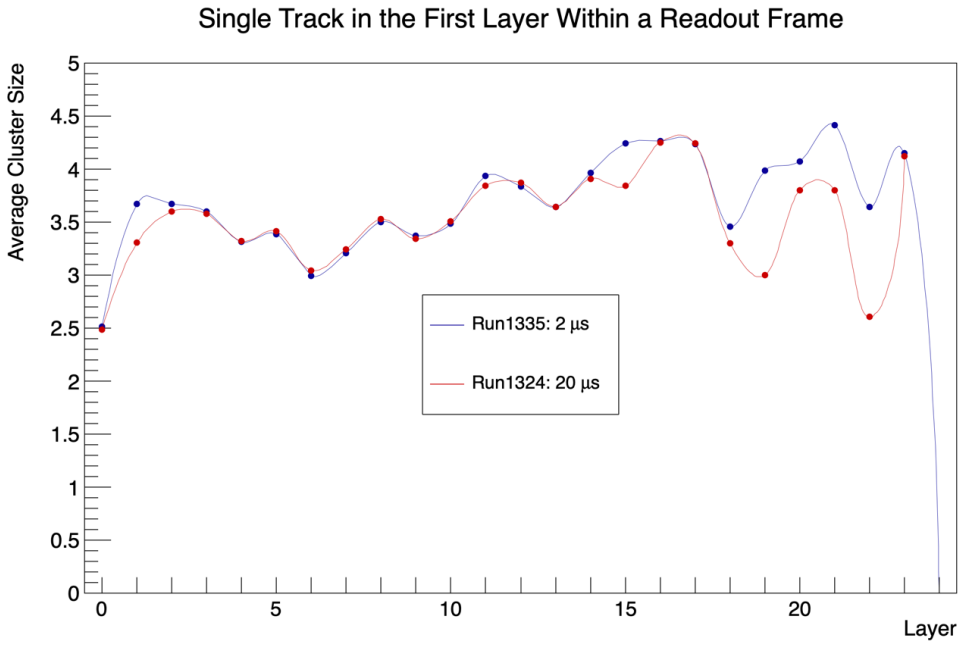


Figure 5.29: Average cluster size versus layer for a single track within a readout frame and different strobe lengths. All the runs used here have an energy of 3 GeV and a temperature of 20°C. The runs are listed in Table 5.2.

5.3.3 Noise

A noisy pixel is defined as a pixel that always produces a signal, or fires significantly more often than the majority of the pixels. This is electrical and not statistical noise, where pixels are producing signals without energy being deposited in the epitaxial layer by a traversing particle. During the beam test, pixels producing signals in more than 90% of the events in the pedestal runs were masked using a pixel mask. These pixels are listed in Table 5.6.

Table 5.6: Coordinates of the noisy pixels masked during the beam test, meaning pixels firing in more than 90% of the events. This table also gives an overview of the fraction of events where these pixel fires. The noisy pixels are the same for all of the pedestal runs listed in Table 5.3.

Layer	**Chip ID**	**Coordinate**	**Fraction**
1	44	(24,21)	1
4	18	(737,351)	1
4	18	(901,239)	1
10	13	(953,414)	1
10	13	(954,414)	1
16	6	(937,375)	1
18	4	(359,301)	1

Figures 5.30 and 5.31 displays the number of entries for the different number of hits integrated over all events and layers for Runs 1246 and 1330, respectively. These distributions are approximately Gaussian in form, with a mean of about 9.8 and 15.5, respectively. The distributions also start at seven hits, because of the seven noisy pixels listed in Table 5.6 that produce a signal in 100% of the events.

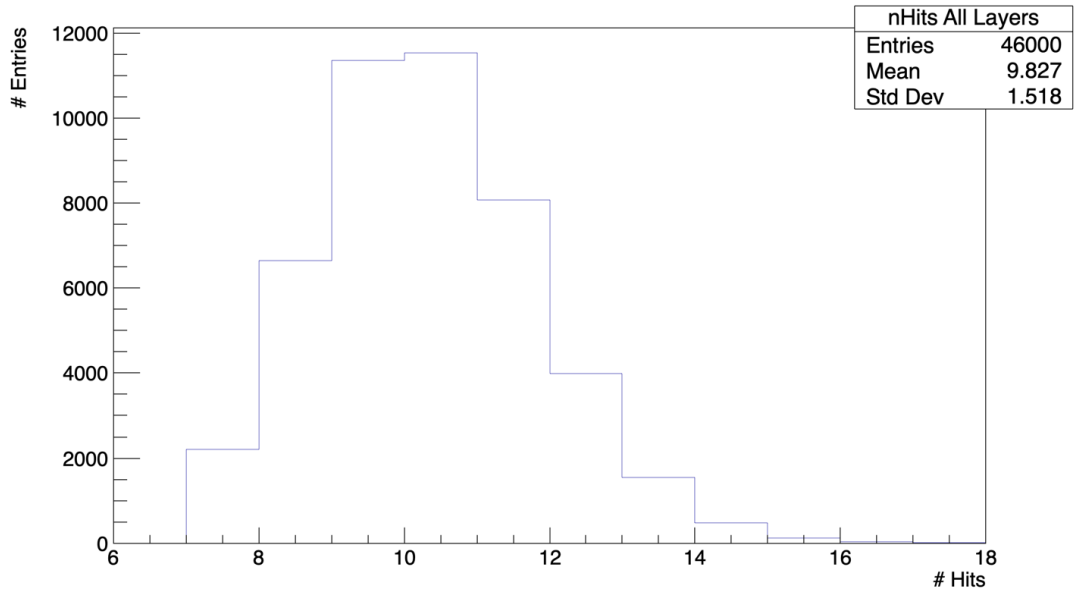


Figure 5.30: Distribution showing the number of entries for the different number of hits, integrated over all layers. **Run 1246**: temperature of 20.1°C and strobe length of 2 μs.

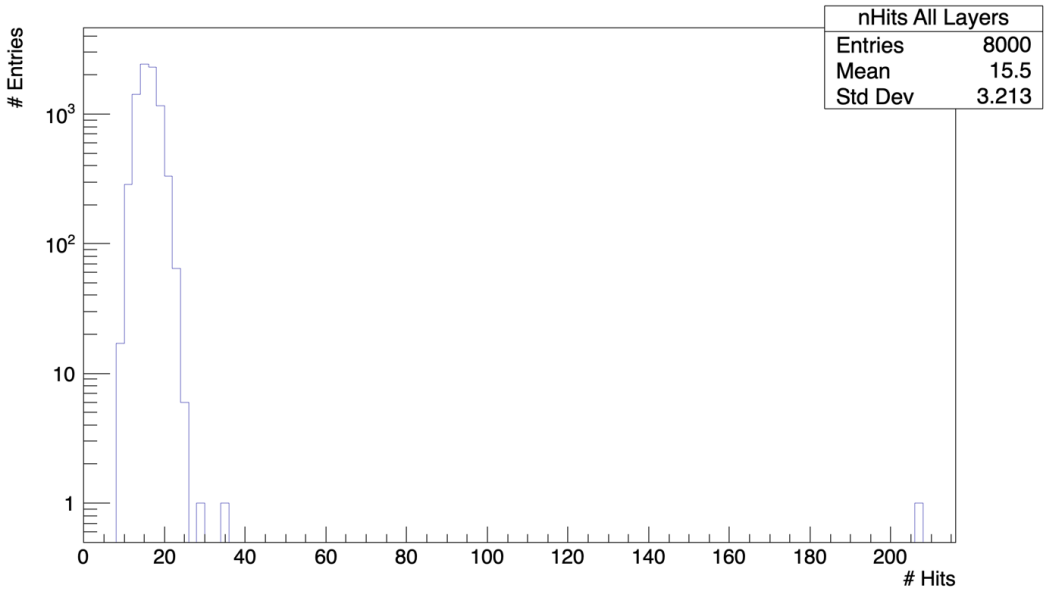


Figure 5.31: Distribution showing the number of entries for the different number of hits, integrated over all layers. **Run 1246**: temperature of 20°C and strobe length of 20 μs.

Table 5.7 gives an overview of the number of hits from single and multiple adjacent pixels present in the pedestal runs listed in Table 5.3. Table 5.7 also gives the percentage and rate of pixels. The percentage of pixels is given by the ratio between the number of firing pixels per event and the total number of pixels in the 24-layered prototype. As mentioned earlier, each ALPIDE chip has 1024×512 pixels, and the EPICAL-2 has 48 ALPIDE chips. Further, the rate is given by the number of firing pixels per readout frame. Table 5.3 lists the number of events and strobe lengths for the different pedestal runs.

As seen in Table 5.7, the percentage of firing pixels is approximately the same for all the runs with a 2 μs strobe length. The percentage is, still, slightly lower for Run 1332 and 1303, compared to Runs 1246 and 1356. However, when looking at the rate of firing pixels, there seems to be no systematic trend relating to temperature for the runs with a 2 μs strobe length. Run 1330 with a 20 μs strobe length, on the other hand, appears to have a higher rate and higher percentage. The higher percentage may be due to the low number of recorded events, while the high rate might be due to the longer strobe length. The rate of pixels firing per readout frame is also consistent with the mean of the distributions displayed in Figures 5.30 and 5.31.

Table 5.7: The number of hits from multiple adjacent and single pixels accumulated over all events and layers for the pedestal runs listed in Table 5.3. This table also gives an overview of the percentage and rate per frame of pixels firing in the pedestal runs.

Run Number	**Water Temperature** [°C]	**Strobe Length** $[\mu s]$	**Number of Hits From Multiple Adjacent Pixels**	**Number of Hits From Single Pixels**	**Percentage** [%]	**Rate**
1246	20.1	2	46040	359975	$3.87 \cdot 10^{-5}$	9.73
1332	20	2	96081	732773	$3.79 \cdot 10^{-5}$	9.53
1303	25	2	46420	352514	$3.80 \cdot 10^{-5}$	9.56
1356	30	2	96609	711510	$3.86 \cdot 10^{-5}$	9.70
1330	20	20	8075	107713	$5.99 \cdot 10^{-5}$	15.07

Figure 5.32 shows the cluster size distribution for Run 1330, illustrating that most of the firing pixels are single pixels. This is also indicated by Figure 5.33, where the firing pixels appear to be uniformly distributed over the chip. There are also a few clusters of size two, seeing that two of the noisy pixel firing in all of the events were neighbouring pixels. There are also clusters of greater size, which can be observed in the spatial distribution in Figure 5.33. The largest clusters appear track-like and might be caused by cosmic muons.

Figure 5.34 shows the same spatial distribution, but zoomed in on the clusters in the lower right quadrant. As indicated by the colour scale and the fraction in Table 5.6, the pixels at $(953, 414)$ and $(954, 414)$ are firing in all of the events. These two pixels are located in Chip ID 13 in Layer 10, as listed in Table 5.6.

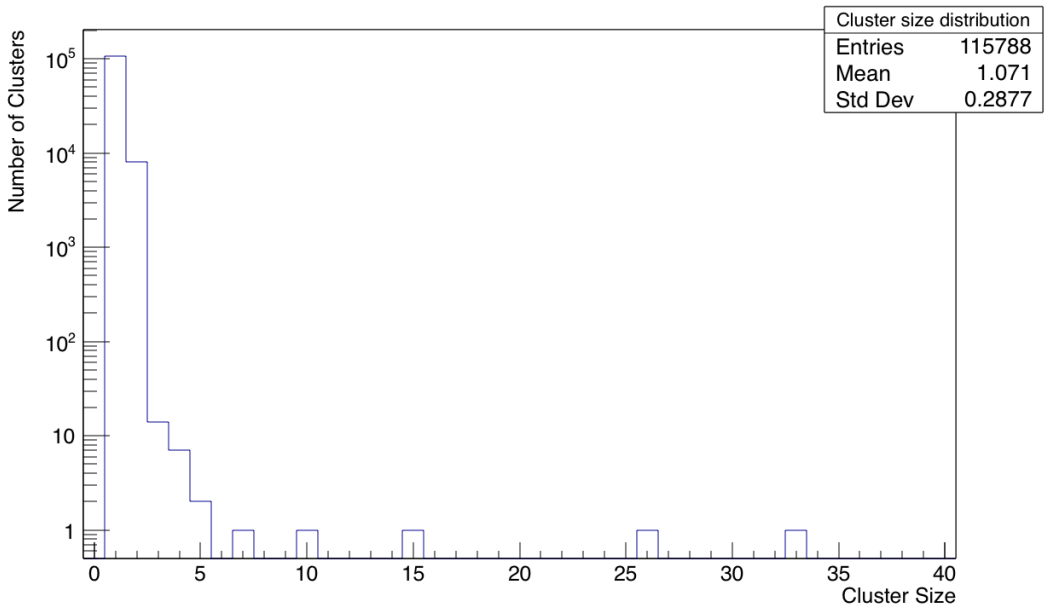


Figure 5.32: Cluster size distribution integrated over all events and layers. **Run 1330**: temperature of 20°C and strobe length of 20 μs.

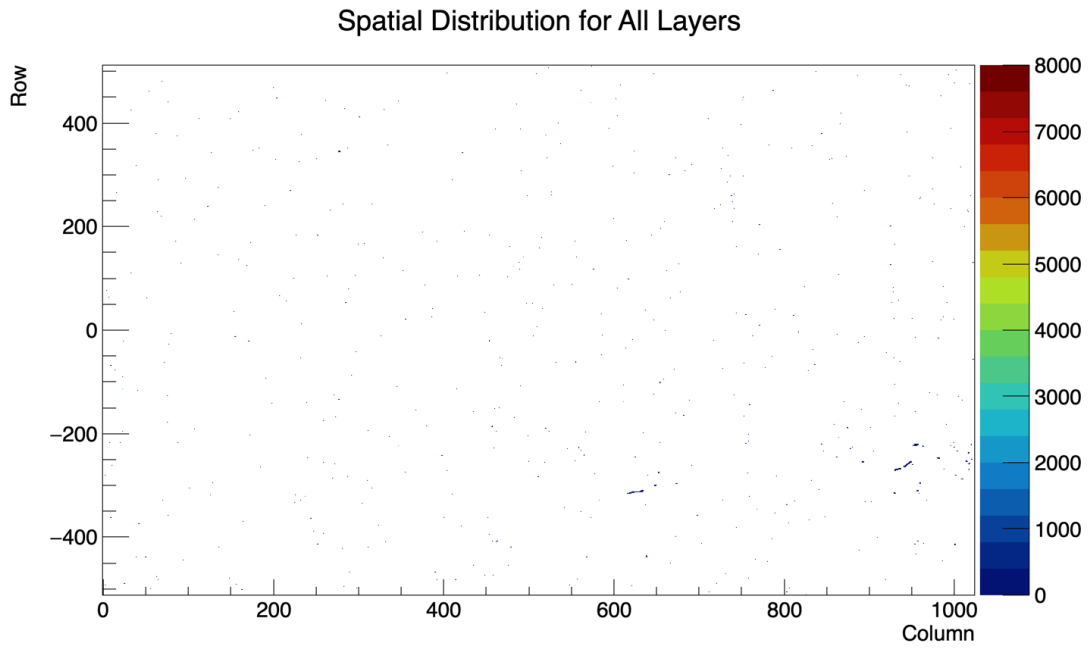


Figure 5.33: Spatial distribution integrated over all events and layers. **Run 1330**: temperature of 20°C and strobe length of 20 μs.

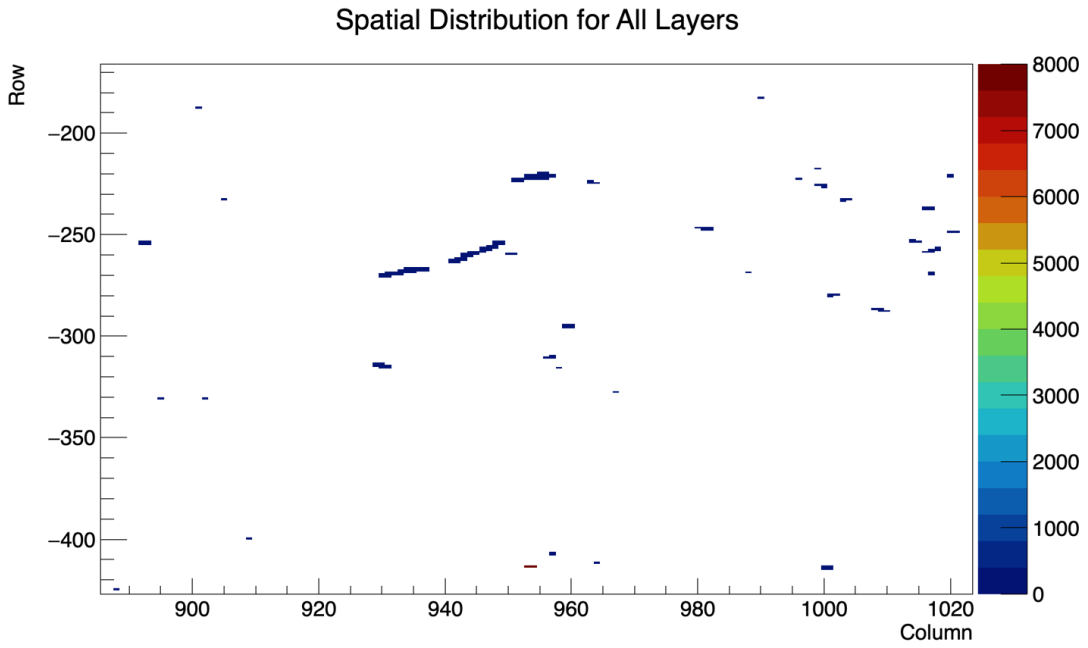


Figure 5.34: Spatial distribution integrated over all events and layers, zoomed in on the clusters in the lower right quadrant. **Run 1330**: temperature of 20°C and strobe length of 20 μs.

5.3.4 Shower Studies

Single Track

The plot in Figure 5.35 shows the number of hits per event versus layer for events with a single track in the first layer, within a 2 μs readout frame. The event selection based on the hit position in the first layer was utilised as an attempt to ensure shower containment. The error-bars are, as described earlier, given by the standard deviation of the quantities. The "wrinkles" seen in the plot might be caused by the difference in threshold between the different chips, as mentioned in context with the average cluster size versus layer. There is also a small dip in Layer 21 as Chip ID 27 was excluded in most of the data runs.

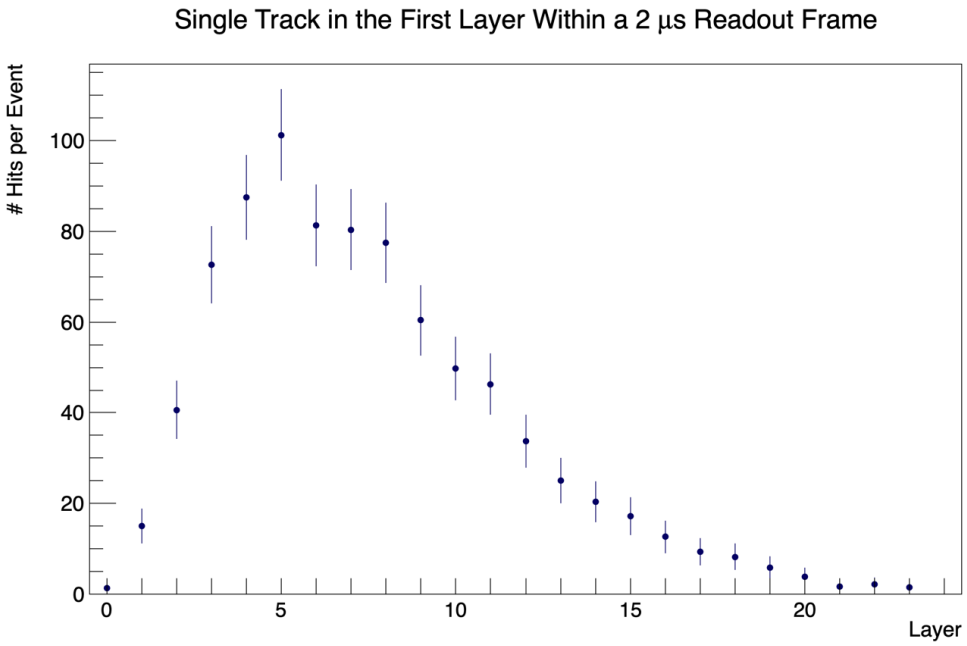


Figure 5.35: Number of hits per event for events with a single track in the first layer, within a 2 μs readout frame, with corresponding error-bars. The cluster in the first layer is also within a 15×15 mm^2 area. **Run 1335**: energy of 3 GeV, temperature of 20°C and strobe length of 2 μs.

Figure 5.36 displays the number of clusters per event versus layer for events with a single track, within a 15×15 mm^2 area around the centre of the first layer and within a 2 μs readout frame. The errors are calculated as described above. The clustering algorithm smooths out the "wrinkles" observed in the previous plot. Clusters of size one have been included, seeing that MIP particles can create single-pixel hits depending on the hit position. The use of MIP particles, thus, leads to difficulties in distinguishing noise from one-pixel clusters created by a traversing particle.

When using a clustering algorithm, there is no need to take the sensitivity of each individual pixel into account, as the traversing particle always leaves behind a cluster. The size of the cluster is, however, dependent on the threshold. Consequently, the clustering algorithm smooths the curve as it is less sensitive to charge diffusion in the epitaxial layer.

As particles pass through a thin layer, there are large Landau fluctuations in energy loss, causing the deposition of energy to fluctuate. This effect, together with the varying thresholds, might result in a different number of hits for the same charge distribution. Thus, the number of hits per event will fluctuate more than the number of clusters per event when combining the effect of the Landau fluctuations and the threshold variations. This effect is visible in Figures 5.35 and 5.36. Here one observes, for instance, a drop in the number of hits in Layer 6 in Figure 5.35 and a smooth curve in Figure 5.36. Thus, providing an argument supporting the use of the number of clusters over the number of hits.

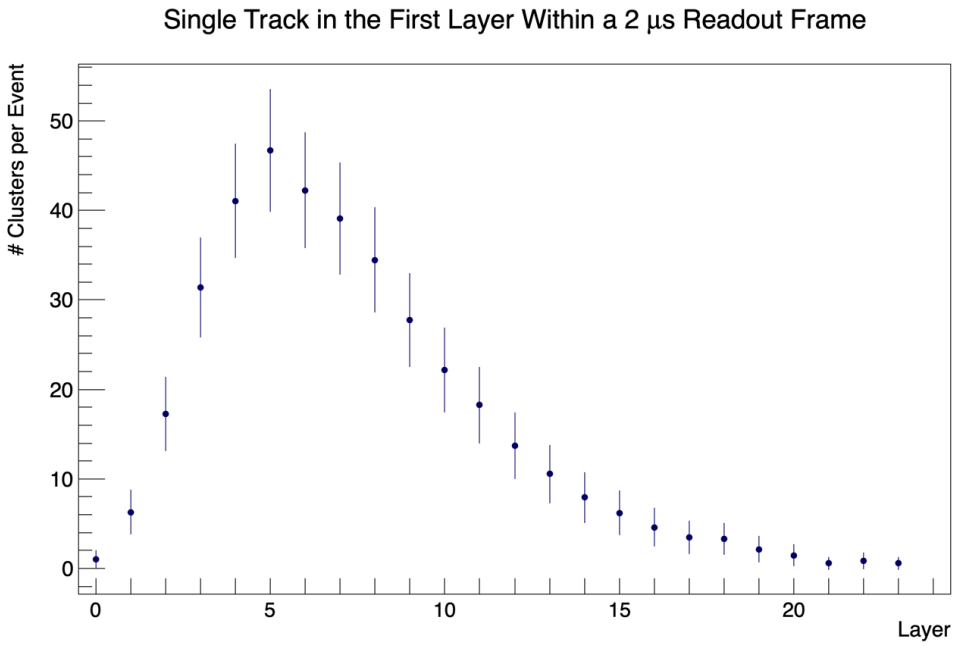


Figure 5.36: Number of clusters per event for events with a single track in the first layer within a 2 μs readout frame, with corresponding error-bars. The cluster in the first layer is also within a 15×15 mm^2 area. **Run 1335**: energy of 3 GeV, temperature of 20°C and strobe length of 2 μs.

Figure 5.37 shows the relationship between the total number of hits in a shower and the energy for a single track, plotted with corresponding uncertainties. The number of hits is integrated over all layers.

The data points are fitted to a polynomial of first degree using `TH1::Fit`. This procedure results in the following linear function:

$$f(x) = (271 \pm 7)x + (12 \pm 20) \tag{5.2}$$

with $\chi^2 = 4.9$.

The relationship between the number of hits per event and energy appears quite linear, as the linear fit falls within the uncertainty interval for most of the data points, except for the energies of 3 GeV and 5.8 GeV. The number of hits per event is slightly too high for 3 GeV and slightly too low for 5.8 GeV for the linear fit to fall within their error-bars.

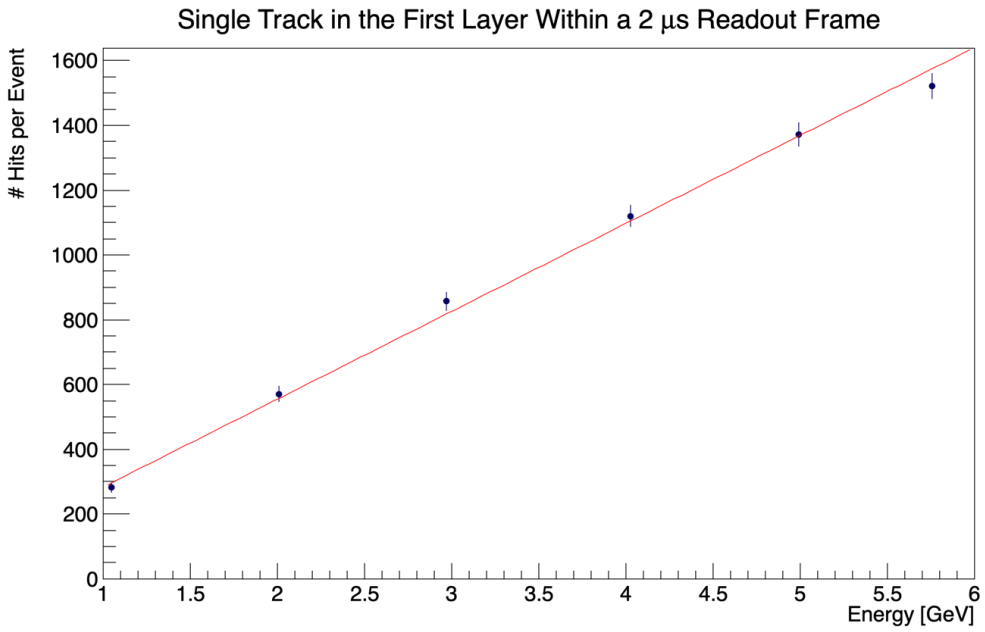


Figure 5.37: Number of hits per event versus energy for a single track within a readout frame, with corresponding error-bars. The data points are fitted with the linear function described in Equation 5.2.

Moreover, Figure 5.38 shows the relationship between the total number of clusters in a shower and the energy for a single track, plotted with corresponding uncertainties. The number of clusters is integrated over all layers. The data is fitted with the following linear function:

$$f(x) = (82 \pm 4)x - (5 \pm 10) \tag{5.3}$$

with $\chi^2 = 0.78$.

In the case of the relationship between the number of clusters per event versus energy, the linear fit seems better. Despite the data point for 3 GeV being slightly too high, and the one for 5.8 being slightly too low, the linear fitting function falls within all the uncertainty intervals. Additionally, it has a lower χ^2-value than for the number of hits per event versus energy.

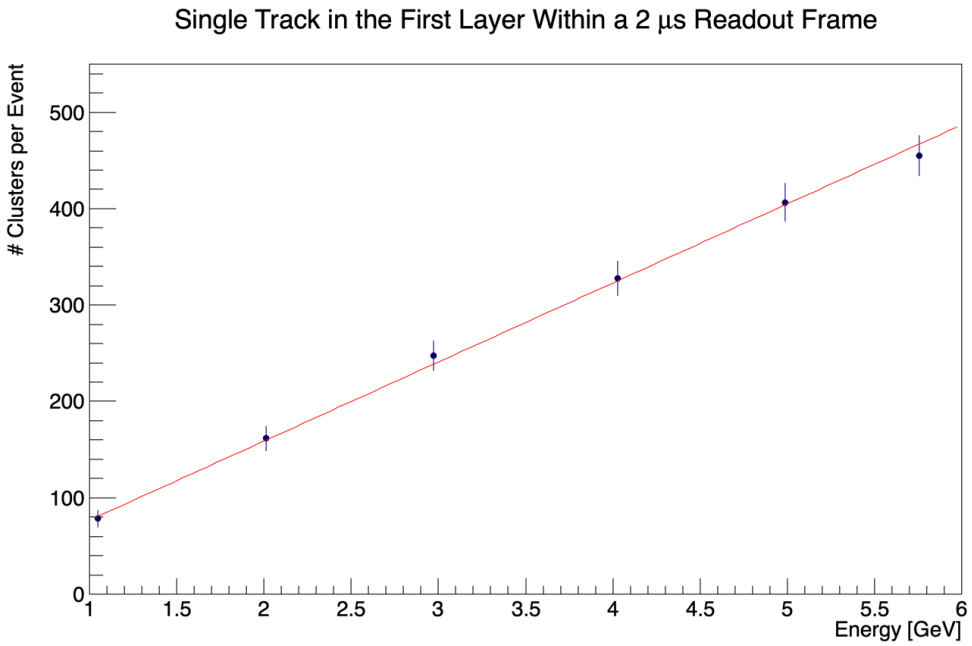


Figure 5.38: Number of clusters per event versus energy for a single track within a readout frame, with corresponding error-bars. The data points are fitted with the linear function described in Equation 5.3.

As the linear fit for the number of clusters per event versus energy has the lowest χ^2-value, where the linear fit falls within the error-bars for all energies, it might be advantageous to use the number of clusters per event instead of the number of hits. Using clusters instead of hits also smooths the distribution versus the different layers of the prototype, resulting in fewer fluctuations.

Figures 5.39, 5.40 and 5.41 display the preliminary results for the transverse hit density distributions, averaged over all events, for single tracks within a 2 μs readout frame. The transverse distributions are plotted for Layers 2, 4 and 18 for 1 GeV, and Layers 2, 5 and 18 for 3 GeV and 5 GeV, where R is the distance from the shower axis. As mentioned earlier, in Section 5.1.2, the results are not corrected for beam divergence, for internal chip misalignment or misalignment of the calorimeter with respect to the beam. Thus, assuming that the beam hits the calorimeter perpendicularly.

The transverse shower development is dominated by multiple scattering, where the primary contribution originates from low energy electrons. The lateral width of the distribution increases with the increasing depth of the longitudinal shower development. The central part of the distribution stems mainly from multiple Coulomb scattering of electrons and positrons, mostly occurring in the early stages of the shower development. The tail comes from low energy photons and electrons produced in photoelectric effect and Compton scattering, which mainly take place late in the shower evolution.

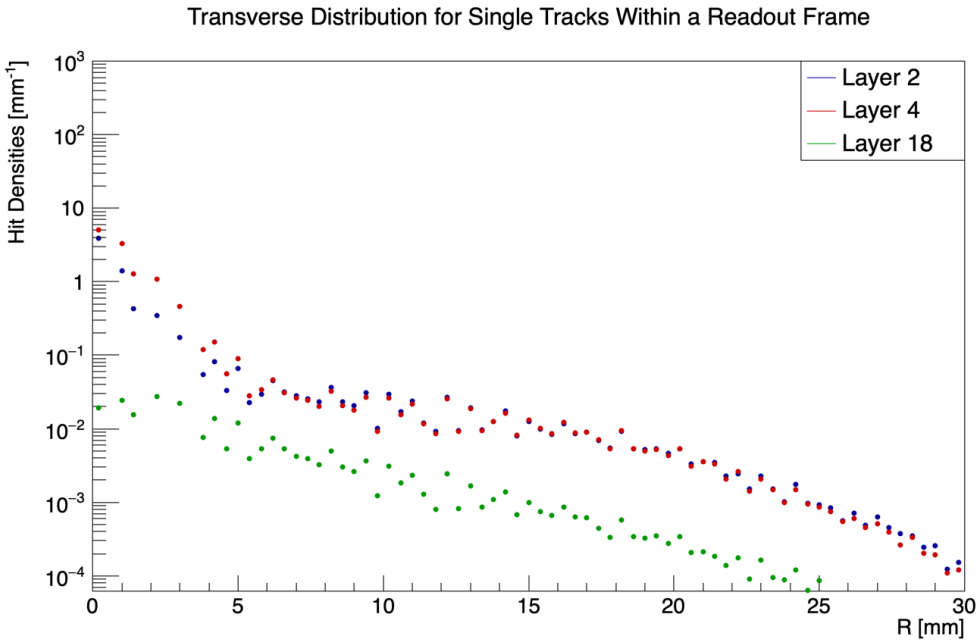


Figure 5.39: Transverse distribution of the hit densities for single tracks within a 2 μs readout frame, averaged over all events, for Layers 2, 4 and 18. The hit in the first layer is also within a 15×15 mm^2 area. **Run 1336**: energy of 1 GeV, temperature of 20°C and strobe length of 2 μs.

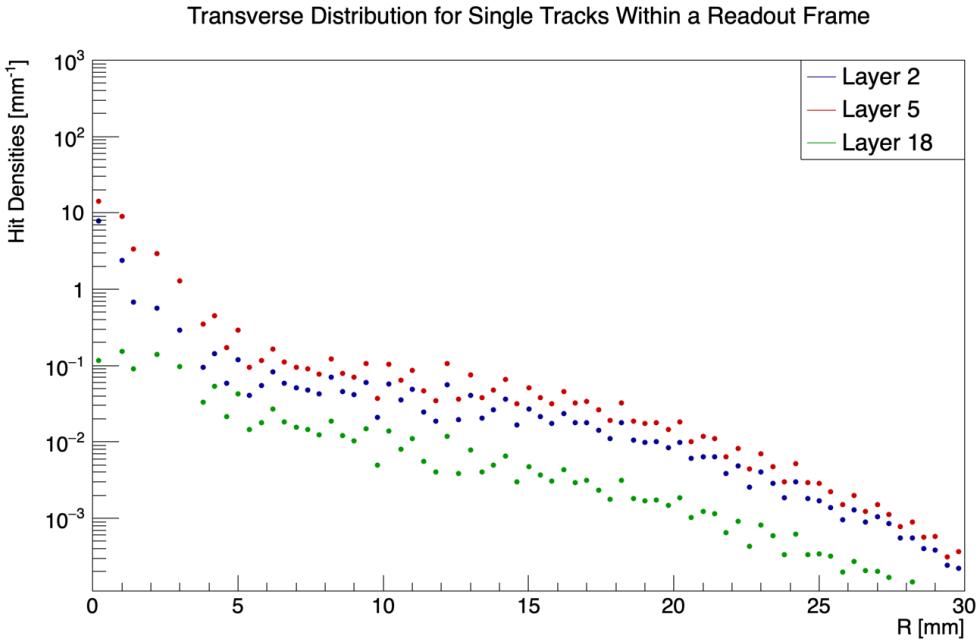


Figure 5.40: Transverse distribution of the hit densities for single tracks within a 2 μs readout frame, averaged over all events, for Layers 2, 5 and 18. The hit in the first layer is also within a 15×15 mm^2 area. **Run 1335**: energy of 3 GeV, temperature of 20°C and strobe length of 2 μs.

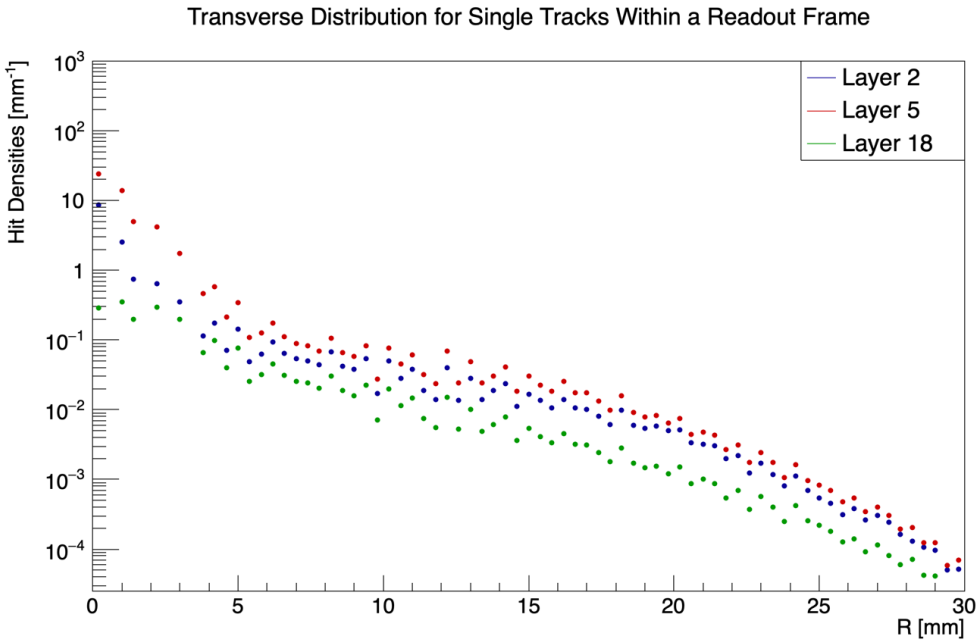


Figure 5.41: Transverse distribution of the hit densities for single tracks within a 2 μs readout frame, averaged over all events, for Layers 2, 5 and 18. The hit in the first layer is also within a 15×15 mm^2 area. **Run 1339**: energy of 5 GeV, temperature of 20°C and strobe length of 2 μs.

Multiple Tracks

Figures 5.42 and 5.43 display the number of hits per event and the number of clusters, respectively, for the different layers in the EPICAL-2 prototype. The plots were retrieved using the first event selection described in Section 5.2.2, which is based on the number of clusters in the first layer, also including no tracks in the first layer and "No Selection". The clusters in the first layer must also be located within a 15×15 mm^2 area around the centre of the layer.

Here, the general trend is the same as for a single track. Several of the event selections have overlapping error-bars, but the graphs for four and five tracks are almost entirely overlapping. This overlap is the case for both the number of hits and the number of clusters per event.

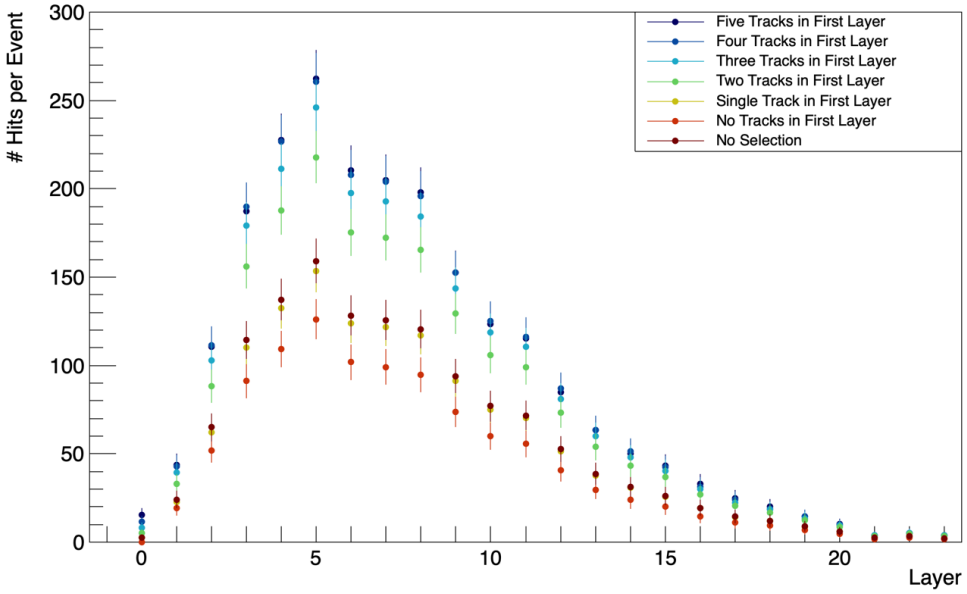


Figure 5.42: Number of hits per event for different event selections, with corresponding error-bars. The clusters in the first layer are also within a 15×15 mm^2 area around the centre of the layer. **Run 1335**: energy of 3 GeV, temperature of 20°C and strobe length of 2 μs.

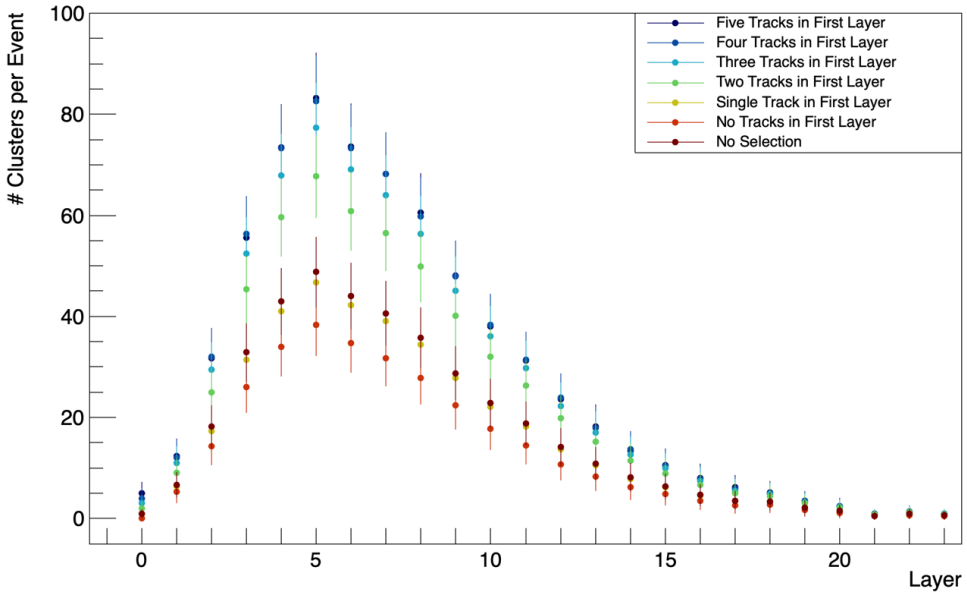


Figure 5.43: Number of clusters per event for different event selections, with corresponding error-bars. The clusters in the first layer are also within a 15×15 mm^2 area around the centre of the layer. **Run 1335**: energy of 3 GeV, temperature of 20°C and strobe length of 2 μs.

Figure 5.44 and 5.45 display the relationship between the number of clusters in the first layer and the number of clusters in shower maximum, averaged over all events. An additional

event selection was also applied to ensure that the identification of clusters in the first layer is correct. This event selection disregards events that do not have a hit in the second layer, directly behind the hit in the first. This additional filtering of events was done to ensure that the hit in the first layer develops into a shower, and is not merely a noisy pixel.

Figure 5.44 shows this relationship for Run 1335 with a strobe length of 2 μs. Here the linear fit falls within the uncertainty intervals for all data points, and the linear fitting function is given by:

$$f(x) = (24 \pm 2)x - (9 \pm 6) \tag{5.4}$$

with $\chi^2 = 1.4$.

The number of clusters in shower maximum for six clusters in the first layer has the largest deviation from the straight line. This data point is also based on the least amount of statistics, with only 107 events, according to the applied event selection. All the other data points have more than 200 events, where there are more events for fewer clusters in the first layer.

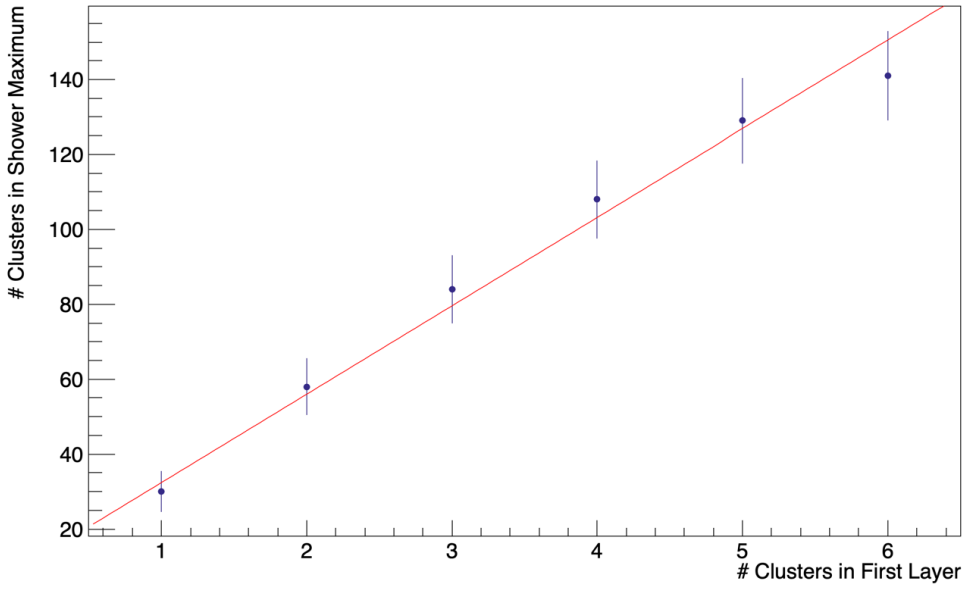


Figure 5.44: The relationship between the number of clusters in the first layer and the number of clusters in shower maximum, averaged over all events with corresponding uncertainty. The red graph is a linear fit to the data given by Equation 5.4. **Run 1335**: energy of 3 GeV, temperature of 20°C and strobe length of 2 μs.

Figure 5.45 shows the same relationship for Run 1324, having a strobe length of 20 μs. Here the linear fit also falls within all the uncertainty intervals, and the linear function is given by:

$$f(x) = (29 \pm 2)x - (1 \pm 6) \tag{5.5}$$

with $\chi^2 = 0.12$.

The relationship for the longer strobe length appears to have a better linear fit, which is also indicated by the lower χ^2-value. Here the linear function intersects close to the centre of the error-bars for almost all data points. The number of clusters in shower maximum for eight clusters in the first layer is the one that deviates the most, but it is also based on only 169 events. All the other data points have more than 200 events, where there are most events for fewer clusters in the first layer.

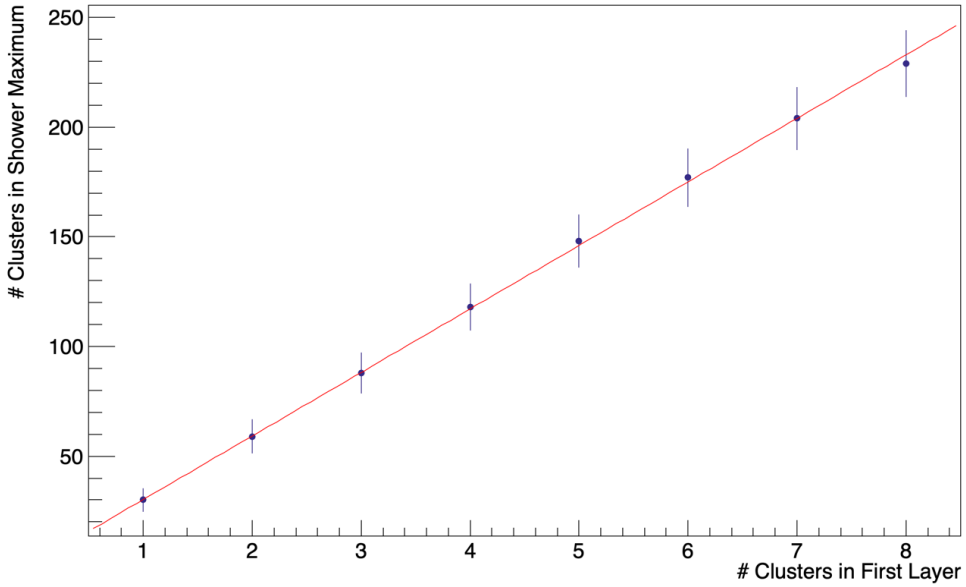


Figure 5.45: The relationship between the number of clusters in the first layer and the number of clusters in shower maximum, averaged over all events with corresponding uncertainty. The red graph is a linear fit to the data given by Equation 5.5. **Run 1324**: energy of 3 GeV, temperature of 20°C and strobe length of 20 μs.

Based on the linear responses shown in Figure 5.44 and 5.45, one might argue that the longer strobe length is a better choice. However, by studying the runs from November taken at different strobe lengths, this conclusion might not be as clear. Figure 5.46 displays the number of hits per event integrated over all the 12 layers of the November prototype for different strobe lengths. Here the number of hits per event seems to peak at 2 μs, having an almost constant value for strobe lengths longer than 2 μs. For a strobe length shorter than 2 μs there is a drop in the number of hits per event for decreasing strobe lengths, so it appears that a strobe length shorter than 2 μs will lead to less sampled data.

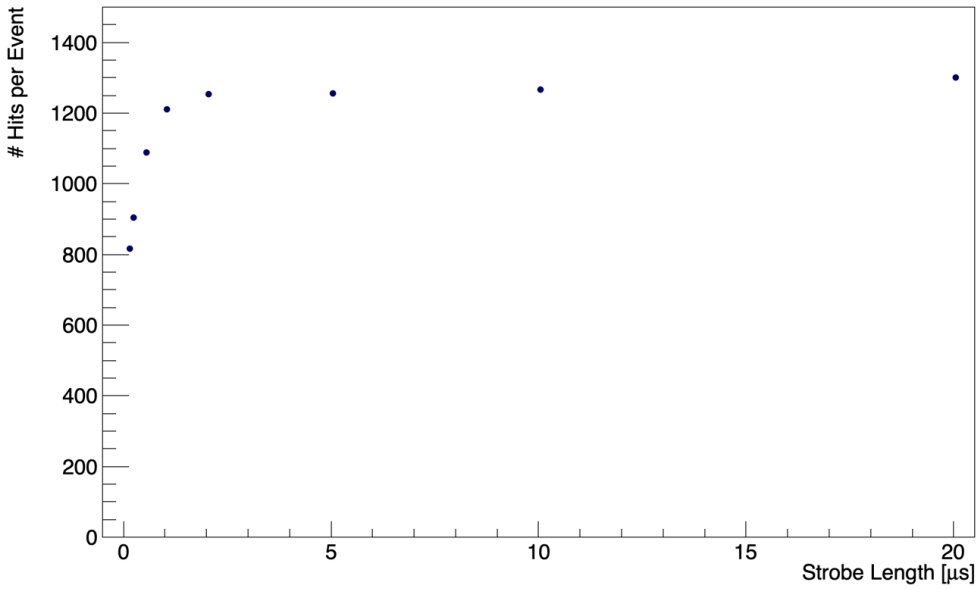


Figure 5.46: Number of hits per event integrated over all layers versus strobe length. This plot is based on the November runs listed in Table 5.4.

Figure 5.47 shows the relationship between the number of clusters per event integrated over all layers, and the number of tracks identified in the first layer, for different energies. The relationship seems approximately linear for the lower energies, especially for the first three points. The slope flattens slightly after three tracks. The graphs do not scale as expected, which is particularly noticeable for the 5 GeV and 5.8 GeV runs. The 5 GeV run seems to have a flattening slope from two tracks and has lower values than the 4 GeV run for four and five tracks. The run with an energy of 5.8 GeV appears to have an almost constant value of clusters per event versus the number of tracks and does not have a linear increase, as one might have expected.

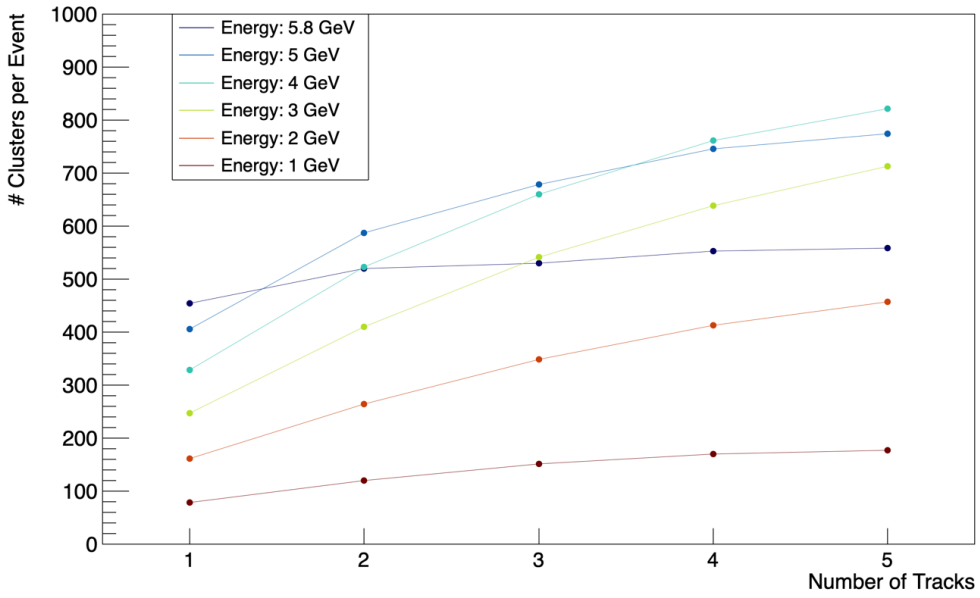


Figure 5.47: Number of clusters per event integrated over all layers versus number of tracks for different energies. All the runs used here have a temperature of 20°C and a strobe length of 2 μs. The runs are listed in Table 5.2.

There is, however, not a guarantee that the cluster identified in the first layer has a shower developing behind it. It could also be a noisy pixel. Hence, the additional cut mentioned above was applied to ensure correct identification of tracks, by disregarding events without a hit in the Layer 1 positioned directly behind the cluster in Layer 0. The result after applying this criterion is shown in Figure 5.48.

The relationship is more linear for certain cases, while others appear less linear. The relationship between the number of clusters per event and the number of tracks appears more linear for the 2 GeV and 3 GeV runs. The 1 GeV run looks approximately identical as before, with the exception of a lower number of clusters for five tracks. The 4 GeV run starts out more linear up to four tracks, but the deviation for five tracks is more considerable compared to in Figure 5.47. The graph for 5 GeV appears more linear up to three tracks, but has a significant deviation at five tracks. Lastly, the behaviour of the 5.8 GeV run appears even more irregular, and has values that are even lower than the 2 GeV run at five tracks.

Most of the runs have relatively low statistics for events with five tracks, especially 1 GeV and 5.8 GeV runs, having 49 and 82 events for five tracks, respectively. The 2 GeV and 3 GeV runs have a considerable statistical basis for all number of tracks.

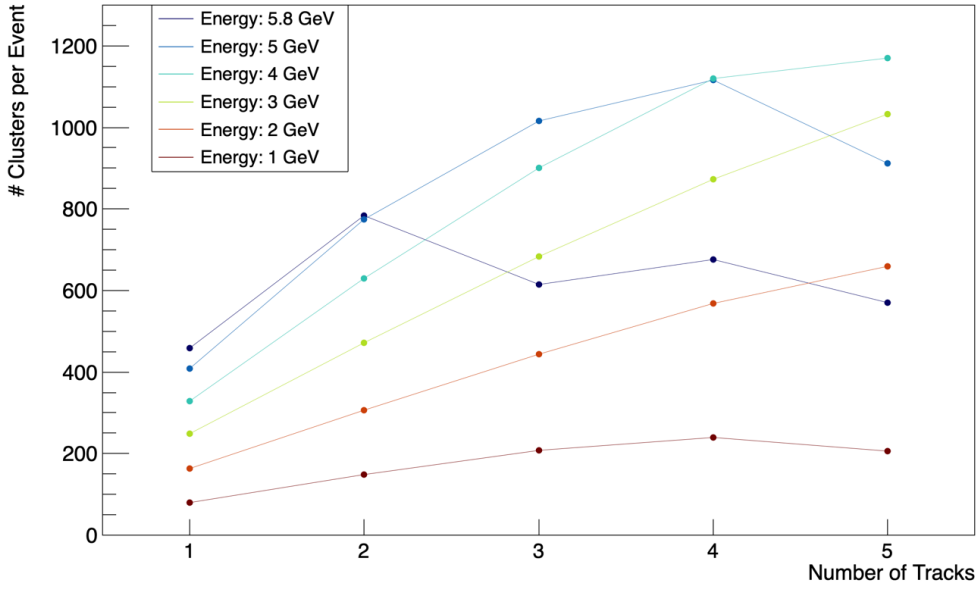


Figure 5.48: Number of clusters per event integrated over all layers versus the number of tracks for different energies, with cuts checking for hits in the second layer. All the runs used here have a temperature of 20°C and a strobe length of 2 μs. The runs are listed in Table 5.2.

A possible explanation of the effects observed in the plots above is that there could be a pile-up in the readout units. If this is the case, there would be no trace of it in the raw data. The trigger is recorded on the ALPIDE, and data is produced. However, data might be lost in the transition between the ALPIDE and the computer. This results in uncertainty as to whether all the frames are read out from the ALPIDE.

The bottleneck could stem from the FIFO memory on the readout units or the offload from the readout units to the computer, where the readout units are connected to the computer via a USB3 connection. Hence, the bottleneck might stem from the readout unit, and not the ALPIDE itself, but it is difficult to determine for certain if, or where, data is lost.

Another possible explanation might be related to the selection criteria for the number of tracks. If the event selection does not perform as intended, the assumption that 1-5 electron tracks start a shower could be wrong.

When studying the beam rates of the runs used in this analysis, the trigger rate seems to be highest for 3 GeV, with 7.5 kHz. The beam rate for 1 GeV and 2 GeV is 4.8 kHz and 4.5 kHz, respectively. For the higher energies, the beam rate is even lower, with 2.7 kHz for the 4 GeV run, 1.5 kHz for 5 GeV and between 0.12 and 0.25 kHz for the 5.8 GeV run. These values do not correspond with the plots above, as the higher trigger rate could cause problems with the scaling. However, the 3 GeV run with the highest trigger rate appears to be the most accurate case, while the 5.8 GeV run with the lowest trigger rate appears to be the least accurate.

Looking at Figures 5.47 and 5.48, the scaling for single tracks does not appear too off for all energies. 5.8 GeV might be a bit low, but as observed in the linearity plots for single tracks shown earlier, the 5.8 GeV run was, also here, slightly on the low side. For multiple tracks, the scaling seems to work for the 2 GeV and 3 GeV run. The other runs do, however, not seem to scale as expected, especially the 1 GeV and 5.8 GeV runs.

The naive expectation that the event selection is able to select for instance three electron tracks developing three showers does not work for all cases. It works in some cases, but in general it proved to correspond poorly with the 5.8 GeV and 1 GeV runs. As mentioned earlier, there are possible issues with the event selection criteria not being sufficient to identify tracks correctly. There might also be some firmware or software problems as the readout units used in this test beam are preliminary versions of the ITS firmware.

Chapter 6

Discussion and Conclusion

As mentioned in Chapter 2, the primary motivation for developing a pCT scanner is to increase the accuracy of the range estimations for the proton beam, so it meets the proton therapy's extensive demands for precision. The Bergen pCT scanner utilises a digital tracking calorimeter; a layer-by-layer pixel detector consisting of ALPIDE pixels sensors, that will operate both as a tracking system and as a range or energy detector.

This thesis aimed to evaluate the performance of the EPICAL-2 prototype for implementation in the DTC prototype for pCT purposes. The characterisation of this prototype is based on experiments conducted at the DESY II Test Beam Facility.

Based on the plots presented in Appendix C, the feature or hot spot observed in the beam profiles appears independent of both beam energy, layer number, beam position and strobe length. As the structure is present in all layers, it is also unlikely to be caused by the mechanical setup. Hence, the assumption is that it is caused by an external factor, such as the scintillators.

By looking at the spatial distributions for different events in Run 1335, the spread of particles in an event appears narrower than the showers, thus to the first-order most of the showers will overlap if multiple tracks are present in the same event. Although the showers may originate from any position in the layer, most of them occur in the central part of the calorimeter.

The six showers developing from the five clusters present in Layer 0 in Event 34495 of Run 1335 gives an argument for the inclusion of the "No Track" event selection. However, this effect may also influence the other event selections, as there may be more showers than indicated from the number of clusters present in the first layer.

One of the cons of the clustering algorithm used in this analysis is that merged and overlapping clusters are treated as a single cluster. This effect can consequently influence the cluster

studies for events with multiple overlapping showers.

However, by looking at the spatial distributions for the largest cluster sizes from the cluster size distribution, these clusters appear to be track-like structures. The tracks may be a result of the shower angle, or effects such as Compton scattering or particles moving in the epitaxial layer. The ratio between the number of large clusters and the total number of clusters in a layer also increases as the shower develops, indicating that the track-like clusters are a result of the shower angle.

Furthermore, by studying the normalised cluster size distribution, the difference between Layer 2 and the layers around shower maximum is minimal in the cluster size range where overlapping clusters are most prominent, thus indicating that the effect of overlapping clusters is small. The average cluster size for the different layers also indicates this effect.

The cluster size appears to have no systematic dependency on energy. There is, however, a systematic dependency on temperature, where the average cluster size increases slightly with increasing temperature. Although the temperature is not a crucial parameter, one can use this dependency to argue for the significance of maintaining a stable temperature during the data acquisition. The importance does not lie in what value the temperature measures, but merely that it be kept relatively constant. This effect can also support the argument for a water-cooling system for the pCT to better maintain a stable temperature during a scan.

When studying the readout frames with a single track, both the number of hits and the number of clusters per event have an approximately linear dependency on the beam energy. As the linear fit for the number of clusters per event has a lower χ^2-value and, thus, a better fit, it might be advantageous to use the number of clusters rather than the number of hits. The number of clusters also have a smoother distribution over the different layers of the prototype.

The linear relationship between the number of clusters in the first layer and the number of clusters in shower maximum for strobe lengths of 2 μs and 20 μs might suggest that a longer strobe length is the superior choice. The study performed on the strobe length's effect on the number of hits per event does, on the other hand, show a minimal difference in the number of hits per readout frame between the 2 μs and the 20 μs strobe length. In addition, the average cluster size versus layer indicates that noisy pixels make a larger contribution to the average cluster size in the later layers, especially for the 20 μs strobe length.

The plots displaying the number of clusters per event versus the number of tracks for different energies still pose certain unanswered questions, especially concerning the scaling of the graphs and the strange behaviour of the higher energy runs. It can be partly explained by low statistics for certain event selections, but there are yet some queries in need of clarification.

Then there is the issue of event selection. As discussed above, the fact that showers may develop without leaving a cluster in the first layer may influence the results produced by the event selection, seeing as there may appear more showers than what is indicated by Layer 0. There may also be noisy pixels misidentified as particle hits. A possible solution may be to filter out noisy pixels by requiring the clusters in Layer 0 to have a minimum cluster size of two pixels. The problem with MIP particles, however, is that they might create single-pixel clusters depending on their position of impact. So by applying a minimum cluster size criterion, there is a risk of filtering out a significant portion of data, in addition to filtering out the noise.

Another solution might be to check the second layer for hits in the area directly behind the hit in the first layer, but here one might end up filtering out too many events to do a proper analysis. However, much work is being done in this area by the EPICAL-2 analysis group, both on developing a suitable event selection algorithm, and on noise studies, figuring out how to deal with dead and noisy pixels.

This beam test is the first conducted on a purely digital calorimeter with a system consisting of 24 layers of ALPIDE chips and absorbers and, hence, an important first step. The results show that the concept works, in terms of being able to read out data, the amount of noise being reasonable and that showers from a single track develop with a linear response, as expected. The attempt to extend this to multiple tracks is, however, not fully understood. The multiplicity is sufficient and is approximately at the level expected for proton CT. The cluster size is, however, larger for protons.

Concerning the future outlook, stress tests should be performed on the system using higher rates and higher energies. The system should also be tested with a proton beam. Apart from this, an improvement should be made regarding the firmware, software and readout. Here, the final versions of both firmware and readout units should be used for both the EPICAL-2 prototype and the proton CT.

All in all, the EPICAL-2 prototype has demonstrated the feasibility of the ALPIDE sensors as the sensitive layers in an electromagnetic sampling calorimeter. The sensors function well within an electromagnetic shower environment, and shower properties can be successfully reconstructed. The data analysis shows that the ALPIDE chip can handle about 100 charged particles per event. All these results are promising regarding the performance of the proton CT system.

Bibliography

[1] Helge Egil S. Pettersen, Dieter Röhrich. “Kreftbehandling med protonterapi og proton-CT”. In: *Fra Fysikkens Verden* 4 (2017).

[2] *Innføring av protonbehandling til Norge*. Nov. 2020. URL: https://helse-bergen.no/om-oss/protonsenter-i-bergen/innforing-av-protonbehandling-til-norge.

[3] Dr John Gribbin. *Q Is for Quantum: an Encyclopedia of Particle Physics*. ReAnimus Press, 2014.

[4] Faiz M. Khan and John P. Gibbons. *Khan's the Physics of Radiation Therapy*. fifth. Philadelphia, PA: Lippincott Williams & Wilkins/Wolters Kluwer, 2014.

[5] William R. Leo. *Techniques for Nuclear and Particle Physics Experiments: A How-to Approach*. Berlin [etc.]: Springer-Verlag, 1994. DOI: 10.1007/978-3-642-57920-2.

[6] Thomas E. Johnson. *Introduction to Health Physics*. 5th ed. McGraw-Hill Education, 2017.

[7] Anton Lechner. “Particle interactions with matter”. In: *CERN Yellow Rep. School Proc.* 5 (2018). Ed. by Bernhard Holzer, p. 47. DOI: 10.23730/CYRSP-2018-005.47.

[8] Stefaan Tavernier. *Experimental Techniques in Nuclear and Particle Physics*. Berlin: Berlin: Springer Berlin / Heidelberg, 2010. DOI: 10.1007/978-3-642-00829-0.

[9] Lucio Cerrito. *Radiation and Detectors : Introduction to the Physics of Radiation and Detection Devices*. 1st ed. 2017. Cham: Springer International Publishing, 2017. DOI: 10.1007/978-3-319-53181-6.

[10] Claus Grupen and Boris Shwartz. *Particle Detectors*. 2nd ed. Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology. Cambridge University Press, 2008. DOI: 10.1017/CBO9780511534966.

[11] Harald Paganetti. *Proton Therapy Physics*. Boca Raton, Fla.: CRC Press, 2012.

[12] Peter Peschke and Günther H. Hartmann. “Molecular Aspects of Radiation Biology”. In: *Radiation Protection in Medical Physics*. Ed. by Yves Lemoigne and Alessandra Caner. Dordrecht: Springer Netherlands, 2011, pp. 17–23. DOI: 10.1007/978-94-007-0247-9_3.

[13] Jacques Balosso et al. “Nuclear Physics for Medicine”. In: Apr. 2014, pp. 11–56. ISBN: 978-2-36873-008-9.

[14] Eugene Surdutovich and Andrey V. Solov’Yov. “Nanoscale Insights into Ion-Beam Cancer Therapy”. In: 2017. Chap. Multiscale Physics of Ion-Beam Cancer Therapy, pp. 1–60. DOI: 10.1007/978-3-319-43030-0_1.

[15] Adeel Kaiser et al. “Proton Therapy Delivery and Its Clinical Application in Select Solid Tumor Malignancies”. In: *Journal of Visualized Experiments* 144 (2019). ISSN: 1940-087X. DOI: 10.3791/58372.

[16] Ana Arrendell. *Treatments and Services: Radiation Oncology*. Aug. 2020. URL: https://www.hopkinsmedicine.org/radiation_oncology/treatments/.

[17] Philip Mayles, Alan E. Nahum, and Jean-Claude Rosenwald. *Handbook of Radiotherapy Physics: Theory and Practice*. New York: Taylor & Francis, 2007.

[18] E. B. Podgorsak. *Radiation Oncology Physics: A Handbook for Teachers and Students*. International Atomic Energy Agency, 2005.

[19] Timothy D. Malouff et al. “Carbon Ion Therapy: A Modern Review of an Emerging Technology”. In: *Frontiers in Oncology* 10 (2020), p. 82. DOI: 10.3389/fonc.2020.00082. URL: https://www.frontiersin.org/article/10.3389/fonc.2020.00082.

[20] J.E. Tepper, R.L. Foote, and J.M. Michalski. *Gunderson & Tepper’s Clinical Radiation Oncology*. Elsevier, 2020.

[21] Nadine Barrie Smith and Andrew Webb. *Introduction to medical imaging: physics, engineering and clinical applications*. Cambridge texts in biomedical engineering. Cambridge: Cambridge University Press, 2011.

[22] A. Boss et al. “Hybrid PET/MRI of Intracranial Masses: Initial Experiences and Comparison to PET/CT”. In: *Journal of Nuclear Medicine* 51.8 (2010), pp. 1198–1205. ISSN: 0161-5505. DOI: 10.2967/jnumed.110.074773. URL: https://dx.doi.org/10.2967/jnumed.110.074773.

[23] Giacomo Luccichenti et al. “3D reconstruction techniques made easy: Know-how and pictures”. In: *European radiology* 15 (Nov. 2005), pp. 2146–56. DOI: 10.1007/s00330-005-2738-5.

[24] Johan Alme et al. “A High-Granularity Digital Tracking Calorimeter Optimized for Proton CT”. In: *Frontiers in Physics* 8.460 (2020). ISSN: 2296-424X. DOI: 10.3389/fphy.2020.568243. URL: https://www.frontiersin.org/article/10.3389/fphy.2020.568243.

[25] Glen Knoll. *Radiation Detection and Measurement (4th ed.)* John Wiley & Sons, Inc., 2010. ISBN: 978-0-470-13148-0.

[26] Leonardo Rossi et al. *Pixel Detectors: From Fundamentals to Applications*. Particle Acceleration and Detection. Berlin, Heidelberg: Berlin, Heidelberg: Springer Berlin / Heidelberg, 2006. DOI: 10.1007/3-540-28333-1.

[27] Lukasz Janusz Maczewski. “Measurements and simulations of MAPS (Monolithic Active Pixel Sensors) response to charged particles: A Study towards a vertex detector at the ILC”. PhD thesis. May 2010. arXiv: 1005.3710 [physics.ins-det].

[28] W. Snoeys. “CMOS monolithic active pixel sensors for high energy physics”. In: *Nuclear Instruments and Methods in Physics Research Section A Accelerators Spectrometers Detectors and Associated Equipment* 765 (2014), pp. 167–171. ISSN: 0168-9002. DOI: 10.1016/j.nima.2014.07.017. URL: https://dx.doi.org/10.1016/j.nima.2014.07.017.

[29] A.P. de Haas, et al. “The FoCal prototype — an extremely fine-grained electromagnetic calorimeter using CMOS pixel sensors”. In: *JINST 13 P01014* (2018).

[30] Monika Varga-Kofarago. “Anomalous Broadening of Jet-Peak Shapes in Pb-Pb Collisions and Characterization of Monolithic Active Pixel Sensors for the ALICE Inner Tracking System Upgrade”. Presented 19 Feb 2018. Oct. 2017. URL: http://cds.cern.ch/record/2310082.

[31] Giacomo Contin. “The MAPS-based ITS Upgrade for ALICE”. In: *PoS* Vertex2019 (2020), p. 003. DOI: 10.22323/1.373.0003. arXiv: 2001.03042 [physics.ins-det].

[32] Luciano Musa. *Letter of Intent for an ALICE ITS Upgrade in LS3*. Tech. rep. CERN-LHCC-2019-018. LHCC-I-034. Geneva: CERN, Dec. 2019. URL: https://cds.cern.ch/record/2703140.

[33] B Abelev et al. *Technical Design Report for the Upgrade of the ALICE Inner Tracking System*. Tech. rep. CERN-LHCC-2013-024. ALICE-TDR-017. Nov. 2013. DOI: 10.1088/0954-3899/41/8/087002. URL: https://cds.cern.ch/record/1625842.

[34] ALICE ITS ALPIDE development team. *ALPIDE Operations Manual*. DRAFT. Version 0.3. July 2016.

[35] Miljenko Suljic. “Study of Monolithic Active Pixel Sensors for the Upgrade of the ALICE Inner Tracking System”. Presented 02 Feb 2018. Nov. 2017. URL: https://cds.cern.ch/record/2303618.

[36] N van der Kolk. “FoCal: A highly granular digital calorimeter”. In: *Nucl. Instrum. Methods Phys. Res., A* 958 (2020), 162059. 4 p. DOI: 10.1016/j.nima.2019.04.013. URL: http://cds.cern.ch/record/2712384.

[37] R Barthel et al. “Performance of a digital calorimeter based on the ALPIDE pixel chip”. unpublished.

[38] R. Diener et al. “The DESY II test beam facility”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 922 (Apr. 2019), pp. 265–286. ISSN: 0168-9002. DOI: 10.1016/j.nima.2018.11.133. URL: http://dx.doi.org/10.1016/j.nima.2018.11.133.

[39] J. van Hoorn. “Study and Development of a novel Silicon Pixel Detector for the Upgrade of the ALICE Inner Tracking System”. PhD thesis. Technische Universität Wien, 2015. URL: http://cds.cern.ch/record/2119197/files/CERN-THESIS-2015-255.pdf?version=1.

Appendix A

Lane Number and Chip ID for the Layers in the EPICAL-2 Setup

Table A.1 gives an overview of the lane numbers and chip IDs corresponding to the different layers of the EPICAL-2 prototype. This information was used in the data analysis, and to construct the coordinate system for the spatial distributions.

Table A.1: Overview of the lane numbers and chip IDs corresponding to the different layers of the EPICAL-2 prototype.

Layer	Lane	Chip ID	Lane	Chip ID
0	32	20	35	21
1	59	45	56	44
2	34	22	33	23
3	57	47	58	46
4	36	18	37	19
5	61	43	60	42
6	38	16	55	17
7	79	41	62	40
8	54	14	53	15
9	77	39	78	38
10	52	12	51	13
11	75	37	76	36
12	50	10	49	11
13	73	35	74	34
14	48	8	47	9
15	71	33	72	32
16	46	6	45	7
17	69	31	70	30
18	44	4	43	5
19	67	29	68	28
20	42	2	41	3
21	65	27	66	26
22	40	0	39	1
23	63	25	64	24

Appendix B

Cluster Size Distribution

Figure B.1 shows the cluster size distributions in more detail, only including four layers and zoomed in on the first ten bins of the histogram. Two layers from the front of the stack are included, Layer 0 and Layer 2, one from shower maximum, Layer 5, and one from the back of the stack, Layer 22. The purpose of this plot is to display how the shape of the distribution develops and to see how the distribution looks in the area with most clusters.

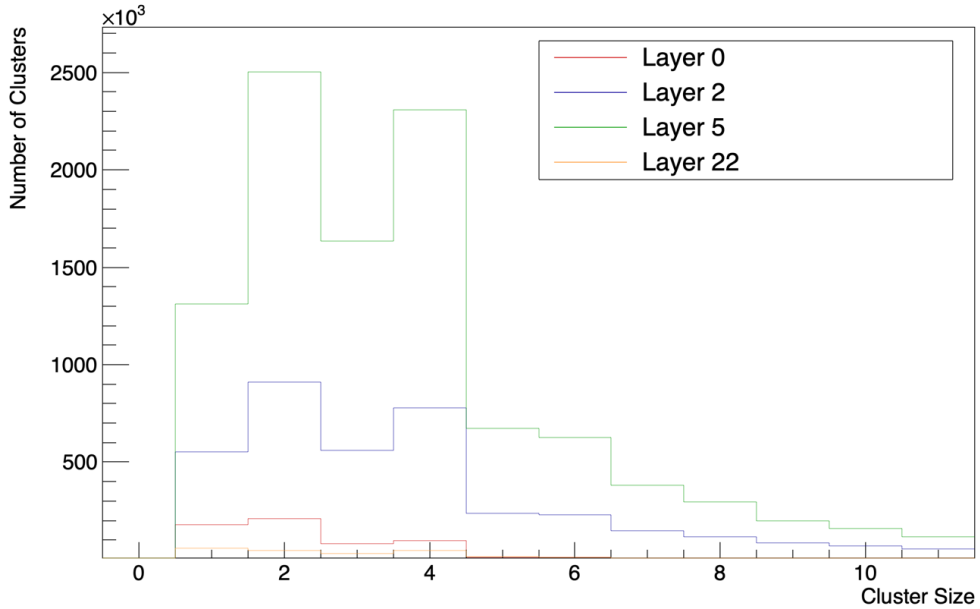


Figure B.1: Cluster size distribution for Layer 0, Layer 2, Layer 5 and Layer 22, zoomed in on the first ten bins of the histogram. **Run 1335**: energy of 3 GeV, temperature of 20°C and strobe length of 2 μs.

As seen in Figure B.1, the distributions seem quite similar for the different layers. There are peaks at cluster sizes two and four, for all layers displayed in the plot, except for Layer 22.

For Layers 0, 2 and 5 the largest peak is the one at a cluster size of two pixels. For Layer 22, the peak at a cluster size of four is about the same as for size two, but here there are more clusters of size one. Furthermore, in Layer 0, there are more clusters of size one than of size four. All the displayed layers have a dip between the two peaks at two and four.

Appendix C

The Feature

The runs listed in Table C.1 are used to study the feature present in the integrated spatial distributions. These data runs are of different energies, water temperatures, strobe lengths, as well as different stage positions and angles of the EPICAL-2 prototype.

Table C.1: Run condition table for electron beam runs at the DESY Test Beam in February 2020. The table gives an overview of beam energy, collimator settings, stage position, water temperature, number of events, strobe length and angle of the EPICAL-2 for the respective run numbers. Chip ID 27 is also excluded in all of the listed data runs.

Run Number	Beam Energy [GeV]	Lead Collimator [mm^2]	Internal Collimator [mm^2]	Stage Position x	Stage Position y	Water Temperature [°C]	Number of Recorded Events	Strobe Length [μs]	EPICAL-2 Angle [°]
1336	1	12×12	14×14	0.3	-448.4	20	594000	2	0
1335	3	12×12	14×14	0.3	-448.4	20	297000	2	0
1339	5	12×12	14×14	0.3	-448.4	20	747000	2	0
1358	3	12×12	14×14	0.3	-448.4	30	522000	2	0
1324	3	12×12	14×14	0.3	-448.4	20	278837	20	0
1280	3	12×12	14×14	-7.7	-441.2	20.1	297000	2	0
1447	3	12×12	14×14	0.3	-400.9	20	445500	2	10
1467	3	12×12	14×14	2.3	-448.4	20	297000	2	180

The first integrated spatial distributions displaying the distribution for each layer, depicted in Figure C.1, are for a run with an energy of 1 GeV. Other than that, it has a temperature 20°C, strobe length of 2 μs, stage position of $(0.3, -448.4)$ and an EPICAL-2 angle of 0°. This run is one of the runs used for analysis in Chapter 5.

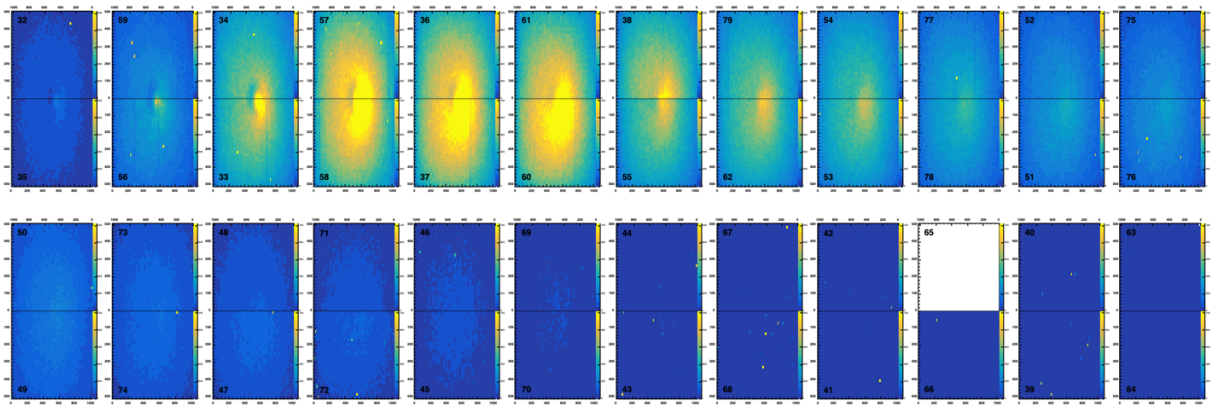


Figure C.1: Integrated spatial distributions, displaying the distribution for each layer separately. The first 12 layers are located in the upper part of the plot, while the 12 last layers are in the bottom. Chip ID 27 in Layer 21 has no registered hits as it was excluded from the data taking. **Run 1336**: energy of 1 GeV, temperature of 20°C, strobe length of 2 μs, stage position of (0.3, −448.4) and EPICAL-2 angle of 0°.

The next integrated spatial distribution, shown in Figure C.2, is for the run used for most of the analysis in Chapter 5. This run has an energy of 3 GeV, a temperature of 20°C, a strobe length of 2 μs, a stage position of (0.3, −448.4) and an EPICAL-2 angle of 0°.

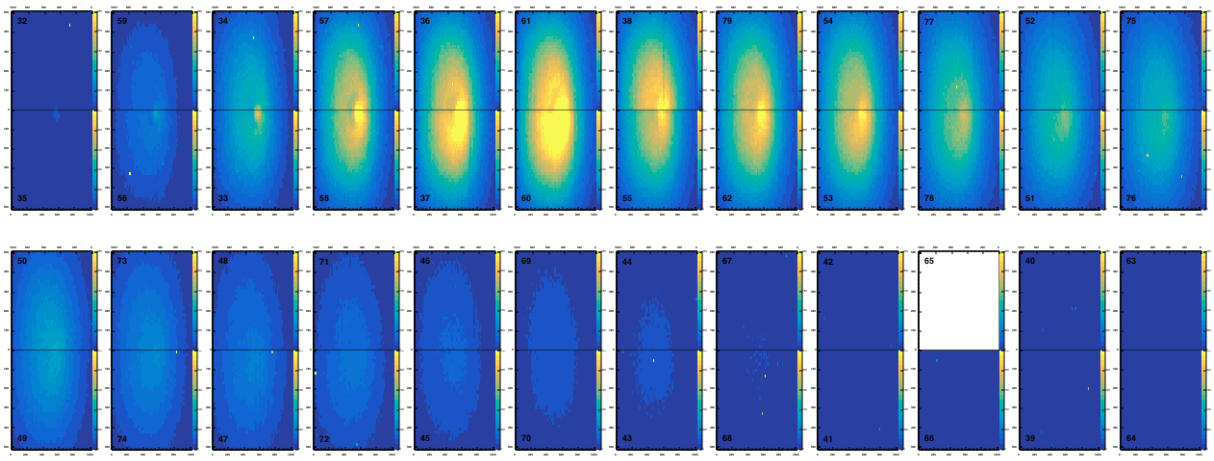


Figure C.2: Integrated spatial distributions, displaying the distribution for each layer separately. The first 12 layers are located in the upper part of the plot, while the 12 last layers are in the bottom. Chip ID 27 in Layer 21 has no registered hits as it was excluded from the data taking. **Run 1335**: energy of 3 GeV, temperature of 20°C, strobe length of 2 μs, stage position of (0.3, −448.4) and EPICAL-2 angle of 0°.

Figure C.3 gives the distributions for a run with an energy of 5 GeV, a temperature of 20°C, strobe length of 2 μs, stage position of (0.3, −448.4) and an EPICAL-2 angle of 0°.

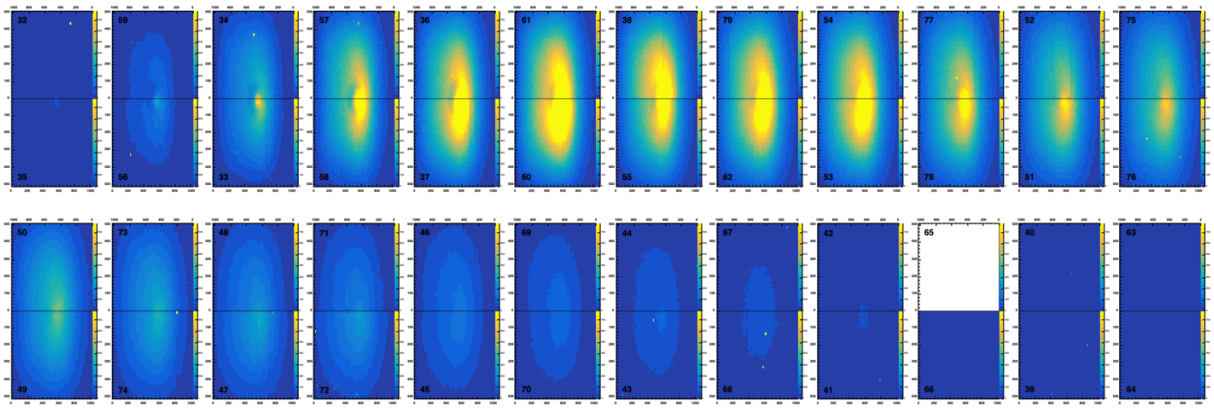


Figure C.3: Integrated spatial distributions, displaying the distribution for each layer separately. The first 12 layers are located in the upper part of the plot, while the 12 last layers are in the bottom. Chip ID 27 in Layer 21 has no registered hits as it was excluded from the data taking. **Run 1339**: energy of 5 GeV, temperature of 20°C, strobe length of 2 μs, stage position of $(0.3, -448.4)$ and EPICAL-2 angle of 0°.

Figure C.4 displays the distributions for a run with the same specifications as the run used in Figure C.2, but it has a temperature of 30°C.

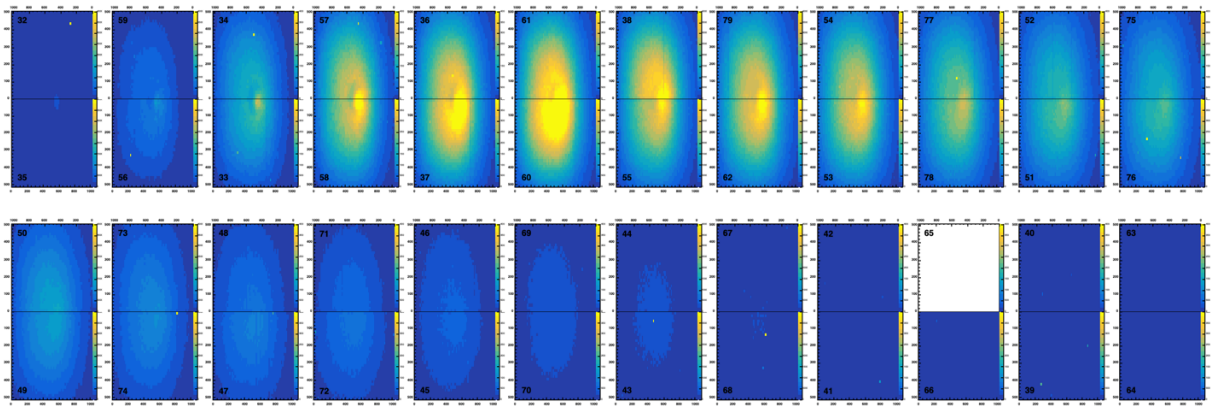


Figure C.4: Integrated spatial distributions, displaying the distribution for each layer separately. The first 12 layers are located in the upper part of the plot, while the 12 last layers are in the bottom. Chip ID 27 in Layer 21 has no registered hits as it was excluded from the data taking. **Run 1358**: energy of 3 GeV, temperature of 30°C, strobe length of 2 μs, stage position of $(0.3, -448.4)$ and EPICAL-2 angle of 0°.

The next run, displayed in Figure C.5, has the same specifications as the run used in Figure C.2, but it has a strobe length of 20 μs.

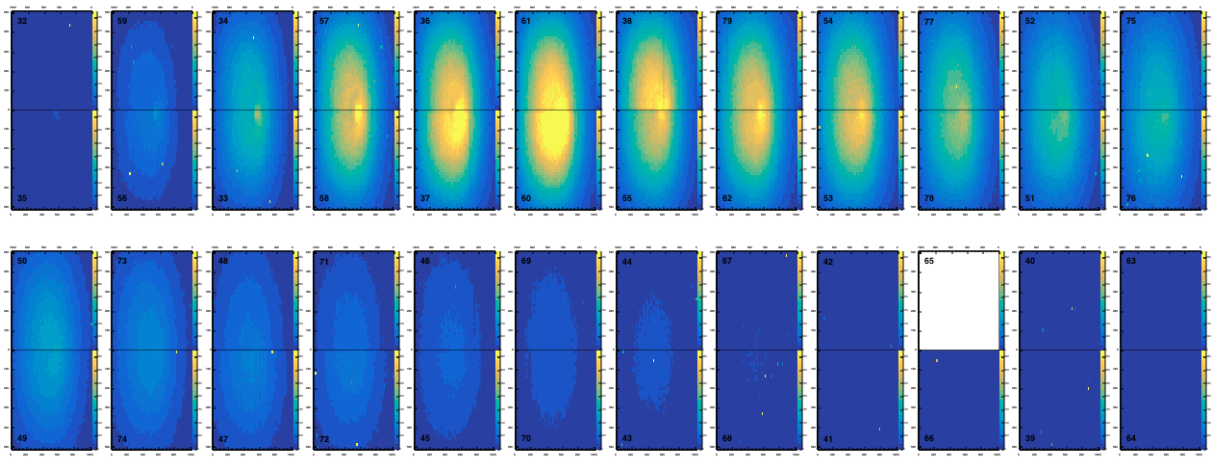


Figure C.5: Integrated spatial distributions, displaying the distribution for each layer separately. The first 12 layers are located in the upper part of the plot, while the 12 last layers are in the bottom. Chip ID 27 in Layer 21 has no registered hits as it was excluded from the data taking. **Run 1324**: energy of 3 GeV, temperature of 20°C, strobe length of 20 μs, stage position of $(0.3, -448.4)$ and EPICAL-2 angle of 0°.

Figure C.6 shows a run with a changed stage position of $(-7.7, -441.2)$, but otherwise the same settings as for Figure C.2.

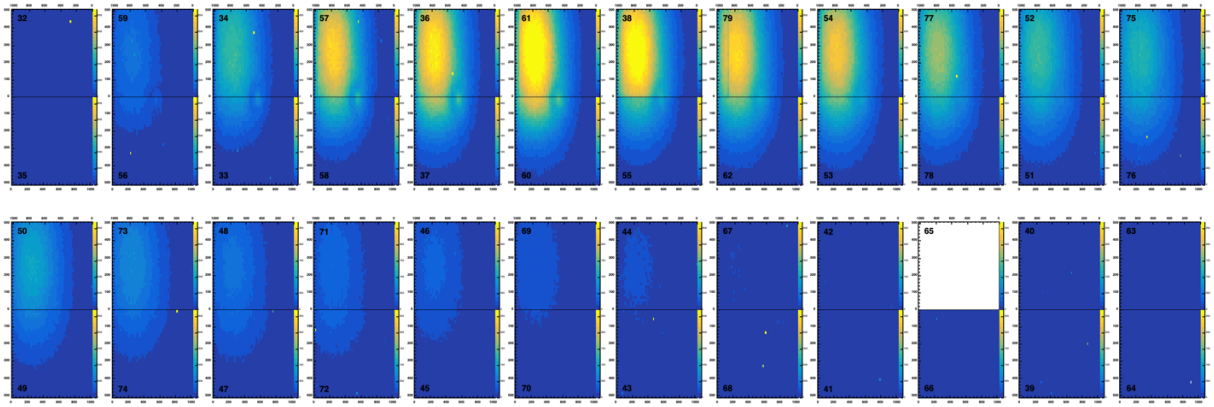


Figure C.6: Integrated spatial distributions, displaying the distribution for each layer separately. The first 12 layers are located in the upper part of the plot, while the 12 last layers are in the bottom. Chip ID 27 in Layer 21 has no registered hits as it was excluded from the data taking. **Run 1280**: energy of 3 GeV, temperature of 20.1°C, strobe length of 2 μs, stage position of $(-7.7, -441.2)$ and EPICAL-2 angle of 0°.

The next distribution, depicted in Figure C.7, is for a run with changed stage position in the y-direction, $(0.3, -400.9)$, and with the EPICAL-2 placed with an angle of 10° with respect to the beam axis.

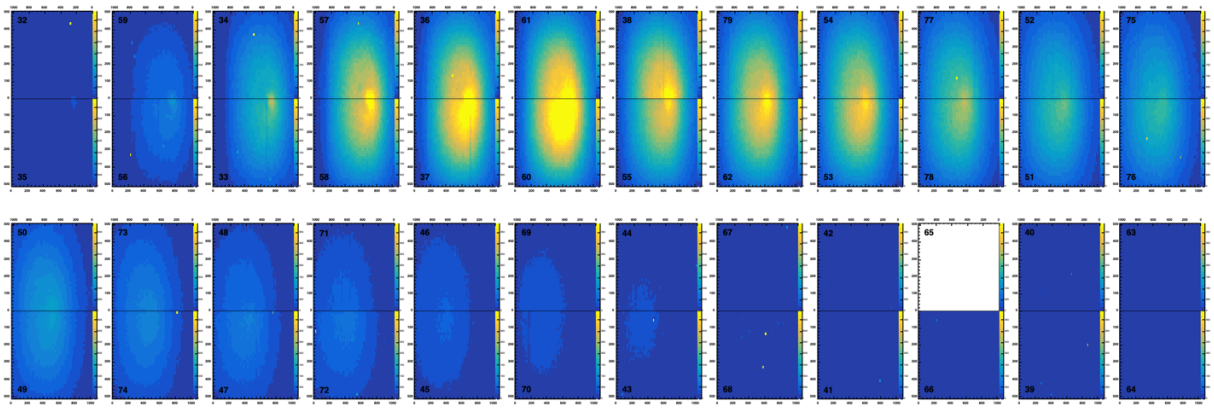


Figure C.7: Integrated spatial distributions, displaying the distribution for each layer separately. The first 12 layers are located in the upper part of the plot, while the 12 last layers are in the bottom. Chip ID 27 in Layer 21 has no registered hits as it was excluded from the data taking. **Run 1447**: energy of 3 GeV, temperature of 20°C, strobe length of 2 μs, stage position of $(0.3, -400.9)$ and EPICAL-2 angle of 10°.

Lastly, Figure C.8 displays a beam run where the stage position is changed in the x-direction, $(2.3, -448.4)$. The EPICAL-2 prototype is also placed so that the back-end faces the beam.

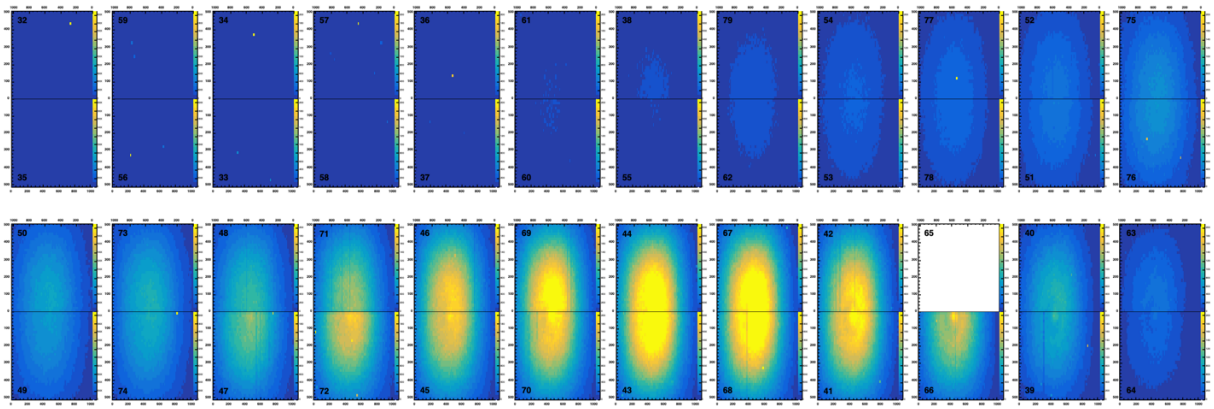


Figure C.8: Integrated spatial distributions, displaying the distribution for each layer separately. The first 12 layers are located in the upper part of the plot, while the 12 last layers are in the bottom. Chip ID 27 in Layer 21 has no registered hits as it was excluded from the data taking. **Run 1467**: energy of 3 GeV, temperature of 20°C, strobe length of 2 μs, stage position of $(2.3, -448.4)$ and EPICAL-2 angle of 180°.