Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

DESY.

Shower Shapes in a Highly Granular SiPM-on-Tile Analog Hadron Calorimeter

DISSERTATION

zur Erlangung des Doktorgrades

an der Fakultät für Mathematik, Informatik und Naturawissenschaften

Fachberiech Physik

der Universität Hamburg

vorgelegt von

OLIN LYOD PINTO

aus Mangaluru, Indien

Hamburg

2022

Gutachter/innen der Dissertation:	Prof. Dr. Erika Garutti Dr. Katja Krüger
Zusammensetzung der Prüfungskommission:	Prof. Dr. Robin Santra Prof. Dr. Erika Garutti Dr. Katja Krüger Prof. Dr. Ingrid-Maria Gregor Prof. Dr. Gudrid Moortgat-Pick
Vorsitzende/r der Prüfungskommission:	Prof. Dr. Robin Santra
Datum der Disputation:	31/08/2022
Vorsitzender Fach-Promotionsausschusses PHYSIK:	Prof. Dr. Wolfgang J. Parak
Leiter des Fachbereichs PHYSIK:	Prof. Dr. Günter H. W. Sigl
Dekan der Fakultät MIN:	Prof. Dr. Heinrich Graener

To My Family...

Abstract

This thesis concerns the development of a hadron calorimeter for future e^+e^- collider experiments such as ILC. The detectors at the ILC are optimized for the particle-flow approach, which aims to reach a jet energy resolution of 3-4% by measuring each particle in a jet using the best sub-detector measurement. This can only be achieved by using highly granular calorimeters. The **CA**lorimeter for **LI**near **C**ollider **E**xperiment (CALICE) collaboration develops different imaging calorimeters with high granularity. The work in this dissertation presents the analysis of shower shapes with data recorded using a CALICE **A**nalog **H**adron **CAL**orimeter (AHCAL) technological prototype. It is a sampling calorimeter made of layers of steel absorbers interleaved with $30 \times 30 \times 3$ mm^3 scintillating tiles with **Si**licon **P**hotoMultipliers (SiPM)-on-tile readout as active material. The prototype has been operated successfully at CERN in 2018 and acquired sizeable datasets using beams of muons, electrons and pions at different beam energies.

Firstly, a successful calibration with more than 99.9% channels of the AHCAL prototype and stable operation over several weeks of testbeam periods is presented. Secondly, `GEANT4` simulations are compared to testbeam data as a cross-check to detector calibration. In addition, the conventional calorimeter properties such as the detector response and single-particle energy resolution is evaluated. Using the obtained AHCAL data, this dissertation presents the first three-dimensional hadronic shower model describing the evolution of the average hadronic shower of a pion in the energy range between 10 and 200 GeV. Furthermore, this model holds potential implications for fast simulations and the **P**article **F**low **A**pproach (PFA). Finally, a comparison is done to validate and understand the picture of an average hadron shower shape by exploiting the Monte Carlo truth information.

Kurzfassung

Diese Doktorarbeit beschäftigt sich mit der Entwicklung eines Hadronkalorimeters für zukünftige e^+e^- Collider-Experimente wie dem ILC. Die Detektoren am ILC sind für den "Particle Flow"-Ansatz optimiert, der darauf abzielt, eine Energieauflösung von 3-4% für Teilchenjets zu erreichen, indem jedes Teilchen in einem Teilchenjet mit der besten Subdetektormessung rekonstruiert wird. Dies kann nur mit hochgranularen Kalorimetern erreicht werden. Die CAlorimeter for LInear Collider Experiment (CALICE) Kollaboration entwickelt verschiedene bildgebende Kalorimeter mit hoher Granularität. Diese Dissertation präsentiert eine Analyse der Form von hadronischen Schauern, mit Daten, die mit dem technologischen Prototyp des CALICE Analog Hadron CALorimeter (AHCAL) aufgezeichnet wurden. Es ist ein Sampling-Kalorimeter aus Schichten von Stahlabsorbern im Wechsel mit Lagen aus $30 \times 30 \times 3$ mm^3 großen Szintillationskacheln als aktivem Material, die durch Silicon Photo Multipliers (SiPM) direkt ausgelesen werden. Der Prototyp wurde 2018 erfolgreich am CERN betrieben und es wurden umfangreiche Datensätze von Myon, Elektron und Pion-Strahlen.

Zunächst wird eine erfolgreiche Kalibrierung von mehr als 99,9% der Kanäle des aufgezeichnet AHCALPrototyps sowie dessen stabiler Betrieb über mehrere Wochen Testbeam präsentiert. Als Nächstes werden `GEANT4`-Simulationen anhand von Teststrahldaten als Gegenprobe zur Detektorkalibrierung validiert. Dazu werden Parameter der Simulation, wie z.B. die Korrektur der SiPM-Sättigung, abgestimmt. Zusätzlich werden die klassischen Kalorimetereigenschaften wie Detektorantwort und Einzelteilchen-Energieauflösung ausgewertet. In dieser Dissertation wird, unter Verwendung der AHCAL-Daten, das erste dreidimensionale hadronische Schauermodell entwickelt, das die Entwicklung des durchschnittlichen hadronischen Schauers eines Pions im Energiebereich zwischen 10 und 200 GeV beschreibt. Es wird mit den Vorhersagen aus `GEANT4`-Simulationen verglichen. Darüber hinaus enthält dieses Modell mögliche Implikationen für schnelle Simulationen und den "Particle Flow"-Ansatz (PFA). Um das Bild einer durchschnittlichen Hadronschauerform zu validieren und zu verstehen, werden schlussendlich noch Details, die unter Verwendung der wahren Teilcheninformation in der Simulation verfügbar sind, verwendet.

Table of contents

Introduction

"Every act of seeing leads to consideration, consideration to reflection, reflection to combination, and thus it may be said that in every attentive look on nature we already theorize."

Johann Wolfgang von Goethe

What is the universe made of? Answering this question is the primary objective of particle physics. Understanding the structure of matter at the smallest accessible scale is the most exciting frontier in particle physics. The current answer is embedded in the **S**tandard **M**odel (SM) involving quarks, leptons, and corresponding anti-particles which interact by exchanging gauge bosons. The photon carries electromagnetism, the W and Z bosons carry the weak force, and the gluons carry the strong force. Moreover, in the background lies the Higgs Boson.

Despite the Standard Model describing all known matter particles and their interactions, it can not be considered an universal law. Nevertheless, the SM has succeeded in explaining a small fraction of the matter around us. Modern particle physics experiments measure the results of high energy particle collisions. At the time of writing, the highest-energy particle accelerator is the **L**arge **H**adron **C**ollider (LHC) at CERN, colliding protons with center of mass energies up to $\sqrt{s} = 14$ TeV. The Higgs Boson discovery at two independent experiments at the LHC in 2012 completed the SM of particle physics. It marked a cornerstone in the research of fundamental particles and their interactions.

Complementary to the physics program at the LHC, proposed lepton colliders like the **I**nternational **L**inear **C**ollider (ILC) open the possibility of precision measurements of the Higgs Boson by searches in final states that are not measurable at the LHC. These future linear electron-positron colliders require a 3 – 4 % jet energy resolution for a wide range of jet energies. This can be achieved by using the PFA for jet reconstruction. Within the CALICE collaboration, several concepts for high-granularity **H**adronic **Cal**orimeter (HCAL) optimised for Particle Flow have been studied and tested.

A key challenge for HCAL is the nature of complex physics processes and their variations within hadronic showers. Highly granular calorimeters offer the opportunity for a detailed spatial resolution of individual particle showers, getting an insight into the physics processes happening within a given shower. In addition, understanding hadronic shower development

and parametrization of the energy density distribution is essential for estimating leakage from calorimeters, validating hadronic shower models in simulations, and improving particle flow algorithms.

This thesis is devoted to the analysis of 2018 testbeam data collected at the CERN-SPS using the CALICE highly granular analog hadron calorimeter. This is a scintillator-steel sampling calorimeter with 3×3 cm^2 tiles, read out using SiPM.

This thesis is structured as follows:

Chapter 1 gives a basic understanding of the SM and briefly discusses the Higgs mechanism as the origin of mass and the top quark. Further, it outlines some of the limitations in the SM with a particular focus on the need for future colliders. It is followed by a short description of the ILC, including its two detectors, the **I**nternational **L**arge **D**etector (ILD) and the **Si**licon **D**etector (SID), based on the Particle Flow concept. Next, the basics of particle interactions and the principles of calorimetry are discussed. Furthermore, a brief overview of the different generations of calorimeter prototypes built within the CALICE collaboration are discussed.

In Chapter 2, before discussing the highly granular AHCAL prototype, the key parameters of SiPM are addressed. Furthermore, this chapter explains the simulation of the AHCAL prototype and the detector effects in the simulation.

Chapter 3 provides the experimental setup and the data recorded during the test beam campaign at CERN **S**uper **P**roton **S**ynchrotron (SPS) in 2018 with different beam energies and particle types.

The calibration of all AHCAL channels performed in order to properly extract the energy and time measurement of particles for the AHCAL detector is described, and the results are reported in chapter 4.

The recorded datasets are contaminated by different particles and have to be separated into samples containing only shower data from one particle type. The algorithm to determine the position of the first hadronic interaction and the event selection criteria applied to improve the purity and the efficiency in the recorded beam data is described in Chapter 5.

Before analysing pion data, the simulations require prior validation with measurements. This validation is an integral part of the CALICE efforts. Chapter 6 evaluates the performance, verifies detector calibration, and validates the detector and simulations by comparing physics models to data using several calorimeter variables.

Once the main parameters of the simulations have been tuned and the agreement between data and simulation has been evaluated, it is then possible to exploit the hadronic showers in detail due to the advantage of a highly granular calorimeter providing detailed information. Chapter 7 discusses the analysis of pion shower shapes in the longitudinal and radial directions with the beam energy ranging from 10 to 200 GeV. Furthermore, this thesis provides a three-dimensional shower model with a description for an average hadronic shower. In addition, fit functions are compared to the truth information provided by `GEANT4`.

The thesis will conclude with a summary and an outlook in chapter 8, providing potential application of this vibrant area of research and development in hadronic shower models for the particle physics community.

Part I

Physics and Calorimetry

Chapter 1

Particle Physics

"Not only is the Universe stranger than we think, it is stranger than we can think."

Werner Heisenberg

PARTICLE PHYSICS intends to learn more about the fundamental forces and particles responsible for the structure of the universe. The origin of the study of particle physics dates back to the discovery of the electron in 1897 [1] and the Rutherford scattering experiment in 1909 [2]. They led to the foundation for the concept of atoms and the forces present in atoms. Subsequent studies and experiments aided us in understanding the particles present and the interacting forces. As the experimental energies increased, the study of fundamental particles became more prominent. Initially, in the 1970s, the attempt to understand the subject was done by understanding the fundamental particles and their interaction. However, it still does not include gravitational force as one among the fundamental forces.

This chapter is divided into four parts. The first part of the chapter gives a short overview of the SM of particle physics and its limitations. Then, one of the possible future high energy e^+e^- experiment and the detector concept of ILC is presented. Afterwards, the general interactions of high energy particles with matter, typically resulting in particle showers, and the simulation of such interactions is discussed. Finally, the CALICE detector concepts are introduced.

1.1 The Standard Model of Particle Physics

The SM of particle physics was formulated in the 1960s [3–5]. Over time and through many experiments, the SM has become established as a well-tested physics theory. An overview of the fundamental particles and their interactions as described by the SM is shown in figure 1.1.

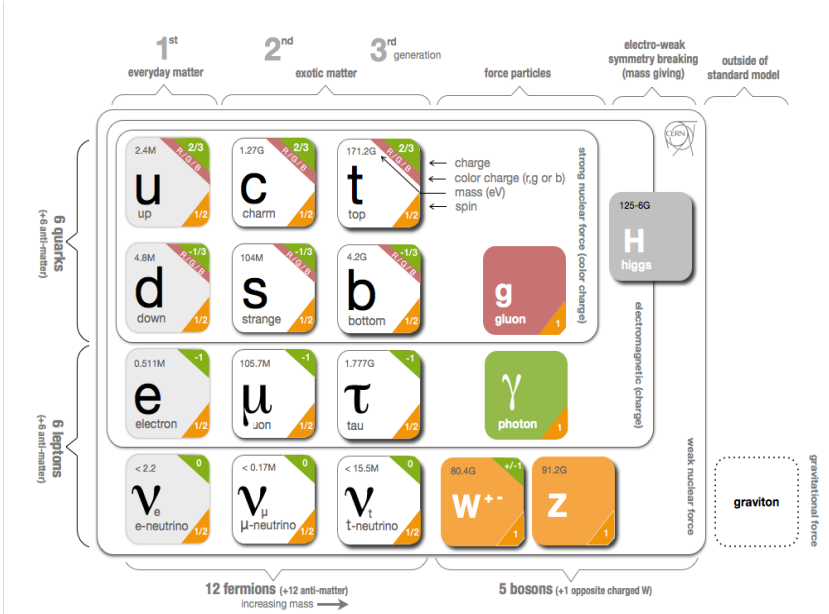


FIGURE 1.1: The Standard Model of particle physics, consisting of the twelve fundamental fermions, four fundamental gauge bosons and the Higgs Boson [6].

The theory classifies all observed elementary particles and describes forces acting between those particles via the electromagnetic force, the weak force, and the strong force. The SM comprises twelve fundamental particles of six leptons and six quarks, also known as fermions that consist of spin-$1/2$ particles[1] and the interaction particle gauge boson with integer spin. Fermions are responsible for all the visible matter around us are classified into three generations, with an increase in mass found with higher generations. The description of gravity is not part of the SM, as it can be neglected due to its weakness on the microscopic scale.

In the SM, leptons are classified as electron, muon and tau with charge -1 and corresponding anti-particles with similar properties but with charge $+1$, and these are associated with their neutral leptons, also called neutrinos, which have negligible mass and zero charge and are therefore hard to detect in particle collider experiments. On the other hand, quark composition gives rise to two or more different hadrons, which was studied under the extended hadronic sector [7], formulating a different theory of **Q**uantum **C**hromo**D**ynamics (QCD). Quarks and gluons carry a colour charge. One with two quark (q,$\bar{\mathrm{q}}$) compositions is called a meson, and one with three quark (qqq) compositions is called baryons. Tetraquarks ($\mathrm{q\bar{q}q\bar{q}}$) and pentaquarks ($\mathrm{qqq\bar{q}q}$) have been observed only recently [8, 9]. The strong force describes the interaction between the quarks, and massless gluons are the force mediators of QCD.

The electromagnetic force is responsible for the interaction of charged particles, and the photons are its massless mediators. Finally, the weak force is mediated by the massive W and Z bosons. Due to the large mass of these bosons, the weak interaction is short-ranged. W boson

[1]The description of the corresponding antiparticle counterparts, with equal mass but opposite quantum numbers, including electric charge, is implied.

carries ± 1 charge while Z boson is electrically neutral with mass of 80.379$\pm$0.012 GeV and 91.1876$\pm$ 0.0021 GeV [10] respectively.

The mathematical formulation of the SM is represented in the form of gauge groups. The gauge structure of the SM is given as $\mathrm{SU(3)_C} \otimes \mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}$. The indices represent the charges of the groups, respectively. The $\mathrm{SU(3)_C}$ represents the gauge group that belong to the strong interaction, mediated by eight massless gluons. Only the quarks and the gluons carry colour charge and participate in the strong interaction. The strong force binds the quarks into nucleons and nucleons into atomic nuclei. $\mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}$ represents the gauge group for the electroweak interaction, which consists of four gauge bosons Z^0, $W^{\pm}$ and photon. Among which triplets correspond to $\mathrm{SU(2)_L}$ and one neutral corresponds to $\mathrm{U(1)_Y}$.

The properties of the Higgs field and Higgs Boson, responsible for mass, are generated through electroweak symmetry breaking. Particle mass generated by Higgs field via Higgs Boson is directly proportional to the strength of the interaction [11–13]. It was proposed in 1964 together with a new spin-zero boson, the Higgs Boson, produced by the quantum excitation of this field. In 2012, the Higgs Boson was finally discovered at the LHC [14, 15][2], by then the last missing piece of the SM, which marks the tremendous success of the SM in describing the interactions of sub-atomic particles.

Brout-Englert-Higgs Mechanism

In quantum field theory, higher-order effects cause divergence leading to infinite mass and charge. Hence to control this divergence, we need a renormalisable theory, and the basic requirement is local-gauge invariance [16]. This principle is fulfilled under the condition that gauge bosons are massless, which agrees with photons and bosons, but $W^{\pm}$ and Z^{-} bosons are heavy particles. Thus, in principle, it violates renormalisation.

Via the process of spontaneous symmetry breaking of the Electroweak Symmetry, the gauge bosons can acquire mass. An additional field is introduced, called the Higgs field, to maintain renormalisation for massive gauge bosons. The process in which electroweak interaction is broken and electromagnetic interaction is retained (this interaction is due to the massless photon for infinite range or strong interaction) is called **B**rout-**E**nglert-**H**iggs (BEH) mechanism. According to the BEH mechanism, a nonzero vacuum value that couples with gauge interaction are introduced to the Higgs field, which gives rise to the mass of all the SM particles. Fermion masses are generated via Yukawa couplings to the Higgs field; hence the coupling experienced is directly proportional to the particle mass. Doublet of the complex field in the Higgs mechanism is,

[2] A proton-proton collider, the abbreviation is derived from the name of the council mandated with the foundation of the organization, **C**onseil **E**uropéen pour la **R**echerche **N**ucléai in Geneva, Switzerland.

$$\Phi = \begin{pmatrix} \Phi^+ \\ \Phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \Phi_2 + i\Phi_1 \\ \Phi_4 + i\Phi_3 \end{pmatrix}. \tag{1.1}$$

This consists of real field $\Phi_i (i = 1 \ldots 4)$, which is called as the Higgs field following the Lagrangian equation

$$\mathcal{L}_{\text{Higgs}} = \partial_\mu \Phi^\dagger \partial^\mu \Phi - V\left(\Phi^\dagger \Phi\right),$$

with a covariant derivative ∂^μ and with a potential of the form

$$\begin{aligned} V\left(\Phi^\dagger \Phi\right) &= \mu^2 \Phi^\dagger \Phi + \lambda \left(\Phi^\dagger \Phi\right)^2 \\ &= \mu^2 |\Phi|^2 + \lambda |\Phi|^4, \quad \lambda, \mu \in \mathbb{R}^+. \end{aligned} \tag{1.2}$$

This potential has a non-zero ground state; thus, the field is maintained even in the lowest energy state. The minimum potential value is also known as the vacuum expectation value of the Higgs field, which is ≈246 GeV. λ is the Higgs self-coupling that has a positive value in the SM. The SM Higgs Boson (H) is the fundamental excitation of the Higgs field manifesting as a spin 0 particle with m_H = (125.7 ± 0.4) GeV [10], which should decay into all SM particles in fractions proportional to their squared Higgs couplings. The exact measurement of the Higgs coupling strengths and the total Higgs decay width are essential cornerstones of the physics programme of current and future particle physics experiments.

The current SM consists of 27 parameters (if neutrino mass and cosmological constants are included) comprising of:

- 10 parameters each in the lepton and quark sector (six masses, three angles and one complex phase)
- three gauge couplings
- the QCD vacuum angle
- the Higgs vaccum expectation value and mass,
- and the cosmological constant

Limitations of the Standard Model

Although the SM is an astounding success and tested with a wide range of experimental results from particle colliders and other particle physics experiments. However, there are some unanswered questions that the SM cannot explain. They are briefly highlighted in the following:

Grand unification: According to **G**rand **U**nified **T**heorie**s** (GUTs) [17], all forces are expected to couple with the same strength at high energies. However, the unification of all three fundamental forces is not feasible within the SM.

Gravity: The SM does not include gravity, the interaction driving the evolution of the universe. Though theorists have proposed some attempts to unify the SM and general relativity, they do

not provide enough prediction power to be experimentally validated.

Dark matter: The SM lacks to explain invisible matter known as dark matter, making ∼85% of the matter in the universe. However, there have been some indications from galaxies observations [18] and the Cosmological Microwave Background (CMB) [19]. Nevertheless, so far, a dark matter candidate has not yet experimentally been observed through direct, indirect or collider searches. If it exists, then it should be massive and interact weakly.

Neutrino masses: According to neutrino flavour oscillations experiments [20], neutrinos carry a mass roughly six orders of magnitude smaller than the masses of the other SM fermions. However, the SM predicts that neutrinos cannot acquire masses.

Electroweak hierarchy problem: The large hierarchy between the weak scale of 10^2 GeV to the Planck scale of 10^{19} GeV is not understood. The hierarchy problem of the SM stems from our inability to answer the question, why the Higgs Boson mass is at a very peculiar value?. In the SM, Higgs Boson mass has many contributions from loop corrections. Considering that the SM is supposed to be valid up to the Planck scale, then these corrections would lead to a higher mass for Higgs Boson than experimentally measured [21]. For Higgs Boson mass to be at the observed value, a "fine-tuning" is one of the explanations which many theorists consider "unnatural". Also there are other hierarchy problem such as why the $\mathrm{m_u}/\mathrm{m_t} < 10^{-4}$?.

Matter-antimatter asymmetry: The SM predicts that matter and anti-matter were created in equal amounts during the Big Bang. However, our universe is composed mainly of matter. This baryon asymmetry is not explained, but Charge Parity (CP) violation (charge conjugation parity symmetry) [22] may answer this asymmetry. Although it is present in the SM, it has only been observed for weak interactions in the quark sector and is not enough to explain the asymmetry.

In addition to the problems mentioned above, several others exist, such as the significant difference between the predicted and observed muon anomalous magnetic moment, the form of the Higgs potential. In addition, effects related to cosmology, like the origin of Dark Energy, the driving force behind the universe's expansion ("inflation"), and the incorporation of quantum gravitational effects, are also eagerly awaiting an answer that the SM cannot provide.

There are theoretical solutions to many of the above-mentioned problems, e.g. Technicolour [23] and Supersymmetry [24] and no compelling experimental indication none have yet been confirmed. This indicates that SM is a work in progress and will have to be extended to the physics Beyond Standard Model (BSM). An experimental approach to solving these open questions is the construction of novel particle colliders and detectors which can probe the fundamental properties of the elementary particles and their interactions with higher precision than currently available.

1.2 The International Linear Collider

Around the TeV scale, new physics discoveries are predicted [25]. The LHC is actively investigating this region with a centre-of-mass energy of up to 14 TeV. Nonetheless, because hadrons are composite states, the LHC suffers from large backgrounds. The hadron machine has identified possible new physics signals, as there is no "real evidence" for BSM at LHC yet. Therefore, in addition to the precision measurements to SM physics, complementary precision measurements at the TeV scale are required to identify whether new physics has occurred. A lepton collider can accomplish this, as it provides a cleaner environment and well-defined initial states. The ILC is the most developed concept planned for the linear electron-positron collider foreseen to be built in Kitakami, Japan. ILC offers tunable centre-of-mass energy. In the first phase, it will operate at centre-of-mass energy energies ranging from 200-500 GeV and with an option to be upgraded to 1 TeV. In addition, the design foresees also running at the 91 GeV Z-pole for precision electroweak measurements.

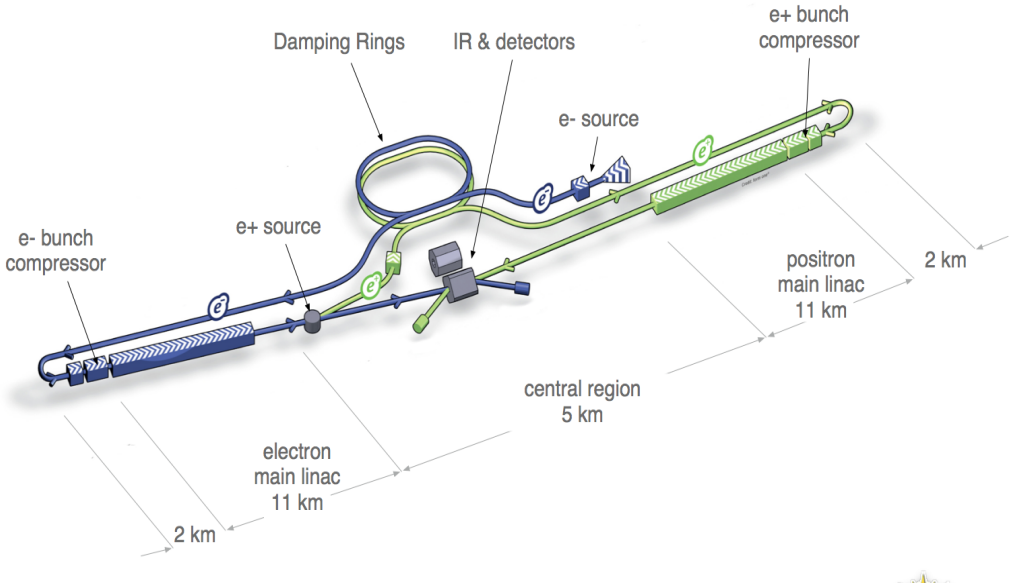


FIGURE 1.2: Schematic layout of the International Linear Collider in the 250 GeV staged configuration, taken from [26]

Figure 1.2, shows a layout of the ILC accelerator with the major subsystems, and its total length is 31 km. The main linear accelerator is based on 1.3 GHz **S**uperconducting **R**adio **F**requency (SCRF) cavities with an average accelerating gradient of $\geq$31.5 MV/m. It is expected to yield an integrated luminosity of $\int \mathrm{Ldt} = 500\ \mathrm{fb}^{-1}$ in the first phase of operation with $\sqrt{s} = 250$ GeV.

The advantage of the ILC is that it can provide beams with longitudinal polarization. By steering the polarization of the incoming beams, it is possible to enhance or suppress specific signals and backgrounds and thus helping precision tests on the SM as well as the search for new particles [27]. The achievable mean polarization for the electron and positron beam is 80% and 30% respectively, as foreseen in the ILC Technical Design Report [28]. The beam structure of

the ILC consists of long bunch-trains separated by around 200 ms in a low repetition rate of 5 Hz. Each bunch-train is separated by 554 ns and contains 1312 bunch crossings with a population of 2×10^{10} particles per bunch.

All key concepts and technologies needed to build the ILC have been demonstrated and documented in the ILC Accelerator Technical Design Report [26].

1.2.1 Physics Program

The ambitious physics program at ILC focuses on precision measurements. After discovering the Higgs Boson, measuring its properties with high precision has become a special interest in particle physics. As the Higgs Boson interacts with all massive particles of the SM, it is likely to be a key to many open questions in particle physics. In addition, measuring at high precision can lead to hints of new physics beyond the SM. A detailed summary of the physics cases at the ILC can be found in [27].

Higgs Physics at the ILC

At the ILC, three dominant techniques are available to determine the characteristics of the Higgs Boson. For center-of-mass energies below 350 GeV the *Higgsstrahlung* ($e^+e^- \rightarrow \mathrm{ZH}$), where the Higgs Boson is produced in association with a Z Boson. For higher energies, the **V**ector **B**oson **F**usion (VBF) process with W fusion ($e^+e^- \rightarrow \nu\bar{\nu}\mathrm{H}$) and Z fusion ($e^+e^- \rightarrow e^+e^-\mathrm{H}$), the Higgs Boson is produced in association with either a neutrino-antineutrino pair or an electron-positron pair. The cross-section of these production processes at the ILC, as well as Feynman diagrams of these processes, are shown in figure 1.3.

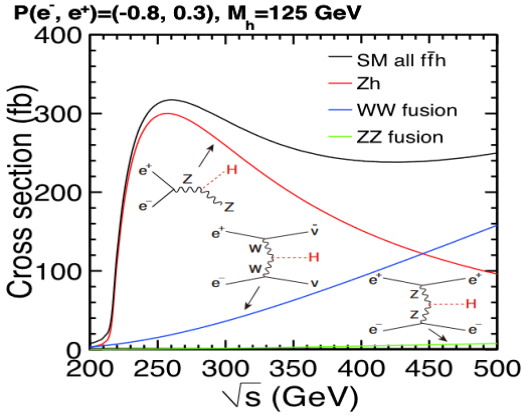


FIGURE 1.3: Standard Model cross-section of the dominant Higgs production processes at the International Linear Collider, as a function of the center-of-mass energy. The polarization is assumed to be 80% for the e^- beam and 30% for the e^+ beam [29].

The Higgsstrahlung production mechanism offers a unique method to reconstruct the mass of Higgs Boson without explicitly measuring the Higgs decay at all. Instead, it aims at reconstructing the Z Boson recoiling against the Higgs Boson.

By using the charged leptonic decay of the Z boson (into e^+e^--pairs or $\mu^+\mu^-$-pairs) provides the highest precision for the reconstruction. Higgs recoil mass, M_{recoil} recoiling against the lepton pair is calculated by [30]:

$$M^2_{recoil} = (p_{CM} - (p_{l^+} + p_{l^-}))^2, \tag{1.3}$$

where, p_{CM} is the four momentum of the initial colliding particles, and $p_{l^\pm}$ are the four-momenta of the two leptons from the Z decay.

This technique provides a model-independent measurement of the Higgs mass as no assumptions are made on the couplings and decays of the Higgs. The approach of the recoil mass measurements additionally provides the Higgs couplings and the Higgs full decay width with an unprecedented precision [31]. Following the same approach, it also offers the possibility to measure directly invisible decays of the Higgs Boson, which opens the door to searches for new physics beyond the SM.

Top Quark

The top quark is the heaviest known elementary particle in the SM with the strongest coupling to the Higgs Boson (called *top Yukawa coupling*) and giving a mass of $\mathrm{m_t} = 172.76 \pm 0.30$ GeV [10] and having a short lifetime of 10^{-25}s. Because it couples strongly, the top mass plays a crucial role in understanding and determining the Yukawa and electroweak coupling constants, which have a great impact on the Higgs sector. In addition, the determination of top properties is also crucial for new physics searches. The top quark gives loop contributions to many precision measurements sensitive to BSM effects.

The measurement of Higgs self-coupling, which has not been observed by LHC but would be possible at ILC with $\sqrt{\mathrm{s}} = 500$ GeV with an integrated luminosity of 2500 fb^{-1}. The production of two Higgs Bosons becomes accessible via the $\mathrm{e^+e^-} \rightarrow \mathrm{ZHH}$ process allowing to measure the self-coupling of the Higgs Boson. This holds large potential for deviations from the SM prediction showing the way to new physics. At $\sqrt{\mathrm{s}} \geq 500$ GeV the coupling of the top quark to the Higgs Boson can be measured via the $e^+e^- \rightarrow \mathrm{t\bar{t}H}$ process. The ILC offers the possibility to scan the threshold region of the top quark mass, allowing for precise measurements of the top quark mass and also to study the left-right asymmetry and forward-backward asymmetry of $\mathrm{t\bar{t}H}$ production.

The ILC expects to measure the top quark mass with a statistical uncertainty of 11% at $\sqrt{\mathrm{s}} = 500$ GeV. This precision can be improved to 6.4% with a higher integrated luminosity of 4000 fb^{-1}. Moreover, a precision of 4-5% is expected at $\sqrt{\mathrm{s}} = 1$ TeV [32].

1.2.2 Concept of Particle Flow

At high-energy physics experiments, colliders typically measure hadron jets rather than single particles. These hadron jets carry a significant fraction of energy. On average a particle jet

comprises, ∼62% of energy carried by charged particles, ∼27% by photons, ∼10% by long-lived neutral hadrons (e.g neutrons and K_L^0) and ∼1.5% by neutrinos [33].

The anticipated precision physics program at ILC drives the requirements for the detectors. Thus, a precise jet reconstruction is essential. Several studies suggest that a jet energy resolution of about 3% is needed to fully exploit the power of the collider. Such a resolution requires an improvement of detector performance. The concept of particle flow is currently believed to be the only practical approach which can reach this level of precision.

In traditional calorimeters, the total energy is reconstructed by summing up the energy deposits in the **E**lectromagnetic **Cal**orimeter (ECAL) and HCAL as illustrated in figure 1.4(left). Unfortunately, this leads to a relatively poor energy resolution of hadronic calorimeter in the order of $60\%/\sqrt{\mathrm{E(GeV)}}$ due to the strong fluctuations in hadron showers. To overcome this issue, a modern approach using PFA for jet energy reconstruction was made possible by the significant development in tracking detectors and high granularity calorimeters over the last decades. Unlike traditional calorimeters, PFA-based calorimeters employ advanced pattern recognition techniques to reconstruct complex events based on the idea that it measures the energy of every single particle in a jet using the subdetector, which gives the best resolution. A simplified working principle of PFA is depicted in figure 1.4(right). Only neutral hadrons of the jet have to be reconstructed in the HCAL and partially with the ECAL if the shower starts early. The charged particles are reconstructed in the tracker system that offers better energy resolution compared to calorimeters, and the electrons (in principle also measured by tracker) and photons are measured in the ECAL.

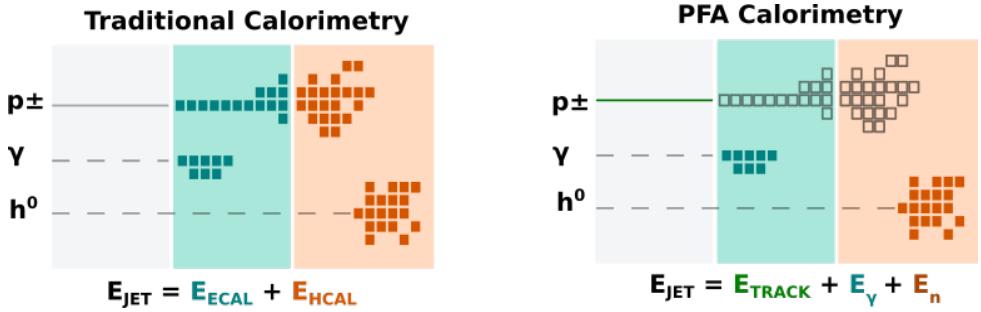


FIGURE 1.4: Schematic depiction of simple calorimetric (left) and Particle Flow jet energy reconstruction (right). In the calorimetric reconstruction, depositions of all the particles are summed up to yield the total jet energy, regardless of particle type and without using the measured momentum information from the tracking system. In the PFA scheme, the tracking system measures the energy of all charged particles instead of their calorimeter depositions. Photon energies are measured in the ECAL and only neutral hadron energies are measured by the HCAL (reconstructed energies are shown in colour).

Previous experiments such as ALEPH at LEP [34], H1 at HERA [35, 36], and currently operating ATLAS and CMS at LHC [37] employs the PFA techniques. However, these detectors were not specifically designed for PFA because the calorimeter granularities needed for such experiments were not achievable at the time.

There are three major known problems encountered while reconstructing the energy. Firstly, when a fraction of energy from a charged hadron is reconstructed as a neutral cluster which leads to double-counting of the energy. Secondly, the problem with missing energy is when a portion or entire neutral energy is associated with a charged energy, leading to an energy loss of neutral particle. Finally, the failure to resolve photons close to a charged hadron track leads to the loss of photon energy.

For a successful implementation of PFA, this imposes an excellent tracking system better than 99% track reconstruction efficiency within a jet. Furthermore, to avoid missing or double counting of energy, matching the reconstructed tracks correctly with the calorimeter clusters is essential to distinguish between charged and neutral hadrons. Apart from the calorimetric performance, this confusion is the limiting factor in PFA, placing stringent requirements on the granularity of the calorimeter systems.

All detector concepts for future colliders are instrumented with highly granular calorimeters and a tracker with good momentum resolution and a high magnetic field to exploit the PFA performance. Currently, PandoraPFA is the best known performing algorithm for the ILC detector concepts, providing a sophisticated pattern recognition for different detector designs. Recent studies [38] for the detectors at ILC have demonstrated the feasibility and performance of a detector design optimised for Particle Flow reconstruction, reaching around 3-4% for di-jet energies in the jet energy range from 45 GeV to 250 GeV. The following section describes the detector concepts at the ILC.

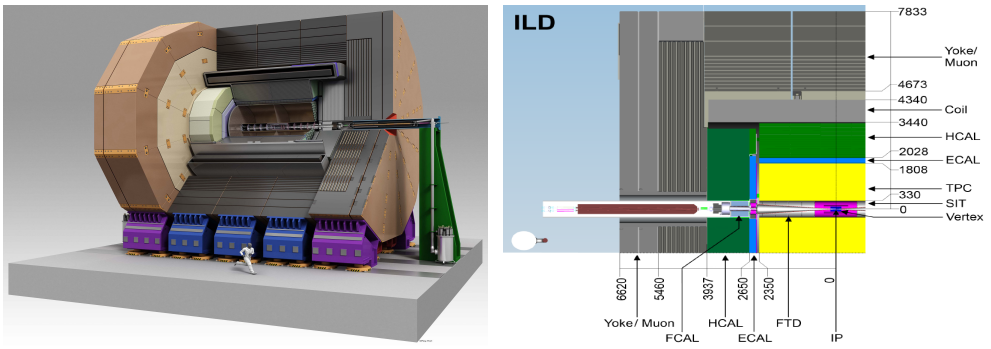
1.2.3 The International Large Detector

Nowadays, collider experiments fulfil the physics community's needs based on the four-eyes principle. However, the community only confirms a discovery when at least two detectors "observe an effect".

For the ILC, two separate detector experiments are planned to be operated. The ILD and SID share the central interaction region in a push-pull configuration. The non-operating detector can be maintained, repaired or upgraded while the other is taking data. These two concepts have been developed and described in detail in the detector volume of the ILC Technical Design Report [39]. High-performance detectors at ILC are optimized for a PFA to reach the best jet energy resolution, as mentioned in the previous section. Here, only the ILD will be discussed.

The ILD is a multi-purpose detector designed to perform precision measurements at the ILC for collision energies from 90 GeV to 1 TeV. The structure of the ILD detector is shown in figure 1.5(a). The vertex detector is the closest to the interaction point, surrounded by a high-efficiency hybrid tracking layout consisting of a silicon tracker and a time projection chamber. This is enclosed by high granular electromagnetic and hadronic calorimeters. The entire sub-detector system is placed within a superconducting solenoid coil to provide a magnetic field of 3.5 T parallel to the beam direction. Outside the coil, an iron return yoke is situated as a muon

system and as a **T**ail **C**atcher (TC) calorimeter. The cross-sections of sub-detectors are shown in the 2D projection of the ILD in figure 1.5(b).



(a) (b)

FIGURE 1.5: Schematic depiction of the ILD detector. (a) 3D view (b) Quadrant view of the ILD detector, the picture shows the interaction point (lower right) with the different sub-detectors. The dimensions are given in mm [26].

The sub-detectors of the ILD is briefly introduced here based on the ILC Technical Design Report:

The **V**ertex **D**etector (VTX) is the innermost part of the tracking detector and is in close vicinity to the interaction point as close as 16 mm to the beam axis. The VTX consists of three-double layers of thin silicon pixel sensors or five single layers of silicon pixel. The vertex detector is optimized towards a point resolution better than 6 μm, and it offers tagging efficiently the long-lived b and c hadrons and a meagre material budget with better than 0.15% X_0 per layer to minimize multiple scattering. Towards the exterior part of the vertex detector, there are two layers of silicon strip detectors which provide a hit position resolution of 7 μm with a material budget of $<0.0065 X_0$ per layer.

The central component of tracking is a large-volume **T**ime **P**rojection **C**hamber (TPC) filled with sensitive gas serving as a detection medium, providing with up to 224 precise measured points along the track of a charged particle. A high electric field parallel to the beam axis is applied between the endplates of the chamber. Charged particles traversing the TPC ionize the gas along their trajectory, creating free electrons along their path, which are drifted by the electric field towards the anode and are detected using a **G**as **E**lectron **M**ultiplier (GEM) [40] or a **Micro-Me**sh **Gas**eous Structure (Micromegas). The combined tracking system is designed to provide high momentum resolution of 2×10^{-5} $/(\mathrm{GeV/c})$ and TPC alone could provide 9×10^{-5} $/(\mathrm{GeV/c})$. In addition, the TPC offers the possibility of particle identification based on dE/dx measurements with a resolution of around 5%.

The calorimeter system follows after the TPC. The two parts of the calorimeter are the ECAL optimized to measure the energy deposit of electrons and photons. In contrast, the HCAL is responsible for measuring electrically neutral hadron showers while the energy of electrically charged hadrons is measured in the tracker. Both ECAL and HCAL are designed as highly

segmented sampling calorimeters optimized for the PFA with very high granularity to achieve the required energy resolution of 3-4% above 45 GeV. The ECAL has 30 active layers using a silicon-based readout with a 5×5 mm segmentation in a tungsten absorber, allowing a very compact design with a total depth of $24X_0$. The HCAL has 48 active layers based on scintillator tiles with 10^8 channels and steel as an absorber material with a total depth of $6\lambda_I$. Different technologies are under development for calorimeters with different active materials and absorbers. The physics and technological prototypes of the ILD detectors have been built and will be introduced in section 1.4. A superconducting coil envelopes the calorimeter system to generate an axial magnetic field. Subsequently, an iron yoke as a final part of the ILD returns the magnetic flux of the solenoid, and the same time serves as a muon filter, muon detector and tail-catcher to measure shower leakage from the HCAL.

1.3 Calorimetry and Shower Physics

From many existing particles, only a few are detectable in a calorimeter. Namely, photons (γ), electrons and positrons ($e^{\pm}$), muons ($\mu^{\pm}$), charged hadrons ($p, \pi^{\pm}$, $K^{\pm}$), and neutral hadrons (n, K_L^0). Charged particles such as $e^{\pm}$, $\mu^{\pm}$ and p interact via the electromagnetic force. Photons only couple to electrically charged particles. Hadrons are composite particles that can be either neutral or electrically charged. All hadrons interact via the strong interaction.

To design a calorimeter it is important to understand the shower processes and particle interaction with matter. This section, briefly introduces the fundamental physics of electromagnetic and hadronic interaction with matter. Then a general overview of calorimeter properties and figures of merit that are relevant for this thesis are discussed. Finally, at the end of this section, different physics lists used for the simulation of particle showers are introduced.

1.3.1 Muons and Heavy Charged Particles

When transversing through matter, muons (which are typically not measured in a calorimeter) and heavy charged particles undergo ionization and bremsstrahlung processes. The energy loss via bremsstrahlung depends on the mass of the particles:

$$-\left\langle \frac{dE}{dx} \right\rangle \propto \frac{1}{m^4} \tag{1.4}$$

Therefore, the bremsstrahlung process is suppressed by the particle mass for muons and heavy charged particles. Thus ionization is the primary process for energy loss of muons and charged hadrons in a wide momentum range between $0.1 \lesssim \beta\gamma \lesssim 1000$ and for intermediate Z materials. The mean ionization energy loss per unit length of a heavy charged particle as a function of its momentum traversing an absorber is well described by the Bethe-Bloch formula [41]:

$$-\left\langle \frac{dE}{dx} \right\rangle = Kz^2 \frac{Z}{A} \frac{1}{\beta^2} \left[\frac{1}{2} \ln \frac{2m_e c^2 \beta^2 \gamma^2 T_{\max}}{I^2} - \beta^2 - \frac{\delta}{2} \right]. \tag{1.5}$$

The constant $K = 4\pi N_A r_e^2 m_e c^2$ is a normalization factor, further details of the variables are provided in table 1.1. Figure 1.6 shows the mean energy loss for positive muons in copper as a function of $\beta\gamma$ in the energy range of MeV to TeV [3].

Symbol	Definition	Unit or Value
A	atomic mass of absorber	$\mathrm{g \cdot mol^{-1}}$
c	speed of light	3×10^8 m/s
E	incident particle energy $\gamma \mathrm{Mc^2}$	MeV
E_c	critical energy for electrons	MeV
M	incident particle mass	$\mathrm{MeV/c^2}$
$m_e c^2$	electron mass $\times \mathrm{c^2}$	0.510998 MeV
N_A	Avogadro's number	6.022141×10^{23} $\mathrm{mol^{-1}}$
r_e	electron radius	2.817940 fm
Z	atomic numberof absorber	
z	charge number of incident particle	
β	ratio of relative velocity and speed of light $\mathrm{v^2/c^2}$	
γ	lorentz factor $1/\sqrt{1-\beta^2}$	
ρ	density of the absorber material	$\mathrm{g \cdot cm^{-3}}$
T_{max}	maximum energy transfer to electron in a single collision	
I	average excitation energy of the material	
δ	relativistic density effect correction	

TABLE 1.1: List of the variables used in this section.

[3] https://meroli.web.cern.ch/lecture_stragglingfunction.html

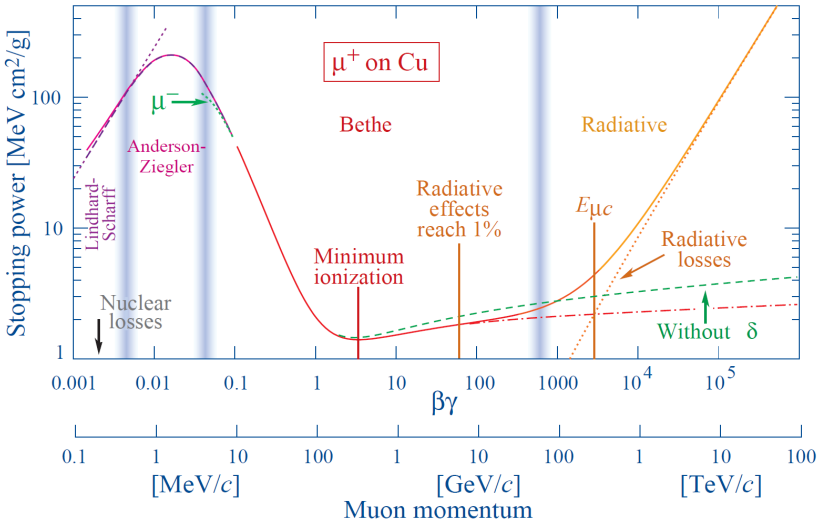


FIGURE 1.6: Average energy loss or stopping power $\left(-\left\langle \frac{dE}{dx} \right\rangle\right)$ of positive muons traversing the material, in this case copper, as function of the muon momentum. The vertical lines indicate boundaries between different theoretical approximations. The solid curve displays the total energy loss. For energies in the range ≈6 MeV to ≈60 GeV, dE/dx is dominated by the ionization process, described by Eq. 1.5. The short dashed lines indicate an idealistic behaviour of the energy loss under certain assumptions. The figure is taken from [10].

With rising energy, a relativistic muon's energy loss diminishes until it approaches a minimum. The term **M**inimum **I**onising **P**article (MIP) refers to a muon near to this minimum in the 'Bethe-region' with $\beta\gamma \approx 4$, losing their energy at a minimum rate, e.g. 1.6 MeV cm^2/g in Cu. A relativistic rise follows the minimum as the energy increases. On the other hand, the mean energy loss does not reveal the whole picture and is challenging to define experimentally since the energy loss distribution for muons traversing moderate amounts of material, such as the CALICE-AHCAL scintillator tiles, is not well-defined. Energy deposited in thin materials is explained by the 'Landau-Vavilov distribution' as shown in figure 1.7. With the distribution function describing such an energy loss must be asymmetric, as the minimum dE/dX can be very small or in principle even zero, but the maximum energy loss can be quite substantial. The skewness of the distribution describes that the **M**ost **P**robable **V**alue (MPV) for energy deposition is below the mean energy. The distribution is less skewed for thick materials, but it never approaches a Gaussian. Due to significant fluctuations that broaden the distribution, the Landau distribution does not accurately explain energy loss in thin absorber layers (e.g. RPC gas gaps). The most probable energy deposition for a charged particle is less dependent on the particle momentum than the mean energy loss, allowing it to be used as a natural and easily accessible deposition scale. For muons, the energy dependence of the ionization is almost constant in this range, and they have the properties of minimum ionizing particles. Muons can therefore be used for calibration purposes.

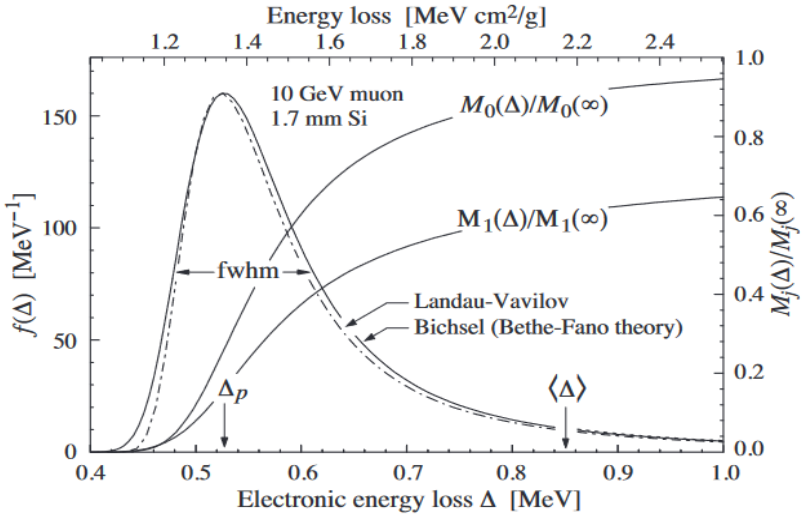


FIGURE 1.7: Energy distribution of a 10 GeV muon passing through 1.7 mm of Silicon. The dashed line corresponds to the Landau-Vavilov distribution, while the continuous line represents the Bethe-Bloch energy loss. $M_0(\Delta)$ and $M_2(\Delta)$ are the cumulative distributions denoting mean number of collisions and mean energy loss in Silicon. Δ_p represents the most probable energy loss, and $\langle\Delta\rangle$ divided by the thickness gives the Bethe $\langle dE/dx\rangle$. The picture is taken from [10].

1.3.2 Electromagnetic Interactions

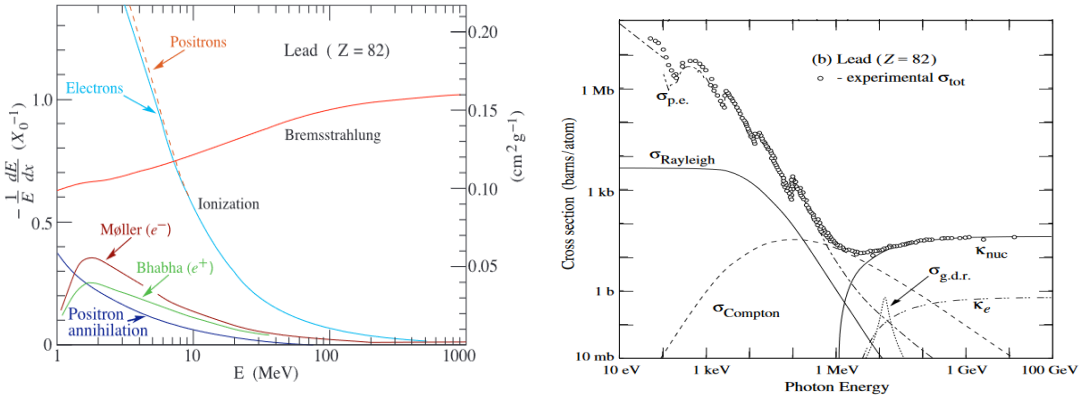
Energy loss of electron and positron in matter

Electrons and positrons, while propagating through matter, lose energy electromagnetically. Depending on the particle energy, different processes contribute to the energy loss. For energies above ∼10 MeV, the energy loss is dominated by the emission via the bremsstrahlung process. This radiative energy loss is caused when electrons (and positrons) interact with the coulomb field of the absorber atoms. At lower energies, most energy is deposited via ionization. Other minor processes contributing to the energy loss at low energies $(< 10\mathrm{MeV})$ are Møller scattering for electrons, Bhabha scattering and annihilation for positrons and multiple scattering on the absorber nuclei. Figure 1.8(a) summarizes these processes and their contribution at different energies for electrons and positrons in the lead. The energy at which the energy loss of electrons (and positrons) via bremsstrahlung equals ionization defines the critical energy, which is about 7.42 MeV for lead and 21.04 MeV for iron [42]. In addition to these processes, electrons (and positrons) can emit Čerenkov light. However, the contribution of this effect to the overall energy loss is negligible.

Energy loss of Photons in Matter

Energy loss by photons passing through matter is either via the photoelectric effect, Coulomb scattering or pair production. For low energy photons (E<100 keV), the energy loss is mainly through the photoelectric effect ($\sigma_{\mathrm{p.e}}$). However, Compton scattering and Rayleigh scattering contribute marginally. In figure 1.8(b) the cross-sections of the different processes in the lead

are summarised for photon energies of 10 eV to 100 GeV. For increasing photon energies above twice the electron mass m_e, (> 1 MeV) photons lose their energy via pair production processes or atomic electrons and minor contributions by nuclear interactions via a photo-nuclear process.



(a) (b)

FIGURE 1.8: (a) Shown is the fractional energy loss per radiation length as a function of electron (or positron) energy for the different processes in lead. Ionisation process occurs if the energy is below 0.255 MeV, while other minor contributions from Molière or Bhabha scattering. (b) Shown is the total photon cross-section as function of energy showing contributions from different photon processes in lead: atomic photo-electric effect ($\sigma_{\mathrm{p.e}}$), Rayleigh scattering ($\sigma_{\mathrm{Rayleigh}}$), Compton scattering ($\sigma_{\mathrm{Compton}}$), pair production in nuclear field (κ_{nuc}), pair production in electric field (κ_{e}), photo-nuclear interactions ($\sigma_{\mathrm{g.d.r}}$). The pictures are taken from [10].

Electromagnetic cascade

P. M. S. Blackett and G. P. S. Occhialini, who researched on high-energetic cosmic rays in 1933, made the first observations of electromagnetic showers [41]. These showers are now explained with high precision using the theory of quantum electrodynamics. The development of such a cascade is a statistical process. The longitudinal and transverse development of an electromagnetic cascade is summarized here.

The combination of bremsstrahlung and energetic photons by electron-positron pair production effects leads to an electromagnetic cascade. A simple schematic view of an electromagnetic cascade is shown in figure 1.9.

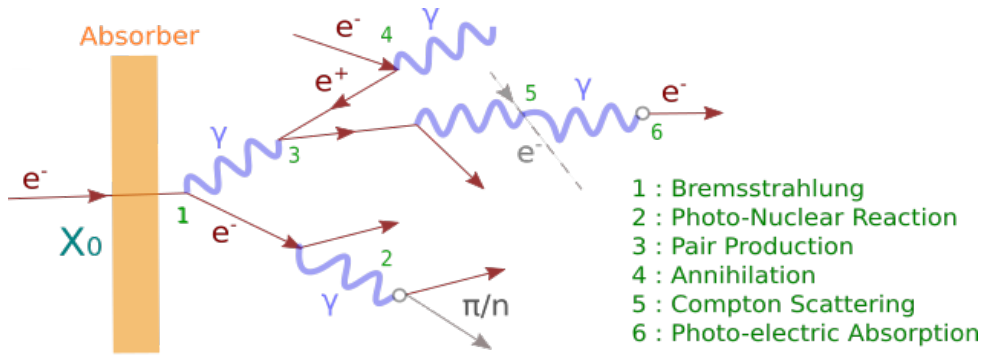


FIGURE 1.9: Schematic depiction of a electromagnetic shower induced by an electron. Electrons and positrons are represented with solid, red line, and photons with wiggly, blue line. Contribution from different processes are marked with numbers.

High-energetic electrons (or positrons) traversing matter start emitting bremsstrahlung immediately. In contrast, a photon entering matter may travel a certain distance before converting to an electron-positron pair. This cascade repeats and reaches its peak and fade out until the energy of the electrons and positrons are below the critical energy. The length of electromagnetic cascades is characterised by the radiation length X_0, which is the mean distance after which a high-energetic electron traversing matter has lost (1-1/e), which is 63.2% of its initial energy by bremsstrahlung. The mean distance a photon travels before pair production is the mean free path length, λ_p and is related to the radiation length via $\lambda_p = 9/7\ X_0$

A simple model describing the shape of the longitudinal profile of an electromagnetic shower was initially proposed by B. Rossi, which is also known as "approximation B" [43]. This model is energy independent and ignores the low-energy process, and the differences between electrons and positrons. Thus a shower initiated by an incident particle with energy E traversing through the matter of depth t expressed in units of radiation length can be described with two parameters ω and b as [44]:

$$\frac{dE}{dt} = At^{\omega} \cdot e^{-bt}. \tag{1.6}$$

The first term depicts the fast rise caused by particle multiplication and the second term with the exponential shower decay represents the absorption taking place. From this parametrisation the maximum position of the shower is calculated using

$$t_{max} = \frac{\omega}{b} \approx ln\frac{E}{E_c}. \tag{1.7}$$

The maximum shower depth increases logarithmically with the energy of the particle initiating the cascade. Alternatively, several other models do exist by now to explain the longitudinal shower development. For example, the Gamma distribution [45]:

$$\frac{dE}{dt} = E_0 b \frac{(bt)^{a-1} e^{-bt}}{\Gamma(a)}. \tag{1.8}$$

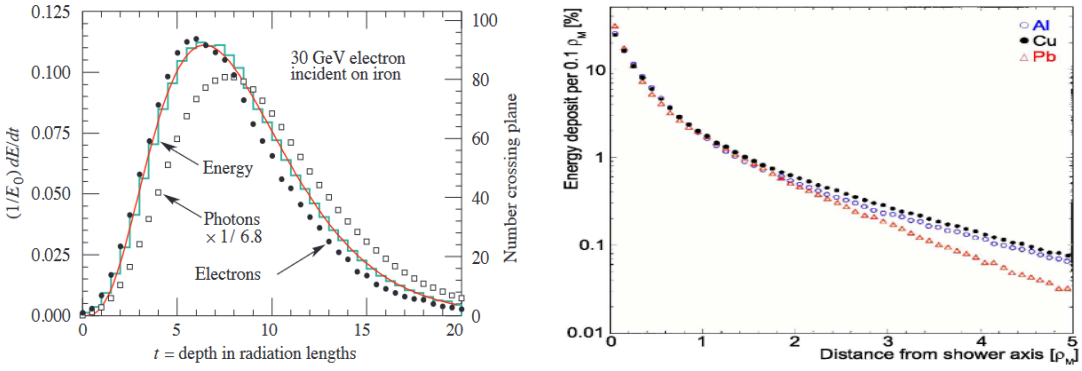
The transverse development of electromagnetic showers is mainly caused by large-angle multiple scattering of the electrons and positrons off the shower axis. A measure for the transverse extension of an electromagnetic shower is the Molière radius ρ_M. It is the radius of a cylinder around the shower axis in which, on average 90% of the total shower energy is contained (3.5 ρ_M correspond to 99% containment of electromagnetic showers). The Molière radius [46] is given by:

$$\rho_{\mathrm{M}} = 21.2\mathrm{MeV}\frac{\mathrm{X}_0}{\mathrm{E}_\mathrm{c}}. \tag{1.9}$$

The transverse development is described with two exponential distributions as a narrow core and a broader halo. The central part is mainly caused by multiple scattering effects produced by relativistic electrons and scales with the Molière radius and is independent of Z. The halo part is mainly caused by soft electrons and photons from Compton scattering and photoelectric absorption that scattered away from shower axis. Thus, it gets steeper as Z decreases. A variety of different functions exists to describe the transverse shower development. A parametrisation that provides a reasonable agreement with experimental data from the H1 experiment at HERA [47], which explained the core and the halo of the electromagnetic shower is given by:

$$f(r) = p\frac{2rR_C^2}{\left(r^2 + R_C^2\right)^2} + (1-p)\frac{2rR_T^2}{\left(r^2 + R_T^2\right)^2}. \tag{1.10}$$

Here p ranges from 0 and 1, which gives the probability to be in the core part of the shower and R_C and R_T are the medians of the core and the halo part, respectively.



(a) (b)

FIGURE 1.10: (a) Simulated average longitudinal profile of 30 GeV electron in iron. The distribution is described using a Gamma function as given in (1.8) shown as a red curve. The picture is taken from [10] (b) Transverse shower profile using simulated 10 GeV electron showers in aluminium, copper and lead [46].

1.3.3 Hadronic Interactions and Showers

Hadrons are composite particles made of quarks and gluons. Neutral hadrons are subjected to strong interaction while traversing through matter. A charged hadron behaves similarly to a muon of the same energy, losing energy by ionization before it undergoes a strong interaction.

In contrast to the radiation length for electromagnetic showers, the length of hadronic showers is described based on the nuclear interaction length, represented as λ_{I}. It is defined as the mean free path travelled by a high-energy incoming hadron before undergoing an inelastic interaction. For any given material, this scale is generally much more prominent than X_0, for example, $\lambda_{\mathrm{I}}/\mathrm{X}_0 \approx 9.5$ for Iron. Additionally, this scale depends on the particle type as the pion interaction length is roughly $3/2$ times larger than the proton interaction length.

After the first hard interaction, the incoming hadron loses its identity. It contributes to various processes leading to a hadronic shower, and the struck nucleus is left in an excited state. An example of a hadron shower initiated by a charged hadron is shown below:

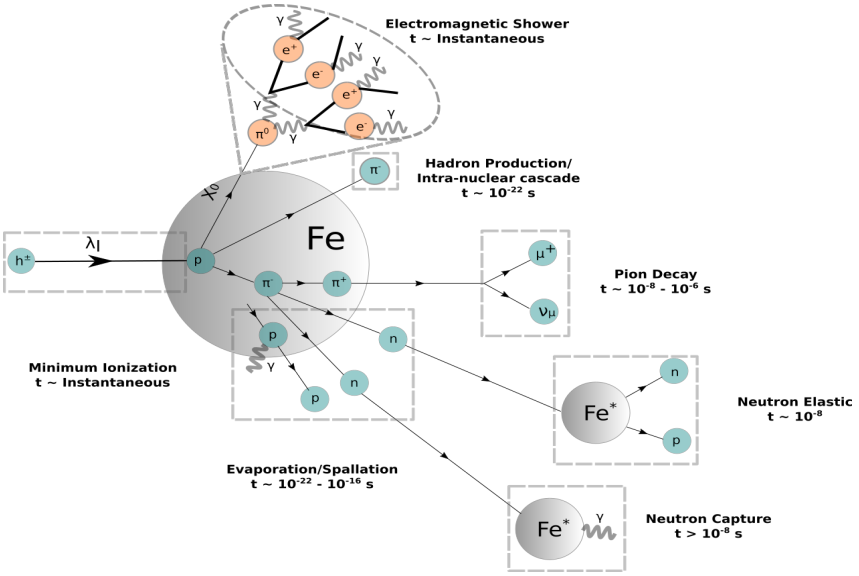


FIGURE 1.11: Schematic of a hadronic cascade. The inelastic scattering of a hadron with a nucleus produces a variety of secondary particles. The generated neutral mesons (e.g. π^0s or ηs) contribute to the electromagnetic fraction due to their decay into 2γ's. The generated hadrons traverse the material until they interact strongly and generate further secondary particles.

From this point on, the particle multiplicity increases, and the shower evolves. However, at some depth, this multiplication is balanced by absorption and the deposited energy decreases. Like electromagnetic showers, the mean shower depth of a hadron shower increases logarithmically with the incident hadron energy. In general, hadron showers have more extensive spatial development and smaller energy density than electromagnetic showers.

While traversing a material, all hadrons undergo elastic or inelastic scattering processes. Inelastic hadronic interactions are the ones that generate hadronic showers by producing secondary particles. Which then lose their energy either by ionization and excitation or undergo further

inelastic interactions such as spallation, fission, and nuclear breakup, which is why hadronic interactions are more complicated than electromagnetic interaction. In the following, general nuclear processes are briefly introduced.

Nuclear spallation

Spallation or nuclear breakup is the most probable process when a high energy hadron interacts strongly with the constituents of a nucleus. Spallation can be described as a two-stage process based on different time scales [48]. In the first step, the fast intra-nuclear cascade occurs as the incoming hadron collides with the nucleons inside the nucleus. The nucleons travel and collide with other nucleons creating a cascade of nucleons. Some particles could reach the boundary or escape from the nucleus. Also, pions are created if the energy is above the pion production threshold. The cascade spans of the order of 10^{-22} s.

The second stage of spallation is the slower evaporation or de-excitation of the intermediate nucleus. After the fast intra-nuclear cascade, the nucleus is in an excited state. It will return to the ground state with secondary processes like slow nucleus evaporation. In this process protons, neutrons, and smaller nuclei like deuteron, tritium, helium-3 and α-particles escape from the excited nucleus. When the energy of the excited nucleus falls below the binding energy of a single nucleon, the energy is released by the emission of photons, typically a few MeV. The nuclear de-excitation happens on a time scale of $10^{-18} - 10^{-13}$ s.

Invisible energy

Invisible energy refers to the part of the deposited energy produced during strong nuclear interactions which do not contribute to the calorimeter signal. Primarily during the spallation process during excitation or recoil of the target nuclei or nuclear binding energy of secondary particles. Nevertheless, neutrons generated within the shower lose their energy either by the decay of the neutron or by neutron capture. In the latter case, the energy it contributes is usually not measured. Furthermore, additional energy might be lost to neutrinos originating from meson decays, resulting in a lower calorimeter response for showers caused by hadrons than for electrons of the same energy. On average, $30 - 40\%$ of the non-electromagnetic shower energy is invisible.

Electromagnetic component

Hadron showers generally contain an electromagnetic component, comprising contribution from neutral pions and etas decaying instantaneously into two photons due to a very short lifetime of about $\approx 10^{-7}s$. Pions are mostly dominant in a hadronic shower and produced in equal parts (π^+, π^-, π^0) after the first nuclear interaction. π^0 decay 99% of the time decay into two photons. The photons are released from neutral mesons deposits via electromagnetic processes. In the second nuclear interaction, again one-third of the remaining purely hadronic fraction will result in π^0 mesons. This process is continued as long as the available energy is higher

than the pion production threshold. This fraction of the total hadron cascade energy that is deposited by local electromagnetic sub-showers is called electromagnetic fraction, f_{EM}. The f_{EM} increases logarithmically with the primary particle's energy which is explained with an empirical model [49] as given by

$$f_{EM} = 1 - \left(\frac{E}{E_0}\right)^{(m-1)}, \tag{1.11}$$

where, E is the energy of incoming hadron initiating the cascade. E_0 is the scale factor, and it corresponds to the pion production threshold (for pions, this is typically in the order of ≈1 GeV). The exponent $(m-1)$ is the multiplicity factor, related to the average number of π^0 mesons produced per hadronic interaction. Both E_0 and m must be determined for a given calorimeter used in an experiment. For example, in iron, the values are $m \approx 0.8$ and E_0 ≈1 GeV [49, 50]. In addition, charge conversion reactions also contribute to f_{EM}, for example, $\pi^- + p \rightarrow \pi^0 + n$ [51, 52]. Typically the values of f_{EM} ranges between 0.3 to 0.8. The electromagnetic fraction has a critical consequence on the energy resolution of hadronic calorimeters, which will be revisited in the following section.

1.3.4 Calorimeters

Calorimeters play a key role in modern particle physics as they are the only way to quantify the energy of a particle by measuring how much energy it deposits when interacting with matter. In addition, calorimeter measurements can also be used for particle identification during data analysis. The basic principles of calorimetry are the following:

1. Ideally, the deposited energy, which is the sum of all signals in a calorimeter cell, also called the visible signal E_{vis}, is linearly proportional to the energy of the incoming particle E_{inc},

   $$E_{vis} \approx c \cdot E_{inc}. \tag{1.12}$$

   The factor of proportionality is the basis of all calorimetric measurements. It is determined using calibration procedures for particles of known energies where the measured energies are compared to the incident energies.
2. The deposited energy is measured either by using scintillating materials such as Silicon Photomultipliers or by using charge sensitive devices such as silicon diodes or gas mixtures like liquid argon.
3. Calorimeters add to our understanding of the position, direction, and type of absorbed particles.

Calorimeters are broadly classified by their construction principles into *homogenous* and *sampling*, ECAL and HCAL. Before explaining the difference between the calorimeters, it is important to understand the key parameters of calorimeters which are the cornerstones for understanding the fundamentals of calorimetry.

Relative Energy Resolution

The *relative energy resolution*, $\sigma_{\mathrm{E}}/\mathrm{E}$ is typically an important figure of merit for most of the calorimeter systems and can be parametrized by:

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c \tag{1.13}$$

The first term, $a/\sqrt{E}$, is the stochastic term due to the Poisson statistics of the shower process. The $1/\sqrt{E}$ dependence is due to the intrinsic signal or sampling fluctuations. Assuming N to be the number of particles that contributed to the signal, the uncertainty is then given by: $\sigma_{\mathrm{N}}/\mathrm{N} = \sqrt{\mathrm{N}}/\mathrm{N} = 1/\sqrt{\mathrm{N}}$.

In addition, for hadronic showers, the fluctuations which arise from invisible energy and the fluctuations in the electromagnetic fraction also contribute to the stochastic term. Typically, for hadronic calorimeters, the stochastic contribution is in the order of $60\%/\sqrt{E}$. While for electromagnetic calorimeter it is up to $10\%/\sqrt{E}$.

The second term b/E, reflects the noise term. It is independent of the particle energy and arises from different effects like the readout electronics. For example, too low thresholds can result in a high noise rate. This term dominates at low energies, which is the limiting factor there.

The third term c, is the constant term and is mainly caused by detector effects such as calibration uncertainties, response inhomogeneities of the sensitive areas, and leakage effects. This term dominates at high energies and is typically in the order of a few percent. The relative contribution of this term is the same for all particle energies.

Calorimeter Response

The calorimeter response is defined as the average calorimeter signal per unit of deposited energy, proportional to the incoming particle's total shower energy. Thus, a linear calorimeter has a constant response. This is true in the case of a homogenous electromagnetic calorimeter since electrons and photons are visible particles and can be detected.

The calorimeter response for hadronic showers is more complicated. Due to significant event to event fluctuations of the electromagnetic sub-shower developing within a hadronic shower, invisible energy leads to intrinsic non-linearity with hadron energy.

The response of a calorimeter for a hypothetical purely hadronic shower h, i.e. without electromagnetic sub-shower (f_{EM} = 0), is different compared to a purely electromagnetic shower (e). To assess the response of a calorimeter, the ratio h/e is used. As in reality hadron showers have an electromagnetic component $f_{EM} \neq 0$. For example, the energy deposited by pion is given by:

$$E_{vis}(\pi) = \langle f_{EM} \rangle \cdot E_{vis}(e) + \langle f_h \rangle \cdot E_{vis}(h),$$

where $E_{vis}(e)$ and $E_{vis}(h)$ are the visible energy deposited by the electromagnetic shower fraction f_{EM} and the purely hadronic fraction f_h. The h/e ratio can only be determined by measurements of the π/e ratio as given by:

$$\begin{aligned} \langle \pi/e \rangle &= \langle f_{EM} \rangle + \langle f_h \rangle (h/e) = 1 - (1 - (h/e)) \langle f_h \rangle \\ \langle f_h \rangle &= 1 - \langle f_{em} \rangle \approx (E/E_0)^{m-1} . \end{aligned} \tag{1.14}$$

Thus one arrives at Eq. 1.11. The h/e signal ratio is energy independent and quantifies the degree of so-called non-compensation.

A compensating calorimeter has a h/e ratio equal to 1. A non-compensating calorimeter ($h/e \neq 1$) has a worse energy resolution than a compensating calorimeter due to significant event to event fluctuations in f_{EM}. Compensation can be achieved either by attenuating the electromagnetic response or by enhancing the hadronic response, for example, by using depleted uranium in the absorber material or increasing the hydrogen content in the active material in the form of organic scintillators. The calorimeter of the ZEUS experiment was designed in this way to be compensating [53]. Besides implementing sampling configurations, compensation can also be achieved later in data acquisition, in so-called offline or software compensation. Other approaches are used and actively investigated, such as a dual readout calorimeter [54].

Types of Calorimeter

Conceptually, there are two types of calorimeters: homogeneous calorimeters or sampling calorimeters. Homogeneous calorimeters are made of a single detector material, which absorbs the energy of the incident particle and, at the same time, generates a measurable signal that is proportional to the energy loss of the particle. Such calorimeters offer excellent single-particle energy resolution. An example of a homogenous calorimeter is the CMS-ECAL detector which provides excellent energy resolution as given in table 1.2.

In sampling calorimeters, the calorimeter is built using alternating layers of a passive medium, the absorber, which is made usually of high-density material, and an active medium that generates the signal to be measured. Sampling calorimeters offer the freedom to choose both the absorber and active material. Nevertheless, it has the disadvantage of worse energy resolution than homogenous calorimeters due to sampling fluctuations, as only a fraction of the energy of a shower is measured. An example of a sampling calorimeter is the ATLAS tile hadronic calorimeter with energy resolution provided in table 1.2.

ECAL are optimized and responsible for measuring the energy of electrons, positrons and photons. In contrast, HCAL measures the energy of neutral particles, such as pions and kaons, but also hadrons in general.

The table below gives examples of both electromagnetic and hadronic calorimeters with their energy resolution from different experiments.

Experiment	Material	Resolution
		ECAL
ATLAS	Lead, LAr	$\frac{10\%}{\sqrt{E(\mathrm{GeV})}} \oplus \frac{0.170}{E(\mathrm{GeV})} \oplus 0.7\%$
CMS	$PbWO_4$	$\frac{2.8\%}{\sqrt{E(\mathrm{GeV})}} \oplus \frac{0.125}{E(\mathrm{GeV})} \oplus 0.3\%$
LHCb	Lead, scintillator	$\frac{10\%}{\sqrt{E(\mathrm{GeV})}} \oplus 1\%$
ALICE	$PbWO_4$	$\frac{3.3\%}{\sqrt{E(\mathrm{GeV})}} \oplus \frac{0.18}{E(\mathrm{GeV})} \oplus 1.1\%$
		HCAL
ATLAS	Steel, scintillator	$\frac{52.9\%}{\sqrt{E(\mathrm{GeV})}} \oplus 5.7\%$
CMS	Brass, scintillator	$\frac{84.\%}{\sqrt{E(\mathrm{GeV})}} \oplus 0.7\%$
LHCb	Iron, scintillator	$\frac{(69\pm5)\%}{\sqrt{E(\mathrm{GeV})}} \oplus (9 \pm 2)\%$

TABLE 1.2: Energy resolutions of electromagnetic and hadronic calorimeters from various experiments [55, 56].

1.3.5 Physics Simulation and Hadronic Lists

Simulation is an indispensable tool to investigate, optimize and validate the experimental results in high energy physics and various other fields of research (e.g. medical and space sciences). In addition, they are helpful for design and optimization studies and serve as a guideline for data analysis. Throughout this thesis, GEANT4 [57, 58] is used. It is a software toolkit to simulate particle interactions and spatial development of electromagnetic and hadronic showers in the matter. GEANT4 offers various tools and models which can simulate both electromagnetic and hadronic cascades. However, this thesis will only focus on the models that have been used to simulate electromagnetic and hadronic interactions.

Electromagnetic showers are generally considered well understood and can be modelled with high precision due to their simple interactions between only electrons, positrons and photons. The complexity lies mainly in the simulation of hadronic interactions due to the composite nature of the incoming particle and the hit nucleus, which also have a quark substructure. Furthermore, additional hadrons may be generated if there is enough energy after interactions, resulting in a large number of final state particles.

Several hadronic models exist, but there is no single model that could describe the entire energy range. For this, a detailed geometry description of the detector model and its material composition is necessary. The physics interactions of the particles with the materials follow the processes implemented in the physics lists.

A detailed list of the physics processes implemented in GEANT4 can be found in [59]. In addition, GEANT4 provides models for each energy range which are combined and are called "physics lists". The principal uncertainty in these physics lists is the choice of the hadronic model. Therefore, it is the user's responsibility to choose which physics processes are simulated and

how they are modelled. Two relevant string model-based on physics lists are introduced and have been chosen for this thesis. They are considered the most stable and well-tuned as recommended from GEANT4 authors [60]. The choices of hadronic models and energy ranges vary between different GEANT4 versions. In this thesis, version 10.03.p02 is used.

The **QGSP_BERT** physics list relies on the Bertini intranuclear cascade model [61] for low energies (<10 GeV). The Bertini cascade models the intranuclear interactions as subsequential particle-particle interactions. For high energies (>25 GeV), the quark-gluon string model is used to describe hadron-nucleon inelastic scattering. In the quark-gluon string model, a string is modelled between two quarks of the scattering particles. These strings can be fragmented (provided they carry enough energy) to generate quark-antiquark pairs. The produced quarks further form hadrons (hadronization). The LEP model connects both models, which is a parameterized model based on fits to experimental data.

The **FTFP_BERT** physics list employs the Bertini cascade as well but only up to 5 GeV. For higher energies, the Fritiof Parton string model is used. This model differs from the quark-gluon string model in the string formation and the string fragmentation process [62].

The **High Precision Neutron Package_HP** serves as an extension to the QGSP_BERT and FTFP_BERT physics lists. It employs a special package to transport neutrons below 20 MeV down to thermal energies, showing better performance of the time structure for hadronic showers [63].

1.4 CALICE Detector Concepts

The CALICE collaboration, is an international collaboration involving around 60 institutes from 19 countries [64]. The collaboration is developing, comparing and validating different concepts foreseen for future **H**igh **E**nergy **P**hysics (HEP) detectors based on innovative technologies to allow unprecedented granularity both in the longitudinal and lateral direction. The CALICE calorimeters are optimised for a linear collider environment driven by the aim of a detector exploiting the PFA. CALICE started to collaborate with the LHC to develop calorimeters that fulfil the required granularity needed for the **H**igh **L**uminosity **L**arge Hadron **C**ollider (HL-LHC) upgrade [65]. Different electromagnetic calorimeter concepts are in development in the CALICE collaboration. For each calorimeter, a physics prototype was developed to improve their performance for the electromagnetic and hadronic shower measurements. Afterwards, the engineering prototypes were developed to improve their design by integrating the front-end electronics and the readout scheme. One example of an adaptation to previous generation prototypes is power pulsing, which allows to minimize the power consumption in the active layers at the ILC. Such that the electronics are switched on during the bunch train (1 ms) while it is off to save power in the long gaps between bunch trains (199 ms) [66].

Besides the AHCAL, which is the leading technology behind the work of this thesis, different other technologies were developed with different absorbers and readout systems, such as the

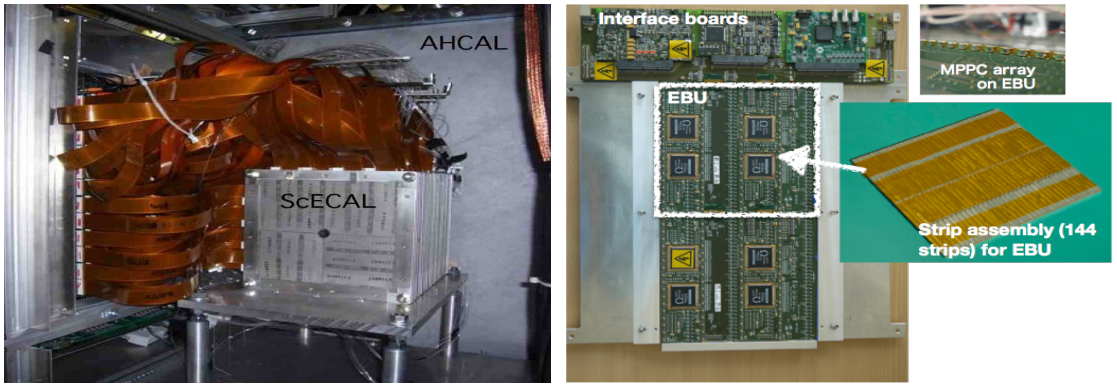
digital and semi-digital calorimeters. A detailed comparative study of different calorimeter concepts within the CALICE Collaboration has been done and can be seen in [67]. These prototypes were tested and exposed to high energy particles in several test beam campaigns since 2006. This section briefly introduces the prototypes developed within the CALICE collaboration.

Electromagnetic Calorimeter Prototypes

The CALICE collaboration has developed two different physics prototypes of electromagnetic sandwich calorimeters - the **Sc**intillator strip-Tungsten **ECAL** (Sc-ECAL) [68] and the **Si**licon-Tungsten **ECAL** (SiW-ECAL) [69]. Both physics prototypes consist of tungsten as absorber material and prototypes have faced several test beam campaign [70].

Scintillator Strip-Tungsten ECAL

The Sc-ECAL physics prototype consists of 30 layers of scintillator strips as active material. The total thickness of the prototype is $\approx 21.5\ X_0$, corresponding to roughly $1\lambda_I$. Each active layer comprises four rows of scintillator strips of $45{\times}10{\times}3\ \mathrm{mm}^3$. The strips are oriented orthogonally in consecutive layers. A sketch of the prototype is shown in figure 1.12(a). Overall, the prototype has 2160 readout channels. This prototype has been operated at Fermilab, and an energy resolution with a stochastic term of $\frac{12.6\%}{\sqrt{E[GeV]}}$ and a constant term of 1.6% has been obtained [68]. Currently, the development of the technological Sc-ECAL prototype, shown in figure 1.12(b), consists of the integration of the front-end electronics in the layer to reduce the dead space in the calorimeter due to cabling. Each **ECAL** **B**ase **U**nit (EBU) has 144 scintillators strips of dimension $45{\times}5{\times}2\ \mathrm{mm}^3$, and it has a transversal dimension of $180{\times}180\ \mathrm{mm}^2$. A new generation of SiPM was used, which allowed the use of scintillators strips without the **W**ave-**L**ength **S**hifting (WLS) due to improvements in SiPM technology.



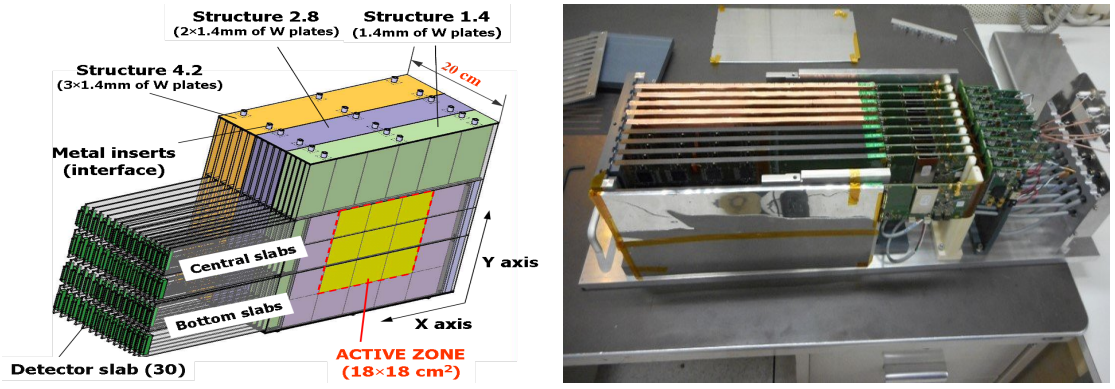
(a) (b)

FIGURE 1.12: (a) Picture of the Sc-ECAL physics prototype. (b) Picture of single layer Sc-ECAL of the technological prototype [71].

Silicon-Tungsten ECAL

The SiW-ECAL physics prototype consists of 30 silicon active layers, shown in figure 1.13(a), with the dimension of 18×18 cm^2 that is segmented into 1×1 cm^2 pad diodes, resulting in 9720 readout channels in total and a total depth of $24X_0$ corresponding to 1 λ_I. The thickness of the passive layers increases with the depth of the prototype. It goes from 1.4 mm for the first ten layers to 2.8 mm for the second group of 10 layers and up to 4.2 mm for the last ten layers. In test beam campaigns at DESY and CERN, this prototype achieved an energy resolution with $\frac{16.53\%}{\sqrt{E[GeV]}}$ stochastic term and 1.07% constant term [72].

The SiW-ECAL technological prototype is shown in figure 1.13(b) illustrating the technological prototype of 7 layers inside the aluminium stack. This prototype is based on sensors that consist of high resistivity silicon wafers with a size of 9×9 cm^2. Each of them is divided into an array of 256 PIN diodes of 5×5 mm^2 in order to improve the pattern recognition of the calorimeter.



(a) (b)

FIGURE 1.13: (a) Schematic 3D view of the SiW-ECAL physics prototype [72]. (b) Picture of SiW-ECAL technological prototype with 7 layers inside the aluminum stack [73].

Hadronic Calorimeter Prototypes

As for the ECAL, the CALICE collaboration has built and thoroughly tested several hadron calorimeter concepts. The major goal is to develop hadron calorimeters with very high granularity, which can vary from 3×3 cm^2 to 1×1 cm^2, for the separation of particles within the hadronic jets, as high precision of the amplitude information.

The **D**igital **H**adron **Cal**orimeter (DHCAL) and **S**emi-**D**igital **H**adron **Cal**orimeter (SDHCAL) calorimeter concepts use digital information of the amplitude by counting the number of hits on one or two bits. On the other hand, the AHCAL calorimeter concept relies on the complete analog information of the amplitude.

Digital HCAL

The DHCAL [74] is a sampling calorimeter based on **R**esistive **P**late **C**hambers (RPC) as active elements with a one-bit readout system. A schematic cross-section of an RPC is shown in figure 1.14. The DHCAL prototype consists of 38 active layers with an active area of 1×1 m^2. The prototype accounts for a total of 350,208 channels. A test beam campaign of this prototype was done at CERN and Fermilab with positron and pions beams. An electromagnetic energy resolution σ_E/E with a stochastic term of $\frac{35.1\%}{\sqrt{E[GeV]}}$ and a constant term of 12.4% and a hadronic energy resolution with a stochastic term of $\frac{64\%}{\sqrt{E[GeV]}}$ and a constant term of 4% has been achieved [75].

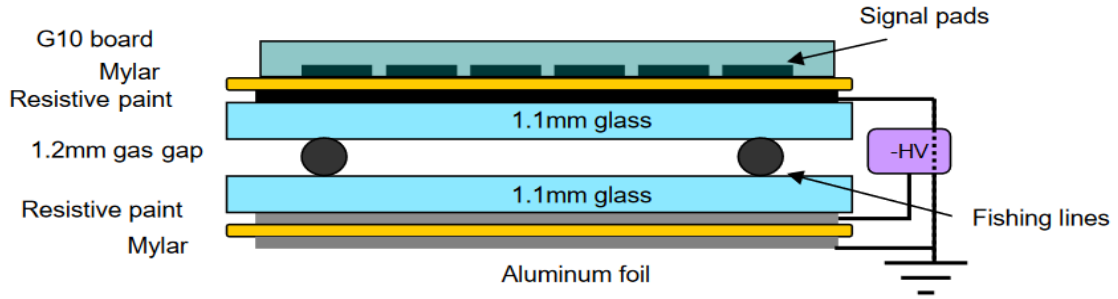


FIGURE 1.14: Schematic of the cross-section of an RPC on the DHCAL prototype [74].

Semi-Digital HCAL

The SDHCAL [76] is a prototype based on **G**lass **R**esistive **P**late **C**hambers (GRPC). It is made of 48 active layers inserted in a steel absorber structure with a layer thickness of 1.5 cm, for a total depth of 5.76 λ_I. It is similar to the DHCAL, the main difference between the two prototypes is the readout. The readout is done by the HARDROC chip [77] with 64 channels. This chip has a 2-bit readout system that corresponds to three thresholds. The chip will store only signals which are approximate by equal to the charged particles traversing the GRPC. The SDHCAL prototype was tested at the SPS in 2012 with various beam particles (muons, electrons and pions). An energy resolution from ∼13.3% for 30 GeV pions and up to ∼10.6% for 80 GeV has been obtained [78]. It is important to mention that these values obtained with the three thresholds readout mode showed an improvement compared with the digital readout mode.

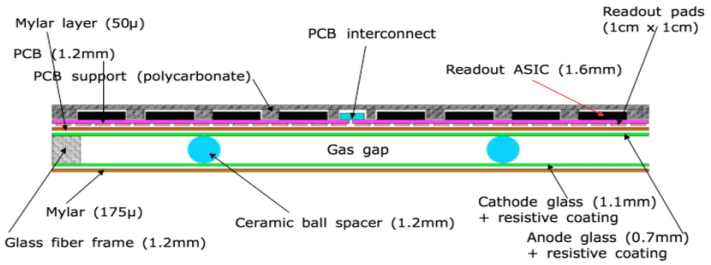


FIGURE 1.15: Schematic of the cross-section of an RPC on the SDHCAL prototype [76].

Analog HCAL

The AHCAL is a sampling calorimeter with scintillator tiles as active material and steel or tungsten as an absorber. The AHCAL technological prototype is described in more detail in Chapter 2. The AHCAL physics prototype [79] consists of 38 active layers interleaved in steel absorber plates of 17.4 mm thickness on average or in tungsten absorber plates which are around 10 mm thick. The calorimeter has a total depth of $\sim$4.2 λ_I. The active layers hold 216 channels for the first 30 layers and 141 channels for the last eight layers. The physics prototype has a total of 7608 channels. Scintillator tiles connected on a PCB is shown in figure 1.16(a). The tiles are 5 mm thick with different sizes of 3$\times$3, 6$\times$6 and 12$\times$12 cm^2. The light produced in each scintillator is guided through a WLS fiber of 1 mm thickness, inserted in the tile to a SiPM. The SiPM has a sensitive area of 1.1$\times$1.1 mm^2 with 1156 pixels. The performance of the physics prototype has been demonstrated in several beam types. A picture of the physics prototype at SPS-CERN is shown in figure 1.16(b). The energy resolution obtained for electrons has a stochastic term of $\frac{21.7\%}{\sqrt{E[GeV]}}$ and $<$1% constant term [80]. For pions, the intrinsic energy resolution has been measured to be $\frac{57.6\%}{\sqrt{E[GeV]}}$ stochastic and 1.6% constant term [80].

(a) (b)

FIGURE 1.16: (a) Picture of an active layer of the AHCAL physics prototype. (b) Picture of the AHCAL physics prototype at the CERN-SPS test beam. The pictures are taken from [79].

Tail Catcher and Muon Tracker

FIGURE 1.17: shows a picture of the TCMT in an orange frame on top of its support structure [81].

The **T**ail **C**atcher **M**uon **T**racker (TCMT) serves as an extension of the hadronic calorimeter. The CALICE collaboration built a prototype of a TCMT [81] as shown in figure 1.17. There are two sampling structures: the first eight layers consist of a 2 cm steel absorber and 0.5 cm scintillator, while the absorber of the last eight layers has a coarse longitudinal segmentation of 10 cm thick steel. The active layers consist of 100×5 cm^2 scintillator strips. The scintillation light from each strip is collected by a WLS and guided to a SiPM in the centre. The scintillator strip layers are alternating between horizontal and vertical orientation and therefore gives the opportunity to reconstruct the hit position. The TCMT has a depth of $6\lambda_I$, and the total number of channels is 320. In the test beam setup, the TCMT is used as an extension of the calorimeter system. Besides identifying muons, it measures the shower energy leaking from the AHCAL.

The Standard Model is the most reliable theory describing the elementary particles and the interactions that govern our universe. The Higgs boson is one of the bases of the theory; therefore, a very precise measurement of its properties, such as mass and couplings, is needed. As discussed in this chapter, ILC would be the perfect experiment to conduct precision measurements.

To fulfil the ILC requirements, calorimeters are an essential tool in jet energy measurement and are designed for the particle flow approach. The basic concepts of calorimetry, simulation models and the particle flow concept are discussed in this chapter. In addition, the development of the calorimeter concepts (electromagnetic and hadronic) within the CALICE Collaboration has been briefly introduced in this chapter.

The AHCAL detector, simulation and digitization procedures relevant to this thesis are discussed in the following chapter.

Part II

Experimental Setup

Chapter 2

Analog Hadron Calorimeter Technological Prototype

"The science of today is the technology of tomorrow."

Edward Teller

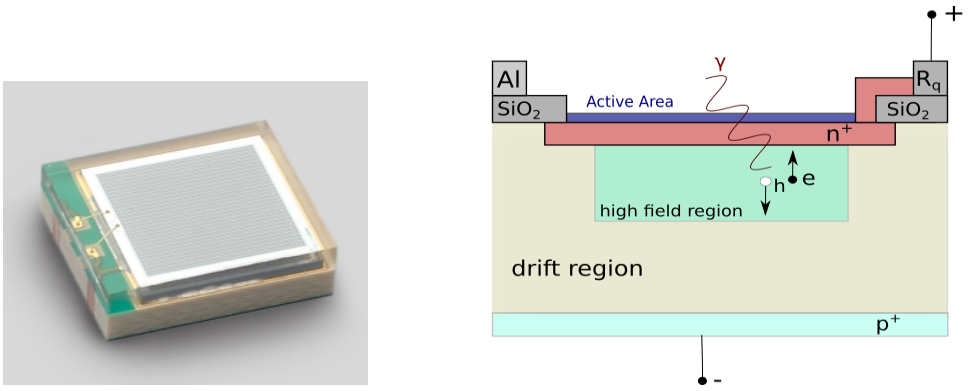
THE AHCAL TECHNOLOGICAL PROTOTYPE is one of the detector concept being developed within the CALICE collaboration for highly granular sandwich calorimeters. Before producing a final, sophisticated detector like the AHCAL, several short-term developments are necessary. In the first stage, a physics prototype was built that demonstrated the physics capabilities and showed successful performance of the PFA in a scintillator tile readout technology as already mentioned in section 1.4. The second stage involved the technological prototype which was developed to demonstrate the scalability and feasibility of the mass production and the implementation in the ILD detector.

This chapter presents a short overview of the CALICE-AHCAL technological prototype. The prototype is based on a steel absorber and is read out directly by the SiPM, which is coupled to scintillator tiles in the active layers of the calorimeter. A detailed description of the AHCAL technological prototype can be found on Ref [82]. Before discussing the prototype, the key element of this detector is introduced. Following that, the digitization process of the simulation is discussed, which includes the modelling of all readout effects resulting from the combination of the active material used, the sensors and the electronics.

2.1 Introduction to Silicon Photomultiplier

A scintillator detector employs indirect detection, where the absorbed energy is converted to visible light. Therefore, the number of scintillating photons generated are in proportional to the absorbed energy. These scintillating photons are detected by photomultipliers consisting of a photo-cathode and an electron multiplier.

A novel type of photodetector is the SiPM. These are semi-conductors used to measure light at the single-photon level. They are widely used due to their excellent properties compared to traditional PMTs. SiPMs consist of a matrix of **A**valanche **P**hoto**d**iodes (APD) operated in Geiger-mode with pixels connected in parallel to a common cathode and anode. This results in an analog signal proportional to the number of photons arriving at the SiPM surface. An example of a SiPM is shown in figure 2.1(a). Each pixel is operated in reverse bias mode, with an applied potential in the order of ∼30-60 V. The applied voltage is more significant than the breakdown voltage (V_{bd}), which is the minimum operating voltage for a self-amplification process to occur. A schematic of a single SiPM pixel is shown in figure 2.1(b). Upon the impact of a photon, an electron-hole-pair may be created in the depletion region by the photo-electric effect. Due to the applied reverse bias voltage, the electric field accelerates the electron. As a result, it creates a self-sustained avalanche or Geiger discharge until the current exceeds a certain threshold.



(a) (b)

FIGURE 2.1: (a) Picture of a SiPM [83]. (b) Schematic of a single pixel in a SiPM.

Silicon Photomultiplier Performance Parameters

A few important properties of SiPMs are highlighted in this section that needs to be taken into account when observing the signal generated in a SiPM. The *Gain* of a SiPM is the amplification of the initial electron-hole-pair. It is proportional to the applied overvoltage V_{ov}, which is defined as the operating voltage, V_b above the nominal breakdown voltage V_{bd}. By definition: $V_{ov} = V_b - V_{bd}$. The capacitance of a SiPM pixel C_p, is typically around a few pF and depends on the geometry and doping profile of the pixel. The gain of the silicon photomultiplier is expressed in units of the elementary charge. It is determined by dividing the accumulated measured charge Q, by the charge of the photo-electron e that triggered the avalanche breakdown.

$$Gain = \frac{Q}{e} = \frac{C_p}{e} \cdot (V_b - V_{bd}) \tag{2.1}$$

Temperature dependence is another important parameter as its dependence on the breakdown voltage which influences the gain of the SiPM. A higher temperature leads to a higher breakdown

voltage of the device, resulting in a lower overvoltage and thus a lower gain. In particular, the SiPMs show a negative dependence on the temperature, typically about -1 K, depending on the value of V_{ov}.

As the amount of pixels in a SiPM is finite, and a single pixel has a dead-time of up to a few hundred ns, *saturation effects* occur if the number of incoming photons within this time scale gets close to the total number of pixels N_t of the SiPM. For N_i number of photons arriving at the SiPM, the number of fired pixels N_f can be approximated to first order by

$$N_f = N_t \left(1 - exp\left(\frac{-\epsilon \cdot N_i}{N_t}\right)\right), \tag{2.2}$$

with ϵ being the **P**hoton **D**etection **E**fficiency (PDE) of an incoming photon leading to a fired pixel:

$$PDE = Q_e \times P_t \times G_F \tag{2.3}$$

The quantum efficiency Q_e is the probability that the photon impinging the SiPM results in electron-hole pair creation. P_t is the Geiger avalanche probability that the generated electron-hole pair initiates a self-sustained avalanche. The PDE also depends on the geometrical fill factor G_F, expressed as the ratio of the active to total SiPM area.

Figure 2.2 shows the measurement of the SiPM (MEPhI[1]/Pulsar) response for different pixel numbers.

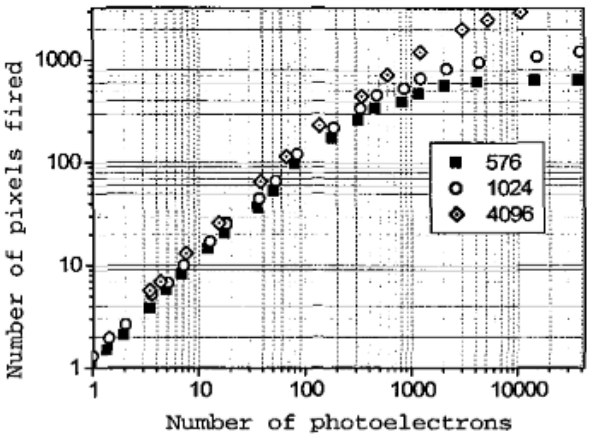


FIGURE 2.2: Non-linear response of MEPhI/PULSAR SiPMs with different pixels numbers. The light signal is produced by a fast laser (40 ps) [84].

The parametrization given in Eq. 2.2 can describe the data very well at low light levels. However, this can significantly differ at higher light intensities, which leads to saturation effects that needs to be corrected offline, which is explained in section 6.2. The dynamic range of the SiPM can be enhanced by choosing a model with a particularly large number of pixels. However,

[1]Moscow Engineering and Physics Institute, Moscow, Russia

this usually comes at the cost of a lower fill factor, i.e. only a certain part of the active area is sensitive to light, and thus a lower PDE.

The SiPM is subjected to thermal noise, which scales with bias voltage and temperature. The thermal noise generates random SiPM signals, by initiating avalanche breakdowns without an actual photon triggering the avalanche, which is called the *dark-count*. The number of dark counts per unit time is the dark count rate. This is the rate at which a single pixel of the SiPM fires without a photon reaching the pixel and it is typically at the order of a few MHz.

When a SiPM produces a signal, two main effects can cause additional pixels to trigger an avalanche. Firstly, a *cross-talk* between pixels on the SiPM can cause neighbouring pixels to trigger an avalanche as well. Secondly, *afterpulses* may occur after a pixel was fired and trigger another avalanche in the same pixel.

There have been remarkable improvements in the manufacturing of SiPMs. The PDE range has been improved to be sensitive down to near-ultraviolet light (300-400 nm) and up to infrared light (800-1000 nm). The dark noise rates have been reduced to as low as tens of kHz at room temperature. In addition, the introduction of trenches between pixels in the substrate has reduced the cross-talk probability.

2.2 AHCAL Technological Prototype

The *AHCAL Technological Prototype* is a highly granular sampling hadron calorimeter designed to absorb and detect incident particles. It is optimized to fulfil the need for a PFA, aimed at the detector systems foreseen at the ILC and other future lepton colliders.

The AHCAL technological prototype is the first large scale particle physics detector using SiPMs. A particular SiPM model built by Hamamatsu, MPPC S13360-1325P, is used and the relevant specifications of this SiPM type is provided in table 2.1. It is sensitive to blue light, which was not the case in the previous SiPM generation used in the AHCAL physics prototype. The dark rate of the new SiPM was dramatically reduced from 500 kHz in the old generation of SiPM to 20 kHz. In addition, the cross-talk was also reduced from 30% to 1% in the new generation. With this improvement, the photon detection efficiency and the signal to noise ratio were significantly increased. As a result, the new SiPMs has excellent uniformity at the operating voltage and gain.

Sensor	Effective area $[\mathrm{mm}^2]$	Pixel size $[\mu\mathrm{m}]$	Pixels	PDE $[\%]$	Gain	DCR $[\mathrm{kHz/mm}^2]$	V_{bd} $[\mathrm{V}]$
S13360-1325PE	1.3×1.3	25	2668	30	7.0×10^5	41	51.1

TABLE 2.1: Manufacturer specifications of the Hamamatsu MPPC S13360-1325PE [85]

The SiPMs are individually coupled to plastic scintillator tiles. The SiPM is placed inside the dimple, which helps to couple the SiPM to the scintillator and significantly enhances the

uniformity of the collected light across different impact positions on the tile. The combination of a scintillator tile and a SiPM is called a single channel. In addition, the scintillator tiles are wrapped in reflective foil, as shown in figure 2.3, to prevent photons from escaping.

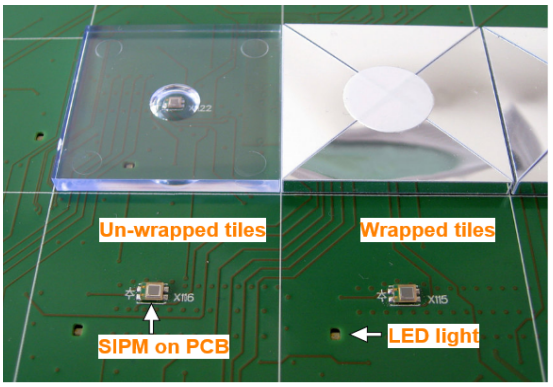
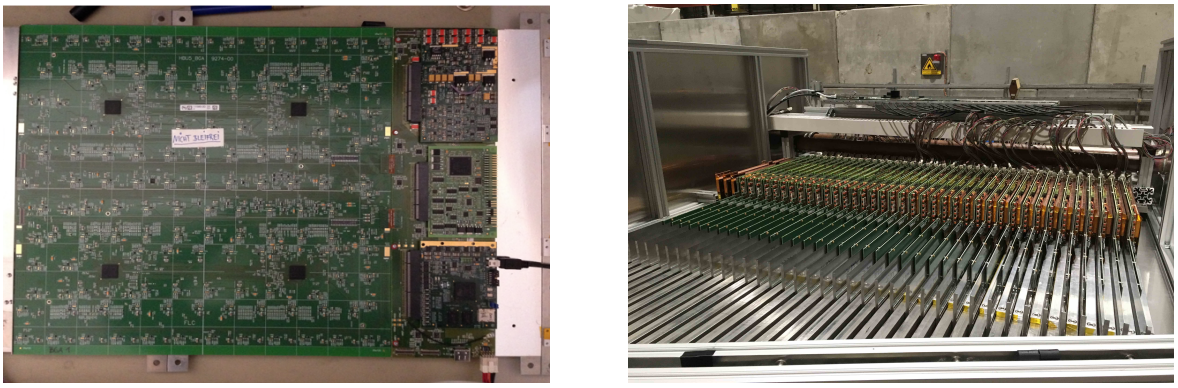


FIGURE 2.3: Picture of the CALICE AHCAL scintillator tiles with central dimple mounted over the SiPM on the HBU [86].

The active layers of the AHCAL consist of **H**adronic Calorimeter **B**ase **U**nits (HBU) with a side length of 36×36 cm^2. A photograph of the front side of an HBU is shown in figure 2.4(a). Each of the HBUs is equipped with Silicon Photomultipliers, soldered onto the front of the boards. The combination of four HBU's makes a full AHCAL layer. In addition, the HBU serves a **L**ight **E**mitting **D**iode (LED), which can inject light pulses directly into the scintillator tile in order to perform the SiPM gain calibration, as described in section 4.2.



(a) (b)

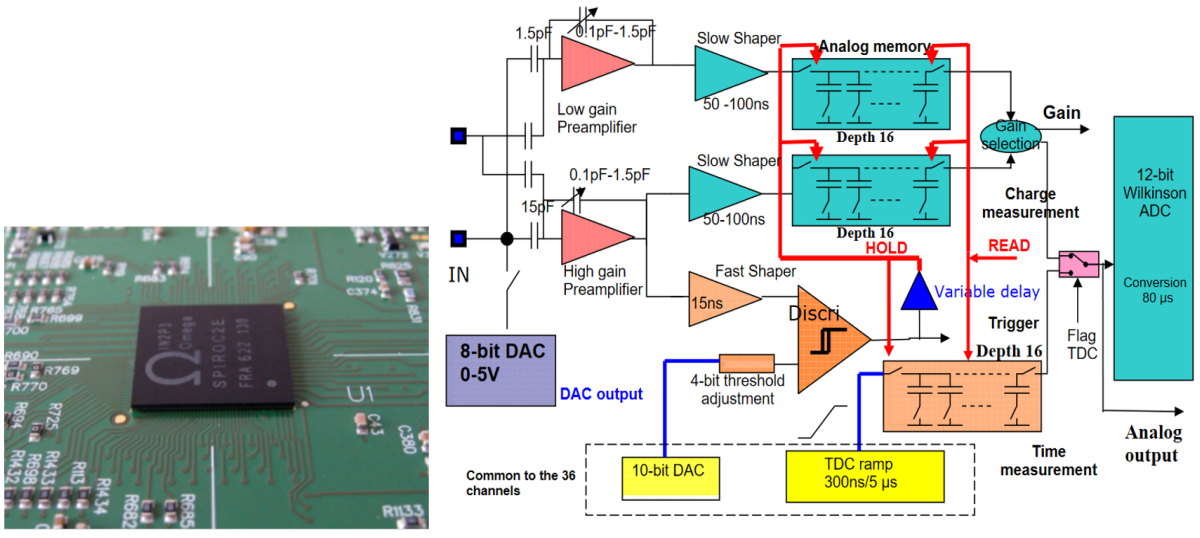
FIGURE 2.4: (a) Frontside of an HBU with the four readout ASICs (black) [86]. (b) AHCAL steel stack with active material placed in gaps between absorbers.

The active layers are interleaved with steel absorber plates which function as a passive layers. The prototype consists of 38 active layers inserted into a structure made of non-magnetic stainless steel absorber plates 1×1 m^2 wide and 17.2 mm thick on average. A picture of the detector with all the active layers inserted into the steel absorber stack is shown in figure 2.4(b). Within the layers, no cooling system is foreseen to keep the calorimeter compact while minimizing the

necessary size of the magnet coil. The calorimeter has a total depth of ∼4.5 nuclear interaction lengths. Active layers consist of a steel cassette housing 576 scintillator tiles connected on a **P**rinted **C**ircuit **B**oard (PCB). The tiles are $3\times3\times0.3$ cm^3 resulting in 21,888 channels in total.

The readout of the SiPMs are mediated by **SiPM I**ntegrated **R**ead-**O**ut **C**hips (SPIROC), which are mounted onto the backside of the HBUs. These SPIROCs provide a readout as well as digitises several SiPMs, recording the measured charge and timing information.

The picture of the SPIROC2E and the schematic layout of one channel of the SPIROC2E chip is shown in figure 2.5(a) and figure 2.5(b), respectively. SPIROC is explicitly designed for a SiPM readout for an HCAL at the ILC. One chip can operate 36 input channels. As active cooling should be minimized in the AHCAL, one of the requirements for the ASIC is a power consumption below 25 μW per channel. To enhance the dynamic range between 1 to several thousand photoelectrons, the SiPM output is amplified by two different configurable low-noise charge sensitive pre-amplifiers. These are **L**ow-**G**ain (LG) pre-amplifier for large input signals and a **H**igh-**G**ain (HG) one for small signals. After passing a slow shaper, that has nominally 50-100 ns shaping time and a fast shaper of 15 ns shaping time, the signals are stored in an analog memory array with a depth of 16 memory cells. They are implemented where the charge measurement and time information are stored independently until a 12-bit Wilkinson ADC digitizes them. Depending on the height of the initial SiPM signal, the high-gain or the low-gain value is selected for digitization in a 12-bit **A**nalog to **D**igital **C**onverter (ADC) and subsequently forwarded to the higher-level **D**ata **A**cquisition (DAQ) systems. Depending on the threshold condition, the chip assigns the hit information either with a hitbit equal to one (physics signal) or zero (pedestal).



(a) (b)

FIGURE 2.5: (a) The picture of SPIROC2E. (b) Schematic layout of one channel of the SPIROC2E chip. Taken from [87]

The chip can be operated in **E**xternal **T**rigger (ET) mode or **A**uto **T**rigger (AT) mode. In an external trigger, the signal of each cell is sampled synchronously to an external signal. This mode is used to measure the SiPM gain using the integrated LED system of the AHCAL, which is described in section 4.2. A fast shaper and a discriminator provide the auto-trigger capabilities of the SPIROC. In addition, an external validation trigger can be provided to the ASIC to reduce SiPM noise. The time measurement is digitized using a 12-bit **T**ime to **D**igital **C**onverter (TDC). The time measurement is stored in another memory cell when there is a trigger. The SPIROC ASIC has two multiplexed TDC voltage ramps to avoid dead time between each clock cycle. Once the 16 memory cells are filled, no further hits can be stored. Hence, the memory cells are readout, digitized, and the data is transferred out of the chip. The chip can be operated in a power-pulsing mode where parts of the chips that are not needed in any given state of operation, can be switched off. The tight power budget of a readout chip for a highly granular calorimeter requires optimal use of the available power. This is realized by exploiting the unique time structure of future linear colliders. With particle bunches organized in bunch trains extending over about 1 ms and a spacing between the bunch trains of about 200 ms, the readout electronics is idle for about 99% of the time. The concept of power pulsing is a rapid power cycling of the active components to power only the parts currently needed for readout and processing of the data.

2.3 Detector Simulation

The detailed detector geometry of the technological prototype is modelled using the most recent DD4HEP framework [88] as a detector description toolkit for High Energy Physics experiments. This toolkit implements a modular and flexible approach to the simulation activities using `GEANT4`. In the DD4HEP framework, a full detector description is provided, including geometry, materials, visualization, readout, alignment, calibration, etc. A right-handed coordinate system is used. The z-axis points along the beam direction and the y-axis is directed upwards. The main parts of the detector model are the absorber plates and the active layers of the AHCAL. The absorber stacks are simulated as a series of steel plates (17.2 mm) for the steel stack. Other elements in the beamline besides the AHCAL are modelled, including the trigger scintillators or the Čerenkov detectors. An additional layer of 2.0 mm thickness of steel (corresponding to half an X_0) is added upstream of the calorimeter in order to account for missing material. This additional material was determined using the electron data and comparing the simulation with the center of gravity distribution in the z-direction. The AHCAL is located in reference to the beamline at 0 m. The three trigger scintillators, with dimensions of 10×10 cm^2, 10×10 cm^2 and 50×50 cm^2, are placed right before the detector to check that the beam particle has deposited energy. The Čerenkov detector has been filled with the Helium gas with a density of 0.0001663 g/m^3 and placed at a distance of -40 m from the AHCAL.

The material description is shown in table 2.2 in terms of thicknesses and the corresponding radiation length and nuclear interaction length.

Material	Thickness (mm)	X_0	λ_I
steel absorber	17.2	0.977	0.101
air gap	0.257	8×10^{-6}	4×10^{-6}
steel cassette	0.5	0.028	0.003
cable mix	1.50	6×10^{-4}	2×10^{-4}
PCB	0.7	0.004	0.001
foil	0.115	3×10^{-4}	3×10^{-4}
Polystyrene	3.0	0.007	0.009
AHCAL layer	26.08	1.05	0.11

TABLE 2.2: AHCAL material description in DD4HEP simulations of the testbeam setup at CERN in June 2018. Composition and properties of the materials are obtained from [10].

In scintillators, simulations of the energy deposited by an incoming particle and the resulting scintillation light produced saturate due to high ionization densities. This results in the light yield L being non-linear, and the effect is taken into account in the simulation via the application of Birk's law. It describes the amount of light yield per unit of length $\mathrm{dL/dx}$ for high ionization densities $\mathrm{dE/dx}$, as described:

$$\frac{dL}{dx} = \frac{S \cdot \frac{dE}{dx}}{(1 + kB \cdot \frac{dE}{dx})}, \tag{2.4}$$

where S denotes the scintillator efficiency, kB is the Birks constant of the medium, k is the quenching parameter, and B is the proportionality constant. In principle, kB is an adjustable parameter to be tuned with the experimental data, for a specific scintillator with S as the absolute normalization. The total light output produced by a particle of an energy E is obtained by integrating the above equation.

The default value for Birk's constant, kB in `GEANT4` simulations is 0.07943 mm/MeV [89], measured for the polystyrene-based scintillator used in the ZEUS experiment. The parameter kB is influenced by the composition of the scintillator material. In the AHCAL simulations done for this thesis, the value for kB = 0.151 mm/MeV was used, which is the value measured for the AHCAL scintillator material [90]. This is almost twice the default value in Geant4. This value is needed for the conversion factor from GeV to MIP, such that the MPV of the muon hit energy distribution peaks at 1 MIP as described in section 2.4.1.

2.4 Digitization

Comparing measurements to theoretical predictions or simulations requires a good understanding of physics signals and detector effects. Hence, it is necessary to include these effects in the simulation. The reconstruction should be identical both for data and simulation. This requires the transformation of the raw simulation output into a structure that is equivalent to the raw data format. This additional step included in the reconstruction chain of the simulation events is digitization. This needs to include appropriate and detailed modelling of detector characteristics such as signal generation processes affecting the measurement. This section describes, in short, the necessary digitization steps needed to resemble the technological prototype design. The digitization is implemented within the **M**odular **A**nalysis and **R**econstruction for the **Lin**ear Collider (MARLIN)[2] framework [91]. After the digitization process, simulated data are treated the same way as real data for subsequent comparison and analysis. The digitization is tuned by comparing the response to muons in data and simulations.

The AHCAL digitization involves several steps which account for the cell dimensions, the conversion from the GeV to the MIP scale, signal shaping, ganging and the non-linearity in SiPM response. A sketch of the digitization chain is shown in figure 2.6.

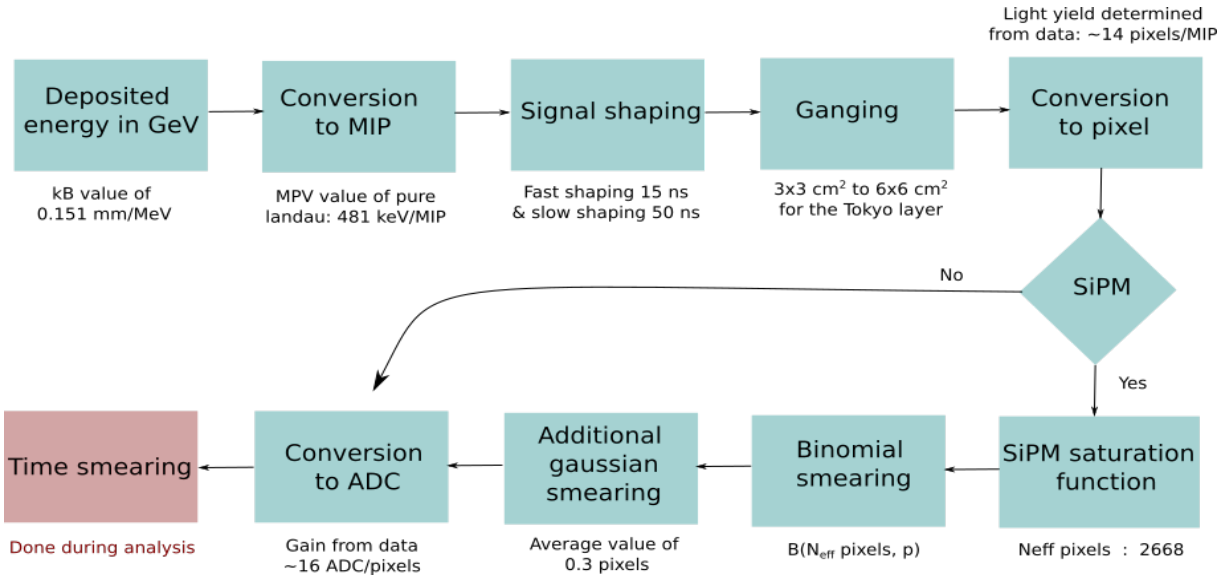


FIGURE 2.6: The AHCAL digitization steps of the simulated data. The details of every step is described in the text.

2.4.1 Conversion to MIP

The first step of digitization is converting hit energy from GeV to the energy scale, equivalent to the energy deposited per hit of a MIP. This is achieved by simulating a muon beam in the calorimeter that is projected perpendicularly onto the calorimeter cell. The hit energy distribution

[2]Marlin is a Linear Collider software framework for analysis and reconstruction code based on LCIO files.

is fitted using a landau distribution and the most probable value is determined. The hit energy distribution of simulated muons is shown in figure 2.7. The extracted maximum value of this fit for the hit energy distribution is 1.002 MIP. The conversion factor which has been determined for the AHCAL prototype using 2018 simulated data corresponds to 488 keV/MIP.

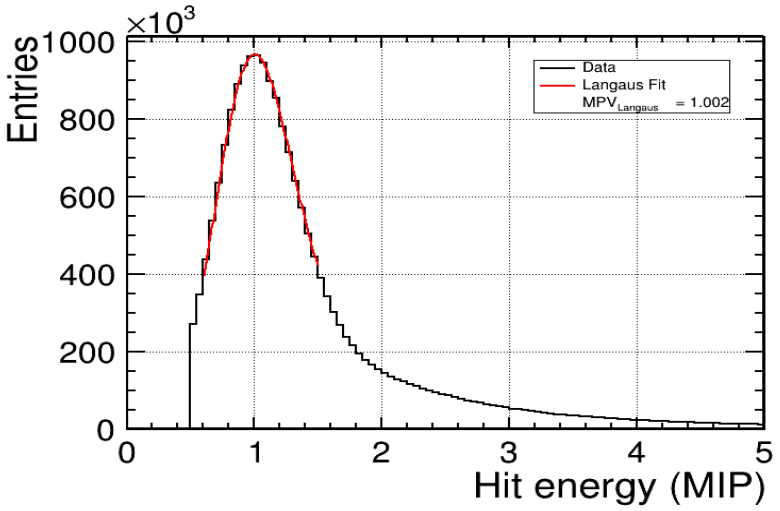


FIGURE 2.7: Landau-Gaussian fit applied to the hit energy distribution using simulated muon data with beam energy at 40 GeV. The fit is used for the adaptation of the GeV to MIP conversion factor in the simulation digitization. The solid red line indicates the range of the final fit. The extracted maximum value of this fit, i.e. the MPV of the hit energy distribution, is 1.002 MIP.

2.4.2 Signal Shaping

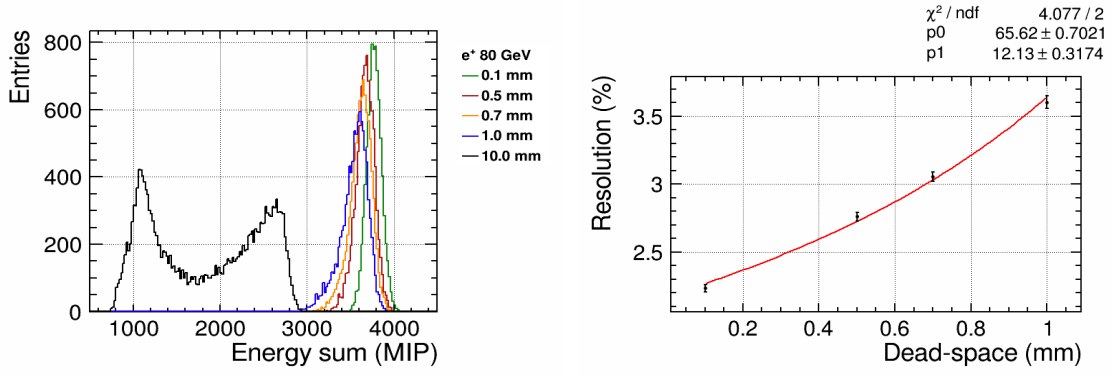
The signal shaping depends on the readout electronics that are used. Signal shaping includes summing up the energy of all sub-hits in a single channel within a time window of 15 ns. If the energy exceeds the trigger threshold, the time of the simulated hit becomes the time of the hit. This step in digitization is mainly necessary as slow neutrons generated in hadronic showers can deposit significant amounts of energy, on much longer time scales up to the order of seconds after the initial particle shower has developed [63]. Energy depositions later than 50 ns after the generation of the primary particle are not included in the output. A function of convoluted Gaussians, further smears the hit time to account for resolution effects in the detector.

Tile Gap Study

A study was done to obtain the optimal tile gap to be included during the digitisation procedure. According to the production tolerances, the tile pitch was 30.15 mm with a gap of 0.5 mm. To determine the tile gap, different tile gaps ranging between 0.1 mm to 1.0 mm were simulated to see the effect on the visible energy and the effect on the energy resolution. The energy distributions extracted for each of these gaps are shown in figure 2.8(a) for a simulated 80 GeV electron. Based on these individual energy distributions, the energy resolution was calculated

using the mean and the width of the distribution within $\pm 2\sigma$. The obtained energy resolution is then plotted versus the tile-gap and is then fitted using the parametrization of the form given by:

$$R = \frac{p_0}{(30.15 - p_1 \cdot x)} \tag{2.5}$$



(a) (b)

FIGURE 2.8: (a) Energy sum distributions for an electron beam energy with 80 GeV, for different tile gap measurements in simulations. (b) Resolution as a function of tile-gap (mm) for electron with beam energy at 80 GeV. A fit is performed using the Eq. 2.5 as shown in red solid curve.

The result is shown in figure 2.8(b), indicating that a tile-gap of 0.7 is the best value to be used during the digitisation, which is the nominal number as measured in [92].

2.4.3 Ganging

The next step in digitization is ganging. The cell size of the main AHCAL detector is 3×3 cm^2 and the cell size of the Tokyo layer is 6×6 cm^2. This step includes the ganging of four cells of size 3×3 cm^2 to 1 cell with a size of 6×6 cm^2. The energy of the new segmented cell is the sum of the four cells, and the timing information is the shortest time over the four ganged cells.

2.4.4 SiPM Modelling

For simulating the SiPM response, the energy depositions in all tiles are converted from the MIP scale to the corresponding number of SiPM pixels, using the light yield measurements of the SiPM-tile systems. The number of pixels fundamentally limits the maximum number of photons detected with a SiPM. The number of effective pixels N_{eff} is corrected with an optimization parameter for the non-linear dependence, using each cell's measured SiPM response function. A MIP-like particle traversing a 3 mm AHCAL tile generates roughly 3000 to 5000 photons inside the scintillator tile. The SiPM does not detect all photons generated. However, because of the low number of detected photons, their distribution should be well modelled, assuming that the light intensity is proportional to the energy deposited in the tile. In the digitization process, the

pixel amplitude is computed for each hit in the calorimeter. With the adequate number of SiPM pixels as an optimization parameter. The total number of pixels N_t in Eq. 2.2 is replaced by N_{eff} and the ϵ for the SiPM's used in AHCAL is assumed to be 100% which results in:

$$N_f = N_{eff} \cdot \left(1 - e^{-N_i/N_{eff}}\right). \tag{2.6}$$

The resulting number of effective pixels is then smeared with a binomial distribution to account for the statistical fluctuation of the photons detected. The number of fired pixels n is randomly generated using a binomial distribution $B\left(N_{eff}, p\right)$. The input parameters are the number of effective pixels of the SiPMs N_{eff}, and the probability p that a pixel is fired, which is given by $p = N_f/N_{eff}$. Furthermore, Gaussian smearing is applied after the Binomial smearing to account for the SiPM noise. The details are explained in section 4.2. Since the prototype has a very low noise level, with a cross-talk below 1%, this is considered negligible, and no additional noise is simulated in the digitization process.

2.4.5 Conversion to ADC

After applying the SiPM features, the energy deposition is converted from the SiPM pixel scale to the electronic ADC scale using the gain calibration and gain inter-calibration factors measured for each channel individually. Alternatively, their default value is in the database if no calibration factor is available. Additionally, time smearing is done using Gaussian smearing of 1 ns.

At this level, the digitized simulation is comparable to the measured data. Furthermore, the same calibration procedures and analysis algorithms can be applied to both data and digitized simulations, excluding all hits with an energy of less than 0.5 MIP. An example of simulated hit energy distributed is shown in figure 2.9, with different levels of digitization steps before arriving at a distribution that is comparable to the reconstructed data. Moreover, digitization uses the inverse of the formulas used in the reconstruction procedure. The calibration values are identical between reconstruction and digitization.

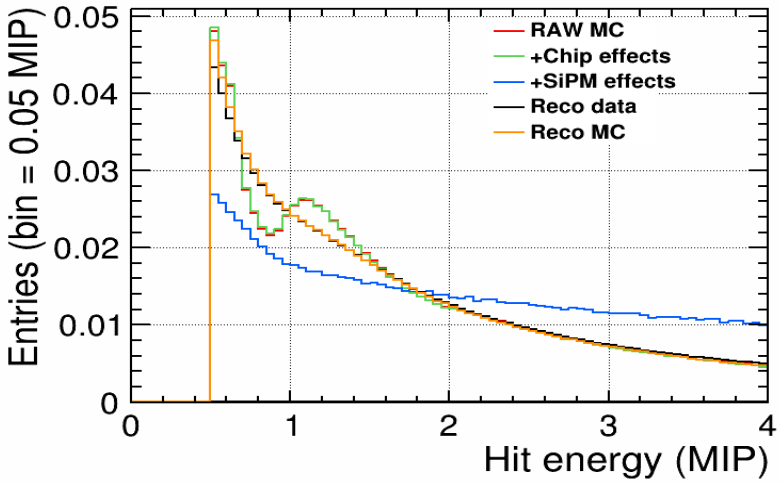


FIGURE 2.9: The simulation of different detector effects. The orange curve shows the final digitized hit amplitude.

The following chapters provide an overview of the analysis chain, which is illustrated in figure 2.10. A block represents a process, while arrows represent a flow of information between processes. This chapter has explained the two processes (generation and digitization) in the analysis chain. In chapter 3 the testbeam campaign will be introduced, followed by the calibration step in chapter 4. Furthermore, event selection will be described in chapter 5. Finally, the analysis of the simulated and testbeam data will be discussed in chapter 7.

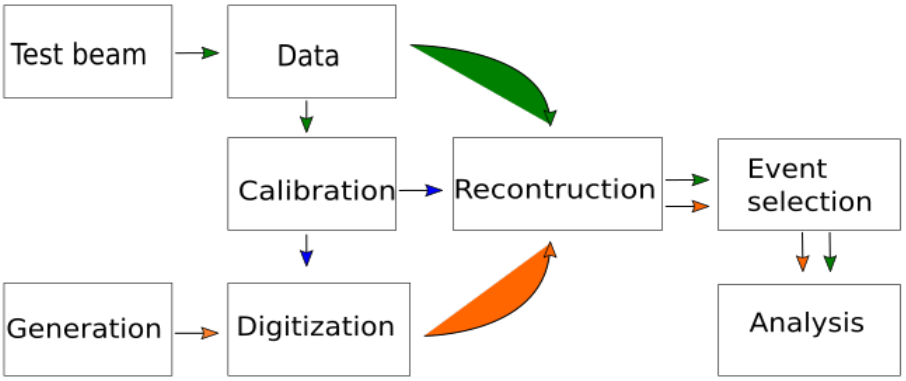


FIGURE 2.10: Overview of the analysis chain. The green lines represents test beam data. The orange lines are Monte Carlo simulated data. The blue lines are calibration data that is used for both reconstruction and digitization.

Chapter 3

The CERN-SPS Test-Beam

"Measure what can be measured, and make measurable what cannot be measured."

Galileo Galilei

THE CALICE-AHCAL introduced in section 2 was built and initially tested with cosmic rays and beams at DESY, subsequently moved to the CERN-SPS facility for two weeks in May and again for one week in June 2018, with an addition week of combined operation with the CMS-HGCAL in October. The main goals of these test beams were:

- Demonstrate not only the performance, but also the scalability and feasibility of the technology to a large detector.
- Exploit the energy and timing capabilities.
- Handling of the large number of channels and investigate hardware performance with new SiPMs.
- Collect muon, electron and pion data to perform the calibration of the detector, understand the behaviour of the detector, validate the **M**onte **C**arlo (MC) simulation, study the hadronic shower shapes and explore the physics potential.

In this chapter, before discussing the calibration procedures used for the detector, a brief description of the testbeam setups and beam instrumentation at the CERN-SPS facility will be introduced before concluding with a summary of the data samples recorded with the AHCAL technological prototype.

3.1 Beamline and Testbeam Setup

The AHCAL prototype was tested in the high energy and high-resolution H2 beam line [93] of the SPS North Area. The SPS accelerator delivers protons with energies up to 450 GeV/c and with a bunch population of approximately 10^{11}.

First, these protons hit a target producing different secondary and tertiary particles such as electrons, muons and hadrons. Then, with the help of dipole magnets, the resulting secondary

particles are deflected according to their momentum, mass, and charge. By using a block of lead, it is then possible to select the desired particles and filter out the rest using collimators along with additional absorbers. Figure 3.1 shows the vertical plan of the momentum selection.

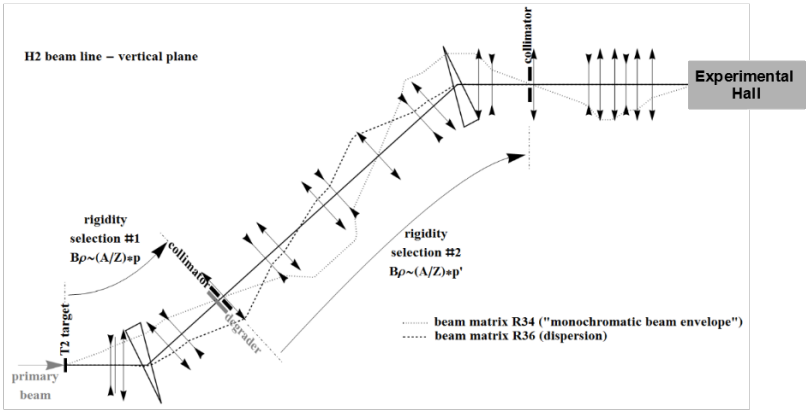


FIGURE 3.1: Schematic view of the vertical plane of the H2 beam line at CERN-SPS facility [94]. The beam line is more than 600 m long and the height difference between T2 and the hall is about 12 m (the picture is not to scale).

A muon beam is obtained via the decay of a pion beam. A beam of pure electrons can be generated from gamma conversion by directing photons onto a lead target. Alternatively, it is also possible to generate an electron beam by directing a charged beam onto a secondary target and filtering it with an absorber. However, this comes with a drawback of lower rates at low energy and contamination of the electron beam with pions. For pions, the momentum is selected via a set of collimators and magnets. The setup starting with the particle gun is set 50 m upstream from the detector and the particles enter in the z-direction, that is perpendicular to the surface area of the AHCAL prototype. The Čerenkov detector is 40 m upstream of the detector and was used to select the desired particles. Further ahead were the delay wire chambers with dimension of $10\times10\ \mathrm{cm}^2$.

It is relatively easy to filter out electrons and pions from a muon beam but rather challenging to filter out muon contamination from electron and pion beams. Although electrons shower at the beginning of the detector volume, some electron contamination still exists in low energy pion beams. Electrons can also be selected from secondary beams using a Čerenkov counter, leading to a momentum dependent purity of $10\%-99.5\%$. A Čerenkov threshold counter detector is installed in the beamline with the possibility to set a threshold needed for particle identification, which can be used for online particle selection and reduction of particle contamination in the recorded beam data. In addition, the beamline is also equipped with many beam instrumentation detectors. These include two large spectrometers for particle selection based on their rigidity, wire chambers and scintillators. They provide information on the beam position and the number of particles.

The three testbeam campaigns had slightly different setups. In the May 2018 testbeam campaign, the AHCAL prototype was equipped with 38 active layers installed between the steel absorber gaps. While the core detector remained the same, for the June testbeam campaign, a layer from Tokyo, which had a larger tile size of $6\times6\ \mathrm{cm}^2$ was installed at the end of the detector to verify the feasibility of a varying layer granularity. Furthermore, an additional layer was installed as a **P**re-**S**hower (PS) consisting of one HBU in the front of the AHCAL prototype without a layer of steel. The effect of installing a PS detector increases the effective depth of the calorimeter. This was mainly installed to reject the early showering particles and for a veto. Aside from this, a previous version of an AHCAL prototype consisting of twelve fully equipped single HBUs of $36\times36\ \mathrm{cm}^2$ alternating with 7.4 cm thick steel absorber plates were installed. The first absorber layer of tail catcher is 1.2 cm thick compared to rest of the absorber layers as shown in figure 3.3. These layers were placed at the end of the main stack and served as a TC to account for the hadronic shower leakage, allowing those events which were not fully contained longitudinally in the layers of the main stack. Finally, in the October testbeam, the AHCAL prototype with 39 active layers served as a TC, combined with the prototype with the silicon part of the CMS-HGCAL in front [95].

For the May and June testbeam campaigns, the entire stack was placed on a movable stage, scanning the whole detector with muons in different x-and y-positions. While for the October testbeam, the detector was placed on a fixed platform. Figure 3.2 shows a picture of the AHCAL technological prototype during the June testbeam in the H2 beamline.

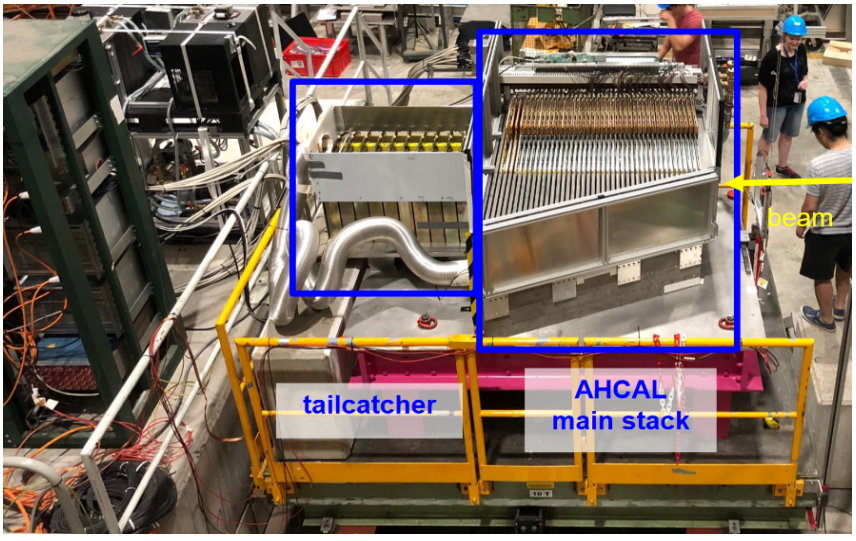


FIGURE 3.2: AHCAL technological prototype setup of the June 2018 testbeam campaign at the SPS-CERN. The AHCAL main stack is placed in front of the TC. The setup is mounted on a movable platform.

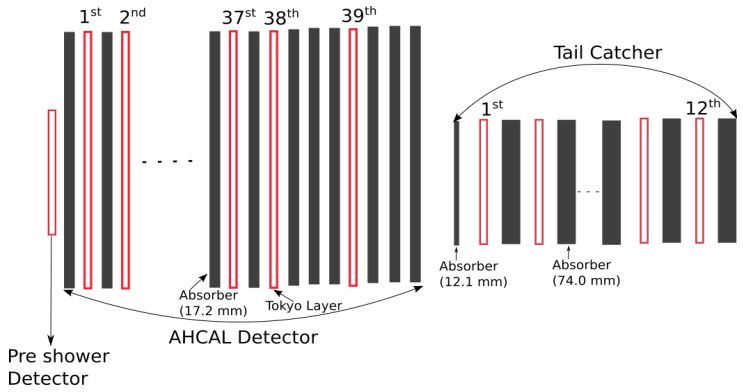


FIGURE 3.3: Schematic of the AHCAL prototype and Tail-Catcher layer configuration of the June 2018 testbeam campaign at the SPS-CERN.

The setup starting with the particle gun is set 50 m upstream from the detector and the particles enter in the z-direction, that is perpendicular to the surface area of the AHCAL prototype. A Čerenkov detector is 40 m upstream of the detector. Further ahead were the delay wire chambers with dimension of $10\times10\ \mathrm{cm}^2$.

Out of six delay wire chambers, four were used to enable positioning of the detector with respect to the beam with a resolution of about an mm, and for position studies during the data analysis. Moreover, four trigger scintillators with a thickness of 1 cm were used while data taking, of which two were small scintillators of side length $10\times10\ \mathrm{cm}^2$, and two large scintillators of roughly $50\times50\ \mathrm{cm}^2$. The coincidence of the $10\times10\ \mathrm{cm}^2$ scintillator plates is used for the electron and pion runs as a trigger signal. The two large scintillators were arranged as one in front and one at the back of the calorimeter. Their coincidence of a validation signal to the readout chip were used for the muon runs[1]. In addition, there were four triggers from the beam line instrumentation which were simulated as $10\times10\ \mathrm{cm}^2$ with 1 cm thickness, which were not used in the analysis.

The beam entered the AHCAL after traversing the last trigger scintillator placed in front of the PS. A schematic view of the beamline configuration is shown in figure 3.4.

[1]Collection of events taken over a period of time under a set of run conditions.

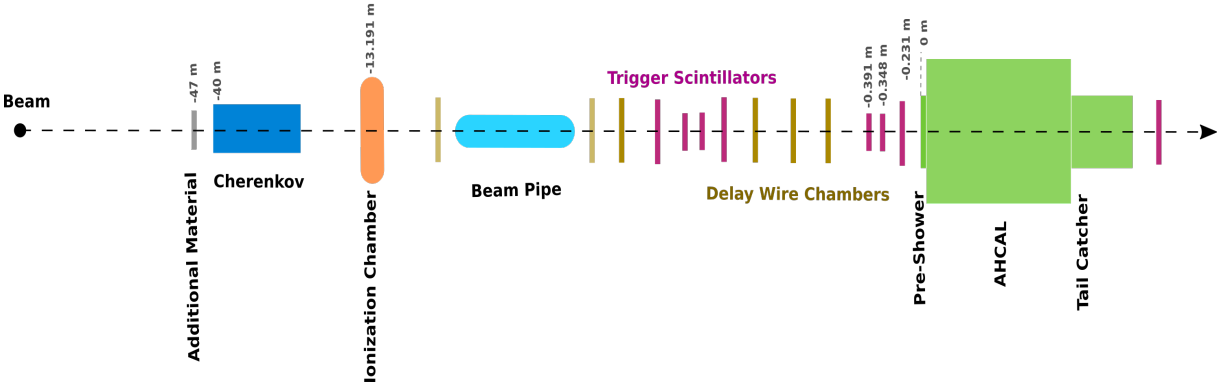


FIGURE 3.4: A schematic illustration of the beamline setup at the CERN SPS H2 beamline in June 2018. The black dashed trajectory indicates the hypothetical path of a particle ($\mu^{\pm}, e^{\pm}, \pi^{\pm}$), which first traverses an additional material, then beam instruments (Čerenkov detector, ionization chamber, delay wire chambers, beam pipe and trigger scintillators) and then lastly enters the pre-shower and analog hadron calorimeter. The pre-shower is the origin where the z-axis is defined along the beamline.

3.2 Dataset

The dataset for the May 2018 test beam included runs with and without power pulsing mode. This was done to validate the power pulsing mode. Several position scans with and without power pulsing for the entire detector were performed using muons with beam energy at 40 GeV and 120 GeV. Alternatively, short and long LED scans also with and without power pulsing were done daily. This was done to calibrate every single channel and monitor the stability of the detector during beam operations. To study the electromagnetic and hadronic response of the detector at different beam energies, a large number of events with positron beam scanning at energy range of 10 to 100 GeV in 10 GeV steps and negative pion beams in the energy range from 10 to 160 GeV were collected.

For the June 2018 and October 2018 testbeam, the entire dataset was collected with power pulsing mode. Muons of 40GeV energy, positrons between 10 GeV and 100 GeV and negative pions between 10 and 200 GeV were collected. In addition, several shifted positions were used for particle separation studies.

During the various testbeam campaigns at the CERN-SPS facility, a large number of datasets in a total of about 10^7 beam events [2] were recorded for particle beams consisting of muons, positrons and negative pions. For the central analysis of this thesis, only beam data from June 2018 testbeam was used, as the dataset from the May 2018 teatbeam was heavily contaminated due to upstream showering events.

The LED runs used in this thesis for gain calibration with the various light intensities can be found in table 3.1. The run numbers for different beam energies from the June 2018 testbeam with different particle species used in this thesis is given below.

[2] Collection of hits that is detected in the calorimeter, and are read out when a trigger occurs.

Dataset	Voltage (mV)	Mode	Run number
May 2018	0, 5400, 5700, 6000	No Power Pulsing	60093-96 60081-84 60211-214 60302-305 60361-364 60413-416 60481-484 60533-536 60543-546 60623-626 60716-719 60776-779
June 2018	0, 5000...7000 in steps of 20 mV	No Power Pulsing	60990-61103
October 2018	0, 5000...6880 in steps of 20 mV	Power Pulsing	62422 - 62523

TABLE 3.1: Run lists used for the LED gain calibration from May 2018 - October 2018 testbeam periods.

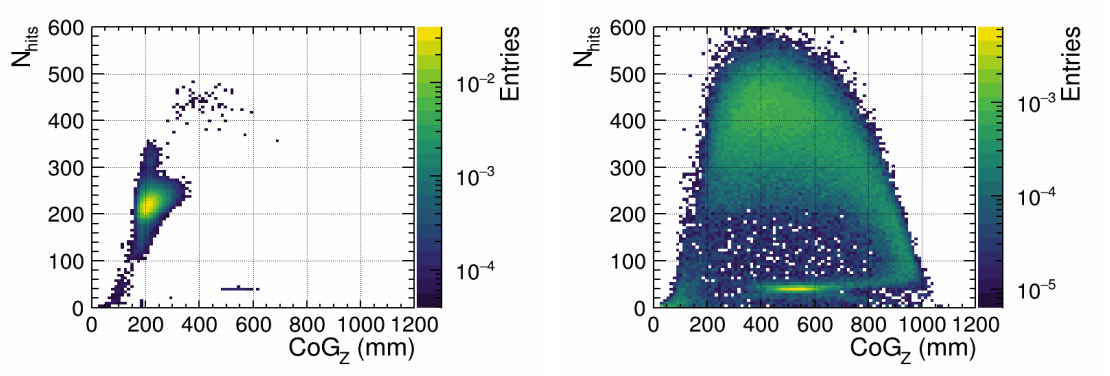
Particle	Energy [GeV]	Runs	Total Events
μ^-	40	61293	158,829
e^+	10	61293, 61202	102,934; 102,977
	20	61296	53,722
	30	61213,	201,984
	40	61212	202,195
	50	61214	202,347
	60	61211	201,970
	70	61215	201,271
	80	61156, 61208, 61210	201,984; 200898; 201984
	90	61216	201,822
	100	61217	168,525
π^-	10	61265	115,053
	20	61273	94,016
	30	61384	49,910
	40	61275	128,519
	60	61262	111,594
	80	61279	99,293
	120	61287	101,083
	160	61222	78,200
	200	61201	95,753

TABLE 3.2: Datasets from June 2018 testbeam recorded at the CERN-SPS facility.

A standard AHCAL data quality monitoring plot is shown in figure 3.5(a) for an electron shower. Figure 3.5(b) shows a pion-induced shower with beam energy at 60 GeV using the data from June 2018 testbeam. Figure 3.5(c) is for a 60 GeV electron-induced shower using the data from the May 2018 testbeam. The data quality plot, displays the number of hits distribution $N_{hits} = \sum_l N_l$. This represents the sum of all hits in all layers as a function of energy-weighted **C**enter **o**f **G**ravity (CoG) in the z-direction. The event-wise CoG of a shower gives the position of the shower's core. The Z coordinate is defined as given below (and similarly for the X and Y coordinates):

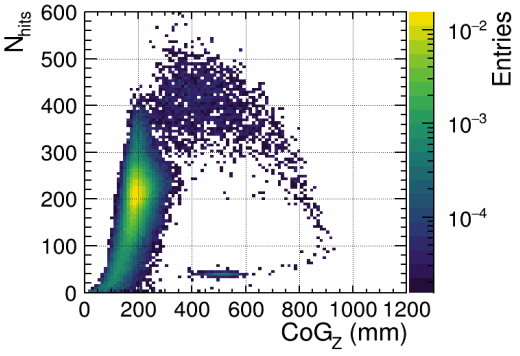
$$\mathrm{CoG_z\ [mm]} = \frac{\Sigma_\mathrm{i} \mathrm{E_i z_i}}{\Sigma_\mathrm{i} \mathrm{E_i}}, \tag{3.1}$$

where, $\mathrm{E_i}$ is the energy of the i-th hit and $\mathrm{z_i}$ is the z-position of the i-th hit.



(a) 60 GeV e^+ June 2018

(b) 60 GeV π^- June 2018



(c) 60 GeV e^+ May 2018

FIGURE 3.5: Distribution of the number of hits, N_{hits} and center-of-gravity along the beam axis, CoG_Z for (a) 60 GeV electron run from June 2018 testbeam data which shower early and with slight contamination from other particle types. (b) 60 GeV pion run from June 2018 testbeam data. Pion run presents a wider distribution on average with more hits and deeper CoG_Z compared to electron run. Muon run in principle don't shower, so have a N_{hits} close to the number of active layers, and a CoG_Z close to the middle of the detector. (c) 60 GeV electron run from May 2018 testbeam data shows a large tail due to the contamination caused by upstream events.

It deals with an admixture of other particles dominating different regions of the calorimeter. Electromagnetic showers with a high energy deposition produce a narrow hit distribution and CoG in the first few layers. A narrow hit distribution with the CoG around the center of the detector is induced by the muons appearing as MIP, passing through the calorimeter. A wide distribution of the CoG-Z and the number of hits results from hadronic showers caused by the pions being spread throughout the calorimeter. A tendency toward energy leakage occurs when the starting point of the high energy shower is deep inside the calorimeter, leading to a decrease in the number of hits. In this case, a TCMT is then added to fully contain the hadronic showers produced by energetic particles.

As mentioned previously, the electron dataset from May 2018 had a clear tail to a smaller number of hits and an earlier CoG. This behaviour was present for all electron energies. However, these tails were not present for the June 2018 testbeam because the beam steering was changed,

and the quality was improved. This is clearly shown by the quality plot in figure 3.5(a), showing a reduced tail in the June testbeam compared to the electron in the May testbeam as shown in figure 3.5(c).

Figure 3.6 shows the event display for a muon, an electron shower and a pion shower in the AHCAL technological prototype.

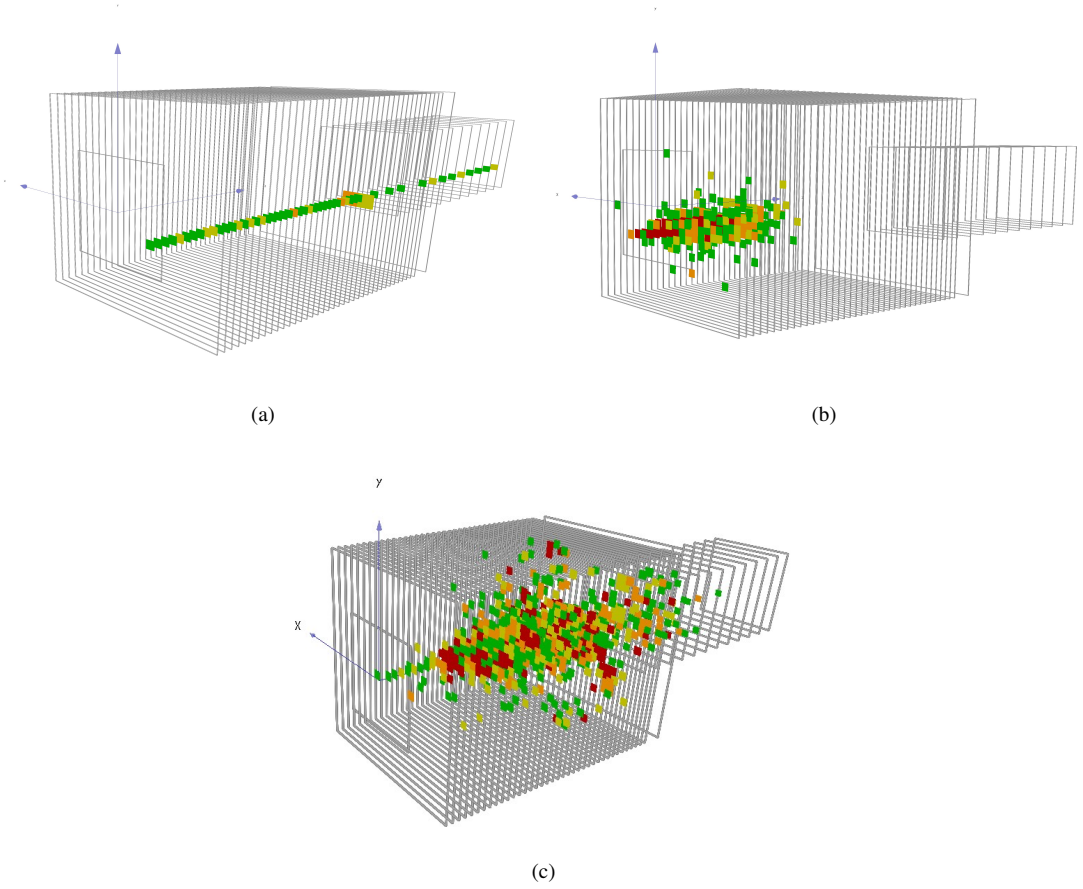


FIGURE 3.6: Examples of event displays for the events taken during the June 2018 testbeam data (a) 40 GeV muon, (b) 80 GeV electron, and (c) 200 GeV pion. The colors corresponds to hit amplitudes, increasing from green (0.5 MIP $\leq$ hit energy $\leqslant$ 3.0 MIP) to red (hit energy $\geq$ 5.4 MIP). The particles enter the detector from the left in the z-direction perpendicular to the layer surface.

In this chapter, the AHCAL experimental setup operated at the SPS-CERN beamline has been described. In addition, the data collected and used in this thesis has been discussed. In the following chapter, the energy calibration of the AHCAL technological prototype is presented.

Part III

Data Analysis

Chapter 4

AHCAL Calibration

"Every line is the perfect length if you don't measure it."

Marty Rubin

THE OBJECTIVE OF THE CALIBRATION is to establish a relation between the cell signals and the energy of the particle to be measured. The calibration of ≈22000 channels in the AHCAL technological prototype is already challenging. Hence this would be a non-trivial task for calibrating a multi-channel calorimeter designed for ILC.

This chapter introduces the calibration procedure developed for the AHCAL technological prototype using robust algorithms to extract calibration data.

Introduction

For trustworthy data measurements with the AHCAL technological prototype, we have to establish a reliable and robust calibration chain. This requires several measurements with beam particles to get the response of a MIP and using light from the LED to get the SiPM gain. Therefore, the calibration of the AHCAL prototype is needed for several purposes:

- The conversion of the hit energy and time from the electronic scale to physical units.
- If necessary, the equalisation of the cell response to allow an efficient auto-triggering in all channels.
- Monitoring the stability of the detector.

Every channel has to be calibrated separately due to non-uniformities such as unequal tile wrapping, glueing or SiPM and ASIC effects. The digitised signal from the SiPM is expressed in units of ADC bins. To perform offline analysis, the data needs to be converted to a physics scale. The chosen physics scale is the MIP, defined as a signal that leaves minimum energy in the detector as it travels perpendicular through the AHCAL tile. This conversion of raw signals is done during the reconstruction step.

Two calibration chains are needed, one for the energy calibration and the other for the time calibration, as is illustrated in figure 4.1.

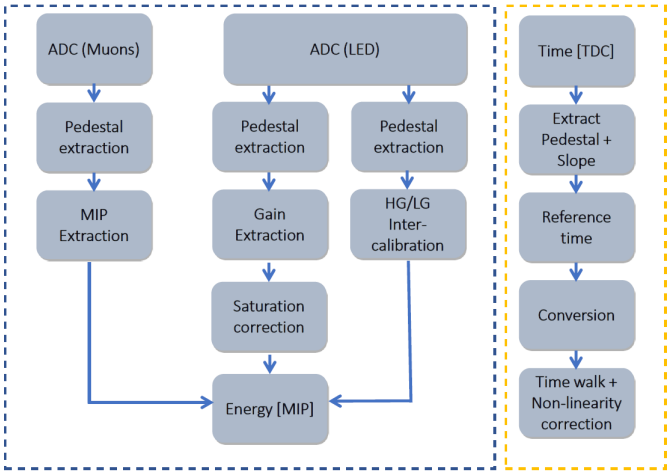


FIGURE 4.1: Overview of steps followed in order to obtain calibrated energy and time. The arrows indicate the direction of steps followed to acquire calibrated data.

The AHCAL calibration chain relates the hit energy or SiPM charge in units of ADC to the sum of visible energy E_i in units of MIP, where one MIP corresponds to the most probable energy deposition by a minimum ionizing particle. In addition, the time is stored as a charge TDC and is calibrated into values of nanoseconds. This relation can be determined by:

$$E_i = \frac{(a_i - p_i) \times I_i}{M_i} \times f_i^{-1}(A_i), \tag{4.1}$$

where,

$$A_i = \frac{(a_i - p_i) \times I_i}{G_i}.$$

Parameter	Meaning
E_i	Calibrated hit energy in units of MIP
a_i	Measured cell amplitude in units of ADC
p_i	Pedestal value in units of ADC
I_i	Inter-calibration factor between HG & LG
M_i	MIP amplitude in units of ADC/MIP
f_i	SiPM response function
A_i	Cell amplitude in units of pixels
G_i	SiPM gain in units of ADC/pixel

TABLE 4.1: Parameters used for the energy calibration

For each channel and memory-cell i, the pedestal p_i is subtracted and is measured during the testbeam in the absence of beam. For the signal conversion to pixels and de-saturation, the signal

is divided by the SiPM gain G_i, which is the average number of ADC counts per pixel extracted from the LED measurements. I_i is the inter-gain calibration factor obtained by measuring the SiPM responses in HG and LG modes. The SiPM response is linearized by applying the function f_i, which is the inverse of the SiPM response function, describing the SiPM output signal as a function of the incoming light intensity, yielding the reconstructed number of photons hitting the SiPM. At small amplitudes, the value of f_i is close to one and increases strongly for large signals. The reconstructed number of photons is converted to the MIP scale using the channel's light yield, which will be discussed in section 4.3. As a final step in calibration, cluster energies will be converted to the GeV scale using electron testbeam data of known energies.

In the following sections for completeness, a set of calibration quantities and their measurements are described in further detail. At the end of the section, the calibration of the hit time measurement will be described. The major contribution of this thesis to the calibration of the detector is associated in performing the gain calibration of the AHCAL using the data from the SPS-CERN 2018 testbeam campaigns which is discussed in section 4.2 and the SiPM saturation correction is discussed in section 6.2.1. All the calibration constants are stored in the official CALICE database are accessible within the collaboration. The location of the individual folders and the database tags used in this thesis are provided in appendix A.

4.1 Pedestal Extraction

The pedestal is a value which determines the baseline of a measurement scale. Measuring the baseline will usually result in a Gaussian distribution with the mean value of the Gaussian corresponding to the pedestal value and the width relating to the electronic noise.

For each channel and memory cell, the pedestal values are extracted using LED and muon runs for gain and MIP calibration, respectively. In the following, the mean value of the pedestal histogram will be used. This value is defined as the pedestal. The width of the pedestal distribution is the noise of the channel.

Pedestal in external trigger mode is extracted from LED measurements using a dedicated pedestal run with the LED voltage set to 0 mV. For each channel and memory cell, a histogram is filled with ADC value larger than 150. The pedestal is computed by fitting a Gaussian function to the ADC distribution. The mean of the Gaussian distribution is defined as a pedestal of that channel. This is used for pedestal subtraction on a channel and memory-cell level. The width of the ADC distribution is the electronic noise. Based on the quality cuts applied on the RMS of the ADC distribution it is possible to filter out the dead or noisy channels with a narrow or wide ADC distribution. The ADC distribution of one AHCAL channel is shown in figure 4.2(a) for the data taken during the CERN testbeam campaign. In some cases it is observed that the ADC distribution is asymmetric and a second peak originates due to a pixel which fires within a SiPM as shown in figure 4.2(b) resulting in a bad pedestal value.

In this approach, the pedestal values are extracted during the three testbeam campaigns. For the AHCAL prototype with 38 layers, this leads to $38 \times 576 \times 16 = 350208$ pedestal values. The result with pedestals extracted from all the AHCAL channel with memory cell-wise is shown in figure 4.2(c) and the RMS value obtained for all the AHCAL channel with memory cell-wise is shown in figure 4.2(d).

The pedestal extracted in external trigger mode during the three testbeam periods for all AHCAL channels, including memory cell-wise results in a mean value of about 497 ADC bins and a width of about 32 ADC bins. In addition, in the case of June and October testbeam periods, it is observed that there exists a long tail in the RMS distribution for values between 5 and 7 ADC bins. This was the case for the pedestal value obtained with no power-pulsing runs. It was found to be widely distributed among the channels of AHCAL and no significant consequences were found.

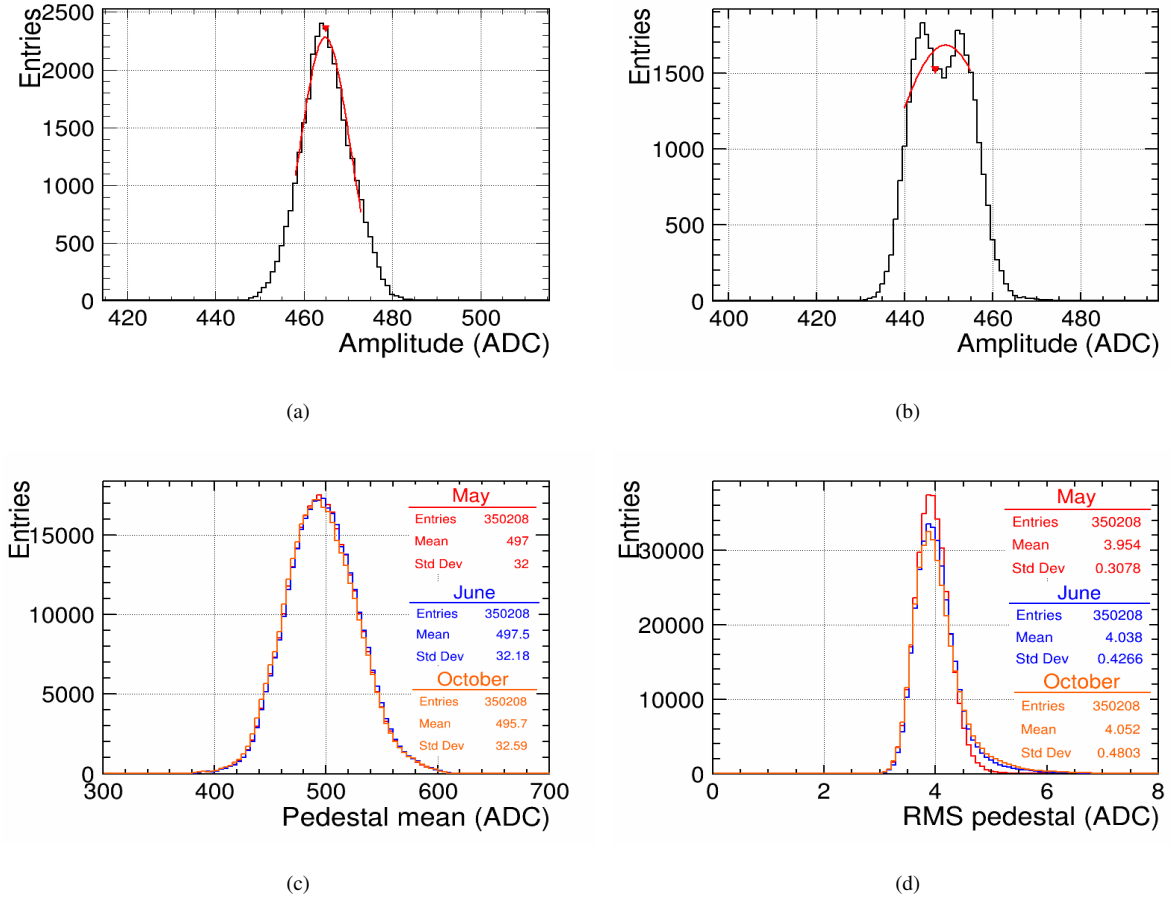
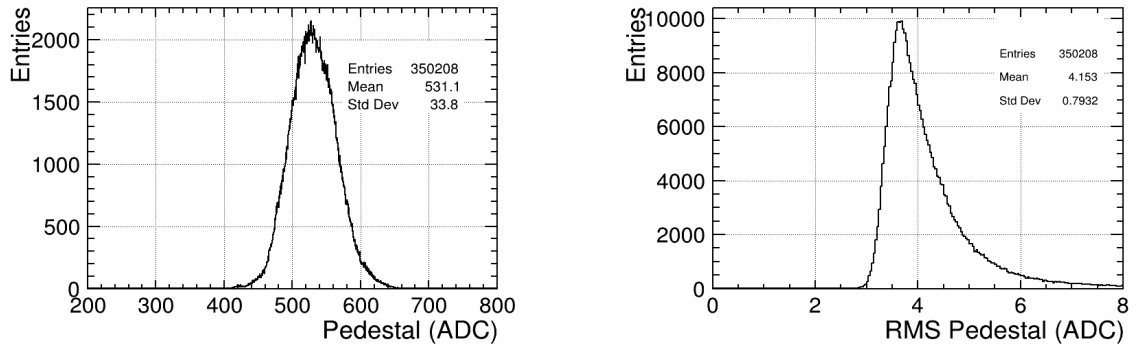


FIGURE 4.2: (a) ADC distribution of a single AHCAL channel. (b) poor ADC distribution of channel due to a second peak originating in the distribution. (c) pedestal values for all channel-memory cell-wise. (d) RMS extracted from the ADC distributions for all channel-memory cell-wise.

For energy calibration, the pedestal value is taken in auto-trigger mode. In auto-trigger mode, the information from all channels of one chip will be read out if at least one channel of the

chip has a signal that passes the trigger threshold. Only information from channels without a hit bit is used to determine the pedestal. This measurement is done using a muon run with the beam centered on only one channel of a chip, with the remaining channels used for pedestal determination.

The pedestal distribution in auto-trigger mode from a power-pulsing muon run during the June testbeam is shown in figure 4.3(a), and it results in the mean value of about 531 ADC and RMS of ≈33 ADC. The tail on the right of the pedestal distribution is due to pedestal jumps in memory cells that are higher than 8. To suppress this behaviour, an average pedestal of memory cells from 0 to 8 was used for all the memory cells higher than 8. The intrinsic widths of the pedestals are shown in figure 4.3(b), and the mean intrinsic width is found to be ≈4 ADC, which is in agreement with values obtained from the external triggered events.



(a) (b)

FIGURE 4.3: (a) Pedestal distribution of all memory cells extracted from muons runs in units of ADC. (b) The width of the pedestal spectrum.

4.2 Gain Calibration and LED Studies

The gain calibration is essential for two main purposes: Firstly, it provides the conversion scale from the ADC bins to a number of SiPM pixels fired in a single cell. Secondly, it allows for monitoring the detector performance and stability during the testbeam. The gain calibration is performed using dedicated LED runs by injecting low amplitude light pulses into each scintillator tile.

An example of a typical single-photon spectrum is shown in figure 4.4. Each peak corresponds to a number of fired pixels in one SiPM. The first peak corresponds to the pedestal (zero photons detected). The peak next to it is from one pixel firing at a time. The third pixel from the left is from two pixels firing simultaneously. Altogether, seven individual peaks are found to be well separated; this indicates good photon counting capabilities of the SiPM. For the pedestal peak, the width accounts for the measure of electronic noise and is independent of the SiPM. The height of each peak and the number of photons hitting the SiPM is Poisson distributed.

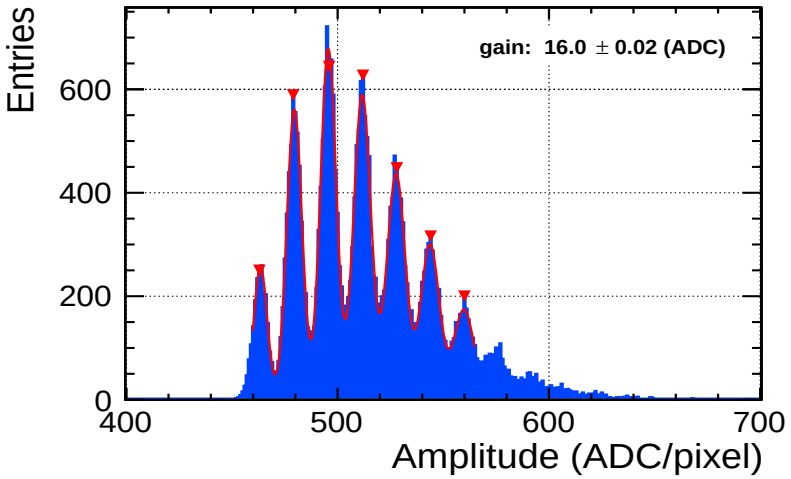


FIGURE 4.4: An example of a single photon spectrum of a single AHCAL cell obtained using LED runs in external trigger mode. The first peak (from left) corresponds to the pedestal. The following peaks correspond to the number of photo-electrons in one SiPM. The superimposed curve is a fit with multiple Gaussian functions. The distance between two consecutive peaks is defined as the SiPM gain.

The standard method of analysing the SiPM spectra uses a multi-Gaussian fit approach:

$$M(x) = \sum_{i=1}^{N_\mathrm{p}} m_i(x) = \sum_{i=1}^{N_\mathrm{p}} a_i \cdot \exp\left\{-\frac{1}{2}\left(\frac{x-b_i}{c_i}\right)^2\right\}, \tag{4.2}$$

where N_p is the total number of peaks fitted in the single photo-electron spectrum. The advantage of this method is, in principle, straightforward as the gain is the distance between two consecutive photo-electron peaks and the peak to peak distance is left as a free parameter. In addition, the amplitude and width of the individual peaks are extracted from the fit.

The limitation of this method is that it ignores the description of background events caused by dark count and after-pulse events in between the peaks. In addition, this approach does not consider variation between the SiPM pixels. Therefore, several alternative models have been proposed [96] that could potentially describe the entire pulse height spectrum. These models involve an excellent description of peaks, including the regions in-between the peaks accounting for the key performance parameters of the SiPM such as gain, the cross-talk probability, dark count rate and correlated noise.

The photo-electron peaks can be broadened by thermal noise, cross-talk, after-pulses or by the inhomogeneity between the response of individual pixels as the charge created in each pixel varies due to the multiplication process. Such effect in broadening of pixels increase with an increasing number of the photo-electrons. The behaviour for a specific AHCAL channel is shown in figure 4.5(a) for an increasing number of pixels.

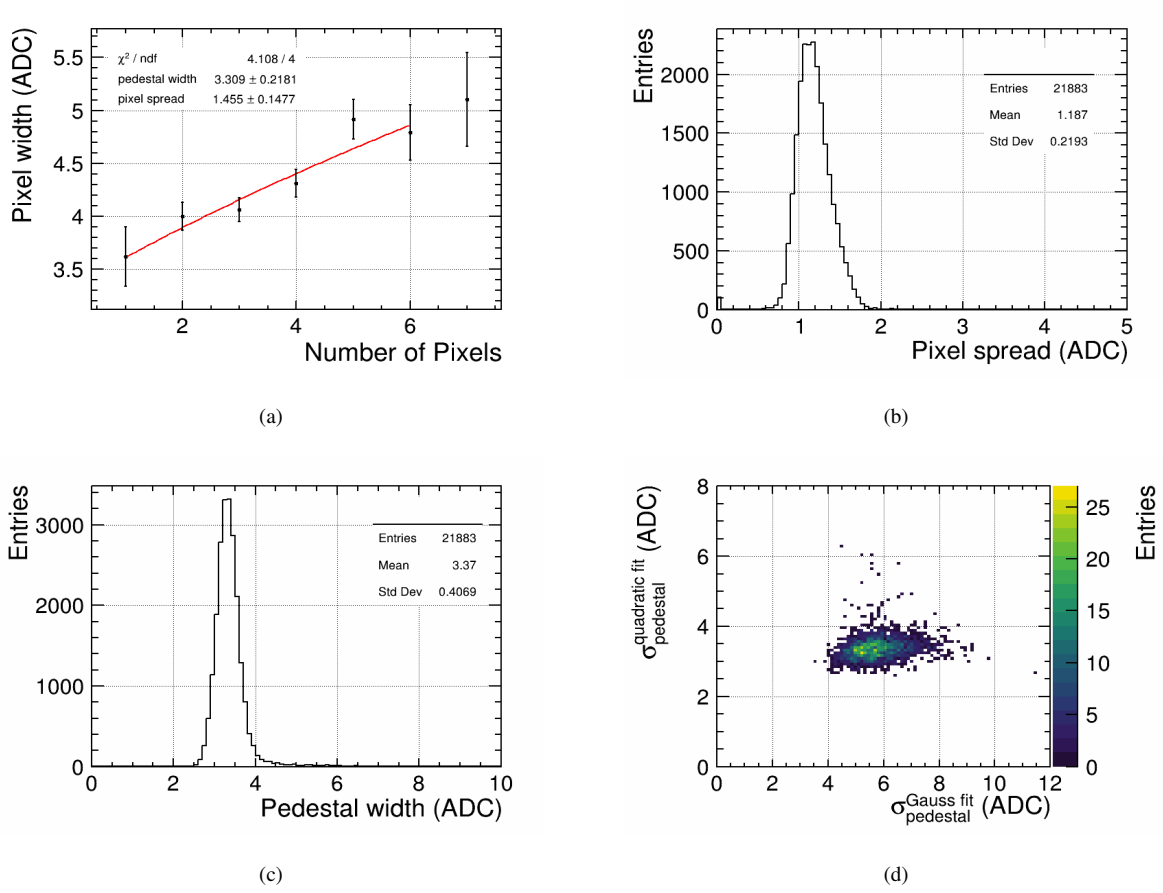


FIGURE 4.5: (a) SiPM pixel width in ADC plotted against the number of pixels for a single SiPM. The width of each peak σ_0 to σ_6 increases from ≈3.6 ADC to ≈5.1 ADC due to inter and intra pixel gain variation. (b) Pixel spread extracted from the quadratic fit. (c) Pedestal width extracted from the quadratic fit. (d) Correlation for the channels between pedestal width extracted from quadratic fit and pedestal width extracted from a pedestal run.

The origin of inter-pixel variations lies in differences in the pixel capacitance among pixels. The intra-pixel variations can be described by changes in the electric field within a pixel. Such a behavior for an increasing width σ_N of N pixels can be explained using a quadratic expression given by:

$$\sigma_N^2 = N \cdot \sigma_G^2 + \sigma_0^2 \quad \text{where } \sigma_G^2 = \sigma_1^2 - \sigma_0^2, \quad \text{and } N = 1, 2, 3 \ldots, \tag{4.3}$$

where σ_G denotes the pixel spread and increases as a square root of the number of prompt Geiger discharges N and σ_0 is the pedestal width. Two distributions are obtained as a result of the fitting performed on all the AHCAL channels. The distribution with pixel spread is shown in figure 4.5(b) and the distribution with pedestal width, is shown in figure 4.5(c).

Among the AHCAL channels, the pixel-to-pixel spread was seen in 21,883 channels, and the remaining five channels showed no pixel-to-pixel spread or pedestal width. The obtained values for the pedestal widths of the fitted channels from LED measurement resulted in a mean value of ≈3.4 ADC and the pixel-to-pixel variation of about 1.2 ADC. Furthermore, a correlation between

the pedestal width obtained from the Eq. 4.3 and the pedestal width obtained from the pedestal extraction procedure as described in section 4.1 was compared. Unfortunately, it was found that there exists no strong correlation between these two parameters as observed in figure 4.5(d), which hints that the two values for the pedestal width are not very consistent.

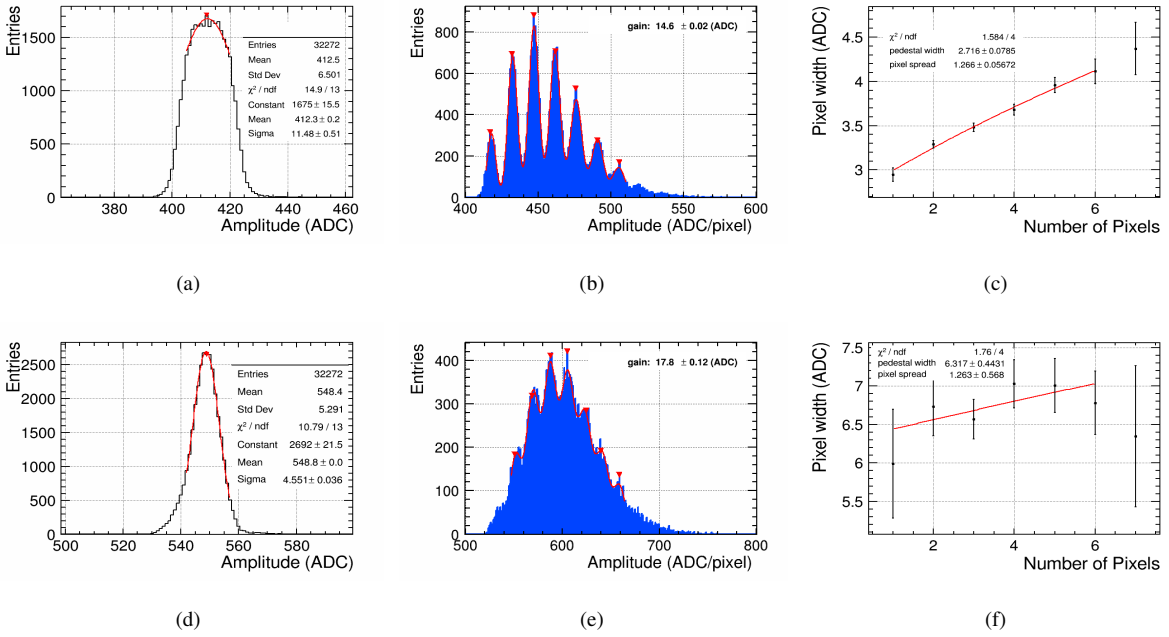
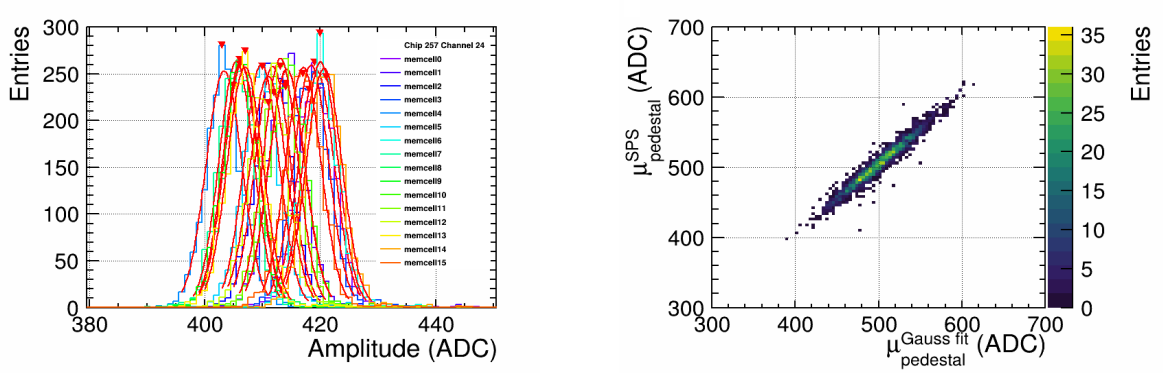


FIGURE 4.6: (a) Pedestal distribution with a broad peak contradicts the result in fitting a SPS spectrum (b) and resulting in a quadratic fit which yields a satisfactory result of pedestal width (d) Pedestal distribution with a narrow peak (e) causing a bad fit to SPS spectrum leading to unsatisfactory result in extraction of pedestal width from quadratic fit (f).

The outliers in the correlation plot from the Gaussian fit method are mainly due to the widening of the pedestal spectrum. An example of such a pedestal spectrum is shown in figure 4.6(a). The effect is mainly governed due to the influence of individual memory cells with shifted mean position as shown in figure 4.7(a) resulting in broadening of the pedestal distribution for the corresponding channel. However, the SPS spectrum for the corresponding channel gives good separation between the peaks as shown in figure 4.6(b) and the pedestal width extracted from the quadratic fit is shown in figure 4.6(c). For the case where the pedestal width extracted from quadratic fit is significant, as seen in figure 4.6(f) is clearly the reason for bad fits to the SPS spectrum as shown in figure 4.6(e). In such cases where the value of the pedestal width is too large, or the SPS peaks are not well separated, it does not reflect the true gain value.

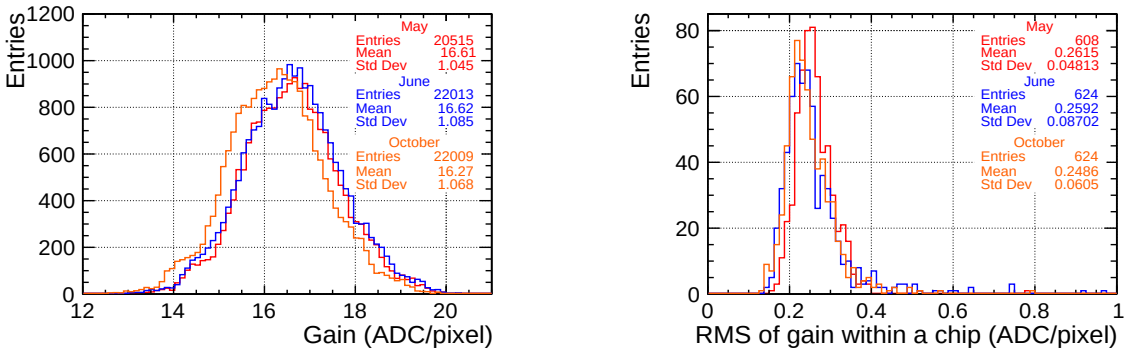
Figure 4.7(b) validates that the pedestal values obtained from the first peak of the SPS spectrum and the mean values obtained from the pedestal extraction procedure with a Gaussian fit to the pedestal distribution are correlated. Furthermore, apart from a few outliers, we see a good correlation between pedestal values for each channel obtained from a dedicated pedestal run and the position of the first peak from the SPS spectra. As a result, there is no large bias in determining the pedestal position obtained from either of the two procedures.



(a) (b)

FIGURE 4.7: (a) Memory cell-wise ADC distribution for a channel showing an outlier from a pedestal with a Gaussian fit larger than 10 ADC. (b) Correlation plot for the mean value of the pedestal extracted from the Gauss fit against the position of the first peak(pedestal) from the fit to SPS peak.

During testbeam operation, LED runs were taken approximately once per day to check the gain stability. The detector was running with the external trigger and with or without power-pulsing mode for the gain calibration. Several long and short LED runs were measured. The long runs took about 8 to 10 hours with LED voltages between 5000 mV to 7000 mV in steps of 20 mV. On the other hand, short runs took roughly less than an hour with LED voltages between 5000 mV to 6000 mV in steps of 500 mV. The gain was measured using long LED runs for 21,888 channels with the extraction procedure described earlier. A total of about 95% of the channels were calibrated. Gain values for the remaining 5% of channels were taken based on the average of the channels of the corresponding chip. Combining several gain runs, more than 99.9% of all channels were calibrated individually. In contrast, a small fraction of channels which failed to fit or showed no clear SPS peaks, for these channels the average of the corresponding chip was considered. The gain calibration was also performed for the May and October 2018 testbeam periods. The distribution of mean gains obtained from all the fitted AHCAL channels for the three testbeam campaigns is shown in figure 4.8(a).



(a) (b)

FIGURE 4.8: (a) Gain extracted for all AHCAL channels during the three testbeam periods, the mean value is ≈16 pixels/ADC. (b) RMS of the gains within a single ASICs for three testbeam periods.

Figure 4.8(b) shows the gain distributions for the three testbeam periods which are centered at ≈16.6 ADC/pixel with a channel-to-channel spread of roughly 1.1 ADC/pixel, which is about 6%. Nevertheless, it is observed that for channels on the same ASIC, the spread is only 1.5%. This indicates that the gain variation between the ASICs contributes significantly to the overall spread. An example of such behaviour of gain variation is shown in figure 4.9 for a single AHCAL layer in its I and J coordinates, clearly showing the pattern from each ASIC.

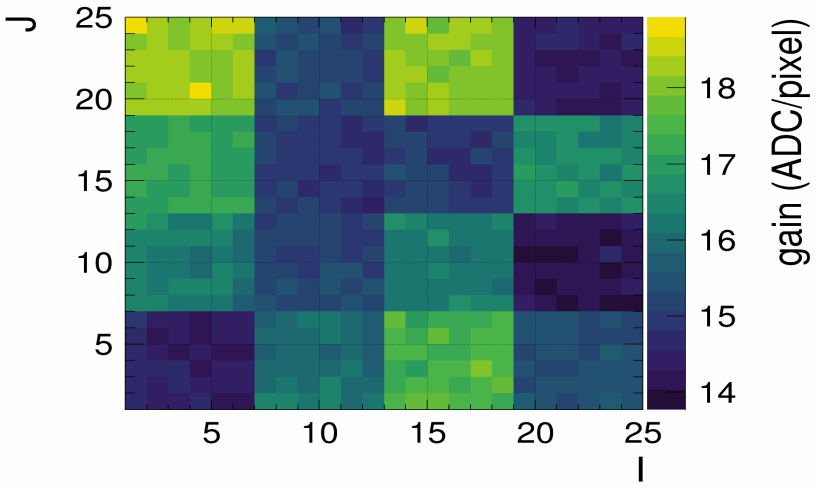
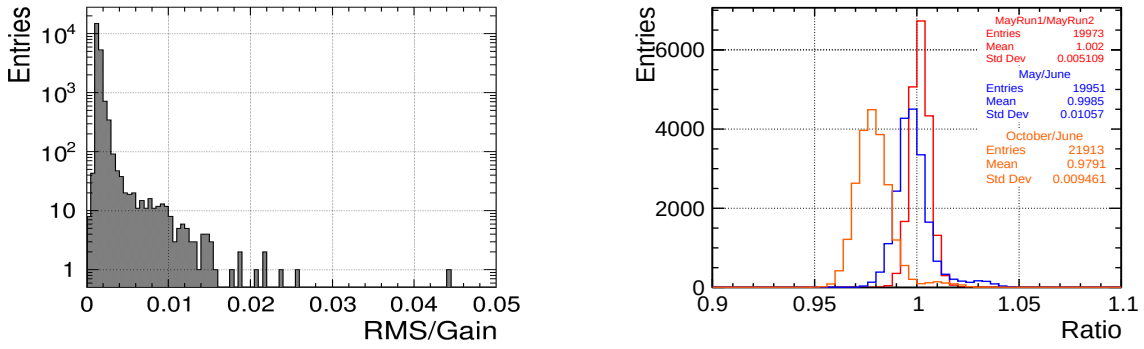


FIGURE 4.9: Gain map in I and J coordinates for layer 12 of AHCAL.

Moreover, the gain is consistent between the May and June 2018 run (both without power pulsing). However, for the October 2018 run (with power pulsing), a slight shift in gain is observed, caused by power-pulsing mode as the detector is cooler and the gain differs by about 2%. A statistical uncertainty on the gain was also determined, mainly contributed from the fitting

procedure. Figure 4.10(a) shows the statistical uncertainty, which is less than 2% for fits that pass the quality criteria.

Gain calibration is also a powerful tool to monitor the performance of the AHCAL prototype during the testbeam. Figure 4.10(b) shows the gain ratios between different testbeam periods and within a testbeam period. The gain ratio indicates that there is good stability in terms of gain between two runs in May and between the testbeam periods due to reliable temperature compensation. In contrast, the shift in gain ratio from October to June is caused by power-pulsing mode as described earlier.



(a) (b)

FIGURE 4.10: (a) Statistical uncertainty in the gain calibration for the channels of AHCAL, obtained from the fit values. (b) Gain of all AHCAL channels between two May runs and between testbeam periods of May and June and between October and June.

Furthermore, the gains were very much stable each day. The gain stability for the May testbeam period extracted from short LED runs is shown in figure 4.11.

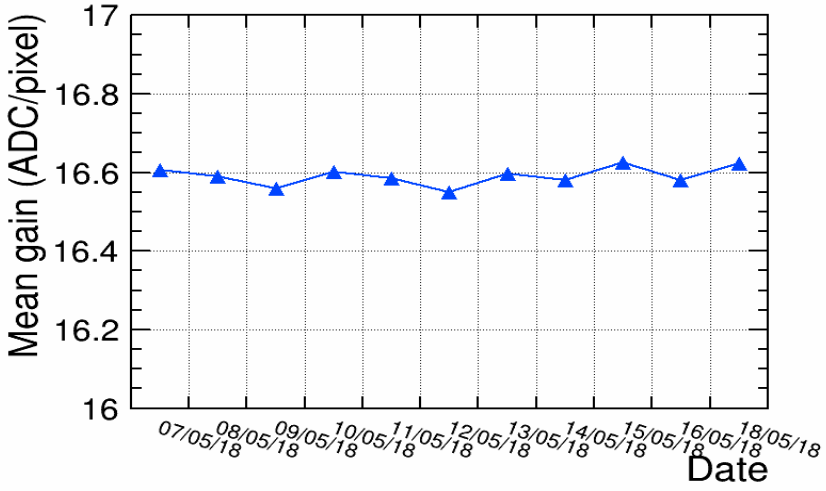


FIGURE 4.11: Mean gains obtained for the May 2018 testbeam period

As discussed in section 2.1, SiPM gain significantly depends on temperature because the breakdown voltage depends on temperature, which needs to be monitored at high temperatures. The breakdown voltage of the SiPM increases, causing a lower over-voltage and thus decreasing the gain of the SiPM. This means that temperatures have to be controlled during measurements. Changes in breakdown voltage with temperature changes determine the temperature coefficients. The temperature coefficients are found to be the same everywhere in the detector. Basically, for 1 Kelvin temperature, which is roughly equal to one **D**igital to **A**nalog **C**onverter (DAC) tick change, the breakdown voltage moves by 54 mV, so a change is made in the bias voltage to 54mV. Moreover, the DAC is used to adjust this bias voltage. Six temperature sensors are installed at various points in the active layers that measure the temperature regularly. The average temperature is measured per layer as some layers have broken sensors. This is then compared with previous temperatures to decide whether an adjustment to the DAC value is needed. This method keeps the gain stable during the test beam periods. Figure 4.12 shows the mean gain as a function of temperature, and the temperature correction works well. During the testbeam operations, the temperatures were monitored at the start of every run and in time intervals between ten minutes and one hour.

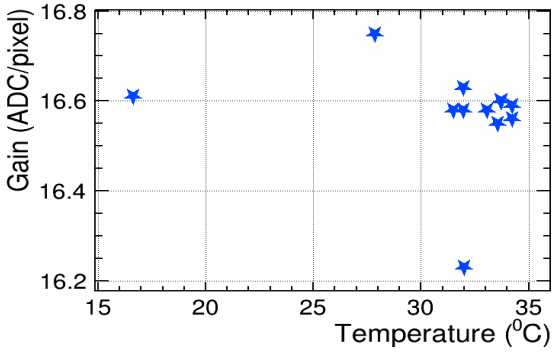


FIGURE 4.12: Mean gain as a function temperature using short LED data measured in May 2018 testbeam.

4.3 MIP Calibration

MIPs are used to perform a cell to cell equalization. As a first approximation, muons are considered to be MIP, as their mean energy loss through the matter is close to the minimum.

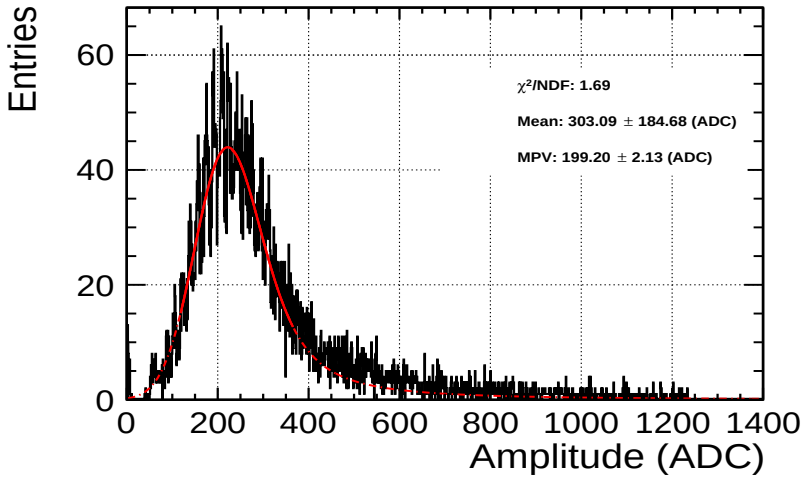
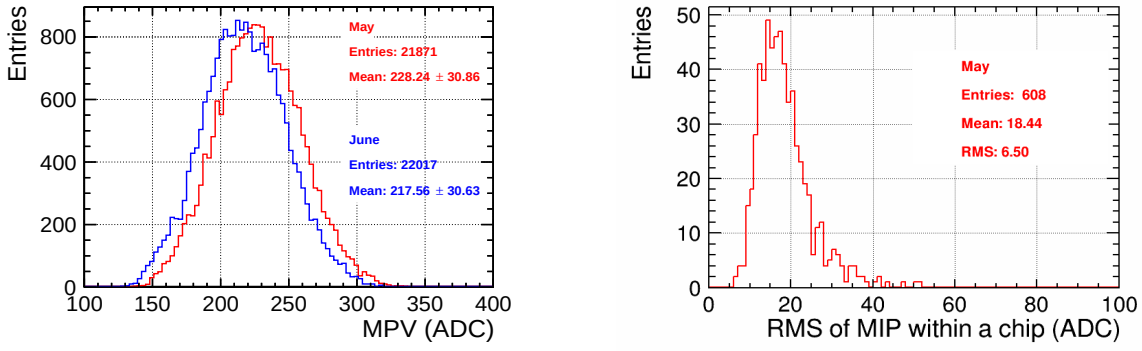


FIGURE 4.13: Energy deposited in an AHCAL cell using 40 GeV muons from May 2018 testbeam data. The distribution is fitted using a Landau convoluted to a Gaussian function shown in a solid red line. The dotted red line is the function extrapolated outside the fitting range.

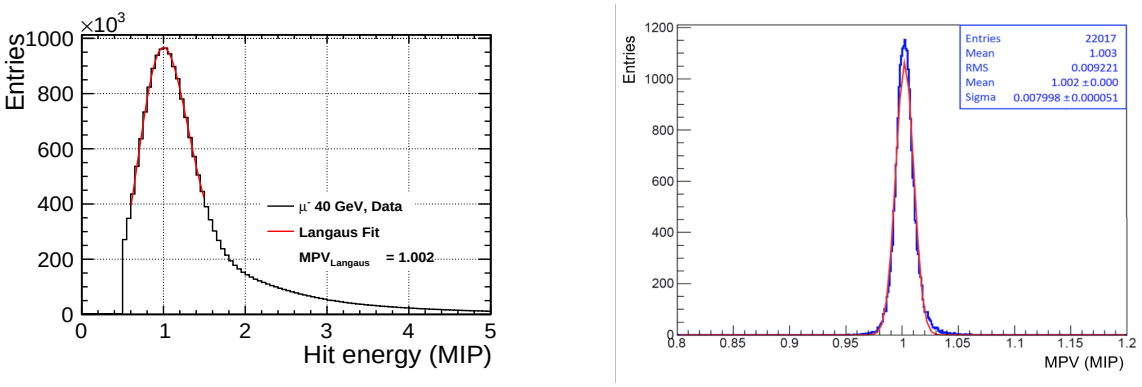
The MIP extraction for the AHCAL technological prototype was performed by D. Heuchel [97], and only the results are highlighted in this section. The MIP calibration was done at a channel level by analyzing the hit amplitude spectrum, using a broad muon beam with known energy. The muon beam is scanned over 16 positions in the detector front face with high statistics of about 10000 events per channel. After a memory cell-wise pedestal subtraction as described in section 4.1 the resulting amplitude distribution in ADC counts is obtained. An example of a typical response of a single AHCAL channel using 40 GeV muons is shown in figure 4.13 which is fitted using a Landau convoluted with a Gaussian function. A distribution when a muon passes through a scintillator tile follows a Landau shape as previously discussed in section 1.3.1.



(a) (b)

FIGURE 4.14: (a) MIP constants extracted for the AHCAL channels using muons runs from May and June 2018 testbeam data. A mean value of roughly 220 ADC is obtained. (b) The spread within a chip is ≈18.4 ADC using muon testbeam data from May 2018 [97]

The Landau part of the fit function considers the energy loss of muons within the scintillator tiles. The Gaussian part considers different deviations from the theoretical point corresponding to photon counting statistics and electronic noise contributions. The MPV extracted from the convoluted function is defined as 1 MIP as the MPV of the convoluted function is more stable than the MPV of the Landau distribution. Moreover, the MPV is extracted from all the fitted AHCAL channels resulting in a distribution as shown in figure 4.14(a). The mean value of the MIP constants is ≈220 ADC with an RMS of ≈30 ADC, which corresponds to ≈14%. Furthermore, the MIP calibration efficiency, i.e. the number of successful fits of muon hit energy spectra is ≈99.92% with less than 0.1% of dead channels or the channels where the fit failed. These channels are neglected during physics analysis. In addition, for the channels on the same ASIC, it is observed that the spread is ≈8.1% as shown in figure 4.14(b), which is good enough for auto-trigger operation with the same threshold for all channels.



(a) (b)

FIGURE 4.15: (a) Hit energy distribution of muon data after the MIP calibration, fitted with Landau convoluted with a Gaussian function, and the MPV is ≈1 MIP. (a) MPV's extracted from the MIP peaks from the June testbeam period. A Gaussian fit to the MPV distribution for all the AHCAL channels gives a mean from the Gaussian fit ≈1 MIP.

The hit energy is converted to MIP units using the Eq. 4.1 by applying all the extracted calibration constants. Figure 4.15(a) shows the hit energy spectra where the MPV is extracted using Landau convoluted with Gaussian function. As a result, after calibration, the hit energy distribution peaks at a value close to 1 MIP. Moreover, the consistency of the MIP calibration procedure can be checked by looking at the MPV of all the channels after applying the calibration constants to the data reconstruction. In figure 4.15(b), as expected, the distribution of the MPV of all the channels is narrow peak at 1 MIP and has a spread of ≈0.008 MIP.

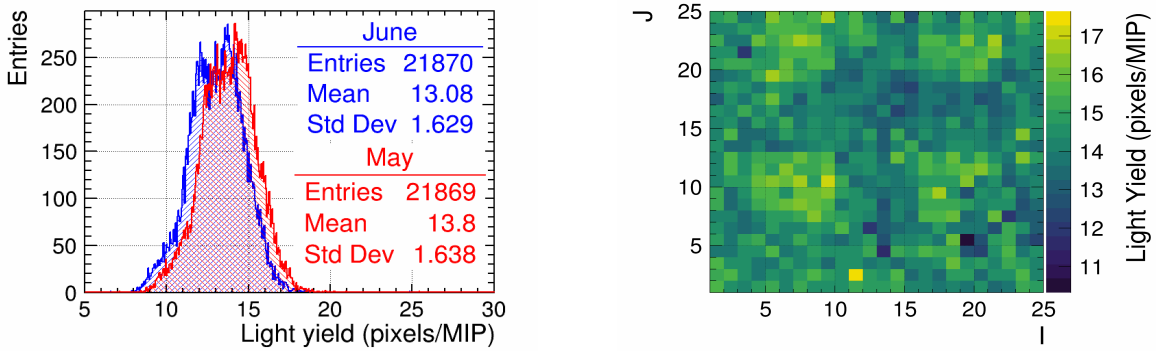
Light Yield

The light yield L_i of a SiPM describes the number of photo-electrons generated by an incident MIP depositing its energy in a tile, i.e. the ratio between MIP response a_i and the gain G_i of a

tile.

$$L_i \left[\frac{\text{pixels}}{\text{MIP}}\right] = \frac{a_i \left[\frac{\text{ADC}}{\text{MIP}}\right]}{G_i \left[\frac{\text{ADC}}{\text{pixels}}\right]} \tag{4.4}$$

Light yield gives a measure of detector performance. The performance of the detector depends on two aspects: the signal to noise ratio and the dynamic range; these are important figures of merit for detectors. The signal to noise ratio is defined as the ratio between the amplitude of a measured signal and the width of the corresponding pedestal. In the case of AHCAL, the signal to noise ratio is the ratio between the most probable muon signal above the pedestal in a single tile and its pedestal width, which corresponds to ≈ 53. The dynamic range is the difference between the largest and the smallest energy deposition detected in a tile. A higher light yield is related to a larger signal to noise ratio but a smaller dynamic range. With the best compromise between a high signal to noise ratio and a broad dynamic range, each AHCAL SiPM-on tile is expected to yield a SiPM response of 15 pixels with an intensity corresponding to the light generated by 1 MIP in a scintillator tile.



(a) (b)

FIGURE 4.16: (a) Light yield measured for AHCAL channels for the two testbeam periods. The mean light yield is ≈13.08 pixels/MIP for the June testbeam period and ≈13.8 pixels/MIP for the May testbeam period and a spread of ≈1.6 pixels/MIP. (b) Light yield distribution of the AHCAL layer number 12 in I and J coordinates.

The estimated value for light yield is about 15 pixels per MIP, whereas the achieved light yield for all AHCAL channels gives a mean value of ≈13.8 pixels/MIP, which is slightly smaller than the desired value and with a relative spread of ≈11.8% as shown in figure 4.16(a). Moreover, the large spreads in the light yield distribution are due to inhomogeneities in the tile. The light yield for the May testbeam period is higher due to larger MIP values obtained using a power-pulsing run. An example of a light yield distribution within a single AHCAL layer in I and J coordinates is found to be homogeneous, as shown in figure 4.16(b).

4.4 Inter-Gain Calibration

As previously stated in section 2.2, the SPIROC2E ASIC operates in two different modes - LG or HG mode. High amplitude signals use the LG mode with less amplification which has a fixed ratio of ∼1:10 compared to HG mode. In contrast, low amplitude signals use the HG mode with large amplification. Inter-gain calibration is required for a smooth transition between HG and LG hit energies. In order to perform the HG and LG inter-calibration, the LED data is taken in both amplification modes [98]. As a result, the HG versus LG variation is linear before reaching the HG saturation level.

The extraction of the inter-gain calibration coefficients from LED runs is possible because the range of linear response of the chip in both modes overlaps. To perform inter-gain calibration, the signal's amplitude is varied within the linear range of HG and LG by changing the LED light intensity. The inter-gain calibration factor of a channel is considered by taking into account the pedestal values and is the ratio of HG to LG. This constant depends on the ratio of capacitors because of the circuit design with 600 fF in the high gain mode and 1200 fF in the low gain mode, which results in an IC factor of 20. Slight variations of the capacitances in the chips can generate variations of this ratio. Therefore it is necessary to be calibrated for each channel.

The following description of HG and LG inter-calibration extraction was based on the work of Y. Sudo [98], and only the results are reported here. Two methods were used for the determination of the inter-calibration factor for the AHCAL prototype:

dLY method− In this method, for each calibration voltage and each AHCAL channel, two histogram with the amplitude values in ADC counts both for the HG, and the LG modes are filled separately. In addition, the pedestal value is subtracted memory cell-wise before filling the histograms. Then, the mean of the histograms is considered for each AHCAL channel and each calibration voltage. The mean amplitude value in HG and LG modes is extracted for each channel. The IC factor is then the ratio of the HG signal's difference and the LG signal's difference.
Slope method− In this method, the inter-calibration coefficient for each channel is extracted from the ratio of the slope of the linear fit.

The limitations of the above two methods are that there needs to be a reasonably large signal, and the total amount of charge in a chip should be less than 60000 ADC. Therefore, the calculation is done in the HG linear region where the LG signal range is between 30 and 100 ADC as shown in figure 4.17 and at least 5 points are available in the fitting range.

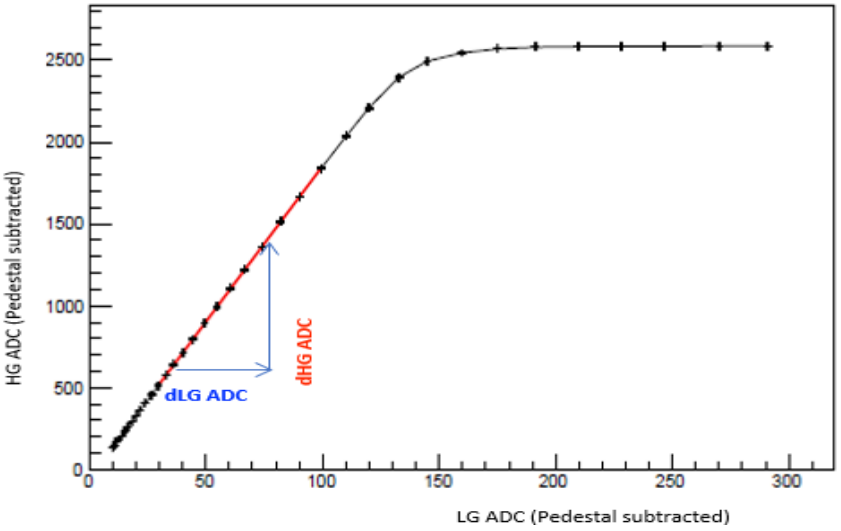
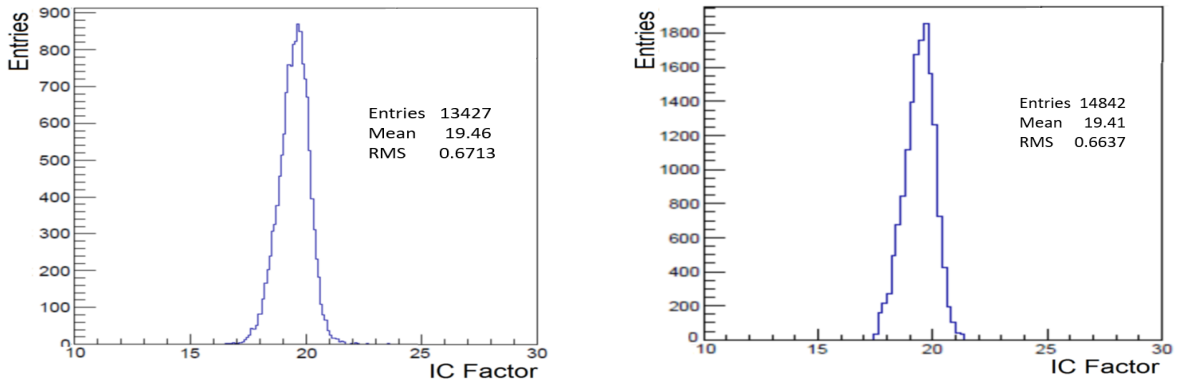


FIGURE 4.17: The response of a AHCAL channel in two amplifier gain modes (Low Gain and High Gain) after the pedestal subtraction for increasing LED light intensity. Each point corresponds to one LED voltage. The IC factor is extracted either by the slope of the linear part or the ratio of dHG/dLG, where dHG (dLG) is the HG (LG) signal difference for two LED voltages [98].

The IC factor extracted from June 2018 testbeam campaign using the dLY method is shown in figure 4.18(a), and the slope method is shown in figure 4.18(b). The IC factor values range between ≈17 and ≈21, with a mean of ≈19.4 and a RMS of ≈0.66, which is a spread of 3.5 to 4%. The extracted IC value is slightly lower than the expected value of 20. No significant differences were observed between the two extraction methods. The IC factor was extracted using the slope method for ≈80% of the AHCAL channels. For the remaining channels, the slope could not be extracted; therefore, the values were taken based on the average from the channels of the corresponding chip.



(a) (b)

FIGURE 4.18: IC factor extracted from the May 2018 testbeam period (a) by the dLY method. (b) by the slope of the linear fit [98].

4.5 Time Calibration

The time calibration is performed using muon runs. Muons are excellent candidates for time calibration as high energy muons deposit energy instantaneously. The SPIROC2E ASIC gives information about the time of the hit in the detector. This information is given in TDC counts; therefore, it needs to be calibrated in physics units. The time calibration has been the core of another study done by L. Emberger. A more detailed discussion on timing measurement with the emphasis on the time calibration is given elsewhere [99, 100]. However, for completeness, it is summarized here. For the AHCAL technological prototype, the time calibration undergoes several steps.

In order to have a standard time reference for each event, the reference time given by the external trigger needs to be calibrated. The ASIC provides information concerning the timing of the hit for every channel in TDC values. Then, with the hit time information stored in TDC values, it is converted to nanoseconds so that it can be interpreted in the unit of time using:

$$T_i^C\,[ns] = t_i^S \left[\frac{ns}{TDC}\right] \times t_i^H\,[TDC] + t_X\,[ns] - T0\,[ns]\,, \tag{4.5}$$

Parameter	Meaning
T_i^C	calibrated hit time in units of ns
t_i^S	slope of linear TDC in units of ns/TDC
T_i^H	measured hit time in units of TDC
t_X	offset of the TDC voltage ramp in units of ns
$T0$	reference time in units of ns

TABLE 4.2: Parameters used for the time calibration

To record time, TDC voltage is applied in a see-saw mechanism in intervals of 4000 ns with a 250 kHz bunch clock. For this purpose, the SPIROC2E chip has two TDC ramps for even and odd bunch crossings forming a combined up and down ramping to avoid dead time. As each channel stores information at a memory cell-level, 16 calibration values and pedestals are essential for each channel.

The slope is common to all channels on a chip and the offset is extracted for every memory cell. This distinguishes between the odd and even bunch crossing parity due to the dual ramps structure of the SPIROC TDC. This is done by plotting the reference time in nanoseconds versus the non calibrated hit time in TDC counts and fitting a linear function to the data points as shown in figure 4.19; this is done per each chip and bxID. Additionally, events with reference time between 500 to 3000 ns are selected to avoid edge effects. Moreover, to avoid time distortions, events with one hit per layer are considered due to rising chip occupancy.

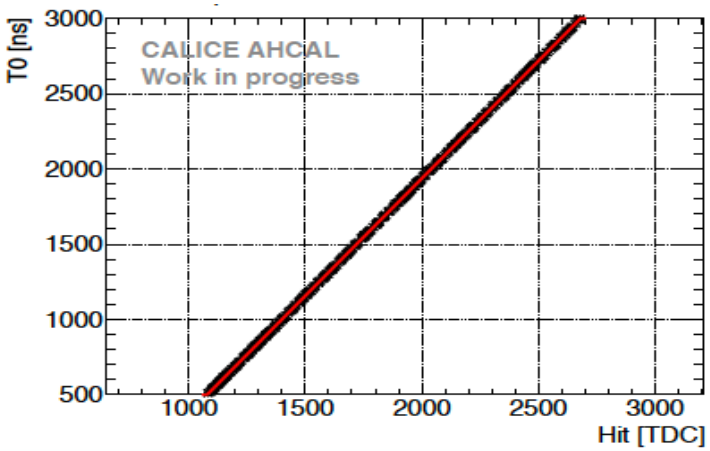


FIGURE 4.19: Linear fit performed between TDC values and the reference time $T0$ for the events of one AHCAL channel [99].

Once a hit occurs, the bunch crossing **ID** (bxID) and the analog values are recorded in one of 16 memory cells. At the end of the read-out cycle, the signal is digitized by the 12 bit TDC. From the slope of the TDC ramp and its pedestal, the time in nanoseconds is calculated for all channels memory cell-wise. After the time has been converted from TDC counts to nanoseconds, certain corrections have to be applied to the data to improve time resolution due to known features of the SPIROC2E. These features are listed below:

- **Non-linearity of the TDC ramp**: The extraction of the slope is performed with the assumption that the voltage ramp of the TDC is linear. Since this condition is not satisfied rigorously, a correction has to be applied.
- **Time-walk correction**: This effect is related to the presence of a fixed trigger threshold. Thus, events with a smaller amplitude will trigger systematically later than events with a larger amplitude. In addition, the time during which the signal is read out is fixed, leading to a more significant effect. A correction for this effect is obtained by looking at the correlation between the hit time and the hit energy deposited in the detector.

After these corrections are applied, the time difference between the time reference and the data should peak at 0 ns, as shown in figure 4.20 for muon with 40 GeV beam energy. Finally, a Gaussian fit is applied to the data resulting in a width of 2.57 ns. With this procedure, about 93% of the AHCAL channels are calibrated.

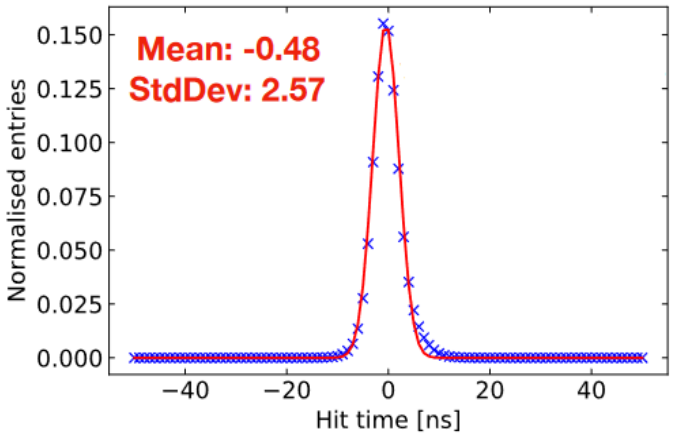


FIGURE 4.20: Distribution of hit time for 40 GeV muon data after time calibration. The red solid line shows the Gaussian fit applied to the data points [99].

This chapter discussed the calibration procedure comprising ≈21888 channels in the AHCAL technological prototype. By applying the calibration procedure, it is possible to convert the energy deposited from the units of ADC to a MIP. In addition, the information on the hit time was converted from units of the TDC to nanoseconds after the start of an event.

Calibrating the data to these units is necessary for the data analysis. Furthermore, it facilitates understanding the physics as well as suppressesing electronic differences between the various single channels and HBUs, thus equalizing their response. Table 4.3 provides a brief summary of those important results presented in this chapter. The next chapter discusses the data samples used in this thesis and the event selections applied.

Calibration type	Parameter name	Parameter value
Pedestal	Mean (ET mode)	∼497.0 [ADC]
	Intrinsic Spread (ET mode)	∼4.0 [ADC]
	Mean (AT mode)	∼531 [ADC]
	Intrinsic Spread (AT mode)	∼4.2 [ADC]
Gain	Mean	∼16.6 [ADC/pixels]
	Global spread	∼1.0 [ADC/pixels]
	ASIC spread	∼0.2 [ADC/pixels]
MIP	Mean	∼228 [ADC/MIP]
	Global spread	∼30.0 [ADC/MIP]
	ASIC spread	∼18.4 [ADC/MIP]
Light yield	Mean	∼13.8 [pixels/MIP]
	Global spread	∼1.63 [pixels/MIP]
Inter-gain factor	Mean	∼19.4
	Global spread	∼0.7
Hit time	Mean	∼0.5 [ns]
	Global spread	∼2.6 [ns]

TABLE 4.3: Summary of key parameters obtained from the AHCAL calibration studies.

Chapter 5

Sample Preparation and Event Selection

"The goal is to turn data into information, and information into insight."

Carly Fiorina

AN EVENT SELECTION is mandatory to filter the contamination of particles like muons or electrons from pions. We have seen in section 3.2 that the particle beam at CERN is not entirely pure even though it runs in a mode that should contain only one particle type. Nevertheless, it consists of a mixture of muons, electrons and hadrons. Therefore, to enhance the quality of the data sample, the most fundamental quality selection is to look into muons, electrons and pions separately.

In this chapter, the first section describes the beam profile calibration. It is followed by a short introduction to the shower start algorithm in section 5.2. Finally, the event selections that are used in the analysis are described in section 5.3. For simplicity, the June setup's additional detector parts (PS and TC) are not considered in this analysis.

5.1 Beam Profile Calibration

A critical parameter for simulation is the simulation of a particle gun. Therefore, it is essential to model the particle beam gun position and emittance in simulation to obtain meaningful predictions. Although it is relatively easy to set the particle energy and type, the beam profile needs to be calibrated by comparing the mean position and the width of the shower. Furthermore, the x-and y-positions of the beam gun must be well modelled. This is performed to ensure that the same detector cells are hit as in the data. More importantly, to avoid bias in comparing testbeam data and simulations. Beam modelling also accounts for the possible transverse leakage and systematic effects of the scintillator geometries in testbeam data and simulations. In the z-direction, the `GEANT4` particle gun is placed before the Čerenkov detector, which is -50 meters away from the AHCAL detector and the momentum uncertainty in the simulation is set to zero.

The most trivial way to estimate the beam emittance and position is by using the mean and the RMS of the event-wise CoG distributions in the x-and y-direction measured in the AHCAL during data collection. The CoG for the x-and y-coordinate is calculated based on Eq. 3.1. The beam calibration is done for every simulated beam energy and particle type for different impact positions. The impact position varies because the AHCAL stage allows moving the detector to different positions. The exact particle gun position has to be estimated iteratively by adjusting the position and the spread until a reasonable agreement in the CoG distributions is achieved. However, these input values do not fully reflect the actual impact position since values could be biased. The particle gun settings used to reproduce the testbeam data runs for each detector position, beam momentum, and particle type in this thesis are given in table B.1.

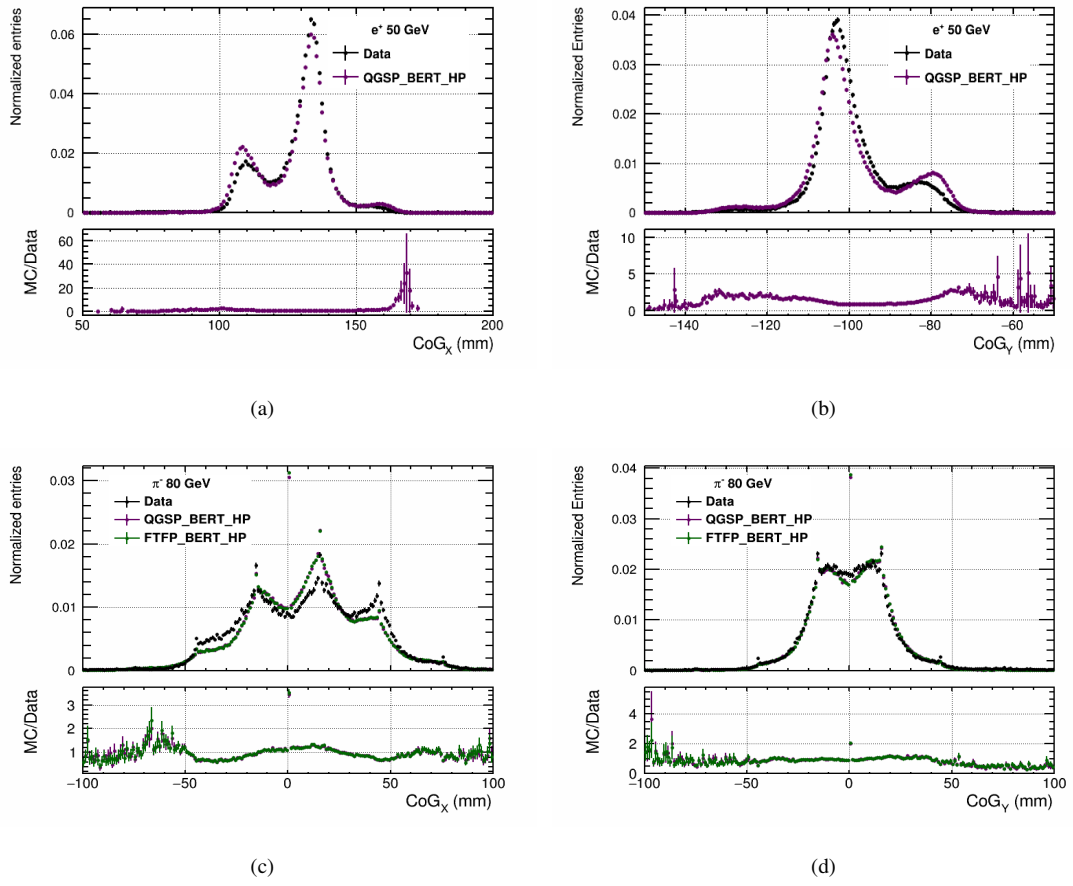


FIGURE 5.1: [
Centre of gravity in x and y-direction]Centre of gravity (a) in x-direction and (b) y-direction, for 50 GeV electron shower compared between testbeam data and simulated data (`QGSP_BERT_HP`). Centre of gravity (c) in x-direction and (d) y-direction for 80 GeV Pion shower compared between testbeam data and simulated data (`QGSP_BERT_HP` and `FTFP_BERT_HP`). Lower window shows the ratio of simluation to testbeam data.

Figure 5.1 shows good agreement between the distributions of the CoG in the x-and y-directions from testbeam data and simulations using electron showers with beam energy at 50 GeV and

pion showers at 80 GeV. The peak structure in the distributions reflect the granularity of the calorimeter, and they correspond to the AHCAL tile size. The distance of the peaks corresponds to the width of a tile, i.e. 30 mm. Each energy deposition in a tile is assigned to the center of the tile. The lower pad shows the ratio of simulations to testbeam data. The agreement looks good, confirming that the beam profiles are simulated reasonably well. This agreement has been found for all the measured energies. The CoG distributions from particle types such as muons, electrons and pions are included in appendix B.

5.2 Shower Start Algorithm

Determining the position of shower start is not only crucial for the study of hadronic shower shapes but also for particle identification, leakage correction and to determine λ_{I} of the detector. The definition of shower starting point is based on topological resolution. The shower start algorithm is utilized in this thesis and is briefly described in this section.

Unlike electrons and photons, neutral hadrons travel a considerable length before interacting via the strong force. Therefore, the particles already lose some energy before showering. This leaves a signal in the detector due to ionization and knock-on electrons creating a MIP track before a hard interaction occurs. The position of shower start is identified at the point where the hadron undergoes its first hard interaction in the detector. Determining the position of the first hard interaction allows disentangling the shower start fluctuations and the hadronic shower fluctuations.

Moreover, to detect the shower start, it has to be insensitive to the detector signal originating from a minimum ionizing particle and the detector's random noise. Based on these considerations, the energy deposition that exceeds a threshold in energy and the number of hits criterium is defined as the shower start origin. The first hard interaction of the showers is found using the `ShowerStartFindingProcessor` in the CALICE software package. The shower start algorithm was initially developed by M. Chadeeva and adapted by D.Heuchel to the current technological AHCAL prototype [97].

The shower starting point is identified using the following relations: Starting from the first AHCAL layer, an average visible energy E_i in MIP with a moving window of five consecutive layers up to i-th layer and the number of hits in the i-th layer, N_i were analyzed on a layer-by-layer basis. The application of moving average for the values E_i helps to minimize the impact of noise.

If the conditions:

$$(E_i + E_{i+1}) > \left(6.0 + 0.1 \cdot \frac{E_b}{\mathrm{GeV}}\right) \mathrm{MIP},$$

and

$$(N_i + N_{i+1}) > \left(3.77 + 1.44 \cdot \ln\left(\frac{E_{\mathrm{b}}}{\mathrm{GeV}}\right)\right),$$

where $E_i = \sum_{k=0}^{i} E_k/(i+1)$ is the accumulated average energy deposit with E_k being the energy deposited in layer k of the AHCAL (layer 1 is the first layer from the calorimeter front). E_b is the incident beam energy, and N_i is the number of calorimeter cells above the 0.5 MIP threshold in layer i. The index i runs from 0 to 39, accounting for the total number of layers in AHCAL. If the above conditions are satisfied, the i-th layer is considered to be the shower starting layer. Else, the layer i-1 is taken as the layer where the shower starts. However, these average values are subjected to large event-to-event fluctuation. Hence both conditions are chosen to minimize the fluctuations. For both criteria, the sum of visible signals in two successive layers is used to distinguish between shower development and local Landau fluctuations. The thresholds are energy-dependent because of fluctuations in the mean values of the energy loss of minimum ionizing particles, which increases with energy. The energy dependence of the threshold value was parametrized by a first-order polynomial. The threshold and limits have to be optimized, and this is done by using the shower start finder on simulated data. In general, the optimization depends on the chosen physics list. For the current data, the thresholds were estimated using pion `QGSP_BERT_HP` physics list for beam energies between 10 GeV and 100 GeV. This was done by minimization of the RMS deviation of the found shower starting layer using the MC truth information. The simulated data have shown that the average difference between the found and true shower start layer does not exceed one layer for 80% of the events and two layers for more than 90% of the events in the studied energy range from 10 GeV to 100 GeV. The performance of the shower start finder is shown in figure 5.2.

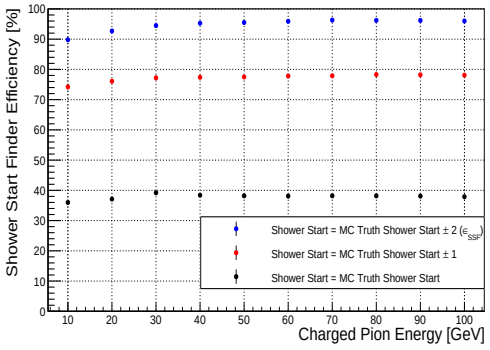


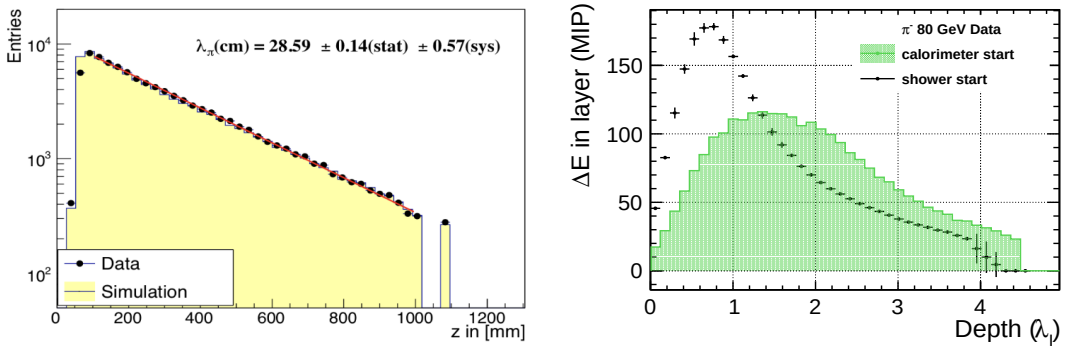
FIGURE 5.2: Efficiency of the shower start finder compared to MC truth information, simulated using `QGSP_BERT_HP` physics list for pion energies between 10 GeV and 100 GeV [101].

Based on the distribution of the shower starting position, it is possible to directly extract the effective nuclear interaction length of pions in the material mix of the AHCAL. This is done by fitting the shower start distribution with an exponential form: $\mathrm{N(z)} = \mathrm{N_0}\ \exp(-\mathrm{z}/\lambda)$ as shown in figure 5.3(a). The major uncertainty introduced by the algorithm is observed to be in the first calorimeter layer.

The fluctuations in the shower starting position is significant due to a single particle's statistical behaviour, which can impact the energy resolution. Therefore, identifying the shower start

position allows disentangling these fluctuations from the intrinsic longitudinal shower development. The distributions of two longitudinal profiles are shown in figure 5.3(b), one shower profile relative to the calorimeter front face and one relative to the shower start. The latter has significantly shorter shower length since only the effective length of the hadronic shower is taken into account, excluding the path of the primary beam particle before the first hard interaction. However, the trend of the agreement with data of the physics lists considered is independent of the definition of the profiles. The unit chosen to express the longitudinal development of showers is the nuclear interaction length λ_{I}. For the AHCAL, it has been estimated to correspond to 231.1 mm.

In this thesis, the results on hadronic shower shapes are discussed for profiles relative to the shower starting position.



(a) (b)

FIGURE 5.3: (a) Determination of AHCAL pion interaction length λ_I with beam energy at 30 GeV using testbeam data [102]. (b) Longitudinal shower profile induced by pion showers using testbeam data with beam energy at 80 GeV, for run relative to the calorimeter front (filled histogram) and relative to the first interaction (black line), which are more peaked than those measured from the calorimeter front.

5.3 Event Selection

The analysis of pion showers is the central part of this thesis. Therefore, it is essential to efficiently select the single-particle events of the correct particle type from the recorded testbeam sample. A pre-selection of electrons and pions is done at the beamline using the information from the Čerenkov detector. This is done by selecting the particles that create a signal in the detector and coincidence from the trigger scintillators. It is possible to apply tags to separate particles based on mass. However, the pion and muon separation can not be reliable due to their similar mass.

As contamination from other particles can always exist in the calorimeter, applying additional event selection without any bias is necessary. Based on the general principles, the events are classified into three groups according to the particle species, with the knowledge that muon

events or MIP-like events that are caused by non-showering charged hadron typically deposit one hit per layer and roughly the same amount of energy. In contrast, events with electrons or positrons produce a dense shower at the beginning of the calorimeter with fewer shower to shower fluctuations. In comparison, pions deposit MIP-like tracks before the first hard interaction, after which the shower evolves throughout the calorimeter with a significant event to event fluctuations.

In this analysis, the offline event selection was applied using a dedicated **P**article **Id**entification (PID) algorithm based on **B**oosted **D**ecision **T**echnique (BDT). This algorithm has been developed by V.Bocharnikov and is used in this thesis to obtain a good classification of the particles and clean event samples. The algorithm uses topological information due to the possibility of high granularity of the AHCAL. The details of this algorithm are found in [103, 104].

The observables, used by the PID algorithm for the classification of particle species are as follows:

- Number of hits per event above 0.5 MIP threshold (nHits)
- Identification of shower start layer per event (st)
- Energy sum (esum)
- Mean shower radius of an event in mm (radius)
- CoG of event in the beam direction in mm w.r.t the beginning of the first layer (zcog)
- Energy fraction in first 22 layers (frac22)
- Energy fraction in central region of shower (fracCentral)
- Energy fraction in shower core (fracCore)
- Number of layers with hits from last 5 (nLastLayers)
- Number of hits after shower start (nShowerHits)
- Mean hit energy after shower start (meanEhitAfterStart)

The electron events are selected by using the combination of the number of hits and the longitudinal CoG of an event to distinguish between muon and pion events. Moreover, the observable sensitive to electromagnetic shower is the fraction of energy in the first 22 layers (since their maximum depth was optimized up to layer 22) corresponding to the absorption of 95% of the average electromagnetic shower. In the transverse direction, the fraction of energy in the shower core with a radius smaller than 30 mm corresponds to one Molière radius with the absorption of 90% of the energy of an average electromagnetic shower. In addition to using the number of hits and the longitudinal CoG of an event, the muons can be well separated from the electron and pion events. Which is done by using the mean hit energy deposited in the cells after shower start as well as using the number of hits in the layers after shower start and the number of layers with hits from the last fice layers. Nevertheless, there are a small fraction of pion events that are similar to muons. These are either pions that decayed into a muon during the flight to the detector or pions that traverse the calorimeter without initiating a shower and have the same calorimeter signature as muons. Therefore, both types of events are removed from the pion samples.

Rejection of multi-particle or empty events − MIP tracks are constructed in a tower with the same x-and-y coordinates for up to the first 5 layers. If the MIP track or clusters are larger than

one, it is considered a multi-particle or an upstream showering event. This will ensure that most showers detected in the AHCAL originated within the AHCAL. Also, empty events caused by fake triggers are rejected for events with hits below a threshold value of 30 cells.

In addition, the particle identification algorithm performs the reconstruction of secondary tracks using minimum number of hits in cluster in a defined 3D volume that groups the neighbours in a volume of $1\times1\times2$, with a requirement of minimum hits in the cluster to be 5.

The PID observables mentioned earlier are trained on simulation data with a priori known particle type, such as applying a multivariate BDT model created by LightGBM [1], on a simulated dataset containing a mixture of all three particle types. This method determines the degree to which an event belongs to the three classes (muon, electromagnetic, hadronic).

The performance of the event classification model using a BDT approach is acquired using simulated data for different particle species as shown in figure 5.4 for the beam energy between 10 GeV and 200 GeV. The purity of the events were acquired with a significant rejection of contamination for all the available pion energies. Hence providing an excellent separation power defined as the area under the **R**eceiver **O**perating **C**haracteristic (ROC)[2]. The separation power between muon and electron is not shown as the possibilities of confusion in identification is minimum. The efficiency of the hadronic signal is better than 99.93% above 30 GeV. While the efficiency of the hadron signal is smaller at lower beam energies due to the similar behaviour in shape and energy as the electron showers. At the same time, the probability of false-negative identification for electron and muon events were 0.3% and 0.2%, respectively.

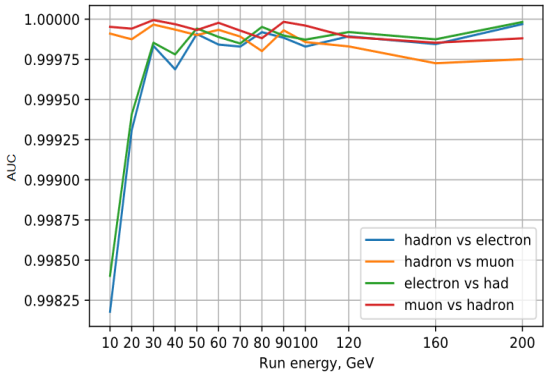


FIGURE 5.4: PID performance using simulated data in the energy range between 10 GeV and 200 GeV for the three classifiers (electron, hadron and muon) based on area under the ROC curve [103].

Furthermore, additional event filtering is applied which are as follows:

- Shower starting in the front part of the AHCAL − a cut on the shower start variable is used to select electrons that interact in the first layers and reject pions that start to shower

[1] https://lightgbm.readthedocs.io/en/latest/

[2] A ROC curve characterizes the performance of a classification model by showing the ratio of false-positive rate to true-positive rate. The integral under the curve evaluates the identification power of the model, where 0 indicates a 100% inaccurate model and 1 indicates a 100% valid model.

later or a particle without interaction (muons/muon-like). In addition, the first physical AHCAL layer is excluded due to uncertainties in shower start identification. Moreover, events with showers starting beyond the sixth AHCAL layer are excluded to minimize leakage of shower energy into the TCMT.

- Requiring a single track for the event and calorimeter hit match in one of the first three AHCAL layers.
- Apply gap rejection of 2.0 mm to require the impact point to not be in between the HBU gaps of calorimeter layers slabs.
- Select events in MC that are within the statistics available in data, this is due to the acceptance area of the trigger scintillator and wire chamber (10×10 cm^2) before the detector. The selection cuts on the CoG applied according to the beam energies used in this thesis are listed in table D.1. These selection cuts are energy dependent as the beam was not always pointing exactly at the center of the detector. Moreover, this selection was applied to shower shapes analysis described in chapter 7.

The event selection procedure is applied to simulated data as well. The results from event selections for the reconstructed energy sum before and after event selection using testbeam data is shown in figure 5.5(a) for beam energy at 40 GeV muon, for 80 GeV electron in figure 5.5(b) and 80 GeV pion in figure 5.5(c). It is observed that after the event selection, the peak due to muons and late pion showers before applying the event selections in the pion energy sum distribution is no longer present afterwards. This leads to an improvement in the width of the energy sum distribution and shows more Gaussian-like distribution. For the electrons, the left tail caused by low energy electrons, coming from electrons starting to shower earlier, is also reduced after applying event selection. In principle, contamination from hadrons can contribute to this low energy tail. Thus, the rejection of these particles is ensured within the particle identification algorithm with a slight improvement in the rejection at 10 GeV between hadron and electron.

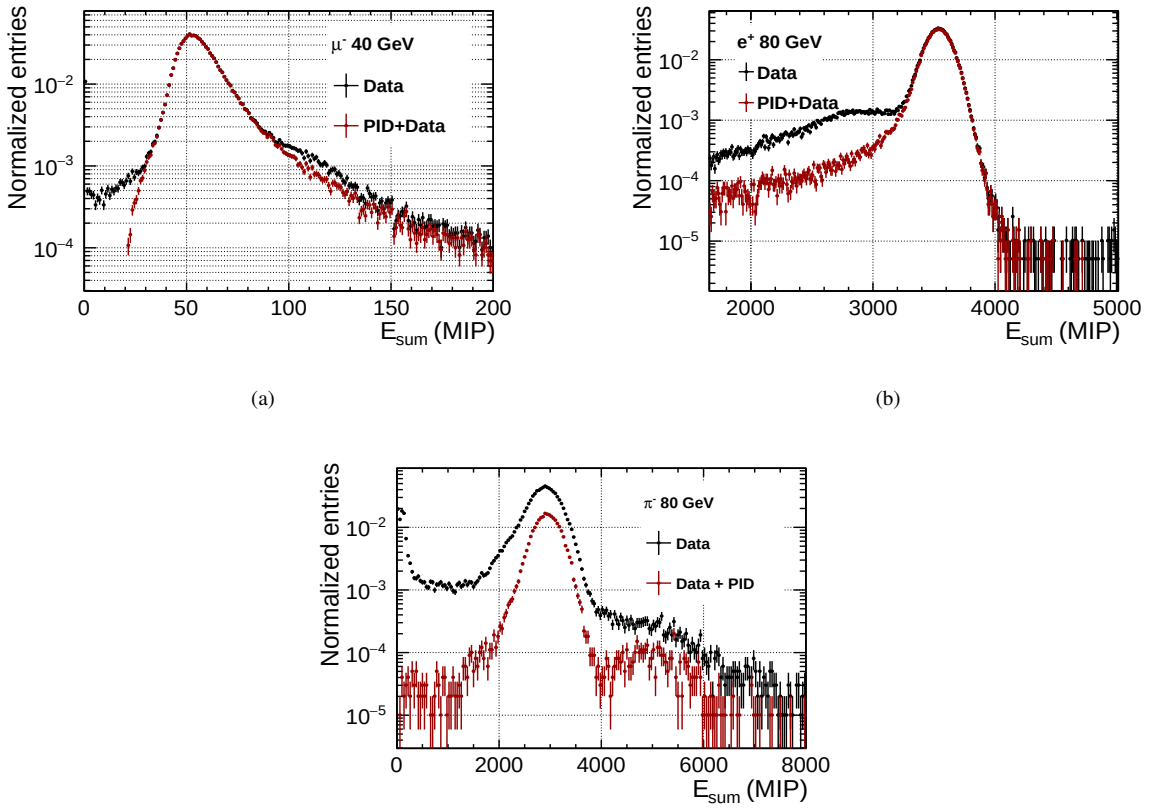


FIGURE 5.5: Shows the reconstructed visible energy sum distribution E_{sum} from testbeam data before event selection (black) and after applying event selection (red) using, (a) muon beam with beam energy at 40 GeV, (b) electron shower with beam energy at 80 GeV and (c) pion shower with beam energy at 80 GeV.

Event selection efficiencies applied on the testbeam data and simulated data for the available beam energies of muon, electron and pions are summarized in table 5.1. Most of the events are rejected after applying the shower start finder. Roughly the same efficiency is obtained between testbeam data and simulations. In general, the efficiency for 10 GeV pions is much lower due to the amount of electrons in the beam being higher, leading to reduced pions statistics.

Particle Type	Run Number	Beam Energy [GeV]	Data All Events	Selected Data Events	Selected Data %	Selected MC QGSP_BERT_HP %	Selected MC FTFP_BERT_HP %
π^-	61265	10	115053	20081	17.5	22.8	23.0
	61273	20	94016	20540	22.0	26.0	26.2
	61384	30	49910	9373	19.0	27.5	27.6
	61275	40	128519	31041	24.2	28.7	28.7
	61262	60	111594	27281	24.4	29.2	29.3
	61279	80	99293	25826	26.0	30.2	30.3
	61287	120	101083	28319	28.0	28.3	28.4
	61222	160	78200	19561	25.0	26.6	26.8
	61201	200	95753	24228	25.3	23.2	23.3
e^+	61293	10	85875	58611	68.3	89.6	✗
	61296	20	30338	23408	77.2	93.1	✗
	61213	30	197284	177873	90.2	94.4	✗
	61212	40	197701	182776	92.5	95.1	✗
	61214	50	198093	183579	92.7	95.4	✗
	61211	60	197781	181128	91.6	95.3	✗
	61215	70	196826	175852	89.3	95.7	✗
	61210	80	196138	165954	84.7	96.0	✗
	61216	90	195573	150686	77.0	96.1	✗
	61217	100	161829	106166	65.6	96.2	✗
μ^-	61293	40	158829	151572	95.4	97.0	✗

TABLE 5.1: Set of runs investigated in this thesis, for events before and after selections from testbeam data and simulated data using different physics lists. All simulations are generated with GEANT4 10.04.p2. The size of each simulated sample is 500,000 events per run.

In this chapter, the beam profile calibration and the agreement of simulations with testbeam was shown using the recorded muon, electron and pion data. The event selection procedure applied to testbeam data and simulated data have been explained. An excellent separation power is obtained classifying the different particle species (muon, electron and pion) over the simulated energy range from 10 to 200 GeV. In the next chapter, the validation of the AHCAL simulation model is presented to confirm the calibration procedure.

Chapter 6

Detector Validation Studies

"The best way to show that a stick is crooked is not to argue about it or to spend time denouncing it, but to lay a straight stick alongside it."

D.L. Moody

DETECTOR VALIDATION provides a valuable basis for checking the understanding of a hadron calorimeter. This study allows us to test the reconstruction chain and the applied calibration factors. This chapter presents the validation of the calibration procedure and the detector performance in terms of energy linearity and detector resolution, as these steps are prerequisites for analyzing pion data.

Muon and Electron data provide a valuable basis for checking the understanding of a hadron calorimeter. For muons and electrons, we can assume that the simulation should describe the data. This tests how well we know the effects of the real detector in the simulation, the implementation of detector features in the digitization and the calibration factors applied in the reconstruction chain. Muon are mainly tested in the low hit energy domain. As discussed in the previous chapter 4, the prototype is calibrated using muons. As in testbeam data, the ADC to MIP conversion factor is obtained by applying the fit to the uncalibrated hit energy distribution. On the other hand, in simulations, it is based on the value of GeV to MIP conversion factor as described in section 2.4. Therefore, the consistency of the MIP calibration needs to be cross-checked.

In contrast, electron data is mainly tested for high hit energies due to the higher cell energy deposits resulting in SiPM saturation. This effect has a significant impact on visible energy. Therefore, it is essential to validate simulated electrons to testbeam data. Once simulation of detector effects with muons and electrons are validated, the hadron shower models in `GEANT4` are then validated using pions.

Five standard quantities have been compared for validation of the simulations with testbeam data. These quantities are the distribution of the number of hits per event, the energy sum per event, which is the total energy measured per event in units of MIP equivalents, the mean number of hits per layer, the mean energy sum per layer and the hit energy distribution. Furthermore, to

judge the quality of the detector, its linearity and single-particle energy resolution for electron and pion data is also discussed.

The set of data runs used in this chapter and the beam profile obtained for muon, electron and pion simulations are given in table B.1. Section 6.1 presents the validation of the simulation with muons between testbeam data and simulation using the QGSP_BERT_HP physics list. In section 6.2, validation of electrons is performed between testbeam data and the simulation using the QGSP_BERT_HP physics list. The QGSP_BERT_HP uses standard electromagnetic physics model with the HP package for neutrons as discussed in section 1.3.

In addition, saturation correction studies are performed for electrons as well as the effect on testbeam data and simulation is investigated and compared. In section 6.3 pions are compared between testbeam data and simulations using the QGSP_BERT_HP and FTFP_BERT_HP physics list. Moreover, for the pion dataset, the linearity and resolution are obtained using the energy information from the tail-catcher (figure 3.3) combined with the AHCAL. Finally, the key parameter values for energy linearity and resolution are summarized.

6.1 Validation with Muons

A detailed study of muon data recorded with the AHCAL was done by S.Huck and is reported in [105]. As a first step towards validating the simulation using QGSP_BERT_HP physics list, the distributions of the mean visible energy deposit (E_{sum}) and the number of hits (N_{hits}) per event, with hits above a 0.5 MIP threshold, are compared between testbeam data and simulated data for muons with a beam energy of 40 GeV.

Figure 6.1(a) shows the number of hits per event. The agreement between simulation and testbeam data is found to be 30-40%. The simulation reproduces the peak position of the hit multiplicity. A clear population can be seen at layer number 39, with one hit per layer in the detector for both data and simulation. However, more events are found with a smaller number of hits in the testbeam dataset compared to the simulation.

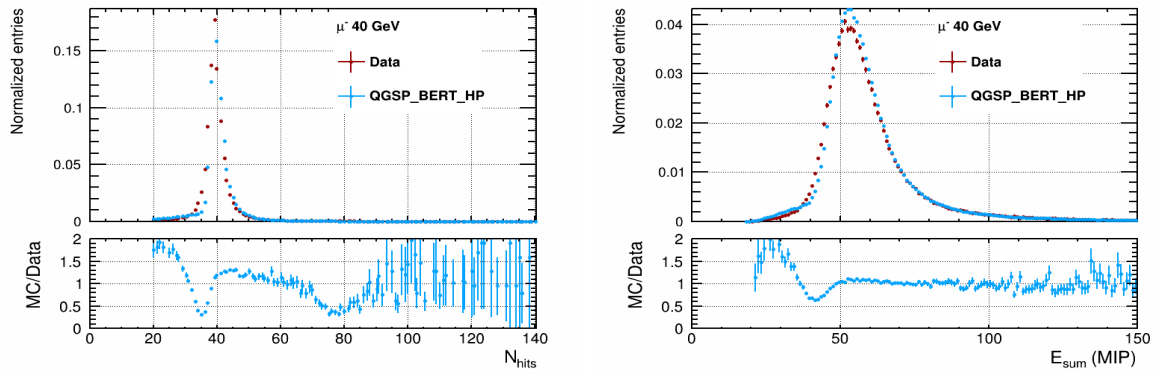
Figure 6.1(b) shows the mean visible energy sum. Although it can be seen that simulation slightly overestimates the data, the simulation reproduces the peak and shape of the energy sum distribution of the testbeam data. The testbeam data and simulation have their maximum slightly above 50 MIP, and the long tail at high energy in testbeam data and simulation is caused by double muon tracks. Almost no tail is seen on the left side of the distributions. On the other hand, the tail on the left of energy sum distribution shows disagreement between data and simulation. This effect is contributed from number of hits and the hit energies of an event.

Figure 6.1(c) shows the profile of the mean energy in a given layer and is expected to be above one, as a characteristic of the Landau distribution. The strong rise seen in the first ten layers goes from approximately 1.4 MIP to 1.65 MIP for the testbeam data and more than 1.65 MIP for the simulated data. The agreement in mean energy per layer between testbeam data and simulation is about 2-5%.

Figure 6.1(d) shows the profile of the mean number of hits per layer. A very similar feature to that seen for the mean energy per layer is also observed in the mean number of hits per layer. The distribution is close to value of one in the first layer. It then increases strongly over the first ten to eleven layers to approximately 1.02 hits per event for data and 1.04 hits per event for the simulation. After that, the mean number of hits per layer remains approximately constant. The increase is much stronger for the simulation than for the data, as is clearly seen from the ratio plot. The agreement between the data and simulation is $\sim$2% throughout the layers. For both the mean energy per layer and the mean number of hits per layer, the trend is well described by simulation, but clearly the plateau region is not. Nevertheless, the agreement is found to be reasonable as no track selection algorithm is considered for the comparisons, which is the most likely explanation for this lack of agreement.

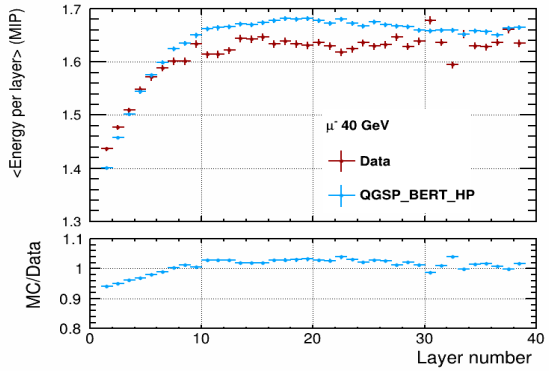
For muons, there is no impact of the saturation correction, because the cell energy deposits are typically at the level of 1 MIP. Figure 6.1(e) shows that the hit energy distributions have a very similar shape for testbeam data and simulation and agree within 10% percent. Nevertheless, when looking at the ratio of the two distributions, it can be seen that the simulation has an excess of hit energies.

Overall the findings agree with the previous analysis [105]: the simulation has slightly too many muons producing delta rays or showers, so the mean hit energy, the mean number of hits (total and per layer) and also the total energy are a bit higher in simulation than in data. All these can be explained by just one effect.

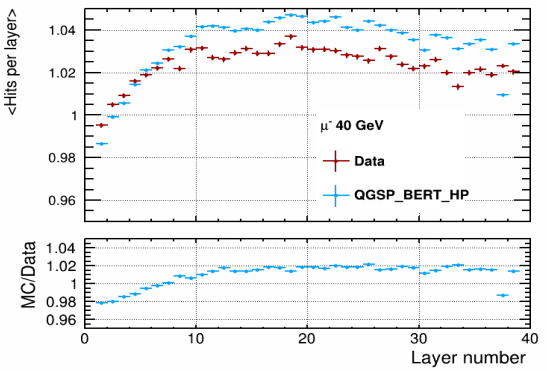


(a)

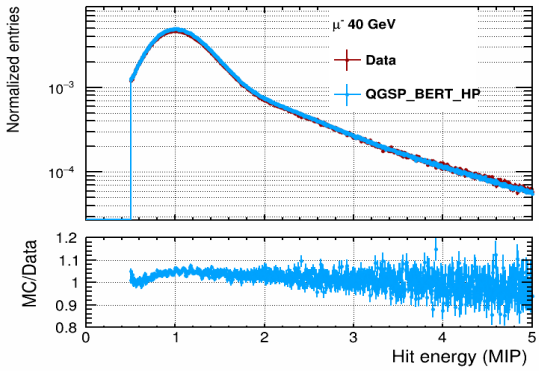
(b)



(c)



(d)



(e)

FIGURE 6.1: Comparision of standard reconstructed quantities for 40 GeV muon using testbeam data (red) and simulated data (blue). Lower window shows the ratio of simulated data to testbeam data. (a) The number of hits distributions per event (N_{hits}). (b) Distribution of the visible energy deposit in MIP equivalents. (c) Mean energy per layer in MIPs. (d) Mean number of hits per layer. (e) Hit energy distribution in MIP equivalents

6.2 Validation with Electron Showers

The physics of electron showers are theoretically well understood and can be simulated with minimal uncertainty. Therefore, electron showers provide an essential tool to validate the detector and the simulation model. Besides using the correct beam profiles, an additional effect must be considered for the detector responses of electron showers, which is the de-saturation of the high hit energy deposits of the electron shower. Currently, no method exists that could verify the saturation correction for all AHCAL channels, as the saturation curves that have been recorded during the testbeam data collection are only for a handful of central cells. However, in these cases, the saturation could be verified since most of the energy is deposited in these cells with high hit amplitude regions. Therefore, the number of effective pixels is estimated for these central cells.

The following section studies the saturation correction before comparing the standard quantities to obtain an optimal number of effective pixels using electron showers. Afterwards, this section presents the validation of electron simulation with testbeam data. Then the energy linearity and resolution for electrons is obtained.

6.2.1 SiPM Non-Linearity Correction

The response function of a SiPM describes the number of fired pixels as a function of the number of incoming photons. An important role is played by the statistical fluctuations of the fraction of pixels fired for very high amplitude hits. These hits lead to large correction factors of saturated response by the SiPMs, amplifying this effect. The saturation effect is expected to be more relevant for hits in which the number of incoming photons exceeds the number of effective pixels of the silicon photomultiplier. Therefore, for complete calibration of a cell, it requires us to account for the non-linearity introduced by the limited number of pixels per SiPM and the finite pixel recovery time.

In the AHCAL digitization chain, the amplitude of each hit A_i is saturated, using an approximated function given in Eq. 2.6. During the AHCAL reconstruction, the signal of every SiPM is corrected for the saturation behaviour by multiplication with an individual correction function. This correction is extracted as the residual to the linearity of the inverted SiPM response function $f_i^{-1}(A_i)$ and is given by:

$$f_i^{-1}(A_i) = A_i^c = -N_{eff} \cdot \ln\left(1 - A_i^s / N_{eff}\right). \tag{6.1}$$

The saturation correction function A_i^c represents the energy of each hit in the detector corrected for the saturation effect. N_{eff} represents the number of effective pixels - a parameter that has to be measured or tuned, and A_i^s is the hit energy uncorrected for saturation effects. In addition, if more than 95% of the pixels are fired, a linear approximation is used to correct the saturation effects. The function is then written as,

$$A_i^c = \frac{1}{1-0.95} \cdot (A_i^s - 0.95 \cdot N_{eff}) - N_{eff} \cdot \ln(1-0.95). \tag{6.2}$$

The studies on saturation behaviour for the SiPMs (MPPC S13360-1325PE), were done by the University of Mainz [106]. The specification and details of this SiPM are given in table 2.1. The measurements for the saturation curves were done for the bare SiPMs, which were not mounted on an AHCAL tile by using laser light. For these measurements, a dedicated set-up was used to ensure that all the pixels were illuminated homogeneously. In this study, a simple exponential response function as given in Eq. 2.6 was used for modelling the saturation of the SiPM response. As a result, the number of effective pixels defined by this study is 2553 pixels. This is $\sim$10% less than the nominal number of pixels (2668 number of pixels). However, no measurements of the saturation curve for the SiPMs used in this prototype were done with SiPM on the tile. Therefore, it was necessary to test the saturation effects after the SiPM was mounted on the AHCAL tile. Moreover, the electronics can introduce additional saturation effects. The current method measures the complete chain of the tile with the contribution from SiPM to ASIC. Therefore, these components cannot be disentangled.

To understand the behaviour of saturation effects from the silicon photomultipliers and to tune the value of the effective number of pixels, the electromagnetic showers are being used, as they are the most sensitive to saturation for high energies and large amplitudes. The variable used to correct the number of effective pixels arises from the hit energy distribution. For this study, simulated electron data have been generated during the digitization, a fixed value for the number of effective pixels is applied with a value of 2533, as proposed by the Mainz study. The simulated data is reconstructed using a saturation correction function with the same number of effective pixels. The simulated data has been compared with the testbeam data applying the saturation correction in the reconstruction procedure. The reconstruction procedure is kept consistent between testbeam data and simulated data.

Furthermore, the hit energy distributions are compared using testbeam data and simulated data for an electron with beam energies of 10 GeV and 80 GeV.

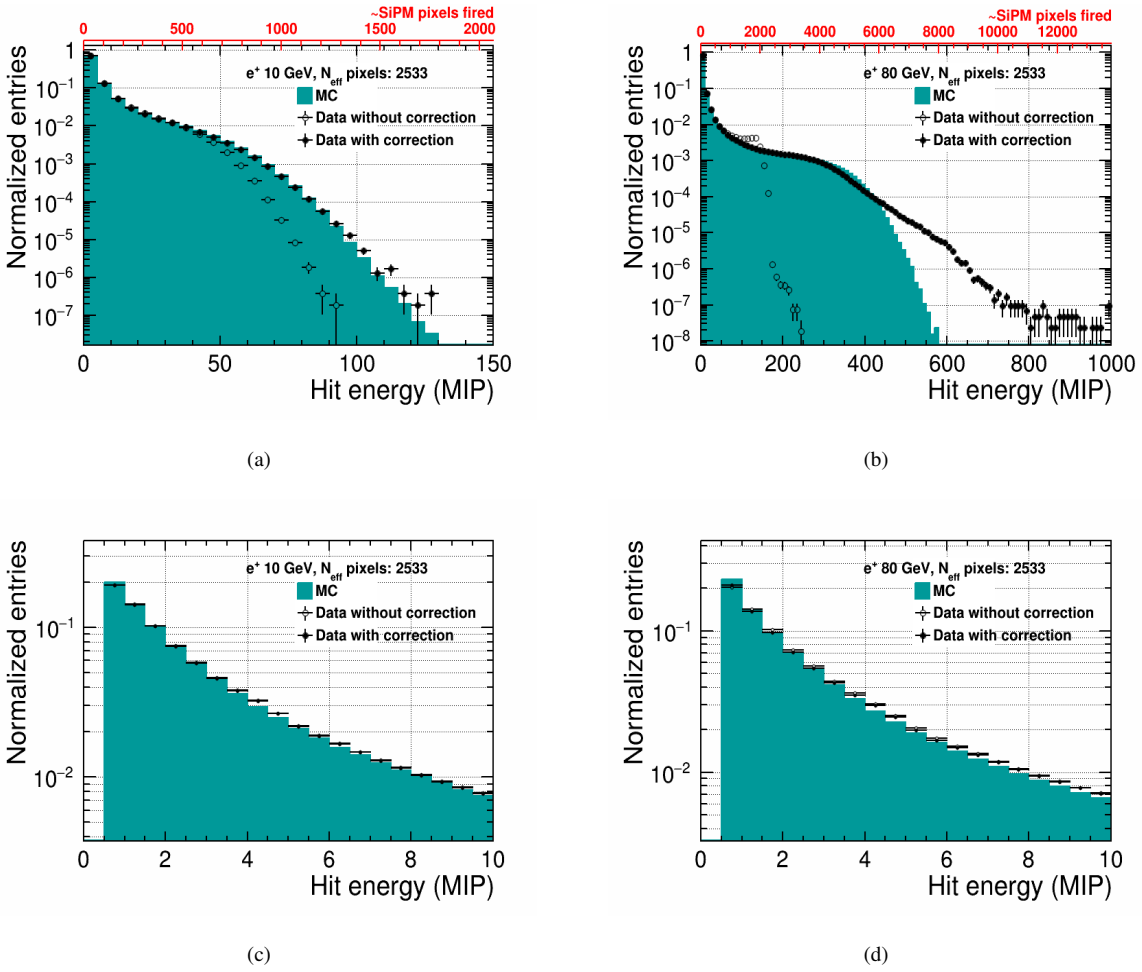


FIGURE 6.2: Comparison of the hit energy distribution for electron showers from simulated data (green) and testbeam data (black) after applying a saturation correction using 2533 as the number of effective pixels. (a) Hit energy distribution for 10 GeV and (b) Hit energy distribution for 80 GeV. Hit energy distribution in the low energy deposits in range betweeen 0-10 MIPs for beam energy at (c) 10 GeV and (d) 80 GeV.

From the first attempt at comparing the testbeam data and the simulation, we observe that saturation does not play a major role at lower energies. An example, for electrons with beam energy at 10 GeV, is shown in figure 6.2(a). However, at higher energies for example at 80 GeV, as shown in figure 6.2(b) a tail at large hit amplitudes can be seen in testbeam data compared to simulation. It is observed that the value of 2553 effective pixels used in the saturation correction function is too low to describe the shape of the hit energy distribution. Moreover, the correction factor for higher energies such as 80 GeV electron shower is roughly a factor 3, which is obtained from the ratio with the maximum amplitude in pixel and the nominal number of pixels of the SiPM. Therefore, it is necessary to define an optimal number of effective pixels for the saturation correction such that the shape of the hit energy distribution is well described between the testbeam data and simulated data.

A detailed study is done to understand the variation in the number of effective pixels from SiPM-to-SiPM. In this study, two values with 2533 and 2668 effective pixels are used in the

reconstruction procedure. Moreover, two datasets are used: first, the dataset after applying the saturation correction and a second dataset without saturation correction applied during the reconstruction procedure. Events (and cells) are selected where the effect of saturation is expected to play a significant role. Therefore, the events in the CoG in X and Y-direction is in the middle of a cell ("tower" of 1×1 cm^2), such that the highest energy deposition in the shower center is contained in one cell (and not spread over two or four cells). Event selection was applied, with the hits that contributed to the cells chosen, and additional wire chamber information was used as it had a precision of better than a mm. This procedure is performed both for testbeam data and the simulated data in a consistent manner. The effect of saturation being highest at the region around the shower maximum, four layers (6,7,8, and 9) around the shower maximum were chosen for the study.

Three tile positions are chosen for study, as given in table 6.1 by applying selections along the X,Y direction in the centre of gravity and similarly to the tracks. The tile selection is made in regions of high hit amplitudes, as shown in figure 6.3. The hit energy spectrum in the corresponding four consecutive layers is investigated. For each position and layer, two distributions are compared between testbeam data and simulated data using 2533 and 2668 as effective pixels.

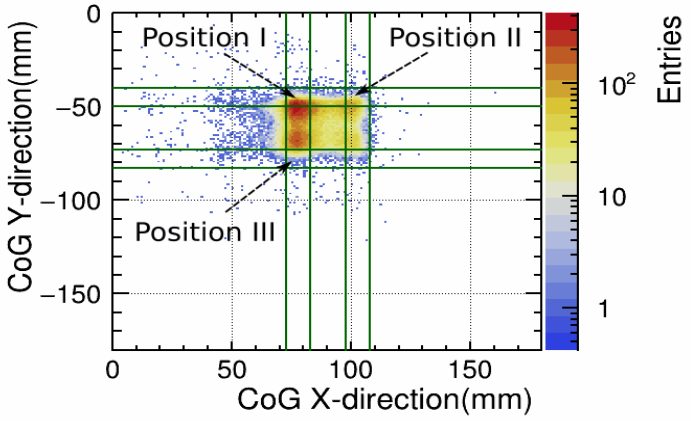


FIGURE 6.3: Centre of gravity plot, depicting position I, II and III tested for saturation correction studies. The green marking shows the region of AHCAL tile acceptance.

Position	X (mm)	Y (mm)
I	[73,83]	[-50, -40]
II	[73, 83]	[-83, -73]
III	[98, 108]	[-50, -40]

TABLE 6.1: Three positions are chosen based on centre of gravity in X-and-Y direction.

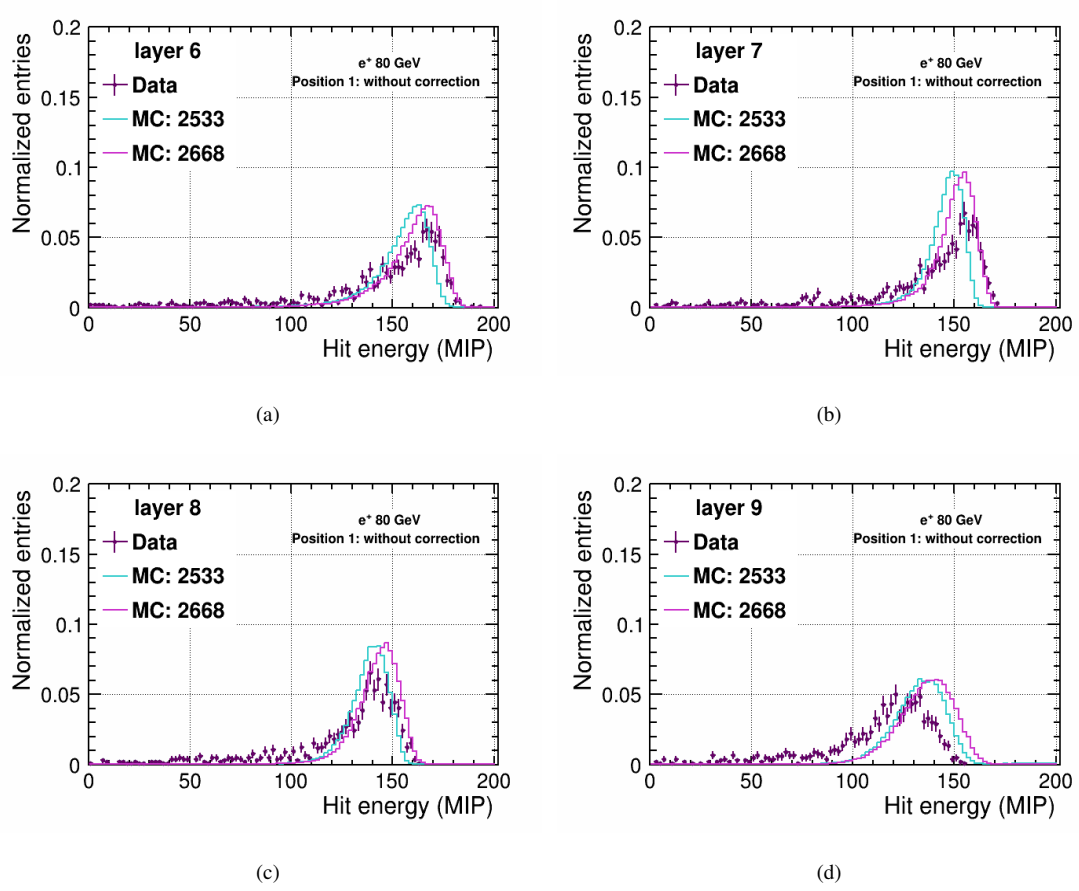


FIGURE 6.4: Layer-wise hit energy distribution without SiPM saturation correction at Position I. (a) AHCAL layer number 6 (b) AHCAL layer number 7 (c) AHCAL layer number 8 and (d) AHCAL layer number 9.

Figure 6.4 shows layer-wise hit energy distributions using 80 GeV electron showers at position I. The comparison is done between testbeam data and simulations with 2533 and 2668 number of effective pixels. The distributions are shown for layer 6 to 9 without applying a saturation correction function. It is observed that the distribution with 2533 effective pixels saturates earlier. Moreover, the distribution suggests a larger number of effective pixels such as 2668 compared to 2533. This feature is clearly visible from the layers 6 to 8.

A similar comparison is done for hit energy distributions after applying saturation correction as shown in figure 6.5. The MC distributions behave nearly the same, due to the linear extrapolation in the de-saturation function. By applying a lower value on the number of effective pixels, the distribution in data gets wider. This broadening is due to finite pixel number. With too small number of pixels in the correction, the large hit energies get a very large correction (as they are in the linear part above the nominal number of pixels), leading to this broadening effect. In general, the peak position of hit energy distributions in simulation shows a shift in mean energy deposition compared to data.

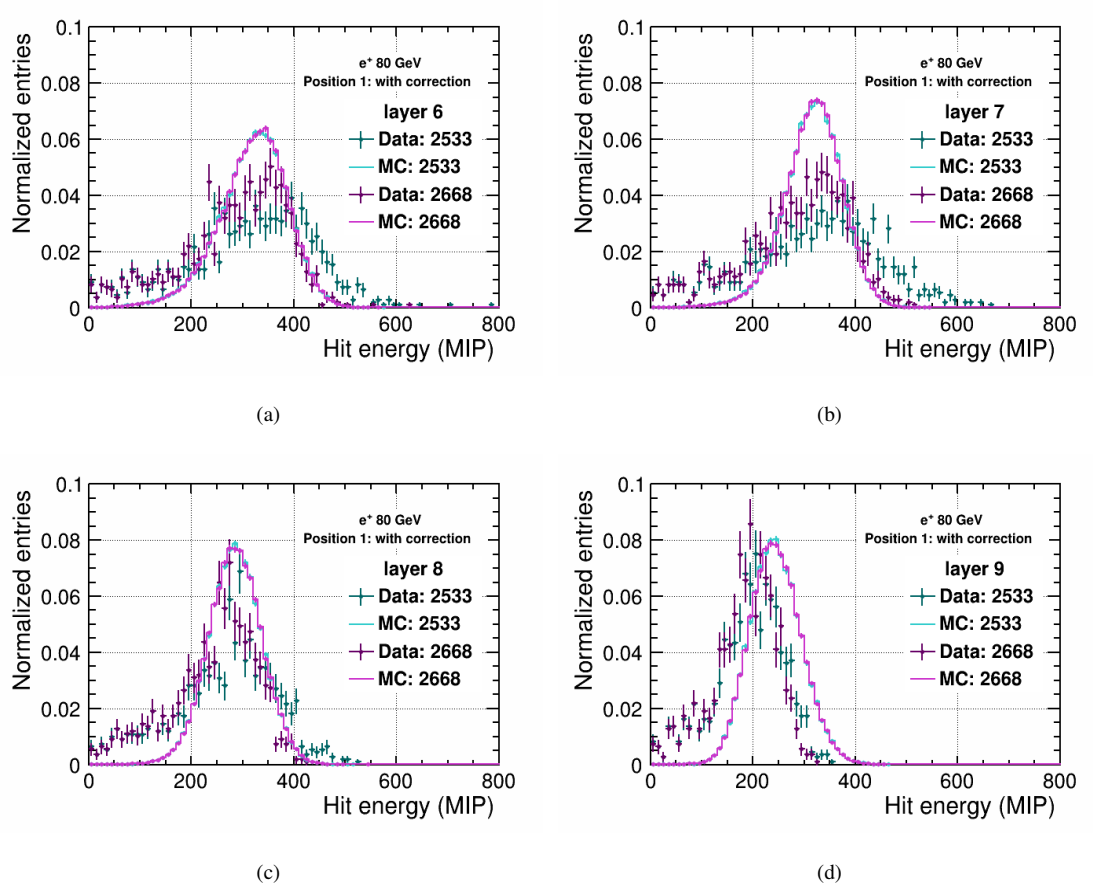


FIGURE 6.5: Layer-wise hit energy distribution after applying SiPM saturation correction at Position I. (a) AHCAL layer number 6 (b) AHCAL layer number 7 (c) AHCAL layer number 8 and (d) AHCAL layer number 9.

In general, the distribution peaks at roughly the same position, but the edge of the distribution is still clearly not captured by using a lower pixel number. It shows a broader distribution and slightly longer tails to the higher edge.

From this position study, it was observed that for layer 6 and 7, the upper edge of hit energies in testbeam data agrees with simulated data by using 2668 as the value for the number of effective pixels. However, for layers 8 and 9, the testbeam data agrees with simulated data by using 2533 as the value for the number of effective pixels. This means that the saturation of the SiPM differs from channel to channel, and applying the saturation correction with a fixed number of effective pixels is not optimal. In this case, the study of the SiPM saturation correction would have to be done per channel, which is not possible due to the lack of statistics available across all channels.

Moreover, a study shown in appendix C was done to check if there were any hints of correlations between the calibration constants and the high energy edge of the hit energy distribution. For this study, the distribution for testbeam data without applying a saturation correction was chosen for the endpoint selection study. This study was done to make sure that not any other mis-calibration is responsible for the channel-to-channel variation.

A threshold value is chosen to avoid that one single high-energy hit determines the endpoint, which could lead to rather large variations of the endpoint value for distributions that look nearly the same otherwise. The high hit edge is chosen if the bin's content is less than 10% of the maximum, such that the bin's lower edge is considered as the endpoint. This procedure is applied at different chosen positions for layers 6, 7, 8 and 9. Figure C.1 shows the calibration constants versus the high hit edge in pixels, each constant corresponds to one channel. With the available statistics, the high hit edge for five channels in each layer was extracted. It is evident that no clear correlation exists between the calibration constants and the high hit edge. This study confirms that there is no problem with the other calibration constants.

In order to have a better estimate of the number of effective pixels and to understand the shape behaviour of the hit energy distribution, the saturation effects are also studied for distributions with different number of effective pixels ranging from 2448 to 2888. The testbeam data are compared to the simulated data with a fixed value of 2668 as the number of effective pixels. Figure 6.6 shows the hit energy distributions for 10, 80 and a 100 GeV electron showers, corrected for different pixels in testbeam data and compared to a fixed number of pixels in simulated data. Different number of effective pixels from testbeam data are compared to the reference value, which is the MC simulation with a value of 2668 as the number of effective pixels. At first glance, all the distributions disagree with simulations. In order to quantify an agreement or disagreement between the distributions, the hit energy distributions are divided into four regions, as given in table 6.2.

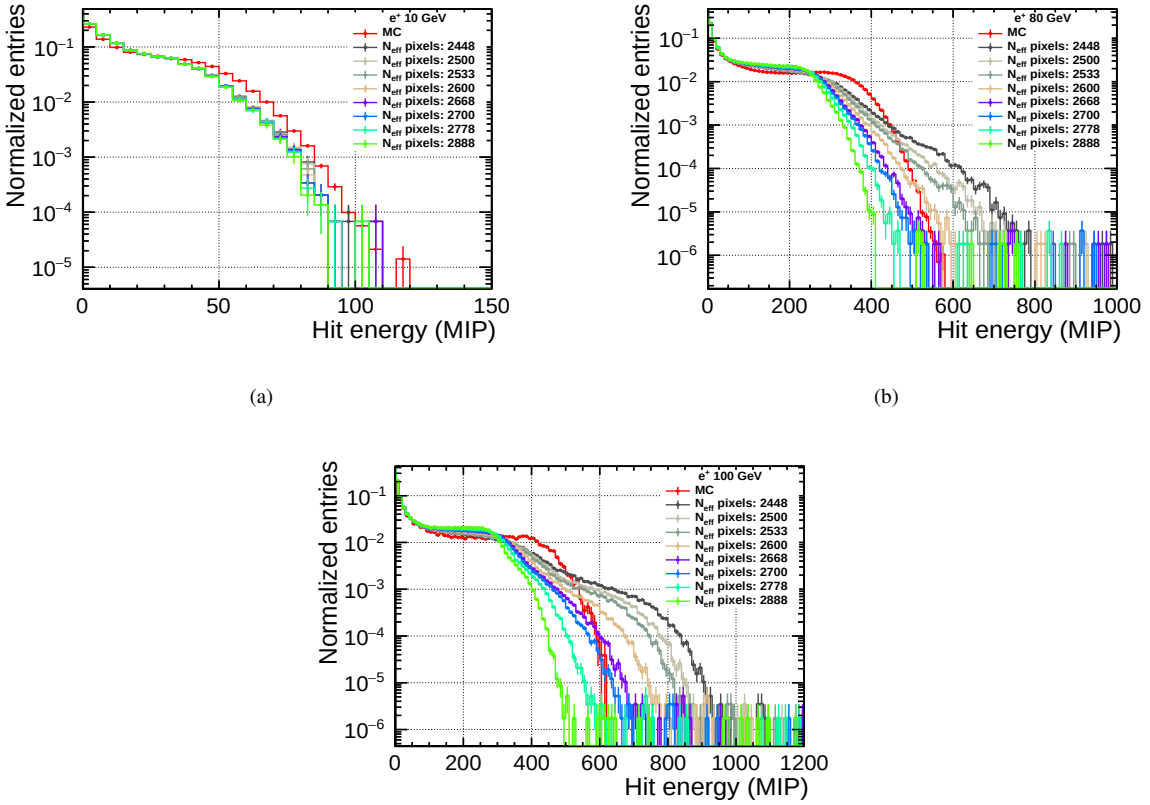
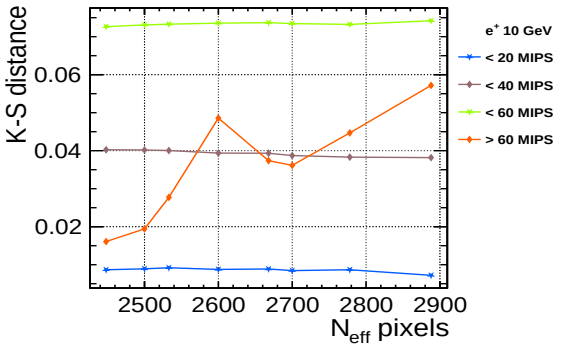


FIGURE 6.6: Comparision of the hit energy distribution for position I from testbeam data for different numbers of effective pixels and simulated electron showers with a fixed number of effective pixels with beam energy at (a) 10 GeV, (b) 80 GeV and (c) 100 GeV. For simulated data, the saturation function is applied during digitization. While the correction function is applied during reconstruction of the datasets.

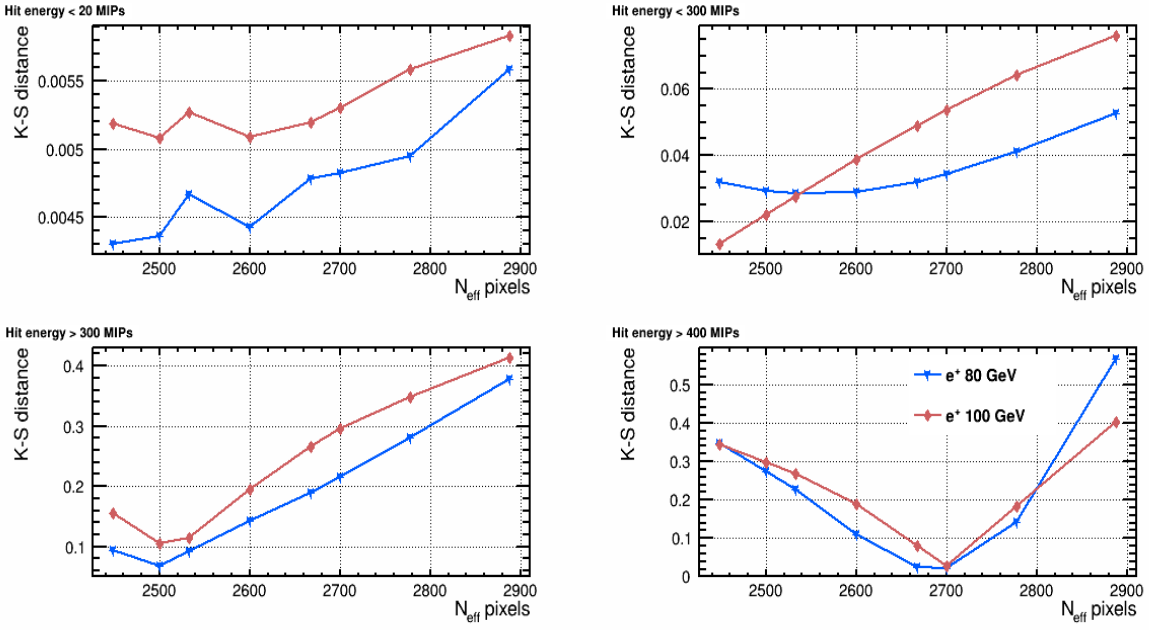
Furthermore, a statistical test is performed using the **K**olmogorov-**S**mirnov **t**est (K-S test) [107]. This method evaluates the compatibility of shape between the distributions. In addition, the K-S test provides better results than a χ^2-test in the case of histograms with low statistics. The K-S test quantifies the distance (K-S distance) between two distributions. Therefore, a lower value of K-S distance denotes better agreement between the two distributions. In the current analysis, a K-S test is applied to the hit energy distributions shown in figure 6.6. The distributions are grouped into different energy regions as given in table 6.2.

Hit region	10 GeV	80 & 100 GeV
I	$<$ 20 MIPS	$<$ 20 MIPS
II	$<$ 40 MIPS	$<$ 200 MIPS
III	$<$ 60 MIPS	$>$ 200 MIPS
IV	$>$ 60 MIPS	$>$ 400 MIPS

TABLE 6.2: Selection cuts applied on different regions of hit energy distribution for beam energy at 10 GeV, 80 GeV and 100 GeV.



(a)



(b)

FIGURE 6.7: K-S test for varying pixel values of 2448 to 2888 (a) for a beam energy of 10 GeV with four region of interest below 20 MIPS (blue), below 40 MIPS (brown), below 60 MIPS (green), and above 60 MIPS (orange) (b) for a beam energy of 80 and 100 GeV in four region of interest.

Figure 6.7 shows the results obtained from the K-S test for 10, 80 and 100 GeV energies. The K-S distance varies strongly depending on the hit energy region for all the beam energies. In the case of 80 and 100 GeV, the hit region I (below 20 MIPs) shows less influence of the number of effective pixels. In contrast, the hit region II (below 300 MIPs) prefers to have a pixel value of 2488 for 100 GeV data and 2600 for the 80 GeV data. For hit region III (above 300 MIPs), the pixel value is at 2500, and for hit region IV (above 400 MIPs), the pixel value is 2700. Therefore, choosing the best pixel value requires a number of effective pixels between 2600 and 2700.

Before choosing the number of effective pixels which will be used for the saturation correction, it is important to do a global study of the SiPM saturation. This serves to validate the key parameters of the calorimeter. For this study, three reasonable values are chosen with 2500, 2668 and 2778 effective number of pixels. These values are above and below the nominal value of 2668.

Figure 6.8 shows the distribution of reconstructed energy with 2500, 2668 and 2778 effective pixels compared to simulated data reconstructed with 2668 effective pixels. Pixel value of 2500 show closer agreement to simulation in mean reconstructed energy. Nevertheless, the distribution is broader compared to the higher pixel values.

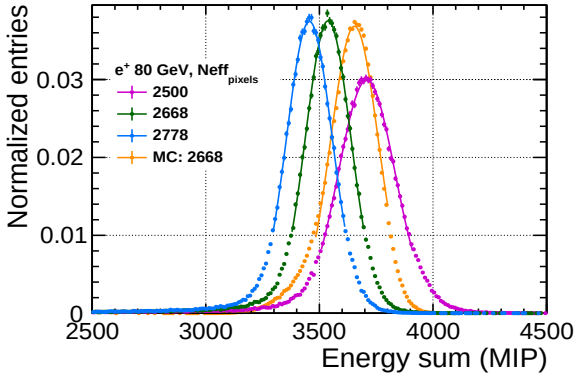


FIGURE 6.8: Distribution of the reconstructed energy using 80 GeV electron showers with pixel values of 2500, 2668 and 2778. The curves show fits with a Gaussian function within a range $\mu \pm 1.5\sigma$.

As the next step, the energy linearity and energy resolution is studied for testbeam data reconstructed with 2500, 2668 and 2778 effective pixels. The distributions are compared to simulated data reconstructed with 2668 effective pixels.

Figure 6.9(a), shows the reconstructed energy as a function of beam momentum for testbeam data reconstructed with varying number of effective pixels (2500, 2668 and 2778). The value of reconstructed energy was defined by using a Gaussian fit function in the central region of the energy sum within a range $\pm 1.5\sigma$. An excess of the energy deposited in MC compared to data is observed. Furthermore, the ratio plot in figure 6.9(a) shows a better agreement of testbeam data with simulation while using 2668 as effective pixels. As the next step, a linear fit is performed to the data points using:

$$\mathrm{E_{reco}[MIP]} = \mathrm{a} \cdot \mathrm{p_{beam}[GeV]} + \mathrm{b}, \tag{6.3}$$

where a is the energy scale factor in $\frac{\mathrm{MIP}}{\mathrm{GeV}}$ and b is the offset in MIP, which takes into account the combined effect of electronic noise and threshold effects. Furthermore, the parameters a and b are determined for each data version and for the simulation.

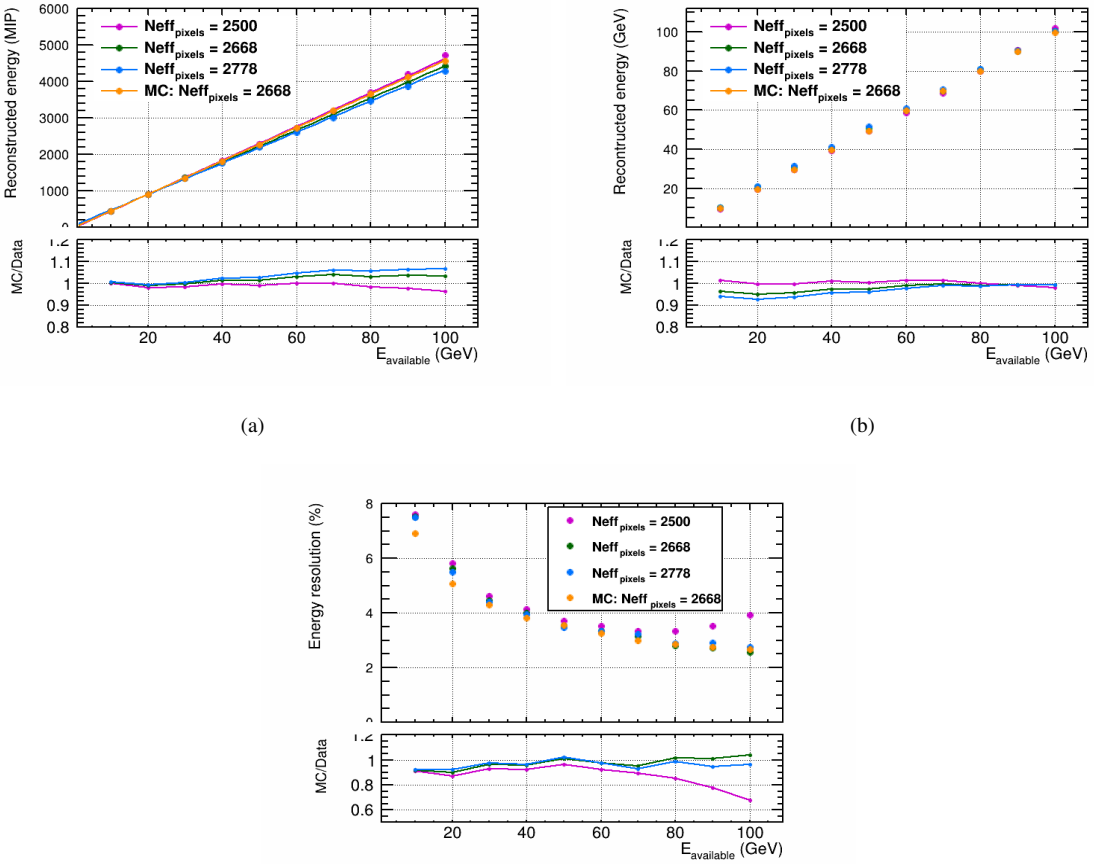


FIGURE 6.9: (a) The upper window is the energy reconstructed as a function of the beam energy from MC and testbeam data with a saturation correction using different numbers of effective pixels. The lower window is the ratio of MC to testbeam data for different numbers of effective pixels. (b) Energy linearity after performing linear fit. (c) The upper window shows the comparison of the shower energy resolution between MC and data with the saturation correction applied using different numbers of effective pixels. The lower window shows the ratio of simulated data to testbeam data for a different number of effective pixels.

Using the parameters (mean and sigma) determined from the fit to the reconstructed energy distribution, the relative energy resolution for 10 to 100 GeV electrons is shown in figure 6.9(c) for the three cases. An improvement in energy resolution can be seen by using a higher number of effective pixels. However, using 2500 as the effective number of pixels, the resolution is worse at high energies. Looking more closely at the ratio plot in figure 6.9(c), it is by no doubt that a

better agreement of testbeam data to the simulated data can be observed by using 2668 and 2778 as effective pixels.

The conclusion from this study is that the variation of the saturation from SiPM-to-SiPM is limiting our knowledge of the SiPM saturation correction. No large difference is found between the two 2668 and 2778 effective pixels, and choosing a lower pixel value is clearly not found to be suitable. Therefore it is decided that the nominal pixel value provides a good compromise in terms of the effective number of pixels. Therefore, 2668 was chosen as the optimal number of effective pixels for the saturation correction and a spread of 100 pixels in the SiPM saturation was included as the systematic uncertainties.

For electron validation, standard observables are compared, as discussed in the beginning of the chapter. The simulated electrons using the `QGSP_BERT_HP` physics list with beam energies from 10 to 100 GeV are validated to the testbeam data.

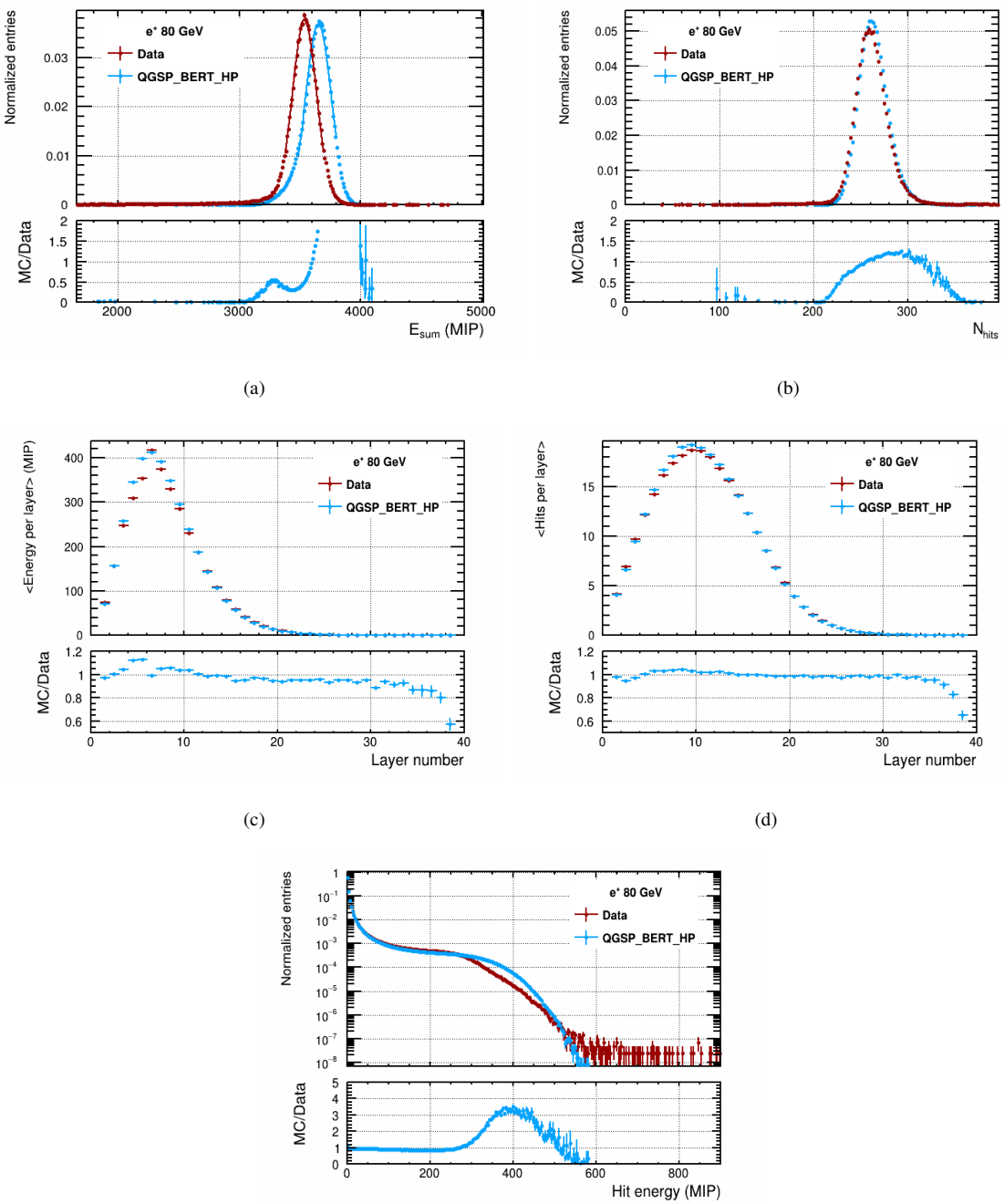
For the number of hits per event, the simulation reproduces the peak position and shape of the hit multiplicity as shown in figure 6.10(b). There is fairly reasonable agreement between testbeam data and simulated data, and the distribution shows no tail towards lower hit multiplicities.

Figure 6.10(a) compares the reconstructed energy from testbeam data to Monte Carlo simulations for 80 GeV beam momentum. The agreement between testbeam data and simulation is satisfactory, although simulation overestimates the testbeam data. In addition, the simulations predict a systematically smaller width (which is reflected in an underestimation of the energy resolution, which will be discussed in section 6.2.3).

Figure 6.10(c) shows the mean energy per hit in a given layer. This observable describes the longitudinal electron shower profiles. It shows that most of the energy is deposited within the first 20 layers, and is almost completely deposited within the first 25 layers. The shower shape and the shower maximum, which is in layer number six, is well predicted by the simulated data. However, when the testbeam data and simulation is compared layer-wise, there is a clear difference between them. Some of the layers in testbeam data are well described. However, up to 10% deviations between testbeam data and simulation is observed due to the limited accuracy of the SiPM saturation correction.

Figure 6.10(e) shows the distribution of the visible hit energy detected in single AHCAL cells for 80 GeV electron testbeam data and the predictions from simulations. The simulations predict more highly energetic signals than observed in data. The deformation in the testbeam data distribution is an artefact of the SiPM saturation, whereas simulations are not affected by this saturation, since the same SiPM response functions are used to simulate the non-linear SiPM response in the digitization procedure, and to correct it during reconstruction. Since the energy density in an electron shower is higher, saturation effects play a role at higher momenta. As discussed previously, almost no effect from the SiPM non-linearity is observed at low energies.

Thus, the validation of simulated electrons using testbeam data gives an estimate of the overall precision with the given status of calibration, simulation, and digitization procedures.



(a) (b)

(c) (d)

(e)

FIGURE 6.10: Comparision of standard reconstructed quantities for 80 GeV electron showers using testbeam data (red) and simulated data (blue). Lower window shows the ratio of simulated data to testbeam data. (a) The visible energy deposition, (b) the number of hits n_{Hits}, (c) Mean energy per layer, (d) Mean number of hits per layer and (e) Hit energy distribution for 80 GeV electron showers using testbeam data and simulation.

6.2.2 Systematic Uncertainties

The systematic uncertainty on the MIP is estimated to be $\sim$1.5%, and this value was given by the RMS of the MPV ratio distribution. For each channel, the MIP constant was smeared randomly with Gaussian smearing by $\sim$1.5%. During the digitization step, each energy deposition was calculated using smeared MIP constants, while in the reconstruction step, the calibrated MIP constant was used. This estimation of systematic uncertainty is done for the entire range of electron energies available. The systematic uncertainty for the gain was estimated in the same way as the MIP uncertainty. The stability of the gain during the testbeam was checked. The systematic uncertainty on the gain was estimated based on two LED runs from the beginning and the end of the testbeam period. It is measured to be around $\sim$1.5% and is given by the RMS of the gain ratio distribution from the two LED runs. The same procedure was followed for the digitization with the energy deposited in the AHCAL calculated using the smeared gain constants. For each channel, the gain constant was smeared with a random Gaussian function with a spread of $\sim$1.5%. However, in the reconstruction procedure, the energy was calculated using the calibrated gain constants. As discussed in the previous section, for the SiPM, a variation of the number of effective pixels by $\approx$100 pixels is observed in the saturation from SiPM to SiPM. Furthermore, in the saturation correction, a fixed number of effective pixels was used for all the channels during the reconstruction of electron data. The saturation function was applied during the digitization with a smeared number of effective pixels, using a random Gaussian function with a spread of 100 pixels. However, in the reconstruction of testbeam data and MC, a fixed number of effective pixels (2668 pixels) was used in the saturation correction function. The total systematic uncertainties used for the electron dataset were calculated by smearing the calibration constants and then by adding the uncertainty in quadrature from MIP, gain and SiPM.

6.2.3 Linearity and Resolution

Finally, the key parameter of the AHCAL calorimeter, i.e., energy linearity and energy resolution, is obtained using the full electron dataset. Gaussian fits are applied to the reconstructed energy sum distributions within a range given by the mean $\pm 1.5\sigma$. This range is applied to remove statistical fluctuations in the tails of the distributions. From these fits, the mean and sigma of the visible energy for electrons are extracted. The average visible energy is given by the mean of the Gaussian fit function, and the width of the energy sum distribution is given by the sigma of the Gaussian fit function.

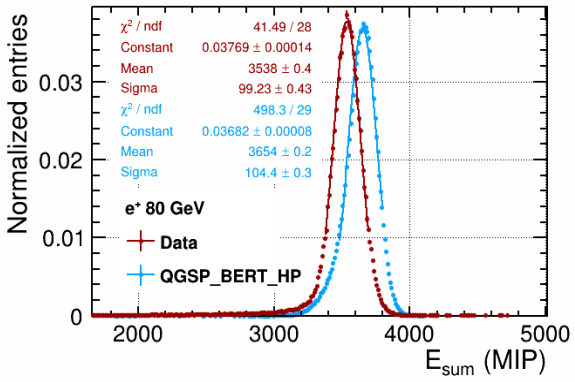
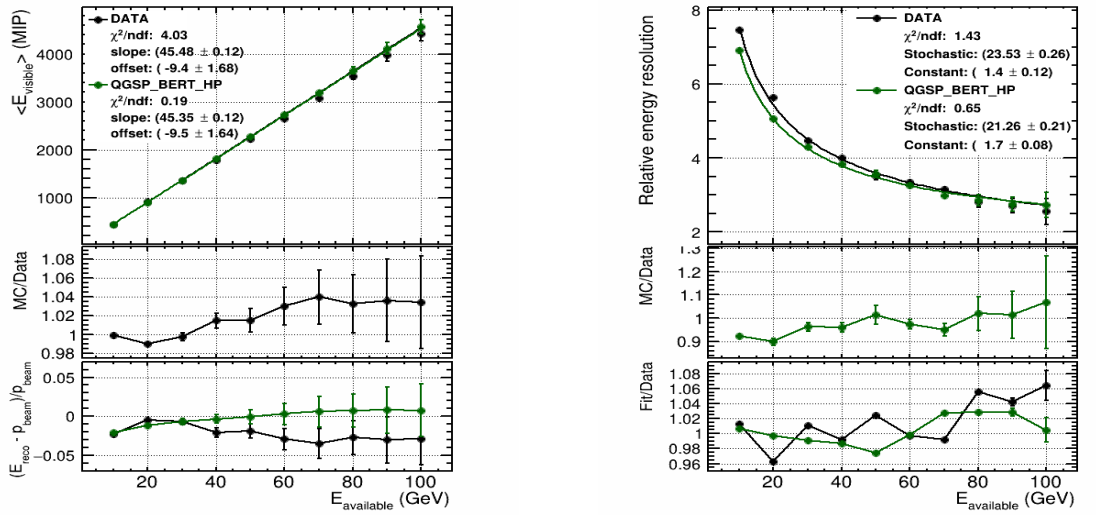


FIGURE 6.11: Reconstructed energy distributions for 80 GeV electrons using testbeam data (red) and simulation (blue). The curves show fits with a Gaussian function within a range $\mu \pm 1.5\sigma$.

Figure 6.12(a) shows the reconstructed visible energy (in units of MIP) for electrons. The uncertainty increases for higher beam energies. This is mainly related to the systematic uncertainty of the SiPM saturation function. As a linear behaviour is observed, the measured testbeam data and simulated data are fitted using Eq. 6.3. A negative offset for the linear response is obtained due to the MIP threshold of 0.5 MIP (loss of energy). This offset and scaling are needed to achieve satisfactory linearity after the energy reconstruction. The middle plot in figure 6.12(a) presents the ratio between MC and testbeam data. We can see that results for low-energy electron data up to 60 GeV agree with the simulation within uncertainties of $\sim$2%, while at higher energies the agreement between testbeam data and simulation are within uncertainties of $\sim$8%. This discrepancy between testbeam data and simulation is mainly due to the uncertainity in SiPM saturation function. The slope a and offset b are in good agreement between data and simulation and both results are consistent within uncertainties. Significant differences in the slope and offset parameters have been observed before in similar analysis of the AHCAL technological prototype [108]. The bottom plot in figure 6.12(a) presents the relative deviation of the reconstructed energy from the true energy. Note that the offset b in testbeam data and simulation is not taken into account. The reason is that for an ideal calorimeter, the offset is zero. This means that if we incident a zero GeV particle, the output should be zero. There are several known effects that introduce non-linearity and if an offset can describe it, this would be the best. For an ideal calorimeter, the offset would be zero. It tells how good the detector is, and of course, one could correct it later. So by not taking into account the offset, the non-linearity you get is the worst-case scenario. One can always do better with correction.

The electron data and simulations agree at the 5% level between 10 GeV and 100 GeV at the MIP scale. The deviation of the electron data points from linearity is within 5%, while for simulation, the deviation is smaller within only 1%. The higher deviation for data might be a sign that there are saturation effects, which are still not described well by the SiPM response function.

The energy resolution σ_{E}/E, which is the width of the visible energy divided by the visible energy, is obtained for electrons at different beam energies. Generally, the relative resolution improves from low to high beam energies. Figure 6.12(b) shows the energy resolution, and the solid lines represent the results of performing a fit using Eq. 1.13 to the electron data and simulation. As the noise contribution in the prototype is considered to be practically negligible, the noise term b is fixed to be zero. Overall, it is observed that the electron simulation tends to underestimate the width of the electron response at low energies, for example below 80 GeV. The stochastic term obtained from simulation is smaller than compared to the data. The deviation as seen in the middle plot in figure 6.12(b) is less than 10%. The lower plot in figure 6.12(b) shows the ratio of the resulting fit to the corresponding data point (testbeam data and simulation). An agreement of $\sim$6% for all energies is observed.



(a) (b)

FIGURE 6.12: (a) The upper window is the mean reconstructed visible electron energy $\langle \mathrm{E}_{\mathrm{visible}} \rangle$ (in MIP) as a function of the beam momentum in the range 10–100 GeV from testbeam data (black circles) and simulations (using the `QGSP_BERT_HP` physics list in green circles). The middle window is the ratio between the mean visible energy from simulation and data. The lower window is the difference between $\langle \mathrm{E}_{\mathrm{visible}} \rangle$ and the true beam momentum $\mathrm{p}_{\mathrm{beam}}$. (b) The upper window is the AHCAL electron relative energy resolution $\sigma_{\mathrm{E}}/\langle \mathrm{E}_{\mathrm{visible}} \rangle$ as a function of the beam momentum in the energy range from 10–100 GeV. The middle window is the ratio between the mean visible energy predicted by simulations and testbeam data. The lower window is the ratio between fit values and the mean reconstructed energy as a function of beam momentum for testbeam data and simulation. The error bars take into account statistical and systematic uncertainties of the response predicted by simulations.

6.3 Validation with Pions

In this section, the validation of GEANT4 physics lists (QGSP_BERT_HP and FTFP_BERT_HP) for beam energies between 40 GeV and 80 GeV pion data is presented. Before going to the shower shape studies, it is important to confirm the agreement between testbeam data and simulations using pion showers. This is necessary since the physics of hadron showers is less understood compared to electron showers and the simulation of hadron showers still undergoes significant improvements. A similar comparison with the global observables between testbeam data and simulation is done for contained pion showers as it was for electron showers.

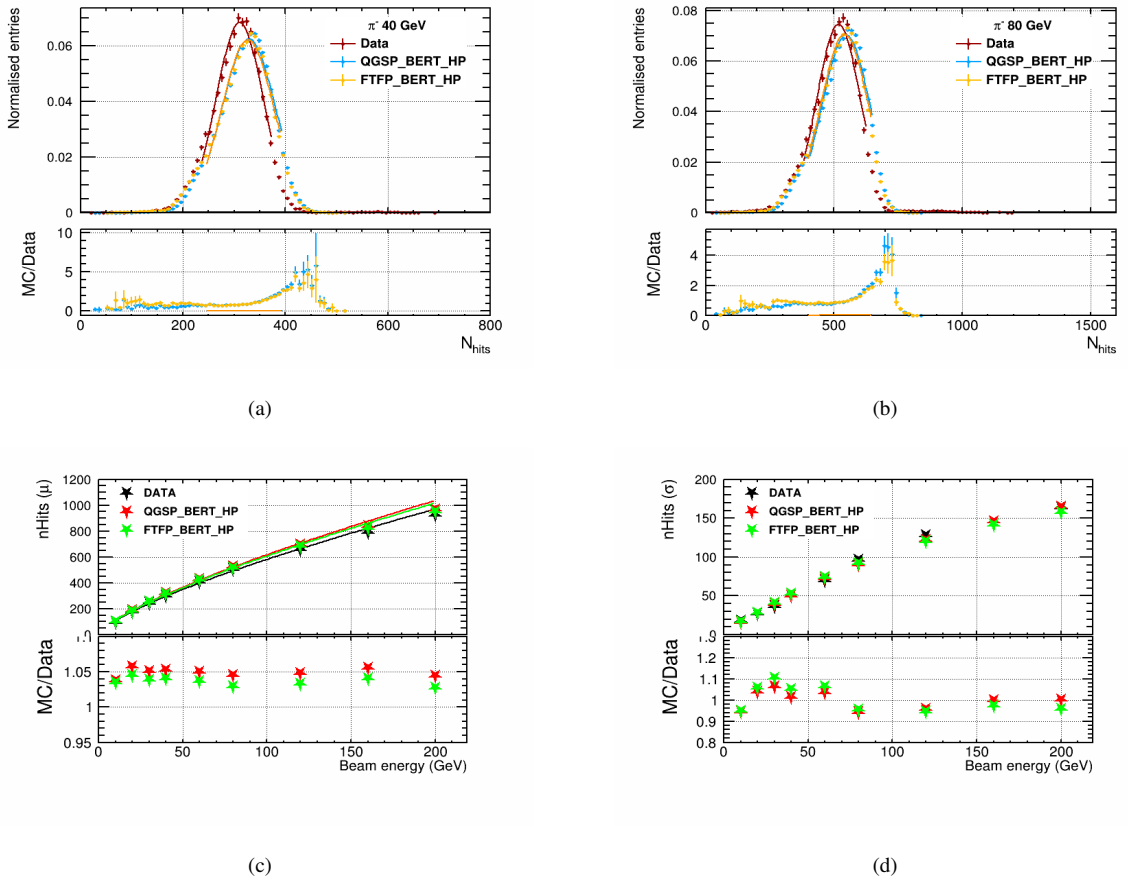


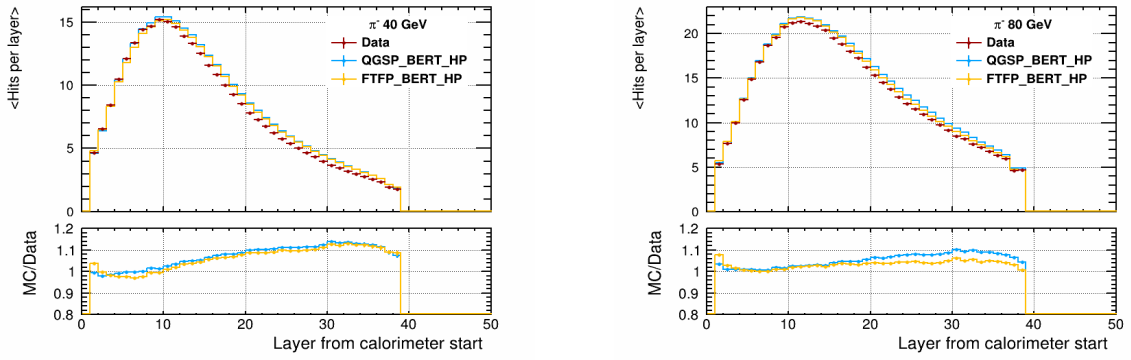
FIGURE 6.13: Comparision of the number of hits distribution for pion showers with beam energy at (a) 40 GeV and (b) 80 GeV using testbeam data and physics lists (QGSP_BERT_HP and FTFP_BERT_HP). The curves represent the Gaussian fit within the range of $\mu \pm 1.5\sigma$. The ratio plots are shown at the lower window. (c) Mean μ extracted from the Gaussian fits within the range of $\mu \pm 1.5\sigma$ to the number of hit distribution as a function of beam energy in the range 10–200 GeV. The solid lines represent the power law function. (d). Width σ as a function of beam energy in the range 10–200 GeV extracted from the Gaussian fits to the number of hit distribution.

Figures 6.13(a) and 6.13(b) show the number of hits distributions for 40 GeV and 80 GeV pions respectively, with the ratio between the distributions shown at the bottom of the plots.

The number of hits distributions are obtained from testbeam data and two physics lists. The solid curves drawn on the distributions are fits using a Gaussian fit function within the range of $\mu \pm 1.5\sigma$ applied on the testbeam data and simulations. The fit provides mean and sigma values, which are obtained for all the available pion beam energies.

The mean and sigma values obtained from the fit are plotted, and the result is shown in figure 6.13(c) and figure 6.13(d). The response to pions was fit empirically with a power law, $a \cdot p^b$, where p is the beam momentum. a and b are free parameters. A linear response would correspond to $b = 1$. However, the power-law is not described well at high energies due to several traversing particles contributing the same information to the reconstructed energy as a single traversing particle. Moreover, we see from the mean response that the simulations and data agree within 8%. On the other hand, the width obtained from the Gaussian fit to the number of hit distribution agrees at 10%.

A comparison of the mean number of hits per layer between testbeam data and simulations for a beam energies of 40 and 80 GeV is shown in figure 6.14(a) and 6.14(b) respectively. The shape and maximum position are well modelled by both the physics lists, with an agreement between testbeam data and simulations within 10%. However, the simulations predict slightly more hits in later layers, as shown in the ratio plot at the bottom of figure 6.14. On the other hand, the physics list `FTFP_BERT_HP` predicts slightly more hits in the first layer. At the same time, it also underestimates the maximum position and the number of hits in the later layers.

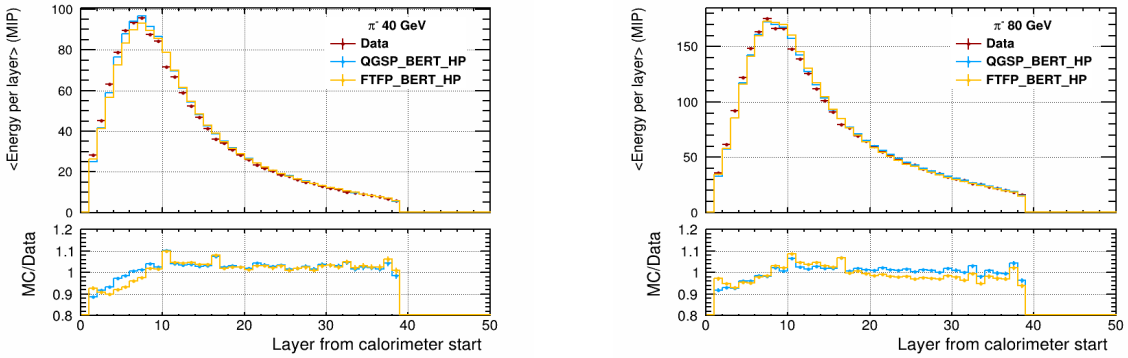


(a) (b)

FIGURE 6.14: Comparision of mean number of hits per layer for (a) 40 GeV and (b) 80 GeV π^- from testbeam data (red) and physics lists (`QGSP_BERT_HP` in blue and `FTFP_BERT_HP` in orange). Lower window shows the ratio of simulated data to testbeam data.

The comparison of mean energy per layer between testbeam data and simulations for 40 and 80 GeV is shown in figure 6.15(a) and 6.15(b) respectively. Similar observations can be made, as in the case of the mean number of hit distributions. Both the physics lists predict the shape and maximum position of the distribution very well. The shower maximum lies between the sixth and eighth layer, depending on the shower energy. The longitudinal pion shower profiles show in general, a good agreement with the simulated data around the shower maximum within

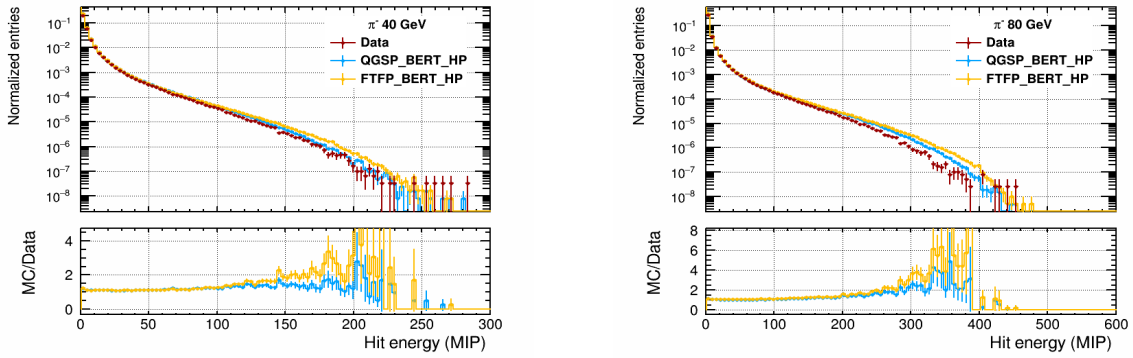
5-10%. Especially in the first seven layers, the energy deposit seems to be underestimated by the simulated data. In contrast, the energy deposits in the later layers seem to be overestimated by the simulations.



(a) (b)

FIGURE 6.15: Comparisons of the longitudinal energy profiles of (a) 40 GeV and (b) 80 GeV π^- from testbeam data (red) and physics lists (`QGSP_BERT_HP` in blue and `FTFP_BERT_HP` in orange). Lower window shows the ratio of simulated data to testbeam data.

The effect of an imperfect saturation correction affects pion energy measurements much less than energy measurements from electron showers. This is because the energy density in pion cascades is, on average, much lower than the density in electron cascades of the same energy. The comparison of the hit energy distribution between testbeam data and simulations for beam energies of 40 and 80 GeV is shown in figure 6.16(a) and 6.16(b) respectively. Both the physics lists describe the shape reasonably well, even at the tails of the distribution, but with a slight overestimation caused by the SiPM correction function. The `FTFP_BERT_HP` physics list predicts more energy hits compared to `QGSP_BERT_HP` and the testbeam data is clearly seen at the tails of the distribution.

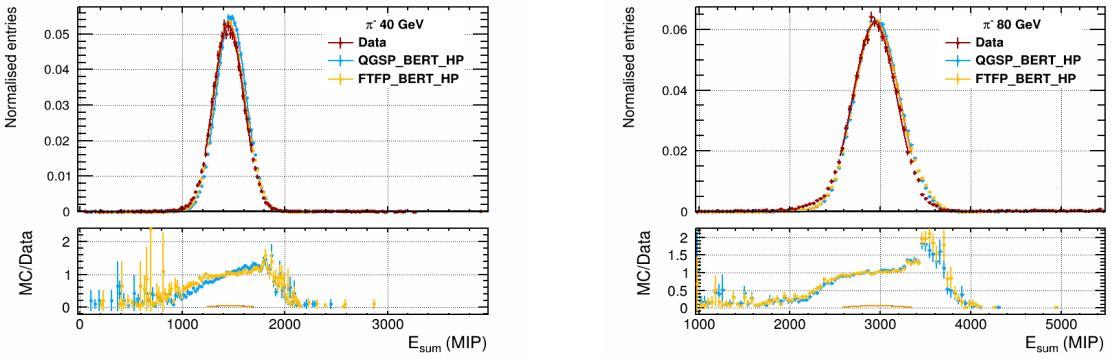


(a) (b)

FIGURE 6.16: Comparison between testbeam data and simulation with the hit energy distribution (a) 40 and (b) 80 GeV π^- data (red) and physics lists (`QGSP_BERT_HP` in blue and `FTFP_BERT_HP` in orange). Lower window shows the ratio of simulated data to testbeam data.

6.3.1 Linearity

The determination of the reconstructed mean position and width from non-Gaussian distributions is not trivial. Asymmetries can occur and originate from different effects, e.g. mainly from the non-compensating device, saturation or energy leakage. The variable used to obtain the energy linearity and resolution of the calorimeter is the energy sum with a similar approach used for electron showers.



(a) (b)

FIGURE 6.17: Comparision of energy sum distributions for (a) 40 GeV and (b) 80 GeV π^- from testbeam data (red) and physics lists (`QGSP_BERT_HP` in blue and `FTFP_BERT_HP` in orange). The curves represents the Gaussian fit within the range of mean $\mu \pm 1.5\sigma$. Lower window shows the ratio of simulated data to testbeam data.

Figure 6.17(a) and 6.17(b), show the energy sum distributions for beam energies of 40 GeV and 80 GeV pions compared between testbeam data and simulated data. Both physics lists describe the shapes reasonably well. The lower window shows the ratio of simulations to testbeam data and is within an agreement of roughly 20%.

Gaussian fits are applied to the energy distributions within a range given by $\mu \pm 1.5\sigma$. With the mean and width obtained from the fit, the linearity and energy resolution as a function of pion beam energy is obtained for energies between 10 and 200 GeV from testbeam data and Monte Carlo simulations using different `GEANT4` physics lists.

Figure 6.18 presents the linearity with a comparison between two physics lists. The uncertainties are the quadratic sum of statistical and systematic errors. As a linear behaviour is observed, the measured values are fitted using a linear function as given in Eq. 6.3. The fit values from pion linearity is summarized in table 6.3. The value a deviates less than 3% between data and simulation. Furthermore, the results are consistent within uncertainties. The offset b on the other hand is a factor 0.8 smaller for data than for simulations, but the uncertainties are a magnitude larger compared to simulations.

The ratio between simulation and testbeam data is shown in the middle part of the figure 6.18. We see that the reconstructed energy of the pion simulations agrees well with testbeam data within 2%. The largest uncertainty is observed at the highest energy, which is mainly due to the

saturation effect. The reconstructed energy E_{reco} in units of GeV can be inferred from Eq. 6.3. The lower plot in figure 6.18 shows the residual $(E_{reco}[GeV] - p_{beam})/p_{beam}$ of the linear fits for testbeam data and simulations. The deviations of the pion data and simulations points from linearity is between 2 to 4%. The higher deviation is found for `FTFP_BERT_HP`. This might be a sign that there are saturation effects, which are still not described well by the response model of the SiPM-tile system. Especially for 200 GeV, the deviation from the linear fit is significant, and the reconstructed energy of testbeam data is higher than simulations. For all beam energies, simulations agree with data within the uncertainties. They are smaller than for electrons showers, because the SiPM saturation has a smaller effect. The residual is mostly negative, as the offset b is not considered, as performed in the case of electron showers. Interestingly, the largest deviation is also found at the lowest beam energy of 10 GeV.

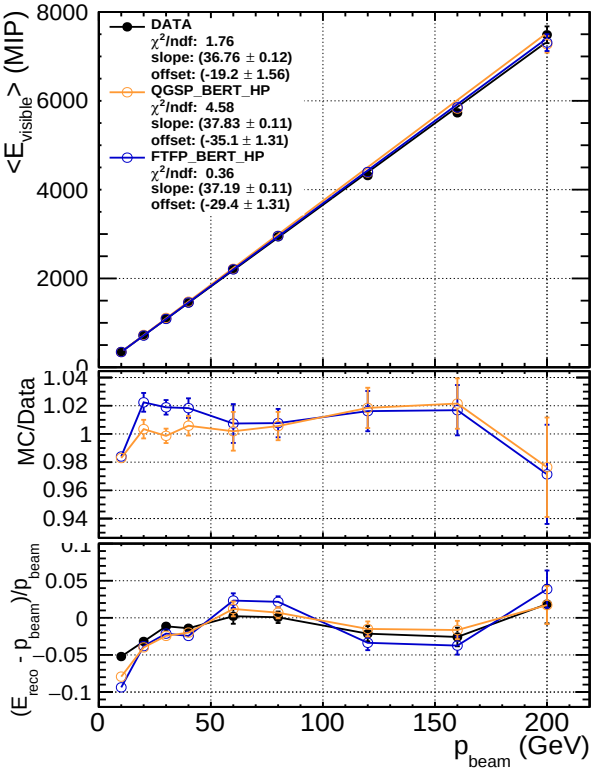


FIGURE 6.18: The upper window is the mean reconstructed visible pion energy $\langle E_{visible} \rangle$ (in MIP) as a function of the beam momentum in the range 10–200 GeV from testbeam data (black circles) and simulations (using the `QGSP_BERT_HP` and `FTFP_BERT_HP` physics list in open circles). The data points are described using a linear fit function. The middle window is the ratio between the mean visible energy from simulations and data. The error bars take into account statistical and systematic uncertainties of the response predicted by simulations. The lower window is the difference between the mean reconstructed energy $\langle E_{visible} \rangle$ and the true beam momentum p_{beam} as a function of beam energy using a scaling factor a (according to Eq. 6.3) to convert visible energy to reconstructed energy for data and simulations.

6.3.2 Energy Resolution

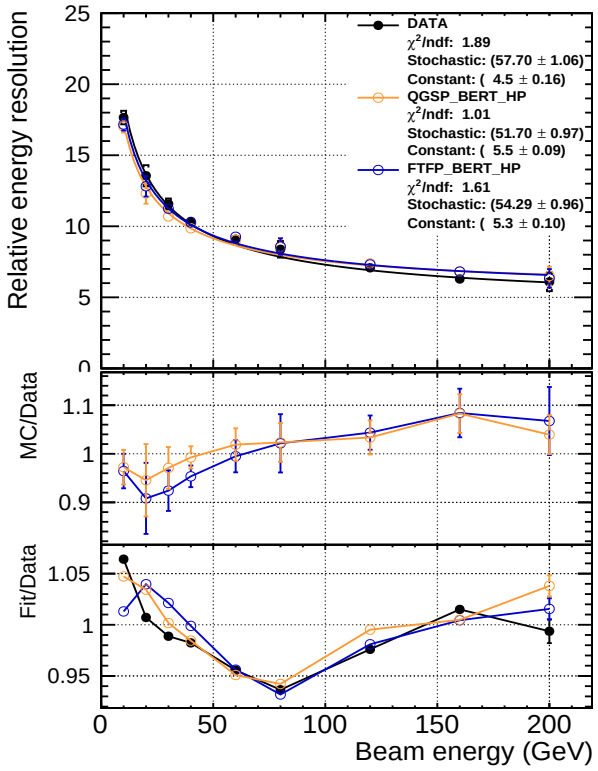


FIGURE 6.19: The upper window is the AHCAL pion relative energy resolution $\sigma_{\mathrm{E}}/\langle \mathrm{E}_{\mathrm{visible}} \rangle$ (σ_E and $\langle \mathrm{E}_{\mathrm{visible}} \rangle$ are the values from Gaussian fits to the distributions in the range of ± 1.5 standard deviations) as a function of the beam momentum 10–200 GeV from testbeam data (black circles) and simulations (using the `QGSP_BERT_HP` and `FTFP_BERT_HP` physics list in open circles). The data points are described using the Eq. 1.13. The middle window is the ratio between the mean visible energy predicted by simulations and data. The error bars take into account statistical and systematic uncertainties of the response predicted by simulations. The lower window is the ratio between fit values and the mean reconstructed energy as a function of beam energy for data and simulations.

The single pion energy resolution $\frac{\sigma_E}{E}$ is plotted as a function of the beam energy in figure 6.19. Energy dependence of the pion energy resolution is parametrized in the energy range between 10 and 200 GeV using the quadratic sum as given in Eq. 1.13 and is represented by the solid curves to the measured values. The noise parameter b is fixed to zero. The energy resolution fit results with a stochastic term of approximately 55 to 60%, which agrees reasonably well with the simulated data. The fit values from pion resolution is summarized in table 6.3. The stochastic term found for both physics lists underestimates the value found in testbeam data. However, as shown in the middle plot in figure 6.19, the obtained energy resolution is well reproduced by the

simulation within the uncertainties about roughly 10%. The agreement with the fit and energy points as seen in the bottom plot in figure 6.19 is fairly reasonable at the level of 7%.

The stochastic term for pion data from the previous physics prototype with 5 mm scintillator is 57.6% [109]. This is in good agreement with the current technological prototype with a stochastic term for contained pion showers is approximately 57%. As expected, the energy resolution is worse for pion showers compared to electrons. Moreover, the estimated h/e signal ratio is found to be ≈ 0.81. This is consistent with the expectations that the technological prototype is a non-compensating calorimeter.

6.4 Conclusion

In this chapter, the simulation for muon, electron and pion has been validated along with several standard reconstructed variables. In the first part, simulated muons were validated using the testbeam data. The validation of muons proves that the detector is well calibrated for energy response. The electron analysis concludes that the precision of the detector simulation is limited by the de-saturation of the high energetic cells, and it plays a major role in systematic studies. Since electron showers are small in size, only a few cells contribute significantly to the detector signal, making it necessary to describe these few cells with the utmost precision. The systematic of ± 100 pixels is estimated as an uncertainty in the SiPM saturation correction. However, besides these constraints, the disagreement between recorded and simulated electron showers does not exceed a few per cent. Within uncertainties, the linearity of the AHCAL in the covered energy range is verified. Furthermore, the analysis of electron data recorded between 10 GeV and 100 GeV beam momentum yields a conversion factor. The resolution extracted from these datasets has a stochastic and a constant term which is summarized in table 6.3. Thus, it confirms that the detector calibration and simulation are well enough under control to analyze pion datasets and validate pion simulations.

After validating the electrons, the pion showers are validated. Pion showers show a good performance in terms of agreement with the simulations. An overview of linearity and energy resolution fit parameters for pion datasets recorded between 10 GeV and 200 GeV is summarized in table 6.3.

	Particle	Parameter	Data	QGSP_BERT_HP	FTFP_BERT_HP
Linearity	e^+	slope (MIP/GeV)	45.48 ± 0.12	45.35 ± 0.12	✗
		offset	-9.4 ± 1.68	-9.5 ± 1.64	✗
	π^-	slope (MIP/GeV)	36.76 ± 0.12	37.83 ± 0.11	37.19 ± 0.11
		offset	-19.2 ± 1.56	-35.10 ± 1.31	-29.4 ± 1.31
Resolution	e^+	stochastic ($\%/\sqrt{[GeV]}$)	23.53 ± 0.26	21.26 ± 0.21	✗
		constant (%)	1.4 ± 0.12	1.7 ± 0.08	✗
	π^-	stochastic ($\%/\sqrt{[GeV]}$)	57.70 ± 1.06	51.70 ± 0.97	54.29 ± 0.96
		constant (%)	4.5 ± 0.16	5.5 ± 0.09	5.3 ± 0.10

TABLE 6.3: Fit parameters of linearity and energy resolution for electrons and pions compared between data and Monte Carlo simulation using QGSP_BERT_HP and FTFP_BERT_HP physics list.

Chapter 7

Shower Shapes

"There are two possible outcomes: if the result confirms the hypothesis, then you've made a measurement. If the result is contrary to the hypothesis, then you've made a discovery."

Enrico Fermi

THE STUDY OF SHOWER SHAPE characteristics of electromagnetic and hadronic showers and validating the shower models used for MC simulation is one of the goals of the CALICE-AHCAL testbeam. The extremely high granularity of the AHCAL allows detailed studies of hadron showers at an unprecedented level of precision. A complete understanding of the shower shapes is vital for several reasons. Firstly, one can use the distinctive shower shapes of electromagnetically and hadronically interacting particles to recognize and distinguish between the two types of particles. Secondly, the knowledge of the profiles enables one to understand the energy overlap of showers initiated by particles that are closely spaced. Shower shapes provide a crucial factor for PFA efficiency to obtain the possibility to separate showers produced by charged and neutral hadrons, as the achievable jet energy resolution depends on the capability to separate nearby showers. Shower profiles can also improve the energy resolution of calorimeters by exploiting the fact that the shapes depend upon particle energy. Information about the shower shapes can be used as a reliable estimate to define the shower containment in the longitudinal direction. It is crucial that a fast simulation correctly models the radial and longitudinal shower development in the calorimeter to gauge effects on resolution and signal that depend on the specific calorimeter configuration, energy leakage out of the detector, variations in the thickness of absorber or active material. The shower parametrization used in fast simulations helps to determine such effects and, therefore, must be chosen with care.

For physics analysis and feasibility studies, a vast number of MC programs exist, which simulate in full detail the development of electromagnetic and hadronic showers. However, they are of limited use to calorimetry problems because of the following reasons:

1. They are time-consuming: with individual particle tracking, the computing time needed for

such simulation increases approximately linearly with the energy absorbed in the detector and might quickly become excessive.

2. They are not easy to adapt to a complicated calorimeter geometry.

So using parameterizations for electromagnetic and hadronic showers we can speed up the simulations considerably without sacrificing precision.

A detailed analysis of the electron and pion shower shapes observed in the CALICE-AHCAL test beam data will be discussed in this chapter. The main interest will be on pion showers, but the electromagnetic showers are also investigated briefly before looking into pion showers. In order to understand pion showers, we first need to see how electromagnetic showers look like, which is the first part of this chapter in section 7.1. In order to understand pion showers, we look at electromagnetic showers first in section 7.1 because the shower physics is much simpler, so it's a way to cross check the implementation of the AHCAL in the simulation. In addition, it makes sure that we have a well-working electromagnetic shower parametrisation that we can use for the EM sub-shower in hadron showers. Furthermore, the pion showers are studied in detail in section 7.2. The results are compared with the previous physics paper as presented in [110]. The idea is to reproduce the results with the large technological prototype with the high spatial granularity throughout the detector, slightly finer longitudinal sampling and practically no noise. This chapter presents an improved radial parametrization. Furthermore, the study is extended to fit the longitudinal and radial shower profiles simultaneously. Moreover, this chapter presents a three-dimensional shower model to exploit the hadronic shower shapes. In addition, before the summary of this chapter, the hadronic shower shapes are analyzed based on the history of true Monte Carlo particles. Throughout this chapter, an identical analysis is performed on simulated data samples, the comparisons of the results from simulated data and from testbeam data will also be presented in this chapter. Fitting method and the details is introduced in appendix E.

Measured Quantities

The study of shower shapes presented in this chapter deals with the following two quantities: The longitudinal shower shape $-$ for hadronic showers, which is defined as the average energy deposited $\Delta\mathrm{E(z)}$ in each layer relative to the shower start position. Where the energy $\Delta\mathrm{E}$ is given in units of MIP, and the longitudinal depth z is measured in units of nuclear interaction length λ_I. The first bin corresponds to the physical layer in which the shower start is identified and the profile is binned with a size of 0.11 λ_I. The longitudinal profile for electromagnetic showers is measured from calorimeter front face with z in units of radiation length X_0, and is binned with a size of 1.05 X_0. The value of interaction and radiation lengths are calculated for the AHCAL prototype based on the material properties from table 2.2.

The radial shower shape $-$ is the distribution of the mean energy density, $\frac{\Delta\mathrm{E}}{\Delta\mathrm{S}}(\mathrm{r})$, as a function of radial distance r from the shower axis. The visible energy $\Delta\mathrm{E}$ in units of MIP is measured in the ring of width $\Delta\mathrm{r}$ and of area $\Delta\mathrm{S} = 2\pi\mathrm{r}\Delta\mathrm{r}$, assuming the integration along the longitudinal direction.

Apart from particle energy and type, the impact point in the calorimeter cell is the most relevant parameter for radial shower profiles. To study the radial shower development of particle showers, the highly granular segmentation of the AHCAL layers are exploited. For each AHCAL cell i, the radial distance to the incoming particle track is determined as:

$$r_i = \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2}$$

where (x_i, y_i) is the coordinate of the cell centre and (x_0, y_0) is the position of the energy weighted shower centre given by,

$$x_0 = \frac{\sum_i E_i \cdot x_i}{\sum_i E_i} \text{ and } y_0 = \frac{\sum_i E_i \cdot y_i}{\sum_i E_i}$$

The radial profiles are binned with a size of 10 mm based on the concept of cell virtualization.

Cell Virtualization

While performing shower shape studies, the detector configuration is essential. The detector is equipped with square cells, but the shower evolution is best described in the cylindrical coordinate system. When a particular radial width is chosen, some cells will intersect more than one bin. So to analyze radial profiles, finer width is chosen. In this method, the energy deposited in the physical cell is equally distributed over the virtual cells covering its area and use the cell-center to assign the energy to a ring. Thus, it avoids having entire energy deposition in a single space point. This procedure is called cell virtualization as shown in figure 7.1, and it aims to homogenize the detector by choosing finer cells of 10×10 mm^2 size.

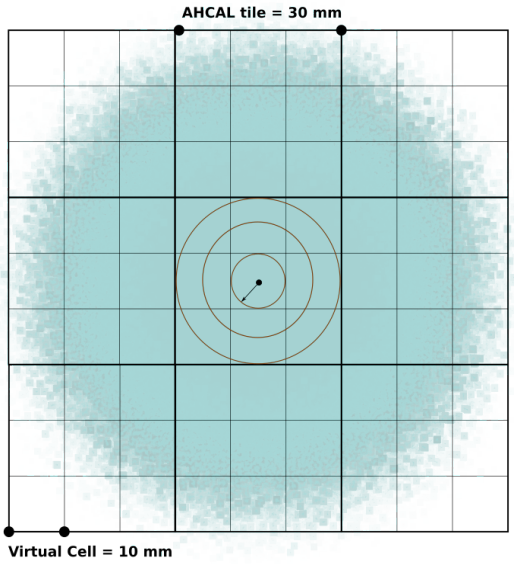


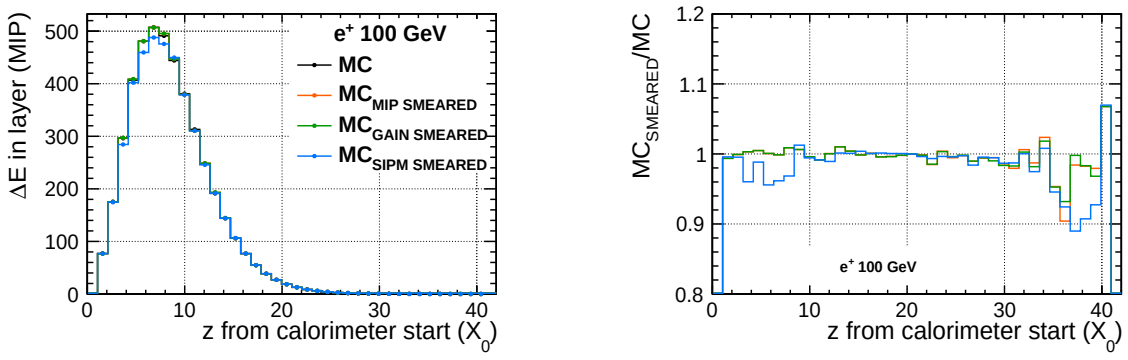
FIGURE 7.1: Schematic illustration of virtual cells. The rings of width 10 mm are shown. The AHCAL cell is represented in thick outlines and sub-virtual cells in finer lines.

7.1 Electromagnetic Shower Shapes

In the following section, an analytical parametrization that describes the longitudinal and radial shower development of an average electromagnetic shower is discussed.

7.1.1 Longitudinal Shower Profiles

The longitudinal shower profiles for electrons are measured from the start of the calorimeter as electrons start to shower immediately when they enter the calorimeter and measured in units of radiation lengths X_0. The systematic uncertainties that affect the longitudinal shower profiles for electrons are estimated from the calibration constants from MIP, Gain and SiPM saturation with the same approach as described in section 6.2. The uncertainties are estimated for simulations, and the resulting longitudinal profiles with different systematic effects for simulated 100 GeV electron showers is shown in figure 7.2(a).



(a) (b)

FIGURE 7.2: (a) Longitudinal shower profile from a 100 GeV electron, overlaid with the shower profiles from smeared calibration constants (MIP, Gain and SiPM). (b) Ratio of longitudinal shower profile for 100 GeV electron from the smeared calibration samples to the calibrated samples.

As previously discussed, the dominant source of uncertainty is due to SiPM saturation behaviour and the effect being prominent at high energies. The effect of various uncertainties can be seen in figure 7.2(b), which shows the ratio of smeared constants to the calibrated constants for a 100 GeV simulated electron shower. The region of large discrepancy is seen near the maximum shower position. Nevertheless, significant differences are also found around layer 35 and above. The statistical uncertainties are large because the energy depositions are rare.

Figure 7.3 shows the longitudinal shower shapes from electron showers for beam energy between 10 to 100 GeV and are compared between testbeam data and simulations using `QGSP_BERT_HP` physics list. The longitudinal shower development for the different energies predicts the rise and fall behaviour as expected. Moreover, the maximum position is well described between the testbeam data and simulation. The showers penetrate deeper into the material and the position where the maximum energy deposition (shower maximum) occurs also moves deeper into the detector, which is well reproduced by the simulation.

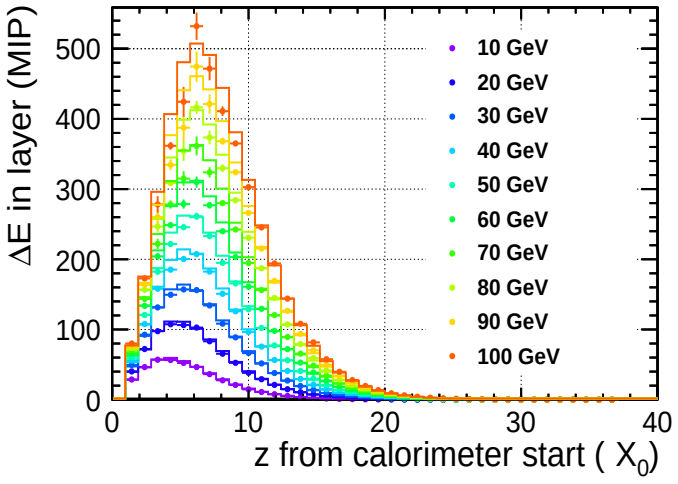


FIGURE 7.3: Longitudinal shower profiles from the calorimeter start obtained for all available electron energies with beam energy between 10 to 100 GeV. The shower profiles are compared between testbeam data (points) and simulations using `QGSP_BERT_HP` physics list (histograms).

The quality of agreement describing the longitudinal shower development is shown in figure 7.4 as the ratio of longitudinal shower extracted from the `QGSP_BERT_HP` physics list to the testbeam data at all studied energies. The simulation predicts slightly lower energy deposition in the first layer of the calorimeter. However, it significantly underestimates the amount of energy in the last layers showing discrepancies of more than 20% at all energies. Moreover, the simulations predict lower energy deposits at the shower maximum position compared to testbeam data by up to 20% at 100 GeV.

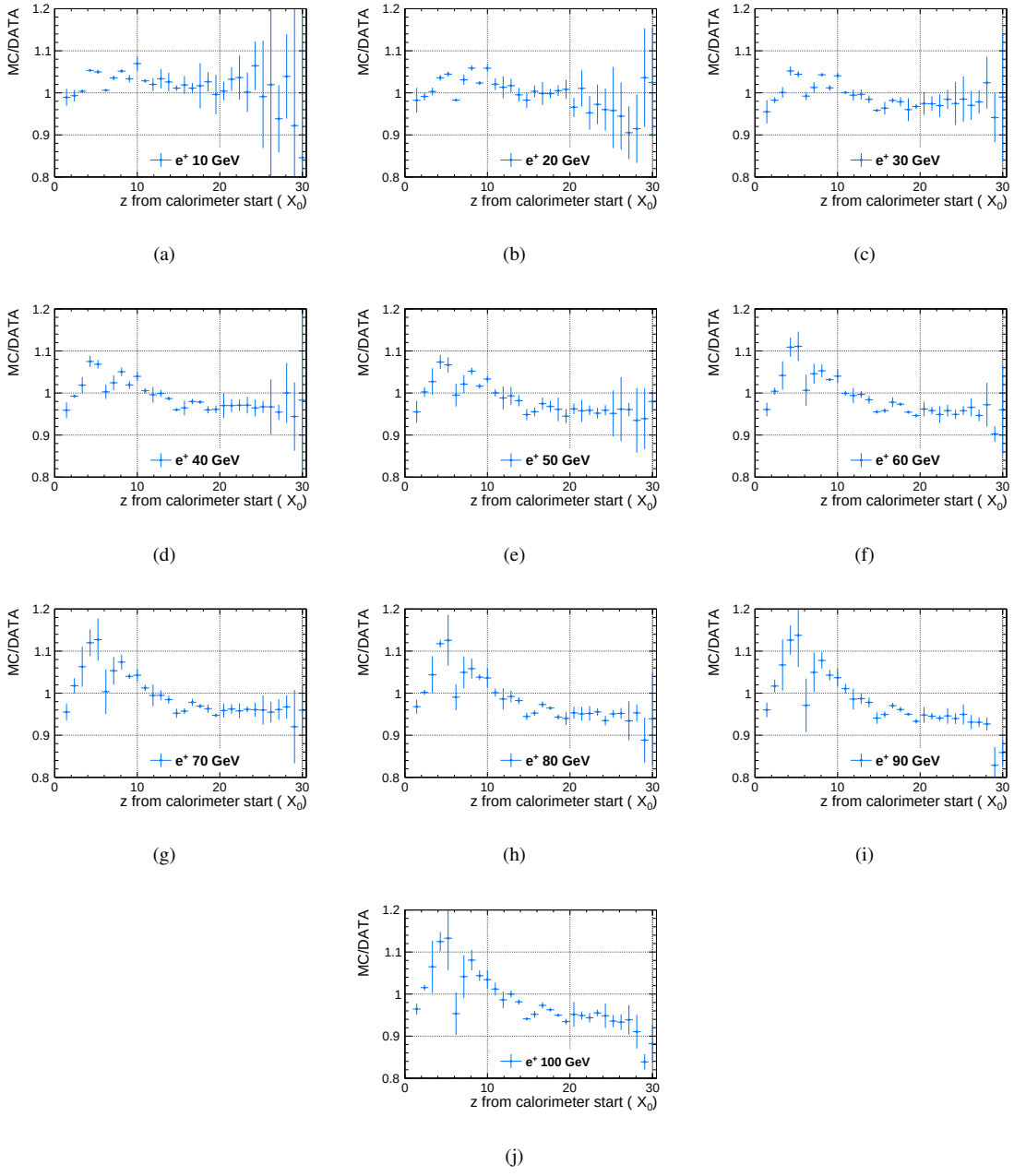


FIGURE 7.4: Ratio of longitudinal shower profiles induced by electron showers from simulated data using `QGSP_BERT_HP` physics list to those from testbeam data. The errors bars for data and simulations indicate the statistical and systematic uncertainties added in quadrature.

A general description of the averaged longitudinal profile for electromagnetic showers can be achieved with the help of a Gamma distribution, as given by:

$$\Delta E(z) = E \cdot \left\{ \frac{1}{\Gamma(\alpha)} \cdot \left(\frac{Z[X_0]}{\beta} \right)^{\alpha-1} \cdot \frac{e^{\frac{-Z[X_0]}{\beta}}}{\beta} \right\}, \tag{7.1}$$

where E is the total energy in units of MIP, and corresponds to the integral of the function up to infinity. α and β correspond to the shape and slope parameter respectively.

Furthermore, the longitudinal profiles are obtained for different energies. As discussed above, a fit is performed on these longitudinal shower profiles using a single Gamma function. An example of the fit and its parameters is shown in figure 7.5 for the longitudinal shower profile of 10 GeV and 100 GeV electron showers using testbeam data and simulations. The model describes the shape of the electromagnetic shower reasonably well for the studied energy range from 10 GeV to 100 GeV but certainly with a large χ^2/NDF as shown in figure 7.6(a).

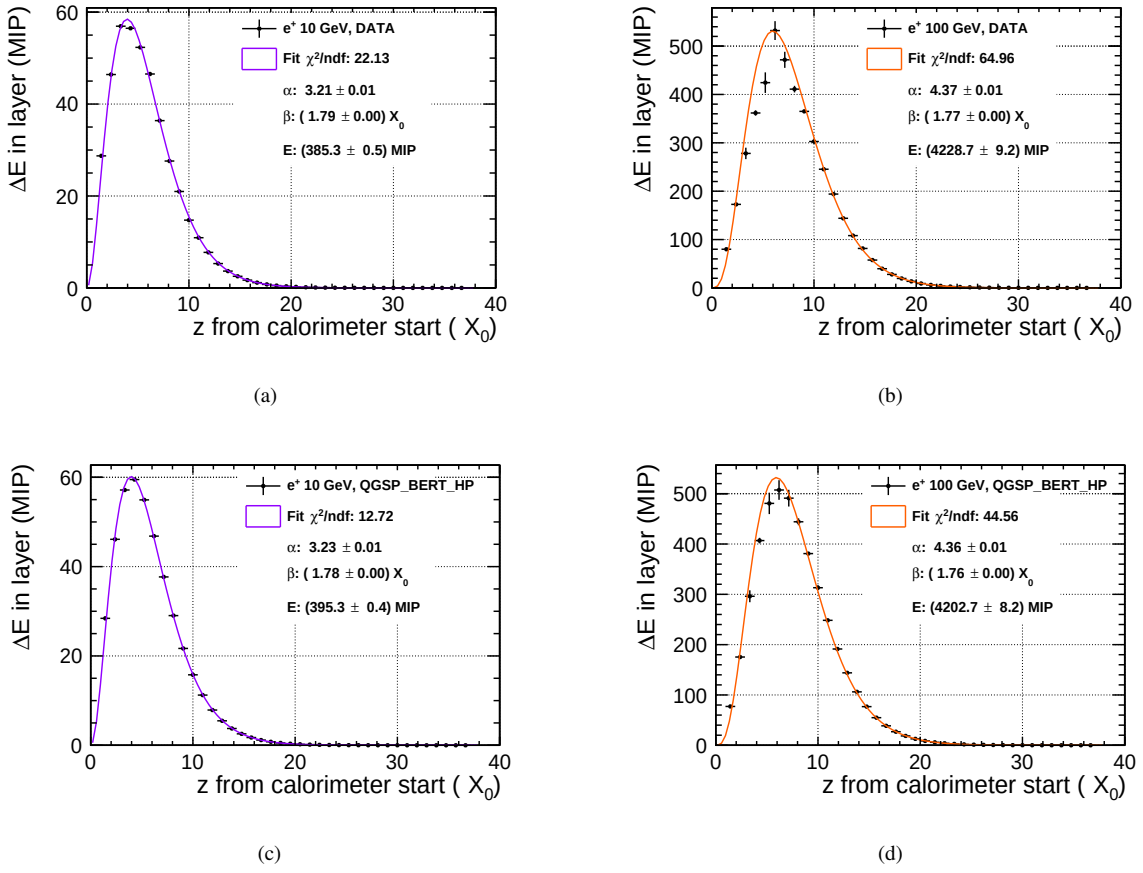
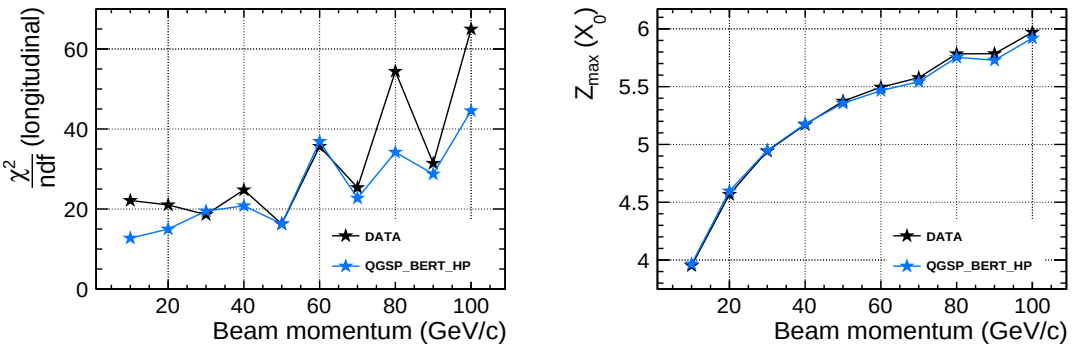


FIGURE 7.5: Fits performed to longitudinal shower profile shown by solid line using Eq. 7.1 (a) for 10 GeV electron shower and (b) for 100 GeV electron shower from testbeam data, and (c) for 10 GeV electron shower and (d) for 100 GeV electron shower from simulated data using `QGSP_BERT_HP` physics lists.

Based on the result of fit parameters that are obtained for different energies, the maximum of the shower is obtained using:

$$Z_{\mathrm{max}} = (\alpha - 1) \times \beta, \tag{7.2}$$

Figure 7.6(b) compares the maximum shower position between testbeam data and MC simulation. Both testbeam data and simulation follows the theoretical expectation of a logarithmic increase in energy.



(a) (b)

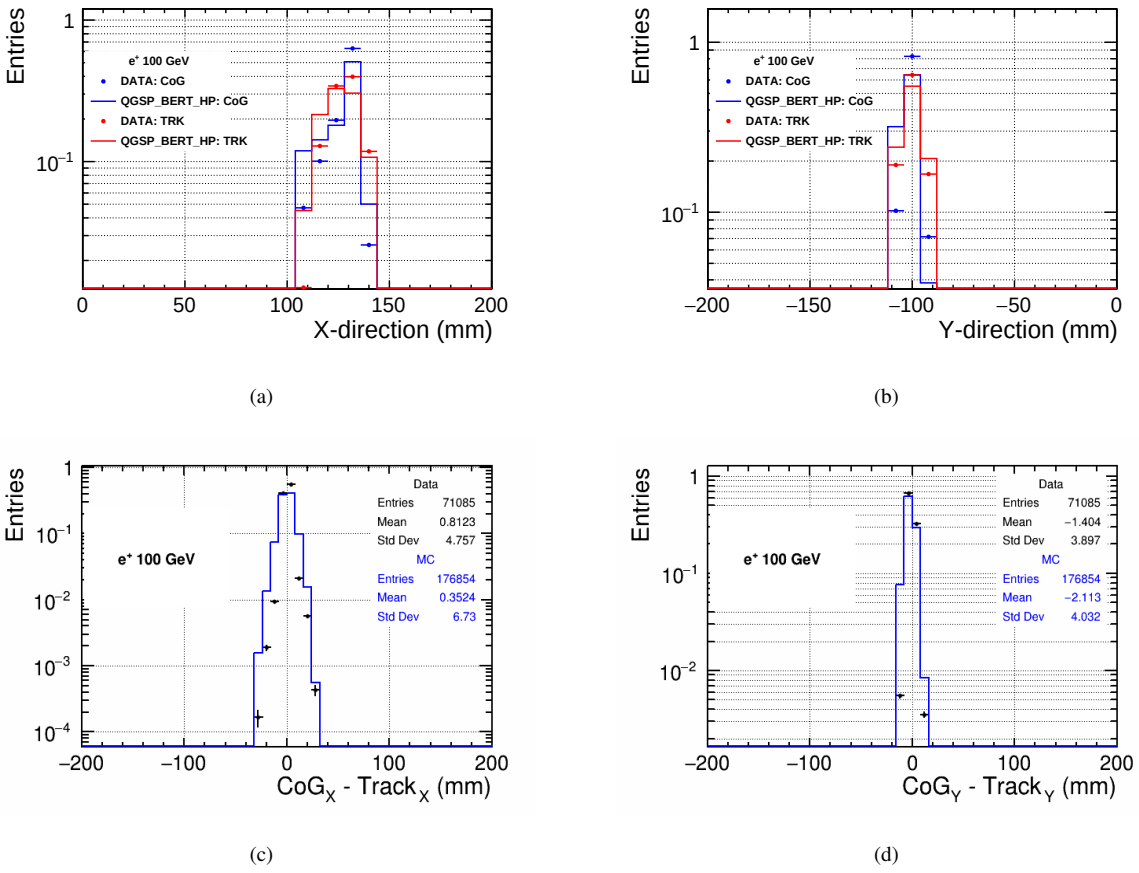
FIGURE 7.6: (a) χ^2/NDF obtained from longitudinal fits for all available electron energies from 10 GeV to 100 GeV both for testbeam data (black) and QGSP_BERT_HP physics list (blue). (b) Shower maximum position for testbeam data and QGSP_BERT_HP physics list obtained from the resultant of the fit to longitudinal shower profiles using Eq. 7.2.

7.1.2 Radial Shower Profiles

Reconstruction of shower axis

The radial shower profiles are calculated based on two approaches of shower axis reconstruction. The first method uses the centre of gravity of an event, which is currently chosen as the reference axis. The second method uses the information from the incoming tracks from the delay wire chambers.

The distributions of reconstructed track and event centre of gravity in X and Y-direction is shown in figure 7.7 as a comparison between the testbeam data and simulation. The agreement between testbeam data and simulation for the two methods is nearly consistent. Figure 7.7(c) shows the shower axis uncertainty of roughly less than a millimetre in the X-direction. Figure 7.7(d) shows shower axis uncertainty of about 2 mm in the Y-direction. The comparison is done within the range accessible in the data due to the acceptance area of the trigger scintillators and wire chamber, which is 10×10 cm^2.

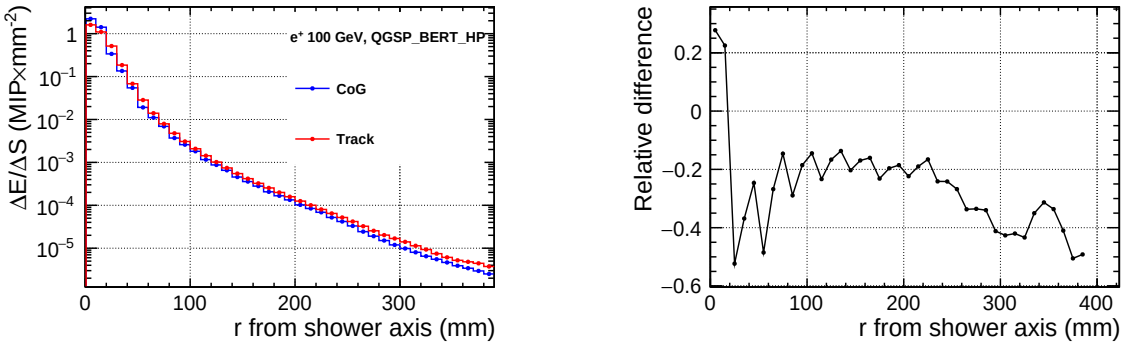


(a) (b)

(c) (d)

FIGURE 7.7: Distribution of track (red) and event centre of gravity (blue) for 100 GeV electron shower (a) in X-direction. (b) in Y-direction. Distribution of difference in Centre of Gravity events and tracks in (c) X-direction and in (d) Y-direction for 100 GeV electron. The y-axis is normalized to the number of events.

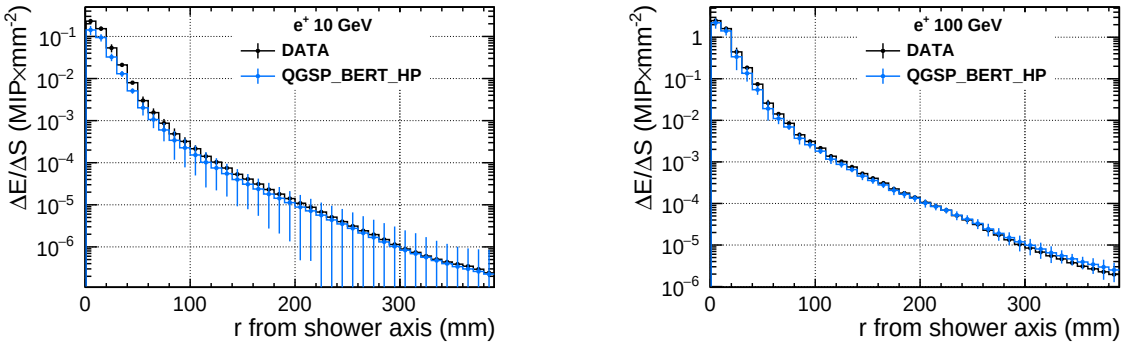
Based on the two approaches of shower axis reconstruction, the radial profiles are obtained as shown in figure 7.8(a). The radial profile is the average energy deposited in the calorimeter as a function of the distance r to the incoming track or from the event centre of gravity. The reconstruction of the shower axis from the event centre of gravity estimates a higher density in the core than the reconstruction of the shower axis from the tracks. While for larger radii, the energy deposits extracted using tracks information as shower axis is dominated. In the current analysis, the difference between the two methods of shower axis reconstruction is taken into account as a systematic uncertainty. The figure 7.8(b) shows dramatic disagreement between the two methods of shower axis reconstruction of $\sim$20% in the radius of 20 mm while $\sim$20-40% for larger radii. A similar level of disagreement is observed for the rest of the electron energies.



(a) (b)

FIGURE 7.8: (a) Radial shower profile for 100 GeV simulated data, with shower axis estimated from track and centre of gravity method. (b) The relative difference between shower axis estimated from track and centre of gravity method.

A comparison of radial shower profile using testbeam data and simulation is shown in figure 7.9(a) and figure 7.9(b), obtained using electron showers of beam energy at 10 GeV and 100 GeV, respectively. The simulation using `QGSP_BERT_HP` physics list reproduces the radial shower development reasonably well as compared to the testbeam data.



(a) (b)

FIGURE 7.9: (a) Radial shower development using electron shower is compared between simulated data and testbeam data for beam energy at (a) 10 GeV and (b) 100 GeV.

Furthermore, the quality of simulated data for radial shower profiles in the studied energy range between 10 and 100 GeV is shown in figure 7.10. It shows the ratios of radial shower profiles extracted from simulated data to those extracted from testbeam data. The energy deposition near the shower axis is underestimated by `QGSP_BERT_HP` up to 30% at all energies. In contrast, the overestimation of deposited energy by the physics list in the tail of the shower increases with energy up to $\sim$40% at 60 GeV.

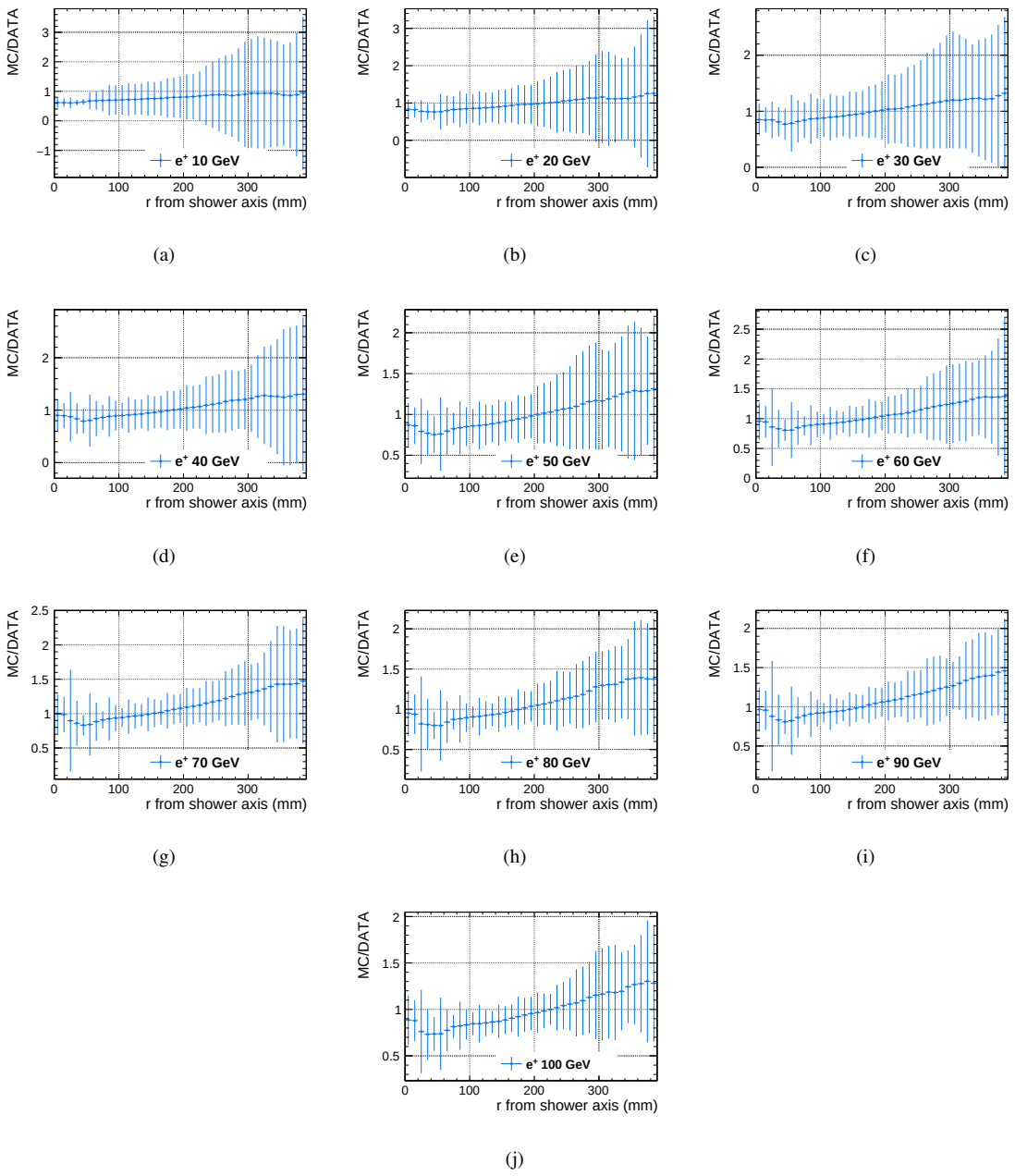


FIGURE 7.10: Ratio of radial profiles induced by electron showers from simulated data using QGSP_BERT_HP physics list to those from testbeam data. The error bars for data and simulation indicate the statistical and systematic uncertainties added in quadrature.

The radial shower profile is parametrized with the sum of two exponential distributions. The "core" component that is close to the shower axis and a "halo" component is distant from the shower axis is given by:

$$\frac{\Delta E}{\Delta S}(r) = \frac{E}{2\pi} \cdot \left\{ f_c \cdot \frac{e^{\frac{-r}{\beta_c}}}{\beta_c^2} + (1 - f_c) \cdot \frac{e^{\frac{-r}{\beta_h}}}{\beta_h^2} \right\}, \tag{7.3}$$

where E is the total energy in units of MIP is measured in the ring of width Δr and of area $\Delta S = 2\pi\Delta r$, assuming the integration along the longitudinal direction. f_c is the fractional contribution of the "core" component or the fraction of total energy in the core. β_c and β_h are the slope parameters in units of mm. The radial shower development is illustrated in figure 7.11 with an inner core part and an outer halo region.

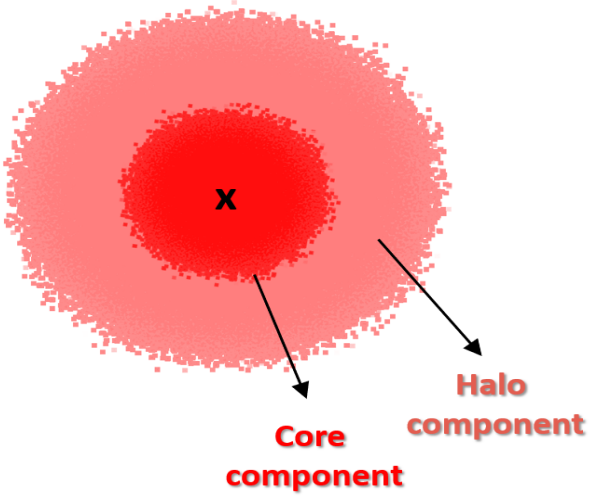
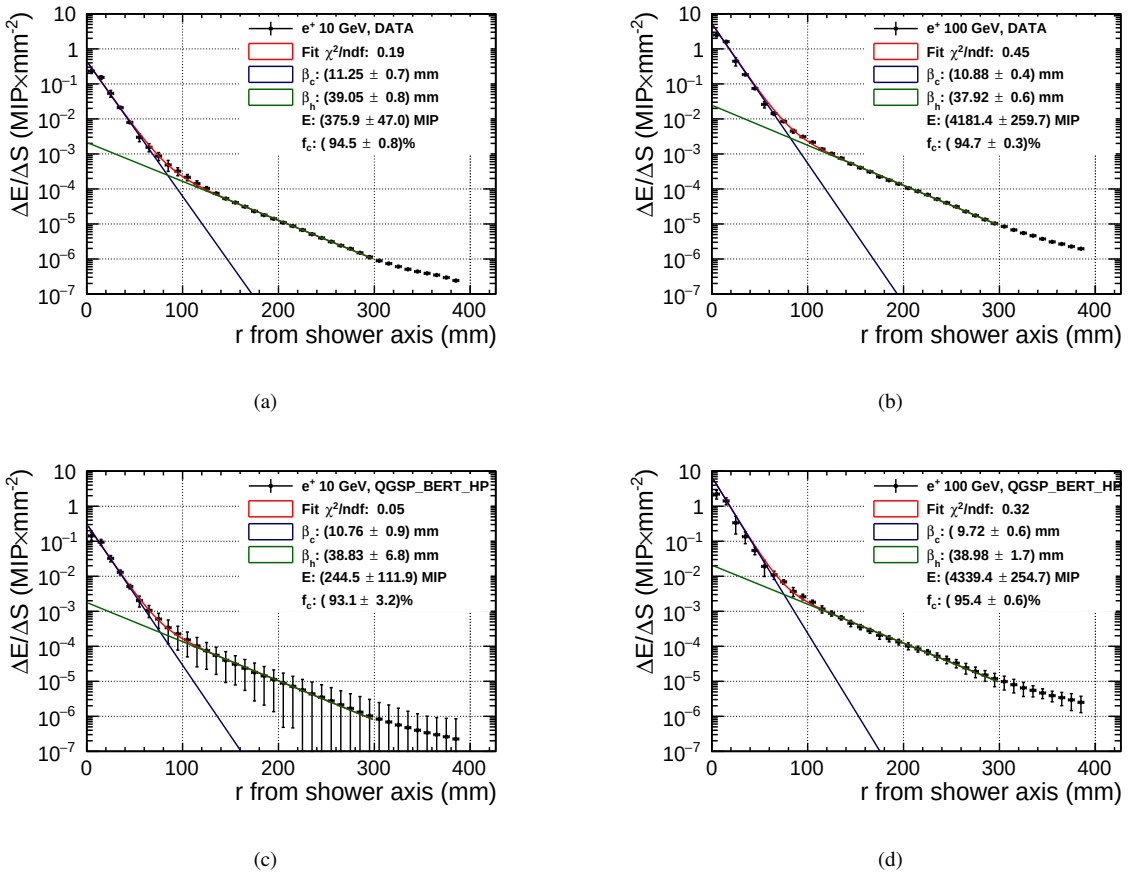


FIGURE 7.11: Visualisation of a radial core component representing EM-core and the halo component representing the late development of EM-showers.

A fit to radial profiles is performed using the radial parametrisation for different energies. An example of the radial shower profile for electron showers with energy 10 GeV and 100 GeV for testbeam data and simulation is shown in figure 7.12. The two components of the radial fit function are well represented, indicating contributions from the electromagnetic core and the halo components. The error bars indicate the uncertainty estimated from the two methods of shower axis reconstruction. An error of $\frac{\text{binwidth}}{\sqrt{12}}$ is used as uncertainty in the X-axis of the radial distribution, as the energy is assumed to be distributed uniformly between the virtual cells. The uncertainties from gain, MIP and SiPM saturation correction are negligible for radial shower shapes and are therefore not considered in the following.



(a) (b)

(c) (d)

FIGURE 7.12: Fits performed to radial shower profile shown by solid line using Eq. 7.3 (a) for 10 GeV electron shower and (b) for 100 GeV electron shower from testbeam data, and (c) for 10 GeV electron shower and (d) for 100 GeV electron shower from simulated data using QGSP_BERT_HP physics lists. The blue line represent the "core" component and the green line represents the "halo" part of the electron shower.

Figure 7.13 shows the reduced χ^2 obtained from the radial fit to electron showers. The χ^2 for low energy $<$0.3. Moreover, for the simulation, the χ^2 is always lower than data and always $<$0.5. This is a clear indication that the errors are overestimated for simulations.

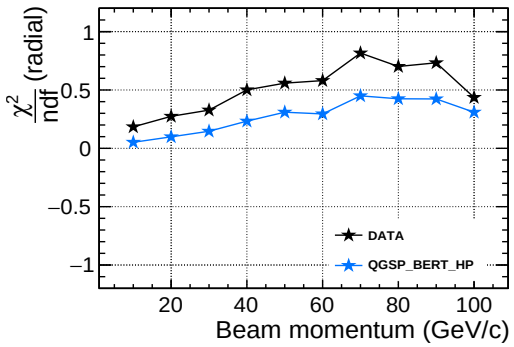


FIGURE 7.13: χ^2/NDF obtained from radial fits for all available electron energies from 10 GeV to 100 GeV both for data (black) and QGSP_BERT_HP physics list (blue).

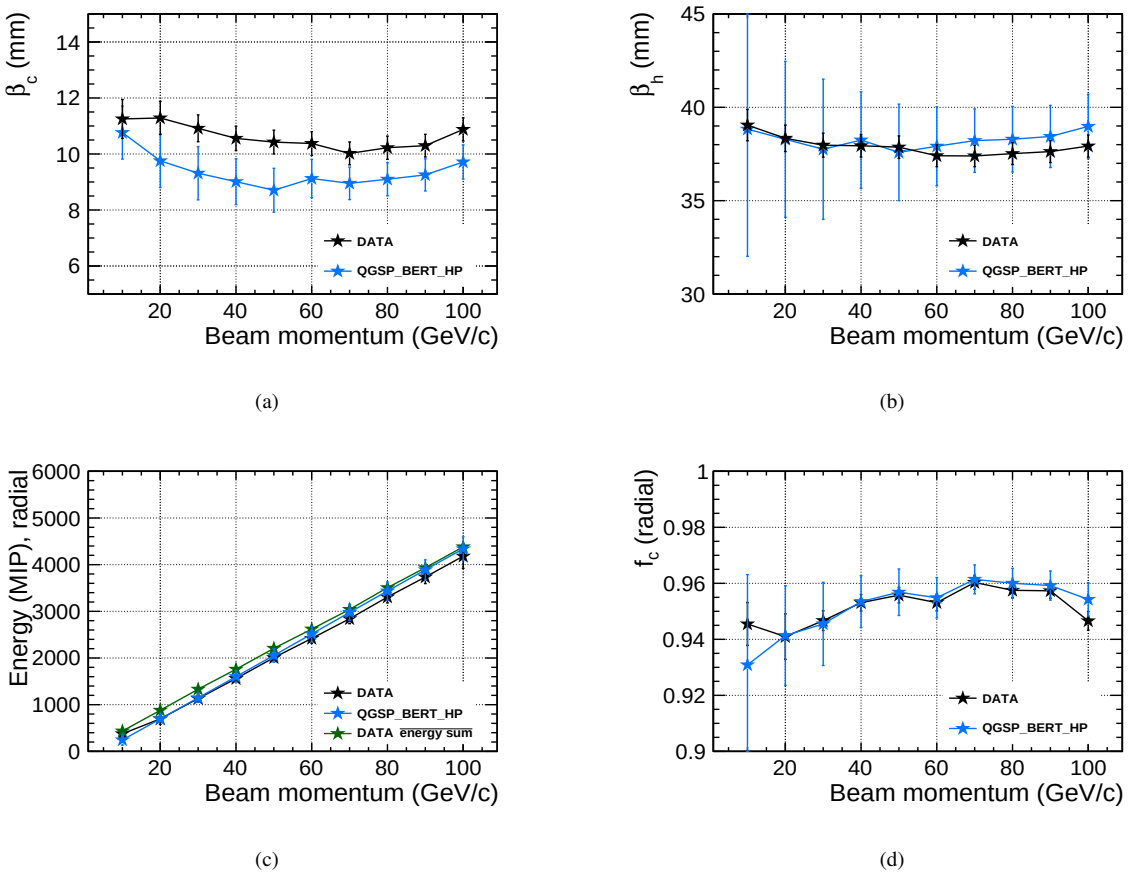


FIGURE 7.14: Comparision of fit parameters obtained from the fit to radial electron shower shapes between testbeam data (black) and simulation using QGSP_BERT_HP physics list (blue) as a function of beam energy (a) the core parameter β_{core} (b), the halo parameter β_{halo}, (c) total energy parameter and mean visible energy (green) as a reference for energy comparison, (d) and the fraction parameter f_c.

Furthermore, the fit results obtained for different electron energies are studied. The obtained parameters for β_c and β_h as a function of beam energy is shown in figure 7.14(a) and 7.14(b), respectively. These parameters show almost no energy dependence, and large uncertainties in the parameters are obtained. The physics list QGSP_BERT_HP (the default EM models are switched on in QGSP_BERT_HP for the electron shower as described in section 1.3) shows a slightly smaller value for β_c compared to testbeam data. In contrast, the β_h is seen to be higher in QGSP_BERT_HP for the higher beam energies and is completely within uncertainties. The parameter energy is plotted as the function of beam energy and is shown in figure 7.14(c). As for the electron dataset, the showers are well contained. Therefore, the E parameter from the fit, which integrates until infinity, and the energy sum, which covers "only" the calorimeter volume, agree well. Moreover, the testbeam data and simulation show a good agreement. The extracted core fraction f_c is about 94% and is almost independent at all the incident beam energies.

7.2 Analysis of Hadronic Shower Shapes

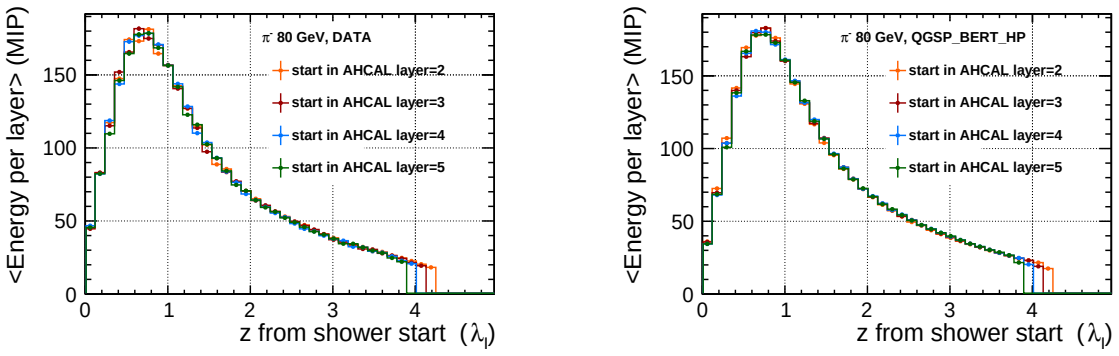
Monte Carlo simulation of a hadron shower evolution is challenging because up to now, there is no single model that describes hadronic interactions over the full range of energies. For completeness, these gaps must be bridged to describe the energy deposition in a calorimeter. Therefore, several models are combined in physics lists to cover the energies of particle showers taking place in detectors. As mentioned previously in section 1.3 two physics lists, `QGSP_BERT_HP` and `FTFP_BERT_HP`, have been picked for this analysis and evaluated with the shower shape variables. The first part of this section describes the one-dimensional longitudinal and radial shower profiles and the systematic uncertainties. Furthermore, in section 7.2.4 a simultaneous fit is performed from the two projections of the hadron shower. Subsequently, a three-dimensional model which describes an average hadronic shower is discussed in section 7.2.6. Finally, in section 7.2.7 the shower shapes are studied by using the Monte Carlo truth information.

7.2.1 Systematic Uncertainties

Longitudinal Shower Profile

The longitudinal shower development for showers induced by hadrons is measured from the shower start position, which is the relevant quantity for a hadronic shower. Figure 7.15(a) and 7.15(b) presents longitudinal shower profiles for 80 GeV pion from testbeam data and from simulated data. The profiles are obtained for events with shower start in a specific physical layer of the AHCAL, and all the profiles are normalised by the number of events. In principle, an ideal calorimeter, these shower profiles should not depend on the shower start position apart from the shower that starts in later layers, which are shorter in length. Therefore, the difference between profiles with different shower start positions is used to estimate the systematic uncertainty of the longitudinal shower shape measurement from detector effects that are not perfectly corrected (e.g. SiPM saturation)

The following procedure is used for simulated and testbeam data to quantify the systematic uncertainty. The longitudinal profiles for different fixed shower start layers have been evaluated. With a sample of size n, where n represents the number of individual profiles. In the current calculation, $n = 4$ as shown in figure 7.15. Each profile has p bins. The content of the i-th bin of the k-th profile is e_{ik}, where $1 < i < p$ and $1 < k < n$. The bin content of the mean profile, averaged over the single profiles, is $E_i = \sum_{k=1}^{n} e_{ik}/n$. The variance s_i of the bin content for the sample of the profiles is $s_i = \sum_{k=1}^{n} (e_{ik} - E_i)^2 / (n-1)$. The mean of profiles is assumed to be an estimate of the "true" profile and is used in this analysis. The uncertainty associated with the energy measurement in the i-th bin caused by variations between the single profiles is calculated as $\sqrt{\mathrm{s_i}/\mathrm{n}}$ and is used for the profiles.



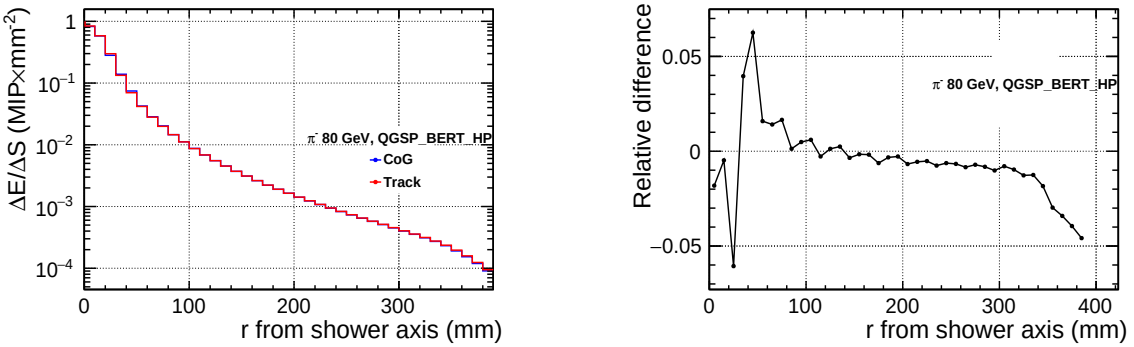
(a) (b)

FIGURE 7.15: Longitudinal shower profiles measured from specific shower start layer from layer 2 to layer number 5 for 80 GeV pion shower (a) from testbeam data and (b) from simulated data using QGSP_BERT_HP physics list.

Moreover, the uncertainties for the longitudinal shower profiles from MIP and Gain uncertainties are much smaller than those mentioned above and are therefore neglected.

Radial Shower Profile

The systematic uncertainty for the radial shower profiles is based on the same approach as described in section 7.1.2 as a difference between the two methods of shower axis reconstruction. For example, figure 7.16(a) shows the radial shower profiles from an 80 GeV pion shower using QGSP_BERT_HP physics list, obtained from CoG and track method. The resulting relative difference between the two methods is within ∼5% as shown in figure 7.16(b).



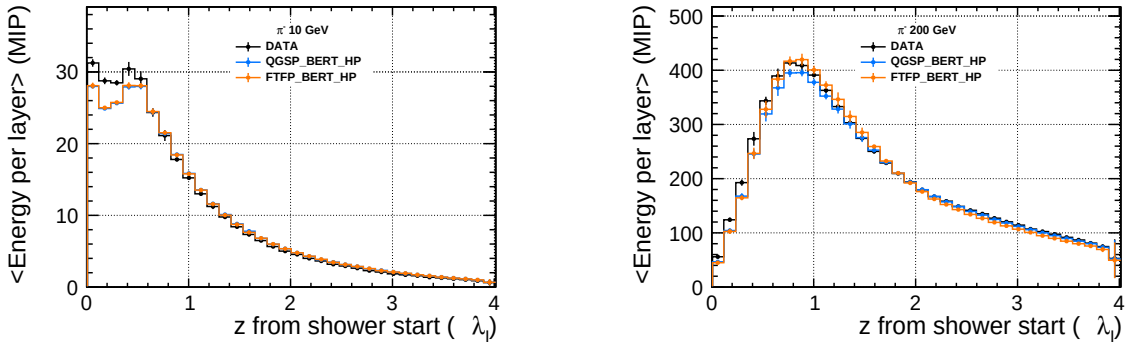
(a) (b)

FIGURE 7.16: (a) Radial shower profile for 80 GeV simulated sample extracted based on two approaches in shower axis reconstruction, Centre of Gravity (blue) and track (red). (b) The relative difference between two method of shower axis reconstruction.

Results

Longitudinal Shower Profile

The longitudinal shower profiles measured from shower start position is compared between data and GEANT4 physics lists (QGSP_BERT_HP and FTFP_BERT_HP) as shown in figure 7.17(a) for 10 GeV pion showers and 200 GeV pion showers in figure 7.17(b). Although overall behaviour in the longitudinal shower profiles from simulations shows a good prediction, the testbeam data suggest larger energy deposition in the first few. At the same time, both the physics lists at 10 GeV predict to have larger energy deposits roughly beyond the shower maximum compared to QGSP_BERT_HP at 200 GeV showing slightly lower energy deposits.



(a) (b)

FIGURE 7.17: Measured averaged longitudinal shower profiles in comparision between data (black) and simulations (QGSP_BERT_HP in blue and FTFP_BERT_HP in orange) with beam energy at (a) 10 GeV and (b) 200 GeV.

Figure 7.18 shows the ratio from physics lists to testbeam data for all available pion energies. The agreement of simulations to testbeam data shows an improvement from 20% by up to 5% in the first few interaction lengths and thereafter roughly remains constant. In addition, the testbeam data at all beam energies show an increase below 1λ. The same behaviour was observed in the electron data and also in the muon data [105]. This arises mainly due to the shower start identification. The variation between the two physics lists is not significant. However, the QGSP_BERT_HP shows slightly better agreement to testbeam data than the FTFP_BERT_HP physics list for the tail regions at all the energies. Furthermore, there is a small difference observed between the two physics lists in this respect. The QGSP_BERT_HP is stabilized to 1 after $\approx 1\lambda$. In contrast, the FTFP_BERT_HP oscillates.

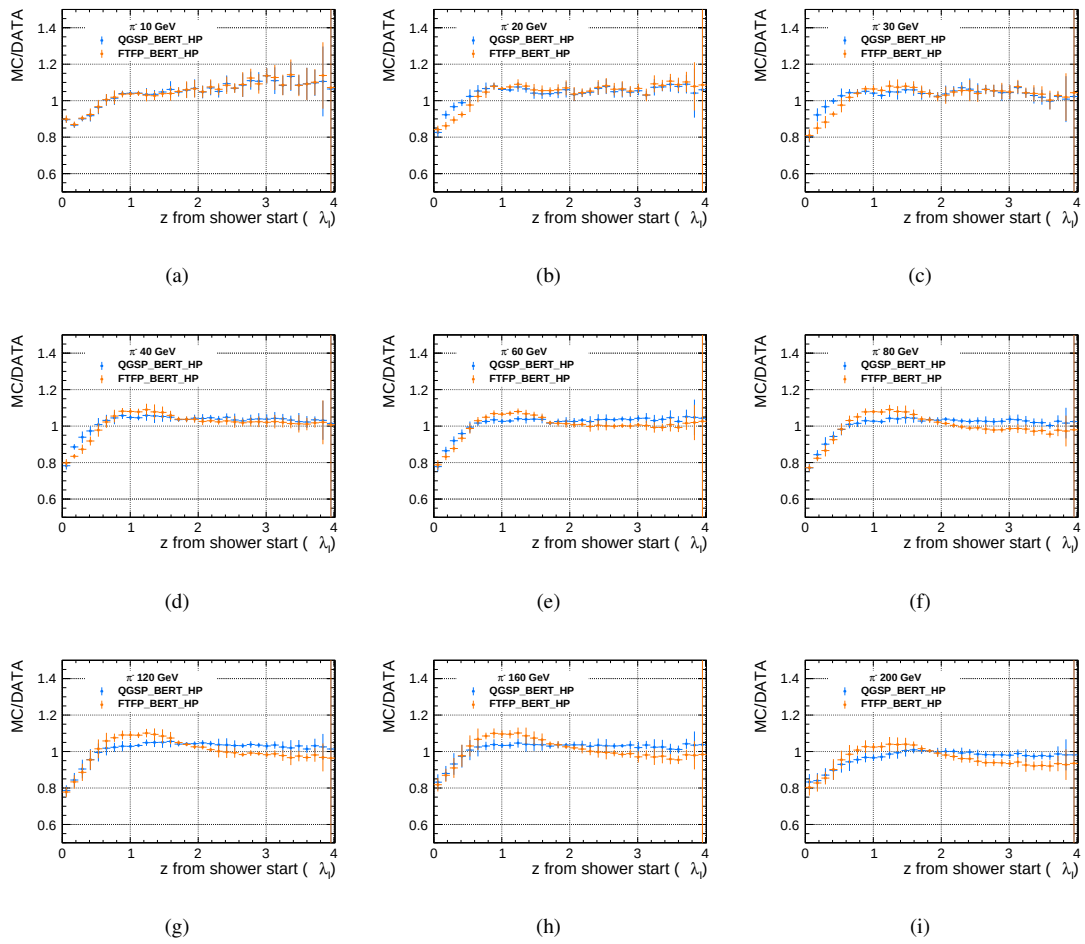


FIGURE 7.18: Ratio of radial shower profiles induced by pion showers from simulated data using QGSP_BERT_HP (in blue) and FTFP_BERT_HP (in orange) physics lists to those obtained from testbeam data. The error bars for data and simulations indicate the statistical and systematic uncertainties added in quadrature.

7.2.2 Longitudinal Parametrisation

The fine segmentation of the CALICE-AHCAL into 38 layers gives the unique possibility to obtain longitudinal shower profiles measured from the shower start, and the longitudinal shower development can be studied in great detail. The parametrisation of the longitudinal shower development for hadronic showers is described with a sum of two gamma distributions as proposed from the analysis of the data obtained using the AHCAL physics prototype [110]. The parametrisation is written as:

$$\Delta E(z) = E \cdot \left\{ \frac{f_{EM}}{\Gamma(\alpha_s)} \cdot \left(\frac{Z[X_0]}{\beta_s}\right)^{\alpha_s - 1} \cdot \frac{e^{\frac{-Z[X_0]}{\beta_s}}}{\beta_s} + \frac{1 - f_{EM}}{\Gamma(\alpha_l)} \cdot \left(\frac{Z[\lambda_I]}{\beta_l}\right)^{\alpha_l - 1} \cdot \frac{e^{\frac{-Z[\lambda_I]}{\beta_l}}}{\beta_l} \right\}. \tag{7.4}$$

The first summand is associated with the electromagnetic part within a hadronic shower called the "short" component. The second summand is the "long" component. It describes the truly hadronic component of the shower. A simple visualisation of the longitudinal shower is shown in figure 7.19. The interpretation of the short component as an electromagnetic component, therefore X_0 is used as a relevant length scale, and call the fraction as f_{EM}. Several parameters are obtained from the longitudinal parametrisation. E is the mean visible energy in units of MIP, which is calculated as the integral up to infinity. f_{EM} is the fractional contribution of the "short" component. α_s is the shape parameter and $\beta_s$is the slope parameter for the "short" component. α_l and β_l are the shape and the slope parameters of the "long" component. The two parameters α (a dimensionless coefficient) and β contain the energy dependence of the shape of the shower, $\Gamma(\alpha)$ is the gamma function ensuring the normalisation.

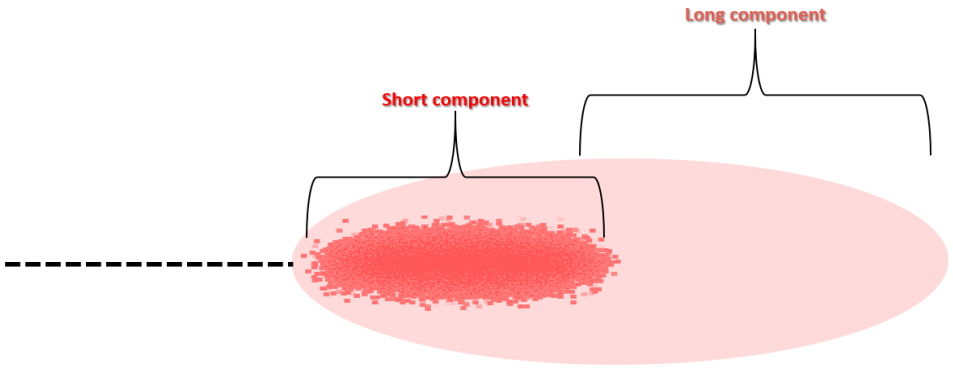
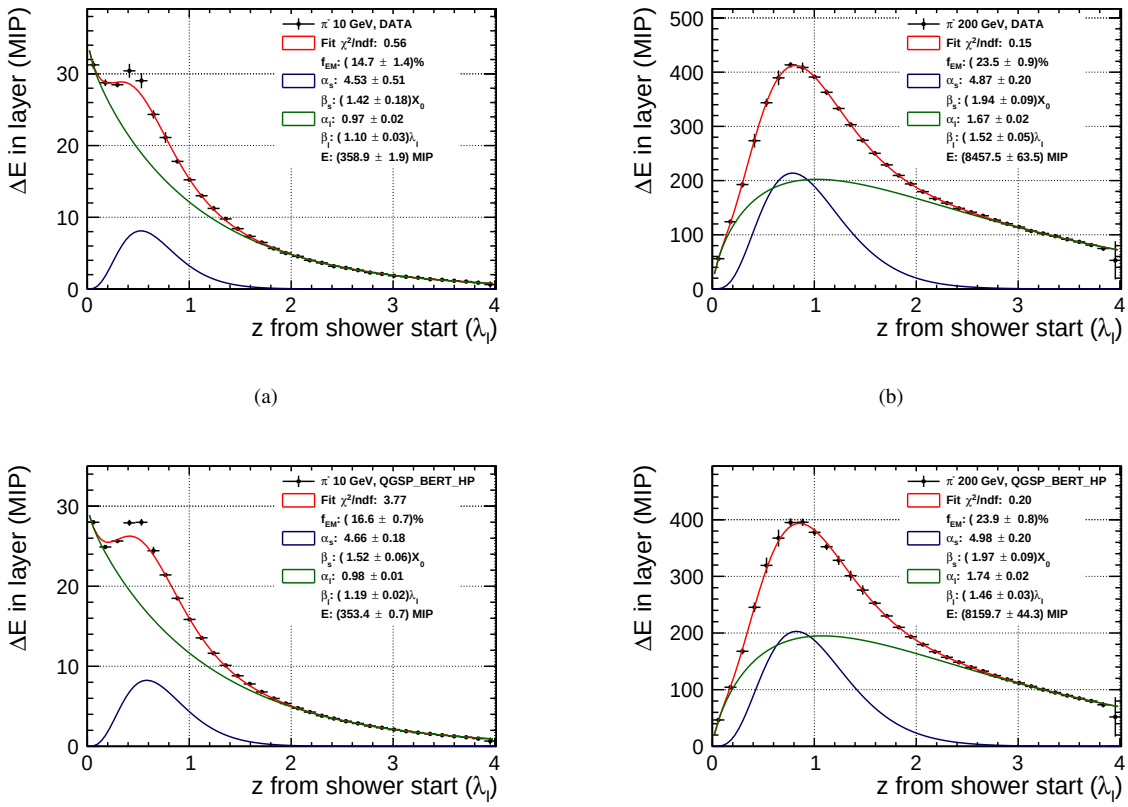


FIGURE 7.19: A sketch of a longitudinal shower depicting two parts of a hadron shower. The short component at the beginning relates to the electromagnetic part of the hadronic shower, and the long component relates to the showering hadrons.

The longitudinal parametrisation is applied to the shower profile for different pion energies. The longitudinal fit range corresponds to a depth of 4 λ_I. An example of longitudinal shower profiles for pion with beam energy of 10 GeV and 80 GeV using testbeam data is shown in figure 7.20(a) and 7.20(b) and using Monte Carlo in figure 7.20(c) and 7.20(d), respectively. The systematic uncertainties are estimated as described in section 7.2.1 and are added in quadrature to the statistical uncertainties.



(a) (b)

(c) (d)

FIGURE 7.20: Fits performed to longitudinal shower profile using Eq. 7.4 depicted in red (a) for 10 GeV pion shower and (b) for 200 GeV pion shower from testbeam data, and (c) for 10 GeV pion shower and (d) for 200 GeV pion shower from simulated data using QGSP_BERT_HP physics lists. The blue curve represents the "short" component and green curve represents the "long" component.

The shape is well described by the model, explaining the rise due to the production of secondaries and the fall due to the absorption of secondaries. The short component is well contained within the calorimeter as the electromagnetic part of the hadronic shower does not penetrate in greater depth, and the energy loss is at a faster rate. On the other hand, the long component is not contained. The fit quality to longitudinal shower profile for the studied pion energy range is shown in figure 7.21 using testbeam data and simulated data. Apart for beam energy at 10 GeV, the reduced-χ^2 obtained from longitudinal fit is below 1.

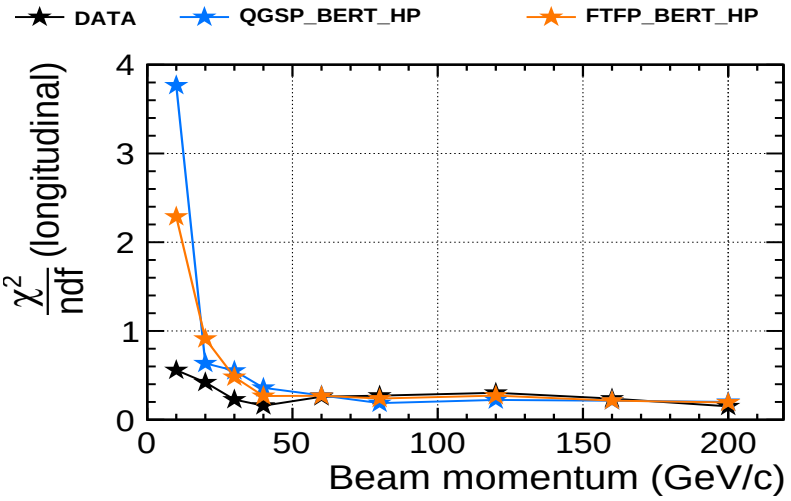


FIGURE 7.21: The χ^2/NDF from longitudinal fits for data (black), and simulations with QGSP_BERT_HP (blue) and FTFP_BERT_HP (orange) physics list.

The resulting parameters from the fit, set in comparison as a function of beam energy both for testbeam data and simulations using QGSP_BERT_HP and FTFP_BERT_HP physics list. In addition, the testbeam data points from the 2007 physics prototype is plotted as a reference. The major differences between the current technological and the previous physics prototype are listed in the table below:

Element	**Physics Prototype (2007)**	**Technological Prototype (2018)**
Sensor	scintillator + WLS fiber + SiPM	SiPM on scintillator tile readout
Total active layers	38	38
Total channels	7,608	21,888
Absorber thickness	21.0 mm	17.2 mm
Tile thickness	5 mm	3 mm
Cell-size	varying (3×3, 6×6, 12×12 cm^2)	homogeneous (3×3 cm^2)
Total depth	$\sim 5.3\lambda_I$ ($\sim 0.14\ \lambda_I$/layer)	$\sim 4.5\ \lambda_I$ ($\sim 0.11\ \lambda_I$/layer)
	$\sim 47.16\ X_0$	$\sim 40.8\ X_0$
Molière radius	25.5 mm	24.9 mm

TABLE 7.1: Comparison of geometry description for 2007 physics prototype and 2018 technological prototype. Composition and material properties for the AHCAL prototype is obtained from [10].

Fit Parameters from Longitudinal Shower Shape

The short parameters α_s and β_s from the fit to longitudinal shower profile are plotted as a function of beam momentum and is shown in figure 7.22(a) and 7.22(b).

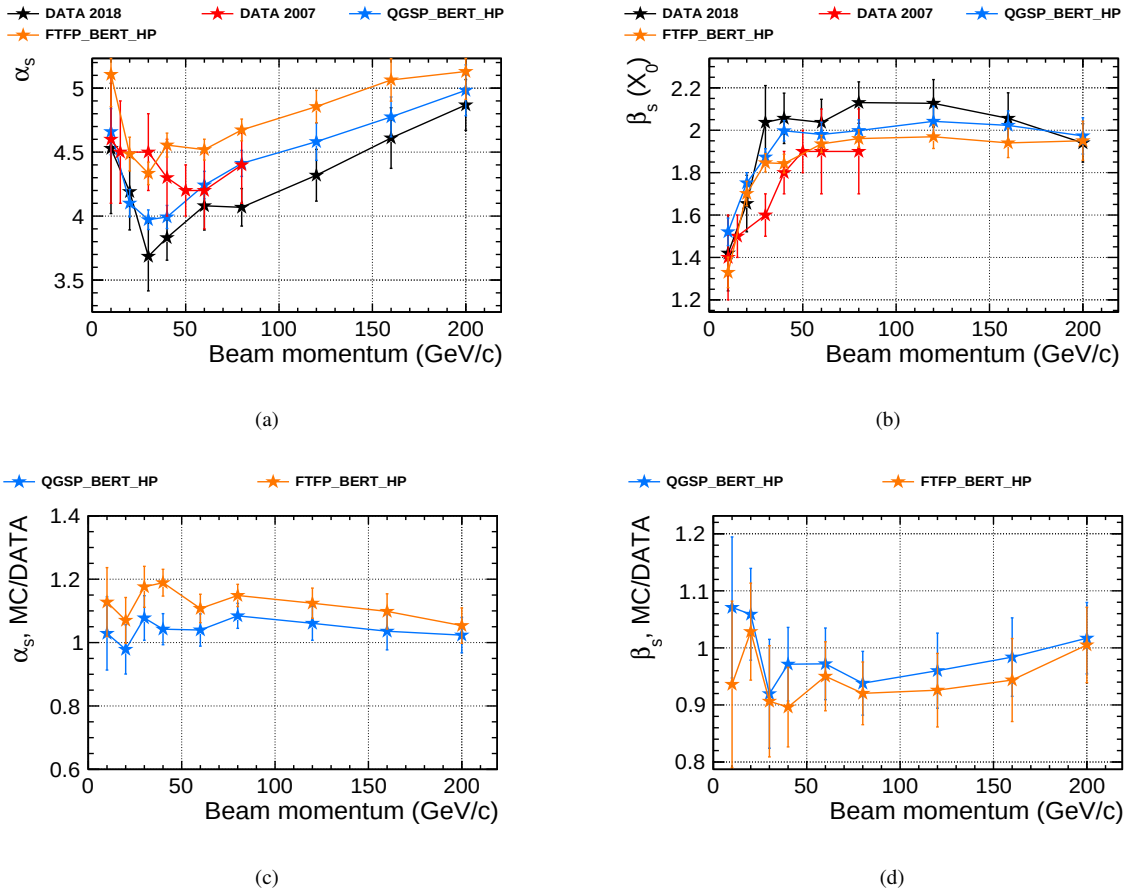


FIGURE 7.22: Energy dependence of shape parameters from 2018 testbeam data (black), and simulations using `QGSP_BERT_HP` (blue) and `FTFP_BERT_HP` (orange) physics lists, along with comparision to the testbeam data (red) from the 2007 physics prototype for (a) α_s and (b) β_s. The ratio of simulation to testbeam data for (c) α_s and (d) β_s.

The two parameters correlate with each other. As one parameter increases, the other shows a decreasing behaviour. The behaviour of α_s shows an increase above 30 GeV, and the parameter β_s shows a slight decrease at higher energies. The fit parameters α_s and β_s show a similar trend between the two physics lists. The shape and values are comparable to the 2007 testbeam data results. Figure 7.22(c) and 7.22(d) show the ratio of short parameters from simulations to those from testbeam data. The agreement between the two sets of physics lists for the parameter α_s and β_s is roughly within 10% for all the available energies.

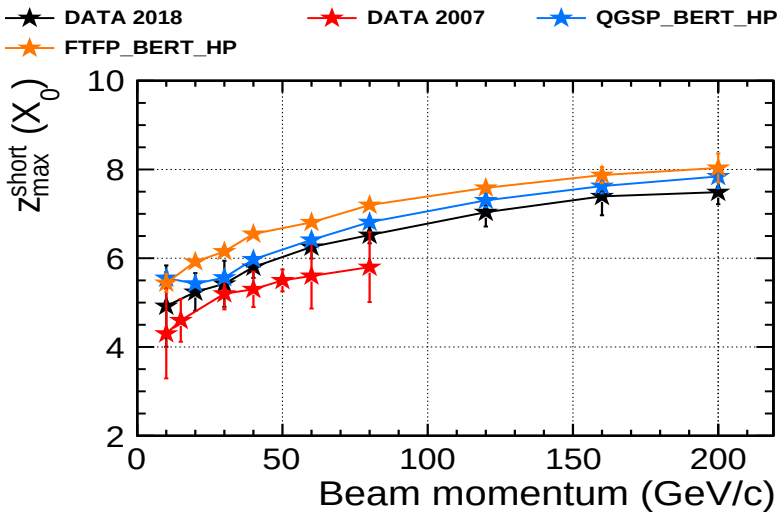


FIGURE 7.23: Maximum of the "short" component in units of radiation length X_0 as a function of pion beam energies. The shower maximum is obtained for 2018 testbeam data (black) and simulations with `QGSP_BERT_HP` (blue) and `FTFP_BERT_HP` (orange) physics list, along with the comparison to the 2007 testbeam data (red) from the physics prototype [110].

The maximum position of the short component z_{max}^{short} using Eq. 7.2 shows a logarithmic dependence with energies as shown in figure 7.23. For α_l, the results from the 2018 testbeam data and testbeam data from 2007 coincide within uncertainties.

Furthermore, the long parameters α_l and β_l are plotted as a function of pion beam energies and is set to a comparison between physics lists and testbeam data from 2007. Figure 7.24 shows the comparison of long parameters. The behaviour shows a slow increase with energy, although this is not entirely true for the slope parameter β_l. The 2018 data points show slightly lower values compared to simulations and testbeam data. Also, the 2007 testbeam data [110] shows slightly different behaviour for the β_l compared to the current fit results from the 2018 datasets. The ratio of α_l and β_l from simulation to those from testbeam data agrees within $\sim$10%.

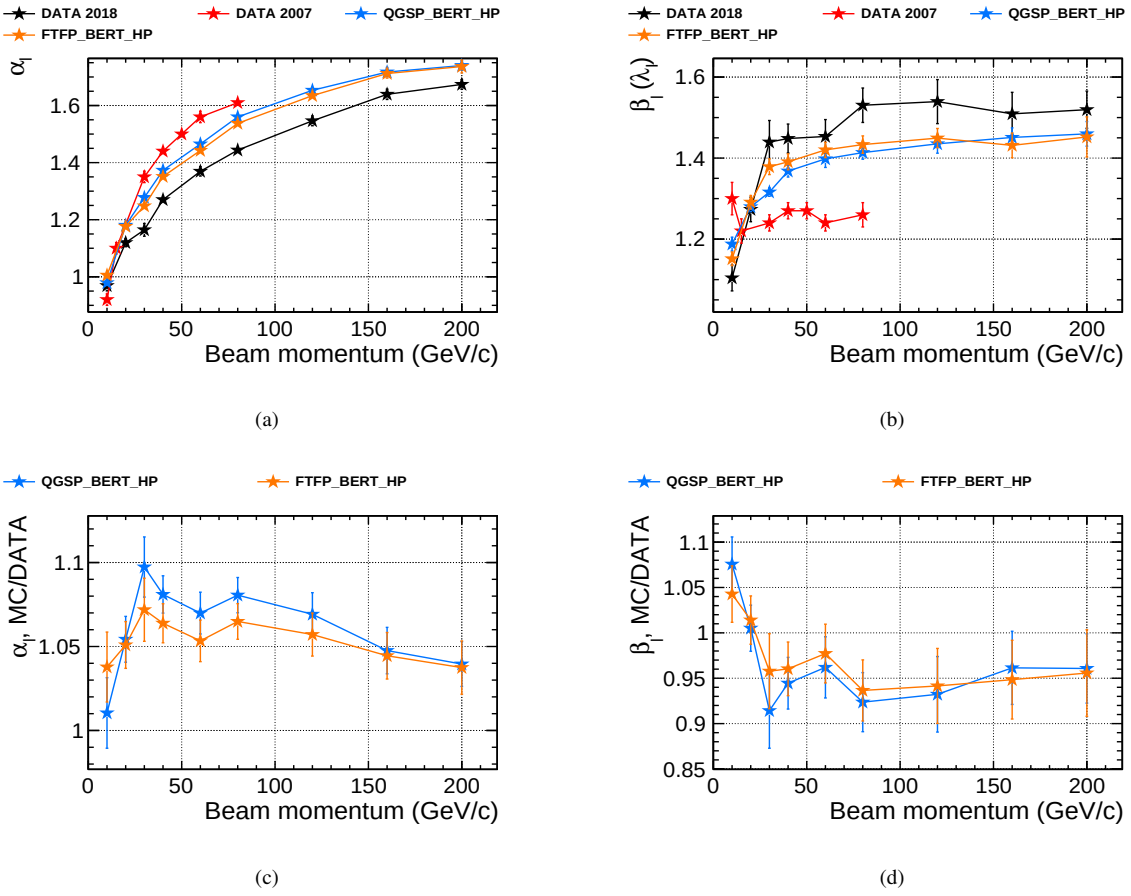


FIGURE 7.24: Energy dependence of shape parameters from 2018 testbeam data (black), and simulations using `QGSP_BERT_HP` (blue) and `FTFP_BERT_HP` (orange) physics lists, along with comparision to the testbeam data (red) from the 2007 physics prototype for (a) α_l and (b) β_l. The ratio of simulation to testbeam data for (c) α_l and (d) β_l.

h/e signal ratio

Signal non-linearity is a ubiquitous feature for hadron shower detection. As discussed in section 1.3.4, the response $\mathcal{R}$ of a calorimeter to a hadron shower is the sum of electromagnetic E_e and non-electromagnetic (hadronic) E_h components of the hadron shower energy, $E = E_e + E_h$.

$$\mathcal{R} = e \cdot E_e + h \cdot E_h, \tag{7.5}$$

where the quantity e and h are responses to the electromagnetic and non-electromagnetic components of a hadron-induced shower. The h/e signal ratio is an important source of signal non-linearity, as the varying f_{EM} introduces an energy dependence in the response to hadrons It is an intrinsic property of all non-compensating calorimeters and gives the degree of non-compensation which cannot be measured directly.

Usually, h/e is inferred from e/π signal experimentally measured at several energies. Both components E_e and E_h in Eq. 7.5, measured in the electromagnetic scale can be expressed in

terms of mean electromagnetic fraction f_{EM}, mean hadronic fraction $1 - f_{EM}$, the beam energy E_b and is written as follows [110]:

$$E_e = f_{EM} \cdot E_b, \quad E_h = h/e \cdot (1 - f_{EM}) \cdot E_b. \tag{7.6}$$

In the current analysis, the h/e signal ratio is estimated from the fit to longitudinal profiles. The method and formalism of h/e is stated in the previous analysis [110] and is repeated for the current technological prototype. The decomposition of longitudinal profiles into electromagnetic and hadronic components gives the possibility to estimate the h/e ratio. Based on this method, the mean visible energy in each component (E_e and E_h) in MIP equivalent can be calculated as the integral up to infinity under the corresponding curve (short and long components) from the fit to longitudinal profile, where, $E_e = E_e^{fit}$ and $E_h = E_h^{fit}$.

The estimates of the deposited energy in units of GeV for E_e and E_h is obtained by multiplying each integral by the electromagnetic calibration factor, C_{em}. This factor was extracted from dedicated positron runs, and the value of C_{em} is 0.02198 GeV/MIP, obtained from figure 6.12(a).

The expression for h/e is then derived from Eq. 7.6 and can be rewritten as:

$$\frac{h}{e} = \frac{E_{\text{h}}^{\text{fit}}}{E_{\text{b}} - E_{\text{e}}^{\text{fit}}}, \tag{7.7}$$

where, $E_h^{fit} = E_{reco} \cdot (1 - f_{EM}) \cdot C_{\text{em}}$ and $E_e^{fit} = E_{reco} \cdot f_{EM} \cdot C_{\text{em}}$.

Figure 7.25 shows the estimated values of h/e extracted from longitudinal fits using Eq. 7.7.

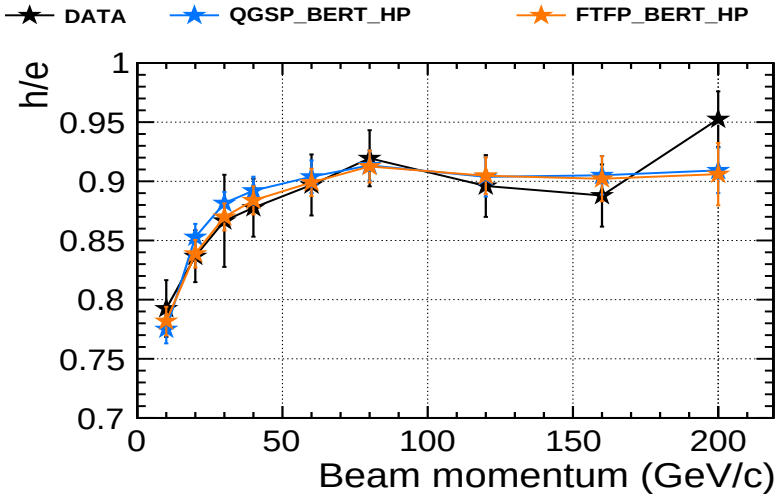


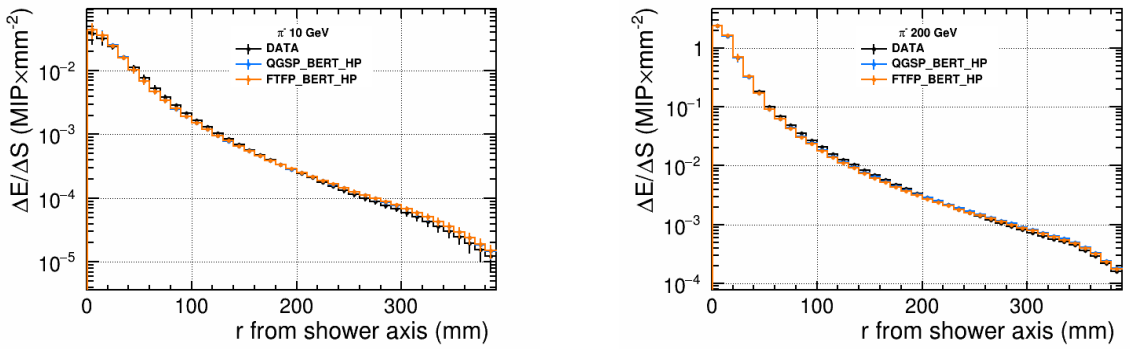
FIGURE 7.25: Estimates of h/e signal ratio from the fit to longitudinal profiles acquired using Eq. 7.7 for testbeam data (black) and simulated data using `QGSP_BERT_HP` (blue) and `FTFP_BERT_HP` (orange) physics lists.

The value of h/e is lower than 1, but certainly not very far from it. Thus, the AHCAL is non-compensating calorimeter. Traditionally h/e value scales with a power law. The values of

h/e are well predicted by both the simulations and agree with data within 5%. The behaviour of the h/e ratio extracted from a fit to longitudinal profiles shows a very slow dependence and almost no dependence at higher energies. This is because the energy spectrum of the secondaries that dominate in the shower are energy independent. The values extracted for the h/e ratio might be over-estimated. This could be due to the simplified description of the longitudinal fit function. By all means, the structure of the longitudinal distribution of energy density is more complex. With increasing incident hadron energy, the probability of neutral hadron production in secondary interactions increases. In a case where electromagnetic sub-showers are produced far from the shower starting point, they will contribute more likely to the "long" component, thus leading to this overestimated value of h/e at higher energies.

Radial Shower Profile

Next, the radial shower profiles are investigated. An example of a radial shower profile overlaid with simulations for comparison is shown in the figure 7.26(a) and 7.26(b) with beam energy at 10 and 200 GeV respectively. The ratio plot of radial shower profiles from two physics lists and from testbeam data for all energies is shown in figure 7.27.



(a) (b)

FIGURE 7.26: Measured averaged radial shower profiles in comparision between testbeam data (black) and simulations (`QGSP_BERT_HP` in blue and `FTFP_BERT_HP` in orange) of showers induced by pions with beam energy at (a) 10 GeV and (b) 200 GeV.

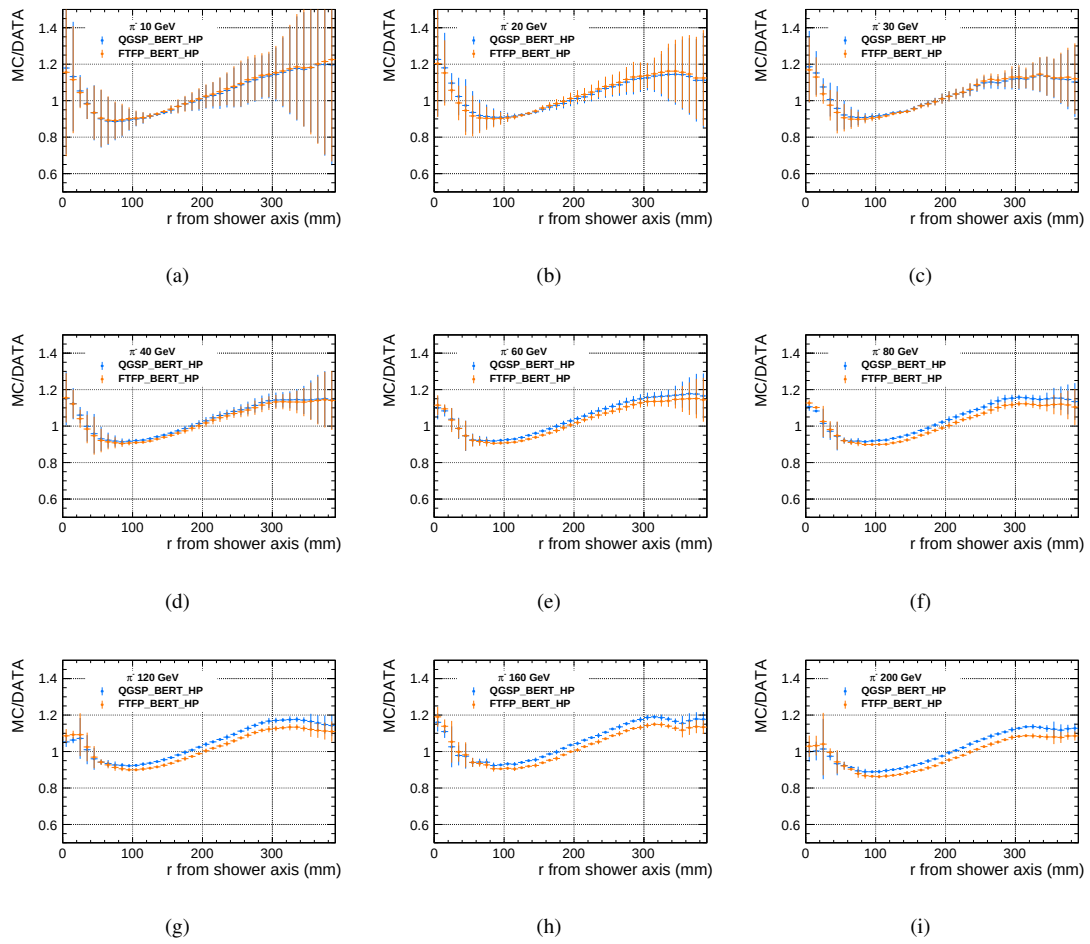


FIGURE 7.27: The ratio of radial shower profiles induced by pion showers from simulated data using `QGSP_BERT_HP` (in blue) and `FTFP_BERT_HP` (in orange) physics lists to those obtained from testbeam data. The error bars for data and simulations indicate the statistical and systematic uncertainties added in quadrature.

From the ratio of radial shower profiles, the two physics lists predict to have excess energies at the core and at the tails of the showers for most of the energies. The `QGSP_BERT_HP` shows excess energy at the tails compared to `FTFP_BERT_HP` for energies above 40 GeV.

Radial Parametrisation

The functional form of the radial parametrization looks slightly different than the parametrization used in the physics prototype paper. The function is now normalized with the integral up to infinity equal E. The radial parametrization is given by:

$$\frac{\Delta E}{\Delta S}(r) = \frac{E}{2\pi} \cdot \left\{ f_{\mathrm{EM}} \cdot \frac{e^{\frac{-r}{\beta_{\mathrm{c}}}}}{\beta_{\mathrm{c}}^2} + (1 - f_{\mathrm{EM}}) \cdot \frac{e^{\frac{-r}{\beta_{\mathrm{h}}}}}{\beta_{\mathrm{h}}^2} \right\}, \tag{7.8}$$

where the fit parameter E corresponds to the total energy in the unit of MIP, which is the normalisation factor, fraction f_{EM} corresponds to the electromagnetic fraction. The radial slopes β_c and β_h represents the "core" and "halo" components expressed in units of mm.

As for electron shower profiles, to describe the radial shower behaviour, the sum of two exponential distributions is needed. Unlike the electron shower, for pion showers, the two parts of the shower denote a different sense. For pion radial shower profiles, the first term, called as the "core" component, represents the electromagnetic part of the hadronic shower. The second term, called as "halo" component, describes the truly hadronic part.

The radial parametrisation based on Eq. 7.3 is then applied to the shower profile for different pion energies. The peripheral regions of radial profile where the calorimeter does not cover the full ring with that radius are excluded.. An example of radial shower profiles for pion with beam energy of 10 GeV and 80 GeV using testbeam data is shown in figure 7.28(a) and 7.28(b) and using Monte Carlo in figure 7.28(c) and 7.28(d), respectively. The systematic uncertainties are estimated as described in section 7.2.1 and are added in quadrature to the statistical uncertainties.

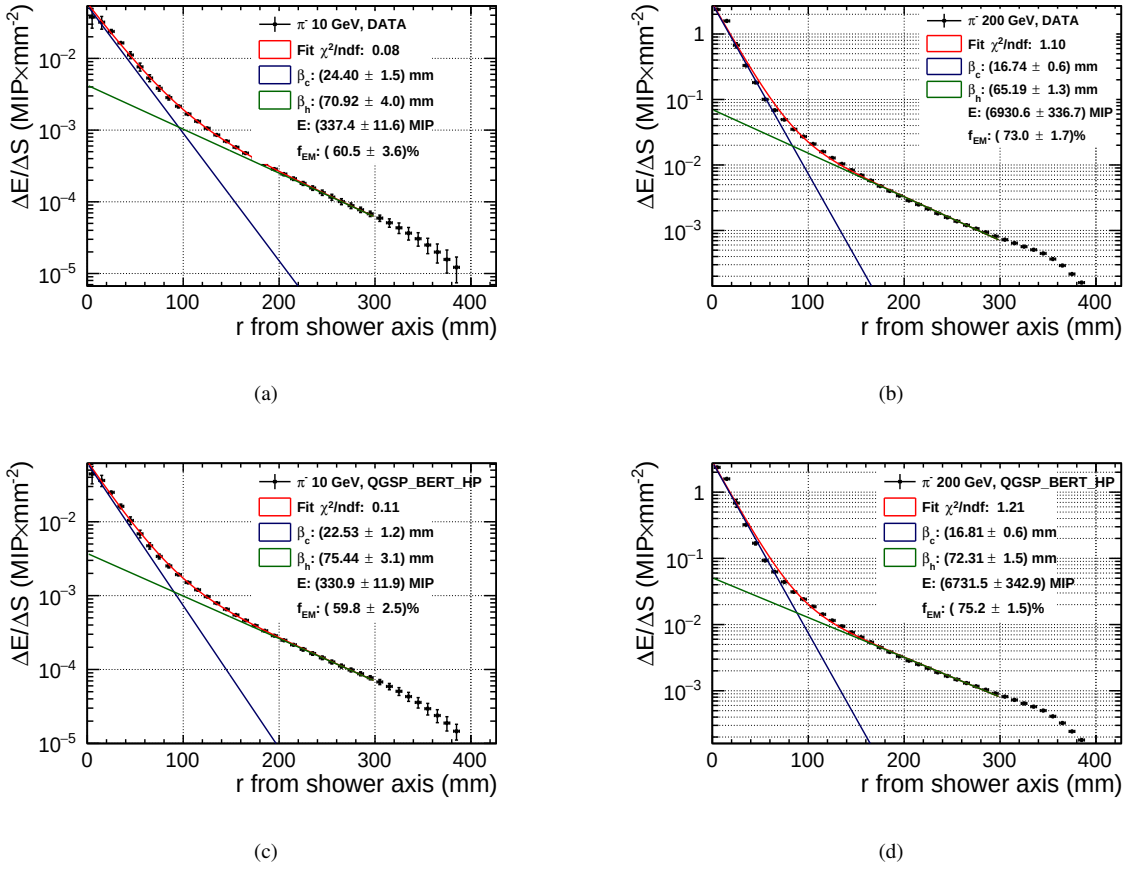


FIGURE 7.28: Fits performed to radial shower profile using Eq. 7.8 depicted in red (a) for 10 GeV pion shower and (b) for 200 GeV pion shower from testbeam data, and (c) for 10 GeV pion shower and (d) for 200 GeV pion shower from simulated data using `QGSP_BERT_HP` physics lists. The blue curve represents the "core" component and green curve represents the "halo" component.

The fit to radial shower profiles results in a good description. For the entire pion energy range both for testbeam data and simulations show a reasonable χ^2/NDF for energies above 10 GeV as shown in figure 7.29.

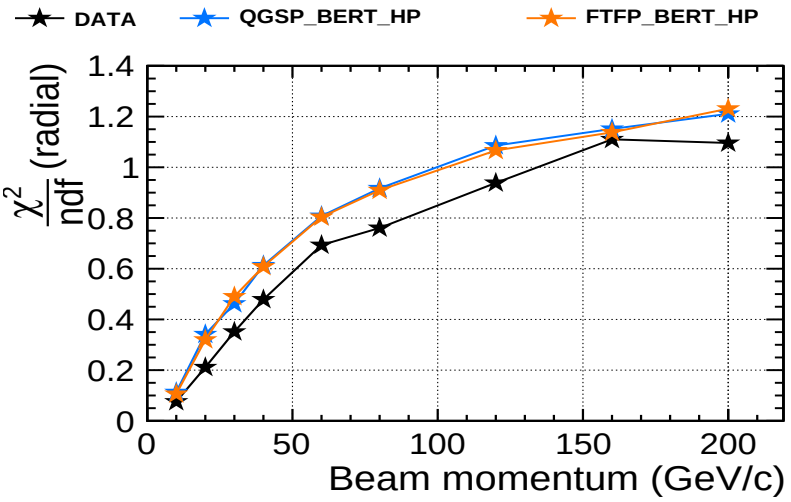


FIGURE 7.29: χ^2/NDF values from radial fits for testbeam data (black) and simulated data with `QGSP_BERT_HP` (blue) and `FTFP_BERT_HP` (orange) physics list.

Fit Parameters from Radial Shower Shape

The parameter obtained from the fit to radial shower profiles is studied as a function of beam energy. Figure 7.30 shows the radial core β_c and halo β_h parameters, which is compared between physics list (`QGSP_BERT_HP` and `FTFP_BERT_HP`) and testbeam datasets from 2007 and 2018.

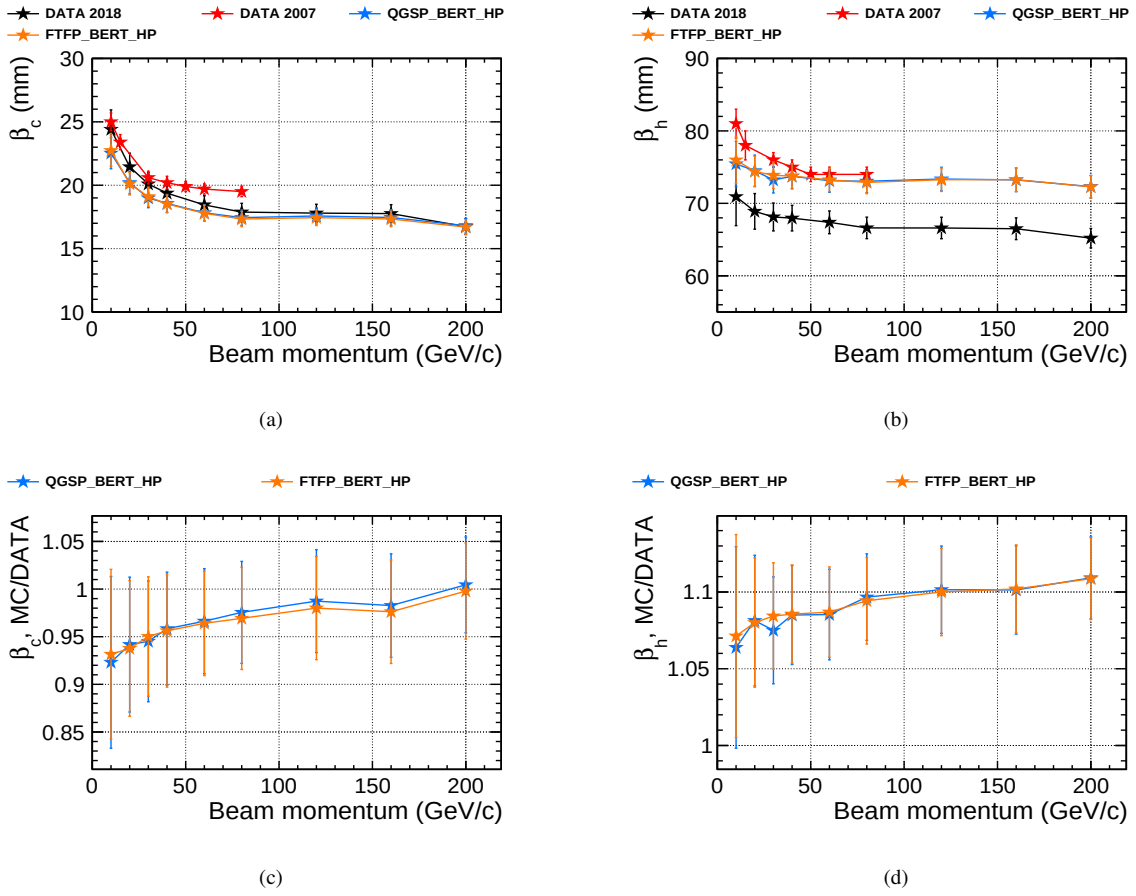


FIGURE 7.30: Energy dependence of slope parameters (a) β_{core} and (b) β_{halo} for 2018 testbeam data (black), and simulations with `QGSP_BERT_HP` (blue) and `FTFP_BERT_HP` (orange) physics lists, along with comparision to the 2007 testbeam data (red) from the physics prototype. The ratio of simulation to data for β_{core} in (c) and β_{halo} in (d).

The shape behaviour of the two parameters decreases with increasing energies, which is consistent between the two datasets (2007 and 2018) and the simulations. The results from the 2018 testbeam data are slightly lower than the 2007 testbeam. The reason is caused by the thickness of the scintillator tile, which is different between the two prototypes as given in table 7.1 and that makes the Molière radius different. Therefore, a calorimeter with a smaller Molière radius shows a smaller value in β_c. Moreover, in the physics prototype tested in 2007, there were different segmentation from the outer region to the central point, leading to a rough estimate for the halo region.

Both parameters (core and halo) are highly correlated; if not very reliable estimates on halo, this will affect the core component. Furthermore, it is observed that the simulations predicted to have a slightly higher halo component closer to the values obtained from 2007 compared to the data values from 2018. This is expected due to differences in the prototype where the Molìere radius plays an important role. The agreement between testbeam data and simulation for radial core and halo parameters is roughly 5-10%.

Reproducibility

As the longitudinal parameters are expressed in units of X_0 or λ_I, so this is a valid comparison, whereas for the radial fit parameters are expressed in units of mm and have different Molière radii. In addition, the two prototypes vary in sampling fractions, about 0.17 for the current prototype while ≈ 0.24 for the physics prototype.

To provide an impression of how similar the two prototypes are, the plot in figure 7.31 shows the ratio of shower shape parameters. The result is shown for both longitudinal and radial shower profiles as a function of beam energy from 10 GeV to 80 GeV between the testbeam data from the 2007 physics prototype and the datasets from the current technological prototype. Furthermore, it confirms that the agreement between the shower shape parameter values obtained from two data measurement periods (between 2007 and 2018) is nearly consistent up to a level of 30% even though the geometry between the two prototypes was not the same as mentioned in the table 7.1.

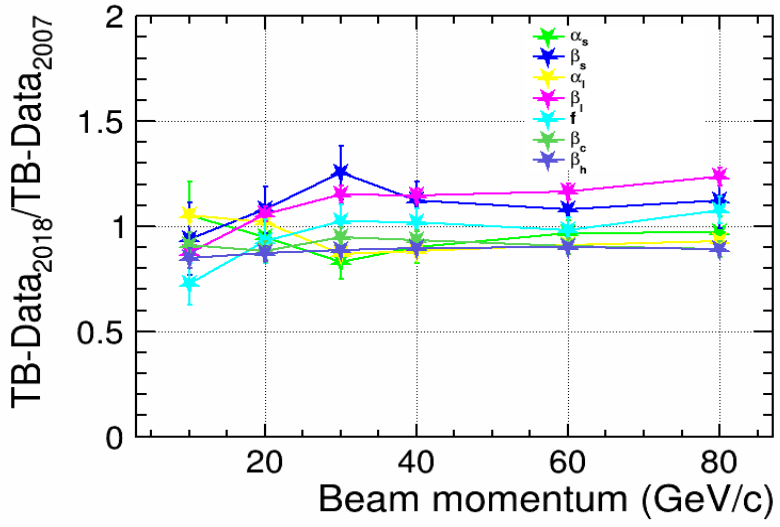


FIGURE 7.31: Ratio of shower shape parameters obtained from radial and longitudinal shower profiles between the testbeam data from 2018 and the testbeam data from 2007.

Comparision of Energy and Fraction from Longitudinal and Radial Shower Shape Fits

A comparative study for the energy and fraction parameters is done between testbeam data and simulated data from the two projections of the shower. Figure 7.32(a) shows the energy parameter obtained from the fit to longitudinal shower profile, and figure 7.32(b) shows the energy parameter obtained from the fit to radial shower profile. In addition, this includes the points from 2007 testbeam data. As a reference point, the mean visible energy sum from the measured data is also compared with the energy fit parameter.

The value of the energy parameter extracted from the longitudinal fit is higher in general than the reference value from the mean visible energy, which is due to leakage. The longitudinal

parametrisation is evaluated up to a depth of infinity. Therefore, the reconstructed energy is corrected for longitudinal leakage. Whereas in the case of energy parameter obtained from the radial shower profile, the energy is well contained within the calorimeter, and no large effects of leakage are observed in case of radial leakage even at higher energies.

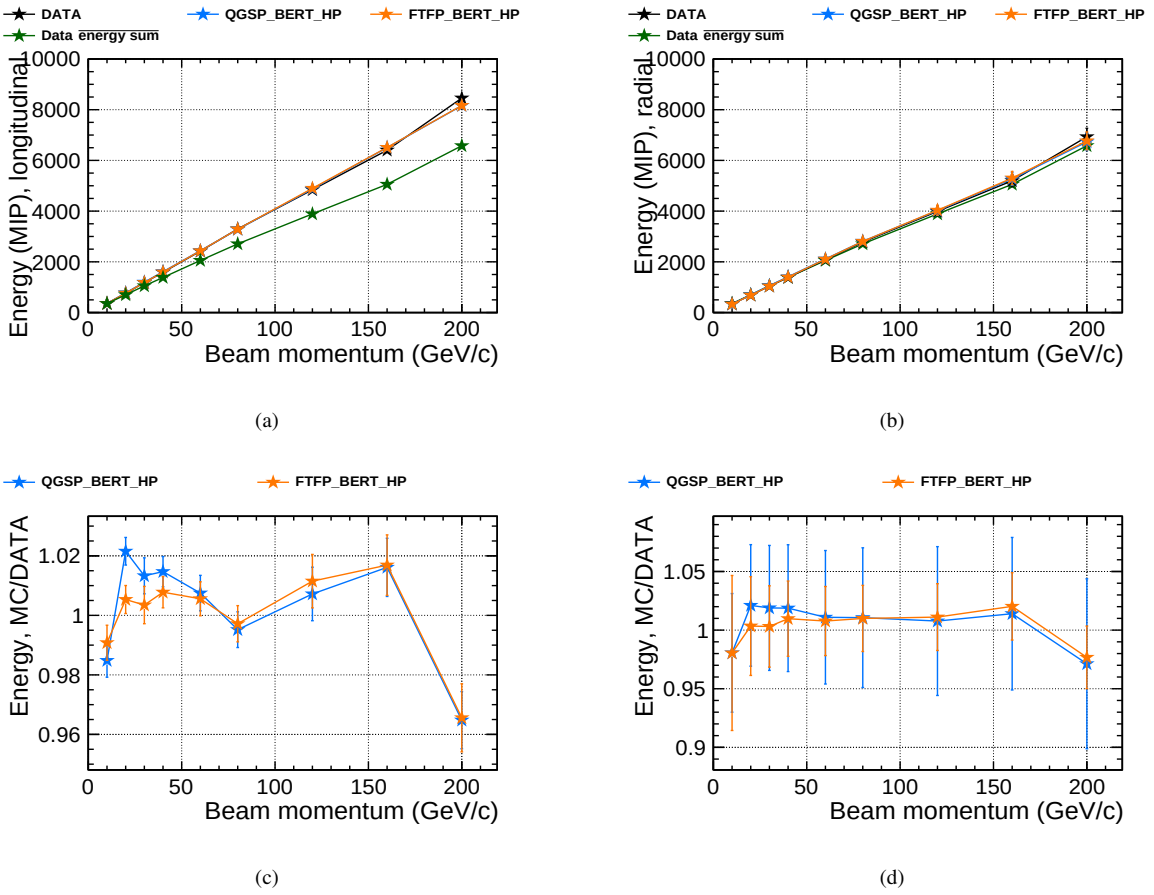


FIGURE 7.32: Energy dependence of energy parameter, (a) from the fit to longitudinal shower profile (b) from the fit to radial shower profile for 2018 testbeam data (black), and simulations with QGSP_BERT_HP (blue) and FTFP_BERT_HP (orange) physics lists. The ratio of simulation to data with energy parameter from longitudinal fit in (c) and with energy parameter from radial fit in (d).

The fraction of the "short" or "core" component is interpreted as it could be the average electromagnetic fraction f_{EM}. It is more or less consistent with results from the physics prototype. The parameter f_{EM} obtained from the longitudinal shower profile is much lower than the value obtained from the radial shower profile. The part of it might be that what is measured is the core fraction, and we assume that this is the point where the shower starts and the first interaction happens. The fact that the large f_{EM} value for high energy pion showers is because, in the first hard interaction it produces on average $\frac{1}{3}\pi^-$, $\frac{1}{3}\pi^+$ and $\frac{1}{3}\pi^0$. Then these $\pi^\pm$ will travel a bit and make a second interaction and then again a $\frac{1}{3}\pi^-$, $\frac{1}{3}\pi^+$ and $\frac{1}{3}\pi^0$ are produced. The charged pion has travelled a bit, and there is a distance, which means this will probably not be contributed into the f_{EM} of the fit because it is not at $r = 0$ and probably at a later shower start.

In addition, the parameter f_{EM} obtained from longitudinal fit is smaller in value as the parametrization considers up to a depth of infinity. As a result, the extracted fraction is smaller by construction. The fractional contributions of the core component related to the electromagnetic fraction can also be calculated from the integral under the fit to radial profiles. However, the estimated uncertainties from these integral values are much higher compared to that of the f_{EM} parameter extracted from the fit to longitudinal profiles.

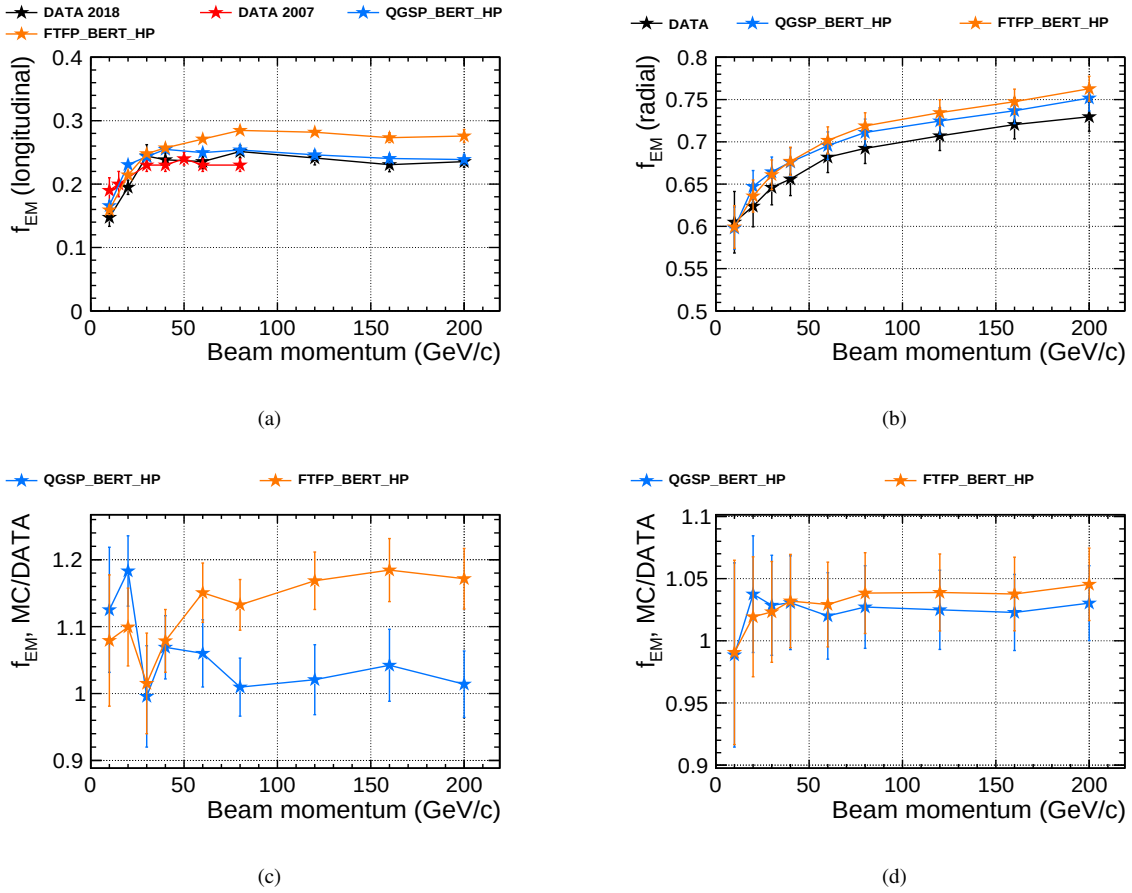


FIGURE 7.33: Energy dependence of fraction parameter f_{EM} (a) from the fit to longitudinal shower profile (b) from the fit to radial shower profile for 2018 testbeam data (black), and simulations with `QGSP_BERT_HP` (blue) and `FTFP_BERT_HP` (orange) physics lists, along with comparision of f_{EM} to the 2007 testbeam data (red) from the physics prototype. The ratio of simulation to data with f_{EM} from longitudinal fit in (c) and f_{EM} from radial fit in (d).

7.2.3 Improved Radial Parametrisation

From figure 7.33(a) and 7.33(b) we see that the parameter f_{EM} extracted from the radial fit is roughly a factor three times larger compared to the f_{EM} obtained from the longitudinal fit. One reason is that the short or core part is well contained in both longitudinal and radial directions of the shower. Therefore, we can assume that the integral under the short or core part is similar in radial and longitudinal profiles. The hadronic or long component in the longitudinal profile is significant compared to the radial part. This is because, in longitudinal parametrization, it

considers an ideal calorimeter with an infinite depth. As a consequence, the f_{EM} obtained is smaller by the construction model. However, a correction is needed for the integral under the long component in the longitudinal profile. To correct for leakage effects, the radial profile is considered. In this case, the integral over the radial distribution will not correct for longitudinal leakage. As expected, the integral over the radial distribution is to be smaller than the longitudinal distribution. Instead of directly using the f_{EM} parameter, an effective f_{EM} is used, and this is corrected for the tail in the longitudinal profile with a K factor. The K factor is extracted from the longitudinal fit, which is the ratio of the hadronic energy integrated up to infinity to the hadronic energy visible in the calorimeter.

Furthermore, an improved radial fit function is used. We had seen in section 7.1 that electromagnetic showers require two components to describe the shower shape. This means in the hadronic shower, in addition to the electromagnetic core component caused by neutral pions, for the charged hadron, there may also be a narrow contribution that is purely hadronic. This is caused by the relativistic part of the hadronic shower that does not spread and is followed by a hadronic halo component far away from the shower center, mainly caused by neutrons. The behaviour of the hadronic shower can roughly be visualized as schematically shown in figure 7.34.

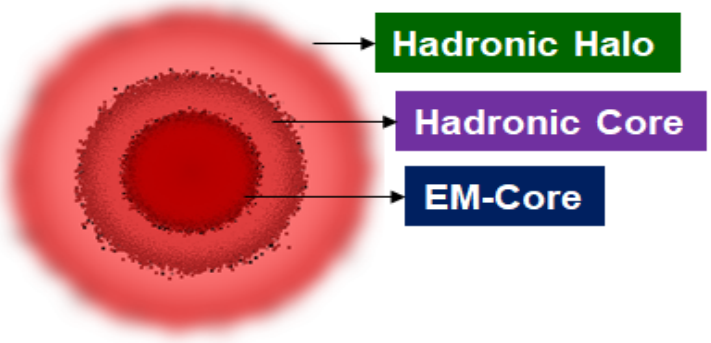


FIGURE 7.34: Schematic depicting the three main contributions in an hadronic shower. The central EM-core from the hadronic showers surrounded by the small part of hadronic core and later developing hadronic halo.

Mathematically, an "effective" radial parametrisation is derived, which is an extension to the previously defined radial parametrisation given in Eq. 7.8. The parametrisation makes the following assumptions:

- The electromagnetic component is contained both in longitudinal and radial direction
- The hadronic component is contained in radial direction but not in longitudinal direction

and makes use of the following definitions:

- The electromagnetic fraction is the ratio of electromagnetic energy over the total energy upto to infinity written as:

$$f_{EM} = E_{EM}/E_t^{\infty} \tag{7.9}$$

- The total energy is defined as the sum of electromagnetic and hadronic energy and is written as:

$$E_t^{\infty} = E_{EM} + E_{HAD}^{\infty} \tag{7.10}$$

- The quantity K, is called the correction factor or the leakage fraction. It is given by the ratio of total energy upto infinity to the total energy visible within the calorimeter which chosen to be within $4\lambda_I$,

$$K = E_t^{\infty}/E_t^{z<4}. \tag{7.11}$$

From the above assumption and definition we can deduce the following set of equations. As from longitudinal profile we have,

$$\Delta E(z) = E_{EM} \cdot p_s + E_{HAD}^{\infty} \cdot p_l, \tag{7.12}$$

where p_s refers to the Γ-function with the short component and the p_l refers to the Γ-function with the long component based on Eq. 7.4 of the longitudinal parametrization. Using Eq. 7.9 and Eq. 7.10 the above equation can be expressed in terms of corresponding fractions as follows:

$$\Delta E(z) = E_t^{\infty} \cdot \{f_{EM} \cdot p_s + (1 - f_{EM}) \cdot p_l\}. \tag{7.13}$$

From the radial parametrization given in Eq. 7.8 we can obtain the following:

$$\frac{\Delta E}{\Delta S}(r) = E_{EM} \cdot p_c + E_{HAD}^{z<4} \cdot p_h, \tag{7.14}$$

where p_c refers to the exponential distribution with core component and p_h refers to the exponential distribution with the hadronic component of the radial function. Further, using Eq. 7.9 and Eq. 7.11, $E_{HAD}^{z<4}$ can be written as follows:

$$\begin{aligned} E_{HAD}^{z<4} &= E_t^{z<4} - E_{EM}, \\ \text{where } E_t^{z<4} &= \frac{E_t^{\infty}}{K} \\ E_{HAD}^{z<4} &= \frac{E_t^{\infty}}{K}(1 - K \cdot f_{EM}). \end{aligned} \tag{7.15}$$

Using Eq. 7.15 and replacing E_{EM} and $E_{HAD}^{z<4}$ in Eq. 7.14. It can be re-written as:

$$\begin{aligned} \frac{\Delta E}{\Delta S}(r) &= E_t^{\infty} \cdot f_{EM} \cdot p_c + \frac{1}{K} \cdot E_t^{\infty} \cdot (1 - K \cdot f_{EM}) \cdot p_h \\ &= \frac{1}{K} \cdot E_t^{\infty} \{K \cdot f_{EM} \cdot p_c + (1 - K \cdot f_{EM}) \cdot p_h\} \end{aligned} \tag{7.16}$$

As mentioned earlier that the p_c corresponds to the electromagnetic core of the hadronic shower which is similar to the core component from the pure electromagnetic shower.

$$p_c = \frac{1}{2\pi} \cdot \frac{e^{\frac{-r}{\beta_c^{EM}}}}{(\beta_c^{EM})^2}, \tag{7.17}$$

where $\frac{1}{2\pi}$ is needed for the normalization. As two components are required to describe the electron showers, describing the electromagnetic core of the pion shower would then require two components and, in addition, a third component for the "pure" hadronic shower. However, involving three exponential distributions and too many free parameters would lead to bad fit quality. Therefore an effective method is considered for the hadronic part of the radial parametrization. Whereby the "pure" hadronic component would then comprise two parts: a hadronic core and a hadronic halo. Hence p_h is expressed as the sum of the hadronic core and the hadronic halo in the form of an exponential distribution.

$$p_h = \frac{1}{2\pi} \cdot \left(f_c \cdot \frac{e^{\frac{-r}{\beta_c^{HAD}}}}{(\beta_c^{HAD})^2} + f_h \cdot \frac{e^{\frac{-r}{\beta_h^{HAD}}}}{(\beta_h^{HAD})^2} \right) \tag{7.18}$$

Using the Eq. 7.17 and Eq. 7.18, the Eq. 7.16 can be rewritten and is called as an "improved radial parametrization" which is normalized to the total energy and is given by:

$$\frac{\Delta E}{\Delta S}(r) = \frac{E}{2\pi K} \cdot \left\{ f_{EM} \cdot K \cdot \frac{e^{\frac{-r}{\beta_c^{EM}}}}{(\beta_c^{EM})^2} + (1 - f_{EM} \cdot K) \left(f_c \cdot \frac{e^{\frac{-r}{\beta_c^{HAD}}}}{(\beta_c^{HAD})^2} + f_h \cdot \frac{e^{\frac{-r}{\beta_h^{HAD}}}}{(\beta_h^{HAD})^2} \right) \right\} \tag{7.19}$$

The meaning of individual parameters involved in the radial parametrisation are defined below:

parameter	**meaning**
r	radial distance from shower axis (mm)
ΔS	$2\pi r \Delta r$, area of the ring of width Δr (mm^2)
E	mean visible energy (MIP)
K	hadronic leakage fraction
f_{EM}, f_c, $f_h = 1 - f_c$	fraction for electromagnetic core, hadronic core, and hadronic halo, respectively
β_c^{EM}, β_c^{HAD}, β_h^{HAD}	shape parameters for electromagnetic core, hadronic core, and hadronic halo, respectively (mm)

TABLE 7.2: Definition of parameters of the improved radial parametrisation.

The "improved radial parametrisation" is applied to the radial shower profile for all the pion energies. Moreover, based on the knowledge from the pure electron showers, the parameter β_c^{EM} is independent of energy as shown in figure 7.14(a). Therefore, the parameter value β_c^{EM} is fixed

with the values from the fit to the EM showers as 11.30 mm for testbeam data and 10.55 mm for both the physics lists. Moreover, the parameter K value is fixed here from the longitudinal fits.

An example of the fit performed to the radial shower profile using the parametrisation given in Eq. 7.19 for 10 GeV and 200 GeV pion showers is shown in figure 7.35 for the testbeam data and simulations. The three components of the parametrisation are well described for the different parts of the hadronic shower.

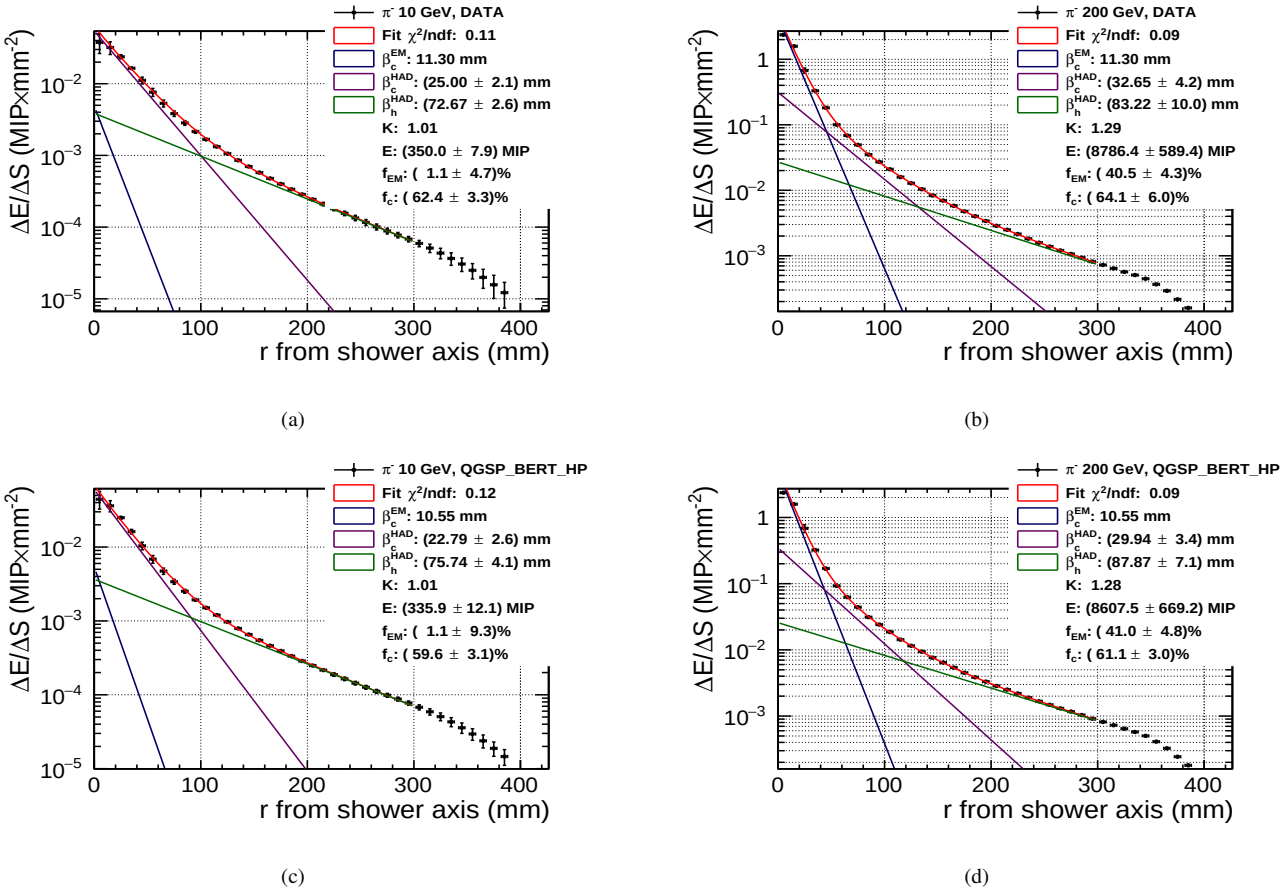


FIGURE 7.35: Fit performed to radial profiles using Eq. 7.19. For showers initiated by pions with beam energy of (a) 10 GeV from testbeam data and (b) 200 GeV from testbeam data. (c) 10 GeV from `QGSP_BERT_HP` physics list (d) 200 GeV from `QGSP_BERT_HP` physics list. The blue, green and magenta curves show the contributions of "EM-core", "Hadronic core" and "Hadronic halo" components, respectively.

7.2.4 Simultaneous Fitting

Based on the assumption that the total energy E and fraction f_{EM} from the longitudinal and radial development of the shower should agree with each other, it is then possible to fit the two parameters simultaneously. Performing simultaneous fit with sharing parameters E and f_{EM} allows us to obtain an average value of f_{EM}. The method to perform the simultaneous fitting for both longitudinal and radial distributions is done using a statistical method based on χ^2 minimization as given below:

$$\tilde{\chi}^2_{L,R} = \sum_{i=1}^{n_l} \frac{(\mu_i^l - F_l(E,\alpha,\beta,f_{EM}))^2}{\sigma_i} + \sum_{i=1}^{n_r} \frac{(\mu_i^r - F_r(E,\beta,f_{EM}))^2}{\sigma_i} \tag{7.20}$$

The value obtained for energy and fraction is the resultant from the best χ^2 with the simultaneous fits to longitudinal and radial shower shapes. An example of a simultaneous fit is shown in figure 7.36 for pion showers of beam energy at 10 GeV and 200 GeV, performed on longitudinal (left) and radial (right) profiles using testbeam data and simulations.

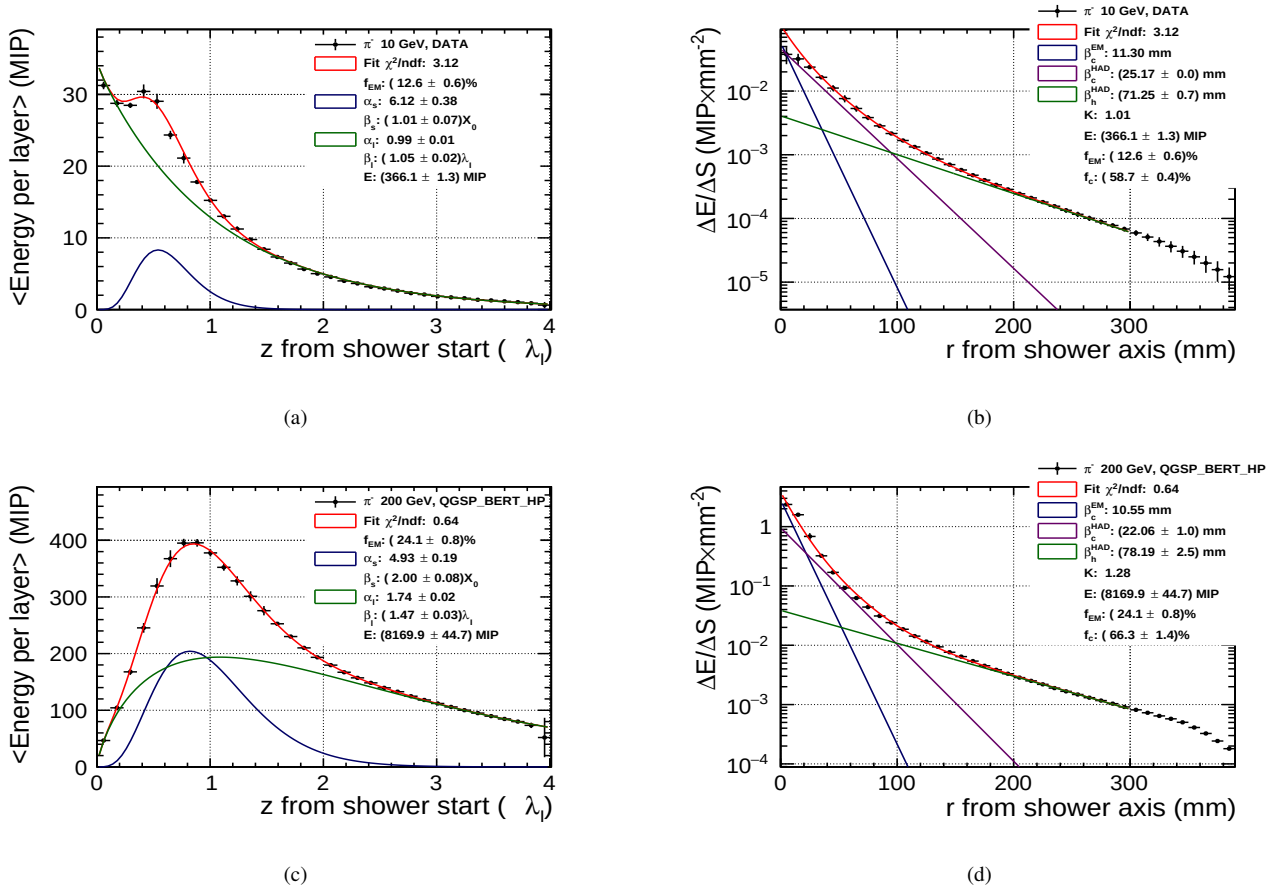


FIGURE 7.36: Simultaneous fitting of the longitudinal and radial distribution. For showers initiated by pions with beam energy of (a) 10 GeV from testbeam data for longitudinal profile (b) 200 GeV from testbeam data for radial profile, (c) 10 GeV from `QGSP_BERT_HP` physics list for longitudinal profile (d) 200 GeV from `QGSP_BERT_HP` physics list for radial profile. The blue, green and magenta curves show the contributions of "EM-core", "Hadronic core" and "Hadronic halo" components, respectively.

The simultaneous fit describes the longitudinal and radial shower profiles with a χ^2/NDF of approximately 1 for the studied pion beam energy above 10 GeV as shown in figure 7.37 for testbeam data and simulations.

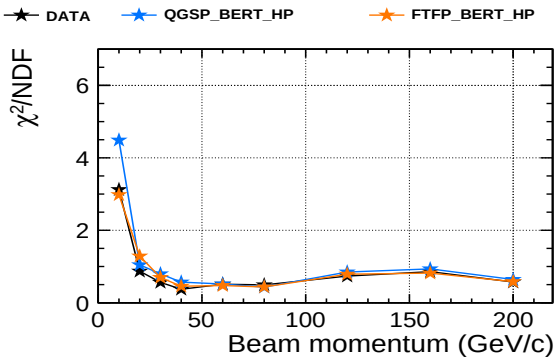


FIGURE 7.37: Energy dependence of χ^2/NDF obtained by simultaneous fitting of longitudinal and radial distribution by sharing energy E and f_{EM} parameter.

It is instructive to compare the fit parameters extracted from the simultaneous fit to test the level of agreement with the fit parameter obtained from the independent fits. For example, figure 7.38 shows the comparison of longitudinal fit parameters (α_s, β_s, α_l, β_l, z^{max}_{short}) to the fit parameters from the independent fit.

Fit Parameters Related to Longitudinal Shower Shape

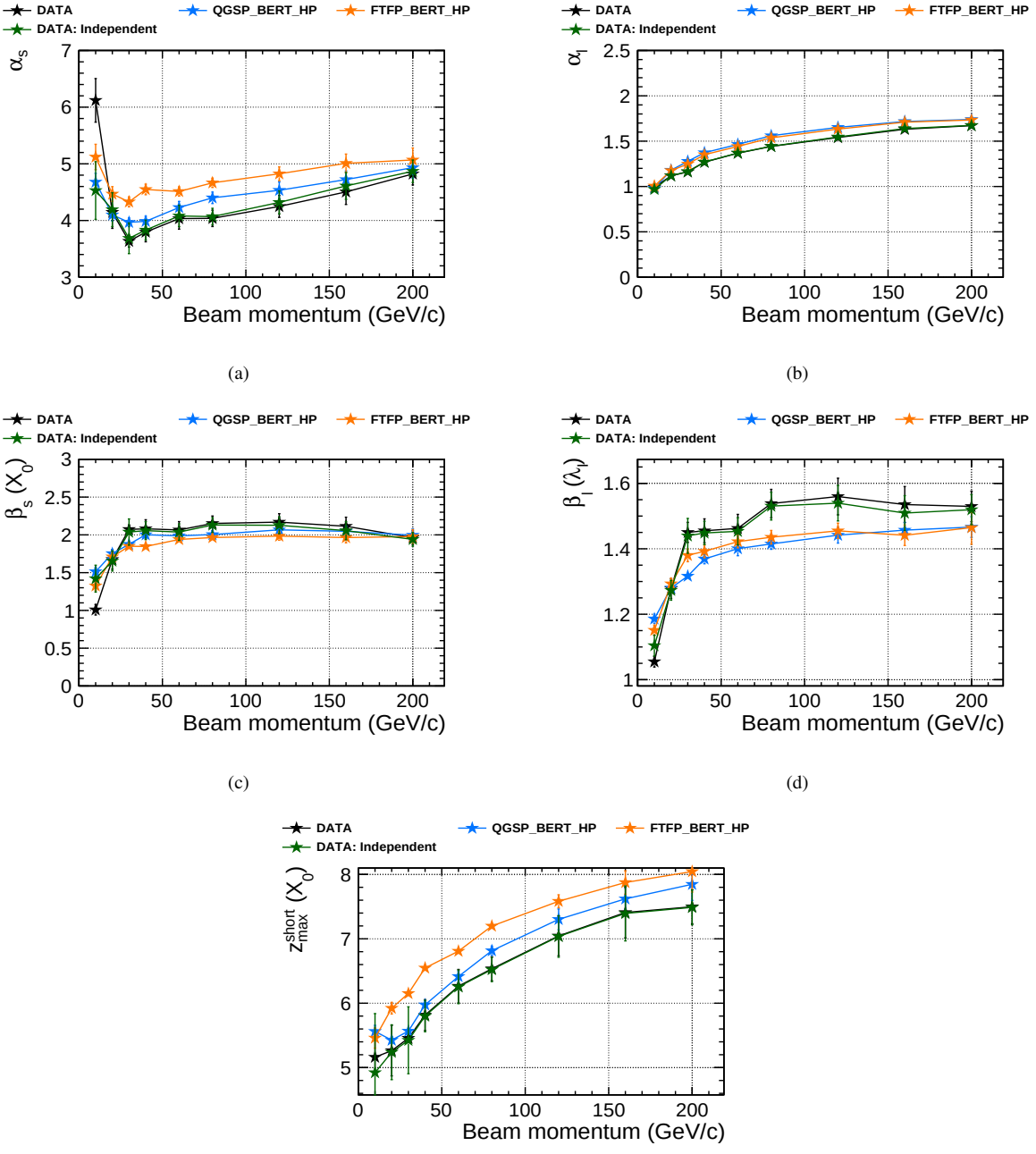


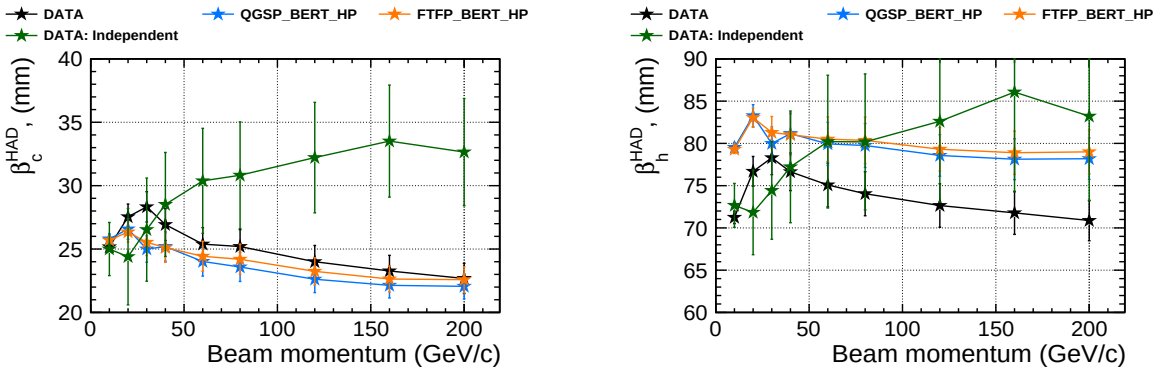
FIGURE 7.38: Energy dependence of longitudinal fit parameters for (a) α_s (b) α_l (c) β_s (d) β_l and (e) maxima of the short component z_{short}^{max}. The result of testbeam data from independent fits shown in green are compared to simultaneous fits for testbeam data in black and two physics lists, `QGSP_BERT_HP` (blue) and `FTFP_BERT_HP` (orange).

However, it appears that no significant change in trend is observed in the short and long parameters. The parameters from the simultaneous fit are comparable to data points from the independent fits. Although two physics lists show some discrepancies, their behaviour is still

fairly consistent. For example, the parameter with a maximum of the short component z_{short}^{max} agrees very close to the data points from independent fits. Moreover, the rising trend is still followed for data points at different energies.

Fit Parameters Related to Radial Shower Shape

Similarly, the fit parameters extracted from the radial showers are compared between the values from independent fit and the values from simultaneous fit. Figure 7.39(a) and 7.39(b) shows the hadronic core and hadronic halo component respectively.



(a) Energy dependence of radial parameters obtained from simultaneous fit

(b)

FIGURE 7.39: Energy dependence of radial parameters for (a) β_c^{HAD} and (b) β_h^{HAD} in comparison to data from independent fits in green and data from simultaneous fits and the different physics lists.

The fit result from the independent fit shows a different trend compared to the fit values from the simultaneous method. At the same time, the hadronic halo component shows a reasonable agreement within the uncertainties. On average, the core of the hadronic shower β_c^{HAD} is about 20-30 mm wide and the hadronic halo β_h^{HAD} is roughly a factor 3 larger for the beam energy between 10 GeV and 200 GeV.

Energy, Fraction and h/e Signal Ratio

Figure 7.40(a) shows the energy parameter obtained from the simultaneous fits in comparison to simulations and the values from the independent fit. Moreover, the correction parameter K shown in figure 7.40(b) gives the impression of the leakage fraction. An energy leakage of 20% or more is observed for energies above 80 GeV.

The resulting fit parameters are in reasonable agreement between testbeam data and simulations to the testbeam data from independent fits. More importantly, the simultaneous fit pulls the value of the parameter f_{EM} towards the longitudinal distribution such that the best fit value is obtained. The value of f_{EM} obtained from the simultaneous fit agrees closely with the f_{EM} obtained from the longitudinal fit, as shown in figure 7.40(c). Furthermore, the acquired f_{EM} value is found

to be around 10-40%. The hadronic core fraction shown in figure 7.40(d) remains to be less dependent on energy.

Figure 7.40(e) shows the h/e signal ratio using Eq. 7.7 is plotted for parameters obtained from simultaneous fit. The h/e signal ratio is found to be close to the values obtained from the independent fit to longitudinal profile, at the same time in good agreement between testbeam data and simulations within the uncertainty.

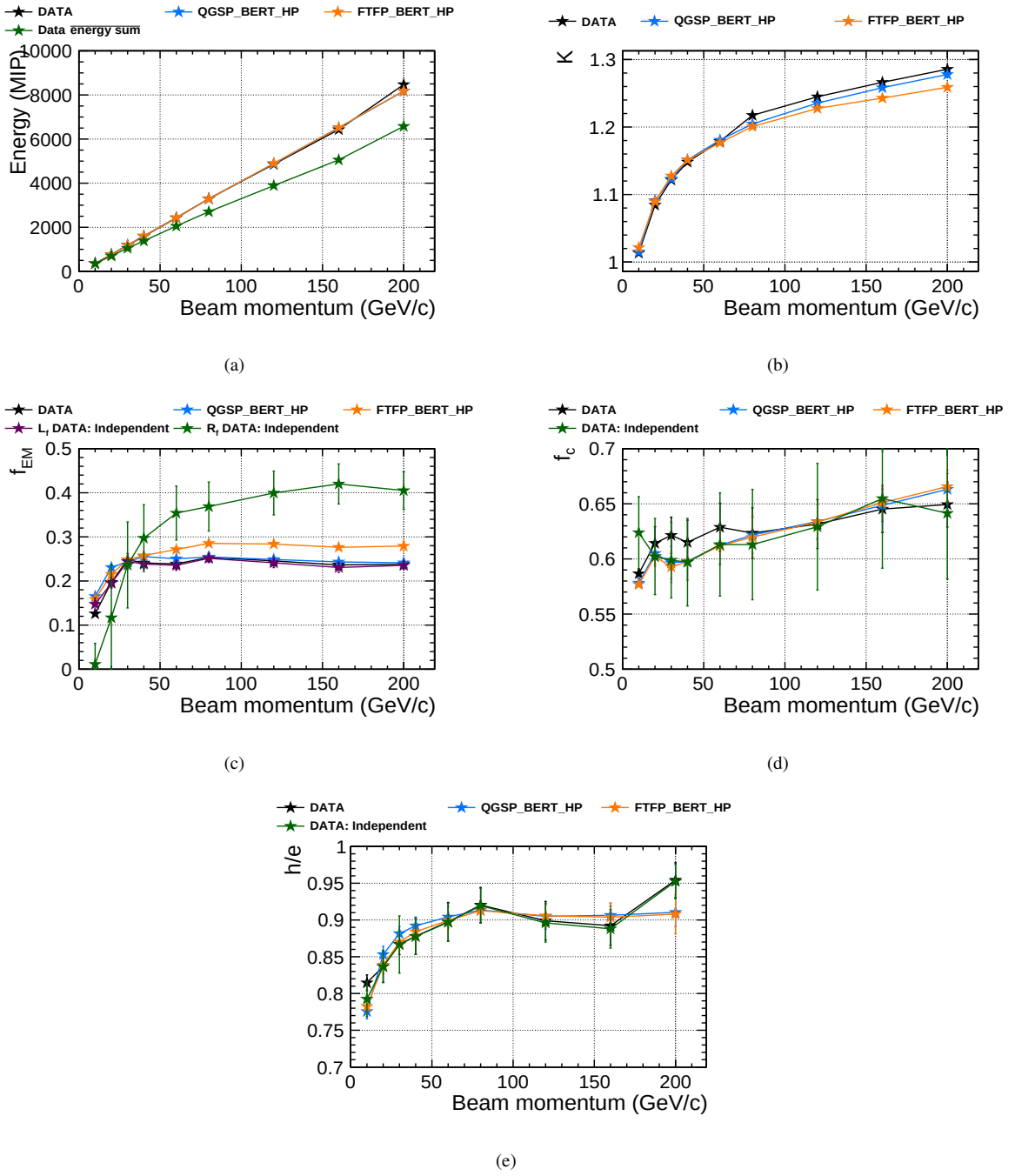


FIGURE 7.40: Energy dependence of parameters from the physics lists using `QGSP_BERT_HP` (blue) and `FTFP_BERT_HP` (orange), in comparision to testbeam data (black) from simultaneous fits and independent fits (green) for (a) Energy, (b) leakage fraction, (c) electromagnetic fraction, (d) hadronic core fraction and (e) h/e signal ratio.

The correlation of nine free parameters obtained from simultaneous fit is shown in figure 7.41. There exists a strong correlation between the radial parameters, β_c^{HAD} and β_h^{HAD} with a positive correlation. At the same time, there exists a moderate correlation between the energy parameter E and fraction f_{EM}. Looking more closely, the alpha's and beta's show an anti-correlation with each other.

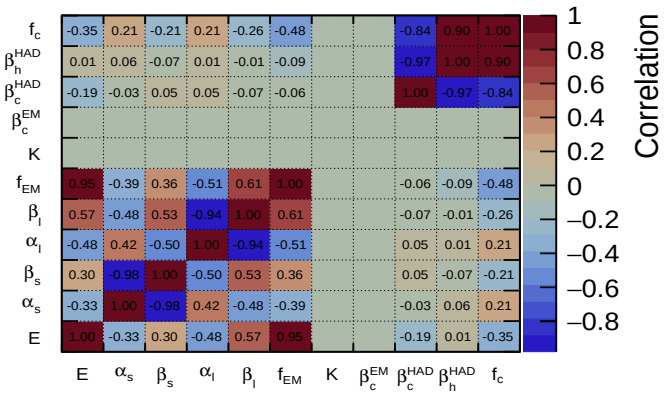


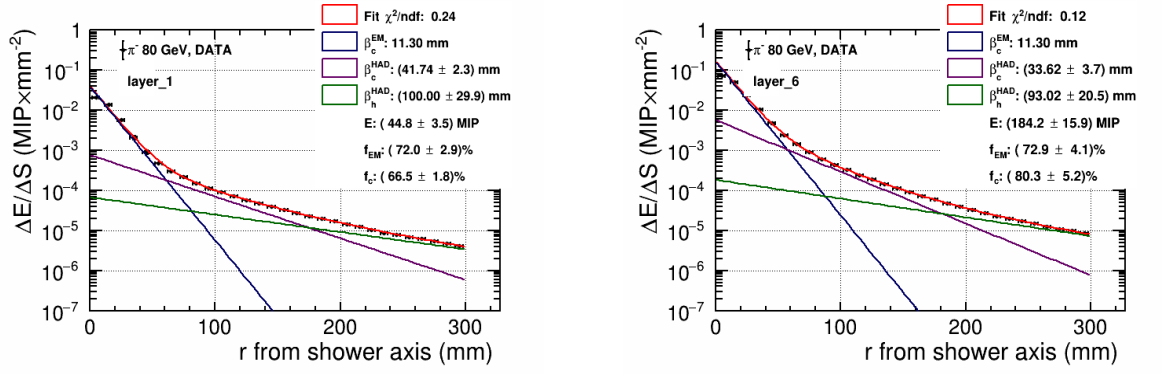
FIGURE 7.41: Correlation of fit parameters obtained from simultaneous fit to longitudinal and radial distribution by sharing energy and fraction parameter obtained for the testbeam data with pion shower of beam energy of 200 GeV. The colour scale represents the strength of correlation.

7.2.5 Two-Dimensional Fitting of Radial Profiles Layer-wise

To understand the change in behaviour of radial shapes, this would require looking for the dependence of radial shower shapes along the depth of the calorimeter. While dealing with radial shapes through the calorimeter depth, the parameter K in Eq. 7.19 does not play a role as there is no such "energy leakage" in a single layer. As a result, this would result in a parametrization without the K term as given below:

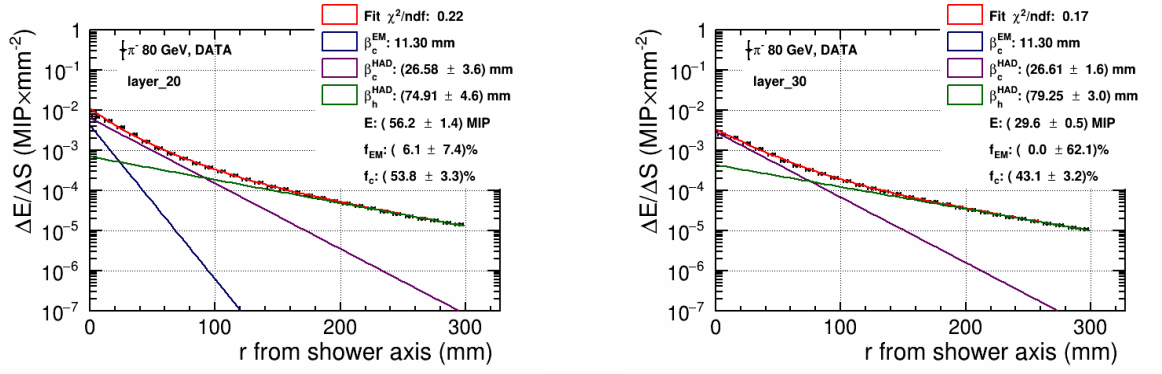
$$\frac{\Delta E}{\Delta S}(r) = \frac{E}{2\pi}\left\{ f_{EM}\cdot\frac{e^{\frac{-r}{\beta_c^{EM}}}}{(\beta_c^{EM})^2} + (1-f_{EM})\left(f_c\cdot\frac{e^{\frac{-r}{\beta_c^{HAD}}}}{(\beta_c^{HAD})^2} + f_h\cdot\frac{e^{\frac{-r}{\beta_h^{HAD}}}}{(\beta_h^{HAD})^2}\right)\right\} \tag{7.21}$$

The above parametrization is applied for radial shapes along the calorimeter depth for layers beyond the shower start using a fixed value of β_c^{EM} from the knowledge of EM showers. An example of fit performed to the following four parts of the radial shower using testbeam data with a beam energy of 80 GeV are: first layer at the beginning of the shower as shown in figure 7.42(a), a layer at the shower maximum, which is layer number 6 from shower start is shown in figure 7.42(b), an intermediate layer which is layer number 20 from shower start is shown in figure 7.42(c), and a layer nearly at the end of the shower which is layer number 30 from shower start is shown in figure 7.42(d).



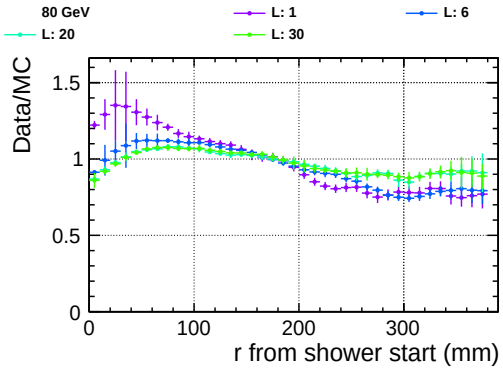
(a)

(b)



(c)

(d)



(e)

FIGURE 7.42: Fits performed using Eq. 7.21 with fixed β_c^{EM} for radial shower shapes at various depths with beam energy at 80 GeV using testbeam data (a) at the beginning of the shower, (b) at the maximum of the shower, (c) at an intermediate layer (d) and layer at the end of the shower. (e) Shows the ratio between data and QGSP BERT HP for 80 GeV pion shower.

In general, a fair agreement between testbeam data and physics lists for 80 GeV pion shower is found to be within ∼20% for layers beyond layer one, as shown in figure 7.42(e). The first layer showing large disagreement is due to the uncertainty in the shower start identification.

Furthermore, the dependence of fit parameters obtained from radial fits layer-wise are compared. Figure 7.43 shows the parameters obtained for 80 GeV pion energy compared between

testbeam data and the physics lists. On a first attempt of the fit, the hadronic slopes (β_c^{HAD} and β_h^{HAD}) show no large dependence through the layers as well as this behaviour is agreed by both the physics lists. Meanwhile, the fraction parameters, f_{EM} and f_c as well as the energy parameter E show dependence through the layers, with an increase up to the shower maximum and falling steadily afterwards. This behaviour is well reproduced by the simulations as shown in figure 7.43(c), 7.43(d) and 7.43(e). Figure 7.43(f) shows the χ^2/NDF value which describes the radial shape behaviour to a reasonably good extent.

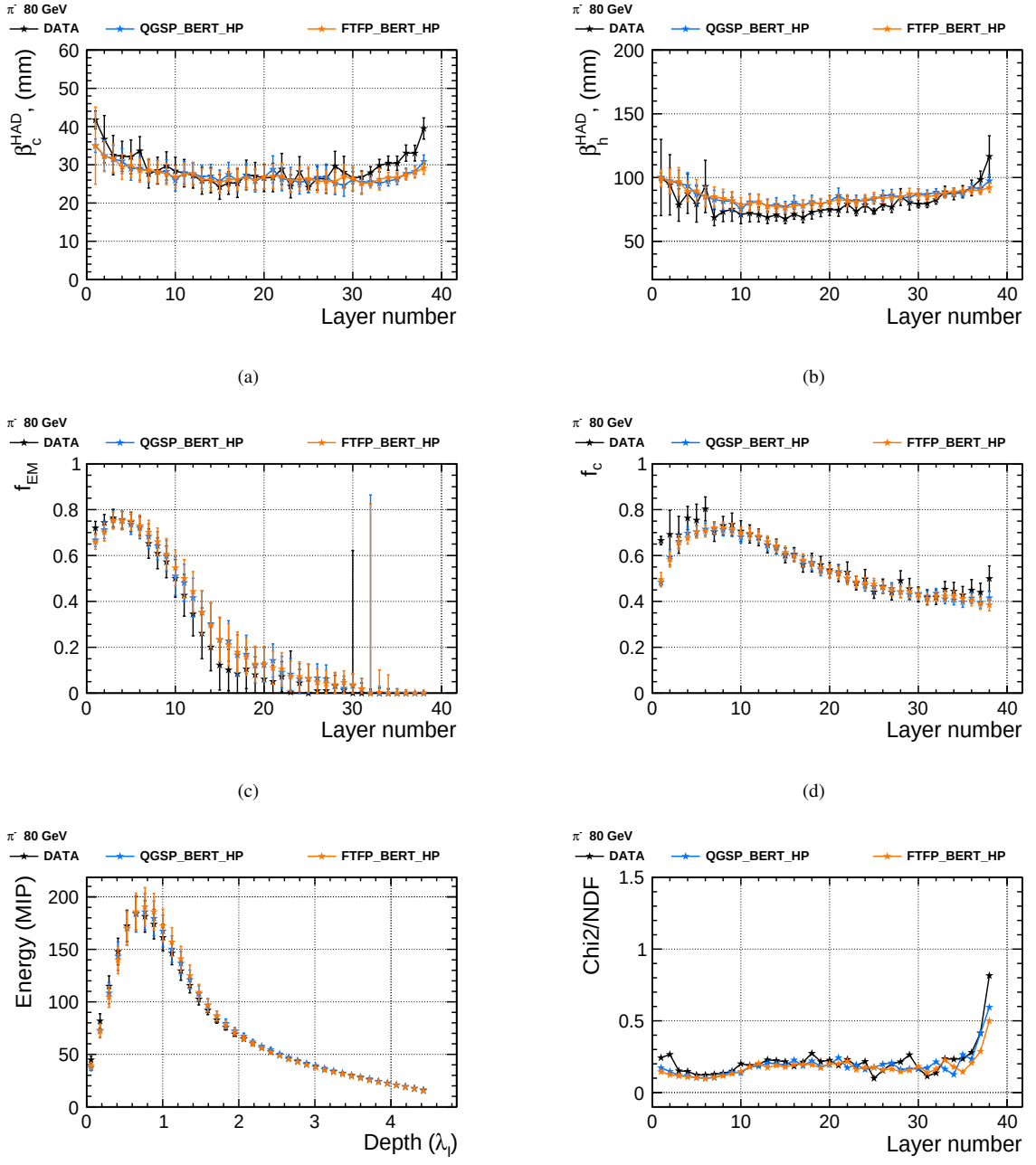


FIGURE 7.43: Dependence of radial parameters for increasing shower depth while keeping the radial slope parameters (β_c^{HAD} and β_h^{HAD}) free during the fit. In comparison to testbeam data and simulations using `QGSP_BERT_HP` and `FTFP_BERT_HP` physics list using pion beam energy at 80 GeV.

It was therefore decided to fix the parameter values of β_c^{HAD} and β_h^{HAD} to the values from simultaneous fits. As a result, this reduces to a total of four free parameters in the parametrization as described in Eq. 7.21. An example of the fits performed on the same four chosen layers (1, 6, 20 and 30) are shown in figure 7.44 with fixed beta's: β_c^{EM}, β_c^{HAD} and β_h^{HAD}.

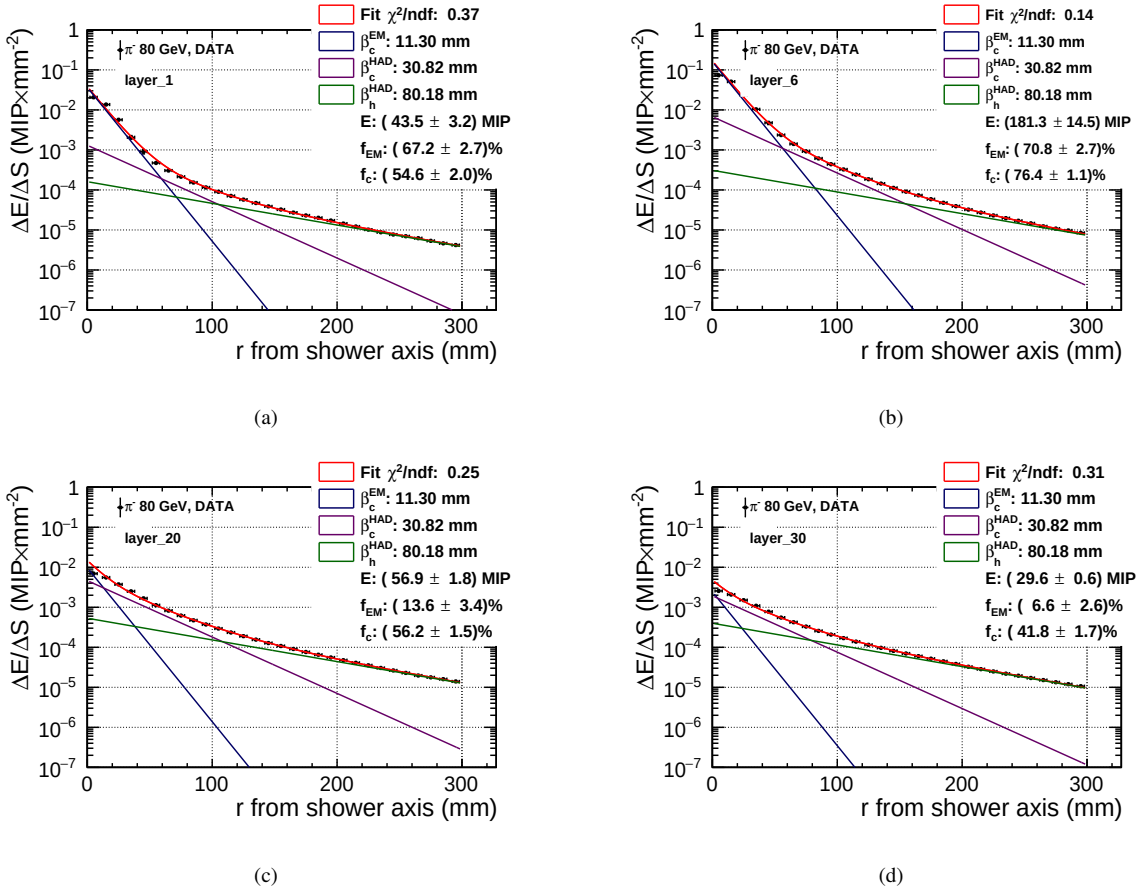


FIGURE 7.44: Fits performed using Eq. 7.21 using fixed β_c^{EM}, β_c^{HAD} and β_h^{HAD} on radial shower shapes at various depths for beam energy at 80 GeV using testbeam data (a) at the beginning of the shower, (b) at the maximum of the shower, (c) at an intermediate layer (d) and layer at the end of the shower.

With the fixed beta's in the radial fit function, the free parameters now obtained are the energy and fractions as shown in figure 7.45, described with a reasonable value of χ^2/NDF for testbeam data with a beam energy of 80 GeV.

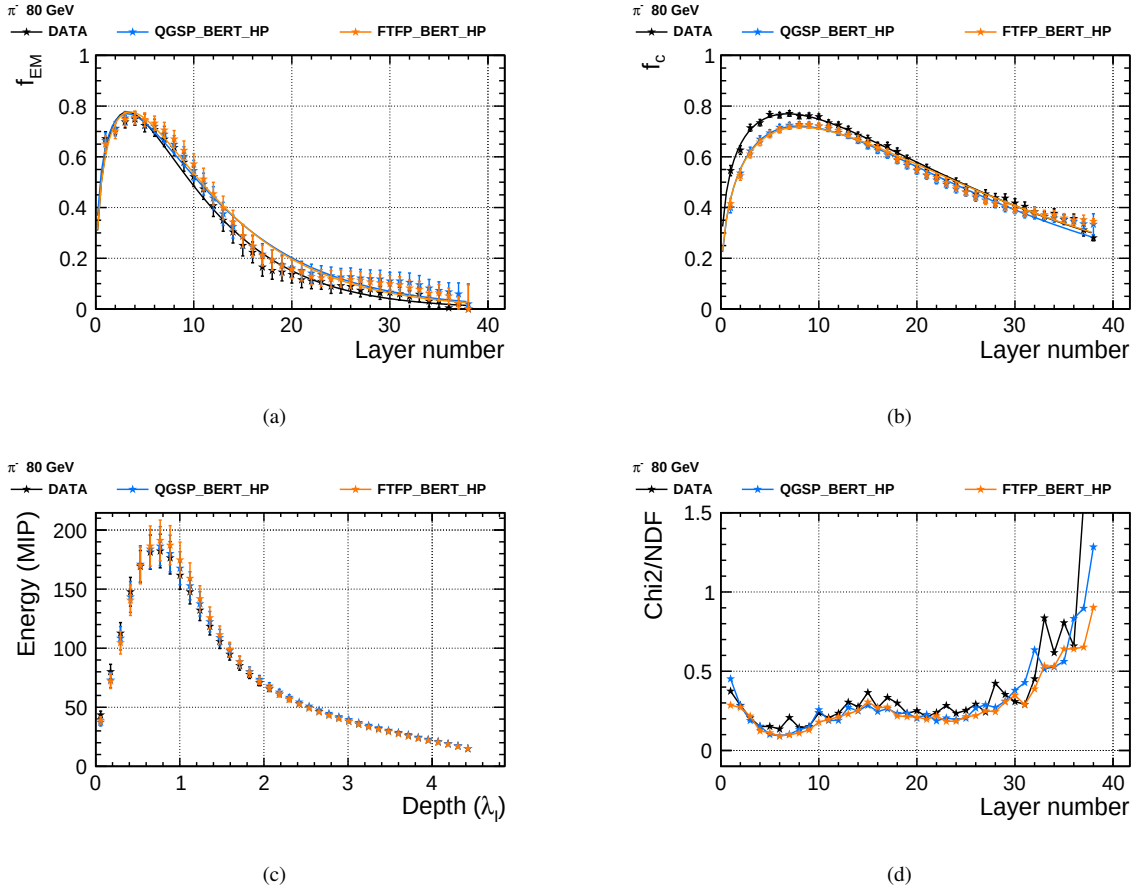
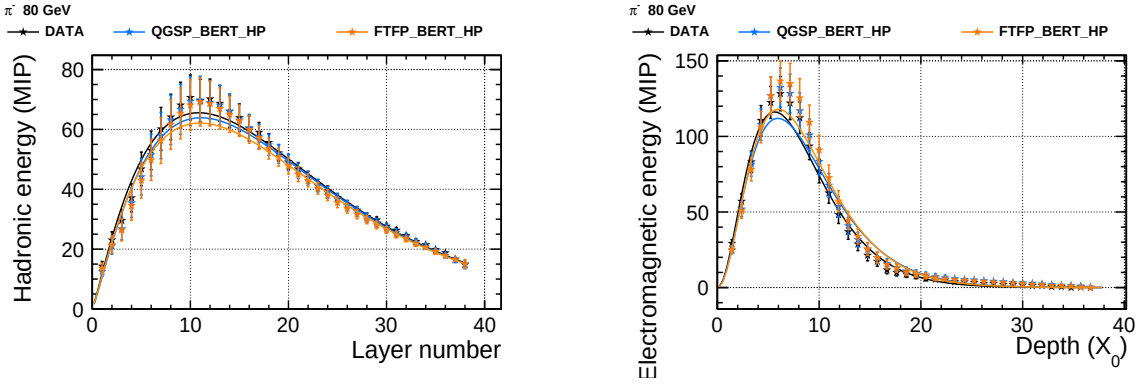


FIGURE 7.45: Dependence of radial parameters for increasing shower depth while keeping the radial slope parameters (β_c^{HAD} and β_h^{HAD}) fixed during the fit based on the values from simultaneous fits. In comparison to testbeam data and simulations using `QGSP_BERT_HP` and `FTFP_BERT_HP` physics list using pion beam energy at 80 GeV.

As we put an assumption in the longitudinal fit that the short part relates to the electromagnetic component and the long part relates to the "pure" hadronic component of the hadron shower, which means the total energy could be split into electromagnetic energy and hadronic energy. As an ansatz, these energy parts should individually be described using a Gamma function based on the Eq. 7.1. Figure 7.46(a) and 7.46(b) shows the fit performed using the Gamma distribution to the electromagnetic and hadronic energy. In addition, the same functional form is used to describe both the fractions (f_{EM} and f_c). It is observed that the Gamma distribution can reasonably describe the energy and fraction behaviour.



(a) (b)

FIGURE 7.46: Hadronic and electromagnetic energy for 80 GeV pion shower using testbeam data and simulations using `QGSP_BERT_HP` and `FTFP_BERT_HP` physics list. The distributions are described with a fit using a Gamma function and represented with a smooth curve.

Moreover, a physics check is done with an idea to test the assumption of a simplified picture of the hadronic shower consisting of a core part that comes from a shower initiated by neutral pions and a later developing halo component that is from "truly" charged hadrons. To prove this, the electromagnetic energy of the longitudinal shower profiles is compared to the longitudinal profiles of true electron showers.

An example in figure 7.47 shows a direct comparison of longitudinal shower shape from the EM part of pion shower of beam energy 120 GeV to the longitudinal shower profile of electron with beam energy at 40 GeV. The two distributions are parameterized using the Gamma distribution.

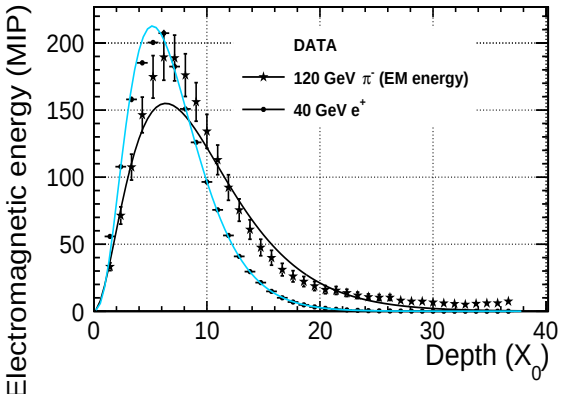


FIGURE 7.47: Distribution of electromagnetic energy with beam energy at 120 GeV using testbeam data extracted from pion shower in comparison to 40 GeV pure electrons. The solid lines are the fit performed to the longitudinal shower shapes using Eq. 7.1.

It is observed that the maximum position of the pion shower is shifted to a higher value, and there exists a long tail in the pion shower, which is clearly due to electromagnetic energy which is

deposited by later shower generations and is deeper in the calorimeter. Moreover, we observe that the simple Gamma function does not fully describe the data. However, as a first approximation, it looks reasonably good.

Furthermore, the fit parameters from the Gamma distribution obtained from electrons and pions showers for all energies are shown in figure 7.48. The values for pion energy are plotted at one-third of its corresponding beam energy. This relation reflects the fact that about one-third of all pions produced in hadronic cascade are neutral pions, which immediately decay into two photons. These, in turn, induce electromagnetic showers. Therefore, the intrinsic electromagnetic component of each hadronic shower is compared to pure electrons.

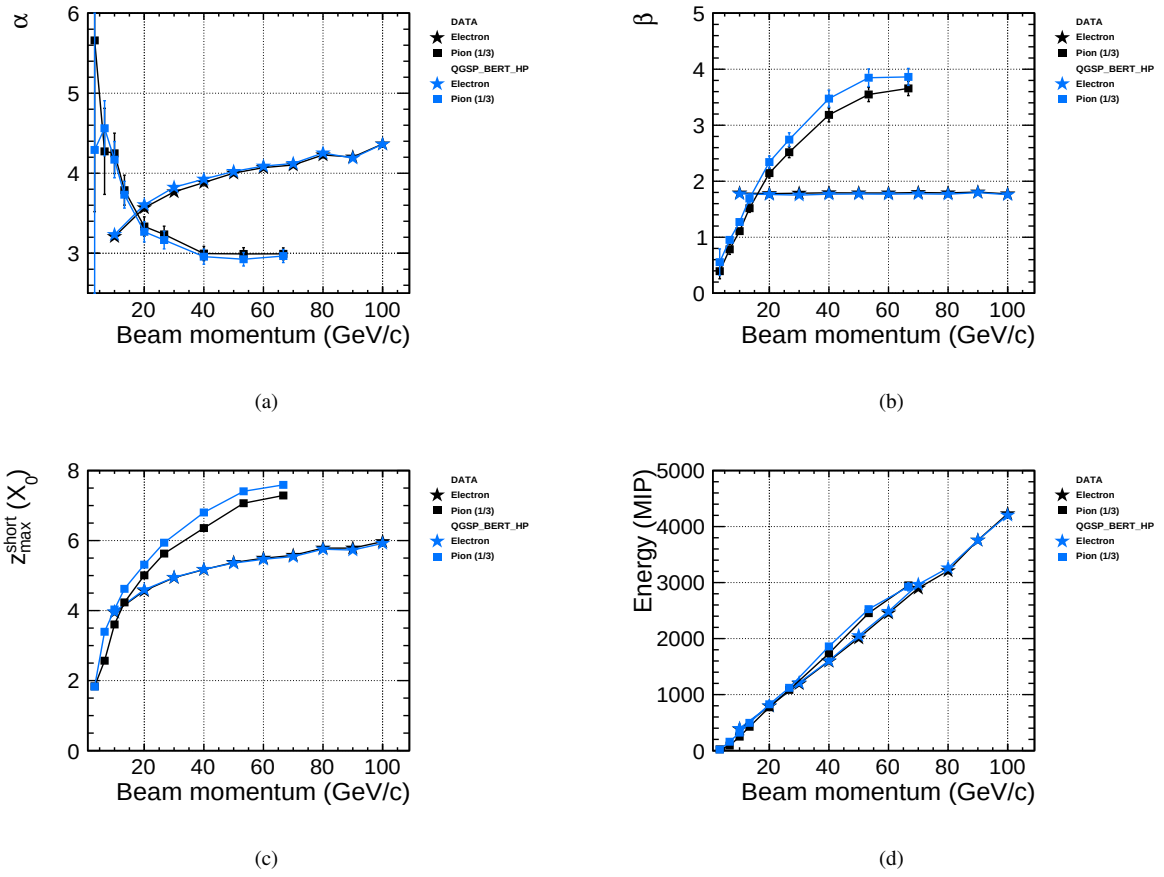


FIGURE 7.48: Energy dependence of fit parameters extracted from the fit using Eq. 7.1 to the electromagnetic energy from pion showers and in comparison to electron showers. The pion energy is plotted to one-third of its corresponding energy and compared to electrons.

The parameter shows a reasonable agreement for lower energies. However, there is more disagreement in the shape behaviour between pion and electron showers with increasing energies. This discrepancy is clearly seen by the parameter from the maximum of the short component (z_{max}^{short}). In addition, the alpha parameter shows an opposite trend for energies above 10 GeV electrons and 30 GeV pions. It shows the limits of our simple picture, which assumes that the short component looks like an electromagnetic shower with ≈1/3 of the beam energy.

7.2.6 Three-Dimensional Modelling of Hadronic Showers

A three-dimensional parameterization describing the hadronic shower shapes in the AHCAL is discussed in this section. This is the first time performed using a highly granular calorimeter.

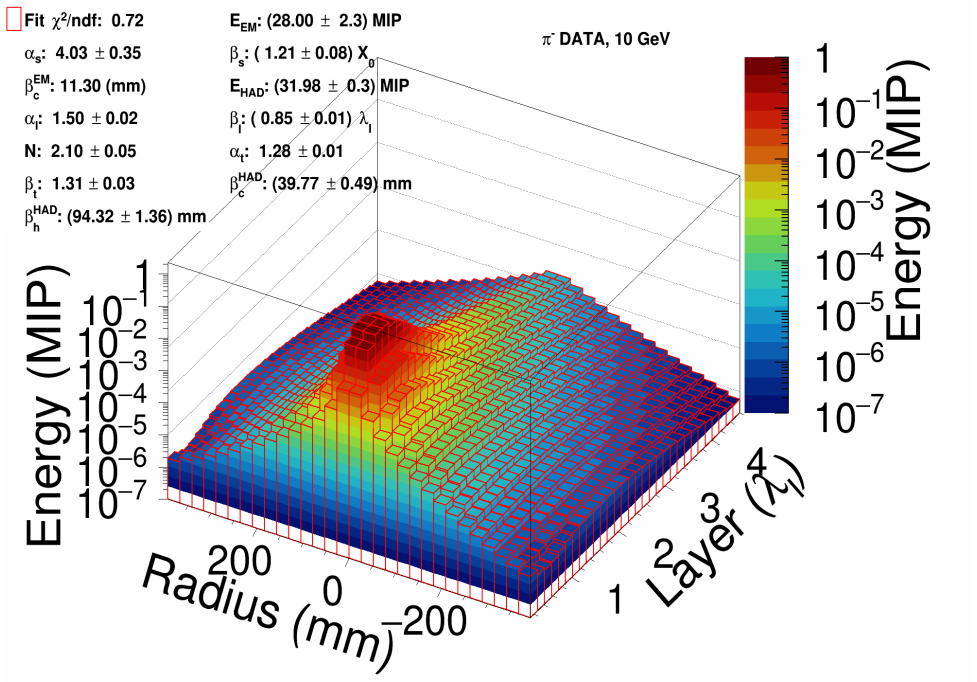
As explained earlier, from the assumption that the total energy E can be expressed as the sum of electromagnetic energy $E_{EM}(z)$ and hadronic energy $E_{HAD}(z)$. These, in addition, to the fraction parameter $f_c(z)$, are reasonably explained with one Gamma function of the form as given in Eq. 7.1. Therefore, the parameters E, f_{EM}, f_c and f_h from Eq. 7.21 can be expressed in depth z using the Gamma function. Then the Eq. 7.21 will arrive at a three-dimensional parametrization described in r and z and is written as:

$$\frac{\Delta E}{\Delta A}(r,z) = \frac{1}{2\pi}\Bigg\{\frac{E_{EM}}{\Gamma(\alpha_s)}\cdot\left(\frac{z}{\beta_s}\right)^{\alpha_s-1}\cdot\frac{e^{\frac{-z}{\beta_s}}}{\beta_s}\cdot\frac{e^{\frac{-r}{\beta_c^{EM}}}}{(\beta_c^{EM})^2}+\frac{E_{HAD}}{\Gamma(\alpha_l)}\cdot\left(\frac{z}{\beta_l}\right)^{\alpha_l-1}\cdot\frac{e^{\frac{-z}{\beta_l}}}{\beta_l}$$
$$\left(\frac{N}{\Gamma(\alpha_t)}\cdot\left(\frac{z}{\beta_t}\right)^{\alpha_t-1}\cdot\frac{e^{\frac{-z}{\beta_t}}}{\beta_t}\cdot\frac{e^{\frac{-r}{\beta_c^{HAD}}}}{(\beta_c^{HAD})^2}+\left(1-\left(\frac{N}{\Gamma(\alpha_t)}\cdot\left(\frac{z}{\beta_t}\right)^{\alpha_t-1}\cdot\frac{e^{\frac{-z}{\beta_t}}}{\beta_t}\right)\right)\cdot\frac{e^{\frac{-r}{\beta_h^{HAD}}}}{(\beta_h^{HAD})^2}\right)\Bigg\} \tag{7.22}$$

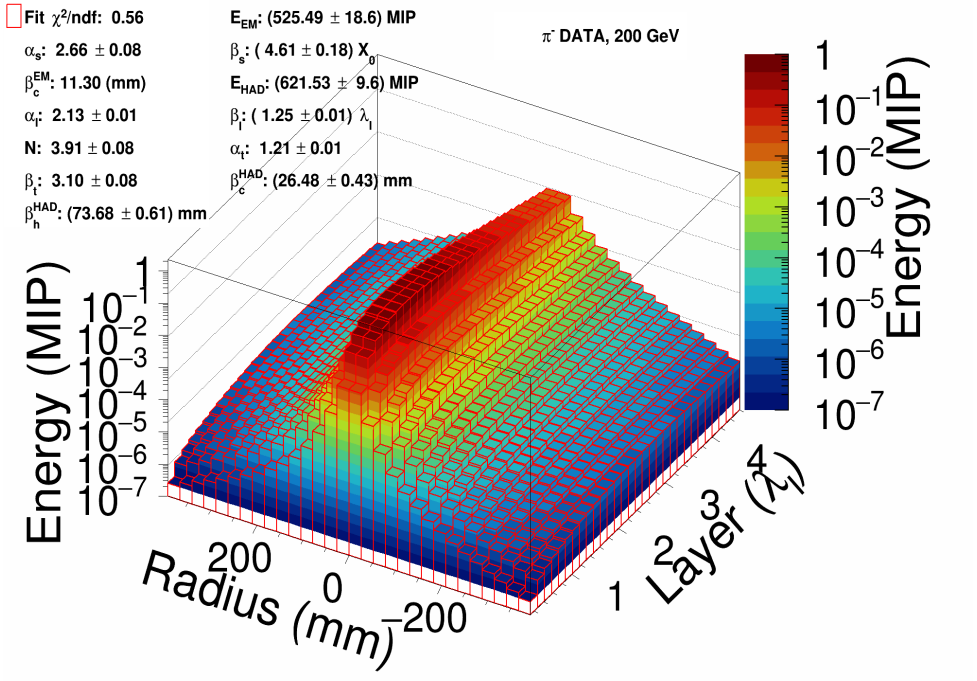
The above three-dimensional parametrisation consists of three parts: electromagnetic core with a z dependent function and an exponent in r. Furthermore, the hadronic component is written as a sum of the hadronic core and the hadronic halo with z times the function of r. The parametrisation is then the sum of two hadronic components and the electromagnetic component, which are factorised in r and z.

The parameters involved in the above parametrization consists of three radial slopes – β_c^{EM}, β_c^{HAD} and β_h^{HAD} are expressed in units of millimetres. The normalization from the three Gamma functions includes – the energy from the electromagnetic core E_{EM}(MIP), the energy from the hadronic component E_{HAD}(MIP), and the normalization from the hadronic core N. In addition, the parameters from the three Gamma functions – α_s, β_s, α_l, β_l, α_t and β_t make altogether twelve parameters that would determine the three-dimensional shape of an average hadronic shower.

To validate the three-dimensional model, if it could explain the shape of an average hadronic shower, it is applied to pion showers using the testbeam data and two physics lists for all the available energies. An example of the three-dimensional model applied to pion showers using testbeam data is shown in figure 7.50(a) at a beam energy of 10 GeV and figure 7.50(b) at a beam energy of 200 GeV. The energy density per event is plotted versus the radius of the radial ring and its longitudinal position with respect to the first hard interaction.

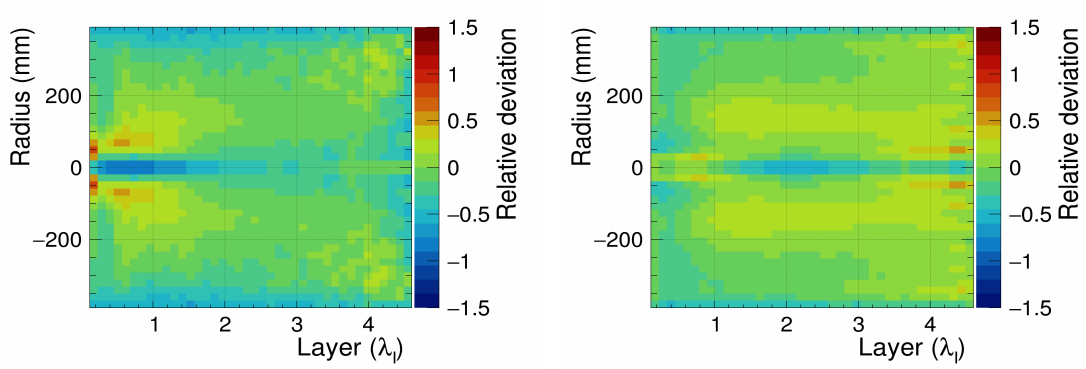


(a)



(b)

FIGURE 7.49: Three-dimensional parametrization applied to the pion showers using the AHCAL testbeam data with beam energy at (a) 10 GeV (b) 200 GeV. The red solid lines shows the parametrization applied using Eq. 7.22 describing the shower shapes.



(a) (b)

FIGURE 7.50: The residual of the shower shape from testbeam data to the parametrization for the beam energy at (a) 10 GeV and (b) 200 GeV.

An evident broadening of the shower is observed as the shower develops through the calorimeter. The shape distribution of hadron showers in the AHCAL is well described by the three-dimensional parameterization, which has parameters that allow one to determine the 3D shape of an average hadronic shower in the AHCAL. The relative deviation of fit to testbeam data for beam energies at 10 GeV and 200 GeV is shown in figure 7.50(a) and 7.50(b), respectively. Most regions are described much better, while those close to the shower core are less described, especially at the shower start for small energies.

The model performs well for the entire pion energies from 10 GeV to 200 GeV, using testbeam data and physics lists. The performance of fit is shown in figure 7.51.

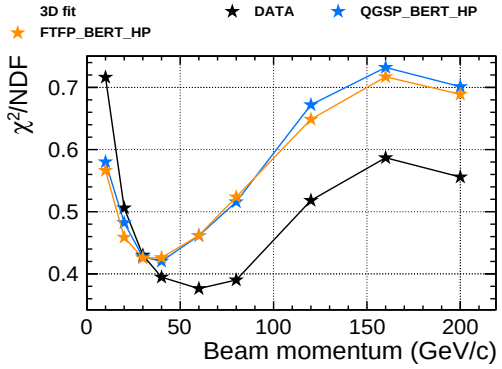
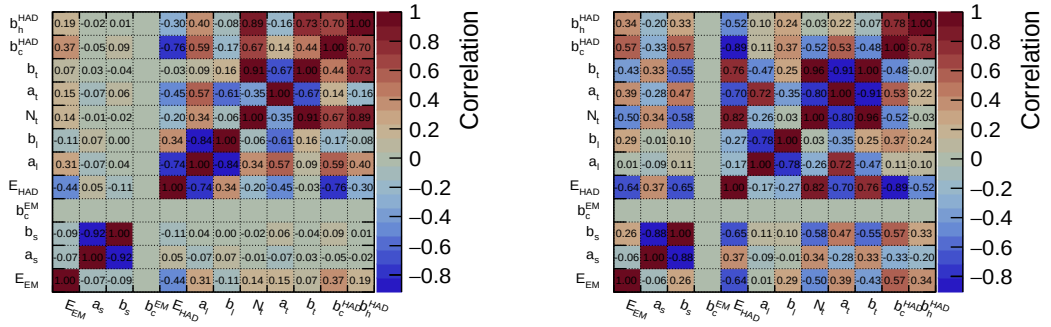


FIGURE 7.51: χ^2/NDF as a function of the beam energy between 10 GeV to 200 GeV pion showers using testbeam data and physics lists.

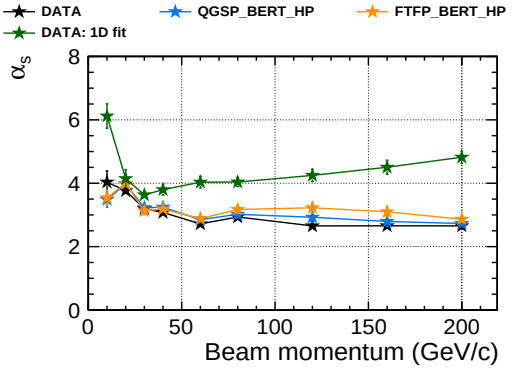
The relationship between different parameters from the three-dimensional parametrization is shown in the correlation plot for a beam energy of 10 GeV in figure 7.52(a) and 200 GeV in figure 7.52(b) for pion showers from testbeam data. The α and β parameters show anti-correlation, while the β's from the hadronic component are highly correlated and have a positive correlation. Moreover, it is also seen that there exists a correlation between the β parameter from the hadronic core and the hadronic energy.



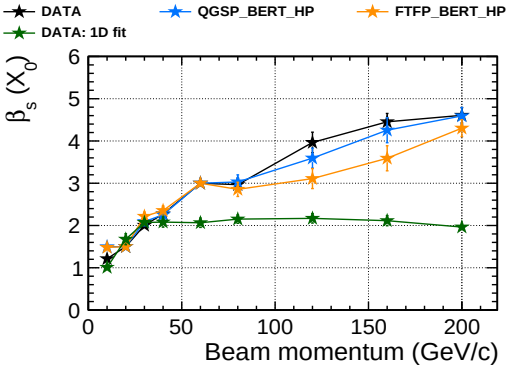
(a) (b)

FIGURE 7.52: Correlation between parameters obtained from three-dimensional fits using testbeam data from pion shower with beam energy at (a) 10 GeV and (b) 200 GeV.

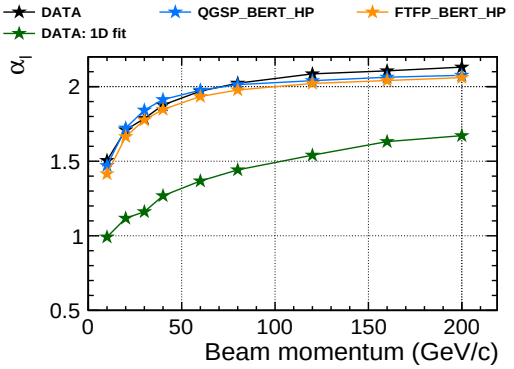
Furthermore, a comparison of fit parameters is done to check for the differences between the parameters obtained from the three-dimensional fits and the parameters from the simultaneous fits of longitudinal and radial shower shapes as a function of pion beam energy. The short parameters (α_s and β_s) is shown in figure 7.53(a) and 7.53(b) respectively. While the long parameters (α_l and β_l) is shown in figure 7.53(c) and 7.53(d), respectively. The parameters describing the maximum from the short (z^{max}_{short}) and long (z^{max}_{long}) components is shown in figure 7.53(e) and 7.53(f), respectively. The hadronic core and hadronic halo parameters are shown in figure 7.53.


DATA
QGSP_BERT_HP
FTFP_BERT_HP
DATA: 1D fit
α_s
Beam momentum (GeV/c)

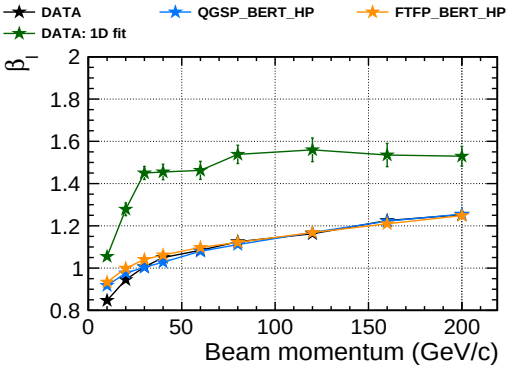
(a)


DATA
QGSP_BERT_HP
FTFP_BERT_HP
DATA: 1D fit
β_s (X_0)
Beam momentum (GeV/c)

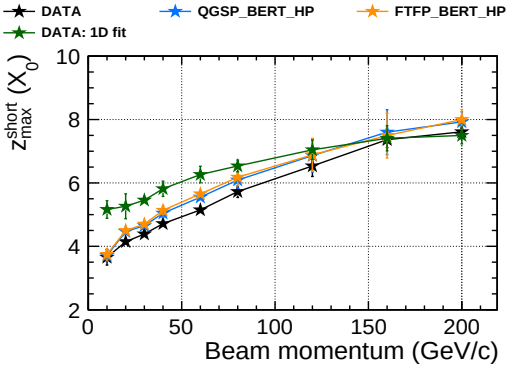
(b)


DATA
QGSP_BERT_HP
FTFP_BERT_HP
DATA: 1D fit
α_l
Beam momentum (GeV/c)

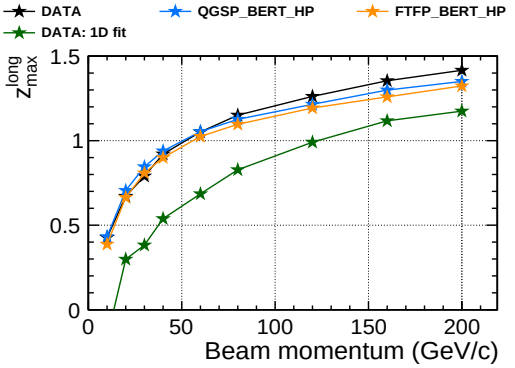
(c)


DATA
QGSP_BERT_HP
FTFP_BERT_HP
DATA: 1D fit
β_l
Beam momentum (GeV/c)

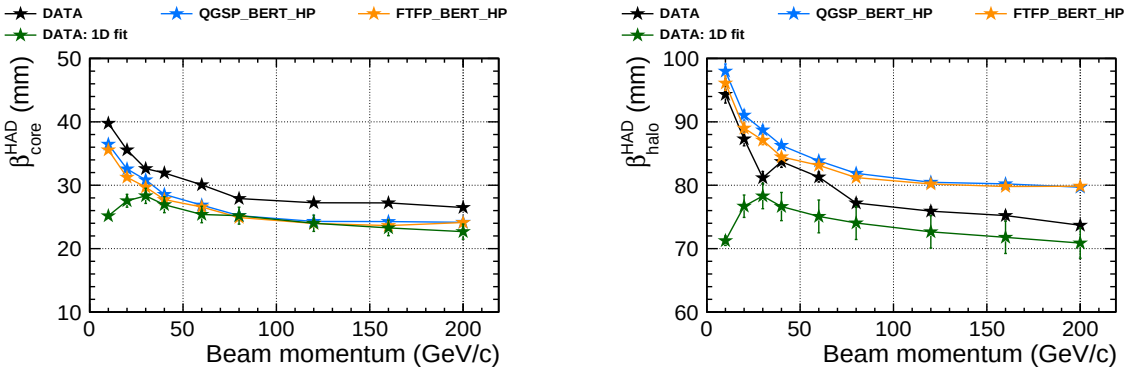
(d)


DATA
QGSP_BERT_HP
FTFP_BERT_HP
DATA: 1D fit
z_{max}^{short} (X_0)
Beam momentum (GeV/c)

(e)


DATA
QGSP_BERT_HP
FTFP_BERT_HP
DATA: 1D fit
z_{max}^{long}
Beam momentum (GeV/c)

(f)



(g) (h)

FIGURE 7.53: The energy dependence of fit parameters obtained from three-dimensional fit using testbeam data and the two physics lists in comparison to the fit parameters using the testbeam data from one dimensional simultaneous fits to longitudinal and radial shower shapes.

The majority of the parameters obtained from three-dimensional fits show energy dependence. Moreover, the parameter's shape behaviour is undoubtedly different from parameters from one-dimensional fits. This difference is obtained as more parameters are involved, which gains additional information not present in the one-dimensional fit. Nevertheless, the three-dimensional parametrization proves to describe the hadronic shower shape for all the available pion energies with a reasonably good value of $\chi^2/\mathrm{NDF} \approx 1$.

7.2.7 Neutron and EM tagging

The motivation of this study aims to understand the hadronic shower structure. The major contribution to the electromagnetic component of hadron showers is from the decay of π^0 and η particles. In contrast, the non-electromagnetic component is connected to neutrons.

As discussed previously in the longitudinal and radial fit functions, we relate the short or core and long or halo components, respectively, to the showers originating from the electromagnetic and pure hadronic components of the showers. The distinction between the two components and the behaviour can be verified by exploiting the Monte Carlo truth information. In addition, this study provides direct confirmation of the energy deposition obtained from the GEANT4. Furthermore, we have previously observed that by specifically looking into the longitudinal profile and comparing the electromagnetic component of pion shower to electrons. It showed some inconsistency in describing with Gamma fit function. Therefore, this behaviour needs to be investigated.

Generally, in testbeam data, we can investigate indirectly if or how well the short or core component agrees to an EM-shower. However, in the simulated data, we can attempt to access the information on particles causing the shower by looking into the history of the particle shower.

A dedicated algorithm is written for this purpose to investigate the particle hits originating from neutrons or EM-like particles. A relation between reconstructed hit and simulated hit is built to

extract all the Monte-Carlo particles in the shower and their properties such as energy, momentum, PDG and time stamps. It gives the relation of all the Monte Carlo particles contributing to that hit. The concept of the algorithm is to search upwards in the Monte Carlo tree, as illustrated in the sketch shown in figure 7.54. It searches from the end of the MC particle hit that deposited its energy in the active layer and points out if the corresponding hit particle has a neutron in its ancestor. Likewise, it also searches if the hit particle has π^0 or η in its ancestor. Then the quantity *Neutron tagged* or *EM tagged* is defined as the energy-weighted contributions of Monte Carlo particles with a neutron or EM particles (π^0 or η) in its ancestor compared to all contributions to the simulated hit. Mathematically written as:

$$Neutron\ tagged\ or\ EM\ tagged(hit) = \frac{\sum_{contrib:Neutron\ or\ EM Ancestors} Energy(contrib)}{\sum_{contrib} Energy(contrib)} \tag{7.23}$$

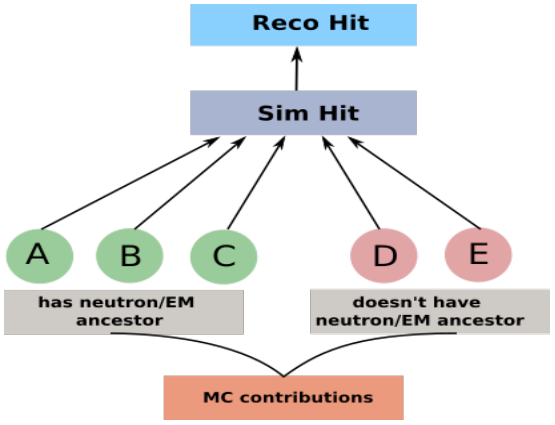
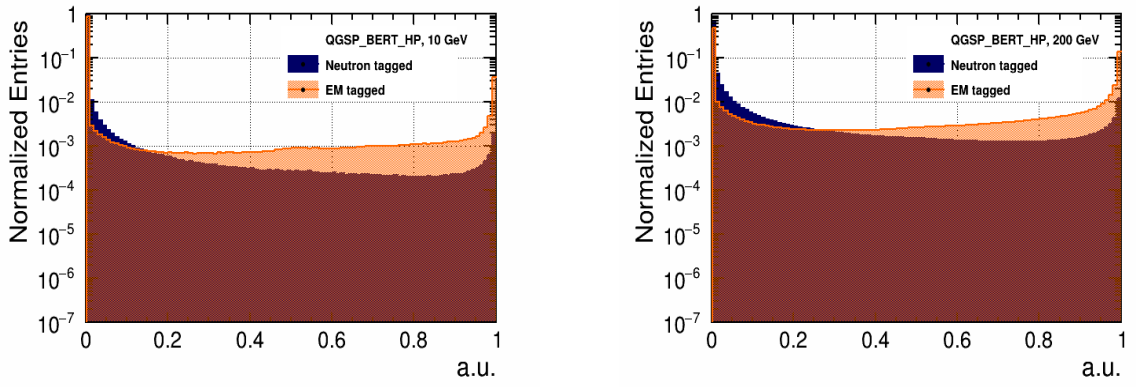


FIGURE 7.54: The schematic of the MC particle tree with MC contributions from the simulated hit is related to the reconstructed hit. Green circles represent the MC particle having a neutron or EM particle in its ancestor. The red circles represent those MC particles that do not have a Neutron or EM particle in their ancestors.

Based on the Eq. 7.23, the distributions of Neutron content and EM content from pion showers are obtained using the `QGSP_BERT_HP` physics list for beam energy at 10 and 200 GeV as shown in figure 7.55(a) and figure 7.55(b), respectively. The quantity Neutron tagged and EM tagged range between 0 and 1 by the construction of the formula. A large population close to one in the distributions means only particles with Neutron or EM ancestors contributed to the hit.



(a) (b)

FIGURE 7.55: Distributions of Neutron tagged (Blue) and EM tagged (Orange) obtained based on Eq. 7.23 using `QGSP_BERT_HP` physics list for pion showers with a beam energy at (a) 10 GeV and (b) 200 GeV.

At first, a validation is done using neutron tagged and EM tagged hits by selecting events larger than 0.9. These events are mainly contributed by neutrons or EM particles. Figure 7.56 shows the hit energy deposition and hit time using neutron tagged and EM tagged hits for 10 and 200 GeV pion showers. The test, confirms the physics that the nuclear processes happen at long timescales up to several microseconds after the initial collision. At the same time, neutrons deposit low energy compared to electrons that deposit high energy relatively at a concise time scale, roughly within 30 ns.

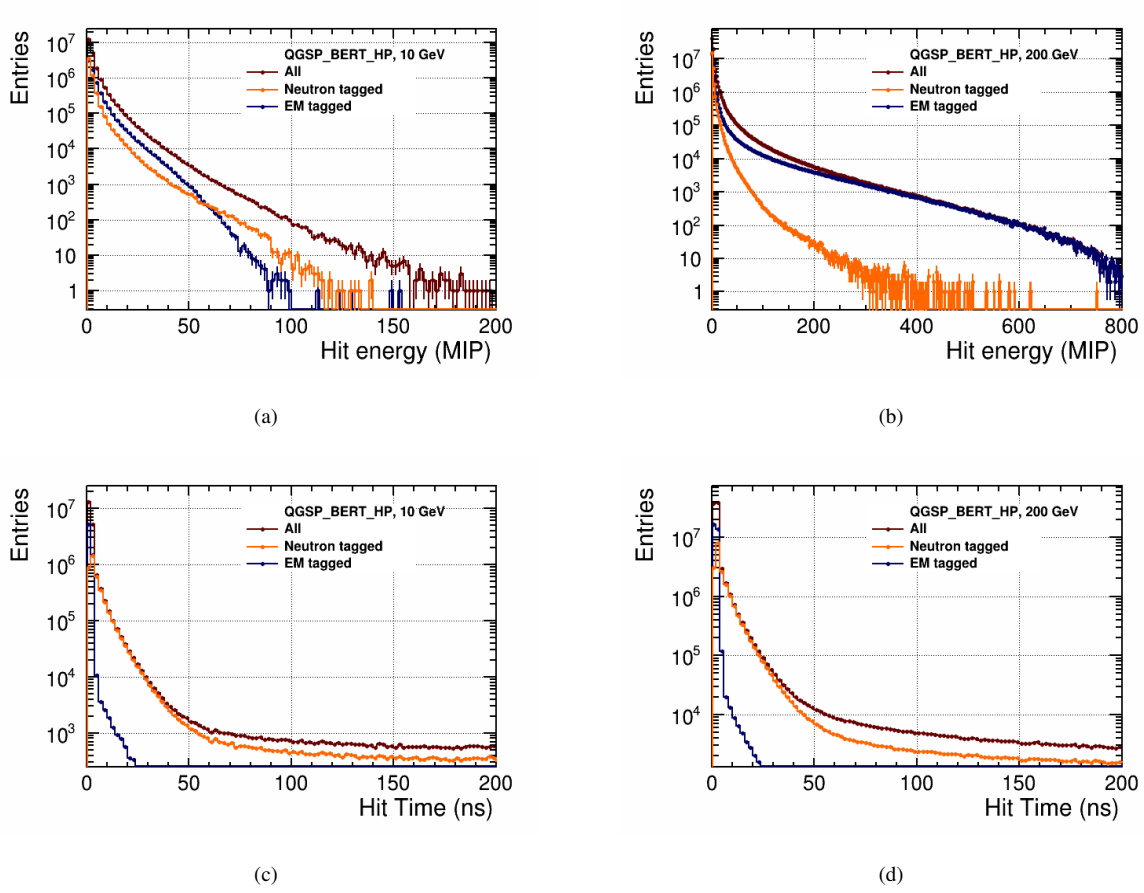


FIGURE 7.56: Hit energy distribution using pion showers with the beam energy at (a) 10 GeV (b) 200 GeV. Hit time distribution using pion showers with the beam energy at (c) 10 GeV and (d) 200 GeV. EM tagged distribution (in blue) describe the hits with the EM tag > 0.9 and Neutron tagged (in orange) distribution describe the hits with the Neutron tag > 0.9. The overall distribution (in red) describes the contribution from all hits.

The observables neutron tagged and EM tagged are used for further study. The study focuses on the primary goal of describing the shape of an average hadronic shower and how far we are from understanding the picture of a hadronic shower. Moreover, to address which region in the hadronic shower do electromagnetic and the "true" hadronic components belong. To investigate this picture, it is essential to look into longitudinal and radial shower shapes. The longitudinal and radial shower profiles are extracted using the method obtained earlier. In addition, the longitudinal and radial distribution is weighted with the corresponding neutron fraction or EM fraction for each hit, as given below:

$$E_N(hit) = E(hit) \times Neutron\ tagged(hit)$$
$$E_{EM}(hit) = E(hit) \times EM\ tagged(hit)$$

Figure 7.57 shows the behaviour of longitudinal and radial shower shape for the EM tagged, Neutron tagged and the overall distribution (without any weights) using QGSP_BERT_HP physics list for beam energies at 10 GeV and 200 GeV.

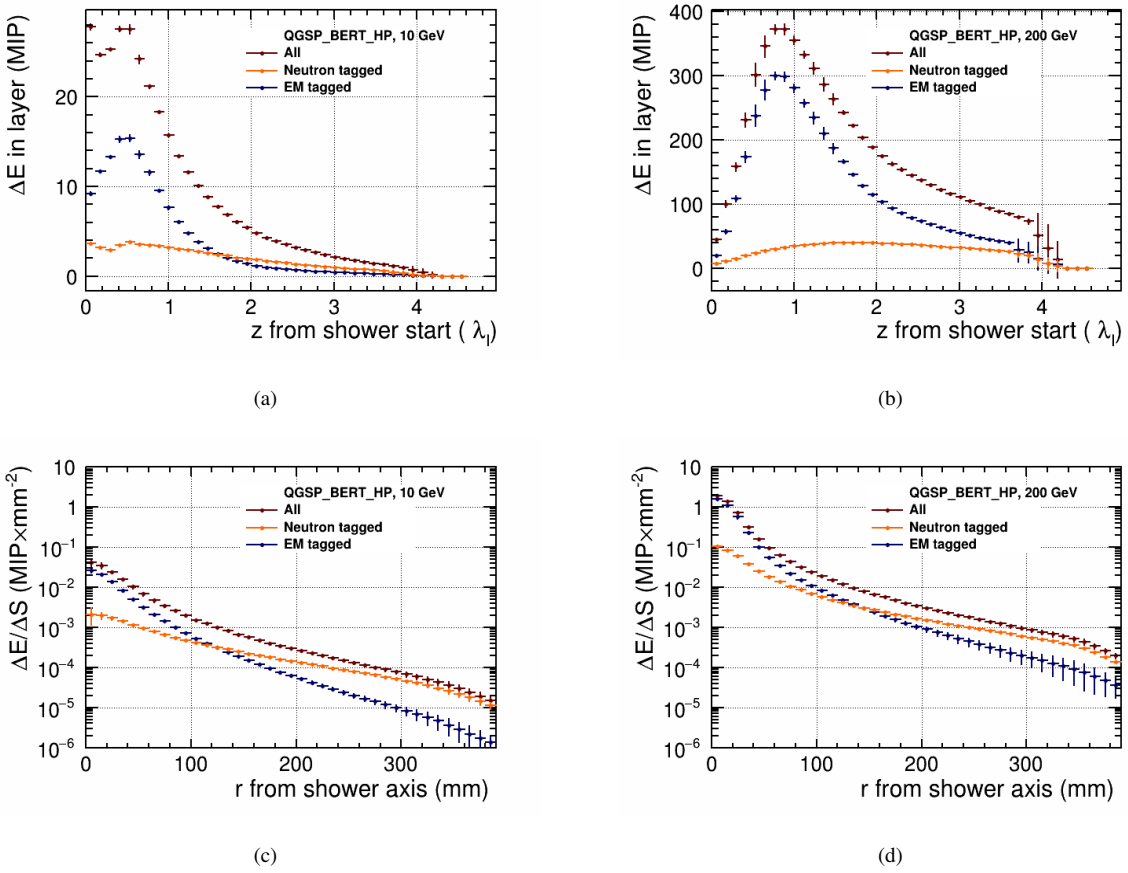


FIGURE 7.57: Longitudinal distribution from shower start using pion showers with the beam energy at (a) 10 GeV (b) 200 GeV. Radial distribution using pion showers with the beam energy at (c) 10 GeV and (d) 200 GeV. Both longitudinal and radial distributions are obtained using QGSP_BERT_HP physics list. EM distributions (in blue) describes the hits weighted with the EM content and Neutron distributions (in orange) describe the hits weighted with Neutron content. The overall distribution (in red) describes the contribution from all hits.

Following observations are noted from the above longitudinal and radial distributions: From the radial profile, it is clear that the neutrons persist in the entire calorimeter region and the neutrons are mainly dominating in the outermost regions of the calorimeter. In contrast, this behaviour is partially suppressed in the longitudinal profiles. Besides, looking more closely at the core of the shower by no doubt is dominated by the electromagnetic part of the hadronic shower.

The understanding is that the EM part has a different spatial resolution than the rest of the shower. This is related to the fact that there is isospin invariance, and one-third of all pions produced is π^0s. However, pions are not produced everywhere. They are only produced above

the pion threshold (≈300 MeV). As emphasized earlier, there is a smaller number of subsequent generations in hadronic showers and significant fluctuations than in the electromagnetic shower.

In the peripheral region of the shower, the energy is mainly carried by neutrons being ejected from the nuclei of the calorimeter material. They travel deeper into the calorimeter than neutral pions but below the threshold for producing further pions. Therefore, mainly the evaporation processes carries the energy to the peripheral region of the shower without producing any new pions. As a result, the particle composition in the radial periphery or the shower tail is different from the shower core. However, as long as it is above the threshold, the particle production in the core still has significant fluctuation.

Therefore, there is not only one electromagnetic shower at the beginning, but it forms a small electromagnetic shower in the core of the overall hadronic shower. This is more likely to be collinear with a small transverse momentum and strongly boosted in the forward direction. As a result, fluctuating in the longitudinal direction between different π^0 productions.

As mentioned earlier, in order to confirm the assumption that the fit functions relate to the electromagnetic and hadronic parts within the shower. The fit functions are directly compared to the truth information using neutron and EM tagged shower profile. Therefore, the longitudinal profile with EM tagged in hadron shower is compared to the Gamma function from electromagnetic longitudinal parametrization as given in Eq. 7.1. Likewise, in the radial profile the EM tagged in hadron shower is compared to the radial parametrization as given in Eq. 7.3.

Figure 7.58(a) and 7.58(b) shows the comparison of EM tagged longitudinal profile using pion showers to the Gamma function with beam energy at 10 and 200 GeV, respectively. Figure 7.58(c) and 7.58(d) shows the comparison of EM tagged radial profile using pion showers to the exponential distribution with beam energy at 10 and 200 GeV, respectively.

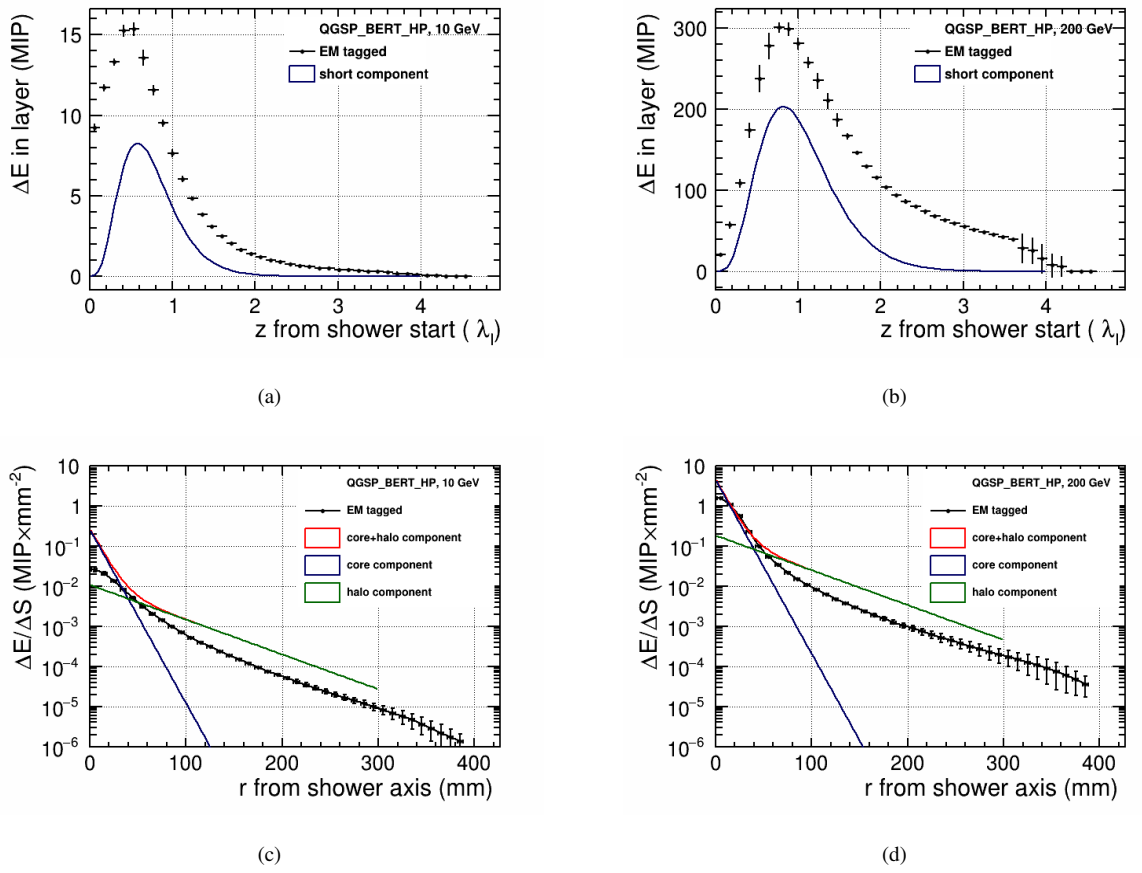


FIGURE 7.58: Short component (in blue) compared to the EM tagged longitudinal distribution using pion showers with beam energy of (a) 10 GeV and (b) 200 GeV. Comparision of exponential distribution to EM tagged radial distribution using pion showers with beam energy at (c) 10 GeV and (d) 200 GeV. The distributions are obtained using `QGSP_BERT_HP` physics list.

It is clear from the above longitudinal distributions that the maximum position of the longitudinal profile agrees reasonably well using the Gamma function. Besides, the EM tagged radial distributions have a similar shape. Nevertheless, it does not give an excellent description of the shower. In addition, the normalization is different between the fit function and the shower profiles in both the longitudinal and the radial distributions. This points out that more electromagnetic energy is deposited as obtained from the truth information than what we can "capture" in the short component.

Similarly, the Neutron tagged longitudinal and radial distributions are compared to the fit functions as given in Eq. 7.4 and Eq. 7.19, respectively. The results are shown in figure 7.59(a) and 7.59(b) for pion showers using `QGSP_BERT_HP` physics list with beam energies at 10 GeV and 200 GeV respectively. It is observed that the radial slope of the halo distribution and the neutron component is reproduced. The truth information from the longitudinal shower profile significantly underestimated the energy deposition. Therefore, this hints at the understanding that we place in the fit functions. The short or core component related to electromagnetic shower and the long or halo component related to the pure hadronic shower structure is too simple.

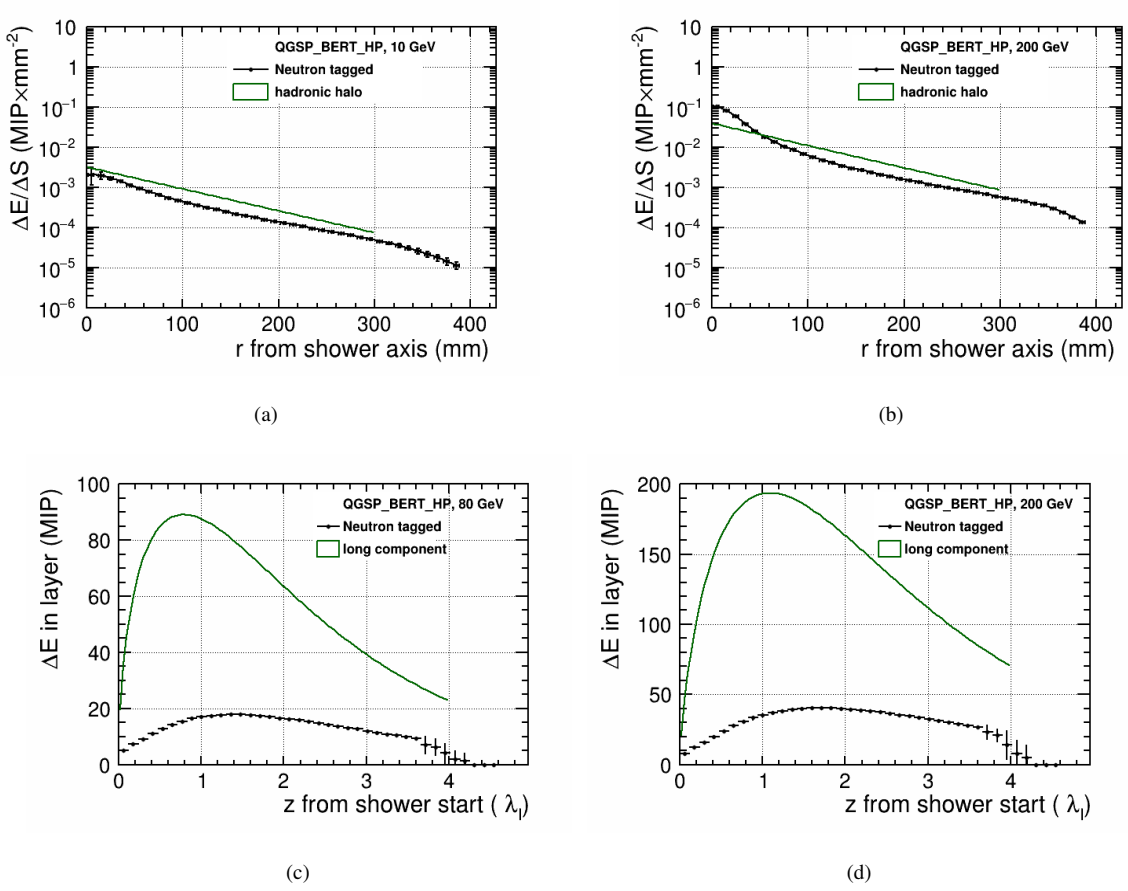


FIGURE 7.59: Halo component from the parametrization given in Eq. 7.19 is compared to the neutron tagged radial distribution using pion showers with beam energy at (a) 10 GeV and (b) 200 GeV. Long component from the parametrization given in Eq. 7.4 is compared to the neutron tagged longitudinal distribution using pion showers with beam energy at (c) 10 GeV and (d) 200 GeV. The distributions are obtained using `QGSP_BERT_HP` physics list.

7.3 Summary

This chapter has described the analysis of electromagnetic and hadronic shower shapes based on longitudinal and radial development of the shower from the data collected in the June 2018 campaign with the AHCAL technological prototype. The shower shapes were compared between testbeam data and `GEANT4` simulations using `QGSP_BERT_HP` physics list for the electron showers and `FTFP_BERT_HP` physics lists in addition for the pion showers.

Before dealing with a complex object, like the hadronic shower, it was essential to understand the agreement of electromagnetic showers. From the study of electromagnetic showers, the longitudinal shower shapes were parametrized using a Gamma distribution, while the radial shower shapes were parametrized with the sum of two exponential distributions. This study confirmed that using two components was necessary to fully describe the radial electromagnetic showers. Furthermore, the fit parameters obtained were compared between the testbeam data and simulations as a reasonable agreement was observed. The maximum from the longitudinal

distribution increased with energy as expected, while the slope parameters from the fit to radial shapes were independent of beam energy.

For pion showers, a comparison of longitudinal and radial shower shapes was done with the previous physics prototype, using the same parametrization for longitudinal shower shape, while a normalized parametrization was obtained for describing radial shower shapes. The resulting fit parameters were consistent with the previously obtained values. However, the study showed that the fit parameter values E and f_{EM} obtained from the longitudinal and the radial part of the shower showed a disagreement. Therefore, this study led towards introducing an improved parametrization for the radial shower shape. This improved parametrization partially solved the difference of f_{EM} between longitudinal and radial fit parameters by introducing a K as a correction factor in the radial parametrization. In addition, three components were used to fully describe the radial behaviour of a hadronic shower. The improved radial parametrization showed a good agreement between the testbeam data and simulations with a good fit quality.

Furthermore, a simultaneous fitting of longitudinal and radial shower shapes was performed using testbeam data and simulations. From this method, an average electromagnetic fraction f_{EM} for beam energies between 10 GeV to 200 GeV was obtained to be $\approx$10% to $\approx$ 30%. The approach was relatively successful at higher energies, while the parameterization is slightly less described at lower energies.

Afterwards, a layer-wise fitting of radial shower shapes was performed using testbeam data and simulations. From the layer-wise fits, the value of the slope parameters from the radial shapes showed no dependence through the layers. Therefore, with the fixed radial slopes, the evolution of showers could still be described. Based on this parameterization, a physics cross-check was also studied to understand the picture of an average hadronic shower structure involving a "core" and a "halo" component. It was verified using testbeam data and simulations that the picture of the hadronic shower was too simplistic. This statement is supported by the study that the electromagnetic part of the pion shower was observed to be similar but was longer and deeper than the pure electron showers.

In light of the understanding of electron and pion showers, the first analysis of a three-dimensional model was discussed to explain an average hadronic shower shape using a highly granular calorimeter. This model considers twelve energy-dependent parameters and a good description of the evolution of hadronic shower within an agreement of $\sim$5% between testbeam data and simulations. Furthermore, the resulting fit parameters from the three-dimensional model differed from those obtained from the one-dimensional, thus gaining additional information.

Finally, a Monte Carlo particle-based study was done to look for the 'truth' information from particle showers. This provides a tool to analyze the shower shapes giving an approximate idea of the particle composition. Monte Carlo particle shower was correlated with the showers that originated from neutral hadrons that contributed to the tails of the showers. In contrast, the particles originating from the EM-like contributed to the core of the hadronic shower. As already emphasized, these studies show that the picture we had in mind when we chose the fit function,

namely that there is an electromagnetic core that looks nearly identical to an electron shower of $\approx$1/3 energy, and a "neutron halo" that can be described with one exponential function in the radial direction, is too simple. Nonetheless, this demonstrates that the CALICE-AHCAL calorimeter has a sufficient longitudinal and radial granularity to reasonably discriminate between the electromagnetic and hadronic components within a hadron shower.

Part IV

Concluding Remarks

Chapter 8

Conclusions and Outlook

"There is no real ending. It's just the place where you stop the story."

Frank Herbert

The International Linear Collider (ILC) is the proposed future linear e^+e^- collider with center-of-mass energy designed at mainly 250 GeV to 500 GeV. ILC requires an unprecedented jet energy resolution to reach their ambitious goals in the precision measurement of Standard Model parameters and the potential discovery of new physics beyond it. One primary goal is to measure the couplings of the Higgs boson precisely. To achieve this, a jet energy resolution of 3-4% is necessary using a Particle Flow Approach (PFA). It aims to combine the tracking and calorimeter measurements into the best jet energy measurement by measuring each particle in a jet using the best suited sub-detector resolution to measure the energy. In addition, PFA needs to associate energy depositions in the calorimeters and tracks correctly. This requires an unprecedented spatial resolution which can be achieved with high granular calorimeters.

The CALICE collaboration develops highly granular calorimeters for detectors based on the PFA. These detectors are expected to meet the jet energy resolution requirements of experiments at future e^+e^- linear colliders. One of its kind is the highly granular analog hadron calorimeter (AHCAL), equipped with plastic scintillator tiles of 3×3 cm^2 area, read out by Silicon Photo-multipliers (SiPMs) with a total of $\sim$22,000 single channels. The aim is to test the hardware and software performances of the AHCAL design and to investigate the scalability of the detector to a full calorimeter for a Particle Flow detector system of a future linear lepton collider.

Overall, the three testbeam campaigns at CERN-SPS in 2018 with the AHCAL technological prototype were successful. A large set of data with different momentum and particle types such as muons, electrons and pions were recorded. Before comparing the testbeam data to simulated data, the calibration of the detector was performed. The detector calibration was performed using LED runs and muon data. Less than 0.1% of channels could not be calibrated.

In this thesis, the experimental setup is described, including the beam generation for the data samples, which are simulated using `QGSP_BERT_HP` and `FTFP_BERT_HP` physics lists

from GEANT4 10.03. For electron or hadron shower analysis, event selections are utilized to select pure showers. The selections have been optimized with the simulation using Particle Identification based on Boosted Decision Technique. The efficiencies are better than 99% in the studied range from 10 to 200 GeV.

The study of the topological structure of the hadron showers, with the pion data collected during the June 2018 testbeam campaign, represents the core of this thesis. For this, the understanding of the detector and its performance is crucial. Therefore, muon and electron data are used to cross-check and validate the calibration to have a reliable description of the detector. One of the main parameter in the digitization process, the parameter for the saturation correction of the silicon photomultipliers, have been tuned using electron showers. A comparison between testbeam data and simulation using the QGSP_BERT_HP physics list has been performed. A variation of the SiPM saturation has been observed among the samples. This leads to a discrepancy at high beam energies between testbeam data and simulation. An optimal number of effective pixels is 2668 with an uncertainty of ± 100 pixels have been determined after detailed studies of SiPM saturation correction. Furthermore, the simulated data was validated using electromagnetic showers and concluded that a good energy calibration is achieved at the cell level. The main figure of merits for a calorimeter is the energy linearity and energy resolution, which were obtained using electron and pion showers. An excellent linear response of the AHCAL was achieved between testbeam data and simulation for electrons with a discrepancy of up to 4% and up to 10% for pions. An energy resolution with a stochastic term of $\approx 23\%$ for electrons and $\approx 58\%$ for pions was achieved using testbeam data and simulations. The stochastic term obtained for pions using the data from the current AHCAL technological prototype is better compared to the current hadronic calorimeters such as ATLAS which has the stochastic term of: $\frac{52.9\%}{\sqrt{E(\mathrm{GeV})}} \oplus 5.7\%$ and CMS with the stochastic term of: $\frac{84.\%}{\sqrt{E(\mathrm{GeV})}} \oplus 0.7\%$ [55, 56] which are obtained after corrections. Moreover, the stochastic term obtained from the physics prototype is 57.6 ± 0.4 [109] and is in good agreement to the current technological prototype.

After the simulation, parameters were tuned, and the agreement between testbeam data and simulations was evaluated. It was then possible to proceed with the shower shape analysis. Before analyzing the pion shower shapes, the electron shower shapes were studied. The longitudinal shower profile for electromagnetic showers was parametrized using a Gamma distribution. The parameters obtained from the Gamma distribution, such as the maximum position of the shower, showed good agreement between testbeam data and simulations. From the study of the electromagnetic shower, it was observed that two exponential distributions were required for the radial shower in describing the "core" and a "halo" of an electromagnetic shower. Moreover, the radial slope parameters were found to be independent of beam energy.

Furthermore, the analysis aimed to understand the structure of hadronic shower shapes in great detail and validate the simulation of hadron physics. Longitudinal and radial shower shapes are analyzed as they provide vital clues as to the nature of the particle(s). At first, independent fitting of longitudinal and radial shower profiles was performed. The parameterization for longitudinal

and radial shower shapes was obtained from the previous physics prototype. The longitudinal profile was parametrized using the sum of two contributions (Gamma distributions) called "short" and "long". In contrast, the radial profile was parametrized with the sum of two exponential distributions, which describe the behaviour near the shower axis ("core" region) and at the shower periphery ("halo" region). From the radial parametrization, the energy parameter E and fraction parameter f_{EM} were extracted. Qualitatively, the fit results are reproduced, and the results are consistent between the physics prototype and technological prototype, which marked a reasonable agreement within $\approx$20%.

The calorimeter characteristic h/e signal ratio (the ratio of the responses to the non-electromagnetic and electromagnetic components of a hadron-induced shower) is estimated for the AHCAL from a fit to the longitudinal profiles using the extracted parameters. Both physics lists tend to give better predictions and overestimate the value of h/e with increasing energy due to the simplified representation of the longitudinal shower development.

While analyzing hadron showers, two major problems were encountered when using the longitudinal and radial parameterization. Firstly, the value of the fit parameter f_{EM} extracted from the radial parametrization was overestimated roughly by a factor 3, which was also consistent between the values from testbeam data and simulations. Secondly, the value of the fit parameter E corresponding to total energy obtained from the fit to longitudinal shower profile did not agree with the total energy obtained from the fit to the radial shower profile. Therefore, a correction in the radial parametrization was done to account for leakage effects. This, resulted in a reasonable agreement of the parameters f_{EM} and E obtained from the fit to the longitudinal and radial shower profiles. The radial parametrization for hadron showers was improved. It was best described using three components: the electromagnetic component of the hadron shower, the hadronic core and the hadronic halo. Furthermore, the fitting procedure was extended to perform simultaneous fits with the assumption that the total energy and fraction from the two projections of the shower should agree. Thus an average electromagnetic fraction between 10 to 30% was obtained for all available pion energies. Moreover, a physics cross-check was performed with the core of the hadron shower to be comparable to one-third of its electrons energy. The assumption in the shape of a hadronic shower was a simplistic picture, with the EM part of the pion shower being similar but longer and more profound than the electron shower.

The ground-breaking achievement of this thesis is the construction of the three-dimensional hadronic shower model. This model allows describing the shape of an average hadronic shower. For the first time, a three-dimensional model is applied to the CALICE-AHCAL testbeam data and simulations. The model involves twelve energy-dependent parameters, which gain additional information compared to one-dimensional fits. Moreover, the model is reasonably explained for energies greater than 10 GeV.

Two variables called Neutron tagged and EM tagged have been described as an energy weighted quantity based on Monte Carlo truth information. The variables offer to disentangle the major contributions to a hadronic shower shape composition. Neutron tagged and EM tagged

information was used for shower shape studies to determine which shower region was contributed by pure electromagnetic showers and the remaining part arising from the neutral hadrons. In addition, to compare the fit function to the truth information provided by GREANT4.

The two variables, Neutron and EM tagged, might provide scope for "offline" software compensation techniques to further improve on calorimeter energy resolution [109]. For most hadronic calorimeters, the response, i.e. the ratio of the signal generated in the active part to the total deposited energy, is different for electromagnetic and for hadronic showers. Therefore event-to-event-fluctuations in the electromagnetic content of the shower contribute to the hadronic energy resolution unless the difference is compensated. Therefore compensation can be achieved by weighting each hit in the pion shower depending on the EM or neutron content.

Data analysis in experiments at the future high energy colliders like ILC requires the detector response simulation via the Monte Carlo technique for a large number of events. The existing Monte Carlo programs provide a detailed and reliable simulation of the detector response in electromagnetic and hadronic calorimeters. However, these programs do have drawbacks. There are many studies where a detailed simulation of showers with their full complexity is not needed, but just some approximate description of the energy distribution and fluctuation. In this case, using a "fast Monte Carlo" seems to be more efficient and adequate. A certain number of fast simulation programs do exist, such as GFlash [111]. A fast Monte Carlo simulation aims to reproduce global rather than detailed shower properties. It is generally obtained from the parametrization of showers generated with detailed Monte Carlo programs or directly with experimental data. The current proposed three-dimensional model is suitable in the areas of fast simulation [112]. However, individual shower profiles deviate considerably from average profiles, especially towards low energies. In this thesis the model has not been attempted on the event-to-event shower fluctuations. This could be interesting to look into shower shapes on an event-by-event basis. One could potentially reach to include two kinds of fluctuations in fast Monte Carlo a) global fluctuations, by fluctuating the parameters around their mean values, and b) local fluctuations, by fluctuating the position of the energy hits.

The key aspects of detailed level of shower analyses are essential for two-particle separation in particle flow algorithms, which is an important constraint on the energy resolution that can be achieved by future experiments. Moreover, the three-dimensional hadronic model has potential usage in particle flow algorithms such as in Pandora [113], to determine what is the probability that a hit belongs to a particle shower or not. Shower shapes in Pandora currently use one Gamma function to describe the longitudinal profile, while no parametrization involves radial profiles.

The AHCAL detector development and testbeam studies are ongoing. Finally, a combined testbeam with the large technological AHCAL prototype and the SiW-ECAL prototype is planned for June 2022. This will allow in testing the three-dimensional model and its flexibility to a complete calorimeter system which has the potential to lead to even better models in future and maybe extremely beneficial. In addition, to perform shower shape studies at different angles of incident beam, would automatically have a different effective sampling fraction. The

parametrization for instance could be made universal, i.e. if the parameters are expressed not in mm but in X_0, λ or R_M. This could be applied to calorimeters with different sampling fractions or different materials to obtain faster simulations to optimize the calorimeter design. Furthermore, the use of wider range of energies in smaller steps and extending the study to different particle types such as protons and kaons would provide further insight into the details of hadronic shower interactions in the calorimeters. Moreover, using hit time information for the shower shape studies might significantly improve in separation between electromagnetic and hadronic components within a shower.

Appendix A

Database Tags

This appendix presents the database folder and tag names used for the geometry and calibration in this analysis for standard data reconstruction. Table A.5 provides the tags used for the different number of effective pixels applied during the reconstruction step to correct for the SiPM saturation, and table A.6 provides the systematic uncertainty applied to electron and pion data analysis.

Calibration set	Database folder & tag
Ahc2ModuleDescription	/cd_calice_Ahc2/TestbeamMay2018/ModuleDescription (HEAD)
Ahc2ModuleConnection	/cd_calice_Ahc2/TestbeamMay2018/ModuleConnection (ahc2_ModuleConnection_180822)
Ahc2ModuleLocationReference	/cd_calice_Ahc2/TestbeamMay2018/ModuleLocationReference (ahc2_ModuleLocationReference_180822)
Ahc2HardwareConnection	/cd_calice_Ahc2/TestbeamMay2018/Ahc2HardwareConnection (ahc2_HardwareConnection_180822)
Ahc2DetectorTransformation	/cd_calice_Ahc2/TestbeamMay2018/DetectorTransformation (HEAD)

TABLE A.1: The CALICE geometry database folder names of the May 2018 CERN testbeam campaign. Tags are given in brackets.

Calibration set	Database folder & tag
Ahc2ModuleDescription	/cd_calice_Ahc2/TestbeamJune2018/ModuleDescription (HEAD)
Ahc2ModuleConnection	/cd_calice_Ahc2/TestbeamJune2018/ModuleConnection (HEAD)
Ahc2ModuleLocationReference	/cd_calice_Ahc2/TestbeamMay2018/ModuleLocationReference (HEAD)
Ahc2HardwareConnection	/cd_calice_Ahc2/TestbeamJune2018/Ahc2HardwareConnection (HEAD)
Ahc2DetectorTransformation	/cd_calice_Ahc2/TestbeamJune2018/DetectorTransformation (HEAD)

TABLE A.2: The CALICE geometry database folder names of the June 2018 CERN testbeam campaign. Tags are given in brackets.

Calibration set	Database folder & tag
E4DPedestal	/cd_calice_Ahc2/TestbeamMay2018/Pedestal (ahc2_pedestal_180906)
E4DPedestalMemoryCellOffset	/cd_calice_Ahc2/TestbeamMay2018/PedestalMemoryCellOffset (ahc2_pedestalmemorycelloffset_181216)
E4DLowGainPedestal	/cd_calice_Ahc2/TestbeamMay2018/LowGainPedestal (ahc2_lg_pedestal_190114)
E4DLowGainPedestalMemoryCellOffset	/cd_calice_Ahc2/TestbeamMay2018/LowGainPedestalMemoryCellOffset (ahc2_lg_pedestalmemorycelloffset_190227)
E4DGainConstants	/cd_calice_Ahc2/TestbeamMay2018/gain_constants (ahc2_gainconstant_180827)
E4DGainSlopes	/cd_calice_Ahc2/TestbeamMay2018/gain_slopes (HEAD)
E4DMipConstants	/cd_calice_Ahc2/TestbeamMay2018/mip_constants (ahc2_mip_constants_180925)
E4DMipSlopes	/cd_calice_Ahc2/TestbeamMay2018/mip_slopes (HEAD)
E4DDeadCellMap	/cd_calice_Ahc2/TestbeamMay2018/DeadCellMap (HEAD)
E4DSaturationParameters	/cd_calice_Ahc2/TestbeamMay2018/SaturationParameters (ahc2_SaturationParameters_180824)
E4DIntercalibration	/cd_calice_Ahc2/TestbeamMay2018/Intercalibration (ahc2_Intercalibration_180824)
E4DPhysicsCalibIntercalibration	/cd_calice_Ahc2/TestbeamMay2018/PhysicsCalibIntercalibration (HEAD)
E4DTimeOffset_Timeout	/cd_calice_Ahc2/TestbeamMay2018/TimeOffset_Timeout (ahc2_timeOffsetsTimeout_190506)
E4DTimeSlopes_Timeout	/cd_calice_Ahc2/TestbeamMay2018/TimeSlopes_Timeout (ahc2_timeSlopesTimeout_190506)
E4DTimeOffsetMemCell_EvenTimeout	/cd_calice_Ahc2/TestbeamMay2018/TimeOffsetMem_Even_Timeout (ahc2_timeOffsetsTimeout_Even_190506)
E4DTimeOffsetMemCell_OddTimeout	/cd_calice_Ahc2/TestbeamMay2018/TimeOffsetMem_Odd_Timeout (ahc2_timeOffsetsTimeout_Odd_190506)
E4DTimeOffsetMemCell_BufferEvenEventEven	/cd_calice_Ahc2/TestbeamMay2018/TimeOffsets_BufferEven_EventEven (ahc2_timeOffsetsBufferEven_EventEven_190517)
E4DTimeOffsetMemCell_BufferEvenEventOdd	/cd_calice_Ahc2/TestbeamMay2018/TimeOffsets_BufferEven_EventOdd (ahc2_timeOffsetsBufferEven_EventOdd_190517)
E4DTimeOffsetMemCell_BufferOddEventEven	/cd_calice_Ahc2/TestbeamMay2018/TimeOffsets_BufferOdd_EventEven (ahc2_timeOffsetsBufferOdd_EventEven_190517)
E4DTimeOffsetMemCell_BufferOddEventOdd	/cd_calice_Ahc2/TestbeamMay2018/TimeOffsets_BufferOdd_EventOdd (ahc2_timeOffsetsBufferOdd_EventOdd_190517)
E4DOccupancyBxidEvenHighGain	/cd_calice_Ahc2/TestbeamMay2018/Occupancy_ConstantsEvenHigh (ahc2_occupacyConstantsEvenHigh_200421_1)
E4DOccupancyBxidEvenLowGain	/cd_calice_Ahc2/TestbeamMay2018/Occupancy_ConstantsEvenLow (ahc2_occupacyConstantsEvenLow_200421_1)
E4DOccupancyBxidOddHighGain	/cd_calice_Ahc2/TestbeamMay2018/Occupancy_ConstantsOddHigh (ahc2_occupacyConstantsOddHigh_200421_1)
E4DOccupancyBxidOddLowGain	/cd_calice_Ahc2/TestbeamMay2018/Occupancy_ConstantsOddLow (ahc2_occupacyConstantsOddLow_200421_1)

TABLE A.3: The CALICE calibration database folder names of the May 2018 CERN testbeam campaign. Tags are given in brackets.

Calibration set	Database folder & tag
E4DPedestal	/cd_calice_Ahc2/TestbeamJune2018/Pedestal (ahc2_pedestal_180926)
E4DPedestalMemoryCellOffset	/cd_calice_Ahc2/TestbeamJune2018/PedestalMemoryCellOffset (ahc2_pedestalmemorycelloffset_181216)
E4DLowGainPedestal	/cd_calice_Ahc2/TestbeamJune2018/LowGainPedestal (ahc2_lg_pedestal_190118)
E4DLowGainPedestalMemoryCellOffset	/cd_calice_Ahc2/TestbeamJune2018/LowGainPedestalMemoryCellOffset (ahc2_lg_pedestalmemorycelloffset_190227)
E4DGainConstants	/cd_calice_Ahc2/TestbeamJune2018/gain_constants (ahc2_gainconstant_181212)
E4DGainSlopes	/cd_calice_Ahc2/TestbeamJune2018/gain_slopes (HEAD)
E4DMipConstants	/cd_calice_Ahc2/TestbeamJune2018/mip_constants (ahc2_mip_constants_181001)
E4DMipSlopes	/cd_calice_Ahc2/TestbeamJune2018/mip_slopes (HEAD)
E4DDeadCellMap	/cd_calice_Ahc2/TestbeamMay2018/DeadCellMap (HEAD)
E4DSaturationParameters	/cd_calice_Ahc2/TestbeamJune2018/SaturationParameters (ahc2_SaturationParameters_190628)
E4DIntercalibration	/cd_calice_Ahc2/TestbeamJune2018/Intercalibration (ahc2_Intercalibration_190109)
E4DPhysicsCalibIntercalibration	/cd_calice_Ahc2/TestbeamMay2018/PhysicsCalibIntercalibration (HEAD)
E4DTimeOffset_Timeout	/cd_calice_Ahc2/TestbeamJune2018/TimeOffset_BufferOdd (ahc2_timeOffsetsBufferOdd_190806)
E4DTimeSlopes_Timeout	/cd_calice_Ahc2/TestbeamJune2018/TimeSlopes_BufferOdd (ahc2_timeSlopesBufferOdd_190806)
E4DTimeOffsetMemCell_EvenTimeout	/cd_calice_Ahc2/TestbeamJune2018/TimeOffsets_BufferOdd_EventEven (ahc2_timeOffsetsBufferOdd_EventEven_190806)
E4DTimeOffsetMemCell_OddTimeout	/cd_calice_Ahc2/TestbeamJune2018/TimeOffsets_BufferOdd_EventOdd (ahc2_timeOffsetsBufferOdd_EventOdd_190806)
E4DTimeOffsetMemCell_BufferEvenEventEven	/cd_calice_Ahc2/TestbeamJune2018/TimeOffsets_BufferEven_EventEven (ahc2_timeOffsetsBufferEven_EventEven_190808)
E4DTimeOffsetMemCell_BufferEvenEventOdd	/cd_calice_Ahc2/TestbeamJune2018/TimeOffsets_BufferEven_EventOdd (ahc2_timeOffsetsBufferEven_EventOdd_190808)
E4DTimeOffsetMemCell_BufferOddEventEven	/cd_calice_Ahc2/TestbeamJune2018/TimeOffsets_BufferOdd_EventEven (ahc2_timeOffsetsBufferOdd_EventEven_190806)
E4DTimeOffsetMemCell_BufferOddEventOdd	/cd_calice_Ahc2/TestbeamJune2018/TimeOffsets_BufferOdd_EventOdd (ahc2_timeOffsetsBufferOdd_EventOdd_190806)
E4DOccupancyBxidEvenHighGain	/cd_calice_Ahc2/TestbeamJune2018/Occupancy_ConstantsEvenHigh (ahc2_occupacyConstantsEvenHigh_200416_1)
E4DOccupancyBxidEvenLowGain	/cd_calice_Ahc2/TestbeamJune2018/Occupancy_ConstantsEvenLow (ahc2_occupacyConstantsEvenLow_200416_1)
E4DOccupancyBxidOddHighGain	/cd_calice_Ahc2/TestbeamJune2018/Occupancy_ConstantsOddHigh (ahc2_occupacyConstantsOddHigh_200416_1)
E4DOccupancyBxidOddLowGain	/cd_calice_Ahc2/TestbeamJune2018/Occupancy_ConstantsOddLow (ahc2_occupacyConstantsOddLow_200416_1)

TABLE A.4: The CALICE calibration database folder names of the June 2018 CERN testbeam campaign. Tags are given in brackets.

N_{eff} *Pixels*	Database tag
2448	ahc2_SaturationParameters_190705
2500	ahc2_SaturationParameters_190706
2600	ahc2_SaturationParameters_181212
2700	ahc2_SaturationParameters_190708
2778	ahc2_SaturationParameters_190709
2888	ahc2_SaturationParameters_190710

TABLE A.5: Database tags from June 2018 used for various number of effective pixels N_{eff} applied during reconstruction for saturation correction.

Calibration	Database tag
MIP	ahc2_mip_constants_192709
Gain	ahc2_gain_constants_192709
SiPM	ahc2_SaturationParameters_191203

TABLE A.6: Database tags from June 2018 used for smeared calibration constants needed for systematic studies.

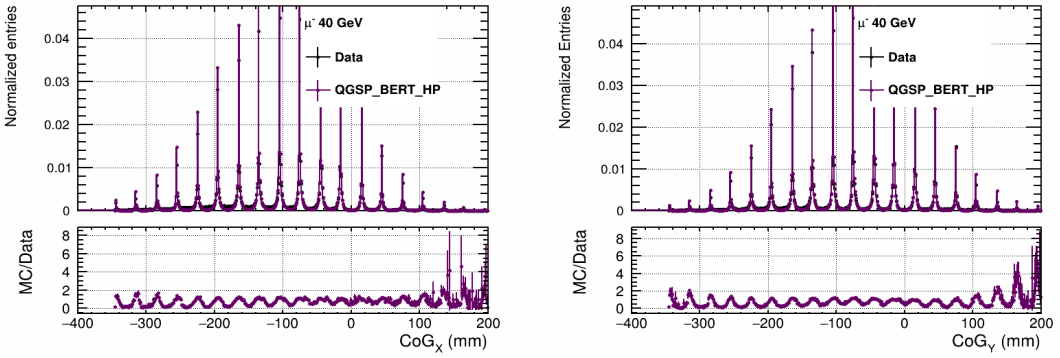
Appendix B

Sample Preparation and Event Selection: Additional Plots

In this appendix, the particle gun settings used to reproduce the testbeam data runs for each beam momentum, and particle type in this thesis is shown in table B.1. Additional plots concerning the center of gravity in X direction (CoG_X) and center of gravity in Y direction (CoG_Y) for electrons is shown in figure B.2 while for pions is shown in figure B.3 with the comparison between testbeam data and simulation for various beam energies. The shape of these distributions are in good agreement between testbeam data and simulations.

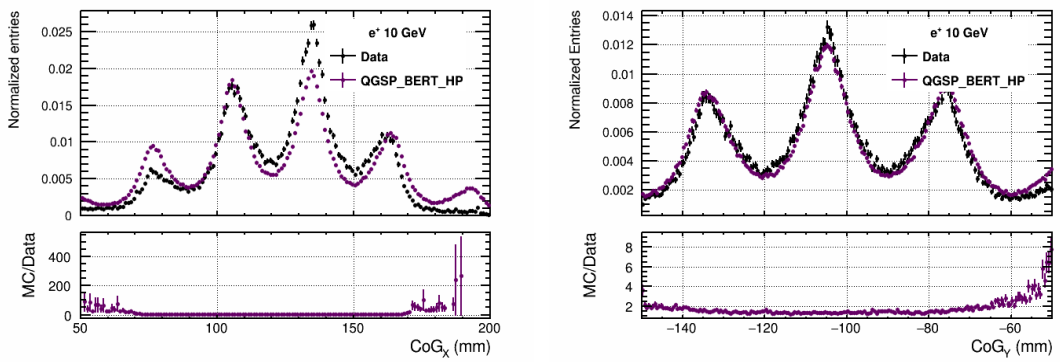
Particle Type	**Energy (GeV)**	**Run Number**	μ_x **(mm)**	μ_y **(mm)**	σ_x **(mm)**	σ_y **(mm)**
μ^-	40	61167	-103.6	-75.27	93.75	94.87
e^+	10	61293	124.2	27.8	-103.9	29.8
	20	61296	129.5	19.5	-101.3	20.0
	30	61213	129.3	14.3	-102.5	14.6
	40	61212	128,9	13,0	-102.9	11.6
	50	61214	128.5	11.4	-99.5	9.5
	60	61211	127.7	10.7	-106.0	8.3
	70	61215	128.0	10.1	-104.3	7.1
	80	61210	124.6	10.4	-103.5	6.3
	90	61216	127.1	9.6	-102.2	6.6
	100	61217	126.3	10.2	-101.0	7.0
π^-	10	61265	5.14	-2.24	10.93	16.96
	20	61273	6.61	-5.42	23.14	21.39
	30	61384	-120.97	-19.77	27.04	21.29
	40	61275	10.37	7.53	27.47	20.54
	60	61262	10.32	-7.9	27.07	16.3
	80	61279	9.93	1.01	27.18	15.52
	120	61287	8.38	0.4	6.77	0.88
	160	61222	5.55	1.77	6.46	2.50
	200	61201	-0.15	9.35	5.47	6.56

TABLE B.1: Simulated particle gun settings applied to reproduce the beam profile for each beam momentum, and particle type used in this thesis.



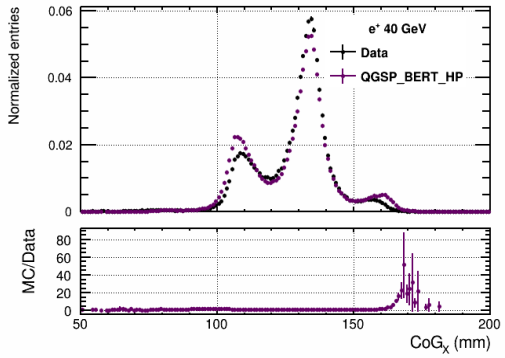
(a) (b)

FIGURE B.1: Center of gravity distributions in x and y-directions using 40 GeV muon from testbeam data and simulated data with `QGSP_BERT_HP` physics list.

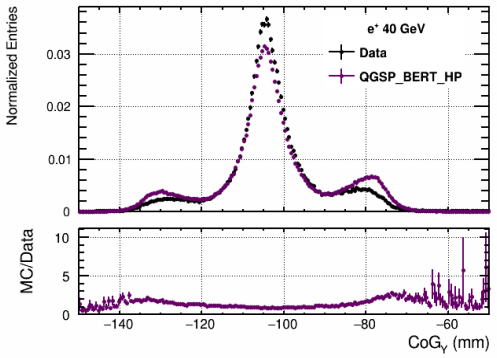


(a)

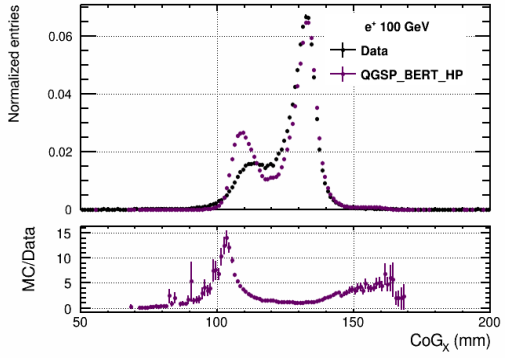
(b)



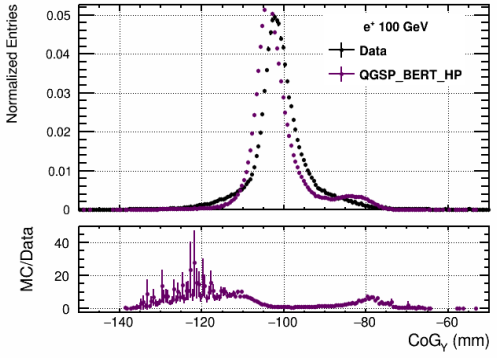
(c)



(d)



(e)



(f)

FIGURE B.2: Center of gravity distributions in x and y-directions using electron showers from testbeam data and simulated data with `QGSP_BERT_HP` physics list.

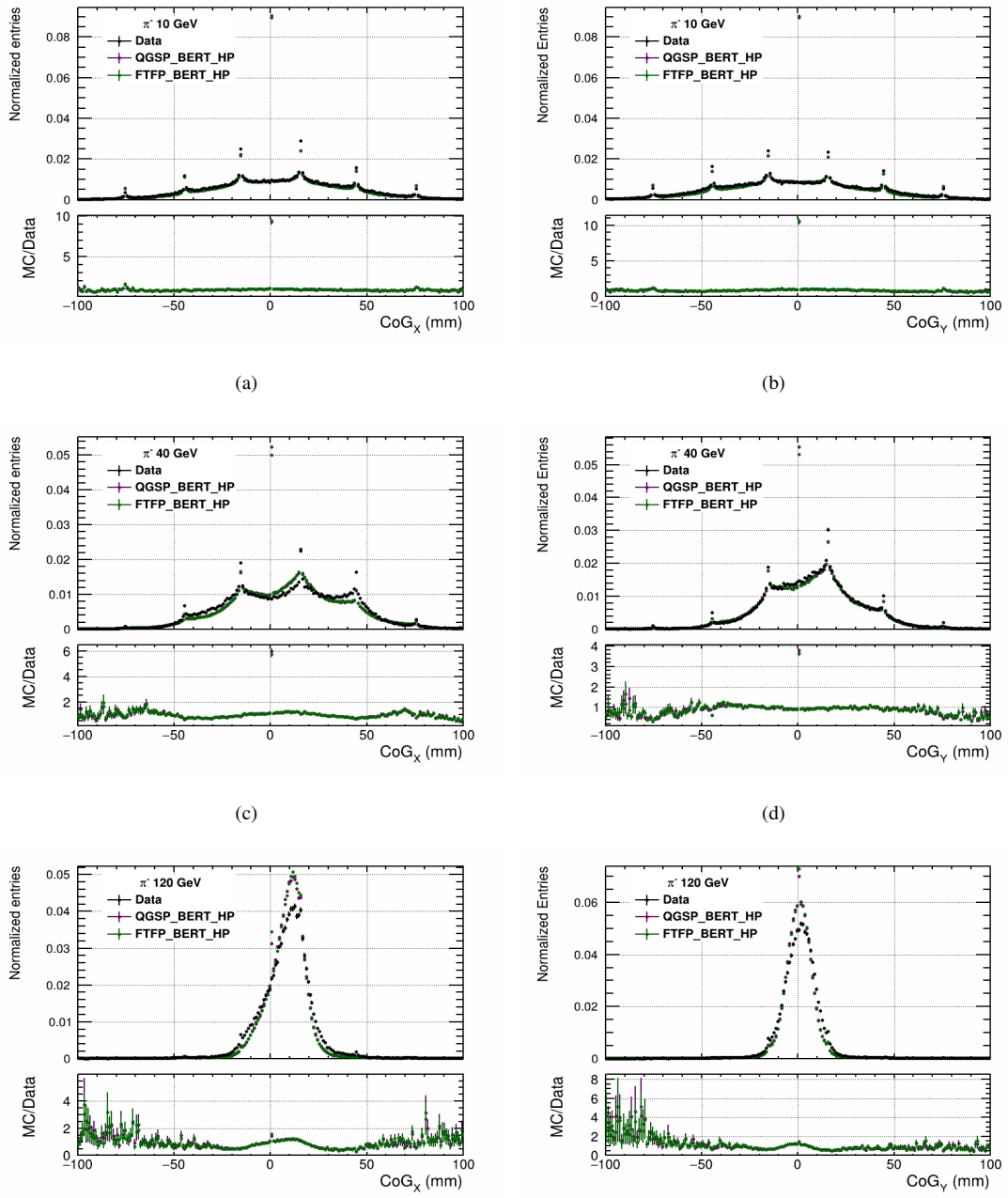


FIGURE B.3: Center of gravity distributions in x and y-directions using pion showers from testbeam data and simulated data with `QGSP_BERT_HP` and `FTFP_BERT_HP` physics lists.

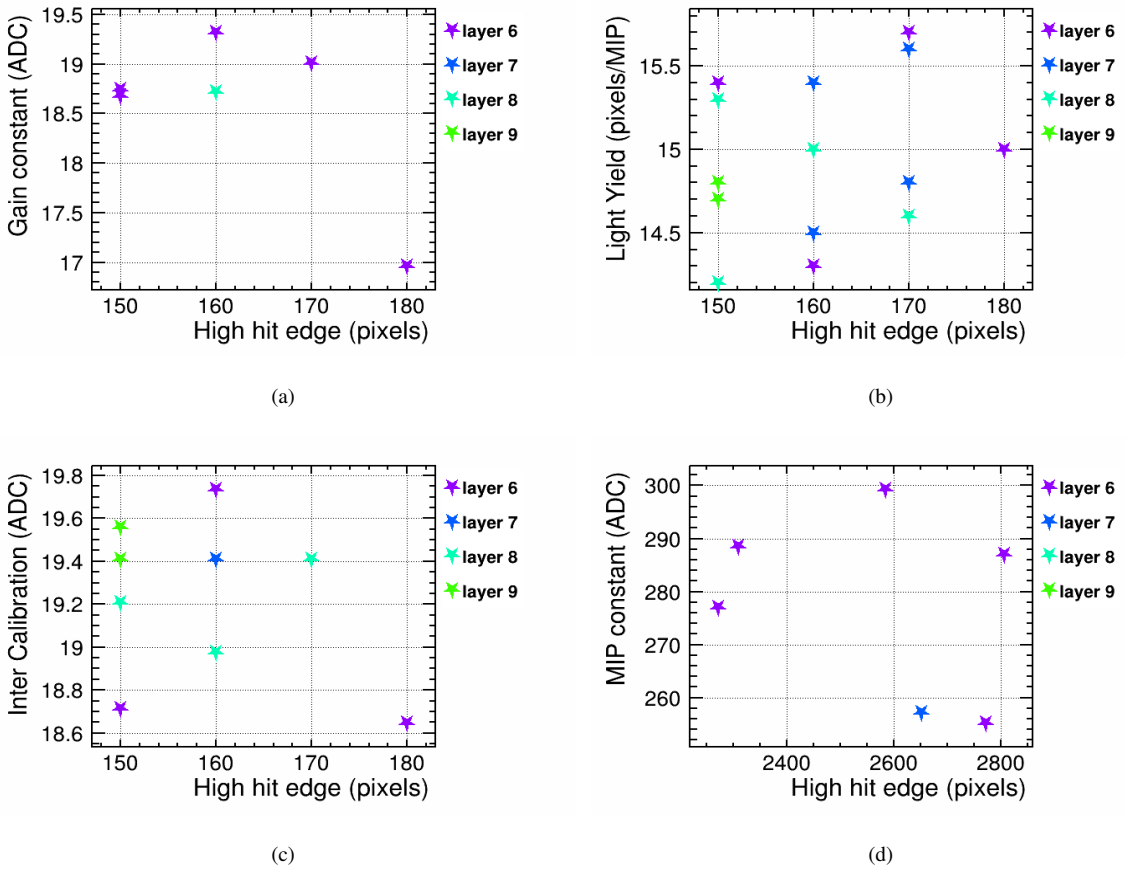
Particle Type	**Energy (GeV)**	**Run Number**	**DWC$_X$ (mm)**	**DWC$_Y$ (mm)**
μ^-	40	61167	-99.54	-57.79
e^+	10	61293	121.33	-98.17
	20	61296	122.14	-98.14
	30	61213	120.15	-98.18
	40	61212	102.34	-97.73
	50	61214	91.44	-97.78
	60	61211	119.13	-99.66
	70	61215	119.57	-97.03
	80	61210	116.90	-102.93
	90	61216	119.61	-97.23
	100	61217	119.87	-99.62
π^-	10	61265	14.18	8.21
	20	61273	14.35	8.00
	30	61384	32.82	-6.29
	40	61275	14.36	8.48
	60	61262	14.22	8.15
	80	61279	14.46	8.27
	120	61287	14.44	8.03
	160	61222	8.97	6.16
	200	61201	6.83	13.74

TABLE B.2: Delay Wire Chamber alignment offsets for June 2018 testbeam data at SPS CERN.

Appendix C

Saturation Studies: Additional Plots

In this appendix, additional plots on the correlation of calibration constant and the high hit edge in pixels for different tile positions of layers 6, 7, 8 and 9 is shown in figure C.1. The values are calculated based on hit energy distributions after applying SiPM saturation correction.



(a) (b)

(c) (d)

FIGURE C.1: Correlation of calibration constants versus the high hit edge in pixels for different positions from layers 6, 7, 8 and 9 using the hit energy distributions after applying SiPM saturation correction. The high hit edge corresponds to the upper edge of the hit energy deposited in 1×1 cm^2 tower of a AHCAL tile. (a) Gain constant in ADC counts versus high hit edge. (b) Light yield versus high hit edge. (c) Inter-Calibration versus High hit edge. (d) MIP constant in ADC counts versus high hit edge.

Appendix D

Shower Shapes: Additional Plots

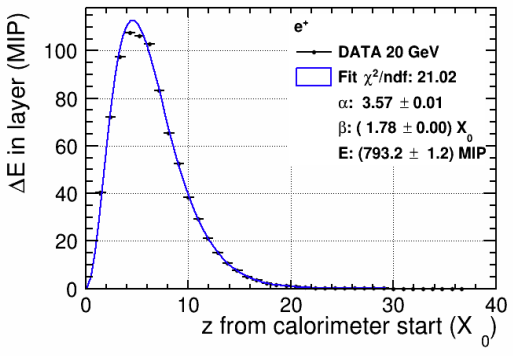
This appendix provides the table with selections applied on beam profiles along x-and y-direction used in shower shape analysis. Furthermore, the distributions of electromagnetic and hadronic shower shapes for the remaining energies is shown here.

Particle Type	**Energy [GeV]**	COG_X **(mm)**	COG_Y **(mm)**
e^+	10	(76, 164)	(-156,-60)
	20	(100, 164)	(-132, -68)
	30	(108, 156)	(-132, -76)
	40	(108, 140)	(-116, -76)
	50	(108, 140)	(-108, -84)
	60	(108, 140)	(-116, -100)
	70	(108, 140)	(-108, -100)
	80	(108, 140)	(-108, -100)
	90	(108, 140)	(-108, -100)
	100	(108, 140)	(-108, -92)
π^-	10	(-52, 68)	(-68, 68)
	20	(-44, 60)	(-52, 52)
	30	(-20, 84)	(-60, 20)
	40	(-44, 60)	(-36, 52)
	60	(-36, 60)	(-36, 20)
	80	(-36, 52)	(-28, 28)
	120	(-12, 28)	(-12, 12)
	160	(-12, 20)	(-12, 12)
	200	(-20, 20)	(4, 20)

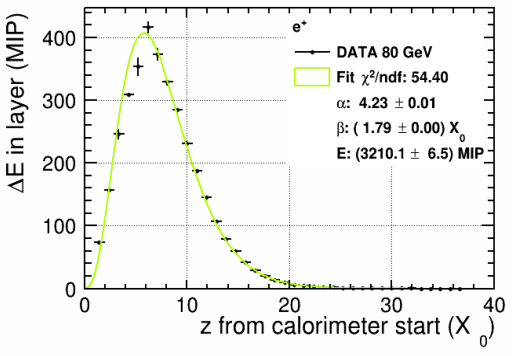
TABLE D.1: Center of gravity range considered for electron and pion data, according to the beam energy.

Electromagnetic Shower Shapes

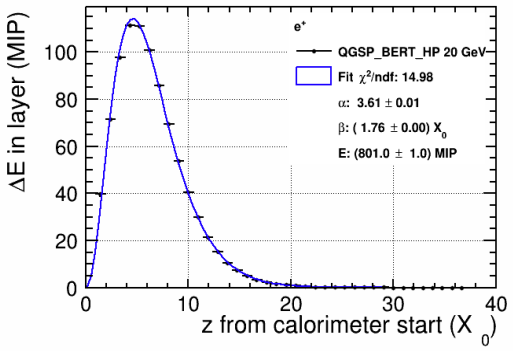
Longitudinal Shower Profiles



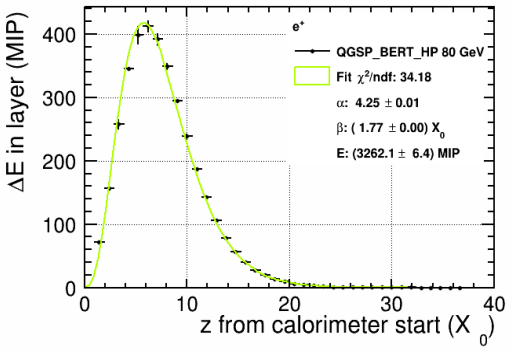
(a)



(b)



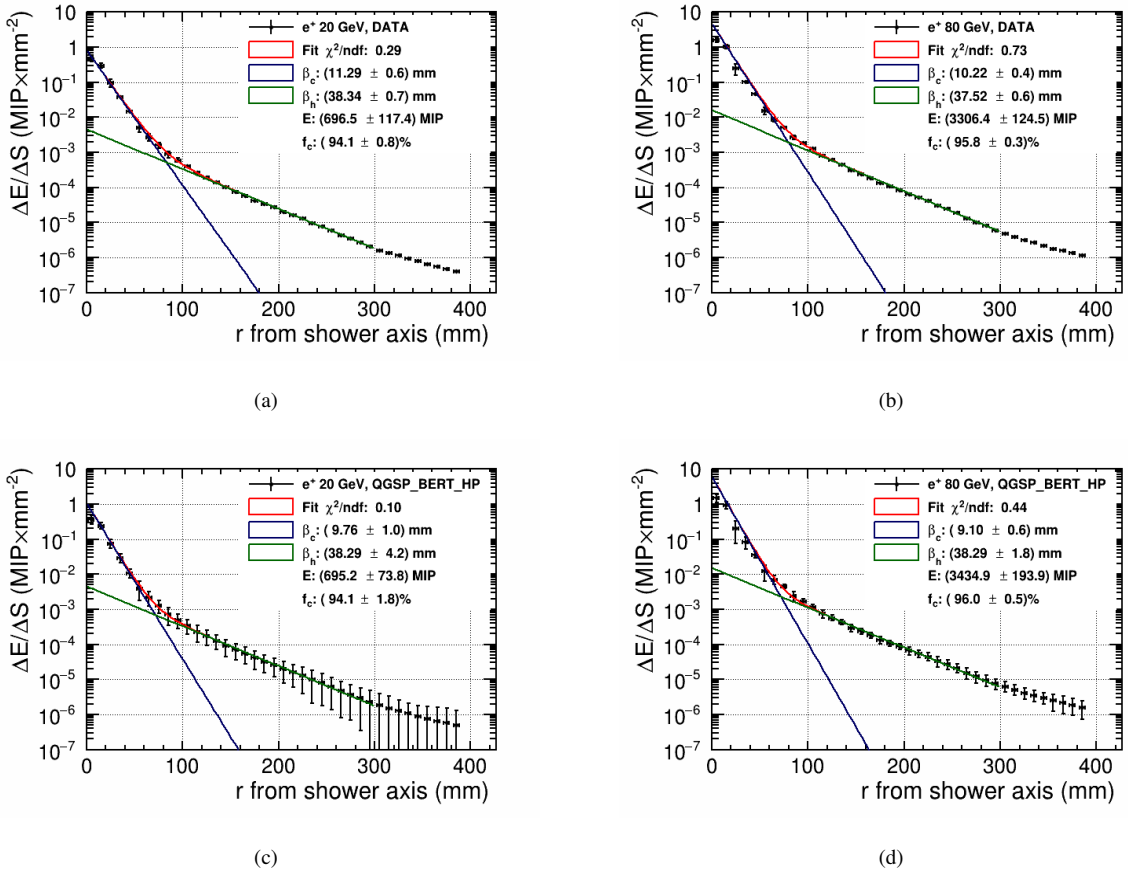
(c)



(d)

FIGURE D.1: Fits performed to electron longitudinal shower profile using Eq. 7.1 shown by solid line. (a) 20 GeV and (b) 80 GeV electron shower from testbeam data. (c) 20 GeV and (d) 80 GeV electron shower from simulated data using QGSP BERT HP physics lists.

Radial Shower Profiles



(a) (b)

(c) (d)

FIGURE D.2: Fits performed to electron radial shower profile using Eq. 7.3 shown by solid line. (a) 20 GeV and (b) 80 GeV electron shower from testbeam data. (c) 20 GeV and (d) 80 GeV electron shower from simulated data using QGSP BERT HP physics lists.

Hadronic Shower Shapes

Longitudinal Shower Profiles

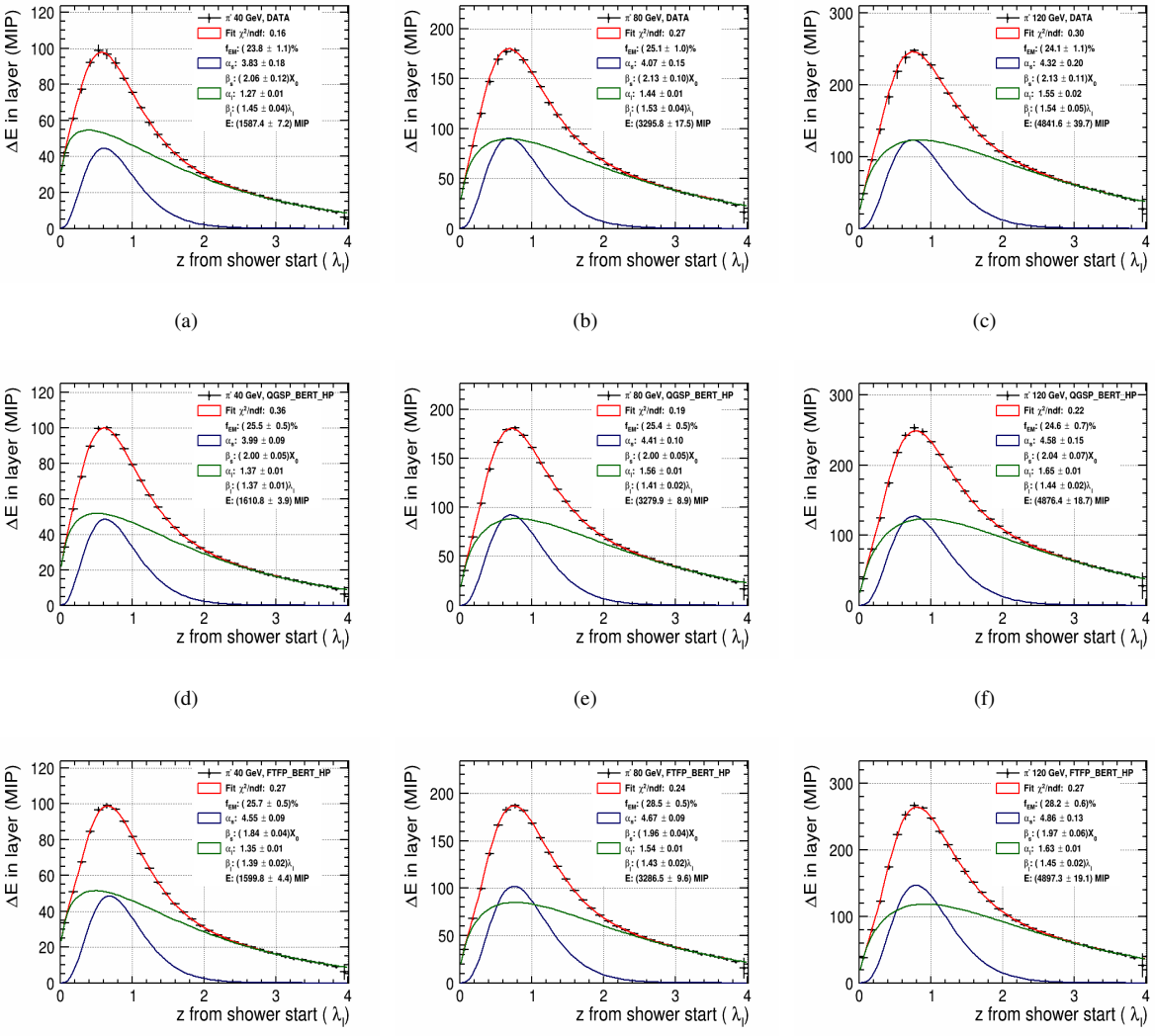


FIGURE D.3: Fits performed to longitudinal shower profile using Eq. 7.1. Testbeam data shown in (a) for 40 GeV, (b) 80 GeV and (c) from 120 GeV pion shower. QGSP_BERT_HP physics list shown in (d) for 40 GeV, (e) 80 GeV and (f) from 120 GeV pion shower. FTFP_BERT_HP physics list shown in (g) for 40 GeV, (h) from 80 GeV and (i) from 120 GeV pion shower. The blue curve represents the "short" component and green curve represents the "long" component.

Radial Shower Profiles

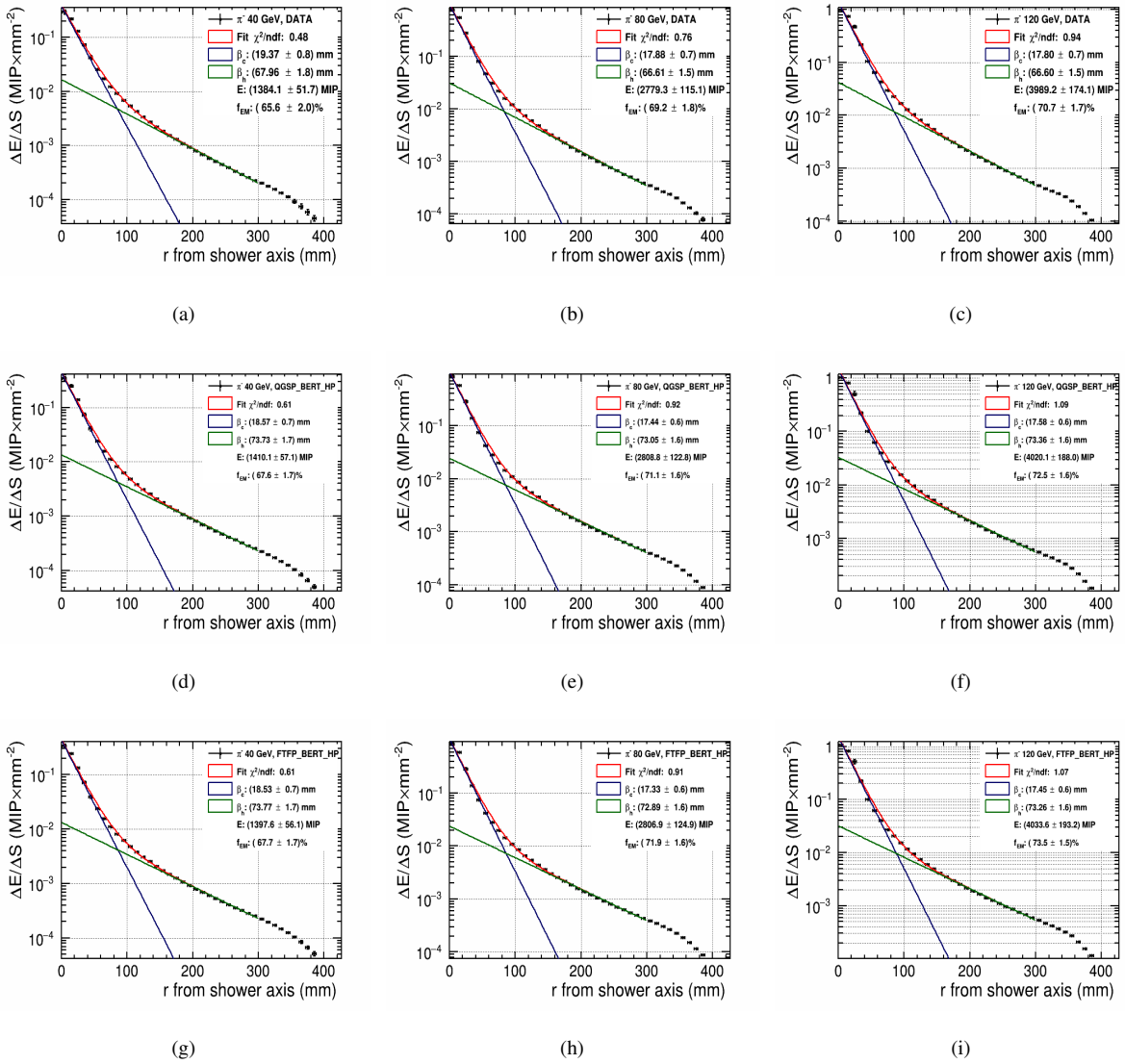
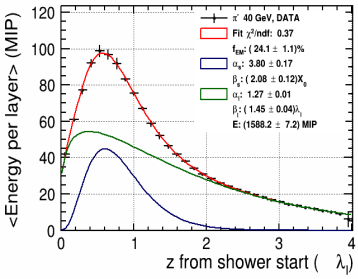
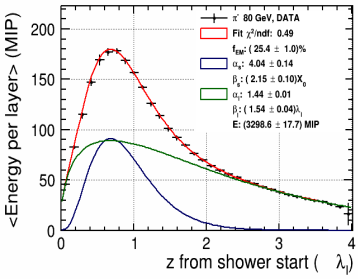


FIGURE D.4: Fits performed to radial shower profile using Eq. 7.8.: Testbeam data shown in (a) from 40 GeV, (b) from 80 GeV and (c) from 120 GeV pion shower. QGSP_BERT_HP physics list shown in (d) from 40 GeV pion shower, (f) from 120 GeV pion shower. FTFP_BERT_HP physics list shown in (g) from 40 GeV, (h) from 80 GeV and (i) from 120 GeV pion shower. The blue curve represents the "core" component and green curve represents the "halo" component.

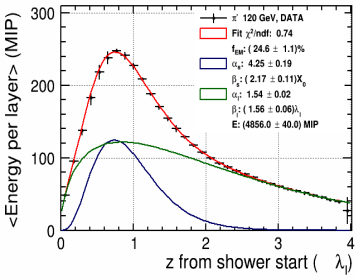
Simultaneous Fitting of Longitudinal and Radial Shower Profile



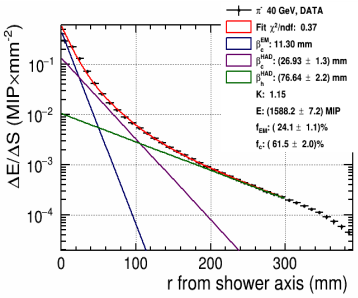
(a)



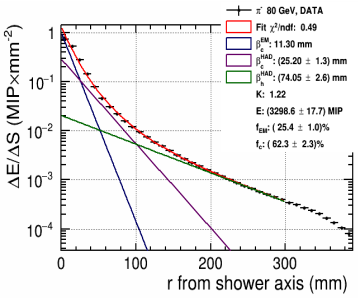
(b)



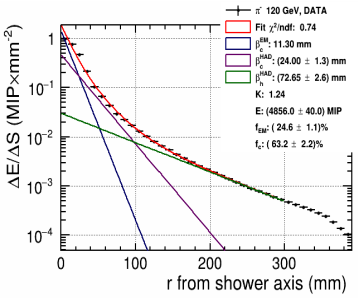
(c)



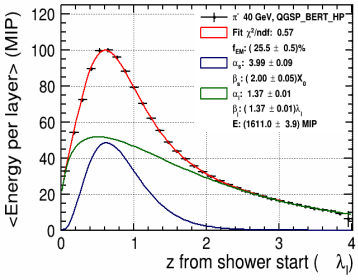
(d)



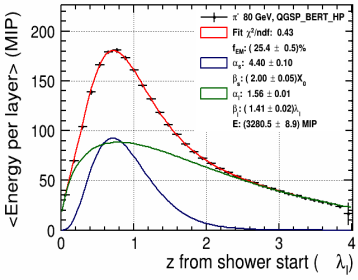
(e)



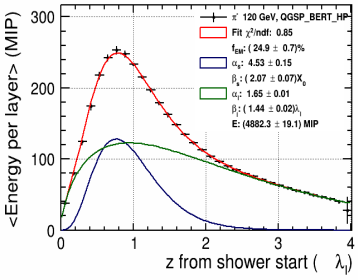
(f)



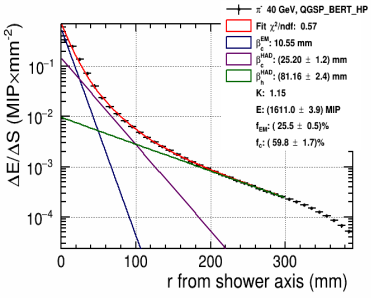
(g)



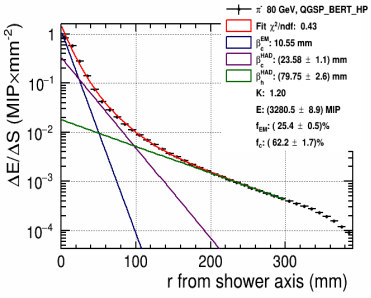
(h)



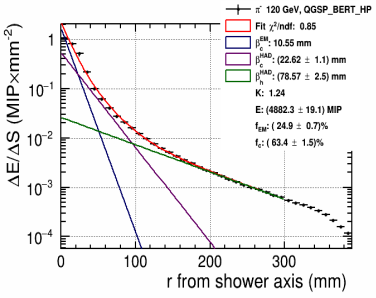
(i)



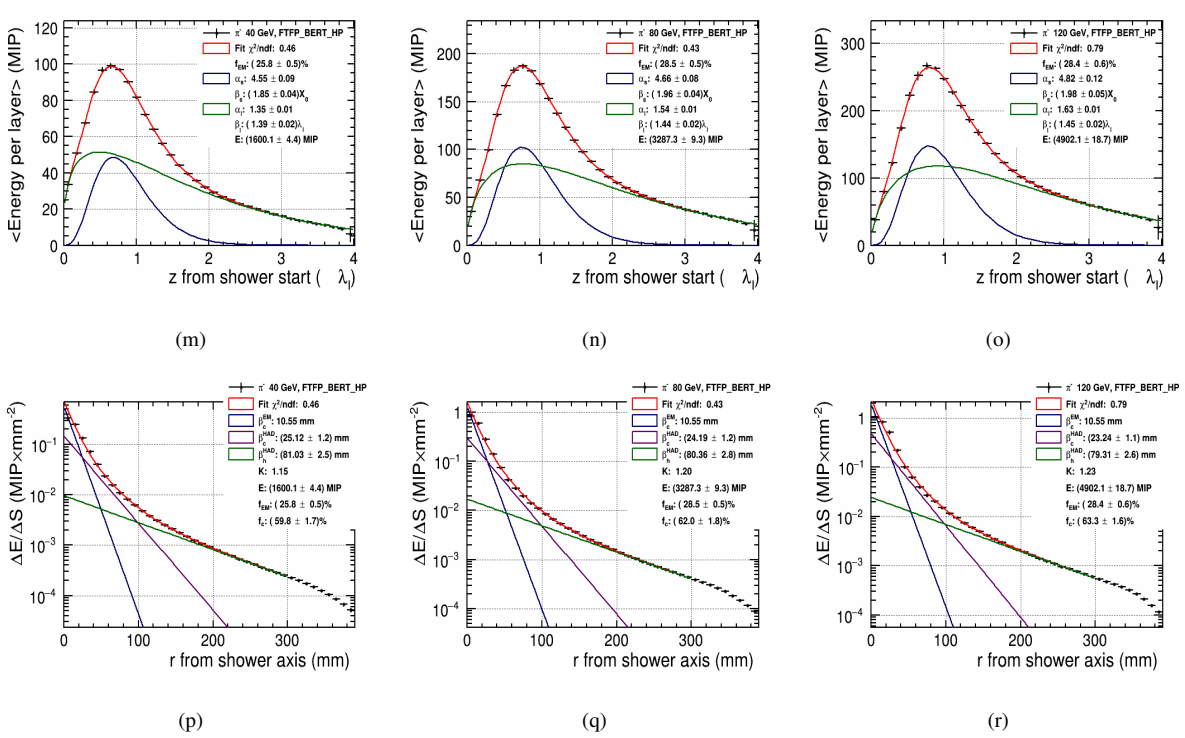
(j)



(k)



(l)

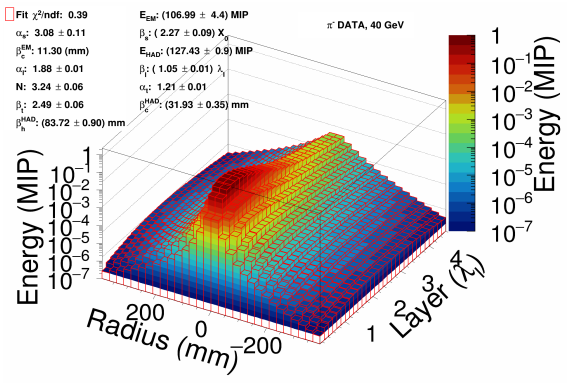


(m) (n) (o)

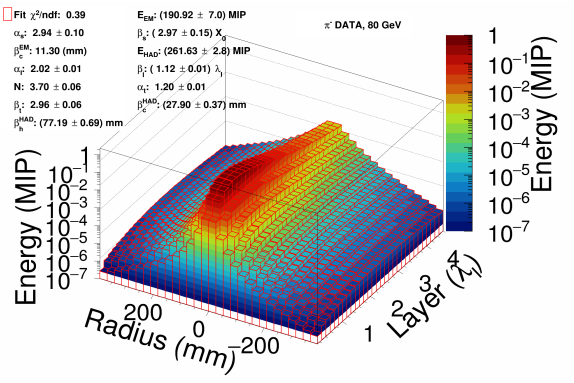
(p) (q) (r)

FIGURE D.5: Simultaneous fits performed to longitudinal and radial shower profile using Eq. and sharing fit parameters E and f_{EM}. Longitudinal profile: shown in (a) from 40 GeV, (b) from 80 GeV and (c) from 120 GeV pion shower using testbeam data. Radial profile: shown in (a) from 40 GeV, (b) from 80 GeV and (c) from 120 GeV pion shower using testbeam data. Longitudinal profile: shown in (g) from 40 GeV, (h) from 80 GeV and (i) from 120 GeV pion shower using QGSP_BERT_HP physics list. Radial profile: shown in (j) from 40 GeV, (k) from 80 GeV and (l) from 120 GeV pion shower using QGSP_BERT_HP physics list. Longitudinal profile: shown in (m) from 40 GeV, (n) from 80 GeV and (o) from 120 GeV pion shower using FTFP_BERT_HP physics list. Radial profile: shown in (p) from 40 GeV, (q) from 80 GeV and (r) from 120 GeV pion shower using FTFP_BERT_HP physics list. The blue curve represents the "core" component and green curve represents the "halo" component.

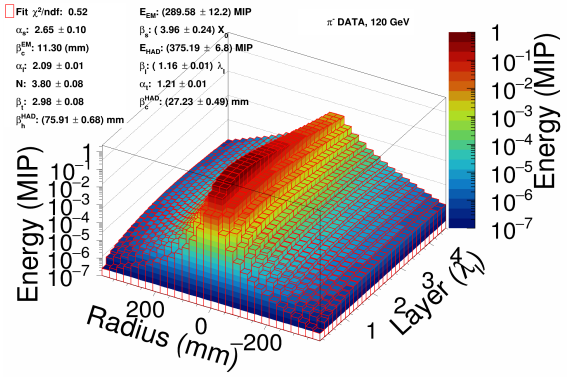
Three-Dimensional Fitting of Shower Shapes



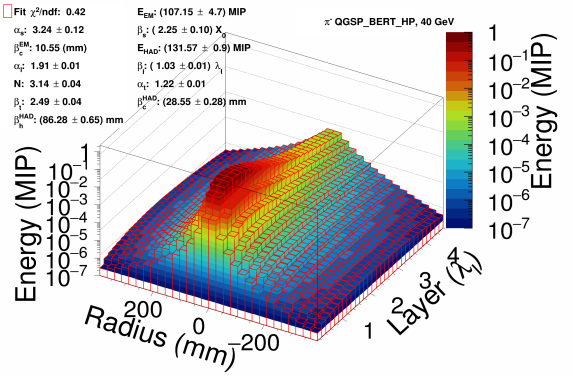
(a)



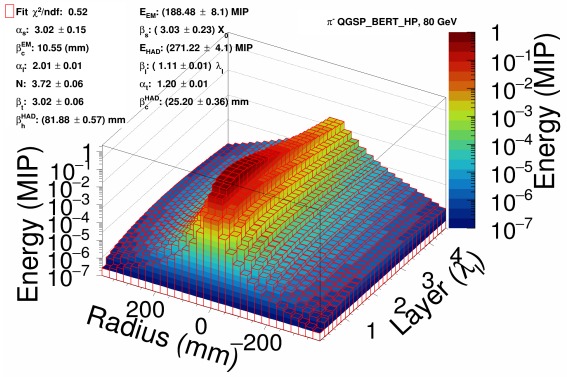
(b)



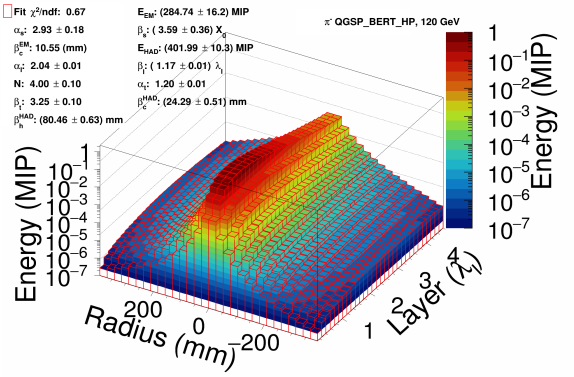
(c)



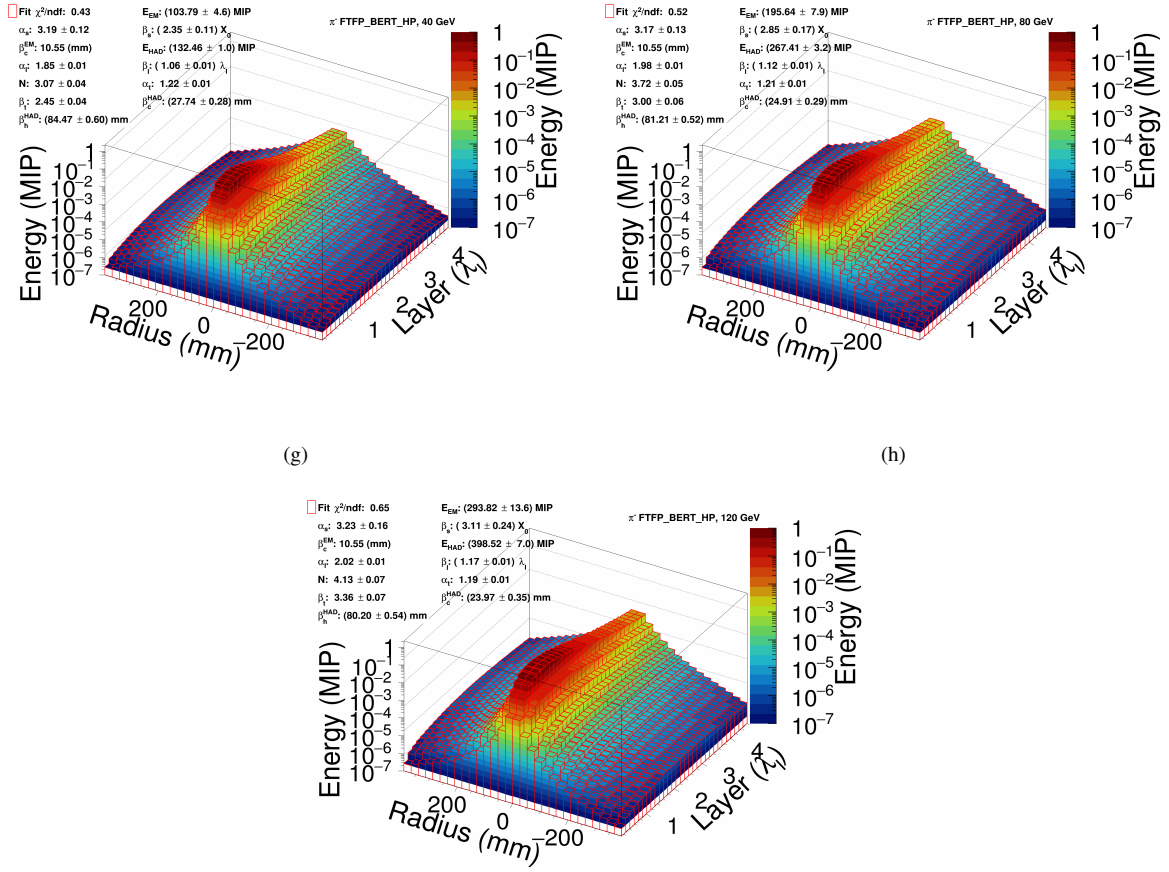
(d)



(e)



(f)



(g) (h)

(i)

FIGURE D.6: Three-dimensional fits performed to pion showers using Eq. 7.22 shown in red lines. Testbeam data shown in (a) at 40 GeV, (b) 80 GeV and (c) 120 GeV. QGSP_BERT_HP physics list shown in (d) at 40 GeV, (e) 80 GeV, (f) 120 GeV. FTFP_BERT_HP physics list shown in (g) at 40 GeV, (h) 80 GeV and (i) 120 GeV.

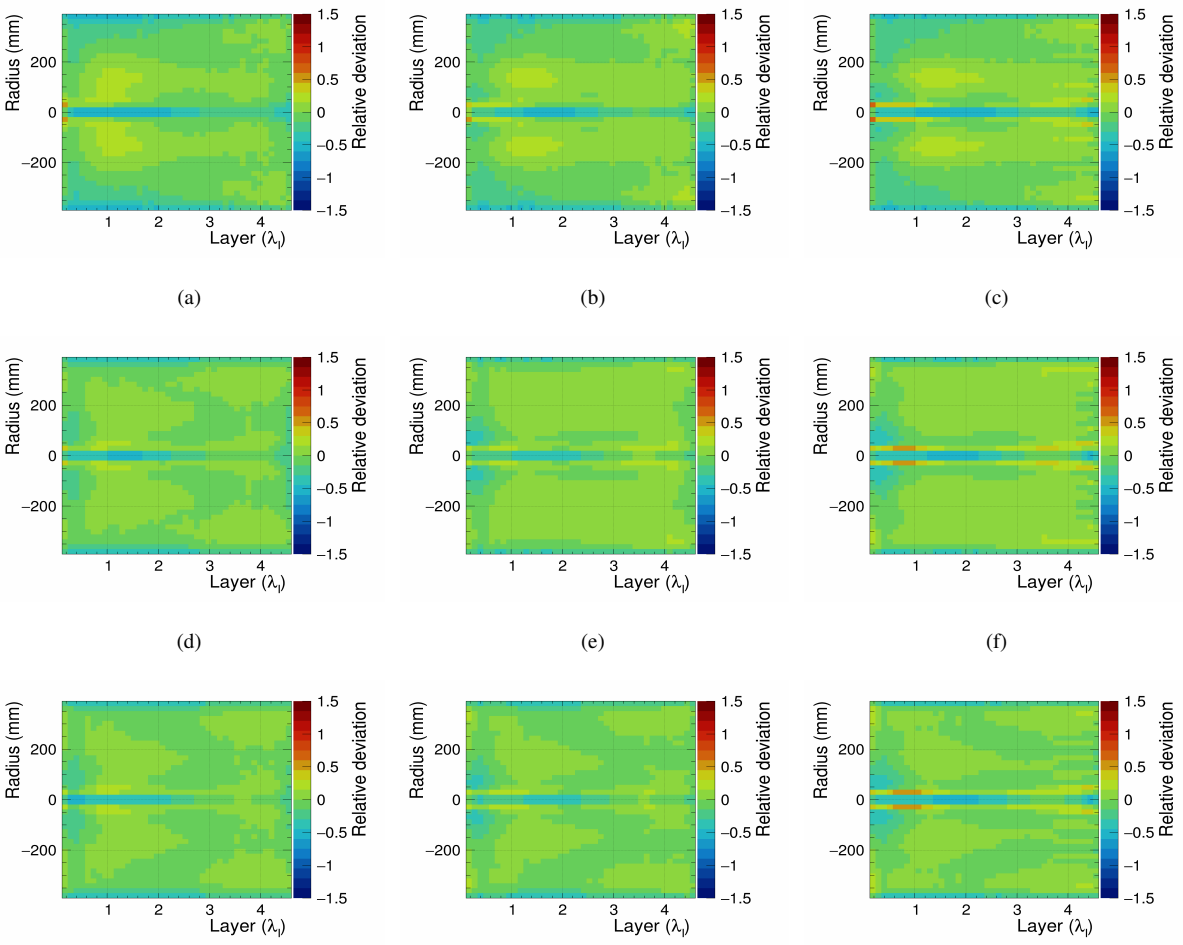


FIGURE D.7: Residual of three-dimensional fits to pion data. For testbeam data shown in (a) at 40 GeV, (b) 80 GeV and (c) 120 GeV. QGSP_BERT_HP physics list shown in (d) at 40 GeV, (e) 80 GeV, (f) 120 GeV. FTFP_BERT_HP physics list shown in (g) at 40 GeV, (h) 80 GeV and (i) from 120 GeV.

Appendix E

Fitting Tool

When fitting a function that involves many parameters is cumbersome, finding the optimal start value that gives the best fit to data is often tricky and tedious. In this thesis, to determine the optimal parameters, a fitting program provided by the ROOT [114] was used. In this fitting method, the TH1::Fit requires the user to provide the fitting functions and starting values for the variables used. The analysis follows a Chi-squared (χ^2) minimization, where parameters are varied in small steps until convergence is achieved, which minimizes the χ^2 between the data and the values obtained from the fit function.

A C++ code written by the author was developed to perform a high number of iterations quickly. The snippet of the code and the fit function used can be found in[1]. The optimal parameters are first found using the values from previous physics prototype studies and further optimized by a trial and error approach. The output value after each iteration is used as the initial value for the next iteration. This is repeated until a stable χ^2 is obtained. The χ^2 minimization is performed using the MINUIT2 [115] package offered by ROOT. In between the iterations, the parameters occasionally converge to a non-optimal limit. In such cases, slightly modifying the optimal parameters is a further trial-and-error approach. When the ranges are not set, the parameters can diverge to infinity. Once a convergence is found, the output values are returned by this method.

[1] https://github.com/olinpinto/Shower_Shapes

List of Tables

List of Figures

Acronyms

ADC **A**nalog to **D**igital **C**onverter
AHCAL **A**nalog **H**adron **CAL**orimeter
APD **A**valanche **P**hoto**d**iodes
AT **A**uto **T**rigger

BDT **B**oosted **D**ecision **T**echnique
BEH **B**rout-**E**nglert-**H**iggs
BSM **B**eyond **S**tandard **M**odel
bxID **b**unch crossing **ID**

CALICE **CA**lorimeter for **LI**near **C**ollider **E**xperiment
CMB **C**osmological **M**icrowave **B**ackground
CoG **C**enter **o**f **G**ravity
CP Charge Parity

DAC **D**igital to **A**nalog **C**onverter
DAQ **D**ata **Acq**uisition
DHCAL **D**igital **H**adron **Cal**orimeter

EBU **E**CAL **B**ase **U**nit
ECAL **E**lectromagnetic **Cal**orimeter
ET **E**xternal **T**rigger

GEM **G**as **E**lectron **M**ultiplier
GRPC **G**lass **R**esistive **P**late **C**hambers
GUTs **G**rand **U**nified **T**heorie**s**

HBU **H**adronic Calorimeter **B**ase **U**nits
HCAL **H**adronic **Cal**orimeter
HEP **H**igh **E**nergy **P**hysics
HG **H**igh-**G**ain
HL-LHC **H**igh **L**uminosity **L**arge Hadron **C**ollider

ILC **I**nternational **L**inear **C**ollider
ILD **I**nternational **L**arge **D**etector

K-S test **K**olmogorov-**S**mirnov **t**est

LED **L**ight **E**mitting **D**iode

LG **L**ow-**G**ain
LHC **L**arge **H**adron **C**ollider

MARLIN **M**odular **A**nalysis and **R**econstruction for the **Lin**ear Collider
MC **M**onte **C**arlo
Micromegas **Micro**-**Me**sh **Gas**eous Structure
MIP **M**inimum **I**onising **P**article
MPV **M**ost **P**robable **V**alue

PCB **P**rinted **C**ircuit **B**oard
PDE **P**hoton **D**etection **E**fficiency
PFA **P**article **F**low **A**pproach
PID **P**article **Id**entification
PS **P**re-**S**hower

QCD **Q**uantum **C**hromo**D**ynamics

ROC **R**eceiver **O**perating **C**haracteristic
RPC **R**esistive **P**late **C**hambers

Sc-ECAL **Sc**intillator strip-Tungsten **ECAL**
SCRF **S**uperconducting **R**adio **F**requency
SDHCAL **S**emi-**D**igital **H**adron **Cal**orimeter
SID **Si**licon **D**etector
SiPM **Si**licon **P**hotoMultipliers
SiW-ECAL **Si**licon-Tungsten **ECAL**
SM **S**tandard **M**odel
SPIROC **S**i**P**M **I**ntegrated **R**ead-**O**ut **C**hips
SPS **S**uper **P**roton **S**ynchrotron

TC **T**ail **C**atcher
TCMT **T**ail **C**atcher **M**uon **T**racker
TDC **T**ime to **D**igital **C**onverter
TPC **T**ime **P**rojection **C**hamber

VBF **V**ector **B**oson **F**usion
VTX **V**er**tex** **D**etector

WLS **W**ave-**L**ength **S**hifting

Bibliography

[1] J.J. Thomson. XL. Cathode rays. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 44(269):293–316, 1897.

[2] E. Rutherford. The scattering of α and β particles by matter and the structure of the atom. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 21 (125):669–688, 1911.

[3] S. L. Glashow. Partial-symmetries of weak interactions. *Nuclear physics*, 22(4):579–588, 1961.

[4] A. Salam and J. C. Ward. Electromagnetic and weak interactions. *Physics Letters*, 13(2): 168–171, Nov 1964. ISSN 0031-9163. URL https://doi.org/10.1016/0031-9163(64)90711-5.

[5] S. Weinberg. A model of leptons. *Physical review letters*, 19(21):1264, 1967. URL https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.19.1264.

[6] A. Purcell. Go on a particle quest at the first CERN webfest. Le premier webfest du CERN se lance à la conquête des particules, Aug 2012. URL https://cds.cern.ch/record/1473657. (visited on 17/02/2022).

[7] H. Fritzsch et al. Advantages of the color octet gluon picture. *Physics Letters B*, 47(4): 365–368, 1973.

[8] V. M. Abazov et al. Evidence for a $B_s^0\pi\pm$ State. *Physical Review Letters*, 117(2):42, 2016. ISSN 10797114. URL http://dx.doi.org/10.1103/PhysRevLett.117.022003.

[9] R. Aaij et al. Observation of J/ψ p Resonances Consistent with Pentaquark States in $\Lambda_b^0 \to$ J/ψ K^- p Decays. *Physical review letters*, 115(7):072001, 2015.

[10] P.A. Zyla et al. (Particle Data Group). Review of particle physics. *Progress of Theoretical and Experimental Physics*, 2020(8):083C01, 2020.

[11] Gerald S Guralnik, Carl R Hagen, and Thomas WB Kibble. Global conservation laws and massless particles. *Physical Review Letters*, 13(20):585, 1964.

[12] P. W. Higgs. Broken symmetries and the masses of gauge bosons. *Physical Review Letters*, 13(16):508, 1964.

[13] François Englert and Robert Brout. Broken symmetry and the mass of gauge vector mesons. *Physical Review Letters*, 13(9):321, 1964.

[14] The CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Physics Letters, Section B: Nuclear, Elementary Particle and High-Energy Physics*, 2012.

[15] The ATLAS Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *arXiv preprint arXiv:1207.7214*, 2012.

[16] M. Veltman et al. Regularization and renormalization of gauge fields. *Nuclear Physics B*, 44(1):189–213, 1972.

[17] S. Dimopoulos and H. Georgi. Softly broken supersymmetry and SU (5). *Nuclear Physics B*, 193:150–162, 1981.

[18] E. S. Battistelli et al. Galaxy clusters as probes for cosmology and dark matter. *International Journal of Modern Physics D*, 25(10):1630023, 2016.

[19] Gaëlle Giesen et al. CMB photons shedding light on dark matter. *Journal of Cosmology and Astroparticle Physics*, 2012(12):008, 2012.

[20] Ubaldo Dore and Domizia Orestano. Experimental results on neutrino oscillations. *Reports on Progress in Physics*, 71(10):106201, 2008.

[21] A Rodrigues Vieira et al. Naturalness and theoretical constraints on the Higgs boson mass. *International Journal of Theoretical Physics*, 52(10):3494–3503, 2013.

[22] Andrei D Sakharov. Violation of CP-invariance, C-asymmetry, and baryon asymmetry of the Universe. In *In The Intermissions... Collected Works on Research into the Essentials of Theoretical Physics in Russian Federal Nuclear Center, Arzamas-16*, pages 84–87. World Scientific, 1998.

[23] C. T. Hill and E. H. Simmons. Strong dynamics and electroweak symmetry breaking. *Physics Reports*, 381(4-6):235–402, 2003.

[24] J. Wess and B. Zumino. Supergauge transformations in four dimensions. *Nuclear Physics B*, 70(1):39–50, 1974.

[25] Shreyashi Chakdar. *New physics at the TeV scale*. PhD thesis, Oklahoma State University, 2015.

[26] T. Behnke et al. The international linear collider technical design report-volume 1: Executive summary. *arXiv preprint arXiv:1306.6327*, 2013.

[27] H. Baer et al. The international linear collider technical design report-volume 2: physics. *arXiv preprint arXiv:1306.6352*, 2013.

[28] Chris Adolphsen et al. The international linear collider technical design report-volume 3. ii: Accelerator baseline design. *arXiv preprint arXiv:1306.6328*, 2013.

[29] K. Fujii et al. Physics case for the international linear collider. *arXiv preprint arXiv:1506.05992*, 2015.

[30] J. Yan et al. Measurement of the Higgs boson mass and $e^+e^- \to$ ZH cross section using Z$\to \mu^+\mu^-$ and Z$\to e^+e^-$ at the ILC. *Physical Review D*, 94(11):113002, 2016.

[31] T. Han, Z.. Liu, and J. Sayre. Potential precision on Higgs couplings and total width at the ILC. *Physical Review D*, 89(11):113006, 2014.

[32] Tony Price, Philipp Roloff, Jan Strube, and Tomohiko Tanabe. Full simulation study of the top Yukawa coupling at the ILC at $\sqrt{s}$= 1 TeV. *The European Physical Journal C*, 75: 1–9, 2015.

[33] M.A. Thomson. Particle flow calorimetry and the PandoraPFA algorithm. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 611(1):25–40, 2009.

[34] D. Buskulic et al. Performance of the ALEPH detector at LEP. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 360(3):481–506, 1995.

[35] M. Peez. Search for deviations from the Standard Model in high transverse energy processes at the electron-proton collider HERA. Technical report, Aix-Marseille-2 Univ., 2003.

[36] B. Portheault. Determination of electroweak parameters at HERA with the H1 experiment. In *AIP Conference Proceedings*, volume 792, pages 563–566. American Institute of Physics, 2005.

[37] CMS Collaboration, CMS Collaboration, et al. Particle-flow event reconstruction in CMS and performance for jets, taus and MET, 2009.

[38] H. L. Tran et al. Software compensation in particle flow reconstruction. *The European Physical Journal C*, 77(10):1–13, 2017.

[39] T. Behnke et al. Ilc reference design report volume 4-detectors. *arXiv preprint arXiv:0712.2356*, 2007.

[40] R. Diener, LCTPC Collaboration, et al. Development of a TPC for an ILC Detector. *Physics Procedia*, 37:456–463, 2012.

[41] B Rossi and High-energy Particles. Prentice-Hall physics series. *New York*, 1952.

[42] Particle Data Group et al. Constants, units, atomic and nuclear properties. *Physics Letters B*, 667(1-5):103–115, 2008.

[43] Bruno Rossi and Kenneth Greisen. Cosmic-Ray Theory. *Rev. Mod. Phys.*, 13:240–309, Oct 1941. doi: 10.1103/RevModPhys.13.240. URL https://link.aps.org/doi/10.1103/RevModPhys.13.240.

[44] Claude Leroy and Pier-Giorgio Rancoita. Physics of cascading shower generation and propagation in matter: Principles of high-energy, ultrahigh-energy and compensating calorimetry. *Reports on Progress in Physics*, 63(4):505, 2000.

[45] Egidio Longo and Ignazio Sestili. Monte Carlo calculation of photon-initiated electromagnetic showers in lead glass. *Nuclear Instruments and Methods*, 128(2):283–307, 1975.

[46] R Wigmans. Calorimetry: Energy measurement in particle physics, Published in Int. *Ser. Monogr. Phys*, 107:1–726, 2000.

[47] G. Grindhammer, M Rudowicz, and S Peters. The fast simulation of electromagnetic and hadronic showers. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 290(2-3):469–488, 1990.

[48] R. Wigmans. New Calorimeter Techniques. In *Calorimetry*. Oxford University Press, 2000.

[49] T. A. Gabriel et al. Energy dependence of hadronic activity. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 338:336–347, 1994.

[50] D. E. Groom. Energy flow in a hadronic cascade: Application to hadron calorimetry. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 572:633–653, 2007.

[51] A.V. Stirling et al. Small-Angle Charge Exchange of π- Mesons between 6 and 18 GeV/c. *Physical Review Letters*, 14:763, 1965.

[52] A.V. Barnes et al. Pion charge-exchange scattering at high energies. *Physical Review Letters*, 37, 1976.

[53] A. Bernstein et al. Beam tests of the zeus barrel calorimeter. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 336:23–52, 1993.

[54] N. Akchurin et al. Dual-readout calorimetry. *arXiv:1307.5538*, 2013.

[55] F. Cavallari. Performance of calorimeters at the LHC. In *Journal of Physics: Conference Series*, volume 293, page 012001. IOP Publishing, 2011. URL https://iopscience.iop.org/article/10.1088/1742-6596/293/1/012001/pdf.

[56] M. Livan. XIII International Conference on Calorimetry in High Energy Physics (CALOR 2008). In *Journal of Physics: Conference Series*, volume 160, page 011001. IOP Publishing, 2009.

[57] S. Agostinelli et al. GEANT4—a simulation toolkit. *Nuclear instruments and methods in physics research section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250–303, 2003.

[58] Hadronic Physics in Geant4 Physics Reference Manual 10.7 documentation, Sept. 2021. URL https://geant4-userdoc.web.cern.ch/UsersGuides/PhysicsReferenceManual/html/hadronic/index.html.

[59] User Documentation geant4.web.cern.ch, Sept. 2021. URL https://geant4.web.cern.ch/support/user_documentation.

[60] R. Pöschl. Private Communication to CALICE Collaboration, 2019.

[61] Aatos Heikkinen, Nikita Stepanov, and Johannes Peter Wellisch. Bertini intra-nuclear cascade implementation in Geant4. *CHEP*, 2003. URL http://arxiv.org/abs/nucl-th/0306008.

[62] G Folger and Wellisch. J P. Parton string models in geant4. In *Monte Carlo 2005 Topical Meeting*, pages 515–524, 2005.

[63] C. Adloff et al. The time structure of hadronic showers in highly granular calorimeters with tungsten and steel absorbers. *Journal of Instrumentation*, 9(07):P07022, 2014.

[64] The CALICE Collaboration, September 2021. URL https://twiki.cern.ch/twiki/bin/view/CALICE/WebHome. CALICE Website.

[65] Burkhard Schmidt. The High-Luminosity upgrade of the LHC: Physics and Technology Challenges for the Accelerator and the Experiments. *Journal of Physics: Conference Series*, 706(2):022002, 2016.

[66] P Göttlicher. System aspects of the ILC-electronics and power pulsing, 2007.

[67] C. Neubuser. *Comparison of Two Highly GranularHadronic Calorimeter Concepts*. PhD thesis, University of Hamburg, Hamburg, 2016. https://bib-pubdb1.desy.de/record/314943/files/Thesis_pub.pdf.

[68] The CALICE Collaboration. Construction and Response of a Highly Granular Scintillator-based Electromagnetic Calorimeter. *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 887:150–168, 2018. ISSN 01689002. doi: 10.1016/j.nima.2014.06.039. URL https://arxiv.org/pdf/1707.07126.pdf.

[69] The CALICE Collaboration. Design and electronics commissioning of the physics prototype of a Si-W electromagnetic calorimeter for the International Linear Collider. *Journal of Instrumentation*, 3(08):P08001, 2008. URL https://arxiv.org/pdf/0805.4833.pdf.

[70] Oskar Hartbrich. *Scintillator calorimeters for a future linear collider experiment*. PhD thesis, Bergische Universität Wuppertal, Wuppertal, 2016. DOI:10.3204/PUBDB-2016-02800.

[71] The CALICE Collaboration, September 2021. URL https://twiki.cern.ch/twiki/bin/view/CALICE/ScEcal. CALICE Website.

[72] C Adloff and the CALICE Collaboration. Response of the CALICE Si-W Electromagnetic Calorimeter Physics Prototype to Electrons. *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 2008. URL https://arxiv.org/pdf/0811.2354.pdf.

[73] T. Suehara et al. Performance study of SKIROC2/A ASIC for ILD Si-W ECAL. *Journal of Instrumentation*, 13(03):C03015, 2018.

[74] B. Bilki. Construction and Testing of the CALICE Digital Hadron Calorimeter. In *Proceedings, International Conference on Calorimetry for the High Energy Frontier (CHEF 2013): Paris, France*, page 168–174, 2013.

[75] B. Bilki et al. The CALICE digital hadron calorimeter: calibration and response to pions and positrons. In *Journal of Physics: Conference Series*, page 012038. IOP Publishing, 2015. doi: 10.1088/1742-6596/587/1/012038.

[76] G. Baulieu et al. Construction and commissioning of a technological prototype of a high-granularity semi-digital hadronic calorimeter. *Journal of Instrumentation*, 10(10): P10039, 2015.

[77] Frederic Dulucq et al. Hardroc: readout chip for calice/eudet digital hadronic calorimeter. In *IEEE Nuclear Science Symposuim & Medical Imaging Conference*, pages 1678–1683. IEEE, 2010.

[78] The CALICE Collaboration. First results of the CALICE SDHCAL technological prototype. *Journal of Instrumentation*, 11(04):P04001, 2016.

[79] C. Adloff et al. Construction and commissioning of the CALICE analog hadron calorimeter prototype. *Journal of Instrumentation*, 5(05):P05004, 2010.

[80] N. Feege. Analysis of low-energetic electron and pion data collected with the AHCAL prototype at Fermilab. *CALICE Analysis Note*, 34, 2011.

[81] K. Francis. *Results of beam tests of a prototype calorimeter for a Linear Collider*. PhD thesis, Northern Illinois University, 2010.

[82] The CALICE Collaboration. Design, Construction and Commissioning of a Technological Prototype of a Highly Granular SiPM-on-tile Scintillator-Steel Hadronic Calorimeter. *JINST*, 2022. In Collaboration review.

[83] *MPPC (Multi-Pixel Photon Counter), S13360 series*, 2016. URL http://www.hamamatsu.com/resources/pdf/ssd/s13360_series_kapd1052e.pdf. Photonics, HAMAMATSU.

[84] B Dolgoshein and SiPM Collaboration. Silicon photomultipliers in particle physics: possibilities and limitations. In *Innovative Detectors for Supercolliders*, pages 442–456. World Scientific, 2004.

[85] *MPPC S13360-1325PE | Hamamatsu Photonics:MPPC (SiPM) S13360-1325PE*, June 2021. URL https://www.hamamatsu.com/us/en/product/type/S13360-1325PE/index.html.

[86] Felix Sefkow, Frank Simon, CALICE Collaboration, et al. A highly granular SiPM-on-tile calorimeter prototype. In *Journal of Physics: Conference Series*, page 012012. IOP Publishing, 2019.

[87] LAL/IN2P3 Omega. *SPIROC2 User Guide*, 2009. URL https://indico.cern.ch/event/232082/contributions/493652/attachments/385835/536707/Spiroc2_userGuide2009_datasheet.pdf.

[88] M. Frank et al. DD4hep: A Detector Description Toolkit for High Energy Physics Experiments. In *Journal of Physics: Conference Series*, volume 513, page 022010. IOP Publishing, 2014.

[89] M. Hirschberg et al. Precise measurement of Birks kB parameter in plastic scintillators. *IEEE Transactions on Nuclear Science*, 39:511–514, 1992.

[90] K. A Tadday. *Scintillation light detection and application of silicon photomultipliers in imaging calorimetry and positron emission tomography*. PhD thesis, Ruperto-Carola-University of Heidelberg, Germany, 2011.

[91] F Gaede. Marlin and LCCD—Software tools for the ILC. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 559:177–180, 2006.

[92] S. Martens. Tile reflector wrapping. *technical staff of the group Particle Physics & Detector Development at University Hamburg*, December 2017. https://agenda.linearcollider.org/event/7807/contributions/40519/attachments/32551/49482/TileWrapping-CollMeeting2017.pdf.

[93] *H2 SPS beamline*, 2017. URL https://sba.web.cern.ch/sba/BeamsAndAreas/H2/H2_presentation.html.

[94] N Abgrall et al. NA61/SHINE facility at the CERN SPS: beams and detector system. *Journal of Instrumentation*, 9(06):P06005, 2014.

[95] B Acar et al. Construction and commissioning of CMS CE prototype silicon modules. *Journal of Instrumentation*, 16(04):T04002, 2021. URL https://arxiv.org/abs/2012.06336.

[96] V. Chmill et al. On the characterisation of SiPMs from pulse-height spectra. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 854:70–81, may 2017. ISSN 01689002. doi: 10.1016/j.nima.2017.02.049. URL https://linkinghub.elsevier.com/retrieve/pii/S0168900217302334.

[97] D. Heuchel. *Particle Flow Studies with Highly Granular Calorimeter Data*. PhD thesis, Universität Heidelberg, Heidelberg, 2022. PhD thesis in progress.

[98] Y. Sudo. HG/LG Inter-Calibration and Gain Stability. *CALICE AHCAL main meeting DESY*, December 2018. https://agenda.linearcollider.org/event/7799/contributions/42262/attachments/33592/51488/TempComp_HGLGIC_201809_calice.pdf.

[99] L. Emberger. Timing calibration. *CALICE Meeting CERN*, October 2019. https://agenda.linearcollider.org/event/8213/contributions/44356/attachments/34776/53688/EmbergerTimingCERN.pdf.

[100] F. Simon C. Graf. Time Analysis of the Partially Equipped CALICE Analog Hadronic Calorimeter Technological Prototype with Tungsten Absorber. *CALICE-CAN-2019-002*, October 2019. http://cds.cern.ch/record/2692839/files/CALICE-CAN-2019-002.pdf.

[101] D. Heuchel. Shower Start Finding and Pandora PFA Studies - AHCAL 2018 Prototype. *CALICE Meeting CERN*, October 2019. https://agenda.linearcollider.org/event/8213/contributions/44345/attachments/34810/53754/DH_calice_02102019_ssf_pandora_v2.pdf.

[102] Jonas M. Mikhaeil. Topological Shower Reconstruction in a Highly Granular Calorimeter with Optical Readout. Bachelor thesis, Universität Heidelberg, 2019.

[103] V. Bocharnikov. Particle ID in the AHCAL + Single Shower Substructure. *CALICE Meeting CERN*, September 2020. https://agenda.linearcollider.org/event/8608/contributions/46466/attachments/35887/55715/VB_AHCAL_PID_CALICE_collab.pdf.

[104] V. Bocharnikov. *Particle Identification with Multi-Class Gradient Boosted Decision Tree in the CALICE AHCAL*. PhD thesis, NRNU MePhI, Moscow, 2022. PhD Thesis in progress.

[105] S.Huck. Investigations of muon detection with the CALICE Analog Hadron Calorimeter. Master's thesis, Universität Hamburg, July 2019. URL https://www.physik.uni-hamburg.de/en/iexp/gruppe-garutti/publikationen/bachelor-master-theses/_files/2019-master-huck-saiva.pdf.

[106] S.Krause. *Studies of the Response of Silicon Photomultipliers and Testbeam Data Analysis of a Highly Granular Analog Hadron Calorimeter Prototype*. PhD thesis, Johannes Gutenberg-Universität Mainz, May 2020. http://doi.org/10.25358/openscience-1642.

[107] Frank J Massey Jr. The Kolmogorov-Smirnov test for goodness of fit. *Journal of the American statistical Association*, 46(253):68–78, 1951. https://doi.org/10.2307/2280095.

[108] Amine Elkhalii. *Analog Hadronic Calorimeter for a Future Linear Collider*. PhD thesis, Universität Wuppertal, Fakultät für Mathematik und Naturwissenschaften . . . , 2020.

[109] Catherine Adloff, J Blaha, J-J Blaising, C Drancourt, A Espargilière, R Gaglione, N Geffroy, Y Karyotakis, J Prast, G Vouters, et al. Hadronic energy resolution of a highly granular scintillator-steel hadron calorimeter using software compensation techniques. *Journal of instrumentation*, 7(09):P09017, 2012.

[110] Eigen. G. et al. Hadron shower decomposition in the highly granular CALICE analogue hadron calorimeter. *Journal of Instrumentation*, 11(06):P06013–P06013, feb 2016. ISSN 1748-0221. doi: 10.1088/1748-0221/11/06/P06013. URL https://arxiv.org/abs/1602.08578.

[111] D. Jang. Parametrized simulation of the CMS calorimeter using GFlash. In *2009 IEEE Nuclear Science Symposium Conference Record (NSS/MIC)*, pages 2074–2080. IEEE, 2009. doi: 10.1109/NSSMIC.2009.5402114.

[112] F. Gaede. Private Communication, 2021.

[113] J. Marshall. Implementation of the backscattered tracks 2 algorithm class., 2015. URL https://github.com/PandoraPFA/LCContent/blob/aabb475ab87a538826befb36e79e27bc977e321f/src/LCPlugins/LCShowerProfilePlugin.cc.

[114] ROOT. User's Guide Webpage, 2018. URL https://root.cern.ch/root/htmldoc/guides/users-guide/ROOTUsersGuide.html.

[115] F. James and M. Winkler. MINUIT2 Fitting Tool, 2018. URL https://root.cern.ch/root/htmldoc/guides/minuit2/Minuit2.html.

Acknowledgements

This thesis represents the end of an intensive and remarkable journey. Thanking everybody who contributed to my academics and my personal growth during this time is nigh on impossible. Nevertheless, I will try to mention the most important people here.

First and foremost, I would like to extend my deepest gratitude to my supervisor *Prof. Dr. Erika Garutti*, for giving me the opportunity to do my PhD thesis under her supervision. I thank her for her guidance, her motivating nature, many exciting and fruitful discussions. I have grown so much as a scientist under your guidance. Grazie di tutto!

I would like to express my sincere gratitude to *Dr. Katja Krüger*; this thesis would not have been finished without your enormous efforts. I am thankful for all the ideas, the questions, the guidance, positive attitude. I truly enjoyed the extensive whiteboard discussions we had throughout my thesis, which gave me more interest in particle physics and shaped me as a young scientist. Thank you for staying so nice and willing to help me proofread presentations, abstracts, papers and this thesis.

A special thank you goes to Prof. Dr. Elizabetta Gallo for being my DESY mentor, continuously supporting me, and for all the discussions.

I would like to thank all current and former FTX group members at DESY for their advice, help, and discussions. I would like to thank my group leader Dr. Ties Behnke. Special thanks go to our secretary Andrea Schrader for all her help concerning administrative tasks. I would like to thank Dr. Sergej Schuwalow for his helpful tips and physics discussions. I am very grateful to Dr. Felix Sefkow for many discussions and suggestions.

Of course, I have to thank all the current and former members of the AHCAL group at DESY and beyond for their assistance at every stage of the research project, especially Eldwan, Jiri, Yuji, Ambra, Remi, Lan, Matthias, Ole, Malinda, Merijn, Jack, Saiva, Erik, Antoine, Anna, Sebastian, Amine, Lorenz, Christian G., Christian W., Konrad, Wataru, Linghui, etc. Thank you for the fantastic collaborative atmosphere and the wonderful and unforgettable test beam experiences at DESY and SPS CERN. Special thanks go to my dear DESY PhD student colleagues and friends, Daniel and Vladimir, for all the memorable times, with lots of fun outings and entertainment apart from work activities.

I would like to thank all members of the CALICE collaboration who actively supported me and for all the fruitful discussions during many collaboration and analysis meetings. Regarding this, I want to thank Marina, Frank, Roman and Francois for the excellent collaboration and the great time at the meetings.

I would like to extend my sincere thanks to Thomas, Peter, Ralf, Anna and Paul for their time extensively proofreading the manuscript for this thesis.

A very personal thank you to Meghashree C.H for her continuous assistance. The most understanding and relaxed person I know who made the challenging phases of my Ph.D. definitely easier and nicer. Thank you very much for everything.

Last but not least, I would like to especially thank my wife, Vinita Baptist, for her unwavering support and her firm trust in me and my capabilities. With your patience and kindness, you have made the downs less dire and the ups even more enjoyable. Finally, I would like to thank my mother and brother for the years of support. I am forever grateful to all of you and love you so much.

Eidesstattliche Versicherung

Hiermit versichere ich an Eides statt, die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebenen Hilfsmittel und Quellen benutzt zu haben.

Declaration on Oath

I hereby declare, on oath, that I have written the present dissertation by my own and have not used other than the acknowledged resources and aids.

Hamburg, May 2022 Olin Lyod Pinto