

Calibration and Analysis of Data taken with the Technological Prototype of the Analog Hadron Calorimeter (AHCAL) for a Detector at the International Linear Collider

Dissertation zur Erlangung des Doktorgrades
an der Fakultät für Mathematik, Informatik und Naturwissenschaften
Fachbereich Physik
der Universität Hamburg

vorgelegt von
AMBRA PROVENZA

Hamburg
2018

Gutachterinnen der Dissertation:

Dr. Katja Krüger
Prof. Dr. Erika Garutti

Zusammensetzung der Prüfungskommission:

Prof. Dr. Dieter Horns
Dr. Katja Krüger
Prof. Dr. Erika Garutti
Prof. Dr. Gudrid Moortgat-Pick
Prof. Dr. Elisabetta Gallo

Vorsitzender der Prüfungskommission:

Prof. Dr. Dieter Horns

Datum der Disputation:

19.12.2018

Vorsitzender Fach-Promotionsausschusses Physik:

Prof. Dr. Wolfgang Hansen

Leiter des Fachbereichs Physik:

Prof. Dr. Michael Potthoff

Dekan der Fakultät MIN:

Prof. Dr. Heinrich Graener

Abstract

This thesis presents results obtained with the high granularity *Analog Hadron Calorimeter* (AHCAL) technological prototype. It studies the structure of hadronic showers combining for the first time the time information and the energy information of each hit in the detector.

High granularity is a stringent requirement for the calorimeter system of an experiment at the *International Linear Collider* (ILC), where the application of the *Particle Flow Algorithm* is foreseen to achieve an unprecedented jet energy resolution.

The CALICE Collaboration is studying several technologies satisfying this requirement. One of them is the AHCAL which consists of $30 \times 30 \times 3$ mm³ scintillating tiles, read out by Silicon Photomultipliers (SiPMs). A technological prototype based on this technology, built for proving the scalability of the calorimeter to a full linear collider detector, has been operated at CERN in July 2015 in muon, electron and pion beam. The muon and electron data have been analyzed to cross check the calibration of the amplitude measurements and to tune the main parameters in the simulation.

The analysis of the pion data focuses on the correlation between the time of the hit in the detector and the radial distance of the hit from the shower's center of gravity. For a better understanding of this dependence, the same correlation has been studied in defined hit energy ranges. The delayed component is observed only for small hit energies, while large hit energies show, within uncertainties, no delayed component. This is consistent with the interpretation that the late component is predominantly caused by neutrons, which mainly lead to small hit energies.

Zusammenfassung

In dieser Arbeit werden Ergebnisse präsentiert, die mit dem hoch-granularen technologischen Prototypen des *Analog Hadron Calorimeter* (AHCAL) gewonnen wurden. Die Struktur von hadronischen Schauern wird untersucht, indem zum ersten Mal die Zeit- und Energieinformation jedes Hits im Detektor kombiniert werden.

Hohe Granularität ist eine zwingende Anforderung an ein Kalorimetersystem an einem Experiment am *International Linear Collider* (ILC), wo die Verwendung von *Particle Flow Algorithmen* vorgesehen ist, um eine bisher unerreichte Jet-Energieauflösung zu ermöglichen.

Die CALICE Kollaboration untersucht verschiedene Technologien, die diese Anforderung erfüllen können. Eine davon ist das AHCAL, das aus $30 \times 30 \times 3 \text{ mm}^3$ großen Szintillatorkacheln besteht, die mit *Silizium Photomultipliern* (SiPMs) ausgelesen werden. Ein auf dieser Technologie basierender technologischer Prototyp, gebaut um die Skalierbarkeit des Kalorimeters auf die Größe eines Detektors an einem Linearbeschleuniger zu demonstrieren, wurde im July 2015 am CERN mit Myonen, Elektronen und Pionen getestet. Die Myon- und Elektrondaten wurden analysiert, um die Kalibration der Amplitudenmessung zu bestätigen und die Simulation zu justieren.

Die Analyse der Piondaten konzentriert sich auf die Korrelation der Zeit eines Hits im Detektor und seinem radialen Abstand von dem Schwerpunkt des Schauers. Um ein besseres Verständnis dieser Abhängigkeit zu erreichen, wird diese Korrelation in mehreren definierten Energiebereichen der Hits untersucht. Eine zeitlich verzögerte Komponente ist nur bei niedrigen Hitenergien zu sehen, während bei hohen Hitenergien, innerhalb der Messunsicherheiten, keine verzögerte Komponente beobachtet wird. Dies ist konsistent mit der Interpretation, dass die verzögerte Schauerkomponente vor allem von Neutronen erzeugt wird, die hauptsächlich geringe Energiedepositionen erzeugen.

Contents

Introduction	9
1. Particle Physics & the ILC	11
1.1. Status of Particle Physics	11
1.1.1. The Origin of Mass in the Standard Model	12
1.1.2. Open Questions of the Standard Model	16
1.2. Physics Motivation for an e^+e^- Collider	18
1.2.1. Higgs Boson at the ILC & Model Independent Measure- ments of its Properties	18
1.3. The International Linear Collider (ILC)	24
1.3.1. The ILC Detectors	27
2. Calorimetry and Particle Flow	31
2.1. Interaction of Particles with Matter	31
2.1.1. Electromagnetic Shower	31
2.1.2. Heavy Charged Particles Interactions	35
2.1.3. Hadronic Shower	37
2.2. Simulation of Particle Cascades	39
2.2.1. Simulation of Hadronic Cascade	39
2.2.2. Physics List	42
2.3. Calorimeters	43
2.4. Particle Flow Approach	46
2.5. CALICE Detector Concepts	48
3. The AHCAL Technological Prototype	53
3.1. Silicon Photomultiplier	54
3.2. The SiPM Integrated Read-Out Chip 2	57
3.3. The AHCAL Technological Prototype	59
3.3.1. The Hadronic Calorimeter Base Unit	59
3.3.2. The AHCAL Technological Prototype	60
3.4. The CERN 2015 Test Beam Campaign	62
3.5. Calibration	65
3.5.1. Gain Calibration	67
3.5.2. Pedestal Extraction	68
3.5.3. High Gain - Low Gain Intercalibration	69
3.5.4. MIP Calibration	77
3.5.5. Time Calibration	80
3.6. Digitization in the Simulation	83
3.6.1. Optical Cross Talk	83
3.6.2. Silicon Photomultiplier Response	84

3.6.3. Timing	85
3.6.4. Noise Implementation	85
3.7. Summary	86
4. Analysis of Electron Beam Data	87
4.1. Event Selection	88
4.2. Simulation	93
4.2.1. Tuning of the Optical Cross Talk	94
4.2.2. Tuning of the Number of Effective Pixels for The Saturation Correction	97
4.3. Systematic Uncertainties	103
4.3.1. Uncertainties on the Simulation	103
4.3.2. Uncertainties on the Data	109
4.3.2.1. Gain Uncertainty	109
4.3.2.2. MIP Uncertainty	109
4.3.2.3. High Gain - Low Gain Intercalibration Uncertainty	111
4.4. Results	111
4.5. Conclusion	124
5. Analysis of Pion Beam Data	125
5.1. Event Selection & the Shower Start Finder	126
5.2. Amplitude Analysis of Pion Beam Data	130
5.3. Correlation between Energy and Time	142
5.4. Conclusion	156
Conclusion	161
Acknowledgments	164
Appendix A. Simulation Input Parameters	165
Appendix B. Runs Used	167
Appendix C. Database Tags	169
Bibliography	171

Introduction

The *Standard Model* of particle physics provides the best description of the fundamental particles and of the fundamental forces through which they interact. The discovery of the Higgs boson at the CMS and ATLAS experiments at the *Large Hadron Collider* (LHC) at CERN in 2012 added the last missing component and completed the model.

The LHC is currently the most powerful accelerator, able to accelerate protons to a center of mass energy up to $\sqrt{s} = 14$ TeV. A high energy lepton collider would complement the LHC both in precision measurements and in possible discovery of new physics beyond the Standard Model. The most mature concept for a lepton collider is the *International Linear Collider* (ILC), a project currently under consideration in Japan. The ILC is foreseen to accelerate electrons and positrons to the center of mass energy of 250 GeV, extendable to 500 GeV. In order to provide best precision measurements, a jet energy resolution of 3%–4% is required. This will be feasible by applying *Particle Flow Algorithms* (PFAs). PFAs aim to combine measurements from the tracker system and the calorimeter system to obtain the desired jet energy resolution. This is possible by reconstructing each single particle in a jet, using the detector that provides the best energy resolution for this particle. This places stringent requirements on the detector system: a tracking system with an excellent track reconstruction efficiency and high granularity electromagnetic and hadronic calorimeters.

A high granularity hadronic calorimeter prototype developed within the CALICE collaboration is discussed in this thesis, from its hardware components to its physics capability. It has been tested in a test beam campaign at CERN in July 2015, where muon, electron and pion data have been collected.

The goal of this thesis is the study of the structure of hadronic showers combining, for the first time, the time information and the energy information of each hit in the detector. For this purpose, pion data have been analyzed. In order to do that, a good understanding of the detector and its performance is crucial. Therefore, muon and electron data were analyzed in order to cross check and validate the calibration of the detector and to tune the main parameters of the simulation, in order to have a reliable description of the prototype. Based on this, it has been possible to study the hadronic showers. In particular, the dependence of the time of a hit in the detector on the radial distance of the hit from the shower's center of gravity has been studied in defined hit energy ranges.

In the first chapter, the Standard Model and its open questions are introduced, including a short description of *Supersymmetry* as a possible answer to these questions. Furthermore, the Higgs boson as the origin of the mass in the Standard Model is discussed, with a particular emphasis on the properties that can be studied in a model-independent way and with high precision at the International Linear Collider, the most mature concept for a lepton collider. After motivating the need of a linear collider, a description of the ILC including its two detectors is given.

In the second chapter, the interaction of particles with matter and how this is implemented in the simulation is described. The principle of calorimetry is introduced, followed by a description of the Particle Flow approach.

In Chapter 3, the high granularity hadronic calorimeter technological prototype (AHCAL) is described, with a focus on the SiPM on tile technology, including the calibration chain performed in order to properly extract the energy and time measurement of particles in the detector.

The analysis of the data collected during the test beam campaign at CERN in 2015 is described in Chapter 4 and Chapter 5.

In Chapter 4, the analysis of electron data is presented. During the test beam campaign, electron data in the energy range between 10 GeV and 50 GeV have been collected. The main purpose of the analysis is to study and cross check the amplitude measurement and to tune the main parameters of the simulation. This is a crucial step in order to have a reliable description of the detector and its properties and a good understanding of its performance.

Once the main parameters of the simulations have been tuned and the agreement between data and simulation has been evaluated, it is possible to study the hadronic showers. This is described in Chapter 5. An advantage of the AHCAL technological prototype is the possibility to have the information of the time of a hit in the detector which can help to identify the different components of the hadronic shower, thus providing a better understanding of such a complex object. The time structure of pion showers is studied looking at the mean time of a hit as a function of its radial distance from the shower's center of gravity in different hit energy ranges.

1. Particle Physics & the ILC

The currently best description of all fundamental particles and forces through which they interact is given by the Standard Model [1–3]. Its validity has been confirmed with the discovery of the last particle missing, the Higgs Boson [4][5].

At the beginning of this chapter, a summary of the Standard Model with its shortcomings is presented. The need for an e^+e^- linear collider is motivated, with a particular emphasis on the precision measurements possible at this kind of experiment. At the end, a description of the most mature linear collider project, the International Linear Collider, is provided.

The book in [6] from D. Griffiths has mainly been used as a guideline for the summary of the Standard Model, while for the description of the International Linear Collider and its detectors, the *Technical Design Report* (TDR) [7][8][9] has been mainly used as a reference, if not stated otherwise.

1.1. Status of Particle Physics

The best description of the subatomic world is given by the Standard Model. Developed in the early 1970s, it successfully explains almost all experimental results and precisely predicts a wide variety of phenomena. The Standard Model consists of seventeen fundamental particles, classified in fermions with spin equal to $1/2$ and bosons with integer spin. Fermions are divided into two groups: quarks and leptons. Each group consists of six particles, which are related in pairs, or *generations*. The six quarks are paired in three generations: up and down, charm and strange, top and bottom. The up, charm and top are the *up-type* quarks with electric charge equal $+2/3$ ($-2/3$ for the corresponding anti-quarks) while the down, strange and bottom are the *down-type* quarks, with electrical charge equal to $-1/3$ ($+1/3$ for the corresponding anti-quarks). Quarks are never observed isolated. They always come in multi-quark bound states, the hadrons. This *confinement* is described introducing an additional *color charge* and imposing that they only bind in a way that only colorless objects are formed. Bound states consisting of a quark-antiquark pair are called *mesons*, while bound states consisting of three quarks are called *baryons*. All known mesons are unstable, as well as baryons, except protons (made of uud quarks) and neutrons (made of udd quarks). All the stable matter in the universe is made of particles belonging to the

first (lightest) generation. Any heavier particles quickly decay to the next lighter particle.

The six leptons are also organized in three generations: the electron with the corresponding electron neutrino, the muon with the muon neutrino and the tau with the tau neutrino. The electron, the muon and the tau have all sizable mass and electric charge equal to -1 , whereas the neutrinos are electrically neutral and have a very small mass.

Four fundamental forces regulate the interaction between particles in the universe: the strong force, the weak force, the electromagnetic force and the gravitational force. They work over different ranges and have different strength. These forces are mediated through the exchange of bosons. Each force has its own corresponding boson. Photons are responsible for the electromagnetic interaction between electrically charged particles. Since photons are massless, the range of the electromagnetic interaction is in principle infinite. It is however limited from its potential that goes as $1/r$. The bosons W^\pm and Z^0 are the mediators of the weak force. All particles can interact through the weak force and this represents the only way of interaction for neutrinos. The weak gauge bosons are massive with $M_W = 80.4 \text{ GeV}$ and $M_Z = 91.2 \text{ GeV}$. The range of the interaction is limited from the Yukawa potential to around 10^{-18} m [10]. The strong force is carried by gluons which couple to color charge and therefore exclusively to quarks. Gluons mediate the exchange of color charge between quarks and they carry the color charge themselves being self-interacting. This corresponds to an increase of the potential with increasing distance of two objects color charged. Once the potential energy exceeds the pair production threshold a quark-antiquark pair is produced, limiting the range of interaction of the strong force even though gluons are massless. This hadronization process generates a cascade of mixed particles called jets. The only way to reconstruct the energy and momentum of quarks is the reconstruction of the energy of the jets. Although not yet found, the graviton should be the corresponding force-carrying particle of gravity. Gravity is not included in the Standard Model since the quantum field theory used to describe the subatomic world and the general relativity used to describe the macroscopic world are not easy to fit in a single framework. Luckily gravity is negligible at the scale of subatomic particles [11]. A summary of the quark and lepton generations together with the mediators of the fundamental forces described in the Standard Model is shown in Figure 1.1.

1.1.1. The Origin of Mass in the Standard Model

In the Standard Model the origin of the mass of particles is explained through the Higgs mechanism introducing a scalar field, the Higgs boson [12][13]. The problem appeared when gauge theory is applied to quantum field theory. Requiring that

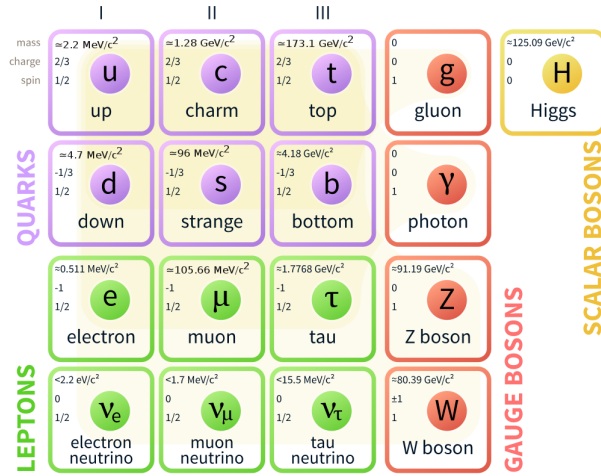


FIGURE 1.1.: Fundamental particles of the Standard Model and their classification: the three generations of quark (in violet) and leptons (in green) and the mediators of the fundamental forces (in red) are shown. Picture from Wikipedia.

the Lagrangian describing a free particle is invariant under local gauge transformation introduces in the Lagrangian itself new field(s), interacting with the field considered. In order to preserve the invariance of the Lagrangian, this field turns out to be massless. The principle of local gauge invariance works perfectly for the electromagnetic and the strong interactions, since the mediators of these two forces (photons and gluons) are massless. The application to the weak interaction was stymied by the fact that the gauge fields have to be massless. This was in contrast to the results from the UA1 [14][15] and UA2 [16][17] experiments at CERN, where the W^\pm and the Z^0 were discovered and resulting to be indeed massive, with a mass respectively of $M_W = 80.379 \pm 0.012 \text{ GeV}$ and $M_Z = 91.1876 \pm 0.0021 \text{ GeV}$ [18]. The mass of the W^\pm and the Z^0 bosons was explained applying the spontaneous symmetry breaking and the Higgs mechanism. The solution was formulated by Higgs [12], Englert and Brout [13]. They proposed that all of space is filled with a field (the Higgs field) that interacts with the weak particles to give them mass. This happens because the Higgs field is assumed to have a vacuum expectation value different from zero. This non-zero ground state violates the symmetry, allowing the mediators of the weak force to acquire mass.

In order to better explain it, let's consider the Lagrangian of a real scalar field ϕ

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)^*(\partial^\mu \phi) + \frac{1}{2}\mu^2(\phi^* \phi) - \frac{1}{4}\lambda^2(\phi^* \phi)^2 \quad (1.1)$$

with μ and λ (real) constants and the field $\phi = \phi_1 + i\phi_2$ complex field such that $\phi^* \phi = \phi_1^2 + \phi_2^2$. Comparing this Lagrangian, with the *Klein-Gordon* Lagrangian

for a scalar field

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)(\partial^\mu \phi) - \frac{1}{2}\left(\frac{mc}{\hbar}\right)^2 \phi^2 \quad (1.2)$$

if μ represents the mass, the sign in the Lagrangian 1.1 is wrong. This implies in fact that the mass is imaginary (that has no physical sense). In order to understand this Lagrangian it has to be considered that to apply the Feynman calculus, the field configuration of minimum energy has to be determined. The Feynman calculus is in fact a perturbation procedure that allows to treat the fields as fluctuations about that ground state (vacuum). For the Lagrangian 1.2, $\phi = 0$ represents the field configuration of minimum energy. This is not true for the Lagrangian 1.1. To determine the ground state, the Lagrangian 1.1 can be written as the difference between the kinetic and the potential term. From Eq. 1.1, the potential can be defined as

$$\mathcal{U}(\phi) = -\frac{1}{2}\mu^2(\phi^* \phi) + \frac{1}{4}\lambda^2(\phi^* \phi)^2. \quad (1.3)$$

Since the Lagrangian 1.1 is a function only of $(\phi_1^2 + \phi_2^2)$, it is invariant under rotation. The minima of the potential lie on a circle $\phi_1^2 + \phi_2^2 = \mu^2/\lambda^2$ of radius μ/λ . In order to express the fields as fluctuations around a minimum, a specific minimum has to be chosen. One possibility is $\phi_1^{min} = \mu/\lambda$ and $\phi_2^{min} = 0$. Requiring the local gauge invariance, meaning the invariance of the Lagrangian under transformations $\phi \rightarrow e^{i\theta(x)}\phi$ and replacing the derivatives in the Lagrangian 1.1 with the *covariant derivatives* ($\mathcal{D}_\mu = \partial_\mu + i\frac{q}{\hbar c}A_\mu$, with A^μ massless gauge field), the Lagrangian 1.1 can be written as

$$\begin{aligned} \mathcal{L} = & \frac{1}{2}\left[\left(\partial_\mu - \frac{iq}{\hbar c}A_\mu\right)\phi^*\right]\left[\left(\partial^\mu + \frac{iq}{\hbar c}A^\mu\right)\phi\right] \\ & + \frac{1}{2}\mu^2(\phi^* \phi) - \frac{1}{4}\lambda^2(\phi^* \phi)^2 - \frac{1}{16\pi}F^{\mu\nu}F_{\mu\nu}. \end{aligned} \quad (1.4)$$

As a reminder, $F^{\mu\nu}$ is defined as $F^{\mu\nu} \equiv \partial^\mu A^\nu - \partial^\nu A^\mu$.

Defining two new fields $\eta = \phi_1 - \mu/\lambda$ and $\xi = \phi_2$, the Lagrangian becomes:

$$\begin{aligned}
\mathcal{L} = & \left[\frac{1}{2}(\partial_\mu \eta)(\partial^\mu \eta) - \mu^2 \eta^2 \right] + \left[\frac{1}{2}(\partial_\mu \xi)(\partial^\mu \xi) \right] \\
& + \left[-\frac{1}{16\pi} F^{\mu\nu} F_{\mu\nu} + \frac{1}{2} \left(\frac{q}{\hbar c} \frac{\mu}{\lambda} \right)^2 A_\mu A^\mu \right] \\
& + \left[\frac{q}{\hbar c} [\eta(\partial_\mu \xi) - \xi(\partial_\mu \eta)] A^\mu + \frac{\mu}{\lambda} \left(\frac{q}{\hbar c} \right)^2 \eta (A_\mu A^\mu) \right. \\
& + \frac{1}{2} \left(\frac{q}{\hbar c} \right)^2 (\xi^2 + \eta^2) (A_\mu A^\mu) - \lambda \mu (\eta^3 + \eta \xi^2) - \frac{1}{4} \lambda^2 (\eta^4 + 2\eta^2 \xi^2 + \xi^4) \left. \right] \\
& + \left(\frac{\mu}{\lambda} \frac{q}{\hbar c} \right) (\partial_\mu \xi) A^\mu + \left(\frac{\mu^2}{2\lambda} \right)^2.
\end{aligned} \tag{1.5}$$

The first line of Eq. 1.5 describes a scalar particle (η) of mass $m_\eta = \sqrt{2}\mu\hbar/c$ and a massless Goldstone boson (ξ); the second line describes the free gauge field A_μ added in order to maintain the gauge invariance of the Lagrangian and now it has acquired a mass term $m_A = 2\sqrt{\pi}(\frac{q\mu}{\lambda c^2})$. The other lines represent interactions between the involved fields η , ξ and A_μ . It is interesting to notice where the mass term for A_μ comes from: the original Lagrangian (Eq. 1.4) contains a term like $\phi^* \phi A_\mu A^\mu$ which, in absence of spontaneous symmetry breaking corresponds to a point like interaction between the fields ϕ and A . But when the ground state moves “off center” this piece of the Lagrangian emerges as a mass term. Choosing a specific gauge, the massless Goldstone boson also disappears and the Lagrangian assumes the form

$$\begin{aligned}
\mathcal{L} = & \left[\frac{1}{2}(\partial_\mu \eta)(\partial^\mu \eta) - \mu^2 \eta^2 \right] + \left[-\frac{1}{16\pi} F^{\mu\nu} F_{\mu\nu} + \frac{1}{2} \left(\frac{q}{\hbar c} \frac{\mu}{\lambda} \right)^2 A_\mu A^\mu \right] \\
& + \left[\frac{\mu}{\lambda} \left(\frac{q}{\hbar c} \right)^2 \eta (A_\mu A^\mu) + \frac{1}{2} \left(\frac{q}{\hbar c} \right)^2 \eta^2 (A_\mu A^\mu) - \lambda \mu \eta^3 - \frac{1}{4} \lambda^2 \eta^4 \right] \\
& + \left(\frac{\mu^2}{2\lambda} \right)^2.
\end{aligned} \tag{1.6}$$

By choosing a proper gauge the Goldstone boson has been eliminated and only a single massive scalar η (the Higgs boson) and a massive gauge field A_μ are left, sacrificing the symmetry. This Lagrangian is not invariant anymore under rotation, therefore the symmetry has been broken. The interpretation of this result is the following: a massless vector field carries two degrees of freedom (transverse polarization). When the gauge field A_μ acquires mass, it picks up a third degree of freedom (longitudinal polarization), absorbing the Goldstone boson. This is the *Higgs mechanism*. The W^\pm and Z^0 bosons acquire their mass through the Higgs mechanism. The Higgs boson is responsible also for the masses of electrically charged leptons and quarks. They acquire mass having a Yukawa couplings with

the Higgs particle [6][19].

1.1.2. Open Questions of the Standard Model

The discovery of the Higgs boson at the *Large Hadron Collider* (LHC) [20], the last particle predicted from the Standard Model, represents a milestone in the history of particles physics [4][5]. As already mentioned above, the Standard Model is the theory that gives the best description of the subatomic world. Nevertheless there are still questions that it is not able to answer. Here these shortcomings are briefly summarized.

Grand Unification

The fact that the weak interaction and the electromagnetic interaction can be seen as the representation of the same force led to the idea that also the unification with the strong interaction is possible. This implies that at a certain (high) energy all the forces couple with the same strength. The Standard Model equations that describe the evolution of the coupling constants at different energy scales show a divergency [21]. Therefore the Standard Model cannot make feasible the unification of all three fundamental forces. Moreover, as mentioned before, gravitation also cannot be included.

Hierarchy Problem

The Higgs boson mass has many contributions from quantum loop corrections. Considering that the Standard Model is supposed to be valid up to the Planck scale $\Lambda_{Planck} \approx 10^{19}$ GeV, these corrections become orders of magnitude larger than the experimentally measured Higgs boson mass equal to $m_H = 125.18 \pm 0.16$ GeV [18]. In order to keep the Higgs boson mass at the observed value, an unnatural degree of fine-tuning that enables cancellations of these contributions would be required.

Matter-Antimatter Asymmetry

The Standard Model assumes that at the beginning of the Universe, an equal amount of matter and anti-matter were produced. Instead of annihilating, matter (i.e baryons) dominates. A possible explanation involves the CP symmetry violation (charge conjugation parity symmetry) [22][23]. However, up to now, CP violation has been observed only in the context of the weak interactions in the quark sector. The effect is not large enough to be able to explain the matter-antimatter asymmetry.

Neutrino Masses

In the Standard Model neutrinos are assumed to be massless. However this is in contrast with the results obtained from neutrino flavour oscillations experiments. From these experiments it results that neutrinos have to be massive, with a mass about six orders of magnitude smaller than the masses of the other Standard Model fermions. Only an upper limit on the neutrino masses, at the level of $O(10^{-1})$ eV, can be set and their ordering [24].

Dark Matter

The amount of “visible” matter in the Universe corresponds only to $\approx 4\%$ of the energy density in our Universe. Many cosmological and astrophysical experiments, like the rotational velocity of galaxies [25] or the measurements of the anisotropy in the cosmic microwave background [26], indicate the presence of some unknown and invisible matter in the Universe, called Dark Matter. Since it has not yet been observed, any potential dark matter candidate should be massive and interact weakly.

One of the possible theories answering these questions is *Supersymmetry*. It carries the idea that every fermion has a bosonic partner and every boson has a fermionic partner. If supersymmetry is an exact symmetry, the masses of the supersymmetric particles should be the same as the masses of the corresponding Standard Model particles. But supersymmetric particles have not yet been observed, therefore the symmetry has to be broken and the supersymmetric particles have to be much heavier.

Supersymmetry could make possible the unification of the coupling constants as hypothesized in the Grand Unified Theory (GUT) and solve the hierarchy problem.

In the Standard Model, the coupling constants themselves are functions of energy and this is because of renormalization. Their dependence is shown by the dashed lines in Figure 1.2. Supersymmetry changes the renormalization group. In this way, the dependence of the coupling constants on the energy is modified, making possible their convergence at the GUT scale around 10^{16} GeV [27]. This is depicted by the solide lines in Figure 1.2.

As mentioned above, supersymmetry also offers a “natural” solution to the hierarchy problem. The Higgs mass is in fact renormalized by various loop diagrams. But the loop corrections for boson and leptons are of opposite sign. Therefore, by pairing particles with sparticles (supersymmetric particles) these loop corrections cancel each other. In addition to that, in most of the supersymmetric models, the lightest supesymmetric particle is colorless, neutral and stable, being therefore an excellent candidate for Dark Matter [6][27].

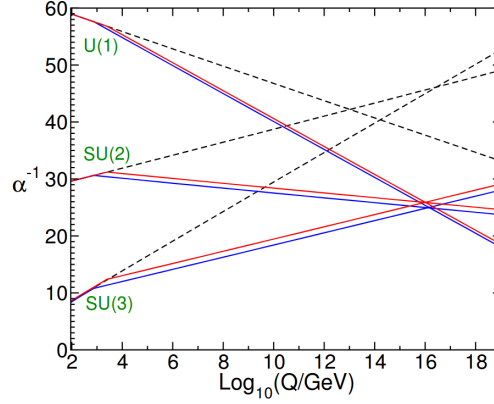


FIGURE 1.2.: Evolution of the coupling constants with the energy in the Standard Model (dashed lines) and in Supersymmetry (solid lines)[21].

1.2. Physics Motivation for an e^+e^- Collider

As explained in the previous section, there are several theoretical models which aim to address and find answers to the open questions of the Standard Model and predict the existence of new particles. The LHC collider gives access to high energy proton-proton collisions for direct production of new particles. Unfortunately production rates of these particles are about $10^{-11} - 10^{-12}$ of the total proton-proton cross section. Even after the event selection these events represent roughly the 10% of the total yield over the huge background due to the many standard model interactions. This limits the range of new processes that can be observed at the Large Hadron Collider and the reachable precision. In an e^+e^- collider these events represent a large fraction of the total e^+e^- cross section. The event selection gives high purity over less background. For the study of heavy particles, almost all the decay modes can be observed and the systematic errors are at the level of 1 ‰. This is a powerful capability that can be applied to the study of the properties of the Higgs boson and the top quark, the two particles most involved in the questions left open from the Standard Model.

The International Linear Collider is the most mature concept for a future linear e^+e^- accelerator. Therefore in the next section, a description of the properties of the Higgs boson that will be possible to study at the International Linear Collider is given, while an overview of the experiment and its detectors is given in Section 1.3.

1.2.1. Higgs Boson at the ILC & Model Independent Measurements of its Properties

As explained in Section 1.1, in the Standard Model the Higgs field is responsible for the masses of charged leptons, quarks and vector bosons. 10^{-10} seconds after the

Big Bang, empty space-time made a transition to a new state filled with a uniform field, the Higgs field. This field breaks the symmetries of the Standard Model that prevents the masses of the Standard Model's particles. Its uniform field value is therefore responsible for the different masses. The discovery of the Higgs boson at CERN was a big step forward in the understanding of the origin of masses within the Standard Model. However this evidence only answers the question how the weak interaction symmetry is broken, but not why this happens, or why the Higgs field acquires this non zero value. With the measurement of the Higgs mass known at the level of 0.2% all the parameters of the Standard Model are fixed. The study of the properties of the Higgs boson, in particular the couplings, are therefore a crucial point: any deviation from the value predicted from the Standard Model will be a clear hint of the existence of new physics beyond the Standard Model. The measurement of the Higgs couplings will in addition give information whether the Higgs is a fundamental scalar particle, the first one ever observed, or a composite particle. The deviations predicted by several models describing physics beyond the Standard Model are at the level of 5% or less. Therefore a precision better than 5% is needed in order to study these properties. At LHC, the precision reachable is limited by the statistical and systematic uncertainties, both already at the level of 5%. The goal is therefore to perform the measurements of the individual couplings of the Higgs at the level of 1%, and this can be obtained only at an e^+e^- collider. This is possible using the properties of an e^+e^- collider: a well defined initial state, absence of strong interaction backgrounds and controlled and calculable backgrounds from electroweak interaction processes.

Higgs Production Modes at the ILC

At the ILC, the production modes of the Higgs boson are: *Higgsstrahlung* ($e^+e^- \rightarrow ZH$), *W fusion* ($e^+e^- \rightarrow \nu\bar{\nu}H$) and *Z fusion* ($e^+e^- \rightarrow e^+e^-H$). The Feynman diagrams of these processes are illustrated in Figure 1.3. All the major Higgs decays, such as $H \rightarrow b\bar{b}, WW^*, c\bar{c}, \tau\tau, gg$ can be separately identified with high efficiency.

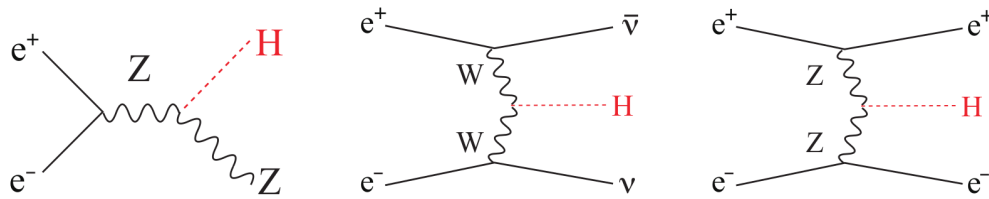


FIGURE 1.3.: Feynman diagram of the Higgs production processes at the International Linear Collider: Higgsstrahlung (left), W-fusion (center), Z-fusion (right) [7].

The cross section of these production processes at the ILC is shown in Figure 1.4. The possibility of controlling the polarization of electrons and positrons at the ILC represents a big advantage. The W-fusion process $e^+e^- \rightarrow \nu_e \bar{\nu}_e H$ for example occurs only in collisions between a left-handed electron and a right-handed positron. The cross section can be increased by a factor of 2 using the polarized beam available at the ILC. Figure 1.4 is obtained with the polarization combinations $P(e^-, e^+) = (-0.8, +0.3)$.

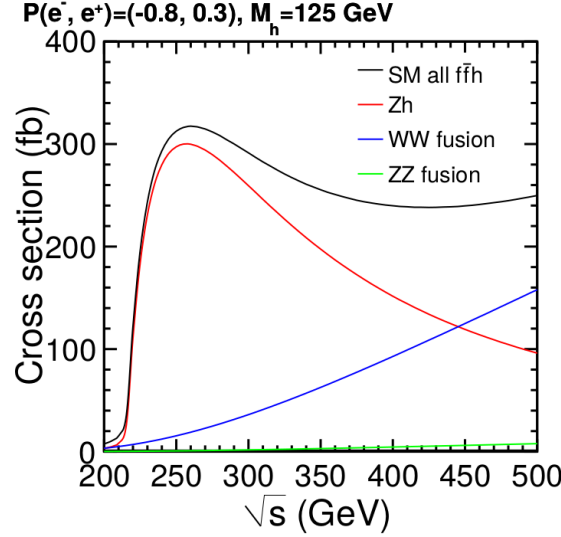


FIGURE 1.4.: Cross section of the Higgs production processes at the International Linear Collider, as a function of the center of mass energy. This plot assumes polarized beams, as indicated on top of the plot [7].

The Higgsstrahlung process $e^+e^- \rightarrow ZH$ has a big advantage: it allows the identification of the Higgs boson, without looking at its decays. This has three important consequences: it allows to determine the total width of the Higgs boson and the absolute values of the Higgs couplings, it allows to observe Higgs decays in invisible or exotic modes and gives a precise determination of the Higgs mass. It is important to point out that the total width of the Higgs Γ_H is given by Eq. 1.7

$$\Gamma_H = \frac{\Gamma_{ZZ}}{BR_{ZZ}} = \frac{\Gamma_{WW}}{BR_{WW}} \quad (1.7)$$

where Γ_{ZZ} (Γ_{WW}) is the partial decay width to ZZ^* (WW^*). Γ_{ZZ} (Γ_{WW}) is related to the coupling of the Higgs boson to the Z (W) boson and the cross section through the relation 1.8

$$\Gamma_{ZZ} \propto g_Z^2 \propto \sigma_{ZH} \quad (\Gamma_{WW} \propto g_W^2 \propto \sigma_{\nu\nu H}). \quad (1.8)$$

The branching ratio BR_{ZZ} is only measured to 6.7%. If only the first part of Eq. 1.7 is used, all the Higgs boson couplings (except g_z) would have an uncertainty of 3%. The branching ratio BR_{WW} is roughly 10 times larger than BR_{ZZ} [28],

therefore can be measured more precisely. For this reason, the measurement of the WW fusion cross section $\sigma_{\nu\nu H}$ together with the branching ratio BR_{WW} is crucial for the measurement of Γ_H and all the coupling constants g_A (with $A \neq Z$).

How the Higgsstrahlung process is used to study the Higgs properties and the consequences are illustrated in the following sections.

Higgs Recoil Mass Measurement

The Higgsstrahlung production mode for the Higgs boson offers the possibility to identify an Higgs event without looking at the Higgs decay at all. This is possible by identifying the Z boson at a well-defined laboratory energy corresponding to the kinematics of the recoil against the 125 GeV Higgs boson. The mass of a particle recoiling against a lepton pair is given by Eq. 1.9

$$M_X^2 = (p_{CM} - (p_{l^+} + p_{l^-}))^2 \quad (1.9)$$

where p_{MC} is the four momentum of the annihilating electron-positron system. The distribution of the expected recoil mass for a Higgs boson with a mass of 125 GeV with an integrated luminosity of 250 fb^{-1} and at the center of mass energy $\sqrt{s} = 250 \text{ GeV}$ is shown in Figure 1.5. This measurement allows the determination of the Higgs boson's mass to better than 30 MeV and a cross section at a sub-percent level of precision [29].

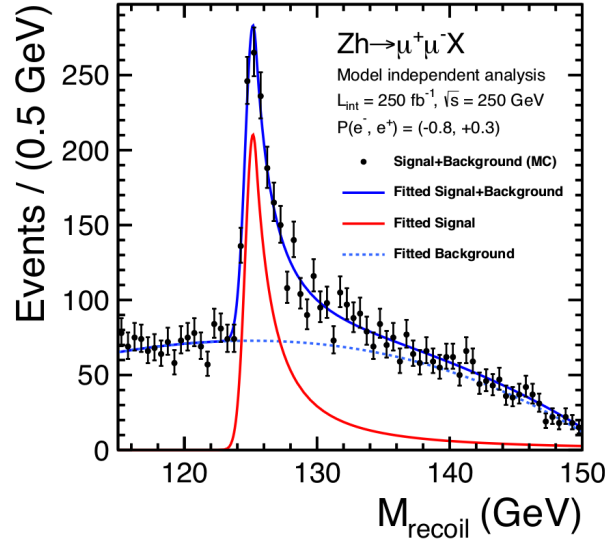


FIGURE 1.5.: Recoil mass distribution for a Higgs boson with a mass of 125 GeV for the process $e^+e^- \rightarrow ZH \rightarrow \mu^+\mu^-H$ with integrated luminosity $L = 250 \text{ fb}^{-1}$ and at the center of mass energy of $\sqrt{s} = 250 \text{ GeV}$ [29].

Higgs Coupling from Leptonic Recoil Mass Measurement

The precise measurement of the Higgs coupling strengths is very important, since deviations to the Standard Model Higgs couplings in the order of 1% are expected in many generic models of new physics beyond the Standard Model [30][31]. What is usually measured at a collider is the event rate of a given final state, corresponding to the product of the production cross section and the branching ratio. The branching ratio of a decay channel ($BR(H \rightarrow A\bar{A})$) is related to the partial decay width ($\Gamma(H \rightarrow A\bar{A})$) and the full Higgs decay width (Γ_H) as given by Eq 1.10

$$BR(H \rightarrow A\bar{A}) = \frac{\Gamma(H \rightarrow A\bar{A})}{\Gamma_H} \quad (1.10)$$

with

$$\Gamma_H \propto g_{HAA}^2 \quad (1.11)$$

In the Standard Model, Γ_H is too small to be measured directly (~ 4 MeV). At the LHC the determination of Γ_H requires model-dependent assumptions, since invisible Higgs decays cannot be disentangled from background processes. At the ILC it is possible to use the fact that all Higgs decays modes can be identified in the Higgsstrahlung process, using the recoil mass measurement to measure certain Higgs couplings directly. The total rate for the Higgsstrahlung is proportional to the ZZH coupling. The WWH coupling can be determined from the rate of the W fusion process

$$e^+e^- \rightarrow H\nu\bar{\nu} \text{ with } H \rightarrow b\bar{b}$$

divided by $BR(H \rightarrow b\bar{b})$ obtained from the Higgsstrahlung. With these two measurements the partial decay widths of the two decay processes $\Gamma(H \rightarrow ZZ)$ and $\Gamma(H \rightarrow WW)$ are determined. Combining these two measurements with the measured branching ratios given by Eq 1.10, a model-independent measurement of the full decay width Γ_H of the Higgs boson is provided. The uncertainty on the measurement of Γ_H directly propagates into the precision of each absolute coupling. Therefore the measurements of all cross sections times the branching ratios are included into the measurement in a global fit, in order to optimise the uncertainty on Γ_H . The precision achievable with a model independent measurement at the ILC with both initial and full statistics of the ILC program is shown in Figure 1.6(a). Here also a possible combination with the LHC for the $\gamma\gamma$ decay channel is shown. The reason why the ILC alone reaches a low precision in this specific decay channel is due to the fact that the decay of the Higgs boson in two photons is rare, with a branching ratio of 0.2%. In Figure 1.6(b), a comparison of the precision reached in the measurements of the Higgs couplings between ILC and LHC using a model dependent fit is shown [32].

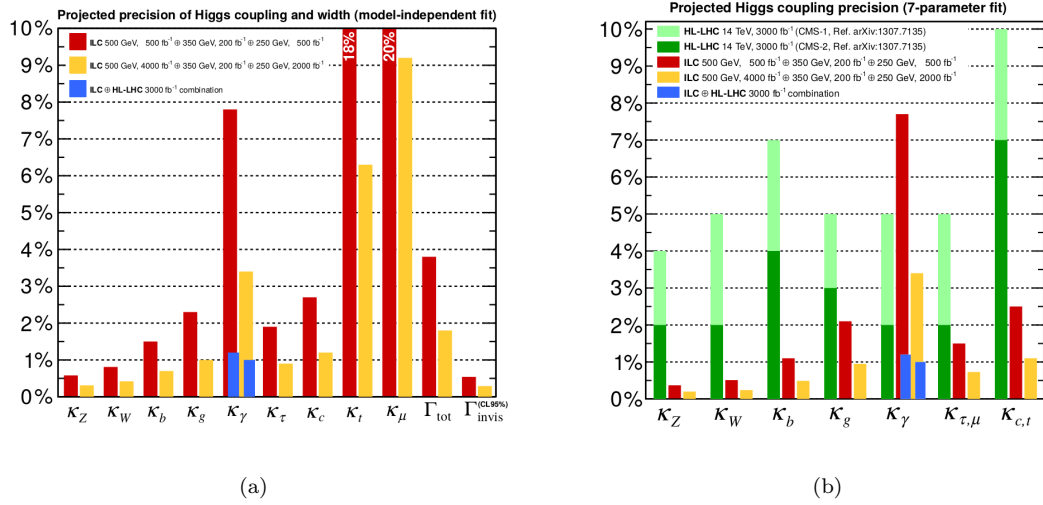


FIGURE 1.6.: Higgs couplings measurement [29].

- (a): Relative precision of the various Higgs couplings extracted from a model independent fit to expected data from the ILC.
- (b): Relative precision of the Higgs couplings using a model dependent fit applied to expected data from high luminosity HL-LHC and ILC. κ_A represents the ratio of the $A\bar{A}H$ coupling to the Standard Model expectation. The blue band in the $\gamma\gamma$ channel is the effect of a combined analysis of HL-LHC and ILC.

Higgs full Width from Hadronic Recoil Mass Measurement

In the previous section it has been shown that the Higgsstrahlung production mode provides a model-independent measurement of the Higgs coupling using the recoil mass technique. This has been shown considering the lepton decays of the Z boson, that means $\sigma(e^+e^- \rightarrow ZH)$ with $Z \rightarrow l^+l^-$ where $l = e, \mu$. A near model-independent measurement recoil mass technique can be extended to the hadronic decays of the Z boson. This has a big advantage, because the branching ratio for $Z \rightarrow q\bar{q}$ is roughly ten times greater than $Z \rightarrow l^+l^-$. A precision of roughly $\pm 1.8\%$ can be reached at the ILC operating at $\sqrt{s} = 350$ GeV [33].

Considering the Z hadronic decays, the selection efficiency will depend much more on the Higgs decay mode. For example, considering the following events $H \rightarrow b\bar{b}$ and $Z \rightarrow q\bar{q}$, the reconstruction of the Z boson is complicated by mis-association of particles to jets and ambiguities in associating jets to the Higgs and the Z boson. These ambiguities increase with the number of jets. Therefore it is difficult to construct an event selection based only on the reconstruction of the $Z \rightarrow q\bar{q}$ events without looking at the Higgs decays, but it is possible to minimize this dependence. It follows that an excellent jet energy resolution is a key requirement to extract Γ_h from hadronic recoil events with sufficient accuracy. For the event selection, events are classified in visible and invisible Higgs decays using the topology of the events themselves. The invisible Higgs decays should appear in a typical di-jet topology and nothing more. The two identified jets are

used to reconstruct the recoil mass. In the visible Higgs decays, the two jets with an invariant mass closes to the Z boson mass are used to reconstruct the recoil mass. The recoil mass distribution obtained from invisible Higgs decay modes and visible Higgs decay modes are shown in Figure 1.7. The two distributions show both a clear peak around the Higgs mass.

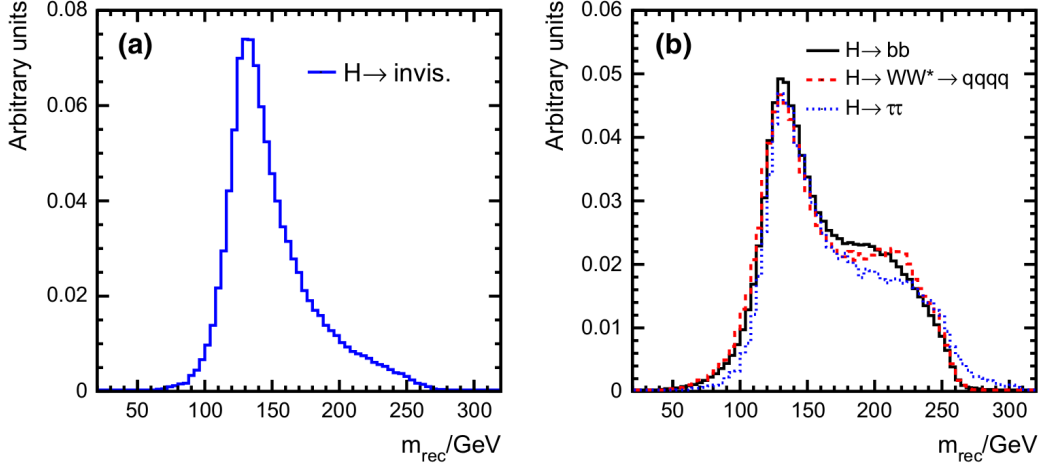


FIGURE 1.7.: Reconstruction of the recoil mass for events with invisible Higgs decays (a) and visible Higgs decays (b). In both cases the distributions have been normalized to unit area [33].

In order to extract the cross section of the process $\sigma(e^+e^- \rightarrow HZ)$ Likelihood functions for events classified as signal and for events classified as background are defined. The variables used in the likelihood selection are the recoil mass of the considered di-jet system (m_{rec}) and the invariant mass of the di-jet system ($m_{q\bar{q}}$). The correlation between these two variables for events classified as signal and events classified as background is shown in Figure 1.8. In Figure 1.8(a), corresponding to the signal events, the expected peak at $m_{\text{rec}} = m_H = 125$ GeV and $m_{q\bar{q}} = m_Z = 91$ GeV is visible. The width of the signal peak corresponds to the jet energy resolution in both dimensions, since the natural widths of the Z and Higgs are comparably narrow. From this plot is clear how a worse jet energy resolution would decrease the separation between signal and background in these variables, reducing the precision of about $\pm 1.8\%$ reachable in the measurement of the cross section $\sigma(e^+e^- \rightarrow ZH)$ for a center of mass energy of $\sqrt{s} = 350$ GeV.

1.3. The International Linear Collider (ILC)

As already mentioned, the International Linear Collider is the most mature concept for a future linear e^+e^- collider. In the previous section the physics motivation and the physics capabilities of such a collider have been summarized. Here a description of the accelerator itself and the proposed detectors is presented.

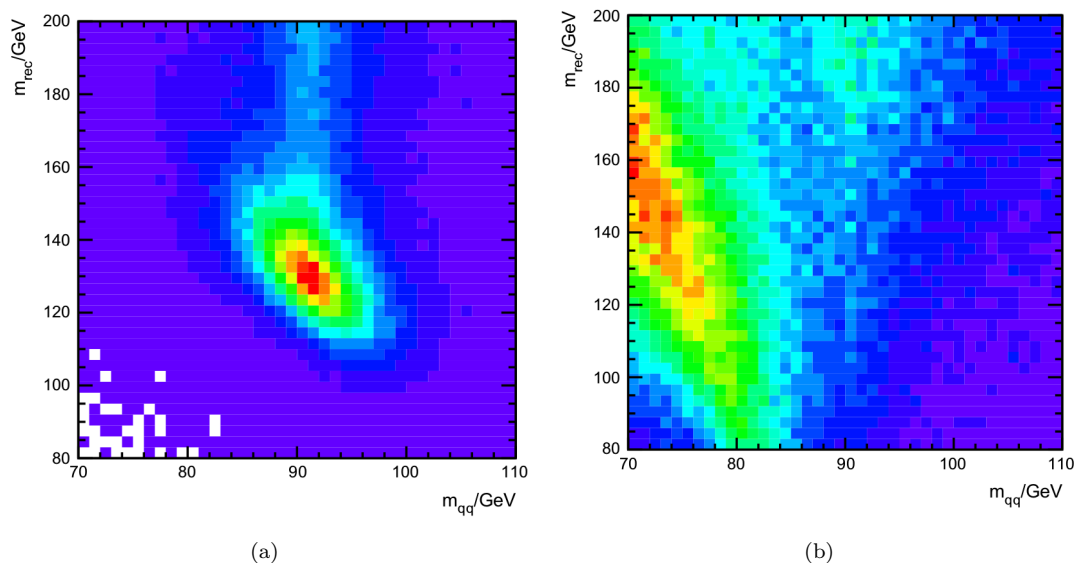


FIGURE 1.8.: Distribution of the reconstructed invariant mass $m_{q\bar{q}}$ $Z \rightarrow q\bar{q}$ versus the hadronic recoil mass m_{rec} [33].

(a): for signal events

(b): for background events

The ILC is a linear collider experiment, foreseen to be built in Japan. Currently a 250 GeV machine is under consideration [34]. In the TDR its length is planned to be 31 km and it is supposed to operate at the center of mass energy between $\sqrt{s} = 250$ GeV and $\sqrt{s} = 500$ GeV with a possible upgrade at $\sqrt{s} = 1$ TeV. The main linear accelerators are based on 1.3 GHz *superconducting radio-frequency* (SCRF) acceleration technology with an average accelerating gradient of ≥ 31.5 MV/m. The mass production and operation of SCRF cavities has been already demonstrated at the linear SCRF-driven synchrotron sources FLASH [35] and XFEL [36] at DESY. A schematic view of the ILC accelerator is shown in Figure 1.9.

The ILC is foreseen to collide longitudinally polarized beams. The mean polarization is expected to be $P_{e^-} = \pm 0.8$ for the electron beam and $P_{e^+} = \pm 0.3$ for the positron beam. The electron beam is produced by a laser illuminating a strained GaAs photocatode providing the necessary bunch train with 90% polarisation. Normal-conducting structures are used for bunching and pre-acceleration to 76 MeV, after which the beam is accelerated to 5 GeV in a superconducting linac. Before injection into the damping ring, superconducting solenoids rotate the spin vector into the vertical and a separate superconducting RF cryomodule is used for energy compression.

The positron source system is shown in Figure 1.10. After being accelerating in the main linac, the primary electron beam is transported through a superconducting helical undulator 147 m long. Here photons with an energy from ~ 10 MeV to

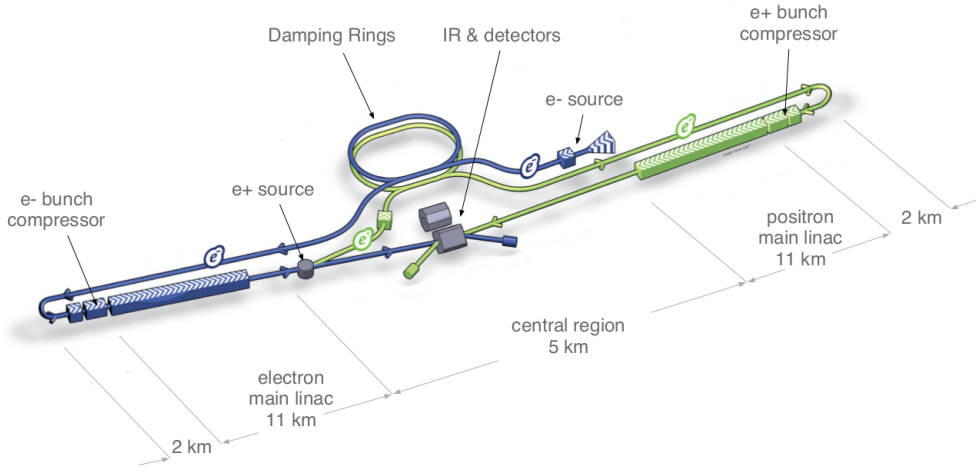


FIGURE 1.9.: Schematic layout of the International Linear Collider (not to scale) [8].

~ 30 MeV (depending on the electron beam energy) are produced. The electron beam is separated from the photons and displaced horizontally by 1.5 m. The photons from the undulator are directed onto a rotating 0.4 radiation-length Ti-alloy target placed ~ 500 m downstream, producing a beam of electron-positron pairs. This beam is accelerated up to 125 MeV using normal RF cavity. The electrons and remaining photons are separated from the positrons and dumped. The positrons are accelerated to 400 MeV in a linac with solenoidal focusing. Similar to the electron beam, the positron beam is accelerated to 5 GeV in a superconducting linac, the spin is rotated into the vertical, and the energy spread compressed before injection into the positron damping ring. The baseline design provides a polarisation of 30%. The beams are then injected into the main accelerator structure and boosted up to the full design energy.

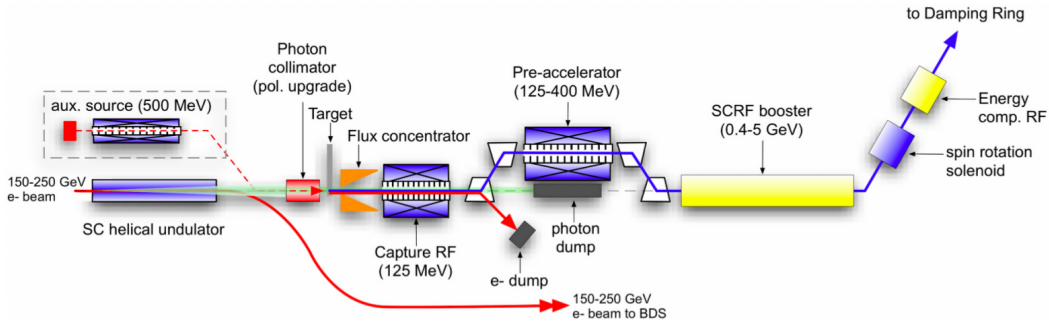


FIGURE 1.10.: Layout of the positron source at the ILC [8].

One ILC bunch train consists of 1312 bunches of 2×10^{10} particles each. The repetition frequency of bunch trains is 5 Hz (up to 10 Hz at $\sqrt{s} = 250$ GeV). Electron and positron beams are focussed to down to $474 \text{ nm} \times 5.9 \text{ nm}$ at the

interaction point, providing an average delivered instantaneous luminosity between $0.8 \times 10^{34} \frac{1}{\text{cm}^2\text{s}}$ and $1.8 \times 10^{34} \frac{1}{\text{cm}^2\text{s}}$ depending on the center of mass energy.

In order to optimise the precision reachable on the Higgs measurement and top physics while looking for new particles beyond the Standard Model, an ideal scenario has been identified. This is called *H-20* [37] and consists in operating the ILC for at least 20 years, recording at the beginning a total of 500 fb^{-1} at 500 GeV, then 200 fb^{-1} at 350 GeV and then 500 fb^{-1} at 250 GeV. After a shutdown for the luminosity upgrade after eight years of operation, additional 3500 fb^{-1} at 500 GeV and finally 1500 fb^{-1} at 250 GeV are planned to be recorded. The plan foreseen and the relatively integrated recorded luminosity are shown in Figure 1.11.

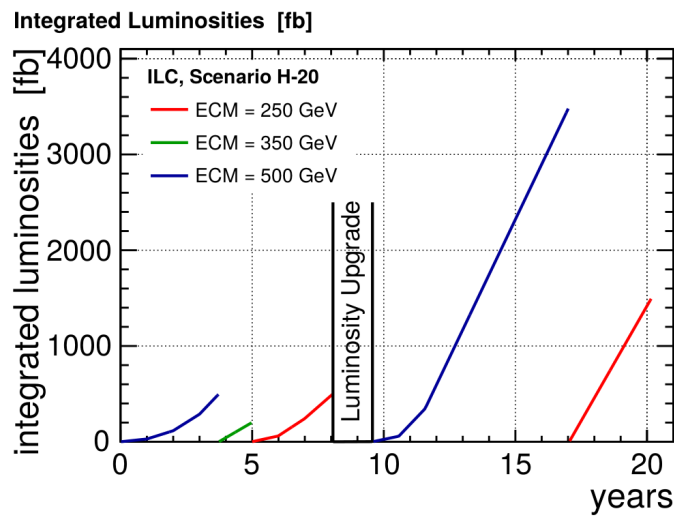


FIGURE 1.11.: Accumulation of integrated luminosity versus real time in the currently preferred ILC running scenario H-20 [37].

1.3.1. The ILC Detectors

The ILC is supposed to have two detectors, the *International Large Detector* (ILD) and *Silicon Detector* (SiD), sharing one interaction region using a push-pull approach. In this operation scheme, when one detector is acquiring data the other is out of the beam in a close-by maintenance position. At regular intervals, the detector taking data is pushed laterally out of the beam line, while the other is pulled in. These intervals are short enough in order to ensure a good data taking for both the detectors. The choice of having two detectors is motivated by the enhanced scientific productivity of past collider facilities which benefited from independent operation of multiple experiments, providing complementary strengths, cross-checking reliability and confirmation of results. Both detectors are foreseen to be multi purpose detectors, equipped with a highly granular calorimeter optimised for particle-flow analysis (see Section 2.4). The SiD is a compact detector

forseen to operate in a 5 Tesla magnetic field and using silicon tracking. Silicon enables time-stamping on single bunch crossings providing robust performance. The ILD is a large detector with robust and stable performance over a wide range of energies. The concept uses a tracking system based on a continuous-readout *time-projection chamber* (TPC) combined with silicon tracking for excellent efficiency and excellent pattern-recognition performance. The high granularity calorimeter system is placed in a 3.5 Tesla magnetic field, providing very good particle-flow reconstruction.

For both detectors, a tracking system with a momentum resolution of $\sigma(\frac{\Delta p}{p^2}) = 2 \times 10^{-5} \text{ GeV}^{-1}$ is required, driven by the reconstruction of a Higgs boson recoiling from the associated Z boson decaying to a lepton pair in the Higgsstrahlung process. The goal of the calorimeter system is to be able to provide a relative jet energy resolution of $\leq 3\%$ across the whole relevant jet energy range from 45 GeV up to 250 GeV.

To achieve this unprecedented required performance, the inner detectors of both ILD and SiD must accommodate very low mass detectors and supports. This was largely simplified by the ILC time structure of 1 millisecond bunch trains at 5 Hertz. This allows the detector subsystems to be switched off between bunch trains (so-called *power pulsing*), reducing the heat load and the need for cooling.

The International Large Detector (ILD)

The ILD is a multi-purpose detector, forseen for the application of the Particle Flow approach. It is planned in a roughly cylindrical shape of 13 m length and 7.8 m radius. The tracking system and the calorimeter system are placed within a superconducting solenoid coil of 3.4 m inner radius, generating a magnetic field of 3.5 Tesla, oriented parallel to the beam axis. A schematic view of the layout of ILD is shown in Figure 1.12.

The innermost part of the tracking system of the ILC is the vertex detector. It is realised as a multi-layer pixel vertex detector (VTX), with three superlayers each comprising two layers, or as a 5 layer geometry, as close as 16 mm to the beam axis. The vertex detector is optimised for a point resolution of less than 6 μm and minimum material thickness, less than 0.0015 radiation length per layer.

Outwards of the vertex detector, two layers of silicon strip detectors placed at 153 mm and 300 mm radius provide a hit position resolution of 7 μm with a material budget of less than 0.0065 X_0 per layer.

The main tracking detector is a large-volume TPC covering 330 mm to 1808 mm in radius. Particles traversing the TPC ionise the gas along their trajectory. The generated electrons are accelerated by an electric field applied parallel to the beam axis. The electrons coming from the ionisation of the gas are multiplied and detected by a *gas electron multiplier* (GEM) or *micromesh gaseous structure*

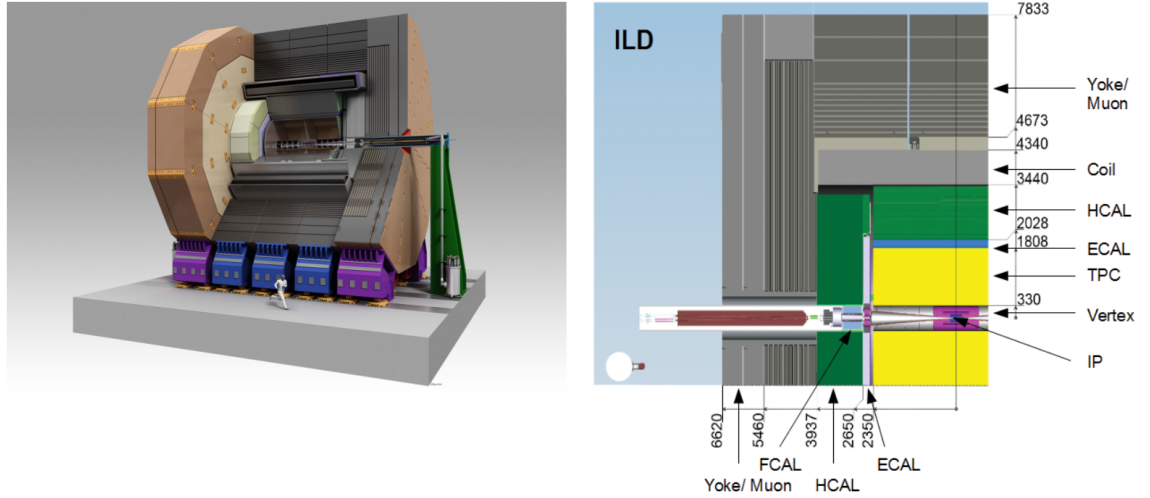


FIGURE 1.12.: Schematic view of the International Large Detector layout. On the right, the interaction point on the lower right corner is visible. Dimensions are in mm.[8].

(Micromegas) readout in the endcaps. The TPC is optimised for 3-dimensional point resolution, measuring up to 224 points per track, and minimum material in the field cage and in the end-plate. The spatial point resolution of the TPC is around $100 \mu\text{m}$ in the $r\Phi$ direction and about 1.4 mm in the z direction. It also allows particle identification based on deposited energy measurement dE/dx , from the measured ionisation strength (see Section 2.1.2).

The combined ILD tracking system is designed to achieve a momentum resolution of $\sigma(\frac{\Delta p}{p^2}) = 2 \times 10^{-5} \text{ GeV}^{-1}$ (the TPC alone could provide a momentum resolution of $\sigma(\frac{\Delta p}{p^2}) = 10^{-4} \text{ GeV}^{-1}$).

A highly segmented electromagnetic calorimeter (ECAL) provides up to 30 samples in depth and small transverse cell size, split into a barrel and an end-cap system. The absorber material foreseen is tungsten with $5 \times 5 \text{ mm}^2$ readout segmentation. For the sensitive area, silicon diodes, scintillator strips or a combination are under discussion.

The ECAL is followed by a highly segmented hadronic calorimeter (HCAL) with up to 48 longitudinal samples and small transverse cell sizes. Two options are considered, both based on a steel-absorber structure. One option uses scintillator tiles of $3 \times 3 \text{ cm}^2$, which are read out with an analogue system. The second uses a gas-based readout which allows a $1 \times 1 \text{ cm}^2$ cell geometry with a binary or semi-digital readout of each cell. More details about the current status of the hadronic calorimeter prototypes built according to these two technologies can be found in Section 2.5 and 3.3.

The ILD calorimeter system is design and optimised for the applicaiton of the Particle Flow jet energy reconstruction. The goal is to achieve a relative jet energy

resolution of $\leq 3\%$ across the whole relevant jet energy range from 45 GeV up to 250 GeV.

The magnetic field return iron yoke is instrumented with scintillator strips or resistive plate chambers (RPCs) in order to be used as a muon detector and a tail-catcher calorimeter for very high energy jets

The Silicon Detector (SiD)

The Silicon Detector concept (SiD) is very similar to the ILD concept. The main differences are in the smaller dimension and in the strenght of the magnetic field, equal to 5 Tesla. SiD is a compact multi-purpose detector, 11 m long and with a radius of 6.6 m. The tracking system is made exclusively of silicon. The calorimeter system is forseen to have a silicon electromagnetic calorimeter and a scintillator hadronic calorimeter. A schematic view of the layout of SiD is shown in Figure 1.13.

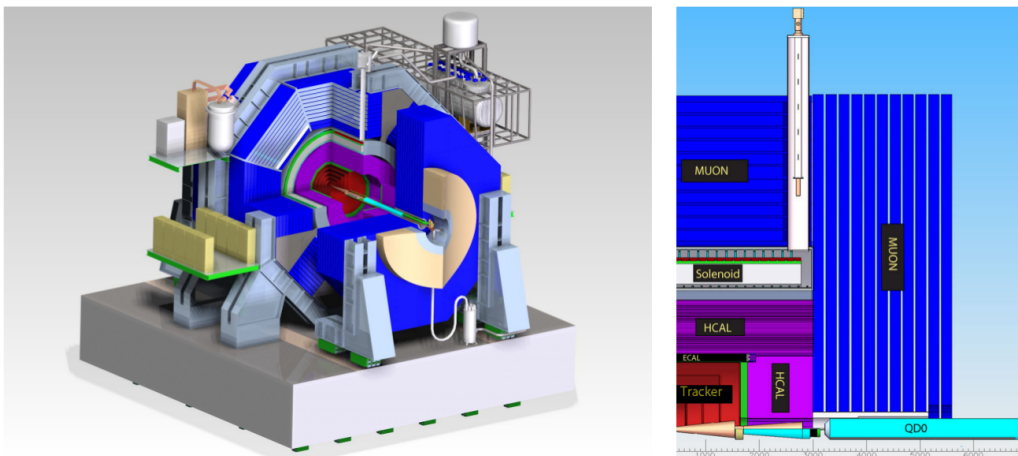


FIGURE 1.13.: Schematic view of the Silicon Detector layout. On the right, the interaction point on the lower right corner is visible [8].

2. Calorimetry and Particle Flow

High energy physics investigates a wide energy range (from the MeV to the PeV scale) through accelerator-based experiments and astronomical experiments. In accelerator-based experiments, the particles produced in each collision are studied, considering the four momenta of the created particles. In this chapter the principal mechanisms of the interaction of particles with matter and how they are simulated are described. The idea of the Particle Flow Algorithm (PFA) and its influence on the design of calorimeters for future experiments is presented.

2.1. Interaction of Particles with Matter

The way a particle deposits its energy depends on the particle itself, its energy and on the properties of the matter it interacts with. Photons and charged leptons interact mostly electromagnetically with the atoms of the material, while neutral hadrons mostly scatter with the absorber nuclei. Charged hadrons interact with the atoms of the absorber material through both the electromagnetic force and the strong force. The secondary particles created in this processes can again interact with the atoms of the absorber material, generating a cascade.

2.1.1. Electromagnetic Shower

The interaction of an electron traversing a medium depends on its initial energy. For electrons up to few MeV the dominant process is ionization, while from ~ 10 MeV, bremsstrahlung is the dominant process. In their passage through matter, electrons and positrons radiate photons as a result of the Coulomb interaction with the electric field generated by the atomic nuclei of the medium. The energy spectrum of the emitted photons falls off as $1/E$. While the energy loss due to ionization rises logarithmically with the energy of the incident electron, the energy loss for bremsstrahlung increases linearly, as shown in Eq. 2.1

$$-\left[\frac{dE}{dx}\right]_{\text{Bremsstrahlung}} = \frac{E}{X_0}. \quad (2.1)$$

The quantity X_0 represents the *radiation length*. It is characteristic of each material and defined as the mean distance over which a high energy electron reduces its energy to $1/e$ of its initial energy. A common parametrization of the

radiation length as a function of the atomic number of the material Z and its mass number A is given in Eq. 2.2 [38]

$$X_0 = \frac{716.4 \text{ [g/cm}^2\text{]} A}{Z(Z+1)\ln 287/\sqrt{Z}}. \quad (2.2)$$

In a mixture or compound it can be approximated by

$$\frac{1}{X_0} = \sum \frac{w_j}{X_j} \quad (2.3)$$

where w_j and X_j are the weight and the radiation length for the j th element in the compound. The radiation length is usually expressed in [cm] (taking into account the density of the material) or in [g/cm²]. The energy loss per radiation length in a lead absorber as a function of the electron energy is shown in Figure 2.1.

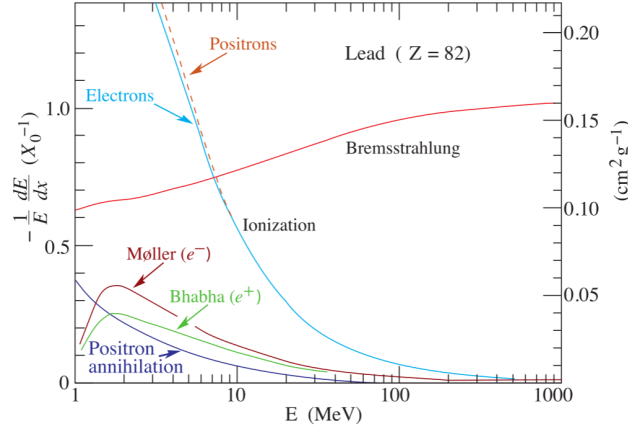


FIGURE 2.1.: Fractional energy loss per radiation length in lead as a function of the electron (positron) energy [39].

The energy at which the energy losses from radiation processes equal those from ionization is called *critical energy* ε_c . The critical energy is material dependent and can be defined as in Eq. 2.4 for a material in solid or liquid state

$$\varepsilon_c = \frac{610 \text{ MeV}}{Z + 1.24}. \quad (2.4)$$

Photons interact with matter mainly through the three different processes listed below.

Photoelectric effect

The photoelectric effect is the dominant process at low energies, up to the MeV scale. In this process, an atom absorbs the photon. The atom, that is therefore in an excited state, emits an electron or X-rays to return to the ground state. The cross section of the process is highly dependent on the number of electrons

available and therefore on the Z of the material. In particular, it is proportional to E^{-3} , therefore the relevance of this process rapidly decreases as the energy of the incident particle increases.

Coherent (Rayleigh) scattering

The Rayleigh scattering is also dominant at low energy. In this effect, the photon is deflected by the atomic electrons, but doesn't lose energy. Therefore, it affects only the spatial distribution of the energy deposition, but not the energy deposition process itself.

Incoherent (Compton) scattering

In the Compton scattering, a photon is scattered by an atomic electron in an unbound state. As for the photoelectric process, the cross section of this process decreases with the increasing energy, but less rapidly: $\sigma_{compton} \propto 1/E$ in the range from few MeV up to hundred GeV.

Electron-positron pair production

At energies higher than twice the electron rest mass ($2 \cdot 0.511$ MeV), a photon can create, in the field of a charged particle, an electron-positron pair. These particles produce then bremsstrahlung as well as ionization along their path.

The cross section of this process rises with the energy and reaches an asymptotic value at very high energy ($E > 1\text{GeV}$), related to the radiation length of the material X_0 , as shown in Eq. 2.5

$$\sigma(E \rightarrow \infty) = \frac{7}{9} \frac{A}{N_A X_0} \quad (2.5)$$

where X_0 is expressed in g cm^{-2} and the ratio between the atomic weight A and the Avogadro's number N_A corresponds to the number of atoms per gram of material. This implies that the mean free path of a very high energy photon is equal to $\frac{9}{7}X_0$. The cross section of the processes described above as a function of the energy is shown in Figure 2.2.

Electromagnetic Cascade

When a high energy photon and positron or electron enter a dense material, the combination of bremsstrahlung by electrons and pair production by photons generate a cascade. The number of particles produced in the shower almost doubles for each X_0 of traversed material, but only for a certain number of X_0 . The cascade ends when the energy of the electron is low enough and the ionization process becomes the dominant one, meaning when its energy equals ε_c defined in 2.4. The number of generated particles is therefore $N \approx E/\varepsilon_c$.

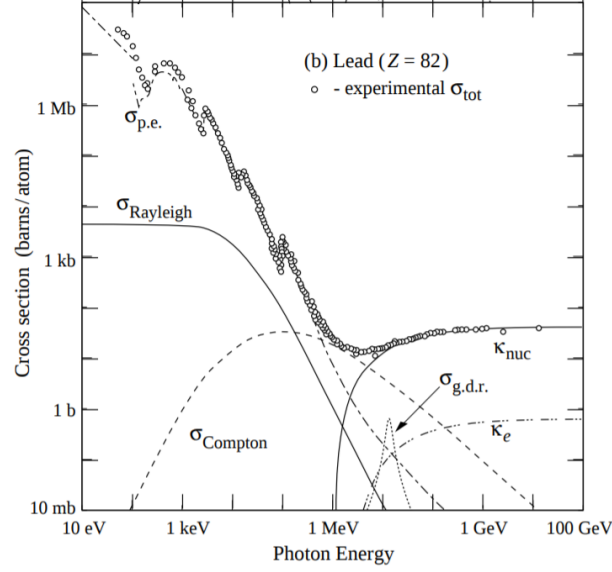


FIGURE 2.2.: Photon total cross sections as a function of energy in lead, showing the contributions of different processes. Picture from [39].

Longitudinal profiles of simulated electrons with an energy of 30 GeV incident on iron are shown in Figure 2.3. Here a cutoff energy is applied, which has been chosen as a total energy of 1.5 MeV for both electrons and photons [39]. Here, the

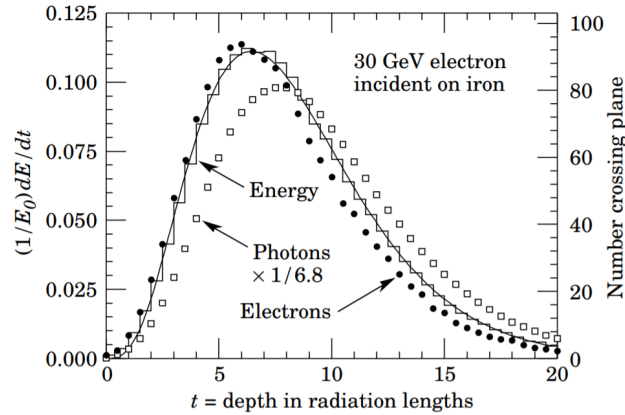


FIGURE 2.3.: Averaged longitudinal profile of 30 GeV electron-induced cascade in iron. The histogram shows the longitudinal energy deposition, fitted with a gamma function. The number of electrons (closed circles) and photons (open squares) with an energy above 1.5 MeV crossing planes at $X_0/2$ intervals (scale on the right) is also shown [39].

number of electrons, depending on the cutoff energy, falls off more quickly than the energy deposition. This is because, with increasing depth, a larger fraction of the cascade is carried by the photons.

The mean longitudinal profile of the energy deposition in an electromagnetic shower is well described by a gamma function, as given in Eq. 2.6

$$\frac{dE}{dt} = E_0 b \frac{(bt)^{(a-1)} e^{-bt}}{\Gamma(a)}. \quad (2.6)$$

The maximum of the shower occurs at $t_{max} = (a - 1)/b$. Studies have been made using different materials and in the energy range from 1 GeV to 100 GeV and the result is that the depth of the shower maximum is well approximated by

$$t_{max} = \ln\left(\frac{E}{\varepsilon_c} + C_j\right) \quad \text{with } j = e, \gamma \quad (2.7)$$

with $C_e = -0.5$ for an electron-induced cascade and $C_\gamma = +0.5$ for a photon-induced cascade.

The transverse development of electromagnetic showers in different materials scales with the *Molière radius* given by Eq. 2.8

$$R_M = 21 \text{ MeV} \frac{X_0}{\varepsilon_c}. \quad (2.8)$$

On average 90% of the energy lies in a cylinder of one Molière radius and about 99% is contained in a cylinder of roughly $3.5 R_M$. In a composite material containing several materials with a weight fraction w_j and a critical energy ε_{cj} and radiation length X_j , the Molière radius is given by

$$R_M = \frac{1}{21 \text{ MeV}} \sum \frac{w_j \varepsilon_{cj}}{X_j}. \quad (2.9)$$

2.1.2. Heavy Charged Particles Interactions

The energy loss of charged particles via bremsstrahlung depends on the mass of the particles, as shown in Eq. 2.10

$$-\left[\frac{dE}{dx}\right] \propto \frac{1}{m^4}. \quad (2.10)$$

For heavy charged particles, the energy loss due to bremsstrahlung is therefore suppressed. When traversing a material, they rather ionise and excite the atoms in the material. Ionization is therefore the main electromagnetic energy loss process for charged hadrons and muons in a wide momentum range of around $0.1 < \beta\gamma = p/Mc < 1000$. The average energy loss as a function of the momentum of the particle is described by the Bethe-Bloch equation given by Eq. 2.11

$$\left\langle -\frac{dE}{dx} \right\rangle = K z^2 \frac{Z}{A} \frac{1}{\beta^2} \left[\frac{1}{2} \ln \frac{2m_e c^2 \beta^2 \gamma^2 W_{max}}{I^2} - \beta^2 - \frac{\delta(\beta\gamma)}{2} \right]. \quad (2.11)$$

Here, z denotes the projectile charge, Z and A the atomic number and the atomic mass of the material, I the mean excitation energy of the absorber and $\delta(\beta\gamma)$ a function correcting for relativistic effects. K is a constant given by $K = 4\pi N_A r_e^2 m_e c^2$ with N_A the Avogadro's number, r_e and m_e the classical electron radius and rest mass of the electron. W_{max} corresponds to the maximum kinetic energy transferred in a single collision and is given by $W_{max} = \frac{2m_e c^2 \beta^2 \gamma^2}{1 + 2\gamma m_e/M + (m_e/M)^2}$, where M is the mass of the incoming particle. The Bethe-Bloch formula describes the mean rate of energy loss in the above mentioned range $0.1 < \beta\gamma < 1000$ for intermediate Z -materials with an accuracy of few percent. Above this energy, radiative processes dominate the energy loss. The mean energy loss for muons in copper as a function of the muon momentum $\beta\gamma$ is shown in Figure 2.4. In the "Bethe-region" the energy loss is a function of β only and it shows a broad minimum around $\beta\gamma \approx 4$. Particles that loose their energy at the minimum rate of the stopping power, e.g. $1.6 \text{ MeV cm}^2/\text{g}$ in Cu, are called *Minimum Ionising Particles* (MIPs).

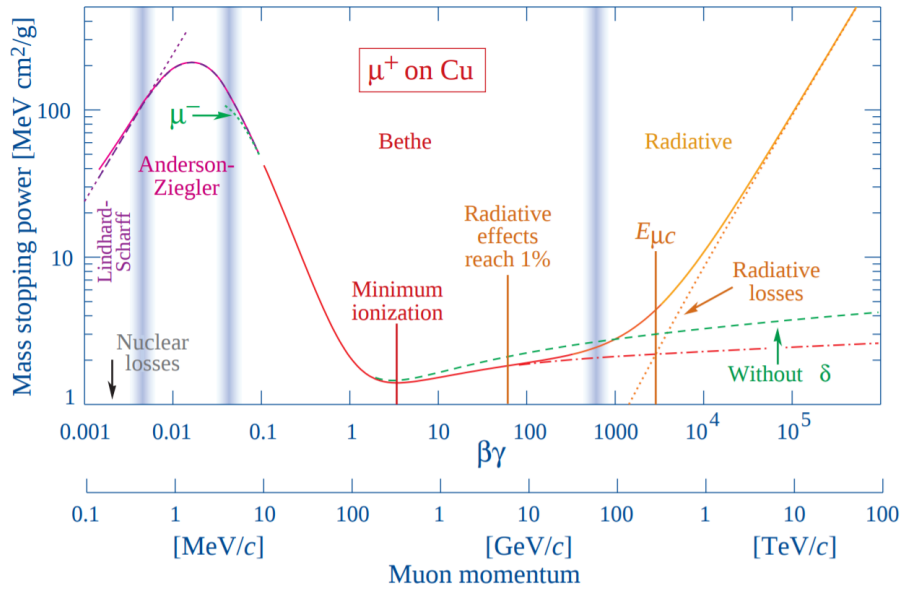


FIGURE 2.4.: Mean energy loss for muons in copper as a function of $\beta\gamma = p/MC$. The solid line indicates the total stopping power. Picture adapted from [39].

For detectors of moderate thickness x (e.g. scintillators or LAr cells), the energy loss probability distribution is reasonably described by the *Landau-Vavilov* distribution [39]. The Bethe-Bloch energy loss and the Landau-Vavilov energy loss are shown in Figure 2.5.

Because of the skewness of the distribution, in this case the most probable energy deposition is roughly 60% of the mean of the energy deposition. The mean of a Landau function is well defined only if a cut-off value at high energy is introduced. Therefore, for practical reason, the most probable energy deposition

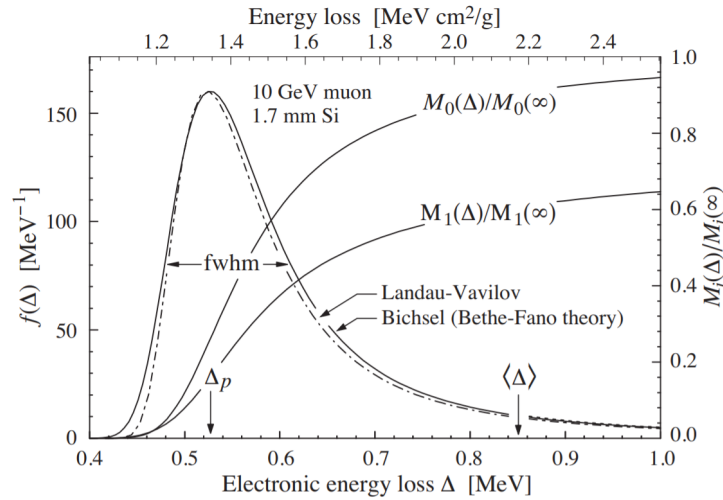


FIGURE 2.5.: Energy deposit distribution for 10 GeV muon traversing 1.7 mm of silicon, stopping power corresponding to 0.3 cm of PVC scintillator. The dot-dashed line represents the corresponding Landau-Vavilov distribution, while the continuous line the Bethe-Bloch energy loss. Picture from [39].

of single charged particles (meaning the maximum of the function) is used as a deposition scale. For very thick absorbers the distribution is less skewed but never approaches a Gaussian. The energy loss in thin absorber layers (e.g. RPC gas gaps) is not described correctly by the Landau distribution due to larger fluctuations leading to a broadening of the distribution.

2.1.3. Hadronic Shower

Charged hadrons interact with matter as described in the previous Section 2.1.2. All hadrons interact through elastic or inelastic scattering with the nuclei of the traversed material. In the elastic scattering hadrons transfer part of their energy to the hit nucleus and change their trajectory. In the inelastic process, the hit nucleus, the incident hadron, or both change their identity and multiple secondary particles can be generated. Most of them are charged or neutral pions (π^\pm or π^0) or η mesons. Nuclear reactions release protons (p) and neutrons (n) from the involved nucleus. The secondary particles generated can themselves again interact elastically or inelastically on passing the absorber material. In this way, the particle multiplicity increases and a cascade is generated. On average, about 20% of the energy transferred into the target nucleus is invisible, going into excitation or recoil of the target nucleus, or nuclear binding energy of secondary particles. In addition, neutrinos originating from the decay of charged hadrons in the cascade contribute to this invisible energy. At a certain depth, the cascade reaches a maximum of multiplicity and then the shower dies out. The mean distance a high energetic hadron passing through a medium before interacting strongly with

an atomic nucleus is given by the nuclear interaction length λ_n and defines the scale of hadronic cascades. For materials with mass number $A > 7$, it can be parametrized as in Eq. 2.12 [40]

$$\lambda_n = (20A^{0.4} + 32) \frac{\text{g}}{\text{cm}^2}. \quad (2.12)$$

Secondary π^0 and η particles produced in the interaction immediately decay in two photons, generating an electromagnetic sub-shower, which develops as explained in Section 2.1.1. The amount of energy deposited via these electromagnetic cascades in a hadron shower is referred as electromagnetic fraction f_{em} . The average electromagnetic fraction increases with the energy of the hadrons initiating the cascades according to a power law as given in Eq. 2.13

$$f_{em} = 1 - \left(\frac{E}{E_0} \right)^{(k-1)} \quad (2.13)$$

where E is the energy of the hadron initiating the cascade, E_0 is the minimum energy needed to produce a π^0 and $(k-1)$ is related to the multiplicity of π^0 mesons produced in single hadronic interactions. Typical values of E_0 are close to 1 GeV and k lies usually between 0.7 and 0.9 [41][42]. A typical process of a hadron cascade is shown schematically in Figure 2.6.

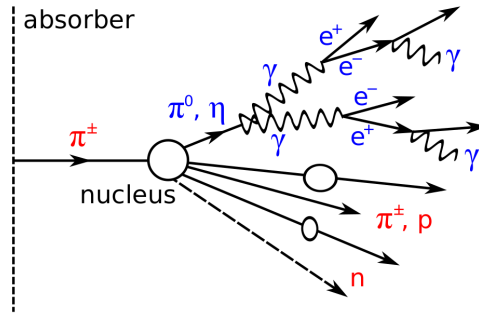


FIGURE 2.6.: Schematic representation of a hadronic cascade. A charged π^\pm traversing a material strongly interacts with a nucleus of the material. Secondary particles are produced. π^0 and η mesons decay instantaneously in two photons, generating an electromagnetic sub-shower (blue). The other secondary hadrons (red) traverse the material until interacting further. Picture from [43].

Compared to electromagnetic showers, hadronic showers develop longitudinally much further, so that the ratio $\lambda_n/X_0 \gg 1$. In addition, they are also much more complex, since they are characterized by a few hard hadronic interactions with a strongly variable number of secondaries per interaction. Therefore, there are large statistical fluctuations in the longitudinal as well as transverse shower extensions, in the invisible energy fraction and in the fraction of energy in electromagnetic

sub-showers. To longitudinally contain $\sim 95\%$ of the energy deposited by a 100 GeV charged pion, a calorimeter of $\sim 8\lambda_\pi$ is needed. Transversely, a cylinder of radius $\sim 1\lambda_\pi$ is needed to contain roughly on average 90% of the full shower energy. Here λ_π is the pion interaction length and it is typically larger by a factor of $\frac{3}{2}$ compared to λ_n .

While electromagnetic showers develop instantaneously due to the prompt generation of particles with relativistic energies, this is not the case for hadronic showers. They also contain a hadronic component that can be delayed by a few microseconds due to nuclear excited states which de-excite on this timescale. In addition, thermal neutrons are emitted, causing a delayed energy deposition from few microseconds up to a few seconds. Neutrons can in addition travel a significant distance through the detector, before being absorbed. These effects have an impact on the reconstructed energy and spatial development of the shower, especially when taking into account the time acceptance of the readout.

2.2. Simulation of Particle Cascades

Simulations of physical processes are a fundamental tool to understand and validate experimental results. In addition they are useful for design and optimization studies and serve as a guideline for data analysis.

The simulations in this thesis have been performed using the GEANT 4 framework [44] [45] [46] [47] for the simulation of the physics processes involved in a cascade, together with MOKKA [48], that provides a realistic description of the geometry of the detector, its material and the interaction of the particles traversing it.

Electromagnetic interactions are considered well understood, due to the theoretically well described interaction of electrons, positrons and photons. In GEANT 4, the electromagnetic interactions are simulated with a standard EM package, that reproduces observables in sampling calorimeter with a precision better than 1% [49][50].

Hadronic interactions are much more complex, therefore their simulations are treated in more detail in the next section.

2.2.1. Simulation of Hadronic Cascade

The simulation of a hadronic cascade is complicated, due to the fact that both the incident particle and the nuclei of the target have an internal structure. Therefore, many secondary particles can be produced in a very large phase space of possible final states. Typically, this cannot be described analytically and models, parametrizations and approximations of the involved processes are used. The basic assumption is that the De Broglie wavelength $\lambda = \frac{h}{p}$ represents the scale of

which the internal substructure of the projectile and the target has to be taken into account in the interaction. At low energy, up to few hundred eV, it is sufficient to model the incident particle and the nucleon in the target without taking into account their sub-structure. At high energy, meaning from few GeV, and at small λ , the quarks forming both the projectile and the nucleon have to be taken into account. Therefore, for different energy ranges different simulation models are applied. In a typical hadronic interaction model, in the first step the interactions of the projectile with the target's nucleus are computed, giving an excited nucleus plus eventually secondary particles. In the second step, the de-excitation of the nucleus from the production of other secondary particles up to its full fragmentation is simulated.

Cascade Models

If the energy of the incident hadron is between few hundred eV and few GeV, the De Broglie wavelength λ is around the distance between nucleons but larger than the quark sub-structure. Therefore an interaction model based on individual nucleons is used. In the cascade model provided by GEANT 4, the path of the incoming particle and of the secondary particles generated in the interaction is tracked. The path length between interactions inside the nucleus is computed from the modeled nucleon densities and the parametrised cross sections. A schematic picture of the cascade model is shown in Figure 2.7. Cascade models treat the

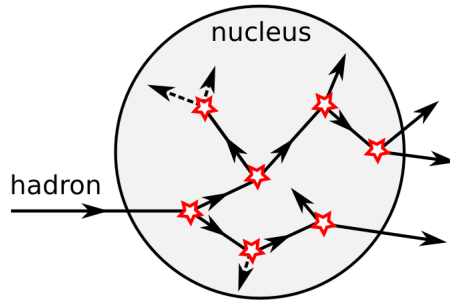


FIGURE 2.7.: Schematic representation of the cascade model implemented in GEANT 4. Picture from [43].

nucleus as a Fermi gas, i.e. the nucleons occupy all states of the system up to the Fermi energy. Due to the Pauli exclusion principle, a minimum energy larger than the Fermi energy is needed so that secondary particles can be produced in the interactions. Within a collision, the nucleus is excited. GEANT 4 describes this excitation through excitons, i.e through the number of excited particles and holes. The cascade stops when all the particles are either absorbed or leave the nucleus, or the remaining nuclear fragments are transferred to an equilibrium state and de-excited.

In the Bertini cascade model [51], the nucleus is described as a spherical shell of constant nucleon density. The nucleons inside the shell are assumed to have a momentum distribution as in a Fermi-gas. For each collision, the model determines the momentum of the hit nucleon, the type of reaction, the products of the reaction and their four momenta. The Bertini cascade model includes a pre-equilibrium model to describe evaporation using the exciton configuration. It also describes the de-excitation of the nucleus via Fermi-breakup, a simple explosion model, a phenomenological fission model, and an evaporation model at equilibrium. More details about the Bertini cascade model implemented in GEANT 4 can be found in the GEANT 4 Physics Reference Manual [44].

String Parton Models

Parton string models [52] describe the scattering of a hadron on nuclei at high energies, usually above 5 GeV. At these energies, the interaction between the quarks of the incident hadron and the nucleons governs the scattering process. The struck nucleus is modeled with protons and neutrons, using the potential of a harmonic oscillator for light nuclei (with $A < 16$) or the Wood-Saxon potential for heavier nuclei. A random momentum between zero and the Fermi energy is assigned to each nucleon. The projectile interacts with each single nucleon. The interaction is determined using the impact parameter of the incident hadron, the center of mass energy of the interaction and inelastic and diffractive cross sections. From the interaction of two quarks, one from the projectile and one from the nucleon, a string is formed. It is described by its four momentum and the two quarks on its ends. The string stretches due to the movement of the constituents. The model splits then the string into quark-antiquark pairs and new strings. The quarks produced form hadrons (hadronization process). A schematic picture of the working principle of the string parton model is shown in Figure 2.8.

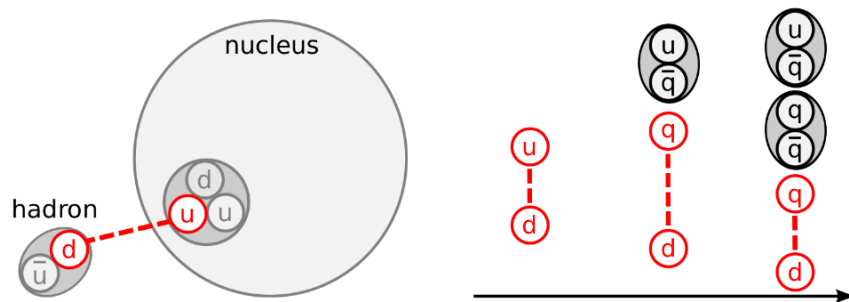


FIGURE 2.8.: Schematic representation of string parton model provided by GEANT 4. On the left, string formation via the interaction of one quark from the projectile and one quark from the nucleon. On the right, string fragmentation via the generation of quark-antiquark pairs and hadronization of the produced quarks. Picture from [43].

In GEANT 4, two different models have been implemented to model the parton string model. These are the *Fritiof* and the *quark-gluon string* model. They differ in the treatment of the string formation. In the Fritiof model, only the momentum exchange between the projectile and the nucleon in the diffractive hadronics interaction is considered [53]. The quark-gluon string model includes inelastic scattering processes mediated by pomerons [54][55].

Other more simple models are based on fits to experimental data to predict the production of secondary particles in hadronic interactions. In these models, the interactions are not modeled in detail and the energy is conserved only on average and not in each single interaction. GEANT 4 provides the Low Energy Parametrization (LEP) and the high energy parametrization (HEP) models, both adapted from the GHEISHA package [44][56].

2.2.2. Physics List

A physic list in GEANT 4 is a set of physics models for all particles involved in an interaction. Each model acts in a defined energy range, in order to realistically describe the interaction of a hadron interacting with the nuclei of a certain material over a wide energy range. A smooth transition between the models is realized overlapping the energy and choosing randomly between one of the two models for each interaction [57].

In this thesis, the QGSP_BERT_HP physics list is used in GEANT 4 version 10.1, since it is considered the most stable and well tuned [58]. It uses

- the quark-gluon string model for protons, neutrons, pions and kaons interactions with nuclei at energies above 12 GeV;
- the LEP model for protons, neutrons, pions and kaons interactions with nuclei at energies between 9.5 GeV and 25 GeV;
- the Bertini cascade (BERT) model for protons, neutrons, pions and kaons interactions with nuclei at energies below 9.9 GeV;
- parametrised models (LEP + HEP) for all remaining hadrons (i.e. hyperons);
- the HP package (High Precision) low energy neutron transport model;
- the standard electromagnetic physics list.

A schematic drawing showing the overlap between the hadronic models of the QGSP_BERT physics list, for incident protons, neutrons, pions and kaons, is shown in Figure 2.9.

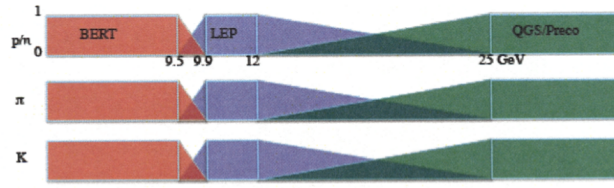


FIGURE 2.9.: Schematic representation of the overlap between hadronic models for incident protons, neutrons, pions and kaons in the QGSP_BERT physics list. Picture from [57].

2.3. Calorimeters

In nuclear and particle physics, calorimetry refers to the detection of particles and the measurements of their properties through the total absorption of the particles themselves within a block of material, the calorimeter. Calorimeters fulfill several crucial aspects, ranging from event selection and triggering to precision measurements of the properties of individual particles as well as jets. In data analysis, their contribution is mainly focused on particle identification (electrons, photons, muons) and on the energy measurements of particles that generate hadronic and electromagnetic showers (e, γ, π^0). In $ep, p\bar{p}, e^+e^-$ experiments, the measurements of single hadrons and jets is a fundamental task for the calorimeter (see Section 2.4).

The energy deposited by a particle shower is usually proportional to the energy of the incident particle. The energy deposited in the calorimeter can be measured in two ways. Using scintillating materials which emit photons from atomic de-excitation, or with charge sensitive media that allow amplification and collection of free charges generated from ionization in the material, like silicon diodes, liquid argon or specific gas mixtures.

Calorimeters are distinguished in two classes:

- *Homogeneous* calorimeters, in which the absorber and the active material are one and the same.
- *Sampling* calorimeters, in which the absorber is distinct from the active material.

For a sampling calorimeter, in which the absorber material and the active layer are alternated, the fraction of energy measured in the active layers can be quantified using the electromagnetic sampling fraction. It is defined as the fraction of the full shower energy deposited in the sensitive material by a purely electromagnetic shower, which to first order is equivalent to Eq. 2.14

$$f_s = \frac{[X_0/layer]_{sens}}{[X_0/layer]_{abs} + [X_0/layer]_{sens}} \quad (2.14)$$

where $[X_0/\text{layer}]_{sens}$ and $[X_0/\text{layer}]_{abs}$ correspond to the radiation length of one layer of sensitive material and the radiation length of one layer of absorber material.

Calorimeter Response

The response of a calorimeter is defined as the average calorimeter signal per unit of deposited energy. The response is therefore expressed in terms of photo electrons per GeV or pico-Coulombs per MeV. From this definition follows that a linear calorimeter has a constant response.

Electromagnetic calorimeters are in general linear, since all the energy carried by the incoming particle is deposited through processes that generate signals (excitation/ionization of the absorbing medium). Non-linearities are an indication of instrumental deficiencies, like saturation or non-containment of the shower (leakage). This is not true for hadronic calorimeters. The non-linearity of the signal in such a calorimeter is mainly due to the invisible energy characterizing the hadronic shower and the electromagnetic sub-shower developing within a hadronic shower as explained in Section 2.1.3. The response of the calorimeter to the hadronic component and the electromagnetic component is not the same and the difference can be up to 20% – 30%. The calorimeter is therefore called a *non-compensating* calorimeter. The ratio between the electromagnetic and the non-electromagnetic responses is known as the e/h ratio of the calorimeter and for a non-compensating calorimeter it differs from one. This ratio cannot be directly measured. Considering the response of the calorimeter to a pion initiating a shower, $\pi = \langle f_{em} \rangle e + (1 - \langle f_{em} \rangle)h$, the ratio e/h can be derived from the e/π signal ratio. The relation between the e/h and the e/π ratios is given by Eq. 2.15

$$e/\pi = \frac{e/h}{1 - \langle f_{em} \rangle + \langle f_{em} \rangle (e/h)} \quad (2.15)$$

where f_{em} represents the average electromagnetic shower fraction. Calorimeters for which $e/h > 1$ are called *undercompensating*, while calorimeters for which $e/h < 1$ are called *overcompensating*. Most calorimeters used are undercompensating, with a typical value of e/h between 1.5 and 2. Compensation can be achieved either using specific materials as explained below or off-line in the data analysis [59].

Most important in this effect are the neutrons. Low energy neutrons ($E_n < 20$ MeV) mostly lose their energy through elastic interactions in which they transfer their energy to the nuclei of the absorber material. The energy transferred is maximal when $\Delta m = m_n - m_{target} = 0$. In absorber materials with very high Z , these neutrons will deposit very little energy. If the active material is rich of hydrogen (e.g plastic scintillators), low energy neutrons mostly transfer their momentum to the hydrogen nuclei. Because the range of MeV protons is small in plastic scintillators,

the neutron fraction of the full shower energy is deposited almost completely in the plastic scintillator. Compensation in the calorimeter can therefore be reached by optimizing the absorber material, the active material, their respective thicknesses and the sampling factor of the calorimeter. In this way, the deposited energy loss due to invisible energy can be compensated by amplifying the deposition of low energy neutrons. As an example, the calorimeter of the ZEUS experiment was designed in this way to be compensating. It was a sandwich calorimeter, where layers of uranium ($Z = 82$) were alternated with layers of plastic scintillator. The thicknesses of the absorber material of 3.2 mm and active material of 3.0 mm were carefully optimized in order to achieve a ratio $e/h = 1.00 \pm 0.03$ [60].

Calorimeter Energy Resolution

The precision of a calorimeter is limited by fluctuations. These include: signal quantum fluctuations (i.e photoelectron statistics); shower leakage fluctuations; fluctuations resulting from instrumental effects (electronic noise, light attenuation, structural non-uniformities); sampling fluctuations. In a sampling calorimeter, the latter one are the dominant one. These are governed by the Poisson statistics and therefore they contribute to the energy resolution through a term that is proportional to $1/\sqrt{E}$. In practice, the energy resolution of a calorimeter is subjected to different sources of fluctuations, each with its own energy dependence. For a full description of the relative calorimetric energy resolution a noise term, describing the resolution effects due to the fluctuating pedestal level of individual readout channels, and a constant term, describing the minimum achievable energy resolution due to detection inefficiencies and calorimeter inhomogeneities, are added in quadrature. The energy resolution of a calorimeter can therefore be expressed as in Eq. 2.16

$$\frac{\sigma}{E} = \frac{A}{\sqrt{E}} \oplus \frac{B}{E} \oplus C \quad (2.16)$$

where A represents the stochastic term, B the noise term and C the constant term.

The sampling fluctuations degrade the electromagnetic energy resolution by a factor of $1/f_s$, with f_s defined in 2.14. Sampling fluctuations are larger for hadron showers than for electromagnetic showers. This is due to the fact that the number of different shower particles $N_{meas} = f_s N$ that contribute to the calorimeter signal is smaller for hadron showers. The hadronic signals are in fact dominated by the contribution of spallation protons, carrying few MeV of energy. Contrary to the Compton and photo-electrons, these protons can travel several active layers of the calorimeter.

Examples of Calorimeters

Here, few examples of the energy resolution achieved with electromagnetic and hadronic calorimeters are presented.

The CMS electromagnetic calorimeter is an example of a homogeneous calorimeter. It is made of fine grained lead tungstate (PbWO_4) crystals and reaches a resolution of $\frac{\sigma}{E} = \frac{2.8\%}{\sqrt{E[\text{GeV}]}} \oplus 0.3\%$ [61]. The ATLAS calorimeter is an example of an electromagnetic sampling calorimeter. It is a lead-liquid argon sampling calorimeter with a sampling fraction of roughly 4% – 6% and it reaches a resolution of $\frac{\sigma}{E} = \frac{9.2\%}{\sqrt{E[\text{GeV}]}} \oplus 0.2\%$ [62][63].

For hadronic calorimeter, fluctuations in the visible energy play an important role and they represent the ultimate limit to the achievable energy resolution. In non-compensating calorimeters, the non-Poissonian fluctuations in the electromagnetic fraction tend to dominate the hadronic performance. The resolution of a compensating calorimeter is limited by fluctuations in the visible energy. The importance of these fluctuations depend on how compensation is achieved. In plastic scintillator calorimeters, the signal from neutrons is correlated with the nuclear energy losses, especially for materials with high atomic number. In this case, the fluctuations are reduced. A typical value of the energy resolution for a hadronic calorimeter is $\sim 60\%/\sqrt{E[\text{GeV}]}$. The compensating ZEUS hadronic calorimeter could reach an energy resolution of $\frac{\sigma}{E} = \frac{35\%}{\sqrt{E[\text{GeV}]}} \oplus 2\%$ [60]. In comparison, the ATLAS Tile hadronic calorimeter reached an energy resolution of $\frac{\sigma}{E} = \frac{52\%}{\sqrt{E[\text{GeV}]}} \oplus 3\%$ for single pions [64], while the CMS hadronic calorimeter can reach a energy resolution of $\frac{\sigma}{E} = \frac{84\%}{\sqrt{E[\text{GeV}]}} \oplus 7\%$ [65].

2.4. Particle Flow Approach

In high-energy physics experiments single particles are measured. In the hard interaction quarks and gluons are produced. They are fragmenting in several hadrons, that are boosted into the original direction of the quark and are called jets. In a jet, $\sim 62\%$ of the energy is carried by charged particles, around $\sim 27\%$ by photons, about $\sim 10\%$ by long lived neutral hadrons (e.g n, \bar{n}, K_L) and $\sim 1.5\%$ by neutrinos. Roughly 72% of particles forming a jet reaches the calorimeter, therefore $\sim 72\%$ of the energy of a jet is measured by the calorimeter system (ECAL and HCAL). For this reason, the energy resolution is limited by the relatively poor resolution of the hadronic calorimeter, usually $\sim \frac{60\%}{\sqrt{E[\text{GeV}]}}$.

To remedy this effect, the *Particle Flow Approach* (PFA) has been developed. The Particle Flow consists in measuring the energy of every single particle in a jet, using the subdetector measurement yielding the best resolution for each particle. The momenta of charged particles are measured in the tracking detectors, while

the energy measurements for photons and neutral hadrons are obtained from the calorimeters. In this manner, the HCAL is used to measure only $\sim 10\%$ of the energy in the jet. The reconstructed jet energy is then the sum of the energies of the individual particles. In Figure 2.10, a sketch explaining the Particle Flow approach is shown.

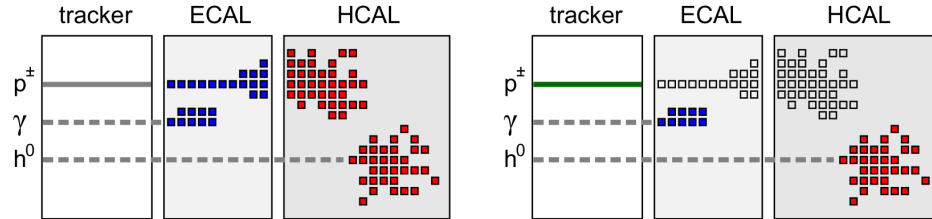


FIGURE 2.10.: Schematic picture of a purely calorimetric and the Particle Flow jet energy reconstruction. In the purely calorimetric reconstruction (left) only the information from the ECAL and HCAL are used. In the Particle Flow energy reconstruction approach (right) the information from the tracker are used to reconstruct the energy of charged particles, instead of their calorimeter deposition. Picture from [43].

In practice, it is not possible to perfectly associate all energy deposits with the correct particles. For example, the wrong reconstruction of two close tracks as a single one can cause the full loss of the energy of one or both particles. Or if the calorimeter hits from a photon are not distinguished from a charged hadron shower, the photon energy is not considered. Similarly, if part of a charged hadron shower is identified as a separate cluster, the energy is double counted, since it is already considered in the tracking. These effects contribute to the so called *confusion* term. This confusion term represents the limiting factor of the Particle Flow Approach, rather than the poor calorimetric performance of the hadronic calorimeter. This places stringent requirements on the detector system:

1. The tracking system needs to have an excellent track reconstruction efficiency, higher than 99%.
2. The calorimeter has to have a high granularity in order to assign calorimeter energy depositions to single particle tracks and to identify neutral hadrons energy depositions, not associated to any track.

The improvement in the jet energy resolution by using the *PandoraPFA* [66] implementation of Particle Flow can be seen in Figure 2.11. Here the jet energy resolution as a function of the jet energy itself is shown. A jet energy resolution of about 3% in the jet energy range between 45 GeV and 250 GeV (solid black line), as foreseen for ILD, can be reached using the PandoraPFA algorithm. This demonstrates the performance of a detector design optimised for Particle Flow reconstruction.

The purely calorimetric approach (blue dotted line) is also shown. Comparing it with the Particle Flow approach, it is relevant to notice that the Particle Flow approach significantly improves the resolution, even for the highest energies. The degradation of the energy resolution for higher energy for the PFA and the realistic (ILD) calorimeter is due to the non-containment of hadronic showers. Here the performance of a Particle Flow calorimeter is also compared to the $50\%/\sqrt{E[\text{GeV}]} \oplus 3.0\%$, a typical resolution achievable with a traditional calorimetric approach (red dotted line). The confusion term (black dotted line) is also shown.

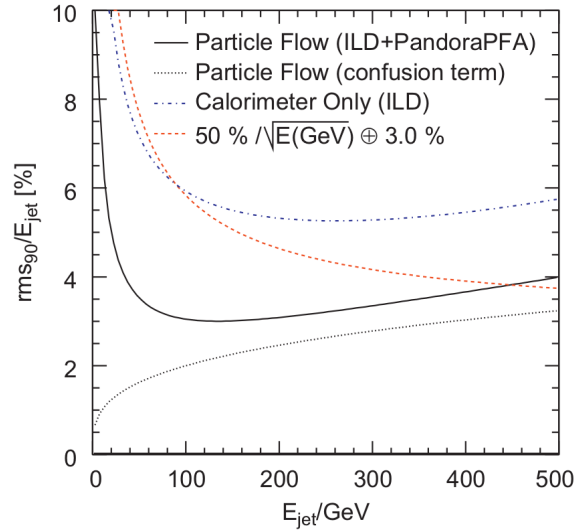


FIGURE 2.11.: Jet energy resolution obtained from Particle Flow calorimetry (PandoraPFA and the ILD concept, full line). The dotted line represents the estimate confusion term. The dot-dashed curve shows a parametrization of the jet energy resolution obtained from the total calorimetric energy deposition in the ILD detector. The dashed $50\%/\sqrt{E[\text{GeV}]} \oplus 3.0\%$ curve is shown to indicate the achievable jet energy resolution using the traditional calorimetric approach [66].

2.5. CALICE Detector Concepts

The CALICE collaboration (CALorimeter for LInear Collider Experiment) develops and validates concepts for highly granular sandwich calorimeters, both hadronic and electromagnetic, foreseen for the application of the Particle Flow approach. The collaboration consists nowadays of 57 institutes from 17 countries [67]. The CALICE calorimeters are optimised for a linear collider environment such as ILC [68][69] or CLIC [70][71]. Recently, cooperation with the Large Hadron Collider (LHC) upgrade group started [72].

The physics performance of the different calorimeter concepts has been demonstrated with the physics prototypes, while the technological prototypes aim to prove the scalability of the calorimeters to a full linear collider detector.

A detailed description of the Analog Hadron Calorimeter (AHCAL) prototype is given in Chapter 3. Here, a short overview of the other possible options is presented. More details of these can be found in [73].

CALICE ECALs

The CALICE collaboration has developed two different electromagnetic sandwich calorimeters: the scintillator-based ECAL (Sci-ECAL) [74] and the silicon-based ECAL (Si-ECAL) [75], both using tungsten as absorber material. Both physics prototypes have been used in several test beam campaign [76].

The Sci-ECAL physics prototype consists of 30 active scintillator layers alternated with 30 passive tungsten absorber plates of 3.5 mm thickness. Each active layer is made of four rows of scintillator strips of $45 \times 10 \times 3$ mm³. The strips are oriented orthogonally in consecutive layers. A sketch of the prototype is shown in Figure 2.12(a). The light produced is collected through a wavelength-shifting fibre and read out with silicon photomultipliers (see Section 3.1). Overall, the prototype has 2160 readout channels. Its total thickness is $20 X_0$ corresponding to roughly $1 \lambda_n$. The Molière radius is approximatively 22 mm [74]. This prototype has been operated at Fermilab and an energy resolution with a stochastic term of $12.5\%/\sqrt{E[\text{GeV}]}$ and a constant term of 1.2 % has been obtained [77].

The Si-ECAL physics prototype is also made of 30 active layers and 30 absorber layers. A sketch of the prototype is shown in Figure 2.12(b). The silicon active layers of the dimension of 18×18 cm² are segmented into 1×1 cm² pad diodes, resulting in 9720 readout channels in total. The thickness of the passive layers increases with the depth of the prototype. It goes from 1.4 mm for the first 10 layers to 2.8 mm for the second group of 10 layers and up to 4.2 mm for the last 10 layers. The total thickness of the Si-ECAL prototype is $24 X_0$ corresponding to $1 \lambda_n$. In test beam campaigns at DESY and CERN this prototype achieved an energy resolution with $16.53\%/\sqrt{E[\text{GeV}]}$ stochastic term and 1.07 % constant term [75].

The Si-ECAL technological prototype and the Sci-ECAL technological prototype with fully integrated readout electronics, that aim to prove the scalability of the calorimeters to a full linear collider detector are under development.

CALICE HCALs

In addition to the Analog Hadronic Calorimeter (AHCAL), described in more detail in Chapter 3, the CALICE collaboration has developed two other hadronic

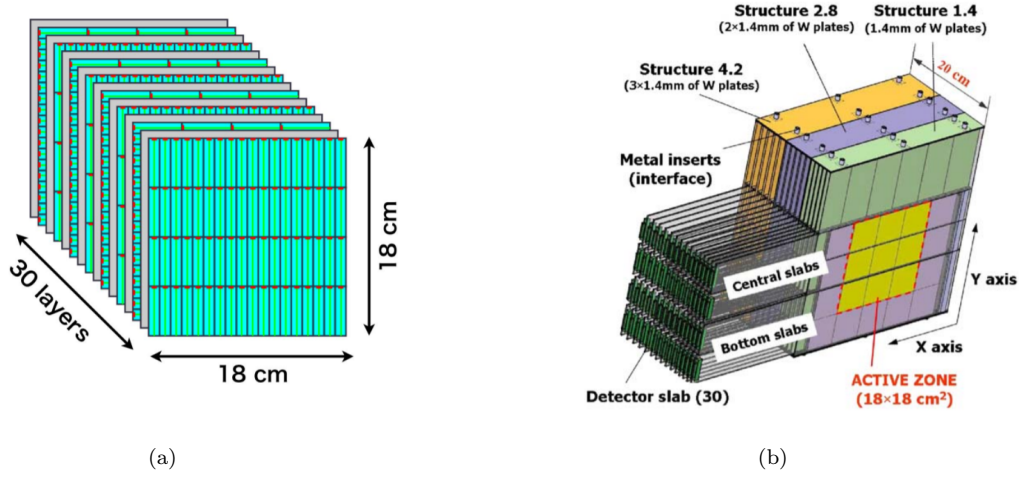


FIGURE 2.12.: CALICE ECAL prototypes.

(a): structure of the scintillator ECAL test module in Sep 2008 at FNAL [77].

(b): Schematic view of the silicon ECAL [75].

sandwich calorimeters: the Digital Hadronic Calorimeter (DHCAL) and the Semi-Digital Hadronic Calorimeter (SDHCAL).

The Digital Hadronic Calorimeter (DHCAL) [78] is a sandwich calorimeter with Resistive Plate Chamber readout (RPC). The RPCs used in this prototype are made of two glass plates with a 1.15 mm gap in between filled with gas. They are read out by pad diodes of $1 \times 1 \text{ cm}^2$ size, placed on the back of the plates. The incident charged particles traverse the gas and ionise it. The ionisation is amplified through avalanche processes induced by the high bias voltage of around 6.4 kV applied between the plates. The charge multiplication is quenched by the high resistivity of the glass of about $4.7 \cdot 10^{13} \Omega\text{cm}$ [78]. If the charge is above the threshold a hit is measured [79].

The DHCAL prototype consists of 38 active layers. Each layer is made of three RPC chambers of $32 \times 96 \text{ cm}^2$ stacked on top of each other, creating an active area of $1 \times 1 \text{ m}^2$. The chambers are placed in a cassette. The front plate consists of 2 mm copper, while the back plate consists of 2 mm steel. The thickness of the RPC used in this prototype is in total 8.3 mm thickness [80]. They are filled a mixture of three gas: tetrafluoroethane $R134A$ (94.5%), isobutane C_4H_{10} (5.0%) and sulfur hexafluoride SF_6 (0.5%). A schematic picture of the RPC is shown in Figure 2.13. Each RPC is read out by two front-end boards of $32 \times 48 \text{ cm}^2$. Therefore, each layer contains three RPCs and six front-end boards. Each board has 24 chips, meaning a total of 144 chips per layer. Each chip reads out 64 pads of $1 \times 1 \text{ cm}^2$ size. This sums up to 3072 channels per RPC, 9216 channels per layer and 350,208 for the prototype. As absorber material, both tungsten and steel have been tested, covering respectively $4.84\lambda_n$ and $5.3\lambda_n$. An energy resolution of $\frac{\sigma}{E} = \frac{64\%}{\sqrt{E[\text{GeV}]}} \oplus 4\%$ has been achieved in test beam with pion data [82].

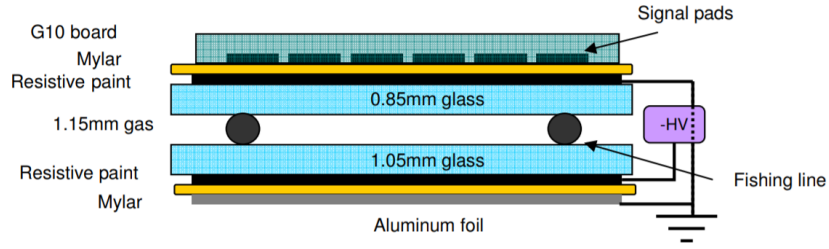


FIGURE 2.13.: Schematic of the cross section of an RPC of the DHCAL prototype [81].

The Semi-Digital Hadronic Calorimeter (SDHCAL) prototype is made of 48 active layers inserted in a steel absorber structure with a layer thickness of 1.5 cm [83], for a total length of $5.76\lambda_n$. Each active layer consists of 2 glass RPCs with a gas gap of 1.2 mm each, which is protected by ceramic ball spacers, and a readout electrode segmented in $1 \times 1 \text{ cm}^2$ pads. A cross section of the RPCs used in the SDHCAL is shown in Figure 2.14. The gas used is a mixture of tetrafluoroethane $R134A$ (93%), carbon dioxide CO_2 (5%) and sulfur hexafluoride SF_6 (2%). The high voltage applied to the RPCs was set during test beam to a fixed value of 6.9 kV. The principle of detecting particles in the SDHCAL is the same as in the DHCAL. The main difference between the two prototypes is the readout. In the SDHCAL three different thresholds can be set. This doesn't allow to estimate the deposited energy, but to distinguish whether the recorded charge is the result of one, few or many charged particles traversing one cell. The SDHCAL prototype has been tested at CERN during several test beam campaigns. An energy resolution from $\sim 12.9\%$ for 30 GeV pions and up to $\sim 7.7\%$ for 80 GeV pions has been reached. It is important to mention that these values, obtained with the three thresholds readout mode shown an improvement compared with the digital readout mode. In this case an energy resolution of $\sim 13.3\%$ for 30 GeV pions and up to $\sim 10.6\%$ for 80 GeV has been obtained [84].

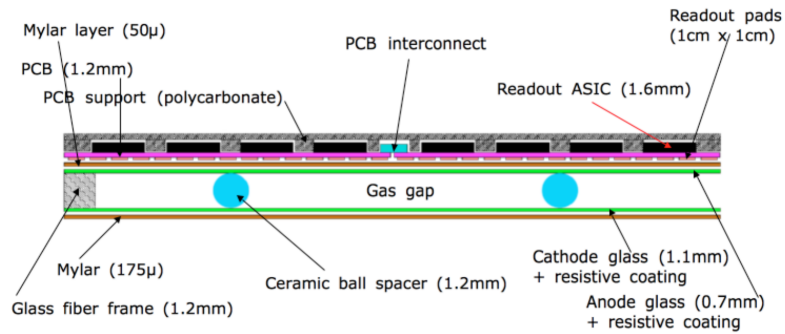


FIGURE 2.14.: Schematic cross section of an RPC of the SDHCAL prototype [84].

3. The AHCAL Technological Prototype

As already mentioned in Chapter 2, the CALICE collaboration develops and validates concepts for highly granular sandwich calorimeters, both hadronic and electromagnetic, foreseen for the application of the particle flow approach.

The CALICE Collaboration has developed several concepts for highly granular calorimeter and built prototypes in order to prove the feasibility of such a detector, the physics capability and the scalability to a large detector. The *Analog Hadron Calorimeter* (AHCAL) is one of these detector concepts. It is a sandwich calorimeter with small scintillator tiles as active material and steel or tungsten as absorber material. Each scintillator tile is read out by a silicon photomultiplier.

The physics performance of such a high granularity calorimeter has been proved with the AHCAL *Physics Prototype*, which has been used in several CALICE test beam campaigns from 2006 to 2012 (more details in [85]).

The AHCAL *Technological Prototype* has been developed since several years in order to prove the scalability of the calorimeter to a full linear collider detector. First test of large multi-layer system have been done in 2014 – 2015.

In this chapter the AHCAL technological prototype, its hardware and its calibration procedures are explained.

First a description of the silicon photomultiplier photodetectors and their readout chips are presented. The chapter then proceeds with the characterization of the different kinds of calibrations that have to be performed and the procedures used. These concern the gain calibration of the silicon photomultipliers and the MIP calibration needed to convert the signal output of the silicon photomultipliers to Minimum Ionizing Particle (MIP) units. In addition the timing calibration, in order to have the time output expressed in nanoseconds and the high gain - low gain intercalibration due to the dual mode working principle of the readout chips are presented.

After that the digitization process of the simulation is discussed, which includes the modeling of all readout effects resulting from the combination of the active material used, the sensors and the electronics.

3.1. Silicon Photomultiplier

A Silicon Photomultiplier (SiPM) is a solid state photodetector that in response to the absorption of an incident photon measures a pulse of few nanoseconds length, corresponding to 10^5 - 10^6 electrons. This is defined as the gain of a SiPM and it is comparable to the gain of a photomultiplier tube (PMT)[86].

The reason why SiPMs are becoming very popular and attractive with respect to the standard PMT is that they have a small size (in the range of mm^3), are completely insensitive to the magnetic field, compact and robust, and their low cost. In comparison to the standard APD (Avalanche Photo Diode) they offer higher intrinsic gain that simplifies the readout electronics and provide a more stable response to voltage or temperature changes. They also offer a fast response (~ 100 ps FWHM for a single photon) and capability to count single photons [87].

The concept of a highly granular calorimeter, like the one foreseen for the ILC, is feasible only due to the SiPMs.

A silicon photomultiplier consists of several *Single Photon Avalanche Diodes* (SPAD) connected in parallel.

A photodiode consists of a silicon p-n junction, which creates a depletion region, free of mobile charge carriers. When an incident photon is absorbed by the silicon, it creates an electron - hole pair. When a reverse bias voltage is applied, an electric field is created, accelerating the charge carriers towards the anode (holes) or cathode (electrons). Therefore, an absorbed photon corresponds to a current in the photodiode.

When a high electric field ($> 5 \times 10^5 \text{V/cm}$ [88]) is generated in the depletion region of the silicon, the charge carriers are accelerated up to the point where they have enough kinetic energy to create secondary charge electron - hole pairs, through the impact ionization process. In this way, a single incident photon generates a self-sustaining avalanche. At this point, the silicon becomes conductive and amplifies the initial electron - hole pair into a macroscopic charge flow. This process is called *Geiger discharge*. A photodiode operated in Geiger mode is called *Single Photon Avalanche Diode*.

The high electric field needed to operate the photodiode in Geiger mode can be achieved by applying a value of the bias voltage beyond the breakdown voltage. The breakdown voltage represents the minimum reverse bias voltage that makes the diode conduct. A sketch of the correlation between the current through the diode and the bias voltage applied is shown in Figure 3.1, where V_{BR} represents the breakdown voltage.

It is common to refer to the overvoltage (U_{OV}) at which the SPAD operates, defined as $U_{OV} = U_{Bias} - U_{BD}$, i.e. the difference between the bias voltage applied (U_{Bias}) and the nominal breakdown voltage (U_{BD}). Once the current is flowing, it has then to be stopped or *quenched*. This is realized using a series of resistors R_Q .

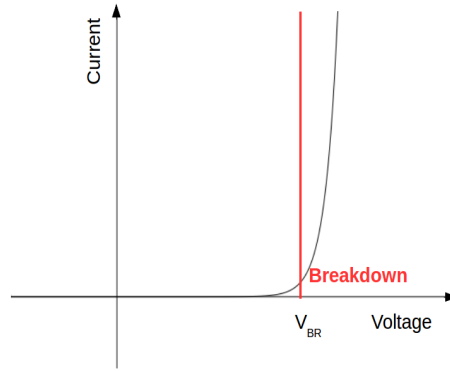


FIGURE 3.1.: Sketch of the correlation between the current through a diode and the voltage applied. V_{BR} represents the breakdown voltage, after which the diode conducts.

This lowers the voltage seen from the diode back to a value below the breakdown voltage, stopping the avalanche. The diode then recharges and is then again able to detect other photons. This happens with a time constant τ that can range between few 10 ns and few μs [87], and it is depending on the quenching resistor (R_Q) and the capacitance (C_{px}) of the based diode $\tau = R_Q \times C_{px}$. The capacitance of the diode depends on the size and geometry of the diode and on the doping profile of the substrate.

A SPAD operating in Geiger mode is therefore a binary device: independently to the number of photons absorbed, it will produce a signal equal to the signal corresponding to the absorption of a single photon. In order to recover this proportionality, several single photon avalanche diodes, each with its quenching resistor, are connected in parallel, forming a silicon photomultiplier. Each SPAD with its R_Q represents a micro cell or pixel of the SiPM. A silicon photomultiplier can have from 100 up to a few thousand pixels on a few mm^2 area, depending on the dimensions of the pixels themselves. An image of a SiPM surface is shown in Figure 3.2.

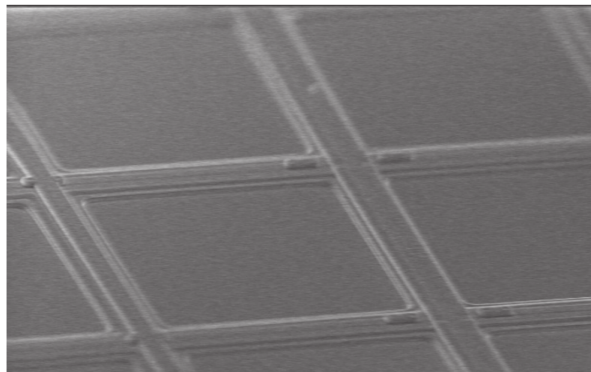


FIGURE 3.2.: Structure of each individual pixel on a silicon photomultiplier [88].

The charge Q expelled from the diode when firing depends on the overvoltage applied and the capacitance of the diode itself, as shown in Eq. 3.1

$$Q = U_{OV} \times C_{px}. \quad (3.1)$$

This corresponds to the gain of the silicon photomultiplier expressed in units of elementary charge. It follows, from Eq. 3.1, that the gain of the SiPM, depending on U_{OV} , depends on the temperature. U_{OV} depends on the temperature, depending on U_{BD} , which depends on the temperature because of the vibrational modes in the semiconductor lattice. In particular, the SiPMs show a negative dependence on the temperature, typically about $\frac{-1\%}{K}$, depending on the value of U_{OV} .

The probability that a single electron - hole pair initiates a Geiger discharge (*quantum efficiency*) increases with the overvoltage, therefore it shows also a negative dependence on the temperature, similar to the gain.

The signal of a silicon photomultiplier is the sum of the charge expelled from all the pixels fired. With uniform pixels capacitance and the amplification needed to resolve the single fired pixels, the gain can be extracted in - situ from the separation of the peaks in a *Single Photon Spectra (SPS)*. This is described in Section 3.5.1.

Because the number of pixels in a silicon photomultiplier is finite, and a recovery time τ is needed before a pixel can be fired again, the SiPM is a non linear device (*saturation effect*).

The response of the silicon photomultipliers is defined as the correlation between the number of pixels fired and the number of photo - electrons generated. This correlation can be approximated with the function shown in Eq. 3.2

$$N_{pix} = N_{tot} \cdot (1 - e^{-N_{pe}/N_{tot}}) \quad (3.2)$$

where N_{pix} indicates the observed number of pixels fired, N_{tot} the maximum number of pixels that can be fired and N_{pe} the effective number of photo-electrons generated.

Signals equivalent to 80% of the total number of pixels in a SiPM are underestimated by 50% with respect to the measurement of a linear device [87]. An example of the silicon photomultipliers response is shown in Figure 3.3 for MEPhI/PULSAR devices with various number of pixels. This effect limits the dynamic range. Nevertheless, it can be corrected offline, as described in Section 3.6.2 and in Section 4.2.2.

There are several effects that have to be taken into account when looking at the signal generated in a SiPM. These are the dark rate, the after pulse and the inter-pixel cross talk.

In absence of light, carriers thermally generated in the depletion region can generate a signal in the silicon photomultiplier. This *dark rate* depends on the

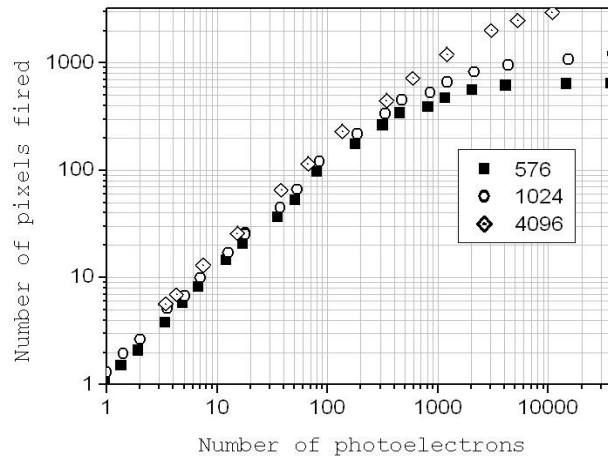


FIGURE 3.3.: *Non linear response of MEPhI/Pulsar SiPMs, made of different number of pixels, as indicated on the legend. The light signal is produced by a fast laser (40 ps) [89].*

temperature and for an increase of eight degrees it almost doubles [87]. In general, the dark rate of SiPMs can largely vary between $10^5 - 10^7$ Hz [87].

The after pulse happens when secondary avalanches follow the primary one, due to the release of carriers trapped by impurities of the silicon substrate. Since the second avalanche happens when the pixels are recharging, the output will consist of the sum of the signal corresponding to the first avalanche plus a second pulse which is lower and delayed with respect to the main one. The size of this effect depends on the recovery time of the pixels. The shorter the recovery time is, the larger the contribution from the after pulse is. The probability for after pulses depends on the structure of the silicon implants/doping in the SiPM.

Also the inter-pixel cross talk needs to be considered. During the avalanche process, photons in the visible range are emitted. If these photons fire a neighboring cell, a second avalanche can be generated.

The last generations of silicon photomultipliers have reached a value of dark rate noise of about 20 kHz at room temperature and a value of the inter-pixel cross talk less than 1% [90].

3.2. The SiPM Integrated Read-Out Chip 2

The silicon photomultiplier technology is applied to the hadron calorimeter prototype foreseen for the International Linear Collider.

A specific chip has been designed in order to be able to read out the signal from the silicon photomultipliers. The two main requirements that need to be fulfilled for such a calorimeter are the minimization of any space for infrastructure (the

calorimeter has to be as dense as possible and as homogeneous as possible) and therefore a lower power consumption.

The *SiPM Integrated Read-Out Chip 2 (SPIROC)* is an *Application Specific Integrated Circuit* (ASIC) in development by Omega [91].

The SPIROC chip is designed to be able to read out the output of 36 silicon photomultipliers and has the advantage of offering an already digitized output.

A schematic diagram of the signal path in the SPIROC2 is shown in Figure 3.4.

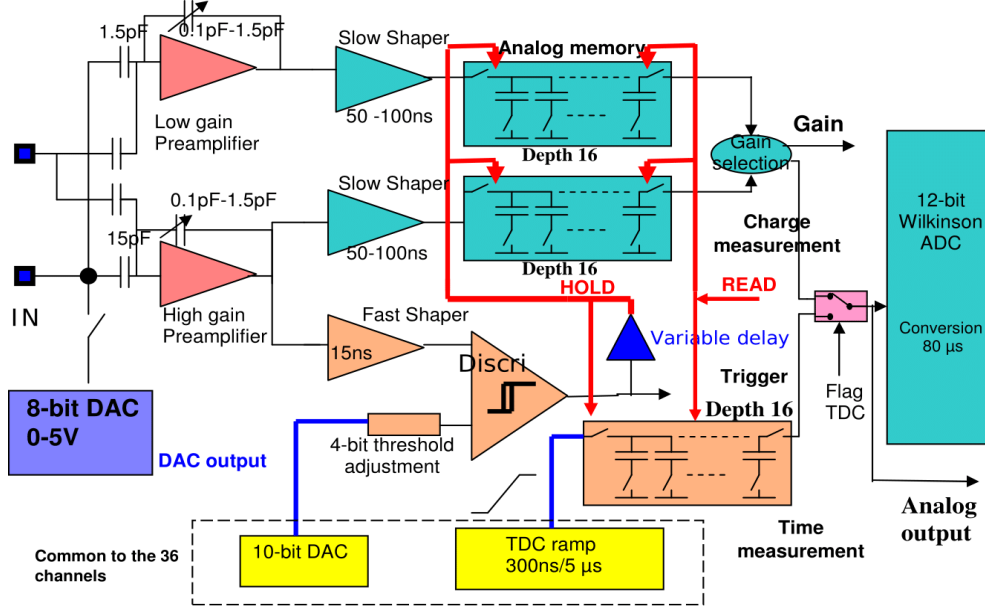


FIGURE 3.4.: Schematic of the signal path in the SPIROC chip for a single channel [92].

In order to have an optimum performance of the silicon photomultiplier, each has to be operated at its individual bias voltage, indicated by the SiPM manufacturer and verified with measurements. With the SPIROC2 it is possible to adjust the individual bias voltage for each SiPM with an 8 bit Digital-Analog Converter (DAC), connected directly to the chip input pin (IN, in Figure 3.4) through an RC filter to ground (not shown in the picture) with a resistor $R = 50 \Omega$ and capacitance $C = 100 \text{ nF}$ ([87]). The DAC can generate a voltage between $0 - 5 \text{ V}$ ([93]).

The chip can be operated in *External Trigger* mode (ET) or in *Auto Trigger* mode (AT). In auto trigger mode, a global threshold can be set for the whole chip. When the signal of one channel exceeds the threshold in a given cycle, all the 36 channels are stored. An external trigger can also be directly fed into the chip. In addition, an external validation trigger can be used in combination with the auto trigger mode to reduce the SiPM noise. In this way only events that are coincident with an external trigger are saved.

The signal is split into a *High Gain* (HG) path and a *Low Gain* (LG) path. The high gain mode is used for small signals and calibration up to 100 fired pixels,

corresponding to around 6 MIPs. The low gain mode is expected to have a fixed 1:10 ratio compared to the high gain mode, extending the dynamic range up into SiPM saturation.

The auto trigger signal is generated from the high gain path using a fast shaper (with a time constant of 15 ns). Only if the signal exceeds a given threshold, the high gain path and the low gain path are then stored in 16 analog buffers (*memory cells*) after a slow shaper. Once the 16 buffers are full, no further signal can be stored, therefore the signal is then digitized through a 12-bit *Wilkinson* ADC.

The measurement of the time in the SPIROC2 is performed by a *Time to Digital Converter* (TDC) implemented as a time to amplitude converter. A linear voltage ramp is generated and it is sampled and stored, similar to the amplitude, in 16 memory cells. When the signal exceeds a fixed trigger threshold, the corresponding value on the TDC voltage ramp is stored, providing the time measured for a certain AHCAL channel in TDC counts. At the beginning of each cycle, the ramp has to be reset, introducing in this way a dead time in the TDC. In order to circumvent this problem, two different linear ramps are used (*dual ramp TDC*), which are activated at alternating clock cycles. Each ramp has therefore the time to reset while the other one is ramping up. The TDC voltage ramp has a designed time resolution of 100 ps if operated in ILC-like conditions, corresponding to a ramp length of 200 ns. In test beam, the time resolution is ~ 1.9 ns, corresponding to a ramp length of 4 μ s.

In auto trigger mode all the three values (TDC, high gain and low gain) can be read out. A *Gain Bit* can be set, distinguishing between the high gain mode (Gain Bit = 1) and low gain mode (Gain Bit = 0). In external trigger mode this is not possible. The chip allows to read only two of the three values.

3.3. The AHCAL Technological Prototype

3.3.1. The Hadronic Calorimeter Base Unit

The *Hadronic calorimeter Base Unit* (HBU) is a $36 \times 36 \times 0.008$ cm³ printed circuit board (PCB). It contains four SPIROC2B ASICs, each of which can read out 36 silicon photomultipliers. Each SiPM is coupled to a $30 \times 30 \times 3$ mm³ plastic scintillator tile mounted on their bottom side. Therefore each tile is read out by one silicon photomultiplier and a HBU contains a total amount of 144 tiles. The total thickness of a fully assembled HBU is less than 5 mm. A picture of the top and bottom view of a HBU is shown in Figure 3.5.

In addition, the HBU hosts *Light Emitting Diodes* (LEDs), which can inject light pulses directly into the scintillator tiles in order to perform the gain calibration of the silicon photomultipliers, as described in Section 3.5.1. More details about the description of the system can be found in [93].

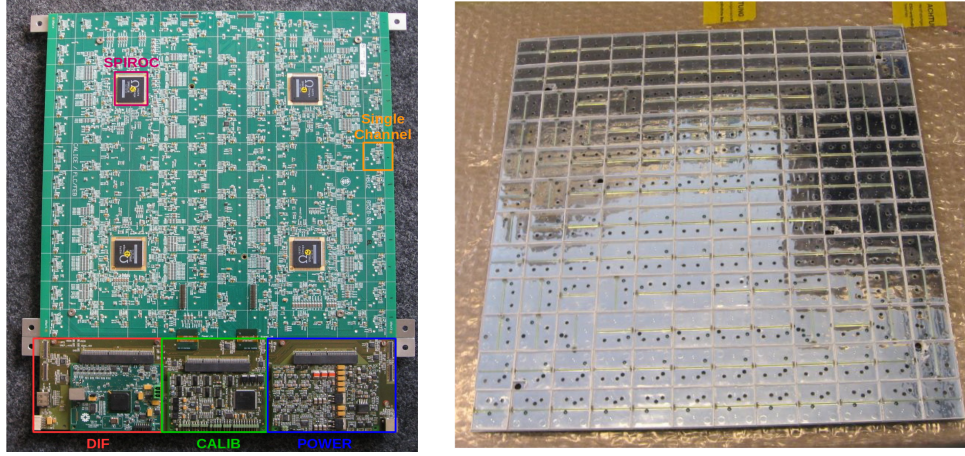


FIGURE 3.5.: The HCAL Base Unit. On the left, the top view with the 4 SPIROC2B chips is shown, on the right, the bottom view with the 144 scintillating tiles is shown.

Six HBUs can be connected and readout in series, forming a *slab*. Up to three slabs are connected in parallel to common DAQ interface boards forming an *ILD AHCAL layer*.

The DAQ interface boards consist of three modules, shown on the left in Figure 3.5, mounted on a *Central Interface Board* (CIB) and connected to the HBU via FlexLead connectors. These modules, starting from the left, are: the *Detector Interface* (DIF), the *CALIB Board* and the *Power Board*.

The DIF is responsible of controlling the ASICs (voltages, acquisition state) and sending the data to the *Link Data Aggregator* (LDA) with an internal clock of 250 kHz in test beam mode; the Power Board controls the power supply needed by the SiPMs, the SPIROC2B and the CALIB board; the CALIB board manages the LED calibration system, providing trigger pulses with voltages up to 10V.

3.3.2. The AHCAL Technological Prototype

The *AHCAL Technological Prototype* is a prototype of the sandwich hadronic calorimeter foreseen for the International Linear Collider. It was operated in beam at CERN in July 2015. The setup consists of 12 active layers in a steel absorber structure. The thickness of the passive material is 17 mm.

A picture of the detector is shown in Figure 3.6, while a sketch of the arrangement of the layers is shown in Figure 3.7. Here, the position occupied from the layers in the stack is shown, as well as the chip number assigned to the different chips in each HBU. On the bottom side of the picture, the module naming convention that will be used throughout this thesis is shown. The names express the place where they were assembled or the kind of SiPMs they contain.

The prototype was able to cover roughly 46.5 radiation lengths, but was fully equipped only for ~ 15 radiation lengths (until layer 11). Therefore only the

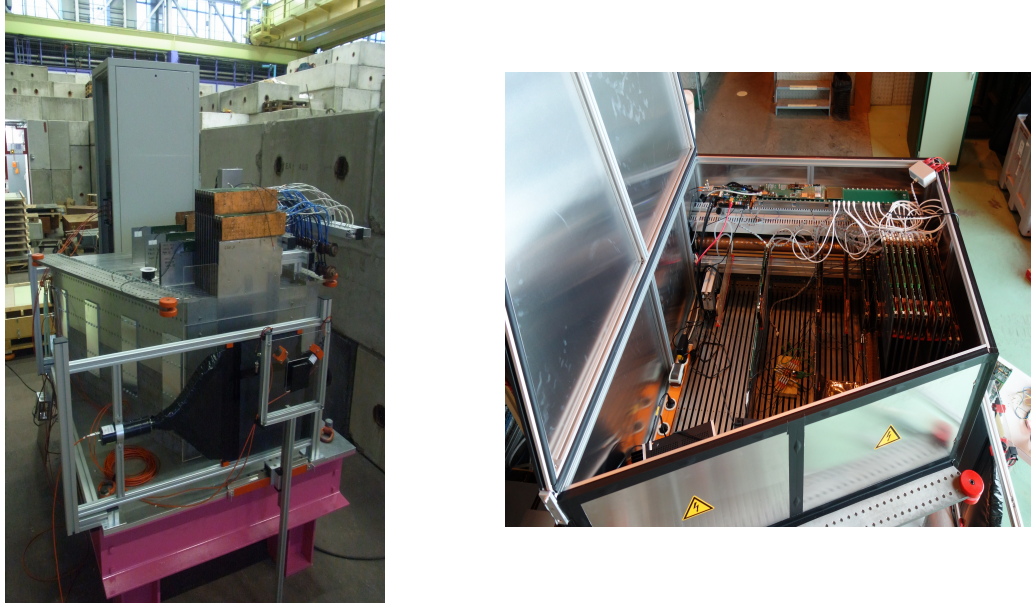


FIGURE 3.6.: Picture of the AHCAL Technological Prototype in a steel absorber structure at the CERN test beam. On the right a zoom of the layers arranged in the stack is shown.

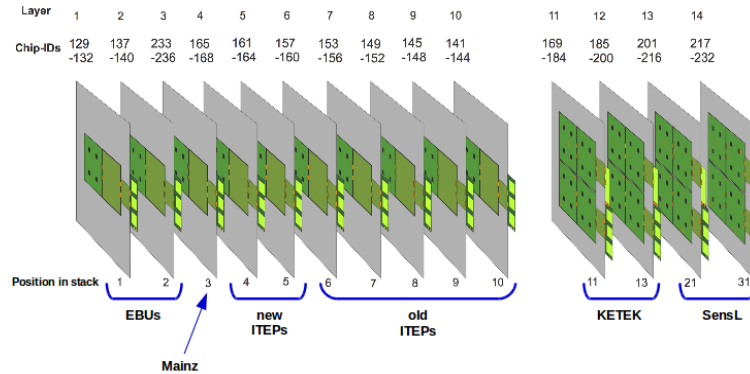


FIGURE 3.7.: Sketch of the layout of the AHCAL Technological Prototype.

electromagnetic shower is fully contained in the equipped part of the prototype, but not the hadronic shower.

The first two layers consist of *Electromagnetic Base Units* (EBUs). Each EBU consists of 144 scintillating strips on a surface of $18 \times 18 \text{ cm}^2$. Since these two layers were very noisy and was not possible to properly calibrate them, they are not considered in this thesis and more details about them can be found in [94].

The first eight layers (3 - 10) are single HBUs, while the remaining four consist of 2×2 HBUs. The single HBUs are meant to be used as a shower start finder (see Section 5.1). This means they are used to identify the first hadronic interaction

and therefore the layer where the hadronic shower starts. The last four big layers will be used to study the spatial and temporal evolution of the shower, that is one of core studies of this thesis and addressed in Chapter 5.

This AHCAL prototype is made of different kind of tiles and silicon photomultipliers. Some of the silicon photomultipliers used are sensitive only to the green light. For this reason the tiles are equipped with a wavelength shift fiber. Furthermore some of the tiles are not individually wrapped in reflector foil, therefore the optical cross talk has to be taken into account in the simulation (see Section 4.2.1).

The main properties of AHCAL layers forming the prototype are listed in Table 3.1.

Type	Layer	Position	Chns	SiPMs	Pixels	Wrapping
Mainz	3	3	144	MPPC	1600	✓
new ITEP	4	4	144	KETEK	12000	✓
new ITEP	5	5	144	KETEK	12000	✓
old ITEP	6	6	144	CPTA	800	✗
old ITEP	7	7	144	CPTA	800	✗
old ITEP	8	8	144	CPTA	800	✗
old ITEP	9	9	144	CPTA	800	✗
old ITEP	10	10	144	CPTA	800	✗
KETEK	11	11	576	KETEK	2300	✓
KETEK	12	13	576	KETEK	2300	✓
SensL	13	21	576	SensL	1300	✓
SensL	14	31	576	SensL	1300	✓

TABLE 3.1.: *Properties of the tiles and SiPMs of the AHCAL technological prototype.*

3.4. The CERN 2015 Test Beam Campaign

In July 2015, the AHCAL technological prototype was moved to CERN and tested with beam for two weeks. The main goals of this test beam campaign were:

- collect muon data, in order to perform the calibration of the detector;
- take electron data, to understand the behaviour of the detector and tune the MonteCarlo simulation (Chapter 4);
- collect pion data in order to study the hadronic showers (Chapter 5).

The test beam took place in Hall 2 (H2) in the North Area. The beam is provided from the *Super Proton Synchrotron* (SPS). It accelerates protons from an initial

energy of 25 GeV up to 450 GeV. Hadrons, electrons and muons in the energy range from 10 GeV to 360 GeV can be produced.

The extracted beam from the SPS is transported for about 1 km, through bending and focusing magnets. It is then split into three parts, each of them directed on a target where secondary particles are produced. There is a slow extraction of the beam from the SPS onto the test beam target, such that 5 s long spills with nearly continuous beam and several seconds without beam in between are provided. The H2 secondary line emerges from the T2 primary target, hosting several beryllium (Be) plates of different thickness. The target plate is chosen in order to optimize the yield of the requested secondary particles and type.

A further optimization can be achieved using several upstream dipole magnets that can modify the incident angle of the primary proton beam on the target and therefore the production angle of the secondary particles emitted into the H2 beam line.

The momentum selection of the particles is done in the vertical plane, as shown in Figure 3.8.

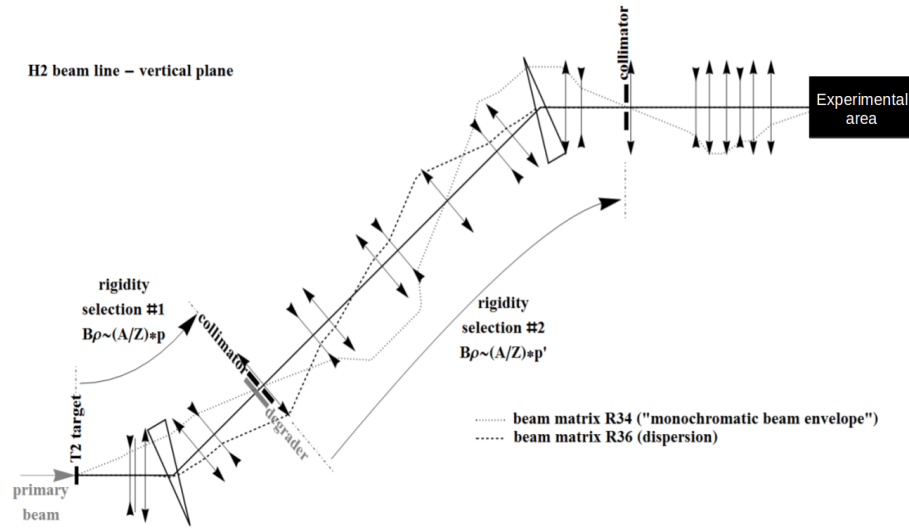


FIGURE 3.8.: Vertical section of the H2 beam line at CERN SPS. The picture is not to scale: the beam line is more than 600 m long and the height between T2 and the Hall is ~ 12 m [95].

The beam line consists essentially of two large spectrometers that select particles according to their rigidity R , defined as $R = B\rho \approx 3.33 p_b/Z$, where B is the magnetic field (in Tesla), ρ the Larmor radius (in meters) due to the magnetic field, p_b is the momentum of the beam particles (in GeV/c) and Z the charge of the particle (in proton charge unit).

In addition to the magnetic spectrometers the beam line is equipped with dedicated devices which provide information on the beam position, profile and intensity

at various locations, as well as particle identification detectors like a Cherenkov threshold counter [95].

As already mentioned, from the interaction of the proton beam onto a target, hadrons, electrons and muons are produced.

A wire chamber and scintillators are installed in the beam line, providing information on the position of the beam. Unfortunately the wire chamber was not connected to the AHCAL detector, therefore this information was not stored. Furthermore the knowledge of the exact upstream material is limited (as described in section 4.2, this is then tuned in order to perform a realistic simulation of the test beam setup).

A Cherenkov threshold counter detector is also installed in the beam line, in order to select the desired particles and reduce contamination.

The generated Cherenkov light is guided with a mirror to a photomultiplier and the amplified signal is discriminated with a fixed threshold and reads out in coincidence with a scintillator particle counter. The output signal is directly sent to 4 of the AHCAL channels, by charge injection into the SPIROC chips. The Cherenkov channels are listed in Table 3.2. The broken channels are not used in the analysis. The probability to detect a signal depends on the operating condition

Layer	Chip	Channel	Status
10	141	29	OK
12	195	12	broken
14	227	6	broken

TABLE 3.2.: *Channels of the AHCAL prototype receiving the Cherenkov signal (See also Figure 3.7).*

of the Cherenkov detector and on the particle type and momentum. Therefore the Cherenkov detector can be used to distinguish particles of different masses in a mixed beam.

Cherenkov light is emitted when a particle traverses an optical medium with a relativistic velocity $\beta = v/c$ higher than the speed of light in the medium c/n , where n is the refraction index of the medium [96].

The Cherenkov light is emitted in a cone with an opening angle θ with respect to the direction of motion of the particle given by:

$$\cos(\theta) = 1/\beta n \approx 1 - \theta^2/2.$$

The relation $\beta^2 = 1 - m^2/p^2$ correlates the opening angle θ with the momentum and the mass of the considered particle [97]. Considering particles with the same momentum p , only particles below a threshold mass m_{th} , defined in Eq. 3.3, emit Cherenkov light

$$m_{th} = \frac{p}{c} \sqrt{n^2 - 1} \quad (3.3)$$

with an angle [98]

$$\theta_{cone} \propto \Delta M = m_{th} - m.$$

Due to the dependence of n on the gas pressure, both θ_{cone} and m_{th} are directly dependent on the gas pressure. The Cherenkov detector operated at CERN was filled with Helium gas, and the pressure could be chosen, performing pressure scan in order to select the desired particles.

As already mentioned earlier, four trigger scintillators were also installed in the beam line. They are plastic scintillator paddles coupled to a PMT.

For the muon beam, the coincidence between two scintillator plates of $50 \times 50 \text{ cm}^2$ placed in front of and behind the prototype is required. For electron and pion beams, the coincidence between two small scintillator plates of $10 \times 10 \text{ cm}^2$ placed in front of the prototype is required.

The trigger signal was directly injected in six AHCAL channels (T_0 channels), in order to have a time reference. A list of these channels can be found in Table 3.3. The broken channels are not used in the analysis.

Layer	Chip	Channel	Status
11	169	29	noisy
11	177	23	broken
12	185	29	OK
13	201	29	OK
13	211	6	broken
14	217	23	OK

TABLE 3.3.: Channels of the AHCAL prototype receiving the trigger scintillator signal (See also Figure 3.7).

3.5. Calibration

The analog output signal of the silicon photomultiplier is digitized by the SPIROC2B ASICs and expressed in *Analog to Digital Converter* (ADC) units. Therefore, the raw amplitude has to be calibrated into a common physics signal scale in order to be used in offline analysis.

Within the CALICE collaboration, the reference physics unit is the *Minimum Ionizing Particle* (MIP), meaning the amplitude of the signal measured from each channel of the calorimeter has to be converted into MIPs unit. A MIP is defined as the signal that a minimum ionizing particle leaves in the detector when crossing a tile perpendicular to the large tile surface. The calibrated energy $A[MIP]$ (or $E[MIP]$) in MIPs unit of each channel i of the AHCAL prototype is extracted

according to Eq. 3.4,

$$E[MIP]_i = A[MIP]_i = \frac{(A[ADC]_i - P[ADC]_i) \cdot IC_i}{C^{MIP}[ADC]_i} \cdot f_i^{-1}(A[pixels]_i) \quad (3.4)$$

where $A[ADC]_i$ represents the raw energy of the channel i in ADC counts, $P[ADC]_i$ the pedestal value (see Section 3.5.2), IC_i the intercalibration value due to the dual gain mode of the SPIROCs (see Section 3.5.3), $C^{MIP}[ADC]_i$ the calibration constant extracted in order to obtain the energy in MIP. It is defined as the most probable value of the distribution of the energy deposition of a MIP (see Section 3.5.4). $A[pixels]_i$ represents the signal amplitude expressed in number of SiPM's pixels fired and $f_i(A[pixels])$ is the function used to correct for the non linear response of the silicon photomultipliers, due to the finite number of pixels of which the SiPMs are made of.

The saturation correction is applied to the amplitude expressed in number of pixels fired. The amplitude in terms of pixels fired is obtained dividing the raw amplitude (in ADC counts) by the gain of the SiPMs (see Section 3.5.1), as shown in Eq. 3.5

$$A[pixels]_i = \frac{(A[ADC]_i - P[ADC]_i)}{Gain[ADC]_i}. \quad (3.5)$$

The calibration procedures have been the core of several studies within the CALICE collaboration [93, 99, 100]. In the following sections, for completeness, these procedures are described.

The major contribution of this thesis to the energy calibration of the detector is associated to:

- performing the gain calibration (3.5.1) for the CERN 2015 test beam campaign;
- checking validation, robustness and portability of the MIP calibration (Section 3.5.4);
- performing the high gain - low gain intercalibration, for which a new procedure had to be developed for this thesis, since up to now a default value has been used (Section 3.5.3).

If one of the mentioned calibration constants couldn't be extracted for one of the AHCAL channel, this specific channel is defined as *dead* and excluded from the analysis. The calibration constants and the dead channels are stored in a database. In Appendix C the CALICE database tags used in this thesis are provided.

3.5.1. Gain Calibration

The saturation behaviour of the silicon photomultipliers appears at the pixel scale. Therefore it is necessary to convert the amplitude of the signal, in ADC counts, into pixels fired. This is done through the gain of the SiPMs, as shown in Eq. 3.5.

As mentioned in Section 3.3.1, each HBU hosts a LED system (one per HCAL channel), that can inject low intensity light directly into the channels. For each channel, the *Single Photon Spectrum* (SPS) is considered and fitted with a multi-Gaussian function. An example of this kind of spectrum is shown in Figure 3.9.

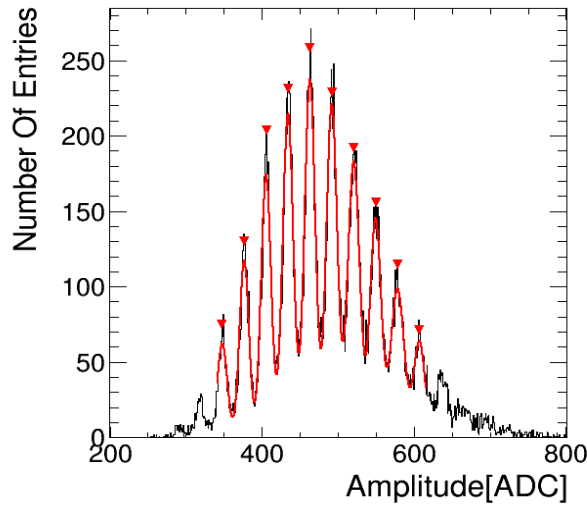


FIGURE 3.9.: *Single Photon Spectrum* acquired with a SensL silicon photomultiplier readout with the SPIROC2B ASIC (layer 14, see Table 3.1). Each peak corresponds to one more photon detected. The spectrum is fitted with a multi-Gaussian function. The mean distance between two consecutive peaks is defined as the gain of the silicon photomultiplier in ADC unit.

The gain value is obtained as the difference between the means of two consecutive single photon peaks. It is possible to notice that the width of the fitted peaks doesn't visibly increase for higher numbers of fired pixels, indicating a good pixel capacity uniformity of the sensor.

The gain calibration is not only important because it gives the scale needed to express the amplitude read out from the SPIROCs in ADC to number of pixels fired, it is also a powerful tool to monitor the performance of the AHCAL prototype during the test beam.

During the test beam campaign, LED data have been taken almost every day. These data were immediately analyzed offline and a feedback on the gain calibration was provided, in order to guarantee a correct operation of the detector. In Figure 3.10, on the left the gain distribution for all the channels of layer 13 is shown for one specific LED data set, while on the right the mean of this distribution for each day of test beam is shown, monitored in order to guarantee the

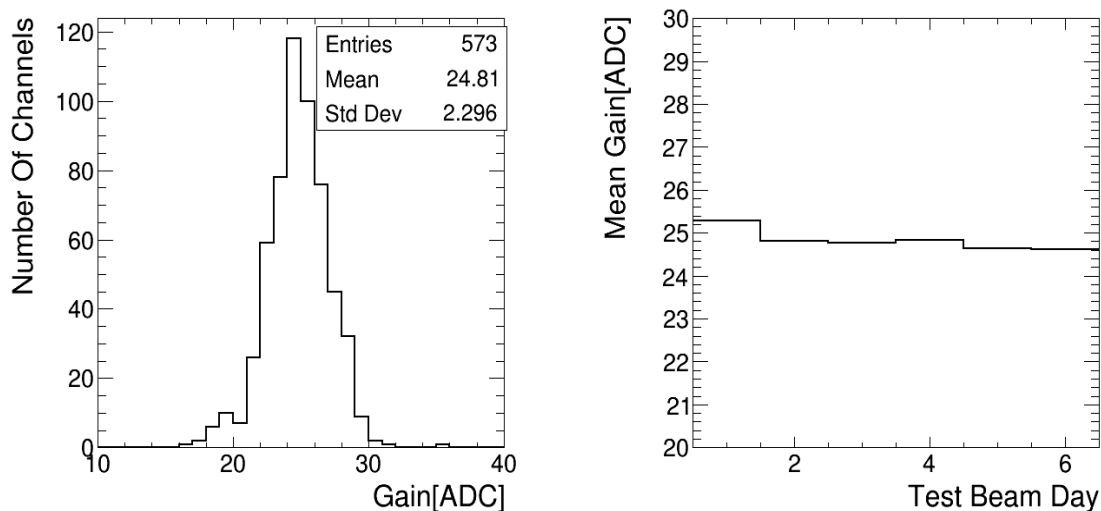


FIGURE 3.10.: Gain values obtained for SensL silicon photomultiplier readout with the SPIROC2B (layer 13, Table 3.1). On the left the distribution of the gain values for all the channels for a specific LED data set is shown. On the right, the mean of this distribution for each day of the test beam campaign is shown. No temperature compensation is applied.

stability of the gain and therefore a proper operation mode of the detector. Here it is possible to see the stability of the gain during the test beam period, indicating a good stability of the prototype. It is important to notice that here no temperature compensation is applied.

3.5.2. Pedestal Extraction

As explained in Section 3.2, the data acquisition from the SPIROC is performed in auto trigger mode: when the signal of one of the 36 channels exceeds a certain trigger threshold, set for the whole chip, the signal of that channel is stored as signal above threshold (a *HitBit* equal to 1 is set). The signal is stored also for the other channels of the same chip (*OR36 logic*), and if the signal doesn't exceed the threshold, this is stored as signal below threshold (*HitBit* equal to 0).

In order to properly perform the energy calibration of the detector, the pedestal value has to be subtracted. This is extracted using muon runs, considering only the channels for which the *HitBit* equal to 0 is set.

For each channel and memory cell, a histogram with the ADC values is filled. The mean of the distribution is then taken. In order to avoid the noise of the silicon photomultiplier (assuming a Poisson distribution for it), with an iterative procedure, the range is restricted to 3RMS around the mean.

Since in the reconstruction process, the pedestal subtraction is implemented per channel, and not per memory cell, the final pedestal for each channel is obtained

computing the mean of the pedestals obtained for that channel for each memory cell.

In Figure 3.11 the ADC spectrum for one AHCAL channel, and the corresponding pedestal are shown.

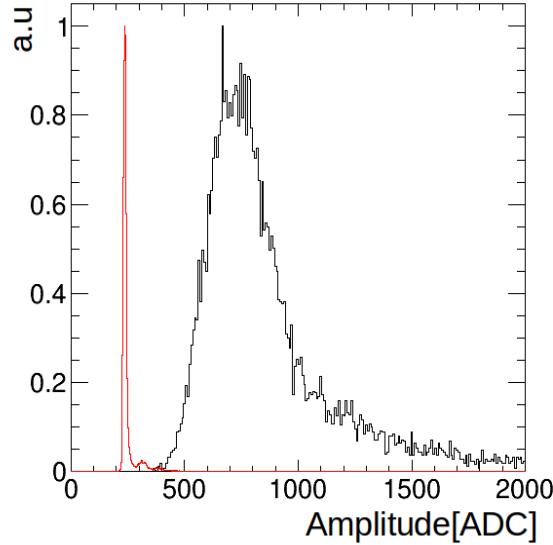


FIGURE 3.11.: *Signal amplitude and corresponding pedestal (in red) for one channel of the AHCAL prototype readout with MPPC SiPM (layer3, see Table 3.1).*

3.5.3. High Gain - Low Gain Intercalibration

As explained in Section 3.2, the SPIROC chips used to read out the output signal of the silicon photomultipliers have a dual gain operation mode. The high gain mode is used for small signals and calibration up to ~ 100 fired pixels, corresponding to roughly 6 MIPs. The low gain mode is expected to have a fixed ratio $\sim 1 : 10$ compared to the high gain mode, extending the dynamic range up into the SiPMs saturation region. This ratio is not exactly 10. Small variations of the capacitances in the preamplifier path of the SPIROC chips generate variations of this ratio. Therefore it has to be calibrated for each AHCAL channel.

In order to perform the high gain - low gain intercalibration, usually special LED runs, in high gain - low gain mode are taken. This means that in these runs, both the high gain measurement and the low gain measurement are stored. The high gain mode is stored instead of the TDC information.

Since during the CERN test beam campaign only few runs in intercalibration mode have been taken, these data couldn't be used, because the statistics were not enough in order to perform the calibration. Therefore the LED data from another test beam campaign, that took place at DESY in August 2016, are used. In this

test beam campaign only the layers 3, 13 and 14 were used (in combination with other new layers, not used during the CERN test beam). For the other layers, a different method, using electron data, has been used [101]. In this section, the two methods are described.

Intercalibration using LED data

In order to perform the high gain - low gain intercalibration for layer 3, layer 13 and layer 14 (see Figure 3.7), the LED data taken in external trigger mode and in intercalibration mode during the DESY 2016 test beam campaign are used. The LED scan has been performed on the 01.08.2016 and the calibration voltage applied is from 3500 mV to 8000 mV in steps of 50 mV. The calibration voltage defined the amplitude of the emitted LED light.

The method used is based on the following steps:

1. for each calibration voltage and each AHCAL channel, a histogram with the amplitude values in ADC counts separately for the high gain mode and the low gain mode is filled. The pedestal value, memory cell wise, is subtracted before filling the histograms. An example of these distributions is shown in Figure 3.12(a) for the high gain mode and in Figure 3.12(b) for the low gain mode.
2. The mean of the histograms in Figure 3.12 is considered for each AHCAL channel and for each calibration voltage. Therefore for each channel the mean amplitude value in high gain mode ($\langle (A[ADC] - P[ADC])_{HG} \rangle$) and in low gain mode ($\langle (A[ADC] - P[ADC])_{LG} \rangle$) is defined.
3. For each channel, the correlation between the quantities in step 2 is studied, as shown in Figure 3.13. Each point corresponds to one calibration voltage.
4. In the linear region, a linear fit is performed.

The point 3 and the relative fit explained in 4 are shown in Figure 3.13. The slope (p_1) represents the intercalibration constant. As it is possible to notice, the value of p_1 is considerably larger than the expected value equal to 10, as described in Section 3.2. The procedure explained above is done for all channels. The linear fit has been performed choosing an arbitrary range for $(A[ADC] - P[ADC])_{LG}$ between 40 and 200. The lower value has been chosen to avoid any non-linearity, that could compromise the outcome of the fit. The upper value has been chosen in order to avoid the saturation region of the low gain mode. As it is possible to see from Figure 3.13 the value of the χ^2 is large, due to the fact that the

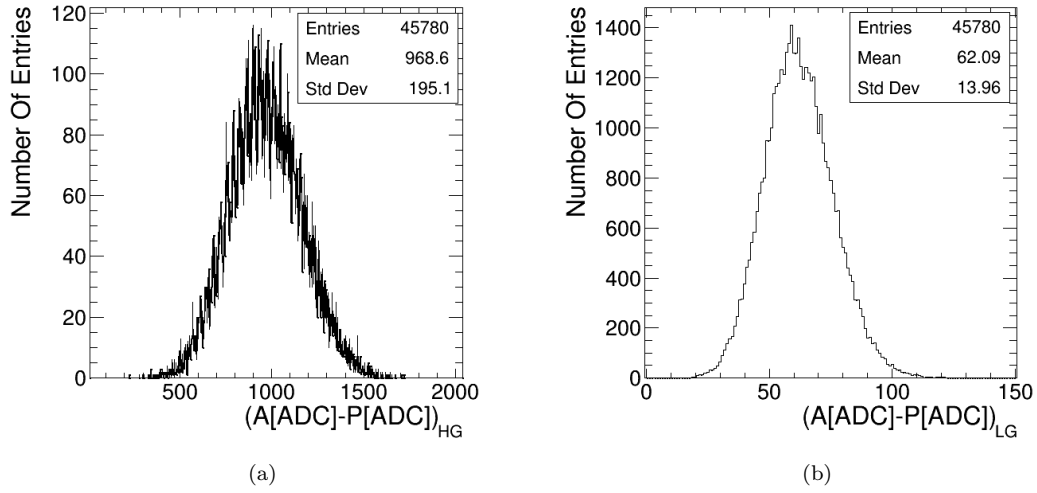


FIGURE 3.12.: Example of the amplitude distribution, pedestal subtracted, for the high gain and low gain mode for one of the AHCAL channel read out with the SensL SiPM read out with the SPIROC2B (channel 15 on chip 213 of layer 13, see Figure 3.7). **(a)**: high gain mode. **(b)**: low gain mode.

statistical errors on the points are small, that the range is arbitrary and due to the assumption that the chips behave linearly. Therefore, the χ^2 couldn't be used to check the result of the fit. The distribution of the slopes has been considered, the outliers have been checked and the range has been adjusted by hand.

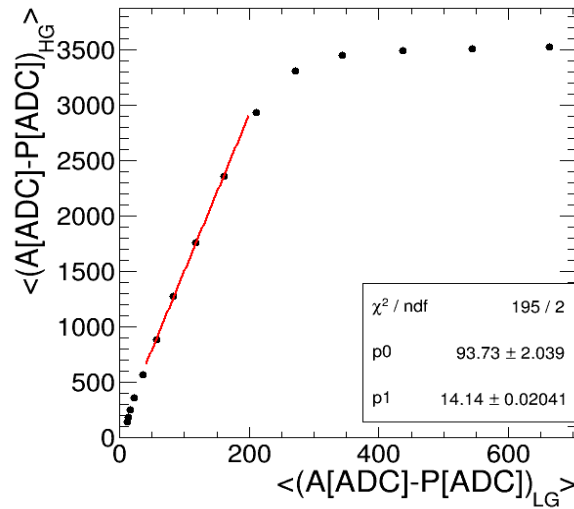


FIGURE 3.13.: High gain - low gain intercalibration extraction for one channel of the AHCAL read out with the SensL silicon photomultiplier read out with the SPIROC2B ASIC (channel 15 on chip 213 of layer 13, see Table 3.1). The mean of the amplitude in ADC counts in high gain mode against the mean of the low gain with a linear fit in the linear region is shown. Each point corresponds to one calibration voltage.

In order to validate the performance of the method applied, the parameter p_1 and p_0 extracted from the fit are used to correct the amplitude value in low gain mode, as expressed in Eq. 3.6

$$A[ADC]_{LG}^{Corr} = (A[ADC] - P[ADC])_{HG} = p_1 \cdot (A[ADC] - P[ADC])_{LG} + p_0. \quad (3.6)$$

The offset p_0 is expected to be zero, since the pedestal has been subtracted. Therefore, in a first step, the constant term has been ignored and a correction to the intercalibration constant extracted in this way will be applied in a second step. The difference between the amplitude value in high gain mode and the $A[ADC]_{LG}^{Corr}$ obtained from Eq. 3.6 (without p_0) is plotted against the amplitude value in low gain mode, not corrected. This is shown in Figure 3.14.

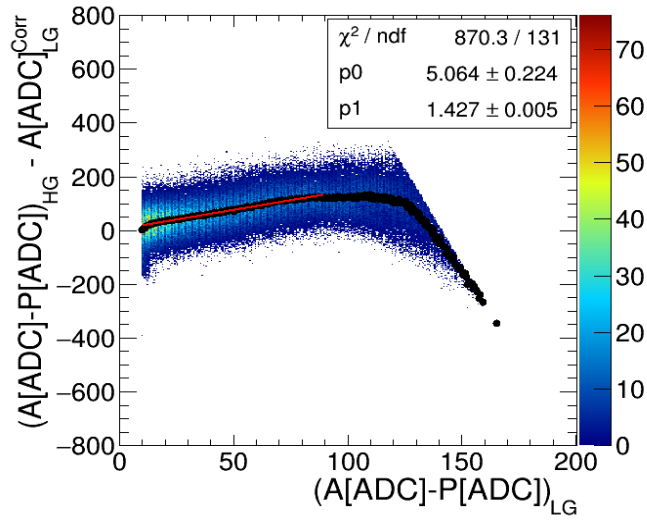


FIGURE 3.14.: Difference between the amplitude values in high gain mode $((A[ADC] - P[ADC])_{HG})$ and the corrected amplitude values in low gain mode $(A[ADC]_{LG}^{Corr})$ versus the non corrected amplitude values in low gain mode $((A[ADC] - P[ADC])_{LG})$, for one channel of the AHCAL, read out with the SensL silicon photomultiplier readout with the SPIROC2B ASIC (layer 13, see Table 3.1). Black points represent the mean value of the difference, while the red line the linear fit performed to extract the correction for the intercalibration constant.

In principle, the difference on the y axis on Figure 3.14 is expected to be 0. The deviation from 0 is taken as a correction needed to be applied to the intercalibration constant. In order to obtain this correction, the mean value of this difference is plotted (black points on Figure 3.14) and a linear fit, in the linear region, is performed.

The final high gain - low gain intercalibration constant (IC) is therefore given by the sum of the intercalibration constant extracted (p_1 in Eq. 3.6), plus the correction extracted, ΔIC (p_1 in Figure 3.14), as described in Eq. 3.7

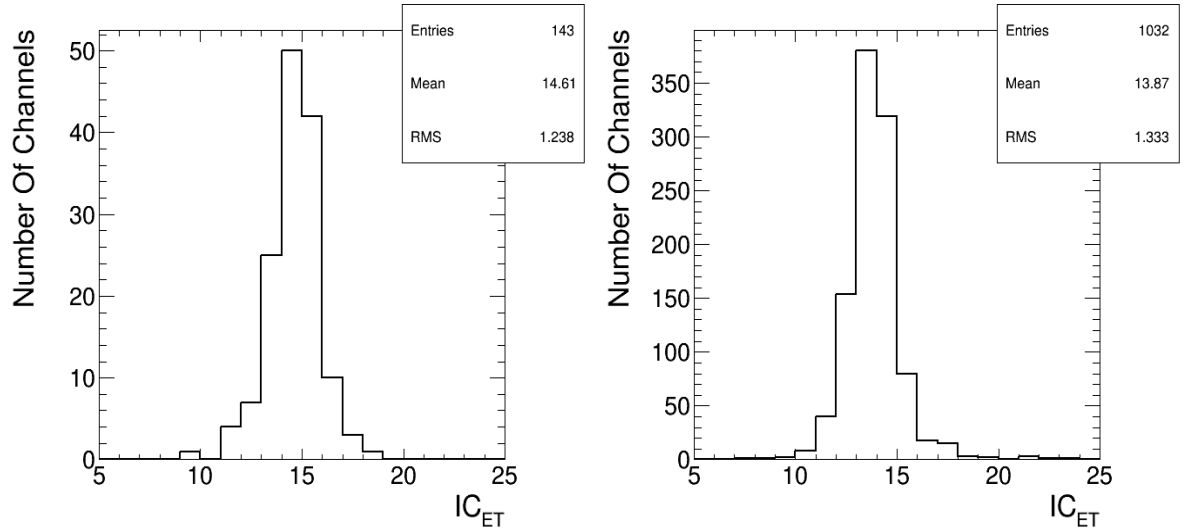


FIGURE 3.15.: Distribution of the intercalibration constants obtained for layer 3, on the left and for layer 13 and 14 on the right (see also Figure 3.7).

$$IC = p_1 + \Delta IC. \quad (3.7)$$

From Figure 3.14 it is possible to notice that the correction to the intercalibration constant is a remarkable effect, at the level of $\sim 10\%$, and therefore not negligible.

The distribution of the intercalibration constants obtained with the procedure described above is shown in Figure 3.15, for layer 3 on the left, and for layers 13 and 14 on the right. It is possible to notice that the mean of the two distributions is slightly different, and this is because the operation settings for these layers were not the same. In the distribution of layer 3 (left in Figure 3.15) it is possible to see also few outliers. The corresponding plots in Figure 3.13 and in Figure 3.14 for these channels have been checked and it has been verified that these values are correct.

From the distribution on layers 13 and 14 it can be seen that 10% of the channels are missing (the total number of channels is equal to 1152). For these channels, the mean value of the intercalibration constants of the chip to which the channel belongs to is taken. For layer 3 only one channel is missing.

During the same test beam campaign, LED data have been collected also in auto trigger mode. Therefore the same procedure has been applied using the LED data taken in auto trigger mode and intercalibration mode on the 03.08.2016 and 04.08.2016. Looking at the correlation between the signal in high mode and the signal in low gain mode, a strange behaviour in the LED scan performed in auto trigger mode has been observed. An example for one of the AHCAL channel of layer 13 (see Figure 3.7) is shown in Figure 3.16(a). For comparison, in Figure 3.16(b), the same plot obtained using the LED data taken in external trigger mode

is shown. A possible explanation of this problem is the pedestal shift effect. The pedestal shift is an already known problem of the SPIROC2B [93]. Depending on the total charge injected into the inputs of the chip, the global pedestal level is shifted. In the LED scan, the LEDs don't start to radiate all at the same time. When some channels get the first small signal other channels can already have a huge signal, causing therefore a large pedestal shift.

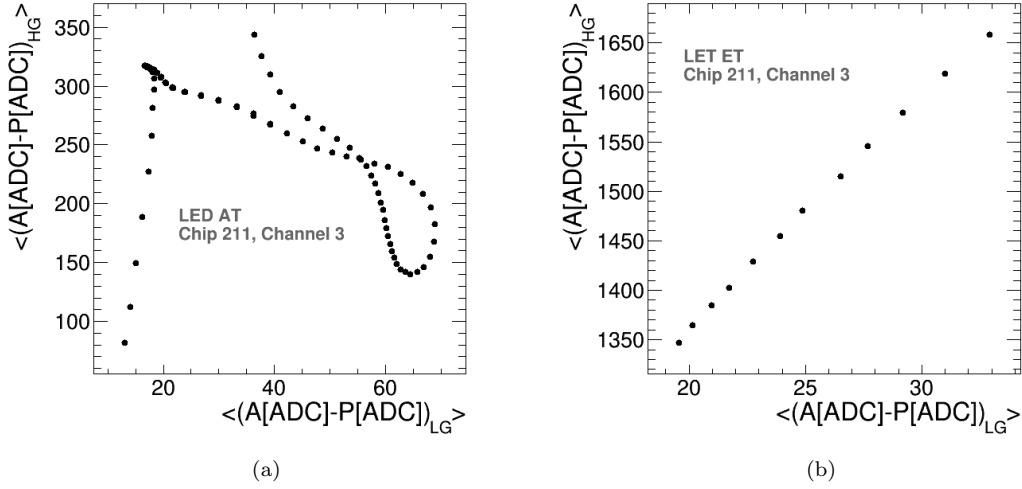


FIGURE 3.16.: Correlation between $(A[ADC] - P[ADC])_{HG}$ and $(A[ADC] - P[ADC])_{LG}$ for the extraction of the high gain - low gain intercalibration using LED data in external trigger and auto trigger mode. The channel 3 on chip 211 has been taken as an example (see Figure 3.7). **(a)**: LED data in auto trigger mode (AT). **(b)**: LED data in external trigger (ET) mode (see Figure 3.7).

Since the problem is not fully understood and it was not possible to further investigate it, the LED data taken in external trigger mode have been used in order to perform the high gain - low gain intercalibration. In addition, a situation in which almost all the channels have a large amount of light and only few channels a small signal is very unlikely while taking "real" data.

Intercalibration using electrons data

For the layers not used during the DESY 2016 test beam campaign, no LED data in intercalibration mode are available, therefore another method has been developed and used within the CALICE collaboration [101]. In this method electron data are used.

Looking at the amplitude spectrum of one AHCAL channel, an incorrect high gain - low gain intercalibration constant is reflected in a step in this spectrum, as shown in the sketch in Figure 3.17.

This method consists on fitting the edge of the two distributions with two complementary functions in such a way that the two edges coincide. The two

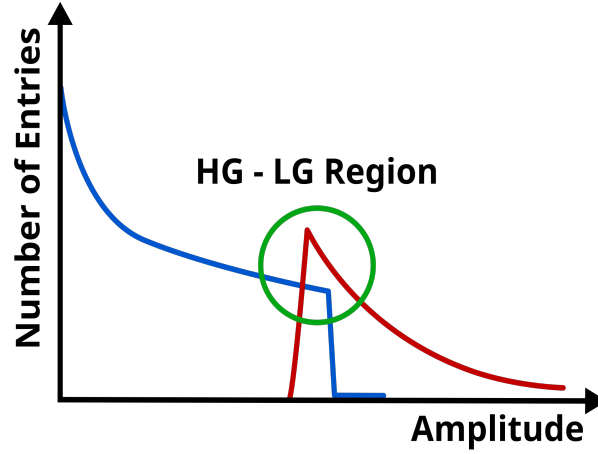


FIGURE 3.17.: Sketch to explain the effect of a non proper high gain low gain intercalibration. In blue the high gain signal is shown, in red the low gain. The green region highlights the transition between the high gain and low gain mode.

functions are defined in Eq.3.8, where $Erfc$ is the error function, and the subscripts are used only to indicate which function is applied to the high gain spectrum and which one to the low gain.

$$\begin{aligned} y_{HG} &= (p_0)_{HG} \cdot (Erfc(x_{HG} - (p_1)_{HG}) / ((p_2)_{HG} \sqrt{2})) \cdot exp((p_3)_{HG} \cdot x_{HG}) \\ y_{LG} &= (p_0)_{LG} \cdot (2 - Erfc(x_{LG} - (p_1)_{LG}) / ((p_2)_{LG} \sqrt{2})) \cdot exp((p_3)_{LG} \cdot x_{LG}). \end{aligned} \quad (3.8)$$

The intercalibration constant is extracted from the ratio $IC = (p_1)_{HG} / (p_1)_{LG}$.

An example of this method applied to one of the AHCAL channels is shown in Figure 3.18. Here on the left, the high gain spectrum is shown, with the relative fit function and a zoom in the region of interest. On the right, the low gain spectrum for the same channel, with the associated fit function is shown. In Figure 3.19 the same channel, after the right intercalibration constant is applied, is shown.

In order to extract the intercalibration constant, this procedure has been applied for different beam energies. For the channels for which the intercalibration could be extracted for all the energies, the average has been taken. This has been done since the intercalibration constants don't depend on the energy.

In this way, only the central channels close to the center of the detector, where the beam was shot, can be calibrated. For all the other channels, the mean value of the intercalibration constants of the layer to which the non calibrated channel belongs to is taken.

This represents the biggest disadvantage of this method, reason for which it has not been applied to the other layers too, since it has been verified that a high gain - low gain intercalibration per channel is needed.

The distribution of the intercalibration constants obtained combining the two

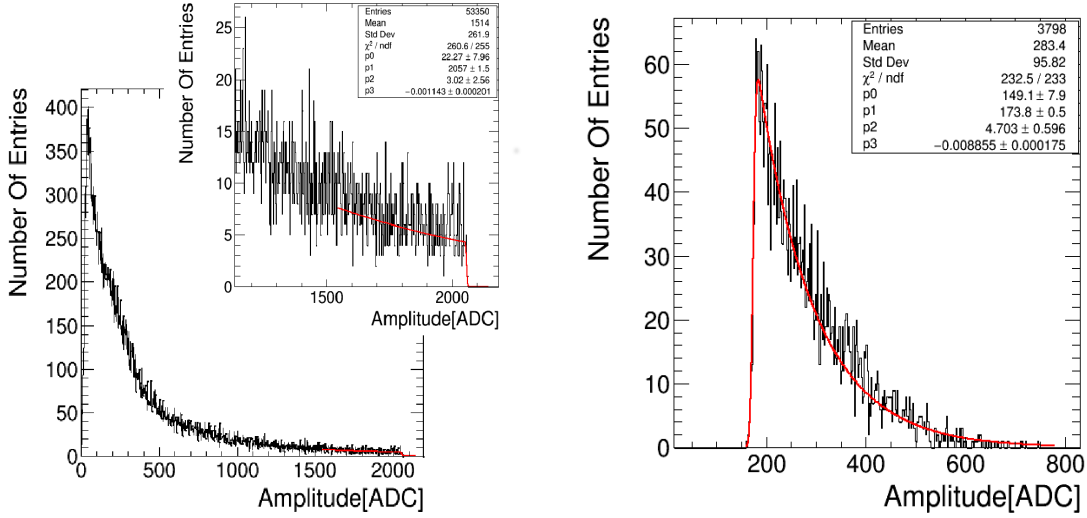


FIGURE 3.18.: *Extraction of the high gain - low gain intercalibration constant for one channel of the AHCAL prototype, using KETEK SiPMs, read out with the SPIROC ASICs (layer 4, see also Figure 3.7). On the left, the HG spectrum with the corresponding fit function and a zoom of the region of interest is shown. On the right, the LG spectrum with the corresponding fit function is shown. Here 10 GeV electron data have been used.*

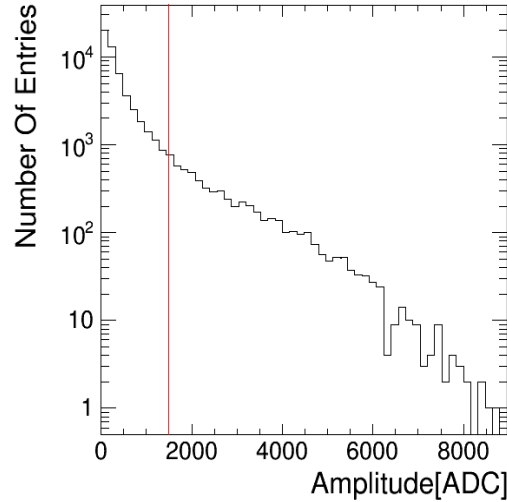


FIGURE 3.19.: *Signal amplitude of the same channel as in Figure 3.18 in high gain mode, after applying the intercalibration constant. The red line indicate the transition region between high gain and low gain mode.*

methods for the 3456 channels of the prototype is shown in Figure 3.20(a). The peak at ~ 15 corresponds to the average value used for layers 11 and 12. For these two layers, neither the method using LED data nor the method using electron data could be used, because not enough statistics were available. In Figure 3.20(b), the ratio between the intercalibration constants obtained with the LED data and the intercalibration constants obtained with the electron data is shown, considering only the channels for which the intercalibration constant could be extracted with

both methods. The two methods agree within $\sim 10\%$. The uncertainties due to the method used to extract the intercalibration constants is considered as a systematic uncertainty, as explained in Section 4.3.2.3.

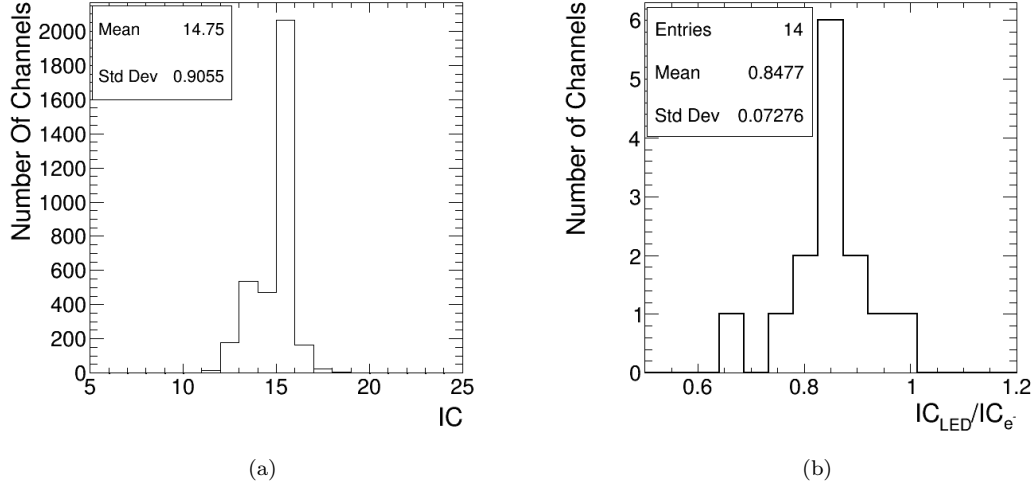


FIGURE 3.20.: (a): Distribution of the intercalibration constants for the 3456 AHCAL channels.

(b): Comparison between the intercalibration constants obtained with LED data and intercalibration constants obtained with electron data.

3.5.4. MIP Calibration

A MIP is defined as the energy deposited in each single AHCAL channel from a Minimum Ionizing Particle. In first approximation, muons are considered minimum ionizing particles, therefore the energy deposited from a muon in one AHCAL channel is 1 MIP.

In order to convert the raw amplitude signal from ADC counts, to MIPs, muon data of 50 GeV and 150 GeV energy have been taken during the test beam campaign.

To reject the noise an offline cut at 0.5 MIPs is applied on the hit energies.

Since the beam was not a pure muon beam, a simple selection has to be performed. A muon track finder algorithm has been developed [76, 99]. The aim of the algorithm is to identify a track in the detector, since muons are expected to travel straight through the detector. In order to do that, the number of hits in the same x and y position is counted. The minimum number of hits needed to identify a track is set to 7. In addition to that, in order to reject pions starting to shower later, it is required to have only one hit per layer. If more than one hit per layer is counted, the event is rejected. This is shown in Figure 3.21.

In order to see the energy deposited in a tile from a MIP like particle, the amplitude spectra of each AHCAL tile in muon events are considered and fitted

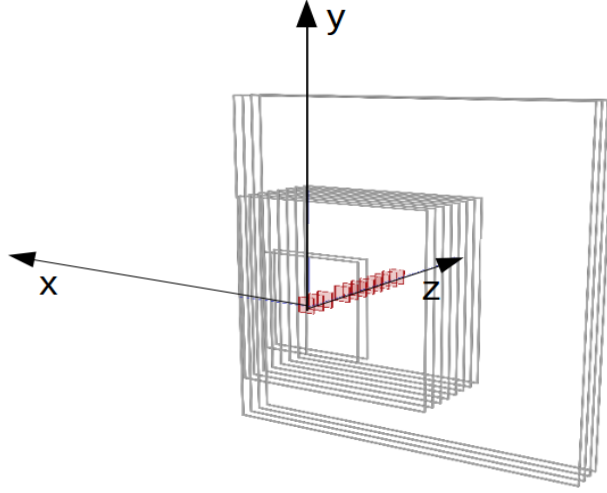


FIGURE 3.21.: Picture of a muon traversing the AHCAL prototype and identified with the track finder algorithm explained above.

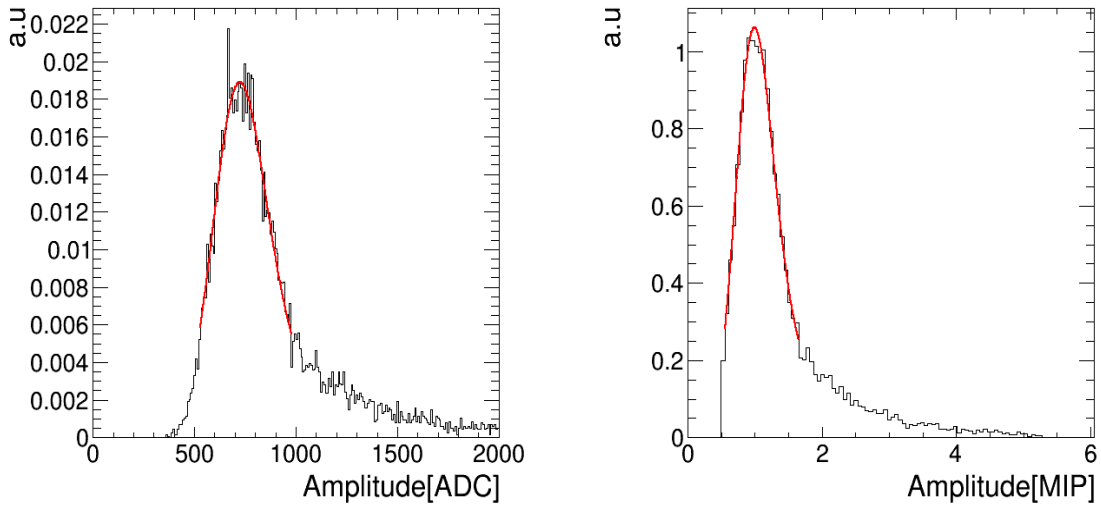


FIGURE 3.22.: Amplitude Spectrum of one AHCAL channel, readout with MPPC SiPM (layer 3, Table 3.1). On the left, the spectrum in ADC count and before pedestal subtraction is shown, together with the fitting function resulting from the convolution of a Gaussian and Landau function. On the right, the same spectrum in MIP is shown. As expected for a well calibrated detector, after the calibration the spectrum is peaking at 1 MIP.

with a function that is a convolution of a Gaussian and Landau distribution. An example of such a spectrum is shown on the left plot in Figure 3.22.

The MIP calibration constant is then defined as the difference between the *Most Probable Value* (MPV) of the fitting function subtracted from the pedestal value, as expressed in Eq. 3.9.

$$C^{MIP}[ADC]_i = MPV[ADC]_i - P[ADC]_i \quad (3.9)$$

In order to check the consistency of the procedure, the set of the obtained calibration constants $C^{MIP}[ADC]_i$ has been applied to the same data set. For a well calibrated detector, a peak at 1 MIP is expected. This is shown for one AHCAL channel on the right in Figure 3.22. The distribution of the most probable value per channel, after applying the calibration constants $C^{MIP}[ADC]_i$, is shown in Figure 3.23. The expected peak at 1 MIP can be seen. The sigma of the distribution, at the level of $\sim 1\%$, gives an estimation of the uncertainty on the fit procedure.

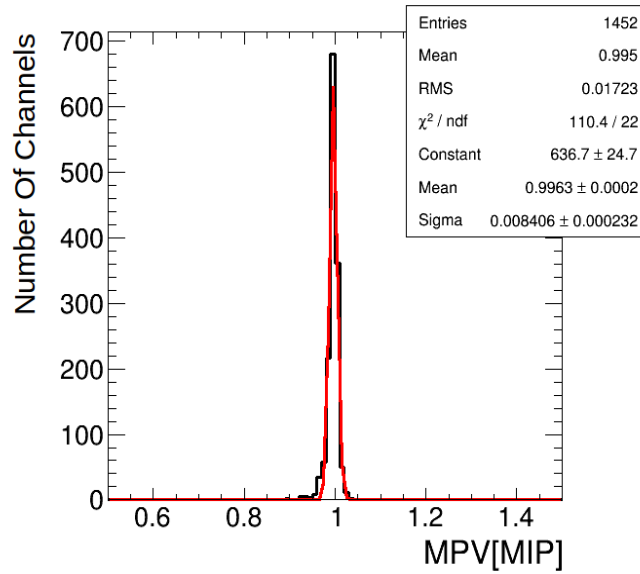


FIGURE 3.23.: *Distribution of the most probable value after applying the calibration constants extracted from the July muon data to the same data set. A Gaussian fit is performed, and as expected the distribution is peaking at 1 MIP.*

Since it was not possible to move the detector during the test beam, the outer channels of the big layers didn't have enough statistics to perform the calibration (more than 1000 entries are required in order to perform the fit, and a typical value per channel is ~ 10000 entries). For these channels data collected during the DESY test beam campaigns in April and May 2015 are used.

In order to evaluate the robustness and the portability of the calibration, the calibration constants obtained have been applied to another set of data, taken with the same prototype, but in a different test beam campaign (CERN test beam campaign occurred in August 2015 in a tungsten absorber structure). The distribution of the most probable value obtained after applying the calibration constants $C^{MIP}[ADC]_i$ is shown in Figure 3.24. As it can be noticed, the distribution is not peaking exactly at 1 MIP and the spread of the distribution is larger. In addition

few outliers are visible. It is important to notice that no temperature correction, which is expected to be of the order of 1.5% per degree, has been applied.

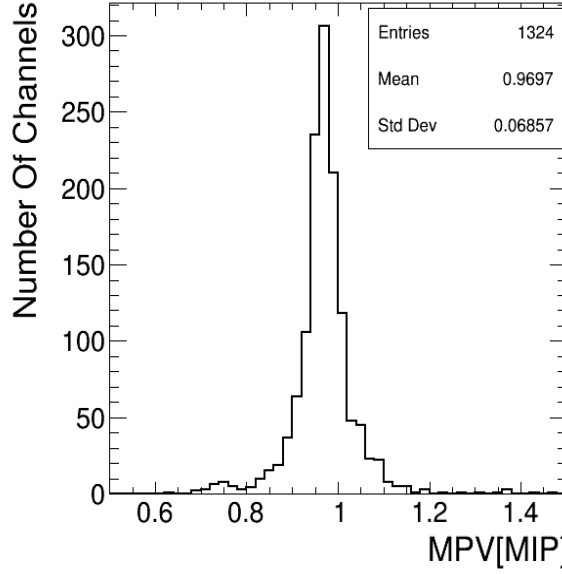


FIGURE 3.24.: *Distribution of the most probable value, after applying the calibration constants extracted from the July muon data to a different data set (August muon data).*

3.5.5. Time Calibration

The SPIROC2B ASICs give information about the time of the hit in the detector. This information is given in TDC counts, therefore it also needs to be calibrated in physics units. The time calibration has been the core of another study within the CALICE collaboration (see [99]). For completeness, it is here summarized.

The calibration of the time from TDC counts to nanoseconds is also done using the muon runs. The conversion from TDC units to nanoseconds is given in Eq. 3.10

$$time[ns] = slope_{(chip, BXID)} \times (time[TDC] - P_{mem=1}) \quad (3.10)$$

where $time[TDC]$ is the time of the hit in TDC counts and $P_{mem=1}$ is the pedestal of the first memory cell. Ideally, the extraction of the pedestal for each memory cell would be necessary. However, due to the limited statistics available, only the first memory cell is used. Any offset introduced by the other memory cells is corrected in a later stage. A typical TDC spectrum for a SPIROC2B chip of the AHCAL prototype is shown in Figure 3.25.

The slope is extracted for each chip of the prototype, distinguishing between the odd and even bunch crossing ID, due to the dual ramps structure of the SPIROC TDC (see Section 3.2). It is defined in Eq. 3.11 as the ratio between

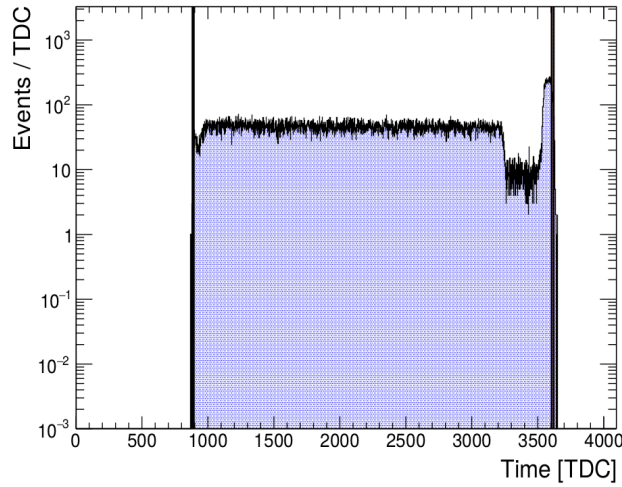


FIGURE 3.25.: *Typical TDC spectrum of a SPIROC2B ASICs, used to extract the slope needed for the calibration of the time, as expressed in Eq. 3.11 [99].*

the ramp length of the TDC in test beam mode and the difference between the maximum value of the TDC spectrum and the pedestal value (minimum value of the spectrum).

$$slope_{(chip,BXID)}[ns/TDC] = \frac{3920 \text{ ns}}{Max_{(chip,BXID)} - P_{(chip,BXID)}}. \quad (3.11)$$

The ramp length of the TDC is 3920 ns instead of the expected 4000 ns in test beam mode due to a dead time of $\sim 2\%$ introduced by a multiplexer that switches between the two TDC ramps [102].

The pedestal value (left edge) and the maximum value of the TDC spectrum (right edge) are identified by the vertical lines in Figure 3.25. In order to extract the lower edge, a threshold μ is introduced, defined as the mean of the vertical axis of the TDC spectrum in Figure 3.25. The lower edge is then given by the value of the first bin above 30% of the threshold μ . The maximum value is extracted as the value of the last bin above 50% of the maximum bin content of the TDC spectrum.

In order to obtain a meaningful value of the time of the hit in the detector, the measured value converted in nanoseconds using Eq. 3.11 has to be compared to a reference time. For this purpose, the signal of the trigger scintillators is directly injected into four AHCAL channels, as shown in Table 3.3. The calibration of these channels is done as described above, with the difference that, due to the availability of huge statistics, for these channels the calibration can be performed per memory cell.

The time of the hit in the detector ($time_{Final}$) is therefore obtained as the difference of the converted TDC value into nanoseconds ($time[ns]$), subtracted by the reference time ($time_{Ref}$), as expressed in Eq. 3.12

$$time_{Final}[ns] = time[ns] - time_{Ref}[ns] \quad (3.12)$$

where $time_{Ref}$ is the average time measured from the three working T0 channels.

After the time has been converted from TDC counts to nanoseconds, some corrections have to be applied to the data, due to some known features of the SPIROC2B. These features are listed below.

- The non linearity of the TDC ramp.
This is an already known characteristic of the SPIROC chip [103]. The extraction of the slope, as described before, is performed with the assumption that the voltage ramp of the TDC is linear. Since this is not exactly the case, a correction for this effect has to be applied. This is done by plotting the calibrated time in nanoseconds versus the non calibrated one in TDC counts and it is done per each chip and bunch crossing ID. It corresponds to an improvement of $\sim 5\%$ in the time resolution (from 5.65 ns to 5.36 ns) [99].
- The time walk effect.
This effect is related to the presence of a fixed trigger threshold. Events with a smaller amplitude will trigger systematically later in time than events with a larger amplitude. In addition, the time during which the signal is read out is fixed, leading to a larger effect. A correction for this effect is obtained looking at the correlation between the time of the hit and the energy of the hit deposited in the detector. A difference up to 6 ns is observed between small and large signal amplitudes and an improvement of $\sim 3\%$ on the time resolution is obtained [99]. The correction is assumed to be the same for all the chips, independently from the position of the threshold in each chip.
- The dependence on the number of triggered channels in a chip.
Looking at the relation between the mean time of the hit as a function of the number of triggered channels in a chip, a correlation has been observed: the mean time of the hit can increase up to 40 nanoseconds for a number of triggered channels above 15. In order to correct for this effect, this correlation has been studied using electron data. It is assumed that the effect is the same for all the SPIROC chips. This assumption is motivated by the fact that not enough statistics were available for the last layers of the prototype, being these layers placed much further (after $\sim 15X_0$ and up to $\sim 46X_0$). The observed differences between the ASICs are taken into account in the systematic uncertainties. After applying the correction, an improvement of $\sim 12\%$ on the hit time resolution for electron data is obtained [99].

More details about the time calibration and the corrections applied can be found in [99].

3.6. Digitization in the Simulation

A simulation of the AHCAL prototype has been performed, using the MOKKA framework v08-05, based on GEANT 4 v10.1[45]. It provides the simulation of the detector according to the test beam setup. The simulation's output given from GEANT 4 consists of a list of energy depositions of particles in a region marked as active material.

What is actually measured in a real detector is a list of digitized measured voltages in hardware units (ADC). These values are then converted in physics units, in the offline reconstruction process.

The reconstruction process and therefore the analysis chain applied to data and simulation should be exactly the same. In order to do that, the effects of the detector, meaning all the readout effects coming from the combination of the active material, the sensor and the electronics, have to be modeled in the simulation.

In this section, the implementation of the optical cross talk, the saturation effect, the timing and finally the noise implementation are shortly described, while the simulation setup is described in Section 4.2.

3.6.1. Optical Cross Talk

As mentioned in Section 3.3.2 and shown in Table 3.1, the AHCAL prototype is composed of different kinds of tiles. Some of the tiles are not individually wrapped in reflector foil. Such a difference affects the cross talk, therefore, for these tiles, the optical cross talk has to be taken into account in the simulation. For the unwrapped tiles, part of the scintillating light produced from an incoming particle in the tile can propagate into the adjacent tiles. This signal will be then added to the signal measured from these tiles themselves.

In the simulation, the optical cross talk effect is simulated by adding a fraction of the energy measured in one tile to the 4 adjacent tiles. A sketch of the implementation of the cross talk in the simulation is shown in Figure 3.26. Here the arrow in the central tile represents the incoming particle depositing most of its energy in this tile, and a fraction of it (lighter blue) propagates in the adjacent tiles.

The intuitive procedure would then be to reduce the energy of the central tile by the amount of energy added in the adjacent tiles. In practice, this couples the MIP calibration parameter to the optical cross talk parameter, impeding to study in an independent way these two aspects. Therefore the energy of the central tile

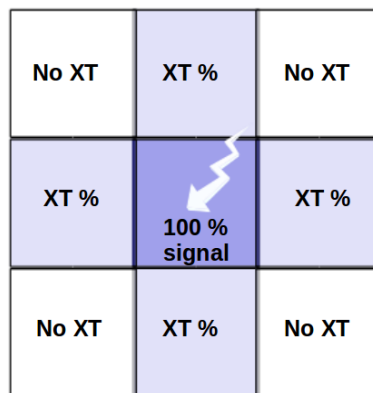


FIGURE 3.26.: Sketch of the implementation of the cross talk (XT) in the simulation. Here a matrix of three by three tiles is shown.

is not reduced by the amount of energy spread in the adjacent tiles, and the MIP calibration is independent from the cross talk parameter [104].

Since no measurements of the cross talk were available for this prototype, the cross talk parameter has been tuned, as explained in Section 4.2.1.

3.6.2. Silicon Photomultiplier Response

A MIP-like particle traversing a 3 mm AHCAL tile generates around 5000 scintillation photons inside the tile [76], and only 10 to 15 of them are detected by the SiPM. Because of the low number of detected photons, their distribution is well modeled by a Poisson distribution, leading to a relative resolution contribution of $1/\sqrt{a}$ for a number of incident photons a .

The major contribution of the sensor to the single channel amplitude resolution is represented from the finite number of pixels of the silicon photomultiplier. In particular, an important role is played from the statistical fluctuations of the fraction of pixels fired for very high amplitude hits. These hits lead to a large correction factors to the saturated response of the SiPMs, amplifying this effect. This effect is expected to be more relevant for hits in which the number of incoming photons exceed the number of effective pixels of the silicon photomultiplier. These effects are modeled for an ideal SiPM, where no optical cross talk, after pulse and recovery time are present.

The model adopted from the CALICE collaboration in simulating test beam data, is a binomial model [43, 76, 104].

In the digitization process, the amplitude energy in pixels is computed for each hit in the calorimeter. Taking into account the fact that the SiPMs have a limited number of pixels, the energy amplitudes of each hit are saturated, with a function

defined as in Eq. 3.13

$$A^{sat}[pixels]_i = N_{effPixels} \cdot (1 - e^{A^{noSat}[pixels]_i/N_{effPixels}}) \quad (3.13)$$

where $A^{sat}[pixels]_i$ is the saturated amplitude of the hit, $N_{effPixels}$ is the number of effective pixels of the silicon photomultipliers (parameter that has to be measured or tuned, as explained in Section 4.2.2), $A^{noSat}[pixels]_i$ the amplitude of the hit not saturated.

After that, the binomial model is applied. The number of fired pixels n is randomly generated using a binomial distribution $B(N_{effPixels}, p)$, that has as input parameters the number of effective pixels of the SiPMs, $N_{effPixels}$, and the probability p that a pixel is fired, defined as $p = A_{sat}/N_{effpixels}$.

3.6.3. Timing

The effects of the readout electronics have to be applied also to the time measurement. The calibration and determination of the resolution has been done using muon data. The distribution of the hit time in the detector is considered and fitted with a sum of two Gaussians. The parameters extracted from the fit have been used to smear the time of the simulated hits.

The effect described in Section 3.5.5 arising when many channels trigger within one chip has also been implemented. A more detailed description can be found in [99].

3.6.4. Noise Implementation

The noise frequency and amplitude of the silicon photomultiplier depend on the specific model and type of SiPM as well as on external factors, such as the temperature and the applied voltage.

Even though, in order to reject the noise in the analysis, an offline cut at 0.5 MIPs is applied, the remaining noise has to be taken into account in the simulation.

In order to properly simulate it, the muon runs taken during the test beam campaign are used. The muon track is identified, requiring to have a number of hits in a row in the detector higher than 7 and only one hit per layer, in order to reject pions. The identified track, shown in Figure 3.21, is removed from the event, and the remaining hits in the detector are considered as noise. They are overlaid event by event to the simulated events. The effect of the noise has been checked on one of the 10 GeV electron runs. It corresponds on average to two hits more per event, and to roughly six MIPs on the average energy sum per event. In average, the hit energy associated to one noise hit corresponds to ~ 2 MIPs. Looking at the hit energy distribution for the noise hits, a tail towards higher hit energies (up to ~ 50 MIPs) is visible. This is most likely due to delta electrons

from the muon tracks [99] and represents the reason of the larger contribution of the hit noise on the average energy sum.

3.7. Summary

In this chapter, a description of the AHCAL technological prototype has been given.

In the first part a description of the silicon photomultiplier in general has been presented. Later, the silicon photomultiplier - on tile technology, and the specific developed readout electronics, the SPIROC2 ASIC chips, have been described.

After this introduction, an overview of the several calibration procedures, needed to have the output signal of the electronics in physics units, have been described. These include the MIP calibration, needed to have the energy in MIP units, the gain calibration needed to convert the hit amplitude to the pixel scale, the high gain - low gain intercalibration, due to the dual gain mode of the read out chips and the time calibration, to have the time of the hit in the detector converted into nanoseconds.

In order to have a good understanding of the data, these have to be compared with a simulation. In this chapter the simulation of the AHCAL prototype is described. In particular the digitization effects, meaning the conversion of the list of energy depositions of particles in the active material provided by GEANT 4 into a list of digitized measured voltages in hardware units (ADC), are illustrated.

4. Analysis of Electron Beam Data

The main goal of this thesis is the study of hadronic showers.

In order to do that, a good knowledge and understanding of the detector are a crucial step. It was in fact the first time that this prototype, with this hardware and in particular with fully integrated electronics, scalable to the full detector, was tested. Therefore it has been needed to develop an analysis able to give a good knowledge of the detector and its performance.

Due to the good knowledge of the electromagnetic shower, electron data allow to do that, and allow the tuning of the most important parameters of the simulation, like the cross talk and the implementation of the saturation corrections, more complicated with pions data, due to the complexity of the hadronic showers.

For this purpose, during the test beam campaign at CERN in July 2015, electron data in the energy range from 10 GeV to 50 GeV (in steps of 10 GeV) were collected.

As mentioned above, the electron data are used only with the purpose of tuning the main parameters of the simulation. Electromagnetic showers are much denser than hadronic showers, therefore they are a very powerful tool to investigate and understand the performance of the detector, and investigate the major effects. A precision and agreement of 20% - 30% between data and simulation on this analysis is therefore acceptable, in order to be able to perform the analysis of the pion data. Being the hadron showers less denser than the electromagnetic showers, a precision of 20% – 30% of electron data is expected to be translated also on pion data. In addition, large fluctuations in the hadronic showers are expected, such that this precision will not be a limit in the comparison between data and simulation for the pion data.

This chapter is focused on the description of the analysis of the electron test beam data. In particular the event selection, the tuning of the main parameters of the simulation and the comparison between data and simulation will be presented.

Even though the electron data are used to tune the parameters of the simulation, the plots shown in Section 4.1 are performed using the final tuned parameters. The tuning of these parameters is then illustrated in the Section 4.2.

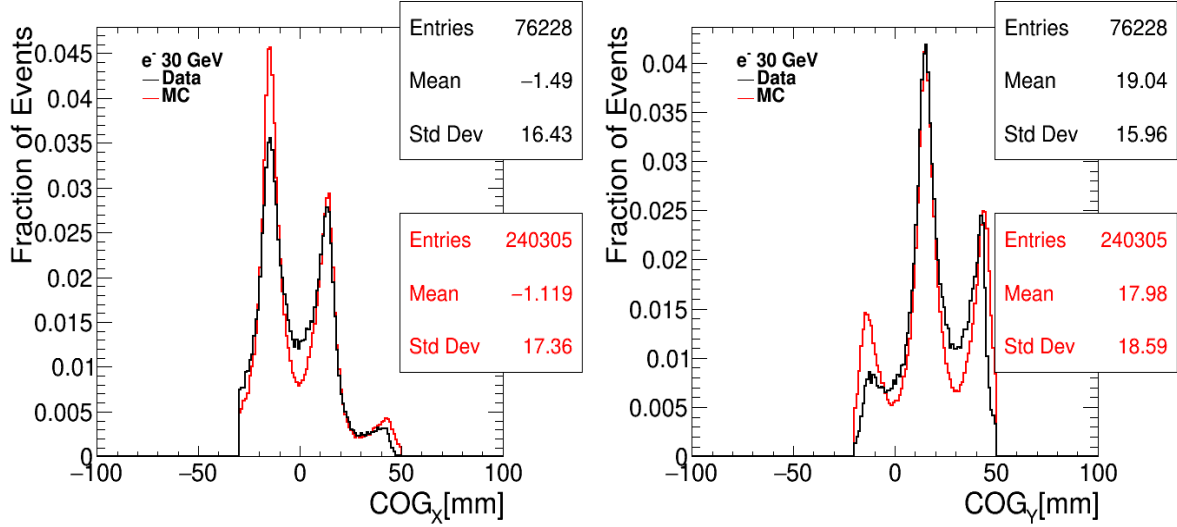


FIGURE 4.1.: Center of gravity distribution for 30 GeV electrons in X (left) and Y (right) for data (black) and simulation (red).

4.1. Event Selection

The electron beam at CERN is not a pure electron beam, therefore it is needed to perform an event selection in order to be sure to consider only single electrons in the analysis.

On the data, a preselection using only the beam instrumentation is performed (see Section 3.4). The Cherenkov detector has been used in order to identify the electrons: only particles that generate a signal in the Cherenkov are considered.

In addition to that, due to the presence of two trigger scintillators in front of the detector, the smallest of the dimension of $10 \times 10 \text{ cm}^2$, a cut on the center of gravity of the shower has been applied, requiring to be within this region.

The center of gravity (COG) of the shower per event indicates the position of the core of the shower and its X coordinate is defined in the relation 4.1 (and similarly in the Y and Z coordinates)

$$COG_X = \frac{\sum_i E_i \cdot X_i}{\sum_i E_i}. \quad (4.1)$$

The sum is performed over the number of hits per event, and E_i and X_i represent respectively the energy and the position of the hit i in the event.

All the information related to one channel of the calorimeter (the energy deposition, the time, the coordinates) are stored in what is called hit, therefore in one event there is more than one hit.

The distribution of the center of gravity in the X and Y directions for 30 GeV electrons is shown in Figure 4.1.

Since the beam was not always pointing through the center of the detector, contrary to the trigger scintillator placed always at the same position, in order

to hit the center of the detector, this cut is different for each energy, as shown in Table 4.1.

From Figure 4.1 appears that the simulation doesn't reproduce exactly the distributions obtained from the data. This is due to the fact that it was not possible to measure the exact position of the beam during the test beam campaign, since the wire chambers were not connected. In order to simulated the beam profile in the simulation, the mean and the RMS of the distributions showed in Figure 4.1 are used as X and Y position of the beam and the spread of the beam. The spread of the beam is assumed to follow a Gaussian distribution. These values don't correspond to the exact position of the beam, but to the position of the beam as it is seen in the detector, after traveling several meters in the beam line and interacting with the several elements in the beam line. Nevertheless, these values give a good approximation. The peaks in the distribution in Figure 4.1 reflect the granularity of the calorimeter, and they correspond to the $3 \times 3 \text{ cm}^2$ size of the scintillating tiles.

Then the electron selection is based on cuts on the total number of hits in an event and the energy sum.

In the analysis, a number of hits higher than 15 is required, in order to avoid low energy electrons contamination and muons contamination, and the energy sum per event in the last two layers is required to be less than 1% in order to avoid pions contamination.

The distribution of the number of hits per event and the energy sum per event for 30 GeV pre-selected electrons are shown in Figure 4.2.

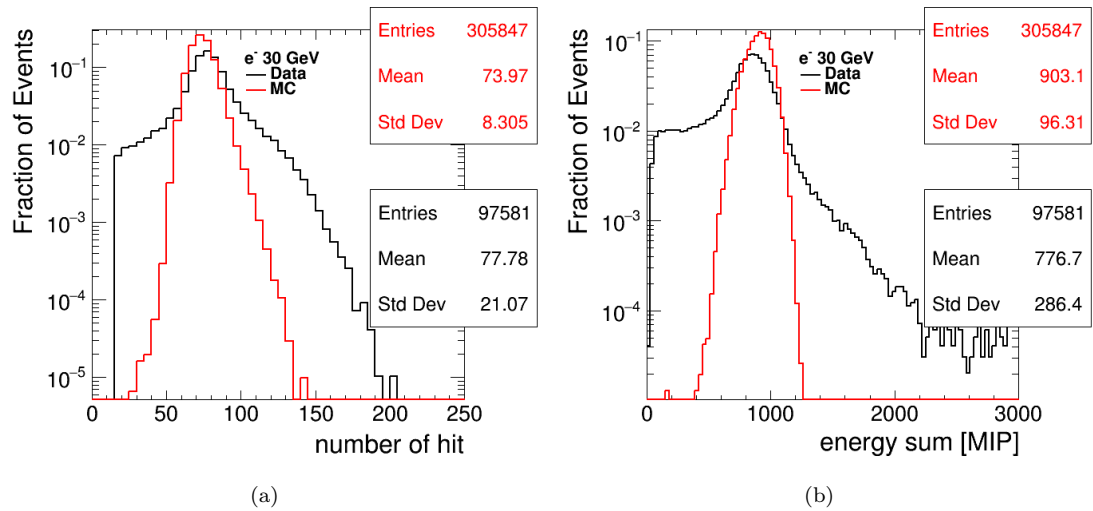


FIGURE 4.2.: Comparison between data and simulation for 30 GeV electrons.

(a): Distribution of the number of hits per event.

(b): Distribution of the energy sum per event.

In Figure 4.2(a) and in Figure 4.2(b) is possible to notice the persistence of a considerable tail both on the left and on the right of the distributions. The right

tail can be due to multiparticles events, i.e. more than one electron showering at the same time, very hard to identify, while the left tail can be caused by low energy electrons, coming from electrons starting to shower earlier. In principle, also hadrons contamination can contribute to this low energy tail. The rejection of these particles is ensured by requiring that the energy sum in the last two layers of the calorimeter is less than 1%.

In order to improve the selection and reducing the left tail of the distribution the quantity R , defined from the relation 4.2 has been considered

$$R = \frac{(\sum_i E_i)_{FirstModule}}{(\sum_i E_i)_{tot}}. \quad (4.2)$$

These low energy electrons are in fact expected to lose all their energy in the first AHCAL module, that correspond to a value of R close to 1.

The distribution of the R quantity for 30 GeV electrons is shown in Figure 4.3.

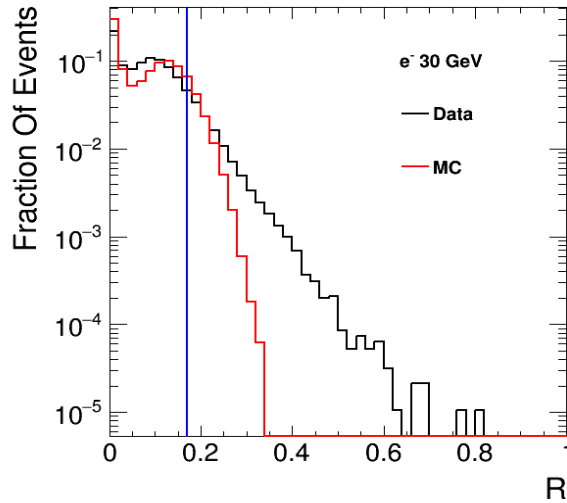


FIGURE 4.3.: Distribution of the R quantity defined in Eq.4.2 for 30 GeV electron data (black) and simulation (red). The blue line indicates the cut applied.

The distribution of the number of hits per event and of the energy sum per event, with all the cuts here described and without the cut on the R quantity, are shown in Figure 4.4 for the data and in Figure 4.5 for the simulation. From Figure 4.4 is possible to see a slight improvement in the rejection of the low energy electrons, while almost no impact can be observed in the simulation, looking at Figure 4.5.

All the values of the cuts applied for the different energies are listed in Table 4.1.

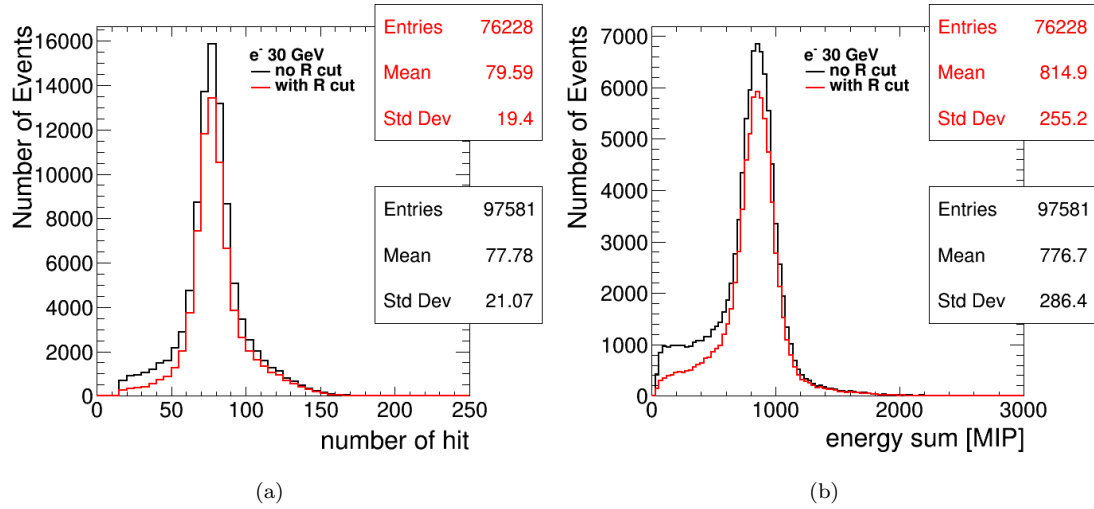


FIGURE 4.4.: 30 GeV electrons data.

- (a): Distribution of the number of hits per event with the cut on the R quantity (red) and without (black).
- (b): Distribution of the energy sum per event with the cut on the R quantity (red) and without (black).

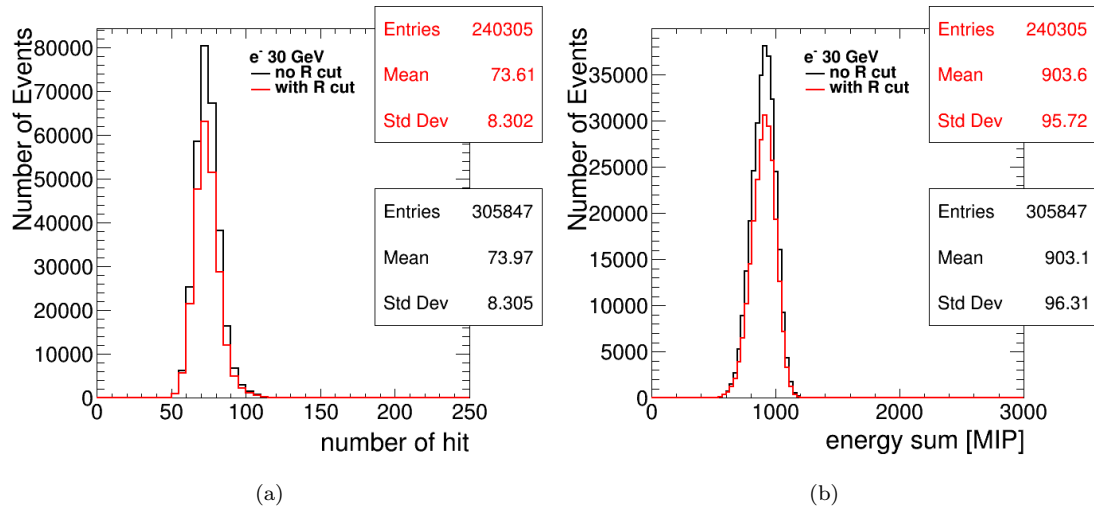


FIGURE 4.5.: 30 GeV simulated electrons.

- (a): Distribution of the number of hits per event with the cut on the R quantity (red) and without (black).
- (b): Distribution of the energy sum per event with the cut on the R quantity (red) and without (black).

Energy	COG_X[mm]	COG_Y[mm]	R
10GeV	(−45, 45)	(−45, 45)	< 0.30
15GeV	(−45, 45)	(−45, 45)	< 0.25
20GeV	(−45, 45)	(−45, 45)	< 0.20
30GeV	(−30, 50)	(−20, 50)	< 0.17
40GeV	(−45, 45)	(−20, 50)	< 0.15
50GeV	(−50, 20)	(−30, 50)	< 0.12

TABLE 4.1.: *Value of the cut according to the beam energy.*

4.2. Simulation

As mentioned in Section 2.2 and in Section 3.6, the simulation of the AHCAL prototype uses the MOKKA framework v08-05 which is based on GEANT 4 v10.1 and provides the simulation of the detector according to the test beam setup.

The description of the beam line, very complex, is not fully implemented. The two trigger scintillators, of the dimensions of $10 \times 10 \text{ cm}^2$ and $50 \times 50 \text{ cm}^2$ have been included, as well as the Cherenkov detector. This has been placed 90 m in front of the detector. It is 11 m long and filled with Helium gas, the pressure of which varies according to the beam energy (see Appendix A).

The description of the upstream material has been simulated placing 12 mm of steel in front of the detector. This value has been tuned looking at the center of gravity of the shower in Z, and it will be kept as a systematic uncertainty.

A sketch of the beam line, as it is implemented in the simulation is shown in Figure 4.6.

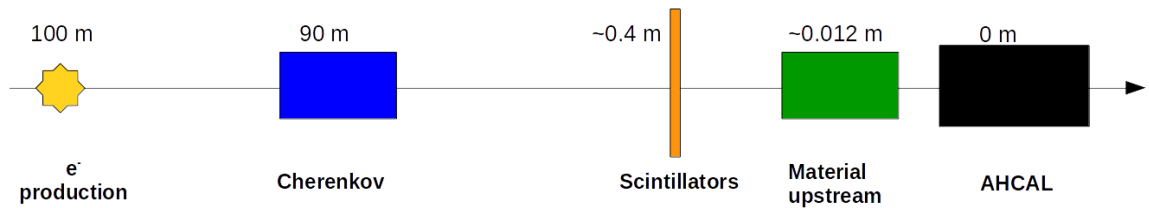


FIGURE 4.6.: *Sketch of the beam line, as described in the simulation. The numbers indicate the distance from the detector.*

The need of the introduction of the additional upstream material in the simulation appears looking at the comparison of the distribution of the center of gravity in the Z direction, shown in Figure 4.7 for one of the run at 10 GeV.

Here the simulation seems to start to shower later than the data. After verified the absence of bugs in the simulation, this led to the conclusion that the amount of material used in the simulation didn't correspond to the amount of material presents in the beam line at CERN.

After adding the Cherenkov detector, three different amounts of additional steel material in the simulation have been considered in order to simulate the upstream material in the beam line (10 mm, 12 mm and 14 mm of additional material before

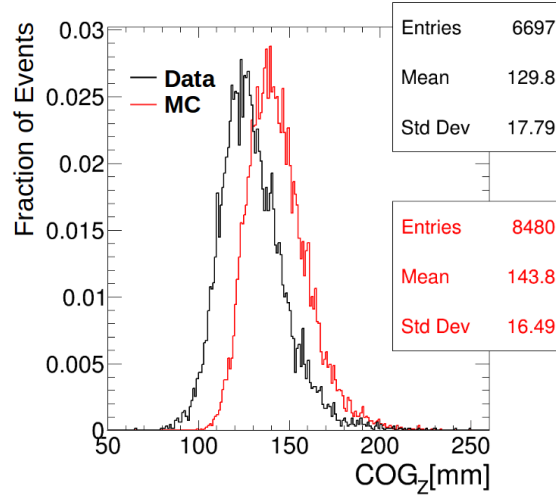


FIGURE 4.7.: *Center of gravity distribution in the Z direction for 10 GeV electrons in black for data and in red for the simulation. Here no Cherenkov detector and additional material are added in the simulation.*

the detector) and the center of gravity distributions in Z have been compared (Figure 4.8).

Here it can be seen that higher energy would prefer more material, but this still will not help in matching the left tail of the distribution, due to the low energy electrons contamination described in Section 4.1. Therefore a value of 12 mm of steel before the prototype is chosen, and the difference between the analysis obtained adding 14 mm more of steel and the analysis obtained adding 10 mm of steel before the detector is considered as systematic uncertainty.

4.2.1. Tuning of the Optical Cross Talk

As explained in Section 3.6.1, the AHCAL prototype is composed of different kind of tiles, some of them not individually wrapped in reflector foil. For these tiles the optical cross talk has to be taken into account in the simulation.

The modules involved are: the two new ITEP boards and the old ITEP boards (layers from 4 to 10, see Figure 3.7 and Table 3.1).

For all the other modules, with tiles individually wrapped, laboratory measurements show a value of the cross talk less than 5% per tile and therefore negligible.

To tune this parameter, a simulation with different values of cross talk (10%, 12% and 15%) have been performed. The number of hits distribution has been studied for tuning this parameter in the simulation for electron energies from 10 GeV to 50 GeV and the mean value has been plotted as a function of the beam energies. The simulation results are then compared to the test beam data. The distribution of the number of hits per event is shown in Figure 4.9 for 10 GeV and 30 GeV electrons. Due to the presence of the left and right tails in the distribution

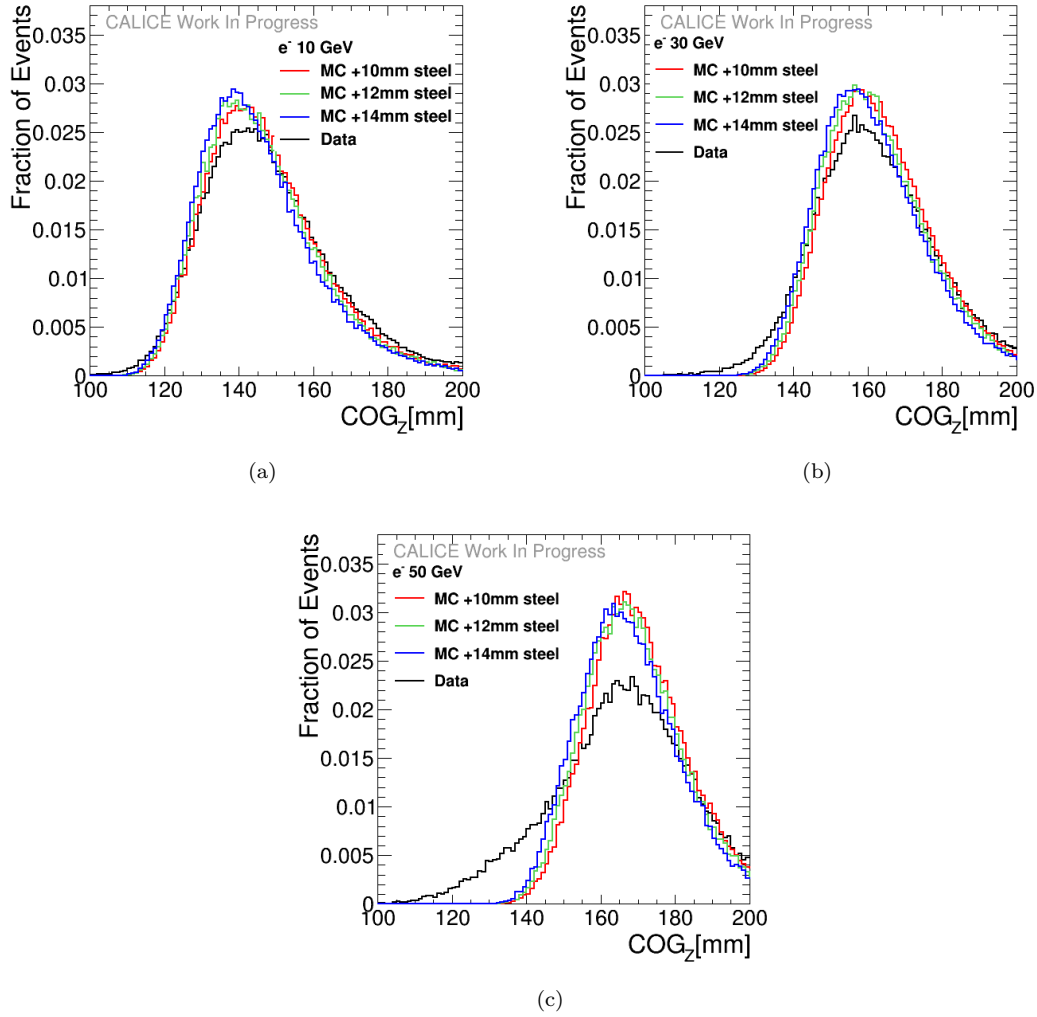


FIGURE 4.8.: Distribution of the center of gravity of the shower (COG) in the Z direction for 10 GeV (a), 30 GeV (b) and 50 GeV (c) electrons. The black histogram represents the data, while different colors are used for the simulation with different amount of steel before the prototype.

of the number of hits per event (see for example Figure 4.4), a cut on the number of hits per event has been applied in order to have a better estimation of the mean of the distribution.

In order to be able to tune better this parameter in the simulation, the distributions in Figure 4.9 have been studied for all the energies and the mean of these distributions has been plotted as a function of the beam energy. This is shown in Figure 4.10. The different colors represent the different values of the cross talk used in the simulation. Figure 4.10 shows also the ratio between the mean number of hits per event in the data and the mean number of hits per event in the simulation as a function of the beam energy.

The value of the cross talk does not depend on the beam energy. The trend seen in Figure 4.10 is due to the fact that the mean of the distribution is influenced

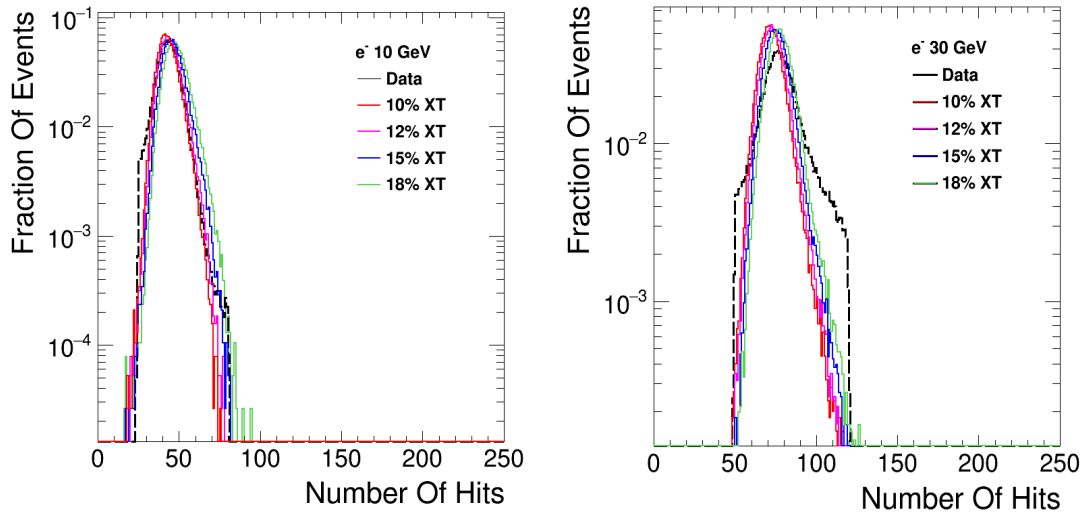


FIGURE 4.9.: Distribution of the number of hits per event for 10 GeV (left) and 30 GeV (right). The black line represents the data, while the different colors, the simulation obtained with different value of cross talk per tile, varied between 10% and 18%. Here a lower and upper cut on the number of hits per event have been applied.

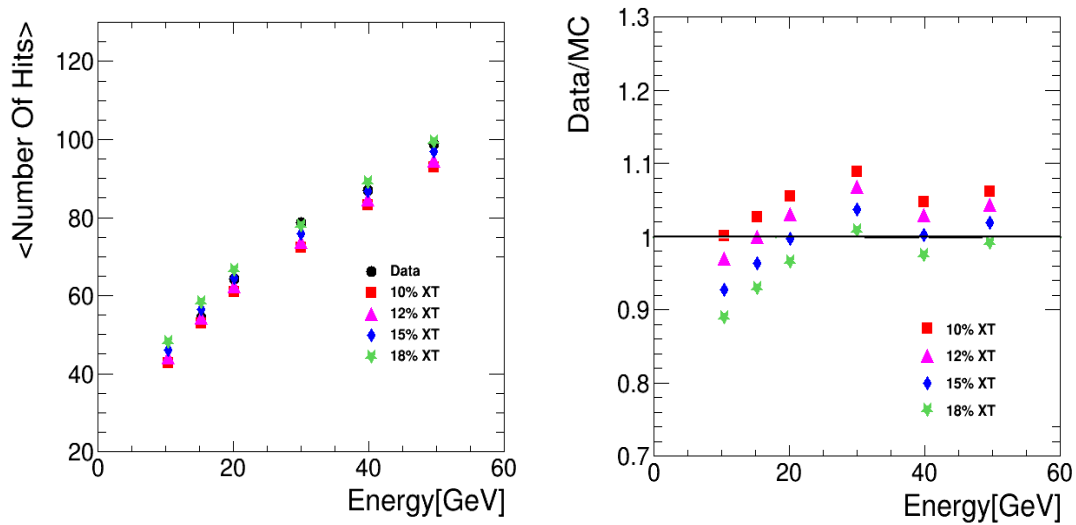


FIGURE 4.10.: On the left, the mean of the number of hits as a function of the beam energy has been shown, while on the right the ratio R between data and simulation as a function of the beam energy is presented.

by the left tail of the number of hits distribution, that is much more pronounced for high energies (see for example Figure 4.4), even though a cut on the number of hits has been applied in the estimation of the mean of the histogram.

A value of 12% cross talk per tile is chosen, which described the lower energies best, where the influence from other effects like low energy electrons contamination, SiPMs saturation and high gain - low gain intercalibration is smallest.

Finally the difference between the two analysis obtained using a value of the cross talk equal to 10% per tile and a value of the cross talk equal to 15% per tile is considered as a systematic uncertainty.

4.2.2. Tuning of the Number of Effective Pixels for The Saturation Correction

As explained in Section 3.1, silicon photomultipliers are non linear devices and this effect can be corrected offline.

Since no measurements of the saturation curve for the SiPMs used in this prototype were done, the number of effective pixels needed to correct for the saturation effect of the silicon photomultipliers has to be tuned, and this has been done using the electron data.

The function used to correct for the saturation effect is given in Eq. 4.3

$$A^{noSat}[pixels]_i = -N_{effpixels} * \ln(1 - A^{Sat}[pixels]_i / N_{effpixels}) \quad (4.3)$$

where $A^{noSat}[pixels]_i$ represents the energy of each hit in the detector corrected for the saturation effect, $N_{effpixels}$ the number of effective pixels used and $A^{Sat}[pixels]_i$ the hit energy not corrected for saturation effect. If more than 95% of the pixels are fired, a linear approximation is used to correct for the saturation effect. The function is defined in Eq. 4.4

$$A^{noSat}[pixels]_i = \frac{1}{1 - 0.95} * (A^{Sat}[pixels] - 0.95 * N_{effpixels}) - N_{effpixels} * \ln(1 - 0.95). \quad (4.4)$$

It is possible to notice that the function defined above used to correct for the saturation effect is not exactly the inverse of the function used in the digitization process, defined in Eq. 3.13. This is due to the fact that this function has a maximum. In data, higher values can be reached, therefore, after a certain threshold (more than 95% pixels fired) a linear approximation is used.

The parameter $N_{effpixels}$ needs to be tuned in order to correct for the saturation effect. For this study, simulated electron data have been produced, applying a saturation function (that means saturating the simulation) using different number of effective pixels, and they have been compared with the data where no saturation correction is applied.

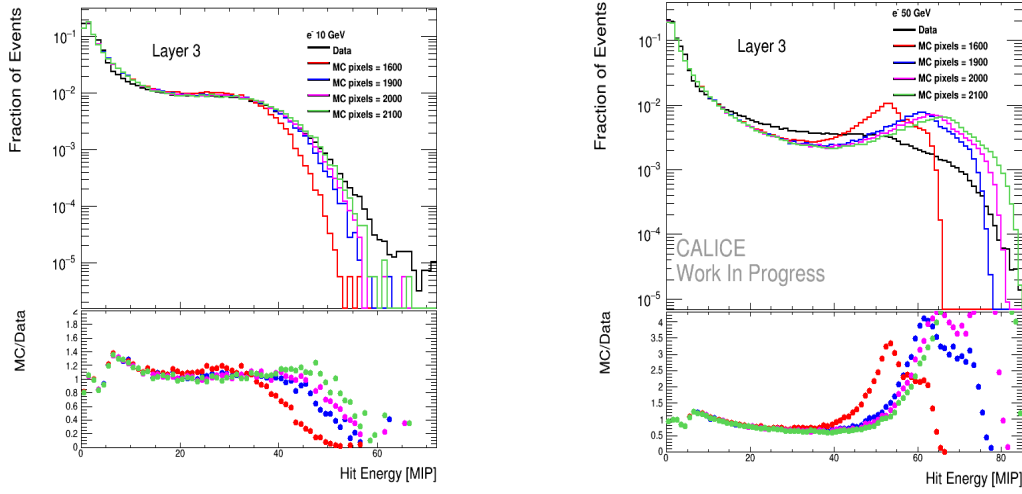


FIGURE 4.11.: *Tuning of the number of effective pixels for the Mainz module. Here the hit energy distribution for the data and simulated events (according to Table 4.2) are shown, on the left side for 10 GeV electrons and on the right side for 50 GeV electrons. Below the ratio between the simulation and the data is plotted.*

The number of effective pixels tested are shown in Table 4.2, while the hit energy distributions for the different modules are shown in Figure 4.11 for the Mainz module and in Figure 4.12 for the old ITEP boards, for the lowest and the highest energy (10 GeV and 50 GeV). Here the black line represents the data, while the different colors the simulation obtained using a different value of the number of effective pixels. Below, the ratio between the hit energy in simulation and data is shown. It is possible to notice that:

Layer	Real pixels	Pixels _{test1}	Pixels _{test2}	Pixels _{test3}
Mainz board(Layer 3)	1600	1900	2000	2100
New ITEP boards (Layers 4,5)	12000	✗	✗	✗
Old ITEP boards (Layers 6 → 10)	800	900	1000	1100
KETEK (Layers 11,12)	2300	✗	✗	✗
SensL (Layers 13,14)	1300	✗	✗	✗

TABLE 4.2.: *Number of pixels of the SiPMs (Real pixels) and number of effective pixels tested in order to tune this parameter in the simulation. For the new ITEP boards the real number of pixels is used, because very large, while for the KETEK and SensL layer the real number of pixels is kept, because very far and not enough statistics were available.*

- a number of effective pixels equal to the real number of pixels of the silicon photomultipliers is not good enough to describe the behaviour of the data

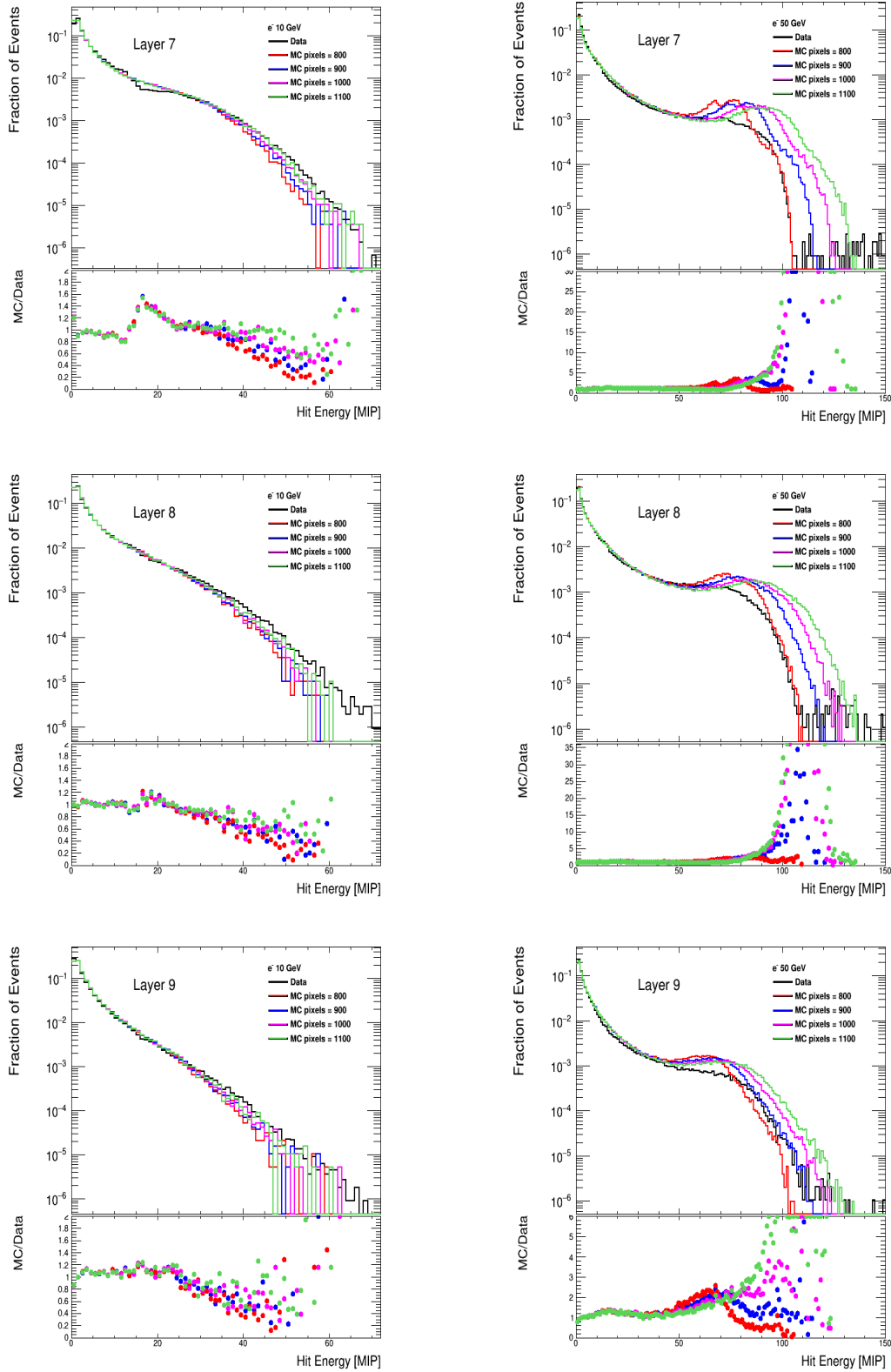


FIGURE 4.12.: Tuning of the number of effective pixels for the old ITEP modules. Here the hit energy distribution for the data and simulated events according to Table 4.2 are shown. The ratio between simulation and data is also plotted. The top plots represent module 7, the two middle plots module 8 and the two bottom plots module 9. The left plots are for 10 GeV electrons and the right one for 50 GeV electrons.

- for low energies the effect is small, as expected, and the saturation function used to saturate the simulation describes the behaviour of the data rather well
- for higher energies the saturation function doesn't give a good description of the shape of the data.

From Figure 4.12 it also possible to notice that above the intercalibration region (~ 5 MIP), a deviation up to 30% is observed. This is due to the difficulties in performing a proper high gain low gain intercalibration (see Section 3.5.3). This is effect is anyway well taken into account in the systematics. As mentioned at the beginning of this chapter, the goal of this thesis is to perform a pion analysis and use the electron data to tune the main parameters in the simulation. The agreement reached with the electron analysis is within 20% – 30%. Hadron showers are less denser than electromagnetic showers, therefore the effect of the saturation correction together with the high gain low gain intercalibration will be less important. In addition, large fluctuations in the hadronic showers are expected, such that these effects visible in the electromagnetic showers will not limit the comparison between data and simulation for pion data.

The number of effective pixels chosen and used for the saturation effect is listed in Table 4.3.

Layer	Effective pixels
Mainz board (Layer 3)	2000
New ITEP boards (Layers 4,5)	12000
Old ITEP boards (Layers 6 \rightarrow 10)	900
KETEK boards (Layers 11,12)	2300
SensL boards (Layers 13,14)	1300

TABLE 4.3.: *Number of effective pixels used in the simulation for all the layers.*

It can be seen that for the two new ITEP boards and for the four big layers, the number of effective pixels is kept equal to the real number of pixels. This was due to the fact that the 4 big layers where too far and consequently not enough statistics were available, to perform a similar study. For the two new ITEP boards, the reason is attributed to the fact that these modules have already a large number of pixels. The corresponding plots of Figure 4.11 and Figure 4.12 for these modules are shown in Figure 4.13 and 4.14.

The same study couldn't be performed on layers 6 and 10. The large amount of not working channels on these two layers made the tuning of the number of effective pixels of the SiPMs with this method not feasible. For this reason, these two layers are not considered further in the analysis.

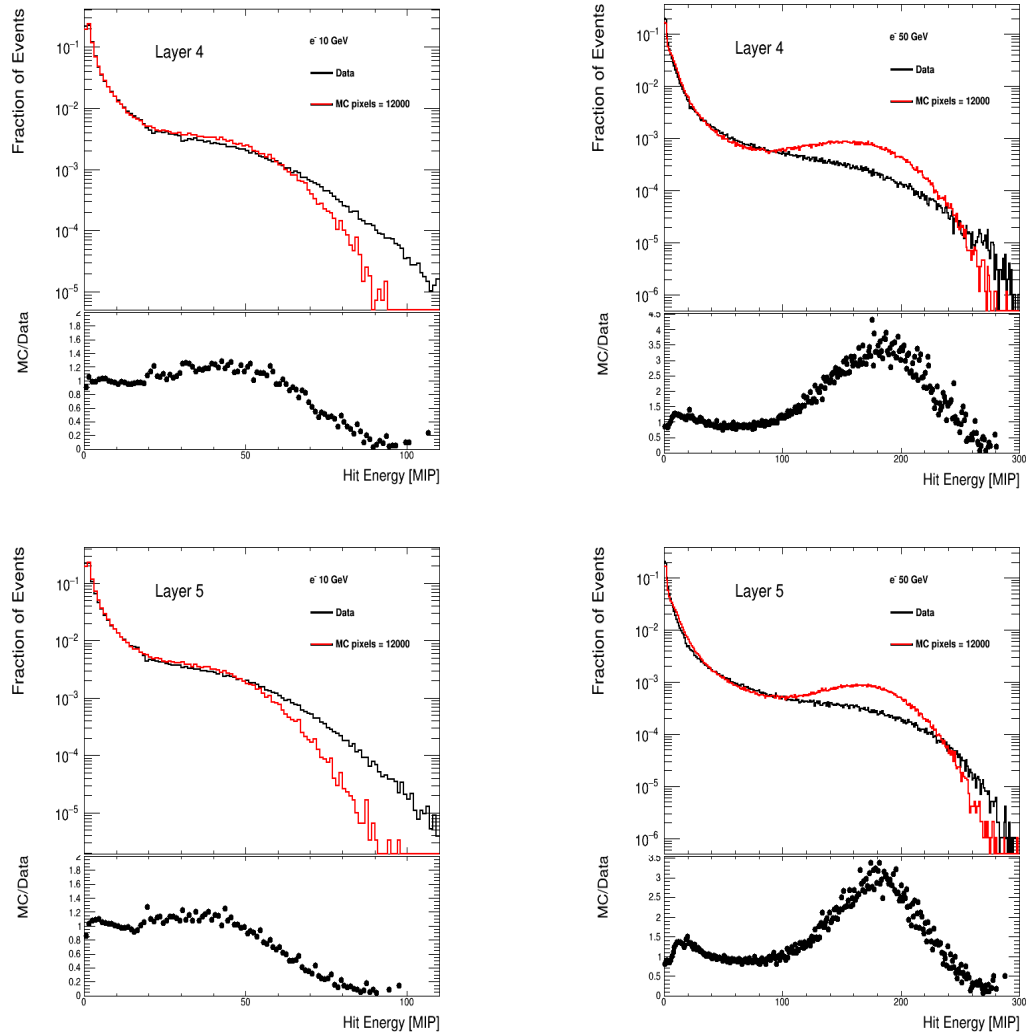


FIGURE 4.13.: Comparison between the hit energy distribution for data and simulation saturated using a number of effective pixels equal to the real number of pixels for the two new ITEP boards. On the left side the 10 GeV electrons are shown, while on the right the 50 GeV electrons

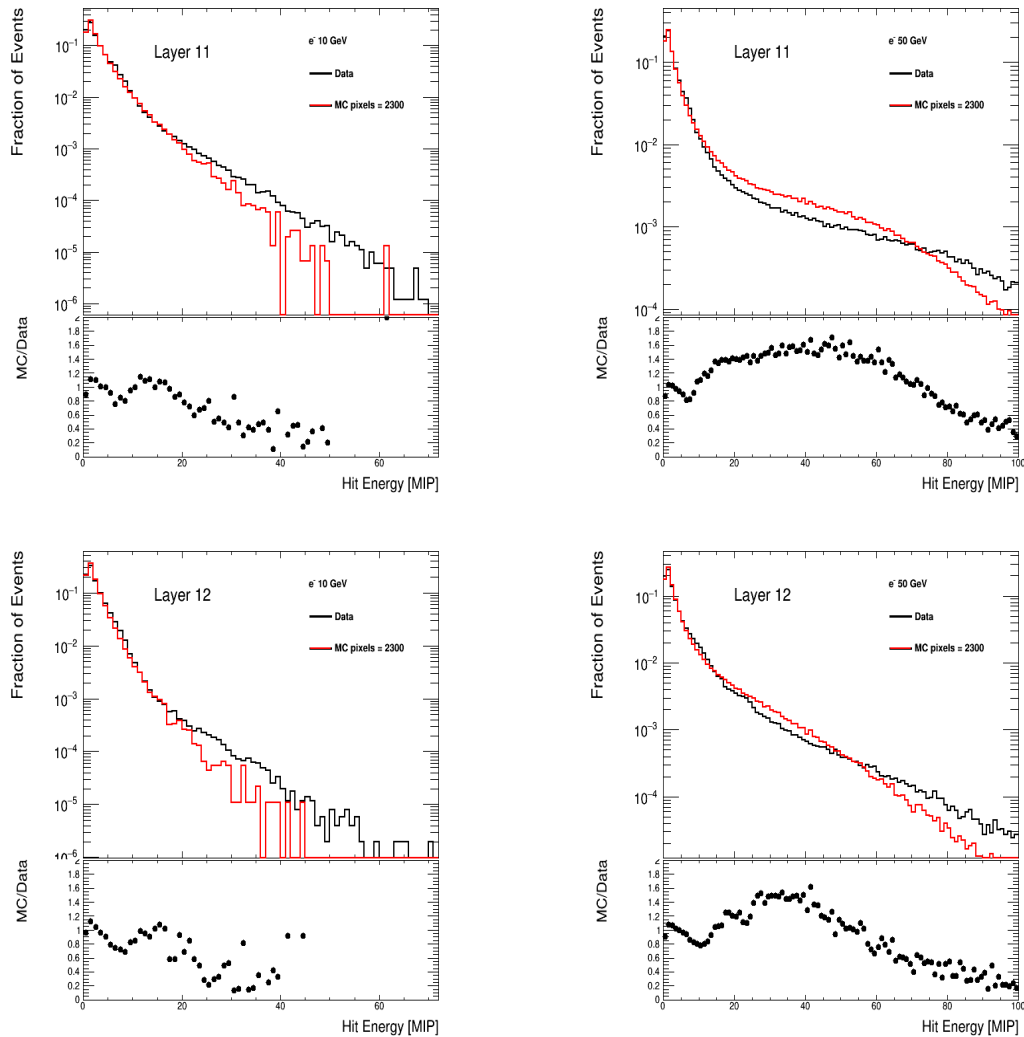


FIGURE 4.14.: Comparison between the hit energy distribution for data and simulation saturated using a number of effective pixels equal to the real number of pixels for the two KETEK modules. On the left side the 10 GeV electrons are shown, while on the right the 50 GeV electrons

4.3. Systematic Uncertainties

As already discussed at the beginning of this chapter, a precision of 20% – 30% would be adequate to perform the pions analysis. This is for several reasons, such as the fact that the first modules of the calorimeter will be used only as a shower start finder (see Section 5.1), or that the electromagnetic showers are denser than the pion showers, therefore some effects (like the saturation effect or the high gain - low gain intercalibration) will have a smaller impact on the hadronic analysis.

In order to know the precision reached, the systematic uncertainties have to be studied.

The contributions that affect the simulation and the data are different.

For the simulation, the following effects have to be taken into account:

- the value of the cross talk per tile chosen
- the number of effective pixels used to correct for the saturation effect
- the amount of additional material placed before the detector in order to simulate the upstream material.

For the data, in addition to the uncertainty due to the number of effective pixels used to correct for the saturation effect, the uncertainties on the extraction of the different calibration constants have to be taken into account. These include:

- the uncertainty due to the gain calibration
- the uncertainty due to the MIP calibration
- the uncertainty due to the high gain - low gain intercalibration.

The uncertainty due to the saturation correction has to be considered both on data and simulation. This is due to the fact that in the simulation the same number of effective pixels is used before to saturate the simulation, in order to simulate the SiPMs response (as explained in Section 3.6.2), and then to desaturate it, correcting for the effect. Therefore in the simulation the effect due to saturation correction is largely canceled, and the residual effect mainly describes the different functions used to saturate and desaturate (see Eq. 3.13, Eq. 4.3 and Eq. 4.4).

4.3.1. Uncertainties on the Simulation

The uncertainties due to the introduction of the additional material in the simulation, the tuning of the number of effective pixels and the cross talk are computed all in the same way.

In each case a lower (min) and an upper (max) value have been defined and the associated uncertainty is equal to half of the difference between the analysis

performed using the min value and the analysis obtained using the max value, as illustrated in Eq. 4.5, where Δ is the uncertainty and Analysis(max)(Analysis(min)) the analysis obtained using the upper (lower) value in the tuning process

$$\Delta = \frac{\text{Analysis}(\text{max}) - \text{Analysis}(\text{min})}{2}. \quad (4.5)$$

This difference is computed for each distribution, for each bin of the distribution.

For the cross talk, the lower and upper values are chosen according to Figure 4.10, where it's shown how lower energy seems to prefer a smaller value of cross talk ($\text{min} = 10\%$) while the higher energy an higher value ($\text{max} = 15\%$).

For the saturation correction, as a lower limit the number of effective pixels used minus 100 is assumed, while as an upper value the number of effective pixels plus 100. These values depend on the module. It needs also to be noticed that if the number of effective pixels used in the simulation is equal to the real number of pixels, this number is also considered as the lower value. These values are summarized for each module in Table 4.4.

Layer	Tuned	Min	Max
Mainz (Layer 3)	2000	1900	2100
New ITEP boards (Layers 4,5)	12000	12000	12100
Old ITEP boards (Layers 6 \rightarrow 10)	900	800	1000
KETEK (Layers 11,12)	2300	2300	2400
SensL (Layers 13,14)	1300	1300	1400

TABLE 4.4.: Values used to extract uncertainty on the tuning of the number of effective pixels in order to correct for the saturation effect.

The lower limit and the upper limit for the tuning of the additional material are chosen according to Figure 4.8, where the distribution of center of gravity in the Z direction for 10 GeV, 30 GeV and 50 GeV simulated electrons is shown assuming different amount of material before ECAL. According to this figure the lower and upper limit are set to $\text{min} = 10$ mm and $\text{max} = 14$ mm, respectively.

In order to quantify the systematic uncertainties affecting the simulation, the distributions obtained running the standard analysis and the analysis obtained using the lower (upper) value have been considered and the ratio defined in Eq. 4.6 has been considered

$$\left(\frac{MC_{\text{tuning}}}{MC} \right)^i = \frac{MC_{(\text{max})}^i}{MC^i}. \quad (4.6)$$

The ratio $\left(\frac{MC_{\text{tuning}}}{MC} \right)^i$ represents the analysis obtained using the upper value ($MC_{(\text{max})}^i$) normalized to the analysis obtained using the final parameter (MC^i), for each bin of the distribution. The same ratio has been calculated for the analysis obtained using the lower value.

In Figure 4.15 these distributions and the associated ratio are shown for 50 GeV electrons.

The top plots represent the distribution of the number of hits (left) and the distribution of the energy sum (right) while the bottom plots, the mean number of hits per layer (left) and the mean energy sum per layer (right).

The black line represents the distributions obtained using the standard analysis, while the red one and the blue one, the analysis obtained using respectively the lower (10% of cross talk per tile) and the upper value (15% of cross talk per tile) used to extract the uncertainty.

Below the distributions, the ratio defined in Eq. 4.6 is plotted, and this ratio gives an estimate of the uncertainty associated to the tuning of the cross talk parameter.

Similar plots have been done in order to estimate the uncertainty associated to the tuning of the number of effective pixels, needed to correct for saturation correction, and the tuning of the additional material used in order to simulate the upstream material in the beam line. These can be seen respectively in Figure 4.16 and in Figure 4.17.

Comparing Figure 4.15, Figure 4.16 and Figure 4.17 it can be seen that the largest contribution to the uncertainties on the simulation is due to the combination of the tuning of the optical cross talk and the tuning of the additional material.

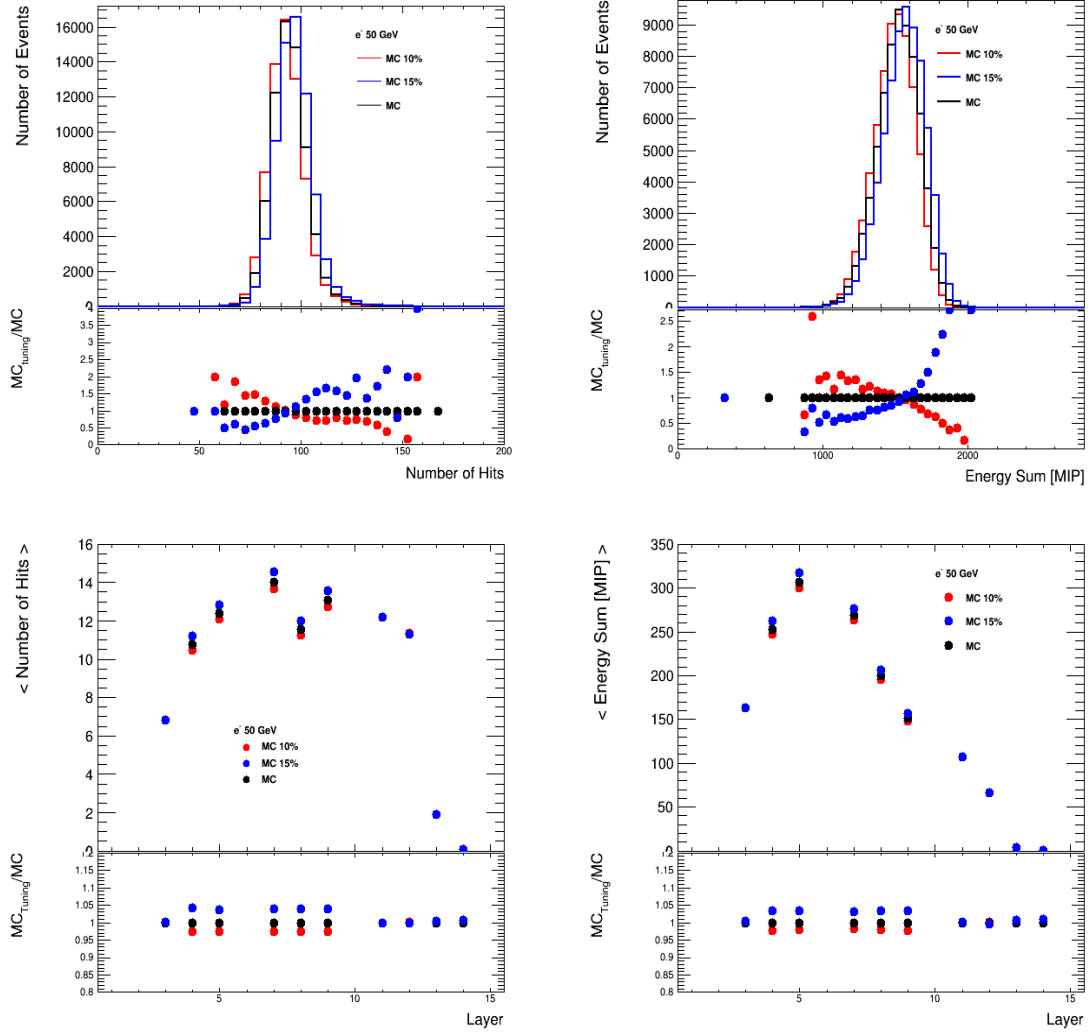


FIGURE 4.15.: Top plots: number of hit distribution (left) and energy sum distribution (right) in black for the standard analysis, in red using a value of 10% cross talk per tile and in blue using 15% cross talk per tile. Bottom plots: mean number of hits per layer (left) and mean energy sum per layer (right) in black for the standard analysis, in red using a value of 10% cross talk per tile and in blue using 15% cross talk per tile. Below the ratio $(\frac{MC_{tuning}}{MC})^i$ defined in Eq.4.6 is shown. Only the layers from 4 to 9, with tiles not individually wrapped are affected.

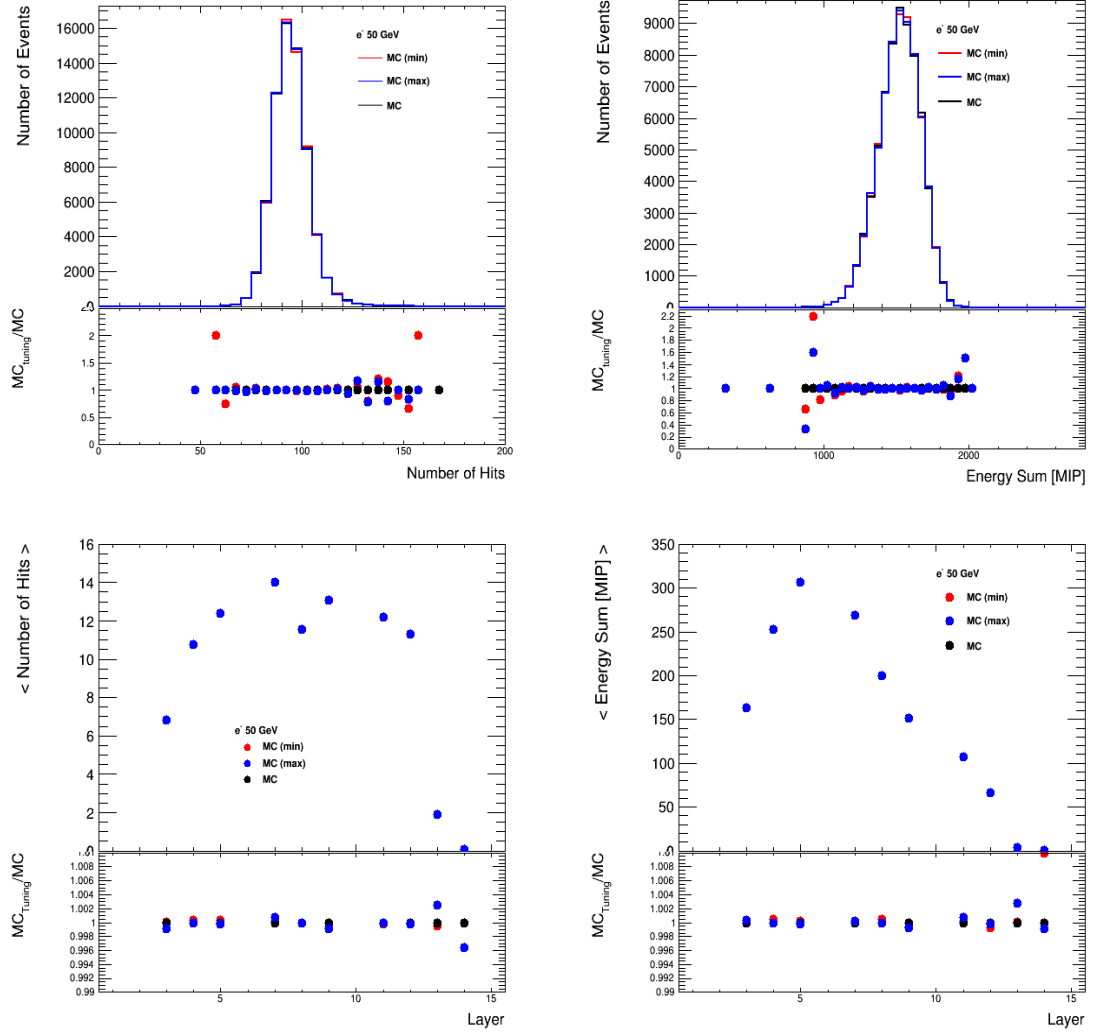


FIGURE 4.16.: Top plots: number of hit distribution (left) and energy sum distribution (right) in black for the standard analysis, in red using the lower value and in blue using the upper value, both listed in Table 4.4. Bottom plots: mean number of hit per layer (left) and mean energy sum per layer (right) in black for the standard analysis, in red using the lower value and the in blue using the upper value, listed in Table 4.4. Below the ratio $(\frac{MC_{tuning}}{MC})^i$ defined in Eq.4.6 is shown.

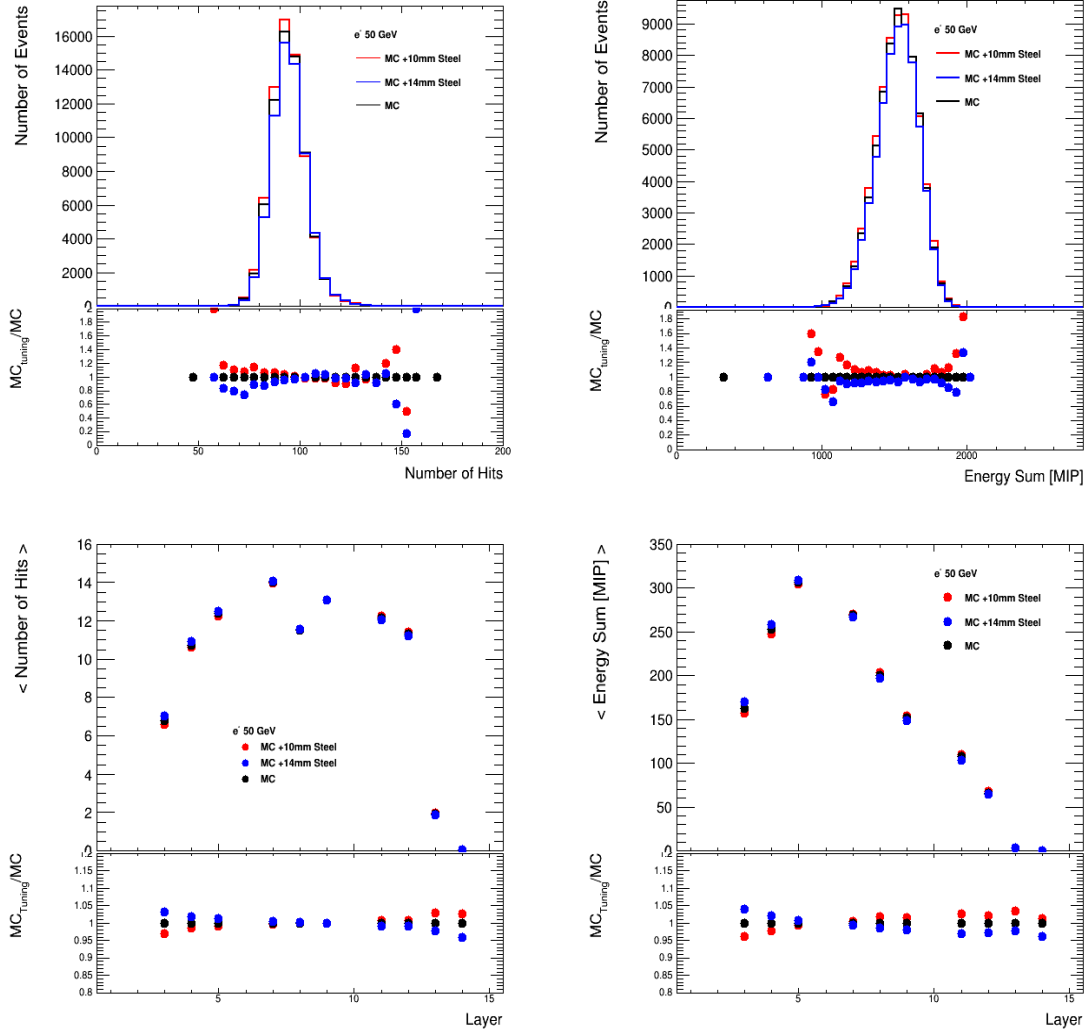


FIGURE 4.17.: Top plots: number of hit distribution (left) and energy sum distribution (right) in black for the standard analysis, in red using 10 mm of steel more before ECAL and in blue using 14 mm of steel more before ECAL. Bottom plots: mean number of hit per layer (left) and mean energy sum per layer (right) in black for the standard analysis, in red using 10 mm of steel more before ECAL and in blue using 14 mm of steel more before ECAL. Below the ratio $(\frac{MC_{tuning}}{MC})^i$ defined in Eq.4.6 is shown.

4.3.2. Uncertainties on the Data

In addition to the uncertainty due to the tuning of the number of effective pixels needed for the saturation corrections, computed as it has been described for the simulation, for the data the uncertainties due to the several calibration steps performed have to be taken into account as well.

4.3.2.1. Gain Uncertainty

As it has been described in Section 3.5.1, the gain calibration is performed using LED data, taken in different days during the test beam period. In order to estimate the effect of the gain calibration procedure on the analysis, two different LED data sets of gain have been considered. These are the gain extracted from the LED scan performed the 4th of July (G_1), and the gain extracted from the LED scan performed the 5th of July (G_2). The ratio G_1/G_2 between the gain obtained for each channels using the two different data sets has been considered and the distribution has been plotted, see Figure 4.18. The analysis has then been run using a gain value shifted by the RMS of the distribution, that means

$$Gain_{new} = Gain * (1 + RMS) \quad (4.7)$$

where $Gain_{new}$ is the new gain value used in the reconstruction process and $Gain$ is the gain used in the standard analysis.

The uncertainty is then obtained by computing the difference between the standard analysis and the analysis using $Gain_{new}$ (see Eq. 4.8)

$$\Delta G = \frac{Analysis(G) - Analysis(G_{new})}{2}. \quad (4.8)$$

4.3.2.2. MIP Uncertainty

The effect of the MIP calibration procedure used to convert the raw energy, in ADC, to MIP, has been evaluated using the muon beam runs. They are divided in two groups, one containing only data with an even run number, and the second one containing only the data with an odd run number.

The calibration constants used in the reconstruction of the data are applied to these two different sub samples. The MIP spectra are then fitted (if the detector is well calibrated it should peak at 1 MIP), and the distribution of the most probable value for each sub sample is obtained.

The two distributions are fitted with a Gaussian function and the sigma of the fit is used to shift the energy of each hit in the calorimeter as

$$E_{new} = (1 + \sigma_{MIP}) * E \quad (4.9)$$

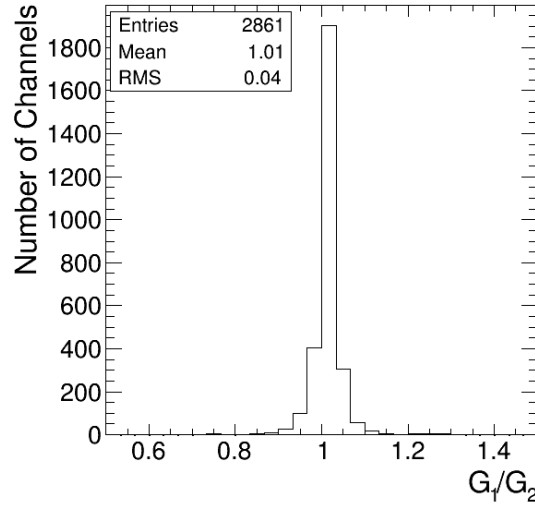


FIGURE 4.18.: *Distribution of the ratio between the gain obtained using two different LED data sets.*

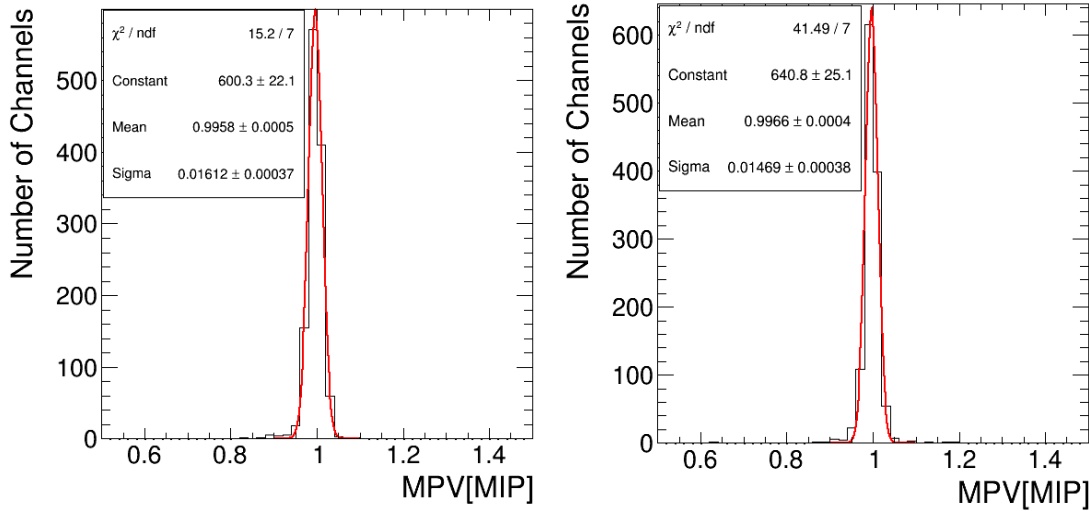


FIGURE 4.19.: *Most probable value distributions with in red the Gaussian fit. On the left, distribution obtained from the even runs, with $\mu = 0.996$ and $\sigma = 0.016$; on the right distribution obtained from the odd runs, with $\mu = 0.997$ and $\sigma = 0.015$*

where E_{new} represents the shifted quantity, $\sigma_{MIP} = \frac{\sigma_1 + \sigma_2}{2}$ is the mean of the two sigma obtained from the two Gaussian fits and E is the energy of the hit.

Finally the uncertainty associated to the MIP calibration is obtained considering half of the difference between the standard analysis and the analysis obtained shifting the energy of each hit, as indicated from E_{new} (similar to Eq. 4.8). The distributions of the most probable value, one for each sub sample are shown in Figure 4.19.

4.3.2.3. High Gain - Low Gain Intercalibration Uncertainty

As explained in Section 3.5.3, two different methods have been used within the CALICE Collaboration, in order to extract the intercalibration constants, one using LED data and the other one using electron data. Only for the channels for which the intercalibration constant could be extracted with both methods, the ratio R between the intercalibration constant obtained with LED data (IC_{LED}) and the intercalibration constant obtained from electrons data (IC_{e^-}) can be determined. This distribution is plotted in Figure 3.20(b).

In this case the uncertainty σ_{IC} is defined as

$$\sigma_{IC} = 1 - \mu = 0.15$$

where μ represents the mean of the distribution shown in Figure 3.20(b).

The data are then reconstructed, shifting the intercalibration constants by a value equal to

$$IC_{new} = IC * (1 + \sigma_{IC}) \quad (4.10)$$

and the uncertainty associated is computed as for the gain and MIP calibrations, that means as half of the difference of the standard analysis and the analysis obtained using the intercalibration constants defined from IC_{new} (similar to Eq. 4.8).

In order to quantify the uncertainties on the data, a plot similar to Figure 4.16 has been obtained for the different sources of uncertainties. Here this is shown for 50 GeV electrons, representing the highest value these uncertainties can reach. Figure 4.20 represents the uncertainty related to the gain calibration, Figure 4.21 the uncertainty related to the MIP calibration, Figure 4.22 the high gain-low gain intercalibration and Figure 4.23 the uncertainty due to saturation.

The σ_{IC} associated to the high gain - low gain intercalibration corresponds to a value of $\sim 15\%$, while the σ_{Gain} associated to the gain and the σ_{Mip} associated to the MIP calibration correspond respectively to $\sim 4\%$ and $\sim 1.5\%$. Therefore, it is expected that the uncertainty related to the high gain - low gain intercalibration is the dominant one. It can reach a value up to 30% on the energy sum distribution. This is confirmed comparing figures 4.20, 4.21 and 4.22. The saturation uncertainty also plays an important role with increasing beam energy. It can reach a value up to 20% as shown in Figure 4.23.

4.4. Results

In order to obtain the total uncertainty, the uncertainties obtained in the previous section are added in quadrature and applied respectively to the data and the

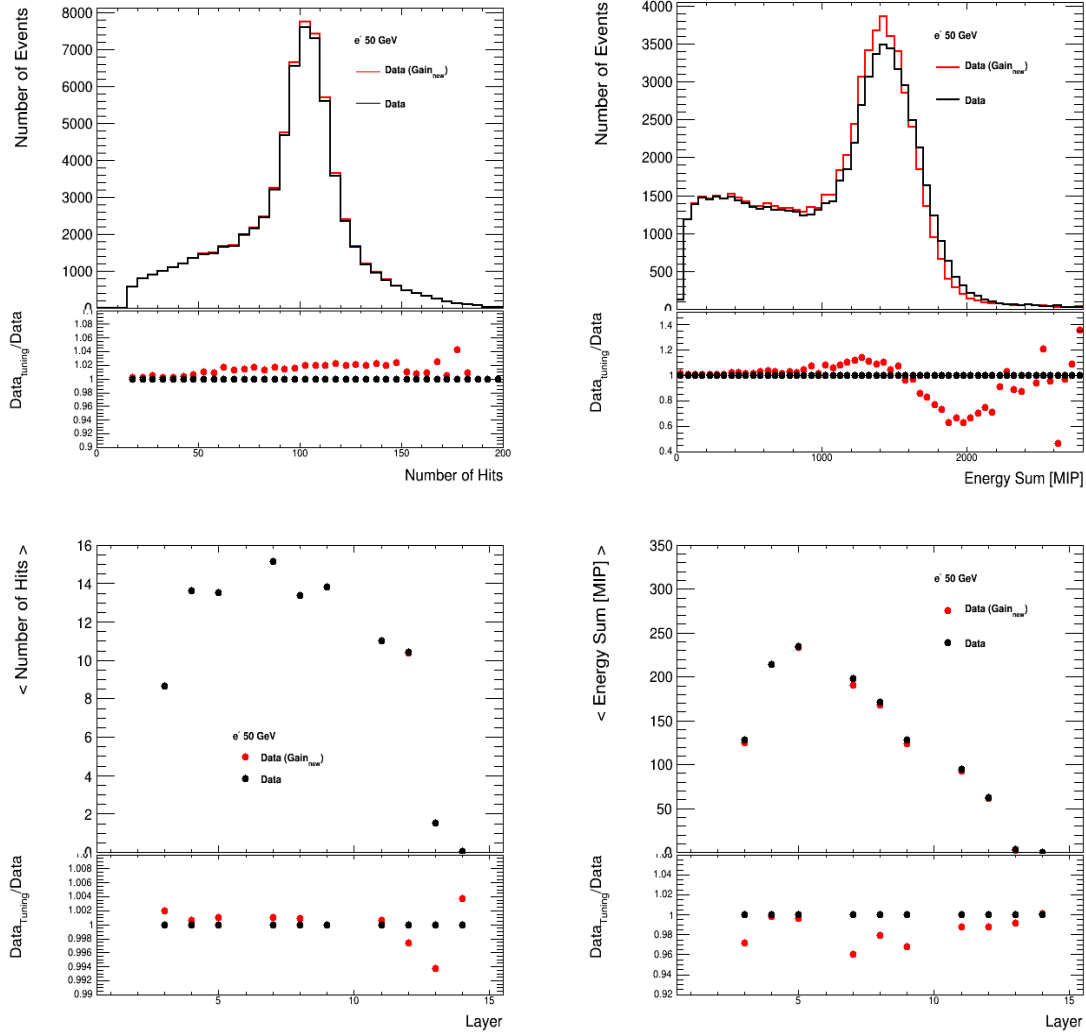


FIGURE 4.20.: Top plots: number of hit distribution (left) and energy sum distribution (right) in black for the standard analysis, in red using $Gain_{new}$ defined in 4.3.2.1. Bottom plots: mean number of hit per layer (left) and mean energy sum per layer (right) in black for the standard analysis, in red using $Gain_{new}$ defined in 4.3.2.1. Below the ratio $(\frac{Data_{tuning}}{Data})^i$, defined in a similar way as defined in Eq.4.6 for the simulation, is shown.

simulation. It is now possible to make a comparison between data and simulation. Here the comparison for all the energies is shown.

In Figure 4.24, is possible to see the comparison for the 10 GeV electrons. In Figure 4.24(a) the comparison of the distribution of the number of hits per event between data (in black) and simulation (in red) is shown, while in Figure 4.24(b) the distribution of the energy sum per event for the same beam energy is presented. The distributions have been normalized to the maximum value of the distribution.

As mentioned before, the dominant uncertainty on the data is due to the high gain - low gain intercalibration. This effect is visible looking at the energy sum distribution, since it has no effect on the number of hits in the detector. In the

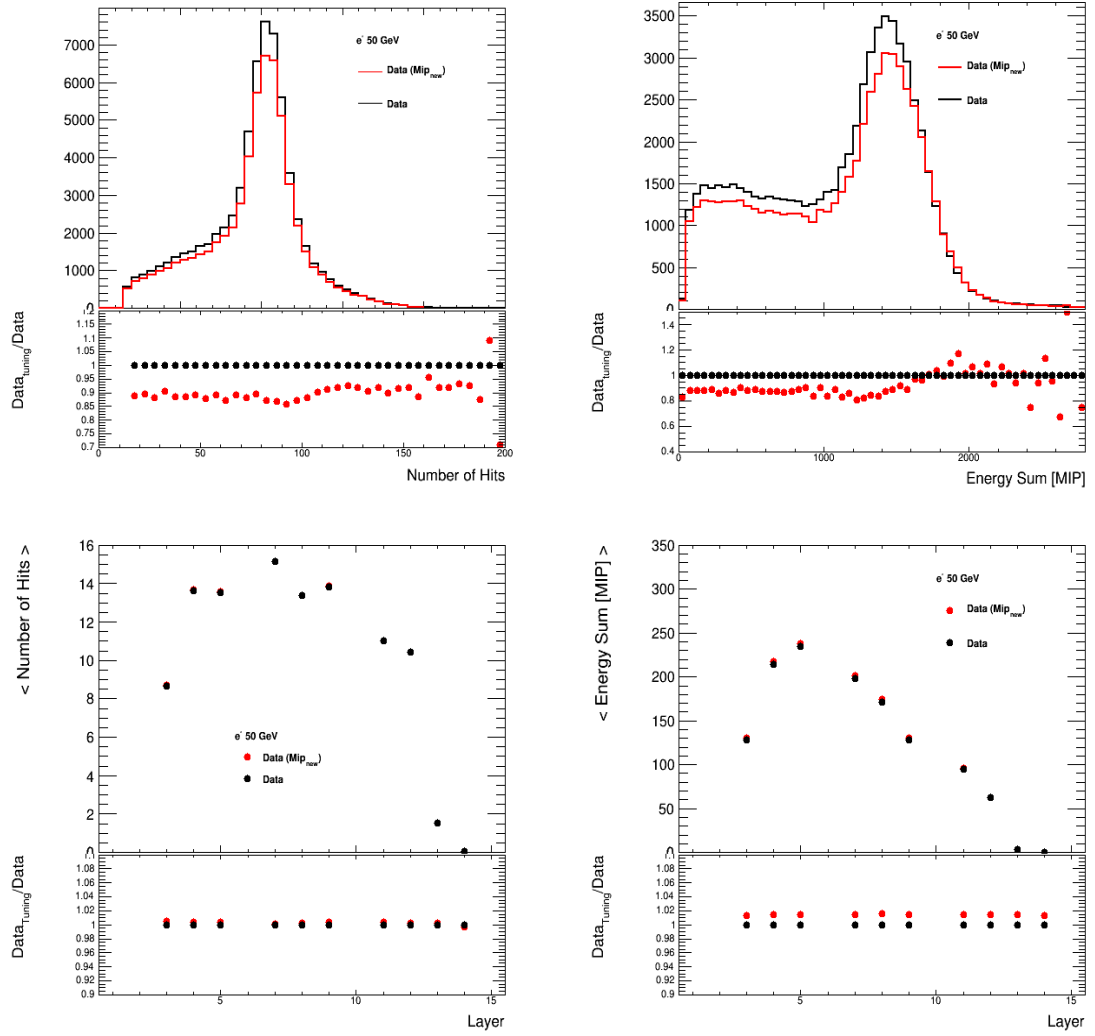


FIGURE 4.21.: Top plots: number of hit distribution (left) and energy sum distribution (right) in black for the standard analysis, in red using MIP_{new} defined in 4.3.2.2. Bottom plots: mean number of hit per layer (left) and mean energy sum per layer (right) in black for the standard analysis, in red using MIP_{new} defined in 4.3.2.2. Below the ratio $(\frac{Data_{tuning}}{Data})^i$, defined in a similar way as defined in Eq.4.6 for the simulation, is shown.

bottom plots in Figure 4.24, the shower profiles are shown, on the left (Figure 4.24(c)) considering the mean of the number of hits as a function of the AHCAL layer position, and on the right (Figure 4.24(d)) considering the mean of the energy sum per each layer.

As shown in the Figure 4.24(c), the uncertainties on the data are smaller than the uncertainties on the simulation. This is because the uncertainties on the data, associated mostly to the calibration of the detectors, influence the energy, and not much the number of hits per event. This might slightly vary due to the event selection cuts relevant to the energy, as shown in Figure 4.24(d).

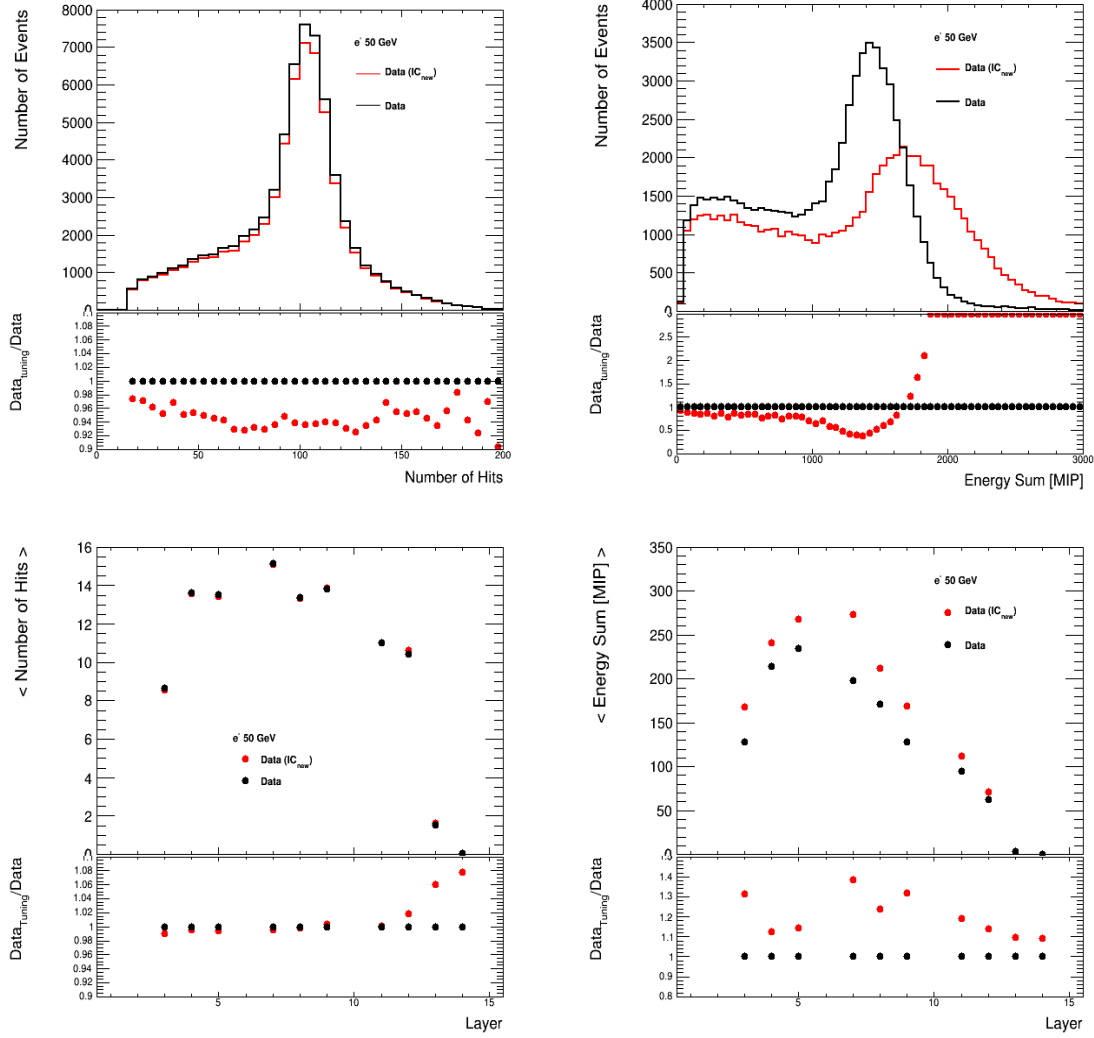


FIGURE 4.22.: Top plots: number of hit distribution (left) and energy sum distribution (right) in black for the standard analysis, in red using IC_{new} defined in 4.3.2.3. Bottom plots: mean number of hit per layer (left) and mean energy sum per layer (right) in black for the standard analysis, in red using IC_{new} defined in 4.3.2.3. Below the ratio $(\frac{Data_{tuning}}{Data})^i$, defined in a similar way as defined in Eq.4.6 for the simulation, is shown.

For the simulation the opposite effect is observed. The most important uncertainty is the tuning of the cross talk, which influences the number of hits per event, but not the energy, as shown in Figure 4.24(c). In addition, the uncertainty is more pronounced for the central modules (from 4 to 9) because the cross talk effect affects only these modules, as shown in Figure 4.24(c).

In principle also the uncertainty associated to the tuning of the number of effective pixels needed to correct for saturation effects plays an important role, but this is not dominant at 10 GeV.

The same plot for 15 GeV, 20 GeV, 30 GeV, 40 GeV and 50 GeV are shown respectively in Figure 4.25, 4.26, 4.27, 4.28 and 4.29.

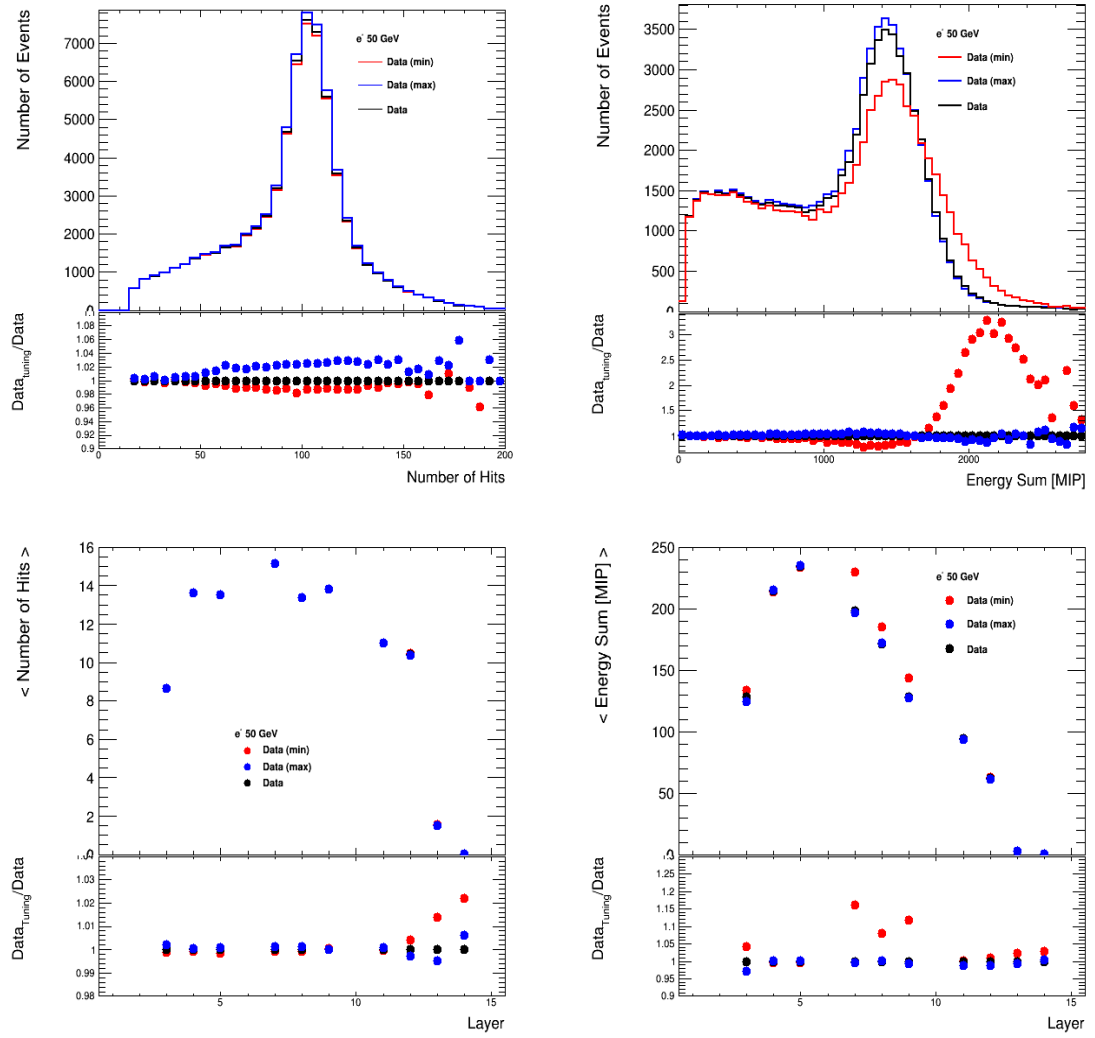


FIGURE 4.23.: Top plots: number of hit distribution (left) and energy sum distribution (right) in black for the standard analysis, in red using the lower value and in blue using the upper value, both listed in Table 4.4. Bottom plots: mean number of hit per layer (left) and mean energy sum per layer (right) in black for the standard analysis, in red using the lower value listed in Table 4.4 and in blue using the upper value listed in Table 4.4. Below the ratio $(\frac{Data_{tuning}}{Data})^i$, defined in a similar way as defined in Eq.4.6 for the simulation, is shown.

Looking at these figures, a similar behaviour to the 10 GeV electrons can be seen. However it is possible to notice that the agreement between data and simulation get worse. This is due to the left low energy tail, due probably to background, that increase with the energy, and it was not possible to completely remove it, because of the limited information. This effect can be seen comparing Figure 4.25(b), 4.26(b), 4.27(b), 4.28(b) and 4.29(b) respectively for 15 GeV, 20 GeV, 30 GeV, 40 GeV and 50 GeV.

In addition there is also a high increase of the uncertainties on the data for the energy sum per event distribution. This is due to the uncertainty related to the

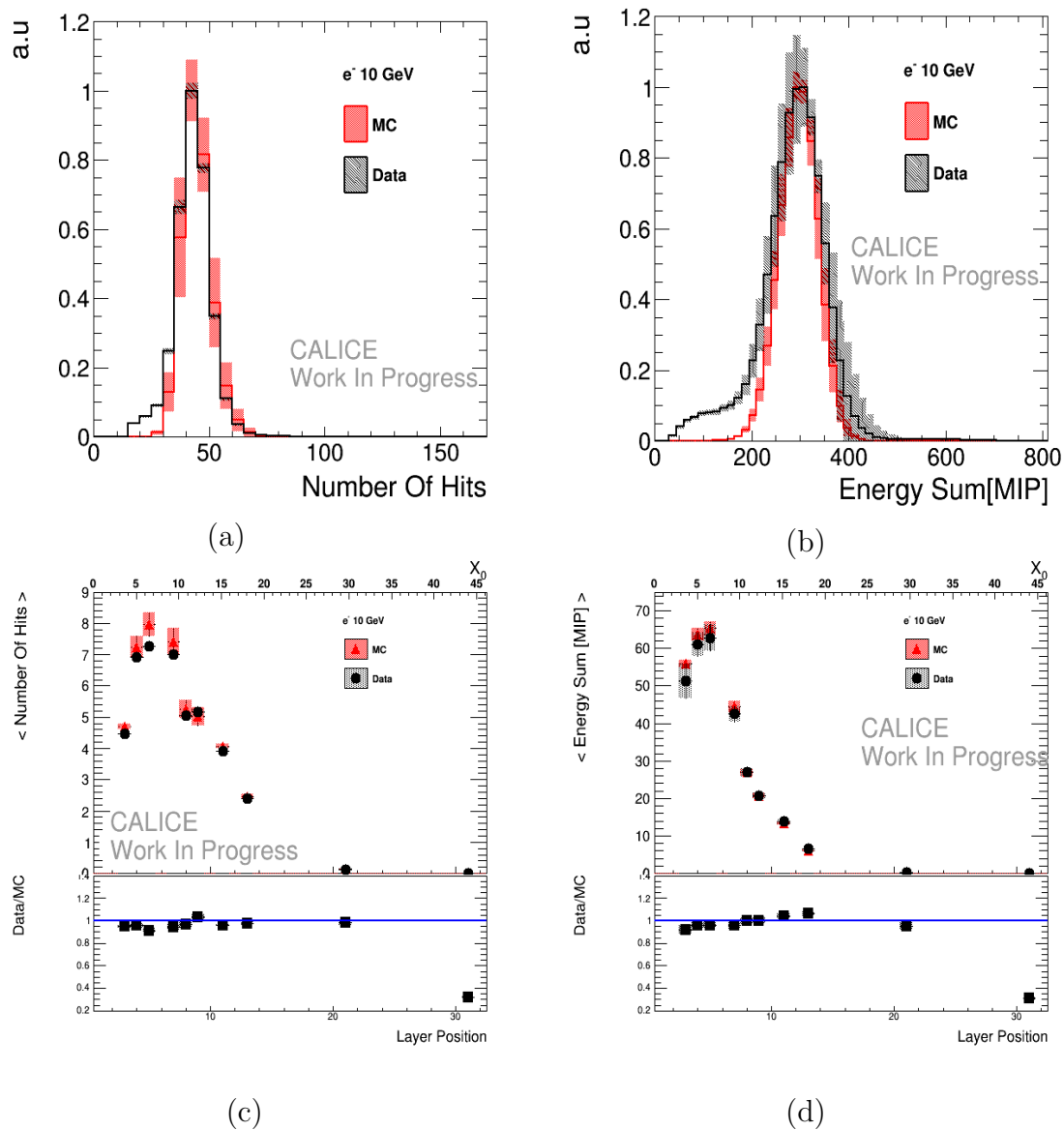


FIGURE 4.24.: Top plots: distributions of the number of hits per event (a) and of the energy sum per event (b) for 10 GeV electrons. The distributions have been normalized to their maximum. Bottom plots: mean number of hits as a function of the AHCAL layer position (c) and mean energy sum as a function of the AHCAL layer position (d) for 10 GeV electrons. Black represents the data and red simulation.

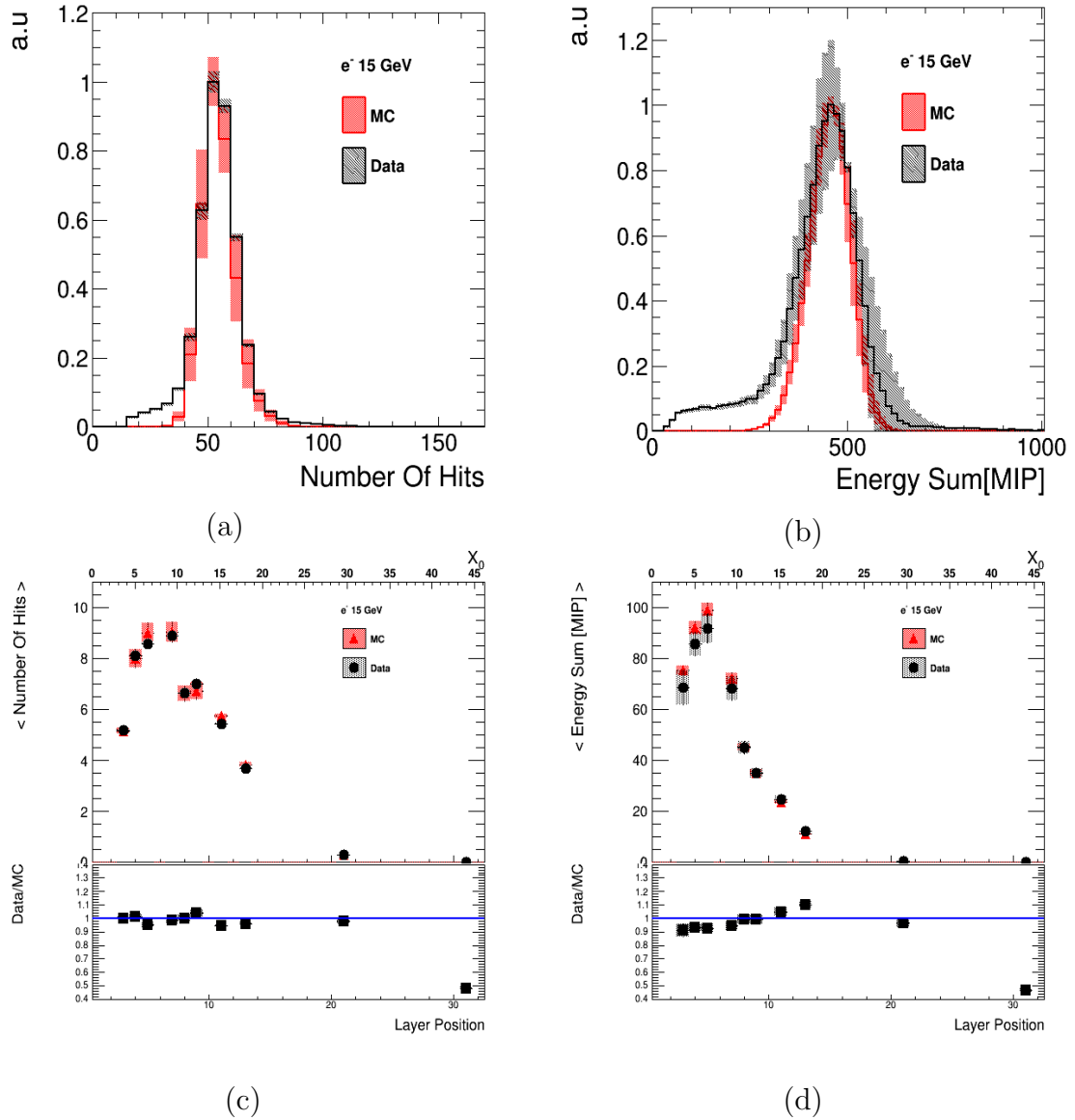


FIGURE 4.25.: Top plots: distributions of the number of hits per event (a) and of the energy sum per event (b) for 15 GeV electrons. The distributions have been normalized to their maximum. Bottom plots: mean number of hits as a function of the AHCAL layer position (c) and mean energy sum as a function of the AHCAL layer position (d) for 15 GeV electrons. Black represents the data and red simulation.

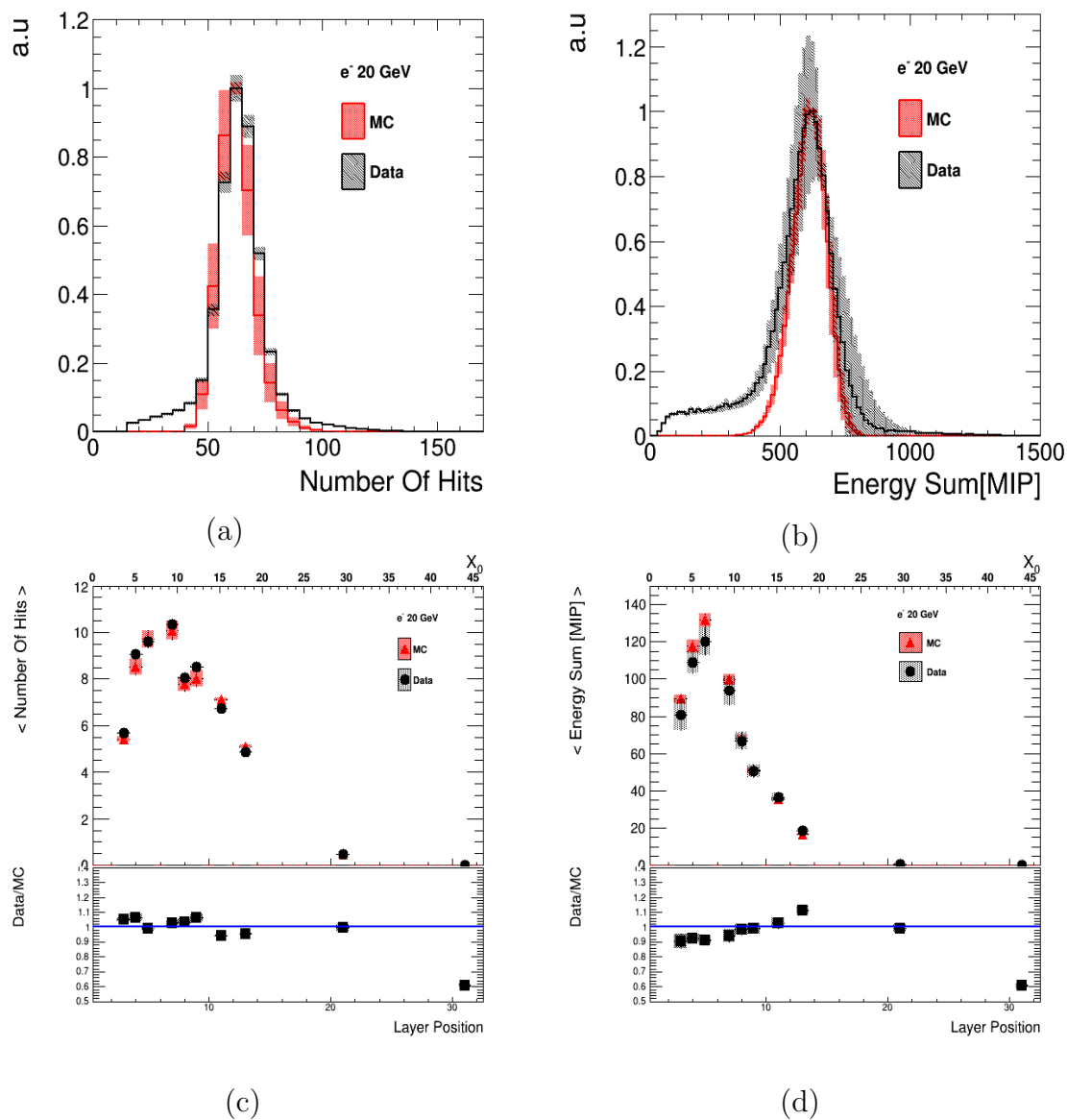


FIGURE 4.26.: Top plots: distributions of the number of hits per event (a) and of the energy sum per event (b) for 20 GeV electrons. The distributions have been normalized to their maximum. Bottom plots: mean number of hits as a function of the AHICAL layer position (c) and mean energy sum as a function of the AHICAL layer position (d) for 20 GeV electrons. Black represents the data and red simulation.

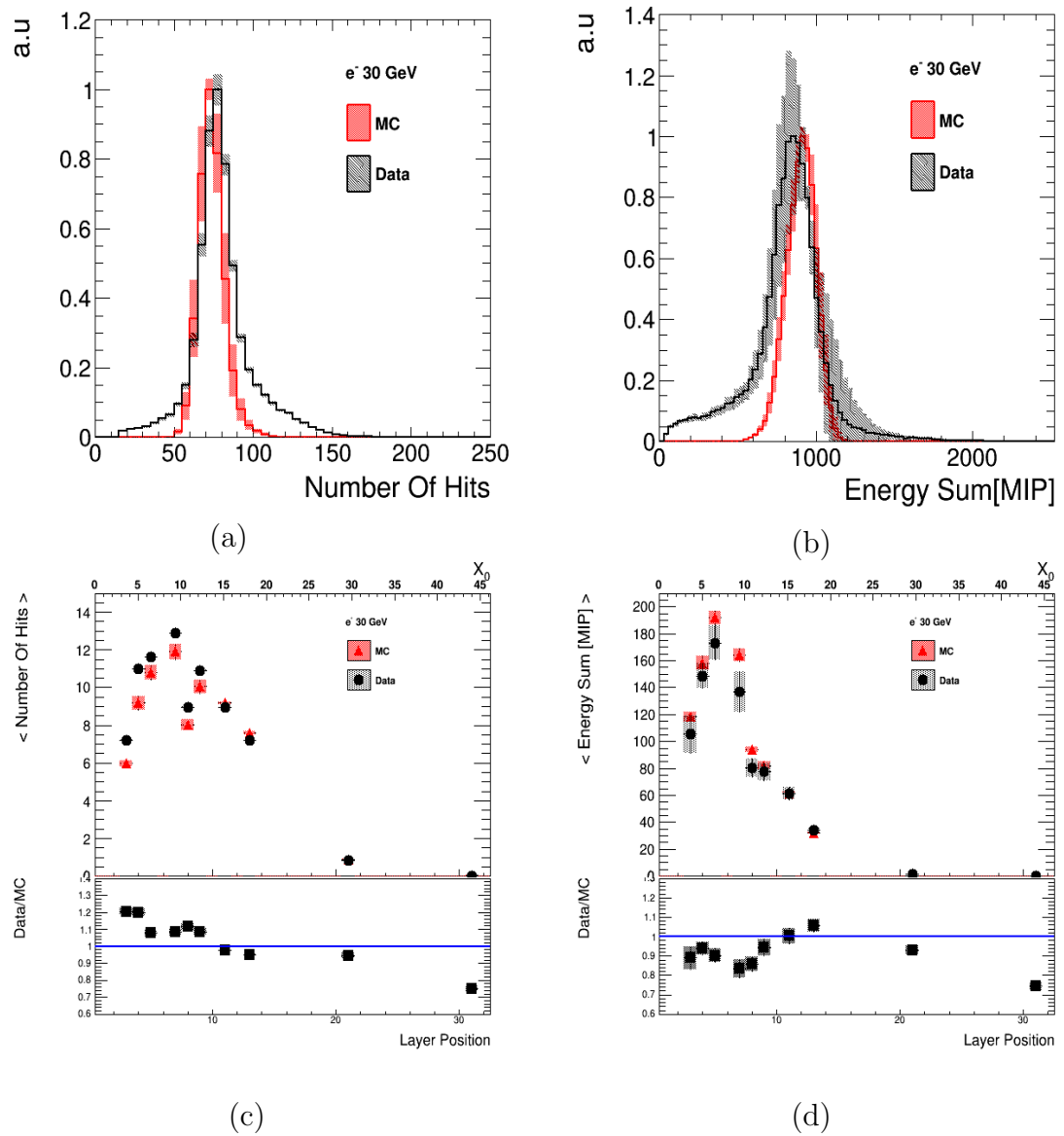


FIGURE 4.27.: Top plots: distributions of the number of hits per event (a) and of the energy sum per event (b) for 30 GeV electrons. The distributions have been normalized to their maximum. Bottom plots: mean number of hits as a function of the AHCAL layer position (c) and mean energy sum as a function of the AHCAL layer position (d) for 30 GeV electrons. Black represents the data and red simulation.

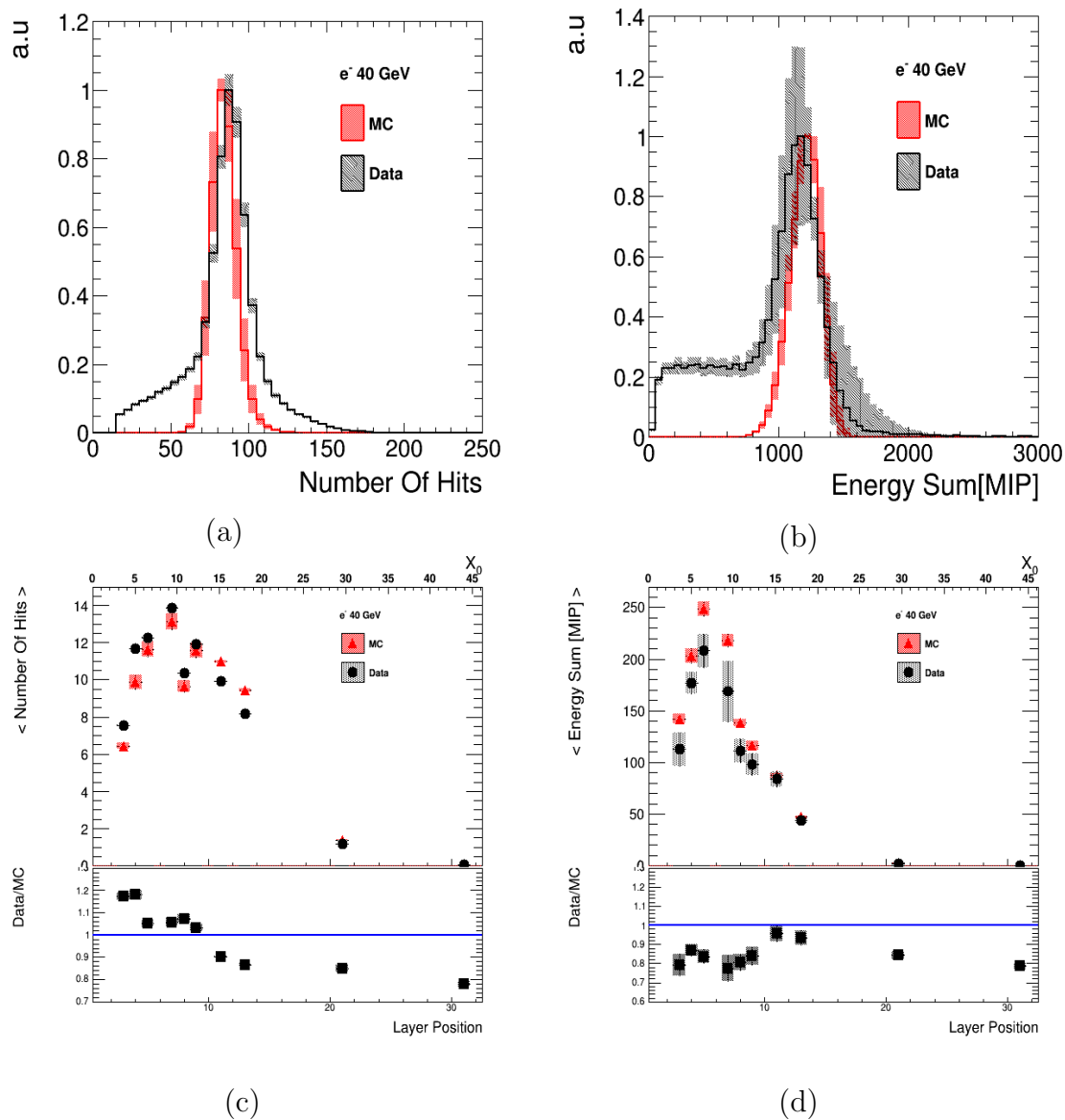


FIGURE 4.28.: Top plots: distributions of the number of hits per event (a) and of the energy sum per event (b) for 40 GeV electrons. The distributions have been normalized to their maximum. Bottom plots: mean number of hits as a function of the AHICAL layer position (c) and mean energy sum as a function of the AHICAL layer position (d) for 40 GeV electrons. Black represents the data and red simulation.

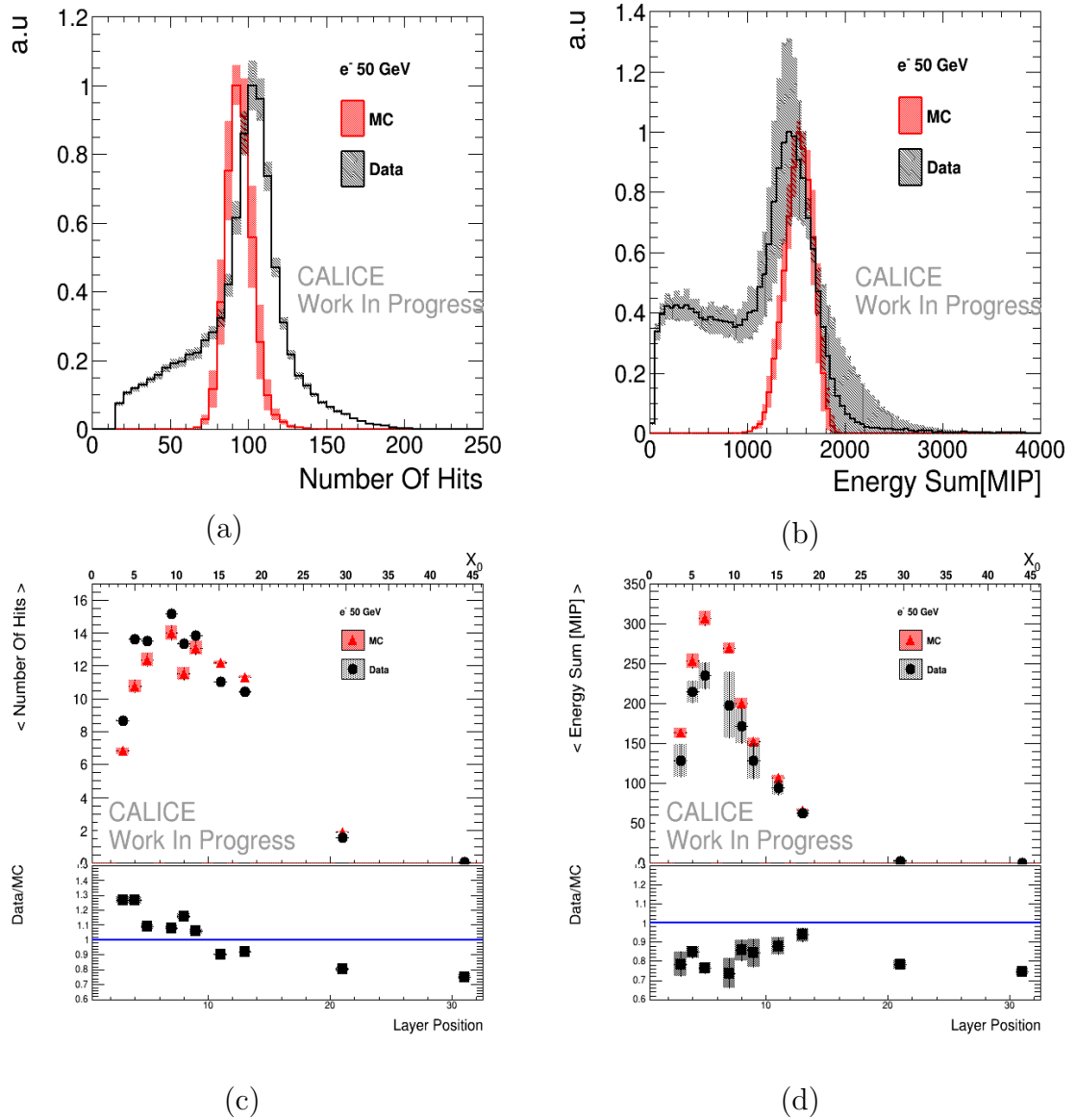


FIGURE 4.29.: Top plots: distributions of the number of hits per event (a) and of the energy sum per event (b) for 50 GeV electrons. The distributions have been normalized to their maximum. Bottom plots: mean number of hits as a function of the AHCAL layer position (c) and mean energy sum as a function of the AHCAL layer position (d) for 50 GeV electrons. Black represents the data and red simulation.

saturation correction that starts to be dominant, together with the high gain - low gain intercalibration at higher energies, because the shower becomes denser. This can also be seen in the mean number of hits per event as a function of the layer and at the mean of the energy sum per event as a function of the layer, respectively shown in figure 4.25(c) and 4.25(d) for 15 GeV, in figures 4.26(c) and 4.26(d) for 20 GeV, in figures 4.27(c) and 4.27(d) for 30 GeV, 4.28(c) and 4.28(d) for 40 GeV and in figures 4.29(c) and 4.29(d) for 50 GeV.

It is important to point out that, in performing the shower profiles, a cut on the number of hits per event has been applied, in order to be less sensitive to the left tail of the distributions, like in Figure 4.29(a).

The cut applied for the different beam energies is listed in Table 4.5.

Energy	Hit Range
10 GeV	(25, 80)
15 GeV	(25, 100)
20 GeV	(25, 125)
30 GeV	(25, 145)
40 GeV	(40, 150)
50 GeV	(60, 150)

TABLE 4.5.: *Range of number of hits per event considered to perform the shower profiles.*

A summary of the systematic uncertainties applied and their values is shown in Table 4.6 for simulation and in Table 4.7 for data.

Source Of Uncertainty	Method	Variable	Value (min-max)
Saturation	according to Table 4.4 and Eq.4.5	Number Of Hits	1% – 3%
		Energy Sum	2% – 4%
		Mean Number Of Hits	‰
		Mean Energy Sum	‰
Cross Talk	10% – 15%, according to Eq.4.5	Number Of Hits	10% – 30%
		Energy Sum	20% – 30%
		Mean Number Of Hits	2% – 5%
		Mean Energy Sum	1.5% – 4%
Upstream Material	$\pm 2\text{mm}$, according to Eq.4.5	Number Of Hits	5% – 15%
		Energy Sum	4% – 10%
		Mean Number Of Hits	1% – 4%
		Mean Energy Sum	3% – 5%

TABLE 4.6.: Overview of the uncertainties applied to the simulation. Details about how the uncertainties are extracted can be found in Section 4.3. Here a lower and upper value for these uncertainties are given, according to Figures 4.15, 4.16, 4.17 .

Source Of Uncertainty	Method	Variable	Value (min-max)
Saturation	according to Table 4.4 and Eq.4.5	Number Of Hits	‰–3%
		Energy Sum	1% – 20%
		Mean Number Of Hits	‰
		Mean Energy Sum	1.5% – 15%
Gain Calibration	according to Eq.4.7 and Eq.4.8	Number Of Hits	‰–2%
		Energy Sum	2% – 5%
		Mean Number Of Hits	‰
		Mean Energy Sum	0.2% – 4%
MIP Calibration	according to Eq.4.9 and Eq.4.8	Number Of Hits	2% – 8%
		Energy Sum	5% – 8%
		Mean Number Of Hits	‰
		Mean Energy Sum	1% – 2%
HG-LG Intercalibration	according to Eq.4.10 and Eq.4.8	Number Of Hits	2% – 10%
		Energy Sum	10% – 30%
		Mean Number Of Hits	‰–1%
		Mean Energy Sum	10% – 30%

TABLE 4.7.: Overview of the uncertainties applied to the data. Details about how the uncertainties are extracted can be found in Section 4.3. Here a lower and upper value for these uncertainties are given, according to Figures 4.20, 4.21, 4.22, 4.23.

4.5. Conclusion

In this chapter the analysis of the electron beam data taken during the test beam campaign at CERN in July 2015 has been illustrated.

This work has been performed in order to tune the main parameters in the Monte Carlo simulation. Therefore the upstream material present in the beam line at CERN, the value of the optical cross talk per tile for the unwrapped tiles and the number of effective pixels needed to correct for the saturation effect have been tuned.

The systematic uncertainties have been studied and added both for data and for simulation. Finally a comparison between data and simulation of the characteristic quantities of the electromagnetic shower has been performed. A precision within 10% and 30% has been reached. Since this work has been performed in order to tune the main parameters of the simulation, to then proceed with the analysis of the pion data taken during the same test beam campaign, the precision reached is considered adequate, in order to proceed with the study of the hadronic showers.

5. Analysis of Pion Beam Data

As already mentioned, the core of this thesis is the analysis of the pion data taken during the CERN test beam campaign. A very important and new feature of this prototype is in fact the possibility to store the time information of the hit in the detector.

The study of the time structure of the hadronic shower has already been performed. The T3B experiment [105], using a setup made of 15 small scintillating tiles readout with SiPMs, investigated the distribution of the measured time of the hit and the radial shower timing profile. In this way the two different components of the hadronic shower could be identified, meaning the prompt component due to relativistic hadrons and to the electromagnetic sub-showers and the delayed component mainly due to neutrons. The same study has been performed with the AHCAL technological prototype that can provide the time information for each cell of the detector [99]. The new aspect of the analysis presented in this chapter is the possibility of including the energy information of the hit in the study of the time development of the hadronic shower. Thanks to the electron analysis illustrated in Chapter 4, a properly calibrated energy and a well tuned simulation are available. Therefore it is possible to study the correlation between the energy and the time of each hit in the detector. This is a good tool for a better identification of the different components of the hadronic shower, thus providing a better understanding of such a complex object. This can also potentially be a powerful tool for improving, in the future, the jet energy resolution [106].

For this purpose, pion data from 10 GeV to 90 GeV were collected during the test beam campaign at CERN. This thesis focuses on the lowest (10 GeV) and highest (90 GeV) pion energy in order to understand the structure of the hadronic shower. In addition to that, an intermediate beam energy (50 GeV) is used in order to confirm the behaviour observed.

In this chapter the analysis of the pion data is presented. First of all a description of the event selection with a particular stress on the start finder algorithm of the shower is presented. Then, the amplitude analysis is shown, similar to the analysis of the electron data performed in Chapter 4. At the end, the study of the time of the hit in the detector correlated with its energy information is described.

5.1. Event Selection & the Shower Start Finder

The pion beam at CERN is not a pure beam, therefore a simple event selection has to be applied, in order to reject contamination due to other particles such as electrons and muons.

As it has been done for electron data, a pre-selection using the beam instrumentation is performed. As discussed in Chapter 3.4, the pressure of the Cherenkov detector installed in the beam line could be configured in order to select the desired particles. Therefore a signal in the Cherenkov detector is required. Then, due to the presence of a trigger scintillator before the prototype, of the dimension of $10 \times 10 \text{ cm}^2$, a cut on the center of gravity of the shower is applied. The center of gravity is required to be within the dimension of the trigger scintillators. The cuts on the center of gravity applied according to the beam energies used in this thesis (10 GeV, 50 GeV, 90 GeV) are listed in Table 5.1. The reason why this cut is not the same for all the energies is that the beam was not always pointing exactly through the center of the detector. In Figure 5.1, a map of the center of gravity in the X and Y coordinates is shown for 10 GeV pions. The cut on the center of gravity requires the beam to hit roughly the central twelve scintillating tiles, being one tile $30 \times 30 \text{ mm}^2$.

Energy	COG _X [mm]	COG _Y [mm]
10GeV	(-45, 45)	(-30, 60)
50GeV	(-45, 45)	(-45, 45)
90GeV	(-45, 45)	(-40, 50)

TABLE 5.1.: Center of gravity range considered for pion data, according to the beam energy.

The pion selection is performed in order to avoid muons, punch-through pions and electrons contamination. Muons and punch-through pions are rejected requiring to have more than 15 hits per event in the detector. A first cut in order to reject electrons is applied: the energy sum per event in the last two layers of the prototype (Figure 3.7) is required to be higher than 1%. Pions are in fact expected to behave like a MIP particle before starting to shower, and lose most of their energy when they start to shower. Because of that, in order to better identify the single pions and suppress the electrons contamination, a *Shower Start Finder Algorithm* has been developed within the CALICE collaboration. In this analysis, only pions starting to shower in the single HBUs of the prototype (layers from 3 to 10 in Figure 3.7) are considered, because this corresponds to the fully equipped part of the prototype. The aim of the algorithm is to identify the layer of the *first hard interaction* (FHI), meaning the layer where the first hard interaction between the primary hadron and a nucleus in the detector takes place, indicating the start of the hadronic cascade. Before the first hard interaction occurs, the

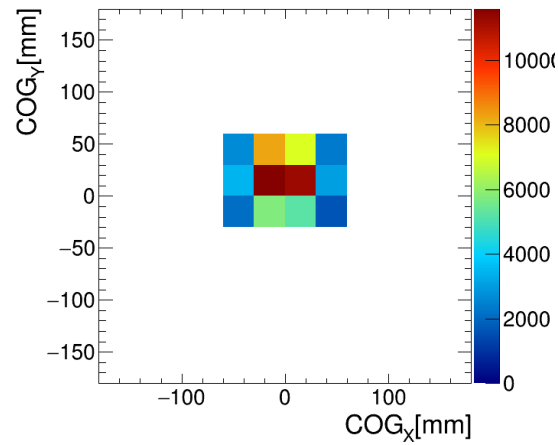


FIGURE 5.1.: Map of the center of gravity in the X and Y coordinates for 10 GeV pions. The cut on the center of gravity corresponds to require the beam to hit roughly the 12 central scintillating tiles of the prototype. Here one square corresponds to one tile

primary hadron typically behaves like a MIP particle, depositing only very small amounts of its energy. In the first hard interaction, typically several secondary particles are generated, causing an increase in the number of hits in the detector and in the deposited energy in the next active layer. The algorithm is based on identifying the layer where this increase in the number of hits happens. The work is based on a previous work realized within the CALICE Collaboration [107], and adjusted according to the layout and performance of the AHCAL technological prototype [99].

The reconstruction of the first hard interaction layer is based on two steps. First the number of hits per event in the same X and Y position is counted, in order to identify the track associated to the MIP-like behaviour of the hadron before it starts showering. This track is identified as it has been done for muons and as it is schematically shown in the sketch in Figure 3.21. The length of this track is required to be at least three, meaning it is required to have at least three hits in the same X and Y position in the detector. Then, in order to identify the layer where the first hard interaction happens, the number of hits in two consecutive layers is also taken into account. If the length of the track is at least 3 and the number of hits in two consecutive layers (layer i and $i+1$) is at least six, the layer of the first hard interaction is assumed to be between layer i and layer $i+1$. This can be understood looking at the event display of a showering 50 GeV pion, shown in Figure 5.2. In addition, in order to pinpoint the correct layer, the energy sums in layer i and in the following three layers ($i+1$, $i+2$ and $i+3$) are also compared. In simulated events, the true first hard interaction layer can be extracted from the detailed information available for each interaction. Each simulated interaction in GEANT 4 gives information about the type of interaction

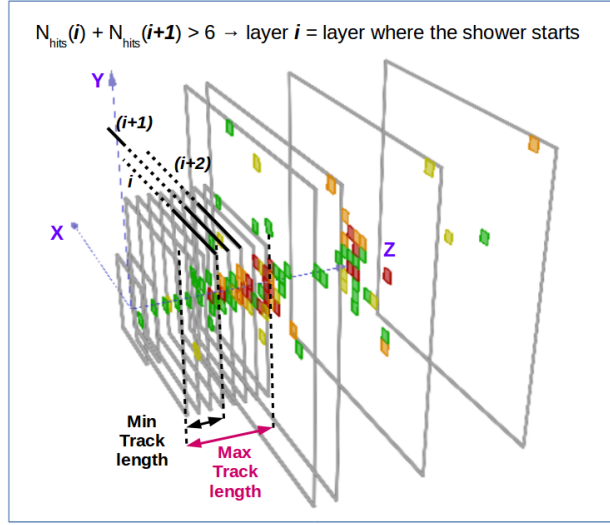


FIGURE 5.2.: Event display of a 50 GeV pion showering in the prototype. Here the minimum and the maximum track length required from the shower start finder algorithm are shown.

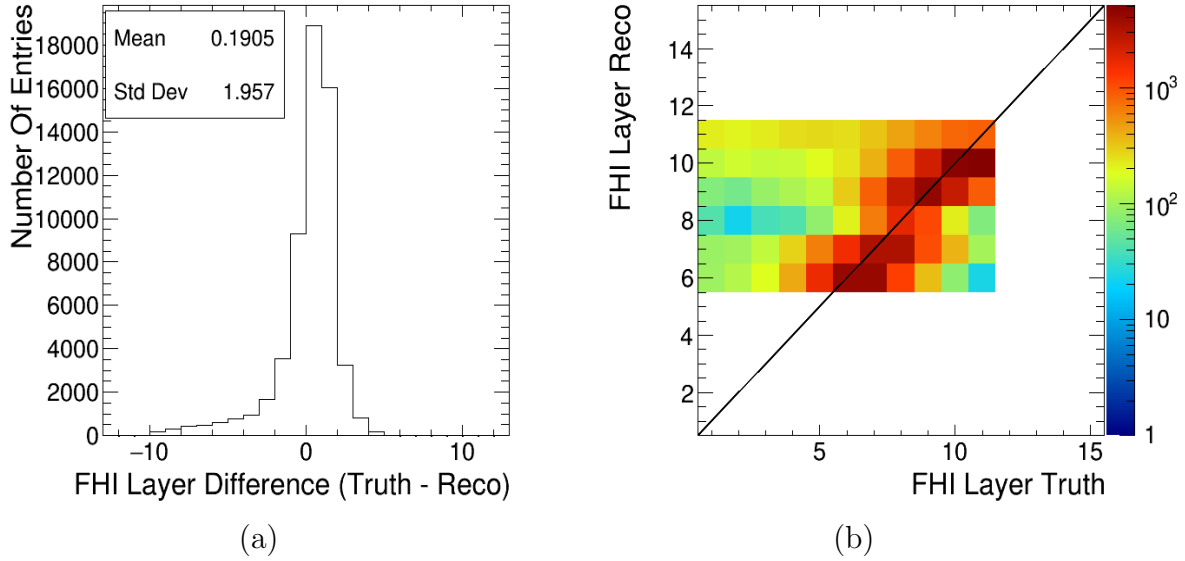


FIGURE 5.3.: Performance of the algorithm for the shower start reconstruction. In Figure a) the difference between the truth and the reconstructed layer where the shower starts is shown, in Figure b) the correlation between the same quantities is shown.

(hadronic, electromagnetic, decay, optical...) as well as the parameters of the initial and final particles before and after the interaction happens (particle ID, momentum, kinetic energy). The accuracy of the algorithm is evaluated comparing the reconstructed layer where the first hard interaction happens with the true Monte Carlo information about the end-point of the incident pion. This can be seen in Figure 5.3 for 50 GeV pions. Here, on the left, the difference between

the true first hard interaction layer and the reconstructed one is shown, while on the right the correlation between the true and the reconstructed is plotted. It is possible to confirm that in general the layer of the first hard interaction is well reconstructed, as shown in Figure 5.3 on the left, where the mean peaks almost at zero, and on the right, where it is possible to see that most of the events lie on the diagonal black line, indicating a good correlation between the two quantities, even though a fraction of events for which the layer of the first hard interaction is not properly reconstructed is still visible. In the following analysis a lower and upper limit on the track length are set, in order to reject electron contamination, punch through pions, and pions which start to shower after layer 10 (see sketch 3.7), since the prototype was not fully equipped after that layer. In Table 5.2, the efficiency of each step of the pion event selection described above is shown, both for pion data and for simulated muons, electrons and pions. The efficiency ε is defined as the ratio between the number of events passing the selection and the initial number of events. Here it is important to notice that the selection applied allows to remove

Energy	Type	$\varepsilon_{\text{Hits}}$	$\varepsilon_{\Sigma E}$	ε_{COG}	ε_{SSF}
10 GeV	π^- Data	67.2%	50.6%	37.1%	15.0%
	π^- QGSP_BERT_HP	60.5%	47.2%	20.1%	11.1%
	e^- QGSP_BERT_HP	99.0%	5.0%	2.3%	< 1‰
	μ^- QGSP_BERT_HP	< 1‰	< 1‰	< 1‰	/
50 GeV	π^- Data	82.2%	80.0%	74.9%	30.9%
	π^- QGSP_BERT_HP	82.3%	80.0%	66.5%	31.0%
	e^- QGSP_BERT_HP	100%	6.4%	4.7%	/
	μ^- QGSP_BERT_HP	< 1‰	< 1‰	< 1‰	/
90 GeV	π^- Data	84.0%	80.6%	72.5%	29.1%
	π^- QGSP_BERT_HP	86.2%	84.6%	69.1%	30.0%
	e^- QGSP_BERT_HP	100%	10.2%	8.6%	/
	μ^- QGSP_BERT_HP	< 1‰	< 1‰	< 1‰	/

TABLE 5.2.: Efficiency of each step of the pion selection. $\varepsilon_{\text{Hits}}$ is the efficiency corresponding to the cut on the number of hits, $\varepsilon_{\Sigma E}$ the efficiency of the energy sum cut, ε_{COG} the efficiency of the center of gravity cut, ε_{SSF} the efficiency of the shower start finder algorithm. The efficiency is defined as the ratio between the number of events passing the selection cut and the initial number of events.

completely electrons and pions contamination. Several events are rejected after applying the shower start finder algorithm. This is due to the combination of two effects. On one side, there is the difficulties in identify a track due to the amount of not properly working channels. In addition to that, in this thesis only pions showering in the fully equipped part of the prototype (from layers 3 to 10) are considered, reducing the number of pions passing the selection. Roughly the same efficiency is obtained in data and simulation. In general, the efficiency for 10 GeV

pions is lower and this is due to the fact that less energy is deposited in the last two layers of the prototype, reducing the number of selected pions. In addition to that, the efficiency obtained for simulated pion events is lower than in data. This is due to the fact that distribution of the center of gravity is not perfectly simulated.

5.2. Amplitude Analysis of Pion Beam Data

Before studying the correlation of the time of the hits in the detector and the energy of the hits, the amplitude analysis has been performed, in order to have a confirmation on the agreement between data and simulation. For this purpose, the standard quantities for the study of a shower have been considered. These are the distribution of the number of hits per event, the energy sum per event and the longitudinal shower profiles, meaning the number of hits per event as a function of the position of the layers of the AHCAL prototype and the energy sum per layer. This can be observed in Figure 5.4 for 10 GeV pions and in Figure 5.5 for 90 GeV pions. In general an agreement within 20% – 30% is observed, as expected from the electron analysis, described in Chapter 4. The simulated pion data are produced using the QGSP_BERT_HP physics list as described in Section 2.2.2.

However, in Figure 5.5, showing the comparison between data and simulation for 90 GeV pions, a worse disagreement can be observed, in particular looking at the distribution of the energy sum per event and its mean as a function of the AHCAL layer's position. In the energy sum distribution, a more pronounced tail to higher values for the data can be seen, reflected in a higher mean energy sum value in the last layers of the shower profile, shown in Figure 5.5 d). This is due to the saturation effect. For these layers in fact, the number of effective pixels could not be tuned using the electron data. The statistics available were not enough, being these modules placed much further (after $\sim 3.2\lambda_n$ and up to $\sim 5\lambda_n$). Therefore, before moving on with the analysis, the number of effective pixels needed to correct for the saturation effect of the silicon photomultipliers has been tuned, according to the procedure described in Section 4.2.2. In this case, the number of effective pixels has been tuned considering only the 90 GeV pions, and then the chosen value has been proven on the 10 GeV pions. The number of effective pixels tested is listed in Table 5.3.

In Figure 5.6 the distribution of the hit energy for the two KETEK layers (layer 11 and 12 in Figure 3.7) and for the two SensL layers (layer 13 and 14 in Figure 3.7) for 90 GeV pions is shown. The different colors correspond to the different number of pixels used to saturate the simulation indicated in the legend, while the black line represents the data, not corrected for the saturation effect.

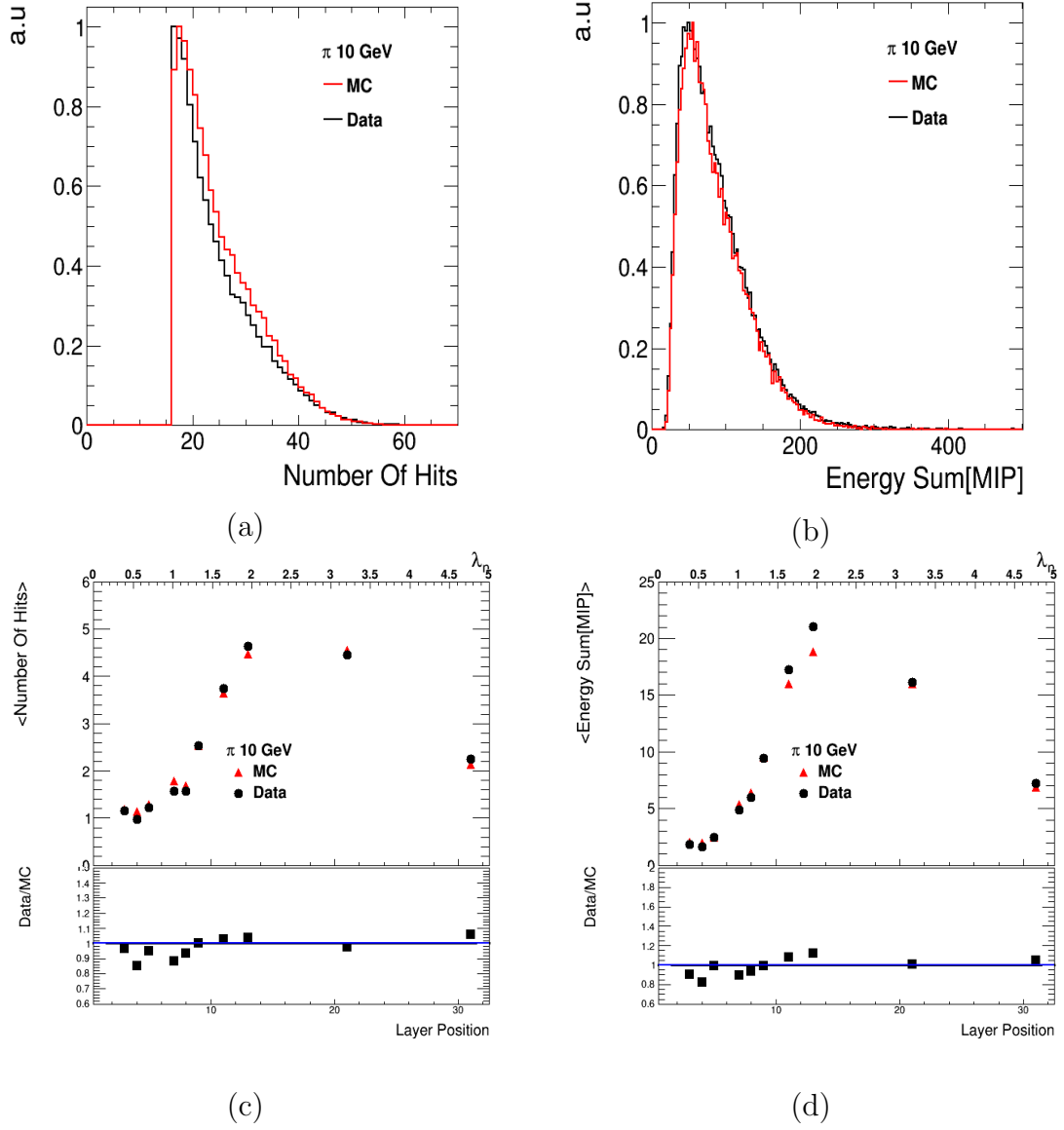


FIGURE 5.4.: Top plots: distributions of the number of hits per event (a) and of the energy sum per event (b) for 10 GeV pions. Bottom plots: mean number of hits as a function of the AHCAL layer position (c) and mean energy sum as a function of the layer position (d) for 10 GeV pions. Black circles represents the data and red triangles the simulation.

Layer	Real pixels	Pixels _{test1}	Pixels _{test2}	Chosen value
KETEK	2300	2700	2800	2800
SensL	1300	2000	2100	2100

TABLE 5.3.: Number of pixels of the SiPMs (Real pixels), number of effective pixels tested in order to tune this parameter in the simulation and number of effective pixels chosen for the KETEK and for the SensL layers.

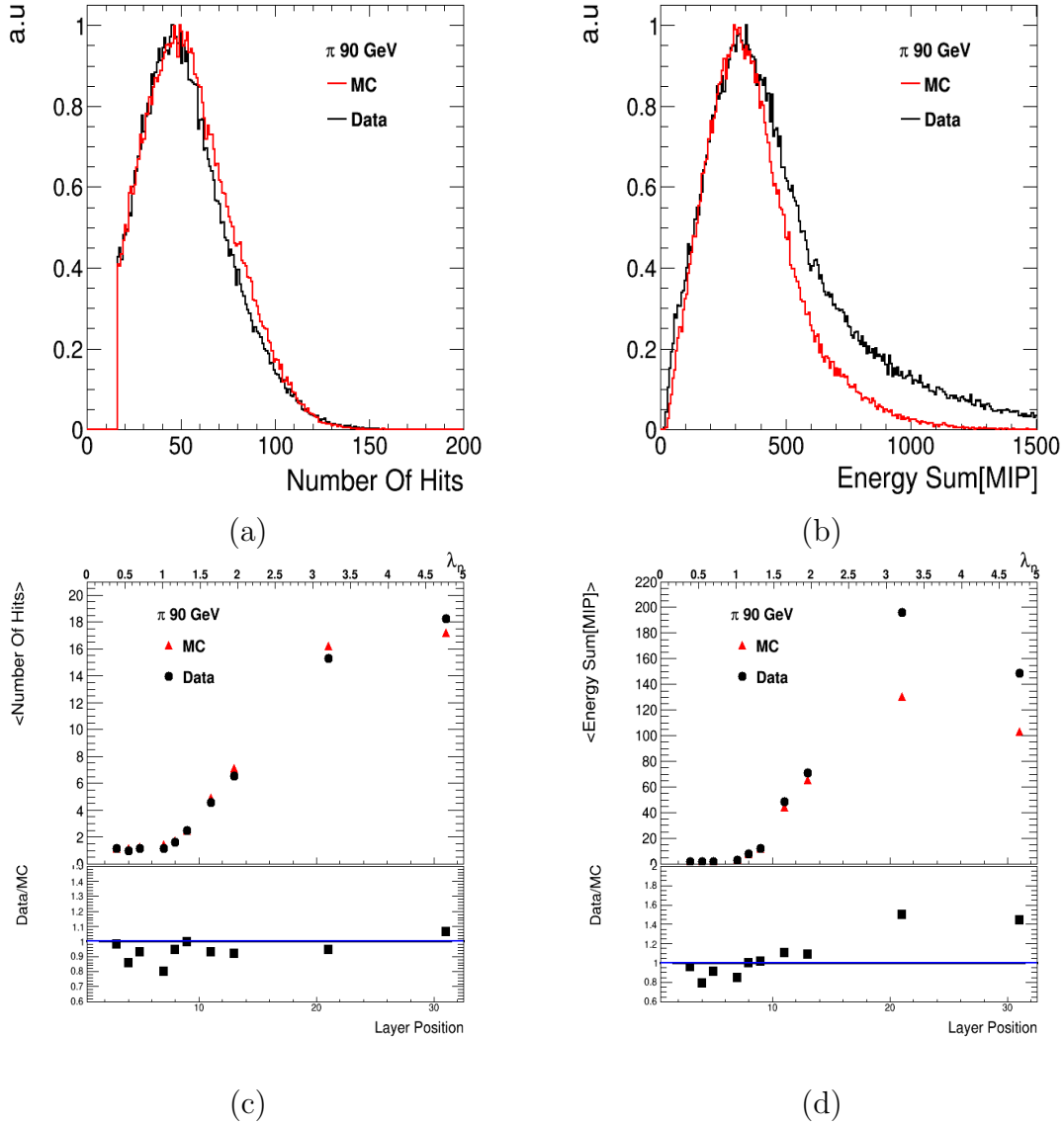


FIGURE 5.5.: Top plots: distributions of the number of hits per event (a) and of the energy sum per event (b) for 90 GeV pions. Bottom plots: mean number of hits as a function of the AHCAL layer position (c) and mean energy sum as a function of the layer position (d) for 90 GeV pions. Black circles represents the data and red triangles the simulation.

From these plots it is clear that a number of effective pixels equal to the real number of pixels of the silicon photomultipliers doesn't give a good description of the hit energy distribution, as already noticed in the electron analysis. According to these plots, a number of effective pixels equal to 2800 has been chosen for the KETEK layers while a number of effective pixels equal to 2100 has been chosen for layers 13 and 14. This is summarized in the last column of Table 5.3.

The same plots have been produced for 10 GeV pions, in order to verify the good agreement also at low beam energy. This is shown in Figure 5.7. Here again the simulation saturated using the real number of pixels of the silicon photo-multiplier

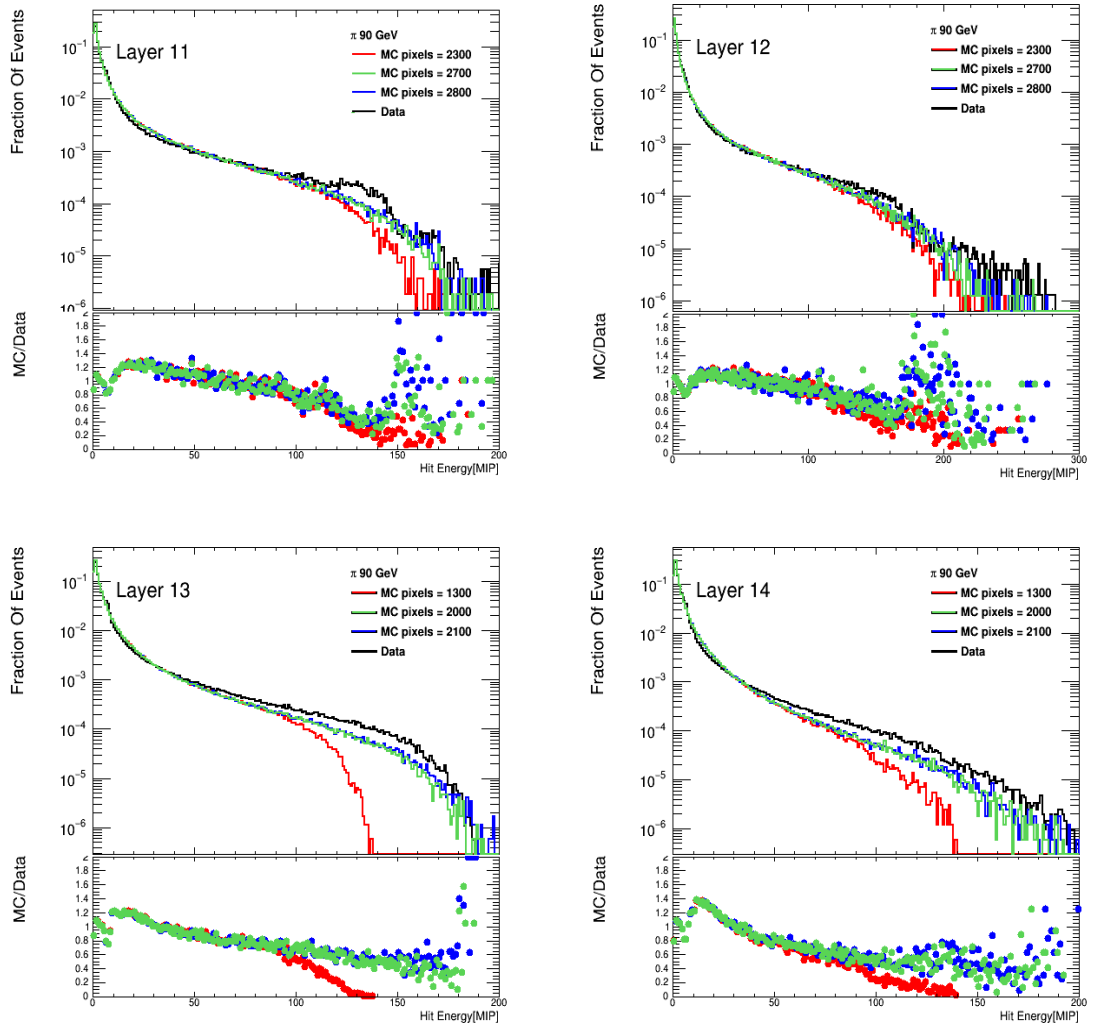


FIGURE 5.6.: *Tuning of the number of effective pixels for 90 GeV pions. Here the hit energy distribution is shown, on the top plots for the KETEK layers, on the bottom plots for the SensL layers. The black line represents the data without correcting for saturation effects, while the different colors, the simulation saturated with a different value of the number of effective pixels, as indicated in the legend.*

is shown. As it can be seen, the difference between using the real number of pixels and the tuned number of pixels in this case is really small. This is due to the fact that the saturation effect for low energy hadrons is not as important as for higher beam energies. In the bottom plots, a step for a hit energy around ~ 10 MIP is visible. This is due to a not proper value of the high gain-low gain intercalibration constants. This effect is nevertheless well taken into account in the systematics. As it has been studied in Section 4.3.2, a non proper high gain-low gain intercalibration can correspond up to 30% error in the energy sum.

The comparison between data and simulation after applying the saturation correction, using the new tuned number of effective pixels for the last four AHCAL

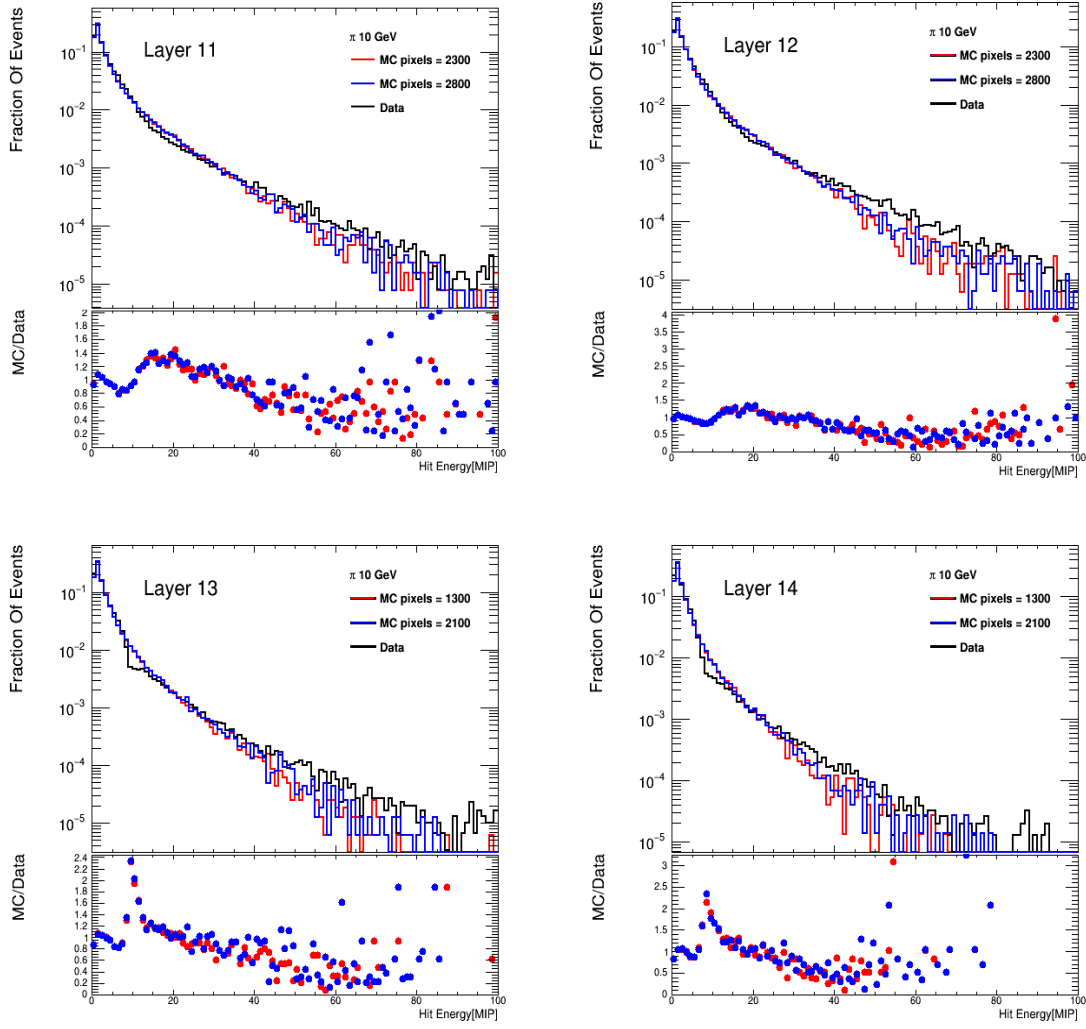


FIGURE 5.7.: Comparison between data and simulation of the hit energy distribution for 10 GeV pions. On the top plots for the KETEK layers (layer 11 on the left, layer 12 on the right) and on the bottom plots for the SensL layer (layer 13 on the left, layer 14 on the right). The black line represents the data, the pink line the simulation, saturated using the chosen number of effective pixels. On bottom left plot, the step around 10 MIPs indicate still a not perfect value of the high gain - low gain intercalibration constants.

layers is shown in Figure 5.8, in the top plots for 10 GeV and in the bottom plots for 90 GeV pions. The quantities where the saturation effect is more visible are shown, meaning the distribution of the energy sum per event and its mean as a function of the AHCAL layer's position. Here a better agreement with respect to Figure 5.4 for 10 GeV pions and Figure 5.5 for 90 GeV pions can be seen.

Finally, in order to properly estimate the level of agreement between data and simulation, the systematic uncertainties, treated in Section 4.3 have to be taken into account. The uncertainties considered for data are:

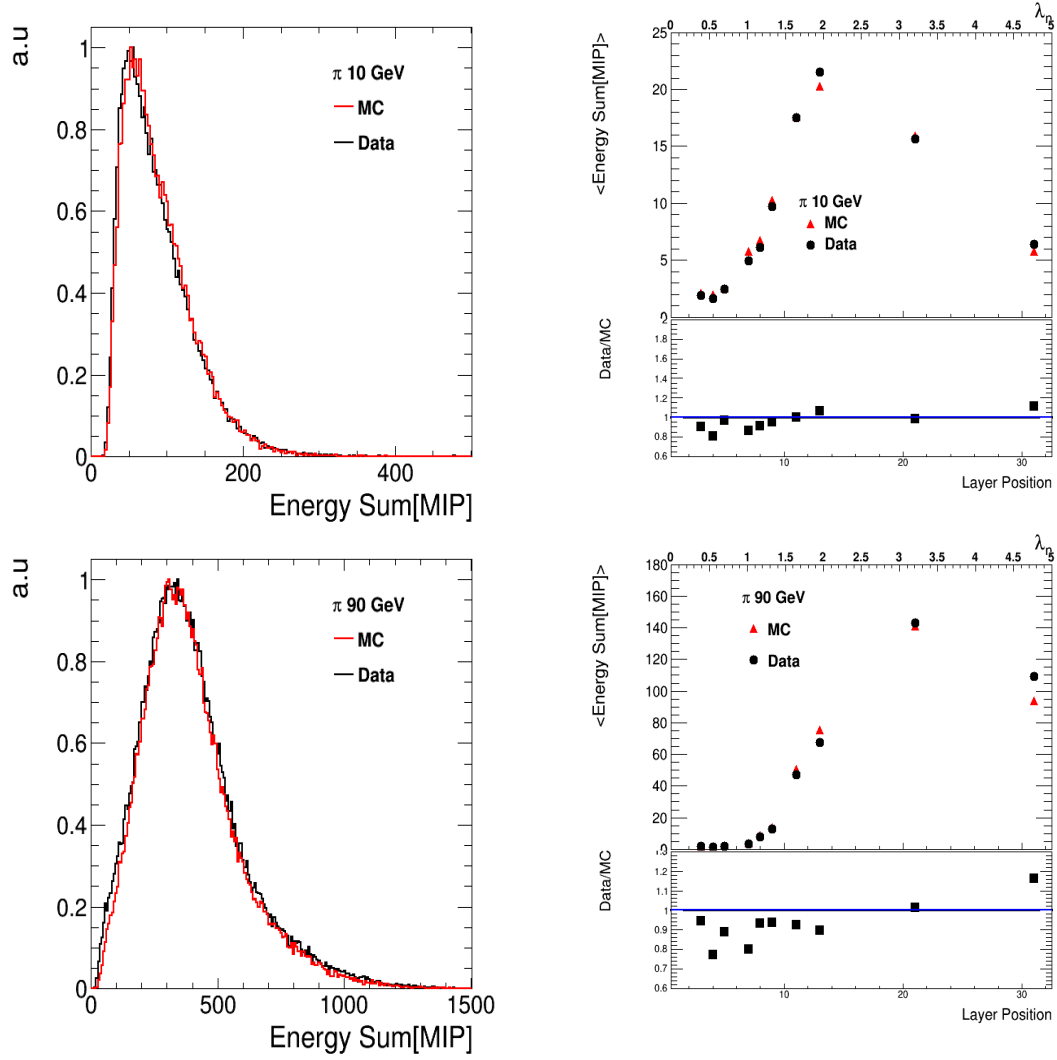


FIGURE 5.8.: Top plots: distribution of the energy sum per event (left) and mean energy sum per each AHCAL layer position(right) for 10 GeV pions. Bottom plots: distribution of the energy sum per event (left) and mean energy sum per each AHCAL layer position (right) for 90 GeV pion. Black represents the data and red the simulation. The plots are performed after the tuning of the number of effective pixels for the last AHCAL layers.

- the uncertainty due to the gain calibration
- the uncertainty due to the MIP calibration
- the uncertainty due to the high gain - low gain intercalibration.
- the number of effective pixels used to correct for the saturation effect listed in Table 4.3 for layers up to 10 and in Table 5.3 for layers from 11 to 14.

For the simulation, the uncertainties considered are:

- the value of the cross talk per tile chosen
- the number of effective pixels used to correct for the saturation effect listed in Table 4.3 for layers up to 10 and in Table 5.3 for layers from 11 to 14.
- the amount of additional material placed before the ECAL in order to simulate the upstream material

An overview of the values of these uncertainties for the quantities considered in the analysis and for the three beam energies is given in Table 5.4 for data and in table 5.5 for simulation.

The comparison between data and simulation is shown in Figure 5.9, in Figure 5.10 and in Figure 5.11 respectively for 10 GeV, 50 GeV and 90 GeV pions. The graphics for 10 GeV and 90 GeV have already been shown in figures 5.4, 5.5 and 5.8. Here they are shown again with the systematic uncertainties in order to estimate the level of agreement achieved. As expected from the electron analysis described in Chapter 4, an agreement within 20% – 30% can be observed.

Comparing the plots in Figures 5.9, 5.10 and 5.11 for pion data and the corresponding plots for the electron data (Figure 4.24, 4.26, 4.27, 4.28, 4.29) it is possible to notice that the agreement between data and simulation for pion data is indeed better than for electron data. This is due to the fact that the electromagnetic showers are denser than the hadronic showers and thus effects like the saturation and the high gain - low gain intercalibration are dominant in electron data, compared to pion data. The dominant uncertainty for data is still the one due to the high gain - low gain intercalibration, that can be up to 20%, as it can be seen in Table 5.4. For the simulation, the dominant uncertainties are the one related to the tuning of the cross talk parameter and the tuning of the upstream material. They can assume a value up to 7%, as it can be seen in Table 5.5. It is also possible to notice that the agreement obtained between data and simulation is within 20% – 30% as expected, and it doesn't deteriorate with increasing beam energy, as observed for electrons. The still not perfect agreement between data and simulation can be explained considering that the position of the beam with respect to the detector is not perfectly reproduced in simulation, as explained in Section 4.1 describing the electron analysis, and taking into account the uncertainties in modeling the hadronic showers.

Source Of Uncertainty	Energy	Variable	Value
Saturation	10 GeV	Number Of Hits	2%
		Energy Sum	2%
		Mean Number Of Hits	‰
		Mean Energy Sum	‰
	50 GeV	Number Of Hits	‰
		Energy Sum	2%
		Mean Number Of Hits	‰
		Mean Energy Sum	3%
	90 GeV	Number Of Hits	‰
		Energy Sum	4%
		Mean Number Of Hits	‰
		Mean Energy Sum	3%
Gain Calibration	10 GeV	Number Of Hits	‰
		Energy Sum	2%
		Mean Number Of Hits	‰
		Mean Energy Sum	‰
	50 GeV	Number Of Hits	‰
		Energy Sum	2%
		Mean Number Of Hits	‰
		Mean Energy Sum	‰
	90 GeV	Number Of Hits	‰
		Energy Sum	3%
		Mean Number Of Hits	‰
		Mean Energy Sum	‰
MIP Calibration	10 GeV	Number Of Hits	3%
		Energy Sum	4%
		Mean Number Of Hits	‰
		Mean Energy Sum	1.5%
	50 GeV	Number Of Hits	2%
		Energy Sum	3%
		Mean Number Of Hits	‰
		Mean Energy Sum	2%
	90 GeV	Number Of Hits	2%
		Energy Sum	4%
		Mean Number Of Hits	‰
		Mean Energy Sum	2%
HG-LG Intercalibration	10 GeV	Number Of Hits	2%
		Energy Sum	10%
		Mean Number Of Hits	‰
		Mean Energy Sum	8%
	50 GeV	Number Of Hits	‰
		Energy Sum	15%
		Mean Number Of Hits	‰
		Mean Energy Sum	15%
	90 GeV	Number Of Hits	‰
		Energy Sum	20%
		Mean Number Of Hits	5%
		Mean Energy Sum	15%

TABLE 5.4.: Overview of the uncertainties applied to the pion data. Details about how the uncertainties are extracted can be found in Section 4.3.

Source Of Uncertainty	Energy	Variable	Value
Saturation	10 GeV	Number Of Hits	2%
		Energy Sum	2%
		Mean Number Of Hits	‰
		Mean Energy Sum	‰
	50 GeV	Number Of Hits	2%
		Energy Sum	3%
		Mean Number Of Hits	‰
		Mean Energy Sum	‰
	90 GeV	Number Of Hits	2%
		Energy Sum	3%
		Mean Number Of Hits	‰
		Mean Energy Sum	‰
Cross Talk	10 GeV	Number Of Hits	5%
		Energy Sum	5%
		Mean Number Of Hits	4%
		Mean Energy Sum	5%
	50 GeV	Number Of Hits	6%
		Energy Sum	6%
		Mean Number Of Hits	5%
		Mean Energy Sum	5%
	90 GeV	Number Of Hits	7%
		Energy Sum	7%
		Mean Number Of Hits	6%
		Mean Energy Sum	5%
Upstream Material	10 GeV	Number Of Hits	5%
		Energy Sum	5%
		Mean Number Of Hits	1%
		Mean Energy Sum	2%
	50 GeV	Number Of Hits	4%
		Energy Sum	5%
		Mean Number Of Hits	1%
		Mean Energy Sum	1%
	90 GeV	Number Of Hits	4%
		Energy Sum	7%
		Mean Number Of Hits	1%
		Mean Energy Sum	1%

TABLE 5.5.: Overview of uncertainties applied to the simulated pion data. Details about how the uncertainties are extracted can be found in Section 4.3.

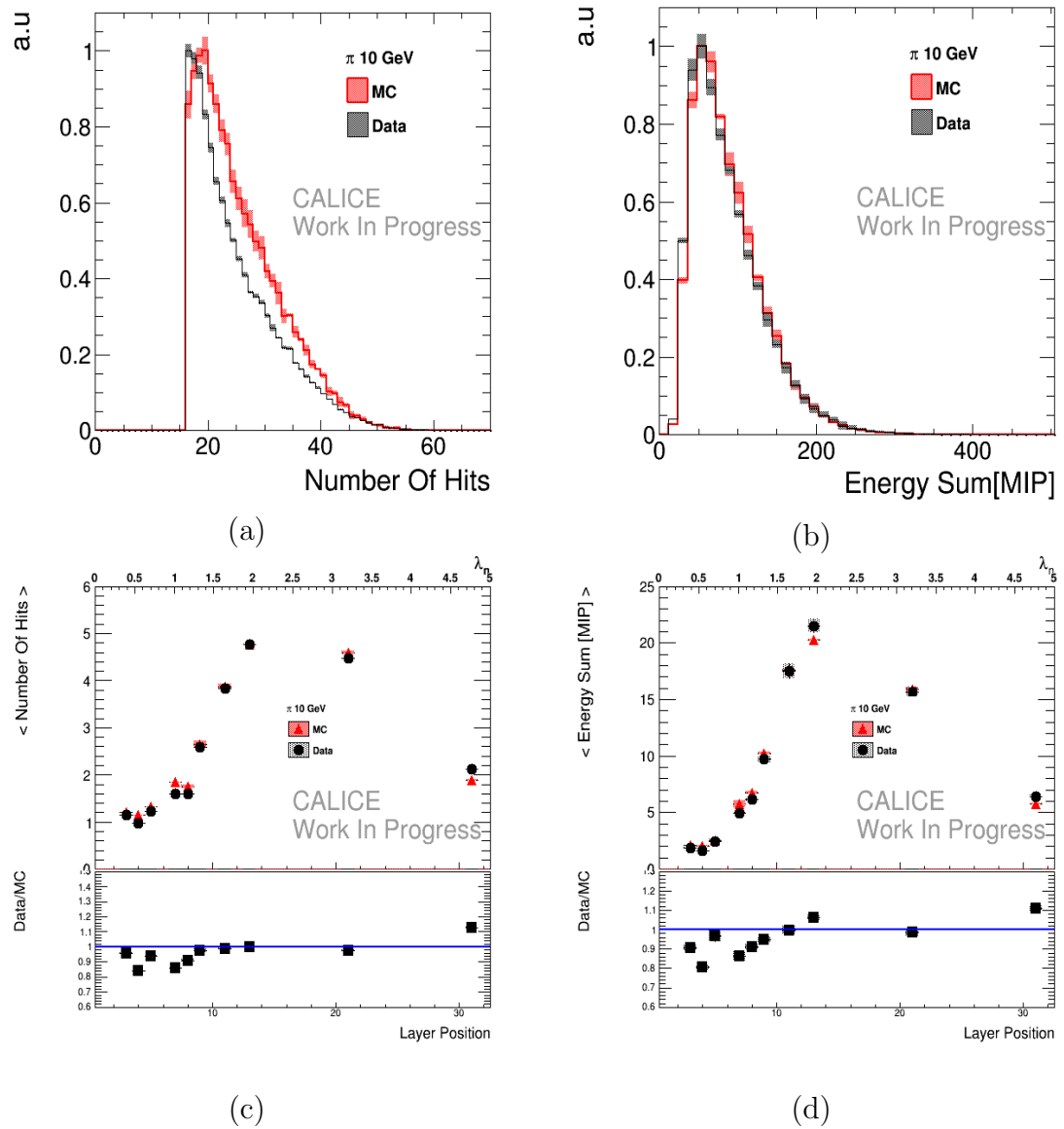


FIGURE 5.9.: *Top plots: distributions of the number of hits per event (a) and of the energy sum per event (b) for 10 GeV pions. Bottom plots: mean number of hits as a function of the AHCAL layer position (c) and mean energy sum as a function of the AHCAL layer position (d) for 10 GeV pions. Black represents the data and red simulation.*

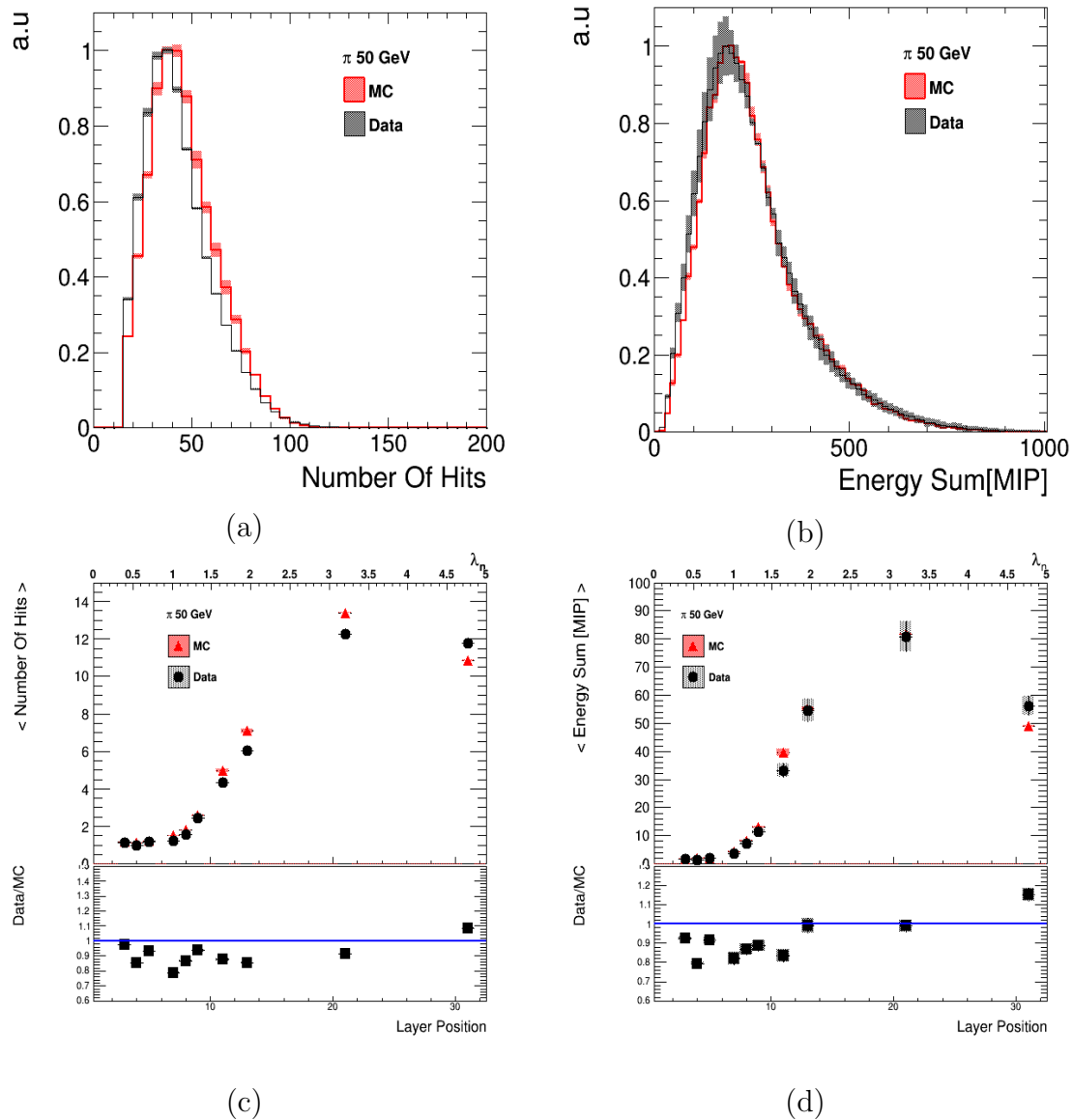


FIGURE 5.10.: Top plots: distributions of the number of hits per event (a) and of the energy sum per event (b) for 50 GeV pions. Bottom plots: mean number of hits as a function of the AHCAL layer position (c) and mean energy sum as a function of the AHCAL layer position (d) for 50 GeV pions. Black represents the data and red simulation.

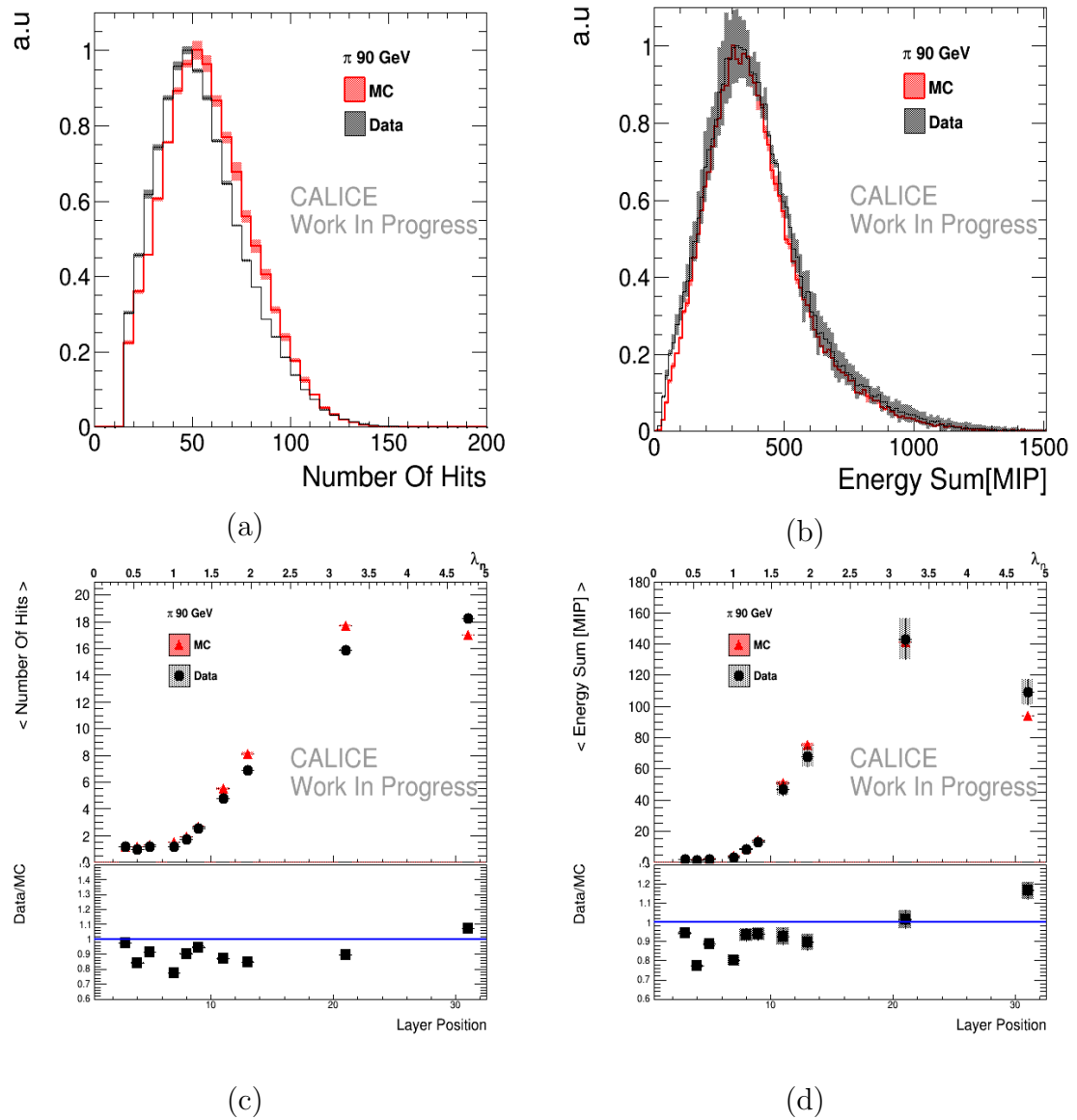


FIGURE 5.11.: Top plots: distributions of the number of hits per event (a) and of the energy sum per event (b) for 90 GeV pions. Bottom plots: mean number of hits as a function of the AHCAL layer position (c) and mean energy sum as a function of the AHCAL layer position (d) for 90 GeV pions. Black represents the data and red simulation.

5.3. Correlation between Energy and Time

As already mentioned, an advantage of the technological prototype is the possibility to have the information of the time of the hit in the detector. This is a very important information. Combined with the information of the energy of the hit in the detector, it allows in fact a better understanding of the composition of the hadronic shower.

Once the agreement of the pion data between data and simulation has been demonstrated by performing the amplitude analysis, the mean time of the hit has been studied. For this study, only layer 3, 7, 13, and 14 could be used (see sketch in Figure 3.7). This is due to difficulties encountered during the calibration procedures. Layers 6 and 10 couldn't be used, because, due to the high number of not working channels, was not possible to properly calibrate the reconstructed energy and was not possible to extract the number of effective pixels needed to correct for saturation effect (see Section 3.5, Section 4.2.2 and [99]). For the other layers, difficulties have been encountered in performing the time calibration. A check on the hit time distribution of each chip has been performed in order to reject outliers. Layer 11 has been completely rejected due most likely to an electronic problem. For more details see [99].

The time of the hit in the detector has been studied as a function of the radial distance of the hit. The hit radial distance, often referred in this thesis as r , is defined in Eq. 5.1

$$r_i = \sqrt{((X_i - X_{COG})^2 + (Y_i - Y_{COG})^2)} \quad (5.1)$$

meaning as the distance of the hit from the shower's center of gravity. In Eq. 5.1, X_i and Y_i represent the X and Y coordinates of the hit in the detector, while X_{COG} and Y_{COG} the coordinates of the center of gravity of the shower. An example of the distribution of the considered quantities is shown in Figure 5.12 for 10 GeV. On the left, the distribution of the time of the hit in the time window $(-50, 200)$ ns (according to [99]) is shown, while on the right a distribution of the distance of the hit from the center of gravity of the shower is shown.

The correlation between the time of the hit and r has been considered. In order to do that, the range of the hit radial distance has been divided in intervals, each of the dimension of one AHCAL tile (30 mm). In each interval, the mean of the time of the hits has been computed in the time window between $(-50, 200)$ ns. The correlation between the mean time of the hit and their distance from the center of gravity is shown in Figure 5.13 for pion data with a beam energy of 10 GeV, 50 GeV and 90 GeV. Here, in order to estimate the level of agreement between data and simulation, the systematic uncertainties have been added. The systematic uncertainties that have been taken into account for the data are:

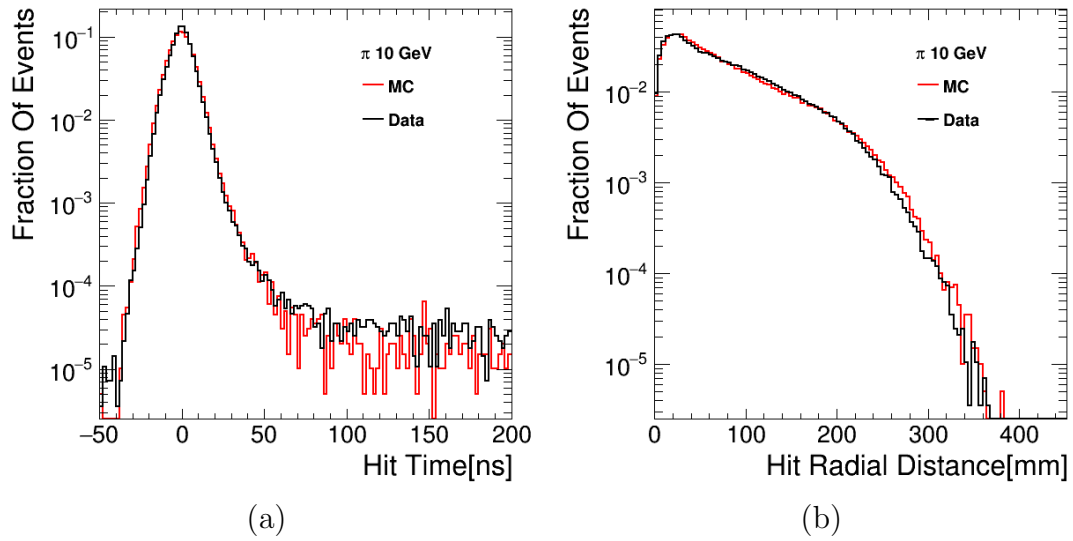


FIGURE 5.12.: *Distribution of the hit time (a) and distribution of the hit radial distance (b) for 10 GeV pions. Here black represents the data and red the simulation.*

- the uncertainty related to the non linearity of the TDC ramp, corresponding to 0.2 ns;
- the uncertainty associated to the time walk effect, corresponding to 0.2 ns;
- the uncertainty connected to the correction for the number of triggered channels in a chip (see Section 3.5.5 and Reference [99]).

In order to estimate this uncertainty, the hit time as a function of the number of triggered channels in a chip after the initial correction for this effect has been obtained and the deviation from 0 has been considered. Three different bins have been identified: a number of triggered channels in a chip up to 5, between 5 and 12 and higher than 12. The deviation of the hit time from 0 in these bins is respectively 2 ns, 5 ns and 7 ns. The final uncertainty has been defined as the sum of the weighted contribution in the three different ranges. This is explained in Eq.

$$\sigma = \frac{n_1}{ntot} \cdot \sigma_1 + \frac{n_2}{ntot} \cdot \sigma_2 + \frac{n_3}{ntot} \cdot \sigma_3 \quad (5.2)$$

where $n_i/ntot$ (with $i=1,2,3$) represents the fraction hits in each bin, and σ_i (with $i=1,2,3$) the deviation of the hit time from zero in each bin. This can correspond up to 4-5 ns and it's the dominant source of systematic uncertainties.

For the simulation, the uncertainties that have been considered are:

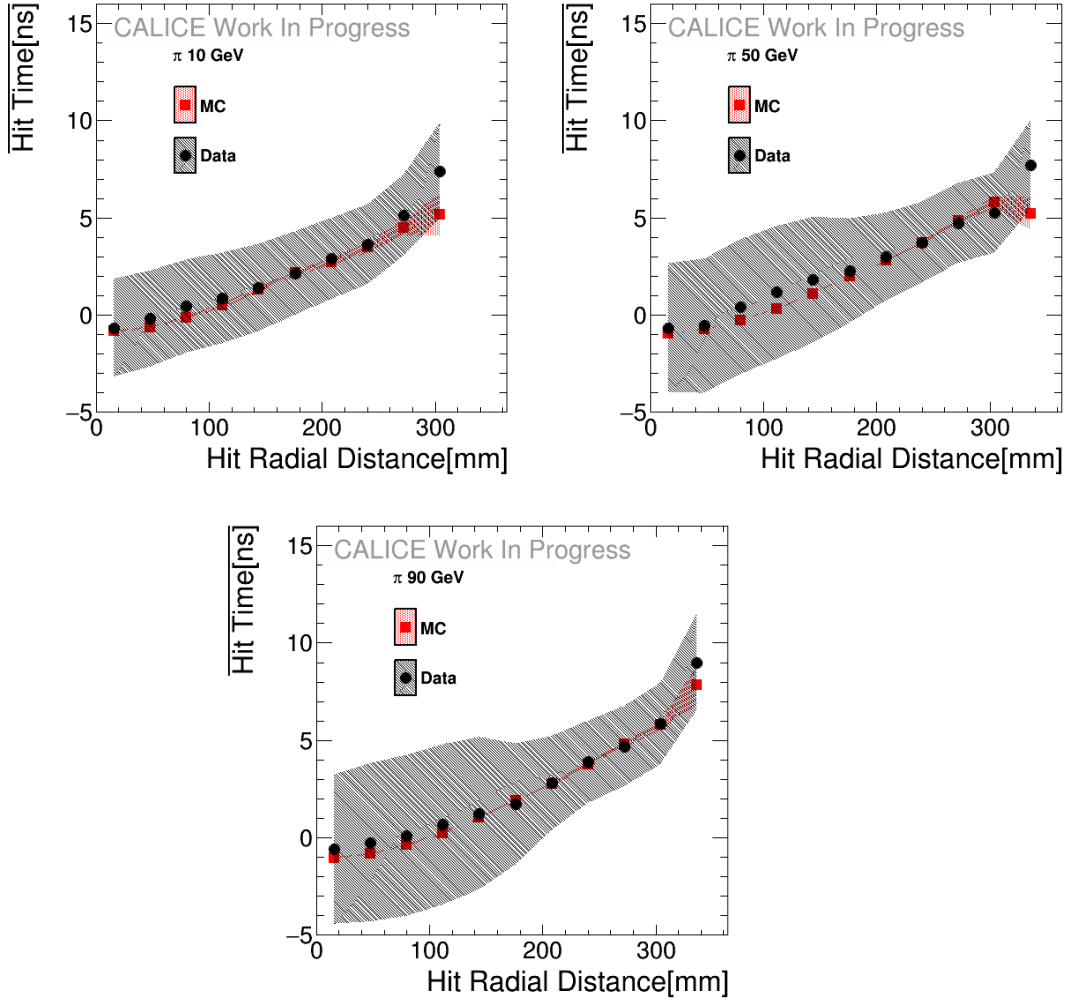


FIGURE 5.13.: Mean time of the hit as a function of its distance from the center of gravity of the shower, for all the three beam energies considered. Here black circles represent the data and red squares the simulation.

- the uncertainty due to the extraction of the offset, in order to shift the peak of the time distribution of each layer to 0, corresponding to 0.01 ns;
- the uncertainty due to the tuning of the cross talk parameter, as explained in Section 4.2.1. The uncertainty has been obtained as explained in Section 4.3.1, meaning the difference between the upper value of 15% and the lower value of 10% of cross talk per tile has been considered as the uncertainty. The contribution of this uncertainty varies between 0.003 ns and 0.1 ns.
- the uncertainty due to the parametrization of the number of triggered channels in a chip correction. It can vary between 0.007 ns and 0.1 ns.

A more detailed explanation on the extraction of the systematic uncertainties due to the time calibration can be found in [99].

A good agreement between data and simulations is observed. A correlation between the mean time of the hit in the detector and its position with respect to the center of gravity of the shower is visible. In addition to that it is possible to notice a larger transversal extension of the shower for higher beam energies, up to 300 mm for 10 GeV and up to 340 mm for 50 GeV and 90 GeV, corresponding to a higher number of hits in the detector. Already here is important to notice that the uncertainties at small radial distance are larger. This will be discussed in more details later in this section. It is important to underline that the center of gravity is computed for the full shower. In principle it could also be possible to define a center of gravity for each layer of the prototype and compute the hit radial distance from this. In this way the radial profile could result distorted. In fact, in presence of a noisy channel in a layer, the center of gravity of this layer would be shifted in that direction. The center of gravity computed for the full shower is much less sensitive to that, providing a non biased interpretation of the obtained results.

In order to better understand the dependence of the mean time of the hit on its distance from the center of gravity, the same correlation can be studied in different energy ranges. For this purpose, six energy ranges have been defined. These ranges have been identified in order to be denser in the range up to 5.5 MIPs. The hits with higher energy are in fact expected to correspond to the instantaneous electromagnetic component of the hadronic shower. The energy ranges are listed in Table 5.6.

Energy Bin	Energy[MIPs]
1	(0.5, 1.0)
2	(1, 1.5)
3	(1.5, 2.5)
4	(2.5, 3.5)
5	(3.5, 5.5)
6	> 5.5

TABLE 5.6.: *Definition of the energy ranges in which the correlation between the mean time of the hit and its distance from the center of gravity of the shower has been studied.*

The correlation between the addressed quantities in the energy ranges defined in Table 5.6 is shown in Figure 5.14 for 10 GeV pions.

Also here a dependence of the mean time of the hit on its radial distance from the center of gravity can be seen. In particular, this dependence is much stronger in the first energy bins and it decreases with increasing hit energy. This is compatible with the two different components of the hadronic shower. The hits with smaller energy correspond to the hadronic component showering in the detector with a certain delay time. The hits with higher energy, therefore the correlation shown in

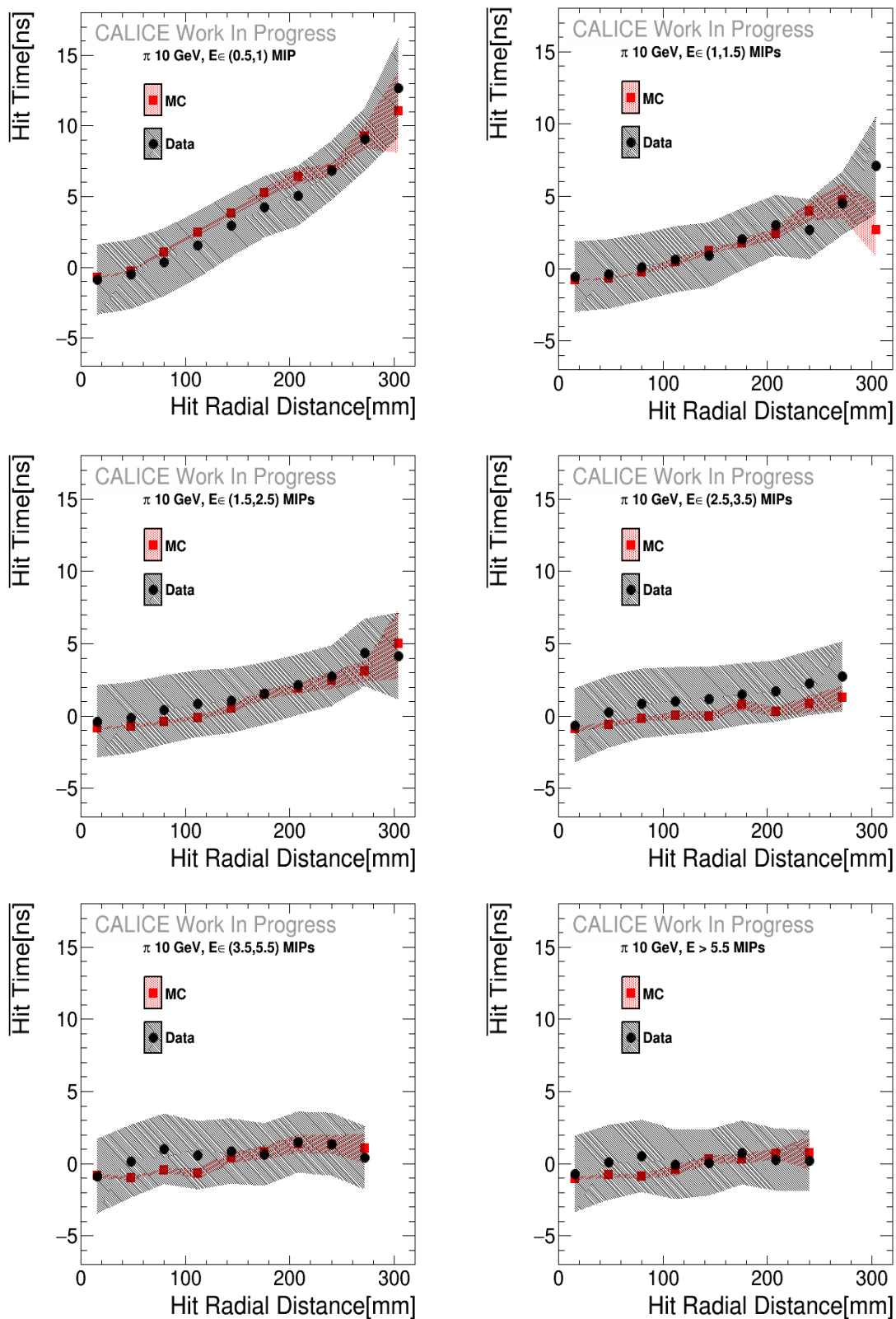


FIGURE 5.14.: Mean time of the hit as a function of its radial distance from the center of gravity of the shower in the defined energy bins, for 10 GeV pions. Here black circles represent the data and red squares the simulation.

the last energy bins, is associated to electromagnetic sub - shower, that propagate instantaneously through the detector. In these bins, the mean time of the hit in the detector as a function of the distance from the center of gravity is compatible with zero. The systematic uncertainties described above are added, together with the statistical. In the last radial distance bins, also the statistical uncertainty plays a more important role. It can reach up to 2 ns for low energies (10 GeV).

The same dependence can be observed for higher beam energies. This is shown in Figure 5.15 for 50 GeV pions and in Figure 5.16 for 90 GeV pions. Here, increasing the hit energy, meaning in the last energy bins, an anomalous behaviour in the region of r between 90 mm and 180 mm can be observed in data. In this region, it is possible to notice also that the uncertainty is larger at small hit radial distance, where the hit density is higher. This is due to the fact that the uncertainty related to the correction of the number of hits in a chip is the dominant one, as mentioned above. In addition, the uncertainties increase both with the beam energy and with the energy of the bin defined, bringing to the hypothesis that the behaviour seen around $r \sim 100$ mm is most likely due to the non optimal correction of the effect caused by the number of triggered channels in a chip.

In order to better understand this behaviour visible only for higher energy beams, and not for the 10 GeV pions, several checks have been done, using the 90 GeV pions.

First of all, the same correlation has been studied individually for each of the remaining layers of the prototype, in the same energy bins. This is shown in Figure 5.17 in the first energy bin ($E \in (0.5, 1)$ MIP) and in Figure 5.18 for the last energy bin ($E > 5.5$ MIPs) for 90 GeV pions. The anomalous behaviour is mainly caused from one of the 2×2 HBUs layers, meaning layer 13 (see Figure 3.7).

In order to further investigate the origin of the effect seen in layer 13, the correlation between the mean time of the hit and its distance from the center of gravity of the shower in the same energy bins is considered for each HBU of the layer. This has been done to investigate if a dependence on the position of the HBUs exists. These plots are shown in Figure 5.19 and in Figure 5.20, always for 90 GeV pions in the first and last energy bin respectively. Here no clear dependence on the position of the HBUs is visible. All the HBUs show an anomalous behaviour, but it seems to not be the same in all the HBUs. Therefore the hit time distribution in each energy bin and in each hit radial bin has been considered. This corresponds to make a distribution of the time of the hit for each points in Figure 5.20. This is shown in Figure 5.21. Here the second HBU has been taken as an example of this proof. In Figure 5.21, as expected, the mean of the time distributions is well centered in zero with some tails, representing the delayed components of the hadronic showers. In Figure 5.22, a shift in the mean

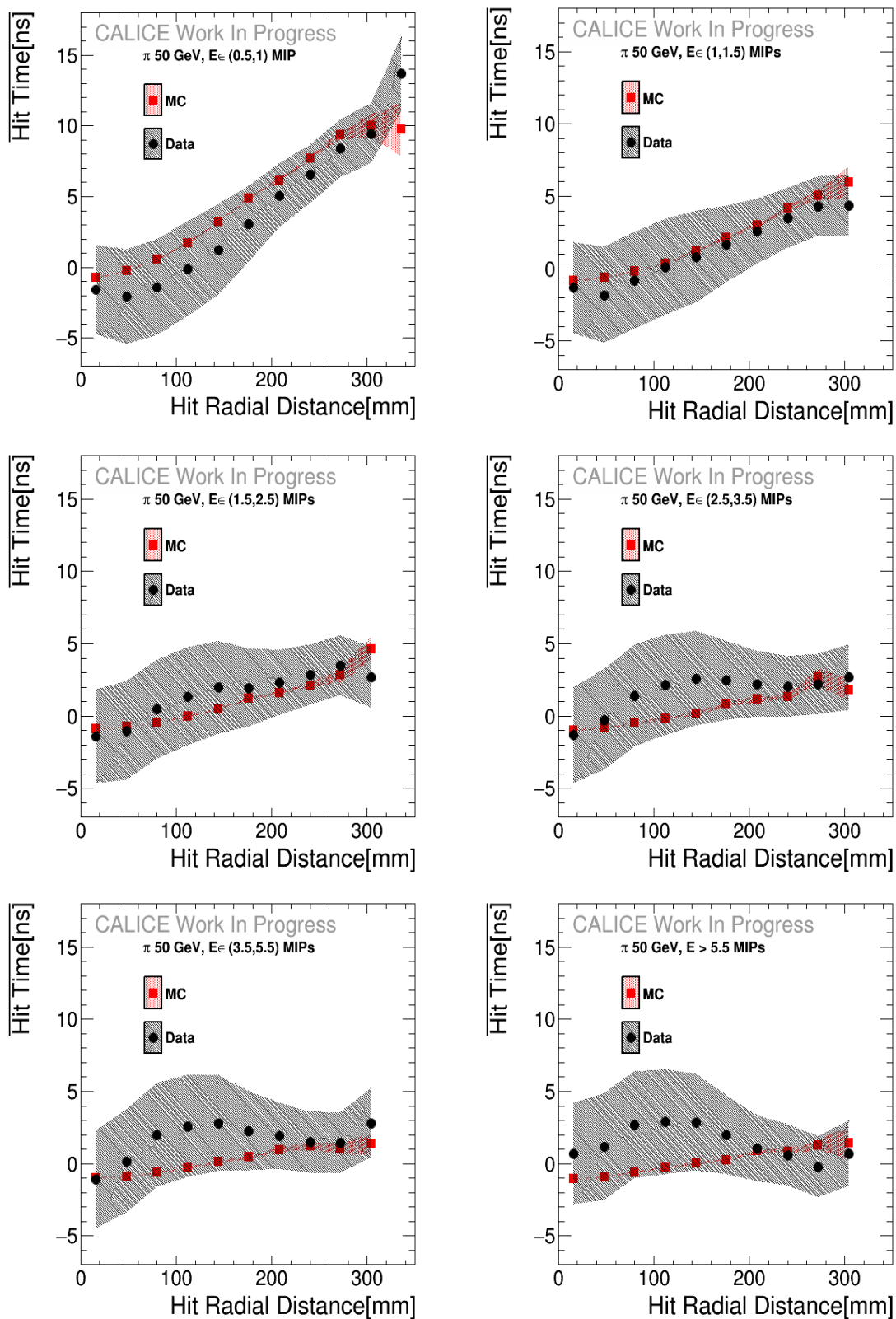


FIGURE 5.15.: Mean time of the hit as a function of its radial distance from the center of gravity of the shower in the defined energy bins, for 50 GeV pions. Here black circles represent the data and red squares the simulation.

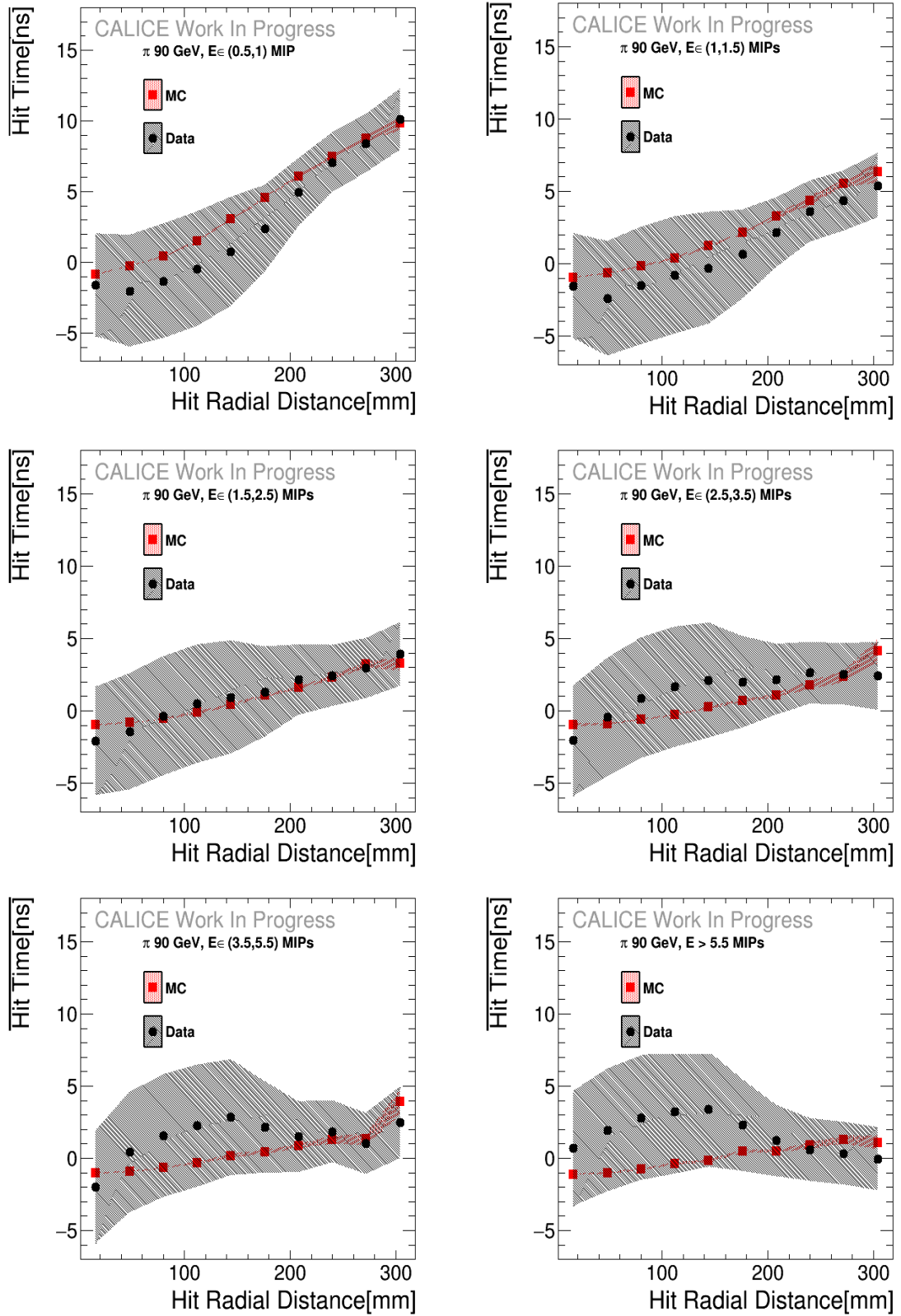


FIGURE 5.16.: Mean time of the hit as a function of its radial distance from the center of gravity of the shower in the defined energy bins, for 90 GeV pions. Here black circles represent the data and red squares the simulation.

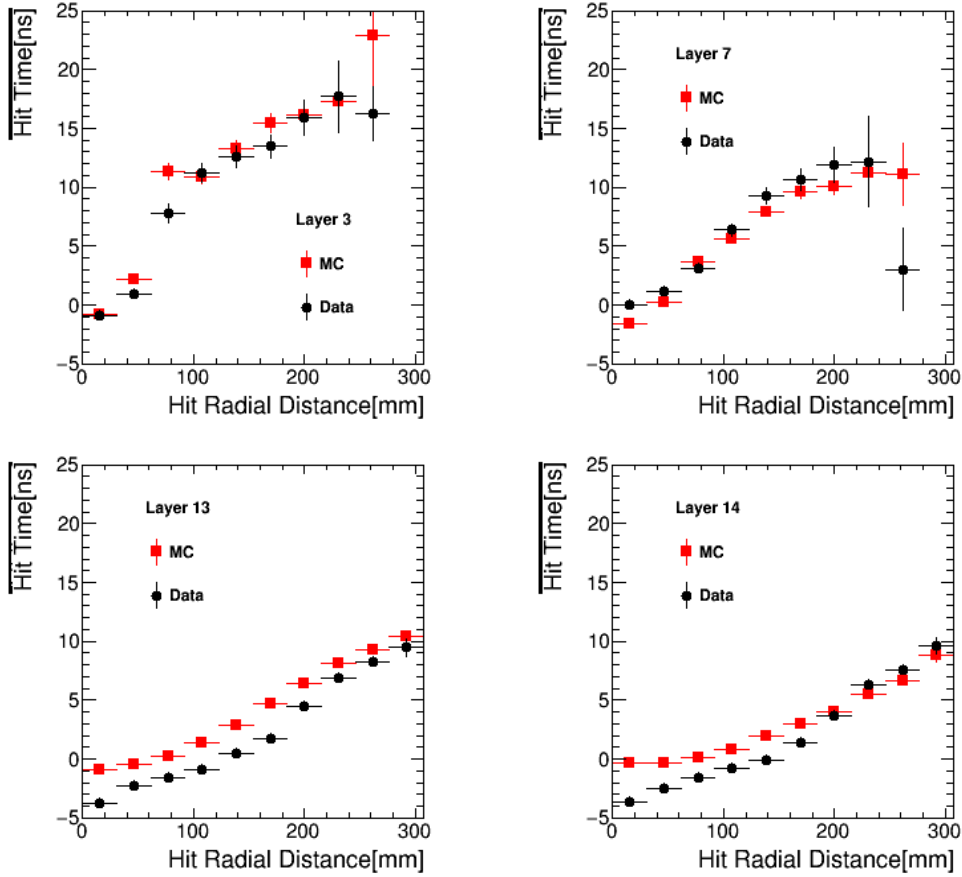


FIGURE 5.17.: Mean time of the hit as a function of the hit radial distance from the center of gravity of the shower in each layer of the prototype, in the first energy bin ($E \in (0.5, 1)$ MIP see Table 5.6), for 90 GeV pions. Here black circles represent the data and red squares the simulation

of the hit time distribution in data can be observed. In addition to that, some distributions (first and second plots in Figure 5.22) result also broader, with a hint of a double peak. This behaviour is a clear hint of some calibration problems, since, as mentioned above, physics effects in the hadronic shower are expected to result in a tail in the time distribution. In particular, this effect seems to be related to the correction related to the number of triggered channels in a chip. As mentioned in Section 3.5.5, a number of hits higher than 15 in a chip can cause an increase up to 40 ns in the time measured from the TDC of the chip. In principle, a correction for this effect has been extracted using electron data by combining all the beam energies since this effect should not be energy dependent. In order to extract the parameters needed to correct for this effect, the correlation between the mean time of the hit and the number of triggered channels in a chip has been studied, considering all the chips together. Due to the amount of not properly working layers, channels, and statistics, reason for which the correction cannot be extracted chip by chip, it is unfortunately not possible to better perform this

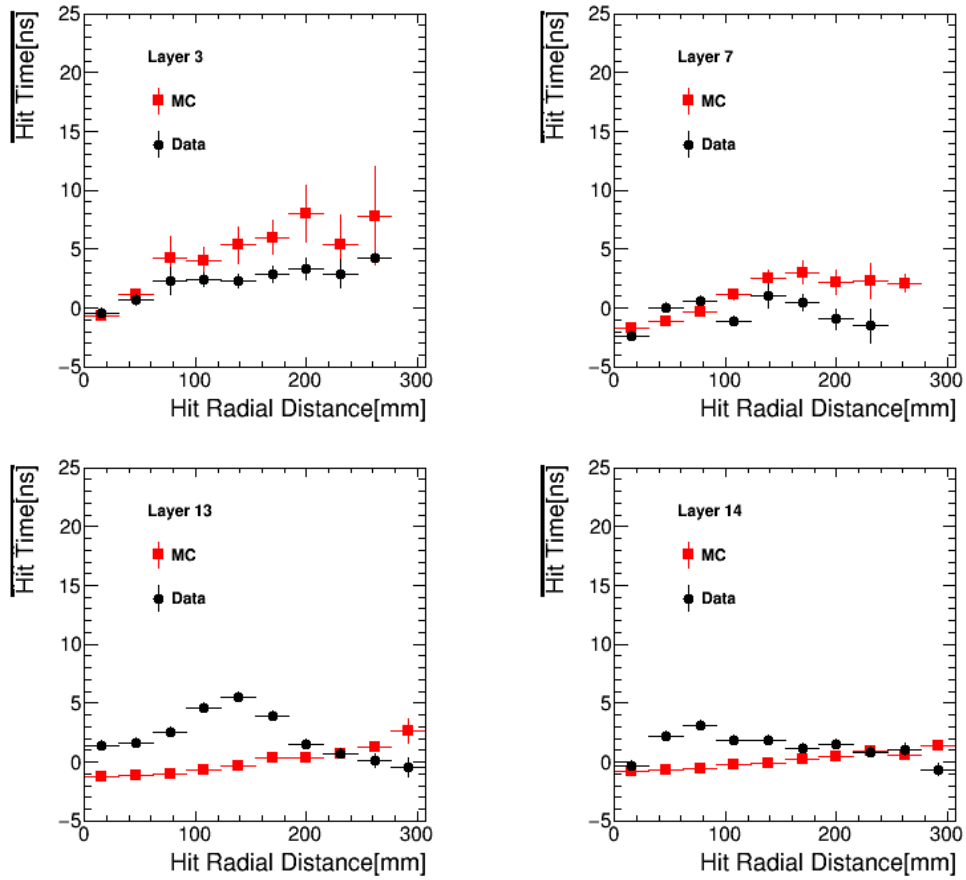


FIGURE 5.18.: Mean time of the hit as a function of the hit radial distance from the center of gravity of the shower in each layer of the prototype, in the last energy bin ($E > 5.5$ MIPs see Table 5.6), for 90 GeV pions. Here black circles represent the data and red squares the simulation

correction with the current data sets [99]. However, this effect is well taken into account in the computation of the systematic uncertainties, as shown in Figure 5.15 and in Figure 5.16.

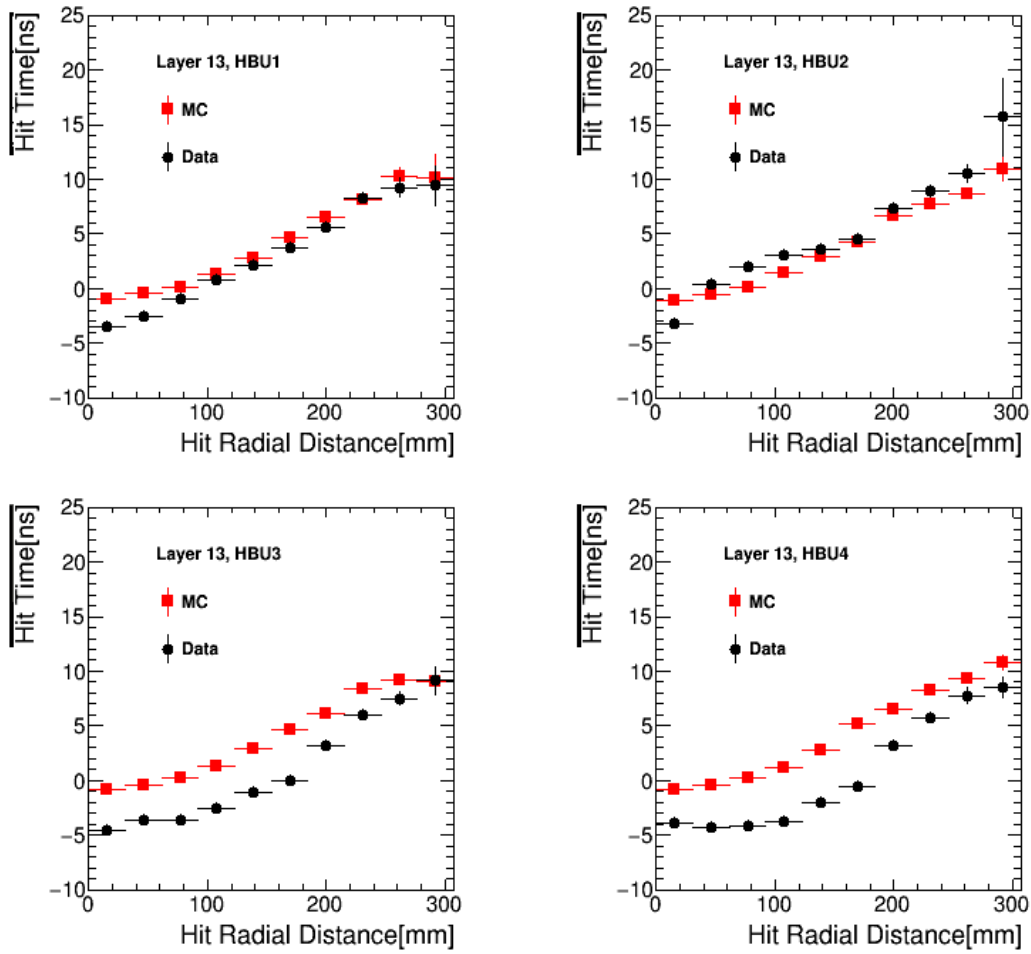


FIGURE 5.19.: Mean time of the hit as a function of the hit radial distance from the center of gravity of the shower in each HBU of layer 13, in the first energy bin ($E \in (0.5, 1)$ MIP see Table 5.6), for 90 GeV pions. Here black circles represent the data and red squares the simulation

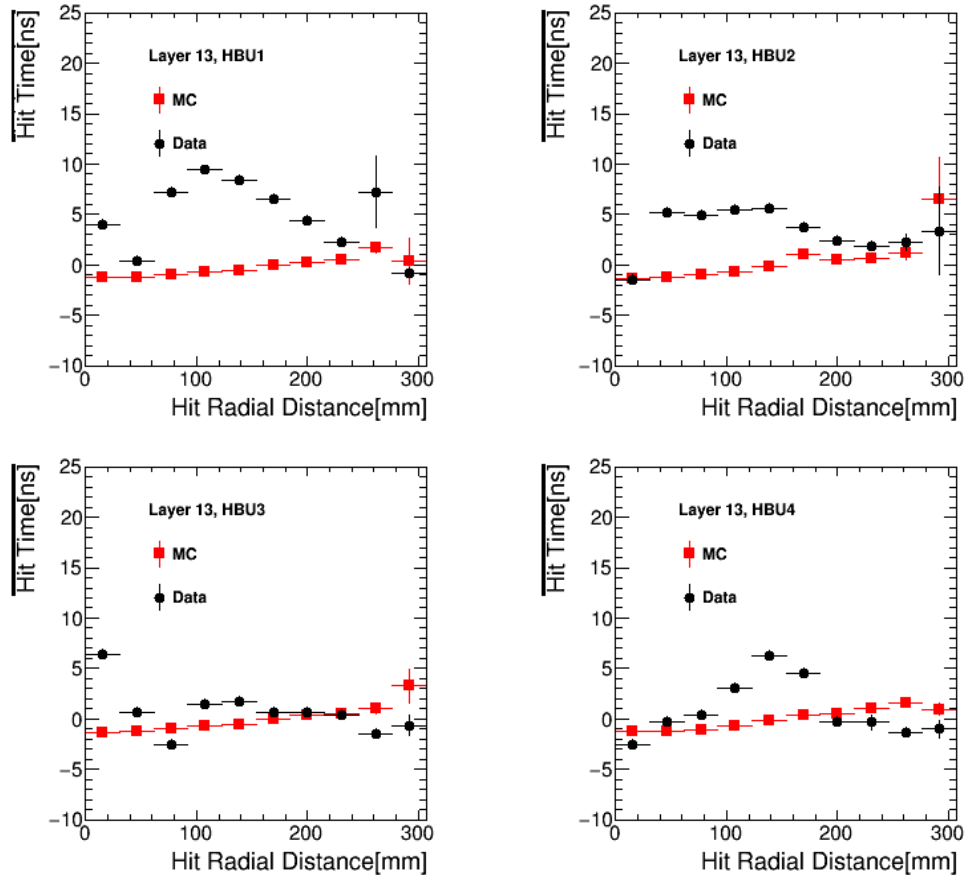


FIGURE 5.20.: Mean time of the hit as a function of the hit radial distance from the center of gravity of the shower in each HBU of layer 13, in the last energy bin ($E > 5.5$ MIPs see Table 5.6), for 90 GeV pions. Here black circles represent the data and red squares the simulation

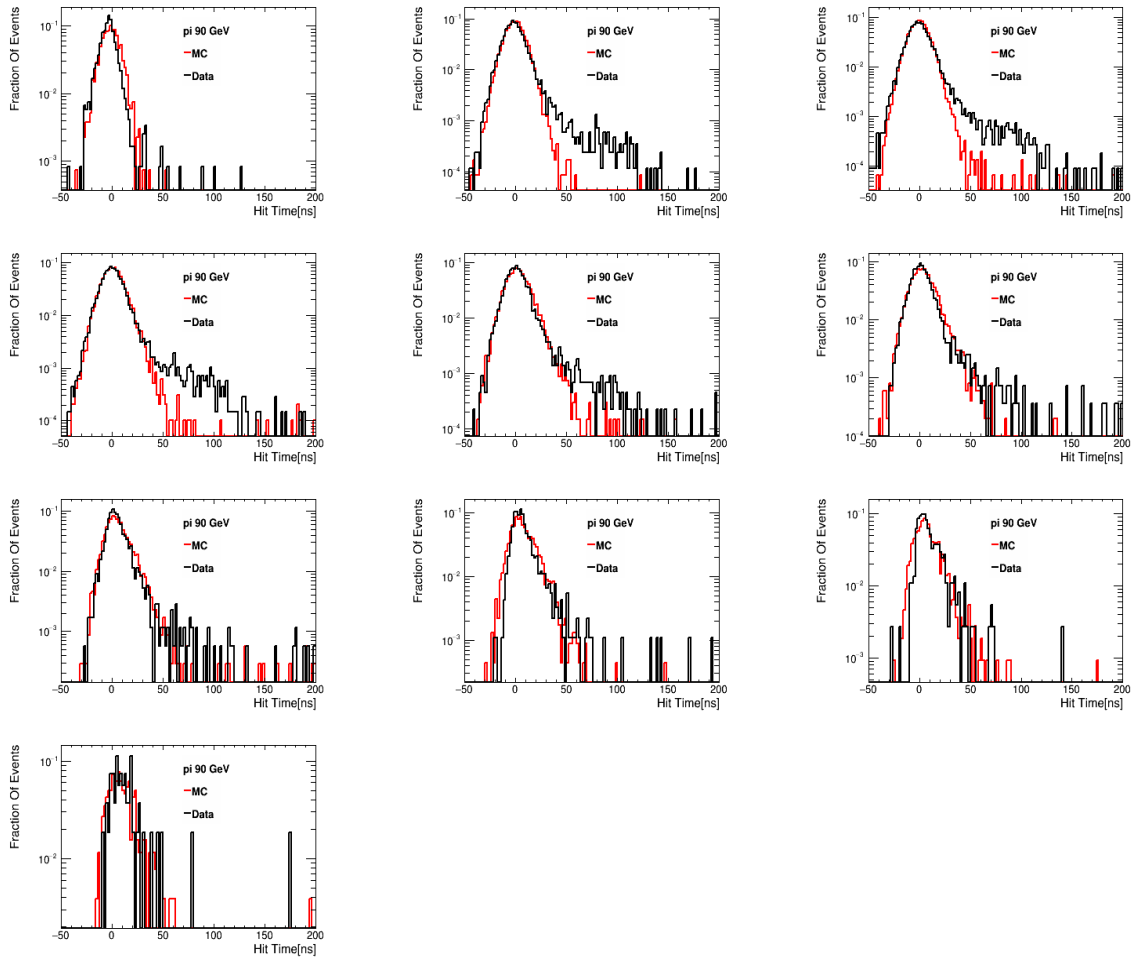


FIGURE 5.21.: *Distribution of the hit time in each hit radial distance bin in the second HBU of layer 13, in the first energy bin ($E \in (0.5, 1)$ MIP, see Table 5.6), for 90 GeV pions. Here each plot corresponds to one point in the top right plot in Figure 5.19 (one radius bin). Black represents the data and red the simulation.*

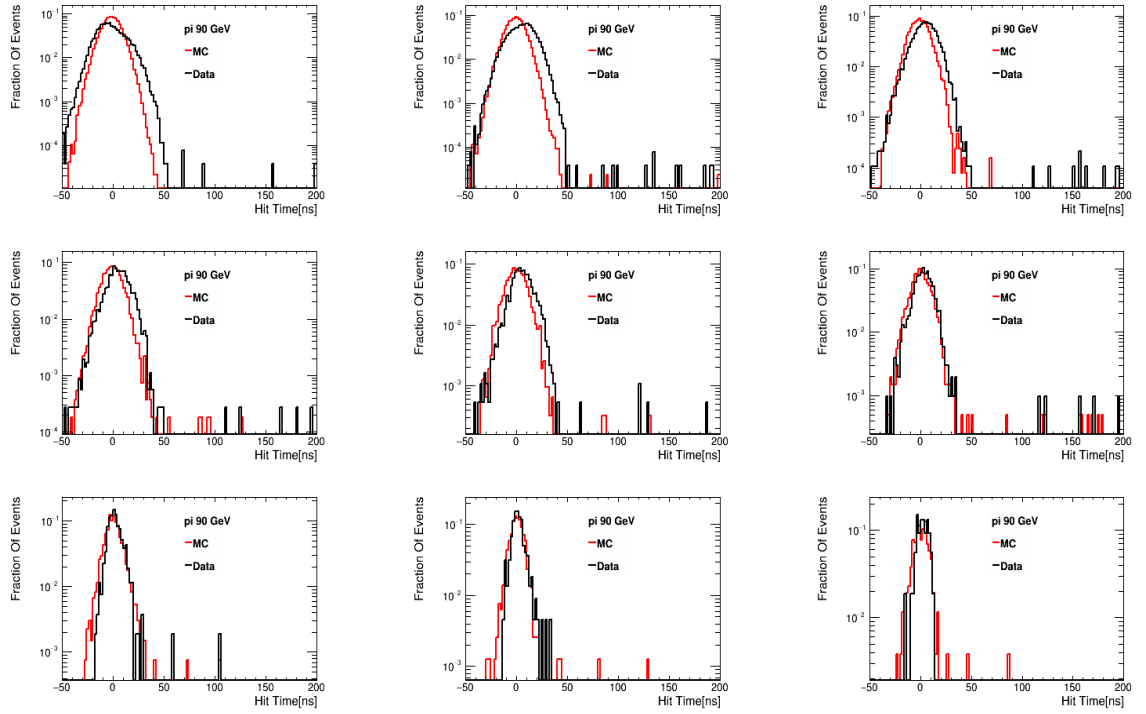


FIGURE 5.22.: Distribution of the hit time in each hit radial distance bin in the second HBU of layer 13, in the last energy bin ($E > 5.5$ MIPs, see Table 5.6), for 90 GeV pions. Here each plot corresponds to one point in the top right plot in Figure 5.20 (one radius bin). Black represents the data and red the simulation

5.4. Conclusion

In this chapter, the analysis of the pion data collected during the test beam campaign at CERN in July 2015, has been described. As a first step, the analysis of the reconstructed energy in the prototype has been illustrated. This has allowed to further tune the input parameters of the digitization process of the simulation. In particular the number of effective pixels needed for correcting for the saturation effect of the silicon photomultipliers for the last four modules has been tuned using pion data. This was not possible with the electron data, since the last layers of the prototype were placed around $15X_0$ and $46X_0$. After that, the characteristic quantities of a shower have been considered and a comparison between data and simulation has been performed, obtaining an additional confirmation about the proper description of the detector in the simulation. Systematic uncertainties have been considered and a data - simulation agreement between 20% – 30% has been obtained.

The main reasons for the 30% disagreement observed in the comparison between data and simulation of the reconstructed energy are difficulties encountered both in data and simulation. In data, difficulties in the calibration have been encountered. These concerned in particular the high gain - low gain intercalibration and they were mostly due to the limited available statistics and unexpected behaviour of the front-end electronics (a linear behaviour has been assumed). In the simulation, the parameters needed in the digitization process, i.e cross talk, number of effective pixels, had to be tuned, since no measurements were available. This limits the achievable precision. The optical cross talk is expected to have a larger impact on the agreement between the number of hits in the detector in data and simulation rather than on the reconstructed energy. Therefore it could have an influence in the timing study. Nevertheless, these effects together don't represent a limit for the timing study of hadronic showers. As it has been discussed in this work, other features related to the front-end electronics will have a larger impact on the timing analysis.

After the agreement between data and simulation of the reconstructed energy in the calorimeter has been validated, the introduction of the time of the hit in the analysis was possible. In particular, the correlation between the energy and the time of the hit measured in the detector has been studied. This study allowed a better understanding of such a complex object, like the hadronic shower. For this purpose this correlation has been studied in selected hit energy ranges. In this way the two different components of the hadronic shower have been identified. In the first energy bins a stronger dependence of the mean time of the hit on its radial distance from the shower's center of gravity is visible. This corresponds to the hadronic components delayed of few nanosecond in time when traversing the detector. In the last energy bins a value of the mean time of the hit compatible

with zero within the systematic uncertainties is observed. This corresponds to the instantaneous electromagnetic sub-shower.

The agreement between data and simulation obtained is within an error of $2 - 4$ nanoseconds, mainly because of an effect of the front-end electronics, that induces an increase of the time measured depending on the amount of triggered channels within a chip.

Regardless of the fact that the prototype was not homogeneous, being made of different kind of tiles and silicon photomultipliers, only few working layers were left and the limited amount of statistics, the study shows that the time information can help in understanding and identifying the different components of the hadronic shower and it looks very promising for the next prototype.

Conclusion

Future e^+e^- collider experiments as the *International Linear Collider* (ILC) are under discussion, in order to complement the *Large Hadron Collider* both in precision measurements and in possible discovery of new physics beyond the Standard Model. The ILC requires an unprecedented jet energy resolution of 3% – 4% to fulfill its ambitious goals in the precision measurement of the Standard Model parameters estimated at the level of 1%. This can be achieved with the application of *Particle Flow Algorithms* (PFAs).

Particle Flow Algorithms aim to measure the energy of every single particle in a jet using the sub-detector measurement yielding the best resolution for each particle. This places stringent requirements on the detector system. The tracking system needs to have an excellent track reconstruction efficiency higher than 99%, while the calorimeter system has to have a high granularity in order to assign calorimeter energy depositions to single particle tracks and to identify neutral hadrons energy depositions, not associated to any track.

The CALICE Collaboration develops prototypes for such novel high granularity calorimeters. One concept consists of scintillator tiles or strips of few cm^2 area, read out by *Silicon Photomultipliers* (SiPMs). Several prototypes have been built and tested in several test beam campaigns.

The *Analog Hadron Calorimeter* (AHCAL) technological prototype has been built to prove the scalability of the calorimeter to a full linear collider detector. It has been operated at CERN in July 2015 in particle beams of muons, electrons up to 50 GeV and pions up to 90 GeV. It consists of $3 \times 3 \text{ cm}^2$ tiles as active material, while the absorber material used was steel. During this test beam campaign, the prototype was only partially equipped with active layers.

The study of the structure of the hadronic showers, with the pion data collected during this test beam 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 and to tune the input parameters of the digitization process applied to the physics simulated events, in order to have a reliable description of the detector. After the main parameters of the simulation have been tuned and the agreement between data and simulation has been evaluated, it has been possible to proceed with the analysis of pion data. In particular, the goal of this work is to combine the time information of a hit in the detector and its energy information to study and improve the

understanding of the structure of hadronic showers. For this purpose, the mean time of the hit as a function of the hit radial distance from the shower's center of gravity has been studied in defined hit energy ranges.

The main challenges in this analysis are the large number of channels to be calibrated in energy and the tuning of the main parameters of the simulation in order to have a reliable description of the prototype, both complicated by the fact that the prototype was made of different kind of tiles and SiPMs and no measurements of the saturation function of the used silicon photomultipliers and of the optical cross talk were available. The event selection, in particular of the electron data, was complicated by the presence of low energy electron contamination.

Before being able to analyze electron and pion data, the calibration of the detector, meaning the gain calibration of the SiPMs and the MIP calibration, has been validated. In particular a method to perform the *High Gain - Low Gain* intercalibration had to be developed.

After the validation of the calibration, the electron data are analyzed. This analysis aims to tune the main parameters in the simulation of the detector, in order to have a reliable description of it. The main parameters in the digitization process, the optical cross talk and the parameters for the saturation correction of the silicon photomultipliers, have been tuned and a comparison between data and simulation has been performed. In particular the distribution of the number of hits per event, the distribution of the energy sum per event and the longitudinal shower profiles have been considered. A precision and agreement of 20% – 30% has been reached. The agreement is good at low beam energy, at the level of 20%, and it get worse, at the level of 30%, for higher energies. This is due to the low energy electron contamination not possible to reduce further with the partially equipped prototype. In addition, there is an increase of the uncertainties on the data on the energy sum per event and on the longitudinal shower profile. This is due to the uncertainty related to the high gain - low gain intercalibration and the uncertainty related to the saturation correction that start to be more important at higher energies, because the shower becomes denser. They can reach values up to $\sim 20\%$. The agreement obtained with the electron analysis is nevertheless considered acceptable for the study of hadronic showers.

The goal of the analysis of pion data is to study the time measurement of the hit in the detector together with the energy information. This can help to disentangle different contributions to the composition of a hadronic shower. For this purpose, the dependence of the time of the hit in the detector has been studied as a function of the hit radial distance from the center of gravity of the shower. A dependence of the mean time of the hit on the hit radial distance from the shower's center of gravity has been observed. In order to understand this dependence, this correlation has been studied in six hit energy ranges, for 10 GeV, 50 GeV and 90 GeV beam energies. An increase of the mean hit time with increasing hit radial distance,

up to about 8 ns at 300 mm has been observed. The effect is stronger in the lowest hit energy ranges, below 1.5 MIP. This corresponds to the slow neutron component of the hadron shower, delayed by a few nanoseconds in time when traversing the detector. At higher hit energy ranges, this effect is less pronounced, and above an energy of 3.5 MIP it is compatible with zero within uncertainties. This corresponds to the instantaneous electromagnetic sub-shower. The agreement between data and simulation is within an error of 2 to 4 nanoseconds, mainly due to an effect of the front-end electronics, that induces an increase of the measured time depending on the amount of triggered channels within a chip.

This analysis presents the first application of the combined use of the time and the energy measurement of each hit in the detector in order to study hadronic showers. Regardless of the fact that the prototype was not homogeneously equipped, being made of different kind of tiles and silicon photomultipliers, and that only few working layers were left after calibration and the limited amount of statistics, the study shows that the time information can help in understanding and identifying the different components of hadronic showers. This can be a very powerful tool. First high granularity together with time information can help in rejecting pile-up, in a collider experiment, where this can represent an issue. This study highlights that the amount of energy rejected if a time cut is applied will be small, having therefore a very small impact on the energy resolution. In addition, the possibility of identifying the prompt and late component of the hadronic shower combining time and energy information could be a powerful tool for improving the jet energy resolution. The time information could be used to correct for energy fluctuations in a hadronic shower, which is the main cause of the limited hadronic energy resolution.

Acknowledgments

This thesis would have not been possible without numerous people I met on the way and whom I would like to thank here.

First of all, I would like to thank my supervisors Katja Krüger, for patiently answering all my questions and for the discussions of presentations, my note and this thesis, and Erika Garutti, for introducing me to the group and for her input to my work. Thanks for giving me the opportunity to do my PhD under your supervision.

I would like to thank all the FLC group for the nice atmosphere. A very special thank you to Andrea for always being kind and always willing to help me with organizational matters.

I would like to thank the old and new members of the AHCAL group for all their help and the great time we had together. I would like to thank Eldwan, who helped me numerous times when my CALICE software was not compiling. And in particular I would like to thank Ali, who had to join night shift at CERN with me talking Italian and getting lost, and Oskar, I couldn't hope for a better office mate.

Ich möchte mich herzlich bei meinem Tandem-/Mittagessenpartner Moritz bedanken. Es war sehr nett von dir, dass du mir von Anfang an mit der deutschen Sprache geholfen hast. Es hat immer viel Spass gemacht und ich bin dir sehr dankbar dafür.

Danach möchte ich mich auch bei Ralf bedanken. Du hast mich nochmal zum Bouldern gebracht, wodurch meine Sonntage schöner geworden sind, und von dir habe ich gelernt, ein Bier mit einem Feuerzeug zu öffnen.

There are two people that have left DESY and that I call Friends: Adrian and Dimitra. I had a lot of fun with both of you, sharing complains about the terrible Hamburg weather or making fun about rules a bit absurd for us or about my sometimes a bit crazy "problems". And I still have every time we chat and we meet. In particular I guess I owe you, Adrian, that I came back from Japan, after being in the weird bar. And Dimitra, you definitely deserve a special thanks, not only for proof-reading my thesis even though you were on holidays on an amazing Greek island, but also for the nice week we spent in December, lying on a beach, and for the nice evenings ending later than 10 pm just because otherwise "who is gonna tell this to Ralf tomorrow?", for your motivation to always try new things,

even though Hamburg was really not your city, and for your always encouraging and positive attitude during my hard days.

A special thanks to my friend Giuseppe, the most easy going and relaxed person I know, to whom you can say no without feeling bad, and who made my arrival and my years here in Hamburg definitely easier and nicer and always introduced me to new people.

Thanks also to my “roman” friends Arianna, Lorenzo and Peter. We don’t meet often, but when we manage, it is like if nothing changed.

Per ultima, ma non per importanza, vorrei ringraziare la mia famiglia. I miei fratelli Alessio e Arianna, che sebbene sempre impegnati e presi con le loro famiglie e le loro vite, so che mi vogliono bene e sono sempre lì per me. Arianna in particolare, che mi ricordi che socializzare a volte è importante, ed Alessio, compagno di concerti belli, che sicuramente hai avuto una grande influenza su di me e sul mio percorso di studi. Un grazie ai miei genitori, che mi hanno sempre lasciato libera nelle mie scelte, senza il loro supporto non avrei potuto fare tutto ciò che ho fatto fino ad ora. La mia mamma merita però un grazie speciale. La persona più forte e coraggiosa che io conosca, che non si lascia abbattere da ciò che accade, che sa sempre trovare il lato positivo in ogni cosa, che ha sempre fiducia in me, pronta a supportarmi ed ad accorrere in qualsiasi momento. Grazie mamma!

A. Simulation Input Parameters

Type & Energy	Beam Pos.[mm]		Beam Spread[mm]		Cherenkov Pres.[bar]
	X	Y	X	Y	
e^- 10 GeV	7.0	3.1	27.6	26.6	1.7
e^- 15 GeV	4.4	3.0	26.0	24.6	1.3
e^- 20 GeV	1.7	-0.9	25.5	24.8	0.75
e^- 30 GeV	-4.5	19.6	21.7	20.0	0.6
e^- 40 GeV	-2.8	10.6	22.6	23.7	0.59
e^- 50 GeV	-19.8	4.8	25.0	26.6	0.6
π^- 10 GeV	0.2	13.6	34.1	30.3	1.0
π^- 50 GeV	2.2	4.7	22.2	19.7	0.11
π^- 90 GeV	2.1	0.6	22.9	18.2	0.03

TABLE A.1.: *Input parameters of the simulation for electrons and pions, for all the energies. The Z coordinate of the beam was always set to -100000 mm.*

B. Runs Used

Type & Energy	Run
e^- 10 GeV	24528 \rightarrow 24577
e^- 15 GeV	24506 \rightarrow 24527
e^- 20 GeV	24476 \rightarrow 24504
e^- 30 GeV	24449 \rightarrow 24475
e^- 40 GeV	24420 \rightarrow 24448
e^- 50 GeV	24404 \rightarrow 24419
π^- 10 GeV	24306 \rightarrow 24397
π^- 50 GeV	24578 \rightarrow 24612
π^- 90 GeV	24233 \rightarrow 24364

TABLE B.1.: *Runs used for the analysis of electron data and pion data, for all beam energies.*

C. Database Tags

Calibration Set	Database Folder & Tag
2AhcModuleDescription	cd_calice_cernSPS2015/TestbeamJuly2015/ModuleDescription (HEAD)
Ahc2ModuleConnection	cd_calice_cernSPS2015/TestbeamJuly2015/ModuleConnection (HEAD)
Ahc2ModuleLocationReference	cd_calice_cernSPS2015/TestbeamJuly2015/ModuleLocationReference (HEAD)
Ahc2DetectorTransformation	cd_calice_cernSPS2015/TestbeamJuly2015/DetectorTransformation (HEAD)
Ahc2HardwareConnection	cd_calice_cernSPS2015/TestbeamJuly2015/HardwareConnection (HEAD)

TABLE C.1.: *CALICE calibration database used for the 2015 CERN July test beam campaign. In brackets the database tags are given.*

Calibration Set	Database Folder & Tag
E4DPedestal	cd_calice_cernSPS2015/TestbeamJuly2015/Pedestal (ahc2_pedestal_010)
E4DGainConstants	cd_calice_cernSPS2015/TestbeamJuly2015/gain_constants (ahc2_gain_constants_004)
E4DGainSlopes	cd_calice_cernSPS2015/TestbeamJuly2015/gain_slopes (ahc2_gain_slopes_003)
E4DMipConstants	cd_calice_cernSPS2015/TestbeamJuly2015/gain_slopes (ahc2_mip_constants_008)
E4DMipSlopes	cd_calice_cernSPS2015/TestbeamJuly2015/gain_slopes (ahc2_mip_slopes_006)
E4DDeadCellMap	cd_calice_cernSPS2015/TestbeamJuly2015/DeadCellMap (ahc2_DeadCells_017)
E4DSaturationParameters	cd_calice_cernSPS2015/TestbeamJuly2015/SaturationParameters (ahc2_Sat_003)
E4DIntercalibration	cd_calice_cernSPS2015/TestbeamJuly2015/Intercalibration (ahc2_IC_009)
E4DPhysicsCalibIntercalibration	cd_calice_cernSPS2015/TestbeamJuly2015/PhysicsCalibIntercalibration (ahc2_PhysicsCalibIC_002)
E4DTimeSlopes	cd_calice_cernSPS2015/TestbeamJuly2015/TimeSlopes (ahc2_time_slopes_001)
E4DTimePedestal	cd_calice_cernSPS2015/TestbeamJuly2015/TimePedestal (ahc2_time_Ped_001)

TABLE C.2.: *CALICE calibration database used for the 2015 CERN July test beam campaign. In brackets the database tags are given.*

Bibliography

- [1] S.L. Glashow. Partial-symmetries of weak interactions. *Nuclear Physics, Volume 22, Issue 4*, 1961. URL [https://doi.org/10.1016/0029-5582\(61\)90469-2](https://doi.org/10.1016/0029-5582(61)90469-2).
- [2] J.C Ward A. Salam. Electromagnetic and weak interactions. *Physics Letters, Volume 13, issue 2*, 1964. URL [https://doi.org/10.1016/0031-9163\(64\)90711-5](https://doi.org/10.1016/0031-9163(64)90711-5).
- [3] S. Weinberg. A Model of Leptons. *Phys. Rev. Lett. 19, 1264*, 1967. URL <https://doi.org/10.1103/PhysRevLett.19.1264>.
- [4] The ATLAS Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Physics Letters B 716 1-29*, 2012. URL <https://doi.org/10.1016/j.physletb.2012.08.020>.
- [5] The CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Physics Letters B 716 30*, 2012. URL <https://doi.org/10.1016/j.physletb.2012.08.021>.
- [6] D. Griffiths. *Introduction to Elementary Particles*. 2010. Wiley.
- [7] H. Baer et al. The International Linear Collider Technical Design Report-Volume 2: Physics. 2013. URL <https://arxiv.org/abs/1306.6352>.
- [8] T. Behnke et al. The International Linear Collider Technical Design Report-Volume 1: Executive Summary. 2013. URL <https://arxiv.org/pdf/1306.6327>.
- [9] T. Behnke et al. The International Linear Collider Technical Design Report-Volume 4: Detectors. 2013. URL <https://arxiv.org/pdf/1306.6329>.
- [10] H.Yukawa. On the Interaction of Elementary Particles.I. *Proceedings of the Phyico-Mathematical Society of Japan. 3rd Series. Volume 17, Pages 48-57*, 1935. URL https://doi.org/10.11429/ppmsj/1919.17.0_48.
- [11] B. Povh, K. Rith, C. Scholz, and F. Zetsche. *Particelle e nuclei*. 1998. Bollati Boringhieri.

- [12] P.W. Higgs. Broken Symmetries and the Masses of Gauge Bosons. *Physical Review Letters*, Volume 13, 508, 1964. URL <https://doi.org/10.1103/PhysRevLett.13.508>.
- [13] F. Englert and R.Brout. Broken Symmetries and the Masses of Gauge Bosons. *Physical Review Letters*, Volume 13, 321, 1964. URL <https://doi.org/10.1103/PhysRevLett.13.321>.
- [14] UA1 Collaboration. Experimental observation of isolated large transverse energy electrons with associated missing energy at $s=540$ GeV. *Physics Letters B*, Volume 122, 1983. URL [https://doi.org/10.1016/0370-2693\(83\)91177-2](https://doi.org/10.1016/0370-2693(83)91177-2).
- [15] UA1 Collaboration. Experimental observation of lepton pairs of invariant mass around 95 GeV/c² at the CERN SPS collider. *Physics Letters B*, Volume 126, 1983. URL [https://doi.org/10.1016/0370-2693\(83\)90188-0](https://doi.org/10.1016/0370-2693(83)90188-0).
- [16] UA2 Collaboration. Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the CERN pp collider. *Physics Letters B*, Volume 122, 1983. URL [https://doi.org/10.1016/0370-2693\(83\)91605-2](https://doi.org/10.1016/0370-2693(83)91605-2).
- [17] UA2 Collaboration. Evidence for $Z^0 \rightarrow e^+e^-$ at the CERN pp collider. *Physics Letters B*, Volume 129, 1983. URL [https://doi.org/10.1016/0370-2693\(83\)90744-X](https://doi.org/10.1016/0370-2693(83)90744-X).
- [18] M. Tanabashi et al.(Particle Data Group). Review of Particle Physics. *Phys. Rev. D* 98, 030001, 2018. URL http://pdg.lbl.gov/2018/listings/contents_listings.html.
- [19] W. Greiner and J. Reinhardt. *Field Quantization*. 1996. Springer.
- [20] The Large Hadron Collider, July 2018. URL <https://home.cern/topics/large-hadron-collider>. CERN.
- [21] S.P. Martin. A Supersymmetry Primer. *arXiv:hep-ph/9709356*, 2016. URL https://doi.org/10.1142/9789812839657_0001.
- [22] A. D Sakharov. Violation of CP invariance, C asymmetry, and baryon asymmetry of the universe. *Soviet Physics Uspekhi*, 34(5):392, 1991. URL <http://stacks.iop.org/0038-5670/34/i=5/a=A08>.
- [23] M. Trodden. Baryogenesis and Leptogenesis. *XXXII SLAC Summer Institute*, 2004. URL <https://arxiv.org/abs/hep-ph/0411301>.

- [24] F. Capozzi, E. Di Valentino, E. Lisi, A. Marrone, A. Melchiorri, and A. Palazzo. Global constraints on absolute neutrino masses and their ordering. *Phys. Rev. D* *95*, 096014, 2017. URL <https://doi.org/10.1103/PhysRevD.95.096014>.
- [25] K.G Begeman, A.H. Broeils, and R.H. Sanders. Extended rotation curves of spiral galaxies: dark haloes and modified dynamics. *Monthly Notices of the Royal Astronomical Society, Volume 249, Issue 3, Pages 523-537*, 1991. URL <https://doi.org/10.1093/mnras/249.3.523>.
- [26] E. Komatsu and C.L. Bennett on behalf of the WMAP science team. Results from the Wilkinson Microwave Anisotropy Probe. *Progress of Theoretical and Experimental Physics, Volume 2014, Issue 6, 06B102*, 2014. URL <https://doi.org/10.1093/ptep/ptu083>.
- [27] M. Robinson. *Symmetry and the Standard Model*. 2011. Springer.
- [28] K. Fujii et al. Physics Case for the 250 GeV Stage of the International Linear Collider. *arXiv*, 2017. URL <https://arxiv.org/abs/1710.07621>.
- [29] K. Fujii et al. Physics Case for the International Linear Collider. *arXiv*, 2015. URL <https://arxiv.org/abs/1506.05992>.
- [30] I. Low, R. Rattazzi, and A. Vichi. Theoretical Constraints on the Higgs Effective Couplings. *Journal of High Energy Physics*, 2010. URL [https://doi.org/10.1007/JHEP04\(2010\)126](https://doi.org/10.1007/JHEP04(2010)126).
- [31] M. Farina, M. Perelstein, and N. Rey-Le Lorier. Higgs couplings and naturalness. *Phys. Rev. D* *90*, 015014, 2014. URL <https://doi.org/10.1103/PhysRevD.90.015014>.
- [32] S. Dawson et al. Higgs working group report. *Proceedings of the APS DPF Community Summer Study*, Snowmass 2013. URL [arXiv:1310.8361v2](https://arxiv.org/abs/1310.8361v2).
- [33] M. A. Thomson. Model-independent measurement of the $e^+e^- \rightarrow HZ$ cross section at a future e^+e^- linear collider using hadronic Z decays. *Eur. Phys. J. C (2016)* *76*:72, 2016. URL <https://doi.org/10.1140/epjc/s10052-016-3911-5>.
- [34] S. Michizono L. Evans. The International Linear Collider Machine Staging Report 2017. Addendum to the International Linear Collider Technical Design Report published in 2013. *arXiv:1711.00568*, 2017. URL <https://arxiv.org/abs/1711.00568>.

- [35] V. Ayvazyan, N. Baboi, and J. Bähr et al. First operation of a free-electron laser generating GW power radiation at 32 nm wavelength. *Eur. Phys. J. D* 37: 297, 2006. URL <https://doi.org/10.1140/epjd/e2005-00308-1>.
- [36] M. Altarelli et al. The European XFEL Technical Design Report. 2007. Technical report.
- [37] T. Barklow et al. ILC Operating Scenarios. 2015. URL <https://arxiv.org/abs/1506.07830>.
- [38] R. Wigmans. *Calorimetry, Energy Measurement in Particle Physics*. 2000. Oxford Science Publications, Oxford University Press.
- [39] K. A Olive and Particle Data Group. *Review of Particle Physics*. Chinese Physics C 38, 2014.
- [40] I. C. Brock and T. Schörner-Sadenius. *Physics at the Terascale*. 2011. ISBN : 9783527410019. DOI : 10.1002/9783527634965.
- [41] T. A. Gabriel et al. *Energy Dependence of Hadronic Activity*. Nucl. Instrum. Meth. vol A338, pp. 336-347, 1994.
- [42] D. E. Groom. *Energy Flow in a Hadronic Cascade: Application to Hadron Calorimetry*. Nucl. Instrum. Meth., vol. A572, pp. 633-653, 2007.
- [43] N. Feege. *Low-energetic Hadron Interactions in a Highly Granular Calorimeter*. PhD thesis, Universität Hamburg, 2011.
- [44] GEANT4 Physics Reference Manual. URL <http://geant4-userdoc.web.cern.ch/geant4-userdoc/UsersGuides/PhysicsReferenceManual/fo/PhysicsReferenceManual.pdf>.
- [45] S. Agostinelli and Et Al. GEANT4: A Simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506.3. DOI 10.1016/S0168-9002(03)01368-8, 2003. URL <http://inspirehep.net/record/593382/files/fermilab-pub-03-339.pdf>.
- [46] J. Allison et al. Geant4 developments and applications. *IEEE Transactions on Nuclear Science*, 53(1):270–278, 2006. doi: 10.1109/TNS.2006.869826.
- [47] J. Allison et al. Recent developments in Geant4. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 835:186 – 225, 2016. doi: 10.1016/j.nima.2016.06.125. URL <http://www.sciencedirect.com/science/article/pii/S0168900216306957>.

- [48] P. Mora de Freitas and H. Videau. Detector simulation with MOKKA / GEANT4: Present and future, in *Linear colliders. Proceedings, International Workshop on physics and experiments with future electron-positron linear colliders, LCWS 2002, Seogwipo, Jeju Island, Korea, 2002.* pp. 623-627.
- [49] V. N. Ivanchenko et al. GEANT4 models for simulation of multiple scattering. *J. Phys.: Conf. Ser.* 219 032045, 2010. URL <http://iopscience.iop.org/article/10.1088/1742-6596/219/3/032045/pdf>.
- [50] V. N. Ivanchenko et al. Recent Improvements in GEANT4 Electromagnetic Physics Models and Interfaces. *Progress in NUCLEAR SCIENCE and TECHNOLOGY, Vol. 2, pp.898-903*, 2011. URL <https://pdfs.semanticscholar.org/55c7/530ea20fdcead1cd2e5de881e9655c7c6a00.pdf>.
- [51] A. Heikkinen, N. Stepanov, and J.P. Wellisch. Bertini intra-nuclear cascade implementation in GEANT4. *[nucl-th]*, 2003. URL <https://arxiv.org/abs/nuc1-th/0306008>.
- [52] G. Folger and J.P. Wellisch. String Parton Models in GEANT4. *Computing in High Energy and Nuclear Physics*, 2003. URL <https://arxiv.org/abs/nuc1-th/0306007>.
- [53] B. Andersson. *The Fritiof Dynamics*. 1987. 2nd International Workshop on Correlation and Fluctuation in Multiparticle Production (IWCF 1997) Jinan, Shandong, China.
- [54] L.P. Csernai N.S. Amelin, E.F. Staubo. Semihard collisions in Monte Carlo quark-gluon string model. *Physical Review D* 46.11, 1992. URL <https://doi.org/10.1103/PhysRevD.46.4873>.
- [55] B. Nilsson-Almqvist B. Andersson, G. Gustafson. A model for low-pT hadronic reactions with generalizations to hadron-nucleus and nucleus-nucleus collisions. *Nuclear Physics, Section B* 281.1-2, 1987. URL [https://doi.org/10.1016/0550-3213\(87\)90257-4](https://doi.org/10.1016/0550-3213(87)90257-4).
- [56] J.M. Quesada et al. Recent Developments In Pre-Equilibrium And De-Excitation Models In GEANT4. *Progress in Nuclear Science and Technology, Volume 2, pages 936-941*, 2011. URL <http://geant4.web.cern.ch/sites/geant4.web.cern.ch/files/geant4/results/papers/deexcitation-MC2010.pdf>.
- [57] J. Apostolakis et al. GEANT4 Physics Lists for HEP. *IEEE Nuclear Science Symposium Conference Record*, 2008. URL <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=47774655>.

- [58] A. Ribon et al. Status of GEANT4 hadronic physics for the simulation of LHC experiments at the start of LHC physics program. *CERN-LCGAPP-2010-02*, 2010. URL <http://lcgapp.cern.ch/project/docs/noteStatusHadronic2010.pdf>.
- [59] H.L.Tran et al. Software compensation in particle flow reconstruction. *arXiv:1705.10363 [physics.ins-det]*, 2017. URL <https://dx.doi.org/10.1140/epjc/s10052-017-5298-3>.
- [60] A. Bernstein et al. Beam tests of the ZEUS barrel calorimeter. *Nuclear Instruments and Methods in Physics Research A 336 (1993) 23-52*, 1993. URL <http://inspirehep.net/record/355922/files/desy93-076.pdf>.
- [61] P. Adzic. Energy resolution of the barrel of the CMS Electromagnetic Calorimeter. *JINST 2 P04004*, 2007. URL <http://iopscience.iop.org/article/10.1088/1748-0221/2/04/P04004/pdf>.
- [62] L. Hervas. Development and construction of large size signal electrodes for the ATLAS electromagnetic calorimeter. *Nuclear Instruments and Methods in Physics Research A 539 558 - 594*, 2005. URL <http://dx.doi.org/10.1016/j.nima.2004.11.005>.
- [63] T. Davidek. Testbeam studies of production modules of the ATLAS Tile Calorimeter. *Nuclear Instruments and Methods in Physics Research A 606 362 - 394*, 2009. URL <http://dx.doi.org/10.1016/j.nima.2009.04.009>.
- [64] A. Henriques on the behalf of the ATLAS Collaboration. The ATLAS Tile Calorimeter. *4th International Conference on Advancements in Nuclear Instrumentation Measurement Methods and their Applications*, 2015. URL <http://dx.doi.org/10.1109/ANIMMA.2015.7465554>.
- [65] The CMS Collaboration. Performance of the CMS Hadron Calorimeter with Cosmic Ray Muons and LHC Beam Data. *JINST 5:T03012*, 2010. URL <https://arxiv.org/abs/0911.4991>.
- [66] M.Thomson. Particle flow calorimetry and PandpraPFA algorithm. *Nuclear Instruments and Methods in Physics Research A 611 25 - 40*, 2009. URL <http://dx.doi.org/10.1016/j.nima.2009.09.009>.
- [67] The CALICE Collaboration, December 2017. URL <https://twiki.cern.ch/twiki/bin/view/CALICE/WebHome>. CALICE Website.
- [68] International Linear Collider collaboration. URL <https://www.linearcollider.org>. ILC Website.

- [69] The ILC Collaboration. Technical Design Report. 2013. Technical Report Volume 4, International Linear Collider.
- [70] Compact Linear Collider (CLIC) collaboration. URL <http://cllc-study.web.cern.ch/>. CLIC Website.
- [71] The CLIC Collaboration. CLIC Conceptional Design Report - Physics and Detectors at CLIC. 2011. URL https://edms.cern.ch/ui/file/1176246/1/CLIC_CDR_PhysDet_SPC-version.pdf. Technical Report Volume 2, 274th CERN SPC.
- [72] P.Azzi, D.Barney, T.Bergauer, P.Bloch, P.Dauncey, A.David, M.Mannelli, J.Mans, A.Marchioro, A. Martelli, S.Moccia, M.Narain, C.Seez, J.Strait, C.de la Taille, and T.Virdee. The Phase-2 Upgrade of the CMS endcap calorimeter. *CERN-LHCC-2017-023*, 2018. URL <https://cds.cern.ch/record/2293646/files/CMS-TDR-019.pdf>. CMS-TDR-019, ISBN 978-92-9083-459-5.
- [73] C.Neubueser. *Comparison of Two Highly Granular Hadronic Calorimeter Concepts*. PhD thesis, Hamburg Universität, 2016.
- [74] K. Francis and the CALICE Collaboration. Performance of the first prototype of the CALICE scintillator strip electromagnetic calorimeter. *Nuclear Instruments and Methods in Physics Research A* 763 278-289, 2014. URL <http://dx.doi.org/10.1016/j.nima.2014.06.039>.
- [75] 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 A* 608 372-383, 2009. URL <http://dx.doi.org/10.1016/j.nima.2009.07.026>.
- [76] O. Hartbrich. *Scintillator Calorimeters for a Future Linear Collider Experiment*. PhD thesis, Bergische Universität Wuppertal, 2016.
- [77] K. Kotera and the CALICE Collaboration. Construction and Response of a Highly Granular Scintillator-based Electromagnetic Calorimeter. *Nuclear Inst. and Methods in Physics Research, A* 887 150-168, 2018. URL <https://arxiv.org/abs/1707.07126v3>.
- [78] C.Adams. Design, Construction and Commissioning of the Digital Hadron Calorimeter - DHCAL. 2016. URL <https://arxiv.org/pdf/1603.01653.pdf>.

- [79] B.Bilki. Calibration of a digital hadron calorimeter with muons. *JINST 3 P05001*, 2008. URL <http://iopscience.iop.org/article/10.1088/1748-0221/3/05/P05001/pdf>.
- [80] G. Drake, J. Repond, D.Underwood, and L.Xia. Resistive Plate Chambers for hadron calorimetry Tests with analog readout. *Nuclear Instruments and Methods in Physics Research A 578 88 - 97*, 2007. URL <http://dx.doi.org/10.1016/j.nima.2007.04.160>.
- [81] J.Repond and the CALICE Collaboration. Analysis of DHCAL Muon Data. 2012. URL <https://twiki.cern.ch/twiki/pub/CALICE/CaliceAnalysisNotes/CAN-030.pdf>. CALICE Analysis Note CAN-30.
- [82] B. Bilki and the CALICE Collaboration. The CALICE digital hadron calorimeter: calibration and response to pions and positrons. *Journal of Physics: Conference Series 587 012038*, 2015. URL <http://dx.doi.org/10.1088/1742-6596/587/1/012038>.
- [83] G. Baulieu et al. Construction and commissioning of a technological prototype of a high-granularity semi-digital hadronic calorimeter. *JINST 10 P10039*, 2015. URL <http://dx.doi.org/10.1088/1748-0221/10/10/P10039>.
- [84] The CALICE collaboration. First results of the CALICE SDHCAL technological prototype. *JINST 11 P04001*, 2016. URL <http://dx.doi.org/10.1088/1748-0221/11/04/P04001>.
- [85] The CALICE Collaboration et al. Construction and commissioning of the CALICE analog hadron calorimeter prototype. *Journal of Instrumentation 5*, 2010. URL <http://iopscience.iop.org/article/10.1088/1748-0221/5/05/P05004/pdf>.
- [86] Hamamatsu technical guide to silicon photomultipliers, 2018. URL https://www.hamamatsu.com/us/en/community/optical_sensors/articles/technical_guide_to_silicon_photomultipliers_sipm/index.html#.
- [87] E. Garutti. Silicon photomultipliers for high energy physics detectors. *Journal of Instrumentation 6.10*, 2011. URL <http://iopscience.iop.org/article/10.1088/1748-0221/6/10/C10003/pdf>.
- [88] SensL sense light. An Introduction to the Silicon Photomultiplier. Technical report, 2011. URL <https://www.sensl.com/downloads/ds/TN%20-%20Intro%20to%20SPM%20Tech.pdf>.

- [89] B. Dolgoshein. Silicon photomultipliers in particle physics: possibilities and limitation. 2004. Proceedings of the 42nd Workshop of the INFN ELOISATRON Project, Erice, Italy, World Scientific Publishing Co. Pte. Ltd., Singapore (2004), pg. 442.
- [90] Frank Simon. AHCAL Status-Focus on Simulations. ILD Meeting, KEK, 2015. URL <https://agenda.linearcollider.org/event/6557/contributions/31838/attachments/26247/40210/AHCALSimulationStatus.pdf>.
- [91] M. Bouchel et al. SPIROC (SiPM Integrated Read-Out Chip): dedicated very front-end electronics for an ILC prototype hadronic calorimeter with SiPM read-out. *Journal of Instrumentation* 6.01, 2011. URL <http://iopscience.iop.org/article/10.1088/1748-0221/6/01/C01098/pdf>.
- [92] Omega. SPIROC2 User Guide. Draft, 2009.
- [93] O. Hartbrich. Commissioning and LED System Tests of the Engineering Prototype of the Analog Hadronic Calorimeter of the CALICE Collaboration. Master's thesis, Bergische Universität Wuppertal, 2012.
- [94] K. Kotera on behalf of the CALICE collaboration. Scintillator Strip ECAL Optimization. 2014. URL <https://arxiv.org/pdf/1404.1672.pdf>.
- [95] N. Abgrall. NA61/SHINE facility at the CERN SPS: beams and detector system. *JINST* 9 P06005, 2014. URL <http://iopscience.iop.org/article/10.1088/1748-0221/9/06/P06005/pdf>.
- [96] K A Olive et al. (Particle Data Group). Review of Particle Physics. *Chinese Physics C* 38, 090001, 2014. URL <http://pdg.lbl.gov/2015/download/rpp2014-Chin.Phys.C.38.090001.pdf>.
- [97] D. Dannheim K. Elsener W. Klempt A. Lucaci Timoce E. van der Kraaij. Particle Identification with Cherenkov detectors in the 2011 CALICE Tungsten Analog Hadronic Calorimeter Test Beam at the CERN SPS. *LCD-Note-2013-006*, 2013. URL <https://cds.cern.ch/record/1545809/files/LCD-Note-2013-006.pdf>.
- [98] J. Litt and R. Meunier. Cerenkov Counter Technique in High-Energy Physics. *Annual Review of Nuclear Science* 23.1, 1973. URL <https://doi.org/10.1146/annurev.ns.23.120173.000245>.
- [99] E. Brianne. *Influence of timing cuts in Testbeam data and Simulation*. PhD thesis, Universität Hamburg.

- [100] S. Schröder. Commissioning of a prototype hadronic calorimeter. Master's thesis, Universität Hamburg, 2015.
- [101] Y. Sudo. Private Communication, 2016.
- [102] E. Brianne. Studies of the front-end electronics of the Analog HCAL. *DESY summer student report*, 2012. URL http://www.desy.de/f/students/2012/reports/Eldwan_Brianne.pdf.gz.
- [103] O. Hartbrich. Investigation of the time measurement capabilities of the SPIROC2b ASIC. 2011. URL <http://www.desy.de/f/students/2011/reports/hartbrich.pdf.gz>. DESY summer students report.
- [104] C. Günter. *Comparison of Iron and Tungsten Absorber Structures for an Analog Hadron Calorimeter*. PhD thesis, Universität Hamburg, 2014.
- [105] C. Adloff and the CALICE Collaboration. The Time Structure of Hadronic Showers in highly granular Calorimeters with Tungsten and Steel Absorbers. *arXiv:1404.6454*, 2014. URL <http://dx.doi.org/10.1088/1748-0221/9/07/P07022>.
- [106] A. Benaglia and E. Auffray. Space-Time Development of Electromagnetic and Hadronic Showers and Perspectives for Novel Calorimetric Techniques. *IEEE Transactions On Nuclear Science*, Vol. 63, No. 2, 2016. URL <http://dx.doi.org/10.1109/TNS.2016.2527758>.
- [107] The CALICE Collaboration. Pion Showers in the CALICE AHCAL Prototype. 2011. URL <https://twiki.cern.ch/twiki/pub/CALICE/CaliceAnalysisNotes/CAN-026.pdf>. CALICE Analysis note CAN-026.
- [108] G. Moortgat-Pick et al. Physics at the e+e- linear collider. *Eur. Phys. J. C (2015) 75:371*, 2015. URL <https://doi.org/10.1140/epjc/s10052-015-3511-9>.
- [109] M. Chera. *Particle Flow: From First Principles to Gaugino Property Determination at the ILC*. PhD thesis, Hamburg Universität, 2018.
- [110] P. Eckert et al. Characterisation Studies of Silicon Photomultipliers. 2010. URL <https://arxiv.org/pdf/1003.6071.pdf>.
- [111] K. Krüger and the CALICE collaboration. Prototype tests for a highly granular scintillator-based hadron calorimeter. *J. Phys.: Conf. Ser. 587 0120331*, 2015. URL <http://iopscience.iop.org/article/10.1088/1742-6596/587/1/012033/pdf>.

- [112] B. Hermberg on behalf of the CALICE collaboration. Commissioning of the Testbeam Prototype of the CALICE Tile Hadron Calorimeter. 2012. URL <https://arxiv.org/pdf/1211.6290.pdf>.
- [113] M. Terwort on behalf of the CALICE collaboration. Realization and Test of the Engineering Prototype of the CALICE Tile Hadron Calorimeter. 2010. URL <https://arxiv.org/pdf/1011.4760.pdf>.
- [114] A. Lucaci-Timoce (on the behalf of the CALICE collaboration). Calibration and monitoring of a scintillator HCAL with SiPMs CALICE scintillator HCAL. *J. Phys.: Conf. Ser.* 160 012027, 2009. URL <http://iopscience.iop.org/article/10.1088/1742-6596/160/1/012027/pdf>.
- [115] H. L. Tran. New ITEP tiles study. Internal Note, 12.09.2014.