# Comparison of Two Highly Granular Hadronic Calorimeter Concepts

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

vorgelegt von Coralie Neubüser
Hamburg, 2016

Datum der Disputation:

26.10.2016

Folgende Gutachter empfehlen die Annahme der Dissertation:

Prof. Dr. Erika Garutti

Dr. Katja Krüger

Vorsitzender des Prüfungsausschusses:

Prof. Dr. Günter H. W. Sigl

Vorsitzender des Promotionsausschusses:

Prof. Dr. Wolfgang Hansen

*To my family*

*Hilka*

*Gerd*

*Annita*

*and Matteo*

## Abstract

The CALICE collaboration develops hadron calorimeter technologies with high granularity for future electron-positron linear colliders. These technologies differ in active material, granularity and their readout and thus their energy reconstruction schemes. The Analogue Hadron Calorimeter (AHCAL), based on scintillator tiles with Silicon Photomultiplier readout, measures the signal amplitude of the energy deposition in the cells of at most $3 \times 3\,\mathrm{cm}^2$ size. The Digital, Resistive Plate Chamber (RPC) based, HCAL (DHCAL) detects hits above a certain threshold by firing pad sensors of $1 \times 1\,\mathrm{cm}^2$. A 2 bit readout is provided by the, also RPC based, Semi-Digital HCAL (SDHCAL), which counts hits above three different thresholds per $1 \times 1\,\mathrm{cm}^2$ pad. All three calorimeter concepts have been realised in $1\,\mathrm{m}^3$ prototypes with interleaved steel absorber and tested at various test beams.

The differences in active medium, granularity and readout have different impacts on the energy resolution and need to be studied independently.

This analysis concentrates on the comparison between these technologies by investigating the impact of the different energy reconstruction schemes on the energy resolution of the AHCAL testbeam data and simulation. Additionally, a so-called software compensation algorithm is developed to weight hits dependent on their energy content and correct for the difference in the response to the electromagnetic and hadronic sub-showers ($e/h \neq 1$) and thus reduce the influence of fluctuations in the $\pi^0$ generation. The comparison of the energy resolutions revealed that it is mandatory for the AHCAL with $3 \times 3\,\mathrm{cm}^2$ cell size to have analogue signal readout, to apply the software compensation algorithm and thus achieve the best possible energy resolution.

The effect of the granularity is studied with a simulation of the AHCAL with $1 \times 1\,\mathrm{cm}^2$ cell size, and it has been found that to achieve the best possible energy resolution the semi-digital energy reconstruction is sufficient.

To study the impact of the active medium, the DHCAL testbeam data was calibrated and the simulation was tuned using the muon and positron data. The energy resolutions, achieved by the DHCAL data and simulation and achieved by the $1 \times 1\,\mathrm{cm}^2$ AHCAL simulation using the digital energy reconstruction, are successfully used to investigate the influence of the active medium.

Finally, the energy resolutions of the data and simulations of the AHCAL, DHCAL and SDHCAL are compared and the influences discussed.

# Kurzfassung

Die CALICE Kollaboration entwickelt hadronische Kalorimeter mit hoher Granularität für Experimente an zukünftigen Elektron-Positron Linear-Beschleunigern. Die Kollaboration arbeitet an unterschiedlichen Kalorimeter-Konzepten, die sich in ihrem aktiven Medium, ihrer Auslese, Auslese-Granularität und ihren Energierekonstruktions-Methoden unterscheiden. Das Analoge Hadronische Kalorimeter (AHCAL) benutzt $3 \times 3\,\mathrm{cm}^2$ große, szintillierende Plastik-Kacheln, welche es mit Silizium-Photomultipliern (SiPMs) analog ausliest. Das Digitale Hadronische Kalorimeter (DHCAL) besteht aus Widerstandsplattenkammern (RPCs), welche mit $1 \times 1\,\mathrm{cm}^2$ großen Pad-Dioden digital ausgelesen werden. Das Semi-Digitale Hadronische Kalorimeter (SDHCAL), welches ebenfalls RPCs als aktives Medium nutzt, liest die Signale der $1 \times 1\,\mathrm{cm}^2$ großen Pad-Dioden mit 3 Schwellen aus. Diese drei Kalorimeter-Konzepte wurden in $1\,\mathrm{m}^3$ großen Prototypen realisiert und mit Stahlabsorber in Teststrahl-Kampagnen getestet.

Diese Arbeit befasst sich mit der Frage, wie die drei Unterschiede: aktives Medium, Auslese und Granularität, die Energieauflösung beeinflussen. Hierzu wurden die unterschiedlichen Energierekonstruktions-Methoden auf die AHCAL Daten und die AHCAL Simulation angewandt und der Effekt auf die Energieauflösung untersucht. Zusätzlich wurde ein Kompensations-Algorithmus entwickelt, der Hits anhand ihrer Energiedichte gewichtet, wodurch die unterschiedlich hohe Sensitivität auf elektromagnetische und hadronische Teilchenschauer-Komponenten korrigiert wird. Hierdurch kann der Effekt von Fluktuationen in der $\pi^0$ Erzeugung unterdrückt werden. Die bestmögliche Energieauflösung des AHCAL kann nur mit analoger Signalauslese erreicht werden und benötigt die Anwendung des Kompensations-Algorithmus'.

Der Effekt einer feineren Granularität der Zellen wurde mit einer Simulation des AHCAL mit $1 \times 1\,\mathrm{cm}^2$ großen Kacheln untersucht. Es konnte festgestellt werde, dass eine semi-digitale Auslese eines HCALs mit $1 \times 1\,\mathrm{cm}^2$ großen Szintilator-Kacheln ausreicht um die bestmögliche Energieauflösung zu erreichen.

Zum Vergleich der aktiven Medien wurden die DHCAL Teststrahldaten kalibriert und die DHCAL Simulation den Myonen- und Positronen-Daten angepasst. Die erzielten Energieauflösungen konnten mit der AHCAL Simulation mit einer Granularität von $1 \times 1\,\mathrm{cm}^2$ verglichen und so der Effekt des aktiven Mediums untersucht werden.

Letztlich wird ein Vergleich der Energieauflösungen der Daten und der Simulation des AHCALs, DHCALs und SDHCALs gezeigt und diskutiert.

# Contents

# Introduction

In order to validate the theoretical models of particle physics, like e.g the Standard Model, High Energy Physics (HEP) experiments are constructed to achieve best possible measurements of the fundamental particle's properties. These measurements are done by large detectors, which demand, since the beginning of large accelerator facilities, the highest technical standards and new technological developments.
Charged particles are usually detected by a tracker within a magnetic field, that forces the particles on a trajectory dependent on their momentum and charge sign. The momentum and energy measurement of neutral particles requires a detector design that first increases the probability of the particle to interact or decay into ionising particles and second measures the decay products with a high precision.
These detectors are called calorimeters and are divided into electromagnetic and hadronic, depending on the particle type that is targeted. While in the past the focus lay on the energy resolution for single particles of these devices, current and future calorimeters are designed for high spacial resolution. The origin of this trend is the development of Particle Flow Algorithms (PFAs), which optimise the energy measurement of jets by combining the measurements of the particles' tracks and showers from the tracker and calorimeters. To match the particle tracks, a high segmentation of the calorimeter is beneficial, even though this can be accompanied by a degradation of the energy resolution.

Electromagnetic Calorimeters (ECALs) and Hadronic Calorimeters (HCALs) have been realised as homogenous and sampling calorimeters and using different active media. Sampling calorimeters allow a more compact design, which is especially important in case of high particle energies, using high Z absorber material between the sensitive layers. The calorimeter designs and technological choices made in HEP experiments are always based on the physics studies that are to be performed. A few concepts and their impact on the physics studies are discussed in the following:
For the current hadron collider experiments at the Large Hadron Collider (LHC), the calorimeter technologies have been refined to suite the requirement of high radiation tolerance. The Compact Muon Solenoid (CMS) experiment e.g. uses an ECAL consisting of scintillating lead-tungstate ($PbWO_4$) crystals [1]. This homogenous electromagnetic calorimeter achieves an energy resolution of 3 to 4 % and played a major role in the discovery of the Higgs boson in the

di-photon channel [2]. The other multi-purpose experiment at the LHC is the A Toroidal LHC ApparatuS (ATLAS) experiment, with an ECAL that uses liquid argon as active material, interleaved by Pb absorbers. This sampling ECAL achieves an energy resolution of $10\,\%/\sqrt{\mathrm{E}}$. However, this ECAL has been designed to achieve hermeticity in $\varphi$, high particle identification capabilities and good angle measurements [3].
A large number of searches for physics beyond the standard model at the LHC require an excellent understanding and suppression of the standard model background to be able to distinguish between the low rate signals and the much larger number of standard model processes. The standard model backgrounds of strong interactions via Quantum Chromodynamics (QCD) processes are usually dominating at a hadron pp collider, which demands an excellent particle identification. Jets, the products of the hadronisation of quarks and gluons, are measured in the electromagnetic and the hadronic calorimeters of the experiments.
Both searches for new particles that do not interact with the detectors (e.g superpartners of the quarks) which result in missing momentum and energy and searches for heavy resonances e.g. $\mathrm{Z}'$ that decays into 2 jets, highly depend on the jet energy resolution.
The HCAL of the ATLAS experiment consists of scintillator tiles interleaved with iron plates and achieves an energy resolution for single hadrons of $58\,\%/\sqrt{\mathrm{E}}$ [4], which is better than the $\sim 85\,\%/\sqrt{\mathrm{E}}$ of the scintillator-brass sampling HCAL of CMS [5]. The jet energy resolution, however, is influenced by the pile up, the jet finding algorithm and the $\eta$ coverage of the detectors. Additionally, PFAs can be applied to improve the sensitivities of the measurements in many searches, as has been shown successfully by CMS [6].

Complementary to the physics program at the LHC, proposed lepton colliders like the International Linear Collider (ILC) or Compact Linear Collider (CLIC) open the possibility of precision measurements of the Higgs boson by searches in final states that are not measurable at the LHC. These future linear electron-positron colliders require a jet energy resolution of $3-4\,\%$ for a wide range of jet energies. This can be achieved by using PFAs for the jet reconstruction. Within the CALICE collaboration, several concepts for high-granularity HCALs optimised for Particle Flow are studied and have been tested with large, $\sim 1\,\mathrm{m}^3$ prototypes: the so-called Analogue, Digital and Semi-Digital HCAL. The prototypes differ in active material for the shower detection, granularity, readout technology and energy reconstruction procedure. This makes it difficult to disentangle the influence of each of these components on the energy resolution of jets as well as of individual particles.
In this thesis, in order to compare these different technologies, the AHCAL is simulated with the cell size of the D- and SDHCAL of $1 \times 1\,\mathrm{cm}^2$ and the impact on the energy resolution is investigated. Additionally the energy reconstruction procedures are studied and compared using the data taken by the Analogue HCAL prototype, which provides the necessary analogue signal information. The effect of the active media can be studied only in simulation. For the comparison with the AHCAL results, the DHCAL testbeam data has been calibrated, electro-

magnetic and hadronic showers have been analysed and, for the first time, the testbeam data is simulated in very detail. In the end the energy resolutions for single hadrons of all three calorimeter prototypes are directly compared.

This thesis is based on the testbeam data of the CALICE hadron calorimeter prototypes. In Chapter 1 the fundamental particle interactions are summarised. The CALICE prototypes are introduced, discussed and compared in Chapter 2. The AHCAL and DHCAL testbeam simulations are based on the particle shower models available by the GEANT4 software package. These models are summarised and discussed in Chapter 3. In Chapter 4 the different energy reconstruction procedures are presented. The impact of the different energy reconstruction procedures on the AHCAL testbeam data as well as the validation of the AHCAL simulations are discussed in Chapter 5. The studies of the electromagnetic and hadronic showers in the DHCAL prototype are presented in Chapter 6. Also the implementation of the simulation is described in detail.
The final comparison of the Analogue, Digital, and Semi-Digital HCAL concepts is shown in terms of the energy resolutions for single hadrons in Chapter 7.

# Chapter 1

# Calorimeters for HEP Collider Experiments

High energy physics explores an energy range of $\sim 9$ orders of magnitude from the MeV to PeV scale by accelerator-based and astronomical experiments.
The high energy collisions of accelerator-based experiments are studied by measuring the four momenta of the particles created in the collisions. In this chapter the basic concepts of the interaction of particles with matter are described and the principles of the momentum and energy measurement are discussed. The idea of Particle Flow Algorithms (PFAs) is introduced and its implications on the detector concepts discussed.

## 1.1 Interactions of Particles with Matter

Particles and radiation are detected exploiting their interaction with matter. The main measurable processes of relativistic particles are ionisation, excitation and bremsstrahlung. Neutral particles are only detectable via there decay products or other charged secondary particles produced in an interaction with matter.
In Table 1.1 all the symbols used in this section are summarised.

### 1.1.1 Ionising Energy Loss

Charged particles, when passing through matter, ionise and excite the material's atoms. The particles transfer their energy by scattering on the electrons of the atoms, which can gain enough energy to leave an ionised atom behind. If the atom gets excited, it can emit photons, which is called luminescence and which is often used for the particle detection. For leptons the energy loss is dominated by ionisation while charged hadrons also interact via the strong force, which will be discussed in Section 1.1.4.
The average energy loss of moderately relativistic charged and heavy particles, in the momen-

**Table 1.1:** Summary of the variables used in this section.

| symbol | definition | value or unit |
|---|---|---|
| A | atomic mass | $\mathrm{g \cdot mol^{-1}}$ |
| c | speed of light | $3 \times 10^8\,\mathrm{m/s}$ |
| E | incident particle energy $\gamma \mathrm{Mc}^2$ | MeV |
| $\mathrm{E_c}$ | critical energy for electrons | MeV |
| M | incident particle mass | $\mathrm{MeV/c^2}$ |
| $\mathrm{m_e c^2}$ | electron mass $\times\, \mathrm{c}^2$ | $0.510998\,\mathrm{MeV}$ |
| $\mathrm{N_A}$ | Avogadro's number | $6.022141 \times 10^{23}\,\mathrm{mol^{-1}}$ |
| $\mathrm{r_e}$ | electron radius | $2.817940\,\mathrm{fm}$ |
| Z | atomic number | |
| z | charge number of incident particle | |
| $\beta$ | ratio of relative velocity and speed of light $\mathrm{v^2/c^2}$ | |
| $\gamma$ | Lorentz factor $1/\sqrt{1-\beta^2}$ | |
| $\rho$ | density of the absorber material | $\mathrm{g \cdot cm^{-3}}$ |

tum range $0.1 \leq \beta\gamma = \mathrm{p/Mc} \leq 1000$, is described by the Bethe equation [7]

$$-\left\langle \frac{\mathrm{dE}}{\mathrm{dx}} \right\rangle \left[\mathrm{MeV\,cm^2/g}\right] = \mathrm{Kz^2} \frac{\mathrm{Z}}{\mathrm{A}} \frac{1}{\beta^2} \left( \frac{1}{2} \ln \frac{2\mathrm{m_e c^2} \beta^2 \gamma^2 \mathrm{T_{max}}}{\mathrm{I^2}} - \beta^2 - \frac{\delta(\beta\gamma)}{2} \right) \tag{1.1}$$

with the mean excitation energy I and the density effect correction function $\delta(\beta\gamma)$. K is a constant given by $4\pi \mathrm{N_A r_e^2 m_e c^2}$, and the maximum energy transfer in one collision $\mathrm{T_{max}} = \frac{2\mathrm{m_e c^2}\beta^2\gamma^2}{1+2\gamma \mathrm{m_e/M} + (\mathrm{m_e/M})^2}$ with the mass of the incoming particle M.
The approximation is precise up to a few percent for particle energies of up to several hundred GeV. Above that, radiative processes dominate the energy loss. For low particle energies the energy loss decreases with $1/\beta^2$ with a minimum around $\beta\gamma \approx 4$, see Figure 1.1. Particles that loose their energy at the minimum rate of the stopping power, e.g. $1.6\,\mathrm{MeV\,cm^2/g}$ in Cu, are called Minimum Ionising Particles (MIPs). The limit of the description of the energy loss for particles with low energies is given by the binding energy of the atom's electrons.
The energy loss of heavy projectiles, like ions, is influenced by higher order photon coupling to the target, for which Equation 1.1 does not account for.

For detectors of moderate thickness (e.g. scintillator tiles), the energy loss probability function is described by the Landau distribution [9]. The energy loss in thin absorber layers (e.g. RPC gas gaps) is not described by the Landau distribution because of larger fluctuations leading to a broadening of the distributions. This will be further discussed in Section 2.3.1.

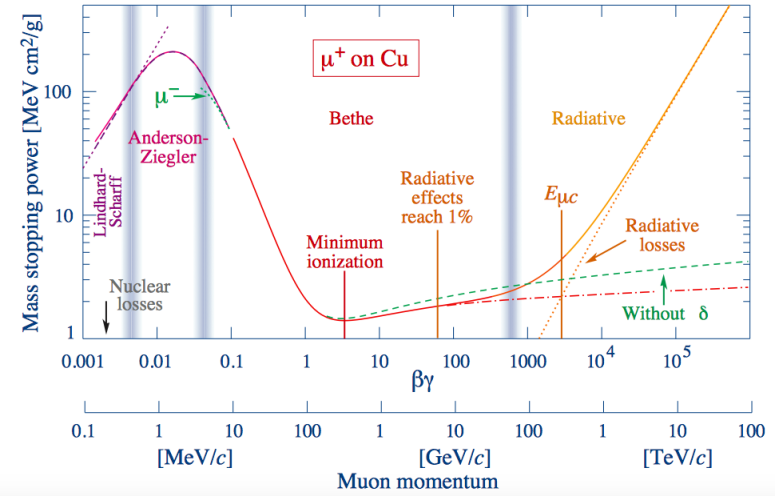



**Figure 1.1:** Stopping power for positive muons in copper as a function of $\beta\gamma = p/Mc$ [8]. The vertical bands indicate the boundaries for different approximation models.

### 1.1.2 Bremsstrahlung and Pair Production

Additionally to ionisation loss, charged particles can interact with the Coulomb field of the nuclei of the traversed medium. The particle decelerates and emits a photon, this process is called bremsstrahlung. Additionally, within the proximity of the Coulomb field of the nuclei electron-positron pairs can be produced via virtual and real photons. This process is called pair production and plays an important role at high energies and especially for muons, which have a higher mass than electrons and loose a larger portion of energy via pair production.
The energy loss due to bremsstrahlung $\left|-\frac{dE}{dx}\right|_{brems}$ is proportional to the particle energy and inverse proportional to its mass square. Therefore highly energetic electrons lose their energy dominantly via bremsstrahlung.
The energy lost by electrons due to bremsstrahlung follows

$$\left|-\frac{dE}{dx}\right|_{brems} = \frac{E}{X_0}, \tag{1.2}$$

and depends on the radiation length $X_0$, which is both the mean distance over which a high-energy electron loses all but 1/e of its energy by bremsstrahlung, and 7/9 of the mean free path for pair production by a high-energy photon. Within an uncertainty of 3 %, $X_0$ can be approximated by [10]

$$X_0 \left[g\,cm^{-2}\right] = \frac{716.4 \cdot A}{Z\,(Z+1)\ln\left(287/\sqrt{Z}\right)}. \tag{1.3}$$

Energy losses due to bremsstrahlung are proportional to the energy, while the ionisation losses are proportional to the logarithm of the energy. The particle energy where both processes occur at the same rate is called the critical energy $E_c$, which is dependent both on the electron

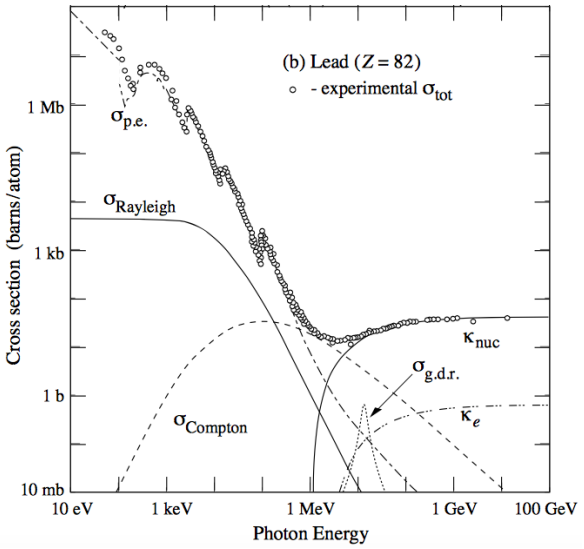



**Figure 1.2:** Photon total cross sections as a function of energy in lead, displaying the contributions of different processes: $\sigma_{\text{p.e.}}$ = photoelectric effect, $\sigma_{\text{Rayleigh}}$ = Rayleigh scattering, $\sigma_{\text{Compton}}$ = Compton scattering, $\varkappa_{\text{nuc}}$ = pair production in nuclear field, $\varkappa_{\text{e}}$ = pair production in electron field, and $\sigma_{\text{g.d.r.}}$ = photo-nuclear interactions [8].

density of the absorber and the mass of the particle [10]

$$\mathrm{E_c} \propto \frac{\mathrm{M}^2}{\mathrm{Z}}. \tag{1.4}$$

### 1.1.3 Interactions of Photons

Photons are usually detected either via the photoelectric effect, Coulomb scattering or pair production. Low energy photons ($\mathrm{E} < 100\,\mathrm{keV}$) loose their energy dominantly via the photoelectric effect, although Compton scattering and Rayleigh scattering also occur. In Figure 1.2 the cross sections of the different processes in lead are summaries for photon energies of 10 eV to 100 GeV. In the energy range of $\mathrm{E} \approx 1\,\mathrm{MeV}$ the Compton effect ($\sigma_{\text{Compton}}$) dominates and for energies higher than 1 MeV, above twice the electron mass $\mathrm{m_e}$, the pair production processes ($\varkappa_{\text{nuc}}$ and $\varkappa_{\text{e}}$) can take place.

For particle energies above $\mathrm{E}_\gamma > 10^{14}\,\mathrm{MeV}$ and $\mathrm{E_e} > 10^{15}\,\mathrm{MeV}$ of photons and electrons respectively the energy loss is dominated by hadronic interactions via photo-nuclear and electro-nuclear processes.

### 1.1.4 Hadronic Interactions

Charged hadrons interact electromagnetically, like described in the previews sections, and can do elastic and inelastic scattering on the nuclei of the absorber. The strong interactions of hadrons are in most cases inelastic and produce secondary particles. The probability for these

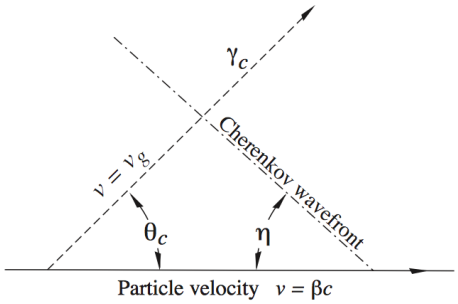



**Figure 1.3:** Schematic of Cherenkov light emission and wavefront angles, where $\eta$ is the opening half-angle and $\nu_g$ is the group velocity of the photons, in a dispersive medium $\vartheta_c + \eta \neq 90°$ [8].

processes to occur is described by the average nuclear interaction length $\lambda_n$ [10]:

$$\lambda_n = \frac{A}{N_A \cdot \rho \cdot \sigma_{inel}}, \tag{1.5}$$

with the cross-section of inelastic interactions $\sigma_{inel}$.
Most generated secondaries are charged or neutral pions ($\pi^{\pm}$ or $\pi^0$) or $\eta$ mesons. Nuclear reactions release protons and neutrons from the nucleus. A large amount of the transferred energy via nuclear interactions is not detected, because of the excitation or recoil of nuclei or its absorption in nuclear binding energy.
More details about hadronic cascades are discussed in Section 1.2.2.

### 1.1.5 Cherenkov Radiation

In case a charged particle traverses a dispersive medium of refractive index n with a velocity $\nu = \beta c$ greater than the local phase velocity of light c/n, Cherenkov light is emitted in a cone with an opening angle of $\vartheta_c$ with respect to the direction of motion. The opening angle is given by

$$\cos\vartheta_c = (1/n\beta). \tag{1.6}$$

In Figure 1.3 the emission direction of the Cherenkov light $\gamma_c$ is illustrated. The contribution of Cherenkov radiation to the energy loss of charged particle is negligible.
Cherenkov threshold counters are used for particle identification in testbeams [11]. These counters are cylindrical gas filled volumes, with photomultiplier readout. The distinction between the particles in a mixed beam is based on the different particle masses. The probability to detect a signal depends on the operating conditions of the Cherenkov counter. The threshold pressure of the gas to observe Cherenkov light for a traversing charged particle of mass M and momentum p is given by

$$P_{th}\ [atm] = \frac{M^2}{2\,(1-n)\cdot p^2}. \tag{1.7}$$

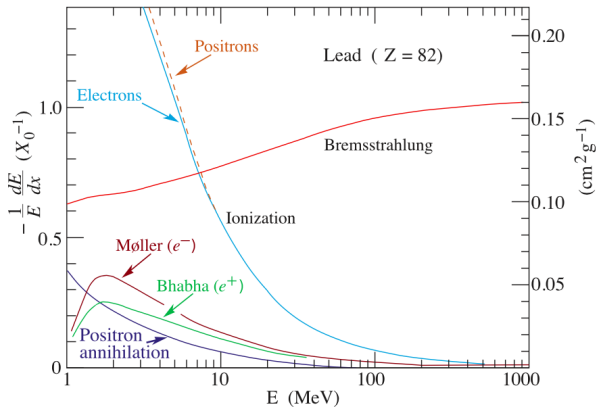



**Figure 1.4:** The fractional energy loss per radiation length in lead as a function of the electron or positron energy [8].

Furthermore, Cherenkov light could be used in hadron calorimetry to measure the electromagnetic content of hadron showers event by event. Which is possible due to the lighter and usually faster particles in the Electromagnetic (EM) part of the hadron showers, see Section 1.2.2. The method of measuring Cherenkov and scintillation light at the same time is called dual-readout and currently under investigation [12]. First tests have been done with the DREAM and RD52 calorimeter which revealed a limitation in the energy resolution due to a very small number of max. $\sim 60$ photoelectrons per GeV in the Photomultiplier Tubes (PMTs) that measure the Cherenkov light [13, 14].

## 1.2 Energy Measurement with Calorimeters

In calorimeters the particle energy is measured by total absorption in the calorimeter volume and the measurement of the deposited energy. In general, calorimeters are grouped in *homogeneous* and *sampling*, *electromagnetic* (E-) and *hadronic* (H-) calorimeters (CALs).
In this section the basic mechanisms of particle shower development and the detection principles are introduced, and the effects that influence the energy resolution are discussed.

### 1.2.1 Electromagnetic Showers

Electrons and photons that traverse matter loose energy due to several electromagnetic processes, mainly ionisation, bremsstrahlung and pair production, introduced in Section 1.1. For low electron and positron energies a small energy fraction is lost via Moeller scattering and Bhabha scattering or annihilation. The fraction of the energy losses in lead for the different interactions are shown in Figure 1.4.
The successive alternation of the emission of bremsstrahlung photons and the $e^+e^-$ pair production leads to the development of an electromagnetic cascade, which is illustrated in Figure 1.5.

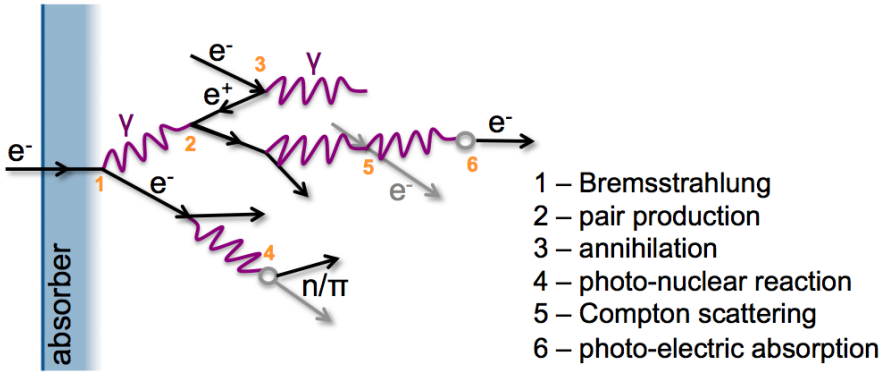



**Figure 1.5:** Schematic of the electromagnetic interactions in electromagnetic cascades: bremsstrahlung (1), pair production (2), annihilation (3), photo-nuclear reactions (4), Compton scattering (5) and photo-electric absorption (6).

The cascade develops and gains particle multiplicity until the critical energy $E_c$ is reached. When the energy of the particles fall below $E_c$, ionisation for electrons and Compton or photoelectric effects for photons dominate.
The electromagnetic shower depth and maximum increases logarithmically with the initial particle energy [7] by

$$t_{max} \propto \ln\left(\frac{E}{E_c}\right), \tag{1.8}$$

and the longitudinal shower development can be described by the Gamma distribution [15]:

$$\frac{dE}{dt} = E \cdot \frac{\left(\frac{t-\mu}{\beta}\right)^{\gamma-1} \cdot \exp\left(-\frac{t-\mu}{\beta}\right)}{\beta\Gamma(\gamma)} \tag{1.9}$$

with the Gamma function $\Gamma$, a shape parameter $\gamma$, a location parameter $\mu$, and a normalisation parameter $\beta$. The shower maximum can be determined as

$$t_{max} = \beta(\gamma - 1) + \mu. \tag{1.10}$$

The lateral extension is mainly determined by multiple scattering and can be best described by the Molière radius which is the radius of a cylinder containing 90 % of the deposited shower energy and can be estimated by [10]

$$\rho_M = 21.2\,\text{MeV} \cdot \frac{X_0}{E_c}. \tag{1.11}$$

The Molière radius has a weaker Z dependence than the radiation length $X_0$, see Equation 1.3. As an example the Molière radius in pure iron is 1.719 cm and in tungsten 0.9327 cm, while the radiation length decreases from 1.757 cm to 0.3504 cm with Z increasing from 26 (Fe) to 74 (W) [16].

### 1.2.2 Hadron Showers

Charged hadrons when traversing matter interact mostly inelastically and in this way generate secondary particles, like described in Section 1.1.4. Secondary hadrons, except for $\pi^0$ and $\eta$ which have a very short lifetime (e.g. $\tau_{\pi^0} = 8.52 \pm 0.18 \times 10^{-17}\,\mathrm{s}$ [7]) and decay into 2 photons, continue their way through the absorber until they interact as well. The particle multiplicity increases and a cascade develops.
Hadron showers have a much larger spatial extension in a given absorber than electromagnetic showers, due to the usually much larger interaction length $\lambda_n$ than radiation length $X_0$. For example in iron the ratio of $\lambda_n/X_0 = 9.5$. The lateral hadron shower dispersion is larger than for electromagnetic showers, due to large transverse energy transfers in nuclear reactions.
A pion e.g., when interacting with an atomic nucleus, produces a number of secondary hadrons of which approximately one third are $\pi^0$s and $\eta$s. These neutral particles have a very short life time and a branching faction of $\Gamma_{\pi^0,i}/\Gamma_{\pi^0} = 99\,\%$ and $\Gamma_{\eta^0,i}/\Gamma_{\eta} = 72\,\%$ into two photons [7]. The rest of the time the neutral $\eta$s decay into $3\pi^0$s or one $\pi^0$ and two $\gamma$s. These processes are the main source for the electromagnetic content of hadron showers. The mean electromagnetic fraction and the similarity between electromagnetic and hadronic cascades increase with the energy of the initial hadron according to a power law

$$f_{EM} = 1 - \left(\frac{E}{E_0}\right)^{k-1}, \tag{1.12}$$

where E is the energy of the hadron initiating the cascade, $E_0$ is the $\pi^0$ production threshold, and $k = 1 + \frac{\ln(1-f_{\pi^0})}{\ln\langle m\rangle}$ is related to the multiplicity of $\pi^0$ mesons $\langle m\rangle$ and the production fraction of $\pi^0$s per interaction $f_{\pi^0}$ [10, 17]. In iron $E_0 = 0.8\,\mathrm{GeV}$ and k is usually of the order of 0.8, which leads to an approximated electromagnetic fraction of around 0.62 (0.56, 0.40) for a 100 GeV (50 GeV, 10 GeV) pion shower.
The fact that neutral pions, generated in the nuclear interactions of charged pions, decay almost instantly into photons and electrons causes an asymmetry in the distribution of the electromagnetic fraction $f_{EM}$ [10].
An example of a hadron shower, initiated by a neutron, is shown in Figure 1.6, where the particles that are invisible for the detector are represented by dashed lines.

The parametrisation of the longitudinal development of hadron showers with a sum of two gamma distributions was proposed in [18] as a natural extension of the parametrisation of electromagnetic shower profiles, see Equation 1.9 in the previous section, following:

$$\frac{dE}{dt} = A\cdot\left(f\cdot\beta_{short}\frac{(\beta_{short}t)^{\alpha_{short}-1}\cdot\exp\left(-\beta_{short}t\right)}{\Gamma\left(\alpha_{short}\right)} + (1-f)\cdot\beta_{long}\frac{\left(\beta_{long}t\right)^{\alpha_{long}-1}\cdot\exp\left(-\beta_{long}t\right)}{\Gamma\left(\alpha_{long}\right)}\right), \tag{1.13}$$

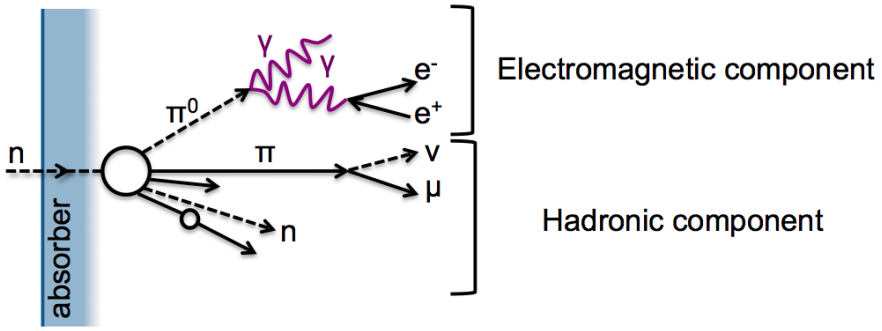



**Figure 1.6:** Schematic of the electromagnetic, strong and hadronic interactions in a hadronic cascade.

with A a scale factor, f the fractional contribution of the short component with shape $\alpha_{\text{short}}$ and slope $\beta_{\text{short}}$ parameters and the shape and slope parameters of the long component $\alpha_{\text{long}}$ and $\beta_{\text{long}}$ [19].
The short component of the parametrisation is related to the electromagnetic sub-showers due to $\pi^0$ decays, while the long component is dominated by secondary particles and has been found to be independent of the type of the incident hadron [19]. The depth of the hadron showers increases with the logarithm of the original hadron energy.

In the previous sections electromagnetic and hadronic shower development have been introduced. The implications on the calorimeters are going to be discussed in more detail in the following.

### 1.2.3 Compensation

The energy measurement of hadrons is compromised by the missing energy due to invisible shower components. These invisible components within a hadron cascade, e.g. neutrinos from $\pi^+$ decays into $\mu^+\nu_\mu$, result in a lower calorimeter response to hadron showers $\pi$ compared to the response to EM showers e of the same energy $e/\pi > 1$. This is called non-compensation. Additionally, this $e/\pi$ ratio decreases with the energy due to the increasing electromagnetic fraction with the energy, see Equation 1.12.
The calorimeter response to the non-electromagnetic shower component of a hadron shower h, e.g. mesons and spallation protons, can be assumed to be constant. While the response to hadrons and electromagnetic showers is experimentally measurable, the ratio of the electromagnetic and non-electromagnetic shower components e/h within hadron showers is not. Both ratios are related via

$$e/\pi = \frac{e/h}{1 - f_{\text{EM}}\,[1 - e/h]}, \tag{1.14}$$

which allows a determination of e/h from experimental data [10].
Homogenous calorimeters, where the constituent material acts both as absorber and as active

medium, have a typical ratio of e/h = 2. Sampling calorimeters, where separate absorbing layers interspersed with active layers, can be designed for compensation e/π = 1. Either the electromagnetic response is suppressed by higher Z absorbers or the response to the non-electromagnetic component is boosted, e.g. by increasing the number of neutrons, which can be achieved by an increase of the hydrogen concentration in the active material.
Another possibility to achieve compensation is to apply software compensation algorithms. These algorithms assign weights to each calorimeter hit dependent on it's energy density to correct for the different calorimeter response to the electromagnetic and hadronic components of the shower. This technique will be further discussed in Section 4.5.

### 1.2.4 Energy Resolution

The energy resolution of a calorimeter is influenced by various factors, but the two reasons for an intrinsic limitation are first the fluctuations in the cascade processes and second the chosen technique to measure the products of the particle cascades.
However, the implications are different for electromagnetic and hadronic calorimeters. The electromagnetic calorimeters do have a natural limitation in the resolution due to fluctuations in the shower development. Hadronic calorimeters on the other hand suffer from even larger fluctuations in e.g. the EM fraction and the fraction of produced invisible particles. However, hadronic calorimeters have the power to recover for some of the fluctuations by the choice of design, readout and algorithms.
The factors contributing to the energy resolution σ(E) are typically split into three terms, differentiated by their energy dependence:

$$\frac{\sigma\,(\mathrm{E})}{\mathrm{E}} = \frac{\mathrm{a}}{\sqrt{\mathrm{E}}} \oplus \mathrm{b} \oplus \frac{\mathrm{c}}{\mathrm{E}}, \tag{1.15}$$

where a is the stochastic term, b is the constant term and c is the noise term. The effects contributing to these terms are discussed in the following.

#### Stochastic term

The number of particles N contributing to the calorimeter signal follow the Poisson statistics, which leads to an uncertainty of $\sqrt{\mathrm{N}}$ with $\sigma(\mathrm{N})/\mathrm{N} = 1/\sqrt{\mathrm{N}}$ and a stochastic uncertainty in the energy resolution of

$$\left|\frac{\sigma\,(\mathrm{E})}{\mathrm{E}}\right|_{\mathrm{stochastic}} = \frac{\mathrm{a}}{\sqrt{\mathrm{E}}}. \tag{1.16}$$

This is true for calorimeters with a signal formation in the active media that is proportional to the number of particles N = kE.
In case of sampling calorimeters only a fraction of the deposited energy is measured, which introduces another uncertainty. See more about sampling fluctuations in Section 1.2.5.
Additional uncertainties arise for hadron showers due to

- fluctuations in the fraction of invisible energy, e.g. energy lost in binding energy, which also scales with $1/\sqrt{E}$.
- fluctuations in the response due to different electromagnetic fractions, which scale with $\sigma(E)/E \propto E^{-j}$, $j < 0.5$, and are avoidable by either compensation, see Section 1.2.3 or by measuring $f_{EM}$ event by event [13].

**Constant term**

A constant term that limits the energy resolution occurs in cases where the shower is not fully contained in the calorimeter. This leads to missing energies and because of the energy dependent shower width and depth to a constant term in

$$\left|\frac{\sigma(E)}{E}\right|_{\text{constant}} = b. \tag{1.17}$$

Additionally, detector inhomogeneities, like e.g uncertainties on the calibration, deteriorate the resolution energy dependently.

**Noise term**

The detectors noise is energy independent and results in an uncertainty on the energy measurement of c, thus

$$\left|\frac{\sigma(E)}{E}\right|_{\text{noise}} = \frac{c}{E}. \tag{1.18}$$

The origins of detector noise are various and arise from different effects. As an example too low thresholds can result in a high noise rate.

Another effect, resulting in an uncertainty on the energy measurement is signal saturation in the calorimeter readout, which will affect the constant and the noise term and occurs due to e.g. a conversion of analogue to digital signals. The effects will be discussed in Section 2.2.2 as well as in Section 6.3.3.

### 1.2.5 Sampling Calorimeters

In sampling calorimeters, where layers of active media are interleaved by absorber layers, typically only a small fraction of the energy carried by the entering particles is deposited in the active medium which generates a signal. The advantage of sampling calorimeters is a more compact design due to a small $X_0$ achieved by high Z absorber layers. The drawback is a degraded energy resolution compared to homogenous calorimeters, because only a fraction of the deposited energy is measured.
The major effects on the energy resolution of EM showers are energy leakage and sampling fluctuations. In contrast to hadron showers, the energy resolution of EM showers has only a

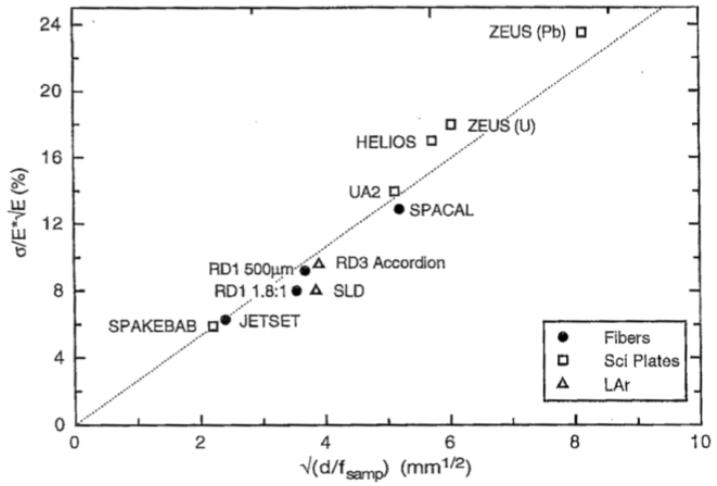



**Figure 1.7:** The EM energy resolution of sampling calorimeters as a function of $\left(\mathrm{d}/\mathrm{f}_{\mathrm{sampling}}\right)^{1/2}$ [10].

small intrinsic limitation due to fluctuations in the shower particles.
The signal of sampling calorimeters is the sum of all shower particles contributing to the signals in the active calorimeter layers. The fluctuation of the number and type of shower particles that contribute to the signals in the active layers affects the energy resolution of the calorimeter. These fluctuations follow the Poisson statistic and therefore the contributions of the sampling fluctuations to the energy resolution is described by Equation 1.16. The impact of these fluctuations on the energy resolution for EM showers is directly proportional to the sampling fraction $\mathrm{f}_{\mathrm{sampling}}$, which is defined as the fraction of the deposited energy measurable in the active medium, determined for a MIP:

$$\mathrm{f}_{\mathrm{sampling}} = \frac{\mathrm{E}_{\mathrm{MIP}}(\mathrm{active})}{\mathrm{E}_{\mathrm{MIP}}(\mathrm{active}) + \mathrm{E}_{\mathrm{MIP}}(\mathrm{passive})}. \tag{1.19}$$

The stochastic term of the resolution can by expressed in $\mathrm{f}_{\mathrm{sampling}}$ by

$$\mathrm{a}_{\mathrm{stochastic}} = \mathrm{a}\sqrt{\mathrm{d}/\mathrm{f}_{\mathrm{sampling}}}, \tag{1.20}$$

where d is the active layer thickness. Thus by reducing the absorber layer thickness by a factor of 2, an improvement of $\sqrt{2}$ is expected. The EM resolution of a number of calorimeters using scintillator plates, scintillating fibres and Liquid Argon (LAr) as active material is summarised in Figure 1.7 and reveals a relation of $\mathrm{a}_{\mathrm{stochastic}} = \frac{2.7\,\%}{\sqrt{\mathrm{mm}}}\sqrt{\mathrm{d}/\mathrm{f}_{\mathrm{sampling}}}$ [20].

The impact of sampling fluctuations on hadrons is considerably larger due to the larger number of traversed layers to deposit the same energy as an electromagnetic shower. With an interaction length of approximately $0.1\,\lambda_{\mathrm{n}}$ per layer, a pion traverses 10 layers before it undergoes a nuclear interaction. The impact of sampling fluctuations on the energy resolution was measured from the ZEUS collaboration as twice as large for hadron compared to EM

showers [21]. However, the hadronic energy resolution of sampling calorimeters is usually dominated by the fluctuations in the visible energy and by fluctuations in the electromagnetic and hadronic shower components if the calorimeter is non-compensating.

## 1.3 Particle Flow Algorithms

In particle collisions of HEP experiments not only single particles are measured, but quantities related to particles produced in the hard interaction, e.g. quarks and gluons. High energetic quarks and gluons hadronise, which means they are fragmenting into several hadrons. The produced particles are boosted into the original direction of the quark and are called jets. The hadron that contains the original quark carries the largest fraction of its energy and can be a charged or a neutral hadron.
Large HEP collider experiments traditionally measure the jet energies as the energy sum of the energy deposits in the calorimeters within a region around the jet axis. Especially for highly energetic particles the calorimeter characteristic of an energy resolution improving with higher energies, following $1/\sqrt{\mathrm{E}}$ see Equation 1.15, is persuasive. However, typically 70 % (62 % charged and 10 % neutral jet particles [22]) of the particle energies is measured in the HCAL, where a typical sampling calorimeter has an energy resolution of $60\,\%/\sqrt{\mathrm{E}}$ that is much worse than for the tracker and the ECAL, see the fourth column in Table 1.2. The benefit of a good resolution for very high particle energies is usually defeated by the distribution of the jet energies to a large number of particles and energy leakage due to the increasing longitudinal expansion of highly energetic hadron showers.
In contrast to a purely calorimetric measurement, Particle Flow Algorithms (PFAs) require the reconstruction of the four-vectors of all visible particles in an event in the sub-detector with the best possible energy resolution. The reconstructed jet energy is the sum of the energies of the individual particles. 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 10 % of the energy in a jet. In this way a jet energy resolution of $19\,\%/\sqrt{\mathrm{E}}$ [23] is achievable assuming a perfect particle identification with typical tracker, ECAL and HCAL performance, summarised in Table 1.2. In reality the energy depositions get not always associated with the correct particles, which is called confusion. This confusion of particles degrades the PFA reconstruction performance and increases the uncertainty. An example of the impact of confusion on the PFA performance in shown in Figure 1.8 on the example of the simulation of the International Large Detector (ILD) detector planned for the ILC. Since confusion has a stronger effect on the jet energy resolution than the actual HCAL energy resolution, the design of calorimeters optimised for PFA is driven by a fine segmentation of the active medium to increase the spatial resolution. Further implications for the calorimeter designs are discussed in Section 1.3.1.

**Table 1.2:** Contribution of the different particle components to the jet energy resolution in GeV [23]. The table lists the approximate fractions of charged particles, photons and neutral hadrons in a jet of energy $E_j$, and the assumed single particle energy resolutions. Whereas the tracker measures the momentum, for high relativistic energies this is equivalent to the energy.

| Component | Detector | Energy Fraction | Energy Resolution | Jet Energy Res. |
|---|---|---|---|---|
| charged particles ($x^{\pm}$) | tracker | $\sim 0.6\,E_j$ | $10^{-4}\,E^2_{x^{\pm}}$ | $< 3.6 \cdot 10^{-5}\,E_j^2$ |
| photons ($\gamma$) | ECAL | $\sim 0.3\,E_j$ | $0.15\,\sqrt{E_\gamma}$ | $0.08\,\sqrt{E_j}$ |
| neutral hadrons ($h^0$) | HCAL | $\sim 0.1\,E_j$ | $0.60\,\sqrt{E_{h^0}}$ | $0.17\,\sqrt{E_j}$ |

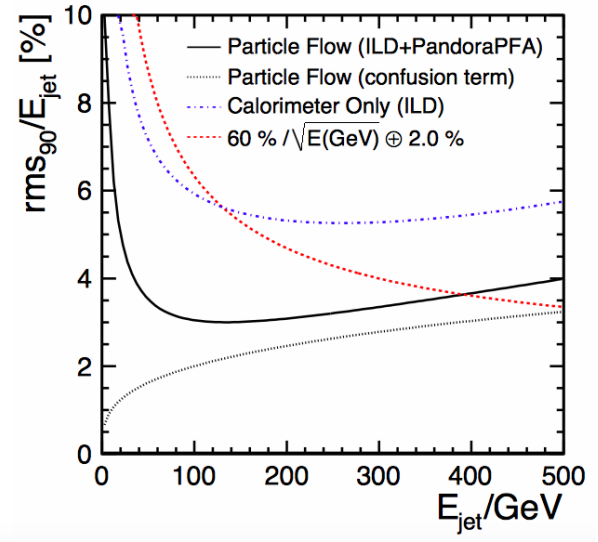



**Figure 1.8:** The empirical functional form of the jet energy resolution obtained from PFA calorimetry (PandoraPFA and the ILD concept). The estimated contribution from the confusion term only is shown (dotted). The dot-dashed curve shows a parameterisation of the jet energy resolution obtained from the total calorimetric energy deposition in the ILD detector. In addition, the dashed curve, $60\,\%/\sqrt{E(\mathrm{GeV})} \oplus 2.0\,\%$, is shown to give an indication of the resolution achievable using a traditional calorimetric approach [23].

The principle of particle flow algorithms, earlier called Energy Flow, has already been successfully used at the Apparatus for LEP Physics (ALEPH) experiment at the Large Electron Positron Collider (LEP) [22], where an improvement of the resolution for hadronic decays of the Z boson from 120 to $65\,\%/\sqrt{\mathrm{E}}$ was achieved. Similar algorithms have been used by many HEP experiments since. Currently the CMS experiment at the LHC uses a particle flow algorithm and achieves e.g. for di-jet events a twice to three times better jet energy resolution than with the traditional particle reconstruction [6, 24].

### 1.3.1 Implications for Detectors of Future HEP Experiments

Future collider experiments like the ILC or CLIC, which are both $\mathrm{e^+e^-}$ colliders, aim for a jet energy resolution of 3-4 % over a wide range of jet energies. This jet energy resolution translates into an energy resolution of the hadron calorimeter of $30\,\%/\sqrt{\mathrm{E}}$ [23], which is 5 % better than the best resolution ever archived by a sampling calorimeter [25]. This high precision is needed to measure the Higgs mass and e.g. the triple Higgs coupling. Especially the Higgs mass measurement in the 4 jet channel $\mathrm{e^+e^-} \rightarrow \mathrm{ZH} \rightarrow \mathrm{qqbb}$ benefits from an excellent jet energy resolution. The unique measurement of the Higgs self-coupling by the reaction $\mathrm{e^+e^-} \rightarrow \mathrm{ZHH} \rightarrow \mathrm{qqbbbb}$ at $\sqrt{\mathrm{s}} = 500\,\mathrm{GeV}$ requires an excellent W,Z and H boson identification with high track multiplicity and 6 jets, and strongly depends on a correct assignment of particles possible by a fine transverse segmentation of the calorimeters [26]. To make these studies possible, the detector designs have to be optimised for optimal PFA performance.
The PFA performance requires the E- and HCAL to be placed within the solenoid providing the magnetic field for the track measurement. In addition to the requirement on the segmentation of the calorimeters, which depends mostly on the Molière radius, the radiation and interaction length of the absorber material, the inner radius of the calorimeters and the magnetic field strength are of importance for the PFA performance. The magnetic field will deflect charged particles from their original trajectory, which leads to a better separation of charged and neutral particles in a jet and thus reduces the confusion. The inner radius of the ECAL and thus the impact on the path length affects also the separation and thus the choice of lateral segmentation. It has been found that a larger outer radius of the calorimeters is favoured over a stronger magnetic field [23].
A definite test of the PFA performance for these optimised detectors is only possible in a full sized experiment. However, test beam data of detector prototypes allow to some extend a test of reconstruction efficiencies by merging two testbeam events and determining the shower separation capabilities of the algorithm [27].

# Chapter 2

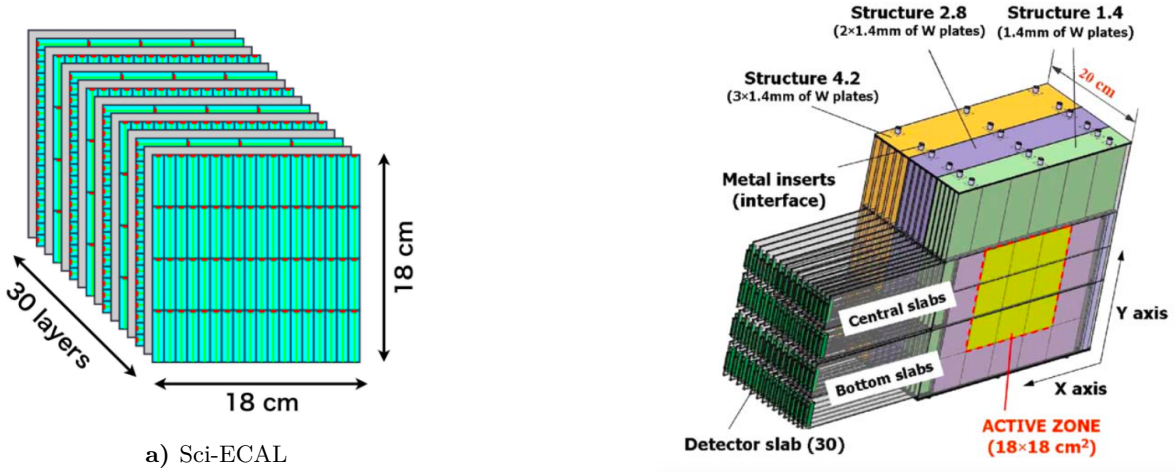
# CALICE Detector Concepts

Within the CALICE collaboration electromagnetic as well as hadronic calorimeter systems are developed and tested. While the CALICE detectors are optimised for a linear collider environment such as the International Linear Collider (ILC) [28, 29] and Compact Linear Collider (CLIC) [30, 31], cooperations between the Large Hadron Collider (LHC) upgrade groups and CALICE members are ongoing [32]. The CALICE calorimeters are imaging calorimeters with high granularity, optimised for the particle flow paradigm. Several prototypes of electromagnetic and hadronic calorimeters were realised and tested, also in joined testbeam campaigns [33–35]. The three large hadronic prototypes of approximately $1\,\mathrm{m}^3$ not only provide different active materials and absorbers, but also different read-out schemes and thus need different energy reconstruction procedures.
In this chapter the electromagnetic calorimeters are shortly introduced, and the three hadronic calorimeter technologies that use steel (Fe) absorber are discussed and compared in detail. The Fe absorber option is currently favoured for ILC and CLIC [36, 37].
Additionally the Tail Catcher and Muon Tracker (TCMT), which is often placed behind the HCAL in testbeam campaigns is introduced.

## 2.1 CALICE ECALs

The CALICE collaboration is developing two different ECAL designs in physics prototypes, whose purpose is to prove their functionality for detailed measurements of Electromagnetic (EM) showers: The scintillator based Sci-ECAL [38] and the silicon based Si-ECAL [39], both using tungsten (W) as absorber material. Both prototypes have been operated in various testbeam campaigns [33–35].
The Sci-ECAL consists of 30 active, scintillator layers and 30 passive, tungsten absorber plates of 3.5 mm thickness. Each active layer consists of four rows of 18 scintillator strips with dimensions $45 \times 10 \times 3\,\mathrm{mm}^3$. The strips are oriented orthogonally in consecutive layers. A schematic of the prototype is shown in Figure 2.1a. The light produced in each strip is guided

by a Wavelength-Shifting (WLS) fibre to a Silicon Photomultiplier (SiPM) from Hamamatsu[1] (Multi Pixel Photon Counter (MPPC)) and read out. The prototype has thus in total 2,160 readout channels. The total thickness of the Sci-ECAL prototype is $20\,X_0$ radiation or around $1\,\lambda_n$ nuclear interaction lengths. The Molière radius is approximatively 22 mm [38]. The prototype was operated in testbeams at Fermilab and shows a relative energy resolution with a stochastic term of $15.15\,\%/\sqrt{E}$ and a constant term of 1.44 % [40].
The Si-ECAL prototype also consists of 30 active and 30 passive layers. The thickness of the tungsten absorber plates increases with depth from 1.4 mm of the first 10 layers to 2.8 mm of the second 10 layers and 4.2 mm thickness of the last 10 layers. The active silicon layers of $18 \times 18\,cm^2$ are segmented into $1 \times 1\,cm^2$ pad diodes, which results in 9,720 readout channels. The total thickness of the Si-ECAL prototype is $24\,X_0$ or $1\,\lambda_n$. In testbeams at DESY and CERN this prototype achieved an energy resolution of $16.53\,\%/\sqrt{E}$ stochastic and 1.07 % constant term [39].




**a)** Sci-ECAL

**b)** Si-ECAL

**Figure 2.1:** The schematics of the scintillator-strip ECAL [40] and the silicon-tungsten ECAL [39] prototypes.

## 2.2 Analogue HCAL

The Analogue Hadron Calorimeter is a scintillator sampling calorimeter, consisting of 38 active layers inserted in a $1\,m^3$ steel structure of 39, 17.4 mm thick, steel absorber plates. One active layer consists of a steel cassette, housing 216 or 141 scintillator tiles, that are connected to a Printed Circuit Board (PCB). A schematic drawing is shown is Figure 2.2a. The square tiles are 5 mm thick and have different sizes of $3 \times 3$, $6 \times 6$ and $12 \times 12\,cm^2$, a picture of these tiles is shown in Figure 2.2b. The tile size was studied in simulations before the realisation of

[1] Hamamatsu Photonics K. K., `http://www.hamamatsu.com/`

the prototype and a $3 \times 3\,\mathrm{cm}^2$ transverse segmentation was found to achieve a satisfying two particle separation [41]. These tiles are made of organic scintillator (BASF 143 polystyrene produced by UNIPLAST[2]). The scintillation light spectrum peaks at 430 nm. The average energy loss of muons in polystyrene is shown in Figure 2.3, which illustrates that muons within the energy range of 0.4 to 100 GeV loose a similar amount of energy via ionisation in polystyrene. Radiative processes are only expected for muons with more than 25 GeV traversing the Fe absorber [8].

A 1 mm thick WLS fibre is inserted in a circular groove on the tile. In this way the photons are captured and lead to SiPMs, that amplify and read out the signals. A mirror is glued at the other end of the fibre to reflect the photons to the SiPM. To keep light crosstalk between neighbouring tiles as small as possible the edges are painted with white colour.

In total the Fe-AHCAL has a thickness of $\lambda_n = 5.3$ nuclear, $\lambda_\pi = 4.3$ pion interaction lengths, and 47.2 radiation lengths $X_0$. The details of the composition are summarised in Table 2.1.

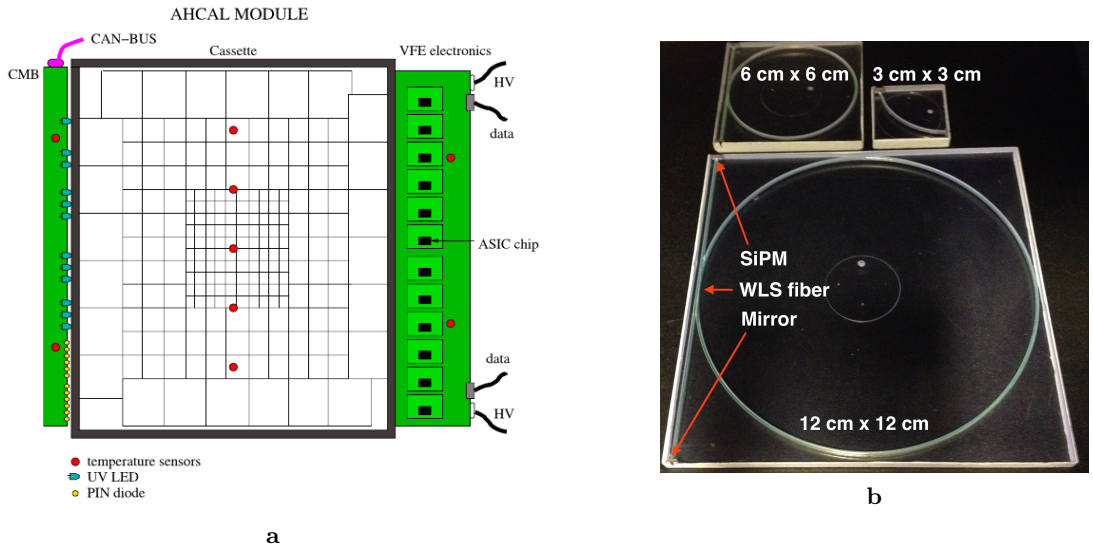



**Figure 2.2:** a) Schematic of one AHCAL module [42]. b) The 3 different tiles of $3 \times 3$, $6 \times 6$ and $12 \times 12\,\mathrm{cm}^2$ with embedded WLS fibres, SiPMs, and mirrors. The edges are painted with white colour.

### 2.2.1 Detection Principle

The scintillation light is produced by ionising radiation. A charged particle traversing the tile excites the scintillator atoms, which relax by emitting low-energetic photons in the visible range, particularly in the wavelengths of blue. The amount of these photons gives a measure of the energy deposited by the incident radiation. Neutrons, produced in hadron showers, scatter predominately on the hydrogen atoms, which gain enough energy to further ionise the scintillator and thus induce scintillation light.

[2] Stahl- und Kunststoff-Produktion GmbH, Krefeld, Germany

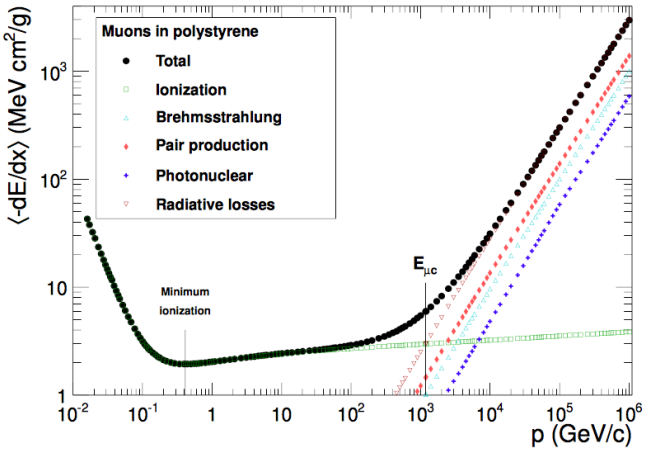



**Figure 2.3:** Average energy loss of muons within the energy range of $10^{-2}$ to $10^{6}$ GeV in Polystyrene [43], computed from [16].

**Table 2.1:** Radiation, nuclear and pion interaction lengths of the Fe-AHCAL [44] per layer and for the whole prototype of 38 layers. Values are taken from the PDG [16], if not differently stated. The air gaps of in total approx. 3 mm thickness per layer and the reflective foil with less than one mm thickness are neglected.

| Material | $X_0$ [g/cm$^2$] | $\lambda_n$ [g/cm$^2$] | $\lambda_\pi$ [g/cm$^2$] | $\rho$ [g/cm$^3$] | thick-ness [cm] | #$X_0$ | #$\lambda_n$ | #$\lambda_\pi$ |
|---|---|---|---|---|---|---|---|---|
| steel absorber [44] | 13.9 | 132.1 | 160.8 | 7.86 | 1.74 | 0.989 | 0.104 | 0.085 |
| steel sheets | 13.9 | 132.1 | 160.8 | 7.86 | $2 \times 0.2$ | 0.226 | 0.024 | 0.020 |
| PCB [44] | 29.8 | 89.45 | 121.4 | 1.7 | 0.1 | 0.006 | 0.002 | 0.001 |
| cable mix | 26.9 | 94.8 | 126.5 | 1.35 | 0.15 | 0.008 | 0.002 | 0.002 |
| tile | 43.8 | 81.7 | 113.7 | 1.06 | 0.5 | 0.012 | 0.007 | 0.005 |
| per layer Fe-AHCAL | | | | | 2.89 | 1.241<br>47.16 | 0.139<br>5.28 | 0.113<br>4.29 |

The light produced is measured by SiPMs.

### 2.2.2 SiPM readout

SiPMs are single photon sensitive devices, composed of an array of Avalanche Photodiodes (APDs) operated in Geiger mode. By applying a reverse bias voltage on the APD, the free electrons generated by the absorption of the photons are accelerated in the electric field of the depletion region and start an electron avalanche via impact ionisation. A quenching resistor of a few MΩ stops the Geiger discharge.
The SiPMs used in the AHCAL are produced by MEPhI [3]/PULSAR [45] and have an area of 1.1 $mm^2$ with 1156 pixels. Each pixel has a size of $32 \times 32\,\mu m^2$. The SiPMs are operated at a bias voltage of 50 V, which is around 3.8 V above the break down voltage. The breakdown voltage is defined as the reverse bias voltage at which the avalanches become self-sustained. The applied 50 V correspond to a gain of approximately $10^6$. The devices have a quantum efficiency of around 80 % and are most sensitive to green light, rendering the use of the WLS fibre necessary. The bias voltage was chosen to correspond to a light yield of 15 firing pixels per Minimum Ionising Particle (MIP), which results in a signal to noise ratio of 9-10 and a detection efficiency of 93 % [46]. Figure 2.4a shows the pulse height spectrum of a SiPM recorded by illuminating it with pulsed low-intensity light from a Light-Emitting Diode (LED). The peaks correspond to different number of pixels firing and their corresponding Analog-to-Digital Converter (ADC) counts. The good separation ability originates from the high resolution of these devices. These spectra are used to determine the SiPM gain from the distance between the peaks. This allows a self-calibration (see more details in Section 2.2.3) for each individual channel.
There are a few more effects that have to be taken into account in the calibration: Since the analogue signal information is extracted by the number of firing pixels, the dynamic range is limited and the signals show saturation effects (see Figure 2.4b). Therefore the calibration includes a correction for the signal loss due to saturation, see Section 2.2.3.
The most important source of noise and dark current in SiPMs operated in Geiger mode are electrons from thermal excitation. The temperature dependence of the SiPM response requires a calibration on the cell level, since the AHCAL experiences different temperatures through the prototype due to the presence of electronic components and particle fluxes in testbeam.
An additional effect that degrades the signal is inter-pixel crosstalk, which occurs due to photons created by Geiger discharges that can trigger avalanches in neighbouring pixels. In newest SiPM technologies the crosstalk is reduced using trenches between pixels [47].
The readout electronics for the data acquisition are based on Application Specific Integrated Circuit (ASIC) chips, following the design for the CALICE Silicon-Tungsten ECAL [48]. Each ASIC reads out 18 channels, amplifying and shaping the SiPM signals. The Digital-to-Analog Converters (DACs) on the ASICs allow a channel-wise voltage adjustment.

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

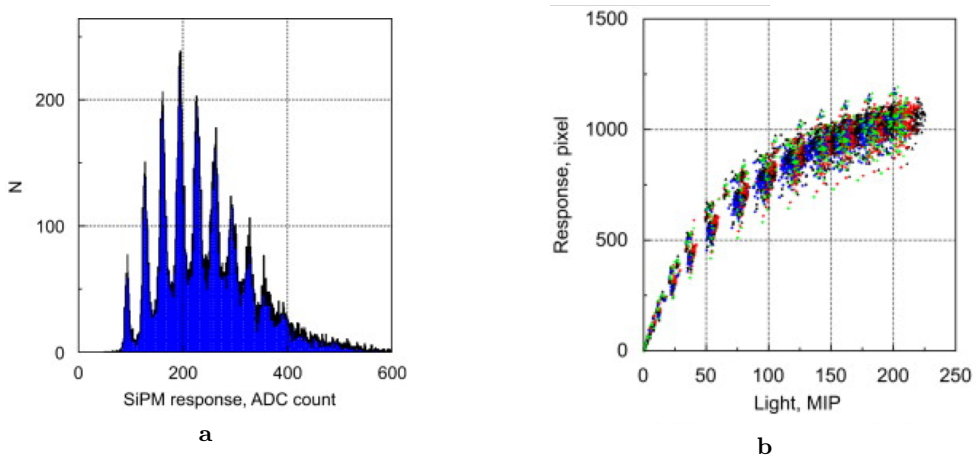



**Figure 2.4:** a) Pulse height spectrum of low-intensity LED light. b) The number of firing pixels as a function of energy deposited in the scintillator (expressed in MIPs) for different SiPMs. The shape differences lie within 15 % [45].

### 2.2.3 Calibration

Each channel of the AHCAL prototype measures signals in ADC counts, which cannot be compared to each other directly. The different cell sizes result in different light yields, because of the variation of the light collection efficiency due to different lengths and positions of the WLS fibres, and the parameter spread of the SiPMs. This makes the application of different over-voltages (voltages above breakdown voltage of a SiPM) necessary. The saturation (non-linearity) of the SiPM response varies from cell to cell and needs to be corrected individually. For the final conversion from ADC to deposited energy, an equalised response is mandatory. The unit of the energy scale is chosen to be in MIP, where one MIP corresponds to the most probable energy deposited by a minimum ionising particle. The calibrated energy measurement E per cell i is given by the equation

$$E_i = f_i^{-1} \left( (A_i - P_i) \cdot \frac{I_i}{g_i} \right) \cdot \frac{g_i}{I_i} \cdot \frac{1}{M_i}, \tag{2.1}$$

with the signal amplitude $A_i$ in ADC counts. In the following this equation is explained following the order in which the corrections are applied.

1. The pedestal $P_i$ is subtracted, which is the baseline of the electronics and measured during testbeam without beam.

2. The signal is divided by the SiPM gain $g_i$, which is the average number of ADC counts per pixel, extracted from the LED light measurement. Thus the signal gets converted to a number of pixels.

3. Another correction for the different amplification modes of the prototype during testbeam is applied by the inter-calibration factor $I_i$. The runs used to study shower physics are

taken with the electronics set into the physics mode, which provides a large dynamic range. Since the gain values are taken with the LED system in calibration mode, with a high amplification factor and short shaping times, the measured gains need to be translated. This is done by measuring the SiPM responses in both modes (for light intensities below saturation) and comparing the slopes: $I_i = \frac{Amp_{calib.mode}}{Amp_{phys.mode}}$.

4. The saturation of the SiPM response can now be corrected using the inverted saturation curve $f_i^{-1}$. And then be re-converted to the ADC scale by multiplying it with $\frac{g_i}{I_i}$.

5. For the final conversion to the MIP-scale, values of $M_i[\frac{ADC}{MIP}]$ are multiplied. These values are the Most Probable Value (MPV) of the Landau function from a convoluted Landau Gaussian fit to the ADC spectrum of 80 GeV muons.

In case that no calibration coefficients are available, default calibration factors are taken for the calibration of that cell.
More details on the calibration procedure and the detector operation during testbeam can be found in [42] and [44].

## 2.3 Digital HCAL

The Digital Hadron Calorimeter (DHCAL) [49] is a sampling calorimeter with Resistive Plate Chamber (RPC) readout. The RPCs used in this detector consist of 2 glass plates and a 1.15 mm gap filled with gas, read out by pad diodes of $1 \times 1\,cm^2$ size placed on the back of the plates. The incoming charged particles traverse the gas and ionise it. The ionisation is amplified through avalanche processes induced by the high bias voltage of around 6.3 kV applied to the RPC. The charge multiplication is quenched by the high resistivity of the glass of around $4.7 \cdot 10^{13}\,\Omega$cm [49]. If the charge exceeds the pad threshold a hit is measured. The applied thresholds varied in the range of 20 to 700 fC [50].
Each of the 38 active DHCAL layers consists of 3 RPC chambers of $32 \times 96\,cm^2$, stacked vertically to create an active area of $1 \times 1\,m^2$. The chambers are contained in a cassette of a 2 mm thick copper front plate and a 2 mm thick steel back plate. Those are inserted in the 1.4 cm wide gap of the same steel absorber structure used for the AHCAL [44], see Section 2.2. The RPCs used for this prototype are roughly 8.3 mm thick [51]. A schematic drawing of the cross section of a RPC is shown in Figure 2.5. The gas gap of 1.15 mm is ensured by horizontally embedded fishing lines placed every 5 cm. The gas mixture flushed through the chambers with ambient pressure, consisted of three components: Forane R134A[4] (94.5 %), isobutane $C_4H_{10}$ (5.0 %) and sulfur hexafluoride $SF_6$ (0.5 %).
The properties of Forane R134a concerning its sensitivity to radiation are summarised and discussed in Section 2.5, see Table 2.5.

[4] tetrafluoroethane $CH_2FCF_3$

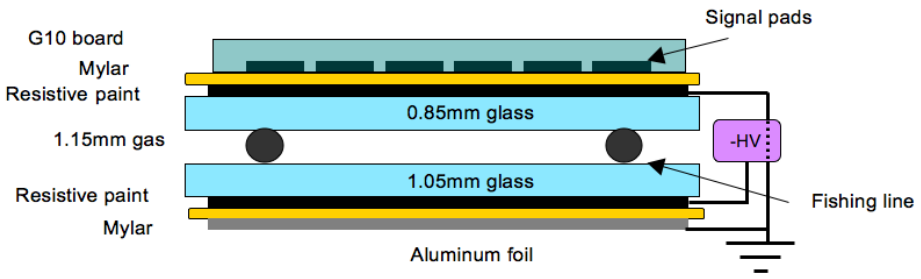



**Figure 2.5:** Cross section of a 2 glass RPC used in the Fe-DHCAL prototype [52].

The bias voltage was adapted during testbeam if necessary, since the variation of environmental conditions like temperature and pressure change the gas flow within the chambers and thus alter the avalanche probability.

Every RPC is read out by 2 front-end boards of $32 \times 48\,\mathrm{cm}^2$. This sums up to 3 RPCs and 6 front-end boards per layer. Each board has 24 chips, which makes in total 144 chips per layer. Each chip reads out $8 \times 8$, $1 \times 1\,\mathrm{cm}^2$ large pads. This sums up to 3,072 channels per RPC, 9,216 channels per layer and 350,208 for the whole prototype. The optimal pad size has been studied in simulation [53] and no strong improvement was observed between $1 \times 1\,\mathrm{cm}^2$ and infinitesimally small cells.
The radiation length of the Fe-DHCAL is calculated by Equation 1.3 and shown in Table 2.2. The mylar foil of in total 0.3 mm thickness, the air gaps of around 3 mm and the resistive coating of max. 0.1 mm thickness are not listed, because their impact is negligible. The corresponding nuclear and pion interaction lengths are given in Table 2.2.

**Table 2.2:** Radiation, nuclear and pion interaction lengths of the Fe-DHCAL per layer and for the whole prototype of 38 layers. Values are taken from the PDG [16], if not differently stated.

| Material | $X_0$ [g/cm$^2$] | $\lambda_n$ [g/cm$^2$] | $\lambda_\pi$ [g/cm$^2$] | $\rho$ [g/cm$^3$] | thick-ness [cm] | $\#X_0$ | $\#\lambda_n$ | $\#\lambda_\pi$ |
|---|---|---|---|---|---|---|---|---|
| steel absorber [44] | 13.9 | 132.1 | 160.8 | 7.86 | 1.74 | 0.989 | 0.104 | 0.085 |
| copper sheet | 12.86 | 137.3 | 165.9 | 8.96 | 0.2 | 0.139 | 0.013 | 0.011 |
| glass | 25.66 | 99.6 | 130.5 | 2.23 | 0.19 | 0.018 | 0.005 | 0.004 |
| Teflon[5](read-out board) | 34.84 | 94.4 | 124.8 | 2.2 | 0.3 | 0.019 | 0.007 | 0.004 |
| steel sheet | 13.9 | 132.1 | 160.8 | 7.86 | 0.2 | 0.113 | 0.012 | 0.010 |
| Forane 134a | 35.15 [54] | - | - | $4.32e^{-3}$ [54] | 0.115 | $\sim$0 | - | - |
| per layer | | | | | 2.75 | 1.278 | 0.140 | 0.115 |
| Fe-DHCAL | | | | | | 48.56 | 5.33 | 4.36 |

### 2.3.1 Detection Principle

The RPCs are operated in saturated avalanche mode, which minimises the streamer probability by achieving a sufficient charge amplification. The electrons and ions produced by ionising radiation get multiplied within the electric field inside the gas gap and drift toward the anode and cathode, respectively. If the free electrons gain enough energy they ionise further and start a Townsend discharge or avalanche multiplication [55]. The probability for this chain reaction to occur follows the Poisson statistic and depends on the path length between collisions and the electric field. The charge multiplication can be approximated by an exponential growth with statistical fluctuations [56]. The probability distribution of the generated charge within the avalanche started by one ion-electron pair is described by the Polya function [57]:

$$P(n) = \left[\frac{n}{N}(1+\vartheta)\right]^{\vartheta} \cdot \exp\left[-\frac{n}{N}(1+\vartheta)\right] \tag{2.2}$$

with the number of produced electrons n, the average generated charge N, which depends on the number of primary generated ion-electron pairs $n_0$ within the path length $l = g - x_0$ and the first effective Townsend coefficient $\eta$. The path length is defined by the gap width g and the position of the primary ionisation in the gap $x_0$. The induced charge of the avalanches on the pad electrodes can be computed following the Ramo theorem [58] [59]:

$$q_{ind} = \frac{q_e}{\eta \cdot g} \cdot \Delta V_W \sum_{j=1}^{n_{cl}} n_{0,j} \cdot M_j \left[\exp\left(\eta(g - x_{0,j})\right) - 1\right], \tag{2.3}$$

with the charge of the electron $q_e$, the weighting potential drop $\Delta V_W$ and $M_j = \frac{P(n)}{N}$ from Equation 2.2 for all generated ion-electron pairs $n_{cl}$. The weighting potential drop is given for a 1 gap RPC by

$$\Delta V_W = \frac{\varepsilon_r}{2d + \varepsilon_r g}, \tag{2.4}$$

with the dielectric permittivity of the resistive glass $\varepsilon_r$, the thickness of the resistive electrode d and the gap width g [60].
The large statistical fluctuations of the measured charge do not allow to reconstruct the original amount of energy deposited in the cell precisely.
The induced charge in a RPC of the DHCAL was measured with an applied voltage of 6.3 kV in a muon beam and is used as input for the DHCAL simulation, see Chapter 6.2.2.
The avalanche multiplication within the gas volume causes the charges to spread over the electrode pad plane and thus generates a hit multiplicity ($\geq 1$ hit per traversing particle) even for MIP-like particles. The multiplicity $\mu$ and the efficiency $\varepsilon$ of a RPC are the key observables used for the performance characterisation.

[5] Polytetrafluoroethylene $CF_2CF_2$

### 2.3.2 Readout

Since the measured charge is not proportional to the original energy deposition, the signal pick-up pads were chosen to be read out in 1 bit [49]. The DCAL ASIC chip reads out 64 pads. The threshold value set between 20 to 700 fC is common to all channels of one chip and is set by an internal DAC with a range of 256 counts. The output of the chip is a hit pattern (64 bits) and a time-stamp per hit with a resolution of 100 ns. The chip can be operated in either triggered or self-trigger mode [50]. The self-triggered mode was used to record so-called "noise runs", otherwise for the testbeam scintillator plates were used to trigger the data aquisition [61].

### 2.3.3 Calibration

During testbeam the operating conditions of the RPCs are influenced by temperature variations, which result in differences in gas pressure. This affects the detection efficiency and requires a tuning of the high voltage. The RPC responses are equalised by applying a calibration factor

$$c_{i,j} = \frac{\varepsilon_0 \cdot \mu_0}{\varepsilon_{i,j} \cdot \mu_{i,j}} \tag{2.5}$$

to each hit dependent on the local efficiency $\varepsilon_{i,j}$ and hit multiplicity $\mu_{i,j}$ of RPC j in layer i. This calibration factor is calculated using the average efficiency $\varepsilon_0$ and multiplicity $\mu_0$ for MIPs, that is determined from all RPCs of all runs. The determination of $\varepsilon_{i,j}$ and $\mu_{i,j}$ can be done with muons or track segments within hadronic showers [50] [62]. The calibrated response follows as

$$N_{cal} = \sum_{i=1}^{38} \sum_{j=1}^{3} \sum_{k=1}^{3,072} h_k \cdot c_{i,j}, \tag{2.6}$$

with the first two sums running over every layer i and RPC j. The third sum running over every pad k per RPC. $h_k$ equals 0, if the pad charge is below threshold and 1 if there is a hit. However, this way of determining the calibration factor implies that every hit in the event gets treated like it originated from a muon. The response to EM and hadronic showers is expected to be overestimated since, in these cases, the hit multiplicity more probably originates from multiple particles. More accurate ways of calibrating the DHCAL have been thought of, but have not been finalised yet [50, 63].
The method used in this analysis and the estimation of systematic uncertainty are described in more detail in Chapter 6.1.2 and 6.1.3.

## 2.4 Semi-Digital HCAL

The Semi-Digital Hadron Calorimeter (SDHCAL) prototype is realised in 48 active layers inserted in a steel absorber structure of 1.5 cm thick plates [64] and has been operated in testbeam at the Super Proton Synchrotron (SPS) at CERN in 2012 [65]. Each active layers consists of 2-glass RPCs with a gas gap of 1.2 mm, which is ensured by ceramic ball spacers,

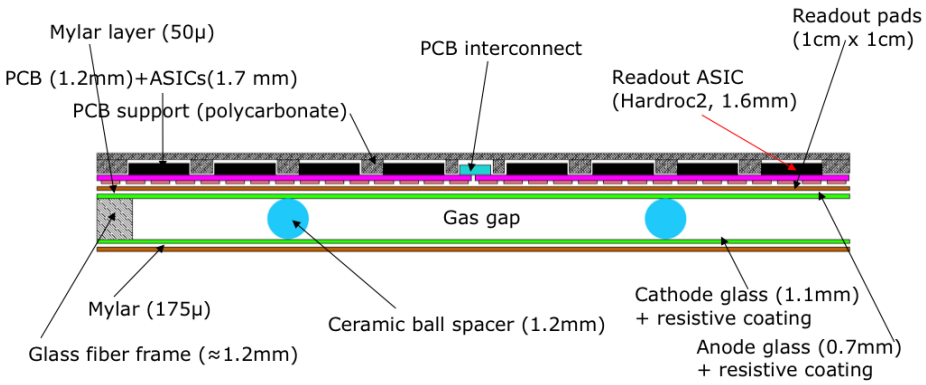



**Figure 2.6:** Cross section of the 2 glass RPC used in the Fe-SDHCAL prototype [65].

and a readout electrode segmented in $1 \times 1$ cm$^2$ pads. The cross section of these RPCs is shown in Figure 2.6. The gas mixture used as sensitive medium consists of Forane R134A[6] (93 %), carbon dioxide $CO_2$ (5 %) and sulfur hexafluoride $SF_6$ (2 %). The interaction and radiation lengths of the active layers inserted in the steel cassette and the absorber plate are shown in Table 2.3[7]. The high voltage applied to the RPCs was set during testbeam to a fixed value of 6.9 kV. To ensure stable temperature and thus gas pressure conditions, the whole detector was operated in power pulsing mode, which places the electronics in an idle state during the time separating two beam spills. Additionally a water cooling system was installed.
The RPC based SDHCAL prototype follows the same detection principle as the DHCAL, see Section 2.3.1. The main difference between the D- and SDHCAL prototype lies in the readout of the pad signals, which is further discussed in Section 2.4.1.
In addition to the RPC based SDHCAL several MicroMEGaS layers as active material for semi-digital calorimetry have been tested. The granularity of the readout of these devices was $1 \times 1$ cm$^2$ and the proof of principle was successful [66].

### 2.4.1 Readout

Three thresholds are used in the SDHCAL readout. This is not an attempt to estimate the deposited energy, but to distinguish whether the recorded charge is the results of one, few or many charged particles traversing one cell. These thresholds are applied by the HAdronic Rpc Detector ReadOut Chip (HARDROC) for groups of 64 channels. The values of the thresholds were 110 fC, 5 pC and 15 pC [64]. These values were not optimised for the semi-digital energy reconstruction and due to the nature of the generated charge spectrum of RPCs, a conversion to the MIP scale is only possible with a large error. Following the conversion factor extracted

[6] tetrafluoroethane

[7] The values calculated here present a few percent difference in comparison to earlier calculations. This is most probably due to the approximation that the readout boards consist of Teflon. Compare $\lambda_n = 0.12$ and $X_0 = 1.14$ with the values in Table 2.3 of one SDHCAL layer [65].

**Table 2.3:** Radiation, nuclear and pion interaction lengths of the Fe-SDHCAL per layer and for the whole prototype of 48 layers. Values taken from the PDG [16], if not differently stated. The air gap between the absorber and the active layer of 2 mm is neglected.

| Material | $X_0$ [g/cm$^2$] | $\lambda_n$ [g/cm$^2$] | $\lambda_\pi$ [g/cm$^2$] | $\rho$ [g/cm$^3$] | thick-ness [cm] | #$X_0$ | #$\lambda_n$ | #$\lambda_\pi$ |
|---|---|---|---|---|---|---|---|---|
| steel absorber [44] | 13.9 | 132.1 | 160.8 | 7.86 | 1.5 | 0.848 | 0.089 | 0.073 |
| steel sheets | 13.9 | 132.1 | 160.8 | 7.86 | $2 \times 0.25$ | 0.283 | 0.030 | 0.024 |
| glass | 25.66 | 99.6 | 130.5 | 2.23 | 0.18 | 0.016 | 0.004 | 0.003 |
| Teflon (read-out board) | 34.84 | 94.4 | 124.8 | 2.2 | 0.3 | 0.019 | 0.007 | 0.004 |
| Forane 134a | 35.15 [54] | - | - | $4.32e^{-3}$ [54] | 0.12 | $\sim$0 | - | - |
| per layer Fe-SDHCAL | | | | | 2.6 | 1.166<br>55.97 | 0.127<br>6.10 | 0.104<br>4.99 |

in [67] $1.5 \pm 0.4$ pC/MIP, the applied thresholds in MIP can be approximated to: $0.07 \pm 0.02$, $3.3 \pm 0.9$ and $10 \pm 2.67$ MIP.
A more detailed description of the SDHCAL hardware can be found in [64].

### 2.4.2 Beam Intensity Correction

RPCs are limited in their rate capability due to the high resistivity of the glass plates [68]. Thus high particle fluxes lead to a significant loss in efficiency. This affected the SDHCAL strongly in the SPS testbeam because of the number of particles per spill being larger than 1,000. The number of hits above the highest threshold is strongly affected and decreases by 25% during a spill [65]. To correct for the loss in hits per time, which degrades the resolution of the calorimeter, the number of hits for each threshold is corrected by a linear fit to the number of hits per time per threshold for each run by

$$N_{corr,j} = N_j - \lambda_j \cdot t, \tag{2.7}$$

with j running from threshold 1 to 3, t is the time from start of a spill and $\lambda$ is the factor extracted from the linear fit [65].
In the DHCAL this effect is significantly lower, due to the lower particle rate at the Fermi National Accelerator Laboratory, Batavia IL, USA (Fermilab) testbeam. Therefore, the DHCAL calibration procedure does not include a correction for this effect, which is estimated to be of the order of 1-2 % [69].

## 2.5 Comparison of the CALICE HCALs

The three hadronic calorimeter concepts of the CALICE collaboration are compared in terms of the main characterisation parameters for sampling calorimeters; radiation length, nuclear interaction length, sampling fraction and e/π ratio.
All three HCALs are very similar in radiation and nuclear interaction length, see Table 2.4. The largest differences are observed in the sampling fraction $f_{sampling}$ and e/π ratio. Here the sampling fraction is approximated by the stopping power for MIPs in the active and passive materials, following Equation 1.19. The values used for the calculation of the sampling fraction are given in Table 2.5. The energy loss in the support structures has been neglected, which results in an uncertainty of a few percent. The values are given in Table 2.4.
The difference in the signal formation and the dependence on the active material is compared by the value W, which estimates the needed energy to generate a signal quantum. This signal quantum is an electron-ion pair in the RPC gas and a photon in the plastic scintillator. The values are compared in Table 2.4 and reveal a lower intrinsic threshold of the RPC gas compared to the plastic scintillator by a factor 2 to 3. The expected number of signal quanta $N_I$ due to a MIP in the gas gap and the scintillator tile are compared in the bottom part of Table 2.5 and reveal a much larger number of photons of the order of ten thousand compared to $\sim 30$ electron-ion pairs in the gas gap. However, only a small fraction of the photons reach the SiPM and again only a fraction of the photons fire a pixel, which is defined by the geometrical efficiency and the particle detection efficiency (PDE) of the SiPMs that are of the order of 50 and 30 % [70]. In this way, in the end a signal of 15 pixels, fired by $\sim 15$ photons, which is less than 1 % of the generated photons per MIP, is measured, see Section 2.2.2.
The e/π ratio is measured comparing the mean response to electrons to the one to pions of the HCAL prototypes. In the AHCAL both responses are linear and therefore one number valid for a large energy range is determined [71]. The S- and DHCAL on the other hand suffer from saturation for high density EM showers and therefore have an e/π ratio smaller than 1 and further decreasing with increasing energy. The effects of this strong non-compensation together with methods for partial recovery will be discussed in Chapters 4 and 6. More details about the mean electron and pion responses in the Fe-DHCAL are discussed in Section 6.3 and 6.4. The e/π ratio of the Fe-DHCAL is determined in Section 6.5.

**Table 2.4:** Radiation $X_0$ and nuclear interaction $\lambda_n$ lengths, as well as the sampling fraction $f_{sampling}$ and $e/\pi$ ratio of the CALICE HCALs.

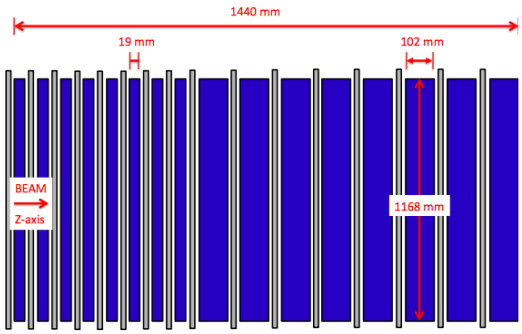
| | Fe-SDHCAL | | Fe-DHCAL | | Fe-AHCAL | |
|---|---|---|---|---|---|---|
| | per layer | 48 layers | per layer | 38 layers | per layer | 38 layers |
| $X_0$ [#] | 1.17 | 55.97 | 1.28 | 48.56 | 1.24 | 47.2 |
| $\lambda_n$ [#] | 0.13 | 6.10 | 0.14 | 5.33 | 0.14 | 5.3 |
| $f_{sampling}$[%] | $5.32 \cdot 10^{-3}$ | | $4.39 \cdot 10^{-3}$ | | 4.91 | |
| $e/\pi$ | $< 1$ [65] | | $< 1$ [72] | | 1.19 [71] | |

**Table 2.5:** The density $\varrho$, the stopping power $-\langle dE/dx \rangle_{min}$ in terms of mean energy loss per length of Forane 134a [54], polystyrene and iron (Fe) [16], and the mean energy to excite an electron-ion pair or a photon W. The values of $-\langle dE/dx \rangle_{min}$ have been determined by Equation 1.1.
The bottom part of the table shows the thickness of the gas gap and scintillator tile d and the estimated energy loss as well as the number of generated ion pairs and photons $N_I$.

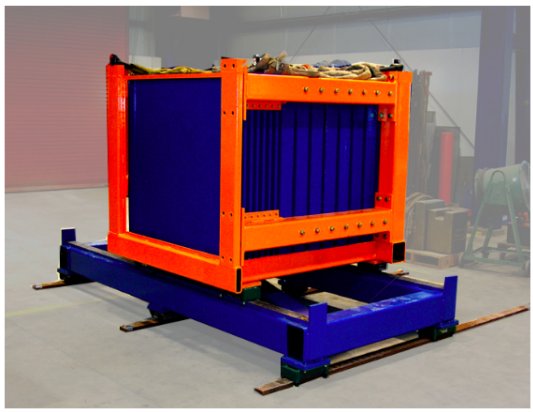
| | Forane R134a | Polystyrene | Fe |
|---|---|---|---|
| $\varrho$ [g/cm$^3$] | $4.32\times10^{-3}$ | 1.05992 | 7.874 |
| $-\left\langle \frac{dE}{dx} \right\rangle_{min}$ [MeVg$^{-1}$cm$^2$] | 1.76439 | 1.936 | 1.451 |
| $-\left\langle \frac{dE}{dx} \right\rangle_{min}$ [MeVcm$^{-1}$] | 0.0076 | 2.052 | 11.43 |
| W [eV] | $\sim 30$ [73] | $\sim 60 - 150$ [74, 75] | - |
| d [cm] | 0.115 (0.12) | 0.5 | - |
| $-\left\langle \frac{dE}{dx} \right\rangle_{min}$ [eV] | 874 (912) | $1.026 \cdot 10^6$ | - |
| $N_I$ | $\sim 29$ (30) | $\sim 6,840 - 17,100$ | - |

## 2.6 Tail Catcher and Muon Tracker

The TCMT is $108 \times 108 \times 142\,\mathrm{cm}^3$ large and is positioned behind the HCAL in order to absorb the tails of the showers leaking out. The TCMT is divided into two sections of 8 active layers each. The first section provides 2 cm thick steel absorber plates and the second section provides absorber plates with a thickness of 10 cm, see the schematic in Figure 2.7a. In total the TCMT has a thickness of $5.8\,\lambda_\mathrm{n}$ nuclear interaction lengths [76].
This absorber structure, displayed in Figure 2.7b, was placed downstream the HCAL in AHCAL and DHCAL testbeams at Conseil Européen pour la Recherche Nucléaire, Geneva, Switzerland (CERN) and Fermilab. In the AHCAL testbeams the active TCMT layers consisted of $5 \times 100 \times 0.5\,\mathrm{cm}^3$ scintillator-strips read out by SiPMs, following the principle of the AHCAL scintillator tiles described in Section 2.2. The scintillator-strips are of alternating horizontal and vertical orientation in adjacent layers.
In the DHCAL testbeam each active layer of the TCMT consisted of three RPCs, as in the DHCAL and described in Section 2.3.




a



b

**Figure 2.7:** The TCMT absorber structure [77]. a) The cross-section of the TCMT in z-direction. The steel absorber plates are shown in blue and active layers in grey. b) The TCMT structure shown in its steel support (orange) on the movable stage for positioning in the testbeam area.

# Chapter 3

# Simulation

The simulation of detector concepts for future High Energy Physics (HEP) experiments plays a key role in technology choices and physics analyses. However the predictive power of simulations requires the validation of the simulation with data.
The CALICE and GEANT4[1] collaborations work together on a regular basis in order to provide each other with crucial information about newest software developments and detailed comparisons between models of electromagnetic and hadronic showers with the data of the CALICE prototypes. A description of the available electromagnetic and hadronic physics lists of GEANT4 can be found in Section 3.2 and 3.3.

## 3.1 Simulations of Particle Interactions with Matter

The software toolkit GEANT4 simulates the interaction of particles with matter. It is widely used in particle physics, radiation protection, medicine and space sciences. The GEANT4 simulations include a large range of physics processes that can be modified by the user.
Electromagnetic (EM) showers only involve electrons, positrons and photons. Therefore the processes are well understood and the descriptions are very accurate, at the percent level [78]. However, low energy EM cascades are described with less accuracy with the electromagnetic standard model [79].
Hadronic showers are much more complex than EM showers. They involve different physics processes, like the strong interaction with composite objects. These processes cannot be described analytically. Additionally, a lot more different secondary particles are involved in the cascade formation, thus GEANT4 uses models that follow assumptions and approximations that are valid only for a certain energy range. To cover a large energy range different models are merged into so-called physics lists.
One important parameter for a GEANT4 simulation is the range cut. This parameter determines the accuracy of the simulation by only tracking a particle if it has enough energy to

[1] `https://geant4.web.cern.ch/geant4/`

travel a distance larger than the range cut, otherwise its energy gets deposited immediately. This parameter is set independently from the chosen physics list.

## 3.2 Geant4 Electromagnetic Physics Lists

The electromagnetic physics lists include different processes in different nuances of accuracy. A detailed list can be found in [80]. GEANT4 provides 9 different options for EM cascade modelling, designed for different applications. The ATLAS experiment at the LHC uses the standard EM list very successfully [81], while the LHCb experiment uses option 2, which provides a few extras like a more detailed description of bremsstrahlung and an improved description of multiple scattering [81]. For the detailed shower studies of the CALICE collaboration, especially for the DHCAL, the GEANT4 working group recommended the EM list option 3 [82]. This option 3 is supposed to have a high accuracy with the standard EM implementation. Furthermore, option 4 is tested in the DHCAL simulation, see Section 6.2.2, because it is supposed to have the highest accuracy for EM showers [81].

## 3.3 Geant4 Hadronic Physics Lists

GEANT4 simulations use different models to describe inelastic scattering of single hadrons with atomic nuclei. These models are either parametrisation based or theory driven. The following analysis concentrates on two physics lists, consisting of different models, that have shown the best agreement with data and are recommended for highly-granular calorimetry studies by the GEANT4 working group [83]: FTFP_BERT and QGSP_BERT. While FTFP is a shortcut for the Fritiof string model, QGSP stands for the quark-gluon string model. Both are parton string models, that simulate each nucleon of the nuclei. The interaction between the incoming hadron and the nucleons is simulated using the impact parameter, the centre of mass energy of the interaction and the inelastic and diffractive cross-sections. Strings are formed between the quarks of the hadron and the nucleons, four-vectors are assigned as well as excitation stages. Then the model splits these strings into quark-antiquark pairs and new strings. The quarks hadronise. Further fragmentation is applied until the energy is too low for further string-splitting. The FTFP and QGSP models differ in the string formation and fragmentation [42]. For low energies the Bertini cascade model (_BERT) is used. Like all cascade models of GEANT4 this model treats the nucleus as a Fermi gas. This implies that secondary particles are produced if the energy exceeds the Fermi energy. The Bertini model describes the nucleus as three concentric spheres with constant nucleon density. The nucleons are assumed to have a Fermi-gas momentum distribution and for each collision the momentum of the hit nucleon is calculated, the reaction is identified and the four momenta of the reaction products are calculated. Furthermore, the excited nucleus is de-excited after all secondaries have left the nucleus or got absorbed.

For the intermediate energies of 10-12 GeV and the transitions regions of 9.5-10 GeV and 12-25 GeV, the QGSP_BERT physics list uses the LEP model, which is a parametrised model. A parametrised model predicts the production of secondaries using fits to experimental data. There is no detailed modelling of interactions and the energy is only conserved on average. Several inaccuracies of these models have been seen, therefore these kind of models are only used if there are no alternatives available in a certain energy range. An overview of the composition of the physics lists used in this analysis is shown in Figure 3.1.

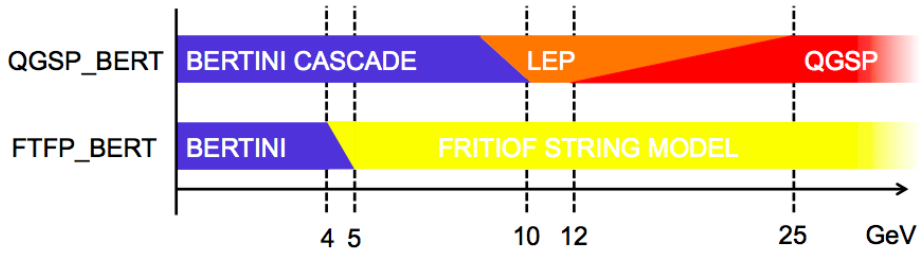



**Figure 3.1:** The energy dependent cascade models used in the GEANT4 physics lists used for hadronic shower simulation.

The simulations of the AHCAL and DHCAL testbeams are performed with GEANT4 version 9.6 and 10.1. The studied physics lists were chosen to be: FTFP_BERT and QGSP_BERT, following the GEANT4 recommendation [83]. Details on the geometric description as well as the digitisation can be found for the AHCAL in Section 5.1 and the DHCAL in Section 6.2.

# Chapter 4

# Energy Reconstruction Procedures

For the three different CALICE Hadron Calorimeters, which use different active material and readout, three different energy reconstruction procedures were developed and are described in the following. In addition, software compensation algorithms have been (re-)designed and will be discussed.
Essential for the energy reconstruction and resolution determination is the fitting method used to extract the mean position and width $\sigma$ of the measured distributions. Therefore, the method chosen for this analysis is introduced and discussed in the following.

## 4.1 Extraction of Response and Resolution

The determination of the mean position and width $\sigma$ of measured or reconstructed distributions is not trivial in case of non-Gaussian distributions. Asymmetries can occur and originate from different effects, e.g. $e/\pi \neq 1$, saturation or energy leakage. This analysis focusses on the comparison between HCAL technologies, therefore saturation effects due to energy leakage are minimised by a tight event selection. More details about this method are discussed in Section 5.2 and 6.1. While the saturation of readout devices, like SiPMs are recovered within the calibration procedures, saturation due to a limited granularity is the main reason for tails on the left hand side of the total number of hits distributions of the HCAL prototypes. An example showing this effect is presented in Figure 4.1, where the total number of hits per event $N_{hits}$ distribution for 40 GeV pions in the AHCAL reveals a tail to smaller number of hits due to the relatively large cell size of $3 \times 3\,cm^2$.
The method that is used to extract the mean position and $\sigma$ of such a distribution follows 3 steps:

1. Fitting a Gaussian in the range: mean $\pm 3$ Root Mean Square (RMS), shown as blue curve in Figure 4.1.

2. Using the Gaussian parameters as input for a fit with the Novosibirsk function

$$f(x) = A \cdot \exp\left(-\frac{1}{2}\left(\frac{\ln^2\left[1+\Delta\cdot\tau(x-\mu)\right]}{\tau}\right)+\tau^2\right) \tag{4.1}$$

with $\Delta = \frac{\sin\left(\tau\cdot\sqrt{\ln 4}\right)}{\sigma\cdot\tau\cdot\sqrt{\ln 4}}$ within the range $\mu_{Gaussian} \pm 3\sigma_{Gaussian}$. Hereby is $\tau$ the tail parameter, A the amplitude and $\mu$ and $\sigma$ the MPV and standard deviation of a Gaussian function. This fit is shown as red curve in Figure 4.1.

3. In the last step a histogram is filled, following the Novosibirsk function in the range from 0 to $\mu + 3\sigma$, and the mean and RMS of that histogram are taken as the mean position and width of the $N_{hits}$ distribution.

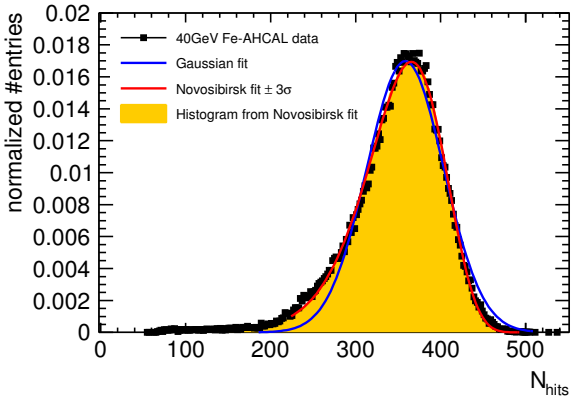



**Figure 4.1:** The total number of hits $N_{hits}$ distribution for 40 GeV pion events in the AHCAL (black points). The Gaussian fit is shown as blue curve, the Novosibirsk fit is shown as red curve and the histogram following the Novosibirsk function is shown in orange.

Following this method, the RMS is secured from outliers, present in the original distributions and the shift of the mean position due to the saturation effect is included, in contrast to the $\mu_{Gaussian}$.

## 4.2 Analogue Energy Reconstruction

The Analogue HCAL physics prototype is a scintillator tile calorimeter with Silicon Photomultiplier (SiPM) readout. A calibration, discussed in Section 2.2.3, is used to convert the measured ADC counts to the response of a Minimum Ionising Particle (MIP) [84]. Thus the visible signal $E_{sum}$ for the reconstructed energy is calculated in units of MIP as a sum of cell signals above a cell energy of 0.5 MIP. The 0.5 MIP threshold is used to reject noise. The MIP scale is converted to a GeV scale using an electromagnetic calibration factor $\omega$ which was determined from dedicated positron runs [85]. The scintillator-steel AHCAL is a non-compensating calorimeter, as its response to electrons is by a factor of $e/\pi = 1.19$ higher than to pions of

the same energy [71]. An additional scaling factor c is included in the energy reconstruction to compensate for the missing energy that is usually added from the Tail Catcher and Muon Tracker (TCMT) [71]. Due to the very high noise levels of the TCMT, energy needs also to be added for the runs using a low energy beam. This factor was estimated by fitting the mean analogue response with

$$\langle E_{sum} \rangle = \frac{E_{beam}}{\frac{e}{\pi} \cdot \omega \cdot c} \tag{4.2}$$

and found to be c = 1.04. Then the reconstructed energy in the AHCAL for each pion event is calculated from the measured energy $E_{sum}$ expressed in MIP as follows:

$$E_{rec,AHCAL} = \frac{e}{\pi} \cdot \omega \cdot E_{sum} \cdot c. \tag{4.3}$$

## 4.3 Digital Energy Reconstruction

The quantity measured by the DHCAL is the total number of hits $N_{hits}$ in the HCAL. The signals amplitude per cell is not measured.
Within the energy reconstruction a correction for the non-linearity of the response is applied. The non-linearity arises from multiple particles traversing the same pad, limited granularity and binary information. Several approaches have been developed to correct for this non-linearity. Here, a simple approach is followed by fitting the mean response versus beam energy $E_{beam}$ with a power law as $\langle N_{hits} \rangle = a \cdot (E_{beam})^b$ and using the extracted parameters a and b for the reconstruction as follows:

$$E_{rec,DHCAL} = \sqrt[b]{\frac{N_{hits}}{a}}. \tag{4.4}$$

## 4.4 Semi-Digital Energy Reconstruction

The principle of the semi-digital HCAL is similar to the digital HCAL, but with a 2-bit readout. This 2 bit readout encodes the information of 3 thresholds. This additional information compared to the DHCAL has the goal to identify multiple particles contributing to the signal of a pad. The energy in the SDHCAL physics prototype is reconstructed as a weighted sum of the number of hits for the 3 thresholds. $E_{rec,SD}$ can be written as a function of $N_1$, the number of hits above the first and below the second; $N_2$, the number of hits above the second below the third; and $N_3$, the number of hits above the third threshold:

$$E_{rec,SD} = \alpha N_1 + \beta N_2 + \gamma N_3, \tag{4.5}$$

with the weights $\alpha$, $\beta$ and $\gamma$ in units of GeV. Hadronic showers change their structure and evolution with energy, which is taken into account by parameterising $\alpha$, $\beta$ and $\gamma$ as second order polynomials of the total number of hits $N_{hits} = N_1 + N_2 + N_3$. In order to find the best

parameterisation of these reconstruction coefficients, a $\chi^2$-like function of the form

$$\chi^2 = \sum_{i=1}^{N} \frac{\left(E_{beam,i} - E_{rec,i}\right)^2}{E_{beam,i}}, \tag{4.6}$$

is minimised, where i runs over all events N.

## 4.5 Software Compensation Algorithms

The purpose of the techniques introduced in the following is the compensation for the differences in the response to electromagnetic and hadronic components in pion showers. In contrast to the regular energy reconstructions (analogue and digital), software compensation techniques apply a correction for each event individually by weighting the hits depending on their initial energy content ($e_j$) or energy density (e.g. in number of neighbouring hits $N_{neighbours,j}$) and the total visible energy, measured by $E_{sum}$ or $N_{hits}$.
In this way, hits of electromagnetic content that typically have a higher energy density and hadrons get weighted individually and thus the response gets equalised and the reconstructed energy optimised. A proof of principle has been shown in [71]. The algorithm developed for this analysis is inspired by the semi-digital weight estimation. Contrary to the previous techniques, the weights for the reconstruction $\omega\left(e_j\right)$ and $\nu\left(N_{neighbours,j}\right)$ are parametrised as second order polynomials of the total visible energy $E_{sum}$ and the total number of hits $N_{hits}$ respectively.

### 4.5.1 Analogue Software Compensation

The Analogue HCAL measures the individual hit energies $e_j$, which allows the software compensation algorithm to use this information for the resolution and linearity optimisation by minimising the $\chi^2$ function of Equation 4.6.
First the hit energy is divided in k = 8 ranges, where this number was chosen to be as small as possible without loosing performance. The impact of the chosen energy range will be discussed in more detail in Section 5.7. Each hit in energy range k is weighted by $\omega_k$, which is parametrised as a 2nd order polynomial of $E_{sum}$, compare Section 4.4. The $\chi^2$ function is then minimised by using the same number of events per beam energy. The resulting weights in units of $\frac{GeV}{MIP}$ are used to reconstruct the energy for each event by

$$E_{rec,A\text{–}SC} = \sum_{k=1}^{8} \omega_k\left(E_{sum}\right) \cdot E_k = \sum_{j=0}^{N_{hits}} \omega\left(e_j, E_{sum}\right) \cdot e_j, \tag{4.7}$$

with the sum of hit energies in the energy range k: $E_k = \sum e_k$.
In this way the beam energy is used for the weight determination, while only the measured energy $E_{sum}$ is needed for the energy reconstruction.

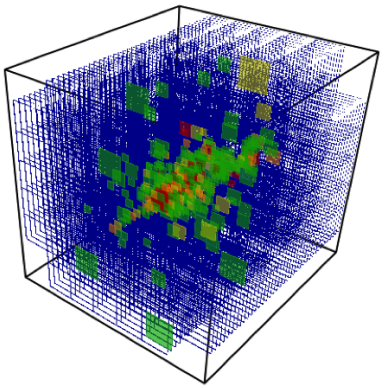
This software compensation algorithm uses analogue hit information and is therefore here called Analogue Software Compensation.

### 4.5.2 Digital Software Compensation

The Digital HCAL only provides the hit information, thus needs a different software compensation algorithm. An optimisation algorithm can be applied, using the number of neighbouring hits $N_{neighbours}$ to estimate the energy density. The $N_{neighbours}$ spectrum is divided into $l = 9$ ranges and the weights for the hits in these ranges $\nu\left(N_{neighbours,l}\right)$ are parametrised as 2nd order polynomials of the total number of hits $N_{hits}$. The same $\chi^2$ function as for the Analogue Software Compensation algorithm is used to determine the weights in units of GeV. Afterwards the energy is reconstructed as

$$E_{rec,D-SC} = \sum_{l=1}^{9} \nu\left(N_{neighbours,l}, N_{hits}\right) \cdot N_l = \sum_{j=0}^{N_{hits}} \nu\left(N_{neighbours,j}, N_{hits}\right), \tag{4.8}$$

with the sum of hits in the $N_{neighbours}$ range l: $N_l = \sum_{j=0}^{N_l} 1$.
The possibility of applying a similar weighting method to the semi-digital reconstruction persists, but because of the weighting nature of the classic semi-digital reconstruction, further improvement is unlikely.



# Chapter 5

# The Analog HCAL at CERN

The Analog HCAL with steel absorber was tested in the particle beam of the CERN Super Proton Synchrotron (SPS) in 2007. The 2007 CERN testbeam setup consisted of 30 layers of CALICE Si-ECAL, 38 layers of the scintillator-steel Analogue HCAL and 16 layers of the scintillator-steel Tail Catcher and Muon Tracker (TCMT). The Si-ECAL has a thickness of $1\,\lambda_n$ nuclear interaction length and is described in Section 2.1. The Fe-AHCAL is described in Section 2.2 and has a total thickness of $5.28\,\lambda_n$. The TCMT, which is not being used in the following analysis, is described in Section 2.6 and has a total thickness of $5.8\,\lambda_n$. Additionally, a Cherenkov threshold counter was placed in front of the calorimeter system and was used to identify protons.
The AHCAL testbeam data with steel absorber from 2007 is well understood and validated by several CALICE analyses [62,71,76,84]. The following analysis studies the impact of the energy reconstruction procedures and the granularity on the energy resolution and has been made public in two CALICE Analysis Notes [86, 87] and was presented at two conferences [88, 89].

## 5.1 Simulation and Digitisation

The testbeam runs are simulated using the software packages GEANT4 version 9.6 patch 1, MOKKA[1] v08_02 and ILCSOFT[2] v01_17_05, followed by the digitisation using calice_soft v04-08 with the conversion coefficient 846 keV/MIP and 15 % optical crosstalk between the AHCAL tiles. The digitisation of the detector effects includes the saturation curves of the SiPMs and the photon statistics. An additional factor of 0.7 is applied to the number of photons guided through the WLS fibre to the SiPM, which accounts for the mismatch between the circular cross section of the WLS fibre and the quadratic SiPM array and additional alignment uncertainties. Only 70 % of the photons are read out by the SiPM, which is included in the digitisation. As

[1] Software package, developed by Laboratoire Leprince-Ringuet (LLR), École polytechnique, Palaiseau (France) `http://ilcsoft.desy.de/portal/software_packages/mokka/`

[2] Software packages for the International Linear Collider (ILC)
`http://ilcsoft.desy.de/portal`

the physics lists FTFP_BERT and QGSP_BERT from GEANT4 9.6 show best performance for hadrons [76], they were chosen for the simulation of the testbeam setup. All testbeam runs listed in Table 5.1 were simulated with 100,000 $\pi^-$ events, the noise being added to the digitised samples from the corresponding runs. Afterwards, the same selection procedure used for the testbeam data was applied. The resulting number of pion events and the percentage of selected events are given in Table 5.1.

## 5.2 Run and Event Selection

The data and simulation samples are selected from $\pi^-$ runs in the energy range of 10 to 80 GeV. The run list and event selection follows the published software compensation analysis [71] and is summarised in Table 5.1. The only difference consists in requirements on the shower start, that was shifted from the first 5 layers to the 2nd to 6th layer in the HCAL in order to clean the data set from showers starting in the last ECAL layer and in the gap between ECAL and HCAL. The runs with the same beam energy are merged and undergo the same $\pi^-$ event selection.
In the $\pi^-$ pre-selection the Cherenkov threshold counter is used to ensure the separation of protons. Additionally, a threshold of 0.5 MIP is applied on every cell to reduce the noise. After the pre-selection of $\pi^-$ events, events are rejected that originate from

- muons and punch-through pions by requiring more than 150 MIP deposited in the AHCAL.
- multi-particles by requiring less than 80 MIP and 13 hits in the first 5 layers of the AHCAL.
- empty events by requiring more than 25 hits in the ECAL and 50 hits in the AHCAL.

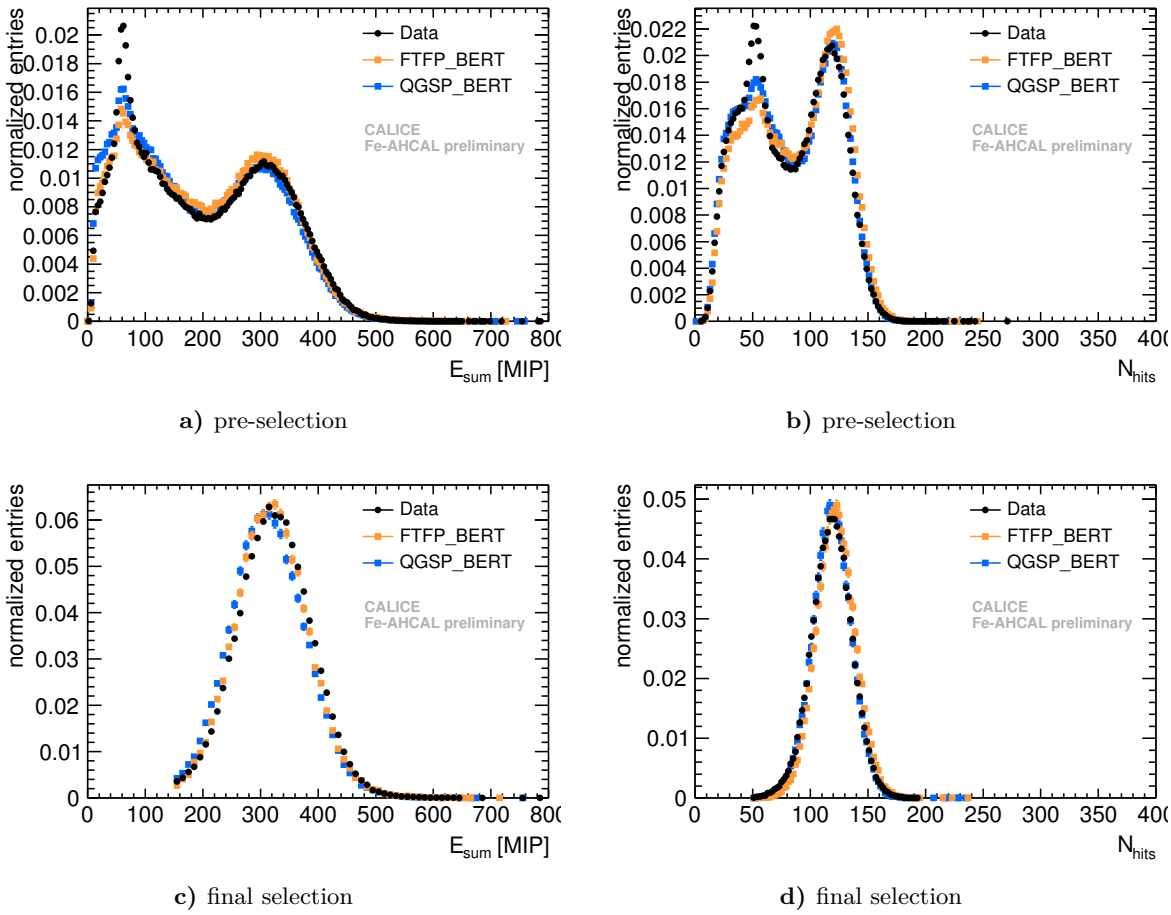
To minimise leakage into the TCMT and fluctuations of the energy depositions in the ECAL, $\pi^-$ events are selected that

- start showering in the 2-6th HCAL layer by the ShowerStartClusterProcessor [90].
- show no hard interaction in the ECAL by requiring less than 50 hits.

The requirement on the shower start has the largest impact on the statistic of the selected $\pi^-$ events, while the multi-particle and empty event contamination is small. This manifests in the very similar percentage of selected events for the data after the pre-selection and the simulations, compare column 5 and 7 in Table 5.1.
The selected pion showers develop predominantly in the AHCAL while the energy leakage into the TCMT is kept as small as possible.
Figures 5.1 and 5.2 show the distributions of the visible energy $E_{sum}$ and the number of hits $N_{hits}$ in the AHCAL for the data and the Monte Carlo simulation (MC) of 10 to 80 GeV

**Table 5.1:** List of data runs used in the analysis and sample statistics. The size of each simulated sample is 100,000 events per run.

| run number | beam energy [GeV] | pre-selection data | selected pions in data | | selected pions in MC (FTFP_BERT/QGSP_BERT) | |
|---|---|---|---|---|---|---|
| | | | $N_{events}$ | in % | $N_{events}$ | in % |
| 330332, 330643, 330777, 330850 | 10 | 587,793 | 95,065 | 16.2 | 67,315/64,807 | 16.8/16.2 |
| 330328 | 15 | 140,441 | 24,044 | 17.1 | 16,963/ 16,245 | 17.0/ 16.3 |
| 330327 | 18 | 148,516 | 25,129 | 16.9 | 16,780/16,094 | 16.8/16.1 |
| 330649, 330771 | 20 | 379,270 | 61,538 | 16.5 | 32,902/31,979 | 16.5/16.0 |
| 330325, 330650 | 25 | 364,170 | 61,037 | 16.8 | 32,250/31,543 | 16.1/15.8 |
| 330551, 330960 | 35 | 404,309 | 57,981 | 14.3 | 31,626/31,460 | 15.8/15.7 |
| 330390, 330412, 330560 | 40 | 509,168 | 83,595 | 16.4 | 47,403/47,367 | 15.8/15.8 |
| 330550, 330559, 330961 | 45 | 520,600 | 84,583 | 16.3 | 47,263/46,836 | 15.8/15.6 |
| 330391, 330558 | 50 | 384,581 | 62,843 | 16.3 | 31,704/31,306 | 15.9/15.7 |
| 331556, 331568, 331655, 331664 | 60 | 787,208 | 133,618 | 17.0 | 62,666/62,302 | 15.7/15.6 |
| 330392, 330962, 331554, 331567, 331654 | 80 | 898,307 | 152,182 | 16.9 | 76,932/79,056 | 15.4/15.8 |




**a)** pre-selection

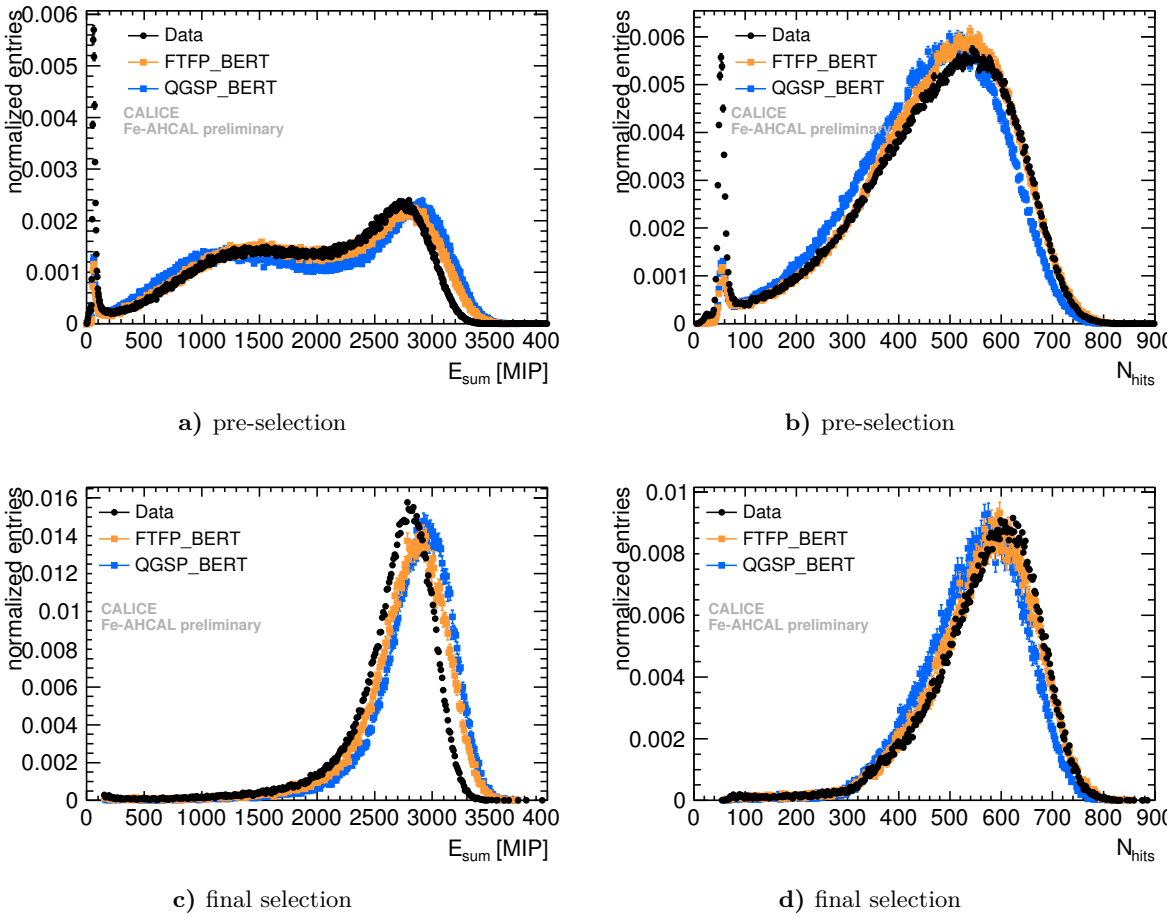
**b)** pre-selection

**c)** final selection

**d)** final selection

**Figure 5.1:** Distributions of the visible energy $E_{sum}$ and the number of hits $N_{hits}$ for $E_{beam} = 10\,GeV$, shown after the pre-selection in a) and b) and the final $\pi^-$ selection in c) and d). The simulated FTFP_BERT data is shown in orange, the simulated QGSP_BERT data in blue, and the testbeam data is represented by black points.

pions. The experimental data is compared to the simulations using the FTFP_BERT and QGSP_BERT physics lists, which shows some distinct differences in the 80GeV distributions. However, for all other beam energies the differences between data and MC are smaller. In all pre-selection plots for the energy sum distributions (see Figure 5.1a, 5.2a) and the number of hits (see Figure 5.1b, 5.2b) a peak is observed around 100 MIP and 40 hits, respectively. The peak is more pronounced in data than in simulation. This difference is due to the muon contamination in data, while in the simulation the peak arises only from punch-through pions. A second difference is a slight overestimation in the FTFP_BERT samples of the number of hits. The largest difference between data and MC is seen in Figure 5.2c, where both physics lists overestimate the AHCAL response. This trend was already seen and studied [42].




**a)** pre-selection

**b)** pre-selection

**c)** final selection

**d)** final selection

**Figure 5.2:** Distributions of the visible energy $E_{sum}$ and the number of hits $N_{hits}$ for $E_{beam} = 80\,GeV$, shown after the pre-selection in a) and b) and the final $\pi^-$ selection in c) and d). The simulated FTFP_BERT data is shown in orange, the simulated QGSP_BERT data in blue, and the testbeam data is represented by black points.

## 5.3 Systematic Uncertainties

The systematic uncertainties in data are estimated following [71] and [62], which both use the detailed analysis of the electromagnetic response [85]. The uncertainty of the beam energy $\Delta E_{beam}$ is taken into account with

$$\frac{\Delta E_{beam}}{E_{beam}} = \frac{12\%}{E_{beam}} \oplus 0.1\%. \tag{5.1}$$

The method to determine these values is described in [91].
The uncertainty on the reconstructed energy is dominated by the MIP to GeV conversion, and is estimated to be 0.9 %. The impact of the SiPM gain and saturation parameters are negligibly small [85].
The detector stability over time was tested by comparing the mean response in terms of the energy sum and the number of hits for all runs separately to the average mean response of all selected pion events. The deviations are found to be 0.9±0.1 % in $\langle E_{sum} \rangle$ and 1.4±0.2 % in $\langle N_{hits} \rangle$, see Figure 5.3.
In the following, for the energies reconstructed from the energy sum, the systematic uncertainty of the beam energy, the uncertainty from the MIP to GeV conversion, the uncertainty arising from the detector stability and the statistical errors are added in quadrature. For the energies reconstructed using the number of hits, the 0.9 % uncertainty from the MIP to GeV conversion is not taken into account.

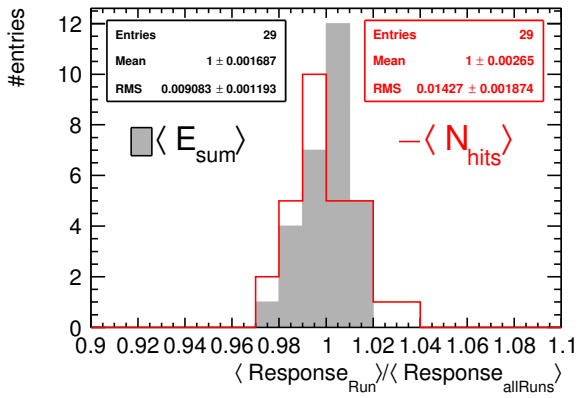



**Figure 5.3:** Detector stability in the observables $\langle E_{sum} \rangle$ and $\langle N_{hits} \rangle$ for pions with initial energies 10-80 GeV.

### 5.3.1 Systematic Uncertainties on the Simulation

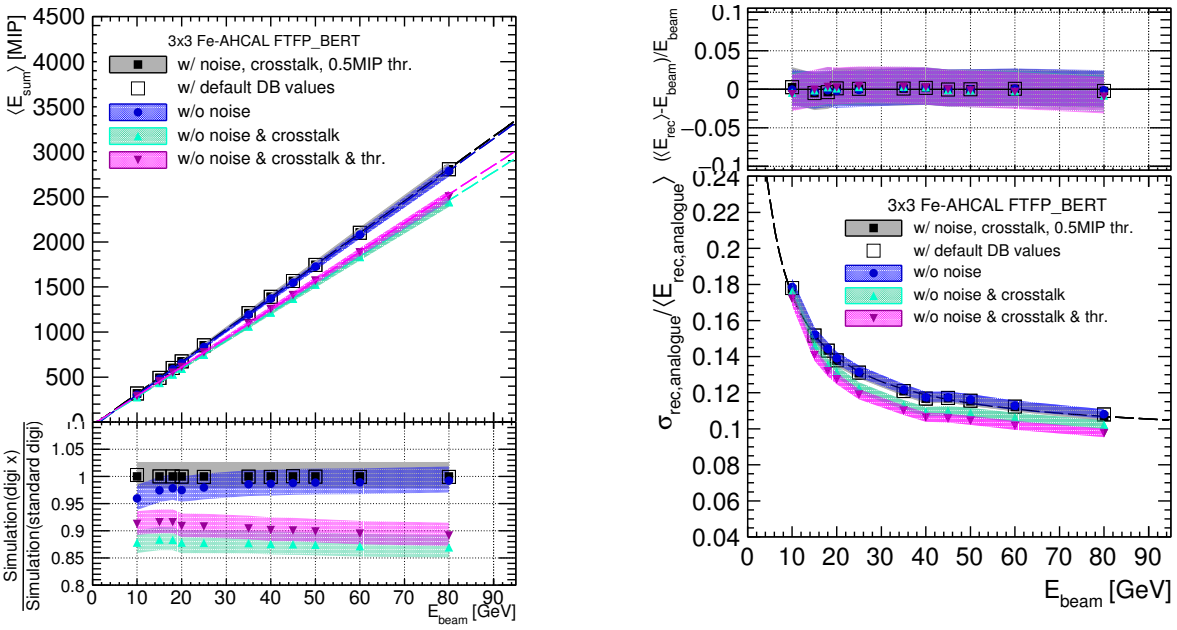
A systematic error on the simulations is estimated by generating two additional simulation samples with a different light crosstalk between neighbouring cells. In the standard digitisation a light leakage of 15 % per tile is assumed, but independent measurements of the crosstalk

showed values of 10 % and 18 % as well [85] [92]. These deviations have been simulated and uncertainties on the mean energy sum (number of hits) of +2.2/-2.6 % (+3.4/-4.2 %) have been found. In the following these uncertainties are added in quadrature to the statistical uncertainties of the simulations.

## 5.4 AHCAL Simulation with $1 \times 1\,\mathrm{cm}^2$ Granularity – A Study of Noise, Crosstalk and Thresholds

In order to study the impact of the granularity on the energy reconstruction and resolution of the AHCAL, a simulation of the AHCAL with a granularity of $1 \times 1\,\mathrm{cm}^2$ is performed. The original GEANT4 and MOKKA simulations of the testbeam setup with the physics lists FTFP_BERT and QGSP_BERT have a granularity of $1 \times 1\,\mathrm{cm}^2$. Within the digitisation the energy depositions are merged into the desired cell size. Therefore, to study the impact of the granularity only, the digitisation is repeated with a cell size of $1 \times 1\,\mathrm{cm}^2$ while the selected events stay the same.
In the standard digitisation of the $3 \times 3\,\mathrm{cm}^2$ AHCAL simulation each cell is calibrated individually with the calibration constants extracted from the data runs. However, for the simulation of a $1 \times 1\,\mathrm{cm}^2$ AHCAL no individual calibration constants are available, therefore average calibration constants are used to mimic the SiPM response.
In the standard digitisation of the $3 \times 3\,\mathrm{cm}^2$ AHCAL, dedicated noise runs are added event by event to the simulation. For the $1 \times 1\,\mathrm{cm}^2$ AHCAL simulation no noise runs have been recorded, but the newest SiPMs show a significantly decreased noise level [47]. Therefore the $1 \times 1\,\mathrm{cm}^2$ AHCAL simulation does not loose its predictive power by neglecting noise.
The standard digitisation of the $3 \times 3\,\mathrm{cm}^2$ AHCAL simulation includes a tile crosstalk of 15 % and the threshold is set to 0.5 MIP. For the $1 \times 1\,\mathrm{cm}^2$ AHCAL digitisation a crosstalk of 15 % is included as well, even though a measurement of the crosstalk between $1 \times 1\,\mathrm{cm}^2$ tiles has not been performed. The currently developed new tile designs includes a wrapping in reflective foil, which eliminates the issue of crosstalk.
The threshold for the $1 \times 1\,\mathrm{cm}^2$ simulation needs to be adjusted since the reduced cell size causes a decreased mean energy deposition per cell.

Due to these differences in the digitisation, all four effects are studied in the simulation of the $3 \times 3\,\mathrm{cm}^2$ AHCAL using the FTFP_BERT physics list. The impact of: the default calibration values ("default DB values"), missing noise ("w/o noise"), crosstalk ("w/o crosstalk") and the threshold ("w/o thr.") on the mean measured energy $\langle E_{sum} \rangle$ for pions are summarised in Figure 5.4a. The impact of using average instead of individual calibration constants is negligible (compare filled black squares with open black squares). The noise increases the response in an energy dependent way, showing the strongest effect on 10 GeV runs of 4 % and less than 1 % on 80 GeV runs (compare blue dots).



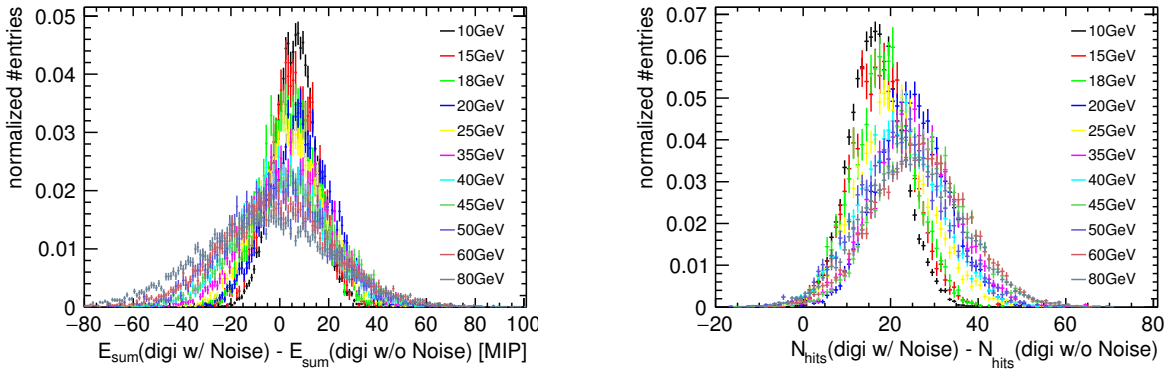

**a)** $\langle E_{sum} \rangle$ **b)** $\frac{\sigma_{rec,analogue}}{\langle E_{rec,analogue} \rangle}$

**Figure 5.4:** a) The top plot shows the mean measured energy $\langle E_{sum} \rangle$ for 10 to 80 GeV pions, the bottom plot shows the ratio of the $3 \times 3\,cm^2$ AHCAL simulations and different digitisation configurations to the standard digitisation. b) The top plot shows the non-linearities of the reconstructed energies and the bottom plot shows the relative energy resolutions of the $3 \times 3\,cm^2$ AHCAL simulations for different digitisation configurations. The black squares show the standard digitisation, the open squares the digitisation using average calibration constants and the blue dots the digitisation without noise. The green and magenta triangles show the results of the digitisation without noise and crosstalk and without noise, crosstalk and threshold. The error bands represent the systematic uncertainties.

This is consistent with the observation of a shift in the mean energy originating from noise, shown in Figure 5.5a. While the added noise mainly increases the measured energy of pion showers below 40 GeV, it has in average no impact on the mean visible energy for high energetic showers. The mean number of hits per event that originate from noise increases with the beam energy, see Figure 5.5b. This can be explained by the increased number of calorimeter cells affected by the particle shower. The conclusion of these observations is that the energy independent noise level of the AHCAL affects the energy measurements of the showers differently and is dependent on the beam energy.

The 15 % crosstalk results into an average increase in the response of around 10 % (compare magenta triangles with blue dots in Figure 5.4a). This does not originate from an increase in measured energy but from the reduced MIP value determined in the calibration. The removal of the threshold increases the mean response by 4 % at 10 GeV and 3 % at 80 GeV.

The straight line fits in Figure 5.4a are used to reconstruct the energy, following the principles explained in Section 4.2, and including an offset parameter N: $\langle E_{sum} \rangle = (E_{beam} - N) / (\frac{e}{\pi} \cdot \omega \cdot c)$. This ensures a good linearity after the energy reconstruction (see the top plot in Figure 5.4b) and thus allows a comparison between the energy resolutions. The reconstruction parameters




**a)** $E_{noise}$ **b)** $N_{noise}$

**Figure 5.5:** a) The total measured energy per event due to noise $E_{noise}$ in MIP in the selected pion events for all beam energies from 10 to 80 GeV. b) The total number of hits due to noise $N_{noise}$ in the selected pion events for all beam energies from 10 to 80 GeV. These values have been determined by comparing $E_{sum}$ and $N_{hits}$ per event for the standard digitisation including noise and the digitisation without noise.

**Table 5.2:** The energy reconstruction parameters used for the $3 \times 3\,cm^2$ AHCAL simulation, digitised in several configurations.

| Simulation (digi configuration) | $\omega$ [GeV/MIP] | N [GeV] |
|---|---|---|
| standard | 0.024 | 0.80 |
| default DB values | 0.024 | 0.77 |
| w/o noise | 0.024 | 1.12 |
| w/o noise and crosstalk | 0.027 | 0.70 |
| w/o noise and crosstalk and threshold | 0.026 | 0.55 |

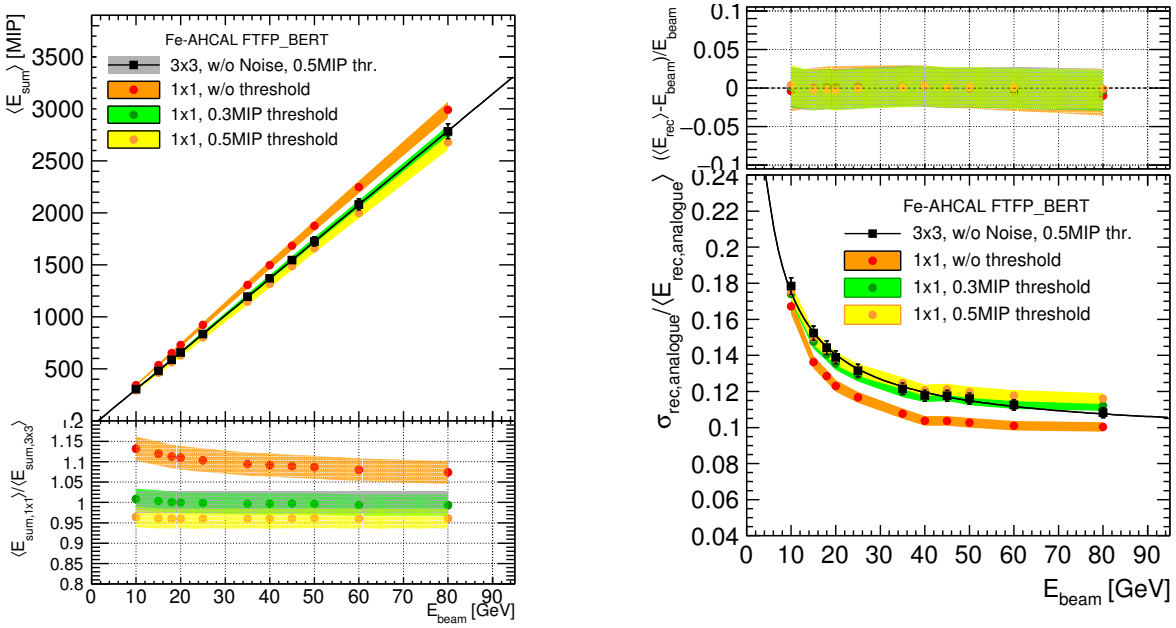
are summarised in Table 5.2 and reveal an increase of the offset when the noise is removed. This leads to the conclusion that noise can decrease the offset parameter by compensating for the negative offset introduced by the threshold.

The impact of average calibration constants, noise, crosstalk and threshold on the resolution for pions is shown in Figure 5.4b. Like already seen in the mean energy response, the average calibration constant does not have an effect on the resolution. In case of a set 0.5 MIP threshold, the same is true for noise. This confirms that a threshold of 0.5 MIP in the $3 \times 3\,cm^2$ AHCAL data is efficient for the suppression of noise.

The energy resolution of the simulation without noise and crosstalk (see green triangles in Figure 5.4b) improves. An additional improvement of the resolution is observed after removing the threshold (compare the magenta triangles).

To determine a realistic threshold for the $1 \times 1\,cm^2$ AHCAL simulation, the mean visible energy in MIP is compared for a series of different thresholds to the values obtained for the $3 \times 3\,cm^2$ AHCAL simulation without noise and applied 0.5 MIP threshold, see Figure 5.6a.

The best agreement is seen for a 0.3 MIP threshold, which is a value realistically achievable. After the individual energy reconstruction, using the parameters determined from the fit to the mean response, the relative energy resolution for 10 to 80 GeV pions is show in Figure 5.6b. It is observed that with a comparable analogue response, the $1 \times 1\,cm^2$ and $3 \times 3\,cm^2$ AHCAL achieve a comparable analogue energy resolution (compare black squares and green dots in Figure 5.6b). Consequentially, in the following a threshold of 0.3 MIP is always applied to the simulations of the AHCAL with $1 \times 1\,cm^2$ cell size.




**a)** $\langle E_{sum} \rangle$ **b)** $\frac{\sigma_{rec,analogue}}{\langle E_{rec,analogue} \rangle}$

**Figure 5.6:** a) The top plot shows the mean measured energy $\langle E_{sum} \rangle$ for 10 to 80 GeV pions, the bottom plot shows the ratio of the $1 \times 1\,cm^2$ AHCAL simulations with different thresholds to the $3 \times 3\,cm^2$ AHCAL simulation without noise and applied 0.5 MIP threshold (black squares). b) The top plot shows the non-linearities of the reconstructed energies and the bottom plot shows the relative energy resolutions of the $1 \times 1\,cm^2$ AHCAL simulation with different thresholds compared to the $3 \times 3\,cm^2$ AHCAL simulation without noise and applied 0.5 MIP threshold (black squares). The error bands represent the systematic uncertainties.

## 5.5 Energy Reconstruction and Linearity

The goal of this analysis is a direct comparison of the reconstruction methods: analogue, digital, semi-digital and software compensation algorithms, applied to the same AHCAL data and the simulation of the AHCAL with a cell size of $1 \times 1\,\mathrm{cm}^2$. The same method, described in Section 4.1, is used on all distributions to extract the mean reconstructed energy and the resolution.
Earlier studies of the testbeam data used the entire setup for the energy reconstruction. The energy in the ECAL and the TCMT complemented the measurements of the HCAL [71]. Here the goal is to study the details of the energy reconstruction in the HCAL. Therefore, to be independent from the reconstruction procedures used by the other sub-detectors, the TCMT measurements are not used, while the information from the ECAL is only used for the event selection. A fixed value of $0.3232 \pm 0.0002(\mathrm{stat.}) \pm 0.0322(\mathrm{syst.})\,\mathrm{GeV}$ is taken as contribution of the track in the ECAL to the total shower energy (see Appendix A). In this section the results of the different energy reconstruction procedures are always shown for the two granularities next to each other; on the left the $3 \times 3\,\mathrm{cm}^2$ AHCAL data and MC comparison and on the right the results from the $1 \times 1\,\mathrm{cm}^2$ AHCAL simulations. In general the parameters needed for the energy reconstructions are extracted for the $3 \times 3\,\mathrm{cm}^2$ AHCAL from the testbeam data and also used for the simulated data samples. For the $1 \times 1\,\mathrm{cm}^2$ AHCAL the parameters are extracted from the simulation using the FTFP_BERT physics list, which is the physics list that describes the data best, and also used for the QGSP_BERT samples.

### 5.5.1 Analogue

The analogue reconstructed energy for the $3 \times 3\,\mathrm{cm}^2$ Fe-AHCAL data and MC is given by

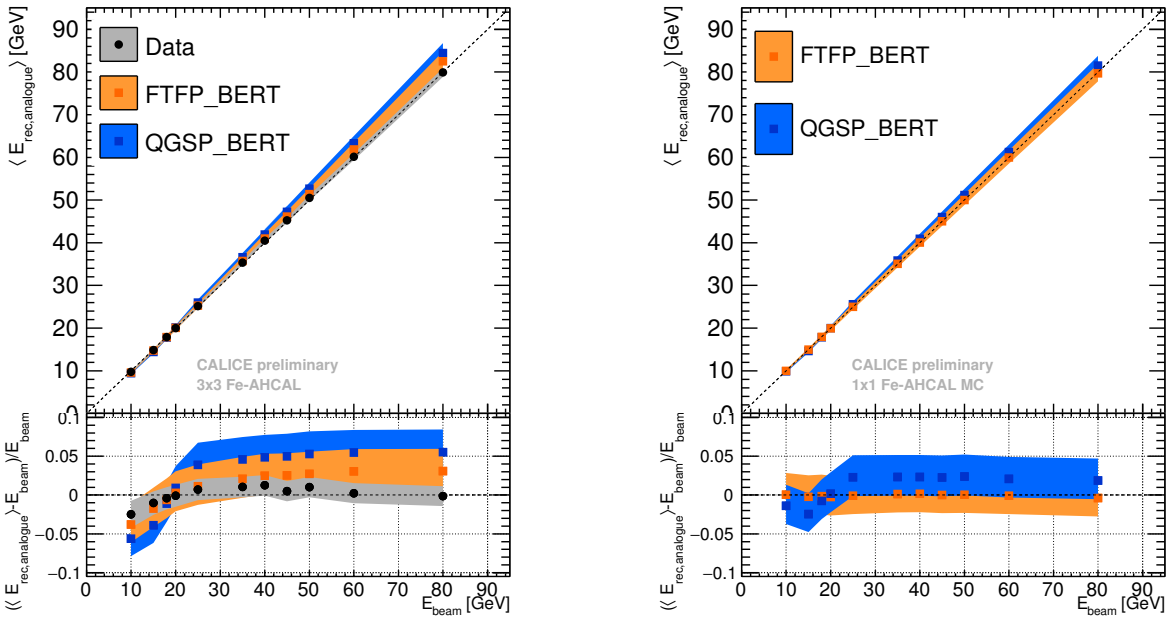
$$\mathrm{E_{rec,analogue}} = 0.3232\,\mathrm{GeV} + \frac{\mathrm{e}}{\pi} \cdot \omega \cdot \mathrm{E_{sum}} \cdot \mathrm{c}, \tag{5.2}$$

with the energy contribution of 0.3232 GeV from the track in the ECAL and the same variables as in Equation 4.3: $\frac{\mathrm{e}}{\pi} = 1.19$, $\omega = 0.02364\,\mathrm{GeV/MIP}$ and $\mathrm{c} = 1.04$. To determine the energy reconstruction parameters for the $1 \times 1\,\mathrm{cm}^2$ AHCAL simulation, the mean energy sum of the FTFP_BERT physics list is fitted, using

$$\langle \mathrm{E_{sum}} \rangle = \frac{(\mathrm{E_{beam}} - 0.3232\,\mathrm{GeV} - \mathrm{n_{1\times1}})}{\frac{\mathrm{e}}{\pi} \cdot \omega \cdot \mathrm{c_{1\times1}}}, \tag{5.3}$$

where $\mathrm{n_{1\times1}}$ is a parameter allowing an offset and $\mathrm{c_{1\times1}}$ is a factor to adjust the conversion factor $\omega$. Assuming $\mathrm{E_{beam}} = \mathrm{E_{rec,analogue1\times1}}$, the energy can be reconstructed as

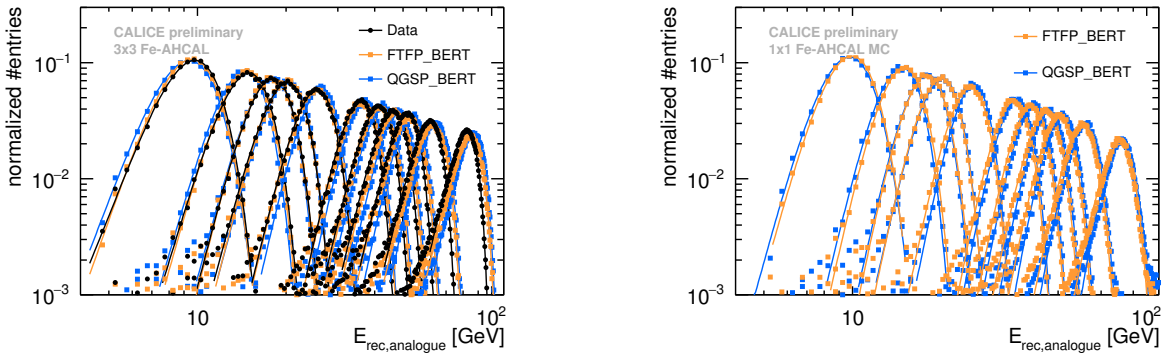
$$\mathrm{E_{rec,analogue1x1}} = 0.3232\,\mathrm{GeV} + \mathrm{n_{1\times1}} + \mathrm{E_{sum}} \cdot \frac{\mathrm{e}}{\pi} \cdot \omega \cdot \mathrm{c_{1\times1}}. \tag{5.4}$$




**a)** $3 \times 3\,\mathrm{cm}^2$ **b)** $1 \times 1\,\mathrm{cm}^2$

**Figure 5.7:** Mean analogue reconstructed energy for pion showers versus beam energy; The testbeam data is represented by black dots, the simulations using the FTFP_BERT and QGSP_BERT physics list in orange and blue squares, respectively. The bottom plots show the residuals to the beam energy with the bands indicating the systematic and statistical uncertainties. The statistical errors are smaller than the markers.

The analogue response is sensitive to two effects: First the applied threshold, which introduces a negative offset for the linear response (less than 0 MIP for a beam energy equal to 0 GeV) by requiring a minimum hit energy. Second the noise, that increases the response and generates a positive offset of the response (more than 0 MIP for a beam energy equal to 0 GeV). Equation 5.3 allows an offset $n_{1\times1}$ and a scaling of the response $c_{1\times1}$, which is needed to achieve a satisfactory linearity after the energy reconstruction, see Figure 5.7b. The positive value of $n_{1x1}$=0.98 GeV can be traced back to the threshold and the lack of noise in the $1 \times 1\,\mathrm{cm}^2$ AHCAL simulation. In the $3 \times 3\,\mathrm{cm}^2$ AHCAL on the other hand, both effects usually compensate each other. The scaling needed in the $1 \times 1\,\mathrm{cm}^2$ AHCAL is found to be one percent, $c_{1x1}$=1.01. The comparison of the analogue reconstructed energy distributions between data and simulation of the $3 \times 3\,\mathrm{cm}^2$ AHCAL is shown in Figure 5.8a. The mean reconstructed energy versus the beam energy and the non-linearity is shown in Figure 5.7a. Similar to the observations in previous analyses, e.g. [62], the FTFP_BERT and the QGSP_BERT predictions lie slightly below the data at low energies and exceed the data by a few percent at large beam energies. The $1 \times 1\,\mathrm{cm}^2$ AHCAL MC samples show the same trend between FTFP_BERT and QGSP_BERT, see Figures 5.8b and 5.7b.



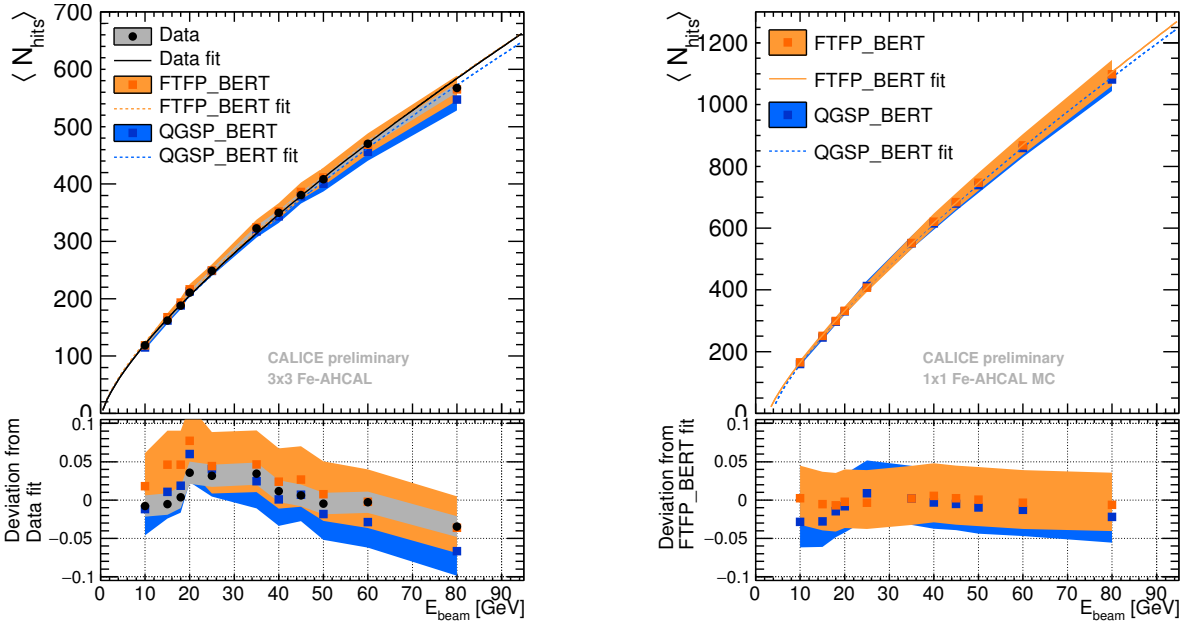

**a)** $3 \times 3\,\mathrm{cm}^2$ **b)** $1 \times 1\,\mathrm{cm}^2$

**Figure 5.8:** Analogue reconstructed energy distributions for the beam energies from 10 to 80 GeV; The black dots show the testbeam data, the orange squares show the FTFP_BERT and the blue squares the QGSP_BERT simulated $E_{rec,analogue}$ distributions. The corresponding Novosibirsk fits are represented by solid lines.

### 5.5.2 Digital

The digital response is reconstructed from the number of hits above threshold. For the $3 \times 3\,\mathrm{cm}^2$ AHCAL the threshold value of 0.5 MIP is usually taken for AHCAL analyses ensuring a minimum contribution of noise to the reconstructed energy. A threshold of 0.3 MIP in the $1 \times 1\,\mathrm{cm}^2$ AHCAL simulation is chosen to mimic the mean analogue response of the $3 \times 3\,\mathrm{cm}^2$ AHCAL, see the discussion in Section 5.4. The mean digital responses $\langle N_{hits} \rangle$ before the correction for non-linearity are shown in Figures 5.9a and 5.9b and fitted with a power law of the form $\langle N_{hits} \rangle = a \cdot (E_{beam} - m)^b$. The corresponding fit parameters are given in the caption and reveal a more linear response in the $1 \times 1\,\mathrm{cm}^2$ granularity by a b parameter closer to 1. However, both responses of the data and the simulation show saturation. This is expected for the AHCAL granularity of $3 \times 3\,\mathrm{cm}^2$ cells, which is not well adapted to the digital reconstruction method, where several particles traversing a cell contribute the same amount to the reconstructed energy as a single particle. In the $1 \times 1\,\mathrm{cm}^2$ AHCAL simulation the saturation is reduced but still visible. The tails to smaller number of hits, see Figure 5.10b, could be a hint that the cell size of $1 \times 1\,\mathrm{cm}^2$ is still not small enough to prevent saturation. In the bottom part of the figures, the relative deviation of the $\langle N_{hits} \rangle$ from the fit function is shown. For the $3 \times 3\,\mathrm{cm}^2$ AHCAL data and simulation, the point at 20 GeV presents the strongest deviation from the fit curve, while the $1 \times 1\,\mathrm{cm}^2$ AHCAL simulations agree within their errors. The non-linearity introduced by the saturation is corrected on an event-by-event basis, assuming $E_{rec,digital} = E_{beam}$, and inverting the fit functions, which leads to

$$E_{rec,digital} = m + \sqrt[b]{\frac{N_{hits}}{a}}, \tag{5.5}$$

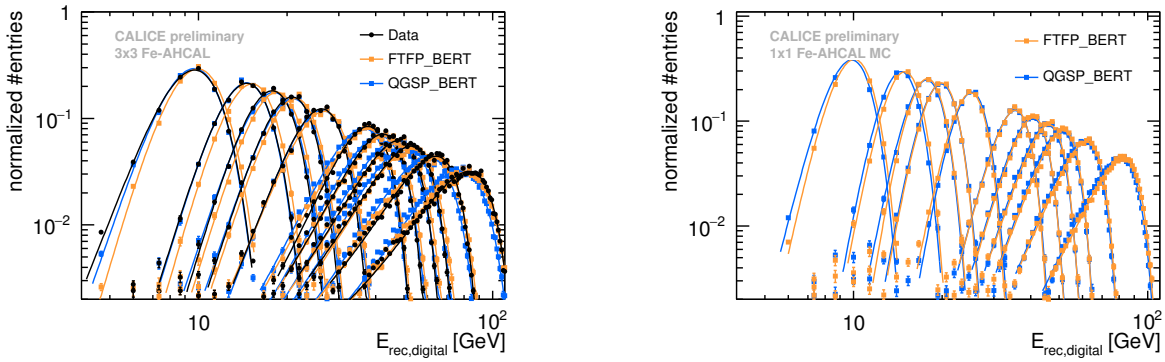



**a)** $3 \times 3\,\mathrm{cm}^2$ **b)** $1 \times 1\,\mathrm{cm}^2$

**Figure 5.9:** Mean digital response before the correction for non-linearity to pion showers, fitted with power law; $3 \times 3\,\mathrm{cm}^2$ AHCAL data: $a = 22.14 \pm 0.5\,\mathrm{GeV}^{-b}$, $b = 0.748 \pm 0.007$, $1 \times 1\,\mathrm{cm}^2$ AHCAL FTFP_BERT: $a = 34.7 \pm 7.5\,\mathrm{GeV}^{-b}$, $b = 0.80 \pm 0.05$. The plots on the bottom show the deviation from the power law fit. The bands indicate the statistical and systematic uncertainty added in quadrature, the statistical error only is smaller than the markers.

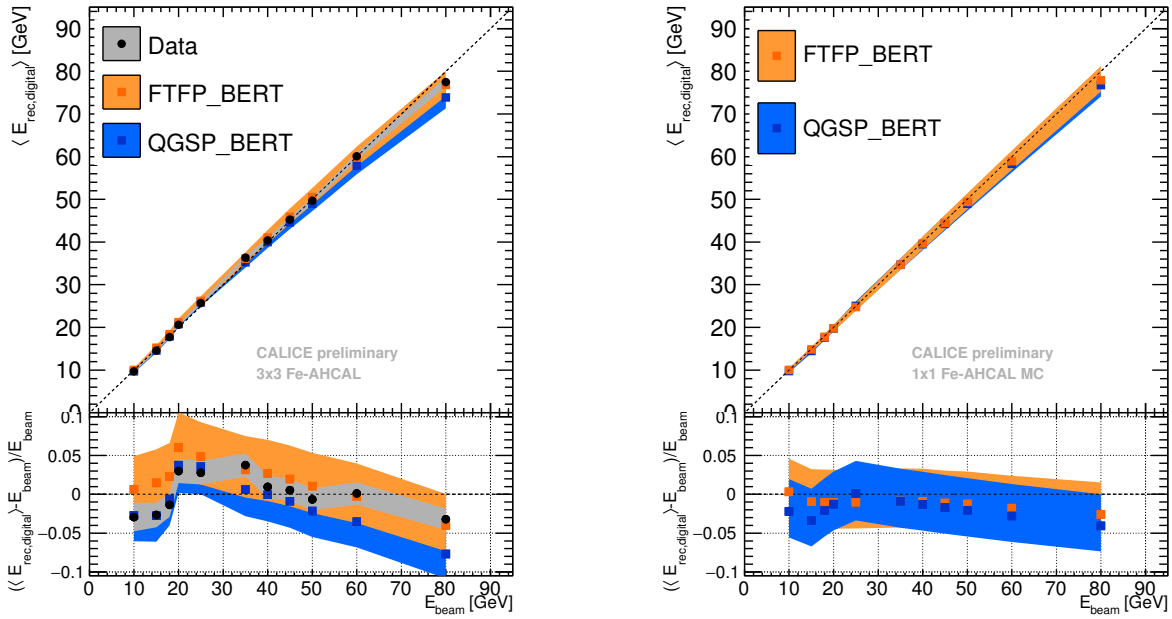
where the value of parameter m is fixed to 0.3232 GeV for the $3 \times 3\,\mathrm{cm}^2$ AHCAL and $m_{1x1} = 3.0 \pm 1.2\,\mathrm{GeV}$ determined by the fit for the $1 \times 1\,\mathrm{cm}^2$ AHCAL simulation. The parameters a and b are extracted from the fit and given in the caption of Figure 5.9. In the following, the parameters determined from the fit to the data are used to reconstruct the energy also for the simulated data of the $3 \times 3\,\mathrm{cm}^2$ AHCAL. In case of the $1 \times 1\,\mathrm{cm}^2$ AHCAL simulation, the values are taken from the FTFP_BERT simulation.

The resulting $E_{rec,digital}$ distributions for data and simulation are compared in Figure 5.10a and 5.10b. These distributions show compared to the analogue reconstructed energy distributions larger tails to the left hand side for the highest beam energies. This is more pronounced in the $3 \times 3\,\mathrm{cm}^2$ than in in the $1 \times 1\,\mathrm{cm}^2$ AHCAL. The Novosibirsk fit functions used to extract the mean and the width of the $E_{rec,digital}$ distribution are also shown. After the correction of the saturation in the mean response $\langle N_{hits} \rangle$, the mean reconstructed energies show a linear behaviour within $\pm 4\,\%$ (Figure 5.11a). Since the simulations of the $3 \times 3\,\mathrm{cm}$ AHCAL are corrected with the same parameters as the data, they show slightly larger deviations from linearity, with the largest deviation for QGSP_BERT of $\sim 8\,\%$ at 80 GeV. For the $1 \times 1\,\mathrm{cm}^2$ AHCAL simulation the non-linearities are b1 elow 5 % for both physics lists, see Figure 5.11b.




**a)** $3 \times 3\,\mathrm{cm}^2$ **b)** $1 \times 1\,\mathrm{cm}^2$

**Figure 5.10:** Digital reconstructed energy distributions for beam energies from 10 to 80 GeV; The black dots show the testbeam data, the orange squares show the FTFP_BERT and the blue squares the QGSP_BERT simulated $E_{rec,digital}$ distributions. The corresponding Novosibirsk fits are represented by solid lines.

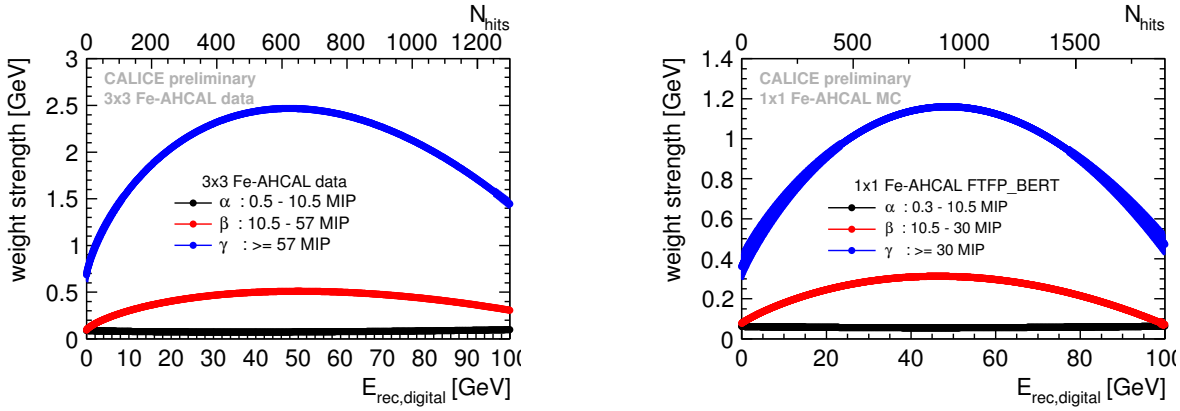



**a)** $3 \times 3\,\mathrm{cm}^2$ **b)** $1 \times 1\,\mathrm{cm}^2$

**Figure 5.11:** Mean digital reconstructed energy for pion showers; The testbeam data is represented by black dots, the simulations using the FTFP_BERT and QGSP_BERT physics list in orange and blue squares, respectively. The bottom plots show the residuals to the beam energy with the bands indicating the systematic and statistical uncertainties. The statistical errors are smaller than the markers.

### 5.5.3 Semi-Digital

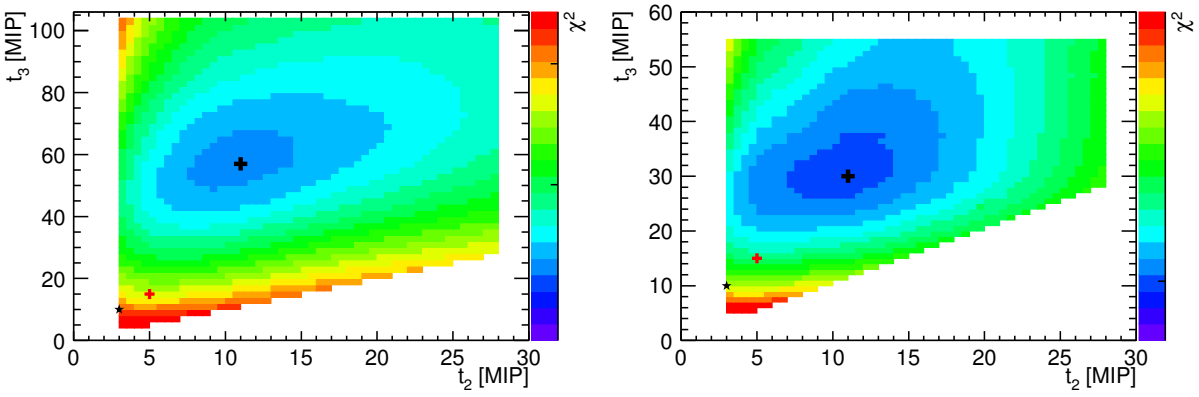
The semi-digital energy reconstruction is done using Equation 4.5, where $N_1$ is the number of hits above the first threshold $t_1$ and below the second threshold $t_2$, $N_2$ is the number of hits above $t_2$ and below the third threshold $t_3$ and $N_3$ is the number of hits above $t_3$. For the determination of the calibration weights $\alpha$, $\beta$ and $\gamma$, 20,000 events are taken from each energy data set. These events have to have a total energy sum within the RMS90 of all $E_{sum}$ to extract best possible weights for the majority of events and to be insensitive to outliers. The $\chi^2$-like function given in Equation 4.6 is minimised by ROOT using TMinuit2. The resulting weights are shown in Figure 5.12. The energy dependence of the weights is observed to be qualitatively the same for both granularities.




**a)** $3 \times 3\,\mathrm{cm}^2$ **b)** $1 \times 1\,\mathrm{cm}^2$

**Figure 5.12:** The weights for the semi-digital energy reconstruction are shown as a function of the total number of hits $N_{hits}$ and the digitally reconstruction energy $E_{rec,digital}$, which is calculated from $N_{hits}$ by using Equation 5.5. The width of the curves correspond to the statistical errors, correlations are taken into account.

In order to preserve the comparability between the energy reconstruction procedures the lowest thresholds are kept at 0.5 and 0.3 MIP for the $3 \times 3\,\mathrm{cm}^2$ and $1 \times 1\,\mathrm{cm}^2$ AHCAL, respectively. The higher thresholds have been optimised for both cell granularities by minimising the $\chi^2$ values, which give an estimate of the reconstruction accuracy, in the $t_2$-$t_3$ plane. The considered thresholds lay between 3 to 28 MIP and 5 to 105 MIP for $t_2$ and $t_3$, respectively. The results are shown in Figure 5.13 and the optimal threshold values are summarised in Table 5.3. The threshold values optimised for the Micro-MEsh Gaseous Structure (MicroMEGaS) SDHCAL prototype [93] and used for the RPC SDHCAL (see Section 2.4.1) are shown for comparison. The semi-digital response in terms of $N_1$, $N_2$ and $N_3$ is shown in Figure 5.14 and reveals a disagreement between data and MC in the number of very high energy hits $N_3$. This observation of 1.5 times higher number of high energy hits in simulation for high beam energies is consistent with the observed analogue hit energy spectra, that are further discussed in Section 5.5.4.
The semi-digital energy reconstruction of the $3 \times 3\,\mathrm{cm}^2$ AHCAL as well as the $1 \times 1\,\mathrm{cm}^2$ AH-




**a)** $3 \times 3\,\mathrm{cm}^2$ **b)** $1 \times 1\,\mathrm{cm}^2$

**Figure 5.13:** The $\chi^2$ values of Equation 4.6 for the semi-digital reconstruction as a function of the thresholds $t_2$ and $t_3$ for the $3 \times 3$ AHCAL data in a) and the $1 \times 1$ AHCAL simulation with FTFP_BERT in b). The black crosses represent the optimal value found, the red crosses mark the thresholds position previously used for the MicroMEGaS based SDHCAL and the black stars show the approximated threshold settings of the RPC based SDHCAL (see Section 2.4.1).

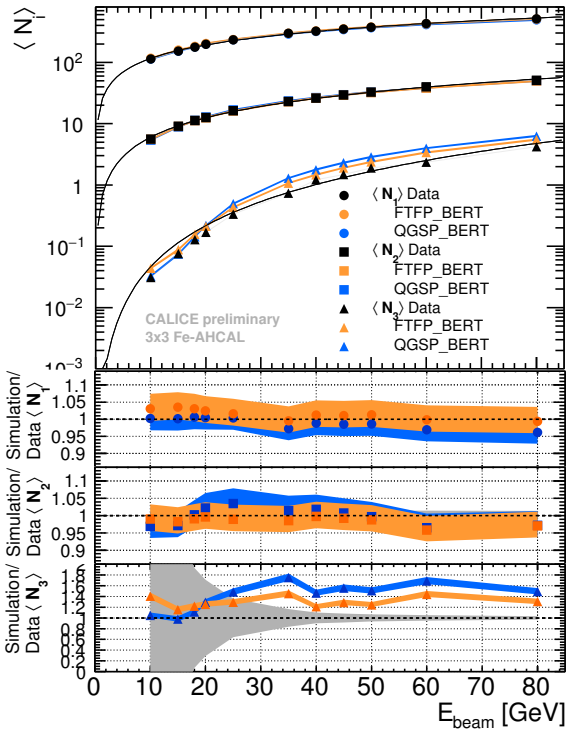
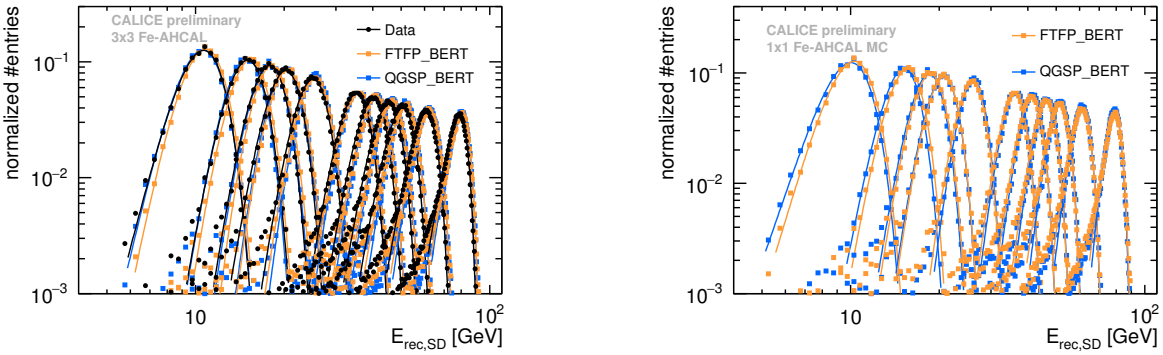



**Figure 5.14:** Mean semi-digital response of the $3 \times 3\,\mathrm{cm}^2$ AHCAL to pion showers; The testbeam data (black markers), FTFP_BERT (orange markers) and QGSP_BERT (blue markers) simulation for hits above the first, below the second threshold $N_1$, hits above the second, below the third threshold $N_2$ and hits passing the third threshold $N_3$. The lines represent fits with a power law and the bands the statistical and systematic uncertainty added in quadrature.

**Table 5.3:** The optimised thresholds used for the semi-digital reconstruction for the AHCAL with both cell sizes and the SDHCAL prototype with MicroMEGaS and RPCs (see Section 2.4.1).

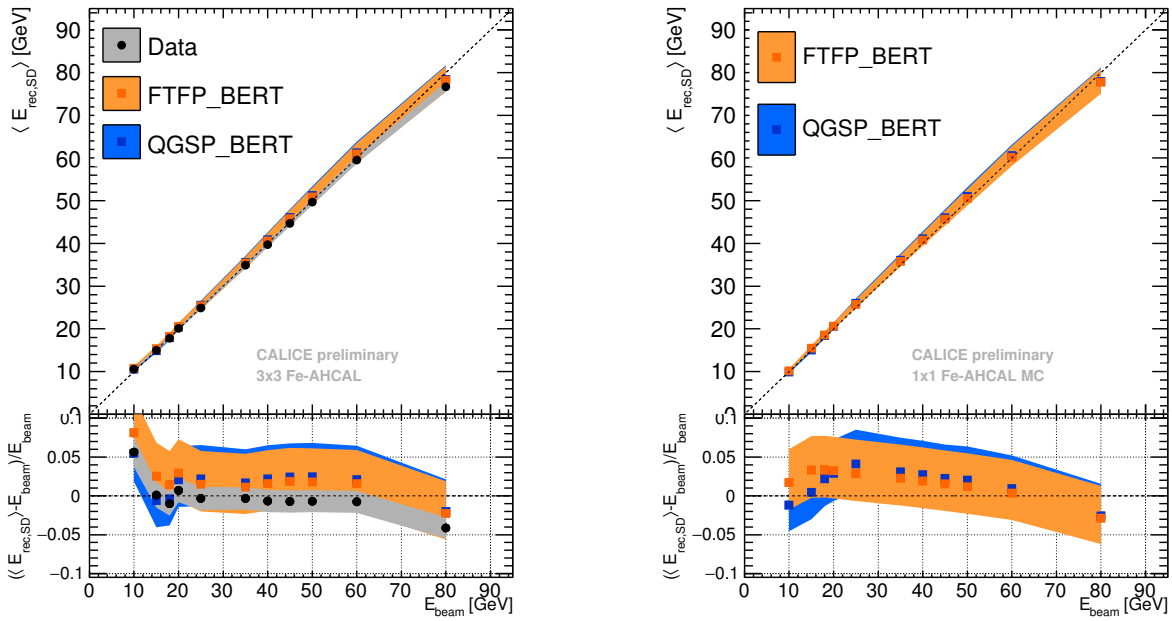
| | $t_1$ | $t_2$ | $t_3$ |
|---|---|---|---|
| $3 \times 3\,\mathrm{cm}^2$ Sci-Fe AHCAL | 0.5 | 10.5 | 57 |
| $1 \times 1\,\mathrm{cm}^2$ Sci-Fe AHCAL | 0.3 | 10.5 | 30 |
| $1 \times 1\,\mathrm{cm}^2$ MicroMEGaS SDHCAL [93] | 0.5 | 5 | 15 |
| $1 \times 1\,\mathrm{cm}^2$ RPC SDHCAL | $0.07 \pm 0.02$ | $3.3 \pm 0.9$ | $10 \pm 2.67$ |




**a)** $3 \times 3\,\mathrm{cm}^2$ **b)** $1 \times 1\,\mathrm{cm}^2$

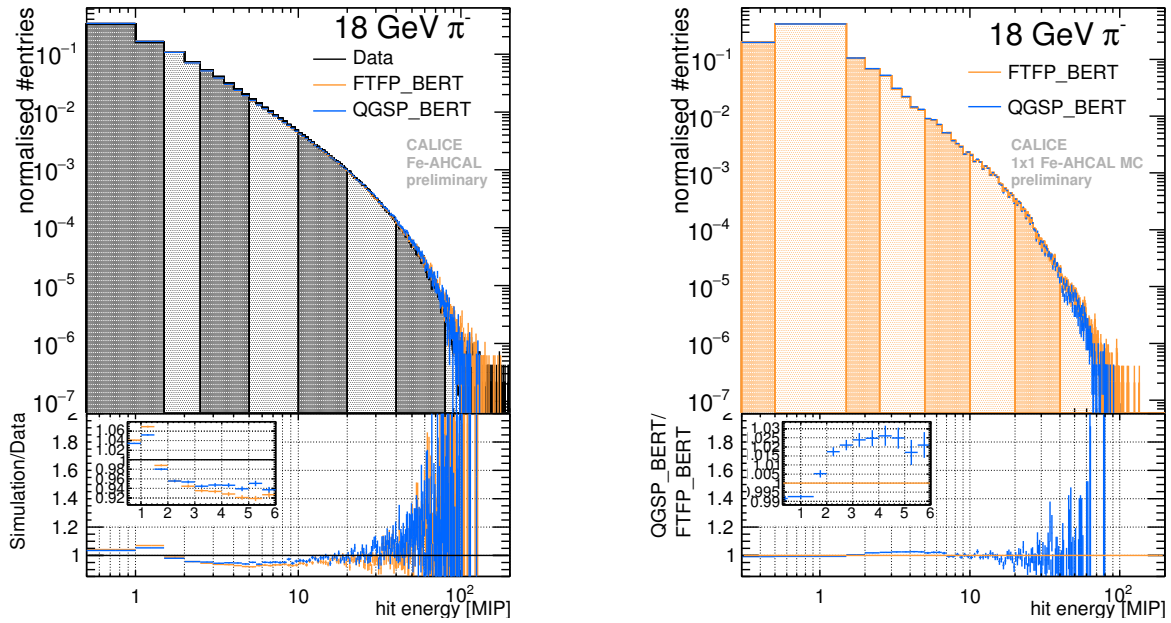
**Figure 5.15:** Semi-digital reconstructed energy distributions for beam energies from 10 to 80 GeV; The black dots show the testbeam data, the orange squares show the FTFP_BERT and the blue squares the QGSP_BERT simulated $E_{rec,SD}$ distributions. The corresponding Novosibirsk fits are represented by solid lines.

CAL data leads to much smaller tails towards low energies compared to the digital energy reconstruction. In Figure 5.15a and 5.15b the distributions of the semi-digital reconstructed energy for the testbeam data and simulated FTFP_BERT and QGSP_BERT events are shown. A good agreement is observed for all energies. The non-linearities are smaller than $\pm 5\,\%$ for all energies except for 10 GeV for $3 \times 3\,\mathrm{cm}^2$, see the bottom plots of Figures 5.16a and 5.16b.




**a)** $3 \times 3\,\mathrm{cm}^2$ **b)** $1 \times 1\,\mathrm{cm}^2$

**Figure 5.16:** Mean semi-digital reconstructed energy for pion showers; The testbeam data is represented by black dots, the simulations using the FTFP_BERT and QGSP_BERT physics list in orange and blue squares, respectively. The bottom plots show the residuals to the beam energy with the bands indicating the systematic and statistical uncertainties. The statistical errors are smaller than the markers.




**a)** $3 \times 3\,\mathrm{cm}^2$ **b)** $1 \times 1\,\mathrm{cm}^2$

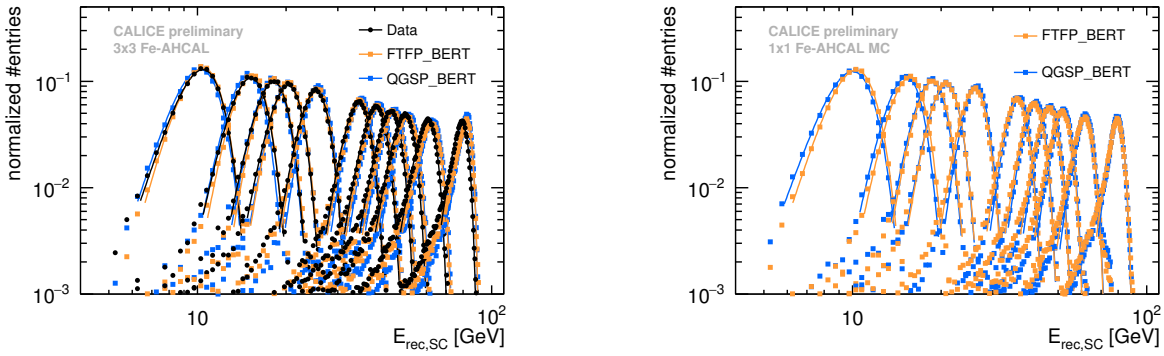
**Figure 5.17:** The hit energy spectra for 18 GeV $\pi^-$ showers in the $3 \times 3$ and $1 \times 1\,\mathrm{cm}^2$ AHCAL, the testbeam data shown in black, compared to simulated FTFP_BERT and QGSP_BERT data samples in orange and blue. At the bottom parts the deviation between data and simulation, and between FTFP_BERT and QGSP_BERT simulations are shown. The errors shown are purely statistical.

### 5.5.4 Analogue Software Compensation

The energy reconstruction with the software compensation algorithm is done following Equation 4.7, where the applied weights are dependent on the individual hit energies $E_i$ and the total energy via the visible energy $E_{sum}$. For practical reasons, the number of hit energy ranges with constant hit energy weights is in this analysis limited to 8, which still requires the determination of $8 \times 3 = 24$ parameters in the $\chi^2$ minimisation, following Equation 4.6. Each weight is described as a 2nd order polynomial of $E_{sum}$. This classification of the hits is visualised by shadowed areas in the hit energy spectra, see Figures 5.17a and 5.17b. The border settings have not been optimised, but chosen to describe a typical MIP and each class is ensured to have enough statistics. The comparison between data and simulations reveals a good agreement, always better than 10 % for beam energies below 30 GeV and in the hit energy range not suffering from low statistics. For higher beam energies the simulations overestimate the number of hits with energy >60 MIP. The differences between the physics lists are smaller than 5 %.

When going from $3 \times 3$ to $1 \times 1\,\mathrm{cm}^2$ the energy range is affected by the threshold lowered from 0.5 to 0.3 MIP, and the significantly smaller hit energy densities in the $1 \times 1\,\mathrm{cm}^2$ cells. Therefore the hit energy ranges of the weights are adapted to the granularity.

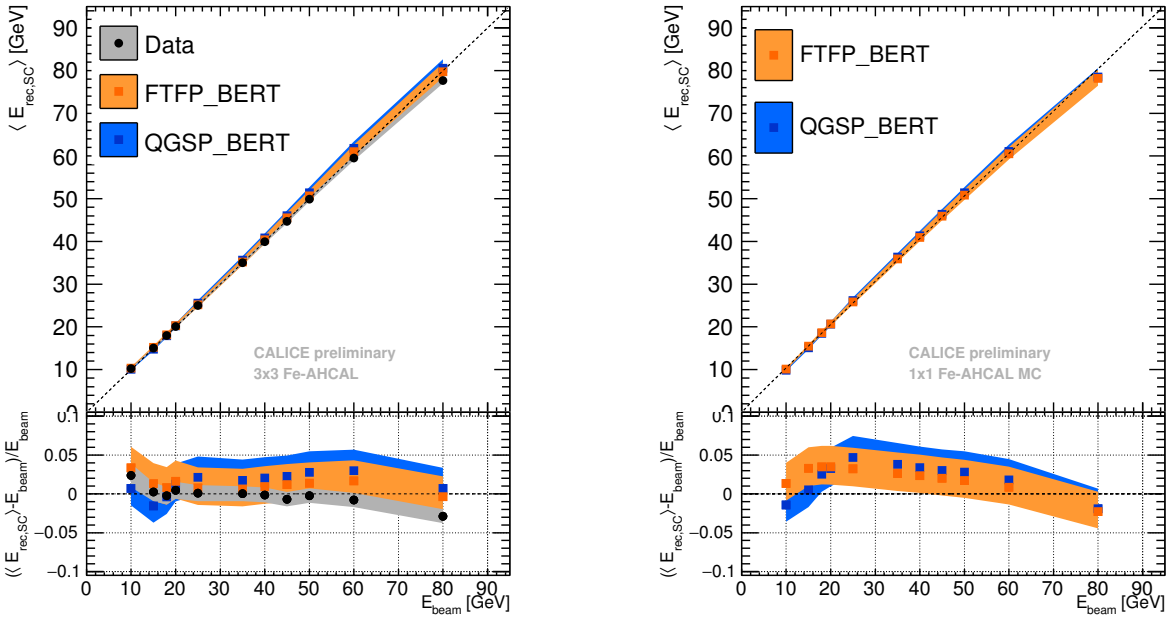
The weights used for the energy reconstruction in the $3 \times 3\,\mathrm{cm}^2$ AHCAL are determined from




**a)** $3 \times 3\,\mathrm{cm}^2$ **b)** $1 \times 1\,\mathrm{cm}^2$

**Figure 5.18:** Reconstructed energy distributions for the beam energies from 10 to 80 GeV after applying the software compensation algorithm; The black dots show the testbeam data, the orange squares show the FTFP_BERT and the blue squares the QGSP_BERT simulated $E_{rec,SC}$ distributions. The corresponding Novosibirsk fits are represented by solid lines.

data and applied to all the samples. In the $1 \times 1\,\mathrm{cm}^2$ AHCAL case, the weights are determined from the FTFP_BERT physics list and applied to both simulation samples. The weights are shown and further discussed in Section 5.7. The reconstructed energy distributions, shown together with the Novosibirsk fits in Figure 5.18a and 5.18b, exhibit narrow peaks. The fit range needed to be limited to $\mu \pm 2.5\,\sigma$ in order to achieve a satisfying $\chi^2$. In the distributions small tails to the left hand side are seen, which is expected due to the limited number of layers (energy leakage into the TCMT). In Figure 5.19a the linearity is shown, and a nice agreement between data and simulation, especially with the FTFP_BERT physics list is seen. The difference in the number of high energy hits discussed earlier, does not have a huge impact on the overall energy reconstruction. In the higher granularity case, the non-linearities for both physics lists do not exceed 5 %, see Figure 5.19b. With the exception of the two lowest beam energies, the physics lists show a nearly identical behaviour.




**a)** $3 \times 3\,\mathrm{cm}^2$ **b)** $1 \times 1\,\mathrm{cm}^2$

**Figure 5.19:** Mean reconstructed energy for pion showers after applying the software compensation algorithm; The testbeam data is represented by black dots, the simulations using the FTFP_BERT and QGSP_BERT physics list in orange and blue squares, respectively. The bottom plots show the residuals to the beam energy with the bands indicating the systematic and statistical uncertainties. The statistical errors are smaller than the markers.

## 5.6 Energy Resolution

All four reconstruction methods show a reasonable linearity, which guarantees a correct determination of the energy reconstruction in all cases. Therefore the resolutions can be compared. The impact of the AHCAL granularity on the digital or semi-digital reconstruction methods and on the resolution is expected to be strong for the highest beam energies. The functional form usually employed to fit the relative energy resolution, which consists of a stochastic, a constant and a noise term, does not accommodate for a degrading resolution at higher energies. Therefore, we introduce a fourth term with variable exponent for the energy dependence, similar to the approach used in [23]:

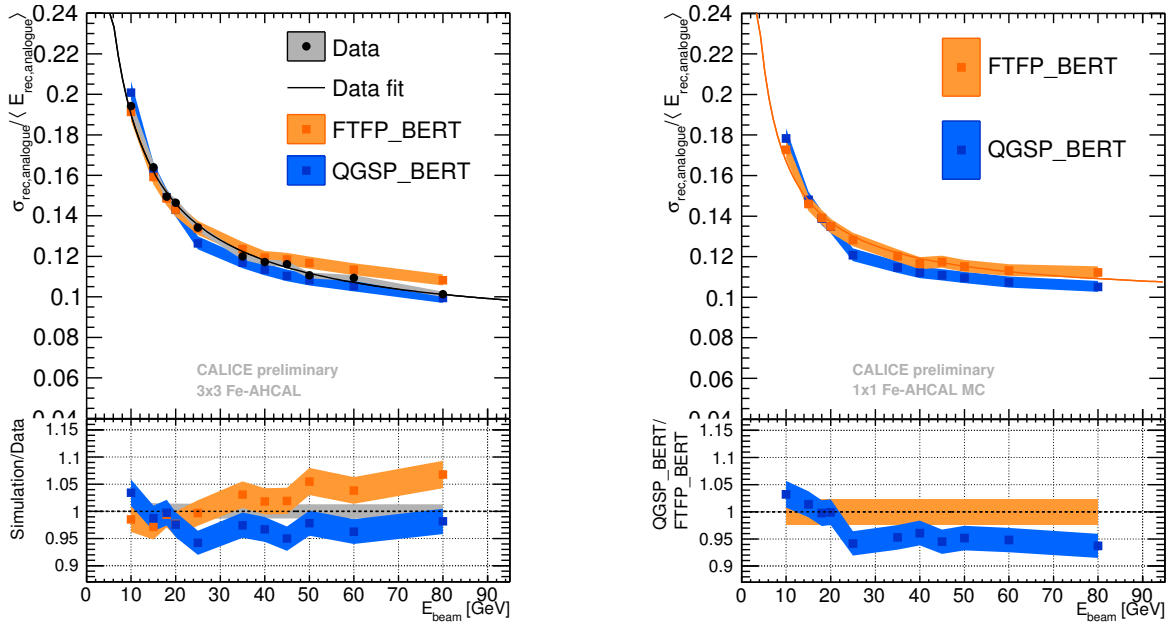
$$\frac{\sigma_{\mathrm{rec}}}{\langle \mathrm{E_{rec}} \rangle} = \frac{\mathrm{a}}{\sqrt{\mathrm{E_{beam}[GeV]}}} \oplus \mathrm{b} \oplus \frac{\mathrm{c}}{\mathrm{E_{beam}[GeV]}} \oplus \mathrm{d} \left( \frac{\mathrm{E_{beam}[GeV]}}{100} \right)^{\mathrm{e}}. \tag{5.6}$$

The fourth term can account for leakage as well as saturation effects. For each reconstruction method only the parameters needed for a reasonable description of the data are left free. A direct comparison of the extracted values between the different methods is therefore difficult, and the fits should mainly guide the eye.

### 5.6.1 Analogue Resolution

The relative resolution for the analogue energy reconstruction of the AHCAL pion data and of the corresponding FTFP_BERT and QGSP_BERT simulations are shown in Figure 5.20a as a function of the beam energy. The FTFP_BERT simulation describes the data quite well for energies below 50 GeV. For higher energies the resolution of the simulated data lies about 5 % above the data. The QGSP_BERT simulation achieves a resolution of up to 5 % better for energies higher than 20 GeV. Testbeam data and simulation show an improvement in the relative resolution with increasing energy, as expected if leakage or saturation play only a minor role. Therefore, in this case, the resolutions can be parametrised without the fourth term in Equation 5.6.

In Figure 5.20b the relative resolution of the analogue energy reconstruction of the AHCAL simulated in $1 \times 1\,\mathrm{cm}^2$ granularity is shown for the two physics lists. Since the FTFP_BERT simulation of the $3 \times 3\,\mathrm{cm}^2$ AHCAL shows the best agreement with data, the points of the FTFP_BERT simulation of the $1 \times 1\,\mathrm{cm}^2$ AHCAL are the only ones fitted. In comparison with the resolution of the $3 \times 3\,\mathrm{cm}^2$ AHCAL data, the resolution improves up to two percent in absolute values for lower energies and reaches approximately the same values for higher beam energies. These small deviations are possibly due to two differences: the lack of noise in the $1 \times 1$ simulations and the different threshold settings.




**a)** $3 \times 3\,\mathrm{cm}^2$ **b)** $1 \times 1\,\mathrm{cm}^2$

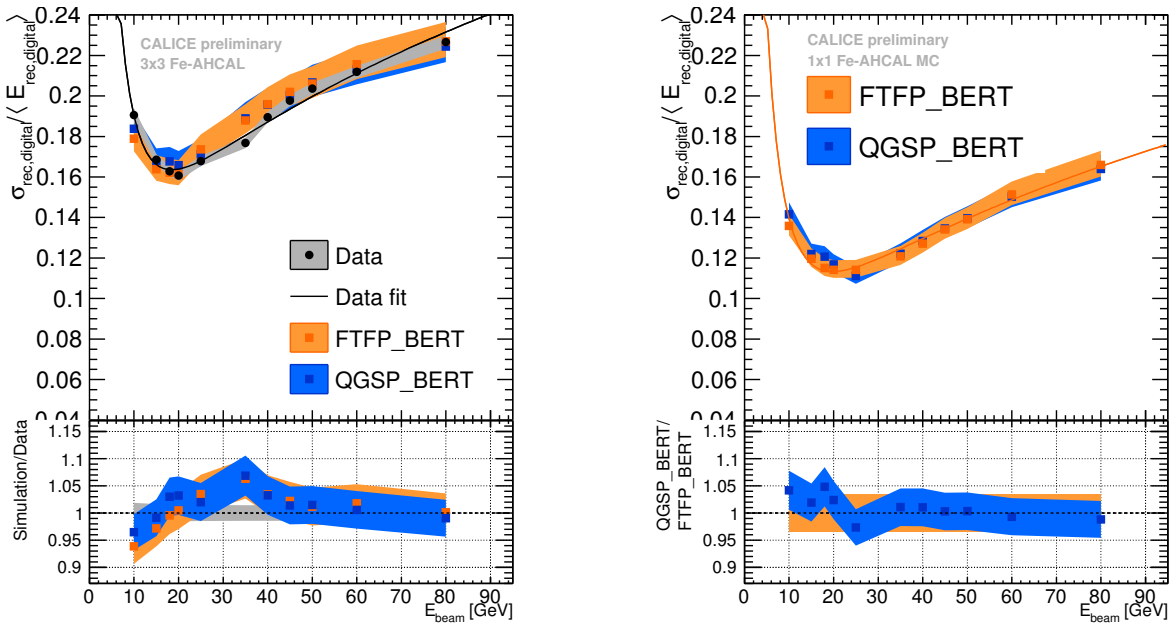
**Figure 5.20:** Analogue energy resolution for testbeam data, FTFP_BERT and QGSP_BERT simulated events, fitted with Equation 5.6. At the bottom the ratios of the testbeam simulations and data, between the QGSP_BERT and FTFP_BERT physics lists are shown.

### 5.6.2 Digital Resolution

In Figure 5.21a and 5.21b, the relative energy resolutions of the digital reconstruction method applied to AHCAL pion data, FTFP_BERT and QGSP_BERT simulations are compared. Both data and simulation of the $3 \times 3\,\mathrm{cm}^2$ AHCAL show a strong worsening of the resolution towards large energies, and a minimum resolution of about 16 % for energies around 20 GeV. Data and simulation agree very well within the errors. The strong rise at larger energies can be fitted when taking into account the fourth term in Equation 5.6. Since the lowest beam energy used in this analysis is 10 GeV, the terms decreasing with increasing energy in Equation 5.6 are not well constrained. For this reason the values for a and b are fixed to zero in the fit to data and simulation.

The finer granularity improves the resolution over the full energy range. However, the behaviour of a degradation from a certain energy onwards stays the same. The best digital resolution of the $1 \times 1\,\mathrm{cm}^2$ AHCAL is achieved around 25 GeV with about 11 %. For larger beam energies the resolution degrades slower than in the $3 \times 3\,\mathrm{cm}^2$ AHCAL.

Both physics lists agree very well within the uncertainties.

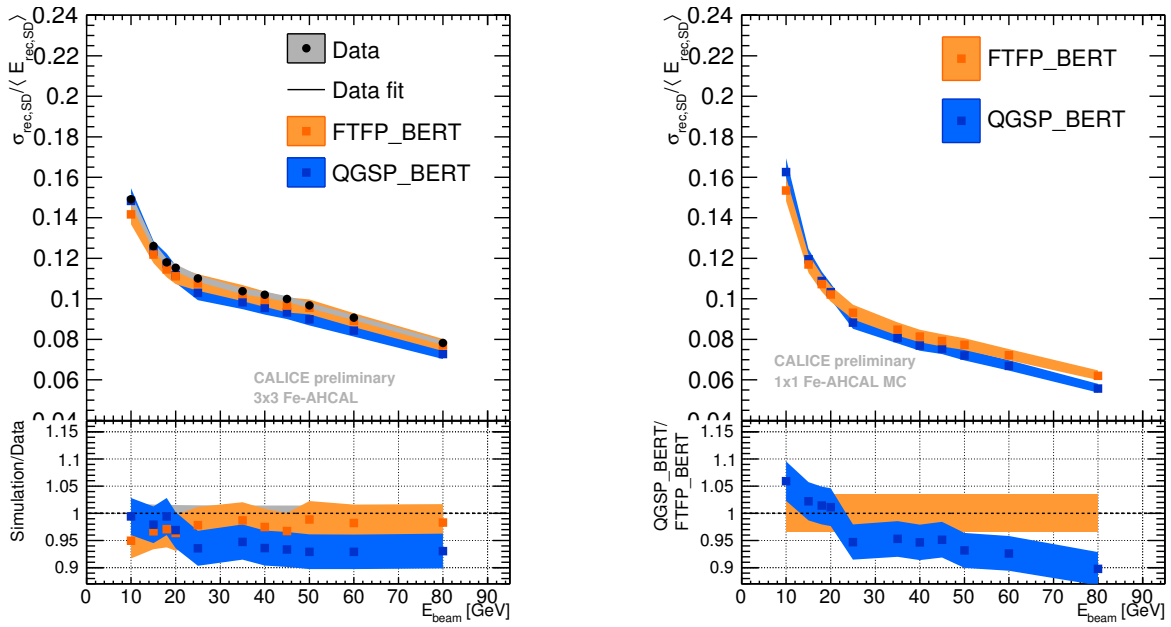



**a)** $3 \times 3\,\mathrm{cm}^2$ **b)** $1 \times 1\,\mathrm{cm}^2$

**Figure 5.21:** Digital energy resolution for testbeam data, FTFP_BERT and QGSP_BERT simulated events, fitted with Equation 5.6. At the bottom the ratios of the testbeam simulations and data, between the QGSP_BERT and FTFP_BERT physics lists are shown.

### 5.6.3 Semi-Digital Resolution

The relative resolution of the semi-digital reconstruction method shows yet a different dependence on the beam energy, see Figure 5.22a and 5.22b. The resolution observed for the $3 \times 3\,\mathrm{cm}^2$ AHCAL shows a linear decrease for energies above 20 GeV, down to 8 % at 80 GeV. This behaviour is not well described by Equation 5.6 and therefore the fit is not shown in Figure 5.22a. Overall the FTFP_BERT simulation agrees well with the data in the analysed energy range, while the QGSP_BERT simulation shows a 5-7 % better resolution in the linear region. The resolution of the simulated $1 \times 1\,\mathrm{cm}^2$ AHCAL shows an improvement with increasing beam energy, also following a nearly linear behaviour for beam energies above 20 GeV, down to about 6 % at 80 GeV.
The weighting method, which is based on a $\chi^2$ function, assumes $\sigma_E$ to follow a $\sqrt{E_{beam}}$ behaviour. This behaviour is not observed in Figure 5.22, which leeds to the assumption that the weights could be further improved by adjusting the $\chi^2$ function.

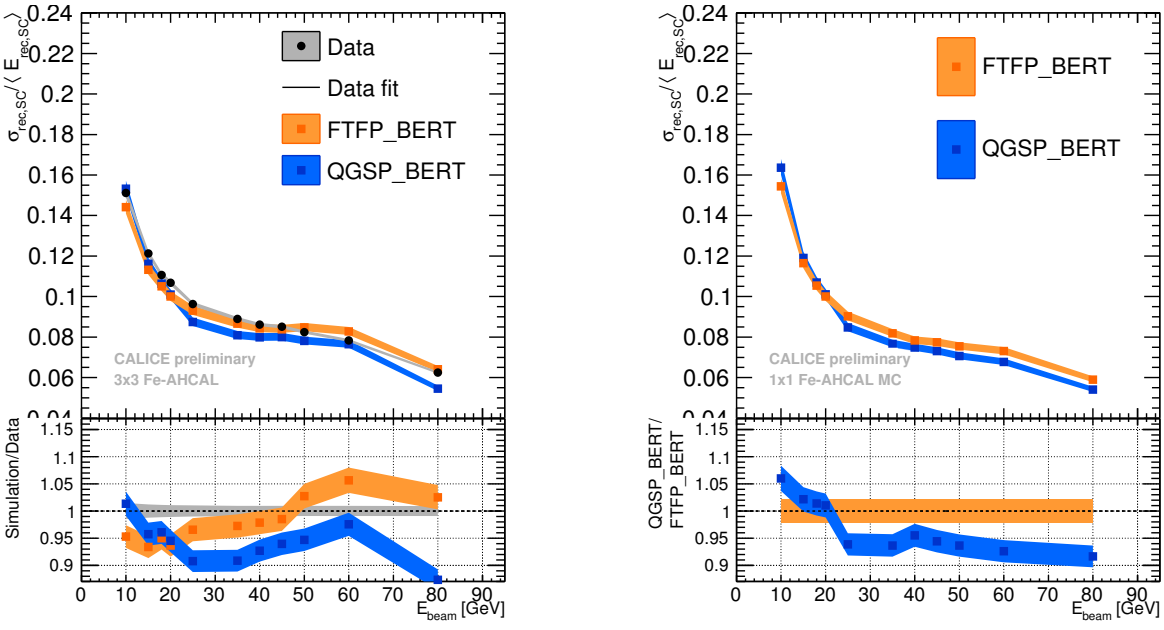



**a)** $3 \times 3\,\mathrm{cm}^2$ **b)** $1 \times 1\,\mathrm{cm}^2$

**Figure 5.22:** Semi-digital energy resolution for testbeam data, FTFP_BERT and QGSP_BERT simulated events. At the bottom the ratios of the testbeam simulations and data, between the QGSP_BERT and FTFP_BERT physics lists are shown.

### 5.6.4 Analogue Software Compensation

The relative resolution of the energy reconstructed with the analogue software compensation technique is shown in Figure 5.23a and 5.23b. The comparison of the testbeam data and simulation reveals the strongest deviations in the resolution for the QGSP_BERT physics list, which exceeds 9 % at 80 GeV. The FTFP_BERT simulation agrees with the data within 5 %. Generally, the resolutions decrease with increasing beam energy and the testbeam data achieve a resolution of 6.5 % at 80 GeV. However, the behaviour can not be well described by Equation 5.6, therefore no fit is included in Figure 5.23a and 5.23b.
The impact of the software compensation algorithm on the resolution of the $1 \times 1\,\mathrm{cm}^2$ AHCAL is shown in Figure 5.23b. The resolution improves compared to the $3 \times 3\,\mathrm{cm}^2$ AHCAL up to 1 % in absolute values in the energy range of 30 to 80 GeV. The deviations between simulations with different physics lists are the largest for this reconstruction method, reaching up to ∼10 %. This difference originates from the differences in the hit energy spectra, discussed in Section 5.5.4.




**a)** $3 \times 3\,\mathrm{cm}^2$ **b)** $1 \times 1\,\mathrm{cm}^2$

**Figure 5.23:** Energy resolution observed applying software compensation algorithms for the testbeam data, FTFP_BERT and QGSP_BERT simulated events. At the bottom the ratios of the testbeam simulations and data, between the QGSP_BERT and FTFP_BERT physics lists are shown.

## 5.7 Comparison of Semi-Digital and Software Compensation Weights

The weight determinations of the semi-digital energy reconstruction and of the software compensation algorithm are very similar. The biggest difference lies in the usage of the hit energy information. The semi-digital weights follow $\alpha_i\,[\mathrm{GeV}] = a_i + b_i \cdot N_{hits} + c_i \cdot N_{hits}^2$, thus using $N_{hits}$ as an estimate of the beam energy, and are used for the reconstruction by

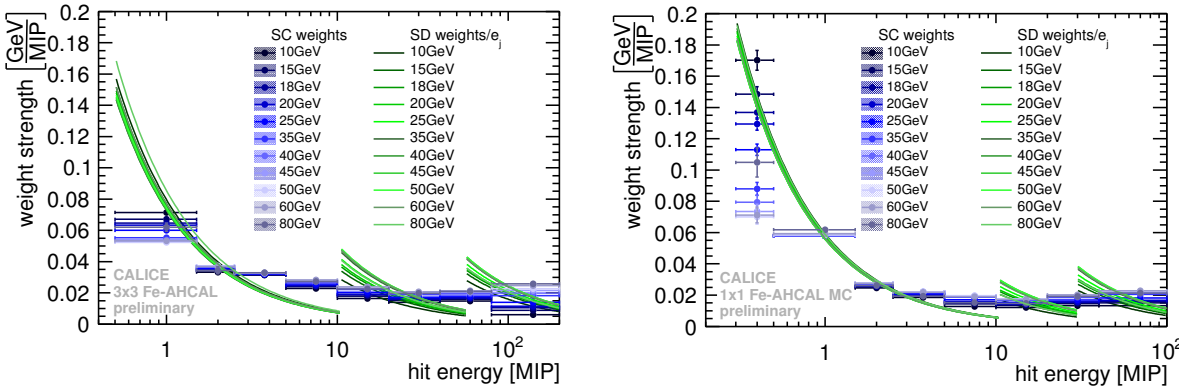
$$E_{rec,SD} = \sum_{i=1}^{3} \alpha_i\,(N_{hits}) \cdot N_i. \tag{5.7}$$

The software compensation weights follow $\omega_j \left[\frac{\mathrm{GeV}}{\mathrm{MIP}}\right] = a_j + b_j \cdot E_{sum} + c_j \cdot E_{sum}^2$ and are in this way determined using a beam energy estimate via $E_{sum}$. In the reconstruction however, each hit energy $e_j$ is taken into account, following Equation 4.7:

$$E_{rec,SC} = \sum_{j=0}^{N_{hits}} \omega\left(e_j, E_{sum}\right) \cdot e_j. \tag{5.8}$$

Both energy reconstructions follow equivalent specifications, by replacing the sum over the thresholds by a sum over all hits in Equation 5.7, $\alpha_i$ is replaced by $\alpha_i/e_j$. In this way the implicit dependence of the semi-digital weights $\alpha_i$ on the individual hit energies is described and can directly be compared to the software compensation weights $\omega$, which is shown in Figure 5.24. The semi-digital weights are represented by the green lines and the software compensation weights are shown as blue points for all beam energies.
The software compensation and semi-digital weights show both a decrease with increasing hit energy. However the software compensation weights show an increase of the weight strength for the last two energy bins for high beam energies. The weights for the first and the last hit energy bin show a strong beam energy dependence. The differences between the software compensation and semi-digital weights are more pronounced for the $3 \times 3\,\mathrm{cm}^2$ AHCAL testbeam data (Fig. 5.24a) than for the $1 \times 1\,\mathrm{cm}^2$ AHCAL simulation with the FTFP_BERT physics list (Fig. 5.24b). This could be an effect of the larger cells sizes in the outer regions of the $3 \times 3\,\mathrm{cm}^2$ AHCAL layers of $6 \times 6\,\mathrm{cm}^2$ and $12 \times 12\,\mathrm{cm}^2$, which is not taken into account in the determination of the hit energy.

Since the digital treatment of hits shows good results for low beam energies (see Figures 5.21 and 5.22), the software compensation algorithm is tested to some extent with counting hits of a certain energy range. Three options have been considered; The hits within the first two hit energy bins ($\sim 1\,\mathrm{MIP}$) are counted and the weights parametrised as 2nd order polynomials of the total number of hits (shown in Figure 5.25 as "SC + 2 digital bins"). The very high energy hits are counted and included in the energy reconstruction digitally (shown as "SC +



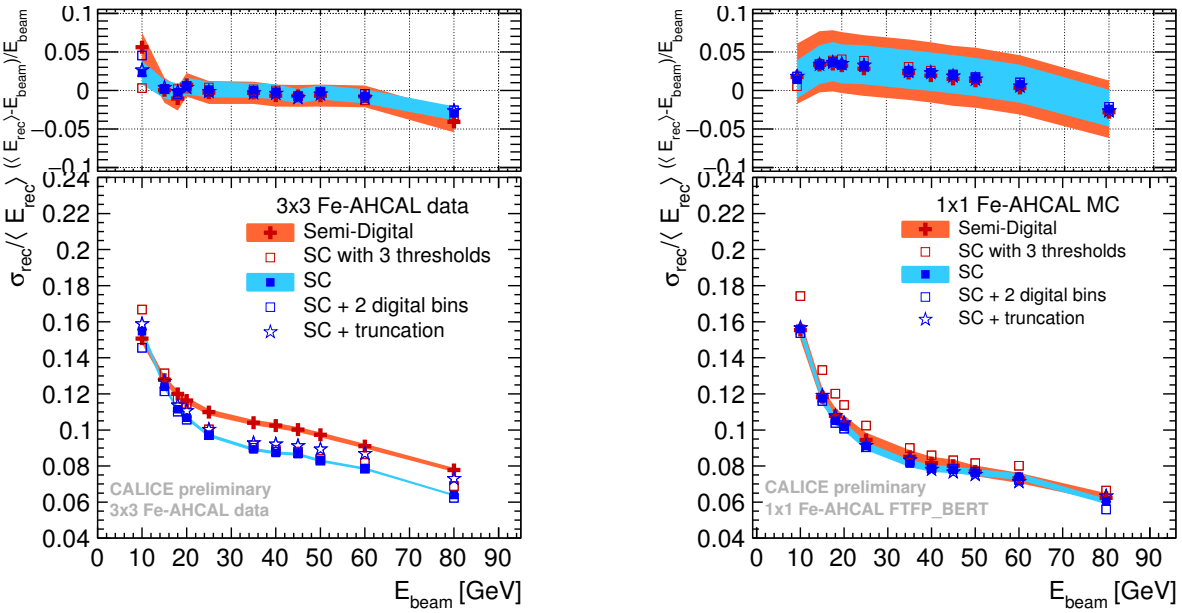

**a)** $3 \times 3\,\mathrm{cm}^2$ **b)** $1 \times 1\,\mathrm{cm}^2$

**Figure 5.24:** The weights used for the software compensation and semi-digital energy reconstruction in the $3 \times 3$ and $1 \times 1\,\mathrm{cm}^2$ AHCAL are shown as a function of the hit energy for the beam energies of 10-80 GeV. The software compensation weights are shown as blue bars covering a certain hit energy range, the semi-digital weights are shown as green lines following a $1/\mathrm{e_j}$ behaviour.

truncation"). This treatment of high energy hits is performed in the PANDORA PFA [23, 94] and has achieved great improvement in the jet energy resolution [95]. Additionally, the software compensation algorithm is tested using the same classification of hits as the semi-digital reconstruction (shown as "SC with 3 thresholds"). For both granularities no strong difference is observed for the different methods of hit treatments within the software compensation algorithm. The linearities achieved are shown in the plots on the top of Figure 5.25.

It has to be mentioned that a positive (negative) deviation from linearity improves (degrades) the relative resolution artificially. However, all observed non-linearities are small enough that the results of the comparison between reconstruction methods of absolute and relative resolutions agree.

A slight improvement is seen in the resolution for 10 and 80 GeV by using two digital energy bins in the 1 MIP range (compare open and filled blue squares). The resolution using the software compensation algorithm with only 3 instead of 8 weights degrades about 0.5 % to 1 % in absolute values. This is observed for both granularities (compare open red squares with filled blue squares). This is most probably an effect of the thresholds optimised for the semi-digital reconstruction and could be further improved by optimising the energy ranges for the software compensation algorithm. The weights for the different methods are shown for the $1 \times 1\,\mathrm{cm}^2$ AHCAL FTFP_BERT simulation in Appendix B.




**a)** $3 \times 3\,\mathrm{cm}^2$ **b)** $1 \times 1\,\mathrm{cm}^2$

**Figure 5.25:** Energy dependence of the relative energy resolution of the AHCAL data and simulation with FTFP_BERT physics list, obtained using different weighting approaches for the energy reconstruction of pions: semi-digital (red cross), with software compensation techniques (blue squares) and variation of the software compensation algorithm (open markers). The plots on the top show the residuals to the beam energy with the bands indicating the systematic and statistical uncertainties. The statistical errors are smaller than the markers.

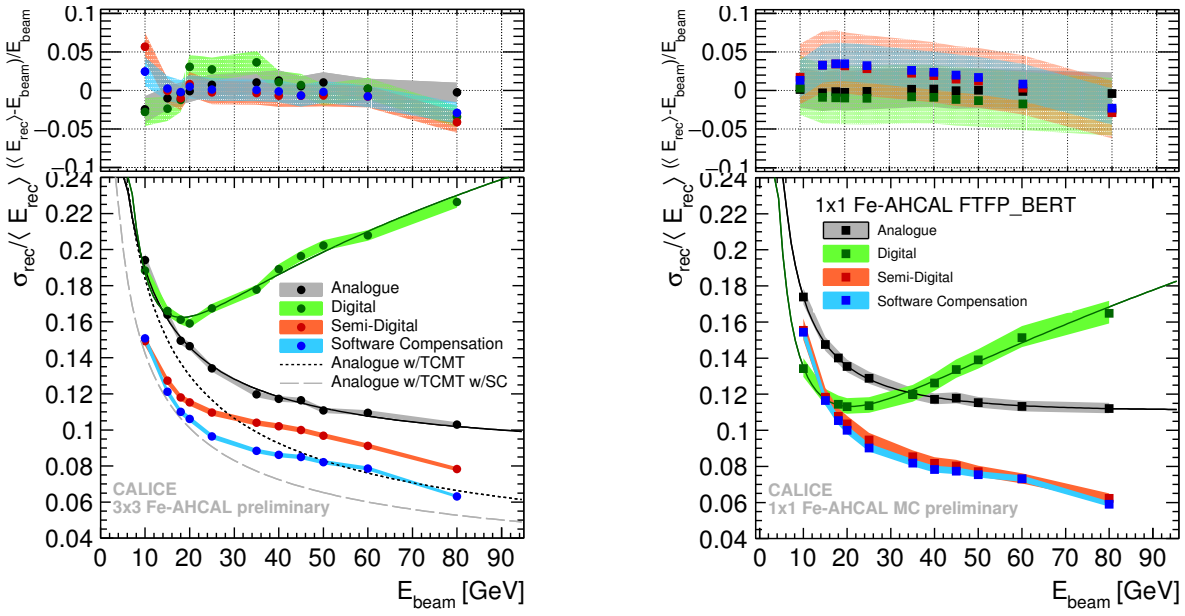
## 5.8 Comparison between Energy Reconstruction Procedures

The resolutions obtained with the different reconstruction methods are compared by applying them on the same data samples. The results obtained for AHCAL data and simulation samples are shown in Figure 5.26. A parametrisation of the best resolution obtained in a previous analysis [71] of the AHCAL data with and without applying software compensation techniques is also shown. In the comparison it is important to keep in mind that in the earlier analysis [71], the TCMT and the ECAL are included in the energy reconstruction. Here a simplified treatment of the ECAL is used and the TCMT contribution is neglected. For energies up to 20 GeV the resolution found for the analogue software compensation algorithm is indeed as good as the resolution reached with the software compensation techniques developed in the past. The difference observed at higher beam energies in Figure 5.26a is caused by energy leakage and is enhanced by the different extraction of the mean and width of the reconstructed energy distributions. This analysis takes tails due to saturation and energy leakage into account by using a Novosibirsk function, while the past analysis only considered the Gaussian peak in the range $\mu \pm 1.5\,\sigma$.

The deviations from linearity of the methods studied in this analysis are also shown in the upper part of Figures 5.26a and 5.26b. For the $3 \times 3\,\mathrm{cm}^2$ granularity, the analogue and digital reconstruction procedures show rather similar resolutions at the lowest energies. For larger energies, the resolution of the analogue reconstruction method continues to decrease, while the digital resolution increases dramatically. The semi-digital reconstruction and the software compensation both apply weights to the energy depositions in a shower depending on the hit energy. The semi-digital reconstruction achieves a resolution similar to the software compensation for the lowest energy, 10 GeV. For higher beam energies the resolution follows a similar shape as for the software compensation but with absolute values 1-2 % worse. The best resolution of all four methods for the whole energy range is found using the analogue software compensation algorithm.

The simulated AHCAL with $1 \times 1\,\mathrm{cm}^2$ cell size is expected to show an improved resolution for the semi-digital and digital readout schemes, which is what is observed in Figure 5.26b. Compared to the classical analogue energy reconstruction the digital reconstruction shows better results for beam energies below 35 GeV. This improvement despite the reduction of information can be explained by the shape of the analogue cell signal, which follows a Landau distribution that is characterised by a long tail to high values. By counting cells above a certain signal amplitude, the signal fluctuations to high values are removed and thus the energy reconstruction is improved, especially for low beam energies where the number of hits is low and these fluctuations play an important role. A degradation due to saturation effects of the digital resolution is only observed above 25 GeV.

The increase of the number of thresholds from 1 to 3, digital to semi-digital, results in a large improvement of the energy resolution of the $1 \times 1\,\mathrm{cm}^2$ AHCAL simulation and in an even larger




**a)** $3 \times 3\,\mathrm{cm}^2$ **b)** $1 \times 1\,\mathrm{cm}^2$

**Figure 5.26:** Energy dependence of the relative energy resolution of the AHCAL testbeam data in (a) and the simulation with $1 \times 1\,\mathrm{cm}^2$ granularity and the FTFP_BERT physics list in (b), obtained using different approaches for the energy reconstruction of pions: analogue (black), digital (green), semi-digital (red) and applying the analogue software compensation algorithm (blue). The dashed and dotted curves in (a) show the resolution achieved in [71] with and without software compensation techniques, using the energy deposits in the TCMT and in the ECAL in addition to the AHCAL. The plots on the top show the residuals to the beam energy with the bands indicating the systematic and statistical uncertainties. The statistical errors are smaller than the markers.

improvement for the $3 \times 3\,\mathrm{cm}^2$ AHCAL. This is different from the observation in the analysis of the SDHCAL data recorded with RPCs, where the resolutions obtained with a digital and the semi-digital reconstruction method are similar up to energies of about 40 GeV, and the semi-digital procedure improves the resolution only for larger energies [65].

The best resolution for the AHCAL with $1 \times 1\mathrm{cm}^2$ cell size is achieved by applying either a weighting by the software compensation algorithm or by the semi-digital energy reconstruction. This result can be understood because both methods apply energy dependent weights, which are determined by a $\chi^2$ minimisation that optimises the resolution. For both methods a decreasing resolution with increasing beam energy is observed. The semi-digital resolution achieves at 80 GeV 6 %, while the software compensation shows a roughly 1 % better resolution in the energy range between 25 and 60 GeV. Otherwise the results are very similar.

## 5.9 Summary

The Fe-AHCAL pion data from 2007 were successfully used to validate the testbeam simulation, to study different energy reconstruction procedures, introduced in Section 4, and to investigate the impact of a granularity of $1 \times 1\,\mathrm{cm}^2$ on the Fe-AHCAL performance.
To validate the $1 \times 1\,\mathrm{cm}^2$ AHCAL simulation all relevant digitisation effects, like noise, crosstalk, and thresholds have been studied. These studies revealed an impact of the detector's noise on the number of hits and the measured energies. This impact depends on the particle energy and the expansion of the shower. The threshold of 0.5 MIP for hits in the $3 \times 3\,\mathrm{cm}^2$ AHCAL is found to be efficient to suppress the noise. The analogue response of the simulated $1 \times 1\,\mathrm{cm}^2$ AHCAL, by applying a threshold of 0.3 MIP, is found to be equivalent to the analogue response in the $3 \times 3\,\mathrm{cm}^2$ AHCAL.
The energy reconstruction methods used for the S- and DHCAL of the CALICE collaboration were tested with the Fe-AHCAL testbeam data and show significant differences in the energy resolution, which is summarised in Section 5.8. The reconstruction procedures using an energy dependent weighting: semi-digital and analogue software compensation achieved very good energy resolutions close to $45\,\%/\sqrt{\mathrm{E}}$, and revealed strong similarities in the weighting parameters, see Section 5.7.

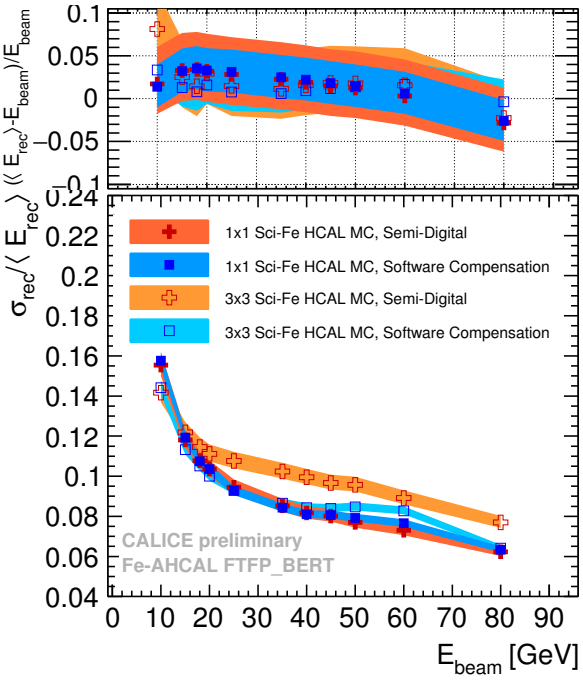
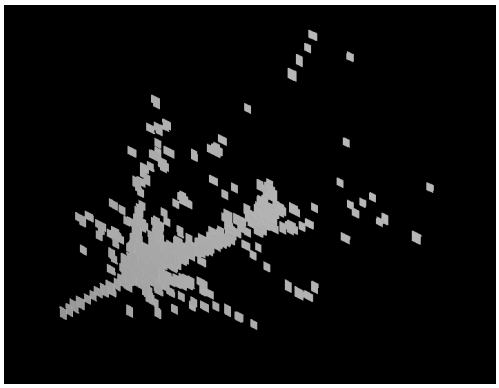



**Figure 5.27:** The energy resolution of the semi-digital and analogue software compensation energy reconstruction procedures of the FTFP_BERT simulations are shown as a function of beam energy for cell sizes of $1 \times 1$ and $3 \times 3\,\mathrm{cm}^2$. The plot on the top shows the residuals to the beam energy with the bands indicating the systematic and statistical uncertainties. The statistical errors are smaller than the markers.

The main simulation-based results of the energy reconstruction and granularity studies are summarised in Figure 5.27, which shows the energy resolutions for the semi-digital and the analogue software compensation reconstruction for different Fe-AHCAL cell sizes. The FTFP_BERT physics list, which showed the best agreement with the data, is used for the comparison. The best possible energy resolution is achieved by the $1 \times 1\,\mathrm{cm}^2$ AHCAL simulation using the semi-digital energy reconstruction, while the analogue software compensation algorithm has to be used for the energy reconstruction in the $3 \times 3\,\mathrm{cm}^2$ AHCAL. Additionally, the resolution achieved by the Fe-AHCAL is similar when using software compensation algorithms for both the studied granularities.



# Chapter 6

# The Digital HCAL at Fermilab

The data sample of the Fe-DHCAL with steel absorber was recorded in 2011 at Fermilab. The testbeam setup consisted of 38 layers of the RPC-steel Digital HCAL, up to 14 layers of RPC-steel Tail Catcher and Muon Tracker (TCMT), a beam Cherenkov threshold counter, and a muon tagger. The Cherenkov threshold counter has been tuned to be responsive to electrons and non-responsive to heavy particles like pions. The Fe-DHCAL has a total thickness of $5.33\,\lambda_n$ and is described in detail in Section 2.3. The TCMT placed downstream of the Fe-DHCAL is used to absorb the tails of the showers leaking out of the DHCAL. Up to 14 active RPC layers were inserted into the steel absorber structure of the TCMT with a total thickness of $5.8\,\lambda_n$. More details are given in Chapter 2.6. The muon tagger consisted of two $1 \times 1\,m^2$ scintillator plates placed 4 meters in front of the Fe-DHCAL and behind the active TCMT layers in the TCMT structure [61]. Both scintillator plates have also been used to trigger the data acquisition.
The analysis of electromagnetic and hadronic showers focusses on the Fe-DHCAL. Since the TCMT has not been always fully equipped with active layers, including it in the analysis would results in inconsistencies due to the variation in the number of active layers. Additionally, omitting the TCMT ensures the comparability with the analysis of the AHCAL data.

## 6.1 Data and Event Selection

The Fermilab testbeam facility provides a mixed beam of muons, positrons and pions, with different compositions depending on the beam energy. The beam energies of the recorded runs range from 2 to 60 GeV. The positron content is negligible for beam energies larger than 32 GeV, while positrons dominate for beam energies smaller than 6 GeV. For the 2, 4, 25 and 32 GeV runs the data from the Cherenkov counter is not available, thus a particle identification based on the shower topology is needed. This is possible due to the imaging capability of the DHCAL.
Before the Particle IDentification (PID) and event selection, a clustering algorithm is applied to the DHCAL data: The hits in each layer are combined into clusters using a nearest-neighbour

algorithm. If two hits share a common edge, they are assigned to the same cluster [96].
Additionally, the first hard interaction and thus the Interaction Layer (IL) is identified using an algorithm that first computes the running Three-Layer Hit Average (TLHA) iteratively through all layers [97]. If over 3 consecutive TLHAs an increase of a factor 2 or larger is observed and the current TLHA is $\geq 4$, the interaction layer is identified as the layer before the current layer. This algorithm shows an improved performance compared to an older algorithm [97, 98].
The Fermilab testbeam not only provided different particle types, but a significant fraction of events recorded more than one particle entering and interacting in the detector. This phenomenon is called multi-particle contamination. In order to remove these events and showers initiating upstream the DHCAL from the data set the following cut is applied:

- requiring exactly one cluster with less than 4 hits in the first layer

The average fraction of excluded events per beam energy is 36 %. A detailed list of the fraction of rejected events for each beam energy can be found in Table 6.1. The impact on the distribution of the number of hits is visualised in Figure 6.1.
Muons are selected by requiring and applying:

- no identified IL (i.e. no shower start).
- at least 20 hits in the HCAL, in order to reject events triggered without a particle traversing the detector
- topological selection:
  - 0.5 < average number of hits per layer < 2.5
  - centre of gravity in beam direction > layer 15
  - every layer has less than 4 clusters

For beam energies between 16 to 32 GeV the muon tagger was used, whereas for larger energies the muon tagger was not efficient anymore, because of the higher probability for punch-through pions.
Positrons are pre-selected via the Cherenkov counters. In addition the final selection requires:

- an identified IL within the first five layers.
- not an identified muon
- topological selection:
  - centre of gravity in beam direction < layer 12
  - electromagnetic showers are narrow < 5 cm in radius, using the RMS of the hit positions per layer
    (Molière radius in steel is 1.8 cm)

**Table 6.1:** Fractions of the multi-particle, muon, positron and pion events in the data set. The positron and pion selections are divided in pre- and final selection, with the pre-selection (using the Cherenkov counters) applied after the removal of multi-particle events. For the beam energies 2, 4, 25 and 32 GeV the Cherenkov information was not available, therefore the pre-selection is required 100 % positrons for beam energies below 6 GeV and 100 % pions for beam energies above 25 GeV. At 25 GeV no pre-selection is applied, which is expressed by the 100 % positions and pions in the 10th row, 4th and 5th column.

| energy [GeV] | multi particle events [%] | muons [%] | pre-selected $e^+$ [%] | pre-selected $\pi^+$ [%] | $e^+$ [% of pre-selected events] | $\pi^+$ [% of pre-selected events] |
|---|---|---|---|---|---|---|
| 2 | 46.9 | 10.1 | 100 | 0 | 57.2 | - |
| 4 | 43.4 | 9.0 | 100 | 0 | 71.9 | - |
| 6 | 42.2 | 5.6 | 58.9 | 41.1 | 92.2 | 47.4 |
| 8 | 34.9 | 14.0 | 38.0 | 62.0 | 96.5 | 41.4 |
| 10 | 33.5 | 11.3 | 32.0 | 38.0 | 96.3 | 45.5 |
| 12 | 31.5 | 17.6 | 18.7 | 81.3 | 96.3 | 44.0 |
| 16 | 29.8 | 19.9 | 11.3 | 88.7 | 93.0 | 43.2 |
| 20 | 29.8 | 17.3 | 6.1 | 93.9 | 91.6 | 46.9 |
| 25 | 30.4 | 13.9 | 100 | 100 | 0.14 | 24.0 |
| 32 | 31.2 | 11.1 | 0 | 100 | - | 41.0 |
| 40 | 35.4 | 4.1 | 0 | 50.6 | - | 63.7 |
| 50 | 40.6 | 2.5 | 0 | 68.8 | - | 64.2 |
| 60 | 48.2 | 1.8 | 0 | 88.2 | - | 63.2 |

Pions are also pre-selected using the Cherenkov signal. Afterwards the final selection requires:

- an identified IL within the first ten layers.
- not an identified muon or positron event

The requirement on the IL ensures a good shower containment of the pions within the DHCAL. The PID based on the event topology is needed due to the missing Cherenkov signals for the 2, 4, 25 and 32 GeV runs, but is applied to the whole data set for all beam energies to correct for the inefficiency of the Cherenkov counters of up to 10 % [96]. The selected runs and the events are summarised in Table 6.2. The impact of the individual cuts is summarised in Table 6.1. The misidentification probability of the pion and positron selection has been determined using Monte Carlo simulation (MC) and found to be always less than 5 %. The cut on the interaction layer has the strongest impact on the pion statistic.

### 6.1.1 Data Preparation

During the operation of the Fe-DHCAL prototype, the temperature, gas pressure and gas flow varied. For yet unknown reasons, some chambers developed a signficant loss of efficiency. The high voltage was adjusted during testbeam to mitigate this problem. These variations caused a different response from different RPCs. To correct for the resulting differences in hit multiplicity and the efficiency of the RPCs, the Fe-DHCAL was calibrated following the approach

**Table 6.2:** List of data runs used in the analysis and sample statistics.

| run number | beam energy [GeV] | selected muons | | selected positrons | | selected pions | |
|---|---|---|---|---|---|---|---|
| | | $N_{events}$ | in % | $N_{events}$ | in % | $N_{events}$ | in % |
| 600139, 600140, 600143, 600145, 600147, 600148, 600149, 600150, 600152, 600153, 600154, 600155, 600156, 600157, 600158, 600159, 600160, 600161, 600162, 600163, 600164, 600165, 600166 | 2 | 18,041 | 5.4 | 98,986 | 29.5 | - | - |
| 600089, 600091, 600092, 600176, 600177, 600178, 600179, 600180, 600181, 600185, 600186 | 4 | 7,909 | 5.1 | 61,842 | 39.8 | - | - |
| 600187, 600193, 600194, 600195, 600196 | 6 | 2,176 | 3.3 | 20,112 | 30.0 | 7,553 | 11.3 |
| 600082, 600083, 600084, 600197, 600198, 600202, 600203, 600204 | 8 | 9,024 | 9.1 | 22,736 | 23.0 | 16,537 | 16.7 |
| 600205, 600206, 600207, 600208, 600209, 600210, 600211, 600212 | 10 | 8,104 | 7.5 | 21,367 | 19.8 | 22,201 | 20.6 |
| 600075, 600076, 600077, 600079, 600080 | 12 | 6,787 | 12.0 | 6,711 | 11.9 | 13,849 | 24.5 |
| 600063, 600064, 600065, 600069, 600070 | 16 | 5,563 | 13.4 | 2,937 | 7.1 | 11,185 | 26.9 |
| 600054, 600055, 600058, 600059, 600062 | 20 | 5,747 | 12.2 | 1,793 | 3.8 | 14,642 | 30.9 |
| 600049, 600050, 600052, 600053 | 25 | 2,915 | 9.7 | 578 | 2.0 | 5,474 | 18.2 |
| 600036, 600038, 600040, 600043, 600044, 600045, 600048 | 32 | 4,096 | 7.7 | - | - | 15,337 | 28.7 |
| 630125, 630126, 630128, 630129, 630130, 630131, 630133, 630134 | 40 | 993 | 2.7 | - | - | 7,745 | 20.8 |
| 630137, 630139, 630141, 630142, 630144, 630145 | 50 | 458 | 1.5 | - | - | 8,108 | 26.3 |
| 630146, 630147, 630148, 630152, 630153, 630154 | 60 | 413 | 0.9 | - | - | 12,681 | 28.9 |

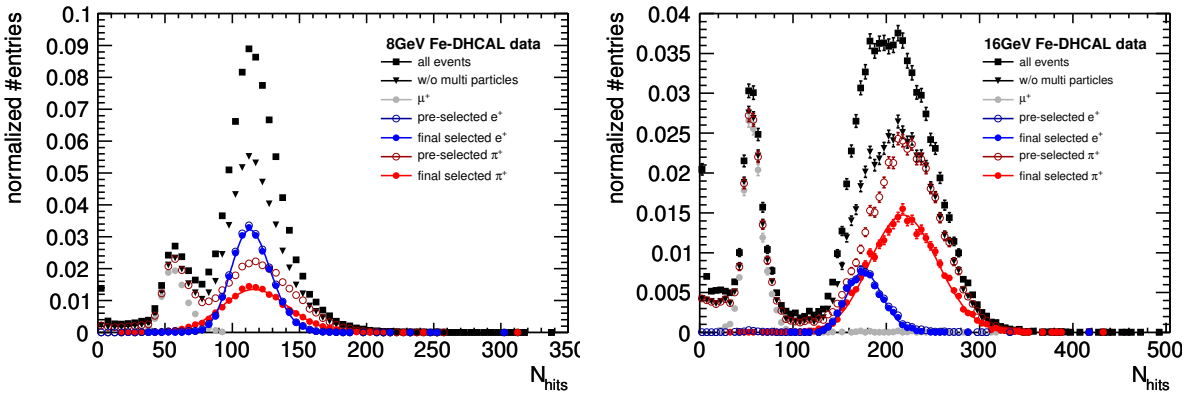



**Figure 6.1:** The distribution of the total number of hits $N_{hits}$ for beam energies of 8 and 16 GeV, shown after various selection cuts. The triangles show the events left after the removal of multiple particle contamination. The grey dots show the selected muons and the open dots mark the pre-selection for positrons (blue) and pions (red), which is based on the Cherenkov PID. The filled dots show the final selected positrons (blue) and pions (red).

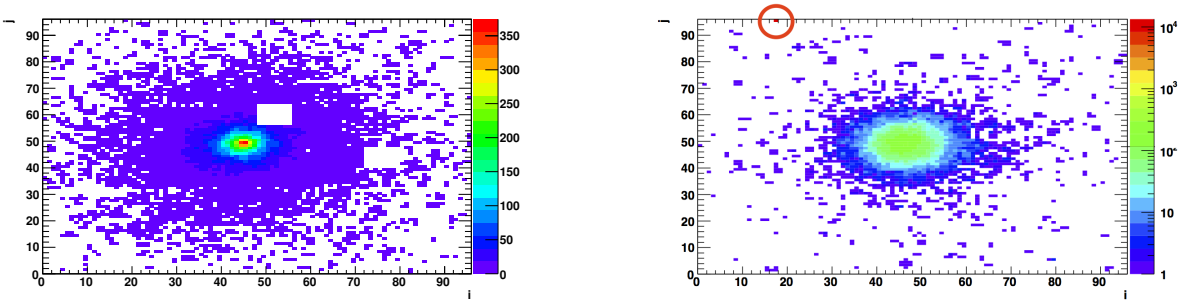
introduced in Section 2.3.3. The detailed calibration procedure and the determination of the according systematic errors are described in Section 6.1.2 and 6.1.3.

Additionally to the variations of the RPC response, hardware parts failed; malfunctioning ASICs, dead front end boards and hot cells occurred. For consistency these channels have been identified for every run and have then been ignored in the further analysis for all runs at the same energy. How these effects manifest in the recorded hit maps is shown in Figure 6.2. 75 non-responding front end boards, with 1,536 channels per board, have been identified in 38 layers of all 101 runs, which corresponds to 0.3 % front end boards averaged over the entire run period. The number of dead ASICs, with 64 channels per ASIC, is 522 in the main stack, this is less than 1 %. The number of removed hot cells is 9. On average 1,472 channels per run are excluded from the analysis, which corresponds to 0.4 % of the channels in all runs.
To ensure that effects in simulation are decoupled from failure in the hardware description, all channels removed in data are ignored in simulation as well.

### 6.1.2 Calibration

The calibration of the Fe-DHCAL data is done using the same data set that is used in the rest of the analysis. The local efficiency ε and multiplicity μ of each RPC is determined, using single particle tracks within the showers in the data runs themselves. In order to define the tracks a cluster-algorithm is used. Hereby hits that share a border are grouped together in one cluster. The calculation of the local ε and μ of each RPC per run is done in 5 steps, repeated for each layer:

1. the algorithm finds clusters of 3 or less hits in the surrounding ±3 layers



**a)** 32 GeV Run600040, layer 36 **b)** 4 GeV Run600186, layer 10

**Figure 6.2:** Hit maps in cell coordinates i-j of 2 layers of 2 different runs, a) featuring 2 dead ASICs, and b) showing a hot cell (highlighted by the red circle).

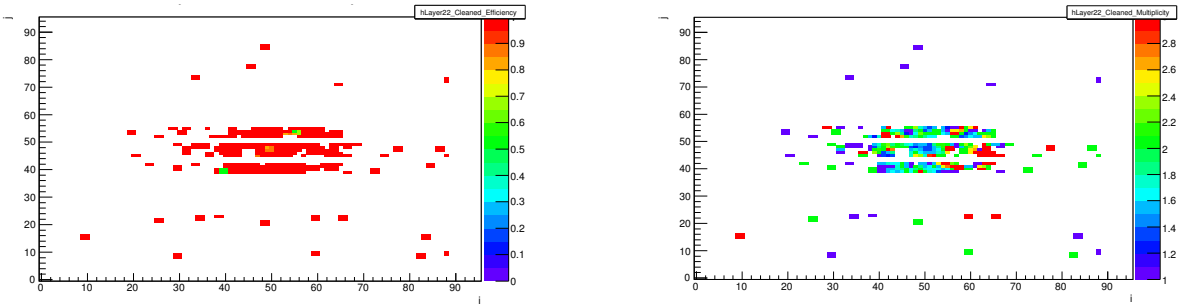
2. if 4 or more clusters that fulfil condition 1 in the layer of interest ±3 layers are found, a straight line fit is applied excluding the layer of interest

3. the intersection point of the fit with the layer of interest is determined, and the algorithm searches for a cluster within 2 cm

4. if a cluster is found (not found) a histogram for the efficiency of that RPC gains an entry of 1 (0) and a histogram for the multiplicity is filled with the cluster size

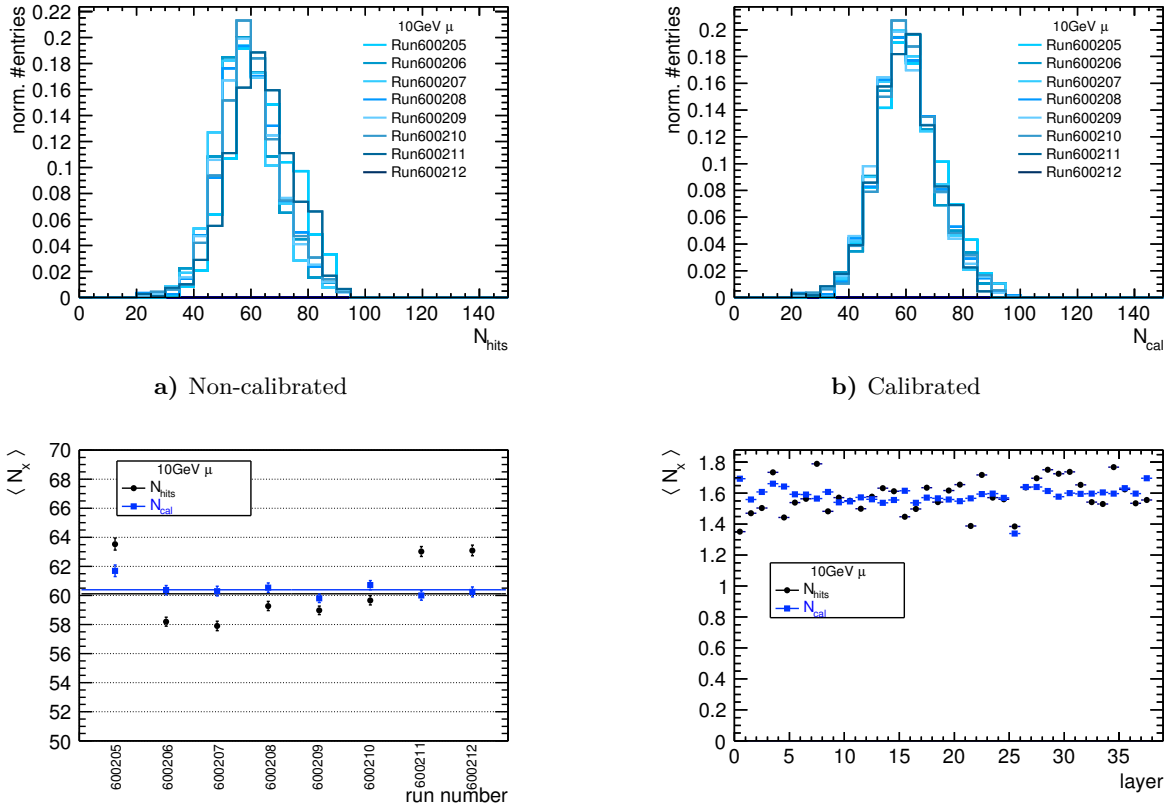
5. the efficiency is given by the fraction of events with a found cluster and the multiplicity is determined by the mean cluster size for events where a cluster is found.

Unfortunately, the beam spread is not always large enough to ensure sufficient statistics in the bottom and top RPCs. Here, sufficient means at least 500 successful track fits per RPC. If this is not obtained, the calibration values determined for the RPC in the middle position are assigned to the outer chambers, which is a good approximation because both the gas flow and the applied high voltage are the same. An example of the extracted efficiencies and multiplicities is shown as a 2D scan of channel-wise efficiencies and multiplicities in Figure 6.3. The calibration coefficients for each channel are calculated using Equation 2.5.
To prevent the calibration procedure from correcting local inefficiencies due to fishing lines and at the borders of the RPCs, these regions are excluded from the calculations. This is important because the fishing lines and the corresponding insensitive areas are already considered in the GEANT4 simulation. The reduced efficiencies at the RPC borders are taken into account at the digitisation level. Following Equation 2.6, the RPC responses become equalised after calibration. This equalisation is demonstrated in Figure 6.4 with the total number of hits for 10 GeV muon runs before a) and after b) calibration. Figure 6.4c shows the mean difference in the total number of hits $N_{hits}$ and the sum of all calibrated hits $N_{cal}$ per event. The reduced layer-to-layer fluctuations after calibration are illustrated by the mean number of hits per layer for all recorded 10 GeV muons in Figure 6.4d.




**a)** Efficiencies **b)** Multiplicities

**Figure 6.3:** Efficiency and multiplicity measurements in layer 22 of Run600185, restricted to the "clean" regions, excluding channels that are affected by the fishing lines and RPC borders. The z-axis show the values determined for each channel individually, for the calibration the values are taken from the whole RPC.




**a)** Non-calibrated **b)** Calibrated

**c)** $\langle N_{hits} \rangle$ and $\langle N_{cal} \rangle$ for 10 GeV muon runs **d)** ($\langle N_{hits} \rangle$ and $\langle N_{cal} \rangle$)/layer for 10 GeV muons

**Figure 6.4:** Hit distributions of muons for all 10 GeV runs, before a) and after b) the calibration is applied. c) The mean number of hits for 10 GeV muons before (black) and after (blue) calibration. d) The average number of hits per layer before (black) and after (blue) calibration for all 10 GeV muon runs.

### 6.1.3 Systematic Uncertainties

The main source of systematic uncertainties on the calibrated number of hits is the error on the calibration factors. One way to determine the uncertainty on the calibration factors is to propagate the statistical uncertainty from the determination of the local efficiency and multiplicity per RPC.
The standard error of the mean cluster size $\sigma_{\mu}^{\pm}$ is symmetric, while the error on the efficiency $\sigma_{\varepsilon}^{\pm}$ is asymmetric and calculated following the binomial error estimation for a $2\sigma$ confidence level. The efficiencies and multiplicities as well as the calibration factors are determined for both cases of $\left(\varepsilon_{i,j} + \sigma^{+}_{\varepsilon,i,j}\right)$, $\left(\mu_{i,j} + \sigma^{+}_{\mu,i,j}\right)$ and $\left(\varepsilon_{i,j} - \sigma^{-}_{\varepsilon,i,j}\right)$, $\left(\mu_{i,j} - \sigma^{-}_{\mu,i,j}\right)$:

$$c^{\pm}_{i,j} = \frac{\varepsilon_0 \cdot \mu_0}{\left(\varepsilon_{i,j} \pm \sigma^{\pm}_{\varepsilon,i,j}\right) \cdot \left(\mu_{i,j} \pm \sigma^{\pm}_{\mu,i,j}\right)}, \tag{6.1}$$

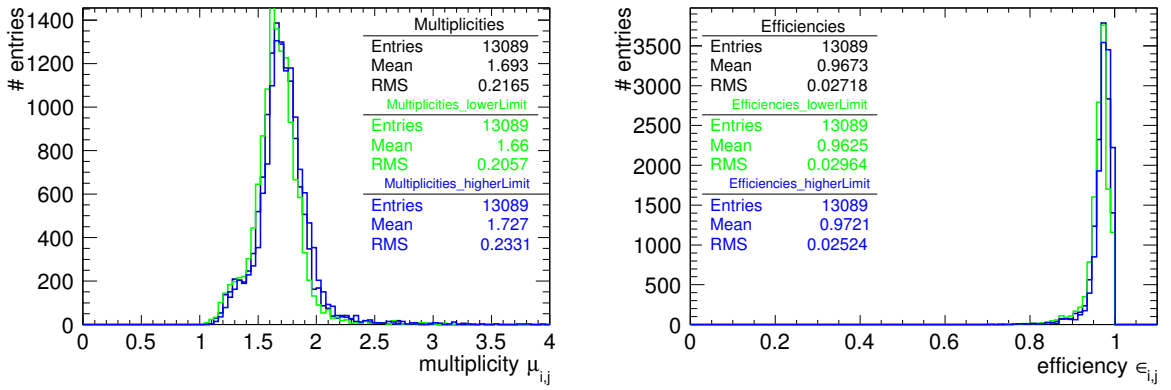
with the average efficiency $\varepsilon_0 = 0.97$ and multiplicity $\mu_0 = 1.69$. The results are shown for all RPCs of all runs in the blue and green histograms, compared to the nominal determinations in black histograms, in Figures 6.5a, b and c. The difference in the resulting average calibration factors $\left\langle c_{i,j} \right\rangle$ give an estimate of the systematic uncertainty on the RPC response equalisation procedure, see the statistics box in Figure 6.5c. The average uncertainty on the calibration factors is found to be +2.6 and -2.4 %. These errors are assigned to the data as systematic uncertainty on the calibration, and are added in quadrature to the statistical uncertainties.
This method has by design a tendency to overestimate the uncertainty and is thus conservative. The results obtained with this method have been compared with the method used in [99] and found to be consistent. The shoulder in the multiplicity distribution (Figure 6.5a) originates from the TCMT-RPCs, which show differences because of missing RPCs and dead front-end boards. This reduces the overall multiplicity and efficiency.

## 6.2 Digital HCAL Simulation

### 6.2.1 Geant4 and Mokka

The simulation of the Fe-DHCAL testbeam setup is done using the software packages GEANT4 version 10.01, MOKKA[1] v08-05 and ILCSOFT[2] v01-17-08. Each testbeam run is simulated with 10,000 events for muons $\mu^+$, positrons $e^+$, and pions $\pi^+$. The additional beam line instrumentation is neglected in the simulation and the particles are generated directly in front of the HCAL prototype. The material description of the Fe-DHCAL is included in the simulation following the description given in Section 2.3. A schematic of the cross section of one active plus one absorber layer is shown in Figure 6.6. The reference system used in the simulation

[1]Software package, developed by Laboratoire Leprince-Ringuet (LLR), École polytechnique, Palaiseau (France) `http://ilcsoft.desy.de/portal/software_packages/mokka/`
[2]Software packages for the International Linear Collider (ILC)
`http://ilcsoft.desy.de/portal`




**a)** Multiplicities

**b)** Efficiencies

**c)** Calibration factors

**d)** Residuals

**Figure 6.5:** Extracted local hit multiplicities $\mu_{i,j}$ in a), efficiencies $\varepsilon_{i,j}$ in b), resulting calibration factors $c_{i,j}$ c) and the residuals of $\mu_{i,j}\varepsilon_{i,j}$ in d). The standard determination is shown in black histograms and the smallest and highest values are shown in green and blue histograms. The statistics boxes show the total number of entries, mean and RMS of the histograms. The average uncertainty on the extracted value of $c_{i,j}$ are found to be +2.6 and –2.4 %.

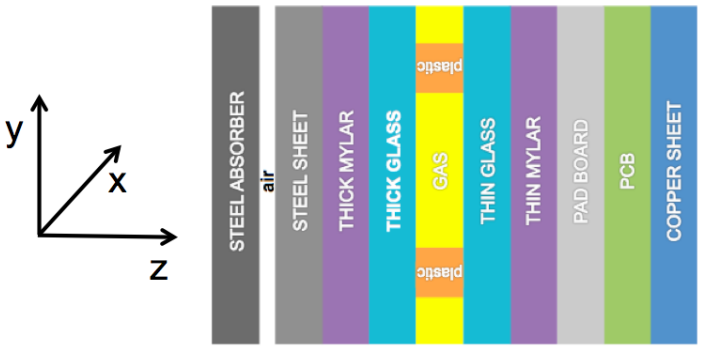
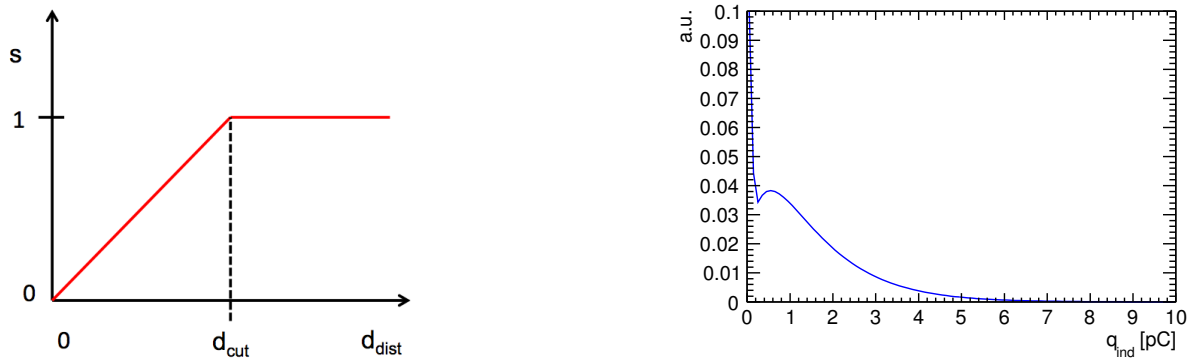



**Figure 6.6:** Material discription of one active RPC layer and absorber plate in the MOKKA geometry of the Fe-DHCAL, shown in the right-handed coordinate system. Dimensions are not to scale.

is also shown. The system is right-handed with the z-axis pointing in the beam direction, the x-axis being horizontally and the y-axis vertically orientated.

### 6.2.2 Simulation of the RPC Response – Digitiser

The digitiser simulates the response of the RPCs. The simulated ionising energy depositions in the gas of the RPCs from GEANT4 are used as input for the DHCAL digitiser.
Within the gas gap, the probability of an electron to gain enough energy to generate a Townsend avalanche decreases in case of an avalanche already developing close by due to the drop in the electric field strength. This limitation in spatial response of the RPCs is simulated by introducing a scaling factor s that is assigned to one energy deposit if it is too close to another deposit and later in time. The timing information of the energy deposits is given by GEANT4. To identify the affected energy deposits, the first step is to calculate the distances $d_{dist}$ between all energy deposits in the same layer. If two deposits are closer than a distance $d_{cut}$, the scaling factor s between 0 and 1, following a linear behaviour with $d_{dist}$, is assigned to the second energy deposit. A schematic of the scaling factor s as a function of the distance $d_{dist}$ is shown in Figure 6.7a.
In the next step the digitiser assigns a charge to each deposit randomly, which gets scaled by s independent of the total energy deposition within the gas gap. The generated charge follows the measured RPC charge spectrum [100], shown in Figure 6.7b. This spectrum was recorded in a muon beam at Fermilab with one of the RPCs, that has been used for collecting the present data set. This RPC had an analogue readout, following [51], and was operated in similar conditions as in the 2007 testbeam period.
The shape of the charge distribution does not follow a Polya function (see Equation 2.2) since the measured charge strongly depends on the distance of the primary ionisation from the readout anode, which defines the induced signal height. The closer a deposit is to the anode the smaller is the probability to generate a Townsend avalanche, the shorter is the path length of an induced avalanche, and the smaller is the induced signal on the pad plane. This effect is




**a)** $d_{cut}$ **b)** $q_{ind}$

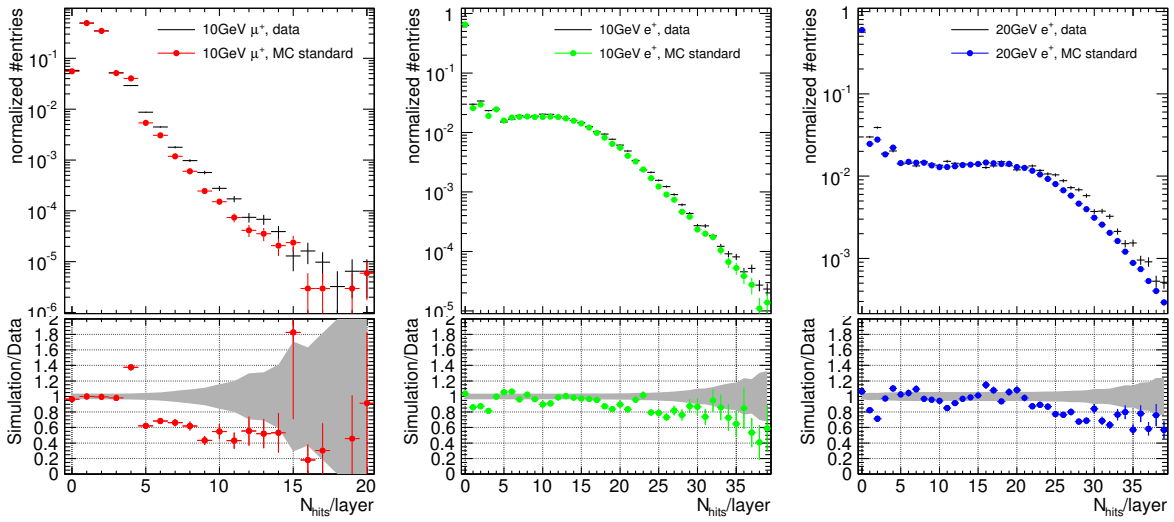
**Figure 6.7:** a) Schematic of the dependence of the scaling factor s on the distance between GEANT4 energy deposits $d_{dist}$. $d_{cut}$ marks the transition of deposits that get assigned a lowered charge by s. b) The analogue measured charge distribution of muons [100].

seen in the steep increase of charges $< 0.2\,\text{pC}$.
Due to the uncertainty on the running conditions a additional factor $q_0$ is introduced to scale the induced charge extracted from the spectrum (in Figure 6.7b).
After a charge is assigned to an energy deposit, this charge is spread on the anode plane following a double Gaussian function depending on the radius r (up to $r = 4\,\text{cm}$) in the x-y plane:

$$f(r) = (1 - R) \cdot \exp\left(-\frac{r^2}{2\sigma_1^2}\right) + R \cdot \exp\left(-\frac{r^2}{2\sigma_2^2}\right), \tag{6.2}$$

with the ratio parameter R and the widths of the Gaussians $\sigma_1$ and $\sigma_2$, that determine the charge spread over the pads. After all charges of all energy deposits are distributed over the readout pads, the charges on each pad are summed up and a threshold T is applied.

The 6 free digitisation parameters ($d_{cut}$, $q_0$, R, $\sigma_1$, $\sigma_2$ and T) are highly correlated and have to be determined from data. The tuning of these parameters is done, matching the simulated number of hits per layer "$N_{hits}$/layer" of muons and positrons to the measured distributions, see Figure 6.8. The parameter space has been explored by assigning to each parameter a value within a reasonable range and testing all possible combinations. The agreement between the data and the simulation is determined for each parameter combination using the $\chi^2$ values between the histograms as a measure of agreement.
This procedure is repeated for three different versions of electromagnetic physics lists of GEANT4 [101]; the "standard", the "option 3" or _EMY, and "option 4" or _EMZ physics lists. These options vary in accuracy, and most important for this analysis, in the step length for which the next ionising energy deposition is calculated for. Since the deposited energies themselves are not taken into account but for each deposition point a charge is assigned and an avalanche is generated, the number of original deposits has a great effect on the generated total




**a)** 10 GeV $\mu^+$ **b)** 10 GeV $e^+$ **c)** 20 GeV $e^+$

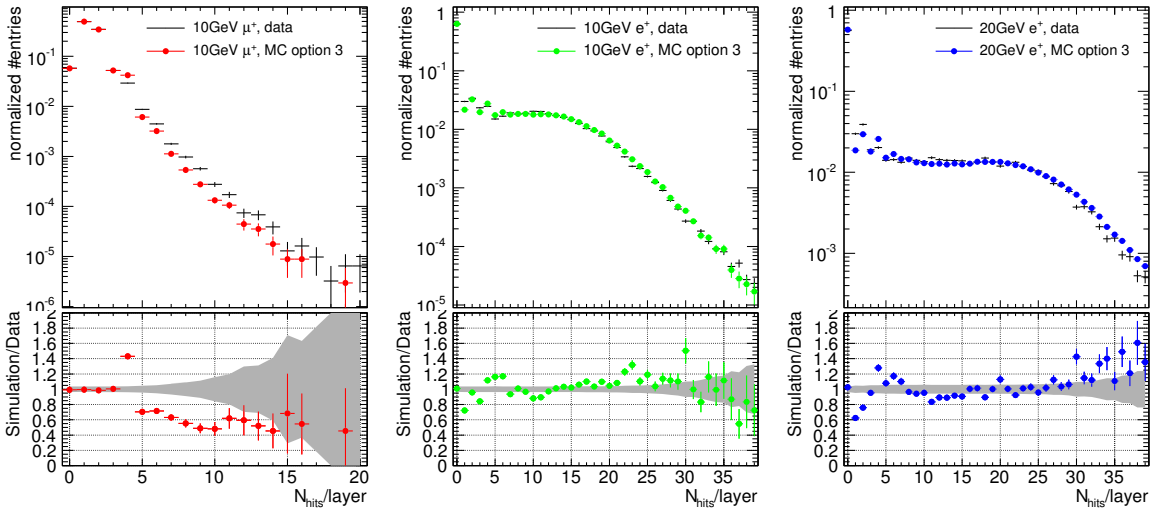
**Figure 6.8:** The top plots show the number of hits per layer $N_{hits}$/layer spectra for a) 10 GeV muons b) 10 GeV positrons and c) 20 GeV positrons. The data is represented in black and the simulation with the standard EM physics list in red, green and blue. The bottom plots show the deviations of the simulation from data, where the grey bands indicate the statistical and systematic uncertainties of the data added in quadrature.

**Table 6.3:** The digitisation parameters for the three EM physics lists of GEANT4 determined from the tuning process. The total $\chi^2$/ndf describes the difference between the data and simulation in the distributions shown in Figure 6.8, 6.10 and 6.9.

| EM physics list | $\sigma_1$ [mm] | $\sigma_2$ [mm] | R | T [pC] | $q_0$ | $d_{cut}$ [mm] | $\left(\chi^2/\text{ndf}\right)_{tot}$ |
|---|---|---|---|---|---|---|---|
| standard | 0.7 | 5.0 | 0.08 | 0.07 | 1.0 | 0.05 | 23.45 |
| option 3 (_EMY) | 0.7 | 4.0 | 0.05 | 0.08 | 1.2 | 0.01 | 23.35 |
| option 4 (_EMZ) | 0.7 | 5.0 | 0.08 | 0.07 | 1.0 | 0.05 | 16.89 |

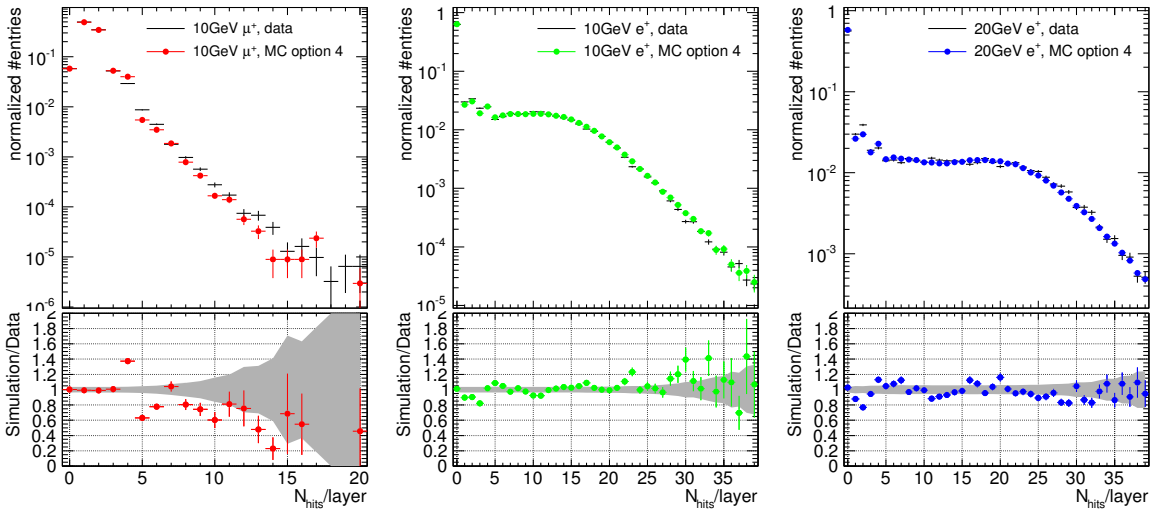
number of hits. From [82] the recommended EM physics list for gases detectors is "option 3". The $N_{hits}$/layer distributions are shown in Figure 6.8 for the simulation with the standard EM physics list. The tuning parameter for all investigated EM physics lists are summarised in Table 6.3. Additionally to the digitisation parameters the sum of $\chi^2$/ndf values from the comparison to the data are given in the table. These values reveal the best agreement of the _EMZ physics list with data. However, the $\chi^2$/ndf values are still quite large, which can be explained by a remaining inaccuracy of the GEANT4 modelling and a limited accuracy of the parameters of the digitiser.

The $N_{hits}$/layer distributions for the _EMY and _EMZ physics lists are shown in Figure 6.9 and 6.10. The _EMZ physics list reveals a better description of the data for high $N_{hits}$/layer compared to the _EMY option and the standard EM physics list.




**a)** 10 GeV $\mu^+$ **b)** 10 GeV $e^+$ **c)** 20 GeV $e^+$

**Figure 6.9:** The top plots show the number of hits per layer spectra for a) 10 GeV muons b) 10 GeV positrons and c) 20 GeV positrons for data in black and coloured for the simulation with EM physics list option 3 (_EMY). The bottom plots show the deviations of the simulation from data, where the grey bands indicate the statistical uncertainty of the data points.




**a)** 10 GeV $\mu^+$ **b)** 10 GeV $e^+$ **c)** 20 GeV $e^+$

**Figure 6.10:** The top plots show the number of hits per layer spectra for a) 10 GeV muons b) 10 GeV positrons and c) 20 GeV positrons for data in black and coloured for the simulation with EM physics list option 4 (_EMZ). The bottom plots show the deviations of the simulation from data, where the grey bands indicate the statistical and systematic uncertainties of the data added in quadrature.

## 6.3 Analysis of Positron Showers

The digitiser described in the previous section is used to simulate the RPC response. The tuning of the parameters of the digitiser is based on the number of hits per layer for 10 GeV muons and 10 and 20 GeV positrons, thus the predictive power of the simulation is limited for positrons. In the following the positron showers are studied for positron energies in the range of 2 to 25 GeV and the data are compared to the simulation with different EM physics lists, introduced in Section 3.2.

### 6.3.1 Positron Showers

The longitudinal and lateral shape and the hit densities of EM showers are measured. In Figure 6.11 these observables are shown for 12 GeV positrons and are compared to the three simulations with different EM physics lists of GEANT4.
The 2D hit density is determined for each hit by counting the number of neighbouring hits in the same layer within a radius of $\sqrt{2}$ pads, see Figure 6.11a. The 3D hit density additionally includes hits in the same x-y positions within $\pm 1$ layer, see Figure 6.11b.
The longitudinal profile is defined as the average number of hits per layer with respect to the shower start by showing the x-axis as "layerNumber-interactionLayer" in Figure 6.11c. The radial shower shape is defined as the standard deviation of the hit positions in the cylindrical coordinate $r = \sqrt{x^2 + y^2}$ per event and layer, by

$$\langle R \rangle = \sqrt{\frac{\left(\sum_i c_i \cdot r_i - \langle r_c \rangle\right)^2}{\sum_i c_i}}, \tag{6.3}$$

with the sums running over all hits per layer $N_{layer}$, c is the assigned calibration factor of the hit, and

$$\langle r_c \rangle = \frac{1}{N_{layer}} \sum_i^{N_{layer}} c_i \cdot r_i \tag{6.4}$$

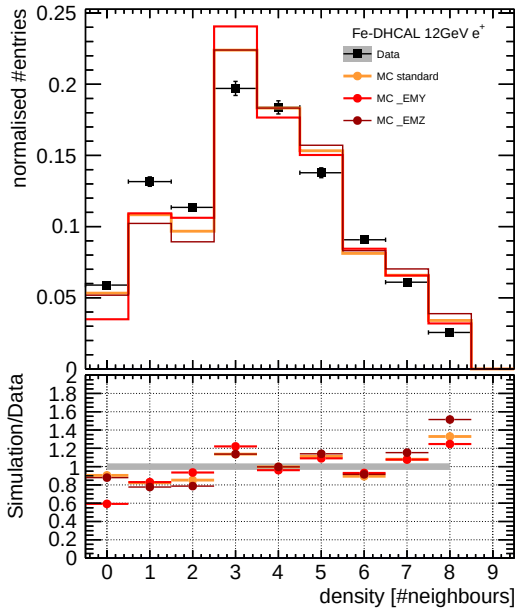
is the mean weighted hit position. The radial shower shape is shown in Figure 6.11d.
The comparison of the data and simulation in the hit densities reveals a good agreement for the standard and _EMZ simulations, and a slightly worse agreement for the _EMY option. However, all simulations show too many entries for hits with 3 neighbouring hits in the 2D and 3D density. This could be an effect due to a problem in the description of the charge spread over the pads. Another observation is a larger number of hits with very high hit densities in the simulations, see the last bins in Figures 6.11a and 6.11b.
The radial shower shape shows the largest deviations between the data and the simulations, particularly at small radii. This tendency is observed over the full energy range, see Appendix C.2.
In the longitudinal profile the simulation with the standard EM physics lists shows the largest

disagreement with the data by too few hits in the layers $> 5$. In Appendix C.1 the longitudinal profiles for all beam energies are shown and reveal the same behaviour for the standard EM physics list as seen in Figure 6.11c. The disagreement between all simulations and the data in the very small number of hits in the tails for layers $> 20$ can be explained by the low noise level of 0.1 hit per event in the data, which is not included in the simulation. The longitudinal profiles are described using a Gamma distribution function introduced in Equation 1.9, within a range determined by the chosen range of shower start positions from 0 to 32. These fits allow an estimation of the shower maximum $t_{max}$ following Equation 1.8, which is shown as a function of the beam energy in Figure 6.12. The standard EM physics list of GEANT4 shows in this observable the best agreement with the data.

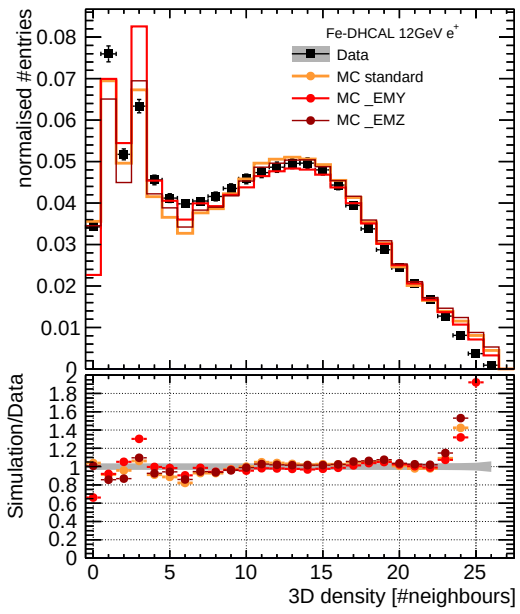
### 6.3.2 Response and Energy Reconstruction

Following the method of the digital energy reconstruction introduced in Section 4.3, the positron response in terms of the mean number of hits per event $\langle N_{hits} \rangle$ is measured. To extract the mean number of hits of every $N_{hits}$ distributions, shown in Figure 6.13a, a Novosibirsk fit is used following the procedure introduced in Section 4.1.
Figure 6.13c shows the mean number of hits as a function of the beam energy, which is fitted with a power law function. The comparison of the data with the three EM physics lists of GEANT4 reveals the best agreement for the "option 4" or _EMZ simulation. The standard EM physics list and the "option 3" or _EMY simulation show a deviation of up to 15 %. The simulation with the standard EM list shows systematically too little and the simulation with _EMY too many hits. The reason for this disagreements could be the too short positron showers of the standard EM simulation, see the orange histogram of the longitudinal profile in Figure 6.11c. The too large number of hits of the _EMY simulation could be related to the too large number hits with high hit density, see Figures 6.11a and 6.11b.

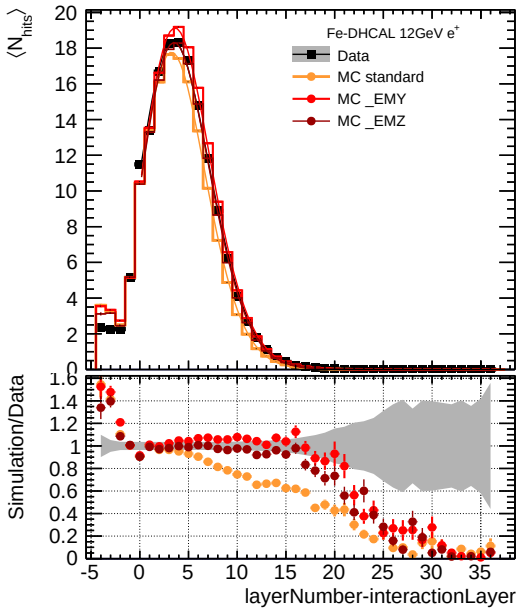
Due to the dense EM showers and relatively large pad size of the DHCAL compared to the core of EM showers, the data as well as the simulation show a saturation in the mean number of hits $\langle N_{hits} \rangle$. In the energy reconstruction procedure this non-linearity is corrected by using the inverse of the power law function $\langle N_{hits} \rangle = a \cdot E_{beam}^{b} + c$, assuming $E_{rec} = E_{beam}$. This procedure allows to reconstruct the energy event by event. An additional offset parameter c is introduced to improve the description of the data.
A satisfactory linearity is achieved for all samples by using the corresponding energy reconstruction parameters for the data and the simulations. The reconstruction parameters are listed in Table 6.4 and the resulting reconstructed energy distributions are shown in Figure 6.13b. The remaining non-linearities of the mean reconstructed energies are always smaller than 5 %, see Figure 6.13d.



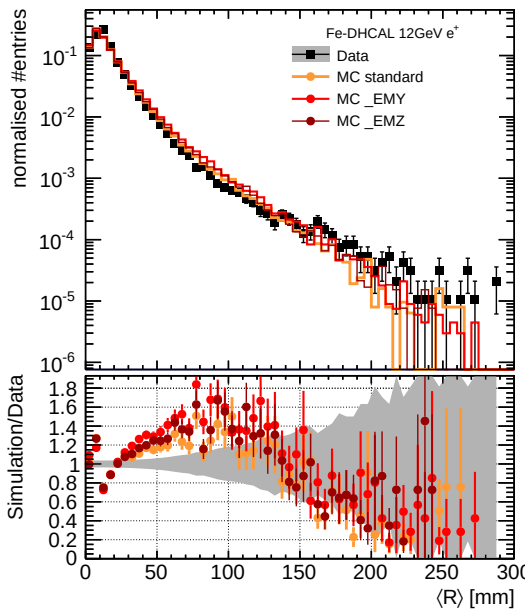

**a)** 2D hit density




**b)** 3D hit density




**c)** Longitudinal profile




**d)** Radial shower shape

**Figure 6.11:** The shower observables for 12 GeV positrons; a) 2D hit density, b) 3D hit density, c) the longitudinal profile and d) the radial shower shape. The data are represented as black squares and the grey error band corresponds to the systematic and statistical uncertainties added in quadrature. The orange histograms show the simulation using the standard EM physics list, the red and dark red histograms correspond to the "option 3" and "4" EM lists, respectively. The longitudinal profiles in c) are fitted with Equation 1.9. The fits are shown as coloured curves.

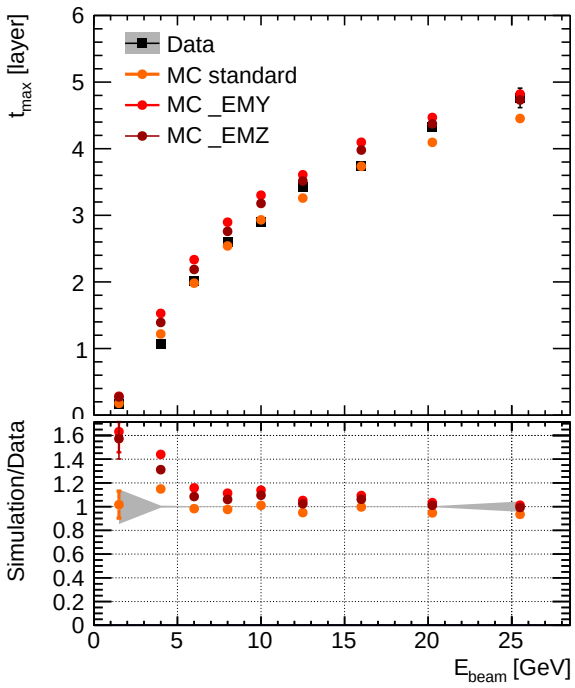
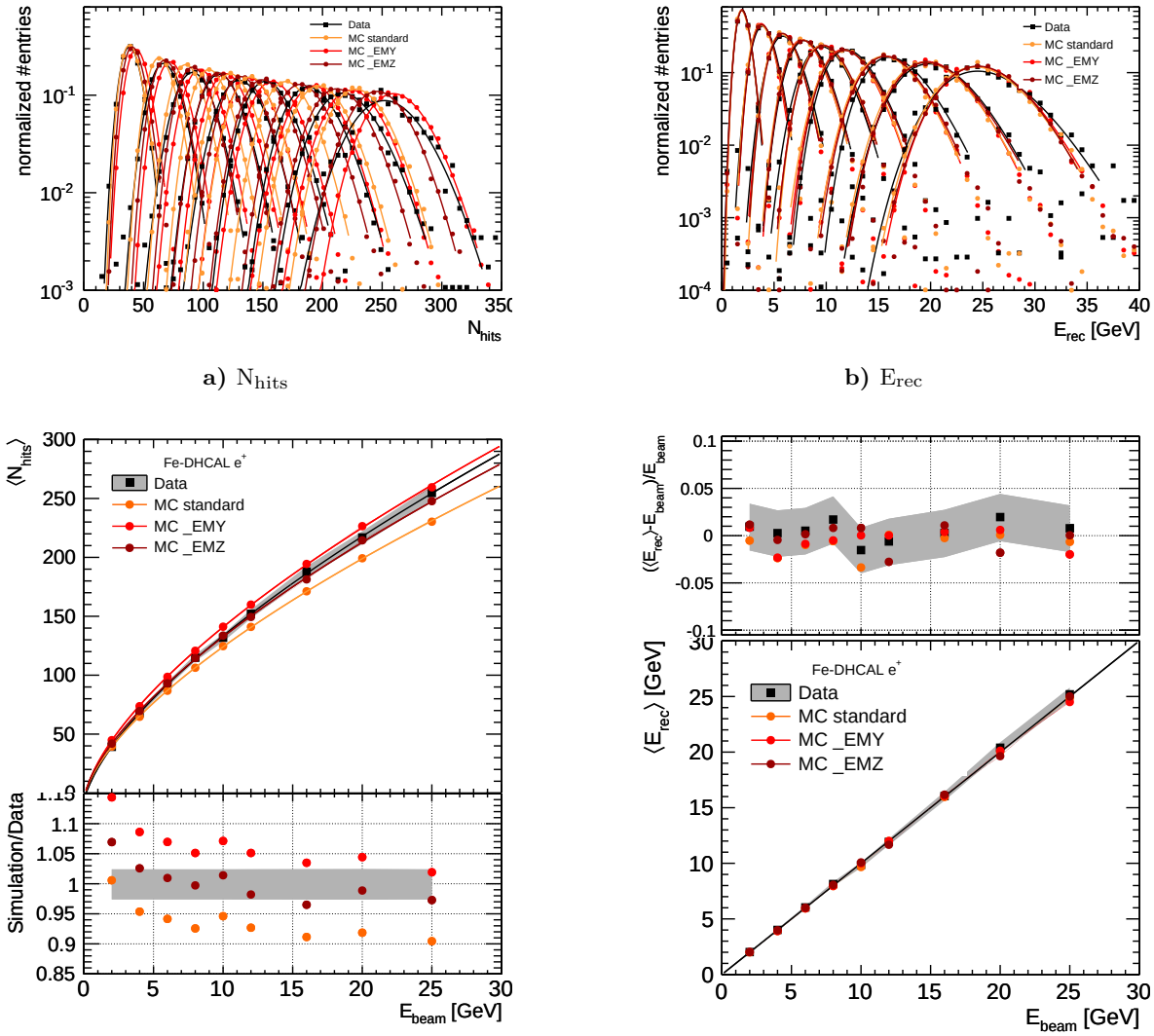



**Figure 6.12:** The positron shower maxima $t_{max}$, calculated from the shower start, for beam energies between 2 and 25 GeV, determined from Equation 1.8. The data is shown as black squares, the standard simulation in orange, the _EMY simulation in red and the _EMZ simulation in dark red histograms. The uncertainty on the shower maximum is shown as error band for the data and coloured error bars for the simulations and has been estimated from the fit function.

**Table 6.4:** The energy reconstruction parameters for $e^+$ events, extracted from the power law fit to the mean response in Figure 6.13c.

| Fit parameter | Data | MC standard | MC _EMY | MC _EMZ |
|---|---|---|---|---|
| a [1/GeV] | $30.6 \pm 4.1$ | $33.4 \pm 0.2$ | $28.9 \pm 0.2$ | $30.5 \pm 0.2$ |
| b | $0.67 \pm 0.04$ | $0.648 \pm 0.002$ | $0.654 \pm 0.002$ | $0.658 \pm 0.002$ |
| c [#] | $9.5 \pm 5.7$ | $7.8 \pm 0.3$ | $6.2 \pm 0.3$ | $6.1 \pm 0.3$ |




**a)** $N_{hits}$ **b)** $E_{rec}$

**c)** $\langle N_{hits} \rangle$ **d)** $\langle E_{rec} \rangle$

**Figure 6.13:** a) The response distributions in number of hits for 2 to 25 GeV positrons for the data (black squares), the standard (orange dots), _EMY (red dots) and _EMZ (dark red dots) simulations. The lines represent the Novosibirsk fits used for the determination of the mean response, shown in c) before the correction for non-linearity to positron showers. The curves show the power law fit function to data (black), the standard (orange), _EMY (red) and _EMZ (dark red) simulations. The fit parameters are summarised in Table 6.4. The plot on the bottom shows the difference between data and simulation. b) The reconstructed energy distributions and d) the mean reconstructed energy. The plot on the top shows the residuals to the beam energy.
The grey bands indicate the statistical and systematic uncertainty of the data. The statistical errors of the simulations are smaller than the markers.

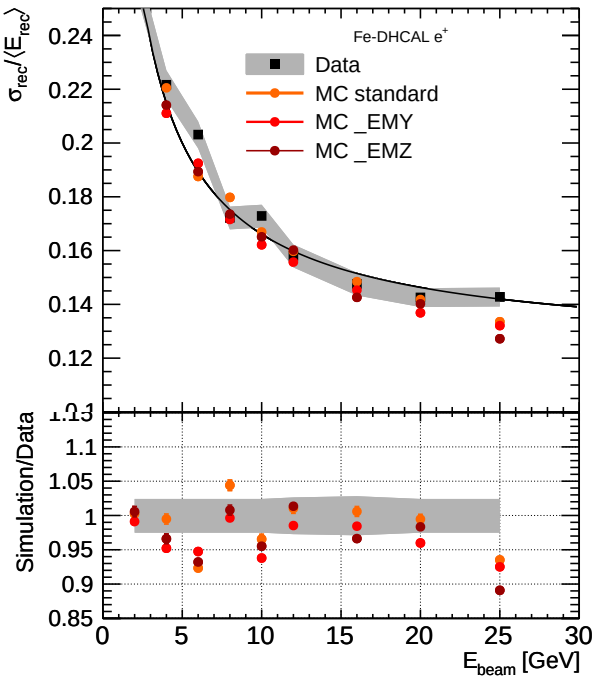



**Figure 6.14:** The positron energy resolution for beam energies of 2 to 25 GeV. The bottom plots shows the ratio of the simulations and data. The data is shown as black squares, the standard simulation in orange, the _EMY simulation in red and the _EMZ simulation in dark red dots. The energy resolution of the data is fitted by Equation 6.5 with $a = 35.1 \pm 0.9\,\%$ and $b = 12.4 \pm 0.3\,\%$. The error bands show the systematic and statical uncertainty added in quadrature. The statistical errors of the simulations are smaller than the markers.

### 6.3.3 Energy Resolution

The energy resolution for positron showers is obtained using the method explained in Section 4.1 and also used in the analysis of the AHCAL data and simulation, see Chapter 5.
The obtained resolution is shown in Figure 6.14, where the data points (black squares) are described using

$$\frac{\sigma_{\mathrm{rec}}}{\langle \mathrm{E_{rec}} \rangle} = \frac{\mathrm{a}}{\sqrt{\mathrm{E_{beam}[GeV]}}} \oplus \mathrm{b}. \tag{6.5}$$

The ratio between the simulation and the data (see the bottom plot in Figure 6.14) shows an agreement within 5 % for the energies of 2 to 20 GeV. The simulated 25 GeV positrons show a better resolution of around 10 %. This disagreement could be a result of the limited statistics of the 25 GeV positron sample in the data, containing only 578 events, see Table 6.2.
This poor resolution of $35.1 \pm 0.9\,\%/\sqrt{\mathrm{E}}$ and a constant term of $12.4 \pm 0.3\,\%$ is mostly due to the saturation caused by the dense EM showers and the digital readout of the $1 \times 1\,\mathrm{cm}^2$ pads. However, by applying a weighting based on the hit densities, following the method described in Section 4.5.2, an improvement is expected and has been achieved in the analysis of the data recorded with the DHCAL without absorbers [69].

## 6.4 Analysis of Pion Showers

The analysis of the positrons revealed a large variation of the simulations using different GEANT4 EM physics lists. Hadron showers underly even larger fluctuations than electromagnetic showers, which requires the use of simplified models for the simulation, see Section 3.3. In the following the $\pi^+$ showers are studied and compared to the simulation, using the hadronic physics lists FTFP_BERT and QGSP_BERT, which showed the best performance for hadrons [76]. The hadronic physics lists are tested for the three different EM physics lists options previously discussed.

### 6.4.1 Systematic Uncertainty on the Simulation

The tuning process of the digitiser is based on the muon and positron data. Therefore there is no meaningful attribution of a systematic error on the simulations for these particles.
The remaining differences in the description of the positrons can be used as an estimate of the uncertainty on the pion simulations. This is done for each shower observable individually; hit density (2 and 3D), longitudinal profile, radial shower shape, shower maximum, mean number of hits and energy resolution.
The uncertainties on the observable x is calculated as the average difference between the positron data and the simulations $\Delta x = |x_{sim}/x_{data} - 1|$ for all beam energies with the three EM physics lists per bin, see the bottom plots of Figures 6.11 and 6.12, following

$$\sigma_x = \frac{1}{N_E N_{bins}} \cdot \sum_{i=1}^{N_E} \sum_{j=1}^{N_{bins}} \Delta x_{i,j}, \tag{6.6}$$

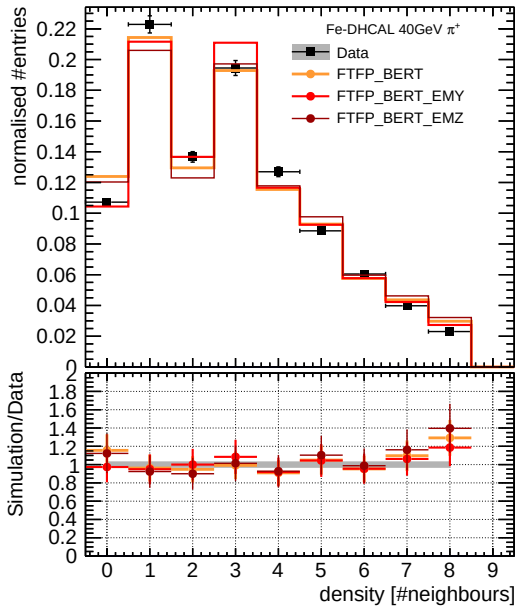
with $N_E$ the number of beam energies, and $N_{bins}$ the number of bins of the compared histograms with sufficient statistics. This is a conservative approach and results in relatively large systematic uncertainties on the pion measurements. The values are summarised in Table 6.5.

**Table 6.5:** The average uncertainty on the longitudinal ($\sigma_{longProfile}$) and radial profile ($\sigma_{radProfile}$), the 2D ($\sigma_{density}$) and 3D hit densities ($\sigma_{3Ddensity}$) as well as on the shower maximum ($\sigma_{t_{max}}$), the mean number of hits ($\sigma_{\langle N_{hits}\rangle}$) and the resolution ($\sigma_{\sigma_{rec}/\langle E_{rec}\rangle}$) in percent for the standard, option 3 (_EMY) and option 4 (_EMZ) EM physics lists.

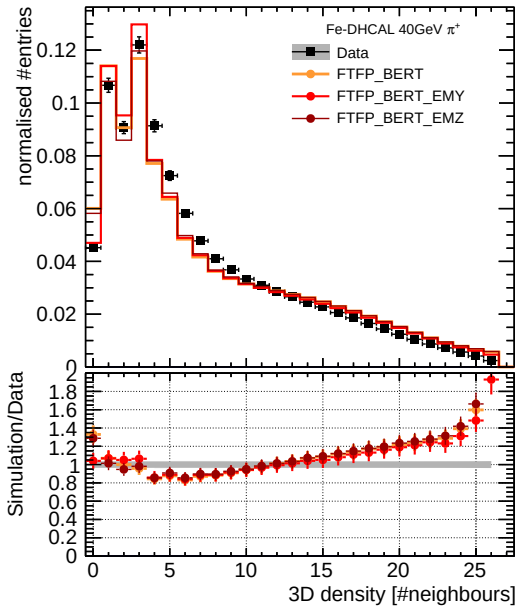
| EM physics list | $\sigma_{density}$ [%] | $\sigma_{3Ddensity}$ [%] | $\sigma_{longProfile}$ [%] | $\sigma_{radProfile}$ [%] | $\sigma_{t_{max}}$ [%] | $\sigma_{\langle N_{hits}\rangle}$ [%] | $\sigma_{\sigma_{rec}/\langle E_{rec}\rangle}$ [%] |
|---|---|---|---|---|---|---|---|
| standard | 10.1 | 24.7 | 29.9 | 22.3 | 0.3 | 6.6 | 4.0 |
| _EMY | 11.9 | 15.3 | 22.8 | 27.1 | 18.0 | 6.3 | 4.4 |
| _EMZ | 13.0 | 26.4 | 20.9 | 20.9 | 14.0 | 0.2 | 4.0 |

### 6.4.2 Pion Showers

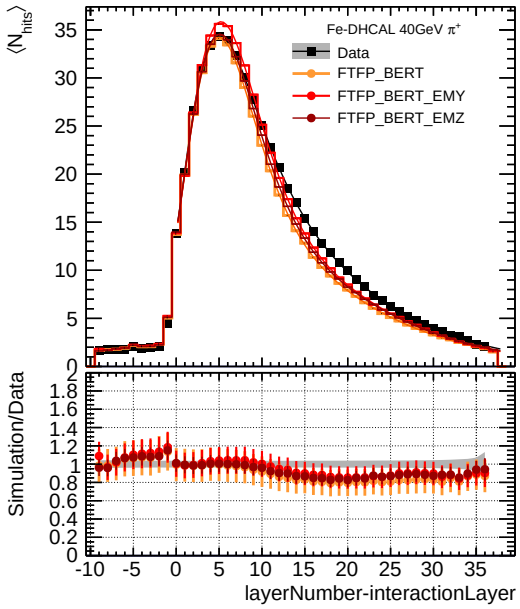
For 40 GeV $\pi^+$ events the comparison of the simulation with data in the shower observables; 2D density, 3D density, longitudinal profile and radial shower shape is shown for FTFP_BERT physics list in Figure 6.15 and the QGSP_BERT physics list in Figure 6.16. Both simulations are repeated using the standard, _EMY and _EMZ EM physics lists, which give similar results. The longitudinal and radial shower profiles for the other beam energies are shown in Appendix D.1 and D.2. A significant disagreement between data and MC is seen in the 3D density. The FTFP_BERT simulations have a larger number of very high density hits (above 18 neighbour hits), whereas the data show a larger number of hits in the intermediate region of 4 to 10 neighbouring hits. This trend is also observed for the QGSP_BERT simulations, see Figure 6.16b. However, the simulations using the _EMY option show an excellent agreement with data in the 2D densities.
The longitudinal profiles, see Figures 6.15c and 6.16c, are well described by the simulations. However, the FTFP_BERT, QGSP_BERT and FTFP_BERT_EMZ, QGSP_BERT_EMZ simulations show a tendency of too little number of hits in the 10 to 35th layer from the shower start. This is observed for energies above 20 GeV, for the beam energies from 10 to 20 GeV the longitudinal profiles show an agreement for all studied physics lists. For 6 and 8 GeV the simulations are in agreement with the data for the first 20 layers after the shower start, compare Figure D.1, for the last 18 layers the simulation produces too many hits. The longitudinal profiles of the data, the FTFP_BERT and QGSP_BERT simulations are shown in Appendix D for all beam energies.
The radial shower shapes of the simulated pion showers, shown in Figures 6.15d and 6.16d, have the largest uncertainties, which have been estimated from the positron simulations, and also show the largest deviations from data over the whole energy range. The simulated showers tend to exhibit larger radial dispersion than the measured showers in data. This effect is slightly enhanced for the _EMY EM physics list.
From the longitudinal shower profiles the shower maximum is determined by fitting Equation 1.13 to the data. The result is shown as a function of the beam energy in Figure 6.17a for the FTFP_BERT simulations and in Figure 6.17b for the QGSP_BERT simulations. In both cases the shower maxima are well reproduced for $\pi^+$ showers of energies above 8 GeV. For the lowest two beam energies the simulations deviate of up to 40 % in the QGSP_BERT simulation.



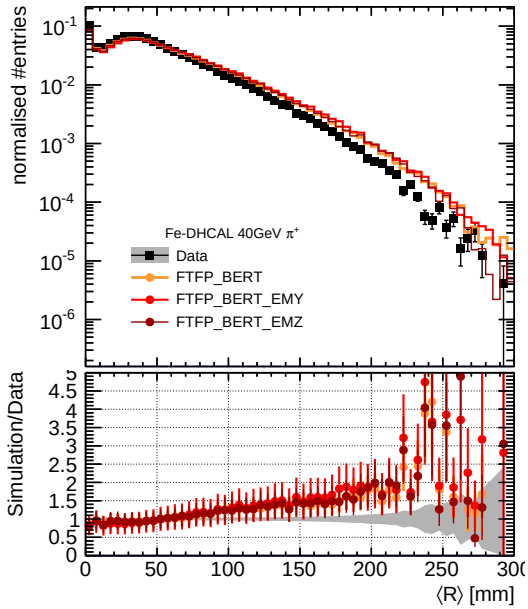

**a)** 2D hit density




**b)** 3D hit density

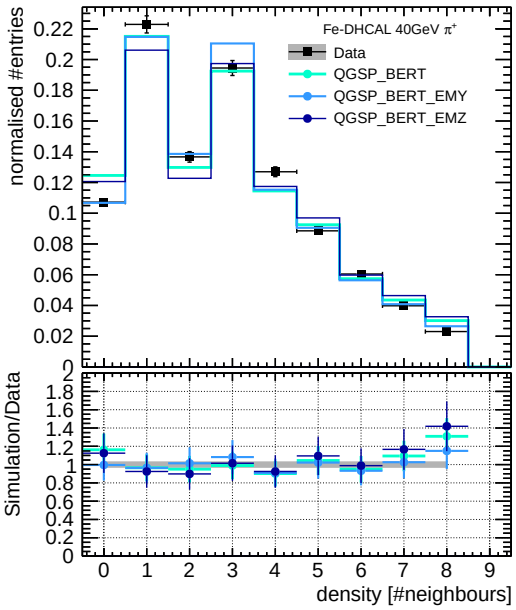



**c)** Longitudinal profile

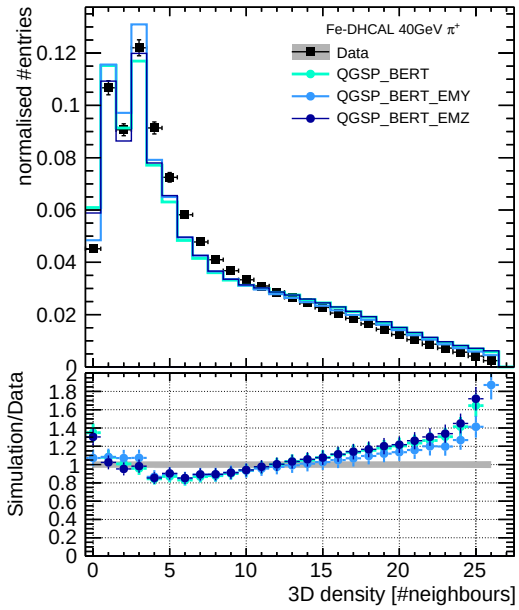



**d)** Radial shower shape

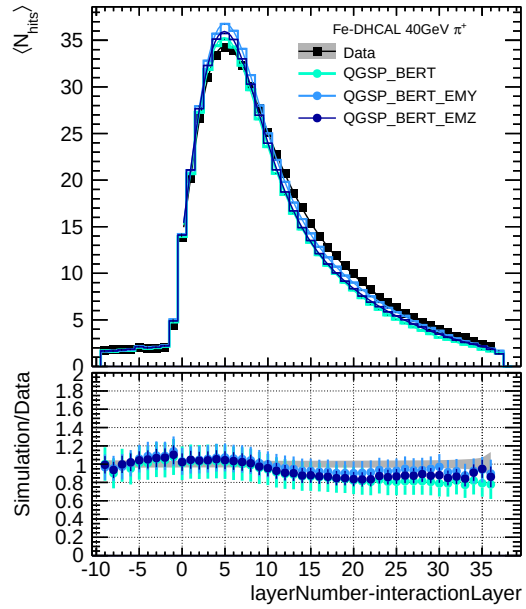
**Figure 6.15:** The shower observables for 40 GeV $\pi^+$ events; a) 2D hit density, b) 3D hit density, c) the longitudinal profile and d) the radial shower shape. The data is represented as black squares and the grey error band corresponds to the systematic and statistical uncertainty added in quadrature. The orange histograms show the FTFP_BERT simulation using the standard EM physics list, the red and dark red histograms correspond to the FTFP_BERT_EMY and FTFP_BERT_EMZ physics lists. The longitudinal profiles in c) are described by Equation 1.13 and the fits are shown as coloured curves. The ratios in the bottom plots show also the systematic uncertainty on the simulations.



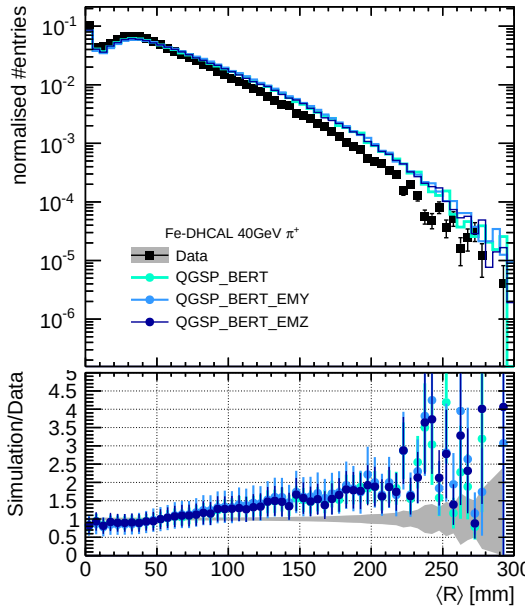

**a)** 2D hit density




**b)** 3D hit density

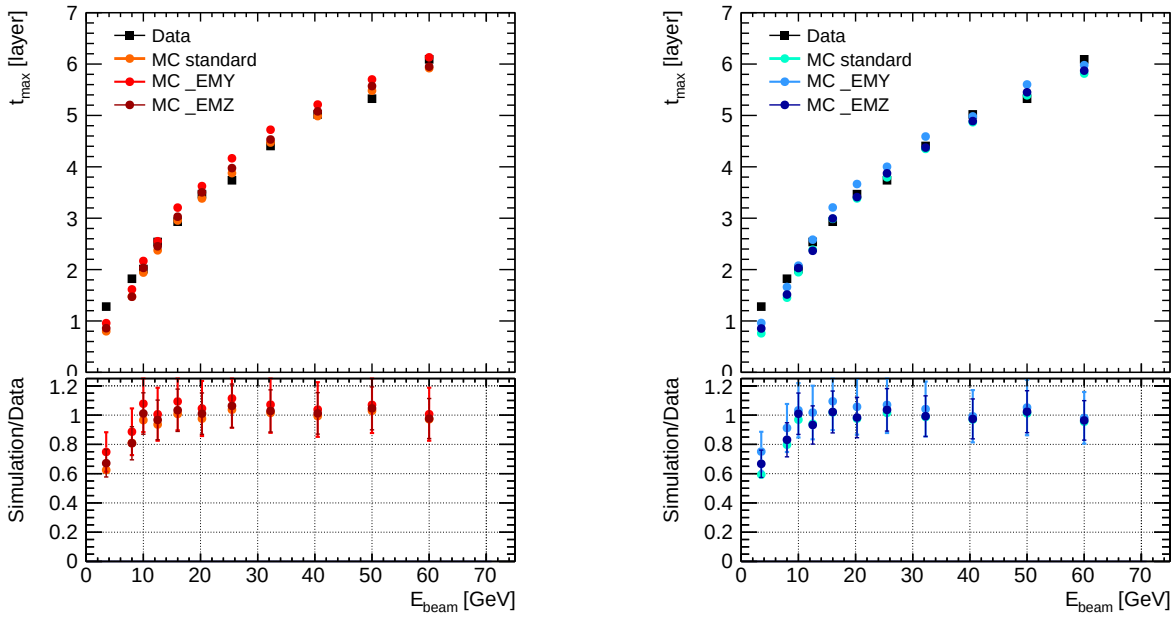



**c)** Longitudinal profile




**d)** Radial shower shape

**Figure 6.16:** The shower observables for 40 GeV $\pi^+$ events; a) 2D hit density, b) 3D hit density, c) the longitudinal profile and d) the radial shower shape. The data is represented as black squares and the grey error band corresponds to the systematic and statistical uncertainty added in quadrature. The cyan histograms show the QGSP_BERT simulation using the standard EM physics list, the blue and dark blue histograms correspond to the QGSP_BERT_EMY and QGSP_BERT_EMZ lists. The longitudinal profiles in c) are described by Equation 1.13 and the fits are shown as coloured curves. The ratios in the bottom plots show also the systematic uncertainty on the simulations.




**a)** FTFP_BERT **b)** QGSP_BERT

**Figure 6.17:** The pion shower maxima $t_{max}$, calculated from the shower start, for beam energies of 6 to 60 GeV, estimated from the longitudinal profiles. The data is shown as black squares in a) compared to the FTFP_BERT simulation and in b) to the QGSP_BERT simulation. a) The results of the FTFP_BERT simulation is shown in orange using the standard, in red using the _EMY option and in dark red using the _EMZ option in the description of the EM processes. b) The results of the QGSP_BERT simulation is shown in cyan using the standard, in blue using the _EMY option and in dark blue using the _EMZ option in the description of the EM shower components. The uncertainty on the shower maxima is shown as coloured error bars for the simulations.

**Table 6.6:** The energy reconstruction parameters for $\pi^+$ events, extracted from the power law fit to the mean response in Figure 6.18c and 6.19c.

| Fit parameter | Data | FTFP_BERT | FTFP_BERT_EMY | FTFP_BERT_EMZ |
|---|---|---|---|---|
| a [1/GeV] | $21.1 \pm 3.0$ | $31.2 \pm 12.1$ | $28.6 \pm 11.6$ | $30.1 \pm 0.5$ |
| b | $0.89 \pm 0.03$ | $0.79 \pm 0.09$ | $0.79 \pm 0.09$ | $0.785 \pm 0.004$ |
| c [#] | $15.0 \pm 9.8$ | $36.9 \pm 31.8$ | $32.6 \pm 30.9$ | $36.1 \pm 1.3$ |
| Fit parameter | Data | QGSP_BERT | QGSP_BERT_EMY | QGSP_BERT_EMZ |
| a [1/GeV] | $21.1 \pm 3.0$ | $30.7 \pm 12.0$ | $30.9 \pm 11.5$ | $31.8 \pm 0.5$ |
| b | $0.89 \pm 0.03$ | $0.79 \pm 0.09$ | $0.80 \pm 0.09$ | $0.782 \pm 0.004$ |
| c [#] | $15.0 \pm 9.8$ | $42.8 \pm 31.3$ | $40.6 \pm 30.9$ | $44.7 \pm 1.3$ |

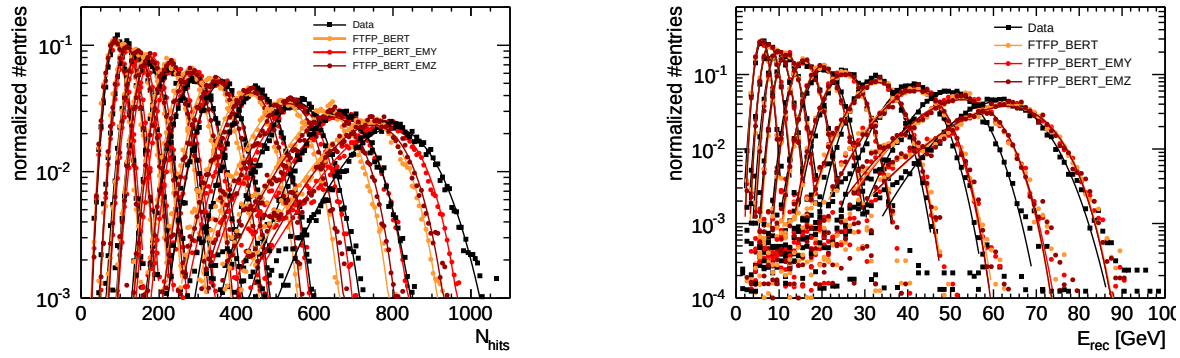
### 6.4.3 Response and Energy Reconstruction

The distributions of the total number of hits for all beam energies are shown for the FTFP_BERT simulations and the data in Figure 6.18a and the QGSP_BERT simulations compared to the data in Figure 6.19a. In both the simulations and the data a tail to smaller number of hits is seen for beam energies larger than 20 GeV due to saturation effects. To include these tails in the estimation of the mean response, Novosibirsk fits are used and shown as curves in these figures. The method to extract the mean response is described in Section 4.1.

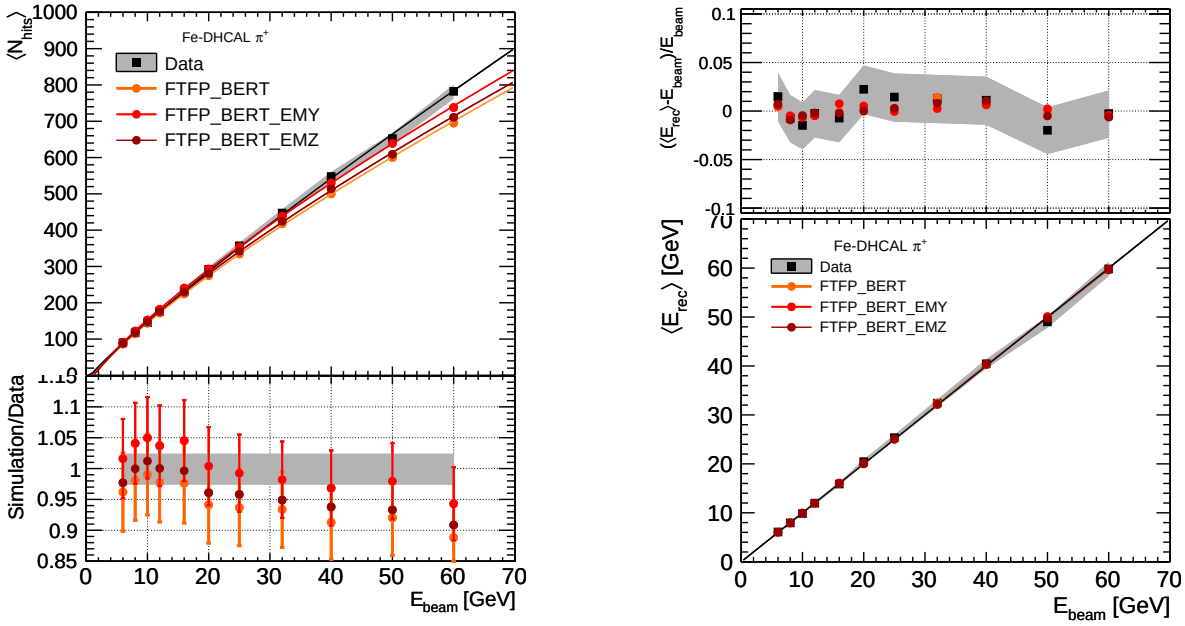
The mean number of hits $\langle N_{hits} \rangle$ is shown in Figure 6.18c for the FTFP_BERT simulations compared to data and in Figure 6.19c for the QGSP_BERT simulations. Both used hadronic physics lists exhibit a stronger saturation than seen in the data. However, the QGSP_BERT simulation using the EM physics list _EMY shows the best agreement with data, as illustrated by the ratio of the simulation with the data in the bottom plot of Figure 6.19c.

To compare the energy resolution of the data and the simulations a satisfactory linearity in the reconstructed energies is required. To achieve this, a power law function $\langle N_{hits} \rangle = a \cdot E_{beam}^{b} + c$ is fitted to the mean response and by requiring $E_{rec} = E_{beam}$, the parameters of this fit are used to reconstruct the event energies. For each data set the corresponding reconstruction parameters are used for the energy reconstruction. These parameters are summarised in Table 6.6, where the stronger saturation in the simulations is expressed by smaller b parameters. The resulting energy distributions are shown in Figures 6.18b and 6.19b.

The determined mean reconstructed energies are shown in Figures 6.18d and 6.19d. The residuals to the beam energy reveal that the non-linearities are smaller than 2 % and smaller than the uncertainty of the data.

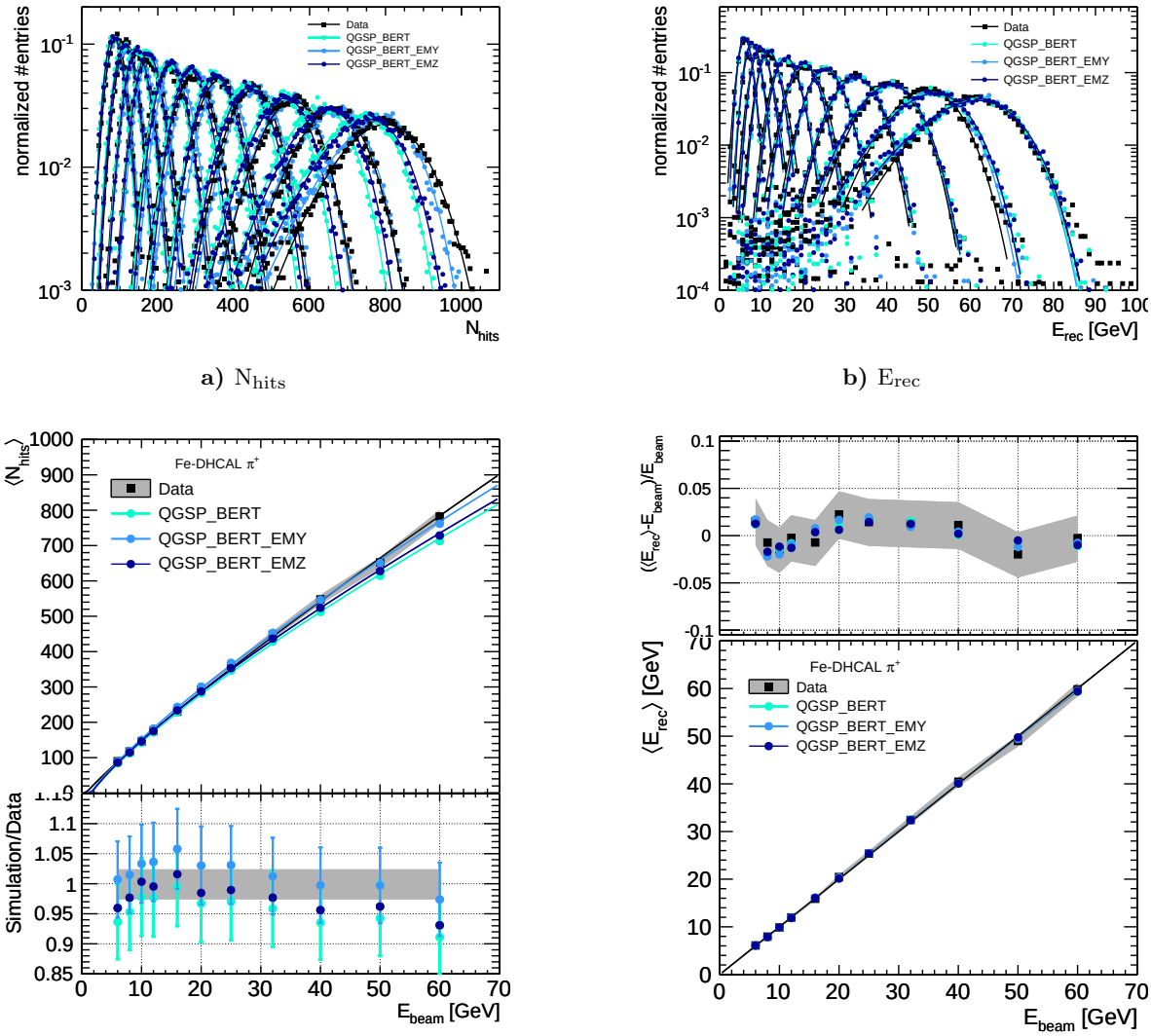



**a)** $N_{hits}$ **b)** $E_{rec}$




**c)** $\langle N_{hits} \rangle$ **d)** $\langle E_{rec} \rangle$

**Figure 6.18:** a) The response distributions in number of hits for 6 to 60 GeV $\pi^+$ for the data (black squares), the FTFP_BERT (orange dots), FTFP_BERT_EMY (red dots) and FTFP_BERT_EMZ (dark red dots) simulations. The lines represent the Novosibirsk fits used for the determination of the mean response, shown in c) before the correction for the non-linearity. The curves show the power law fit function to data (black), the FTFP_BERT (orange), FTFP_BERT_EMY (red) and FTFP_BERT_EMZ (dark red) simulations. The fit parameters are summarised in Table 6.6. The plot on the bottom shows the difference between data and simulation. b) The reconstructed energy distributions and d) the mean reconstructed energy. The plot on the top shows the residuals to the beam energy.
The grey bands indicate the statistical and systematic uncertainty of the data from the calibration. The statistical errors are smaller than the markers.




**a)** $N_{hits}$ **b)** $E_{rec}$

**c)** $\langle N_{hits} \rangle$ **d)** $\langle E_{rec} \rangle$

**Figure 6.19:** a) The response distributions in number of hits for 6 to 60 GeV $\pi^+$ for the data (black squares), the QGSP_BERT (cyan dots), QGSP_BERT_EMY (blue dots) and QGSP_BERT_EMZ (dark blue dots) simulations. The lines represent the Novosibirsk fits used for the determination of the mean response, shown in c) before the correction for non-linearity to positron showers. The curves show the power law fit function to data (black), the QGSP_BERT (cyan), QGSP_BERT_EMY (blue) and QGSP_BERT_EMZ (dark blue) simulations. The fit parameters are summarised in Table 6.6. The plot on the bottom shows the difference between data and simulation. b) The reconstructed energy distributions and d) the mean reconstructed energy. The plot on the top shows the residuals to the beam energy.
The grey bands indicate the statistical and systematic uncertainty of the data from the calibration. The statistical errors are smaller than the markers.

### 6.4.4 Energy Resolution

The pion energy resolution of the Fe-DHCAL is shown in Figure 6.20. The resolution observed in the data is showing a typical $1/\sqrt{E}$ behaviour only for beam energies below 30 GeV, for the beam energies of 30 to 50 GeV a minimum is reached at around 14 % and the energy resolution is degrading for 60 GeV $\pi^+$ events. The black curve in Figure 6.20 shows the fit to the data with [23]

$$\frac{\sigma_{rec}}{\langle E_{rec}\rangle} = \frac{a}{\sqrt{E_{beam}[GeV]}} \oplus b \oplus d \left(\frac{E_{beam}[GeV]}{100}\right)^e, \tag{6.7}$$

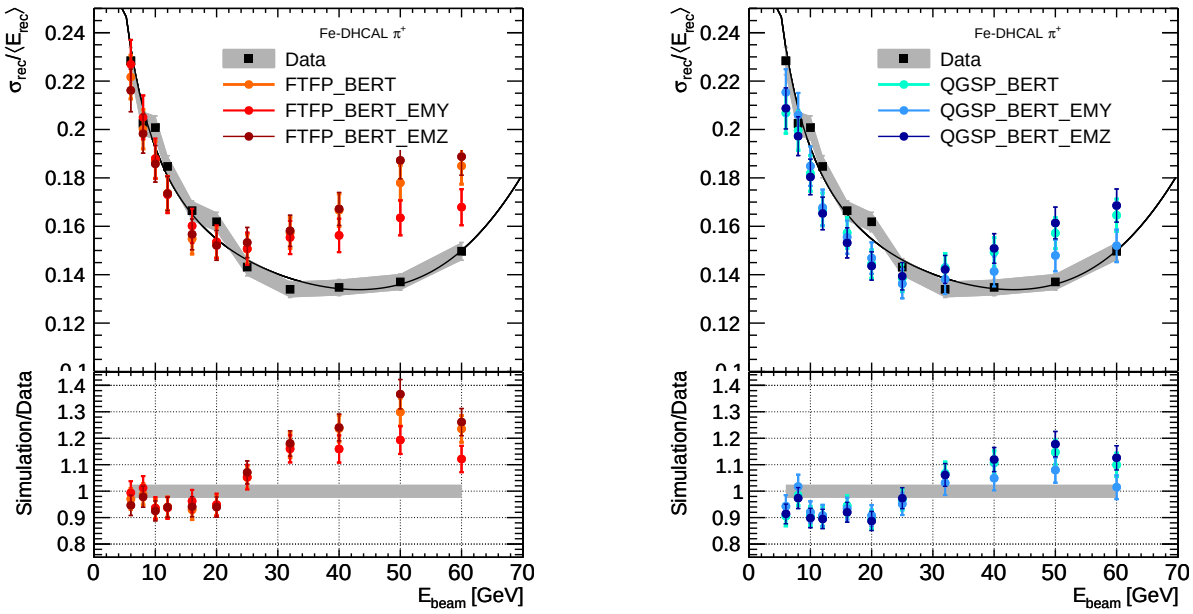
where the noise term of $\frac{c}{E_{beam}[GeV]}$ is neglected because of the low noise rates in the Fe-DHCAL [102] of 0.1 hit/event. The parameters are not meaningful to compare to values obtained with a more standard parametrisation without the last term and therefore not stated here. The main purpose is to show the trend of the resolution with the beam energy.

The behaviour of the resolution is not influenced by leakage, neither longitudinal nor lateral, see the longitudinal profiles and radial shower shapes in Appendix D and D.2. It is due to the saturation in the response and the cell size of $1 \times 1\,cm^2$.

The comparison of the resolutions observed in the simulations reveal a strong dependence on the EM and the hadronic physics lists. While all simulation samples achieve a good linearity, the energy resolutions deviate up to 40 % for the FTFP_BERT_EMZ simulation from the data. This effect seems to originate from the prediction of either too dense shower parts or a too high electromagnetic fraction $f_{EM}$, which results in a stronger saturation than observed in data. This saturation appears in the $N_{hits}$ distributions in large tails towards smaller numbers of hits, especially pronounced in the FTFP_BERT simulations, see Figure 6.18a. The hypothesis is additionally confirmed by the observations in the hit densities, where all simulations predict a larger number of hits with very high hit densities, see Figures 6.15b and 6.16b.

The best agreement in the energy resolution between the data and MC is observed for the simulation using the QGSP_BERT_EMY physics list, with a mean remaining difference of less than 5 %, see the bottom plot in Figure 6.20b.

It has to be mentioned that by including a weighting of hits dependent on their hit density the strong non-compensation of the DHCAL can be corrected to a certain degree and the energy resolution can thus be improved [103]. However, further studies are necessary to determine to which extent this is possible.




**a)** FTFP_BERT **b)** QGSP_BERT

**Figure 6.20:** The $\pi^+$ energy resolution of the Fe-DHCAL for beam energies from 6 to 60 GeV. The bottom plots show the ratio of the simulations and data. The data is shown as black squares and the black curve represents the fit with Equation 6.7. The standard simulation in orange (cyan), the _EMY simulation in red (blue) and the _EMZ simulation in dark red (dark blue) dots for the a) FTFP_BERT and b) QGSP_BERT physics lists. The error bands show the systematic and statical uncertainty added in quadrature. The statistical errors are smaller than the markers.

## 6.5 e/π Ratio of the Fe-DHCAL

The e/π ratio of the Fe-DHCAL is determined from the mean response to positrons and pions, shown in Figure 6.13c and 6.18c, for the data and the simulations. The e/π ratio of the Fe-DHCAL is energy dependent and varies from 1.03 to 0.74 between 6 and 25 GeV and can be parameterised, following Equation 1.12 and 1.14, as:

$$\frac{e}{\pi} = \frac{e/h}{1 - \left[1 - \left(\frac{E_{beam}}{E_0}\right)^{k-1}\right] \cdot (1 - e/h)}, \tag{6.8}$$

with e/h the ratio between the response to electromagnetic and non-electromagnetic shower components, $E_0$ the energy threshold for $\pi^0$ production and the factor k, that is related to the multiplicity of $\pi^0$s. The fit to the data is shown as a black curve in Figure 6.21 and the resulting parameter values are: $e/h = 0.62 \pm 0.04$, $E_0 = 1.1 \pm 0.8$ and $k = 0.74 \pm 0.03$. The values of $E_0$ and k are in agreement with the literature of $E_0 = 0.8$ GeV for iron and $k \sim 0.75$–$0.85$ [10]. All simulations agree within the errors with the data. However, the FTFP_BERT and QGSP_BERT physics lists using the EM physics list option _EMY achieve the best agreement with the data. It is observed that the standard EM physics lists option have a tendency to underestimate the e/π ratio.

The increasing non-compensation of the Fe-DHCAL with higher beam energies degrades the energy resolution for pion (hadron) showers and motivates the development of software compensation algorithms. These algorithms can correct for the lower EM response by weighting hits belonging to EM sub-showers and hits in the hadronic shower parts differently. The classification of hits is possible using the measured hit densities since EM sub-showers are denser than hadronic sub-showers and thus the hits belonging to the EM part have much higher hit densities than the hits in the hadronic shower parts.

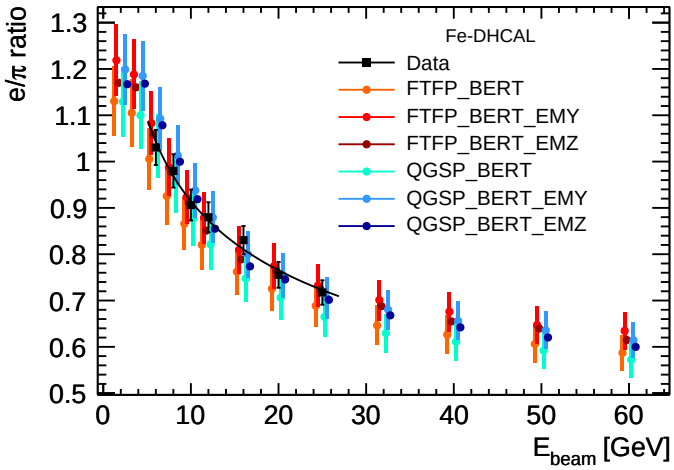



**Figure 6.21:** The e/π ratio determined from the mean response to positrons and pions in the Fe-DHCAL, for the data in black squares, the FTFP_BERT simulations in red and the QGSP_BERT simulations in blue dots. The black curve shows the fit to data with Equation 6.8 and the parameters: $e/h = 0.62 \pm 0.04$, $E_0 = 1.1 \pm 0.8$ and $k = 0.74 \pm 0.03$. The markers of the simulations are shifted in $E_{beam}$ for visibility. The error bars represent the systematic and statistical uncertainties added in quadrature.

## 6.6 Summary and Outlook

The Fe-DHCAL was operated in a mixed particle beam at Fermilab. During the testbeam the environmental conditions varied and a calibration had to be applied. The calibration method used in this analysis showed a satisfactory performance, see Section 6.1.2, and is only limited in the particle track statistics in the bottom and top RPC for each DHCAL layer.
The imaging capabilities of the DHCAL are successfully used in the event selection to separate muon, positron and pion events, using their shower topologies, see Section 6.1, without biasing the data samples.
The simulation of the Fe-DHCAL testbeam setup is implemented within the GEANT4 and MOKKA software framework, which allows the test of a variety of different physics lists provided by GEANT4 and introduced in Chapter 3. The simulation of the RPC response (digitisation) is done, assuming all RPCs to be operated in the same condition. The Fe-DHCAL response to muons and positrons from calibrated data samples is used as reference. The tuning of the digitisation parameters demands some assumptions on the physical range of the parameter values and much computing time, due to the testing of all possible combinations.
The comparison between the data and the simulations reveals a strong dependence of the response and energy resolution on the used EM physics lists for the positron and the pion showers. The simulation of the positrons using the "option 4" or _EMZ physics list shows overall the best agreement with data. The best agreement for pion showers between the data and the simulation is achieved using the QGSP_BERT_EMY physics list.
The difference in the response to electromagnetic and hadronic showers, see Section 6.5, should allow an improvement of the energy resolution by the application of software compensation algorithms. In future, the energy resolution of the Fe-DHCAL of at best 13.5 % at 32 GeV can be improved, following a weighting principle introduced in Section 4.5.
The simulation of the Fe-DHCAL testbeam is the first in detail validated simulation of the prototype calorimeter based on RPCs, which allows future studies to compare the scintillator and RPC based calorimeters in further detail.

# Chapter 7

# Comparison of Calorimeter Technologies

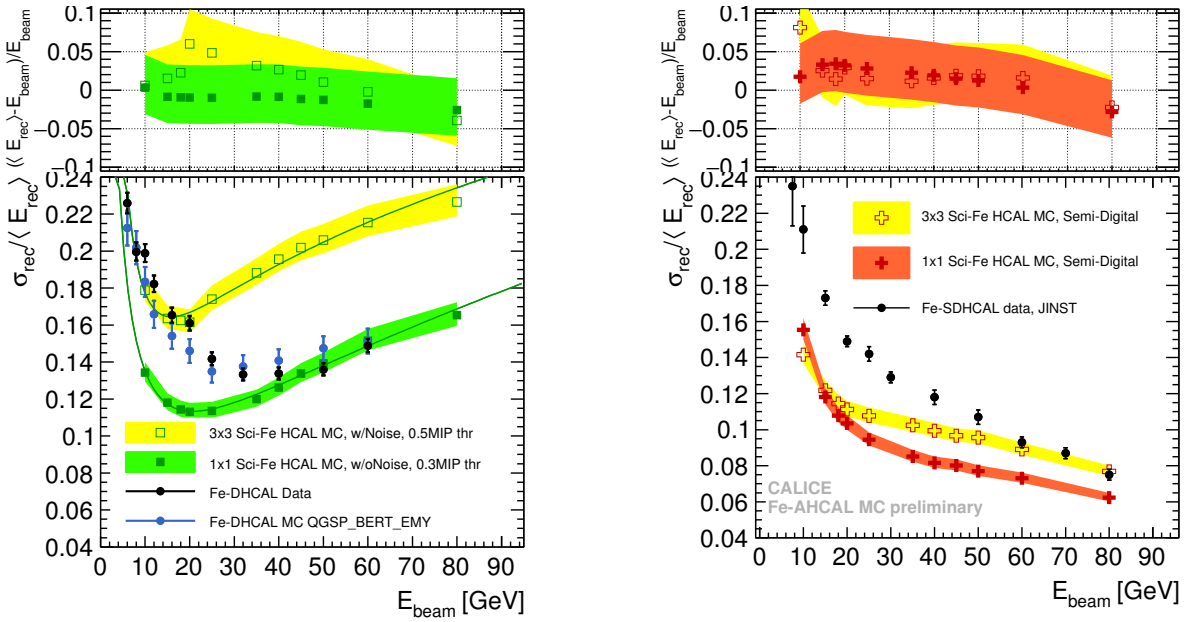
The CALICE collaboration has operated three different hadron calorimeters in testbeams at CERN and Fermilab. These calorimeter prototypes used steel as an absorber and featured a similar $X_0$ and $\lambda_n$ per layer. The different active media (scintillator and gas) require different readout schemes, which again demands different energy reconstruction methods for each calorimeter technology.
This analysis studied the impact of the energy reconstruction methods in the Analogue Hadron Calorimeter (AHCAL) and the impact of a smaller cell size of $1 \times 1\,\mathrm{cm}^2$ on the different energy reconstruction schemes. This allows for the first time a direct comparison between the energy resolutions of all three readout schemes.
The remaining differences in the energy resolution can be attributed to the different active media. This will be explored by comparing the simulations of the scintillator AHCAL to the RPC-based DHCAL and SDHCAL.

The energy resolutions of the Fe-AHCAL simulated with $1 \times 1\,\mathrm{cm}^2$ and $3 \times 3\,\mathrm{cm}^2$ cell size, using the FTFP_BERT physics list of GEANT4, are shown for the digital and semi-digital energy reconstruction in Figure 7.1a and 7.1b. The results are compared to the Fe-DHCAL data and simulation, from Chapter 6, and the testbeam data of the SDHCAL prototype [65]. The energy resolution of the $1 \times 1\,\mathrm{cm}^2$ AHCAL simulation for the digital energy reconstruction is compared to the Fe-DHCAL data and the simulation using the QGSP_BERT physics list, from Section 6.6, see the black and blue dots in Figure 7.1a. The comparison reveals that the digitally read out $1 \times 1\,\mathrm{cm}^2$ AHCAL agrees with the Fe-DHCAL data and simulation for high beam energies $> 40\,\mathrm{GeV}$. However, for energies lower than 40 GeV the $1 \times 1\,\mathrm{cm}^2$ AHCAL simulation achieves a better energy resolution than the Fe-DHCAL with $1 \times 1\,\mathrm{cm}^2$ pad readout. This difference for lower energies can be explained by the difference in the mean energy losses: following a Landau distribution in the scintillator tile and a broader Polya distribution in the

gas gap. This difference is expected to have a stronger effect on the smaller hit number at lower beam energies than on the larger number of hits for the higher beam energies, because at the higher energies the hits originate from more than one traversing particle. This has been seen in the AHCAL and DHCAL data where the mean number of hits shows saturation (see Figure 5.9 and 6.19c). Additionally, the Fe-DHCAL suffers of larger sampling fluctuations due to a much smaller sampling fraction. For beam energies lower than 15 GeV the (Sci-Fe) AHCAL with cell size of $3 \times 3\,\mathrm{cm}^2$ achieves an energy resolution comparable to the one of the Fe-DHCAL.

The increase in the resolution due to saturation effects happens much earlier and much stronger in the scintillator-steel (Sci-Fe) AHCAL with $1 \times 1\,\mathrm{cm}^2$ and $3 \times 3\,\mathrm{cm}^2$ cell size compared to the Fe-DHCAL data and simulation.

For the digital energy reconstruction the energy resolution of the AHCAL is improved between 4.5 and 6.5 % in absolute values by refining the segmentation from $3 \times 3$ to $1 \times 1\,\mathrm{cm}^2$, showing a larger improvement with increasing beam energy.




**a)** Digital Resolutions **b)** Semi-Digital Resolutions

**Figure 7.1:** The digital and semi-digital energy resolutions of the FTFP_BERT simulations are shown as a function of beam energy for cell granularities of $1 \times 1$ and $3 \times 3\,\mathrm{cm}^2$. The results are compared to the test beam data and QGSP_BERT simulation of the Fe-DHCAL (see Chapter 6) and the testbeam data of the SDHCAL prototype [65]. The plots on the top show the residuals to the beam energy with the bands indicating the systematic and statistical uncertainties. The statistical errors are smaller than the markers.

Compared to the SDHCAL data points from [65], the AHCAL simulation with $1 \times 1\,\mathrm{cm}^2$ cells shows a better resolution for all beam energies. Between 60 and 80 GeV, however, the resolutions of the SDHCAL data and AHCAL simulation with $3 \times 3\,\mathrm{cm}^2$ cell size agree. It has to be

mentioned that the SDHCAL data was recorded with 48 active layers, while these simulations are done for only 38 AHCAL layers. However, the SDHCAL event selection is constrained to have the showers started after the first 5 layers to remove the electron contamination, thus reducing the effective number of layers to 43. In addition the threshold settings of the SDHCAL prototype have, in contrast to the AHCAL, not been optimised. Therefore, the semi-digital resolution of the $1 \times 1\,\mathrm{cm}^2$ AHCAL is expected to show a better performance and can not be directly ascribed to the differences in the signal formation and sampling fraction.
The improvement due to a finer granularity on the semi-digital resolutions is approximately 2 % in absolute values for beam energies larger than 25 GeV. For lower beam energies the improvement is smaller and for 10 GeV no difference is observed.

The comparison of the AHCAL with the DHCAL and SDHCAL data shows an advantage for low energies of the scintillator-tile calorimeter. However, this effect can have many different explanations: the different sampling fractions of 4.9 and $4 \times 10^{-3}$ (see Table 2.4 in Section 2.5); the different threshold setting; the difference in the signal readout; or the difference in the energy loss in the scintillator and the gas gap (a Landau [9] versus a Polya function [56]). Further investigations will be necessary to disentangle the different effects. A first step has been done by validating the simulation of the Fe-DHCAL data, see Section 6.3 and 6.4. This validated simulation has the power to give answers to some of the open questions.

# Conclusions and Outlook

This analysis concentrates on the comparison between the different high granularity hadron calorimeter concepts developed by the CALICE collaboration: The Analogue HCAL consisting of $3 \times 3\,\mathrm{cm}^2$ scintillator tiles with analog SiPM readout, the Digital HCAL based on RPCs and binary $1 \times 1\,\mathrm{cm}^2$ pad readout, and the Semi-Digital HCAL that also uses RPCs but including 3 thresholds applied to the $1 \times 1\,\mathrm{cm}^2$ pads. To understand the impact of the different active media, readout schemes and granularities the testbeam data of the Fe-AHCAL and Fe-DHCAL prototypes have been analysed.

The Fe-AHCAL prototype has recorded 10 to 80 GeV pion events from the SPS at CERN in 2007. This testbeam data set has been calibrated and the simulation has been validated by earlier studies [62,71,76,84]. This analysis studied the impact of different energy reconstruction procedures on the energy resolution of the CALICE Fe-AHCAL. To ensure the comparability between these procedures a method to determine the mean and width of asymmetric distributions has been developed, using a Novosibirsk function. The observed energy resolutions have been compared and validated with the results of earlier studies.
Additionally the impact of a finer granularity on the energy reconstruction performance has been investigated and the reconstruction procedures that include an energy dependent weighting (semi-digital and analogue software compensation) have been compared. The best energy resolution for the AHCAL with a granularity of $3 \times 3\,\mathrm{cm}^2$ is achieved by applying the analogue software compensation algorithm. This resolution is comparable to the best energy resolution achieved for the $1 \times 1\,\mathrm{cm}^2$ AHCAL simulation with the semi-digital and analogue software compensation energy reconstruction.
The developed analogue software compensation algorithm is currently tested in the Pandora PFA with the ILD detector which includes a full size Fe-AHCAL. The study shows an absolute improvement in the single particle resolution of 3 % and in the jet energy resolution of around 0.5 %, achieving a jet energy resolution of 3 % for jet energies of 100 to 250 GeV [104]. This study uses simulations of the process $e^+e^- \to (Z/\gamma)^* \to q\bar{q}$, with the Z boson decaying into light quarks.

For the comparison with the CALICE AHCAL results, the Fe-DHCAL testbeam data is used. This data set was recorded at Fermilab in 2011 with beam energies of 2 to 60 GeV. Within

this analysis, the testbeam data was calibrated and simulated. The simulation was found to strongly be affected by the choice of the EM physics list of GEANT4. The sensitivity to the different models originates from the implementation of the digitiser and the signal saturation of the $1 \times 1\,\mathrm{cm}^2$ pads to the dense EM showers. However, positron showers of 2 to 25 GeV are successfully simulated with the _EMZ or "option 4" physics list [101]. The pion showers of 6 to 60 GeV show the best agreement with the simulations using the _EMY or "option 3" EM physics list, which is the option recommended by the GEANT4 working group [82], together with the hadronic models of FTFP_BERT and QGSP_BERT.
The comparison of the Fe-DHCAL response to positrons and pions, revealed an overcompensation that increases with the beam energy, which motivates the application of digital software compensation algorithms. As seen for the Fe-AHCAL which has a $e/\pi = 1.19$, an offline correction of the response has the power to improve the energy resolution by up to 45 %, see Section 5.8. An idea for such an algorithm is introduced in Section 4.5.2 and a first attempt achieved an improvement of up to 21 % in the energy resolution for pion showers [103].

The first comparison of the digital and semi-digital energy resolutions of the $1 \times 1\,\mathrm{cm}^2$ AHCAL simulation to the resolution achieved by the DHCAL and SDHCAL prototypes shows an advantage of the scintillator-tile calorimeter especially at low energies. Further investigations are needed to verify the origin of this behaviour.
A study of the sampling fluctuations in the A- and DHCAL would be possible by the so-called "two interleaved calorimeter" technique, introduced in [21]. By comparing the impacts of the sampling fluctuations on both calorimeter technologies, it would be possible to understand whether the different sampling fractions of the calorimeters are the reason for the better energy resolution, determined with the digital energy reconstruction, of the Sci-Fe AHCAL.
Another interesting study would be to decrease the cell sizes of the Fe-DHCAL in the simulation and study the impact on the saturation effects and the impact of the threshold in comparison to the data taken with a pad size of $1 \times 1\,\mathrm{cm}^2$.

The final choice of the hadron calorimeter technology for a future electron-positron linear collider experiment requires the validation of the capability of RPC based hadron calorimetry for PFAs. This is now possible using the, with testbeam data validated, simulation of the RPC based digital hadron calorimeter. The performance of the DHCAL can be investigated through its simulation within a full detector model for a linear electron-positron collider.

# Acronyms

**ADC** Analog-to-Digital Converter . . . 33
**AHCAL** Analogue Hadron Calorimeter . . . 123
**ALEPH** Apparatus for LEP Physics . . . 27
**APD** Avalanche Photodiode . . . 33
**ASIC** Application Specific Integrated Circuit . . . 33
**ATLAS** A Toroidal LHC ApparatuS . . . 9
**CERN** Conseil Européen pour la Recherche Nucléaire, Geneva, Switzerland . . . 43
**CLIC** Compact Linear Collider . . . 10
**CMS** Compact Muon Solenoid . . . 9
**DAC** Digital-to-Analog Converter . . . 33
**DESY** Deutsches Elektronen Synchrotron
**DHCAL** Digital Hadron Calorimeter . . . 35
**ECAL** Electromagnetic Calorimeter . . . 9
**EM** Electromagnetic . . . 18
**Fermilab** Fermi National Accelerator Laboratory, Batavia IL, USA . . . 40
**HARDROC** HAdronic Rpc Detector ReadOut Chip . . . 39
**HCAL** Hadronic Calorimeter . . . 9
**HEP** High Energy Physics . . . 9
**IL** Interaction Layer . . . 90
**ILC** International Linear Collider . . . 10
**ILD** International Large Detector . . . 25
**LAr** Liquid Argon . . . 24
**LED** Light-Emitting Diode . . . 33
**LEP** Large Electron Positron Collider . . . 27
**LHC** Large Hadron Collider . . . 9

**LHCb** Large Hadron Collider beauty

**MicroMEGaS** Micro-MEsh Gaseous Structure . . . . . 70

**MC** Monte Carlo simulation . . . . . 56

**MIP** Minimum Ionising Particle . . . . . 14

**MPPC** Multi Pixel Photon Counter . . . . . 30

**MPV** Most Probable Value . . . . . 35

**PCB** Printed Circuit Board . . . . . 30

**PDE** particle detection efficiency . . . . . 41

**PDG** Particle Data Group

**PFA** Particle Flow Algorithm . . . . . 9

**PID** Particle IDentification . . . . . 89

**PMT** Photomultiplier Tube . . . . . 18

**QCD** Quantum Chromodynamics . . . . . 10

**RMS** Root Mean Square . . . . . 49

**RPC** Resistive Plate Chamber . . . . . 35

**SDHCAL** Semi-Digital Hadron Calorimeter . . . . . 38

**SiPM** Silicon Photomultiplier . . . . . 30

**SPS** Super Proton Synchrotron . . . . . 38

**TCMT** Tail Catcher and Muon Tracker . . . . . 29

**TLHA** Three-Layer Hit Average . . . . . 90

**WLS** Wavelength-Shifting . . . . . 29

# Bibliography


[1] A. Benaglia. The CMS ECAL performance with examples. *Journal of Instrumentation*, 9(02):C02008, 2014.

[2] F. De Guio for the CMS collaboration. Performance of the CMS electromagnetic calorimeter and its role in the hunt for the Higgs boson in the two-photon channel. *Journal of Physics: Conference Series*, 455(1):012028, International Workshop on Discovery Physics at the LHC 2013.

[3] C. W. Fabjan and F. Gianotti. Calorimetry for Particle Physics. Technical report, CERN, 2003.

[4] S. Akhmadaliev et. al. Hadron energy reconstruction for the ATLAS calorimetry in the framework of the non-parametrical method. *Nuclear Instruments and Methods in Physics Research*, 2002.

[5] The CMS collaboration. Performance of the CMS hadron calorimeter with cosmic ray muons and LHC beam data. *Journal of Instrumentation*, 5(03):T03012, 2010.

[6] The CMS collaboration. Particle–Flow Event Reconstruction in CMS and Performance for Jets, Taus, and ETmiss. CMS PAS PFT-09-001, April 2009.

[7] K.A. Olive et al. (Particle Data Group). Review of Particle Physics. *Chinese Physics C*, 38(9):090001, 2014.

[8] H. Bichsel, D.E. Groom, S.R. Klein. Passage of particles through matters, 2015 `http://pdg.lbl.gov/2015/reviews/rpp2015-rev-passage-particles-matter.pdf`.

[9] L. Landau. On the energy loss of fast particles by ionization. *J.Phys.(USSR)*, vol. 8:pp. 201205, 1944.

[10] R. Wigmans. *Calorimetry*. Oxford Science Publications, 2000.

[11] 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, June 2013.

[12] R. Wigmans. New results from the RD52 project. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 824:721 – 725, 2016. Frontier Detectors for Frontier Physics: Proceedings of the 13th Pisa Meeting on Advanced Detectors.

[13] R. Wigmans. The DREAM Project – Results and plans. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 572(1):215 – 217, 2007. Frontier Detectors for Frontier Physics: Proceedings of the 10th Pisa Meeting on Advanced Detectors.

[14] G. Gaudio, J. Hauptmann, R. Wigmans for the RD52 (DREAM) collaboration. Dual-Readout Calorimetry for High-Quality Energy Measurements. Technical report, CERN-SPSC-2014-010 / SPSC-SR-135, March 2014.

[15] E. Longo, I. Sestili. Monte Carlo Calculation of Photon Initiated Electromagnetic Showers in Lead Glass. *Nuclear Instruments and Methods in Physics Research*, 128(283), 1975.

[16] PDG Atomic and nuclear properties of materials, `pdg.lbl.gov/2011/AtomicNuclearProperties/`.

[17] T.A. Gabriel, D.E. Groom, P.K. Job, N.V. Mokhov, G.R. Stevenson. Energy dependence of hadronic activity. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 338(2):336 – 347, 1994.

[18] R.K. Bock, T. Hansl-Kozanecka, T.P. Shah. Parametrization of the longitudinal development of hadronic showers in sampling calorimeters. *Nuclear Instruments and Methods in Physics Research*, 186(3):533 – 539, 1981.

[19] M. Chadeeva. Parametrisation of hadron shower profiles in the CALICE Sc-Fe AHCAL. CALICE Analysis Note CAN-48, April 2014.

[20] R. Wigmans M. Livan, V. Vercesi. Scintillating-fibre Calorimetry. Cern 95-02, CERN Yellow Report, Geneva, Switzerland, 1995.

[21] G. Drews et al. Experimental determination of sampling fluctuations in uranium and lead hadronic calorimeters. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 290(2):335 – 345, 1990.

[22] D. Buskulic et al. (ALEPH collaboration). Performance of the ALEPH detector at LEP. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 360(3):481 – 506, 1995.

[23] M.A. Thomson. Particle flow calorimetry and the PandoraPFA algorithm. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 611(1):25 – 40, 2009.

[24] F. Beaudette for the CMS collaboration. The CMS Particle Flow Algorithm. *International Conference on Calorimetry for the High Energy Frontier (CHEF 2013)*, arXiv:1401.8155, April 2013.

[25] U. Holm et al. (ZEUS collabration). The ZEUS Detector. Technical Report Volume 1, DESY, February 1993.

[26] T. Behnke, C. Damerell, J. Jaros, A. Myamoto et al. ILC Reference Design Report Volume 4 - Detectors. Technical report, ILC Global Design Effort and World Wide Study, August 2007.

[27] C. Adloff et al. (CALICE collaboration). Tests of a Particle Flow Algorithm with CALICE test beam data. *Journal of Instrumentation*, 6(07):P07005, 2011.

[28] International Linear Collider (ILC) collaboration `https://www.linearcollider.org`.

[29] The ILC collaboration. Technical Design Report. Technical Report Volume 4, International Linear Collider, 2013.

[30] Compact Linear Collider (CLIC) collaboration `http://clic-study.web.cern.ch`.

[31] The CLIC collaboration. CLIC Conceptional Design Report - Physics and Detectors at CLIC. Technical Report Volume 2, 274th CERN SPC, 2011.

[32] D. Contardo, M. Klute, J. Mans, L. Silvestris, J. Butler. Technical Proposal for the Phase-II Upgrade of the CMS Detector. CERN-LHCC-2015-010; LHCC-P-008; CMS-TDR-15-02 ISBN 978-92-9083-417-5, June 1 2015.

[33] K. Seidel. *Top quark pair production and calorimeter energy resolution studies at a future collider experiment.* PhD thesis, Ludwig–Maximilians–Universität München, 2012.

[34] M. Gabriel. Energy reconstruction for hadrons in a highly granular calorimeter system. Master's thesis, Technische Universität München, 2014.

[35] O. Hartbrich. Pion Response and Resolution in a Combined Scintillator Calorimeter System in the FNAL Testbeam. CALICE Analysis Note CAN-56, May 2016.

[36] The International Linear Collider. Technical design report. Technical Report Volume 1, International Linear Collider, 2013.

[37] The CLIC collaboration. CLICdet: The post-CDR CLIC detector model. CLICdp-Draft-2016-XY, to be published, current version: `https://edms.cern.ch/ui/#!master/navigator/document?D:1092704155:1092704155:subDocs` 2016.

[38] K. Francis et al. (CALICE collaboration). Performance of the first prototype of the CALICE scintillator strip electromagnetic calorimeter. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 763:278 – 289, 2014.

[39] C. Adloff et al. (CALICE collaboration). Response of the CALICE Si-W electromagnetic calorimeter physics prototype to electrons. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 608(3):372 – 383, 2009.

[40] K. Kotera. First Stage Analysis of the Energy Response and Resolution of the Scintillator ECAL in the Beam Test at FNAL, 2008. CALICE Analysis Note CAN-16, May 2010.

[41] V. Morgunov and A. Raspereza. Novel 3-D clustering algorithm and two particle separation with tile HCAL. *Proceedings of the International Conference on Linear Colliders LCWS 2004*, April 2004.

[42] N. Feege. *Low-energetic Hadron Interactions in a Highly Granular Calorimeter*. Phd thesis, DESY-THESIS-2011-048, Universität Hamburg, 2011.

[43] C. B. Lam. Interactions of Particles with Momenta of 1-10 GeV in a Highly Granular Hadronic Calorimeter with Tungsten Absorbers. Master's thesis, Faculteit Technische Natuurwetenschappen Universiteit Twente, 2012.

[44] C. Adloff et al. (CALICE collaboration). Construction and commissioning of the CALICE analog hadron calorimeter prototype. *Journal of Instrumentation*, 5(05):P05004, 2010.

[45] M. Danilov for the CALICE collaboration. Scintillator tile hadron calorimeter with novel SiPM readout. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 581(A582):451 – 456, 2007. Proceedings of the 11th International Vienna Conference on Instrumentation.

[46] A. Lucaci-Timoce for the CALICE collaboration. Calibration and monitoring of a scintillator HCAL with SiPMs CALICE scintillator HCAL. *Journal of Physics: Conference Series*, 160(1):012027, XIII International Conference on Calorimetry in High Energy Physics (CALOR 2008) 2009.

[47] K. Krüger for the CALICE collaboration. Prototype tests for a highly granular scintillator-based hadron calorimeter. *Journal of Physics: Conference Series*, 587(1):012033, 16th International Conference on Calorimetry in High Energy Physics (CALOR 2014) 2015.

[48] J. Repond et al. (CALICE collaboration). Design and electronics commissioning of the physics prototype of a Si-W electromagnetic calorimeter for the International Linear Collider. *Journal of Instrumentation*, 3(08):P08001, 2008.

[49] C. Adams et al. Design, Construction and Commissioning of the Digital Hadron Calorimeter - DHCAL. arXiv:1603.01653, March 2016.

[50] B. Bilki et al. Calibration of a digital hadron calorimeter with muons. *Journal of Instrumentation*, 3(05):P05001, 2008.

[51] D. Underwood G. Drake, J. Repond and L. Xia. Resistive Plate Chambers for hadron calorimetry: Tests with analog readout. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 578(1):88 – 97, 2007.

[52] J. Repond. Analysis of DHCAL Muon Data. CALICE Analysis Note CAN-30, March 2011.

[53] V. Ammosov et al. RPC as a Detector for High Granularity Digital Hadron Calorimetry. *DESY 04-057*, March 2004.

[54] L. Benussi et al. Properties of potential eco-friendly gas replacements for particle detectors in high-energy physics. *INFN-14-13/LNF CERN-OPEN-2015-004*, May 2015.

[55] J. Koppitz. Image Converter and Intensiner Investigations on Luminous Fronts (Ionizing Waves) in Nitrogen Discharges in a Homogeneous Field. *Z. Naturforsch.*, 26a(26):700–706, December 1971.

[56] M. Abbrescia et al. The simulation of resistive plate chambers in avalanch mode: charge spectra and efficiency. *Nuclear Instruments and Methods in Physics Research*, A(431):413–427, 1999.

[57] H. Genz. Single electron detection in proportional gas counters. *Nuclear Instruments and Methods in Physics Research*, 112 (1973):83–90, 1973.

[58] S. Ramo. Currents Induced by Electron Motion. *Proc. IRE*, 27:584, 1939.

[59] W. Shockley. Currents to Conductors Induced by a Moving Point Charge. *Journal of Applied Physics*, 9, October 1938.

[60] C. Lippmann. *Detector Physics of Resistive Plate Chambers*. PhD thesis, Johann Wolfgang Goethe-Universität in Frankfurt am Main, 2003.

[61] J. Repond for the CALICE collaboration. Calibration of the DHCAL with Muons. *XI workshop on Resistive Plate Chambers and Related Detectors (RPC2012)*, PoS (2012)(RPC2012):076, 5-10 February 2012.

[62] B. Bilki et al. (CALICE collaboration). Pion and proton showers in the CALICE scintillator-steel analogue hadron calorimeter. *Journal of Instrumentation*, 10(04):P04014, 2015.

[63] B. Bilki for the CALICE collaboration. The CALICE digital hadron calorimeter: calibration and response to pions and positrons. *Journal of Physics: Conference Series*, 587(1):012038, 16th International Conference on Calorimetry in High Energy Physics (CALOR 2014) 2015.

[64] G. Baulieu et al. Construction and commissioning of a technological prototype of a high-granularity semi-digital hadronic calorimeter. *Journal of Instrumentation*, 10(10):P10039, 2015.

[65] Z. Deng et al. (CALICE collaboration). First results of the CALICE SDHCAL technological prototype. *Journal of Instrumentation*, 11(04):P04001, 2016.

[66] C. Adloff et al. Construction and test of a $1\times1$ $m^2$ Micromegas chamber for sampling hadron calorimetry at future lepton colliders. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 729:90 – 101, 2013.

[67] H. Ran. Induced charge signal of a glass RPC detector. *Chinese Physics C*, 38(4):046002, 2014.

[68] M. Affatigato et al. Measurements of the rate capability of various Resistive Plate Chambers. *Journal of Instrumentation*, 10(10):P10037, 2015.

[69] B. Freund et al. (CALICE collaboration). DHCAL with minimal absorber: measurements with positrons. *Journal of Instrumentation*, 11(05):P05008, 2016.

[70] KETEK Silicon Photomultipliers. `http://www.ketek.net/products/sipm/`, 2016.

[71] C. Adloff et al. (CALICE collaboration). Hadronic energy resolution of a highly granular scintillator-steel hadron calorimeter using software compensation techniques. *Journal of Instrumentation*, 7(P09017), September 2012.

[72] B. Bilki. DHCAL Response to Positrons and Pions. CALICE Analysis Note CAN-32, March 2011.

[73] T. D. Strickler. Average energy loss per ion pair produced by alpha particles in binary gas mixtures. *ORNL-3080*, Contract NO. W-7405-eng-26, 1961.

[74] F. H. Attix. *Introduction to Radiological Physics and Radiation Dosimetry*. ISBN: 978-0-471-01146-0. Wiley, 1986.

[75] G. W. Clark, F. Scherb, W. B. Smith. Preparation of Large Plastic Scintillators. *Review of Scientific Instruments*, 28(433), 1957 doi: 10.1063/1.1715900.

[76] C. Adloff et al. (CALICE collaboration). Validation of GEANT4 Monte Carlo models with a highly granular scintillator-steel hadron calorimeter. *Journal of Instrumentation*, 8(07):P07005, July 2013.

[77] C. Adloff et al. (CALICE collaboration). Construction and performance of a silicon photomultiplier/extruded scintillator tail-catcher and muon-tracker. *Journal of Instrumentation*, 7(04):P04015, 2012.

[78] V. N. Ivanchenko, O. Kadri, M. Maire and L. Urban. GEANT4 models for simulation of multiple scattering. *Journal of Physics: Conference Series*, 219(3):032045, 2010.

[79] J. Apostolakis et al. Progress in GEANT4 Electromagnetic Physics Modelling and Validation. *Journal of Physics: Conference Series*, 664(7):072021, 2015.

[80] `http://geant4.cern.ch/support/proc_mod_catalog/processes/`.

[81] A. Dotti. Recent GEANT4 Developments, CALICE Collaboration Meeting in Argonne `https://agenda.linearcollider.org/event/6341/contributions/30037/\attachments/24865/38356/Geant4-Summary-ForCALICE-Dotti-Wright.pdf`. March 21th 2014.

[82] A. Dotti (GEANT4 working group @ SLAC). Private communication.

[83] `http://geant4.cern.ch/support/proc_mod_catalog/physics_lists/useCases.shtml`.

[84] CALICE collaboration and ILD concept group. Calibration of the Scintillator Hadron Calorimeter of ILD. CALICE Analysis Note CAN-18, October 2009.

[85] C. Adloff et al. (CALICE collaboration). Electromagnetic response of a highly granular hadronic calorimeter. *Journal of Instrumentation*, 6(04):P04003, 2011.

[86] C. Neubüser, K. Krüger. Analogue, Digital and Semi-Digital Energy Reconstruction in the CALICE AHCAL. CALICE Analysis Note CAN-49, May 2014.

[87] C. Neubüser. Comparison of Energy Reconstruction Schemes and Different Granularities in the CALICE AHCAL. CALICE Analysis Note CAN-49a, May 2016.

[88] C. Neubüser for the CALICE collaboration. Analogue, Digital and Semi-Digital Energy Reconstruction in the CALICE AHCAL. *3rd International Conference on Technology and Instrumentation in Particle Physics (TIPP 2014)*, (PoS TIPP2014 (2014)), June 2-6 2014.

[89] C. Neubüser for the CALICE collaboration. Comparison of Energy Reconstruction Schemes and Different Granularities in the CALICE AHCAL. *Talk at CALOR 2016 – XVIIth International Conference on Calorimtery in Particle Physics*, proceedings to be published, May 15-20 2016.

[90] B. Lutz. *Hadron Showers in a Highly Granular Calorimeter*. Phd thesis, DESY-THESIS-2010-048, Universität Hamburg, 2010.

[91] D. Boumediene for the CALICE collaboration. Response of the CALICE Si-W ECAL Physics Prototype to electrons. *Journal of Physics: Conference Series*, 160(1):012065, XIII International Conference on Calorimetry in High Energy Physics (CALOR 2008) 2009.

[92] C. Günter. *Comparison of Iron and Tungsten Absorber Structures for an Analog Hadron Calorimeter*. Phd thesis, DESY-THESIS-2015-018, Univeristät Hamburg, 2015.

[93] M. Chefdeville. Off-line compensation of a SDHCAL, a Monte Carlo study, CALICE Collaboration Meeting in Hamburg `http://agenda.linearcollider.org/event/5947/contributions/27139/`. 20-22 March 2013.

[94] M. A. Thomson J. S. Marshall. Pandora Particle Flow Algorithm. *Calorimetry for High Energy Frontiers - CHEF 2013*, arXiv:1308.4537v1, 2013.

[95] S. Green. Detector Optimisation Studies using the Pandora Particle Flow Algorithm, CLIC Workshop 2016 `https://indico.cern.ch/event/449801/contributions/1945361/\attachments/1213774/1771666/CLICWS_DetOpt_Green_19-1-16.pdf`. January 2016.

[96] B. Bilki. The DHCAL Results from Fermilab Beam Tests: Calibration. CALICE Analysis Note CAN-42, May 2013.

[97] C. G. Reichelt. Particle ID Studies in a Highly Granular Hadron Calorimeter. Cern summer student program, CERN - European Organization for Nuclear Research, Switzerland, September 2013.

[98] C. Grefe. W-DHCAL Calibration, CALICE Collaboration Meeting in Annecy `http://agenda.linearcollider.org/event/6050/contributions/28159/\attachments/23352/36289/2013-09-10_CALICE-Workshop.pdf`. 9-11th September 2013.

[99] B. Freund. The Digital Hadron Calorimeter: Analysis of the November 2011 Fermilab Data. Master's thesis, McGill University, July 2015.

[100] B. Bilki (University of Iowa). Private communication.

[101] https://geant4.web.cern.ch/geant4/collaboration/working_groups/electromagnetic/physlist10.1.shtml.

[102] L. Xia. CALICE DHCAL Noise Analysis. CALICE Analysis Note CAN-31, March 2011.

[103] C. Neubüser. Fe-DHCAL: Software Compensation,
Asian Linear Collider Workshop (ALCW2015)
https://agenda.linearcollider.org/event/6557/contributions/31752/\attachments/26182/40131/20150420_CaliceMeetingKEK_SCforDHCAL.pdf.
April 19-24th 2015.

[104] H. Lan Tran. Private communication, July 2016.

# Acknowledgments *Danksagung*

The analysis presented in this thesis has only been possible by close cooperation with a lot of institutes and excellent physicists. The studied data sets have been recorded in testbeams, running around the clock, requiring a lot of men power. I can not mention all the people that have been involved but want to thank the CALICE collaboration as a whole that coordinates the efforts of many institutes and allows a natural competition and supports interesting research.
Personally I want to thank my supervisors: Katja, Erika, and Jose, who are responsible for the success of this project and the reason for the extension of my knowledge and my horizon. I want to thank my mentors Kerstin and Juliette for the support and advice during stressful times.
Additionally I want to thank the groups that have welcome me at the Argonne National Laboratory and CERN. Specifically: Burak, Lei and Benjamin as well as Christian, André and Eva. I am very grateful for your support and patience.
The FLC group at DESY has given me the great opportunity to work with amazing people, many of which I can call friends.

*Zum Schluss möchte ich mich bei meinem Lunch Stammtisch: Moritz, Eva, Janni und Christian bedanken für die seit dem Studium andauernde Freundschaft und die schönen Mit- und Nachmittage. Meine liebste Kollegin und Freundin Claude, mit der ich mir seit Beginn des PhD das Büro teile, hat mir mit Zuspruch und viel Spaß durch viele lange Arbeitstage geholfen.*
*Ich möchte mich bei meiner Familie bedanken die mir immer den Rückhalt und die Liebe gegeben hat um meine Ziele zu erreichen.*
*Meinen beiden besten Freundinnen, Annette und Mama, möchte ich danken für die schönen Töpferabende und Abenteuer die wir zusammen erlebt haben und noch erleben werden.*
*Last but not least, Matteo, grazie mille per il tuo sostegno amore mio.*

# Appendix A

# ECAL Contribution to the Energy Reconstruction

Usually the ECAL contribution to the reconstructed energy is calculated by

$$E_{ECAL} = \sum_{k=1}^{3} \nu_k \cdot M_{ECAL,k} \tag{A.1}$$

with $M_{ECAL,k}$ is the energy sum in the ECAL layers with in region k with sampling fraction $\nu_k$. In this study, for the events with MIP-like tracks in the ECAL, the $E_{ECAL}$ was reconstructed with an average conversion factor, taken from [71], of $\sum_{k=1}^{3} \nu_k/3 = 0.005906\,\mathrm{GeV/MIP}$. Thus the reconstructed energy is given by

$$E_{ECAL} = 0.005906 \cdot \sum_{k=1}^{3} M_{ECAL,k}. \tag{A.2}$$

The resulting $E_{ECAL}$ distribution for all selected pion events, with MIP like tracks in the ECAL, summed up over all energies is shown in Figure A.1.
The mean value of the energy deposited in the ECAL, 0.3232 GeV, is used in the analysis.

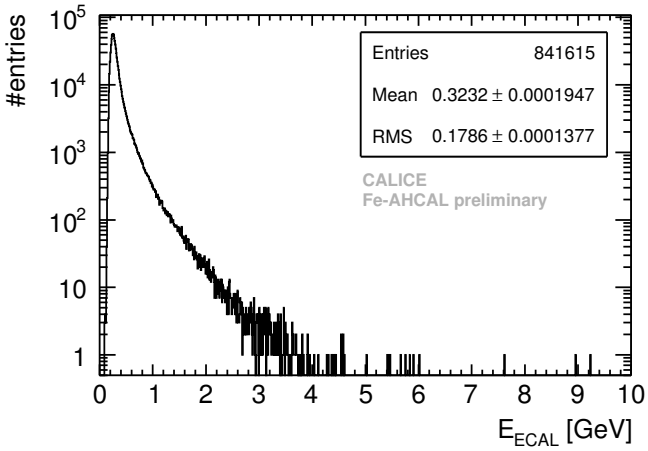
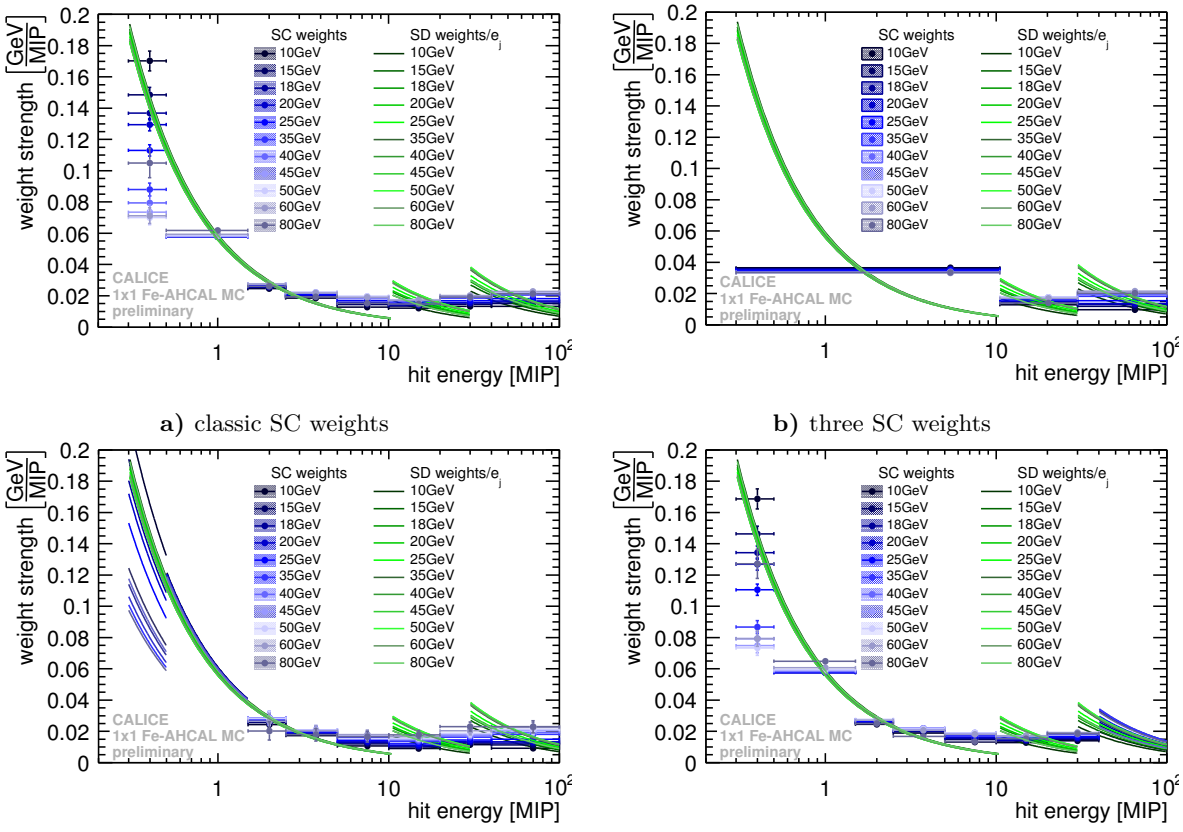



**Figure A.1:** Distribution of the reconstructed energy in the ECAL for selected events with track in ECAL for all runs and energies.

# Appendix B

# Software Compensation Weights




**a)** classic SC weights

**b)** three SC weights

**c)** two digital SC weights

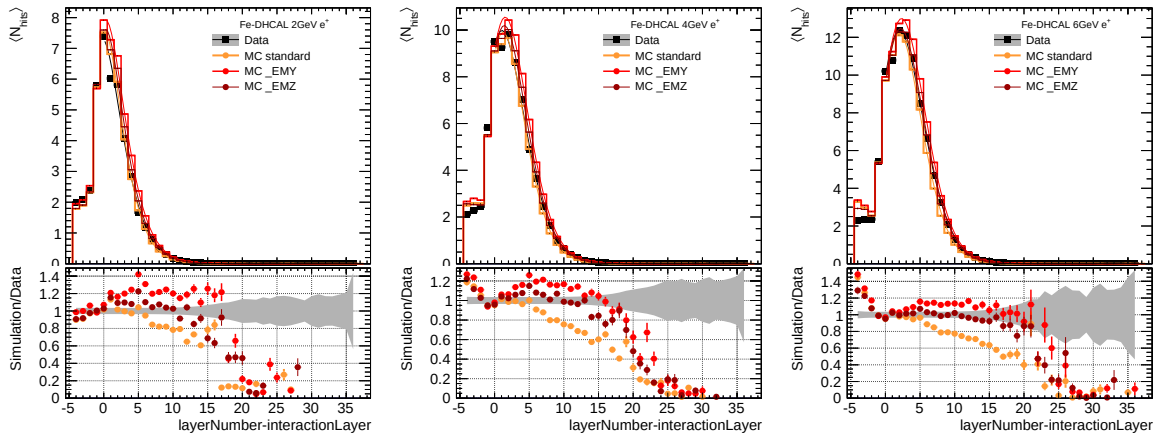
**d)** digital highest SC weight

**Figure B.1:** The weights used for the software compensation compared to the weights of the semi-digital energy reconstruction in the $1 \times 1\,\mathrm{cm}^2$ AHCAL FTFP_BERT simulation are shown as a function of the hit energy for the beam energies of 10-80 GeV. The software compensation weights are shown as blue bars/lines covering a certain hit energy range, the semi-digital weights are shown in green lines following a $1/e_j$ behaviour.

# Appendix C

# $e^+$ Shower Profiles

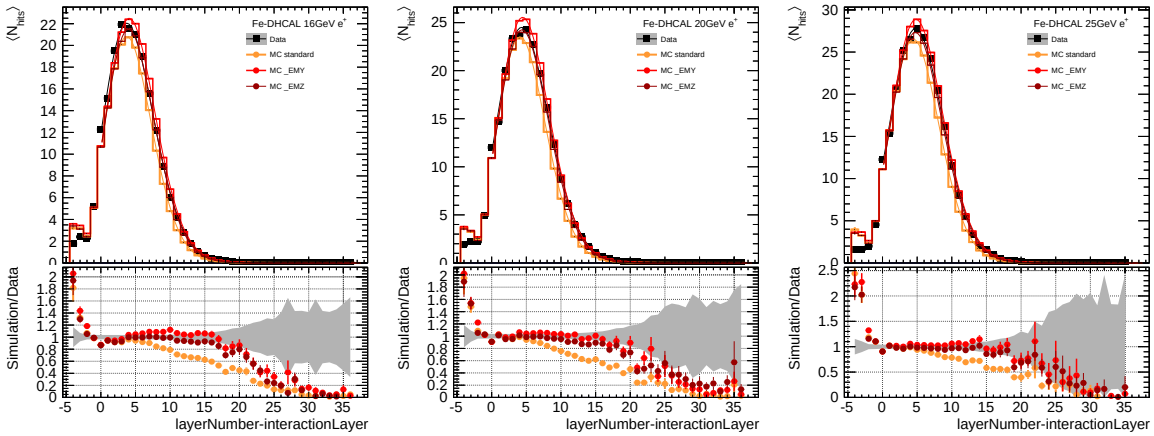
## C.1 Longitudinal Shower Profiles




**a)** 2 GeV $e^+$ **b)** 4 GeV $e^+$ **c)** 6 GeV $e^+$

**d)** 8 GeV $e^+$ **e)** 10 GeV $e^+$ **f)** 12 GeV $e^+$

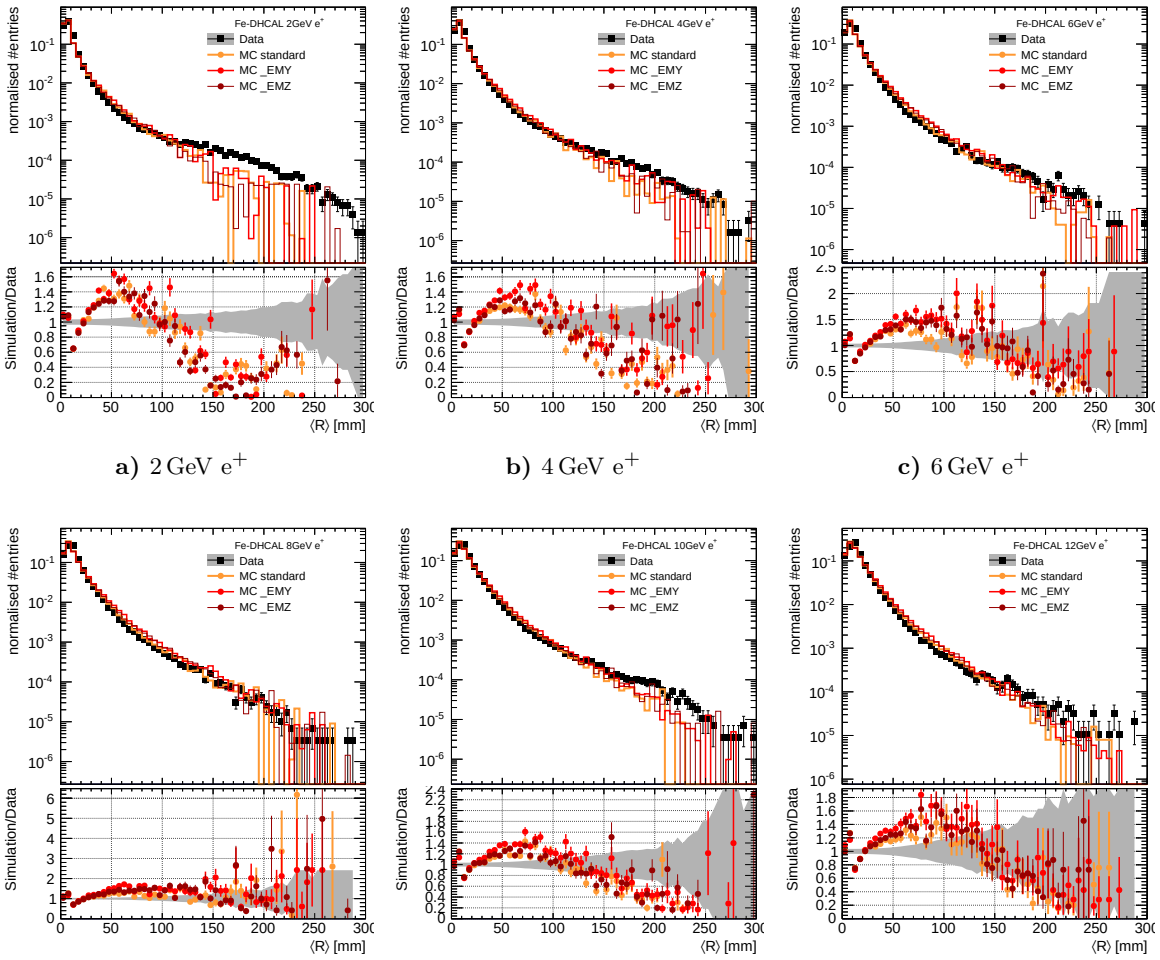
**Figure C.1:** The longitudinal shower profiles for 2 to 12 GeV $\mathbf{e^+}$ compared to the simulations using the standard, _EMY, and _EMZ EM physics lists.




**a)** 16 GeV $e^+$ **b)** 20 GeV $e^+$ **c)** 25 GeV $e^+$

**Figure C.2:** The longitudinal shower profiles for 16 to 25 GeV $\mathbf{e^+}$ compared to the simulations using the standard, _EMY, and _EMZ EM physics lists.

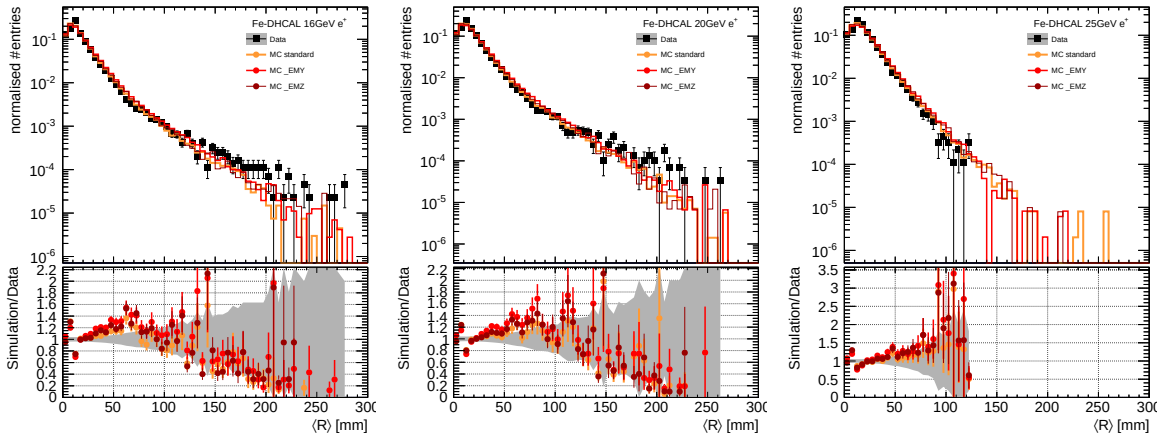
## C.2 Radial Shower Shapes




**a)** 2 GeV $e^+$ **b)** 4 GeV $e^+$ **c)** 6 GeV $e^+$

**d)** 8 GeV $e^+$ **e)** 10 GeV $e^+$ **f)** 12 GeV $e^+$

**Figure C.3:** The radial shower shapes for 2 to 12 GeV $\mathbf{e^+}$ compared to the simulations using the standard, _EMY, and _EMZ EM physics lists.



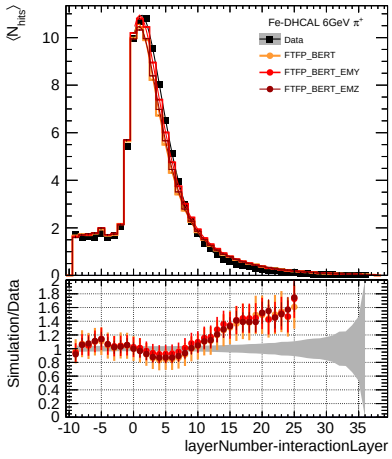

**a)** 16 GeV $e^+$ **b)** 20 GeV $e^+$ **c)** 25 GeV $e^+$

**Figure C.4:** The radial shower shapes for 16 to 25 GeV $\mathbf{e^+}$ compared to the simulations using the standard, _EMY, and _EMZ EM physics lists.

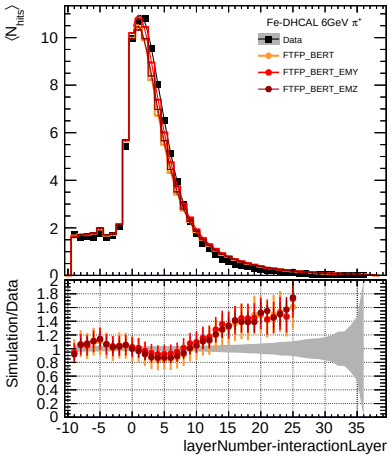
# Appendix D

# $\pi^+$ Shower Profiles

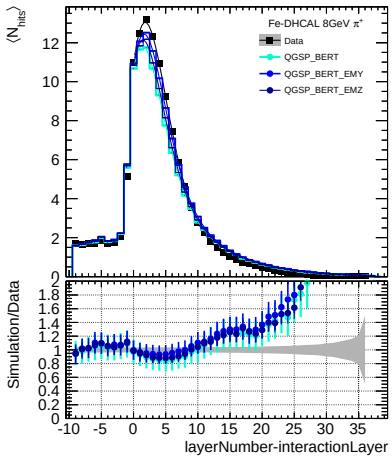
## D.1 Longitudinal Shower Profiles



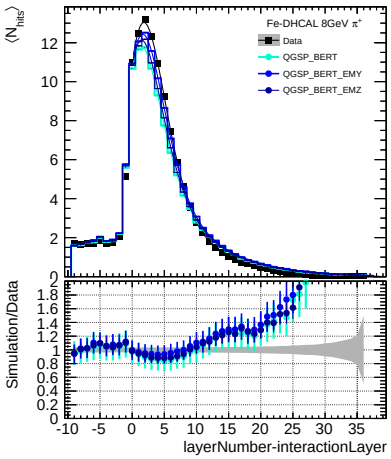

**a)** 6 GeV $\pi^+$ FTFP_BERT




**b)** 8 GeV $\pi^+$ FTFP_BERT

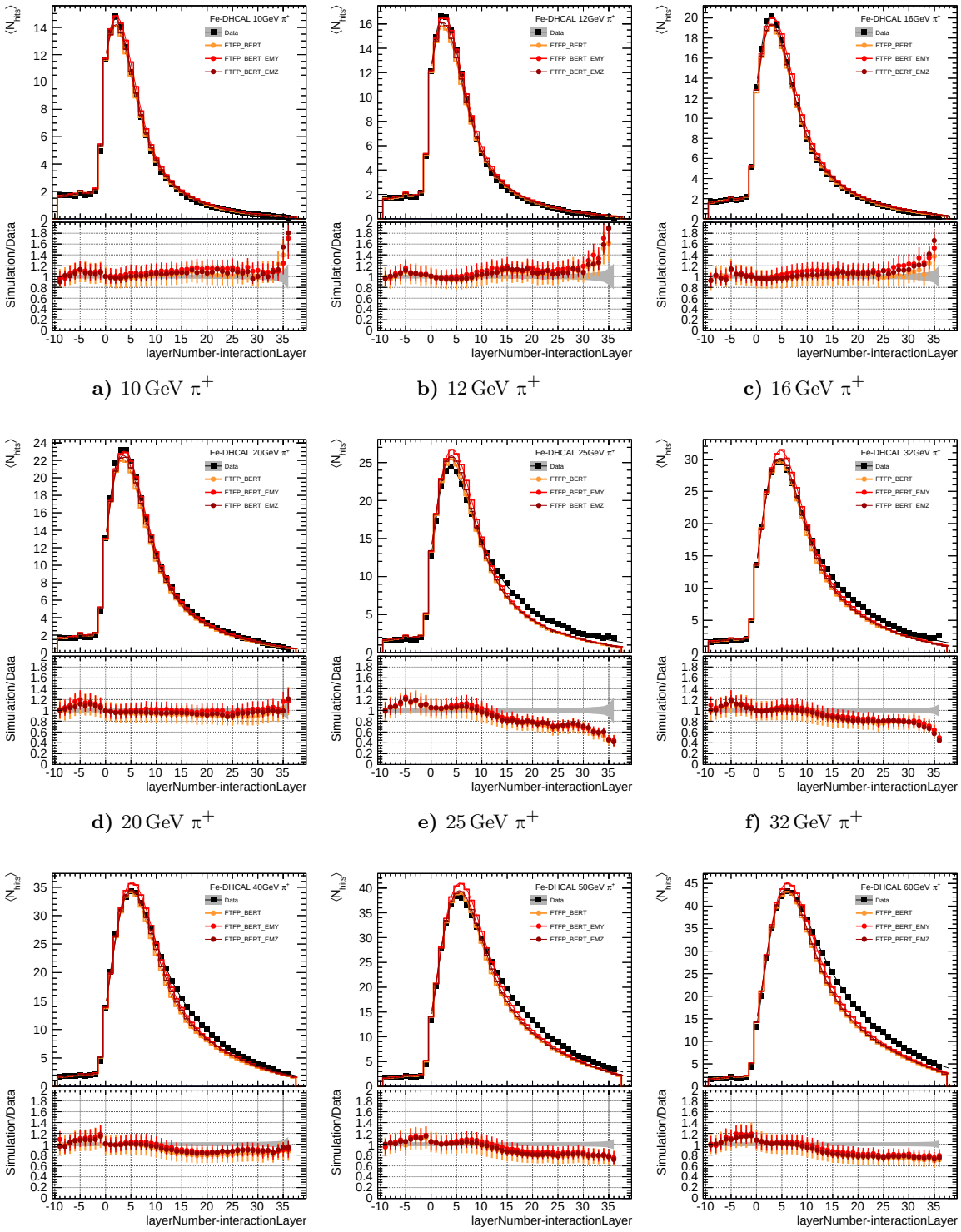



**c)** 6 GeV $\pi^+$ QGSP_BERT




**d)** 8 GeV $\pi^+$ QGSP_BERT

**Figure D.1:** The longitudinal shower profiles for 6 and 8 GeV $\pi^+$.

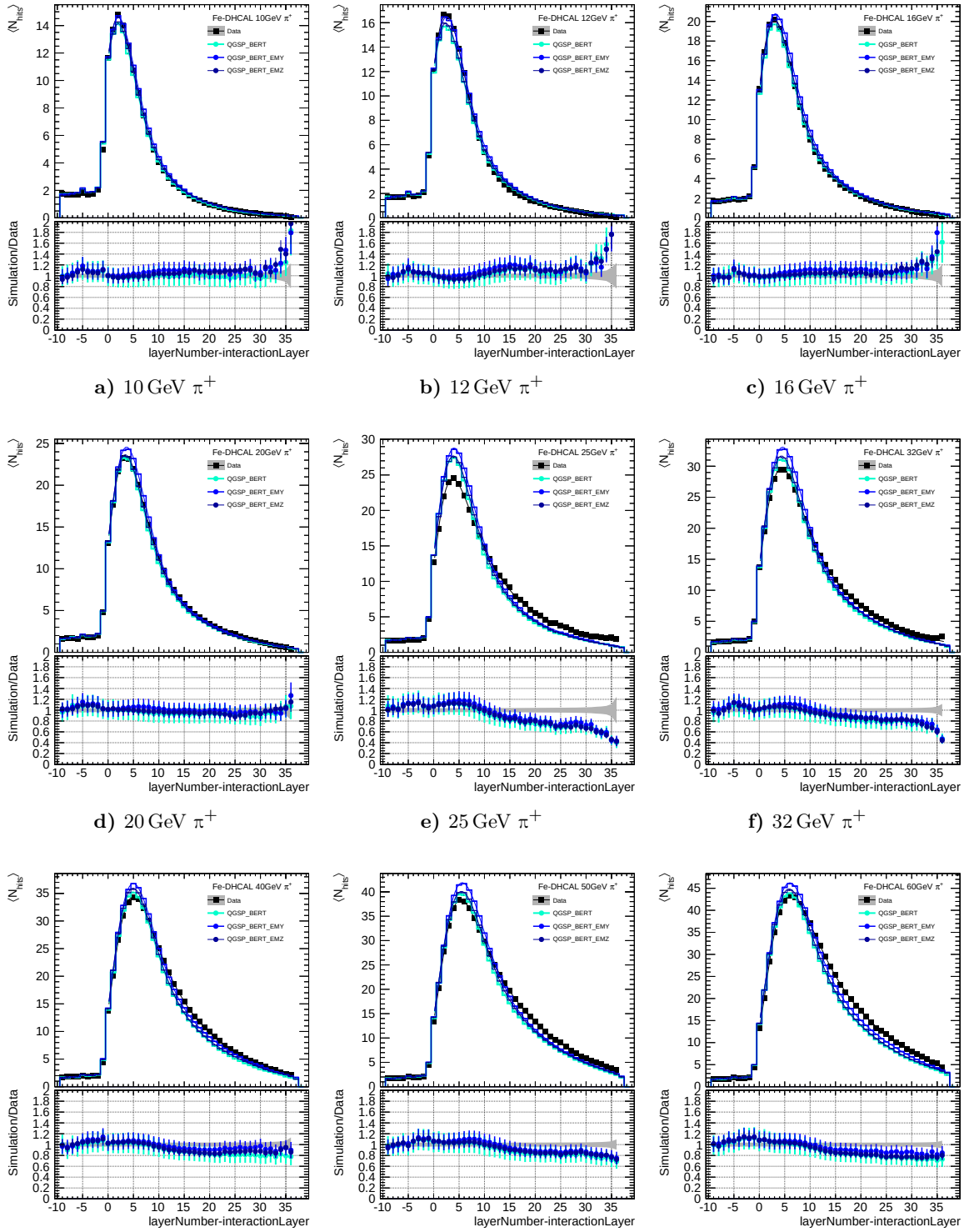



**a)** 10 GeV $\pi^+$ **b)** 12 GeV $\pi^+$ **c)** 16 GeV $\pi^+$

**d)** 20 GeV $\pi^+$ **e)** 25 GeV $\pi^+$ **f)** 32 GeV $\pi^+$

**g)** 40 GeV $\pi^+$ **h)** 50 GeV $\pi^+$ **i)** 60 GeV $\pi^+$

**Figure D.2:** The longitudinal shower profiles for 10 to 60 GeV $\boldsymbol{\pi^+}$ compared to the FTFP_BERT simulations.



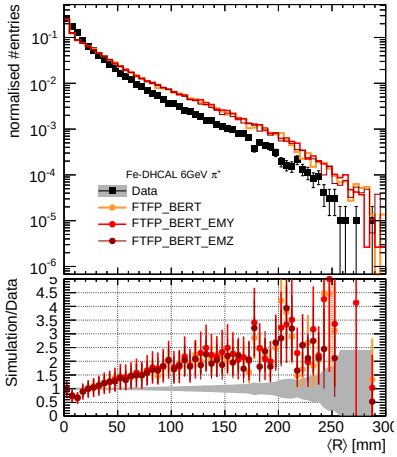

**a)** 10 GeV $\pi^+$ **b)** 12 GeV $\pi^+$ **c)** 16 GeV $\pi^+$

**d)** 20 GeV $\pi^+$ **e)** 25 GeV $\pi^+$ **f)** 32 GeV $\pi^+$

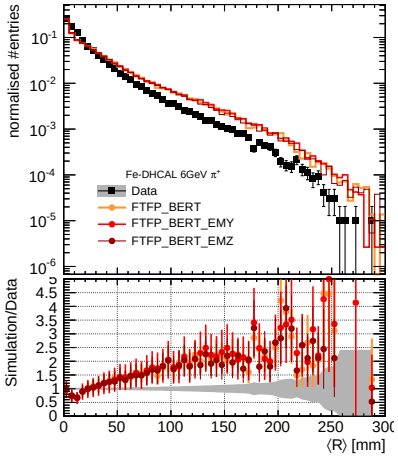
**g)** 40 GeV $\pi^+$ **h)** 50 GeV $\pi^+$ **i)** 60 GeV $\pi^+$

**Figure D.3:** The longitudinal shower profiles for 10 to 60 GeV $\pi^+$ compared to the QGSP_BERT simulations.

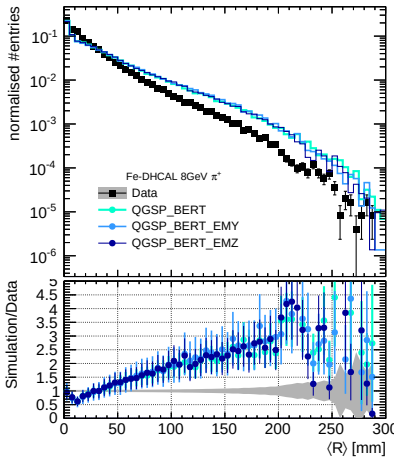
## D.2 Radial Shower Shapes



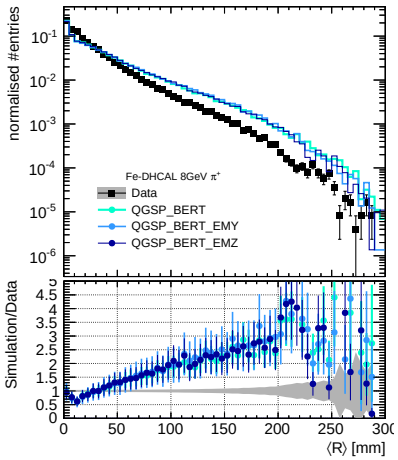

**a)** 6 GeV $\pi^+$ FTFP_BERT




**b)** 8 GeV $\pi^+$ FTFP_BERT

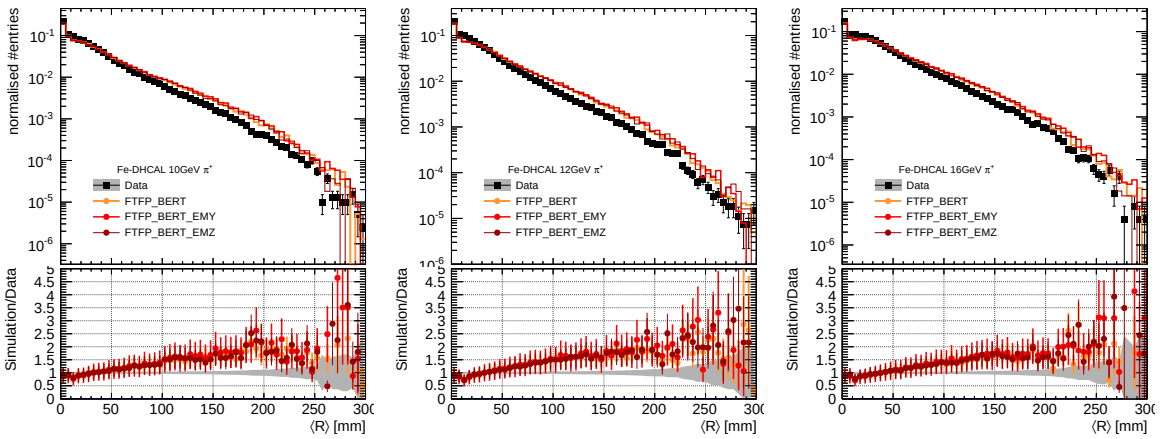



**c)** 6 GeV $\pi^+$ QGSP_BERT

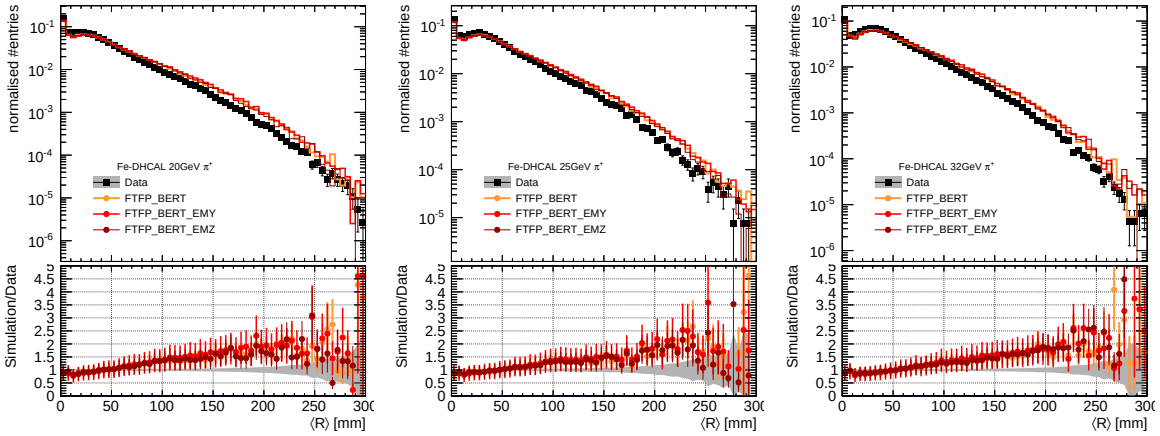



**d)** 8 GeV $\pi^+$ QGSP_BERT

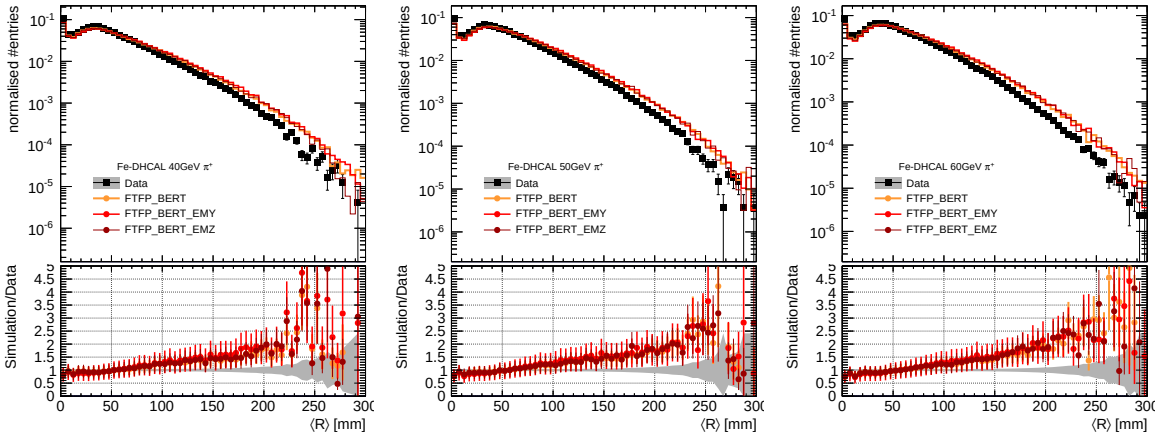
**Figure D.4:** The radial shower shapes for 6 and 8 GeV $\pi^+$.




**a)** 10 GeV $\pi^+$ **b)** 12 GeV $\pi^+$ **c)** 16 GeV $\pi^+$

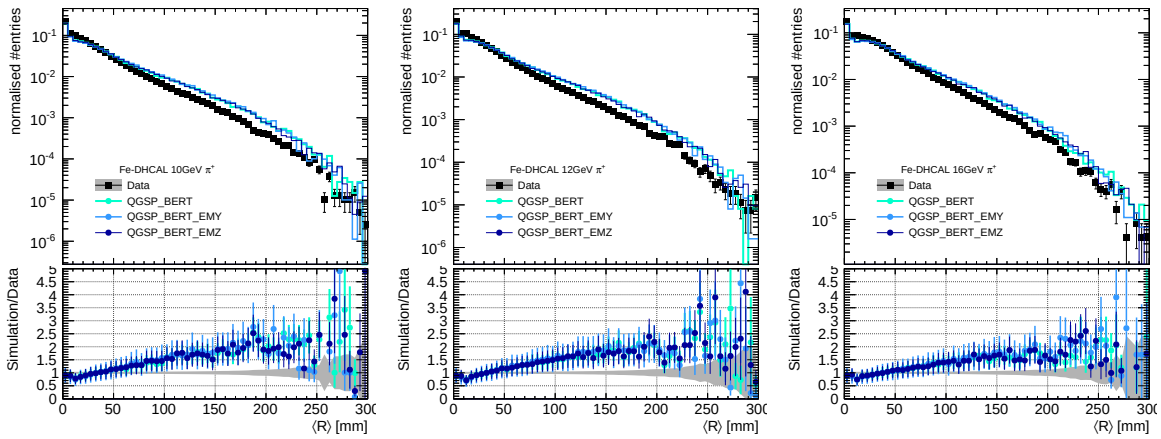



**d)** 20 GeV $\pi^+$ **e)** 25 GeV $\pi^+$ **f)** 32 GeV $\pi^+$

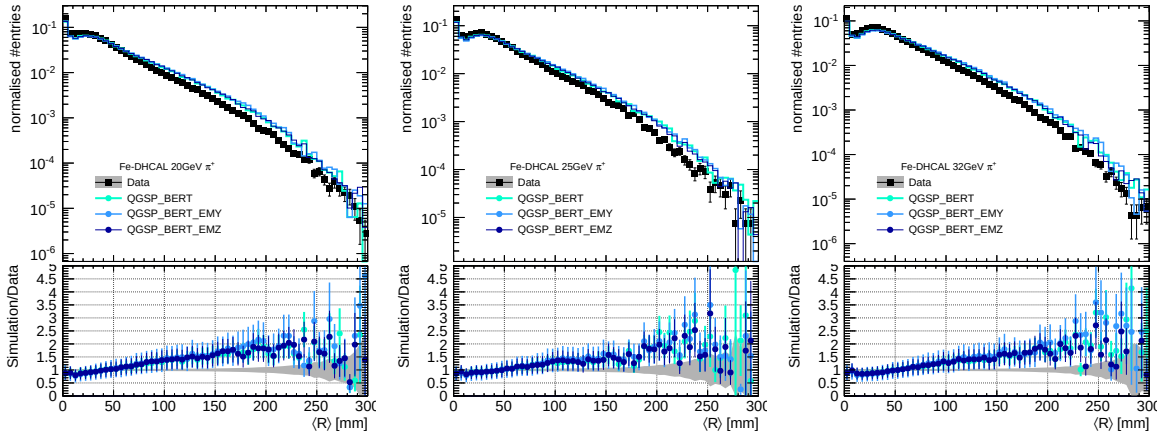



**g)** 40 GeV $\pi^+$ **h)** 50 GeV $\pi^+$ **i)** 60 GeV $\pi^+$

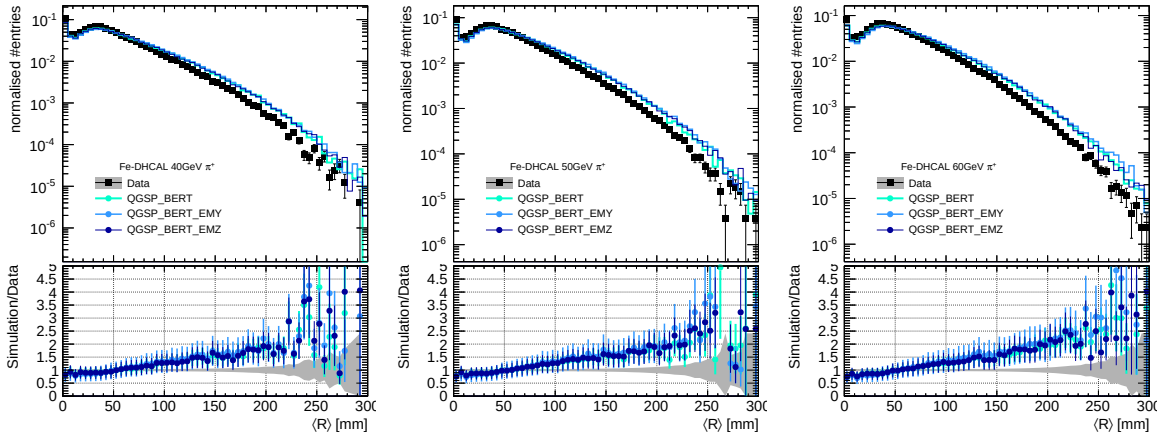
**Figure D.5:** The radial shower shapes for 10 to 60 GeV $\pi^+$ compared to the FTFP_BERT simulations.




**a)** 10 GeV $\pi^+$ **b)** 12 GeV $\pi^+$ **c)** 16 GeV $\pi^+$




**d)** 20 GeV $\pi^+$ **e)** 25 GeV $\pi^+$ **f)** 32 GeV $\pi^+$




**g)** 40 GeV $\pi^+$ **h)** 50 GeV $\pi^+$ **i)** 60 GeV $\pi^+$

**Figure D.6:** The radial shower shapes for 10 to 60 GeV $\pi^+$ compared to the QGSP_BERT simulations.

Eidesstattliche Versicherung *Declaration on oath*
Hiermit erkläre ich an Eides statt, dass ich die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.
*I hereby declare, on oath, that I have written the present dissertation by my own and have not used other than the acknowledged resources and aids.*

Coralie Neubüser
Hamburg, November 21, 2016