

Validation of GEANT4 Monte Carlo Models with a Highly Granular Scintillator-Steel Hadron Calorimeter

The CALICE Collaboration

**C. Adloff, J. Blaha, J.-J. Blaising, C. Drancourt, A. Espargilière, R. Gaglione,
N. Geffroy, Y. Karyotakis, J. Prast, G. Vouters**

*Laboratoire d'Annecy-le-Vieux de Physique des Particules, Université de Savoie, CNRS/IN2P3, 9
Chemin de Bellevue BP110, F-74941 Annecy-le-Vieux CEDEX, France*

K. Francis, J. Repond, J. Schlereth, J. Smith*, L. Xia

Argonne National Laboratory, 9700 S. Cass Avenue, Argonne, IL 60439-4815, USA

E. Baldolemar, J. Li†, S. T. Park, M. Sosebee, A. P. White, J. Yu

Department of Physics, SH108, University of Texas, Arlington, TX 76019, USA

T. Buanes, G. Eigen

University of Bergen, Institute, of Physics, Allegaten 55, N-5007 Bergen, Norway

Y. Mikami, N. K. Watson

*University of Birmingham, School of Physics and Astronomy, Edgbaston, Birmingham B15 2TT,
UK*

G. Mavromanolakis‡, M. A. Thomson, D. R. Ward, W. Yan§

University of Cambridge, Cavendish Laboratory, J J Thomson Avenue, CB3 0HE, UK

D. Benchekrout, A. Hoummada, Y. Khoulaki

Université Hassan II Aïn Chock, Faculté des sciences, B.P. 5366 Maarif, Casablanca, Morocco

J. Apostolakis, A. Dotti, G. Folger, V. Ivantchenko, V. Uzhinskiy

CERN, CH-1211 Genève 23, Switzerland

M. Benyamna, C. Cârloganu, F. Fehr, P. Gay, S. Manen, L. Royer

*Clermont Université, Université Blaise Pascal, CNRS/IN2P3, LPC, BP 10448, F-63000
Clermont-Ferrand, France*

G. C. Blazey, A. Dyshkant, J. G. R. Lima, V. Zutshi

NICADD, Northern Illinois University, Department of Physics, DeKalb, IL 60115, USA

J. -Y. Hostachy, L. Morin

*Laboratoire de Physique Subatomique et de Cosmologie, Université Joseph Fourier Grenoble 1,
CNRS/IN2P3, Institut Polytechnique de Grenoble, 53, rue des Martyrs, F-38026 Grenoble
CEDEX, France*

**U. Cornett, D. David, G. Falley, K. Gadow, P. Göttlicher, C. Günter, B. Hermberg,
S. Karstensen, F. Krivan, A. -I. Lucaci-Timoce[‡], S. Lu, B. Lutz, S. Morozov,
V. Morgunov[¶], M. Reinecke, F. Sefkow, P. Smirnov, M. Terwort, A. Vargas-Trevino**

DESY, Notkestrasse 85, D-22603 Hamburg, Germany

N. Feege, E. Garutti[♠], I. Marchesini^{||}, M. Ramilli

*University of Hamburg, Physics Department, Institut für Experimentalphysik, Luruper Chaussee
149, D-22761 Hamburg, Germany*

P. Eckert, T. Harion, A. Kaplan, H. -Ch. Schultz-Coulon, W. Shen, R. Stamen

*University of Heidelberg, Fakultät für Physik und Astronomie, Albert Uberle Str. 3-5, 2.OG Ost,
D-69120 Heidelberg, Germany*

B. Bilki, E. Norbeck, Y. Onel

*University of Iowa, Department of Physics and Astronomy, 203 Van Allen Hall, Iowa City, IA
52242-1479, USA*

G. W. Wilson

*University of Kansas, Department of Physics and Astronomy, Malott Hall, 1251 Wescoe Hall
Drive, Lawrence, KS 66045-7582, USA*

K. Kawagoe

Department of Physics, Kyushu University, Fukuoka 812-8581, Japan

P. D. Dauncey, A. -M. Magnan

*Imperial College London, Blackett Laboratory, Department of Physics, Prince Consort Road,
London SW7 2AZ, UK*

V. Bartsch^{*}, M. Wing

*Department of Physics and Astronomy, University College London, Gower Street, London WC1E
6BT, UK*

F. Salvatore^{}**

Royal Holloway University of London, Department of Physics, Egham, Surrey TW20 0EX, UK

E. Calvo Alamillo, M.-C. Fouz, J. Puerta-Pelayo

*CIEMAT, Centro de Investigaciones Energeticas, Medioambientales y Tecnologicas, Madrid,
Spain*

B. Bobchenko, M. Chadeeva, M. Danilov^{††}, A. Epifantsev, O. Markin, R. Mizuk^{††}

E. Novikov, V. Popov, V. Rusinov, E. Tarkovsky

*Institute of Theoretical and Experimental Physics, B. Cheremushkinskaya ul. 25, RU-117218
Moscow, Russia*

N. Kirikova, V. Kozlov, P. Smirnov, Y. Soloviev

*P.N. Lebedev Physical Institute, Russian Academy of Sciences, 117924 GSP-1 Moscow, B-333,
Russia*

P. Buzhan, A. Ilyin, V. Kantserov, V. Kaplin, A. Karakash, E. Popova, V. Tikhomirov

*Moscow Physical Engineering Institute, MPhI, Department of Physics, 31, Kashirskoye shosse,
115409 Moscow, Russia*

C. Kiesling, K. Seidel, F. Simon, C. Soldner, M. Szalay, M. Tesar, L. Weuste

Max Planck Institute für Physik, Föhringer Ring 6, D-80805 Munich, Germany

**M. S. Amjad, J. Bonis, S. Callier, S. Conforti di Lorenzo, P. Cornebise, Ph. Doublet,
F. Dulucq, J. Fleury, T. Frisson, N. van der Kolk, H. Li, G. Martin-Chassard,
F. Richard, Ch. de la Taille, R. Pöschl, L. Raux, J. Rouëné, N. Seguin-Moreau**

*Laboratoire de l'Accélérateur Linéaire, Centre Scientifique d'Orsay, Université de Paris-Sud XI,
CNRS/IN2P3, BP 34, Bâtiment 200, F-91898 Orsay CEDEX, France*

**M. Anduze, V. Boudry, J-C. Brient, D. Jeans, P. Mora de Freitas, G. Musat,
M. Reinhard, M. Ruan, H. Videau**

*Laboratoire Leprince-Ringuet (LLR) – École Polytechnique, CNRS/IN2P3, F-91128 Palaiseau,
France*

B. Bulanek, J. Zacek

*Charles University, Institute of Particle & Nuclear Physics, V Holesovickach 2, CZ-18000 Prague
8, Czech Republic*

**J. Cvach, P. Gallus, M. Havranek, M. Janata, J. Kvasnicka, D. Lednicky,
M. Marcisovsky, I. Polak, J. Popule, L. Tomasek, M. Tomasek, P. Ruzicka, P. Sicho,
J. Smolik, V. Vrba, J. Zalesak**

*Institute of Physics, Academy of Sciences of the Czech Republic, Na Slovance 2, CZ-18221
Prague 8, Czech Republic*

B. Belhorma, H. Ghazlane

*Centre National de l'Energie, des Sciences et des Techniques Nucléaires, B.P. 1382, R.P. 10001,
Rabat, Morocco*

T. Takeshita, S. Uozumi

Shinshu University, Department of Physics, 3-1-1 Asahi, Matsumoto-shi, Nagano 390-861, Japan

M. Götze, O. Hartbrich, J. Sauer, S. Weber, C. Zeitnitz

*Bergische Universität Wuppertal, Fachbereich 8 Physik, Gausstrasse 20, D-42097 Wuppertal,
Germany*

♠ *Corresponding author*

E-mail: erika.garutti@desy.de

ABSTRACT: Calorimeters with a high granularity are a fundamental requirement of the Particle Flow paradigm. This paper focuses on the prototype of a hadron calorimeter with analog readout, consisting of thirty-eight scintillator layers alternating with steel absorber planes. The scintillator plates are finely segmented into tiles individually read out via Silicon Photomultipliers. The presented results are based on data collected with pion beams in the energy range from 8 GeV to 100 GeV. The fine segmentation of the sensitive layers and the high sampling frequency allow for an excellent reconstruction of the spatial development of hadronic showers. A comparison between data and Monte Carlo simulations is presented, concerning both the longitudinal and lateral development of hadronic showers and the global response of the calorimeter. The performance of several GEANT4 physics lists with respect to these observables is evaluated.

KEYWORDS: hadronic calorimetry; hadronic shower models; imaging calorimetry.

Contents

1. Introduction	2
2. Experimental Setup	2
2.1 The SiW-ECAL	3
2.2 The AHCAL	3
2.3 The TCMT	4
3. Calibration	5
4. Monte Carlo Simulations	5
4.1 Digitization	5
4.2 Physics Lists	6
5. Event Selection	7
6. Calibration Validation	9
6.1 Systematic Uncertainties	10
7. Energy Response	11
8. Pion Interaction Length	13
9. Longitudinal Development	15
10. Radial Development	19
11. Conclusions	20

*Also at University of Texas, Arlington

†Deceased

‡Now at CERN

§Now at Department of Modern Physics, University of Science and Technology of China, 96 Jinzhai Road, Hefei, Anhui, 230026, P. R. China

¶Also at ITEP

||Also at DESY

**Now at University of Sussex, Physics and Astronomy Department, Brighton, Sussex, BN1 9QH, UK

††Also at MEPhI and at Moscow Institute of Physics and Technology

‡‡Now at LPSC Grenoble

1. Introduction

The next generation of lepton colliders, such as the International Linear Collider (ILC) [1], will require unprecedented detector performances in order to fully exploit the physics potential. A clean separation between the hadronic decays of the W and Z bosons sets the goal for the jet energy resolution. Ideally, the di-jet mass resolution should be comparable to the decay widths of the weak vector bosons. This requires a jet energy resolution σ_E/E_j of the order of 3-4% at the Z peak, which translates approximately into $\sigma_E/E_j \sim 30\%/\sqrt{E_j/\text{GeV}}$. Within the ILC community, Particle Flow calorimetry has been identified as the most promising way to achieve this level of precision [2, 3, 4], as successfully tested on Monte Carlo simulations. For instance, the Pandora Particle Flow algorithm, which has been developed in the context of the detector optimization studies for the ILC, has demonstrated the feasibility of jet energy resolutions of $\sigma_E/E_j \lesssim 3.8\%$ for 40 – 400 GeV jets [4].

The key factor in Particle Flow calorimetry is the single particle separability within a jet, which requires a high granularity of the calorimeters. The CALICE collaboration is performing studies of several designs of highly granular calorimeters, with sensitive layers finely segmented into cells, which are individually read out. The stability over a large number of channels of new readout technologies has been demonstrated and reliable procedures of cell equalization and calibration have been established. Furthermore, the high spatial resolution allowed by the fine granularity provides valuable input to the validation of the shower models used for Monte Carlo simulations. Such studies are of broader interest, which goes beyond the support of simulation studies of Particle Flow algorithms. Uncertainties in the shower modeling still contribute significantly to energy scale uncertainties in high energy physics experiments, for instance at the Large Hadron Collider detectors [5].

Since 2006 several beam tests of the CALICE prototypes have been conducted at the Deutsches Elektronen-Synchrotron (DESY), at the European Organization for Nuclear Research (CERN) and at the Fermi National Accelerator Laboratory (FNAL). This paper focuses on data collected at CERN in 2007, when the experimental set-up consisted of a silicon-tungsten electromagnetic sampling calorimeter (SiW-ECAL) [6], a scintillator-steel hadron sampling calorimeter with analog readout (AHCAL) [7] and a scintillator-steel tail catcher and muon tracker (TCMT) [8].

The analyzed data were collected when the prototypes were exposed to pion beams in the energy range from 8 GeV to 100 GeV, provided by the CERN SPS H6 beam line. The overall response of the calorimeter and the lateral and longitudinal profiles of hadronic showers are compared to the predictions made by several Monte Carlo simulations, which make use of different GEANT4 physics lists (Sec. 4).

The study is completed by an update on the validation of the calibration of the AHCAL performed with electromagnetic showers, already described in a previous publication [9]. This study uses data collected both at FNAL and CERN, exposing the prototypes to electron and positron beams in the energy range from 1 GeV to 50 GeV.

2. Experimental Setup

During the beam tests at CERN the detectors were exposed to muon, positron and pion beams.

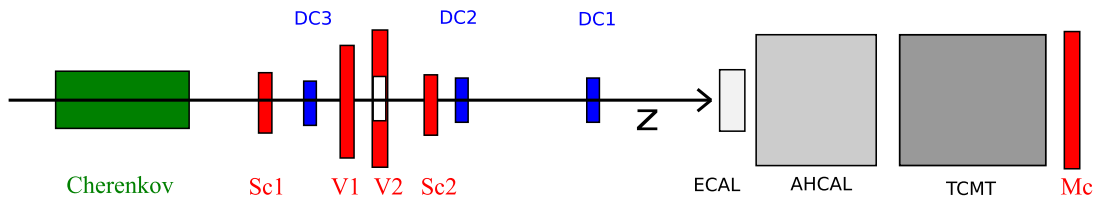


Figure 1. Schematic layout (not to scale) of the CALICE experimental setup at CERN, with calorimeters and beam instrumentation. *Sc1* and *Sc2* are the scintillators for the beam trigger, *V1* is the multiplicity counter, *V2* is the veto trigger and *Mc* is the scintillator for the muon trigger.

The SiW-ECAL and the AHCAL were mounted on a movable stage, providing the possibility to translate and rotate the calorimeters with respect to the beam. However, the studies discussed in the following only make use of data collected with the beam incident in the center of the calorimeters, along their center line.

A schematic overview of the experimental setup is shown in Fig. 1. Three sets of wire chambers were operated upstream of the detectors in order to measure the beam coordinates. The coincidence signal of two upstream $10 \times 10 \text{ cm}^2$ scintillator counters (*Sc1* and *Sc2* in Fig. 1) triggered the readout of the detector (*beam trigger*). One scintillator (*V1* in Fig. 1) with an area of $20 \times 20 \text{ cm}^2$ and analog readout tagged multi-particle events (*multiplicity counter*). A $100 \times 100 \text{ cm}^2$ scintillator with a $20 \times 20 \text{ cm}^2$ hole in the center (*V2* in Fig. 1) rejected beam halo events (*veto trigger*). Muons were identified by a downstream $100 \times 100 \text{ cm}^2$ scintillator (*Mc* in Fig. 1). A Cherenkov counter was also operated in threshold mode to discriminate between electrons and negative pions or between positive pions and protons.

2.1 The SiW-ECAL

The SiW-ECAL is divided into three stacks, each composed of 10 modules of alternating tungsten and silicon layers. Each stack has tungsten layers with different thicknesses: 1.4 mm ($0.4 X_0$) per layer in the first stack, 2.8 mm or $0.8 X_0$ in the second and 4.2 mm or $1.2 X_0$ in the third one. The silicon layers are segmented into PIN diodes of $1 \times 1 \text{ cm}^2$, for a total number of about 9700 read-out cells.

The overall thickness of the prototype is about 20 cm, corresponding to $24.6 X_0$ or 0.9 nuclear interaction lengths λ_I . The lateral size of the SiW-ECAL is $18 \times 18 \text{ cm}^2$.

2.2 The AHCAL

The AHCAL is a sampling structure of 38 modules, each consisting of a $\sim 2 \text{ cm}$ thick steel absorber plate and a sensitive layer instrumented with 0.5 cm thick scintillator tiles. The total depth of the prototype is 1.2 m, translating into about $5.3 \lambda_I$, while the lateral dimensions are approximately $1 \times 1 \text{ m}^2$. Each sensitive layer is composed of scintillator tiles of different sizes (Fig. 2, left). The $30 \times 30 \text{ cm}^2$ core has a granularity of $3 \times 3 \text{ cm}^2$, while the outer region is equipped with tiles of increasing sizes ($6 \times 6 \text{ cm}^2$ and $12 \times 12 \text{ cm}^2$). In the last 8 layers the highly granular core is replaced by $6 \times 6 \text{ cm}^2$ tiles. The scintillation light from each tile is read out individually by a Silicon PhotoMultiplier (SiPM) [10, 11], coupled to the scintillator via a WaveLength Shifting fiber (WLS). The SiPMs employed in the AHCAL have a photosensitive surface of 1.1 mm^2 , which is

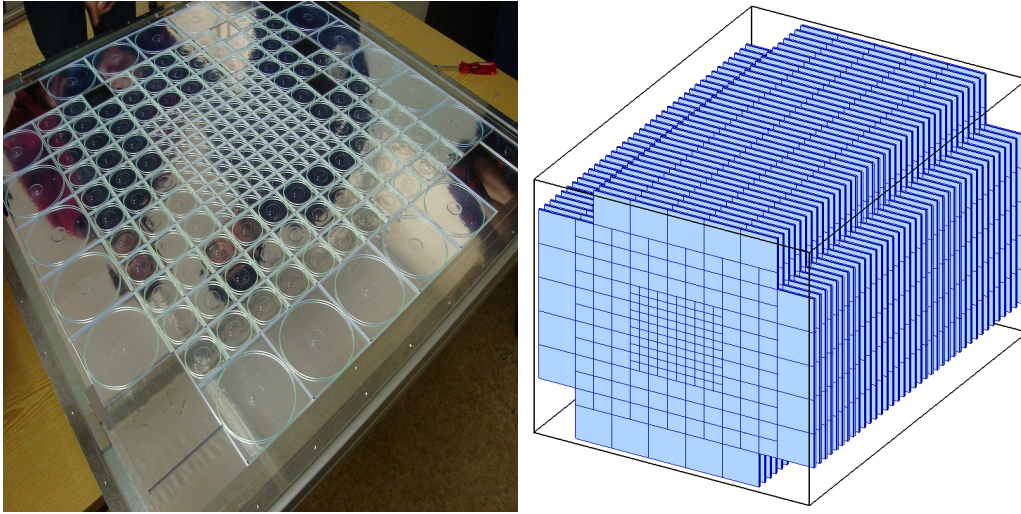


Figure 2. Left: Segmentation of a sensitive layer of the AHCAL. The size of the scintillator tiles increases towards the outer region. Right: schematic view of the active layers of the AHCAL.

divided into 1156 pixels. These pixels are individually equipped with a quenching resistor and mounted on a common substrate, in such a way that the charge signal is proportional to the number of pixels fired.

The readout system of the SiPMs is operated in two different modes, called *calibration* and *physics* modes. For calibration purposes single photons need to be resolved in the SiPM spectrum, since the distance between two consecutive peaks in the spectrum determines the gain of the SiPM. Therefore, a short shaping time of the signal (40 ns) and a high amplification are needed. In contrast, during physics runs large signals are produced and the amplification needs to be reduced by approximately a factor 10 to minimize saturation effects. Furthermore, a longer shaping time of about 180 ns is used to provide sufficient latency for the beam trigger decision.

The response of the SiPMs is measured both in physics and calibration modes, in order to determine the gain and saturation level of the SiPMs and the electronics intercalibration between the two operation modes. The response is monitored using LEDs with tunable intensities.

Temperature sensors placed inside the detector in correspondence to each module monitor temperature changes with an accuracy better than 0.6°C .

2.3 The TCMT

The TCMT is positioned downstream with respect to the AHCAL in order to absorb the tails of the showers leaking out of the AHCAL. The TCMT has a lateral size of $109 \times 109 \text{ cm}^2$ and is 142 cm in depth, corresponding to $5.8 \lambda_I$. It consists of two sections, a *fine* one and a *coarse* one. Each section has 8 sensitive layers, alternating with steel absorbers. The absorber plates are 2 cm thick in the fine section and 10 cm thick in the coarse section. The sensitive layers are 0.5 cm thick and are segmented into 5 cm wide and 1 m long scintillator strips, with alternating horizontal and vertical orientation in adjacent layers. The scintillation light is collected by WLS fibers and detected by SiPMs, as for the AHCAL prototype.

3. Calibration

As typical of semiconductor devices, SiPMs are sensitive to temperature and bias voltage changes that affect most of their performance parameters [7]. Moreover, SiPMs show saturation effects, due to the finite number of pixels and the finite pixel recovery time. One of the main technical aims of the AHCAL prototype is to show that these effects can be handled over a large number of channels.

The calibration chain is described in [9]. It proceeds through the following steps:

- inter-cell equalization of the response;
- calibration of the SiPM signal and correction for the non-linear response;
- conversion of the calibrated signal to the GeV scale.

The equalization of the response of the 7608 AHCAL cells is performed using the reference signal from muons [12]. Muons represent the best approximation of the behavior of a Minimum Ionizing Particle (MIP). The signal measured in a cell expressed in ADC counts can be converted to MIPs, using as unit the Most Probable Value (MPV) of the response to the muon beam.

In order to calibrate the SiPM signals, the gain is measured periodically during data taking using the LED system. A correction is also applied to account for the non-linear response of the SiPMs, based on the specific response curve of each SiPM, measured during dedicated laboratory tests. Since the SiPM properties depend on the temperature and the voltage, a given set of calibration constants is only valid for measurements at the same operation conditions and needs to be extrapolated to the operating conditions to account for possible changes, such as temperature fluctuations [13].

Finally, the conversion of the signal to the GeV scale is derived from the response of the detector to electromagnetic showers (Sec. 6).

4. Monte Carlo Simulations

The Monte Carlo simulations for the CALICE prototypes are carried out in the framework of GEANT4 [14]. The version 9.4 patch 3 of GEANT4 is used as default in the following, apart from explicit comparisons with previous versions. The geometries of the calorimeters are simulated within GEANT4 using Mokka [15, 16]. The detectors upstream of the calorimeters are simulated as well. The origin of the simulated particles is located upstream of the full system, such that any interactions with the Cherenkov counter, the scintillator-triggers, the tracking chambers and the air volumes in between are taken into account. Sensitive and passive materials, gaps and support structures are also considered in detail. The simulated AHCAL active layers have a uniform granularity of $1 \times 1 \text{ cm}^2$ cell size. The realistic geometry of the AHCAL is obtained during the digitization procedure, as described below.

4.1 Digitization

The simulated events are processed further in the so-called digitization step, which models the response of the detector and its electronics. This allows the same treatment for data and Monte

Carlo in the subsequent analysis steps, such as calibration and reconstruction. This procedure takes into account several factors:

- the detector granularity. The signal amplitude of the $1 \times 1 \text{ cm}^2$ virtual cells simulated is summed up to obtain the real geometry with 3×3 , 6×6 and $12 \times 12 \text{ cm}^2$ cells.
- the light sharing between neighboring tiles (known also as *cross-talk*). A 2.5% light sharing from each tile edge is used in simulations [9].
- non-linearity effects of the SiPMs, based on their specific saturation curves.
- Poissonian fluctuations of the photo-electron statistics.
- the noise contribution. The noise is overlaid on the simulation, using real noise measured in dedicated triggers without beam.

Shielding effects in the scintillator material saturate the scintillation process at high ionization densities. This causes a non-linearity of the light yield, which has been simulated in Monte Carlo according to the Birks Law [17].

The read out electronics have a defined time window. Late energy depositions, for instance from neutrons, might escape this time window. In order to reproduce this effect, a time cut of 150 ns is applied in simulations.

The effects of gaps between the calorimeter tiles, as well as the non-uniform response of the tiles have been studied in Monte Carlo events in [18], showing that these type of effects do not have a significant influence on the measurement of hadron showers.

4.2 Physics Lists

The interactions of hadrons with matter cannot be modeled from first principles alone; several phenomenological models, working with different approximations, exist. An overview of the different models is given in [19]. The main features are summarized here:

- **String Parton Models.** These models are employed to simulate the interaction of medium or high energy hadrons with nuclei. As a first stage the interaction of the particles with at least one nucleon of the nucleus is modeled using a string excitation model. Two different approaches are available in GEANT4: the Fritiof (**FTF**) and the Quark Gluon String (**QGS**) model. In the FTF approach the diffractive scattering of the primary particle with the nucleons is realized only via momentum exchange. In the QGS model the hadron-nucleon interaction is mediated via pomerons. The products of an interaction between the primary particle and the nucleus are one or several excited strings and a nucleus in an excited state. The fragmentation of the excited strings into hadrons is handled by a longitudinal string fragmentation model, with differences between FTF and QGS. The interaction of secondaries with the excited nucleus is handled by a shower model or by a precompound model (see below). In the latter case the string parton models get the suffix “P” (FTFP and QGSP). The de-excitation of the excited nucleus is further simulated by nuclear fragmentation, precompound and nuclear de-excitation models.

- **Parameterized Models.** Low Energy Parametrized (**LEP**) and High Energy Parametrized (**HEP**) models are based on fits to experimental data, with little theoretical guidance. They do not conserve energy, momentum, charge or other quantum numbers on an event-by-event basis, but instead seek to conserve these quantities on average. The approach of GEANT4 is to limit the use of these models and replace them by more sophisticated ones where possible.
- **Cascade Models.** The Bertini cascade **BERT** and the binary cascade **BIC** models are employed at medium and low energies, where the quark-substructure of individual nuclei can be neglected. The BERT model describes a nucleus as a sphere with constant nucleon density. Incident hadrons collide with protons and neutrons in the target nucleus and produce secondaries which in turn collide with other nucleons generating a so-called intra-nuclear cascade. At the end of the cascade the excited nucleus is represented as a sum of particle-hole states which is then decayed by pre-equilibrium, nucleus explosion, fission and evaporation models. The BIC model considers the nucleus as a sum of discrete nucleons, at defined positions and with defined momenta. The propagation through the nucleus of the incident hadron and the secondaries produced are modeled by a cascading series of two-particle collisions. Secondaries are created during the decay of resonances formed during the collisions. The decay of the excited nucleus is handled by the precompound model.
- **Precompound Model.** The precompound model generates the final state for hadron inelastic scattering. It describes the emission of protons, neutrons, and light ions before the equilibrium of a nuclear system is reached. The final products are passed to de-excitation models.
- **CHIPS Model.** The Chiral Invariant Phase Space (CHIPS) model is a nuclear fragmentation model acting at the quark level. The model is based on the concept of a quasmon, which is an excited intermediate state of massless quarks that are asymptotically free, formed from the quarks of the projectile hadron and of the target nucleon. The quasmons decay via internal quark fusion or by double quark exchange with neighboring quasmons. This model is still at an experimental phase and several parameters are still being optimized.

Several “physics lists” are available in GEANT4, which combine different models in different energy ranges, with a random choice of which model is used for overlapping energy regions. A summary of the physics lists considered in this paper and of their composition at different energies is given in Fig. 3.

5. Event Selection

This paper presents electron (or positron¹) and negative pion data collected with the experimental set-up described in Sec. 2. While performing measurements with an electron beam, the SiW-ECAL is moved out of the beam line. Though the beam is highly enriched with the nominal particle, some contamination is present and the recorded event samples need to be purified for analysis. The selection applied is summarized in the following and has been tested on Monte Carlo

¹In the following, by “electron beam” we refer to either an electron or positron beam, since no distinction between the two beam particles is relevant for the performed studies.

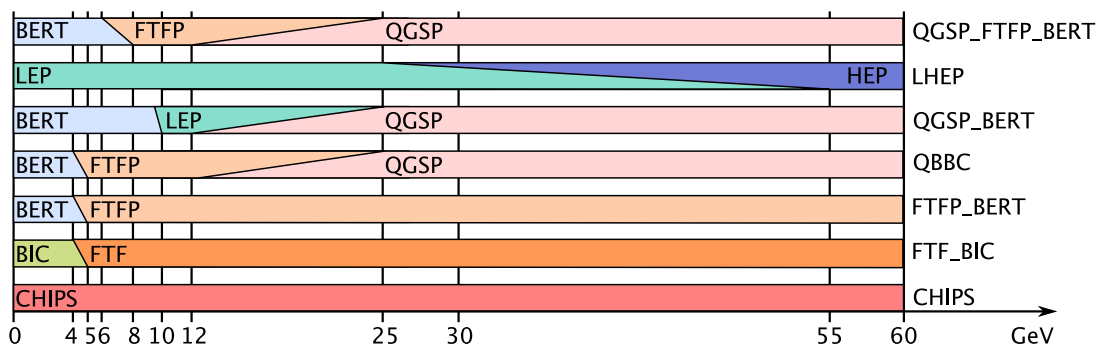


Figure 3. Models applied at different energies in several physics lists for the simulation of pions in GEANT4, version 9.4 patch3.

simulations, comparing the selection efficiencies for simulated muons, electrons and pions. Some of the selection criteria use information from the AHCAL itself. Applying the same requirements to simulated events confirms that no bias is introduced in the energy distributions. More details on the event selection are given in [20].

Beam events are selected requiring the beam trigger. Events where particles generate a beam trigger signal but fail to reach the calorimeters are rejected by requiring a minimum energy deposition of 4 MIPs in the $3 \times 3 \text{ cm}^2$ cells of the first five layers of the AHCAL. Signals of less than 0.5 MIPs in single cells are rejected in all calorimeters to reduce the noise contribution. This requirement has been tested on events collected without incoming beam, since they have a topology similar to that given by spurious trigger events (i.e. only noise). About 96% of the beam-off events are rejected. Some events contain additional particles in the beam halo or particles that initiate a shower before reaching the AHCAL. These events are excluded by requiring no signal from the veto trigger and less than 15 hits in the $6 \times 6 \text{ cm}^2$ cells of the first five layers of the AHCAL. In order to exclude events with more than one particle depositing energy in the AHCAL at the same time, only events with a multiplicity counter amplitude of less than 1.4 MIPs are kept. This selection yields a multi-particle contamination of less than 0.1%, which has been evaluated using the estimated distribution of the signal from double particle events in the multiplicity counter. Such a distribution has been simulated from the known distribution of noise, measured during triggers without beam, and the signal from single particle events. The signal from single particles has been obtained fitting the distribution of the measured amplitude in the multiplicity counter in a low amplitude region, where no significant contribution from double particle events is expected.

In order to improve the purity of the electron data, a signal is required in the Cherenkov detector. In addition, the center of gravity of the energy deposits in the beam direction is required to be in the front part of the AHCAL (less than 360 mm beyond the front face of the AHCAL). At least one cluster of neighboring hits with a total energy of 18 MIPs or higher has to be found in the AHCAL and an energy of less than 5 MIPs has to be deposited in the last 10 layers of the AHCAL.

For minimizing the electron contamination in pion data, events with a signal in the Cherenkov detector are rejected. The contamination from muons is suppressed by requiring the identification in the AHCAL of the position where the pion shower is initiated by the first hard interaction, since muons generally do not produce showers. The position of the first hard interaction is identified

using a cluster-based algorithm, based on the three-dimensional distribution of hits [21]. Seed hits with visible energies of more than 1.65 MIPs are sorted by their z -positions in ascending order. Starting with the first seed hit according to this ordering, each seed hit and all neighboring hits are assigned to a cluster. As long as one or more of the cells added to a cluster meet the seed hit requirement, the clustering continues and all hits adjacent to these cells are assigned to the same cluster. The cluster closest to the point the pion enters the calorimeter consisting of at least 4 hits and having a minimum energy of 16 MIPs is identified as the beginning of the shower. The end of the cluster axis pointing in the direction of the incoming hadron is used as location of the first inelastic scattering. According to Monte Carlo simulations this procedure allows rejection of muons with an efficiency of more than 80%. Since muons can produce shower-like clusters along the track due to Bremsstrahlung, the muon contamination is further reduced by requiring additionally more than 60 hits in the AHCAL, yielding a muon rejection efficiency of more than 95%.

In order to minimize the systematic uncertainties arising from combining the information from different detectors, pion events for which the shower starting point is found in the SiW-ECAL are rejected. The algorithm used to identify the beginning of the hadronic shower in the SiW-ECAL is described in [22]. The beginning of the shower is found by imposing thresholds on the energy measured in the different layers and on the number of hits deposited. This requirement further minimizes the contamination from electrons, which start to shower in the SiW-ECAL.

6. Calibration Validation

Although the AHCAL is designed to measure hadrons, during beam tests it has also been exposed to electron beams. The study of the electromagnetic response serves to prove the understanding of the detector and to validate the calibration procedure, since electromagnetic showers can be precisely modeled. Additionally, the AHCAL is a non-compensating calorimeter and the energy response to hadrons may be non-linear. In contrast, apart from saturation effects, the response to electromagnetic showers is linear and allows for the determination of the conversion of the calibrated signal to the energy scale in units of GeV (referred to as electromagnetic energy scale).

The study of the electromagnetic response of the AHCAL is described in detail in [9]. Electron data of energies between 10 GeV and 50 GeV are compared to Monte Carlo simulations and the systematic uncertainties associated to the different steps of the calibration procedure are estimated.

Overall, the studies in [9] validate the simulation of electromagnetic showers at a level sufficient for the present analysis of the hadronic response. A residual disagreement of the order of 2 mm is found between data and simulation, when comparing the radial development.

In order to confirm that the results in [9] are up-to-date with the most recent calibration and reconstruction software tools and to extend them over a broader energy range, the measurement of the energy response to electromagnetic showers is repeated here. This new study focuses on the determination of the electromagnetic energy scale, since it is applied in the following to pion data. The validation of the simulation of electromagnetic showers for the observables discussed in [9] is not repeated. The data set has been extended with respect to [9] to lower energies down to 1 GeV, using data collected at FNAL exposing the AHCAL to electron beams, with an experimental set-up similar to the one employed at CERN (Sec. 2).

	u [MIP/GeV]	v [MIP]
Data	42.4 ± 0.3	-1.1 ± 0.9
MC	42.8 ± 0.1	-6.6 ± 0.5

Table 1. Measured parameters of the linear fit to the electromagnetic response, as defined in Eq. 6.1, for data and Monte Carlo. The errors reported are the fit uncertainties.

The expected linear response to electrons is described by the following equation:

$$\langle E_{\text{rec}}^e \rangle = p_{\text{beam}} \cdot u + v, \quad (6.1)$$

where $\langle E_{\text{rec}}^e \rangle$ is the mean calibrated response to electrons expressed in MIPs, p_{beam} is the beam momentum, and the parameters u and v are obtained from a linear fit to the dependence of $\langle E_{\text{rec}}^e \rangle$ on p_{beam} . The factor u represents the calibration to the GeV scale of the electromagnetic response. The offset v accounts for the combined effect of the 0.5 MIPs threshold applied to all events (Sec. 5) and of the residual noise above the threshold. The residuals to the fit are shown in Fig. 4 (left). The linearity is better than about 2% at all energies. The errors taken into account in the fit and shown in the figure are both the statistical and the systematic uncertainties. The systematic uncertainty includes a calibration scale uncertainty of 1.6% and the SiPMs saturation uncertainty, which ranges between 0.5% at 1 GeV and 3% at 50 GeV.

The same procedure is applied to Monte Carlo, yielding the results summarized in Tab. 1, which are compatible with the results obtained previously in [9]. The factor u agrees with the results obtained for data within the fit uncertainties, while the offset v disagrees by about 5 standard deviations (5 MIPs or ~ 100 MeV). This remaining discrepancy is attributed to an imperfect implementation of noise, light-sharing and threshold effects. An imperfect modeling of the beam line and of the energy lost in the upstream material could also contribute. Since in electromagnetic showers the signal cells below threshold concentrate on the edges, an inaccurate simulation of the lateral extension of electromagnetic showers, as observed in [9], could also contribute to this effect.

The difference between the simulated response and the measured response is shown in Fig. 4 (right). The discrepancy in v is visible at low energies, while at high energies the Monte Carlo slightly underestimates the SiPMs saturation.

The electromagnetic analysis allows the conclusion that the detector calibration and simulation are sufficiently understood, in order to carry on studies based on hadronic data. The calibration to the GeV scale obtained for electromagnetic data is applied in the following to pion data and to the digitized Monte Carlo.

6.1 Systematic Uncertainties

The results shown in the following focus on the response of the calorimeter to hadrons. The systematic uncertainty on the calibration scale of 1.6%, quoted above for the electromagnetic study, affects the hadronic analysis in the same way. This uncertainty is relevant for the study of the hadronic energy response discussed in Sec. 7. The saturation uncertainty, which ranges between 0.5% at 1 GeV and 3% at 50 GeV for electromagnetic showers, is expected to have a reduced impact on hadronic showers, due to the lower energy density. This is discussed in [9], where it is shown that electromagnetic showers have more hits with high energy deposition than pion showers,

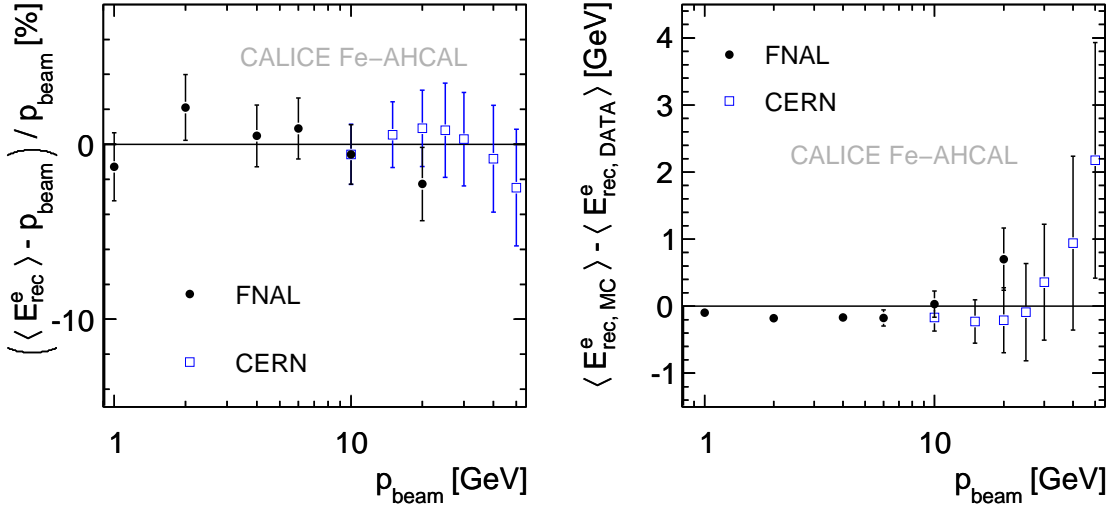


Figure 4. Left: Residuals from linearity for electromagnetic data collected at FNAL (circles) and CERN (squares). Right: Difference between the simulated and the measured electromagnetic response. The error bars take into account both the statistical and the systematic uncertainties, as described in the text.

even when the beam energy is only half that of the pion. Therefore, this source of uncertainty is neglected during the measurement of the hadronic energy response.

The measurement of the pion interaction length is dominated by the fit uncertainty, as discussed in Sec. 8.

The individual response of each layer is relevant for the longitudinal shower profiles discussed in Sec. 9. For this study a dedicated evaluation of the systematics has been performed, as described in Sec. 9. A similar procedure has been performed for radial profiles (Sec. 10), yielding negligible uncertainties.

Systematic uncertainties are assumed to affect in a negligible way the observables concerning average global quantities of the calorimeter, such as the center of gravity and the standard deviation of longitudinal profiles (Sec. 9), the mean energy-weighted shower radius and the standard deviation of radial energy distributions (Sec. 10). As a cross-check, the uncertainties obtained for longitudinal profiles have been propagated to the measurement of the longitudinal center of gravity of showers, obtaining uncertainties lower than 1%, and to the standard deviation of the longitudinal profiles, yielding uncertainties lower than 0.5%. Such a level of accuracy does not affect comparisons between data and Monte Carlo, which involve larger effects.

7. Energy Response

The calibrated response to pion showers in units of MIPs is converted to the GeV scale according to the following equation:

$$E_{\text{rec}} [\text{GeV}] = (E_{\text{ECAL1}} + E_{\text{ECAL2}} \cdot 2 + E_{\text{ECAL3}} \cdot 3) \cdot 0.002953 + (E_{\text{AHCAL}} - v)/u, \quad (7.1)$$

where E_{ECAL1} is the energy in MIPs measured in the first section of the SiW-ECAL. E_{ECAL2} , E_{ECAL3} are defined analogously. E_{AHCAL} is the energy in MIPs measured in the AHCAL. The factors 2 and 3 account for the different sampling structure in the three sections of the SiW-ECAL, with different absorber thicknesses. The factor 0.002953 is obtained from simulations. It has been evaluated using simulated muons, since the selection criteria require pions to start showering in the AHCAL and so to traverse the SiW-ECAL losing energy only by ionization, similarly to muons. The parameters u and v have been defined in Sec. 6 and refer to the electromagnetic energy scale. Additional corrections to account for the different response to electromagnetic and hadronic showers need to be applied, if one is interested in the absolute scale of the hadronic response. This topic is discussed in [23] and is not relevant for this paper, which only concerns comparisons between the measured and the simulated response to pions.

In addition to the requirements discussed in Sec. 5, only events for which the hadronic shower starts in the first five AHCAL layers are considered for the studies presented in this section. This requirement minimizes the leakage of the showers out of the AHCAL and reduces the energy collected by the TCMT. This additional selection can be applied since this paper focuses on the validation of Monte Carlo models, rather than on the estimate of the energy resolution. Studies of leakage for non-contained showers have been performed in [24, 8].

The total energy distribution for pion runs at different energies is shown in Fig. 5 (top). Figure 5 (lower left) presents the response to 10 GeV pions for data and for digitized Monte Carlo simulations based on the `FTFP_BERT` physics list. The shape of the response as well as the position of the peak are rather well simulated by the Monte Carlo. At higher energies, as exemplified by an 80 GeV run in Fig. 5 (lower right), the agreement gets worse. This is due to a worse simulation of the energy response and to an imperfect description of the shower length by the Monte Carlo, which results in a different impact of leakage in data and Monte Carlo. The longitudinal development of showers is described in Sec. 9.

In order to quantify the agreement between data and Monte Carlo models, their mean response as function of beam momentum is compared in Fig. 6. The average energy response $\langle E_{\text{rec}} \rangle$ for the different beam energies is given by the arithmetic mean of the energy distributions, after subtraction of the average contribution of the noise above the 0.5 MIPs threshold. The `QGSP_BERT`, the `QGSP_FTFP_BERT` and the `QBBC` physics lists agree within 4% with data at 8 GeV, where the Bertini model is employed. At higher energies the disagreement with data increases, by about 7-10% at 100 GeV. The difference between `QGSP_FTFP_BERT` and `QBBC` at low energies, where they apply the same model, is attributed to the different treatment of the inelastic scattering of secondary protons and neutrons at very low energies, which is described using the Bertini model in `QGSP_FTFP_BERT`, while `QBBC` uses the BIC model. The Fritiof-based physics lists underestimate the response compared to the data at 8 GeV by up to 2-3%, while they agree with data in the range 10-30 GeV. At higher energies, between 40 GeV and 100 GeV, they overestimate data, though the deviations are below 6%. The performance of the `FTFP_BERT` physics list for previous versions of GEANT4 is shown in Fig. 6 (top, right). The performance has significantly improved from the version 9.2 to the version 9.3. In the most recent version the response has slightly increased with respect to the version 9.3, worsening the agreement with data at high energies. `CHIPS` overestimates data at all energies, by up to 10% at 80 GeV. At low energies, below 20 GeV the description of the energy response improves and the disagreement reduces to 2%. The `LHEP` physics list, based

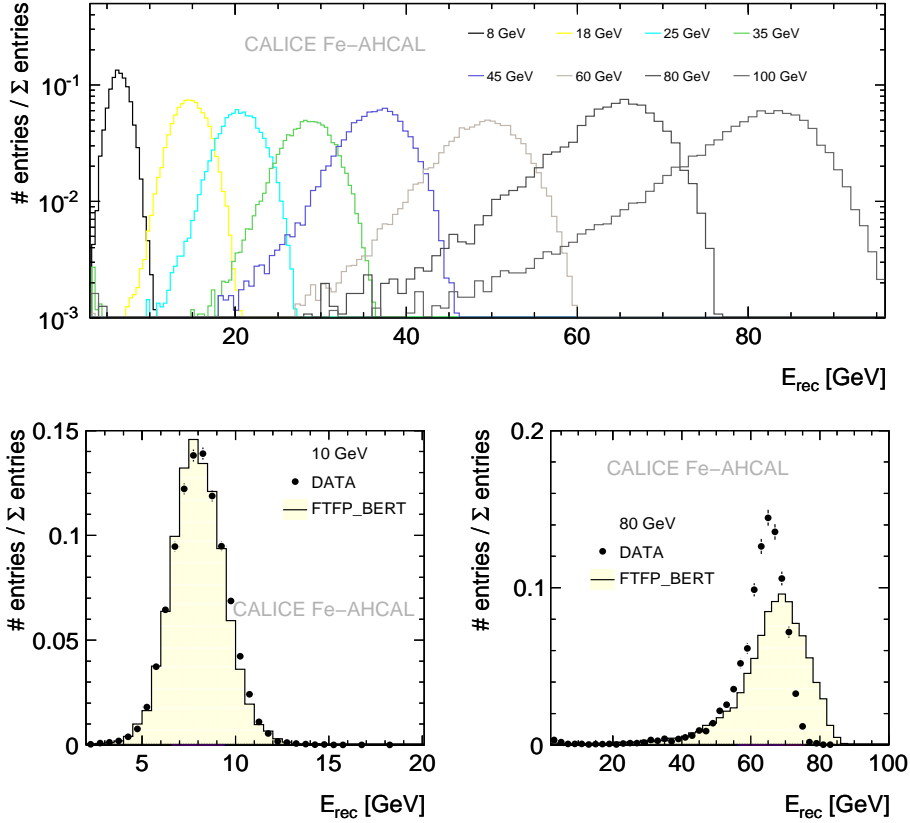


Figure 5. Top: Total energy distribution for selected pion events at different energies. Bottom, left (right): Response to 10 GeV (80 GeV) pions for data (points) and for digitized Monte Carlo simulations using the *FTFP_BERT* physics list.

on parameterized models, largely underestimates the response by up to 8-10% at 8-10 GeV. At high energies above 40 GeV the disagreement between data and Monte Carlo reduces to about 3%.

8. Pion Interaction Length

As long as a pion traversing a material does not interact strongly with a nucleus, it loses energy mainly by ionization. The probability P_I of having an inelastic hadron-nucleus interaction before a distance x is given by:

$$P_I = 1 - e^{-x/\lambda_\pi}, \quad (8.1)$$

where λ_π is the pion interaction length.

The position of the first hard interaction of the primary beam particle is approximated by the reconstructed layer where the shower starts. The measured distribution of the shower starting point follows the exponential behavior expected from Eq. 8.1, as shown in Fig. 7 (left) for 45 GeV pion showers.

From this distribution it is possible to derive the effective interaction length of a pion λ_π in the AHCAL. The interaction length is obtained fitting an exponential to the distribution. The first

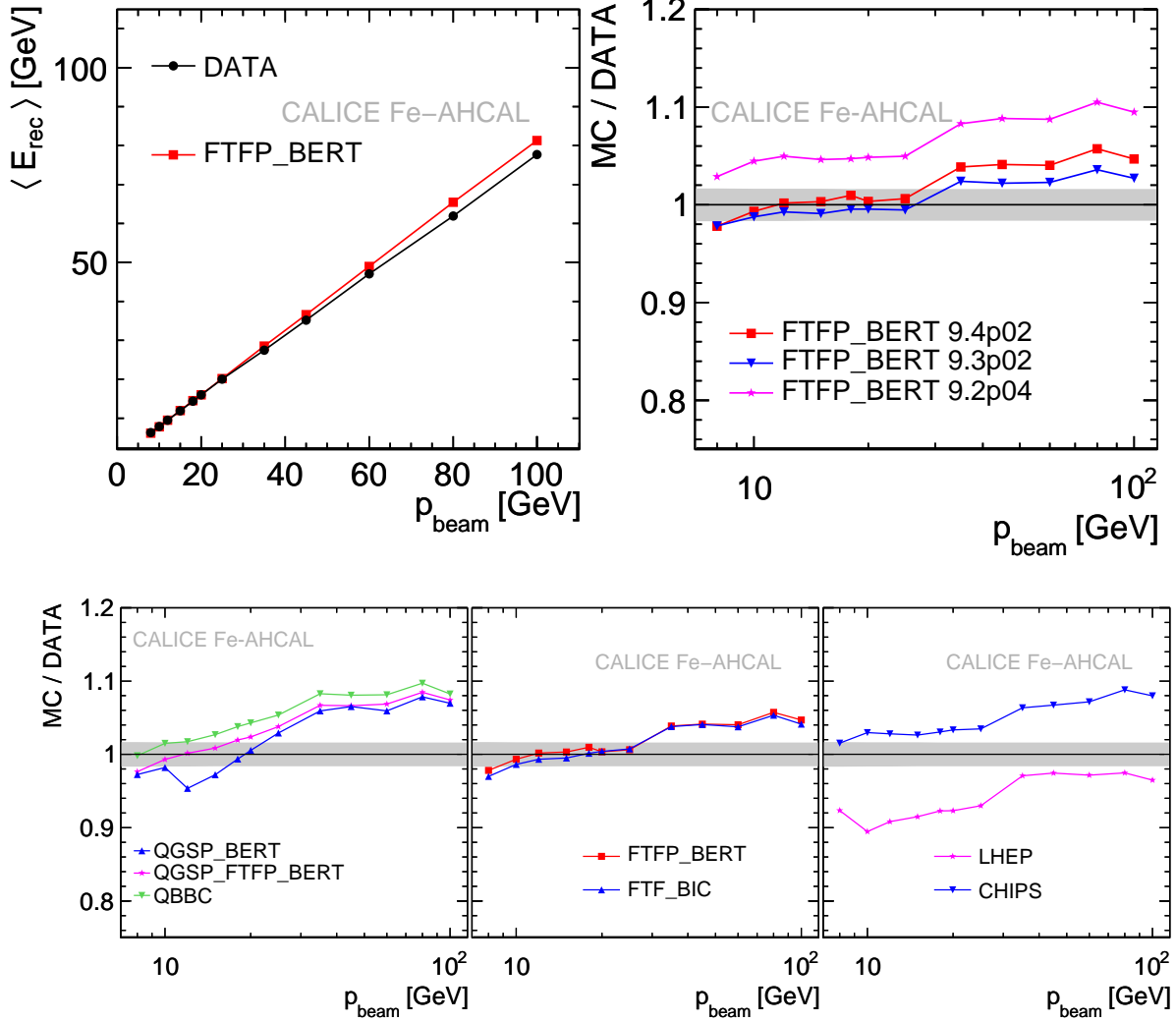


Figure 6. Summary of the energy response to pions. Top, left: For data and for the *FTFP_BERT* physics list. Top, right: Ratio between Monte Carlo and data using the *FTFP_BERT* physics list with different versions of GEANT4. Bottom: Ratio between Monte Carlo and data for several physics list. The gray band in the ratios represents the uncertainty on data (statistical plus a systematic calibration uncertainty of 1.6%).

layers are excluded from the fit, since the uncertainty of the algorithm used to determine the layer of the first hard interaction is larger there, due to the contamination from showers that start to develop in the last layers of the SiW-ECAL. The last layers are also not considered when performing the fit, since the fluctuations are higher due to the lower statistics and the algorithm used to identify the shower starting point has not been optimized for the coarser granularity of the last layers of the AHCAL. Hence, the fit is performed in the range from 120 mm to 800 mm of the longitudinal z coordinate.

The extracted values of the interaction length for data and Monte Carlo are shown in Fig. 8. No energy dependence of the interaction length is observed in data. The measurements yield an average of $\lambda_{\pi} = (26.8 \pm 0.46)$ cm. The error includes the statistical error, determined from the

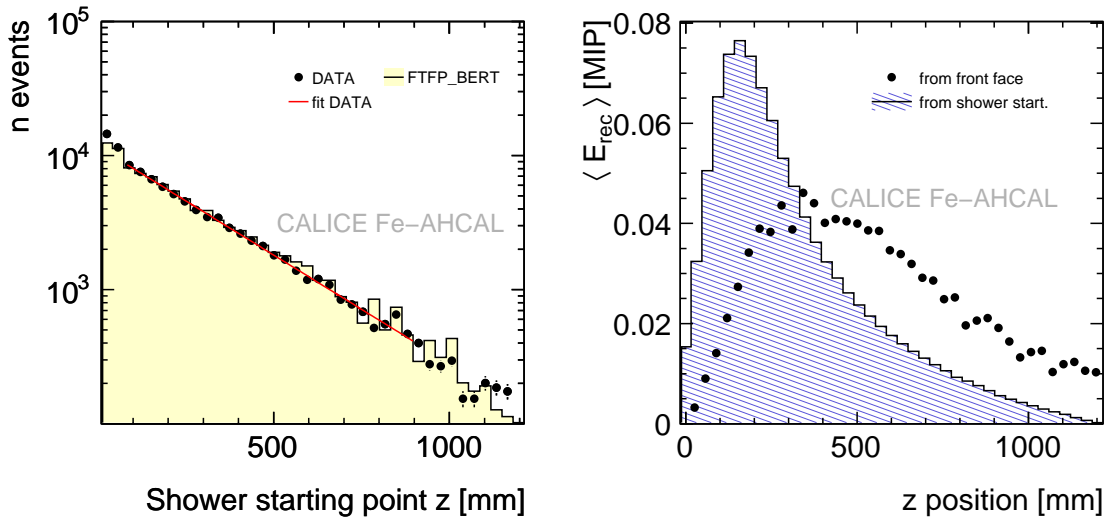


Figure 7. Left: Distribution of the measured layer of the first hard interaction for 45 GeV pions, for data (circles) and for the `FTFP_BERT` physics (histogram). The distributions are normalized to an arbitrary number. The fit to the data distribution is also drawn. Right: Longitudinal profile of 45 GeV pion showers relative to the calorimeter front face (circles) and relative to the first hard interaction (histogram). The profiles are normalized to unity.

standard deviation of the measurements performed at different energies, and the systematic error. The only source of systematics taken into account is the fit uncertainty due to choice of the fit range. Such an uncertainty has been determined by comparing the fit results for two different z ranges, namely [120, 800] mm and [90, 900] mm. The λ_π value obtained is in good agreement with the expectations from the material composition of the detector, which yield a λ_π of approximately 28 cm (from [7]: $4.28 \lambda_\pi$ for an AHCAL thickness of 120.26 cm, which yield a λ_π of 28.1 cm). This serves as a cross-check of the detector modeling used for Monte Carlo simulations, which assumes the material composition of the detector described in [7], and of the algorithm developed for the identification of the shower starting point, used in the event selection.

All Monte Carlo models except CHIPS and LHEP use the same cross-sections for hadron-nucleus interactions, which is reflected in a general agreement between the models themselves within fit uncertainties. The `QGSP_FTFP_BERT` and the Fritiof-based physics lists agree with data at approximately the 4% level at all energies, while `QBBC` agrees with data within 6%. The `QGSP_BERT` physics list is consistent with data within 4% for energies greater than 10 GeV, while at 10 GeV and 8 GeV the disagreement with data increases by up to $\sim 9\%$. CHIPS shows a behavior similar to the other models, with a good agreement with data, better than 4% at all energies. Only LHEP has an energy-dependent trend and systematically underestimates data by up to 14% at 100 GeV.

9. Longitudinal Development

Thanks to the fine longitudinal segmentation of the AHCAL into 38 layers, the longitudinal profile of hadronic showers can be investigated with an excellent accuracy. In particular, the fluctuations

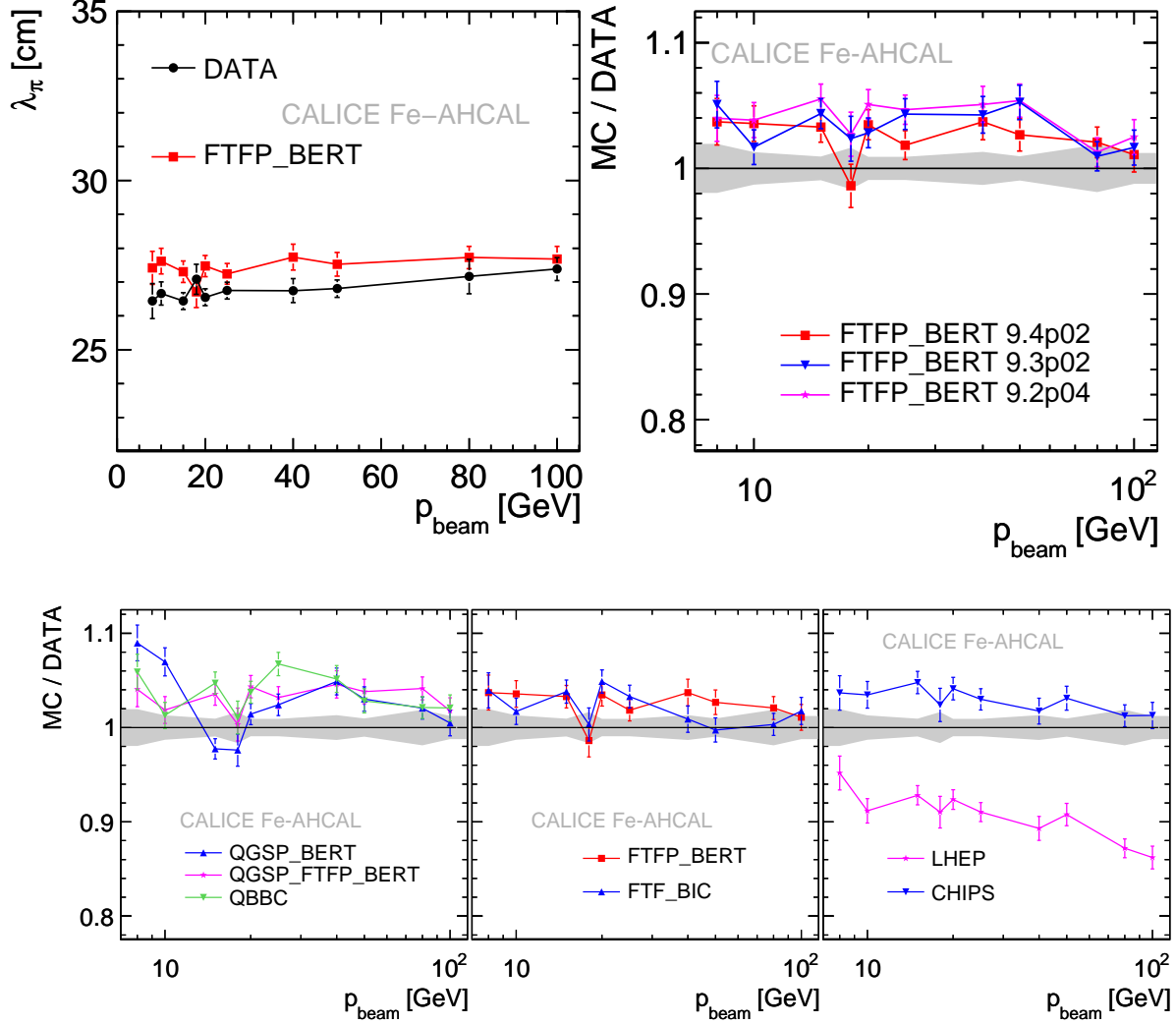


Figure 8. Summary of the measurement of the interaction length of pions in the AHCAL. Top, left: For data and for the *FTFP_BERT* physics list. Top, right: Ratio between Monte Carlo and data using the *FTFP_BERT* physics list with different versions of GEANT4. Bottom: Ratio between Monte Carlo and data for several physics lists. The gray band in the ratios represents the uncertainty on data (statistical plus systematic fit uncertainty). The error bands take into account both the statistical uncertainty and the systematic uncertainty due to the choice of the fit range.

in the shower starting point, which are large since they are due to the statistical behavior of one single particle, can be separated from the intrinsic longitudinal shower development. This is shown in Fig. 7 (right), where the shower profile relative to the calorimeter front face and the shower profile relative to the measured shower starting point are compared. The latter is shorter, since only the effective length of the hadronic shower is taken into account, excluding the path of the primary beam particle before the first hard interaction. In addition, the layer-to-layer fluctuations due to calibration uncertainties are smeared out into a smoother distribution, as different showers extend over different regions of the calorimeter. In the following, the results are therefore discussed

for profiles relative to the measured shower starting point, shown in Fig. 9. However, the trend of the agreement with data of the physics lists considered is independent from the definition of the profiles. The unit chosen to express the longitudinal development of showers is the nuclear interaction length λ_I . For the AHCAL it has been estimated to correspond to 231.1 mm.

Fig. 9 shows the longitudinal shower profiles for pions of 8 GeV, 18 GeV and 80 GeV, in data and Monte Carlo. The shown distributions are average distributions over all selected events. The profiles are normalized to unity. The normalization allows the decoupling of the pure description of the spatial development of showers from the energy response already discussed above. The three energy points have been chosen in order to partially disentangle the different contributions from low, medium and high energy models (Fig. 3). The energy contribution from electrons, positrons, pions and protons produced during the showers can be determined for the simulated events and is also shown in the graphs. The technique used for this decomposition is described in [25]. This additional information helps to understand which physics processes contribute most in which phase of the shower development. In the first layers electrons and positrons contribute about equally to hadrons to the energy deposition. In the region around the shower maximum the deposition of electromagnetic energy dominates while in the tails the energy depositions by hadrons and electrons are again about equal.

The uncertainty on the measurement of the profiles has been evaluated creating 6 profiles, for showers starting in 6 different layers of the AHCAL, namely the first 6 layers. Due to the exponential distribution of the measured layer of the first hard interaction (Fig. 7), the 6 profiles considered dominate the final measurement of the longitudinal shower profiles, which has been performed without any selection on the shower starting point. The first (second, ...) bin of the profiles for showers starting in the first layer of the AHCAL, corresponds to the first (second, ...) layer of the AHCAL. The first (second, ...) bin of the profiles for showers starting in the second layer of the AHCAL, corresponds to the second (third, ...) layer of the AHCAL. An analogous correspondence between the bins of the profiles and the AHCAL layers is true for the remaining profiles measured for showers starting in layers 3-6 of the AHCAL. The variance var_j of the energy deposited in the bin j of the profiles has been evaluated as the variance between the 6 measurements. Some bins in the tails are not covered by all the 6 measurements. For instance, showers starting in the second layer of the AHCAL extend over at most 37 AHCAL layers and give no contribution to the bin 38 of the shower profiles.

The development of showers does not depend on the measured layer of the first interaction and with an ideal calorimeter the 6 considered profiles would not differ. The measured variance is an inclusive measurement of the uncertainties associated to the energy measurement, such as bad calibrations and saturation effects. Thanks to the high statistics of several 10^4 events used to create the profiles, the variance is dominated by systematic uncertainties, while statistical fluctuations are expected to be negligible.

For the final measurement of the longitudinal shower profiles no selection on the measured layer of the first hard interaction has been applied and n_j different profiles are averaged to calculate the average energy deposition in the j -th layer, namely 37 for the first layer, 36 for the second, and so on for the remaining layers. The final uncertainty to be associated to the energy measurement in

the layer j is given by:

$$\sigma_{E_j} = \sqrt{\frac{\text{var}_j}{n_j}}. \quad (9.1)$$

This uncertainty is shown with a grey area in Fig. 9.

The position of the shower maximum is in general quite well reproduced by Monte Carlo models. In the first part of the showers QGSP_BERT, QGSP_FTFP_BERT and QBBC agree with data at the 5-10% level at 8 GeV and 18 GeV. The same is true for the Fritiof-based physics lists and for CHIPS. Larger deviations are present in the tails of the showers, but they remain within systematic uncertainties. At 80 GeV the energy deposition in the shower maximum is generally underestimated by Monte Carlo simulations, with discrepancies up to 20%. At 80 GeV CHIPS shows a moderate tendency to prefer compact showers, overestimating the energy deposition in the first part of the shower and underestimating the tail. A similar trend is exhibited by LHEP at 8 GeV and 18 GeV, though with a more pronounced disagreement with data, up to 50% at 8 GeV.

The energy dependence of the center of gravity and of the standard deviation of the longitudinal shower profiles are shown in Fig. 10 and Fig. 11, respectively. The center of gravity is defined as the energy weighted mean of the hit longitudinal coordinate along the shower axis:

$$\langle z \rangle = \frac{\sum_i E_i \cdot z_i}{\sum_i E_i}, \quad (9.2)$$

while the standard deviation is defined as:

$$\sqrt{\langle z^2 \rangle - \langle z \rangle^2} = \sqrt{\frac{\sum_i E_i \cdot (z_i - \langle z \rangle)^2}{\sum_i E_i}}, \quad (9.3)$$

where z_i is the longitudinal coordinate of the cell i and E_i is the energy measured in that cell.

The center of gravity is located between 0.8 and 1.6 interaction lengths and exhibits the expected logarithmic increase with energy. The standard deviation only mildly increases with energy from about 0.825 to about 0.925 λ_I .

All physics lists except LHEP show a similar energy-dependent behavior in the description of the center of gravity $\langle z \rangle$ of showers. The ratio Monte Carlo over data decreases with energy. Overall the best behavior is shown by the physics lists based on the Fritiof model, which agree with data at about the 4% level. The LHEP parameterized models show also for this observable the worst agreement with data. The disagreement is up to 15% at 8 GeV. A similar message is given by the standard deviation $\sqrt{\langle z^2 \rangle - \langle z \rangle^2}$.

The changes in FTFP_BERT with different GEANT4 versions are significant for both longitudinal observables and the agreement with data changes by a few percent for different versions. The ratio Monte Carlo over data for the center of gravity $\langle z \rangle$ gets generally worse when comparing versions 9.3 and 9.4, apart from the intermediate energy points between 30 GeV and 50 GeV, where the agreement is comparable. Both versions of the physics list show an improvement with respect to the older version 9.2. The agreement with data for the standard deviation $\sqrt{\langle z^2 \rangle - \langle z \rangle^2}$ improves in the version 9.3 of GEANT4, with respect to the older version. The most recent version has a worse performance than the version 9.3 at energies greater than 20 GeV, but is still significantly better than version 9.2.

10. Radial Development

An accurate modeling of the transverse shower profile is particularly important for a successful development of particle flow algorithms, since it affects the degree of overlap between showers and therefore the efficiency in disentangling single particles within jets. Using the AHCAL, it is possible to reconstruct the radial development of showers with high precision, thanks to the fine lateral segmentation of the sensitive layers.

For each cell of the AHCAL, the radial distance to the incoming particle trajectory is determined as:

$$r_i = \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2}, \quad (10.1)$$

where (x_i, y_i) are the coordinates of the center of the cell i and (x_0, y_0) is the position of the energy weighted shower center:

$$x_0 = \frac{\sum_i E_i \cdot x_i}{\sum_i E_i} \quad \text{and} \quad y_0 = \frac{\sum_i E_i \cdot y_i}{\sum_i E_i}, \quad (10.2)$$

E_i being the energy measured in the cell i .

For this study, all physical AHCAL cells are subdivided into virtual cells of $1 \times 1 \text{ cm}^2$ [21]. The energy deposited in the physical cells is equally distributed over the virtual cells covering its area. The dimension of the smallest AHCAL tiles, i.e. $3 \times 3 \text{ cm}^2$, is chosen as natural bin width for the radial shower profiles, which show the average energy deposited in the AHCAL as a function of the coordinate r . The radial shower profiles are shown in Fig. 12, for the same set of beam energies and physics lists considered for the longitudinal profiles.

All physics lists show a similar behavior and tend to underestimate the radial extent of showers, showing a relatively higher energy deposition in the core of the showers. The disagreement is more pronounced at high energies. QGSP_BERT, QGSP_FTFP_BERT and QBBC agree better with data at low energies, in particular QGSP_BERT, which is consistent with data within 5-10%. CHIPS has the best behavior at 80 GeV, describing the data at the 5-10% level. The ratios between data and Monte Carlo generally show a discontinuity at about 300-350 mm, corresponding to the transition between the tiles with a granularity of $6 \times 6 \text{ cm}^2$ and the outer tiles with a granularity of $12 \times 12 \text{ cm}^2$ in the active layers.

As expected, the contributions from electrons, pions and protons show that the electromagnetic component of the showers is concentrated in the core, while in the tails the energy is mostly deposited by protons.

The mean energy-weighted shower radius and the standard deviation of the radial energy distribution are compared for all physics lists and all energies in Fig. 13 and Fig. 14, respectively. As for the longitudinal observables, the mean is defined as:

$$\langle r \rangle = \frac{\sum E_i \cdot r_i}{\sum E_i}, \quad (10.3)$$

and the standard deviation as:

$$\sqrt{\langle r^2 \rangle - \langle r \rangle^2}, \quad \langle r^2 \rangle = \frac{\sum E_i \cdot r_i^2}{\sum E_i}, \quad (10.4)$$

where r_i is defined in Eq. 10.1.

The energy-weighted shower radius ranges between 90 mm and 60 mm and exhibits the expected exponential decrease with energy [26]. The standard deviation only mildly decreases with energy from about 80 mm to about 70 mm.

Monte Carlo models underestimate the radial extent of showers, both in terms of radius and of standard deviation. The majority of simulations exhibit an energy-dependent behavior and the disagreement with data increases with energy. QGSP_BERT reproduces the mean radius of showers at 8-10 GeV, but disagrees with data by up to about 12% at high energies. Similar results are achieved by QGSP_FTFP_BERT and QBBC, though they give a worse agreement with data at low energies. FTF_BIC underestimates the mean shower radius by about 3% at 8 GeV and by up to $\sim 12\%$ at high energies. A similar behaviour is shown by FTFP_BERT, though the disagreement between data and Monte Carlo is about 2% worse at all energies. Above 10 GeV, the best results are achieved by CHIPS, which shows the least energy dependence and underestimates the mean radial shower expansion by 2-4%. At lower energies the performance degrades and the disagreement increases by up to 5% at 8 GeV. LHEP shows a dramatic disagreement with data at all energies.

Similar trends are observed for the standard deviation of the radial energy distributions, though the agreement with data improves for all physics lists with respect to the mean radius of showers (Fig. 14).

The study of the evolution of FTFP_BERT with different GEANT4 versions (Fig. 13 and 14) shows that in the version 9.4 the simulation of the radial shower development went back to the performance achieved for the version 9.2. The intermediate version 9.3 is the closest to the data. This conclusion concerns both the observables considered.

It should be underlined that the residual discrepancies in radial profiles observed in electromagnetic showers [9], discussed above, do not prevent conclusions from being drawn concerning these radial observables and the results shown remain an important input to the process of validation of Monte Carlo models. The absolute uncertainties in the description of electromagnetic radial profiles are of the order of 2 mm, while the deviations observed in hadronic showers are of the order of 10 mm.

11. Conclusions

The response of the CALICE analog hadron calorimeter to pions is measured for energies between 8 GeV and 100 GeV, using data collected at CERN in 2007. The high sampling frequency together with the fine segmentation of the sensitive layers of this detector allows the investigation of the properties of hadronic showers and the validation of Monte Carlo models with an unprecedented spacial accuracy.

The paper covers the measurement of several properties of hadronic showers, such as the global energy response and the longitudinal and radial development of showers. The physics lists based on the Fritiof model yield overall the best simulation of the energy response, with a particularly good agreement with data at energies between 10 GeV and 30 GeV. A good simulation of the energy response at low energies is achieved also by the physics lists based on the quark-gluon string model, such as QGSP_BERT and QGSP_FTFP_BERT, which agree with data at the 2-4% level. At high energies these physics lists overestimate the energy response by up to 10%. A similar trend is exhibited by the CHIPS physics list, while the LHEP physics list underestimates the

deposited energy at all energies by up to 10%. This list has generally the worst performance also with respect to the other observables considered.

The best description of the longitudinal development of showers is achieved by the physics lists based on the Fritiof model and by CHIPS. The physics lists based on the quark-gluon string model agree with data at about the 6-8% level.

Previous publications show limitations in the understanding of the radial development of electromagnetic showers [9]. However, the results published in [22] show that despite these limitations the ability to model the hadronic shower separation is not degraded. Moreover, the discrepancies at the electromagnetic level are one order of magnitude smaller than the average radial extension of hadronic showers and do not prevent conclusions from being drawn concerning radial observables of hadronic showers. Hadronic data indicate broader showers than expected from simulation. The best performing list is CHIPS, which shows an almost energy-independent behavior and underestimates the shower mean radial expansion by less than about 2-5% at all energies. All the other physics lists with the exception of LHEP exhibit a disagreement with data increasing at high energies. The LHEP physics list shows a moderate energy-dependence, but largely disagrees with data by up to 15%. More detailed studies are needed, in order to improve the description of radial shower profiles.

FTFP_BERT and CHIPS are overall the physics lists that best agree with the observables presented. CHIPS is a very recent list, which is still under development. FTFP_BERT is an older list and the presented time evolution reflects changes in both the Bertini and the Fritiof models and their combination. Most observables fluctuate by $\pm 2-6\%$ depending on the GEANT4 version considered and remain in acceptable agreement with data. The considered radial observables present more significant changes of the order of 10% and show a better agreement with data for an older version of the physics list.

In the last decade the LHC experiments have also performed comparison studies of their test-beam data to GEANT4 v9.3.p01 models. The beam energies available in the LHC beam tests were either below 9 GeV or above 20 GeV. They concluded in [27, 28] that the physics list QGSP_BERT was the closest to their pion test-beam data. The agreement was within 2-3%, with QGSP_BERT response higher in data than in simulations. During the ATLAS test beam the lateral spread of pion showers has also been quantified. Simulated showers were found significantly narrower than in data, with a disagreement of about 15% for QGSP_BERT [29]. Since these tests several improvements have been implemented in the simulation. The results presented in this paper have extended the comparison to more recent versions of GEANT4 over a similar energy range.

Acknowledgements

We gratefully acknowledge the DESY and CERN managements for their support and hospitality, and their accelerator staff for the reliable and efficient beam operation. We would like to thank the HEP group of the University of Tsukuba for the loan of drift chambers for the DESY test beam. The authors would like to thank the RIMST (Zelenograd) group for their help and sensors manufacturing. This work was supported by the Bundesministerium für Bildung und Forschung, Germany; by the the DFG cluster of excellence ‘Origin and Structure of the Universe’ of Germany; by the Helmholtz-Nachwuchsgruppen grant VH-NG-206; by the BMBF, grant no. 05HS6VH1;

by the Alexander von Humboldt Foundation (Research Award IV, RUS1066839 GSA); by joint Helmholtz Foundation and RFBR grant HRJRG-002, SC Rosatom; by the Russian Ministry of Education and Science via grants 8174, 8411, 1366.2012.2, P220; by MICINN and CPAN, Spain; by CRI(MST) of MOST/KOSEF in Korea; by the US Department of Energy and the US National Science Foundation; by the Ministry of Education, Youth and Sports of the Czech Republic under the projects AV0 Z3407391, AV0 Z10100502, LC527 and LA09042 and by the Grant Agency of the Czech Republic under the project 202/05/0653; by the National Sciences and Engineering Research Council of Canada; and by the Science and Technology Facilities Council, UK.

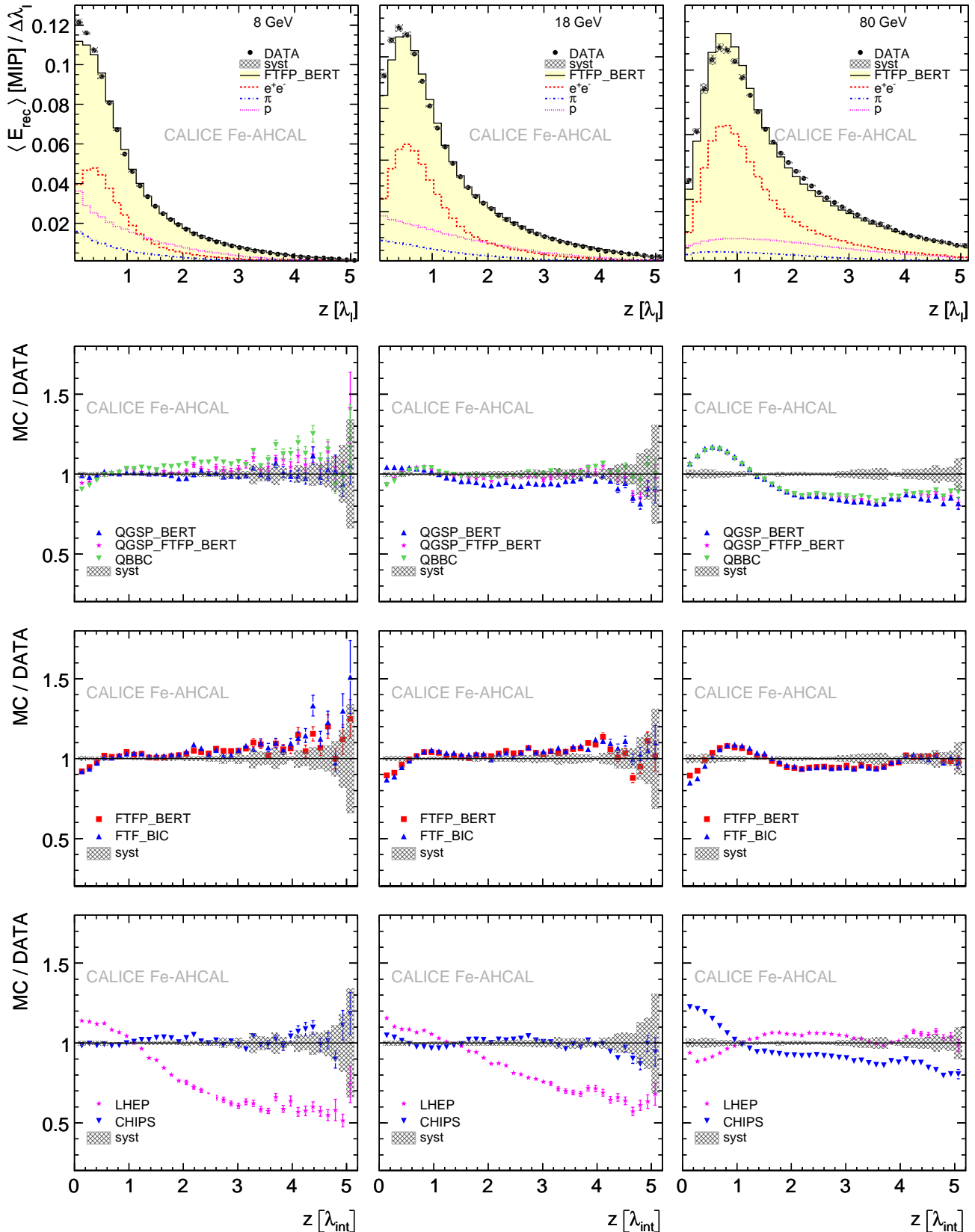


Figure 9. Mean longitudinal shower profiles from shower starting point for 8 GeV (left column), 18 GeV (center column) and 80 GeV (right column) pions. First row: For data (circles) and for the *FTFP_BERT* physics list (histogram). Second to fourth rows: Ratio between Monte Carlo and data for several physics lists. All profiles are normalized to unity. The grey area indicates the systematic uncertainty on data. $\langle E_{\text{rec}} \rangle / \Delta\lambda_1$ is the average deposited energy in a $\Delta\lambda_1$ thick transverse section of the calorimeter. z is the longitudinal

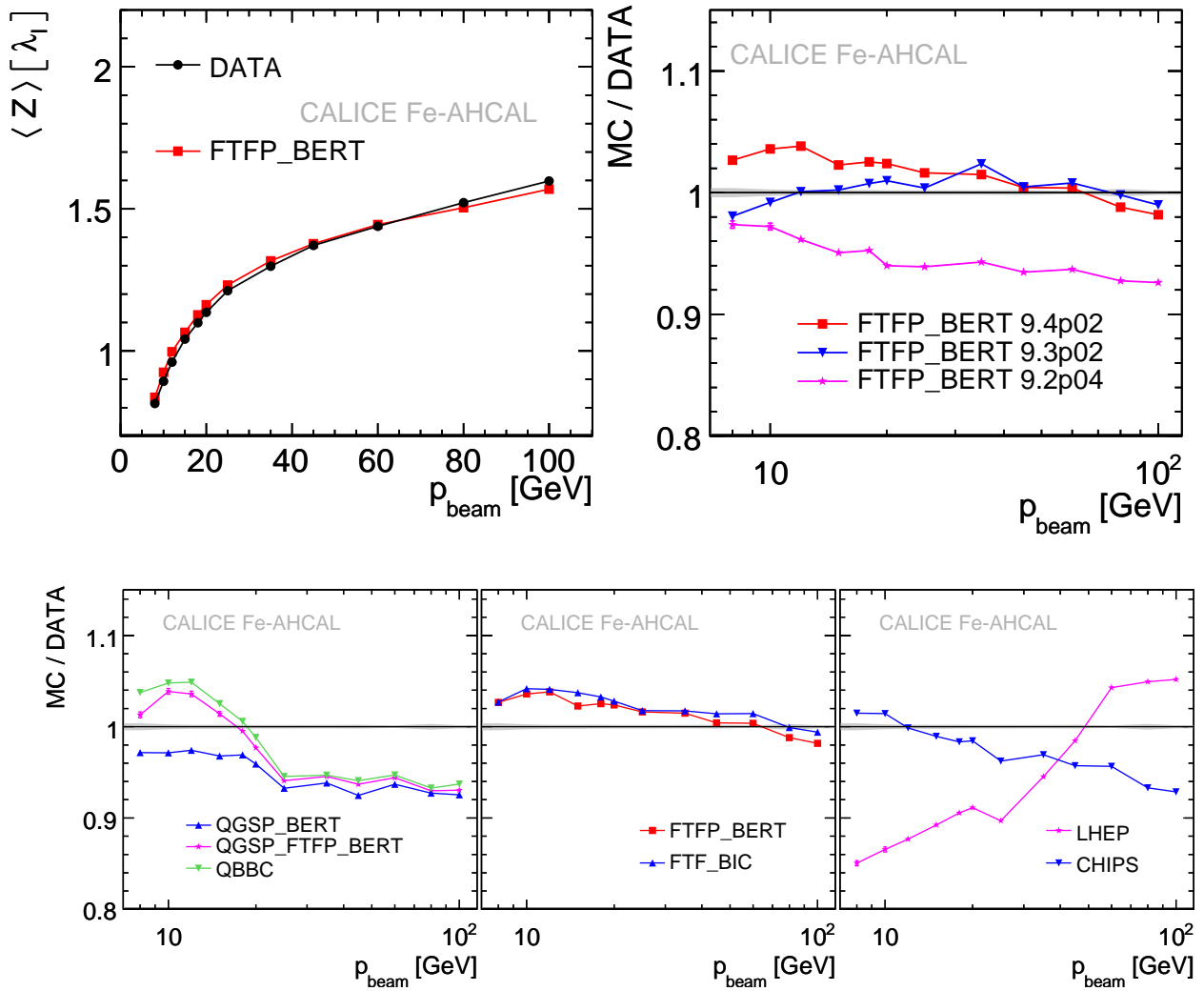


Figure 10. Summary of the measurement of the center of gravity in the longitudinal direction, for pions in the AHCAL. Top, left: For data and for the *FTFP_BERT* physics list. Top, right: Ratio between Monte Carlo and data using the *FTFP_BERT* physics list with different versions of GEANT4. Bottom: Ratio between Monte Carlo and data for several physics lists. The gray band in the ratios represents the statistical uncertainty on data.

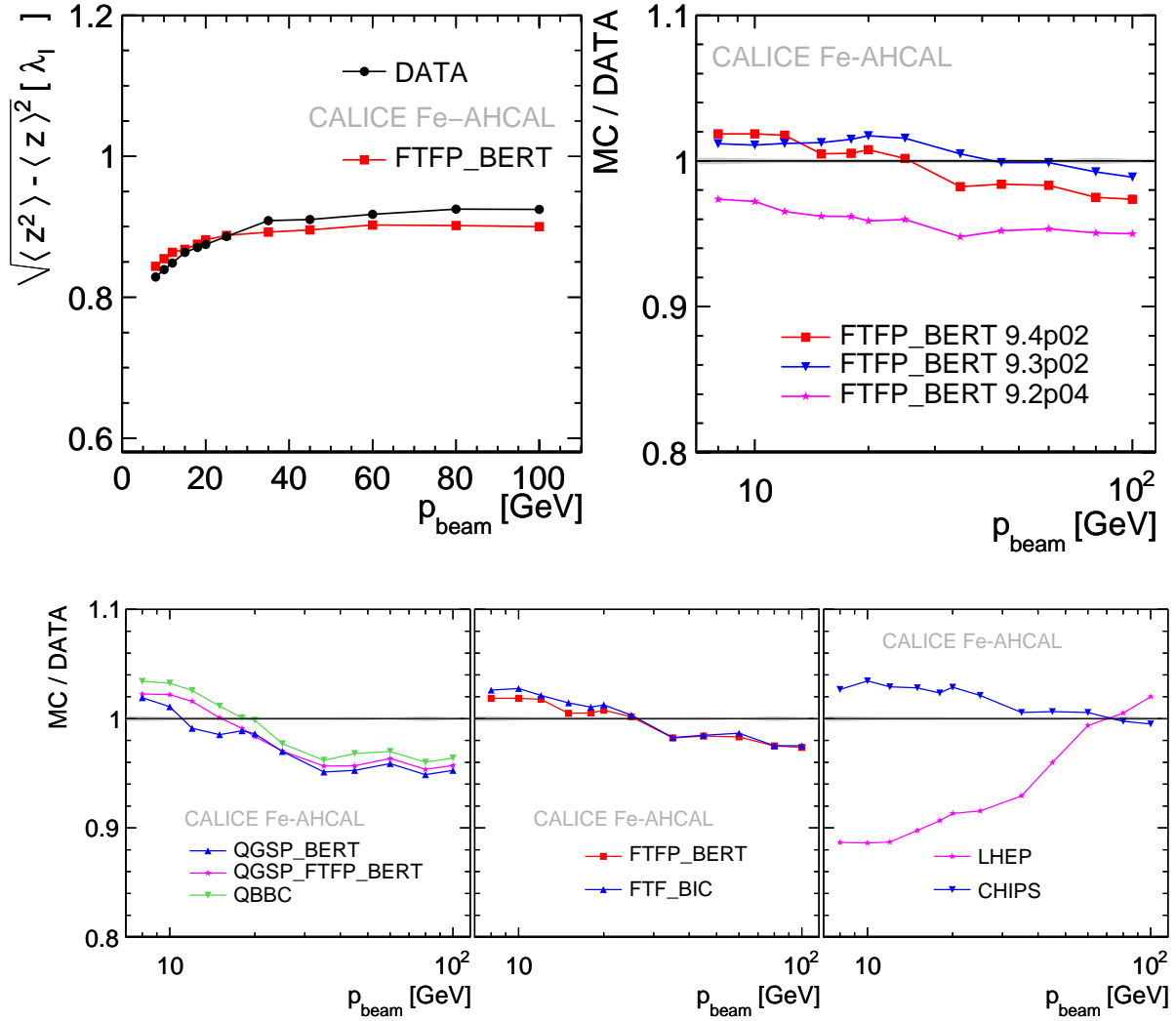


Figure 11. Summary of the measurement of the standard deviation of the longitudinal shower profile, for pions in the AHCAL. Top, left: For data and for the *FTFP_BERT* physics list. Top, right: Ratio between Monte Carlo and data using the *FTFP_BERT* physics list with different versions of GEANT4. Bottom: Ratio between Monte Carlo and data for several physics lists. The gray band in the ratios represents the statistical uncertainty on data.

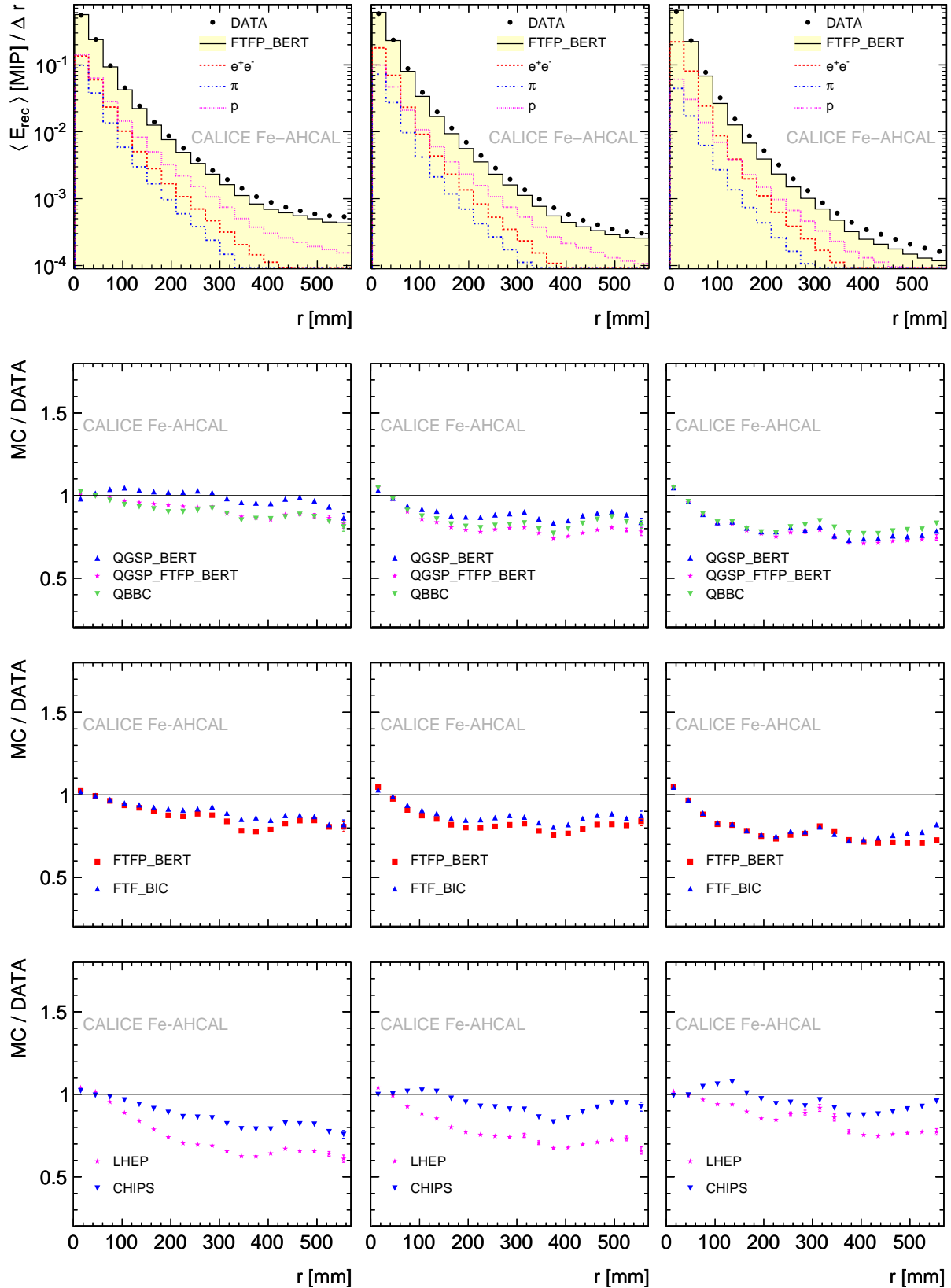


Figure 12. Mean radial shower profiles for 8 GeV (left column), 18 GeV (center column) and 80 GeV (right column) pions. First row: For data (circles) and for the FTFP_BERT physics list (histogram). Second to fourth rows: Ratio between Monte Carlo and data for several physics lists. All profiles are normalized to unity. $\langle E_{\text{rec}} \rangle / \Delta r$ is the average deposited energy in Δr , where r is the radial coordinate.

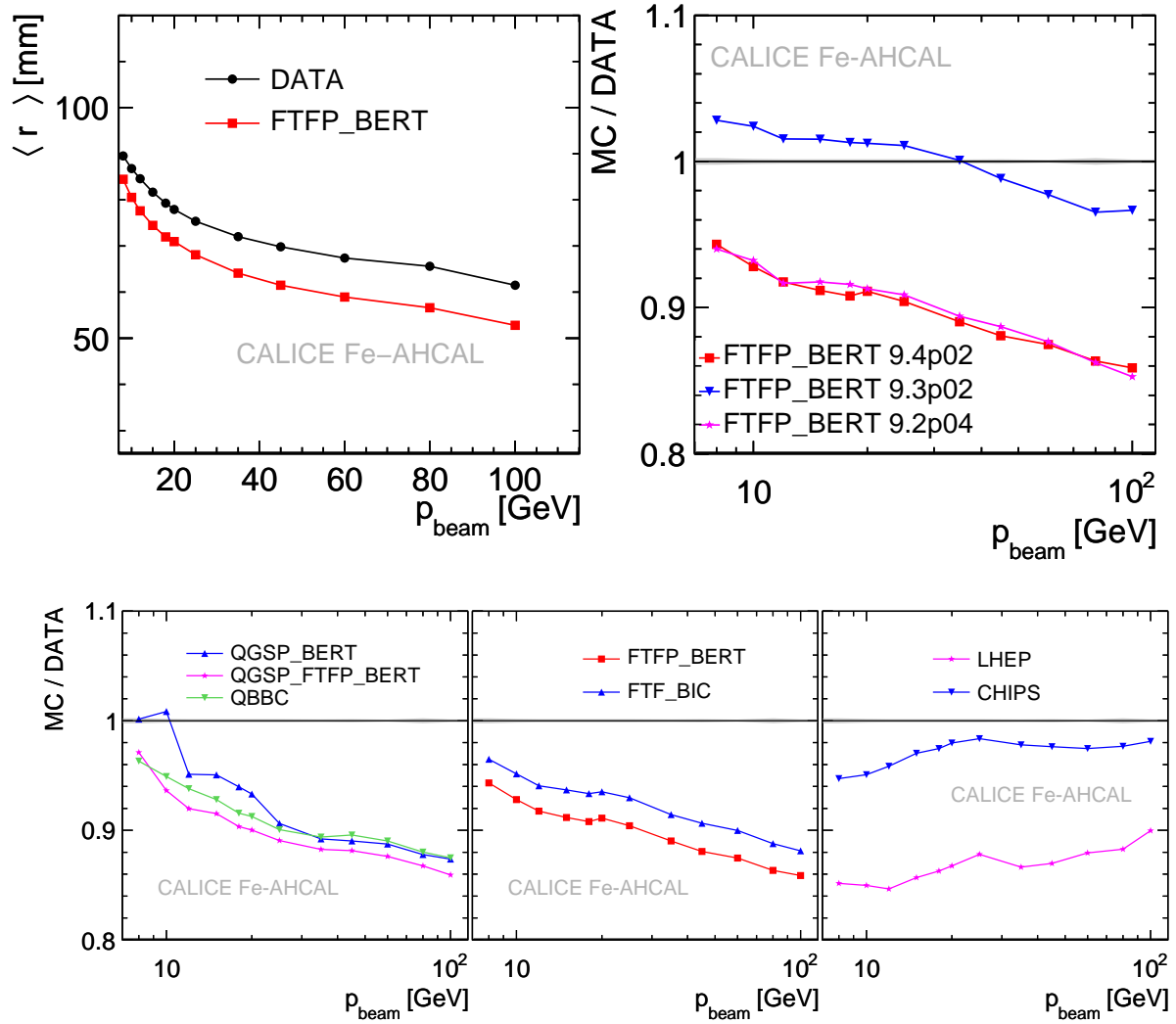


Figure 13. Summary of the measurement of the average center of gravity $\langle r \rangle$ in the radial direction, for pions in the AHCAL. Top, left: For data and for the *FTFP_BERT* physics list. Top, right: Ratio between Monte Carlo and data using the *FTFP_BERT* physics list with different versions of GEANT4. Bottom: Ratio between Monte Carlo and data for several physics lists. The gray band in the ratios represents the statistical uncertainty on data.

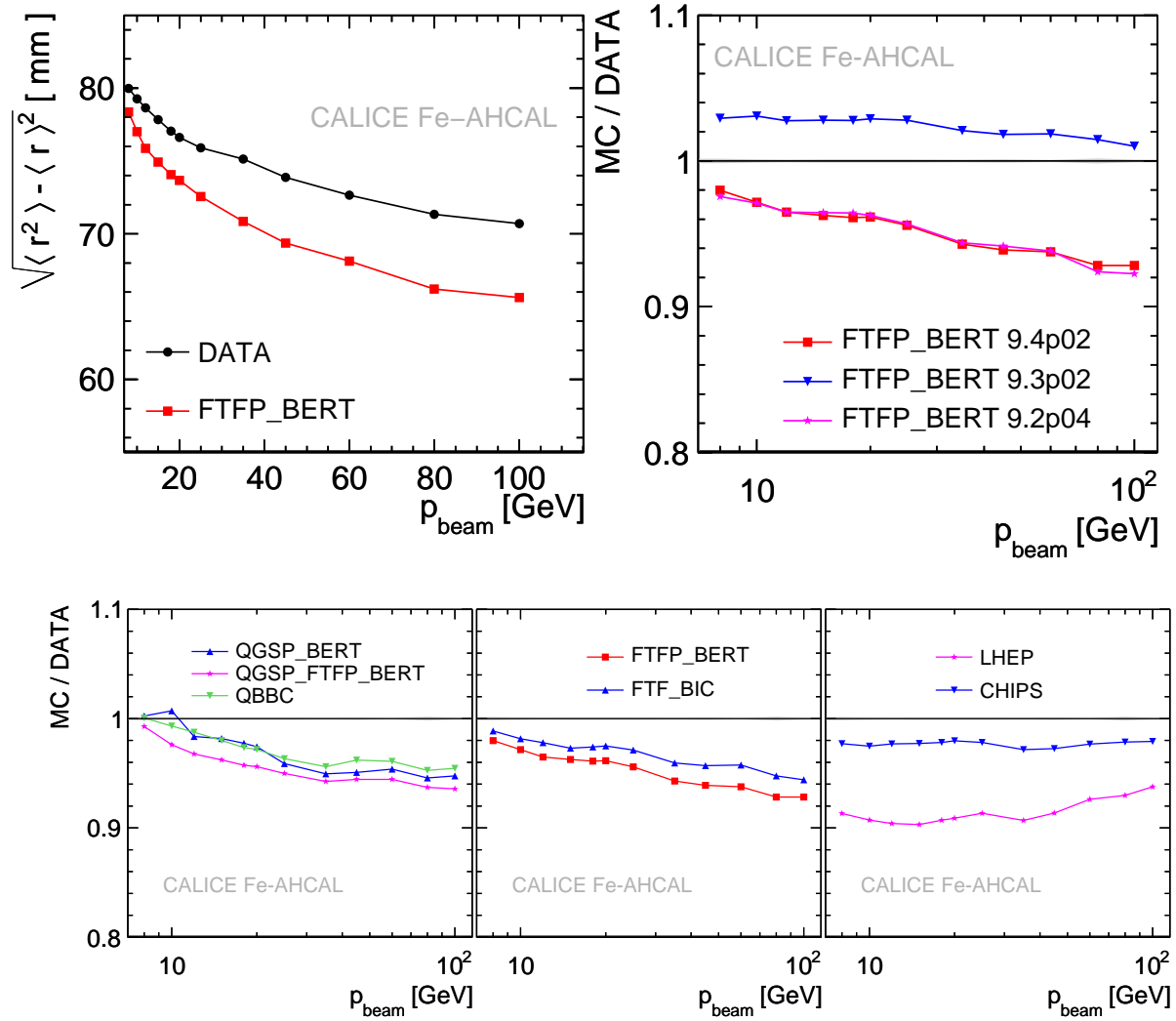


Figure 14. Summary of the measurement of the average standard deviation of the radial shower profile $\sqrt{\langle r^2 \rangle - \langle r \rangle^2}$, for pions in the AHCAL. Top, left: For data and for the *FTFP_BERT* physics list. Top, right: Ratio between Monte Carlo and data using the *FTFP_BERT* physics list with different versions of GEANT4. Bottom: Ratio between Monte Carlo and data for several physics lists. The gray band in the ratios represents the statistical uncertainty on data.

References

- [1] J. Brau, Y. Okada, N. J. Walker, A. Djouadi, J. Lykken, *et al.*, “International Linear Collider reference design report. 1: Executive summary. 2: Physics at the ILC. 3: Accelerator. 4: Detectors”, 2007.
- [2] J.-C. Briant and H. Videau, “The Calorimetry at the future $e^+ e^-$ linear collider”, *eConf*, vol. C010630, p. E3047, 2001.
- [3] V. Morgunov, “Calorimetry design with energy-flow concept (imaging detector for high-energy physics)”, CALOR 2002, Pasadena, California.
- [4] M. A. Thomson, “Particle Flow Calorimetry and the PandoraPFA Algorithm”, *Nucl. Instrum. Meth.*, vol. A611, pp. 25–40, 2009.
- [5] G. Aad *et al.*, “Single hadron response measurement and calorimeter jet energy scale uncertainty with the ATLAS detector at the LHC”, *Eur. Phys. J.*, vol. C73, p. 2305, 2013.
- [6] J. Repond *et al.*, “Design and Electronics Commissioning of the Physics Prototype of a Si-W Electromagnetic Calorimeter for the International Linear Collider”, *JINST*, vol. 3, p. P08001, 2008.
- [7] C. Adloff *et al.*, “Construction and Commissioning of the CALICE Analog Hadron Calorimeter Prototype”, *JINST*, vol. 5, p. P05004, 2010.
- [8] “Construction and performance of a silicon photomultiplier/extruded scintillator tail-catcher and muon-tracker”, *JINST*, vol. 7, p. P04015, 2012.
- [9] C. Adloff *et al.*, “Electromagnetic response of a highly granular hadronic calorimeter”, *JINST*, vol. 6, p. P04003, 2011.
- [10] G. Bondarenko *et al.*, “Limited Geiger-mode microcell silicon photodiode: New results”, *Nucl. Instrum. Meth.*, vol. A442, pp. 187–192, 2000.
- [11] P. Buzhan *et al.*, “Silicon photomultiplier and its possible applications”, *Nucl. Instrum. Meth.*, vol. A504, pp. 48–52, 2003.
- [12] N. D’Ascenzo, “Study of the neutralino sector and analysis of the muon response of a highly granular hadron calorimeter at the International Linear Collider”. DESY-THESIS-2009-004, <http://www-library.desy.de/cgi-bin/showprep.pl?desy-thesis-09-004>.
- [13] N. Feege, “Silicon photomultipliers: Properties and application in a highly granular calorimeter”. DESY-THESIS-2008-050, <http://www-library.desy.de/cgi-bin/showprep.pl?desy-thesis-08-050>.
- [14] G. Simone, “GEANT4: Simulation for the next generation of HEP experiments”, Prepared for International Conference on Computing in High-energy Physics (CHEP 95), Rio de Janeiro, Brazil, 18-22 Sep 1995.
- [15] P. Mora de Freitas and H. Videau, “Detector Simulation with Mokka and Geant4: Present and Future”. LC Note, LC-TOOL-2003-010, <http://mokka.in2p3.fr>, 2003.
- [16] <http://mokka.in2p3.fr>. Mokka Web Site.
- [17] J. B. Birks, “The theory and practice of scintillation counting”, 1964. Published in International series of Monographs on Electronics and Instrumentation, v. 27 Macmillan, New York.
- [18] F. Sefkow and A. Lucaci-Timoce, “Monte Carlo Studies of the CALICE AHCAL Tiles Gaps and Non-uniformities”, 2010. International Linear Collider Workshop 2010.

- [19] J. Apostolakis *et al.*, “Validation of GEANT4 hadronic models using CALICE data”. EUNET-MEMO-2010-15, <http://www.eudet.org/e26/e28/e86887/e109012/EUNET-Memo-2010-15.pdf>, 2010.
- [20] N. Feege, “Low-Energetic Hadron Interactions in a Highly Granular Calorimeter”. DESY-THESIS-2011-048, http://www-flc.desy.de/flc/work/group/thesis/dissertation_feege.pdf, 2011.
- [21] B. Lutz, “Hadron showers in a highly granular calorimeter”. DESY-THESIS-2010-048, <http://www-library.desy.de/cgi-bin/showprep.pl?desy-thesis-10-048>, 2010.
- [22] C. Adloff *et al.*, “Tests of a particle flow algorithm with CALICE test beam data”, *JINST*, vol. 6, p. P07005, 2011.
- [23] C. Adloff *et al.*, “Hadronic energy resolution of a highly granular scintillator-steel hadron calorimeter using software compensation techniques”, *JINST*, vol. 7, p. P09017, 2012.
- [24] I. Marchesini, “Triple Gauge Couplings and Polarization at the ILC and Leakage in a Highly Granular Calorimeter”. DESY-THESIS-2011-044, <http://www-library.desy.de/cgi-bin/showprep.pl?desy-thesis-10-048>, 2011.
- [25] A. Kaplan, “Hadronic Imaging Calorimetry”. PHD Thesis, <http://www.kip.uni-heidelberg.de/Veroeffentlichungen/download.php/4977/ps/2197.pdf>, 2011.
- [26] C. Leroy and P. Rancoita, “Physics of cascading shower generation and propagation in matter: Principles of high-energy, ultrahigh-energy and compensating calorimetry”, *Rept. Prog. Phys.*, vol. 63, pp. 505–606, 2000.
- [27] P. Adragna, C. Alexa, K. Anderson, A. Antonaki, A. Arabidze, *et al.*, “Testbeam studies of production modules of the ATLAS tile calorimeter”, *Nucl. Instrum. Meth.*, vol. A606, pp. 362–394, 2009.
- [28] E. Abat, E. Arik, S. Cetin, J. Abdallah, M. Bosman, *et al.*, “Study of the response of the ATLAS central calorimeter to pions of energies from 3 to 9 GeV”, *Nucl. Instrum. Meth.*, vol. A607, pp. 372–386, 2009.
- [29] P. Adragna, C. Alexa, K. Anderson, A. Antonaki, A. Arabidze, *et al.*, “Measurement of pion and proton response and longitudinal shower profiles up to 20 nuclear interaction lengths with the ATLAS tile calorimeter”, *Nucl. Instrum. Meth.*, vol. A615, pp. 158–181, 2010.