AIDA-PUB-2014-027

AIDA

Advanced European Infrastructures for Detectors at Accelerators

Journal Publication

Pion and proton showers in the CALICE scintillator-steel analogue hadron calorimeter

Bilki,B et al.

11 December 2014

The research leading to these results has received funding from the European Commission under the FP7 Research Infrastructures project AIDA, grant agreement no. 262025.

This work is part of AIDA Work Package **9: Advanced infrastructures for detector R&D**.

The electronic version of this AIDA Publication is available via the AIDA web site <http://cern.ch/aida> or on the CERN Document Server at the following URL: <http://cds.cern.ch/search?p=AIDA-PUB-2014-027>

Preprint typeset in JINST style - HYPER VERSION

Pion and proton showers in the CALICE scintillator-steel analogue hadron calorimeter

The CALICE Collaboration

B. Bilki[1], **J. Repond, L. Xia**
Argonne National Laboratory, 9700 S. Cass Avenue, Argonne, IL 60439-4815, USA

G. Eigen
University of Bergen, Inst. of Physics, Allegaten 55, N-5007 Bergen, Norway

M. A. Thomson, D. R. Ward
University of Cambridge, Cavendish Laboratory, J J Thomson Avenue, CB3 0HE, UK

D. Benchekroun, A. Hoummada, Y. Khoulaki
Université Hassan II Aïn Chock, Faculté des sciences. B.P. 5366 Maarif, Casablanca, Morocco

S. Chang, A. Khan, D. H. Kim, D. J. Kong, Y. D. Oh
Department of Physics, Kyungpook National University, Daegu, 702-701, Republic of Korea

G. C. Blazey, A. Dyshkant, K. Francis, J. G. R. Lima, R. Salcido, V. Zutshi
NICADD, Northern Illinois University, Department of Physics, DeKalb, IL 60115, USA

F. Salvatore[2]
Royal Holloway University of London, Dept. of Physics, Egham, Surrey TW20 0EX, UK

K. Kawagoe, Y. Miyazaki, Y. Sudo, T. Suehara, T. Tomita, H. Ueno, T. Yoshioka
Department of Physics, Kyushu University, Fukuoka 812-8581, Japan

J. Apostolakis, D. Dannheim, G. Folger, V. Ivantchenko, W. Klempt, A. -I. Lucaci-Timoce, A. Ribon, D. Schlatter, E. Sicking, V. Uzhinskiy
CERN, 1211 Genève 23, Switzerland

J. Giraud, D. Grondin, J. -Y. Hostachy, L. Morin
Laboratoire de Physique Subatomique et de Cosmologie - Université Grenoble-Alpes, CNRS/IN2P3, Grenoble, France

E. Brianne, U. Cornett, D. David, A. Ebrahimi, G. Falley, K. Gadow, P. Göttlicher, C. Günter, O. Hartbrich, B. Hermberg, S. Karstensen, F. Krivan, K. Krüger, S. Lu, B. Lutz, S. Morozov[3], **V. Morgunov**[4], **C. Neubüser, M. Reinecke, F. Sefkow, P. Smirnov, H.L. Tran**
DESY, Notkestrasse 85, D-22603 Hamburg, Germany

P. Buhmann, E. Garutti, S. Laurien, M. Matysek, M. Ramilli
Univ. Hamburg, Physics Department, Institut für Experimentalphysik, Luruper Chaussee 149, 22761 Hamburg, Germany

K. Briggl, P. Eckert, T. Harion, Y. Munwes, H. -Ch. Schultz-Coulon, W. Shen, R. Stamen

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

B. van Doren, G. W. Wilson

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

M. Wing[5]

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

C. Combaret, L. Caponetto, R. Eté, G. Grenier, R. Han, J.C. Ianigro, R. Kieffer, I. Laktineh, N. Lumb, H. Mathez, L. Mirabito, A. Petrukhin, A. Steen

Université de Lyon, Université Lyon 1, CNRS/IN2P3, IPNL 4 rue E Fermi 69622, Villeurbanne CEDEX, France

J. Berenguer Antequera, E. Calvo Alamillo, M.-C. Fouz, J. Marin, J. Puerta-Pelayo, A. Verdugo

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

F. Corriveau

Department of Physics, McGill University, Ernest Rutherford Physics Bldg., 3600 University Ave., Montréal, Quebec, Canada H3A 2T8

B. Bobchenko[6], **R. Chistov**[6], **M. Chadeeva**[6⋆], **M. Danilov**[6,7], **A. Drutskoy**[6], **A. Epifantsev, O. Markin**[6], **D. Mironov**[6,7], **R. Mizuk**[6], **E. Novikov, V. Rusinov**[6], **E. Tarkovsky**[6]

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

D. Besson, P. Buzhan, A. Ilyin, E. Popova

National Research Nuclear University MEPhI (Moscow Engineering Physics Institute) 31, Kashirskoye shosse, 115409 Moscow, Russia

M. Gabriel, C. Kiesling, N. van der Kolk, F. Simon, C. Soldner, M. Szalay, M. Tesar, L. Weuste

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

M. S. Amjad[8], **J. Bonis, S. Callier**[9], **S. Conforti di Lorenzo, P. Cornebise, F. Dulucq**[9], **J. Fleury**[9], **T. Frisson**[10], **G. Martin-Chassard**[9], **R. Pöschl, L. Raux**[9], **F. Richard, J. Rouëné, N. Seguin-Moreau**[9], **Ch. de la Taille**[9]

Laboratoire de L'accélerateur Linéaire, Centre d'Orsay, Université de Paris-Sud XI, BP 34, Bâtiment 200, F-91898 Orsay CEDEX, France

M. Anduze, V. Boudry, J-C. Brient, C. Clerc, R. Cornat, M. Frotin, F. Gastaldi, A. Matthieu, P. Mora de Freitas, G. Musat, M. Ruan[11], **H. Videau**

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

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. 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

D. Jeans

Department of Physics, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

S. Weber

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

⋆ *Corresponding author*

E-mail: marina@itep.ru

[1]*Also at University of Iowa.*
[2]*Now at University of Sussex, Physics and Astronomy Department, Brighton, Sussex, BN1 9QH, UK.*
[3]*Now at Institute for Nuclear Research RAS, Moscow, Russia.*
[4]*Also at Institute of Theoretical and Experimental Physics.*
[5]*Also at DESY and Univ. Hamburg.*
[6]*Also at National Research Nuclear University MEPhI.*
[7]*Also at Moscow Institute of Physics and Technology MIPT.*
[8]*Now at COMSATS/Pakistan.*
[9]*Now at Laboratoire OMEGA – École Polytechnique-CNRS/IN2P3, Palaiseau, F-91128 France.*
[10]*Now at CERN.*
[11]*Now at IHEP, Beijing and CERN.*

ABSTRACT: Showers produced by positive hadrons in the highly granular CALICE scintillator-steel analogue hadronic calorimeter were studied. The experimental data were collected at CERN and FNAL for single particles with initial momenta from 10 to 80 GeV/c. The calorimeter response and resolution and spatial characteristics of shower development for proton- and pion-induced showers for test beam data and simulations using GEANT4 version 9.6 are compared.

KEYWORDS: hadronic calorimeter; hadronic shower; calorimeter response; calorimeter resolution.

Contents

1. Introduction

A wide range of highly granular calorimeter prototypes has been developed by the CALICE collaboration to test the particle flow concept as well as new technologies for future particle physics experiments. The particle flow approach (PFA) was proposed in order to achieve the jet energy resolution required for future linear collider experiments [1]. Recently, a PFA was successfully implemented for jet energy reconstruction in the CMS detector [2] and a further increase of granularity is now considered an option for the CMS calorimeter upgrade to provide a PFA-ready detector for the third stage of the LHC experiments. Besides testing the PFA, highly granular electromagnetic and hadronic calorimeter prototypes provide an opportunity to test Monte Carlo models with unprecedented detail. While the development of electromagnetic showers is quite well understood and reproduced by simulations, predictions of hadronic shower development are not so precise and there are no hadronic models which demonstrate agreement with data for all types of hadrons over a wide energy range.

Hadronic showers, produced in a calorimeter after a deep inelastic interaction of an incident hadron with a nucleus, are characterised by a relatively narrow core from the electromagnetic component surrounded by an extended halo. The core is usually formed by electromagnetic cascades

initiated by photons from π^0 decays, while charged mesons and baryons dominate in the radial halo and longitudinal tail of the shower. The complicated structure of hadronic showers results in significant fluctuations of their longitudinal and radial sizes as well as the calorimeter response [3].

Besides fluctuations of the hadronic response, differences in the average calorimetric response for different types of hadrons have been predicted [4] and observed experimentally for pions and protons [5]. This can be mainly explained by the baryon number conservation law and the associated different available energy for mesons and baryons. The response for pions is $\sim$10% larger than that for protons in the energy range 200 to 375 GeV [5]. Such behaviour was confirmed for lower energies (20 to 180 GeV) with the Fe-Scintillator ATLAS Tile calorimeter [6] for which the response to pions was $\sim$4% higher than protons of the same initial energy.

The CALICE scintillator-steel analogue hadron calorimeter (Fe-AHCAL) is the first example of the large-scale application of silicon photomultipliers (SiPM) in the field of high energy physics [7]. This calorimeter has high longitudinal and transverse granularity and provides an opportunity for detailed study of hadronic shower development. Previous studies of pion-induced showers in the Fe-AHCAL include calorimeter response and resolution, as well as comparisons of pion shower profiles and spatial characteristics with simulations using GEANT4 version 9.4 [8, 9]. Identification of the position of the first inelastic interaction (shower start) in the calorimeter disentangles the energy density distribution inside the shower from the distribution of the shower start. Hence, the direct comparison of showers induced by mesons and baryons becomes possible.

This paper describes the analysis of hadronic showers induced by positive pions and protons with initial energies from 10 to 80 GeV in the CALICE Fe-AHCAL. The shower parameters extracted from data are compared with simulations using physics lists from GEANT4 version 9.6 [10]. The experimental setup, event selection procedure, and systematic uncertainties are described in Section 2. Data-MC comparisons of global observables, such as deposited energy, energy resolution, shower radius, and longitudinal centre of gravity, are presented in Section 3.

2. Experimental data and simulations

2.1 Experimental setup

The data analysed here using beams of positively charged hadrons were collected at CERN in 2007 and at FNAL in 2009. The CALICE setup at CERN is described in detail in [8] and comprised the silicon-tungsten electromagnetic calorimeter (Si-W ECAL), the Fe-AHCAL, and the scintillator-steel tail catcher and muon tracker (TCMT). Beams of positive hadrons in the momentum range from 30 to 80 GeV/c were delivered with the CERN SPS H6 beam line. The CALICE setup during the test beam campaign at FNAL is described in detail in [11] and comprised the Fe-AHCAL and TCMT. The data at FNAL for positive hadrons with initial momenta of 10 and 15 GeV/c from the MTest beam line were collected without an electromagnetic calorimeter. Both configurations included Čerenkov counters placed upstream of the calorimeters. The information from the Čerenkov counters was used for off-line discrimination between pions and protons on an event-by-event basis. Data collected with normal incidence of the beam with respect to the calorimeter front plane are used for the current analysis.

The Si-W ECAL is a highly granular sampling electromagnetic calorimeter [12] comprised of 30 layers (three sections with different absorber thicknesses) and has a depth of approximately one

nuclear interaction length $\lambda_I^{\rm eff}$. The Si-W ECAL has a very fine transverse segmentation equivalent to 1×1 cm^2 cells, the transverse size of its active zone is $\sim18\times18$ cm^2. In this analysis, the electromagnetic calorimeter was used in the event selection procedure and as to obtain the track of an incoming hadron.

The Fe-AHCAL is a sampling structure of 38 active layers interleaved with absorber plates (21 mm of stainless steel per layer). The full transverse size of the calorimeter is 90×90 cm^2. Each active layer is assembled from 5 mm thick scintillator tiles of varied transverse sizes: 3×3 cm^2 in the central part, 6×6 cm^2 in the surrounding region, and 12×12 cm^2 in the peripheral region. Each tile is individually read out by a silicon photomultiplier (SiPM). The longitudinal depth of the Fe-AHCAL is $\sim5.3\lambda_I^{\rm eff}$ ($\sim0.14\lambda_I^{\rm eff}$ per layer). The calibration procedure for the Fe-AHCAL is described in [7]. The calorimeter was positioned so that the beam struck the calorimeter centre tiles to minimise lateral leakage.

The TCMT is also a sampling calorimeter with 16 active layers assembled from scintillator strips with SiPM readout [13]. The first section (9 layers) of the TCMT has 2 cm thick absorber plates and the same sampling fraction as the Fe-AHCAL. The absorber thickness of the second TCMT section is larger by a factor of five than in the first section. The total depth of the TCMT amounts to $\sim5.5\lambda_I^{\rm eff}$.

The visible signal in each calorimeter cell is obtained in units of minimum-ionising particle (MIP). Only cells with a signal above 0.5 MIP were considered for further analysis; a cell above threshold is called hit.

2.2 Event selection

The total number of collected events, percent contamination, size of selected samples, and proton sample purities are shown in Table 1. The event selection procedure comprises sample cleaning, shower start and primary track finding, selection by shower start, and pion and proton separation. The sample cleaning includes the rejection of multi-particle events, muons, and positrons.

The fraction of multi-particle events in the CERN test beam does not exceed 0.3% while for the FNAL data samples the fraction is as high as 14% (see Table 1). An additional cut limits the total number of hits in the Fe-AHCAL to less than 165 and less than 220 for 10 and 15 GeV, respectively. This cut reduces the fraction of multi-particle events to the percent level. A cross-check with simulations and the CERN data for negative pions of the same energy shows that these cuts reject less than 0.2% of hadron events.

The muon identification algorithm involves information from all calorimeter sections and is based on the comparison of the energy deposition in the combined ECAL+AHCAL and TCMT. The efficiency of muon identification is better than 99.5% in the energy range studied. The highest muon contamination of $\sim$30% was observed for 30 GeV beams at CERN, while for other energies it did not exceed 5%.

The CERN test beam data (30 GeV/c and above) were taken with the Si-W ECAL in front of the Fe-AHCAL. For these data samples, the requirement of a track in the electromagnetic calorimeter is enough to fully remove positron contamination, which is assumed to be negligible in the selected samples. Additional constraints to reject positrons are required for the FNAL data taken without an electromagnetic calorimeter and without the Čerenkov trigger, which was set to separate pions from protons. The criteria for positron identification involves cuts on the shower radius,

Table 1. Positive hadron data samples used in the analysis.

Beam momentum GeV/c	Total number of events	Fraction of μ^+	Fraction of e^+	Fraction of multi-particle	Number of selected π^+	Number of selected protons	Purity of selected proton sample	with ECAL in front
10	45839	3.0%	37.2%	13.2%	5275	1239	0.74±0.13	no
15	46323	3.8%	19.9%	13.8%	6660	2122	0.80±0.09	no
30	192066	30.1%	n/e*	0.2%	10838	7714	0.95±0.01	yes
40	201069	4.6%	n/e	0.3%	20936	4799	0.84±0.06	yes
50	199829	4.4%	n/e	0.3%	21151	4192	0.79±0.06	yes
60	208997	3.8%	n/e	0.3%	21133	5759	0.85±0.05	yes
80	197062	2.8%	n/e	0.3%	16964	8545	0.83±0.04	yes

* n/e - not estimated because of the requirement of a track in the electromagnetic calorimeter for selected hadron events.

R, (see Section 3.4) and event longitudinal centre of gravity, Z, (see Section 3.3), as electromagnetic showers are known to be more compact than hadronic showers. The shower is considered to be induced by a positron if $R < 37$ mm and $Z < 260$ mm. The efficiency of the criteria was estimated using dedicated positron runs taken at CERN without an electromagnetic calorimeter and was found to be ∼96% at 10 GeV and ∼98% at 15 GeV. The application of these cuts to the negative pion samples extracted from CERN data (taken with the Si-W ECAL in front) results in pion rejection of less than 0.8%. The total estimated positron contamination is shown in Table 1. The cuts mainly affect pion samples and have negligible impact on the proton samples.

The identification of the position of the primary inelastic interaction (shower start) is based on the threshold analysis of the visible energy and the number of hits detected in several consecutive layers [14]. The thresholds were optimised using simulated samples. The algorithm finds the shower start within ±1 layer of the true layer for ∼80% of hadron events in the energy range studied, as determined with MC (See Appendix A). The distributions of the found shower start layer were used to perform a cross-check of the algorithm by extracting the effective nuclear interaction lengths $\lambda_\pi^{\rm eff}$ and $\lambda_{\rm I}^{\rm eff}$ for pions and protons, respectively (see Appendix A for details). The energy dependencies of these estimates are shown in Fig. 1. The systematic uncertainties for $\lambda_{\rm I}^{\rm eff}$ include uncertainties due to pion contamination of the proton samples. The extracted values for both pion and proton interaction lengths are in good agreement with simulations and with the estimates of the nuclear interaction length for the CALICE Fe-AHCAL based on the PDG values [15]. The averaged ratio $\lambda_\pi/\lambda_{\rm I}$ in the studied energy range was estimated to be 1.26±0.02. As expected, no energy dependence for the nuclear interaction length is observed.

A minimum-ionising particle track in the electromagnetic calorimeter is required. In addition, to minimise leakage into the TCMT, showers were selected which start developing at the beginning of the Fe-AHCAL, namely, in any of the physical layers from 2 to 5 (from 3 to 6) for the configuration with (without) an electromagnetic calorimeter. Events with an identified shower start

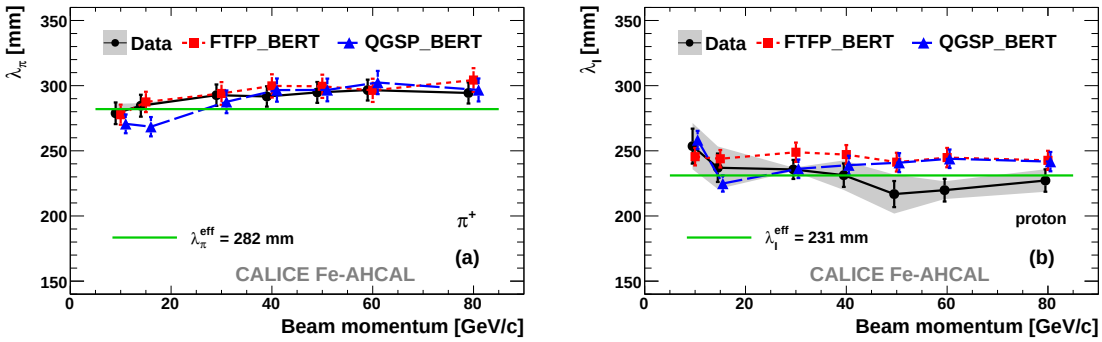

Figure 1. Nuclear interaction length (a) λ_π for pions and (b) λ_I for protons in the CALICE Fe-AHCAL extracted from found shower start for data samples (black circles) and simulations using the FTFP_BERT (red squares) and QGSP_BERT (blue triangles) physics lists of GEANT4 9.6. Error bars show the uncertainties from fit. The green solid line corresponds to the effective nuclear interaction length calculated from PDG data. Systematic uncertainties for data are shown with a grey band (not visible in the left plot).

in the first physical layer of the Fe-AHCAL were excluded from the analysis due to uncertainties associated with the shower start identification. The exclusion of events with a shower start in the first and second AHCAL layer significantly reduces the fraction of remaining positrons in the data samples taken without an electromagnetic calorimeter (to ∼2.5% at 10 GeV and less than 1% at 15 GeV). After requiring a shower start at the beginning of the hadronic calorimeter, the contamination of the selected samples by muons does not exceed 0.1% for all energies and the admixture of multi-particle events is less than one percent. The same procedure of shower start identification and selection by shower start was applied to the data and simulated samples.

Pion identification in the test beam experiments fully relies on the Čerenkov counter efficiency. As the pressure in the gaseous Čerenkov detector used was set well below the proton threshold, we assume here that the probability of proton contamination in the pion samples is negligible. The estimate of the proton sample purity from pions, η, is based on the independent muon identification, as described in Appendix B.

2.3 Monte Carlo simulations

Simulations were done using the physics lists QGSP_BERT and FTFP_BERT from GEANT4 version 9.6 patch 1 [10]. The physics list QGSP_BERT is widely used for simulation in the LHC experiments and has demonstrated the best agreement with data in earlier versions (e.g. version 9.2, as mentioned in [16]). The QGSP_BERT physics list is maintained due to its wide use. The physics list FTFP_BERT was significantly improved in version 9.6 and is now recommended for HEP simulations by the GEANT4 collaboration [17].

The QGSP_BERT physics list employs the Bertini cascade model (BERT) below 9.5 GeV, the low energy parametrised model (LEP) in the intermediate energy region, and the quark-gluon string precompound model (QGSP) above 25 GeV. The transition regions between models are from 9.5 to 9.9 GeV and from 12 to 25 GeV. The FTFP_BERT physics list uses the Bertini cascade model for low energies and the Fritiof precompound model (FTFP) for high energies with the transition region from 4 to 5 GeV. A detailed description of the models can be found in [18].

Simulated samples were digitised taking into account the SiPM response, light crosstalk between neighbouring scintillator tiles in the same layer, and calorimeter noise extracted from data. The digitisation was validated using the electromagnetic response of the Fe-AHCAL [19]. The test beam profile and its position on the calorimeter front face in each data run were well reproduced in simulations.

2.4 Systematic uncertainties

Uncorrected systematics

Calculation of the reconstructed energy and resolution (see Section 3.1) requires a conversion from MIP response to the GeV energy scale. The conversion coefficient from MIP to GeV for the Fe-AHCAL (electromagnetic calibration) was extracted from dedicated positron runs with a systematic uncertainty of 0.9% [19]. Other contributions (such as the saturation correction) were studied by varying the calibration constants within allowed limits and were found to be negligible. The same conversion coefficient was applied to data and simulated samples.

The impact of shower start uncertainties on the observables was found to be negligible for simulated samples. The systematic uncertainty of the ratio of simulation to data is assumed to cancel, as the same algorithm is applied in both cases. This is supported by the fact that the estimates of the nuclear interaction lengths are in good agreement between data and simulations (see Fig. 1).

The spatial observables are still affected by leakage due to the limited Fe-AHCAL depth ($5.3\lambda_{\mathrm{I}}$) in spite of the applied shower start selection. As shown with MC, biases from leakage are negligible below 20 GeV and do not exceed a few percent at 80 GeV. The main impact of the leakage is on the longitudinal and radial dispersions. Again, the bias has negligible impact on the comparison of data and simulation.

Corrected bias due to contamination

The admixture of particle species in the samples introduces a bias on the observables, which can be corrected if the sample purity and parameters of the contaminating sample are available. A measured value, A_{meas}, is obtained from the contaminated sample with purity, η, estimated as described in Appendix B. The value of the parameter for the contaminating admixture, A_{cont}, is determined independently from a pure sample of contaminating particles. The corrected value, A_{corr}, can be calculated as follows:

$$A_{\mathrm{corr}} = A_{\mathrm{meas}}\frac{1}{\eta} + A_{\mathrm{cont}} \cdot \left(1 - \frac{1}{\eta}\right). \tag{2.1}$$

The positron contamination in the pion samples taken without the ECAL results in an overestimation of the mean reconstructed energy because the Fe-AHCAL is a non-compensating calorimeter, while the mean shower depth and width are underestimated. The purity of the analysed pion samples with respect to positrons is 0.975 ± 0.015 and 0.99 ± 0.01 at 10 GeV and 15 GeV, respectively. The values of the observables for positrons in the Fe-AHCAL were extracted from the dedicated positron runs [19]. The most significant correction for bias due to positron contamination is 2.5% for the mean shower radius of 10 GeV pions. For 15 GeV pion samples, the estimated

bias does not exceed 1% for all observables. The positron contamination of the proton samples is negligible, as protons were selected by requiring no signal in the Čerenkov counter.

Pion contamination of the proton samples was caused by inefficiency of the Čerenkov counters. The estimated purity of the proton samples varies from 74% to 95% (see Table 1). The admixture of pions in the proton samples results in an overestimate of the reconstructed energy and an underestimate of the longitudinal and radial sizes of the proton showers. The largest biases due to pion contamination are observed for the sample of 10 GeV protons and the corresponding corrections are 4.8%, 1.3%, and 2.8% for the reconstructed energy, longitudinal centre of gravity, and mean shower radius, respectively. The contamination of the selected samples with muons and multi-particle events is negligible and does not need correction.

3. Comparison of observables

3.1 Calorimeter response and p/π ratio

The total deposited energy of a particle in units of GeV is calculated from the visible signal in units of MIP measured in different detector sections and multiplied by a suitable calibration factor for each detector. The Fe-AHCAL is a non-compensating calorimeter with a lower response to hadrons than to electrons. The electromagnetic calibration factor $w = 0.02364$ MIP/GeV for the Fe-AHCAL is extracted from dedicated positron runs [19]. The scaling factor, which determines the hadronic energy scale for the Fe-AHCAL, is on average 1.19 over the energy range studied [8]. The selection conditions include a track in the Si-W ECAL, hence the energy deposited in the electromagnetic calorimeter $E_{\mathrm{ECAL}}^{\mathrm{track}}$ is that of a minimum ionising particle. The corresponding calibration factor was extracted from data and simulated muon samples in the Si-W ECAL. As the total depth of the Fe-AHCAL is about $\sim 5.3\lambda_{\mathrm{I}}^{\mathrm{eff}}$, the signal from the TCMT is added to calculate the reconstructed energy in each event as follows:

$$E_{\mathrm{event}} = E_{\mathrm{ECAL}}^{\mathrm{track}} + E_{\mathrm{HCAL}} + E_{\mathrm{TCMT}}, \tag{3.1}$$

where E_{HCAL} and E_{TCMT} are deposited energies in the Fe-AHCAL and TCMT calorimeters obtained with the electromagnetic scale. Since the first nine TCMT layers are essentially identical to the Fe-AHCAL layers in terms of absorber and active material, the same electromagnetic calibration factor is assumed. For the last seven TCMT layers, the calibration factor is adjusted according to the increased absorber thickness. The value of $E_{\mathrm{ECAL}}^{\mathrm{track}}$ for the incident energies from 30 GeV and above accounts for less than 1.4% of the reconstructed energy ($\sim$0.35 GeV).

The reconstructed energy distributions were fitted with a Gaussian curve in the interval of ± 2 r.m.s. around the mean value. Hereafter, the parameters of this Gaussian fit at a given beam energy are referred to as the mean reconstructed energy E_{reco} and resolution σ_{reco}.

Two examples of the reconstructed energy distribution are shown in Fig. 2 for pions and protons of 10 and 80 GeV together with the predictions of the FTFP_BERT physics list. In agreement with the earlier published results [4, 5, 6], the reconstructed energy for protons is lower than that for pions. The difference increases with decreasing initial particle energy. This can largely be explained by baryon number conservation that results in lower probability to produce a leading baryon in the interaction of a pion with a nucleus. The measurable energy is different for pions and

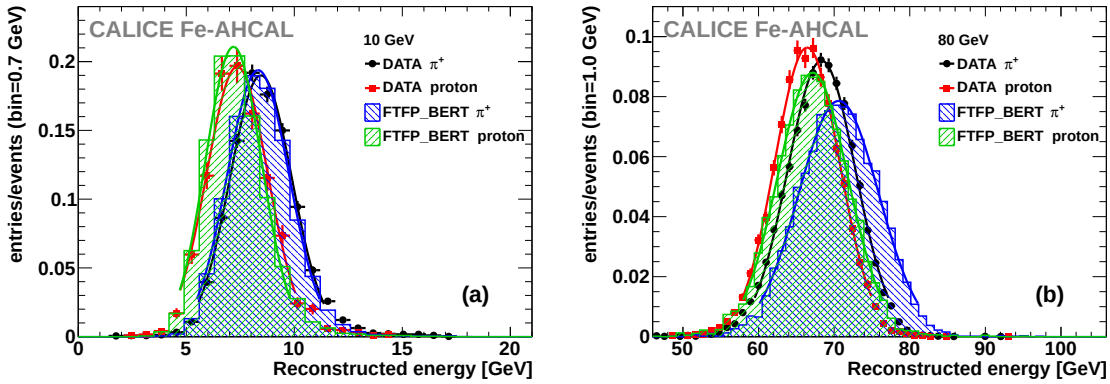

Figure 2. Reconstructed energy distributions for pions and protons with initial energies (a) 10 and (b) 80 GeV for data (points) and simulations using the FTFP_BERT physics list (hatched histograms). The solid curves are Gaussian fits to the data and simulations. Error bars show the statistical uncertainties.

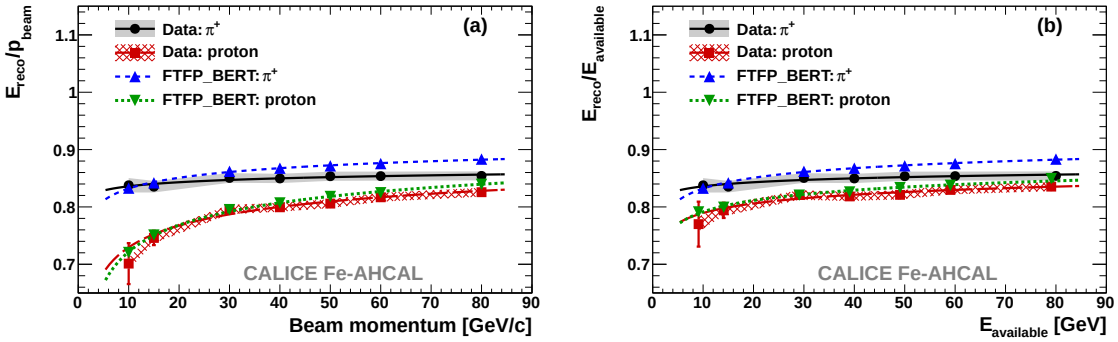

Figure 3. Ratio of the mean reconstructed energy E_{reco} to (a) beam momentum and (b) available energy for data and the FTFP_BERT physics list. The grey and crosshatched red bands show the systematic uncertainties for pion and proton data, respectively. Error bars show the statistical uncertainties. The curves correspond to the power-law approximation. See text for details.

protons and corresponds to the total particle energy in the case of mesons and to the kinetic energy in the case of baryons:

$$E^{\mathrm{proton}}_{\mathrm{available}} = \sqrt{p^2_{\mathrm{beam}} + m^2_{\mathrm{proton}}} - m_{\mathrm{proton}}, \tag{3.2}$$

where p_{beam} is the beam momentum and m_{proton} is the proton rest mass.

Figures 3 and 4 show the ratios of the mean reconstructed energy to beam momentum and to the available energy, respectively. In the latter case, the difference between positive pion and proton response remains at the level of ∼4%, in agreement with the difference observed in [6] for the Sc-Fe Tile ATLAS calorimeter. The FTFP_BERT physics list gives better predictions of the response for pions than QGSP_BERT and very good predictions for protons.

The phenomenological interpretation of the observed response behaviour is usually done in terms of the mean electromagnetic fraction f_{em} within a hadron-induced shower and the mean efficiencies e and h of the reconstruction of the electromagnetic and hadronic energy component, respectively. The factor e also defines the electromagnetic scale and can be extracted from electro-

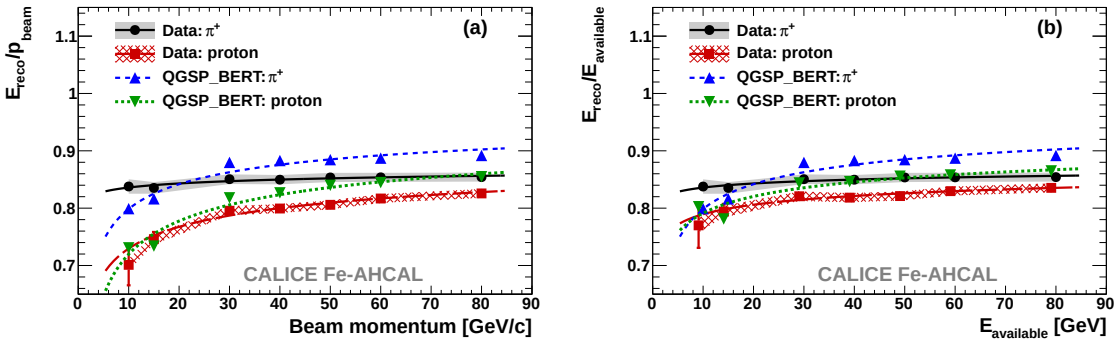

Figure 4. Ratio of the mean reconstructed energy E_{reco} to (a) beam momentum and (b) available energy for data and the QGSP_BERT physics list. The grey and crosshatched red bands show the systematic uncertainties for pion and proton data, respectively. Error bars show the statistical uncertainties. The curves correspond to the power-law approximation. See text for details.

magnetic calibration. The corresponding mean hadronic fraction is $f_{\mathrm{h}} = 1 - f_{\mathrm{em}}$. In the frame of such an approach, the mean reconstructed energy can be expressed as follows:

$$E_{\mathrm{reco}} = \frac{1}{e} \cdot E_{\mathrm{beam}} \cdot (e \cdot f_{\mathrm{em}} + h \cdot f_{\mathrm{h}}), \tag{3.3}$$

where E_{reco} is the mean reconstructed energy of pions (E_{π}) or protons (E_{p}) measured in the electromagnetic scale and E_{beam} is the beam energy. Considerations about the cascade shower development and constraints on π^0 production (see e.g. [20]) lead to the assumption of a "power law" scaling of the mean hadronic fraction: $f_{\mathrm{h}} \approx (E_{\mathrm{beam}}/E_0)^{m-1}$, where E_0 is the energy at which multiple pion production becomes significant and which is expected to be different for pions and protons. Applying this scaling to Eq.(3.3) gives the following representation of the calorimeter response [20]:

$$\frac{E_{\mathrm{reco}}}{E_{\mathrm{beam}}} = 1 - (1 - h/e) \cdot f_{\mathrm{h}} = 1 - a \cdot E_{\mathrm{beam}}^{m-1}, \quad a = (1 - h/e) \cdot E_0^{1-m}, \tag{3.4}$$

where a and m are free parameters to be determined. The energy dependencies of the response shown in Fig. 3 and 4 were approximated with Eq.(3.4). The fits to data and FTFP_BERT in the studied energy range resulted in $\frac{\chi^2}{\mathrm{NDF}} < 1$. The behaviour of the response, predicted by the QGSP_BERT physics list, exhibits variations in the model transition region around 10–15 GeV, hence no good fit can be achieved. The values of the parameters a and m obtained from fits to pion and proton responses are shown in Table 2.

The power-law parametrisation provides enough flexibility to describe a wide range of dependencies with different representations of response, for instance for protons. In the power-law context, the parameter m is expected to be the same for pions and protons [20]. The values of the parameter m, predicted by the FTFP_BERT physics list for pions and protons, agree within uncertainties if we consider the dependence on available energy. The values of m extracted from data tend to be higher than those extracted from simulations, though for protons they are in agreement within uncertainties.

Table 2. Parameters m and a, obtained from the fit of Eq.(3.4) to the energy dependencies of pion and proton response for CALICE test beam data and simulations.

	π^+		proton			
			Beam energy		Available energy	
	m	a	m	a	m	a
Data	0.94±0.04	0.19±0.03	0.78±0.04	0.45±0.06	0.88±0.04	0.28±0.05
FTFP_BERT	0.83±0.03	0.25±0.02	0.74±0.02	0.51±0.03	0.86±0.02	0.29±0.02
QGSP_BERT	0.65±0.03	0.45±0.04	0.67±0.02	0.60±0.04	0.78±0.02	0.34±0.03

The ratio of the calorimeter response of pions to protons of the same initial energy is called the p/π ratio ($\mathrm{p}/\pi = \frac{E_\mathrm{p}}{E_\pi}$). Figure 5 shows the p/π ratio extracted in this study for the Fe-AHCAL together with the results obtained for two other iron-scintillator calorimeters: the CDF End Plug hadron calorimeter [21] and the ATLAS Tile hadron calorimeter [6]. The energy dependence of the p/π ratio is mainly driven by the difference in measurable energy for mesons and baryons, which dominates below 20 GeV and gives way to other effects at higher energies. This behaviour is qualitatively supported by the comparison of Fig. 3(a) and 3(b) (4(a) and 4(b)) and is quantitatively estimated in [21]. The results obtained from the measurements performed with different calorimeters are in good agreement.

The additional difference between pion and proton response, which remains after the exclusion of the available energy effect, amounts to 2–4% (see Fig. 3 and 4) and is related to the lower probability of π^0 production in the interaction of a proton with a nucleus as suggested in [20]. This remaining difference decreases faster with increasing energy in data than is predicted by simulations. Both physics lists tend to underestimate the p/π ratio above 20 GeV. The predictions of QGSP_BERT are closer to the data because both pion and proton response is overestimated by this physics list above 20 GeV. At the same time, abnormal behaviour is visible around the model transition region in the QGSP_BERT physics list. The FTFP_BERT physics list underestimates the p/π ratio due to an overestimate of the pion response while the proton response is reproduced within uncertainties.

3.2 Energy resolution

Absolute and fractional energy resolutions for pions and protons are shown in Fig. 6 and 7. The study of the fractional energy resolution of the CALICE Fe-AHCAL was performed in [8] using test beam data of positive and negative pions in the energy range 10-80 GeV. The energy dependence of the fractional energy resolution was described in [8] by the three-component function

$$\frac{\sigma}{E} = \frac{a1}{\sqrt{E}} \oplus a2 \oplus \frac{a3}{E}, \tag{3.5}$$

where E is in GeV and the parameters $a1$, $a2$, and $a3$ correspond to the stochastic, constant, and noise contributions, respectively. The values of the parameters $a1 = 0.576 \pm 0.004\ \mathrm{GeV}^{\frac{1}{2}}$ and $a2 = 0.016 \pm 0.003$ were obtained in [8] from the fit to pion data, for which the value of parameter $a3 =$

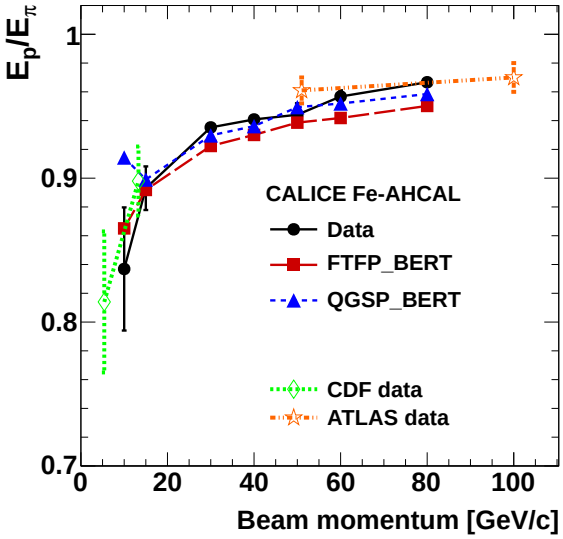

Figure 5. p/π ratio versus beam momentum for data and simulations of the CALICE Fe-AHCAL; error bars show the statistical uncertainties, the systematic uncertainties are assumed to cancel. Green and orange markers show the data obtained with the CDF [21] and ATLAS [6] hadron calorimeters, respectively.

0.18 GeV was fixed to the estimated noise of the combined calorimeter setup (Si-W ECAL+Fe-AHCAL+TCMT). This result from [8] is shown in Fig. 6(b) and 7(b) with the dashed black curve.

The energy dependence of σ_{reco} for pions is not reproduced by the simulations. The fluctuations of the energy deposition in the simulated hadronic showers grow steeper with increasing energy (see Fig. 6(a) and 7(a) where the smooth curves are shown to guide the eyes). The overestimate of σ_{reco} exceeds 15% at 80 GeV, while below 20 GeV σ_{reco} tends to be underestimated. As both response and absolute resolution for pions are overestimated by both physics lists, the behaviour of the fractional energy resolution is well reproduced by simulations.

The absolute energy resolution tends to be smaller for protons relative to pions. The simulations predict a larger difference in σ_{reco} between pions and protons than is observed in data. The fractional energy resolutions for protons and pions are within uncertainties.

3.3 Longitudinal shower depth

The longitudinal centre of gravity is an energy weighted sum of the longitudinal hit coordinates. Two different observables are constructed to characterise the longitudinal development. The first is the event longitudinal centre of gravity, Z, calculated with respect to the calorimeter front plane as follows:

$$Z = \frac{\sum_{i=1}^{N} e_i \cdot z_i}{\sum_{i=1}^{N} e_i}, \tag{3.6}$$

where N is the total number of hits in the Fe-AHCAL, e_i is the hit energy, and z_i is the distance from the hit layer to the calorimeter front plane. The event longitudinal centre of gravity, Z, is correlated with the shower start position and is very helpful for event selection (see Section 2.2).

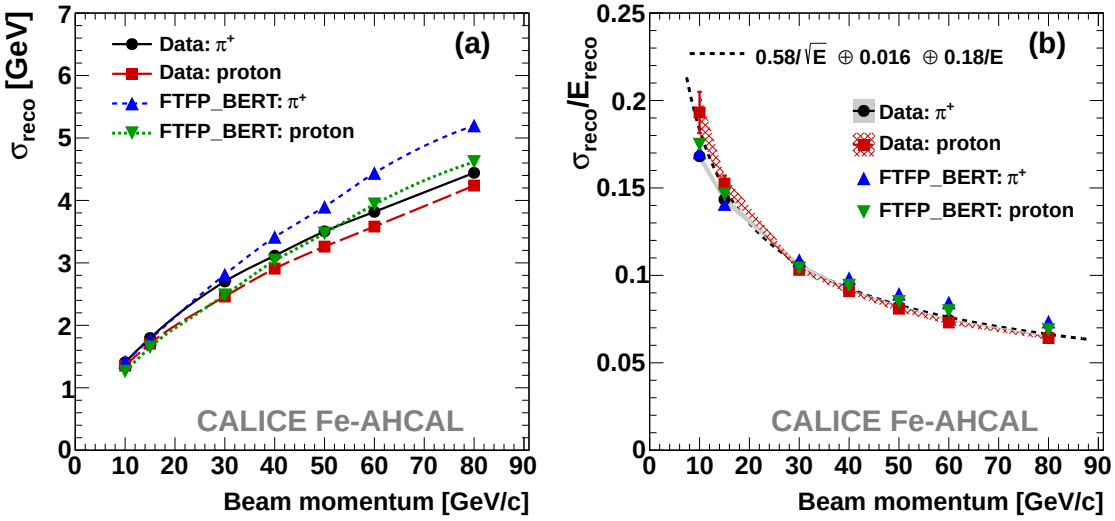

Figure 6. (a) Absolute and (b) fractional energy resolution for pions (black circles), protons (red squares), simulated pions (blue triangles), and simulated protons (green down triangles) using the FTFP_BERT physics list. Error bars show the statistical uncertainties. The grey and crosshatched red bands show the systematic uncertainties for pion and proton data, respectively (not estimated for the absolute resolution).

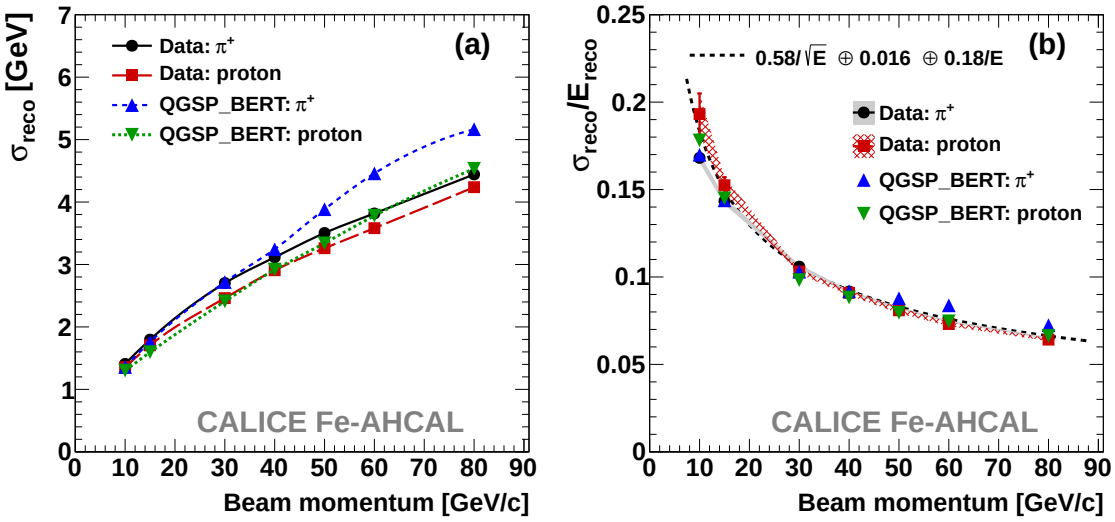

Figure 7. (a) Absolute and (b) fractional energy resolution for pions (black circles), protons (red squares), simulated pions (blue triangles), and simulated protons (green down triangles) using the QGSP_BERT physics list. Error bars show the statistical uncertainties. The grey and crosshatched red bands show the systematic uncertainties for pion and proton data, respectively (not estimated for the absolute resolution).

To study the shower development, we introduce the longitudinal shower depth, $Z0$, which represents a longitudinal centre of gravity calculated with respect to the shower start in each event as follows:

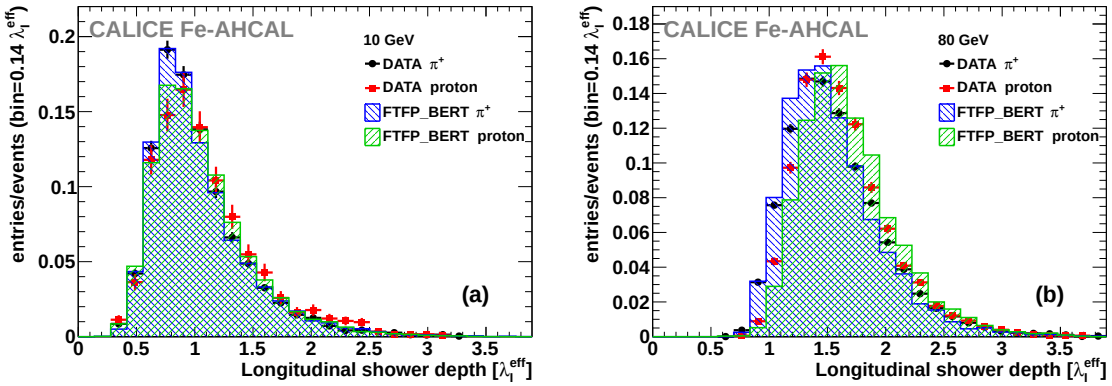

Figure 8. Distributions of the longitudinal shower depth $Z0$ of pion- and proton-induced showers at initial momentum (a) 10 GeV/c and (b) 80 GeV/c for data and simulation (FTFP_BERT). Error bars show the statistical uncertainties.

$$Z0 = \frac{\sum_{i=1}^{N_{\rm sh}} e_i \cdot (z_i - z_{\rm start})}{\sum_{i=1}^{N_{\rm sh}} e_i}, \tag{3.7}$$

where $N_{\rm sh}$ is the number of hits in the Fe-AHCAL from the shower start layer and beyond, e_i is the hit energy, z_i is the distance from hit layer to the calorimeter front, and $z_{\rm start}$ is the distance from the shower start layer to the calorimeter front face. In contrast to Z, the value $Z0$ is independent of the distribution of the shower start position and describes an intrinsic longitudinal shower development. The observable $Z0$ is more convenient for the comparison of showers induced by different types of hadrons with different nuclear interaction lengths. Typical distributions of $Z0$ are shown in Fig. 8 for pions and protons. The longitudinal shower depth of a pion shower tends to be closer to the shower start than that of a proton shower. The shape and width of the distribution is well reproduced by the FTFP_BERT physics list for both pions and protons in the energy range studied.

The mean longitudinal shower depth, $\langle Z0 \rangle$, is extracted from the distributions shown in Fig. 8. The energy dependence of $\langle Z0 \rangle$ in Fig. 9 increases logarithmically with energy from $\sim 1\lambda_{\rm I}^{\rm eff}$ at 10 GeV to $\sim 1.5\lambda_{\rm I}^{\rm eff}$ at 80 GeV. Figure 10 shows the ratios of simulations to data. The QGSP_BERT physics list underestimates $\langle Z0 \rangle$ by ~5-7% for both pions and protons above 20 GeV. The FTFP_BERT physics list gives a very good prediction of $\langle Z0 \rangle$ for pions and slightly overestimates the rate of growth for protons.

The longitudinal dispersion, σ_{Z0}, characterises the scattering of shower hits around the longitudinal centre of gravity and is calculated for each event using the following formula:

$$\sigma_{Z0} = \sqrt{\frac{\sum_{i=1}^{N_{\rm sh}} e_i \cdot (z_i - z_{\rm start})^2}{\sum_{i=1}^{N_{\rm sh}} e_i} - Z0^2}, \tag{3.8}$$

where $Z0$ is from Eq. 3.7. The mean longitudinal dispersion, $\langle \sigma_{Z0} \rangle$, is of the same order of magnitude as the mean $\langle Z0 \rangle$ (Fig. 11) and also increases logarithmically with energy. The values of $\langle \sigma_{Z0} \rangle$ predicted by the FTFP_BERT physics list are in agreement with data for both types of hadrons. The QGSP_BERT physics list underestimates the mean longitudinal dispersion above 20 GeV by ~5% (Fig. 12).

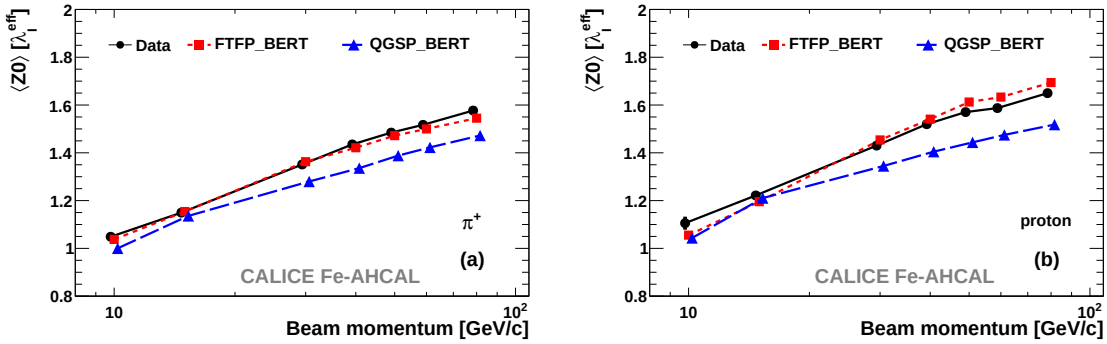

Figure 9. Mean longitudinal centre of gravity of (a) pion and (b) proton-induced showers in units of $\lambda_I^{\text{eff}} =$ 231 mm for data (black circles) and simulations with the FTFP_BERT (red squares) and QGSP_BERT (blue triangles) physics lists. Data points are corrected for bias as described in Section 2.4.

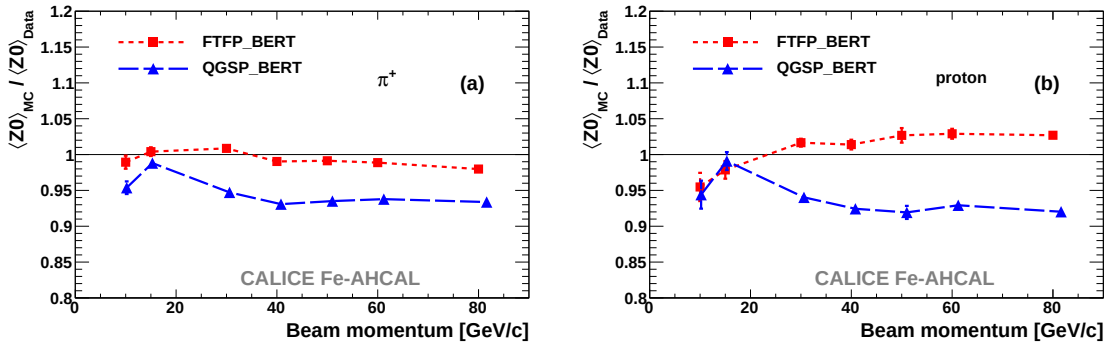

Figure 10. Ratio of the mean longitudinal centre of gravity for simulations using the FTFP_BERT (red squares) and QGSP_BERT (blue triangles) physics list to data for (a) pion and (b) proton-induced showers. Data are corrected for bias as described in Section 2.4.

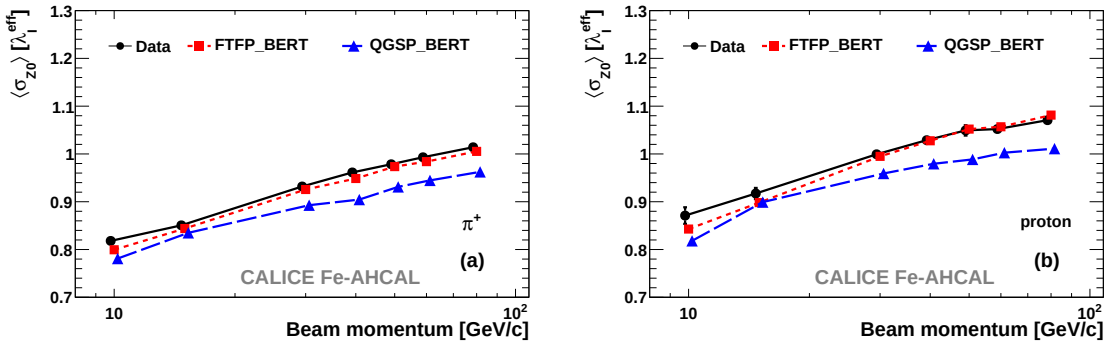

Figure 11. Mean longitudinal dispersion of (a) pion and (b) proton-induced showers in units of $\lambda_I^{\text{eff}} =$ 231 mm for data (black circles) and simulations with the FTFP_BERT (red squares) and QGSP_BERT (blue triangles) physics lists. Data points are corrected for bias as described in Section 2.4.

3.4 Mean shower radius

The shower radius, R, is an energy weighted sum of hit radial distances to the shower axis (in the plane perpendicular to the beam direction) and is calculated as follows:

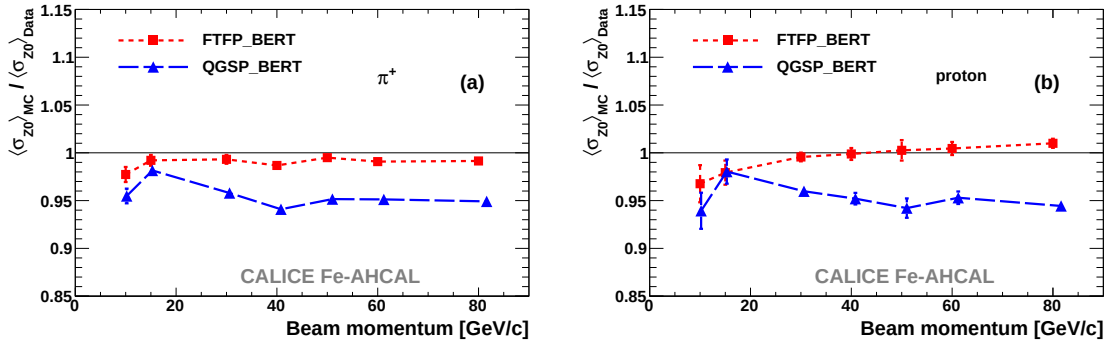

Figure 12. Ratio of the mean longitudinal dispersion for simulations using the FTFP_BERT (red squares) and QGSP_BERT (blue triangles) physics lists to data for (a) pion and (b) proton-induced showers. Data are corrected for bias as described in Section 2.4.

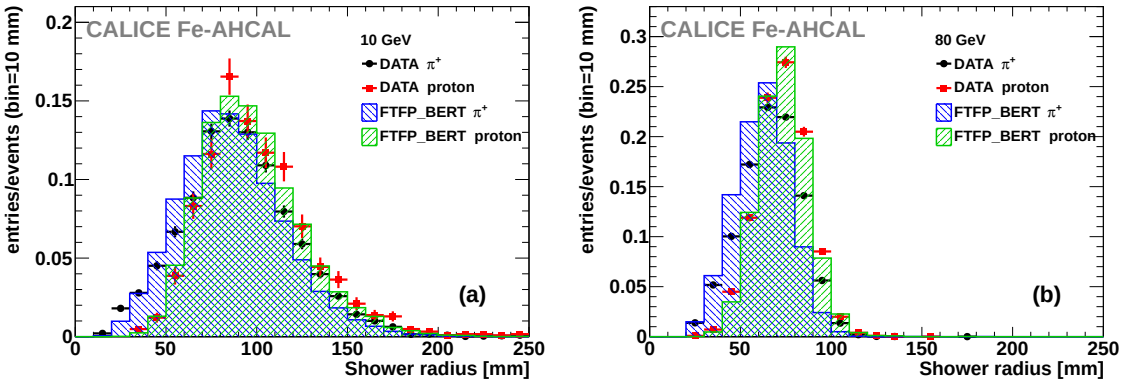

Figure 13. Distributions of the shower radius for pion- and proton-induced showers at initial momentum (a) 10 GeV/c and (b) 80 GeV/c for data and simulation (FTFP_BERT). Error bars show the statistical uncertainties.

$$R = \frac{\sum_{i=1}^{N_{\rm sh}} e_i \cdot r_i}{\sum_{i=1}^{N_{\rm sh}} e_i}, \tag{3.9}$$

where $N_{\rm sh}$ is the number of hits in the Fe-AHCAL from the shower start layer and beyond, e_i is the hit energy, $r_i = \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2}$ is the distance from hit with coordinates (x_i,y_i) to the shower axis with coordinates (x_0,y_0). The shower axis is defined using the primary track coordinates in the Si-W ECAL or the event centre of gravity for the data taken without the electromagnetic calorimeter. Typical distributions of the shower radius are shown in Fig. 13, from which the mean shower radius, $\langle R \rangle$, is extracted for pion- and proton-induced showers. Proton-induced showers tend to be wider than pion showers. For data, the fluctuations of the radius of pion-induced showers are larger than those of proton-induced showers. The shape and width of the distributions are reasonably well reproduced by the FTFP_BERT physics list for both pions and protons in the energy range studied.

The energy dependencies of the mean shower radius are shown in Fig. 14. The values of $\langle R \rangle$ decrease logarithmically with increasing energy and this general behaviour is well reproduced

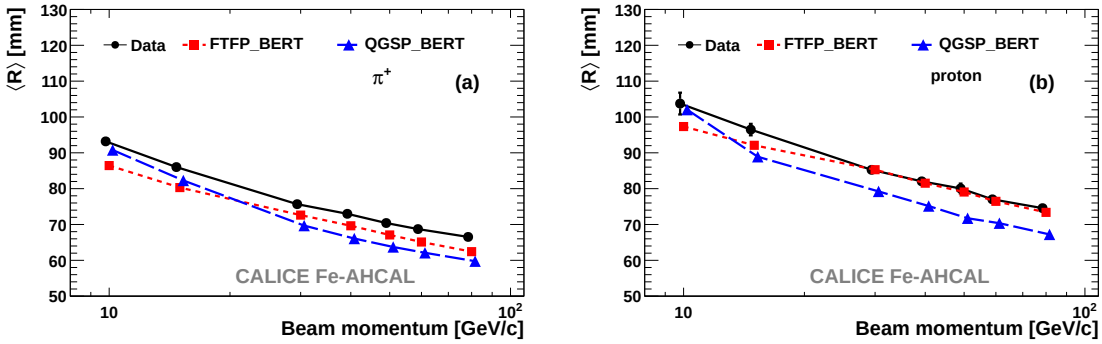

Figure 14. Mean shower radius of (a) pion and (b) proton-induced showers for data (black circles) and simulations with the FTFP_BERT (red squares) and QGSP_BERT (blue triangles) physics lists. Data points are corrected for bias as described in Section 2.4.

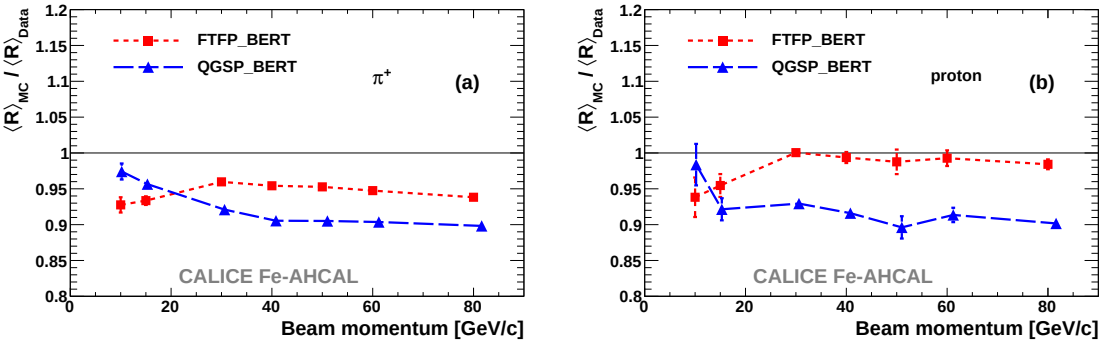

Figure 15. Ratio of the mean shower radius form simulations using the FTFP_BERT (red squares) and QGSP_BERT (blue triangles) physics lists to data for (a) pion and (b) proton-induced showers. Data are corrected for bias as described in Section 2.4.

by all studied physics lists. The pion (proton) showers are observed to be narrower by ∼25% (∼30%) at 80 GeV than at 10 GeV. This is explained by the increase of the electromagnetic fraction in hadronic showers, since electromagnetic sub-showers tend to be more compact. The ratio of simulations to data is shown in Fig. 15. The FTFP_BERT physics list predicts the width of the proton showers within uncertainties and underestimates the width of pion showers by ∼5-7%. The QGSP_BERT physics list demonstrates better agreement with data at 10 GeV but underestimates the shower width at higher energies by ∼10% for both pions and protons.

The radial dispersion, σ_R, characterises the scattering of shower hits around the shower radius and is calculated for each event as follows:

$$\sigma_R = \sqrt{\frac{\sum_{i=1}^{N_{\rm sh}} e_i \cdot (r_i)^2}{\sum_{i=1}^{N_{\rm sh}} e_i} - R^2}, \tag{3.10}$$

where R is from Eq. 3.9.

The mean radial dispersion, $\langle\sigma_R\rangle$, shown in Fig. 16 is of the same order of magnitude as the mean value $\langle R\rangle$ but decreases more slowly with increasing energy. The discrepancy between data and simulation increases with energy but is smaller than for the mean shower radius. Again,

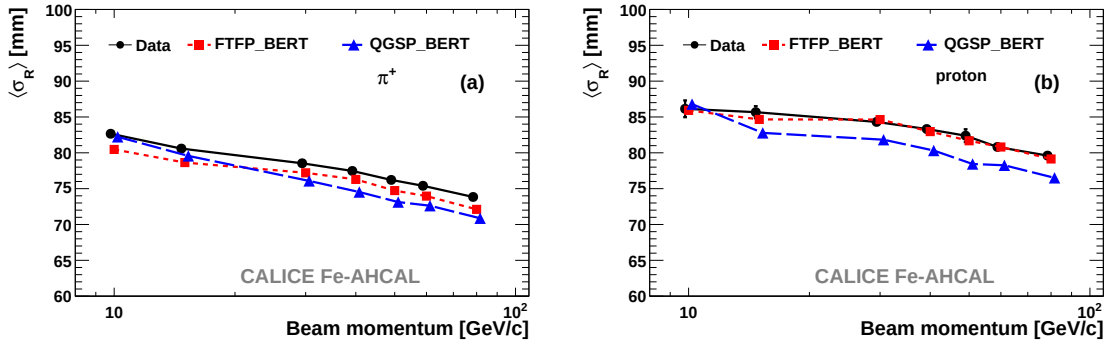

Figure 16. Mean radial dispersion of (a) pion and (b) proton-induced showers for data (black circles) and simulations with the FTFP_BERT (red squares) and QGSP_BERT (blue triangles) physics lists. Data points are corrected for bias as described in Section 2.4.

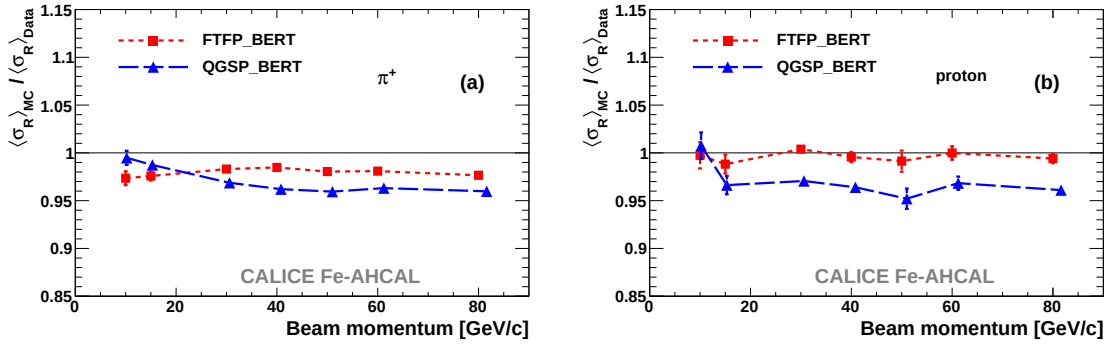

Figure 17. Ratio of the mean radial dispersion from simulations using the FTFP_BERT (red squares) and QGSP_BERT (blue triangles) physics lists to data for (a) pion and (b) proton-induced showers. Data are corrected for bias as described in Section 2.4.

the FTFP_BERT physics list describes the data better than QGSP_BERT, especially for protons (Fig. 17).

4. Conclusion

Global parameters of showers induced by positive hadrons with initial momenta from 10 to 80 GeV/c in the CALICE analogue scintillator-steel hadronic calorimeter have been analysed and compared with simulations using the QGSP_BERT and FTFP_BERT physics lists from GEANT4 version 9.6 patch 01. In general, the detector response to hadrons tends to increase more rapidly with energy in simulations than in data. Of the two physics lists studied, FTFP_BERT gives a better prediction of the response for both pions and protons. The deficiency of the calorimeter response for protons with respect to pions, which cannot be explained by the difference in available energy, is observed to be ∼2–4%. The ratio of the proton to pion response tends to be underestimated by the simulations.

Proton-induced showers tend to be ∼5% longer and ∼10% wider than pion showers at the same energy in terms of the longitudinal centre of gravity and shower radius, respectively. The

spatial shower development for both types of hadrons is much better predicted by the FTFP_BERT physics list, especially above 20 GeV. The event-by-event fluctuations of the spatial characteristics are also quite well reproduced. The simulated showers are still narrower than those observed in data but the FTFP_BERT physics list predicts the mean shower radius with an accuracy of $\sim$5-7%.

The most significant discrepancy between test beam data and simulations is seen in the absolute energy resolution for pions. The simulated width of the energy distribution of pion showers increases faster with pion initial energy than observed in data. The fractional energy resolutions for pions and protons are in good agreement and are well reproduced by simulations with both physics lists studied.

Acknowledgments

We would like to thank the technicians and the engineers who contributed to the design and construction of the prototypes. We also gratefully acknowledge the DESY and CERN managements for their support and hospitality, and their accelerator staff for the reliable and efficient beam operation. 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 (including Research Award IV, RUS1066839 GSA); by the Russian Ministry of Education and Science contracts 4465.2014.2 and 14.A12.31.0006 and the Russian Foundation for Basic Research grant 14-02-00873A; 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; and by the Science and Technology Facilities Council, UK.

References

[1] J. C. Brient and H. Videau, *The Calorimetry at the future e^+ e^- linear collider*, *eConf* **C010630** (2001) E3047;

V. L. Morgunov *Calorimetry design with energy-flow concept (imaging detector for high-energy physics)*, *Proc. of CALOR 2002*, Pasadena, California;

M. Thomson *Particle Flow Calorimetry and the PandoraPFA Algorithm*, *Nucl. Instrum. Meth.* A **611** (2009) 25 [arXiv:0907.3577].

[2] F. Beaudette, *Performance of the particle flow algorithm in CMS*, *PoS*, ICHEP 2010, 002;

The CMS collaboration, *Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET*, CMS-PAS-PFT-09-001.

[3] R. Wigmans, *Calorimetry: Energy Measurement in Particle Physics*, International Series of Monographs on Physics, vol. 107, Oxford University Press, Oxford, 2000.

[4] T. A. Gabriel, D. E. Groom, P. K. Job, N.V. Mokhov and G. R. Stevenson, *Energy dependence of hadronic activity*, *Nucl. Instrum. Meth.* **A338** (1994) 336.

[5] N. Akchurin *et al.*, *On the differences between high-energy proton and pion showers and their signals in a non-compensating calorimeter*, *Nucl. Instrum. Meth.* **A408** (1998) 380.

[6] P. Adragna *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.* **A615** (2010) 158.

[7] The CALICE collaboration, C. Adloff *et al.*, *Construction and commissioning of the CALICE analog hadron calorimeter prototype*, 2010 *JINST* **5** P05004 [arXiv:1003.2662].

[8] The CALICE collaboration, C. Adloff *et al.*, *Hadronic energy resolution of a highly granular scintillator-steel calorimeter using software compensation techniques*, 2012 *JINST* **7** P09017 [arXiv:1207.4210].

[9] The CALICE collaboration, C. Adloff *et al.*, *Validation of GEANT4 Monte Carlo Models with a Highly Granular Scintillator-Steel Hadron Calorimeter*, 2013 *JINST* **8** P07005 [arXiv:1306.3037].

[10] The GEANT4 Collaboration, S. Agostinelli *et al.*, *Geant4–a simulation toolkit*, *Nucl. Instrum. Meth.* **A506** (2003) 250.

[11] 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).

[12] The CALICE collaboration, J. Repond *et al.*, *Design and electronics commissioning of the physics prototype of a Si-W electromagnetic calorimeter for the International Linear Collider*, 2008 *JINST* **3** P08001 [arXiv:0805.4833].

[13] The CALICE collaboration, C. Adloff *et al.*, *Construction and performance of a silicon photomultiplier/extruded scintillator tail-catcher and muon-tracker*, 2012 *JINST* **7** P04015 [arXiv:1201.1653].

[14] The CALICE Collaboration, C. Adloff *et al.*, *Tests of a particle flow algorithm with CALICE test beam data*, 2011 *JINST* **6** P07005 [arXiv:1105.3417].

[15] J. Beringer *et al.* (Particle Data Group) *The Review of Particle Physics*, *Phys. Rev.* **D86** (2012) 010001.

[16] A. Ribon *et al.*, *Status of Geant4 hadronic physics for the simulation of LHC experiments at the start of LHC physics program*, CERN-LCGAPP-2010-02.

[17] A. Dotti, *Simulation of Showers with Geant4*, *Proc. of CHEF 2013*, Paris, France.

[18] J. Apostolakis *et al. GEANT4 physics lists for HEP*, 2008 IEEE Nucl. Sci. Symp. Conf. Rec. (2008) 833.

[19] The CALICE collaboration C. Adloff *et al.*, *Electromagnetic response of a highly granular hadronic calorimeter*, 2011 *JINST* **6** P04003 [arXiv:1012.4343].

[20] D. E. Groom, *Energy flow in a hadronic cascade: Application to hadron calorimetry*, *Nucl. Instrum. Meth.* **A572** (2007) 633.

[21] Jinbo Liu (for the CDF Plug Upgrade Group), *Testbeam Results for the CDF End Plug Hadron Calorimeter*, *Proc. of CALOR 1997*, Tucson, Arizona, USA [FERMILAB-Conf-97/414-E].

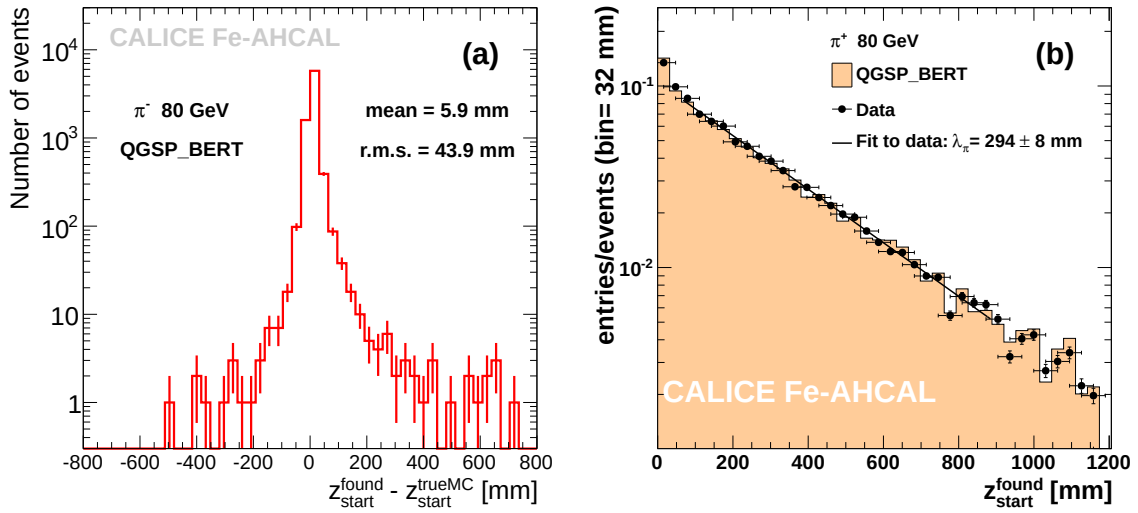

Figure 18. (a) Difference between the found shower start position and the true shower start position for simulated samples of 80 GeV pions. (b) Distributions of the found shower start position for data (black circles) and simulations (filled histogram) for 80 GeV pions. The estimated nuclear interaction length obtained from the fit to data is shown in the legend. See text about fit details.

A. Estimate of the nuclear interaction length

The uncertainty of the shower start is approximately ± 1 AHCAL layer ($\approx$32 mm) and was estimated using the simulated pion samples. The difference between the identified shower start position, $z_{\text{start}}^{\text{found}}$, and the real shower start position, $z_{\text{start}}^{\text{trueMC}}$, is shown in Fig. 18(a) for 80 GeV pions. The values of $z_{\text{start}}^{\text{trueMC}}$ were obtained from the GEANT4 information about the position of the primary inelastic interaction of the incoming particle. Typical distributions of the found shower start position $z_{\text{start}}^{\text{found}}$ with respect to the calorimeter front face are shown in Fig. 18(b) for data and FTFP_BERT physics list. The distributions of the shower start layer for pions $f(z_{\text{start}})$ can be fit with the following exponential function:

$$f(z_{\text{start}}) = A \cdot \exp\left(-\frac{z_{\text{start}}}{\lambda_\pi}\right), \tag{A.1}$$

where A is a normalisation factor and λ_π is the estimated nuclear interaction length. The two first layers with large uncertainty for the shower start finding algorithm as well as several last layers were excluded from the fit, and the fit interval is from 65 mm to 900 mm for all samples.

The proton data samples contain pion contamination that varies from 5% to 35%. The distribution of the shower start position for a mixed sample of hadrons with different inelastic cross sections can be considered as a sum of two independent contributions:

$$f_{\text{mix}}(z_{\text{start}}) = A \cdot \left(\eta \cdot \exp\left(-\frac{z_{\text{start}}}{\lambda_{\text{I}}}\right) + (1-\eta) \cdot \exp\left(-\frac{z_{\text{start}}}{\lambda_\pi}\right) \cdot \frac{\lambda_{\text{I}}}{\lambda_\pi}\right), \tag{A.2}$$

where the normalisation factor, A, and the nuclear interaction length for protons, λ_{I}, are estimated variables, while the purity of proton sample, η, and the nuclear interaction length for pions, λ_π,

are taken as previously determined and known parameters. The purity η is estimated as described in Appendix B. The values of λ_π are extracted from the fit to the distributions obtained from the corresponding pion samples at the same energy.

The following procedure was used to estimate the systematic uncertainty due to the uncertainties of λ_π and purity. The set of parameter values was generated using Gaussian distributions with the variance corresponding to the uncertainty of a given parameter and the generated values were used in the fit to proton data. The r.m.s. of the obtained distribution of $\lambda_{\rm I}$ is taken as a systematic uncertainty. The contributions from both parameters are summed up in quadrature.

B. Estimate of the sample purity

The purity of the sample, η, is estimated using the efficiency, ε, to select contaminating events. The efficiency ε in turn is determined from an independent procedure, which does not involve information from the Čerenkov counter.

In the current study, the pion-proton separation is based on the information from the Čerenkov detector. The independent calorimeter-based muon identification procedure, as introduced in Section 2.2, can be used to measure the efficiency of the Čerenkov counter. Assuming that the efficiency for pions $\varepsilon_{\rm pion}$ is approximately the same as for muons $\varepsilon_{\rm muon}$ the value of $\varepsilon_{\rm pion}$ can be calculated as follows:

$$\varepsilon_{\rm pion} \approx \varepsilon_{\rm muon} = \frac{N^{\rm cher}_{\rm muon}}{N^{\rm total}_{\rm muon}}, \tag{B.1}$$

where $N^{\rm cher}_{\rm muon}$ is the number of identified muons that gave a signal in the Čerenkov detector, $N^{\rm total}_{\rm muon}$ is the total number of muons identified using the calorimeter-based procedure.

The purity of the proton sample, $\eta_{\rm p}$, i.e. the ratio of the true number of protons to the number of identified protons, can be calculated as follows:

$$\eta_{\rm p} = 1 - \frac{N_\pi}{N_{\rm p}} \cdot \left(\frac{1-\varepsilon_{\rm pion}}{\varepsilon_{\rm pion}} \right), \tag{B.2}$$

where N_π ($N_{\rm p}$) is the number of pions (protons) in the selected samples identified using the Čerenkov counter. The uncertainties on $\varepsilon_{\rm muon}$ are estimated from the available statistics of the muon event sample and are propagated to the uncertainty of $\eta_{\rm p}$. As the pressure in the gaseous Čerenkov detector used was set well below the proton threshold, the misidentification of protons is negligible.

The efficiency of positron identification, $\varepsilon_{\rm pos}$, is estimated from the dedicated positron runs taken without an electromagnetic calorimeter as follows:

$$\varepsilon_{\rm pos} = \frac{N_{\rm selected}}{N_{\rm total}}, \tag{B.3}$$

where $N_{\rm selected}$ is the number of selected positrons from the pure positron sample with $N_{\rm total}$ events. Then the purity of the selected pion samples can be calculated as follows:

$$\eta_\pi = 1 - \frac{N_{\rm pos}}{N_\pi} \cdot \left(\frac{1-\varepsilon_{\rm pos}}{\varepsilon_{\rm pos}} \right), \tag{B.4}$$

where N_{pos} (N_π) is the number of positrons (pions) in the selected samples identified using the the positron selection procedure (see Section 2.2) based on shower radius and longitudinal centre of gravity.