Interactions of Pions in the CALICE Silicon-Tungsten Calorimeter Prototype

The CALICE Collaboration *

Abstract

This paper presents a detailed study of hadronic interactions using data recorded with the highly granular CALICE silicon-tungsten electromagnetic calorimeter. Appoximately 750,000 selected π^- at energies between 2 and 10 GeV have been studied. Using algorithms developed to identify the position of the first hadronic interaction within the calorimeter, the predictions of several physics models available within the GEANT4 simulation toolkit are compared to data. Although a good overall description of the data is observed, there are significant quantitative discrepancies in the longitudinal and transverse distributions of deposited energy.

*Corresponding author: Naomi van der Kolk; kolk@lal.in2p3.fr

Contents

1 Introduction

The primary physics goals at a future high energy lepton collider require the precise measurement of the energy of hadron jets. Particle flow algorithms (*PFA*) foreseen to be employed at future linear electron-positron colliders [1, 2], are promising for achieving a jet energy resolution of 3–4% for jets with an energy range from 50 GeV to more than 500 GeV [3].

The PFA approach aims to reconstruct all particles in the final state of the e^+e^- collision. This requires highly segmented calorimeters to disentangle thr contributions from showers created by different types of particle within a jet, i.e. from hadrons and photons. The CALICE collaboration [4] designs, constructs and operates prototypes of calorimeters dedicated to the application of particle flow algorithms.

To optimise particle flow algorithms, the interactions of hadrons must be modelled reliably in Monte Carlo simulations and the detector response to hadrons must be well-understood. In view of this, highly granular calorimeter prototypes provide a unique means to test and to further develop models of hadronic cascades.

In this paper a highly granular silicon-tungsten electromagnetic calorimeter protopype (Si-W ECAL) [5] is used to test hadronic shower models at low energies. The depth of the Si-W ECAL correspods to one interaction length (λ_{I}), which means that, although the complete shower is not recorded, the first hadronic interaction can be studied in great detail because of the fine longitudinal and transversal sampling. The Si-W ECAL was operated in a test beam at FNAL in 2008 with negatively charged pions in the energy range 2–10 GeV. The majority of individual hadrons within high energy jets have energies in this range and therefore it is of considerable interest to test the performance of Monte Carlo simulations with respect to the recorded data. The high granularity of the Si-W ECAL permits a detailed measurement of hadronic interactions in terms of global observables describing both the longitudinal and transverse shower development and these observables should be reproduced correctly by simulations.

This paper is organised as follows: the Si-W ECAL prototype is described in the following section, the data and Monte Carlo simulations, as well as the event selection criteria employed, are presented in Sect. 3. The algorithm used to identify interactions is described in Sect. 4. Results obtained with the prototype during the test beams with negatively charged pions (π^-) and comparisons with Monte Carlo are discussed in Sect. 5. A summary, conclusions and prospects for future studies are given in the last paragraph.

2 The Si-W ECAL prototype

The Si-W ECAL prototype consists of a sandwich structure of 30 layers of silicon as active material, alternating with tungsten as the absorber material. The active layers are made of silicon wafers segmented in 1×1 cm^2 pixels (or pads). As shown in Fig. 1, each wafer consists of a square of 6×6 pixels and each layer is a matrix of 3×3 of these wafers, resulting in an active zone of 18×18 cm^2.

The Si-W ECAL is divided in three modules of 10 layers. The tungsten thickness per layer is different in each module increasing from 1.4 mm in the first module (layers 1–10), to 2.8 mm in the second (layers 11–20) and 4.2 mm in the third (layers 21–30). This corresponds to 24 radiation lengths (X_0) and approximately 1 interaction length, which ensures that more than half of the hadrons traversing the Si-W ECAL prototype will have a primary interaction within its volume.

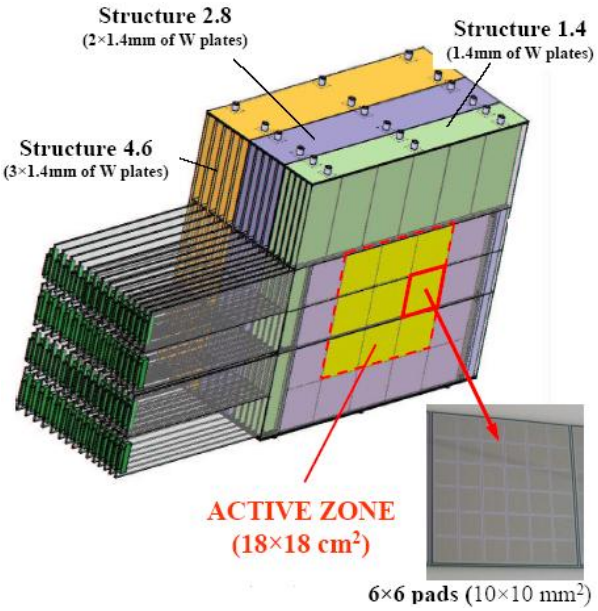


Figure 1: *Schematic view of the Si-W ECAL prototype*

3 Data samples

Test beams were conducted in May and July of 2008 at the Fermilab Test Beam Facility [6] at FNAL. The analysis presented in this paper uses data from runs with π^--mesons at energies of 2, 4, 6, 8 and 10 GeV. This energy range covers typical hadron energies within a jet. The Si-W ECAL was placed in front of two other CALICE prototypes: an analogue hadronic calorimeter (HCAL) [7] and a Tail-Catcher and Muon Tracker (TCMT) [8]. Upstream the beam line was instrumented with two scintillator counters, each with an area of $10x10cm^2$, for triggering on incoming particles and two Cherenkov detectors for particle identification. The choosen coordinate system is right-handed with the z-axis pointing into the beam direction and the x-axis vertical.

Monte Carlo simulations corresponding to the recorded test beam data have been produced using the simulation toolkit GEANT4 [9]. Version 9.6 patch 1 of GEANT4 has been the default for this paper. The full geometry of the CALICE test beam setup is taken into account in the simulation via the MOKKA framework [10] which provides the geometry interface to GEANT4. A detailed description of the detector simulation can be found in [11].

3.1 Simulation with various GEANT4 physics lists

Due to the complicated nature of hadronic interactions it is difficult to achieve an accurate description of hadronic showers in simulations. In GEANT4 several phenomenological hadronic interaction models are available as described in [12].

At low energies, where nucleons can be considered pointlike in nature, two cascade models are considered. The Bertini cascade model treats the nucleus as a sphere with uniform nucleon density. Incident hadrons collide with the nucleons producing secondary particles which also collide with the nucleons in a so-called intra-nuclear cascade. In the binary cascade model the hadron interacts with the nucleus as a set of discrete nucleons, each having a defined momentum and position, and propagates through the nucleus in a cascading series of two-particle collisions.

For medium to high energy hadronic interactions the string parton models are imple-

mented. There are two approaches, the Fritiof and the Quark Gluon String model. In the Fritiof string model diffractive scattering of the primary hadron with the nucleons is via momentum transfer alone whereas in the Quark Gluon String model pomerons are exchanged. An interaction results in several excited strings (and an excited nucleus) that are fragmented to produce secondaries, which interact via a shower model or precompound model. Additionally there are the Low Energy Parameterized and High Energy Parameterized models, which are based on fits to experimental data.

These models are combined into *physics lists* where they are applied in a specified energy range. Where two models are combined in a physics list, a smooth transition is achieved by randomly choosing the model on an event-by-event basis, with a probability that varies linearly with the energy in the interval. The physics list QGSP_BERT, for example, combines the Bertini model at low energies with the Low Energy Parameterized model at intermediate energies and the Quark-Gluon-String Precompound model at high energies.

Four physics lists have been studied to be sensitive to differences in and the effect of the transitions between them. The hadronic models employed in these physics list for the studied beam energy range are illustrated in Fig. 2. Electromagnetic processes for these physics lists all have the same underlying physics model.

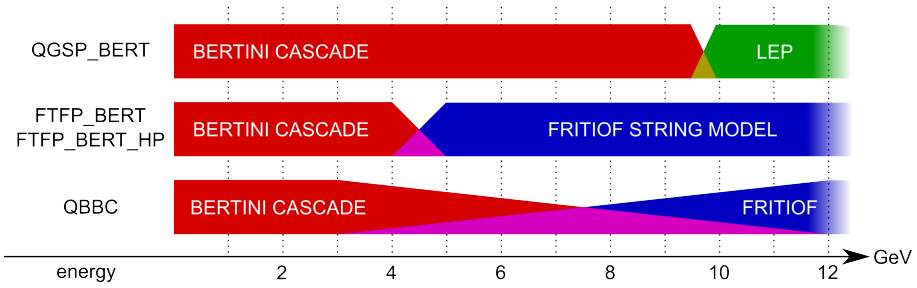


Figure 2: *Model used for hadronic interactions of π^- depending on the physics list and the energy of the interacting particle.*

The physics lists QGSP_BERT and FTFP_BERT allow the effect of the transition from the cascade to the string model to be studied, while QBBC offers an alternative having a larger transition region between the two and by combining the bertini and binary cascode models for neutrons and protons below 1.5 GeV. FTFP_BERT_HP employs a high precision treatment of neutrons with kinetic energies below 20 MeV. FTFP_BERT is currently the recommended physics list for LHC calorimeter results [13] and therefore it is used as the reference in this paper.

3.2 Event selection

Events are triggered using the coincidence of the two scintillator counters and pions are identified with the help of two threshold Cherenkov counters. The gas pressure in these counters is set such that for 2, 4 and 6 GeV neither Cherenkov counters should fire, while for 8 and 10 GeV only the first should fire. The FNAL π^- test beam is contaminated with muons and electrons, in particular at the lowest energies where the beam is dominated by electrons. The Cherenkov trigger reduces the contamination from electrons and muons, with further refinements to this preselection made as summarised below.

Data and simulation are subject to the same selection chain except where otherwise stated. The FTFP_BERT physics list is used as the default for background optimisation studies.

1. An energy threshold of 0.6 MIPs on the deposited energy in an individual pixel (a hit) is chosen to remove hits caused by detector noise.

2. Isolated hits are discarded in the analysis. A hit is called isolated if none of the 26 nearest-neighbour pixels contains a hit. This requirement removes 7 – 10% of the hits on average.

3. The total number of hits in the Si-W ECAL is required to be at least 25 to remove events in which the pion enters the Si-W ECAL at a large angle, towards the acceptance limits. In simulated events between 3% at 2 GeV and 9% of events at 10 GeV are removed by this cut. In the data about 5% of events is removed at all energies.

4. The barycentres (energy weighted mean positions) $\bar{x}$ and $\bar{y}$ of all hits in the Si-W ECAL volume are required to lie in the central part of the detector: -50 mm $< \bar{x} <$ 50 mm and -50 mm $< \bar{y} <$ 50 mm, to avoid selecting events where there may be significant lateral shower leakage.

5. Events in the data in which there are instrumental noise (0.3%) or spurious activity due to known effects are excluded.

6. The contamination from muons in the data is reduced with a selection based on the number of hits in the Si-W ECAL (N_{ECAL}), HCAL (N_{HCAL}) and TCMT (N_{TCMT}). The selection is also applied to simulated events. Based on the distribution of hits in a sample of simulated muon events, the following criteria are used to identify muons:

$$N_{\mathrm{ECAL}} < 50 \ , \ 30 < N_{\mathrm{HCAL}} < 70 \ , \ 10 < N_{\mathrm{TCMT}} < 35 \ . \tag{1}$$

   At 2 GeV, where the energy loss of muons is about 1.4 GeV in the HCAL, a looser requirement on the activity in the TCMT is used, due to the reduced number of counts in the TCMT ($N_{\mathrm{TCMT}} > 5$). If an event passes these requirements, it is counted as a muon.

   The efficiencies to reject muons and to select pions using these criteria are estimated using samples of 500 000 simulated events of each particle type. The results are given in Table 1.

E (GeV)	2	4	6	8	10
Muon rejection efficiency	0.998	0.999	0.997	0.996	0.994
Pion selection efficiency	0.876	0.931	0.924	0.949	0.960

Table 1: *Efficiencies to reject muons and to select pions using criteria 1–6 as described in the main text. These estimates are based on simulated samples generated using the* FTFP_BERT *physics list.*

   Based on the measured rejected fraction of muons in data, the FNAL pion beam is estimated to be contaminated with 10% muons at 2 GeV while at 10 GeV the contamination is less than 1%. The residual muon contamination in the data after the cuts are applied is negligible.

7. The pion beam is also contaminated with events where two primary particles hit the Si-W ECAL simultaneously. The contamination with events where a pion and muon are present is estimated to vary between 10% at 2 GeV and 5% at 10 GeV,

and the contamination with events containing two pions is estimated to be about 20% independent of energy. Such multi-particle events removed from the analysis by rejecting events with more than 11 hits in the TCMT. The distributions of N_{TCMT} for such events is sufficiently distinct from single pion events to allow a very efficient rejection of pion plus muon events and a sufficient rejection of double pion events. Table 2 shows the predicted efficiency to reject events (rejection efficiency) using Monte Carlo samples in which pions were overlayed with either muons or with other pions. The efficiency to select single pion events is also shown.

E (GeV)	2	4	6	8	10
Pion + muon event rejection efficiency	1.00	0.98	1.00	1.00	1.00
Pion + pion event rejection efficiency	0.81	0.775	0.834	0.835	0.848
Single pion event selection efficiency	0.988	0.991	0.970	0.957	0.935

Table 2: *Efficiencies to reject multi-particle events (pion + muon and pion + pion) and to select single pion events estimated using Monte Carlo samples (*FTFP_BERT*) in which pions were overlayed with either muons or with other pions.*

The contamination due to these backgrounds in the data has been estimated by a fit to the TCMT hit distribution. The residual contamination with pion plus muon events is estimated to be less than 0.5% at all energies and the double pion contamination 5% for all energies.

The number of events after the selection criteria are applied to the test beam data is given in Table 3. The number of events in the simulated event samples are of similar size.

E (GeV)	2	4	6	8	10
Events	15594	102660	59679	183346	375143

Table 3: *Number of events remaining after the selection criteria are applied to the FNAL* π^- *data.*

The contamination of the selected data with electrons is high at low energies; as high as 42% at 2 GeV. It decreases with increasing beam energy and at 10 GeV it is negligible. This contamination is reduced by rejecting events in which the incoming particle interacts in the beginning of the Si-W ECAL. Details about this additional event selection are given in the next section.

4 Identifying interacting events

A primary particle traversing the Si-W ECAL can either pass the detector material as an ionising particle or undergo interactions which lead to the creation of secondary particles. In the latter case the ionising track in the first layers is followed by several secondary tracks after the interaction. Figure 3 shows a recorded event where this can be clearly seen.

The bottom right histogram clearly illustrates that the deposited energy in consecutive layers increases significantly at the interaction point (here at layer 11). This change in deposited energy can be used to identify the layer in which the interaction takes place. Two criteria are applied; one based on the absolute energy increase and one based on the relative energy increase.

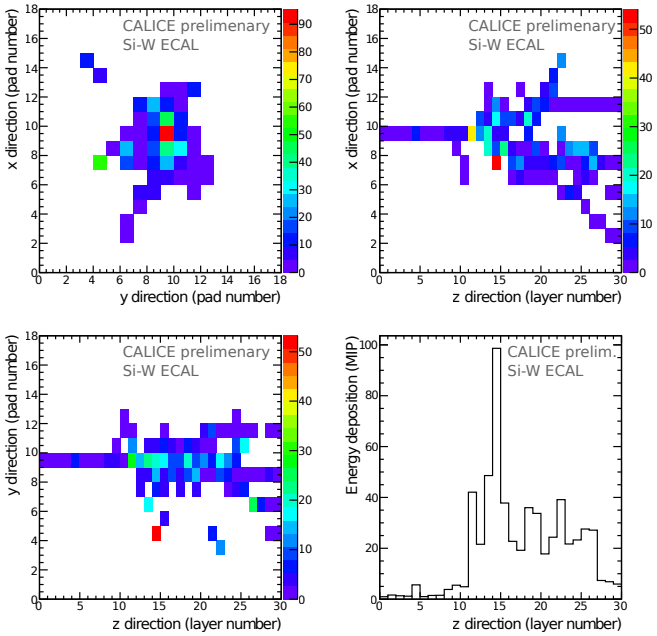


Figure 3: *A hadronic interaction of a pion with an incident kinetic energy of 10 GeV in the Si-W ECAL. Top left: projection in the x-y plane of the deposited energy. Top right: projection on the x-z plane of the deposited energy. Bottom left: projection on the y-z plane of the deposited energy. Bottom right: the energy deposition in each layer of the Si-W ECAL. The energy unit is in MIPs.*

First a cut, Ecut, is applied on the deposited energy in each layer, E_i. If three consecutive layers have an energy higher than Ecut, the interaction layer is assigned as the first of these (layer i). For interactions occuring in the last two layers of the Si-W ECAL this algorithm is not applicable and therefore has zero efficiency.

Although this criterion has also been used for the analysis at higher beam energies [14], it is not efficient at lower beam energies, a particularly interesting region for hadronic modelling. Because particularly at small hadron energies shower fluctuations are expected to be strong, a second criterion based on the relative increase in deposited energy is introduced:

$$\frac{E_i + E_{i+1}}{E_{i-1} + E_{i-2}} > \text{Fcut} \quad \text{and} \quad \frac{E_{i+1} + E_{i+2}}{E_{i-1} + E_{i-2}} > \text{Fcut} \ . \tag{2}$$

This measures a relative increase of energy deposition before and after a given layer i. As two consecutive layers are grouped together the variables are less sensitive to local fluctuations in the energy deposition. When both fractions in Eq. 2 are approximately unity the energy deposit is MIP-like; when they are larger they indicate a hadronic interaction. In case where the relative increase continues over several layers one has to make sure that this increase is not an artifact caused by a backscattered particle that deposits energy several cells away from the incoming primary MIP track. To ensure that the increase is caused by the start of an hadronic interaction, the sum of the energies in the cell of the extrapolated primary MIP track (which is found by clustering hits in the first eight layers of the Si-W ECAL) and in the eight cells in the same layer (i) around it ($E_{\text{around},i}$) should satisfy,

$$\frac{E_{\text{around},i}}{E_i} > 0.5 \ . \tag{3}$$

In this analysis the value of Ecut is choosen to be 8 MIPs, which optimises, for 10 GeV events, the interaction-finding efficiency and the standard deviation on the difference between the true and the reconstructed interaction layer. The optimal value of Ecut varies with a maximum of 1 GeV between different Monte Carlo physics lists. The value of Fcut which minimises the contamination with non-interacting events is 6. This value is largely independant of energy and Monte Carlo physics list.

As mentioned in Section 3.2 the contamination of the test beam with electrons is large even after the Cherenkov triggers are applied. In simulated electron and pion events the rejection of events with an interaction found in the first six layers removes 88% of the electron events while removing 22% of the pion events at 2 GeV. At 10 GeV 98% of the electron events are removed while the percentage of removed pion events is 26%. With this additional event cut the final contamination with electrons is reduced from 42% to 10% at 2 GeV. The contamination decreases with energy and at 10 GeV it is negligible. The estimate of the contamination is based on the rejection efficiency found in simulated events and the fraction of rejected events in the data when applying the selection cut.

The second selection criteria (Eq. 2) accepts a small fraction of delta rays. This fraction is estimated to be between 2.3% at 2 GeV and 4.1% at 10 GeV. The estimate is based on the fraction of accepted interacting events in a sample of 50k simulated muons. Because the mass of the muon and of the pion are very close, their behaviour in terms of electromagnetic interactions is very similar.

Table 4 shows the efficiency, η, to find an interaction. The efficiency is defined as the fraction of real interacting events that are correctly classified as interacting by the algorithm. The second column gives the efficiency of the absolute energy criteria and the third the additional efficiency when also the relative energy criteria (Eq. 2) is used. Clearly at low beam energies the second criterion is needed. The third column shows the efficiency to identify the interaction layer with a maximum difference of one layer from the true interaction layer, while the fourth column shows the same for a maximum difference of two layers. As intuitively expected the efficiency increases with increasing energy.

E (GeV)	η	η_{Ecut}	η_{Fcut}	$\eta_{\pm1}$	$\eta_{\pm2}$
2	0.60	0.35	0.25	0.47	0.50
4	0.77	0.61	0.16	0.62	0.65
6	0.85	0.74	0.11	0.72	0.75
8	0.88	0.80	0.08	0.76	0.79
10	0.90	0.83	0.07	0.78	0.82

Table 4: *The interaction-finding efficiency η, decomposed in the contribution of the absolute energy criteria only, η_{Ecut}, and the improvement by Eq. 2, η_{Fcut}, the efficiency $\eta_{\pm1}$ to find interactions within ±1 layer and within ±2 layers, $\eta_{\pm2}$, found from Monte Carlo events* (FTFP_BERT).

Figure 4 shows the distribution of the difference between the reconstructed interaction layer and the true interaction layer.

The interaction-finding efficiencies for the other studied Monte Carlo physics lists are consistent with those found for FTFP_BERT, their maximum difference is 0.02.

Events that are not identified by the criteria described above are considered as non-interacting events. The event sample classified as interacting, however, contains a contamination with non-interacting events of between 2.5% and 3.2% for all energies and physics lists. This contamination is defined as the fraction of events classified as interacting that

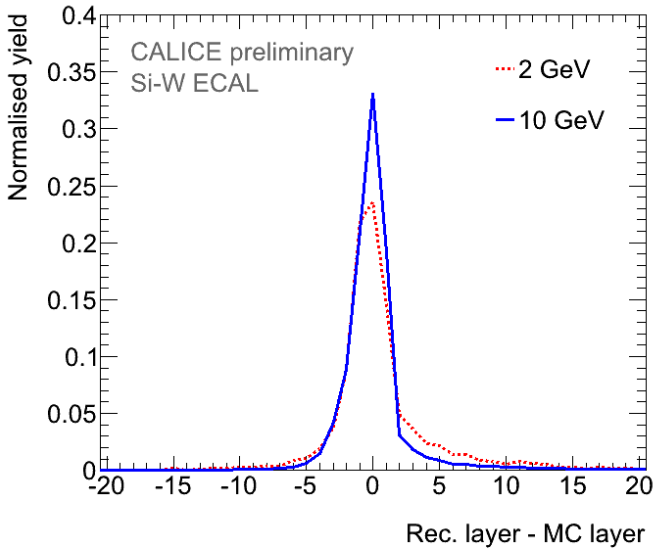


Figure 4: *The difference between the reconstructed interaction layer and the true interaction layer found with the physics list* FTFP_BERT *for 2 and 10 GeV.*

is non-interacting according to the Monte Carlo truth.

Among the events found using Eq. 2, new topologies leading to a smaller number of secondaries appear. An example is shown in Fig. 5. This event features a strong local increase in energy.

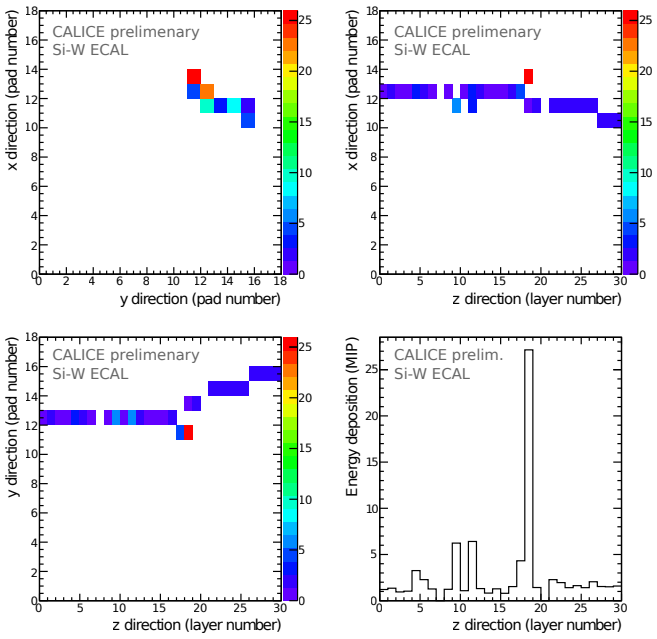
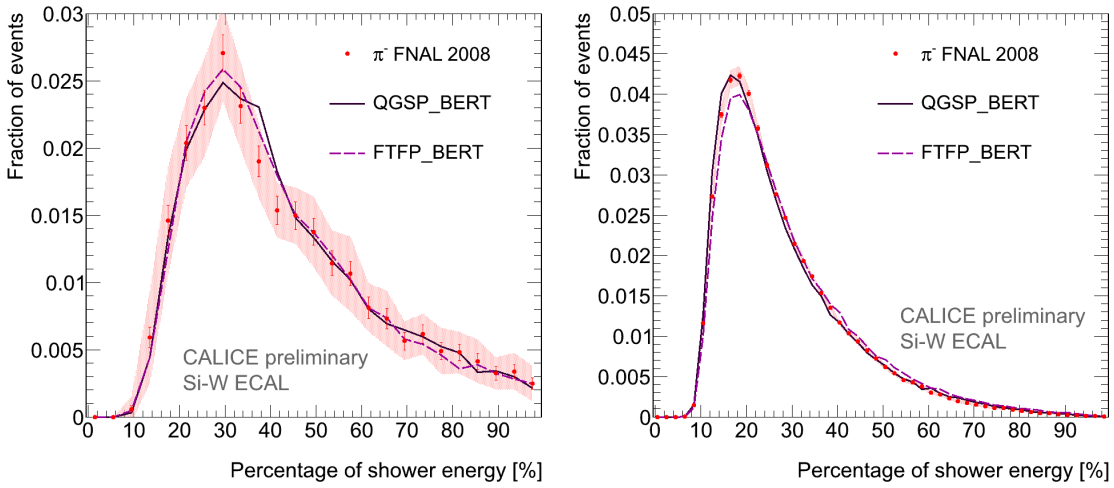


Figure 5: *A hadronic interaction in the Si-W ECAL of a 2 GeV pion.*

The fraction of these kind of events is not negligible at low beam energies, as can be seen in Fig. 6. It shows the highest fraction of the shower energy observed in a single layer for 2 GeV and 10 GeV pions for data and two Monte Carlo physics lists, QGSP_BERT and FTFP_BERT. The other two physics lists, FTFP_BERT_HP and QBBC, give a result

very similar to FTFP_BERT. The data are shown together with their statistical and total uncertainties. The evaluation of the systematic uncertainty as well as a correction applied to the data are described in Section 5.1. At 2 GeV, 20.1% of events in data and between 18.9% and 19.7% of events in Monte Carlo have 60% or more of the energy deposited in one single layer. At 10 GeV this is reduced to 4.2% and between 4.8% and 5.7%.



(a) 2 GeV (b) 10 GeV

Figure 6: The highest shower energy fraction observed in a single layer in 2 GeV (a) and 10 GeV (b) data together with the prediction of event samples simulated using the physics lists QGSP_BERT and FTFP_BERT.

5 Comparing Monte Carlo models with data

After identifying interacting events various Monte Carlo models can be compared to the test beam data in terms of the interaction fraction and radial and longitudinal shower profiles. Events are shown for which the reconstructed interaction layer, i, lies in the range $7 \leq i < 28$.

5.1 Treatment of uncertainties and corrections to the data

The figures shown in the following sections compare the four Monte Carlo physics lists to test beam data. The data are contaminated, especially at low energies, with electrons and also with events having multiple interacting particles. The contamination is reduced by applying triggers and selected cuts (see Section 3.2) and the data have been corrected for the residual contamination. The correction factors have been calculated based on Monte Carlo samples of pions mixed with electrons, single pions mixed with pion plus muon events, and mixed samples of single and double pions events. These have been determined with the physics lists FTFP_BERT and QGSP_BERT and the average correction has been applied to the data. The difference in the correction factors is a contribution to the systematic uncertainty. Furthermore, the systematic uncertainty includes the effect of varying the cut variables to select interacting events, Ecut and Fcut, by one unit and the contamination with non-interacting events. The relative size the systematic uncertainty has been determined with simulated events (FTFP_BERT). The choice of the energy threshold

of 0.6 MIPs on the deposited energy per pixel in the Si-W ECAL has very little influence on the final analysis results: when changed to 0.4 results change by a maximum of 0.6% and when changed to 0.8 the maximum change is 1.2%. This contribution is therefore not included in the systemetic uncertainty.

5.2 Interaction fraction

The interaction fraction is the fraction of interacting events found in the data sample according to the criteria described in Section 4, corrected with the interaction finding efficiency. For the data the efficiency as given by the FTFP_BERT physics list is used. Figure 7 shows the interaction fraction as a function of beam energy for data and the predictions of simulations using the physics lists QGSP_BERT, FTFP_BERT, FTFP_BERT_HP and QBBC. The corrected data are shown with their statistical uncertainties (here smaller than the size of the datapoints) and their total uncertainty (shaded area); statistical and systematic uncertainties added in quadrature, which is dominated by the systematic uncertainty. Some contributions to the systematic uncertainty always increase the interaction fraction while others decrease it (e.g. the opposite effect of changing a cut variable up or down). This is taken into account and results in an asymmetric uncertainty.

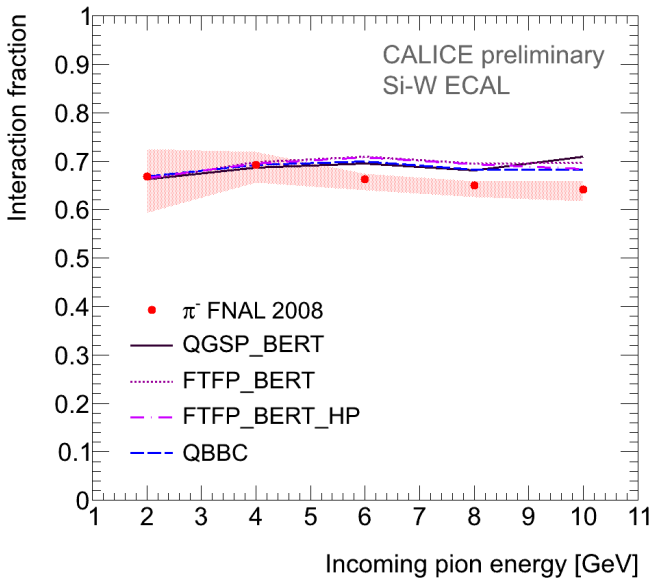


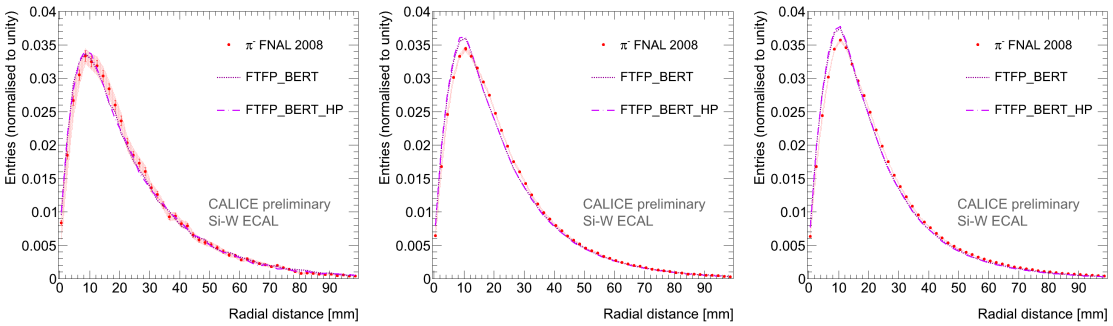
Figure 7: *Interaction fraction for π^- in the Si-W ECAL for data and various Monte Carlo physics lists as a function of beam energy (2 GeV to 10 GeV).*

The interaction fraction is approximately independent of the beam energy and is consistent with the material budget of the Si-W ECAL (one interaction length). For low beam energies the contribution from events with small energy transfer as well as events with high local energy transfer is highest, while at 10 GeV most of the events are selected by the absolute energy threshold criteria. The physics lists are in good agreement with each other, and (at low energies) with the data. At higher energies, all physics lists are found to oberestimate the interaction fraction. On average the difference with the data is of the order of 4%.

5.3 Lateral shower extension

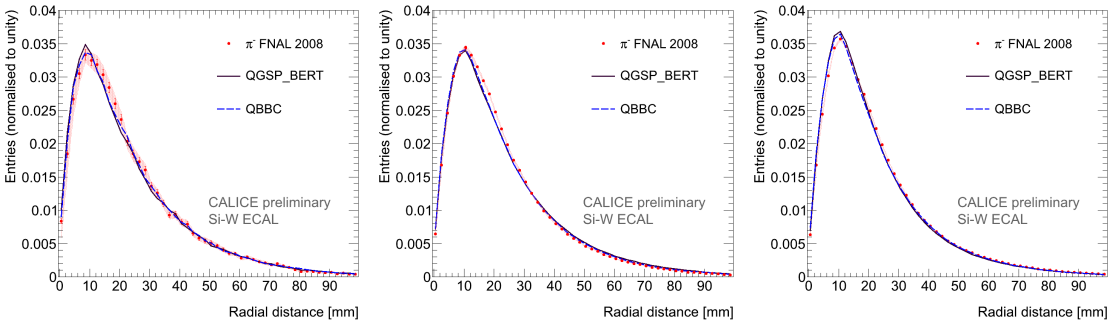
The radial distribution of hits in the shower and the radial energy profile can be used as a measure of the lateral extension of the shower formed as a result of an interaction. Figures 8 and 9 show the radial distance of shower hits with respect to the hit barycentre for beam energies of 2, 6 and 10 GeV. Only hits in the interaction layer and subsequent layers are taken into account. In Fig. 8 the data are compared to FTFP_BERT and FTFP_BERT_HP while in Fig. 9 they are compared to QGSP_BERT and QBBC. The data are shown with their statistical and total uncertainty. All physics lists are within a few percent of the data, however the physics lists FTFP_BERT and FTFP_BERT_HP overestimate around the peak at 6 and 10 GeV.

Figure 10 shows the mean, $< r >$, and standard deviation, $\sqrt{< r^2 > - < r >^2}$, of the radial distance. Both decrease a little with energy. The Monte Carlo is within 10% of the data, however the data are precise enough to distingish between the models. The most pronounced feature in these figures is the clear transition visible for FTFP_BERT and FTFP_BERT_HP between 4 and 6 GeV. This is where the physics lists change from the Bertini cascade to the Fritiof string model. The Bertini model at 2 GeV agrees with the mean radial distance, while at higher energies it is too large. The Fritiof model generates narrower showers than the data. Additionally the high precision treatment of low energy neutrons in FTFP_BERT_HP gives a systematically smaller mean and standard deviation compared to FTFP_BERT.



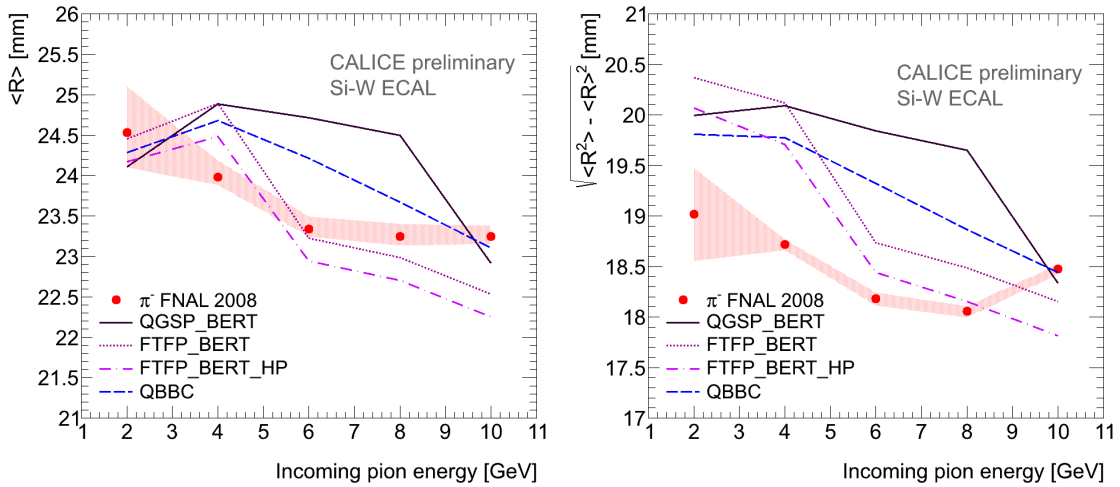
(a) 2 GeV (b) 6 GeV (c) 10 GeV

Figure 8: *The radial distance from the shower centre of hits in the shower for interacting events at 2, 6 and 10 GeV, for data and the Monte Carlo physics lists* FTFP_BERT *and* FTFP_BERT_HP.



(a) 2 GeV (b) 6 GeV (c) 10 GeV

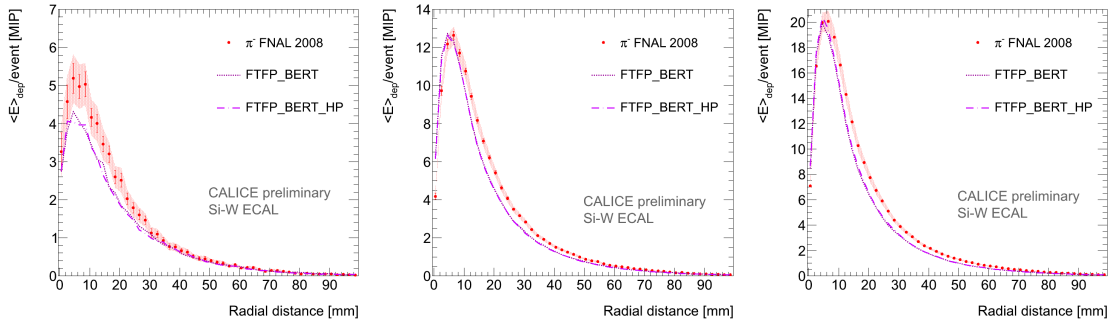
Figure 9: *The radial distance from the shower centre of hits in the shower for interacting events at 2, 6 and 10 GeV, for data and the Monte Carlo physics lists* QGSP_BERT *and* QBBC.



(a) mean (b) standard deviation

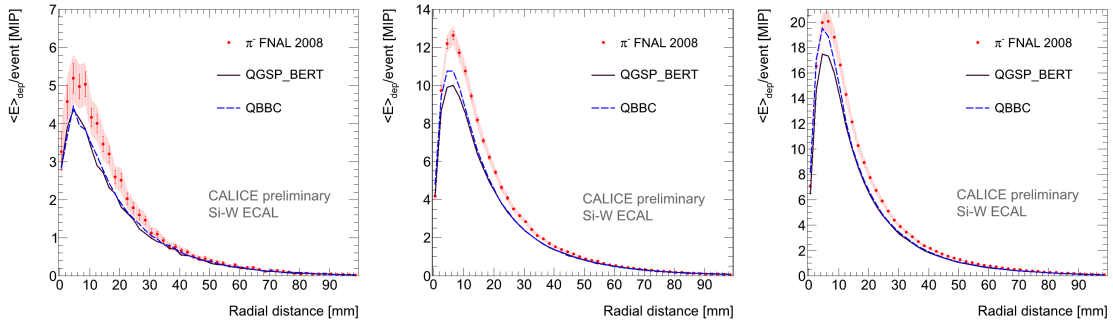
Figure 10: *Mean and standard deviation of the radial distance of hits for interacting events as a function of beam energy (2 GeV to 10 GeV) for data and various Monte Carlo physics lists.*

Figures 11 and 12 show the radial energy profile, defined here as the deposited energy per event as a function of the radial distance to the shower barycentre, for 2, 6 and 10 GeV. In Fig. 11 the data are compared to FTFP_BERT and FTFP_BERT_HP, in Fig. 12 they are compared to QGSP_BERT and QBBC. Overall, the energy deposition is underestimated by all the physics lists, especially QGSP_BERT.



(a) 2 GeV (b) 6 GeV (c) 10 GeV

Figure 11: *The radial energy profile for interacting events at 2, 6 and 10 GeV, for data and the Monte Carlo physics lists* FTFP_BERT *and* FTFP_BERT_HP.

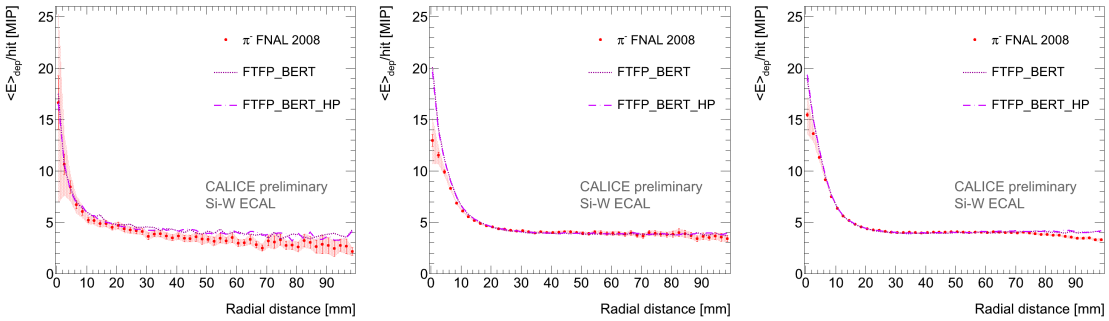


(a) 2 GeV (b) 6 GeV (c) 10 GeV

Figure 12: *The radial energy profile for interacting events at 2, 6 and 10 GeV, for data and the Monte Carlo physics lists* QGSP_BERT *and* QBBC.

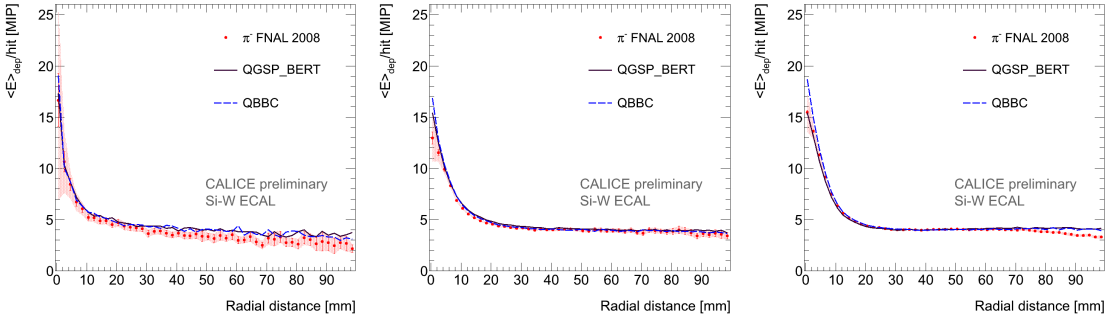
FTFP_BERT and FTFP_BERT_HP hint at an excess for small radii. This excess can be clearly seen when looking at the mean energy per hit as a function of the radial distance, as shown in Fig. 13 and 14. In Fig. 13 FTFP_BERT and FTFP_BERT_HP are compared to the data, which they describe within the uncertainties for 2 GeV while for 6 and 10 GeV they overshoot the data for small radii. This suggests that too much energy is deposited close to the shower axis in the Fritiof model. The effect is much smaller for QGSP_BERT and QBBC, as can be seen in Fig. 14. At 10 GeV QGSP_BERT even underestimates the energy deposition a little possibly due to the admixture of the Low Energy Parameterized model.

In the same way as for the radial distance, Fig. 15 shows the mean and standard deviation of the radial energy profiles as a function of the beam energy. Again the transition in FTFP_BERT and FTFP_BERT_HP is very distinct and it shows clearly that the energy is deposited too close to the shower axis, while for QGSP_BERT the oppsite is true and QBBC follows closest the trend in the data.



(a) 2 GeV (b) 6 GeV (c) 10 GeV

Figure 13: *The radial mean hit energy for interacting events at 2, 6 and 10 GeV, for data and the Monte Carlo physics lists* FTFP_BERT *and* FTFP_BERT_HP.



(a) 2 GeV (b) 6 GeV (c) 10 GeV

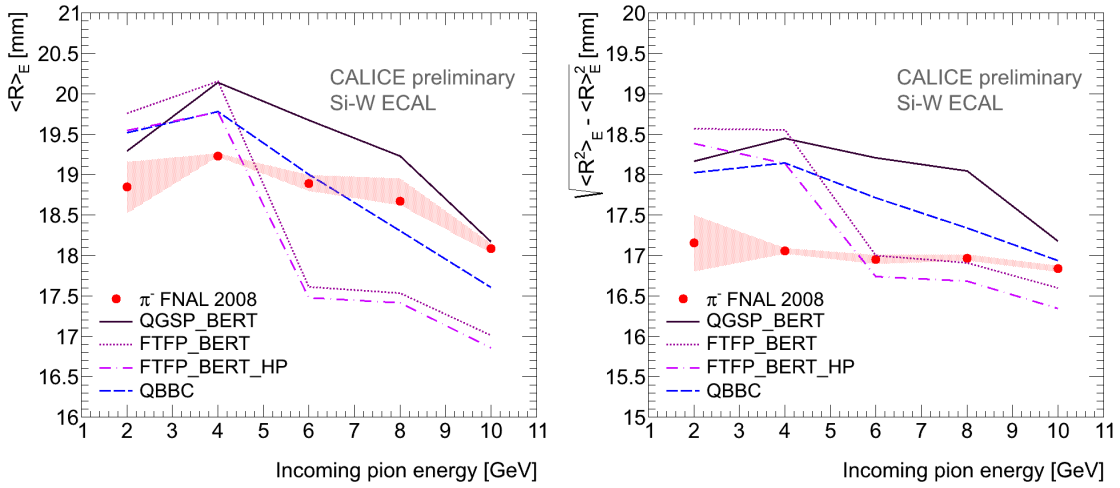
Figure 14: *The radial mean hit energy for interacting events at 2, 6 and 10 GeV, for data and the Monte Carlo physics lists* QGSP_BERT *and* QBBC.

5.4 Longitudinal energy profiles

The next global observable is the longitudinal distribution of hits and that of the deposited energy. Figures 16 and 17 show the hit distribution in the shower as a function of the hit layer. The distributions start at the interaction layer so the x-axis represents the shower depth in layers. Figure 16 shows the distributions at 2, 6 and 10 GeV for the physics lists FTFP_BERT and FTFP_BERT_HP compared to the data while Fig. 17 shows the same for QGSP_BERT and QBBC. The distributions are normalised to 1, so the shape of the distributions can be compared. The longitudinal hit distribution in showers is well modeled by all physics lists. QGSP_BERT and QBBC deviate a little from the data at intermediate energies. At 2 GeV all physics lists overestimate for the first few layers.

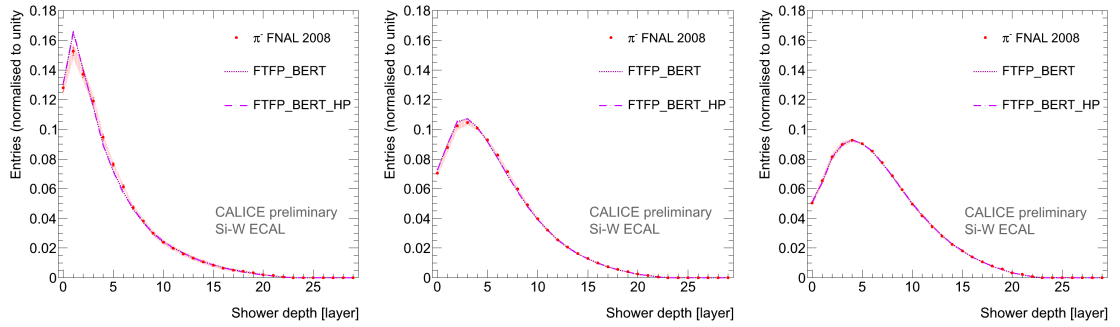
Figure 18 shows the mean, $< z >$, and standard deviation, $\sqrt{< z^2 > - < z >^2}$, of the longitudinal hit distribution for the data and all 4 physics lists. Both are very well described by the Monte Carlo physics lists. The mean increases with beam energy, the standard deviation less strongly.

The longitudinal energy profiles are defined as introduced in [14] and also start from the reconstructed interaction layer. They give the energy in MIPs per *pseudolayer*. Pseudolayers are introduced in order to account for the different sampling fractions in the Si-W ECAL. There is a one to one correspondence between physical layers and pseudolayers in the first module, while each layer in the second module has been subdivided in two



(a) mean (b) standard deviation

Figure 15: *Mean and standard deviation of the radial energy profile for interacting events as a function of beam energy (2 GeV to 10 GeV) for data and various Monte Carlo physics lists.*



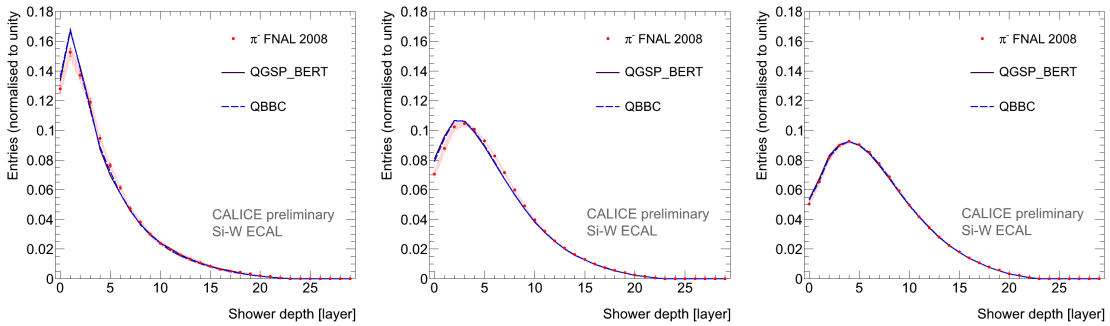
(a) 2 GeV (b) 6 GeV (c) 10 GeV

Figure 16: *The longitudinal shower hit distribution for interacting events at 2, 6 and 10 GeV, for data and the Monte Carlo physics lists* FTFP_BERT *and* FTFP_BERT_HP.

pseudolayers and those in the third module have been subdivided into three pseudolayers. The energy in the added pseudolayers is calculated by interpolating between the energy deposited in the considered physical layer and the energy deposited in the previous physical layer. Figures 19 and 20 show the longitudinal energy profiles for 2,6 and 10 GeV. The Monte Carlo physics lists are again divided over the two figures. The profiles are normalised for each bin separately with the number of entries in that bin.

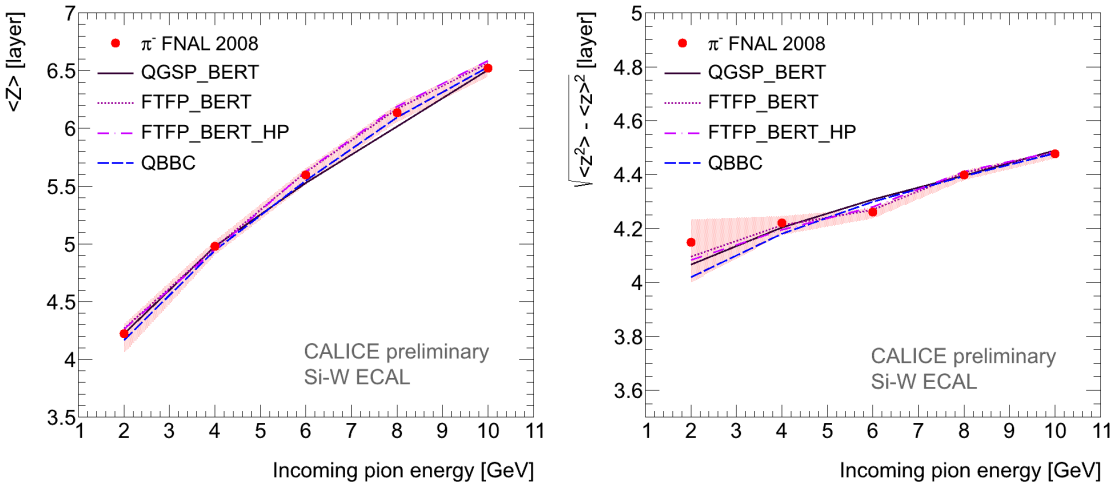
The longitudinal energy profiles are progressively worse described with increasing energy. In general they underestimate the energy deposition, but FTFP_BERT and FTFP_BERT_HP overshoot in the first layers, QBBC less so. This suggest too much energy is being deposited close to the interaction layer. There is very little difference between FTFP_BERT and FTFP_BERT_HP.

Figure 21 shows the mean, $< z >_E$, and standard deviation, $\sqrt{< z^2 >_E - < z >_E^2}$, of the longitudinal profiles for all four physics lists compared to the data. The mean is underestimated at higher energies which supports the observation of too much deposited



(a) 2 GeV (b) 6 GeV (c) 10 GeV

Figure 17: *The longitudinal shower hit distribution for interacting events at 2, 6 and 10 GeV, for data and the Monte Carlo physics lists* QGSP_BERT *and* QBBC.

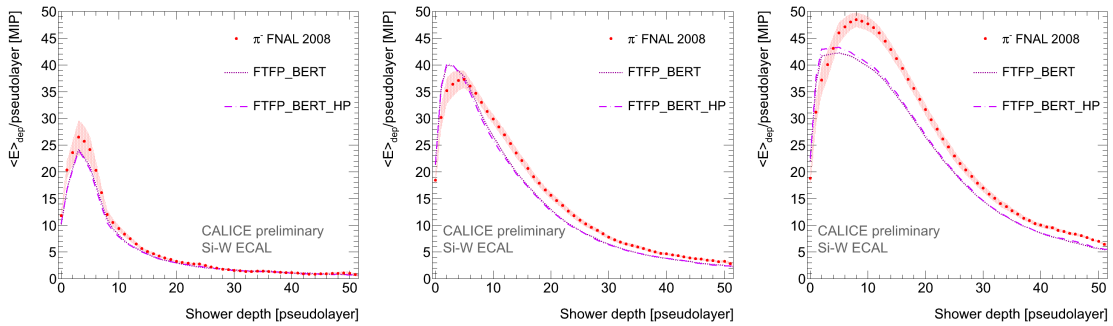


(a) mean (b) standard deviation

Figure 18: *Mean and standard deviation of the longitudinal shower hit distribution for interacting events as a function of beam energy (2 GeV to 10 GeV) for data and various Monte Carlo physics lists.*

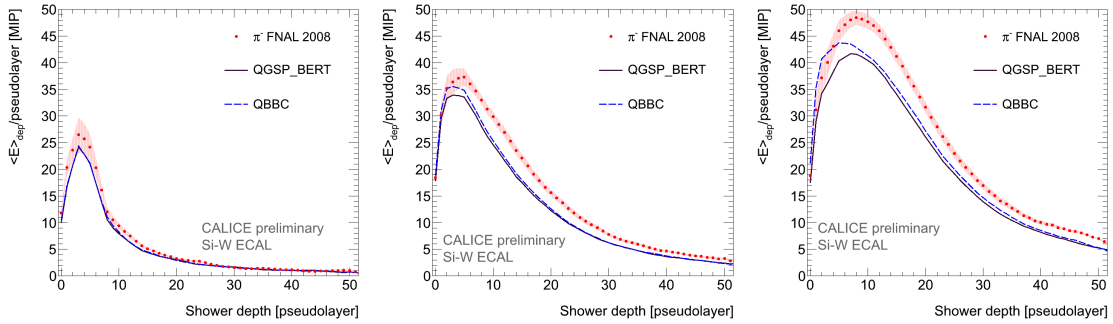
energy near the interaction layer. The standard deviation is also underestimated, except for higher energies, but is compatible with the data within the uncertainties. There is little difference between the physics lists.

The hadronic models implemented in GEANT4 are constantly being revised and improved to describe best the available data. The analysis presented in this paper started out using GEANT4 version 9.3 and later updated to version 9.6. Between these two versions the Fritiof string model has been heavily revised and tuned based on thin target data and LHC test beam data, while the Bertini cascade model has undergone minor revisions. Unfortunately, the changes in the Fritiof model have led to a larger mismatch between the data and the physics list FTFP_BERT in the longitudinal energy profile, as is illustrated in Fig. 22. Version 9.3 FTFP_BERT describes the data reasonably well at 10 GeV, while version 9.6 clearly does not. On the other hand the longitudinal hit distribution is well modeled and the change between the versions is minimal. For QGSP_BERT such a change in the longitudinal energy profile is not seen. The discrepency with the data could be related



(a) 2 GeV (b) 6 GeV (c) 10 GeV

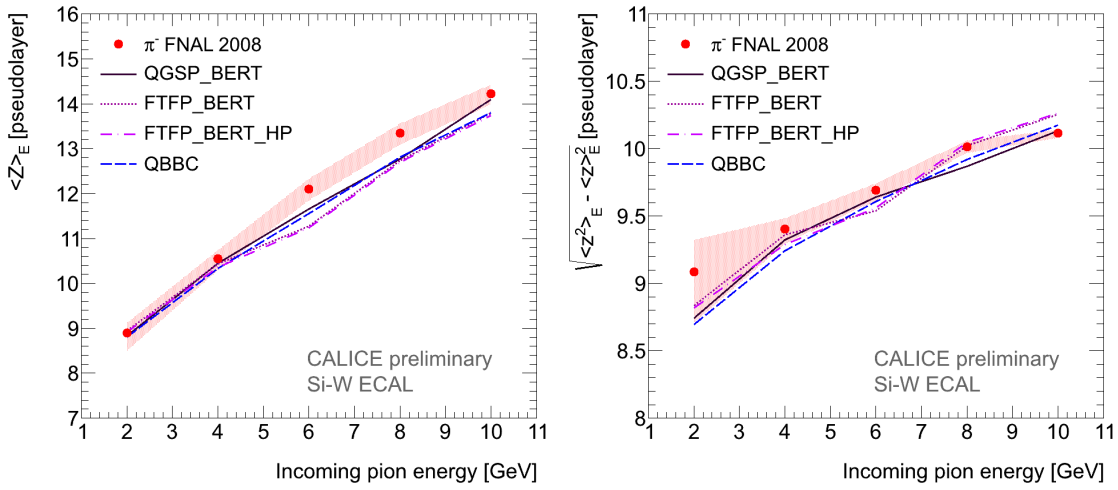
Figure 19: *The longitudinal energy profile for interacting events at 2, 6 and 10 GeV, for data and the Monte Carlo physics lists* FTFP_BERT *and* FTFP_BERT_HP.



(a) 2 GeV (b) 6 GeV (c) 10 GeV

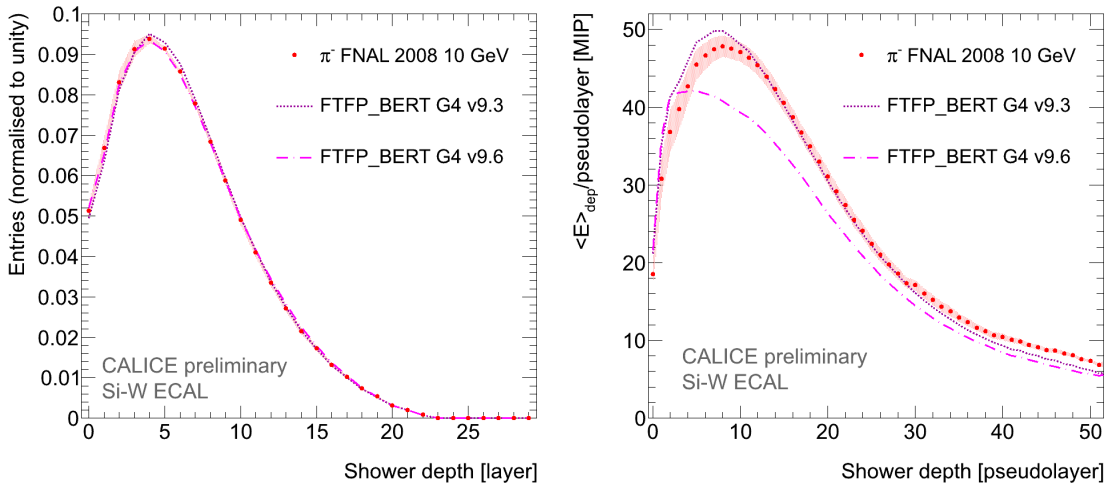
Figure 20: *The longitudinal energy profile for interacting events at 2, 6 and 10 GeV, for data and the Monte Carlo physics lists* QGSP_BERT *and* QBBC.

to the sensitive material of the prototype, silicon for the Si-W ECAL, as the optimisation of the Fritiof model has been mostly done with data obtained from detectors with scintillator as sensitive material. In a recent CALICE publication [15] the longitudinal energy profile of pions in a Scintillator-Tungsten hadronic calorimeter prototype is well described by FTFP_BERT in GEANT4 version 9.6.



(a) mean (b) standard deviation

Figure 21: *Mean and standard deviation of the longitudinal energy profile for interacting events as a function of beam energy (2 GeV to 10 GeV) for data and various Monte Carlo physics lists.*



(a) Longitudinal hit distribution (b) Longitudinal energy profile

Figure 22: *The longitudinal hit distribution (a) and energy profile (b) for interacting events at 10 GeV, for data and the Monte Carlo physics list* FTFP_BERT *for two different* GEANT4 *versions.*

6 Summary, Conclusions and Outlook

A deep understanding of hadronic showers is of general interest and naturally supports the future development of particle flow algorithms for detectors at a future linear collider. This study demonstrates the large potential of the CALICE Si-W ECAL to obtain a detailed image of the early part of hadronic cascades. Data obtained in test beams with negatively charged pions with an energy between 2 and 10 GeV are compared with Monte Carlo predictions employing different physics lists as contained in the simulation toolkit GEANT4. If an hadronic interaction takes place within the Si-W ECAL volume, the start of the shower can be reconstructed with an accuracy of $\pm$ 2 layers at an efficiency of at least 50% at 2 GeV and 82% at 10 GeV. At the low beam energies studied here interactions are selected using not only the absolute energy increase in subsequent layers but also the relative energy increase.

The quality with which the Monte Carlo describe the data varies with the beam energy and the chosen physics observable. There is little difference between the physics lists and none describe the entire set of data. The longitudinal hit distribution is very well described as is the radial hit distribution, although there the mean is shifted. On the other hand the physics observables which take into account the energy deposition are not reproduced well by the Monte Carlo. Combining the longitudinal and radial energy profiles it seems that especially the Fritiof model deposits too much energy near the interaction region.

The radial distributions prove to be sensitive to the different hadronic models implemented in the physics lists. The transition between the Bertini cascade and Fritiof string model in FTFP_BERT and FTFP_BERT_HP is much more pronounced in the mean and standard deviation of the radial observables compared to the longitudinal observables. Additionally the deviations of the phyics lists from the data and each other are larger. The precision treatment of neutrons in FTFP_BERT_HP gives smaller mean and standard deviations. The results for QBBC tend to be in between QGSP_BERT and FTFP_BERT.

In conclusion no preference for a hadronic model is seen as none of the physics lists reliably reproduce the data in detail. The mean deficiencies are in the longitudinal and radial energy profiles. The observables that are well described show very little difference between physics lists.

Future analysis into hadronic showers will attempt to classify inelastic reactions in terms of shower topology. This comprises the determination of size and energy density of the interaction region as well as the measurements of tracks emerging from the interaction region. These steps will exploit further particularly the lateral granularity of the Si-W ECAL which will be even higher in the next prototype. They may form a solid base for the development or the improvement of particle flow algorithms.

References

[1] J. C. Brient and H. Videau. The calorimetry at the future e^+e^- linear collider. In *Proceedings of the APS / DPF / DPB Summer Study on the Future of Particle Physics (Snowmass 2001)*, 2001. arXiv:hep-ex/0202004v1.

[2] T. Behnke et al. The International Linear Collider Technical Design Report - Volume 4: Detectors. 2013. arXiv:1306.6329 [physics.ins-det].

[3] M. A. Thomson. Particle Flow Calorimetry and the PandoraPFA Algorithm. *Nucl. Instrum. Meth. A*, 611:25, 2009. arXiv:0907.3577v1 [physics.ins-det].

[4] CALICE Collaboration webpage. http://twiki.cern.ch/CALICE.

[5] 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. *J. Instrum.*, 3:P08001, 2008. arXiv:0805.4833v1 [physics.ins-det].

[6] Fermilab Test Beam Facility. http://www-ppd.fnal.gov/MTBF-w.

[7] The CALICE Collaboration, C. Adloff, et al. Construction and Commissioning of the CALICE Analog Hadron Calorimeter Prototype. *J. Instrum.*, (5):P05004, 2010. arXiv:1003.2662v1 [physics.ins-det].

[8] The CALICE Collaboration, C. Adloff, et al. Construction and performance of a silicon photomultiplier/extruded scintillator tail-catcher and muon-tracker. *J. Instrum.*, (7):P04015, 2012. arXiv:1201.1653 [physics.ins-det].

[9] The GEANT4 Collaboration. GEANT4 – a simulation toolkit. *Nuclear Instruments and Methods in Physics Research*, A 506:250–303, 2003. http://geant4.web.cern.ch/geant4.

[10] Mokka webpage. http://polzope.in2p3.fr:8081/MOKKA.

[11] The CALICE collaboration, C. Adloff, et al. Response of the CALICE Si-W Electromagnetic Calorimeter Physics Prototype to Electrons. *Nucl. Instrum. Meth. A*, 608:372, 2009. arXiv:0811.2354 [physics.ins-det].

[12] J. Apostolakis et al. Validation of GEANT4 hadronic models using CALICE data. *EUDET-MEMO-2010-15*, 2010. http://www.eudet.org/e26/e28/e86887/e109012/EUDET-Memo-2010-15.pdf.

[13] A. Dotti et al. Description of Hadron-induced Showers in Calorimeters using the GEANT4 Simulation Toolkit. *Proceedings of the IEEE NSS MIC 2011 Conference*, 2011. http://geant4.web.cern.ch/geant4/results/papers/hadronic-showers-IEEE11.pdf.

[14] The CALICE Collaboration, C. Adloff, et al. Study of the interactions of pions in the CALICE silicon-tungsten calorimeter prototype. *J. Instrum.*, 5:P05007, 2010. arXiv:1004.4996 [physics.ins-det].

[15] The CALICE Collaboration, C. Adloff, et al. Shower development of particles with momenta from 1 to 10 GeV in the CALICE Scintillator-Tungsten HCAL. *J. Instrum.*, (9):P01004, 2014. arXiv:1311.3505v2 [physics.ins-det].