

# Pion Response and Resolution in a Combined Scintillator Calorimeter System in the FNAL Testbeam

The CALICE Collaboration <sup>1</sup>

## Abstract

This note presents results obtained with the combined scintillator-SiPM calorimeter system consisting of the *Scintillator Electromagnetic Calorimeter*, *Analogue Hadronic Calorimeter* and *Tail Catcher & Muon Tracker* in the energy range 4–32 GeV at the FNAL testbeam campaign in 2009. The analysis includes longitudinal profiles as well as response linearity and energy resolution of pions for standard and software compensation reconstruction. The results are compared to several GEANT 4 physics lists in software versions 9.6 and 10.1.

*This note contains preliminary CALICE results, and is for the use of members of the CALICE Collaboration and others to whom permission has been given.*

---

<sup>1</sup>Corresponding author:  
Oskar Hartbrich; oskar.hartbrich@desy.de

# Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>The FNAL 2009 Testbeam Setup</b>	<b>3</b>
<b>3</b>	<b>Simulation</b>	<b>4</b>
3.1	ScECAL Simulation Model & Digitisation . . . . .	5
3.2	ScECAL Model Validation . . . . .	5
<b>4</b>	<b>Run &amp; Event Selection</b>	<b>10</b>
4.1	Pion Selection . . . . .	10
<b>5</b>	<b>Pion Energy Reconstruction</b>	<b>13</b>
5.1	Standard Weighting . . . . .	13
5.2	Software Compensation . . . . .	14
<b>6</b>	<b>Results</b>	<b>23</b>
6.1	Profiles . . . . .	23
6.2	Energy reconstruction and linearity . . . . .	24
6.3	Energy resolution . . . . .	25
6.4	Applying Software Compensation Weights from Simulation to Data .	28
<b>7</b>	<b>Summary</b>	<b>28</b>
	<b>Bibliography</b>	<b>31</b>
	<b>Appendix</b>	<b>34</b>

# 1 Introduction

$\alpha_i, \beta_i$  as well as the TCMT weight  $\gamma$ . Experiments at future  $e^+e^-$  colliders require unprecedented jet energy resolutions of 3–4% across the full expected energy range. One concept to achieve such resolutions is *Particle Flow Algorithms* (PFAs) aiming to combine measurements from the tracking system with calorimeter depositions, which needs exceptional granularity in the calorimeters. The CALICE collaboration develops, builds and tests different calorimeters that aim to fulfill the requirements for ideal application of PFAs.

One explored option consists of scintillator tiles or strips of a few  $\text{cm}^2$  size, read out by one individual *Silicon Photomultiplier* (SiPM) per cell. Several prototypes with different absorbers, granularities and sampling structures have been constructed.

In a common beamtest at FNAL in 2009 the *Scintillator Electromagnetic Calorimeter* (ScECAL), *Analogue Hadronic Calorimeter* (AHCAL) and *Tail Catcher & Muon Tracker* (TCMT) have been operated as a combined scintillator-SiPM calorimeter system in beams of muons, electrons and pions in the energy range 1–32 GeV. Studying the characteristics of single pion events is of special interest, as they can deposit significant amounts of energy in all detector parts and their calorimetric energy resolution is a significant contribution to the expected achievable PFA performance.

Apart from PFA reconstruction, the high granularity of the used calorimeters on the scale of electromagnetic shower development ( $\approx 1 X_0$  longitudinal sampling,  $\approx 1 r_M$  transverse segmentation) enables the identification of electromagnetic sub-showers within hadron showers and thus the application of software compensation techniques to improve the energy resolution of the calorimeter system.

This note presents results obtained from the FNAL 2009 testbeam, combining depositions in the ScECAL, AHCAL and the first eight layers of TCMT.

## 2 The FNAL 2009 Testbeam Setup

The data sets used in this note were acquired during a testbeam campaign in May 2009 at the Fermilab Testbeam Facility (FTBF). Its MTest beamline delivers secondary particle beams with particle momenta in the range 1–32 GeV. The beamline is equipped with a two channel differential gas Cherenkov counter as well as various scintillator triggers. The calorimeter prototypes are installed in the order ScECAL, AHCAL, TCMT downstream of the beam instrumentation. More detailed descrip-

tions of the beamline setup are available in [1, 2].

One ScECAL layer consists of 3.5 mm W-based absorber and 72 scintillator strips of  $10 \times 45 \times 3 \text{ mm}^3$  each for a total area of  $18 \times 18 \text{ cm}^2$ . 30 layers are stacked for a depth of around  $20X_0$  ( $\approx 1 \lambda_n$ ) and 2160 channels total [2]. The AHCAL is a  $1 \times 1 \times 1 \text{ m}^3$  calorimeter prototype with Fe absorber plates, instrumented with 7608 5 mm thick scintillator tiles of size  $3 \times 3 \text{ cm}^2$ – $12 \times 12 \text{ cm}^2$  in 38 layers for a full depth of 5.3 nuclear interaction lengths  $\lambda_n$  [3]. The TCMT covers  $1 \times 1 \text{ m}^2$  Fe absorber plates with a total of 320 scintillator strips of  $5 \times 100 \text{ cm}^2$  each in 16 layers, instrumenting a depth of  $5.2 \lambda_n$  [4].

### 3 Simulation

The simulation models for AHCAL and TCMT are used as in [1, 5], the ScECAL is described in the next section. The right handed coordinate system is laid out such that the Z-axis is pointing in the beam direction and the Y-axis is pointing upwards.

Depending on the pressure in the Cherenkov gas tube, the material between final beam energy selection magnet and calorimeter front face amounts to 0.08-0.16  $X_0$  distributed over around 30 m as estimated in [1]. As the helium filled beam transport tubes and most of the beam instrumentation is not represented in the simulation model, a similar amount of material between beam generation and calorimeter front face is achieved by filling the world volume with air, setting the beam gun at 6 m from the calorimeter front face and adding 5 mm of Al directly in front of the ScECAL, amounting to around 0.1  $X_0$  in total. The beam momentum spread is set to 2.7% for runs with beam energy  $\leq 4 \text{ GeV}$  and 2.3% for runs with higher beam energies following [6]. The beam profile is extracted from data and transferred into the simulation for each run individually.

All simulated electron shower events in this note are produced with the QGSP\_BERT physics list in GEANT4 10.1p2, as no significant differences have been found to any other tested physics list. Pion showers are simulated using QGSP\_BERT\_HP in GEANT4 10.1p2 as well as FTFP\_BERT\_HP in GEANT4 versions 10.1p2 and 9.6p3. Per simulated run 100,000 single particle events are generated.

### 3.1 ScECAL Simulation Model & Digitisation

The ScECAL simulation model is constructed out of 30 single layers. Each logical layer is primarily composed of 3.5 mm absorber plate, 3 mm scintillator tile, 1 mm PCB and 1.24 mm air. The thicknesses of absorber, scintillator and PCB were measured precisely. The depth of air between layers follows from the measured total length of the ScECAL [2]. Density and material composition of scintillator and PCB are taken as the defaults for polystyrene and G10 in GEANT 4. X-Ray diffraction (XRD) measurements show that the ScECAL absorber plates consist of WC (tungsten carbide) and Co, possibly a negligible amount of Cr, but no significant amount of elementary W [7]. Assuming no air or vacuum bubbles, as supported by microscopic scale back scattered electron imaging (BSEI), the composition of the absorber plates can thus be calculated from its measured density and the known densities of WC and Co to be 74.6% W, 4.9% C and 20.5% Co in mass fractions.

The digitisation of simulated hits closely resembles the AHCAL algorithm described in [8], using (where available) individual calibration factors per channel obtained from data to extract the light yield which is needed to model the statistical fluctuations of photons hitting the SiPM. Saturation effects are included in the same way, but only one global parameter is available for all cells (see subsection 3.2). Since each scintillator strip in the ScECAL is individually wrapped in reflective foil, optical crosstalk effects between channels are assumed to be negligible and thus not included. Noise is extracted from random trigger events for each data run and overlaid onto simulated events. After the digitisation all simulated hits have the same format as raw data hits and are reconstructed with the same software chain that is used for data. The relative alignment between ScECAL and AHCAL is determined from data using tracks extracted from muon runs. To suppress noise hits, only hits above 0.5 MIP are considered in the analysis in both data and simulation.

### 3.2 ScECAL Model Validation

To validate the simulation and digitisation model described above, comparisons of electromagnetic interactions within the ScECAL are preferred as such interactions should be well described in simulations.

One of the lowest level observables to compare is the spectrum of a MIP-like particle traversing a scintillator strip. A sample of such hits is obtained from straight tracks by muons and pions punching through the whole ScECAL without a hard

interaction. The comparison for a single ScECAL strip in Figure 1(a) shows that the shape of the hit energy spectrum is well described in the digitised simulation. Fitting these spectra for all strips and extracting the most probable value yields the distribution in Figure 1(b). The mean of the distribution for both data and simulation is very close to unity. The width of the data distribution is slightly higher than in simulation, as channel-wise miscalibrations are not included in the digitisation.

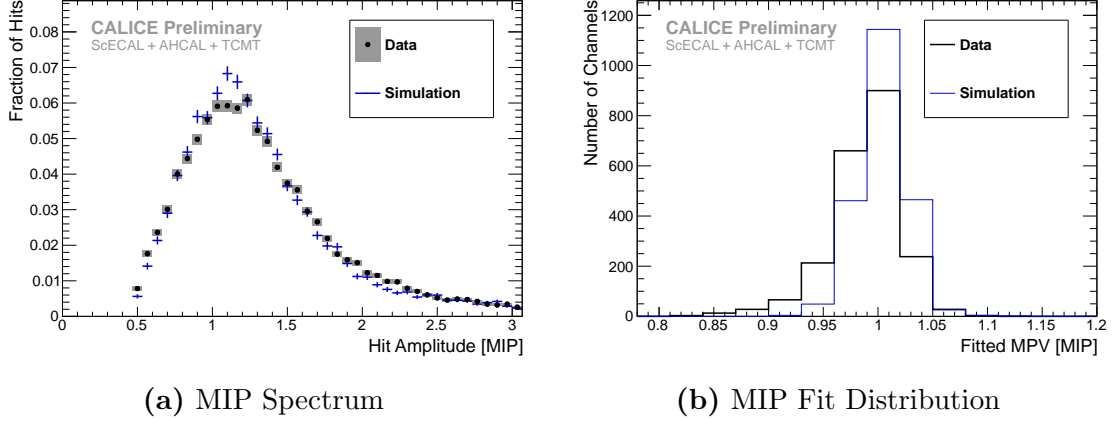


Figure 1: Comparison of MIP-like particles in data and simulation.

(a): Hit energy spectrum of a single ScECAL cell in run 560474 ( $32 \text{ GeV } \pi^-$ ), normalised to the number of hits.

(b): Distribution of fitted most probable values of fits to the MIP spectra in single ScECAL cells from Run 560269 ( $32 \text{ GeV } \mu^-$ ).

$$\mu_{\text{data}} = 0.99, \text{RMS}_{\text{data}} = 0.040, \mu_{\text{MC}} = 1.00, \text{RMS}_{\text{MC}} = 0.022.$$

To perform further comparisons on electron showers, a sample of single electron events is selected from the recorded data runs. A simple electron selection, based on the inversion of the pion selection described in subsection 4.1, incorporating the beamline Cherenkov counters and beam scintillators, is applied. As even electrons of the highest energy available from the beamline are expected to be fully contained in the ScECAL, additionally only events with low activity in the AHCAL are accepted to reject pions. A detailed list of cuts performed is given in Table A2 in the Appendix. As shown in Figure 2 the selection yields a clean sample of single electron events.

The comparison of hit energy distributions in Figure 3(a) shows a significantly longer tail towards higher hit energies in simulated events than data. Assuming a perfect description of the physics of electron showers in the simulation, this could be either caused by a wrong or incomplete treatment of SiPM saturation effects or a wrong material description in the simulation model.

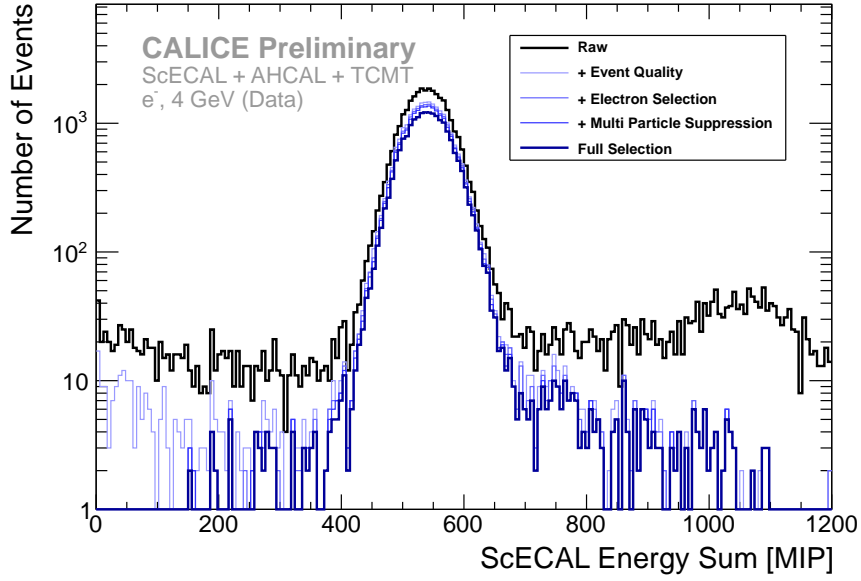


Figure 2: Reconstructed energy spectrum of data run 560332 (4 GeV  $e^-$ ) for different steps of the applied event selection.

The unfolding of SiPM saturation in ScECAL data is performed by assuming the same number of effective pixels per sensor for all cells, using the mean of the number of effective pixels measured from a subset of ScECAL strips (as described in detail in [6]). Re-reconstructing the dataset with this parameter reduced by 20% ( $\approx 2\sigma$  of the measured distribution) yields the spectra and profiles labelled as *Data (Sat. Scale)* in the following. This leads to slightly longer tails in the hit energy distribution in data for all beam energies, although not enough to make data and simulation agree, as shown in Figure 3(a). However the specific behaviour observed in the type of SiPM used in the ScECAL (see [6]) is not included into the desaturation calculation during reconstruction and is still under active investigation [9].

An incorrect description of the materials in the simulation model would lead to a discrepancy in observed shower profiles between data and simulation. The transverse shower profile, as the distribution of energy weighted distances to the reconstructed center-of-gravity of the shower, in Figure 3(b) indeed shows a significantly narrower shower core in simulation compared to data for all beam energies, leading to higher energy deposition densities and thus potentially explaining the harder hit energy spectrum. The longitudinal profile in Figure 4 is however well described, indicating that the amount of absorber material in  $X_0/\text{Layer}$  is well modelled in simulation.

The only way to increase the effective *Molière Radius*  $r_M$  nearly without influencing  $X_0/\text{Layer}$  is increasing the thickness of air between individual ScECAL layers. Separate simulations with modified air gap between individual ScECAL layers (doubled from 1.24 mm to 2.48 mm) labelled *MC (Airgap\*2)* are shown in the figures below. The wider air gap slightly increases  $r_M$  and reduces the tail in the hit energy spectrum, but not enough to make the simulation agree with data. A further increase of the air gap is entirely unphysical, as the doubled air gap already makes the modified ScECAL simulation model  $30 \cdot 1.24 \text{ mm} \approx 3.7 \text{ cm}$  longer than the measured length of the ScECAL prototype. Possible effects from remaining pion contaminations are considered to be small, as the pion fraction should show a strong dependence on the beam energy, while the observed mismatch in the transverse shower profile is We thus consider the possibility that the simulation model used in this study indeed produces slightly too narrow electromagnetic showers. Shifting the saturation scale in data reconstruction has a small influence on the measured transverse shower profile, but does not improve the agreement with simulations.

The studied effects produce shifts in the order of the observed discrepancies between data and simulation in the hit energy spectra while keeping the longitudinal profile compatible to data, but the description of the transverse shower profile does not improve in the same way. No final conclusion whether the discrepancies between data and simulation can be attributed to the shower simulation or the geometry modelling and SiPM saturation can be drawn. The datasets obtained with modified simulations and reconstruction parameters are thus used as estimates on the systematic uncertainties on shower simulation, geometry modelling and SiPM desaturation.

The longitudinal profile of electron showers is generally well described by simulations in both shape and amplitude, see Figure 4. The remaining fluctuations from layer to layer can be accounted to channel-wise miscalibrations (which are not fully included in the simulation) in combination with most energy being deposited in few cells per layer. The systematic effects described in the previous paragraphs show very small influence within the existing fluctuations. Generally the measured response calibrated to the MIP scale in data and simulations agrees within a few %, and the description of electromagnetic showers in simulation is satisfactory for the study of hadron showers.



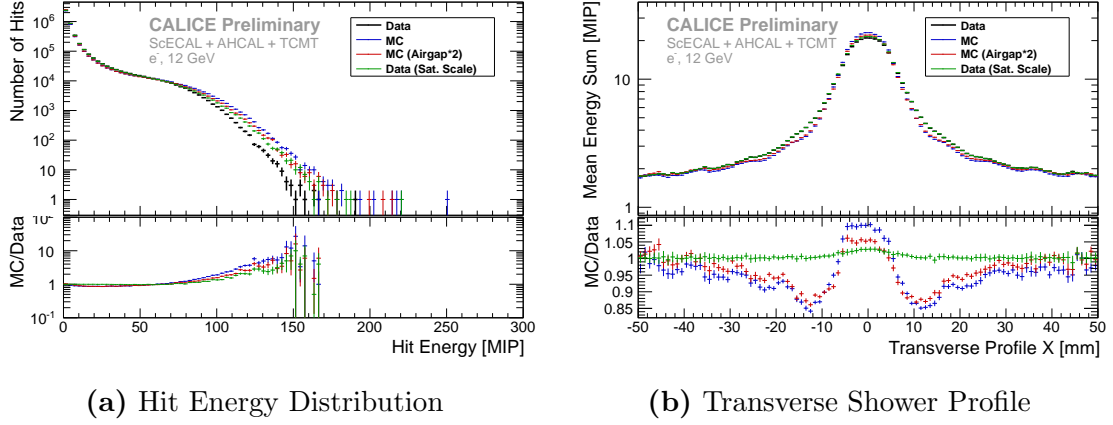


Figure 3: Electron shower hit energy distribution and transverse shower profile of run 560294 (12 GeV  $e^-$ ) for data, simulation and possible systematic effects in simulation and reconstruction. *MC (Airgap\*2)* are simulated events obtained from a simulation model with doubled amount of air between ScECAL layers. *Data (Sat. Scale)* is the same raw data set as *Data*, but with the effective number of SiPM pixels used for saturation unfolding reduced by 20% in reconstruction.

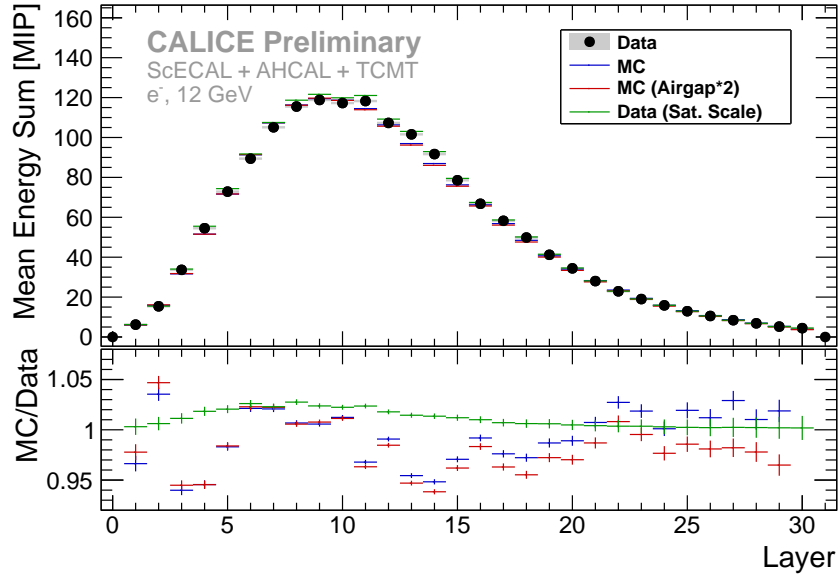


Figure 4: Mean longitudinal shower profile of electron run 560294 (12 GeV  $e^-$ ) for data, simulation and possible systematic effects in simulation and reconstruction.

## 4 Run & Event Selection

During the FNAL testbeam period in 2009,  $\pi^-$ -runs were taken with beam momenta ranging from 2 GeV to 32 GeV. The 2 GeV energy point is omitted for containing a large admixture of electrons and multi-particle events as shown in [1, 10], a very wide beam profile due to multiple scattering in beam instrumentation and air downstream of the final selection magnet as well as inefficient and imprecise determination of the layer of first hard interaction, leading to an inefficient and impure selection of pions for further analysis.

The data and event numbers used for this analysis are listed in Table 1.

Table 1: Data runs used in this analysis, including number of events before and after selecting for single pions.

Run	Energy	Events (raw)	Events (sel.)	$\frac{\text{Events (sel.)}}{\text{Events (raw)}}$
560506	4 GeV	202,943	45,518	22.4%
560498	12 GeV	164,341	50,548	30.8%
560496	15 GeV	225,112	69,058	30.7%
560481	20 GeV	169,131	50,577	29.9%
560474	32 GeV	215,992	60,596	28.1%

### 4.1 Pion Selection

As the MTest beamline at FTBF does not offer a direct selection of particle type apart from its polarity, the delivered particle beam is a mixture of mostly electrons, pions and muons in varying fractions depending on the beam energy. Especially for the lower range of beam energies there is also a large fraction of events with multiple particles hitting the calorimeters at the same time, see [1, 5, 10]. The goal of the event selection described in the following is to efficiently select events containing a single, contained pion shower in the detector system.

The first selection step evaluates readouts from the DAQ and external beam instrumentation, differential Cerenkov counters, the multi-particle counter and other trigger scintillators, as well as the exclusion of empty events (*Event Quality*).

Then a *Pion Selection* is performed to suppress electron and muon events in the sample. Single muons and punch-through pions are suppressed by rejecting events based on number of hits and center of gravity along the beam axis in the AHCAL as

explained in [1, 5]. Electrons are suppressed by requiring the reconstructed layer of first hadron interaction (FHI layer) to be  $\geq 5$ , which also removes events that have started showering upstream of the calorimeters. The FHI layer reconstruction in the combined system is based on the *AHCAL Primary Track Finder* algorithm [11], extended with optimised thresholds to also work in the ScECAL. The distribution of reconstructed FHI layers is very well described by all tested simulations models as shown in Figure 5.

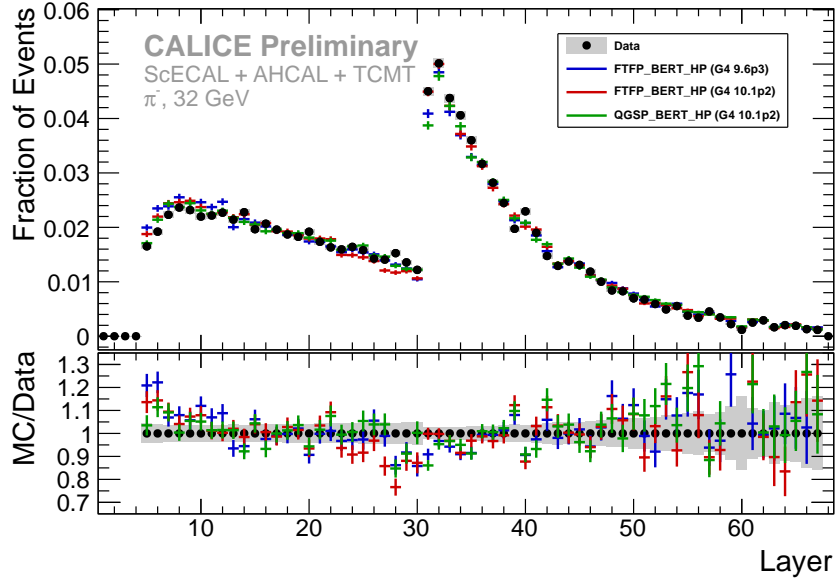


Figure 5: Reconstructed FHI layer for 32 GeV  $\pi^-$  (Run 560474) in data and different simulation physics lists.

The suppression of events with multiple pion showers in the detector is achieved by reconstruction of the primary MIP-like track a pion is producing in the ScECAL before its first hard interaction. Only events with exactly one isolated primary track are selected. To suppress events with additional muons entering the AHCAL outside of the coverage of the ScECAL and multi particle counter, all events with tracks longer than five layers parallel to the beam axis in the outer parts of the AHCAL, reconstructed with the track segment finder algorithm described in [12], are rejected (*Multi particle suppression*).

Finally, to select for showers that are laterally and longitudinally contained, only events with the primary pion track around the center of the ScECAL and reconstructed FHI layer at the latest in the fifth AHCAL layer are accepted into the

analysis (*Containment*). The ratio of mean contribution to the reconstructed energy in ScECAL and AHCAL  $r = E_{\text{rec}}^{\text{AHCAL}}/E_{\text{rec}}^{\text{ScECAL}}$  varies from around unity (at 4 GeV) to around three (at 32 GeV).

An example of the full step-by-step event selection in data and simulation is shown in Figure 6. The resulting reconstructed energy spectrum contains only a minor fraction around  $10^{-2}$  remaining contamination in data, leading to only very small systematic uncertainties in the extraction of response and resolution. A detailed table of cuts performed is given in Table A3 in the Appendix.

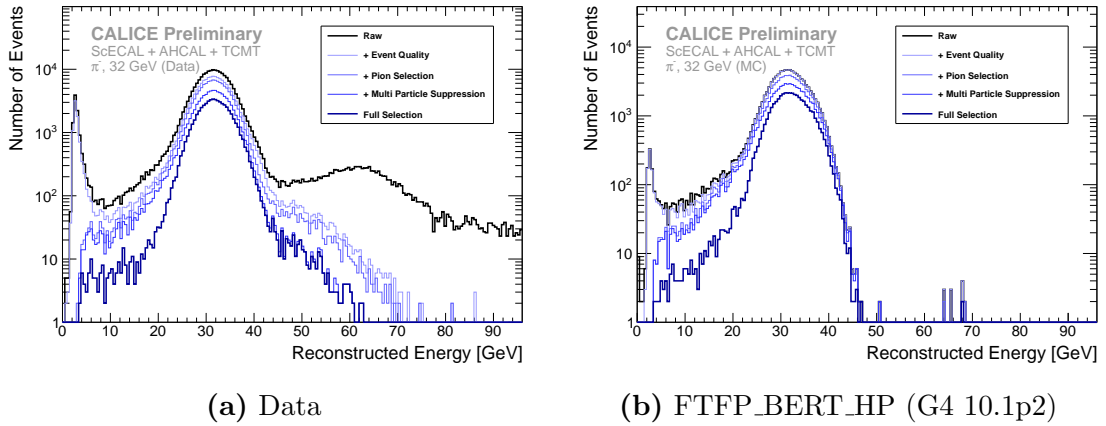


Figure 6: Reconstructed energy spectrum of run 560474 (32 GeV  $\pi^-$ ) in data and simulation for different steps of the applied event selection.

Selection efficiencies and biases are studied from simulated event samples for different physics lists and particle types. Table A4 in the Appendix shows that 45–50% of simulated pions pass the pion selection across all examined physics lists. Most of the excluded pion events are rejected by the FHI layer requirements for electron suppression and shower containment. From the selection requirements of FHI layer and isolated primary track an excellent electron suppression of  $\gtrsim 99.9\%$  is achieved.

Table A4 also shows the fitted response and resolution (see subsection 5.1) to reconstructed energy spectra of simulated pion runs for all used beam energies and physics lists. This is done both for a minimal selection (*FHI*) and the full pion selection (*sel.*). The minimal selection contains the lower and upper cuts on FHI layer to roughly preserve sampling fractions and average containment between the selections. The bias on both response and resolution is  $\lesssim 1\%$  for all examined beam energies and physics lists.

## 5 Pion Energy Reconstruction

This section describes how measured energy depositions in the different calorimeter prototypes are combined into a common reconstructed energy. This is achieved by applying a weight to each deposition according to the inverse sampling fraction of each calorimeter hit. In more complicated energy reconstruction schemes the reconstructed energy can depend on any number of measured event quantities and external parameters (*weighting*). For any chosen energy reconstruction algorithm, external parameters can be optimised by minimising the sum of quadratic distances of the reconstructed event energy to the known beam energy, resembling a  $\chi^2$  function:

$$\chi^2 = \sum_{events} \frac{(E_{rec}^{event} - E_{beam}^{event})^2}{\left(\frac{\sigma}{E}\right)^2 \cdot E_{beam}^{event}} \quad (1)$$

In this formalism parameters can be estimated for multiple runs at multiple beam energies in one optimisation. In the energy range considered here, the variance of the reconstructed energy spectrum is expected to scale with the beam energy (according to the  $1/\sqrt{E}$  stochastic term resolution dependence). To normalise contributions to  $\chi^2$  between different beam energies, each event is thus de-weighted by its known beam energy. The constant factor in the denominator is approximating the expected stochastic term of the calorimeter, using  $\frac{\sigma}{E} = 55\%$  here, for energies given in GeV. This enables the correct parameter uncertainty estimation by the optimisation algorithm, but does not influence the estimated parameter values. In order not to bias the parameter estimation towards a specific beam energy, the same number of events are added to  $\chi^2$  for each run. The first 40,000 selected events of each data run (20,000 in simulated runs) are used in the parameter optimisation. No significant difference in reconstructed energy resolution is observed whether the events used for parameter optimisation are excluded in the reconstruction or not for both energy reconstruction schemes discussed here.

### 5.1 Standard Weighting

In an idealised sampling calorimeter, the reconstructed energy for an incoming particle is directly proportional to the measured depositions. For a calorimeter system consisting of several different sampling ratios, the reconstructed energy is the sum of all hit energies weighted by a constant factor for each calorimeter. As the sampling

fraction within the ScECAL is constant and the AHCAL and first eight layers of the TCMT (depositions in the latter eight TCMT layers are disregarded in this analysis due to the limited beam energy range) have identical sampling fractions, the standard energy reconstruction only has two parameters  $w_{\text{ECAL}}$  and  $w_{\text{HCAL}}$ :

$$E_{\text{rec}}^{\text{classic}} = w_{\text{ECAL}} \cdot E_{\text{sum}}^{\text{ScECAL}} + w_{\text{HCAL}} \cdot (E_{\text{sum}}^{\text{AHCAL}} + E_{\text{sum}}^{\text{TCMT}}) \quad (2)$$

The  $\chi^2$ -optimisation described above inherently assumes Gaussian distributions. The reconstructed energy spectra obtained from the used calorimeter setup can exhibit non-Gaussian features from fluctuations in the electromagnetic fraction of pion showers, leakage effects and sample impurities remaining in data. Thus the optimisation of weights is performed iteratively on the central 90% of reconstructed energies.

Figure 7 lists the weights both optimised for each beam energy on its own and all energies at once. The weight dependence on beam energy is flat, with the smallest beam energy point at 4 GeV preferring slightly different weights. Weights optimised for all beam energies are close to the values obtained for single beam energies. All used physics lists produce very similar weights. Weights obtained from data have slightly ( $\approx 5\%$ ) higher values than weights obtained from simulations. This hints to a general overestimation of depositions in simulations, as can also be observed in the longitudinal shower profile (Figure 13 in subsection 6.1).

The reconstructed resolution and linearity only depend on the ratio of weights, which is very similar between data and simulations, as shown in Figure 7. Using weights obtained from simulation to reconstruct data events (or vice versa) would thus not notably influence resolution, but only result in a shifted energy scale.

## 5.2 Software Compensation

For most calorimeters, the measured response generated by a hadron shower is typically smaller than the response of an electromagnetic shower of the same initial particle energy ( $e/\pi > 1$ ). Within hadron showers energy can be lost to *invisible* processes like recoil, excitation and fragmentation of absorber nuclei and, depending on the active material, neutron emission. This leads to a lower response compared to purely electromagnetic showers. Hadron showers can also develop purely electromagnetic sub-showers from  $\pi^0/\eta$  production in inelastic interactions and subsequent decay to two photons, yielding the full electromagnetic response for the two photons.

Both the loss of measurable energy to *invisible* mechanisms and the creation of

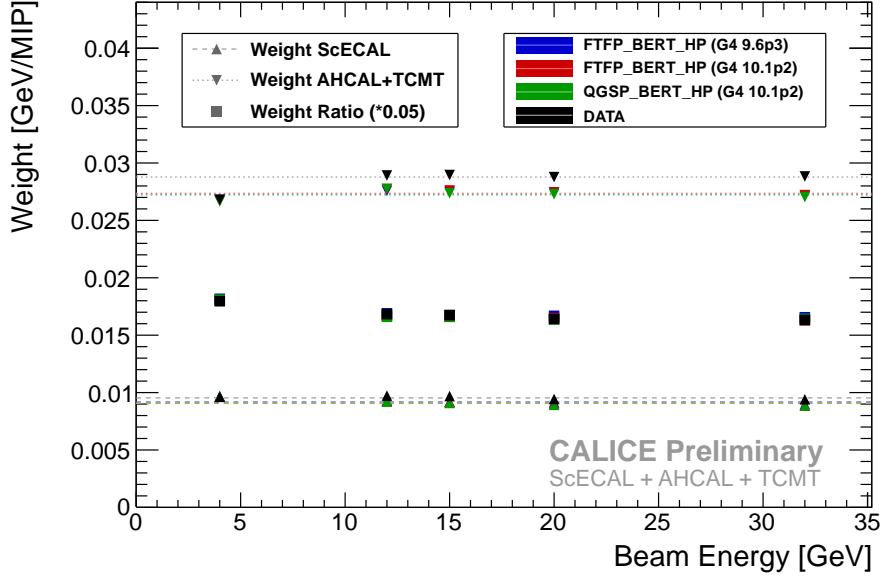


Figure 7: Standard energy reconstruction weights obtained from data and simulations. Markers indicate weights optimised for single beam energies, lines represent weights optimised for all available beam energy runs at once. All statistical uncertainties are smaller than the used markers.

$\pi^0/\eta$  are statistical processes and thus subject to significant fluctuations from event to event, leading to a large fluctuation in measured depositions and thus decreased energy resolution. Furthermore the number of generated  $\pi^0/\eta$  particles depends on the number of inelastic interactions within a hadron shower and thus scales with beam energy, impairing the linearity of the calorimeter response. If it is possible to identify electromagnetic sub-shower contributions within hadron showers, reweighting them down to the purely hadronic deposition scale should lead to an improvement in resolution and linearity of the reconstructed energy.

With the materials used in this setup, the length scales of hadron showers and electromagnetic showers are notably different ( $X_0/\text{layer} \gg \lambda_\pi/\text{layer}$  in both ScECAL and AHCAL). In combination with the high readout granularity of around the feature size of electromagnetic showers, this enables the differentiation between electromagnetic and hadronic shower components by deposited energy density. Each measured cell deposition is thus weighted as a function  $w(\rho, E_{\text{est.}})$  of the local deposition density  $\rho$  and an estimate of the full shower energy  $E_{\text{est.}}$ .

Instead of fully parametrising  $w(\rho, E_{\text{est.}})$ , the dependence in  $\rho$  is divided into fixed

bins, while the dependence on  $E_{\text{est.}}$  is parametrised over the used energy range for each such bin in this note. This scheme differs from the *local software compensation* implementation in [13], which is iteratively parametrising  $w$  in both  $\rho$  and  $E_{\text{est.}}$ , by not enforcing any functional dependence on  $\rho$ , as  $\rho$  is only ever evaluated in the chosen bins of deposition density. The scheme presented here leads to more free parameters and thus more degrees of freedom while improving the stability of the optimisation.

Instead of using the deposition density as the hit amplitude divided by the cell size, the hit energy for each hit is used directly, disregarding the differently sized tiles in the AHCAL, slightly improving the performance of the full algorithm. This analysis uses eight hit energy bins. The obtained resolutions do not critically depend on the number of bins or exact bin boundaries. For the two lowest energy bins in the ScECAL and AHCAL, instead of summing up hit energies, only the number of hits falling into these bins are counted to suppress Landau fluctuations from low particle multiplicity hits (similar to the SDHCAL reconstruction [14]), slightly improving the resolution of the algorithm. All TCMT depositions are treated as falling into the same hit energy bin, effectively parametrising the relative TCMT weight as a function of the beam energy only.

The lowest hit energy bin has significant contributions from the primary track before the first hard interaction, which show nearly no dependence on beam energy. To exclude biasing of the parameter optimisation towards weighting up the primary track hits to the full beam energy, hits on the primary track are excluded from the software compensation weighting. All hits on the axis of the reconstructed isolated primary track (as described in subsection 4.1) from the first ScECAL layer up to two layers before the reconstructed FHI layer are included into the energy reconstruction without hit energy or shower energy dependent weighting. To exclude Landau fluctuations, only the number of such hits is used and multiplied by the mean energy deposition of a MIP-like particle in a single cell of the given calorimeter section.

An example of the distribution of hits into hit energy bins is given in Figure 8. It is apparent that in the lowest hit energy bin in the ScECAL around one quarter of all contributions would come from the primary track if not identified and excluded. The contribution of primary track hits to the AHCAL hit energy spectrum is small, as around 70% of selected events start showering in the ScECAL (see Figure 5).

In this analysis, the bin weights for the  $i$ th hit energy bin  $\alpha_i, \beta_i$  as well as the TCMT weight  $\gamma$  are parametrised as a second order Chebyshev polynomial as a function of the estimated particle energy  $E_{\text{est.}}$ . The full formula to reconstruct the energy



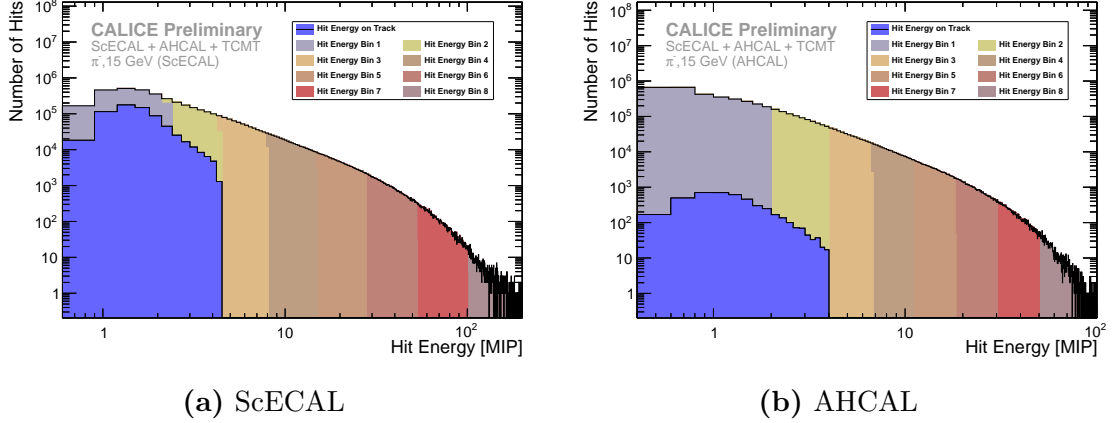


Figure 8: Hit energy spectra of data run 560496 (15 GeV  $\pi^-$ ), colours assigned to track hits and by software compensation bin.

in the combined system, with the sampling weights  $w$ , the software compensation weights  $\alpha_i, \beta_i, \gamma$ , the sum (or count) of energy depositions in the  $i$ th hit energy bin  $E_i$  and the energy depositions on the primary track  $E_{\text{track}}$ , is:

$$\begin{aligned}
 E_{\text{rec}}^{\text{SC}} = & w_{\text{ECAL}} \cdot \left( \sum_i^{\text{bins}} \alpha_i(E_{\text{est.}}) \cdot E_i^{\text{ECAL}} + E_{\text{track}}^{\text{ECAL}} \right) \\
 & + w_{\text{HCAL}} \cdot \left( \sum_i^{\text{bins}} \beta_i(E_{\text{est.}}) \cdot E_i^{\text{HCAL}} + E_{\text{track}}^{\text{HCAL}} + \gamma(E_{\text{est.}}) \cdot E_{\text{sum}}^{\text{TCMT}} \right) \quad (3)
 \end{aligned}$$

The software compensation reconstruction is defined by 51 (8 bins in the ScECAL  $\times$  3 parameters per bin,  $8 \times 3$  parameters for the AHCAL and 3 parameters for the TCMT) parameters in total. In the optimisation of the parameters the known beam energy is used for  $E_{\text{est.}}$ , while during reconstruction the standard reconstruction result is used as an estimate.

Figure 9 shows the beam energy dependence of the bin weights for ScECAL and AHCAL. The slope in the first two bins of ScECAL and AHCAL in Figure 9 (c), (d) corresponds to a  $1/E$  dependence and thus a constant contribution to the reconstructed energy of each hit in these bins, regardless of the hit energy. Assuming a shower of  $E_{\text{est.}} = 4 \text{ GeV}$ , a hit in the AHCAL with a measured hit energy  $E_{\text{hit}} = 1 \text{ MIP}$  would be weighted with a factor of around 1.5 (as given by the blue lines in Figure 9(d)) for a contribution to the reconstructed shower energy of  $1.5 \times 1 \text{ MIP} = 1.5 \text{ MIP}$ . A hit

of measured energy  $E_{\text{hit}} = 0.5 \text{ MIP}$  would be weighted with the double weight, due to the  $1/E$  dependence of the first two hit energy bins, for the identical contribution of  $3 \times 0.5 \text{ MIP} = 1.5 \text{ MIP}$  to the reconstructed shower energy. In hit energy bins  $\geq 3$ , two hits of different hit energy within the same hit energy bin would contribute to the reconstructed shower energy proportional to their hit energy.

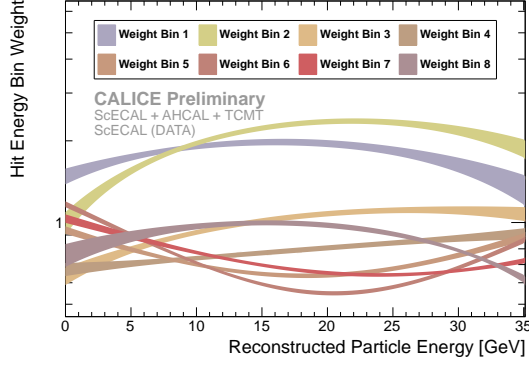
Higher energy bins tend to be weighted below unity, indicating a high energy hit to belong more likely to an electromagnetic sub-shower. Especially in the ScECAL it is clearly visible that bin weights are not monotonically decreasing for increasing hit energies, as is enforced in the *local software compensation* scheme used in [13]. However the hit energy of minimal bin weight is increasing with energy, indicating that the typical hit energy scale for electromagnetic sub-showers increases with the incident pion energy.

Applying the weights shown in Figure 9 to the dataset yields an improvement in energy resolution as shown in Figure 10. Iterations of software compensation reconstruction using the previous result as  $E_{\text{est.}}$  do not further improve the energy resolution. The correlation between standard and software compensation reconstruction shows a clear non-linearity in the central part of the reconstructed energies, causing events with a high hadronic fraction, and thus lower standard reconstructed energy, to be shifted up in software compensation reconstruction. Likewise events with above average electromagnetic shower content, and thus too high standard reconstructed energy, are shifted down when reconstructed with the software compensation reconstruction.

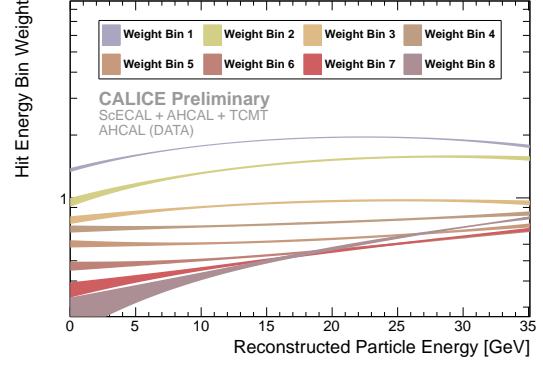
The identical procedure of reconstructing energies and optimising weight parameter is applied on simulated events. The dependence of bin weight to beam energy compared between data and simulation is shown in Figure 11 for selected bins. All used simulation physics lists agree with each other within parameter uncertainties. While the AHCAL generally shows reasonable agreement between weights derived from data and simulations in all bins, the ScECAL shows discrepancies especially for the first two bins and the highest hit energy bin. The TCMT weight also has a large discrepancy between data and simulations, although mostly for low beam energies in which TCMT depositions are expected to be minimal.

The ScECAL pion hit energy spectra show similar behaviour to the electron hit spectra discussed in subsection 3.2, as in high tails of the ScECAL hit energy spectrum being overestimated in simulation, especially for high beam energies.

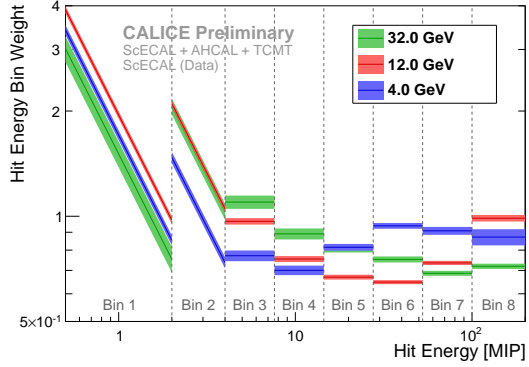
Optimising the software compensation weights on data runs reconstructed with a



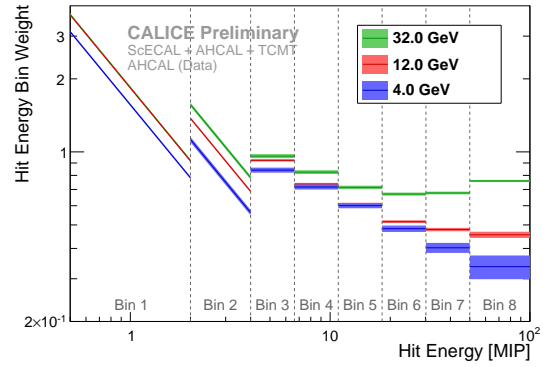
(a) ScECAL bin weights as a function of energy



(b) AHCAL bin weights as a function of energy

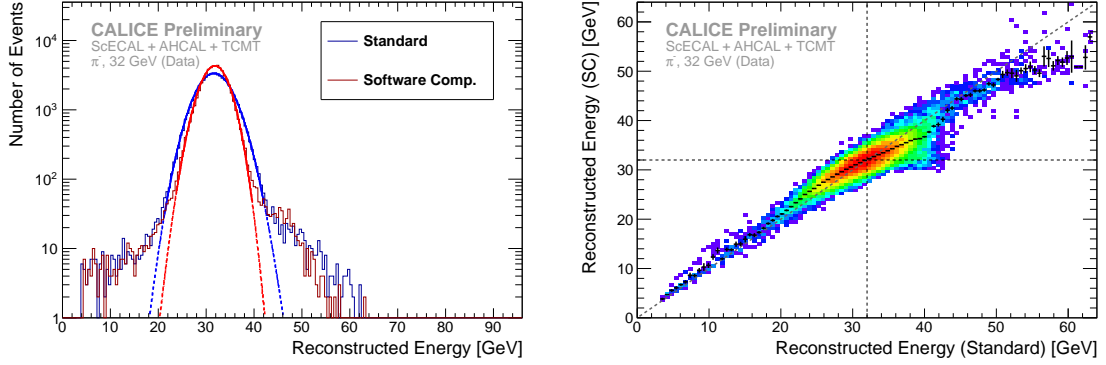


(c) ScECAL bin weights for selected energies



(d) AHCAL bin weights for selected energies

Figure 9: Hit energy bin weights as a function of beam energy, optimised from data runs. In (c) and (d), the hit energy dependent weights of the first two bins correspond to a  $1/E$  dependence and thus a counting of hits in these bins. The width of each plotted line indicates the weight uncertainty propagated from the parameter errors.



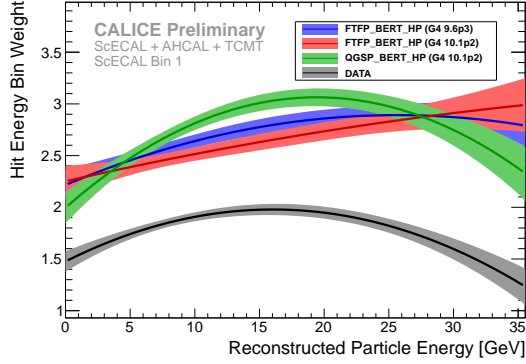
(a) Reconstructed energy spectra for standard and software compensation reconstruction. (b) Correlation standard vs. software compensation reconstructed energy.

Figure 10: Reconstructed energy in standard and software compensation reconstruction for data run 560474 ( $32 \text{ GeV} \pi^-$ ). The black markers in the correlation plot show the profile of mean software compensation reconstructed energy for bins in standard reconstruction energy. The black dashed lines in the correlation plot indicate the beam energy of the run.

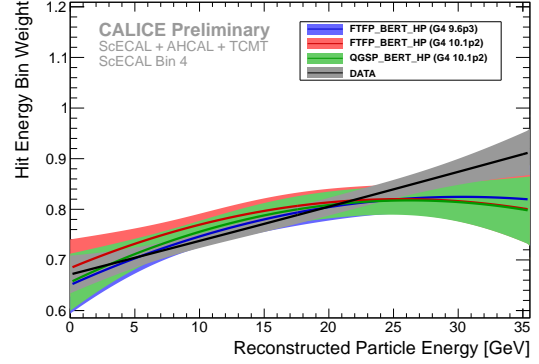
shifted saturation scale in the ScECAL (as described in subsection 3.2) shows only a very small influence on the optimised weights. Optimising weights from simulations with adjusted air gaps between ScECAL layers (also as described in subsection 3.2) yields bin weights with small but significant differences to the standard simulation.

The averaged summed deposition per event for each bin is investigated in order to understand the observed differences in the weights, as shown in Figure 12. The highest hit energy bin in the ScECAL has around twice of the mean deposition in simulation compared to data, with significant differences showing from the sixth bin. For lower energies the effect is less pronounced. In the AHCAL all bins show reasonable agreement between data and simulation, with smaller overestimations in the highest hit energy bin.

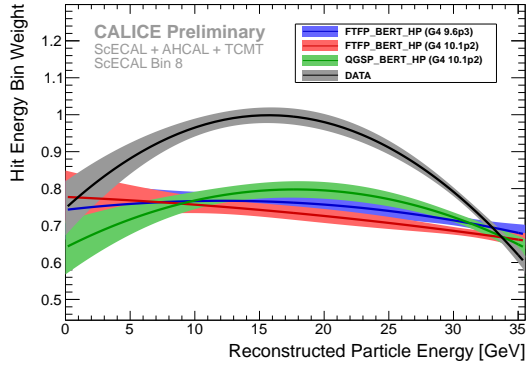
The observed differences in mean hit energy per bin (and thus hit energy spectra) do not sufficiently explain the differences in optimised software compensation weights on their own. Although the highest ScECAL hit energy bin shows discrepancies between data and simulation for all beam energies, its bin weight is only different in the central part of the beam energies (see Figure 11(c)). Likewise the mean energy sum in the first ScECAL hit energy bin is well described for all energies, but the corresponding bin weights are not (see Figure 11(a)).



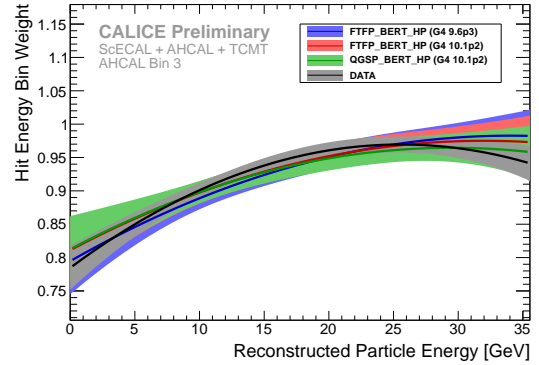
(a) ScEcal Bin 1



(b) ScEcal Bin 4

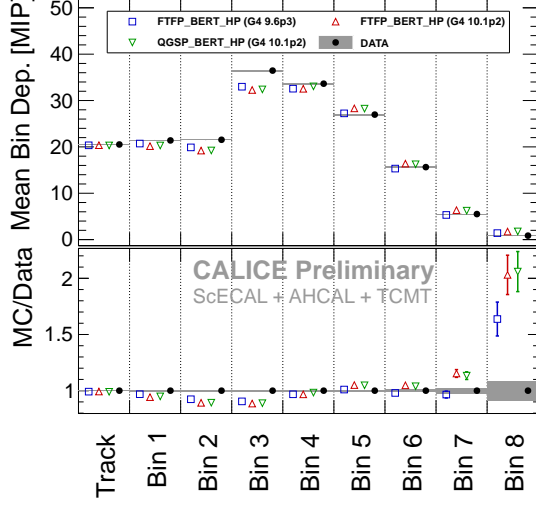


(c) ScEcal Bin 8

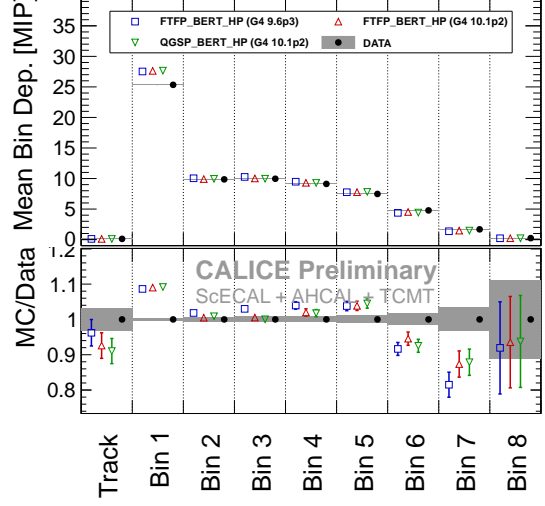


(d) AHCAL Bin 3

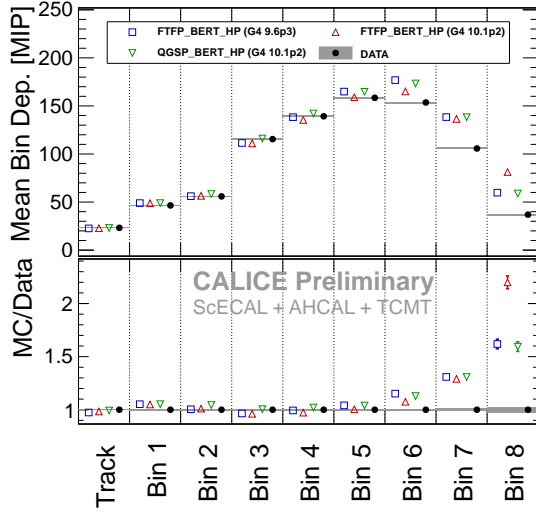
Figure 11: Hit energy bin weights as a function of beam energy for data and different simulations. The width of each plotted line indicates the weight uncertainty propagated from the parameter errors. Low and high hit energy bin weights are not well reproduced in the ScEcal.



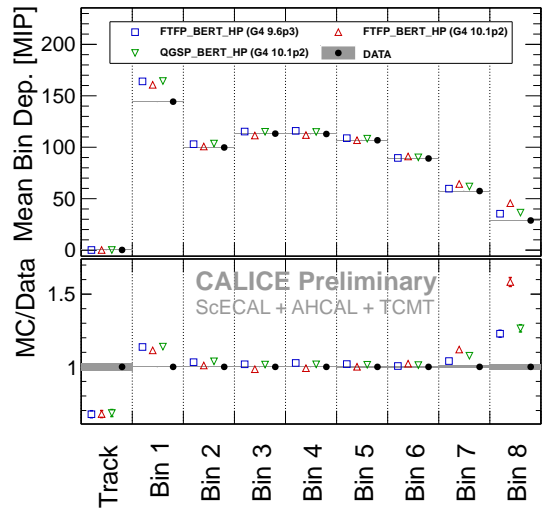
(a) ScECAL, Run 560506 (4 GeV)



(b) AHCAL, Run 560506 (4 GeV)



(c) ScECAL, Run 560474 (32 GeV)



(d) AHCAL, Run 560474 (32 GeV)

Figure 12: Averaged energy deposition sum per hit energy bin per event for data and simulations in ScECAL and AHCAL. For most entries the statistical error is smaller than the used markers.

Remaining instrumental and shower modelling effects could affect the optimised weights even in hit energy ranges where the observed hit energy spectra match reasonably between data and simulation, as bin weights are necessarily anti-correlated to conserve the mean reconstructed energy. Possible instrumental effects are covered by the systematic uncertainties given, so we consider the possibility that the modelling of the pion shower structure is imperfect, especially in high-Z absorbers and on the granularity scale that is studied here.

## 6 Results

This section describes the results obtained from using the techniques described above in comparison between data and simulations, starting with longitudinal profiles of pion showers to energy reconstruction and linearity of the system as well as energy resolution with standard and software compensation reconstruction.

### 6.1 Profiles

The mean longitudinal profile of all pion shower events passing the event selection for 32 GeV beam energy is shown in Figure 13. The low mean deposition in the first five layers is a consequence of the event selection (FHI layer  $>5$ ). The low mean deposition in the last AHCAL layers points to good average shower containment even without including the TCMT layers. The general shape of the profile is well described by all physics lists, including the dip in responses in the ScECAL around the transition region between the ScECAL and AHCAL. GEANT 4 with the FTFP\_BERT physics lists shows a peak in the ratio to data around layer 5 which is not present for QGSP\_BERT\_HP. All studied simulations overestimate the mean depositions by around 5% regardless of beam energy, as already noted in subsection 5.1. This difference is larger than the systematic uncertainty on the MIP calibration or saturation effects.

The longitudinal shower profiles from shower start are shown in Figure 14, separately for showers starting in the ScECAL and in the AHCAL. For events with reconstructed shower start in the ScECAL, the shower profiles predicted by the FTFP\_BERT and QGSP\_BERT simulation models vary significantly, under and over-estimating depositions compared to data in different layers, especially for high beam energies. Most of the data points lie between the predictions of the different physics

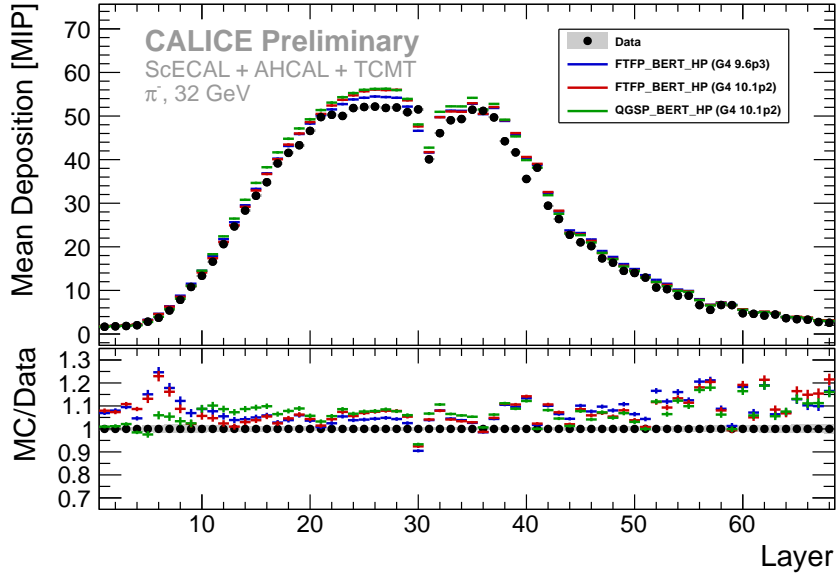


Figure 13: Averaged longitudinal shower profile for 32 GeV  $\pi^-$  events in data and different simulation physics lists. Depositions in the ScECAL are shown in layers 1–30, layers 31–68 are AHCAL layers.

lists used. FTFP\_BERT\_HP in GEANT 4 10.1p2 shows slightly longer showers than GEANT 4 9.6p3.

In events with reconstructed FHI layer in the AHCAL the examined simulation models agree well with each other. The agreement between data and simulation is on a similar level, with similar features in the ratio, as in [15]. In the AHCAL FTFP\_BERT\_HP shows slightly shorter showers in GEANT 4 10.1p2 than in GEANT 4 9.6p3, opposite to the behaviour in the ScECAL.

## 6.2 Energy reconstruction and linearity

For each event the particle energy is reconstructed with the techniques described in section 5, without use of the known beam energy. For each run the spectrum of obtained reconstructed energies is fitted with a Novosibirsk function, of which the response and resolution are calculated using Monte-Carlo integration as explained in [16].

Data and all simulation physics lists are reconstructed with the parameters optimised from their own datasets. A systematic uncertainty of 1% on the MIP scale is added in quadrature to the negligible statistical error of the response. The systematic



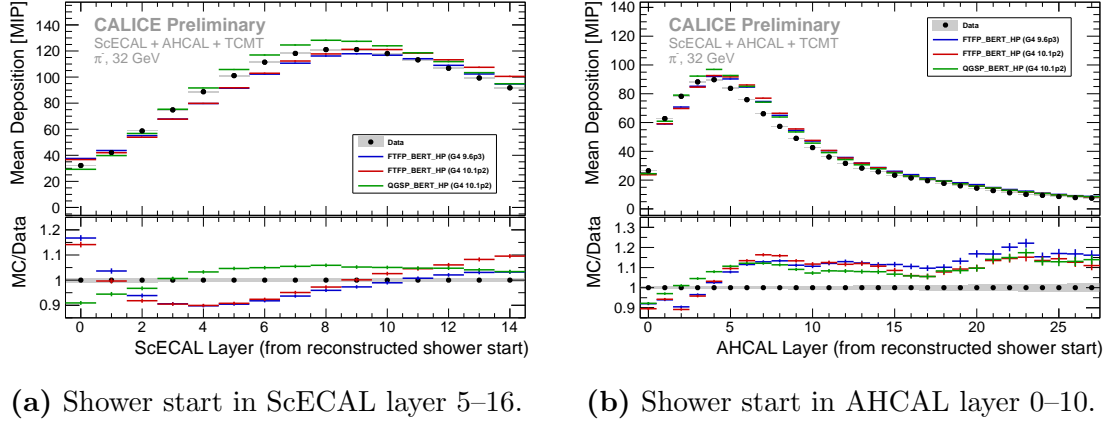


Figure 14: Averaged longitudinal shower profile for 32 GeV  $\pi^-$  events in data and different simulation physics lists. Depositions are plotted as function of distance to reconstructed FHI layer.

uncertainty from the modified air gap in simulations and the shifted saturation scales in data is negligible. The residual is defined as the relative deviation of the mean reconstructed energy from the known beam energy.

The mean reconstructed energies in data agree very well with the beam energy with all deviations better than 3%, as shown in Figure 15. The reconstructed energy observed in simulations shows a small non-linearity with no deviation from the beam energy exceeding 5%.

Reconstructing the particle energy using the software compensation scheme described in subsection 5.2, the linearity of data and simulated events is better than 3%.

### 6.3 Energy resolution

The resolution of each run is calculated from the width and mean of the Novosibirsk function fitted to the reconstructed energy spectrum, using reconstruction parameters as described in the previous section. For simulated runs, the difference in resolution between standard simulation and simulation with modified airgap is taken as fully correlated systematic uncertainty. Shifting the saturation scale has no significant effect on the reconstructed energy resolution in data and is thus omitted as a systematic uncertainty. The statistical uncertainty on each resolution point is small.

In standard energy reconstruction, the data points are generally well described

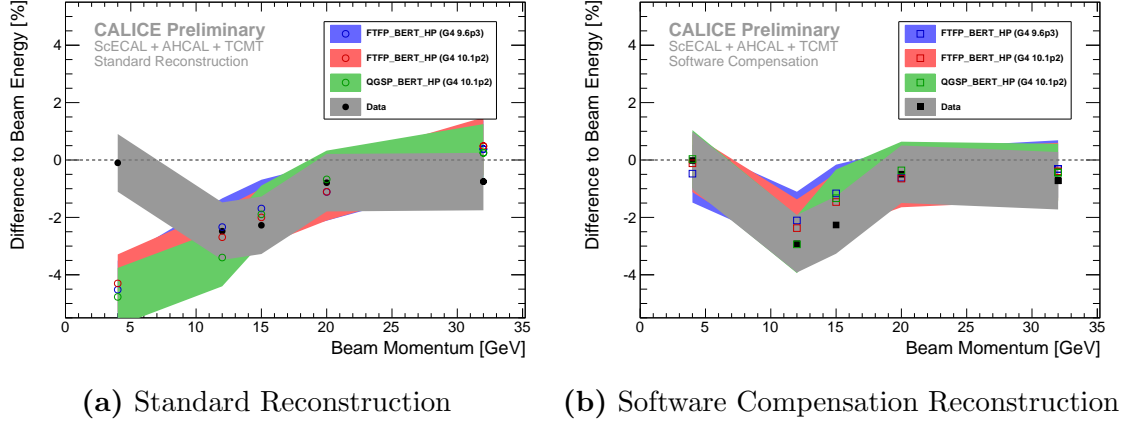


Figure 15: Residual of mean fitted reconstructed energy over beam energy in data and different simulation physics lists.

by the simulations, see Figure 16(a). QGSP\_BERT\_HP produces the best agreement with data while the FTFP\_BERT\_HP samples show deviations either in the low or high beam energy region of the sample.

Software compensation leads to a relative improvement of data point resolutions between 10–20% (2–3% absolute), depending on beam energy, see Figure 16(a). The resolution of the simulated samples agrees between physics lists when using software compensation reconstruction, but generally and significantly overestimates the achievable resolution improvements by relative 5–10% compared to data as shown in Figure 16(b). The used calorimeter system combining several readout granularities, samplings and absorber materials is reasonably described in simulations, even when complex energy reconstructions, utilising the detailed structure of individual showers, are employed.

Figure 17 compares the energy resolution obtained from data in this analysis to the energy resolutions obtained from [13], in which only pion showers with reconstructed FHI layer in the AHCAL are considered. The resolutions obtained for both analyses are in good agreement, indicating that the combined energy reconstruction with different absorber materials and sampling fractions maintains the single pion energy resolution of the calorimeter system.

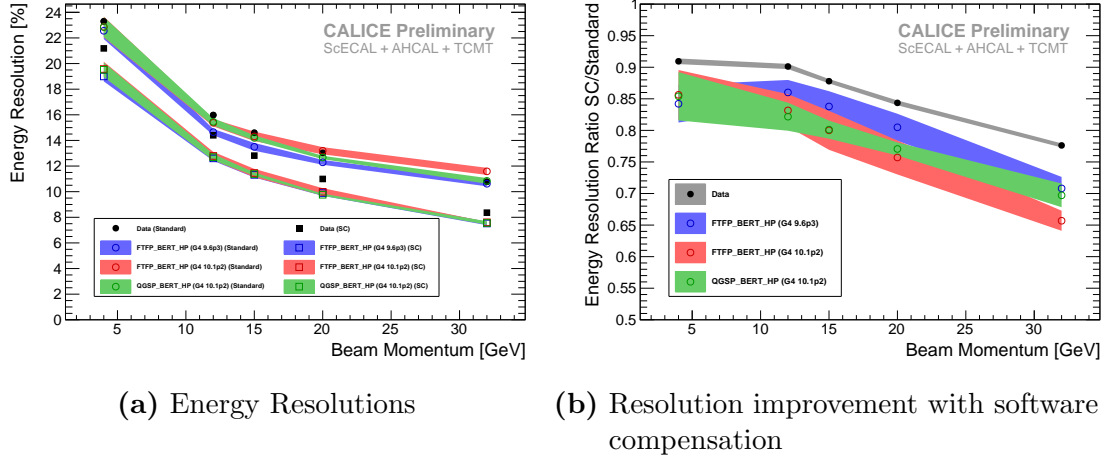


Figure 16: Reconstructed energy resolution as a function of beam energy in data and different simulation physics lists for standard and software compensation reconstruction. All plotted values are tabulated in Table A5.

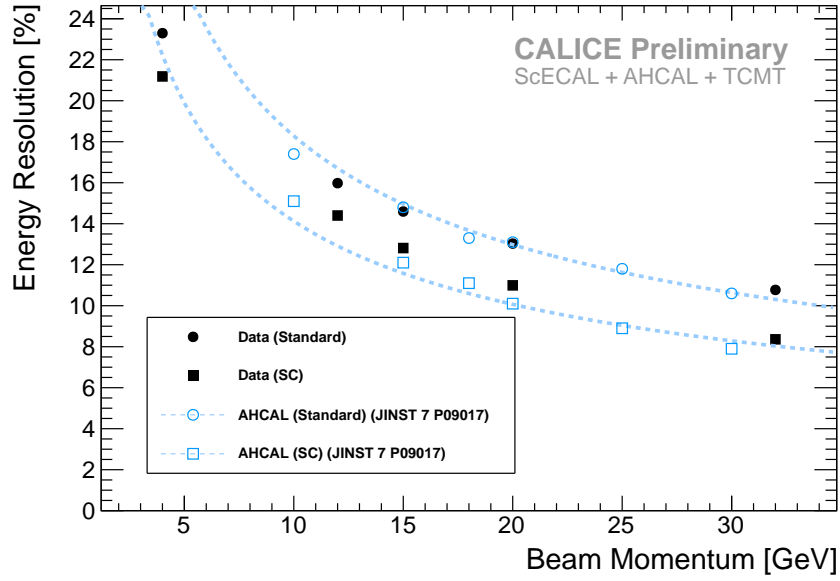


Figure 17: Single pion energy resolutions with standard and software compensation reconstruction from the combined ScECAL+AHCAL+TCMT system compared to resolutions obtained from AHCAL+TCMT in [13].

## 6.4 Applying Software Compensation Weights from Simulation to Data

To estimate the influence of the deviations in optimised software compensation weights between data and simulations, and to test whether software compensation weights obtained from simulated events can be used to reconstruct data events, Figure 18 shows the energy resolution and linearity of data runs reconstructed using weights optimised from both data and simulation. Only the simulation weights optimised from QGSP\_BERT\_HP in Geant4 10.2p1 are used for this comparison, as the difference between weights of different simulation physics lists is small. Furthermore only the software compensation specific weights  $\alpha_i, \beta_i, \gamma$  are used from simulation, while the standard reconstruction weights  $w_{\text{ECAL}}, w_{\text{HCAL}}$  are used from data to set the correct energy reconstruction scale (see subsection 5.1).

Applying software compensation weights obtained from simulations to data events actually improves the energy resolution slightly by relative 1–3%. However the achieved linearity is deteriorated, showing generally larger deviations than for weights derived from data, with fluctuations of magnitude similar as seen in the resolution of 1–4%. Although there are significant differences in the first two and last ScECAL hit energy bin weights between data and simulation, applying weights optimised from simulation onto data events does not deteriorate the energy resolution of the combined system and only very slightly worsens the deviations from linearity to still acceptable  $\leq 4\%$  in the energy range examined in this analysis.

## 7 Summary

This note presents results obtained with the combined scintillator-SiPM calorimeter system consisting of ScECAL, AHCAL and TCMT in the energy range 4–32 GeV.

The simulation model of the ScECAL is validated with electromagnetic interactions from MIP signals in single cells up to full shower profiles. All observables show good agreement between data and simulation except for hit energy spectra which show differences in the high end tails. This is likely due to a combination of underestimated shower radii in simulation and imperfect saturation correction in data. The influence on pion measurements is expected to be small. Systematic uncertainties from these effects are propagated into the pion results.

Longitudinal profiles of pion showers show good agreement between data and sim-

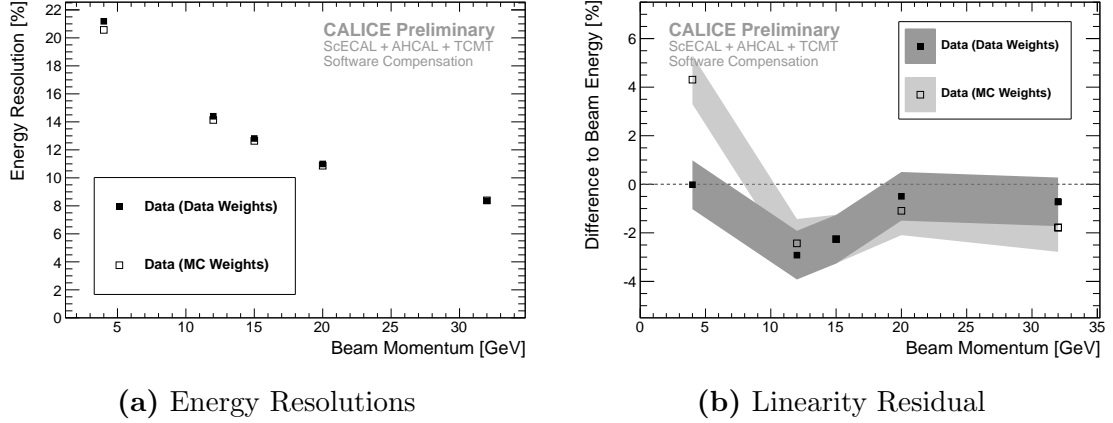


Figure 18: Reconstructed energy resolution and linearity for the data sample when using software compensation weights derived from data and simulation (QGSP\_BERT\_HP in Geant 4 10.2p1) for energy reconstruction. Reconstructing the data sample with software compensation weights derived from simulations, the energy resolution is marginally improved while the linearity is slightly deteriorated.

ulation, with a general deposition overestimation of around 5% in all tested simulation models. In the ScECAL the longitudinal shower profile as a function of distance to shower start shows considerable differences between simulation models, especially for high beam energies.

Weights for standard energy reconstruction agree well in data and simulation apart from the generally overestimated depositions. The linearity deviation of pion energy reconstruction is  $<5\%$  in simulation and  $<3\%$  in data. The energy resolution of standard reconstruction is well described by all simulation models.

The novel implementation of a local software compensation scheme does not enforce any functional dependence of weights between bins in hit energy density, while parametrising the dependence on particle energy. Counting hits instead of summing up depositions in the first two bins suppresses Landau fluctuations and slightly improves the energy resolution. The hit energy and beam energy dependent weights used for software compensation reconstruction are well described in the AHCAL. Significant discrepancies of these weights are observed in some ScECAL bins. This is potentially connected to, but not adequately explained by, the deviations in the hit energy spectra observed. The linearity deviation using software compensation reconstruction improves for data and simulation to  $<3\%$ . In data, the energy resolution improves by 10–20% from applying software compensation. Energy resolution

improvements from software compensation are overestimated in simulation by 5–10% for all beam energies and simulation models.

Both standard and software compensation energy resolutions in data are in good agreement with resolutions obtained from a similar analysis in which only showers starting in the AHCAL are considered. The single pion energy resolution of the AHCAL is thus not altered by adding the ScECAL in front.

Applying the software compensation weights obtained from simulations to data runs results in practically identical performance numbers, slightly improving the achieved energy resolution while similarly slightly deteriorating the linearity.

## References

- [1] N. Feege, Low-energetic hadron interactions in a highly granular calorimeter,  
DESY-THESIS-2011-048
- [2] CALICE Collaboration, K. Kotera, First Stage Analysis of the Energy Response and Resolution of the Scintillator ECAL in the Beam Test at FNAL, 2008  
CALICE Analysis Note CAN-016
- [3] The CALICE collaboration, C. Adloff et al., Construction and Commissioning of the CALICE Analog Hadron Calorimeter Prototype  
2010 JINST 5 P05004  
doi:10.1088/1748-0221/5/05/P05004
- [4] 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  
doi:10.1088/1748-0221/7/04/P04015
- [5] C. Günter, Comparison of Iron and Tungsten Absorber Structures for an Analog Hadron Calorimeter,  
DESY-THESIS-2015-018
- [6] CALICE Collaboration, K. Kotera, Update of the Analysis of the Test Beam Experiment of the CALICE ScECAL Physics Prototype  
CALICE Analysis Note CAN-016b
- [7] K. Kotera, Tungsten Absorber for the CALICE ScECAL Prototype  
Internal Technical Note
- [8] The CALICE Collaboration, C. Adloff et al., Electromagnetic response of a highly granular hadronic calorimeter  
2011 JINST 6 P04003  
doi:10.1088/1748-0221/6/04/P04003

- [9] K. Kotera, W. Choi, T. Takeshita, SiPM Response Functions Representing Wide Range Including Linear Behavior After Saturation  
arXiv:1510.01102 [physics.ins-det]
- [10] FTBF Website,  
<http://ftbf.fnal.gov/beamdeliverypath/>
- [11] The CALICE Collaboration, Pion and proton showers in the CALICE scintillator-steel analogue hadron calorimeter,  
2015 JINST 10 P04014  
doi:10.1088/1748-0221/10/04/P04014
- [12] The CALICE Collaboration, C. Adloff et al., Track segments in hadronic showers in a highly granular scintillator-steel hadron calorimeter,  
2013 JINST 8 P09001  
doi:10.1088/1748-0221/8/09/P09001
- [13] CALICE Collaboration, C. Adloff et al., Hadronic energy resolution of a highly granular scintillator-steel calorimeter using software compensation techniques  
2012 JINST 7 P09017  
doi:10.1088/1748-0221/7/09/P09017
- [14] CALICE Collaboration, First results of the CALICE SDHCAL technological prototype  
2016 JINST 11 P04001  
doi:10.1088/1748-0221/11/04/P04001
- [15] CALICE Collaboration, M. Chadeeva, Parametrisation of hadron shower profiles in the CALICE Sc-Fe AHCAL  
CALICE Analysis Note CAN-048
- [16] CALICE Collaboration, C. Neubüser, Analogue, Digital and Semi-Digital Energy Reconstruction in the CALICE AHCAL  
CALICE Analysis Note CAN-049



- [17] CALICE Collaboration, Pion and proton showers in the CALICE scintillator-steel AHCAL: comparison of global observables  
CALICE Analysis Note CAN-040
- [18] CALICE Collaboration, A. Khan, Temperature Correction on ADC/MIP Conversion Factors for September 2009 Data of ScECAL FNAL Test Beam Module  
CALICE Analysis Note CAN-016a

## Appendix

Table A2: Cuts applied to select single electrons. Cuts marked with <sup>(\*)</sup> are applied only on data, cuts marked with <sup>(†)</sup> are applied only on simulated events.

Name	Beam Energy	Cut
Raw	All	beamBit = 1 <sup>(*)</sup>
	All	a10x10Bit = 1 <sup>(†)</sup>
Event Quality	All	vetoBit = 0 <sup>(*)</sup>
	All	2000 < multiADC < 3800 <sup>(*)</sup>
	All	Energy in first 5 ScECAL layers > 3.5 MIP
	All	cherenkowBit = 0 <sup>(*)</sup>
	All	cherenkow2Bit = 1 <sup>(*)</sup>
Electron Selection	All	reconstructed FHI layer $\leq 4$
	All	$\frac{E_{\text{sum}}^{\text{ScECAL}}}{E_{\text{sum}}^{\text{ScECAL}} + E_{\text{sum}}^{\text{AHCAL}}} > 0.9$
	All	(nHits <sup>AHCAL</sup> < 35    nHits <sup>AHCAL</sup> > 65) &&
	All	(cog <sub>Z</sub> <sup>AHCAL</sup> < 2000 mm    cog <sub>Z</sub> <sup>AHCAL</sup> > 2250 mm)
Multi Particle Suppr.	All	nTracks in AHCAL = 0
	All	nTracks (isolated) in ScECAL = 0
Containment	All	shower CoG in central quarter of the ScECAL

Table A3: Cuts applied to select single pions. Cuts marked with <sup>(\*)</sup> are applied only on data, cuts marked with <sup>(†)</sup> are applied only on simulated events.

Name	Beam Energy	Cut
Raw	All	beamBit = 1 <sup>(*)</sup>
	All	a10x10Bit = 1 <sup>(†)</sup>
Event Quality	All	vetoBit = 0 <sup>(*)</sup>
	All	2000 <multiADC <3800 <sup>(*)</sup>
	All	Energy in first 5 ScECAL layers >3.5 MIP
	$\leq 4$ GeV	cherenkowBit = 0 <sup>(*)</sup>
	$\leq 4$ GeV	cherenkow2Bit = 0 <sup>(*)</sup>
Pion Selection	All	reconstructed FHI layer $\geq 5$
	All	(nHits <sup>AHCAL</sup> <35    nHits <sup>AHCAL</sup> >65) && (cog <sub>Z</sub> <sup>AHCAL</sup> < 2000 mm    cog <sub>Z</sub> <sup>AHCAL</sup> > 2250 mm)
Multi Particle Suppr.	All	nTracks parallel to beam in outer AHCAL = 0
	All	nTracks (isolated) in ScECAL = 1
Containment	All	reconstructed FHI layer $\leq 35$
	All	isolated track in central quarter of ScECAL

Table A4: Selection efficiencies and response/resolution biases for minimal and full pion selection for different physics list and particle types. The *FHI* sample is selected for scintillator trigger and FHI layer cuts only, *sel.* contains the full pion selection as described in 4.1. The efficiency  $\epsilon_{\text{sel.}}$  is given as the fraction of events that pass the full pion selection out of all events that hit the trigger scintillator (*raw cut*). Statistical errors on fit are negligibly small. *Raw* samples have around 100,000 events each, with only around 50,000 events at 4 GeV beam energy. All simulations are done with GEANT 4 10.1p2 except lines marked with (\*), which are GEANT 4 9.6p3.

Run/Particle	Physics List	$\mu_{\text{FHI}}/\text{GeV}$	$\sigma_{\text{FHI}}/\mu_{\text{FHI}}$	$\mu_{\text{sel.}}/\text{GeV}$	$\sigma_{\text{sel.}}/\mu_{\text{sel.}}$	$\epsilon_{\text{sel.}}$
560506 (4 GeV)	$\pi^-$ FTFP_BERT_HP	3.82	22.92%	3.83	22.86%	46.1%
	$\pi^-$ FTFP_BERT_HP(*)	3.81	22.64%	3.82	22.59%	46.5%
	$\pi^-$ QGSP_BERT_HP	3.80	22.95%	3.81	22.87%	45.9%
	$e^-$ QGSP_BERT	-	-	-	-	0.2%
560498 (12 GeV)	$\pi^-$ FTFP_BERT_HP	11.62	15.49%	11.67	15.37%	49.5%
	$\pi^-$ FTFP_BERT_HP(*)	11.68	14.73%	11.71	14.65%	49.6%
	$\pi^-$ QGSP_BERT_HP	11.54	15.54%	11.59	15.45%	49.4%
	$e^-$ QGSP_BERT	-	-	-	-	<0.1%
560496 (15 GeV)	$\pi^-$ FTFP_BERT_HP	14.64	14.43%	14.70	14.30%	48.8%
	$\pi^-$ FTFP_BERT_HP(*)	14.69	13.59%	14.74	13.49%	49.1%
	$\pi^-$ QGSP_BERT_HP	14.66	14.30%	14.71	14.20%	49.2%
	$e^-$ QGSP_BERT	-	-	-	-	<0.1%
560481 (20 GeV)	$\pi^-$ FTFP_BERT_HP	19.69	13.25%	19.78	13.19%	47.5%
	$\pi^-$ FTFP_BERT_HP(*)	19.70	12.35%	19.77	12.28%	47.6%
	$\pi^-$ QGSP_BERT_HP	19.78	12.74%	19.86	12.66%	48.0%
	$e^-$ QGSP_BERT	-	-	-	-	<0.1%
560474 (32 GeV)	$\pi^-$ FTFP_BERT_HP	31.97	11.59%	32.15	11.58%	44.3%
	$\pi^-$ FTFP_BERT_HP(*)	31.97	10.66%	32.12	10.60%	44.3%
	$\pi^-$ QGSP_BERT_HP	31.92	10.96%	32.08	10.82%	45.0%
	$e^-$ QGSP_BERT	-	-	-	-	<0.1%

Table A5: Energy resolutions extracted from pion runs in data and simulations, using the standard energy reconstruction (Std.) and the software compensation reconstruction (SC) as plotted in Figure 16. Errors on data are purely statistical, errors on simulation include systematic modelling uncertainties. All simulations are generated with GEANT 4 10.1p2 except lines marked with (\*), which are GEANT 4 9.6p3.

Run/Particle		Type	$(\sigma/\mu)_{\text{Std.}} [\%]$	$(\sigma/\mu)_{\text{SC}} [\%]$	$\frac{(\sigma/\mu)_{\text{SC}}}{(\sigma/\mu)_{\text{Std.}}}$
560506 (4 GeV)	$\pi^-$	Data	$23.30 \pm 0.08$	$21.19 \pm 0.07$	$0.91 \pm 0.00$
	$\pi^-$	FTFP_BERT_HP	$22.56 \pm 0.65$	$19.00 \pm 0.39$	$0.84 \pm 0.03$
	$\pi^-$	FTFP_BERT_HP(*)	$22.86 \pm 0.81$	$19.57 \pm 0.57$	$0.86 \pm 0.04$
	$\pi^-$	QGSP_BERT_HP	$22.85 \pm 0.78$	$19.51 \pm 0.56$	$0.85 \pm 0.04$
560498 (12 GeV)	$\pi^-$	Data	$15.98 \pm 0.05$	$14.40 \pm 0.05$	$0.90 \pm 0.00$
	$\pi^-$	FTFP_BERT_HP	$14.65 \pm 0.25$	$12.60 \pm 0.20$	$0.86 \pm 0.02$
	$\pi^-$	FTFP_BERT_HP(*)	$15.38 \pm 0.28$	$12.79 \pm 0.32$	$0.83 \pm 0.03$
	$\pi^-$	QGSP_BERT_HP	$15.45 \pm 0.29$	$12.69 \pm 0.25$	$0.82 \pm 0.02$
560496 (15 GeV)	$\pi^-$	Data	$14.59 \pm 0.04$	$12.81 \pm 0.04$	$0.88 \pm 0.00$
	$\pi^-$	FTFP_BERT_HP	$13.50 \pm 0.30$	$11.31 \pm 0.20$	$0.84 \pm 0.02$
	$\pi^-$	FTFP_BERT_HP(*)	$14.32 \pm 0.37$	$11.45 \pm 0.33$	$0.80 \pm 0.03$
	$\pi^-$	QGSP_BERT_HP	$14.20 \pm 0.20$	$11.38 \pm 0.13$	$0.80 \pm 0.01$
560481 (20 GeV)	$\pi^-$	Data	$13.03 \pm 0.04$	$10.99 \pm 0.04$	$0.84 \pm 0.00$
	$\pi^-$	FTFP_BERT_HP	$12.29 \pm 0.19$	$9.89 \pm 0.22$	$0.80 \pm 0.02$
	$\pi^-$	FTFP_BERT_HP(*)	$13.19 \pm 0.28$	$9.99 \pm 0.28$	$0.76 \pm 0.03$
	$\pi^-$	QGSP_BERT_HP	$12.67 \pm 0.13$	$9.76 \pm 0.08$	$0.77 \pm 0.01$
560474 (32 GeV)	$\pi^-$	Data	$10.77 \pm 0.03$	$8.36 \pm 0.02$	$0.78 \pm 0.00$
	$\pi^-$	FTFP_BERT_HP	$10.62 \pm 0.20$	$7.52 \pm 0.13$	$0.71 \pm 0.02$
	$\pi^-$	FTFP_BERT_HP(*)	$11.59 \pm 0.28$	$7.61 \pm 0.03$	$0.66 \pm 0.02$
	$\pi^-$	QGSP_BERT_HP	$10.84 \pm 0.22$	$7.56 \pm 0.13$	$0.70 \pm 0.02$

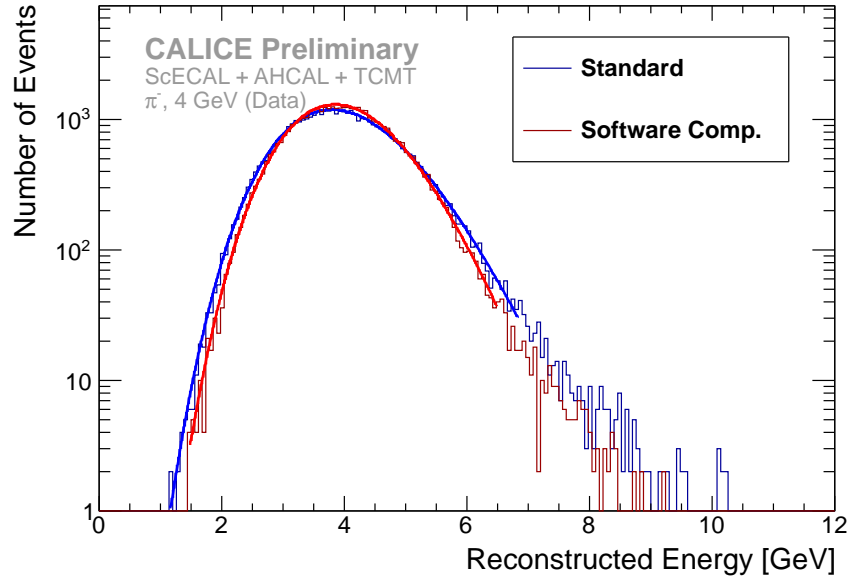


Figure 19: Reconstructed energy spectrum in standard and software compensation reconstruction for data run 560506 ( $4 \text{ GeV } \pi^-$ ).

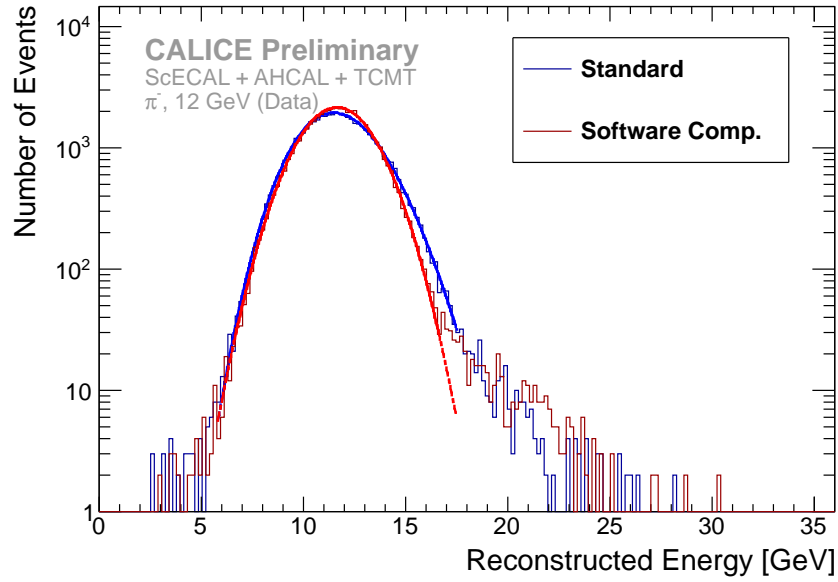


Figure 20: Reconstructed energy spectrum in standard and software compensation reconstruction for data run 560498 ( $12 \text{ GeV } \pi^-$ ).

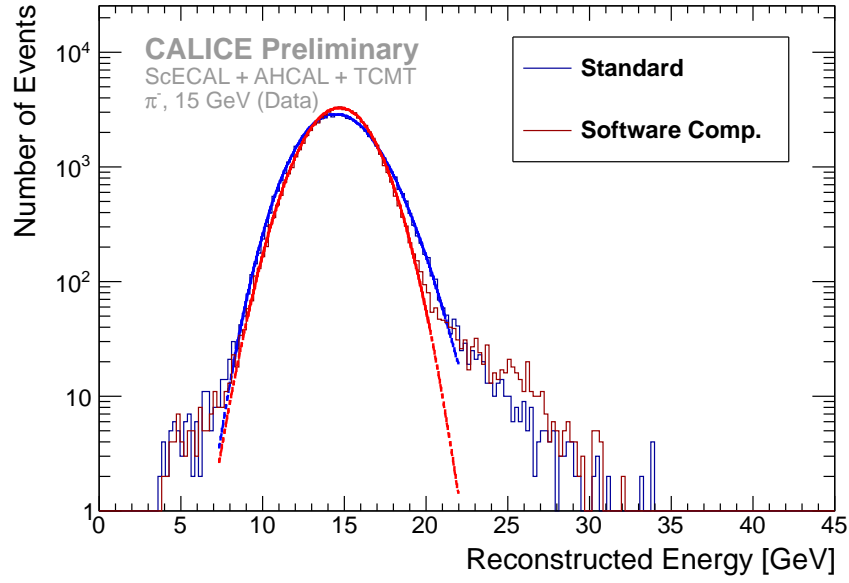


Figure 21: Reconstructed energy spectrum in standard and software compensation reconstruction for data run 560496 (15 GeV  $\pi^-$ ).

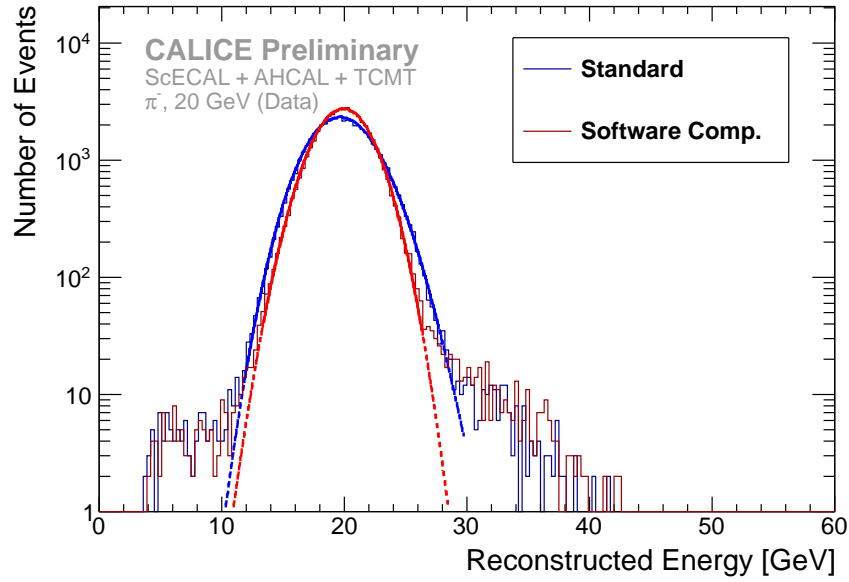


Figure 22: Reconstructed energy spectrum in standard and software compensation reconstruction for data run 560481 (20 GeV  $\pi^-$ ).

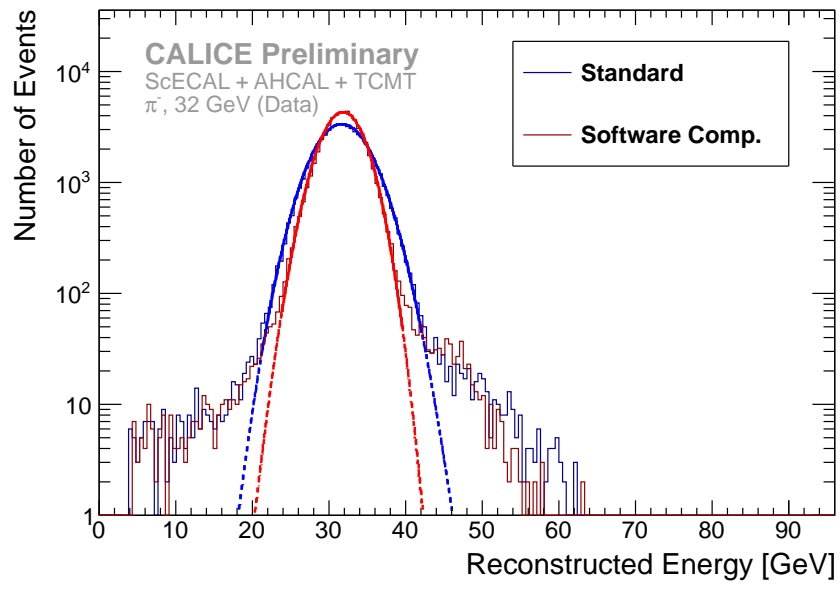


Figure 23: Reconstructed energy spectrum in standard and software compensation reconstruction for data run 560474 (32 GeV  $\pi^-$ ).