CALICE Analysis Note CAN-011
June 3, 2008

Preliminary results from hadron shower data with the CALICE tile AHCAL prototype

E. Garutti, B. Lutz, J. Samson, O. Wendt
for the CALICE collaboration

1 Introduction

This note presents preliminary CALICE results obtained analyzing the data collected during the test-beam running periods in 2006 and 2007 at CERN. The response of the Analog Hadronic Calorimeter (AHCAL) to hadronic showers is investigated and compared to various Monte Carlo models. Studies on the effect of the shower leakage on the reconstructed energy and energy resolution have been performed. First results are presented, which are obtained by applying a naïve particle flow approach to reconstruct clusters in real data.

This note summarizes the developments in the analysis of hadronic showers in the AHCAL and TCMT. Three main topics are covered: the study of the hadronic response of the calorimeter and comparison to MC models, shower leakage studies and two-shower separation studies.

2 The CALICE detector installation at the CERN test beam

In the analysis presented in this note, data from both CALICE installations at CERN during 2006 and 2007 is used. More details on the detector configuration in each installation period can be found in the detector papers (currently being written).

The ECAL

In 2006 the electromagnetic calorimeter (ECAL) was equipped with 30 sensitive layers of silicon pads, but each layer was missing one third of the readout channels, thus reducing the active area to 18×12 cm^2. The total number of instrumented readout channels was 6480 out of the 9720 foreseen. For the 2007 installation the bottom layers were also equipped, apart from the front-most 4 layers.

For the purpose of this analysis the ECAL is used as a track finder to select pion events which have not showered before reaching the AHCAL. This requirement reduces the data sample to about 40%, since the remaining 60% of events interact in the 1 λ_0 of the ECAL. The ECAL is made out of three sampling structures of different thicknesses of Tungsten. The sampling factors for muons of the three structures are fixed in the following analysis to, 33.0, 62.3 and 93.0.

Figure 1: *Picture of one AHCAL active layer (left) and of AHCAL the steel stack support with active layers installed (right).*

The AHCAL

The layout of scintillator tiles in one AHCAL active module is shown in the left picture in Fig. 1. The right picture shows the active modules mounted in the AHCAL sandwich structure made from 2 cm thick steel plates. The scintillator thickness is 0.5 cm.

In September 2006, 23 of the 38 modules had been assembled. To cover a sufficient fraction of the longitudinal shower development, without sacrificing too much of the highly granular information, it was decided to use two different samplings. The first seventeen layers, in which ≈75% of the shower energy is deposited, are fully equipped. The following thirteen layers are equipped with the remaining 6 modules, one every other layer. This configuration covers the part with statistically the highest activity whilst still giving a $\sim$3.5 λ_0 deep caloremeter. The sampling factor, $SF = E^{\mu}_{dep}/E^{\mu}_{vis}$, for one AHCAL layer as obtained from MC is 25.0. The second part of the calorimeter, equipped with coarser longitudinal readout, scaling by the difference in material composition the sampling factor for this part is 45.9. In 2007 all 38 modules of the AHCAL were installed and operated, covering the whole depth of $\sim$4.5 λ_0 (at zero degree rotation angle).

The TCMT

The tail catcher and muon tracker (TCMT) was completely equipped during the CERN data taking. It has 16 layers adding up to $\sim$5.5 λ_0.

The sampling factor of the first layer of the TCMT is taken to be equal to the second part of the AHCAL, 45.9. Layers 2-8 are equivalent to a AHCAL layer and have a sampling factor of 25.0. The last layers, from 9-16, have an absorber thickness a factor 5 larger (10 cm) and a sampling factor is rescaled by the material composition to 129.0.

The beam line instrumentation, i.e. drift chambers and Cerenkov counter, was not used in this analysis.

3 Hadronic shower response in the AHCAL

The analysis of single pion showers offers the unique possibility to validate MC hadronic shower models using a large number of well defined observables. Total energy and longitudinal profiles with very high resolution are studied in detail and compared with the available GEANT4 models.

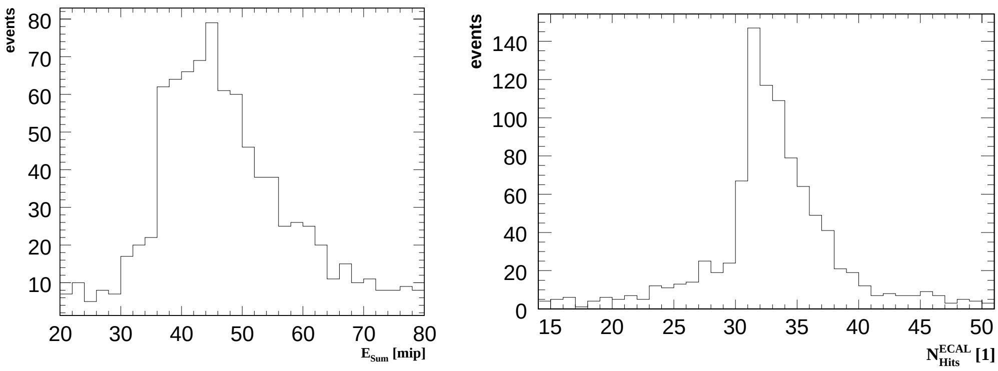

Figure 2: *Energy distribution (a) and hit distribution (b) of muons in the ECAL.*

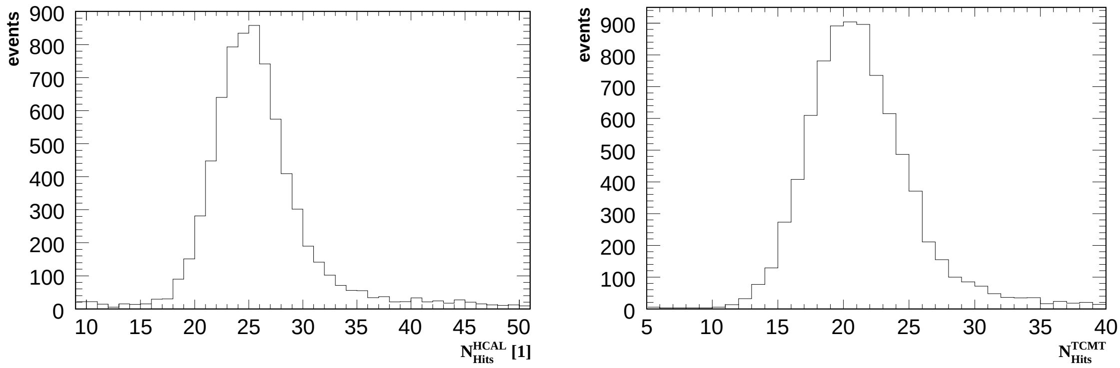

Figure 3: *Hit distribution of muons in the AHCAL (a) and TCMT (b).*

3.1 Data sample

This analysis is based on October 2006 pion runs. Beam energies from 6 GeV to 80 GeV have been analyzed. The range 6-20 GeV is a negative pion sample; the range 30-80 GeV is a positive pion sample.

3.2 Event selection

In order to optimize the AHCAL description of hadronic shower profiles, we reject those events containing muons, double tracks or early/late showers.
The selection criteria used was a beam trigger (10x10 cm coincidence trigger) and no trigger from the muon veto (100x100 cm^2 trigger panel in the back of the TCMT). Events were selected in which the pion does not start to shower in the ECAL. For this, a restriction in the number of hits is imposed to be $20 < N_{hits} < 42$; and a cut on the energy deposited in the ECAL is imposed, 25 mip $< E_{ECAL}^{tot} <$ 70 MIP. Additionally a loose cut on the number of hits in the AHCAL is applied (NhitsHCAL > 8) to exclude nearly empty events in the AHCAL.
These cuts are based on the distributions of the muon deposited energy and number of hits in each calorimeter. The distributions used to set the cuts are presented in Fig. 2a-b and Fig. 3a-b.

A track finder for the AHCAL is applied. The algorithm searches for consecutive hits in z, layer-wise, by a maximal allowed deviation in x and y. If such a sequence is found it is called a track if it consists of more than a certain number of hits. The parameters used for the track finder in this analysis are:

- Radius of track = 55 mm, to account for tracks which are not perfectly centered in one tile or which traverse the AHCAL not perpendicularly.
- Minimum number of hits in a track = 17.
- If 1 or 2 tracks are found the event is a muon candidate. In this case if $N_{hits}^{HCAL} < 43$ and $N_{hits}^{TCMT} < 30$ the Track Finder has found a muon. These events are discarded from the analysis.
- Additionally, to reject very late started showers events are rejected if more than 6 hits are found in the last 4 layers of the TCMT.

3.3 Monte Carlo models and digitization

A selection of the available MC models in GEANT4 (G4) was made according to the studies performed by the GEANT4 group (CERN-LCGAPP-2007-02). The following models have been studied, and the most discrepant ones among them selected for the comparison to data:

LHEP Parameterization of existing data from GHEISHA. List made of two sets of parameterizations. LEP: Low Energy Parameterization for E<55 GeV; HEP: High Energy Parameterization for E>25 GeV. In the common energy region one of the two lists is randomly picked for each interaction.

QGSP Quark-Gluon String model for the primary projectile-nucleon collision plus the Pre-compound model for de-excitation of the nucleus. QGS is used for E>12 GeV for protons, neutrons, pions, kaons and for E>20 GeV for other particles. Outside of this specified energy range LHEP is used.

QGSP_BERT The Bertini cascade model is added to the QGSP model in the range E<10 GeV for Nucleons, pions, kaons and hyperons. This model also includes de-excitation of the remnant nucleus, Fermi breakup and fission, as well as its own routines for nuclear evaporation for neutrons and alpha particles and final gamma emission.

QGSP_BIC The Binary cascade model is added to the QGSP model. This model is initially valid in the range E<3 GeV for protons, neutrons; E<1.5 GeV for pions; and E<3 GeV/A for light ions. In this simulation it is extended to E<10 GeV for all these particles. After the Binary cascade the Pre-compound model is used for de-excitation of the remnant nucleus.

QGSP_BERT_HP The additional High Precision package for neutron transport is used in the energy range E<100 MeV.

QGSC Quark-Gluon String model for the primary projectile-nucleon collision plus Chiral Invariant Phase Space model for de-excitation of the nucleus. QGS is used in the same energy intervals as for QGSP.

For each model a sample of 500000 events was generated for each beam energy. The generated events were digitized using the official CALICE digitization package.In the

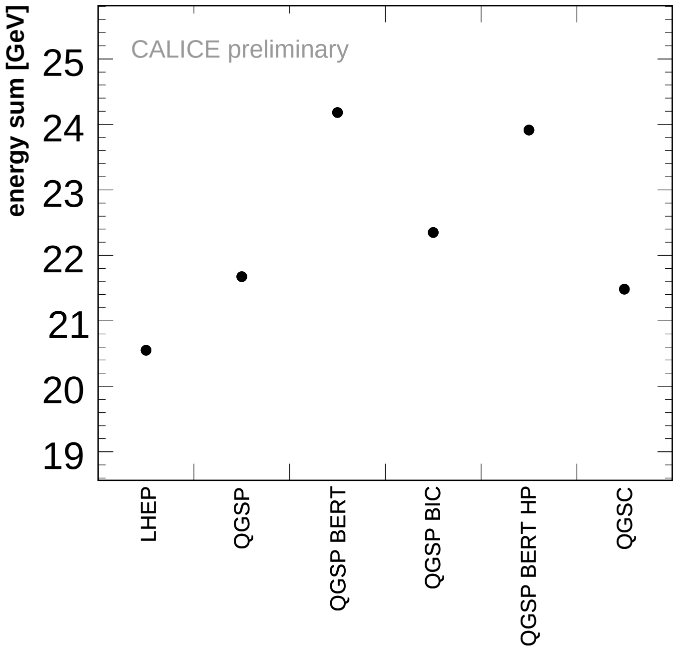

Figure 4: *Energy deposited in the AHCAL by a 20 GeV pion shower according to various MC models in G4. LHEP and QGSP_BERT show the largest discrepancy between models and are therefore taken to compare to data at first. No digitization or threshold cuts have been applied.*

digitization step, dead cells in the detector are removed and noise is add to the MC events according to the data noise distribution, simulating therefore hot cells where they are located in reality. The AHCAL digitization is discussed in more details in [1]. The digitized MC was reconstructed with the same chain as for the data.

Fig. 4 is a compilation of the total energy deposited in the AHCAL for a 20 GeV pion shower for the various MC models. No digitization and no threshold cut on the amplitude in each calorimeter cell is applied for this plot.

For the remainder of the analyses presented, only 10000 events from each of only two hadronic models has been fully digitized and compared to data. The LHEP and QGSP_BERT models have been chosen, for being the two most discrepant models in terms of total energy deposition.

3.3.1 Energy response and resolution

The data was reconstructed according to the official calibration chain already documented in [2] and the visible energy per cell in MIP units is saved for analysis. Identical cuts are applied to data and digitized MC. If not explicitly stated otherwise all errors reported in the following plots are statistical only.

The total energy of a pion shower is obtained as the sum of energy deposited in all three calorimeters multiplied by the appropriate sampling factor (see section 2). Only cells with energy above a threshold of 0.4 MIP are considered. The distributions of the total energy obtained in this way are displayed in Fig. 5. Each distribution is fit with a Gaussian function in a range of $\pm$ 2 σ; the response and resolution at each energy are the mean and sigma values from the fit. The calorimeter response as a function of energy is

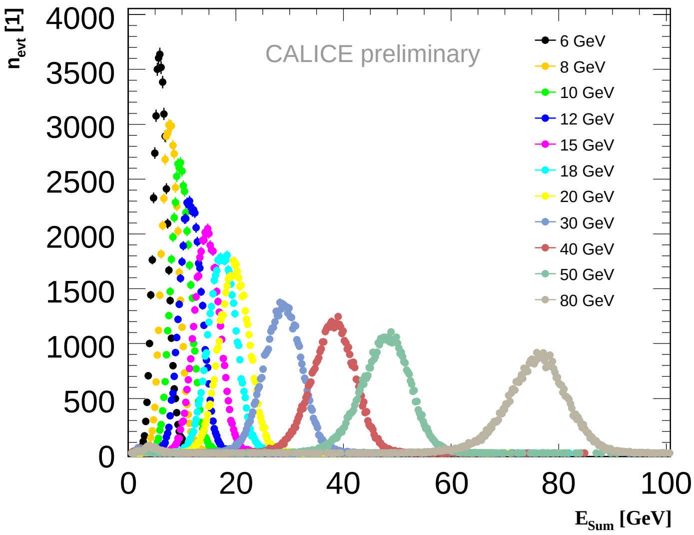

Figure 5: *Energy distribution of pion showers in the AHCAL and TCMT combined for energies between 6 GeV and 80 GeV.*

shown in Fig. 6a, for data compared to the two selected MC models. To the statistical errors a 3% systematic error on the MIP calibration, and a 0.5% beam energy uncertainty are added in quadrature. An additional source of systematic comes from the Gaussian fit to the energy distributions. To determine this uncertainty the fit range was varied between 2.5 σ and 1.5 σ, the resulting error of ??% is added in quadrature. Fig. 6b shows the data residual to a linear response fit. The fit is performed in the range 6-20 GeV, where the response is expected to be linear, and extrapolated to higher energies. At most a 2.5% negative deviation from linearity is observed at 80 GeV. Fig. 7 shows the same residual for the two MC models, LHEP and QGSP_BERT. Also, the MC points are fit in the range 6-20 GeV. Both MC indicate a positive deviation from linearity at high energy, between 3% and 6% depending on the model used. This effect could come from the non-compensating nature of the AHCAL and the increasing values of the e/π ratio at high energies. The fact that data does not reproduce this effect may indicate a remaining non-linearity in the data. This point will have to be checked further.
The results of the three linear fits to data and MC are reported in table 1.

	Slope	**Offset [GeV]**
Data	0.93±0.01	0.06±0.04
LHEP	0.97±0.01	0.14±0.06
QGSP_BERT	1.17±0.01	0.11±0.05

Table 1: Result of the linear fit to data and MC response. The fit is performed in the energy range 6-20 GeV.

The effect of the remaining non-linearity in the data after saturation correction is better seen in the hit energy spectrum per calorimeter cell in Fig. 8. In this plot the spectra of data and MC are renormalized, such as to remove the absolute difference in the energy scale. The trend at high energy indicates a small remaining deviation of data from MC prediction, which can be associated to a non-perfect correction of the non-linearity effect.

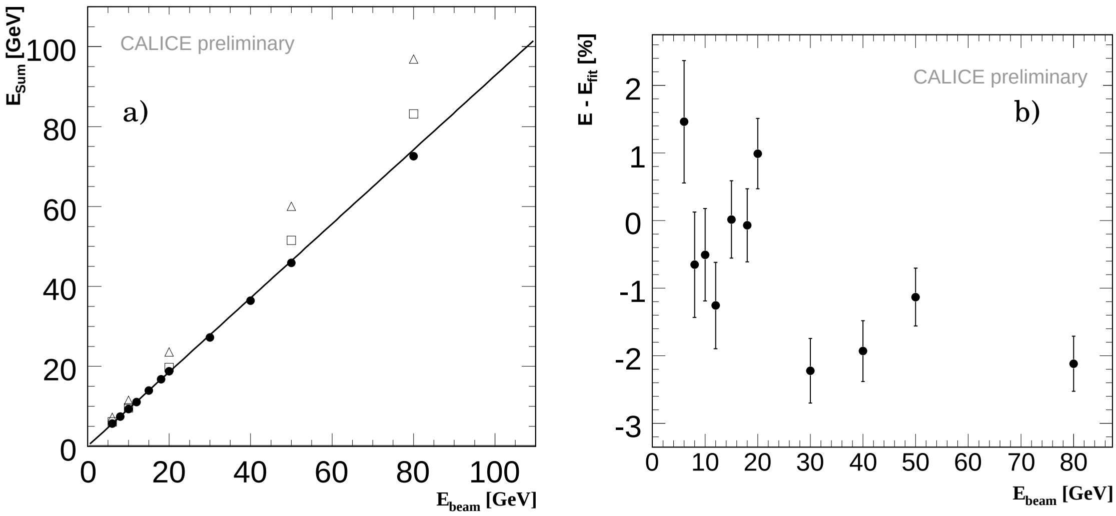

Figure 6: *a) Linearity of the AHCAL and TCMT combined response to pions. Black circles are data, open squares digitized LHEP and open triangles are digitized QGSP_BERT. b) Data residual from a linear fit in the energy range 6-20 GeV.*

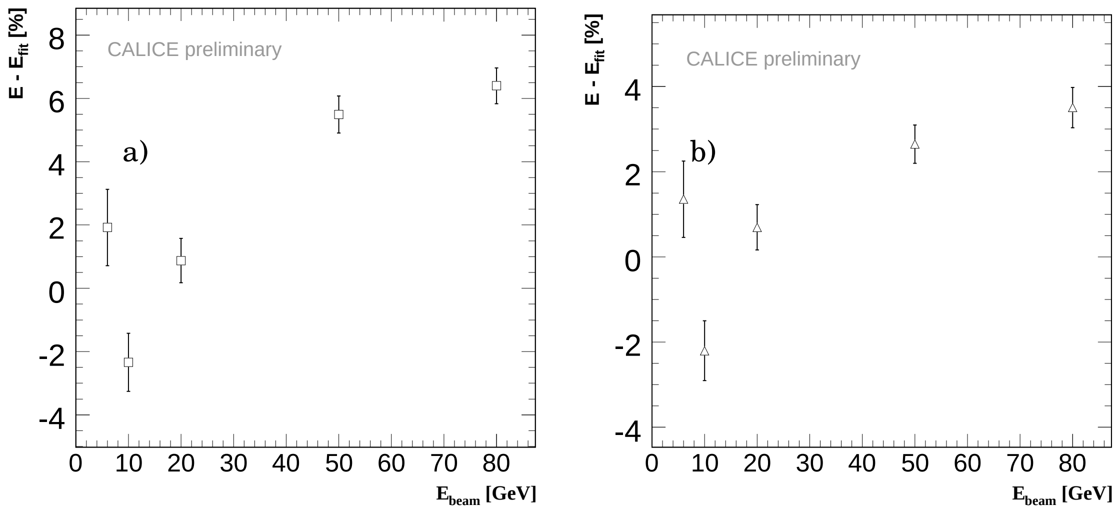

Figure 7: *Residual from a linear fit to the total energy in the range 6-20 GeV for LHEP (a) and QGSP_BERT (b).*

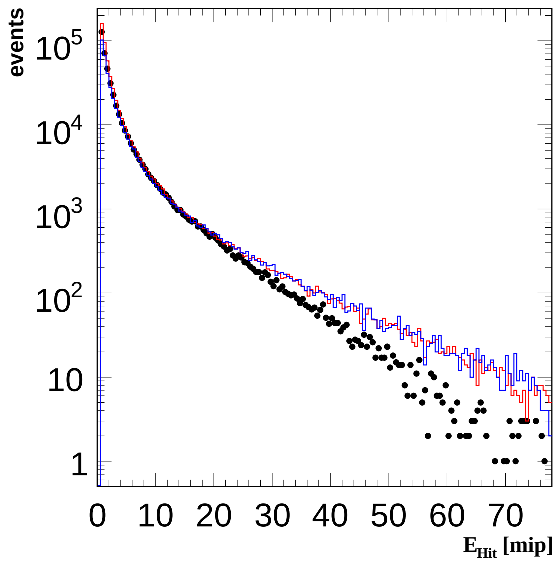

Figure 8: *Hit energy spectrum for 20 GeV pion showers. The black points are data, the red line QGSP_BERT and the blue line LHEP.*

PLEASE NOTE: in the current analysis the latest correction of SIPM saturation, as developed for the electromagnetic analysis, is still not applied. These plots may become available to the Editorial Board in a week time.
The energy resolution obtained with the combined AHCAL and TCMT is presented in Fig. 9 for data and MC. The data lies in between the two models chosen for comparison, illustrating the large variations in the MC predictions. This plot is, at present, only a qualitative comparison.

In order to convert the energy from visible to total the same MIP to GeV factor and sampling factors are taken for data and MC. In addition a mip/π = 1.42 is applied. For the MIP to GeV conversion the factors are obtained from MC where a MIP deposits 175 keV in one ECAL layer and 861 keV in one AHCAL or TCMT layer.

3.4 Shower profiles

The longitudinal shower profiles of a 20 GeV pion shower is displayed in Fig. 10 and compared to two MC models, LHEP and QGSP_BERT.

The shower maximum is determined with a fit to this distribution using the parameterization

$$\frac{dE}{dt} = kt^{a-1} \cdot e^{bt}, \tag{1}$$

were t is the depth in the calorimeter expressed in units of λ_0. The last two layers of the AHCAL are not included in the fit since they are known for being noisy, unreliable layers, which were replaced with new ones in the 2007 installation. The energy dependence of the fit parameters (t_{max} and b) is presented in Fig. 11. The expected logarithmic increase of the shower maximum with energy is observed both in data and in the two MC models. LHEP indicates a somewhat steeper increase of t_{max} with energy than data

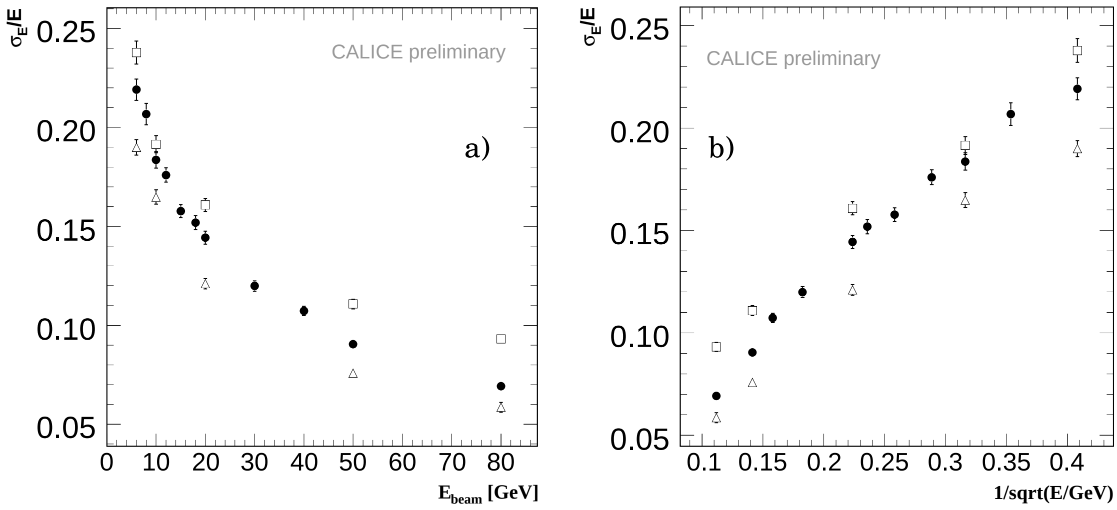

Figure 9: *Energy resolution for pions in the combined AHCAL and TCMT detector as a function of the beam energy, E (a) and of* $1/\sqrt{E}$ *(b). Black circles are data, open squares digitized LHEP and open triangles are digitized QGSP_BERT. The errors are statistical only.*

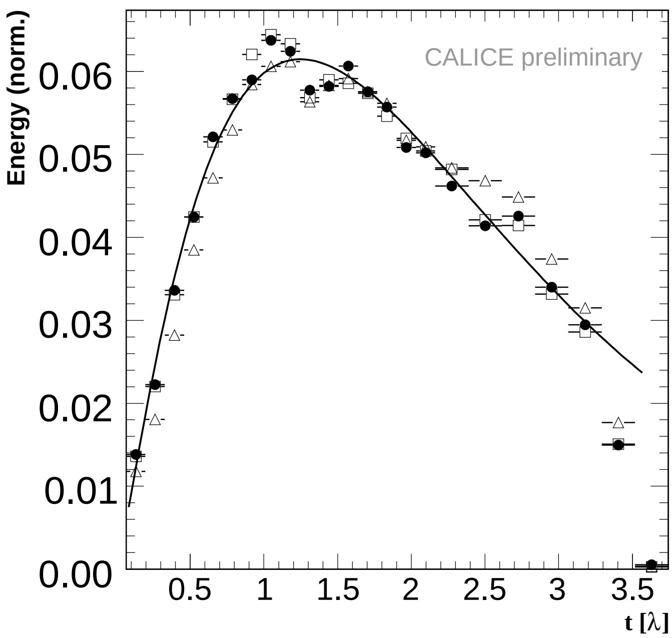

Figure 10: *Longitudinal shower profile of 20 GeV pions. Black circles are data, open squares digitized LHEP and open triangles are digitized QGSP_BERT. The curve is a fit to the data points, excluding the last two layers which are known to have low efficiency.*

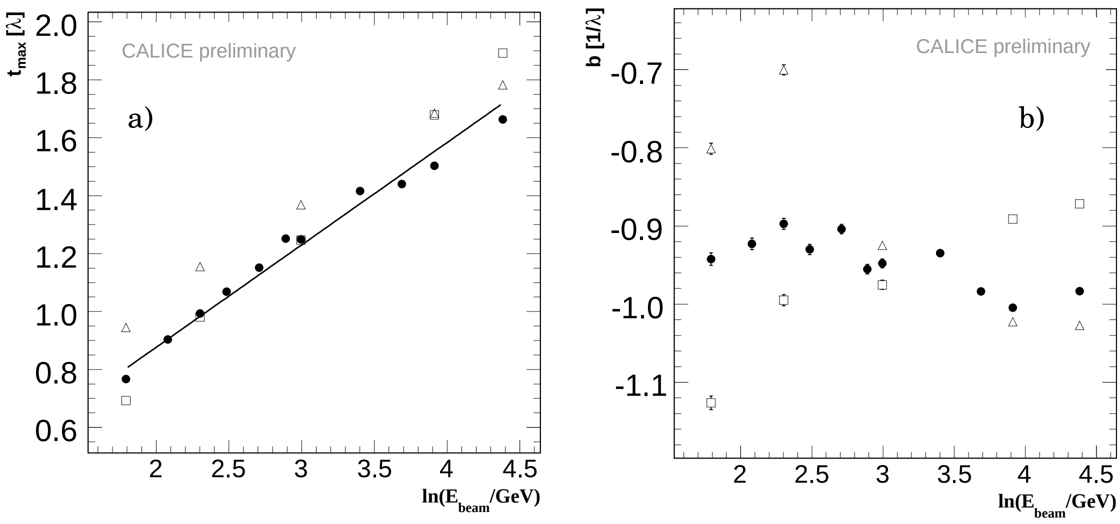

Figure 11: *Energy dependence of the longitudinal shower maximum (a) and shower attenuation parameter (b). Black circles are data, open squares digitized LHEP and open triangles are digitized QGSP_BERT.*

and QGSP_BERT. The theoretical value of the attenuation parameter b is -1λ_0, which is consistent with measurements.

3.5 Conclusions

This analysis shows a first comparison of CALICE hadron data to MC models including digitization. At the time of this note only two models were selected for evaluation. The comparison to all models available in G4/MOKKA with higher statistics will follow.
First results indicate that the digitized MC models agree in trend with the data in all observables studied. For the extraction of the total energy from the visible events it is observed that the two MC models have different MIP to GeV conversion factors or different sampling factors for hadrons. This effect will have to be investigated in more detail.
The deviation from linear behaviour at high energy is different in data and MC. While data is consistent with linearity up to 80 GeV at the 2.5% level, both MC models indicate a positive deviation. This effect could be due to an increasing e/π at high energy, which is either not observed in data or compensated by a remaining non-linearity effect. It should be stated that from the electromagnetic analysis, reported in a parallel note [1], the deviation from linearity at the highest energy is about 4%, where systematic uncertainties on the saturation scale are around 2%. In the pion analysis the effect of non-linearity should be comparable or even smaller than in the electromagnetic case.
From the longitudinal profile comparison, LHEP seems to describe the shower shape better, while QGSP_BERT implies a later start of shower with respect to data. The soft containment cuts in the second half of the TCMT can introduce a mild bias in the longitudinal shower development by enhancing the number of events with high electromagnetic fraction. If on one side this cut is necessary to ensure a Gaussian shape in the total energy sum, it should be removed to study shower profiles in more detail.
These preliminary studies will need to be corroborated by larger statistics and deeper in-

vestigation of the digitization process itself. From the electromagnetic analysis the MC digitization has been shown to have remaining effects of the order of 10%, which need to be understood. Systematic errors on the data are still missing, but they are in large part similar to those in the electromagnetic analysis.

4 Shower leakage

Hadronic shower development suffers from much larger fluctuations than electromagnetic ones. In particular the shower starting point varies over a much larger range in the calorimeter depth from event to event. The determination of the shower starting point is crucial to study hadronic shower shapes in detail and to address the problem of shower leakage. The definition of shower starting point is based the topological resolution of the AHCAL. The track of the incoming pion is followed through the first layers of the detector until its first interaction takes place and an energy larger than minimum ionizing is released. The starting point identification method can be generalized to neutral hadron showers, but this study is restricted to charged hadrons.

4.1 Data sample and event selection

For this analysis pion data collected with AHCAL and TCMT only are used, the ECAL was removed during these runs. Runs from 2007 CERN test beam in the energy range 8-20 GeV have been considered where the effects of the calibration chain are best under control. The runs taken with a detector rotation angle of 30 degrees are chosen to have the maximum possible depth of active detector.
Data was selected which have a beam trigger. No other trigger selections were imposed. For the identification of the shower starting point the detector layers are scanned from front to back in groups of three in depth. The energy deposited and the number of hits in the three consecutive layers are summed. An optimum set of cuts is found to distinguish between track stub and beginning of shower type behavior. The front of the first block of calorimeter layers in which the energy deposited is larger than 8 MIPs and the number of hits larger than 5 is defined as the shower starting point. The cut values have been tuned for 10 GeV pion data and checked to be stable for all energies analyzed.

4.2 Calibration

The shower starting point definition is based on the identification of the minimum ionizing track of the pion before the first interaction. The energy deposited and the number of hits created by a muon in each layer of the calorimeter are used as normalization.

4.3 Results

Fig. 12a shows the occurrence of shower starts as a function of the number of interaction lengths, λ_0.

The decreasing exponential slope agrees with expectations.
Fig. 12b shows the longitudinal profile for 10 GeV pion showers. Knowing for each event the shower starting point, it is possible to shift the single event profile to the same starting point. This is shown by the open symbols. The longitudinal shower profile becomes much shorter after correcting event-by-event for the variation of the shower starting point. Note this data set is different from that in Fig. 10. It refers to 2007 pion data with the calorimeter at an incident angle of 30 derees, instead of 2006 data at zero degrees. In

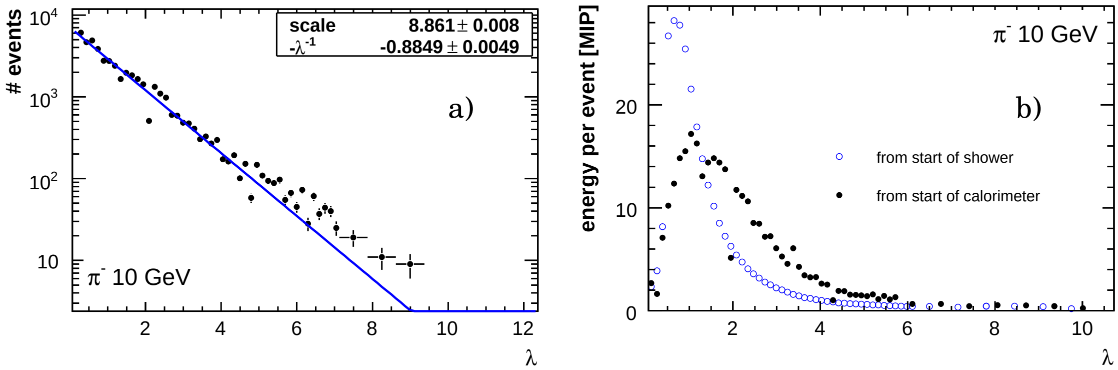

Figure 12: *a) Position of shower start in the calorimeter as a function of the number of interaction lengths. b) Longitudinal shower profile in the calorimeter determined before and after the shift, event by event, to the shower starting point, for 10 GeV pion showers.*

the solid points a remaining effect of wrong calibration of some cells is visible, which is smeared when shifting the shower starting point (open symbols). This will improve with the next release of calibration constants.
For the study of the leakage, data has been binned according to their starting point and the total energy reconstructed in the AHCAL is displayed in Fig. 13a.

For increasing beam energies the effect of losses due to leakage becomes more severe, an effect which also worsens the energy resolution as shown in Fig. 14a and Fig. 13b respectively. Taking the curves shown in Fig. 14a a correction function for the total energy reconstructed in showers started at various depths in the calorimeter is calculated. After correction, the improvement on the energy sum in the AHCAL is seen in Fig. 14b.

4.4 Conclusions

This analysis indicates the possibility to determine the shower starting point from the highly granular information of the AHCAL. The knowledge of the shower starting point can be used to correct the longitudinal shower profile to obtain the true development of a pion shower in depth. It can also be used to correct for longitudinal leakage in the AHCAL. For moderate energies up to 20 GeV as studied in this analysis, the correction does not improve the single hadron energy resolution. This needs to be studied at higher energies and compared to simulations.

5 Two-showers separation

The ultimate goal of highly granular calorimeter is shower separation in a single event. To test the CALICE calorimeter system for shower separation quality two single pion events are overlaid and reconstructed with a certain cluster algorithm (track-wise clustering algorithm). The quality of shower separation is studied as a function of the distance between the two pions. Furthermore, a naive particle flow algorithm can be applied to the reconstructed clusters assuming one of the two pions to be charged and, therefore, having a perfectly determined momentum from the tracking system. The energy resolution of the second pion (playing the role of the neutral hadron here) is investigated.

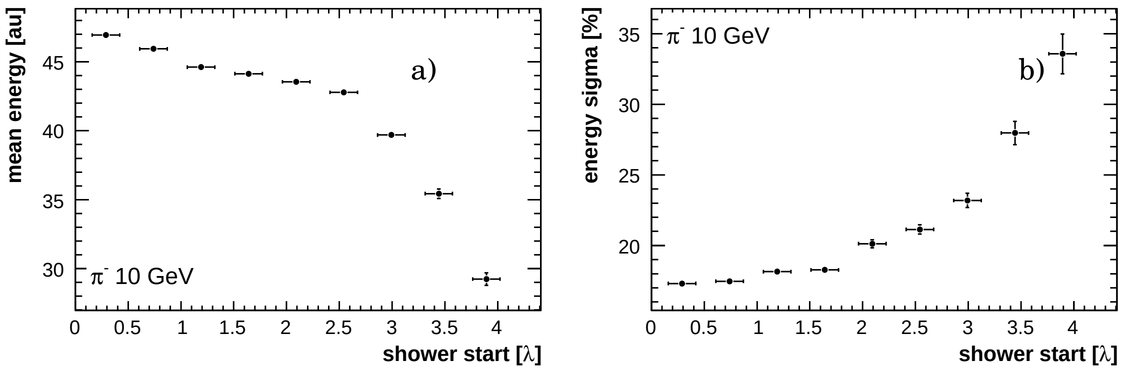

Figure 13: *a) Mean energy deposited in the calorimeter as a function of the shower starting point. b) Energy resolution as a function of the shower starting point, for 10 GeV pion showers.*

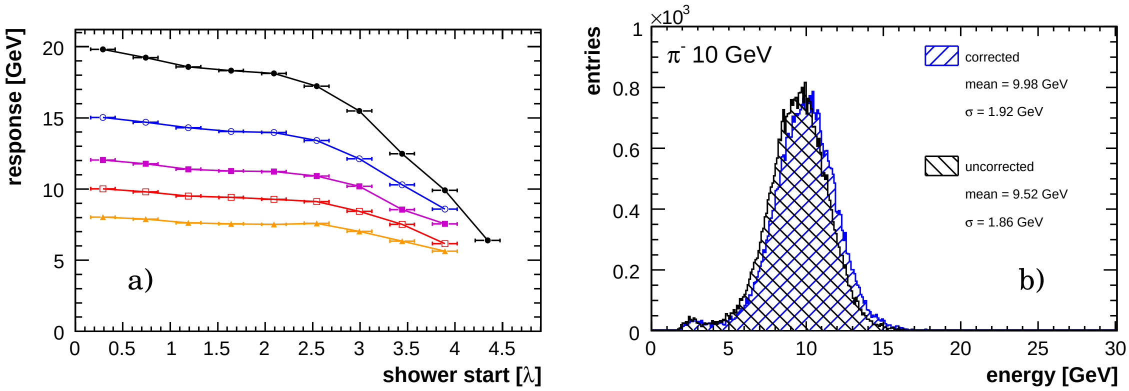

Figure 14: *a) AHCAL response as a function of the shower starting point, for various beam energies from 8 to 20 GeV. b) Total energy distribution in the calorimeter before and after correction for leakage, for 10 GeV pion showers.*

5.1 Data sample

For this analysis, pion data collected with ECAL, AHCAL and TCMT at the CERN test beam in 2006 are used. The runs taken with a detector rotation angle of 0 degrees are analized. It is planned to extend this analysis to include also showers taken at different incident angles.

5.2 Event selection

Data was selected which had a beam trigger (10x10 cm coincident trigger). The desired events are pion events in which the hadronic shower is fully contained in the AHCAL. The signature of these events is a MIP like track in the ECAL and low activity in the TCMT, which is compatible with noise. The position of the MIP like track in the ECAL is determined by searching for a tower of ECAL hits. The position of the tower determines the "reconstructed" position of the pion. If such a tower is found, the tower is extended by adding the neighborhood of the tower. The neighborhood around the tower are the 3×3 cells next to the tower at each layer.
To ensure that the particle does not start to shower in the ECAL, the properties of the extended tower and the hits outside the extended tower are examined. The cuts are the following: There must be less than 8 hits outside the extended tower and none of these hits must have an energy of larger than 15 MIP. At most 4 layers in the extended tower may have more than one hit. No layer within the extended tower must have more than 3 hits The energy of every hit within the tower must be less 20 MIP. Additionally no hit in the last layer of the extended tower must have an energy of more than 8 MIP.

Finally the gaps in the tower are considered. A gap is defined as one or more consecutive layers with no hit in the extended tower. The longest gap must not be larger than 3 layers. Additionally the number of layers, which have hits within the towers neighborhood plus the number of gaps must be larger than 26. This requirement puts a limit on the number of large gaps.

The cuts on the TCMT response which are used are: total visible energy in the TCMT$<$ 8.5 MIP; $E_{vis,TMT}[mip] < N_{hits,TCMT} - 0.3$; number of hits in the first three TCMT layers must be less than 4.

The 100×100 trigger behind the tail catcher is used to veto muon events.
Finally a soft cut on the number of hits in the AHCAL ($N_{hits,AHCAL} >= 30$) is applied to veto muons which get scattered in the AHCAL and miss the TCMT.

Single pions selected with these criteria are then binned into 1cm$\times$1cm bins according the their track position in the ECAL. Pairs of pions can be created using almost all combinations of distances in the range from 3 to 12 cm.

5.3 Event overlay and clustering

All pion events are first calibrated according to the official calibration chain [2] and the amplitudes per calorimeter cell after saturation correction are stored in units of MIP. A threshold of 0.5 MIP is applied to the energy deposited in each calorimeter cell. After calibration, two pion events (selected according to their distance) are overlaid on a cell basis. The energy of the overlaid event in each cell is the sum of the single pion energies. With this method noise is double counted, but the effect of non-linearity is minimized, since the amplitudes after correction are added. An energy cut off of 250 MIPs is applied

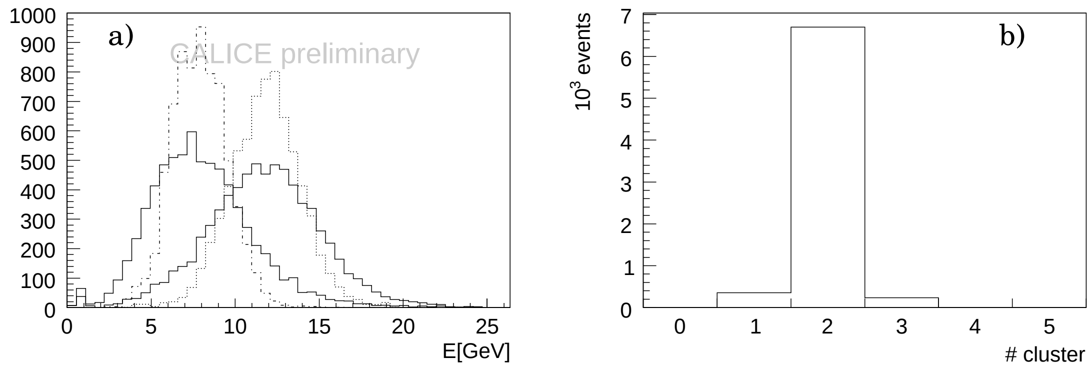

Figure 15: *a) Reconstructed cluster energy and true particle energy for an overlaid event with a 8 GeV and a 12 GeV pion. b) Two clusters separation efficiency for for 8 and 12 GeV pions at 10 cm distance.*

to each cell to simulate the effect of saturation. Eventually, these effects will be included in the simulations.

The overlaid events are then input to a clustering algorithm which identifies the two clusters and measures their total energy. A track-wise clustering algorithm is used, which is an algorithm developed for particle flow in the full ILC detector. The parameters of the code have been tuned for the AHCAL geometry to maximize the number of two identified clusters and minimize the difference between reconstructed and true cluster energy in the calorimeter.

An example is presented here where two pions of 8 GeV and 12 GeV respectively are overlaid. After tuning, the reconstructed cluster energies $E_{cluster}$ are shown in Fig. 15a) as compared to the energy reconstructed in the calorimeter for the single pions, E_{calo}, separately. One can see a broadening of the energy spectra of both pions due to the clustering procedure. The efficiency of two cluster separation in this case is presented in Fig. 15b) and is more than 90% for two particles at a distance of 10 cm.

5.4 Naïve Particle Flow

Given the reconstructed clusters an assumption can be made to test a naïve Particle Flow approach. One of the two pion is considered to be a charged hadron; its energy measured in the calorimeter is replaced by the tracker information, see Fig. 16a. For each overlaid event a "track" is defined by choosing the position of one of the mip stubs in the ECAL. The track's energy, E_{track}, is the beam energy of the corresponding event. The track is assigned to the cluster with cluster position (projection of the center of gravity to the $x-y$ plane) closest to the track.

The charged and neutral hadron energies are then summed to form a "jet"-cluster object. Following this procedure Fig. 16b is obtained which correlates the calorimeter sum of the two pion energies, $E_{calo} = E^1_{calo} + E^2_{calo}$ to the particle flow energy, $E_{PFLOW} = E_{cluster} + E_{track}$. Note that these two quantities are completely uncorrelated and their spread is not expected to agree.

From the comparison of the spread of the two distributions in Fig. 16b) an indication can be drawn: for a 20 GeV "jet" made of two particles at a distance of 10 cm the energy re-

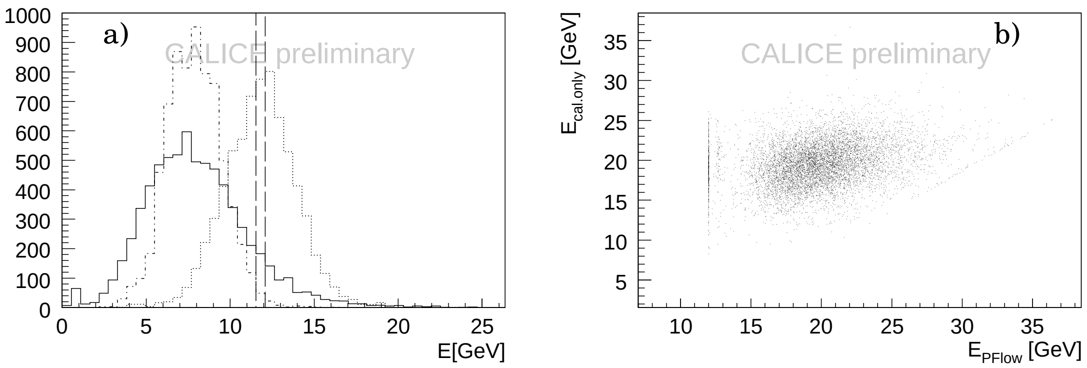

Figure 16: *a) The left heavy dashed line is the energy distribution of $E_{cluster}$, the right heavy dashed line is E_{track}. The dash-dotted and dotted lines are E^1_{calo} and E^2_{calo}. b) Correlation between the particle flow energy E_{PFLOW} and the calorimeter energy E_{calo}, for the sum of a 8 GeV and a 12 GeV pion at 10 cm distance.*

constructed in the AHCAL using a naïve Particle Flow approach has the same resolution as the calorimeter energy.

The study can be performed for all combinations of energies of neutral and charged particle and all distances between 3 cm and 12 cm. Fig. 17 shows the efficiency for resolving two clusters for two energies of the charged particle, 6 GeV and 20 GeV. Neutral particles with energy between 6 and 20 GeV are separated from the charged one. The efficiency for resolving two clusters is larger then 90% for particle at a distance of 10 cm and decreases up to 40-50% for particles at 3 cm distance. This is almost independent of the particls' energies.

To quantify the result of the particle flow approach compared to pure calorimeter energy a quantity has been defined, the efficiency of shower separation. For events where exactly two clusters have been found the cluster energy $E_{cluster}$ in an interval of $\pm\, 3\sigma$ of E^1_{signal} is integrated and compared to E^1_{signal}. The ratio of these two quantities defines the particle flow efficiency,

$$eff_{Pflow} = \frac{\int_{-3\sigma}^{+3\sigma} E_{cluster}}{\int_{-\infty}^{+\infty} E^1_{signal}}, \tag{2}$$

where σ is one standard deviation of the E^1_{signal} distribution. Fig. 18 shows the efficiency of Pflow for two energies of the charged particle (track), 6 GeV and 20 GeV. Neutral particles with energy between 6 and 20 GeV are separated from the charged one. The efficiency increases for larger distances between particles and for lower energy of the two particles to be separated. The test of larger distances is not possible using the currently analyzed data set.
These studies can be directly compared to MC simulation studies [3] performed for the optimization of the AHCAL cell size. Comparing to the 3x3x1 configuration (which is the one the AHCAL has been realized with) there is good qualitative agreement between data and MC trend. It has to be noted that in these studies one neutral and one charged particle were simulated, whereas in data we use two charged pions. In data, shower fully

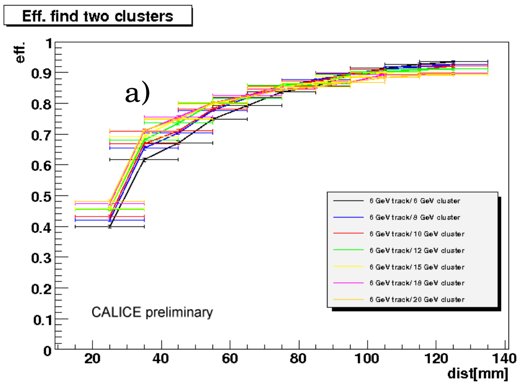

Figure 17: *Two clusters separation efficiency as a function of particle separation for an incident 6 GeV (a) and 20 GeV (b) "charged" particle.*

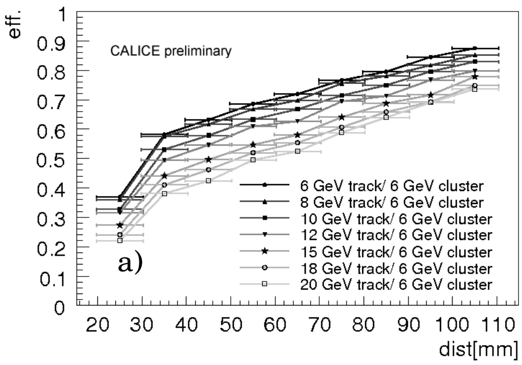

Figure 18: *Efficiency of particle flow (defined in text) as a function of particles separation, for 6 GeV cluster (a) and 20 GeV cluster (b).*

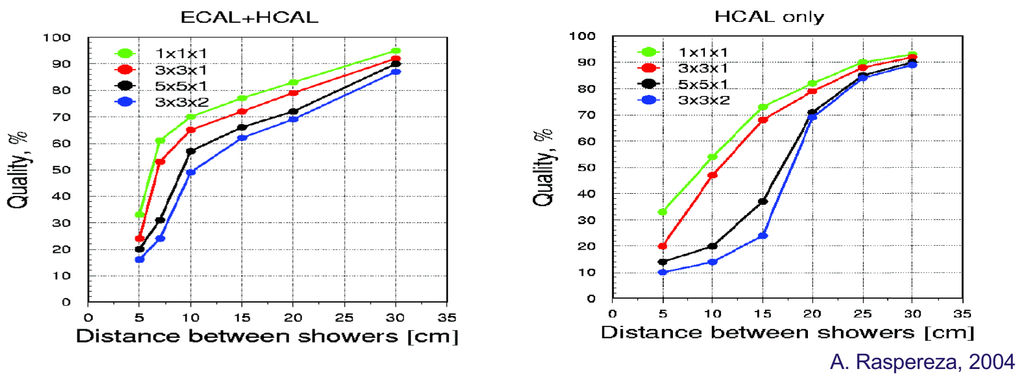

Figure 19: *MC studies of shower separation for various AHCAL tile geometries.*

contained in the AHCAL are used but the ECAL is present and it is used as tracker in the event selection. Detailed comparison with MC is still needed to validate these results.

5.5 Conclusions

This analysis is still very preliminary, but it shows the method and some qualitative results of the first application of Particle Flow type reconstruction to real data.
Be aware that the requirement of fully contained showers in the AHCAL can bias the result, favoring short and broader showers. The requirement of a track-like signal in the ECAL helps the track-wise clustering algorithm to identify the clusters. For a truely unbiased analysis, the track of the "neutral hadron" should be removed before clustering is applied.
One more word of caution, no magnetic field is present in these studies and the particles are perpendicular to the calorimeter face which is not the case at the ILC.
This type of analysis is a powerful tool to compare various clustering algorithms in a realistic hadronic shower environment. Once these results have been validated by comparison to MC prediction they will be an important step in the verification of PFLOW performance estimates for physics events.
For future data taking it is advisable to take data at a larger separation in the AHCAL but keeping the events centered in the ECAL to optimize the track reconstruction this will allow the extension of these studies beyond 11 cm distance between showers.

References

[1] HCAL electromagnetic analysis update 2008, in preparation.

[2] First results from electron data with the CALICE tile HCAL prototype at the CERN test beam, Internal note to the CALICE collaboration, http://www-flc.desy.de/hcal/pictures/cerntestbeam/hcalElectronNote.ps

[3] V. Morgunov, A. Raspereza, ECFA workshop, Montpellier, 13-16 Nov. 2004.