PandoraPFA Tests using Overlaid Charged Pion Test Beam Data

O. Markin¹ on behalf of the CALICE collaboration

¹Institute for Theoretical and Experimental Physics ITEP, Moscow, Russia

E-mail: markin@itep.ru

Abstract. The test beam data obtained with CALICE calorimeter prototypes were used to test the PandoraPFA program. The program capability to recover a neutral hadron energy in the vicinity of a charged hadron was studied. The impact of overlapping of two hadron showers on energy resolution was investigated. The dependence of the confusion error on the distance between a 10 GeV neutral hadron and a charged pion was derived for pion energies of 10 and 30 GeV which are representative of a 100 GeV jet. The comparison of these test beam data results with Monte Carlo simulation using GEANT4 physics lists was performed.

1. Introduction

The Particle Flow Analysis (PFA) algorithm was implemented in the PandoraPFA program [1] as a part of the software for the future International Linear Collider (ILC) and was tested using Monte Carlo (MC) simulated jets. The capability of the PandoraPFA to recover a neutral hadron energy in the vicinity of a charged hadron is of crucial importance because the recovery mistake would degrade the energy resolution. This occurs in case the reconstructing program mixes up hits from showers created by charged and neutral hadrons as a result of shower overlapping. One more factor that degrades the energy resolution is an overlapping of a neutral hadron shower and a photon shower. However to resolve this confusion, contrary to the case of two hadron showers, one can use energy profiles of electromagnetic showers. In case of two hadron showers overlap the task for PFA becomes more complicated because the energy profiles are practically useless and only topological and energy criteria can help to disentangle showers.

The impact of the overlapping of showers on energy resolution for Monte Carlo simulated jets is shown in [1]. However it is known that different available physics lists give noticeably different predictions for hadron shower shapes, that might be important for resolving the overlapped showers. Moreover, the real detector could have not as good performance as that of the idealized MC model. The main goal of this study is to check the PandoraPFA performance using real test beam data and to compare the result with MC predictions. To investigate this issue we use test beam data collected at CERN in 2007 by the CALICE detector prototypes that are very similar to one of the detector technologies proposed for the future ILC and known as the Large Detector Concept (LDC) [2]. The detailed description of CALICE silicon-tungsten electromagnetic calorimeter (ECAL) and high granular analog hadronic calorimeter (HCAL) can be found in [3, 4]. For the first time the procedure of shower overlapping at the CALICE structure was described in [5].

XIV International Conference on Calorimetry in High Energy Physical Science (1997) 2010	ysics (CALOR 2010)	IOP Publishing
Journal of Physics: Conference Series 293 (2011) 012033	doi:10.1088/1742-6	596/293/1/012033

To study the confusion error of overlapped hadron shower recovering, we pass to PandoraPFA two showers: one charged pion shower as it is and one emulating a neutral hadron (e.g. K_L^0 or neutron) shower. For the latter, all hits before the shower starting point were deleted. To confront test beam data with MC, GEANT4 simulation for two physics lists, LHEP and QGSP_BERT, was performed using beam profiles corresponding to the data runs. As the CALICE prototype was tested without magnetic field, the PandoraPFA processor has been adjusted for the case of straight tracks without modifications of the PandoraPFA shower separation algorithm. It is necessary to note that the magnetic field makes it easier for PandoraPFA to separate showers. Therefore our analysis gives a conservative estimate for the PandoraPFA performance.

2. Event preparation

For our study, positive and negative pion induced showers were used. The pion events in the energy range from 10 to 50 GeV were selected from the data collected in 2007 at CERN SPS test beam using complete CALICE setup [3, 4]. The information from the Čerenkov counter was taken into account to exclude beam contamination from electrons in the case of negative pion beam and protons in case of positive one. Besides this trigger, an additional selection procedure was used to purify the sample by excluding muons and trash events including multiparticle ones. The 0.5-MIP threshold was applied to all cell signals to reject noise. The events with MIP-like energy deposition in all calorimeter parts including tail catcher were considered as muons. The same selection procedure was applied to digitized MC samples.

The deposited energy distributions before and after selections are shown in Fig. 1. The blue histograms represent muons. The green histogram corresponds to an electron admixture to the negative pion beam identified by Čerenkov counter on the left plot and to a proton admixture to the positive pion beam on the right one. The resulting selected pion samples (red) are well fitted by Gaussian distribution. Trash and multiparticle contributions are not plotted here because of their low impact comparing to the shown components.



Figure 1. Deposited energy distributions before (shaded) and after selections for 10-GeV negative pion (left) and 30-GeV positive pion (right) beams. Red solid lines show Gaussian fits to resulting pion sample distributions. See text for details.

3. Mixing of showers

The selected pion events were overlaid and passed to PandoraPFA after an energy selection procedure: we selected for the analysis only events with overwhelming containment (more than 95% of deposited energy) in the ECAL plus HCAL, where the fine granularity allows us to

retrieve the shower hit positions. At energies from 10 GeV to 30 GeV showers are not so long; their energy density drops *e* times after about 15 HCAL layers from the shower start. For this reason such a selection does not restrict physics too much. It means that we do not use showers which start in far HCAL layers and it is worth to note that such showers will be better separated due to magnetic field in a future detector and hence the confusion for them will be smaller.

We took two pion events from different runs and for one of them moved all hits by a certain transverse (w.r.t. beam direction) distance from their original position. This distance varied from 5 cm to 30 cm. For this displaced event we kept only hits behind the shower start, i.e. proper shower hits without incident track, and used this shower as an imitation of the neutral hadron shower. This imitated shower represents the main subject of our study. In what follows we will call the energy of this imitated hadron shower the *neutral hadron shower energy*.

The top row of Fig. 2 shows the energy distributions for the 30 GeV charged (left), 10 GeV charged (middle), and 10 GeV "neutral" (right) hadron events after our preparation procedure. These energies were chosen for the following analysis as being typical for 100-GeV jets.



Figure 2. Top row: energy distributions for 30 GeV charged (left), 10 GeV charged (middle) and 10 GeV "neutral" (right) hadron events prepared from data runs for mixing of two showers. Solid lines correspond to Gaussian fit. Bottom row: the same energy distributions after their disentangling by PandoraPFA for the case of 15 cm distance between corresponding charged and "neutral" hadrons.

4. Recovering of showers

Due to shower overlapping in the calorimeter, the energy recovered by PandoraPFA for each of two showers is not accurate. The resulting energy distributions after their recovering by PandoraPFA are shown in the bottom row of Fig. 2. The overlapping is considerable in cases where the charged showering particle appears to be close to the neutral one. Since we know the neutral hadron energy from calorimetric measurements, we can compare it with the energy recovered by PandoraPFA and get the confusion error.

The maximum confusion takes place between the high energy charged hadron and the small energy neutral one (see left plot in Fig. 3). In such a case that high energy is given to the program as a reference point. So the PandoraPFA is to consider a significant number of neutral shower hits as charged shower hits while keeping the difference between the TPC track energy and the energy of charged hadron within the measurement error. Therefore for a large number of events PandoraPFA recovers the charged hadron energy higher than it actually was in the calorimeter, adding to it a significant part of the neutral hadron energy. That mostly happens for events, in which, due to intrinsic shower fluctuations, the difference between the measured charged hadron energy and the beam energy is comparable with the neutral hadron energy and PandoraPFA tries to cover up this difference at the expense of the neutral hadron energy. On the other hand, PandoraPFA can not add to the charged shower the energy much larger than the typical measurement error. Both effects together give a peak around -6 GeV in case of overlapping of 30 GeV charged and 10 GeV neutral hadrons (see Fig. 3).

The shoulder on the right slope of the zero difference peak appears due to the fact that the sufficiently large size of the 30 GeV charged hadron shower gives the PandoraPFA an opportunity to associate many hits of this shower with the neutral hadron and to recover its energy even higher than it actually was. At large distances this confusion vanishes and the mean value of the difference becomes positive since we include into the recovered neutral hadron shower all hits which PandoraPFA has not attached to the charged hadron shower, among them there are isolated hits and small clusters which actually belong to the charged hadron shower but were not recognized. Using the histograms shown in Fig. 3 one can extract a mean value of the difference between recovered energy and measured energy of the neutral hadron. The second characteristic used to estimate the confusion error is the RMS deviation value.



Figure 3. Difference between the recovered energy and the measured energy for the 10 GeV neutral hadron at 5 cm (left) and at 30 cm (right) from 30 GeV charged hadron. Data (red, yellow shaded) is compared to MC predictions for LHEP (blue) and QGSP_BERT (green) physics lists.

The mean value of the neutral hadron energy measured in the calorimeter exceeds the mean energy recovered by PandoraPFA at small distances between particles where shower overlapping is considerable (see Fig. 4). At large distance where confusion vanishes, the mean recovered energy of the neutral hadron becomes even higher than its mean measured energy. That happens because we include in the neutral hadron energy the energy of isolated hits and small clusters which in fact belong to the charged hadron shower but could not be associated with it because of remoteness.

The transverse shower size does not affect the situation appreciably whilst the showers are so close to each other that PandoraPFA tends to attach the small neutral particle shower to the big charged one. As the distance between hadrons grows, the transverse size naturally dominates the shower confusion (see Fig. 5). For this reason LHEP based simulation which gives smaller values for transverse shower sizes, predicts smaller confusion for the distances larger than 15 cm.





Figure 4. Mean difference between the recovered energy and the measured energy for the 10 GeV neutral hadron vs. the distance from the 10 GeV (solid) and 30 GeV (dashed) charged hadron.

Figure 5. *RMS* deviation of the recovered energy of 10 GeV neutral hadron from its measured energy vs. the distance from 10 GeV (solid) and 30 GeV (dashed) charged hadron.

The RMS deviation of the recovered energy of a neutral hadron from its measured energy can be interpreted as a confusion error. It is particularly large for the 30 GeV charged and 10 GeV neutral hadrons overlapping. However, this does not affect the jet energy reconstruction accuracy too much because the probability to find a 30 GeV fragment in a 100 GeV jet is relatively low.

Fig. 6 shows the probability of recovering of the neutral hadron energy within 3 standard deviations from its real energy at different distances from 10 GeV and 30 GeV charged hadrons. For the 10 GeV neutral hadron we take the standard deviation equal to $0.6\sqrt{10 \times 0.86 \times 0.97}$ GeV, where 0.86 is the π/e ratio and the coefficient 0.97 takes account of track fragment loss for the imitated neutral shower. The latter coefficient is just the approximate ratio of the mean value of the right histogram in Fig. 2 to the mean value of the middle one.

If a charged hadron is situated in the vicinity of a neutral one with similar or exceeding energy, the confusion is not so large. The relative confusion is larger for small neutral hadron energy. This results in a lower probability of neutral hadron energy recovering for small neutral hadron energy (see Fig. 7).

5. Summary

To test the PandoraPFA algorithm, we have mapped pairs of CALICE test beam events shifted by the definite transverse distances from each other onto the LDC detector geometry. Then we modified the treatment of tracks in the PandoraPFA processor for the case of straight tracks. In this study we have investigated the hadron energy range for a 100 GeV jet. For jet fragment energies from 10 GeV to 30 GeV we estimated the confusion error for the recovered neutral hadron energy caused by the overlapping of showers. We have confronted our result for test beam data with the result of Monte Carlo simulations for LHEP and QGSP_BERT physics lists. The results for the data and MC are in a good agreement. This fact together with the successful PandoraPFA performance for simulated jets [1] allows us to consider the PandoraPFA program as a good reconstruction tool for a full-size experiment. All the way we were trying not to simplify the task for PandoraPFA to disentangle showers in order to get a conservative estimate



Figure 6. Probability of the 10 GeV neutral hadron energy recovering within 3 standard deviations from its real energy vs. the distance from 10 GeV (solid) and 30 GeV (dashed) charged hadron for data (red) and Monte Carlo simulations using LHEP (blue) and QGSP_BERT (green) physics lists.



Figure 7. Probability of the neutral hadron energy recovering within 3 standard deviations vs. the neutral hadron energy in the vicinity of 10 GeV (red) and 30 GeV charged hadron (blue) for typical distances from the neutral hadron in a 4T magnetic field in LDC.

for the confusion error. In particular we assign all isolated hits to the neutral hadron shower and we also underestimate the separation of showers towards the end of the calorimter because, unlike a full detector, our testbeam apparatus has no magnetic field.

The agreement between the PandoraPFA performance achieved with real calorimeter prototype data and with the MC simulation demonstrates that the extrapolation to the complete detector is reliable. No hidden imperfections in the real data (non-perfect calibration, non-uniformity of tile response, cross talk between tiles, noise, dead or noisy channels) which could deteriorate the PFA performance were found. We point out that in our study LHEP physics list gives worse predictions for test beam data than the QGSP_BERT one.

This work was supported by grants of Russian National Educational Center 02.740.11.0239, Russian Scientific Schools 3270.2010.2, and Russian Foundation for Basic Research - Helmholtz Association 07-02-92281, 09-02-91321.

References

- [1] Thomson M A 2009 Particle Flow Calorimetry and the PandoraPFA Algorithm NIM A 611 25
- [2] Behnke T, Damerell C, Jaros J and Miyamoto A 2007 Detectors (ILC Reference Design Report vol 4 DESY 07-046) p 25
- [3] The CALICE collaboration 2008 Design and electronics commissioning of the physics prototype of a Si-W electromagnetic calorimeter for the International Linear Collider *JINST* **3** 08001
- [4] The CALICE collaboration 2010 Construction and commissioning of the CALICE analog hadron calorimeter prototype JINST 5 05004
- [5] Morgunov V, Raspereza A 2004 Novel 3-D clustering algorithm and two particle separation with tile HCAL (Paris: Linear colliders vol 1) p 431