Summer Student Project Report: Investigating calibration hits in the ATLAS calorimeter

Claire Antel

05/08/2013

supervisor: Dr Heather Gray

Abstract

Particle flow is a process in which information from all subdetectors are combined using various algorithms to reconstruct objects with the most accurate measurements of a particle's momentum and energy. For the ATLAS experiment, the largest improvement would be to reject pileup contributions in the calorimeter by associating those energy deposits to additional vertices using their track measurements. This report summarises an investigation into the clustering behaviour in the ATLAS calorimeter using specially generated files that include calibration hits. It contributes to a larger project in which the aim is to develop and improve the algorithms used to construct particle flow objects, with the overarching object being the rejection of pileup. Initial results show that a significant fraction of particles produce more than one cluster or none at all. These conditions, however, are shown to have an η and p_T dependence, which will help us reject particles that are likely not going to produce a single cluster only.

1 Introduction

The studies presented here are performed in the context of the construction and development of particle flow in ATLAS events. Particle flow is the process of reconstructing and identifying stable particles in an event by drawing upon and associating all information recorded by the various detector components. The process requires development and fine-tuning of algorithms used in each sub-detector in order to afford maximum efficiency and low fake rates, the main benefit being more accurate momentum and energy measurements through the association of tracks to calorimeter clusters. Particle flow has been widely used in the CMS experiment[1].

The focus of these studies is the part of the reconstruction that links up the information from the charged particle tracking system with the information obtained from the calorimeter. The calorimeter in the ATLAS detector measures the energies of charged and neutral particles. A particle travelling through the calorimeter will develop into a shower of particles, which deposit their energies in the cells of the calorimeter to be measured. The ATLAS calorimeter has around 100 000 cells[2]. However, instead of working with individual cells, cells are put together into clusters. This is done in order to reduce the noise in the calorimeter, as only cells with certain significant energy above the noise level are accepted as part of the cluster. Cells are clustered using the Topological Cluster Algorithm. This algorithm first finds a seed, i.e. cells with a certain significant amount of energy above the expected energy due to electronic noise (the first energy threshold). All neighbouring cells (and cells neighbouring these) that lie above some secondary energy threshold are then added to the seed. Adjacent cells are added until no more lie above this threshold. Outer adjacent cells are then added to the cluster if they pass a tertiary energy threshold. The energy thresholds are defined as

$$\Gamma = \left| \frac{E_{cell}}{\sigma_{noise,cell}} \right| . \qquad (1)$$

Here, $\sigma_{noise,cell}$ depends on the expected electronic and pileup noise. ATLAS uses settings of $\Gamma = 4$, 2 and 0 for the primary, secondary and tertiary thresholds respectively.

In addition, it is also attempted to recover a split cluster (the so-called Split Shower Recovery Algorithm), i.e. two clusters are merged into one if certain conditions suggests they both originate from the same particle. [3]

Associating clusters to tracks in the tracking device is constructive as at $\sim 150GeV$ and below the energy measurements in the tracker have a better resolution than the energies measured in the calorimeter. During the particle flow reconstruction process the recontructed track trajection is extrapolated to the calorimeter, whereupon the nearest cluster is sought out. The measured energy of the linked cluster is then replaced with the track momentum measurement.

2 Motivation

The analysis presented here attemps to commission particle flow reconstruction for ATLAS. We aim to reject pileup: the association of clusters to tracks in eflow objects enables us to more succesfully associate the correct vertices to clusters, as we can extrapolate tracks to vertices with much greater precision than linking clusters to vertices directly.
The behaviour of clustering for different particle types, p_T ranges and η regions were investigated using calibration hits. Calibration hits are objects that are produced in specialised Monte Carlo simulations that provide full information on which particles deposited energy in calorimeter cells and clusters (one has the truth energy deposited in cells, and also the invisible and missing energy).
The D3PD file samples were specially created to include calibrated clusters, calibration hits, cell information, mc truth information and eflow objects. To make initial analysis easier single particle events were created (so that it is certain that any clustering in the event is due to one particle). Files were generated for π^0, π^+ and π^- particles at pileup noise suppression thresholds 0, 8 and in the case of π^0 and π^+, 30. The effects of threshold settings on clustering, the clustering dependence on p_T and η as well as the efficiency of the Split Shower Recovery Algorithm were investigated.

3 Analysis

Figures 1, 2 and 3 show the effect of changing the noise threshold for π^0, π^+ and π^-, respectively. Figure 4 summarizes these effects. The plots illustrate the strong dependence of the number of clusters on the noise threshold and as expected, the number decreases with higher noise thresholds. Comparing between the different particle types, one can see that in the case of the π^0 particle, $\sim 38\%$ of the events produced one cluster, $\sim 25\%$ of the events measured no clusters and $\sim 27\%$ of events had a splitting of clusters. The number of clusters peaks at 1 cluster. Meanwhile, for the π^+ particle (and presumably similar for the π^-) the number of clusters peaks at 0 clusters (at 30%). Around 25% have 1 cluster, and cluster splitting happens in $\sim 16\%$ of events. The mean cluster number is shown in Fig. 4. It shows that π^0 has a mean cluster number close to one for threshold 30. For the charged pions the mean number of clusters appears to always be slightly higher.

It is evident from looking at Fig. 1 and 2 that a substantial amount of clusters are still split, or completely lost at threshold 30. What was initially suprising is that the charged pions appear to produce more fragmented clustering than the neutral pion. One could expect the opposite as the neutral pion decays into two gamma-rays and so is more likely to produce a split cluster. This however is most likely due to the fact that the Split Recovery algorithm is tuned for the π^0 decay specifically.

Next, events with a specific number of clusters were looked at in detail to see how this depends on p_T and η. This was done for the neutral and positive pion at threshold 30. The plots show the fraction of the total number of events in which 0 clusters, 1 cluster, 2 clusters and $>$2 clusters occur. Figures 5 and 6 show this as functions of p_T and Fig. 7 and 8 as function of η.
In both the neutral pion and charged pion case, events in the higher p_T region become more frequent the higher the cluster number. It appears that the more energy a particle travels through the calorimeter with, the more likely it is that more than one cluster will be seeded.
Looking at the very low p_T end it is found that for both particle types all events producing 0 clusters have p_T's no larger than $30MeV$, indicating that a low p_T cut will enable us to discard this fraction of unusable events.
At low pT, the $\pi^0 s$ typically produce 1 or 2 clusters and also at a higher p_T range than the π^+. The π^+ particle, on the other hand, typically produces more than 2 clusters, and this well into the higher p_T range.

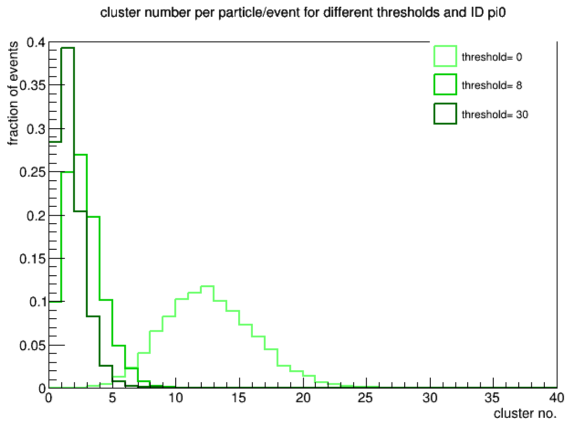

Figure 1: The fraction of number of clusters for π^0 for thresholds 0, 8 and 30 is displayed. The peak number of clusters decreases from 12 down to 1 as threshold values increase from 0 to 30.

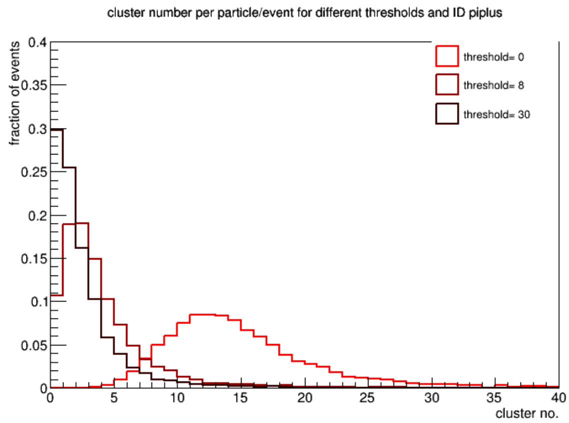

Figure 2: The fraction of number of clusters for π^+ for thresholds 0, 8 and 30 is displayed. The peak number of clusters decreases from 11-12 to 0 clusters for rising threshold values.

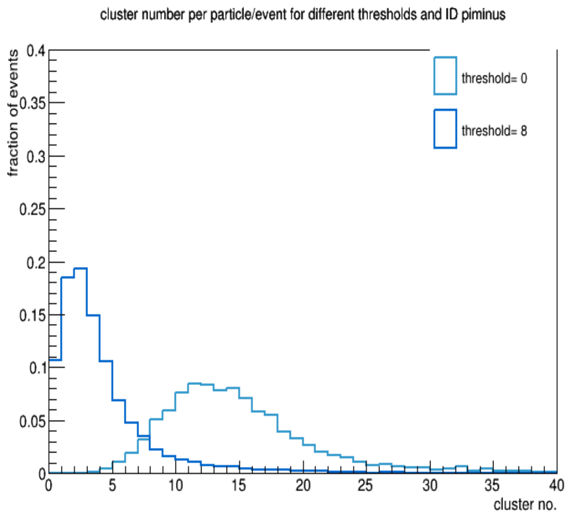

Figure 3: The fraction of number of clusters for π^- for thresholds 0 and 8 is displayed. It is almost identical to the same plot for π^+.

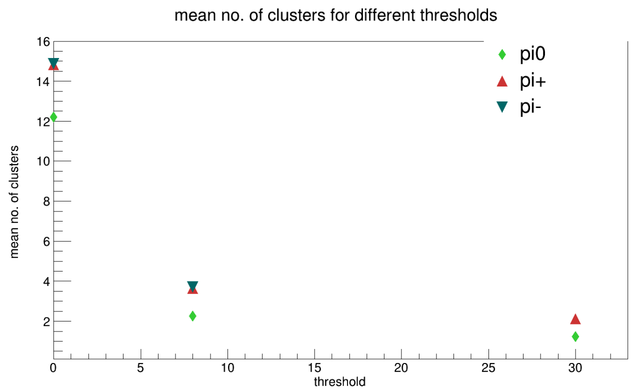

Figure 4: The mean number of clusters for each particle type at each threshold value. The mean cluster number is always lowest for π^0.

Since the π^0 particle decays to two γ's, one may expect that this particle will produce more split clusters than the charged pions, explaining the higher fraction of 2 clusters in an event. Moreover, one would expect this to be more likely in the low p_T range, as the gamma-rays will be emitted further apart from each other, whilst at high p_T these clusters may start merging. The higher fraction of single cluster events at high p_T for π^0 may be an indication of this, though a substantial fraction of split cluster events remains evident at high p_T, also.

In Fig. 7 and Fig. 8, which show the η distributions for π^0 and π^+ repsectively, it is demonstrated that in both cases the particles are more likely to be completely lost (no clusters) in the outer η regions (the distribution has a $\cosh\eta$ shape). Meanwhile, cluster splitting is most likely to occur in the central η region. The value for $\sigma_{noise,cell}$ for each cell depends on the pseudorapidity of the cell, where the dependence goes like $\cosh\eta$. This explains the shape of the η distributions for 0 clusters, as the higher noise level in the outer η regions will decrease the number of clusters formed. Likewise, the distribution shape for 2 clusters becomes the inverse of this for the same reason. A particle breaking into more than 2 clusters is also more common in the central region and two peaks are evident at around $\mid \eta \mid = 1.8$. These peaks are most likely due to the extra material behind the presampler that one finds in this region.

The distribution for 1 cluster differs in the two cases. For the π^0, the distribution is fairly even, whilst for the π^+ particle, single clusters appear to be more common in the outer η regions. Why the η distributions for 1 cluster differ between the two particle types is not understood as yet.

4 Conclusion

A study of the clustering behaviour of cells in the ATLAS calorimeter was presented in the context of the development of algorithms used for the reconstruction of particle flow objects. The results showed that there are significant fractions of events in which no cluster appears or the cluster is fragmented (this more so for the charged pions than the neutral pion). However, the low p_T dependency in the 0 cluster events means that a low p_T cut will be sufficient to discard these events. Some dependencies on p_T and η are not quite understood, and it is clear that events must be looked at in much greater detail.

5 Future work

Further insight may be gained by looking at the energy ratio between the clusters in the event of a cluster split. If the energy of the one generally proves to be insignificant then this one can be discarded leaving one cluster to represent the particle. Also, the spatial splitting of clusters may be worth investigating, to see if clusters tend to split in the θ or ϕ direction or in a longitudinal direction.
The impact of pileup on clusters will be investigated in future studies.

References

[1] CMS Collab. Particle-flow event reconstruction in cms and performance for jets, taus, and met. Technical Report CMS-PAS-PFT-09-001, CERN, 2009. Geneva, Apr 2009.

[2] *ATLAS liquid-argon calorimeter: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1996.

[3] M Hodgkinson, D Tovey, and R Duxfield. Energy flow reconstruction with the eflowrec combined reconstruction software in athena 15.6.9.8. Technical Report ATL-PHYS-INT-2011-031, CERN, Geneva, Apr 2011.

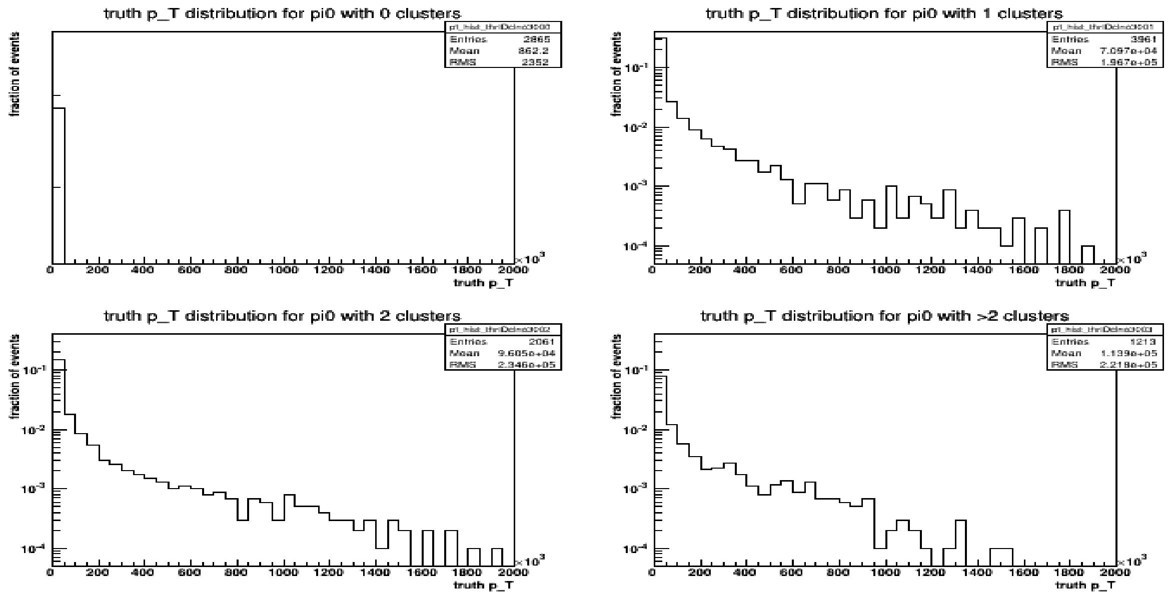

Figure 5: The p_T distribution for 0, 1, 2 and $>$2 number of cluster(s) for the π^0 particle is displayed. The abrupt drop in the 0 cluster plot and the tails in the plots for 1 clusters show that high p_T particles are more likely to form clusters. The mean p_T value is highest for > 2 clusters, suggesting that higher p_T particles are more likely to form fragmented clusters.

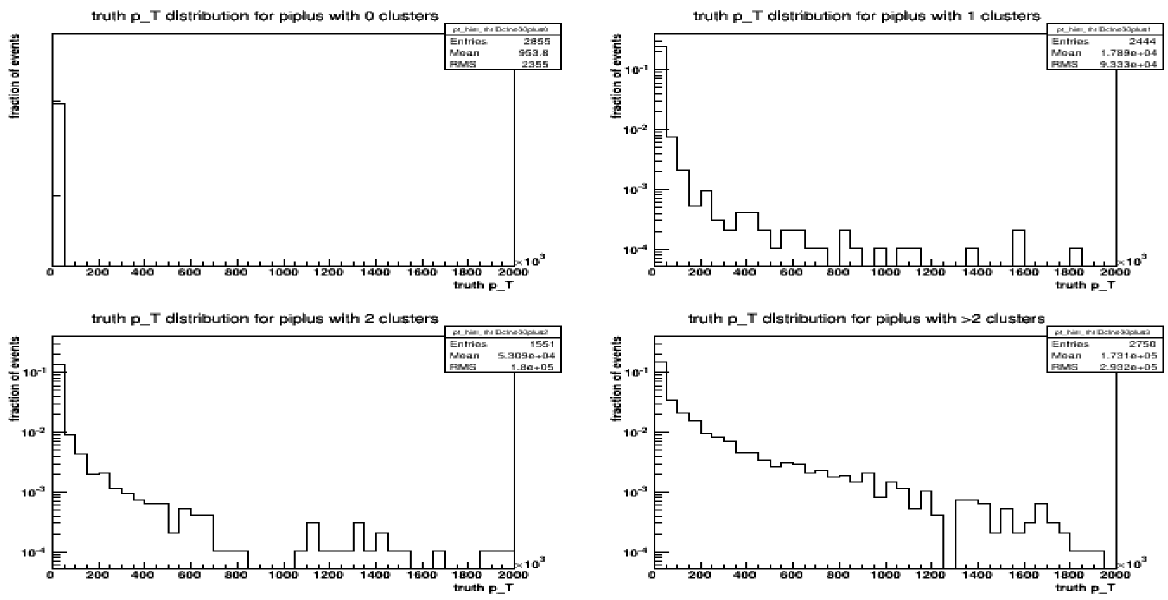

Figure 6: The p_T distribution for 0, 1, 2 and $>$2 number of cluster(s) for the π^+ particle is displayed. The fraction of events with more than 2 clusters appears higher than that of π^0. The fraction of events with 1 or 2 clusters, on the other hand, appear lower than for π^0 and drop off sooner along the p_T axis.

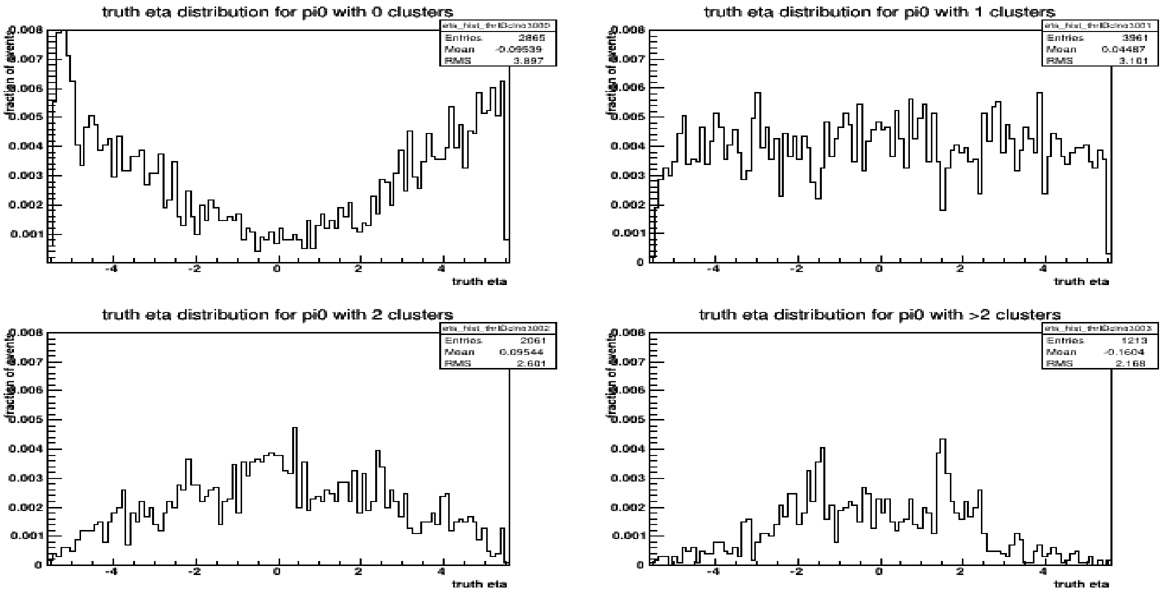

Figure 7: The η distribution for 0, 1, 2 and >2 number of cluster(s) for the π^0 particle. The plot for 0 clusters has a $\cosh\eta$ shape whilst the plot for 2 clusters is the inverse of this. This is due to the fact that the expected noise value in each cell follows a $\cosh\eta$ dependence. The plot for 1 cluster shows an almost even distribution. At >2 clusters, the plot shows a similar shape to the 2 cluster plot, but in addition has two peaks at around $+/-1.8$, this most likely being due to extra material found infront of the calorimeter in these regions.

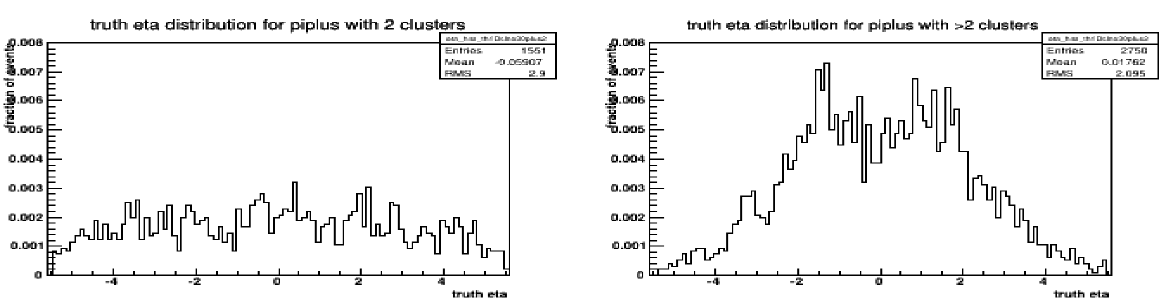

Figure 8: The η distribution for 0, 1, 2 and >2 number of cluster(s) for the π^+ particle are displayed. Just as for π^0, the plots at 0 and 2 clusters display a $\cosh\eta$ dependence, and the plot for >2 clusters contains the two peaks, which are due to extra material infront of the calorimeter. The plot for 1 cluster also carries a $\cosh\eta$ shape - unlike the even distribution found for the π^0.