THE FLORIDA STATE UNIVERSITY

COLLEGE OF ARTS AND SCIENCES

STUDIES IN DEVELOPING A PARTICLE FLOW ALGORITHM FOR THE NEW CMS FORWARD CALORIMETRY UPGRADE

By

HAMPTON BLACK

A Thesis submitted to the
Department of Physics
in partial fulfillment of the
requirements for graduation with
Honors in the Major

Degree Awarded:
Spring, 2014

The members of the Defense committee approve the thesis of Hampton Black defended on April 16, 2014.

Dr. Todd Adams
Honors Thesis Director

Dr. Paul Eugenio
Committee Member

Dr. Tomasz Plewa
Outside Committee Member

TABLE OF CONTENTS

ABSTRACT

This thesis covers comparison studies for the Shashlik calorimeter, one of three possible electromagnetic calorimeter (ECAL) upgrades for the CMS detector at CERN. Currently, the lead tungstate crystals are doing very well; however, when the LHC will increase the total integrated luminosity, the crystals in the forward end-cap will eventually become severely damaged from radiation, which is why scientists are investigating possible replacements.

The methods used involved developing a Particle Flow algorithm, in hopes of getting better results and increased accuracy when the number of interactions per proton bunch crossing (instantaneous luminosity) increases from 30 to 140. The Particle Flow algorithm could give improvement on the resolution in the high intensity environment compared to simpler algorithms.

The Particle Flow algorithm being tested is still under development. Comparisons between three different clustering algorithms for the simplest case of electrons are presented. Initial studies of pions are also shown. The resolution for electrons is found to be around 6-8% for energies greater than 200 GeV.

CHAPTER 1

BACKGROUND INFORMATION

The main motivation for this thesis is to study a promising technique that could mitigate a major challenge faced by the CMS collaboration in the coming decade. During the initial operations period (2010 - 2012), the CMS electromagnetic calorimeter (ECAL) performed very well (note the discovery of the Higgs boson in 2012 as an example [1]); but with the planned increase in luminosity (two orders of magnitude), parts of the current ECAL crystals ($PbWO_4$) will be damaged. The increased instantaneous luminosity will mean a larger number of interactions per recorded event. Reconstruction algorithms will have to sort through this information to learn about the one interaction per bunch crossing of interest.

1.1 The CMS Detector

The CMS detector is one of the two, large general purpose detectors at the Large Hadron Collider (LHC) at CERN. There are five layers of different detectors (Fig. 1.1): the silicon tracker, the electromagnetic calorimeter, the hadronic calorimeter, the superconducting solenoid, and the muon system. These studies rely primarily on the electromagnetic and hadronic calorimeters, as they are the only ones included in the simulations. Each of these components are either located in the "barrel" section, which forms the lateral section of the cylinder, or in the "end-caps", which close off the cylinder, perpendicular to the beam line. The end-caps are the primary area of interest in our studies (Fig. 1.2).

Electromagnetic Calorimeter

The current ECAL is constructed of lead tungstate ($PbWO_4$) crystals. These crystals are extremely dense, yet optically clear, and make an excellent scintillator, which is the primary means of detection [2]. When electrons and photons pass through the crystals, they interact with the atoms – either with elastic/inelastic scattering or with bremsstrahlung – resulting with bursts of light that are proportional to the original particle's energy [3].

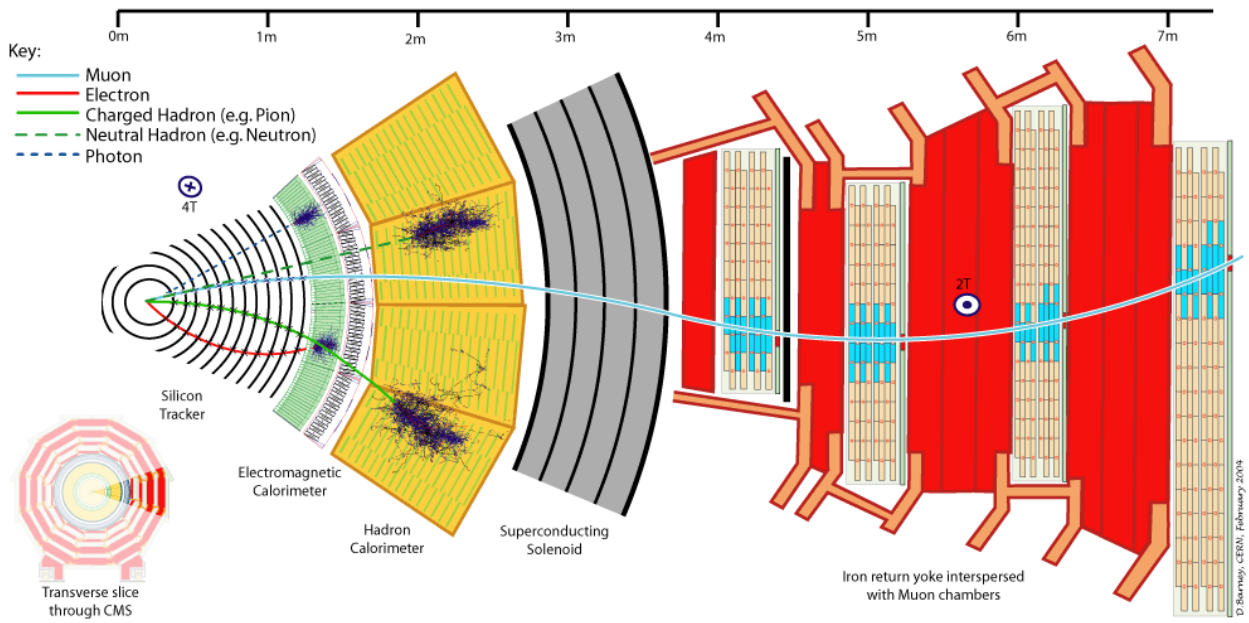


Figure 1.1: A cross section of the CMS detector showing all of the different components of the detector.

Hadronic Calorimeter

The HCAL is very similar to the ECAL, except its purpose is measuring the energy deposited by hadrons. The HCAL consists of layers of dense material, either brass or steel, interweaved with plastic scintillators. The readout detectors are hybrid photodiodes, using wavelength-shifting fibers to measure the small electric signals [4].

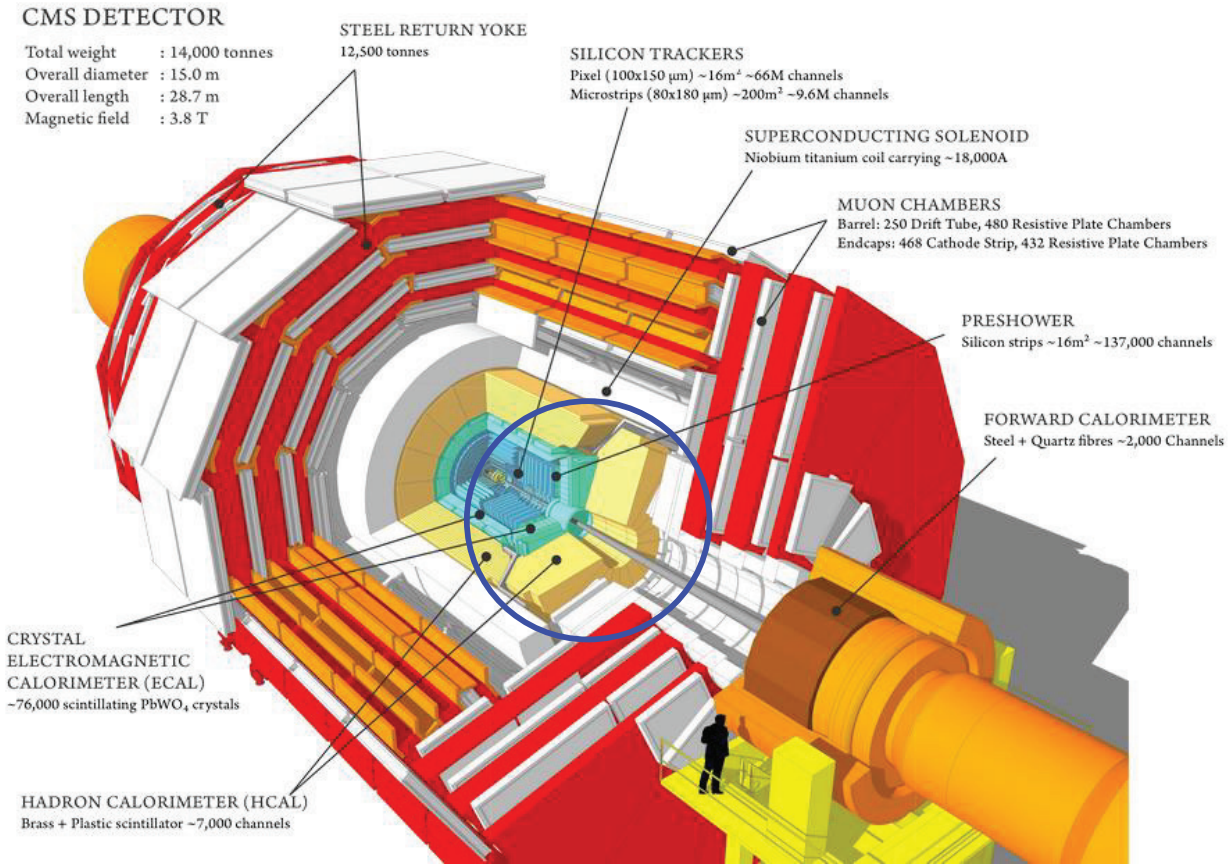


Figure 1.2: Cut out view of the CMS detector. The region inside the blue circle is the forward end-cap.

1.2 The Shashlik Calorimeter

The Shashlik Calorimeter is one of the proposed upgrades for the ECAL, and the primary one investigated by the Forward Calorimetry group at FSU. This new calorimeter is comprised of many layers, alternating between an absorber and an active scintillating piece, with optical fibers running perpendicular to the layers for proper readout against the absorber's opacity (Fig. 1.3) [5].

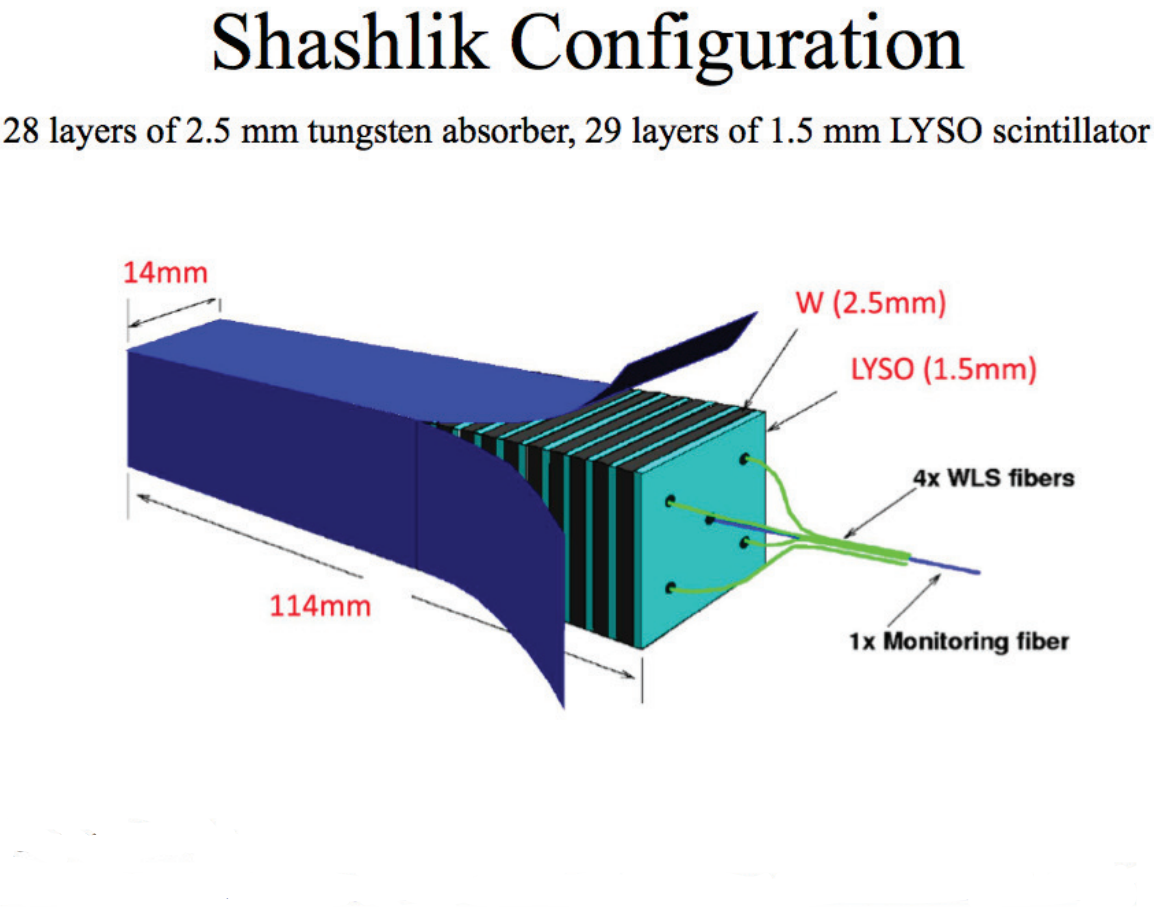


Figure 1.3: Diagram of the proposed Shashlik calorimeter.

The absorber will be either lead or tungsten that absorbs some of the energy from the passing particles. This allows the calorimeter to be more radiation resistant as well as more cost effective. The scintillating layers are made of $LYSO$, Cerium-doped Lutetium Yttrium Orthosilicate, crystals that scintillate with the passage of charged particles. Again, like the current ECAL, charged particles passing through the scintillator will ionize some of the atoms in the crystal, creating a shower of photons. The light is proportional to the energy deposited by the original particle. These studies use the $WLYSO$ simulations, which incorporate $LYSO$ crystals with tungsten absorbers [6].

1.3 Simulations

The simulations use GEANT4 [7] to produce the shell of the CMS detector. It creates the detector space and only includes the ECAL and HCAL. Within this detector space, we simulate some of the possible products of a collision and how they interact with the calorimeters, not the actual collisions themselves. The three different particles we model include: electrons, π^-, and jets. All three of these types are generated with Pythia [8], using various energies and η, a type of angle measurement shown below (Eq. 1.1), over the range: 10 - 1000 GeV, and 1.6 - 2.8, respectively. GEANT4 models the stochastic processes when the particles decay and interact within the calorimeter.

$$\eta \equiv -\ln\left(\tan\frac{\theta}{2}\right) \tag{1.1}$$

where θ is the angle measured from the beam line that the particles are oriented along. Once these simulations are run, we have saved certain information about the particles within our area of interest. We are only working with the end-cap, so any particles that do not end up reaching the front face of the ECAL will be useless. Therefore, we only keep track of particles that have a positive momentum in the $\hat{\mathbf{z}}$ direction. We store the following information per individual particle: momentum in all three dimensions, initial vertex position, and particle identification number in Particle Data Group (PDG) code (a way to identify different particles) [9].

We use a square grid within the ECAL composed of 10000 cells in a 100x100 matrix. Each cell is a square of width 1.4 (1.9) cm for tungsten (lead) absorber. The grid is positioned for each particle/energy/eta combination such that the particle under study will impact the ECAL near the center of the grid. The information we store about this grid include: energy depositions in each cell, and its x and y offset in the detector space (center of ECAL at origin).

CHAPTER 2

METHODS

The primary components developed to study aspects of a Particle Flow-like algorithm include: code that tracks the propagation of particles within the detector, energy deposition clustering programs, and macros that match energy clusters with particles tracked to the respective calorimeters.

2.1 Particle Tracking

Tracking of the particles before they enter the calorimeter is straightforward. The assumptions that are used include: the magnetic field is constant and lies along the z-axis within the detector, the electric field is zero in the space of interest, and the particles are traveling through a vacuum before hitting the ECAL front face. Under these assumptions, the x and y position of a particle is given by,

Charged Particles:

$$x = x_0 + \frac{p_{y0}}{qB_0} - \frac{p_{y0}}{qB_0} cos\left(\frac{qB_0}{\gamma p_{z0}}(z - z_0)\right) + \frac{p_{x0}}{qB_0} sin\left(\frac{qB_0}{\gamma p_{z0}}(z - z_0)\right) \tag{2.1}$$

$$y = y_0 - \frac{p_{x0}}{qB_0} + \frac{p_{x0}}{qB_0} cos\left(\frac{qB_0}{\gamma p_{z0}}(z - z_0)\right) + \frac{p_{y0}}{qB_0} sin\left(\frac{qB_0}{\gamma p_{z0}}(z - z_0)\right) \tag{2.2}$$

Neutral Particles:

$$x = x_0 + \left(\frac{p_{x0}}{p_{z0}}\right)(z - z_0) \tag{2.3}$$

$$y = y_0 + \left(\frac{p_{y0}}{p_{z0}}\right)(z - z_0) \tag{2.4}$$

where z is the distance along the z axis as measured from the origin, B_0 is the magnetic field strength within the detector, x_0, y_0, and z_0 are the initial positions of the vertex in the i dimension, q is the charge of the particle, p_{i0} $(i = x, y, z)$ is the initial momentum of the particle in three

dimensions, and γ is the relativistic factor which can be determined from the momentum. The derivation of this result can be found in Appendix A [10].

2.2 Energy Deposition Clustering

The clustering algorithms were developed by another member of the FSU CMS group. There are three different methods: 5x5 linear energy weighted, 5x5 log energy weighted, and particle flow clusters.

2.2.1 5x5 energy weighted

The 5x5 energy weighted clustering algorithm is the more straight-forward approach to clustering energy depositions. The basic logic organizes energy depositions in terms of seeds. These are the local maxima in the energy deposits in the calorimeters. It starts at the highest energy seed, and then branches outwards searching for other local maxima. Specific criteria are applied to determine if an energy deposition can be a seed: it is a local maximum within the vicinity of the 8 adjacent cells forming a 3x3 matrix of cells, and it is above a certain threshold of energy. It loops over the entire energy distribution, labeling every seed. Once this is done, it starts back at the original seed and creates clusters. These clusters include, at maximum, 25 cells in a 5x5 matrix. As the algorithm continues forward, if a cluster realizes there is another seed inside the cluster it is producing, that cell will not be included. Therefore, you could have an odd shaped cluster towards the end of the program, that may only have 9 cells in it.

For each cluster, the total energy and position of the cluster is calculated using the cells associated with the cluster. The energy is calculated by summing the energy in each cell within the cluster.

Two different techniques are used to calculate the position of the 5x5 cluster: a linear weighted method weighted with the energies of each cell, and a log weighted method that smoothes out the energy weights.

Linear weighted. This is the simpler method. We know the size of the individual cells, so we can use it as a grid in determining position. This method uses the ratio of the specific cell energy to the energy of the entire cluster as a weighting method, shown below in equation 2.5. You

simply loop over all of the clustered cells with this weighting process and add up the positions of these cells.

$$x_{cluster} = \sum_{n=0}^{N-1} \frac{E_{clustercells}}{E_{cluster}} x_{clustercells} \tag{2.5}$$

where $x_{cluster}$ is the position of the 5x5 cluster in the x direction, n is the number of cells in the cluster, N is the total number of cluster cells, $E_{clustercells}$ is the energy deposited into individual cells, $E_{cluster}$ is the energy associated with this cluster, and $x_{clustercells}$ is the position of the cluster cells (each is 1.4 cm wide). A corresponding equation is used to determine the position along the y axis.

Log weighted. The log weighted method is similar. It repeats the process of summing the cell's position with a certain weight attached. However, the weight changes slightly in this case – the final position is determined by normalizing the position with a total summed weight. This method is shown below in equation 2.6.

$$x_{cluster} = \frac{1}{w} \sum_{n=0}^{N-1} \left(4.2 - \ln \left| \frac{E_{clustercells}}{E_{cluster}} \right| \right) x_{clustercells} \tag{2.6}$$

where w is the total summed weight you use to normalize,

$$w \equiv \sum_{n=0}^{N-1} \left(4.2 - \ln \left| \frac{E_{clustercells}}{E_{cluster}} \right| \right)$$

2.2.2 Particle Flow clusters

The Particle Flow clusters use a more elegant approach to clustering the energy depositions in the calorimeters. The topological clustering starts with the maximum energy value in the calorimeter, then does a systematic walk around the detector face, checking if the cells pass a certain energy threshold. Once the program comes across a cell under the threshold, it changes direction until it discovers another region below the threshold. This process continues until it maps out a wide area in which to make a Particle Flow cluster. It repeats this process if there is another seed, or local maximum above a certain threshold, in a region outside of the first topological cluster.

The program repeats this until all seeds are accounted for (Note: most seeds will be within the same topological cluster).

Once the topological clusters are mapped out, the program starts developing the Particle Flow clusters. Each seed gets a Particle Flow cluster that is the same size and shape as the topological cluster it belongs to, in terms of cells. The clustering program puts a bias on cell energies closer to the seed than those further away. Therefore, although each Particle Flow cluster will cluster the same cells, each seed's energy distribution will be different.

2.3 Matching Algorithms

The analysis involves using root macros that combine both the clustering algorithms and the particle tracking code. The goal is to be able to reconstruct accurate energies, as well as match the positions between the clusters and the tracked particles.

The analysis started relatively simple – with electrons. For this case, the ECAL is the main calorimeter of interest. The two quantities studied were: the difference between the transverse radial distance for the calculated track and the cluster (Δr), and the difference between the initial energy of the particle and the reconstructed energy of the cluster (ΔE). Another useful study for this phase is to get a rough estimate of the resolution of the ECAL. One can do this only for electrons with energy $<$ 1000 GeV, as it will guarantee all of the particle's energy will be deposited only in the ECAL.

In analyzing the differences between the radial distance, the algorithm follows some relatively simple steps which include:

- Calculate transverse position (x_{track} and y_{track}) for the particle's theoretical trajectory to the front face of the ECAL,
- Calculate transverse position ($x_{cluster}$ and $y_{cluster}$) for the cluster positions with,

  $$\overrightarrow{r}_{cluster} = \overrightarrow{r}_{grid} + \overrightarrow{r}_{offset}$$

  where $\overrightarrow{r}_{grid}$ is the position within the 100x100 matrix space for both x and y, and $\overrightarrow{r}_{offset}$ is how far that grid is displaced from the origin for x and y.
- Calculate Δr with the following equation,

  $$\Delta r = \sqrt{(x_{track} - x_{cluster})^2 + (y_{track} - y_{cluster})^2}$$

- Run an optimization loop over all of the clusters to determine which cluster gives the smallest Δr. Once determined, the cluster number is recorded and stored .

In determining the differences in energy, a similar process is followed.

- Calculate the generated energy the particle should produce with the scalar product of its 4-momentum vector,

  $$E_{generated} = \sqrt{\mathbf{p}^2 + m^2}$$

  where $\mathbf{p}$ is the relativistic momentum in the ith direction ($i = x, y, z$), and m is the particle's mass.

- Calculate the energy of the cluster associated with the closest particle match by looping over the cells that are in that cluster.
- Calculate ΔE with

$$\Delta E = E_{generated} - E_{cluster}$$

To determine the resolution of the ECAL, we do the following:

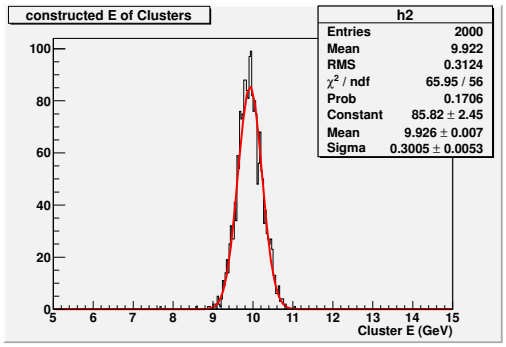
- Determine the standard deviation of the reconstructed energies from the clusters by using a Gaussian fit to the histograms of the energy distribution. Some examples of the fits are shown (Fig. 2.1).
- Create a table of data points for the resolution. The formula is σ/E so once the standard deviation is obtained, divide by the mean energy.
- Use propagation of uncertainties to determine the error bars on the graph. For our equation, σ/E, the following equation was used for the uncertainty,

$$\sigma_{res} = \sqrt{\left(\frac{1}{E}\sigma_\sigma\right)^2 + \left(\frac{\sigma}{E^2}\sigma_E\right)^2}$$

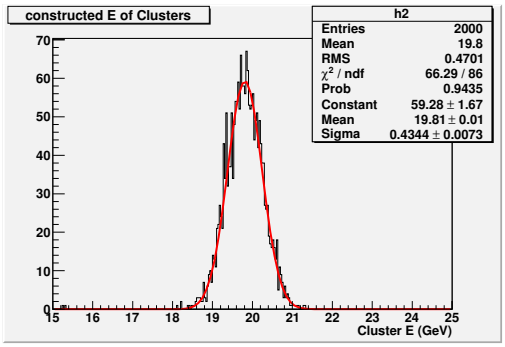
  where σ_{res} is the uncertainty in the resolution, σ_σ is the uncertainty in the standard deviation, and σ_E is the uncertainty in the mean of the energy.
- Fit the resolution graph with the proper form shown below. This is the equation the resolution is modeled for sampling calorimeters, featuring a stochastic term, a noise term, and a constant term.

$$\frac{\sigma}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c$$

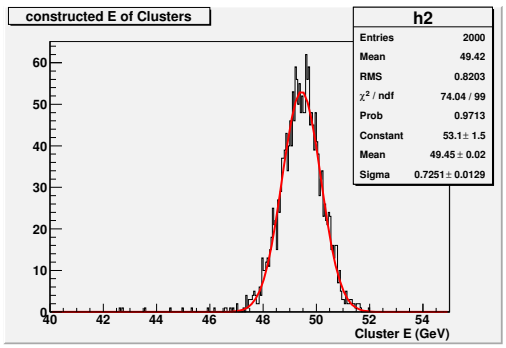
  where σ is the standard deviation of the distribution, a, b, and c are constants, and $\oplus$ refers to quadratic summing [11]. The plot is shown in the Results section.



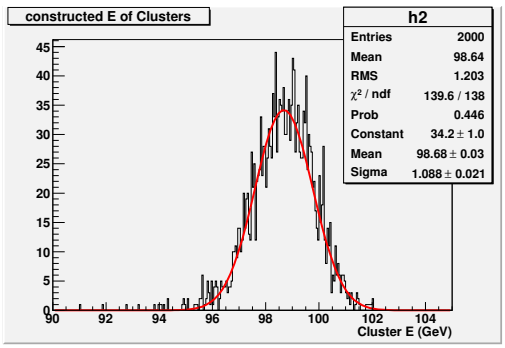
(a) $p = 10$ GeV



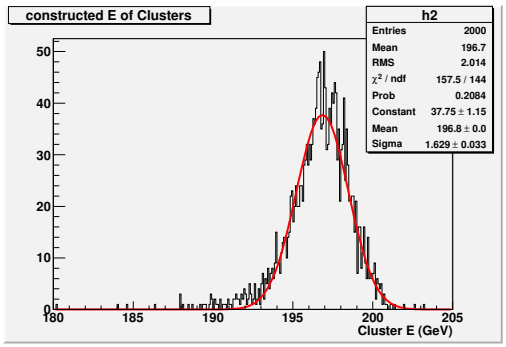
(b) $p = 20$ GeV



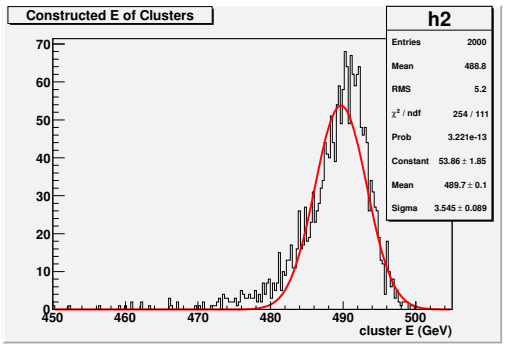
(c) $p = 50$ GeV



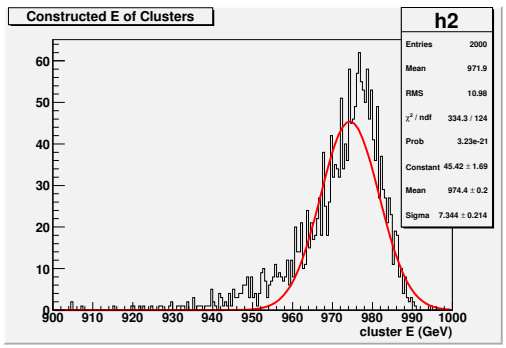
(d) $p = 100$ GeV



(e) $p = 200$ GeV



(f) $p = 500$ GeV



(g) $p = 1000$ GeV

Figure 2.1: Constructed energies of electrons at $\eta = 2.0$ with Gaussian fit

CHAPTER 3

ANALYSIS

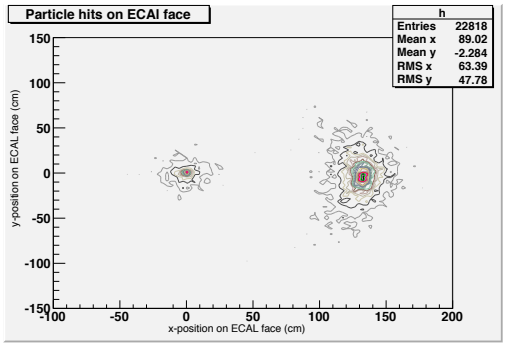
3.1 Visualization

The figures below show some plots illustrating the placement of the particles as they hit the face of the ECAL. The figures display jets with various energies at the same η, and with the same energy but with varying η. The visualizations for electrons and pions were much less interesting, only consisting of a small region of impact.

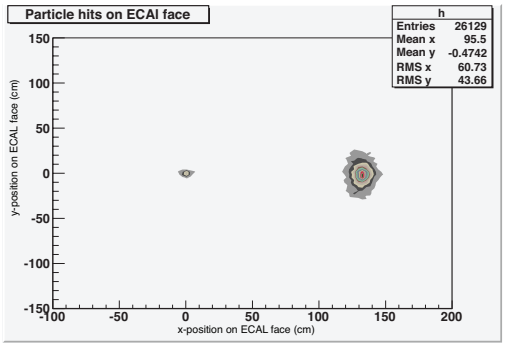
Below in figure 3.1, the plots illustrate jets oriented at an η of 1.6 with varying energies. One interesting thing that occurs is the jet splits into two bunches for lower energies, specifically from 50 - 200 GeV. For lower energy jets, the individual hadrons that form the jet have a lower momentum, resulting in a smaller radius of curvature. Once the jets have high enough energy, the curving is much less noticeable, resulting in one general impact location.

Another thing to note, for these lower energy jets, is the curving charged hadrons bend into the vicinity of the beam line. This means nothing for the simulation since it just calculates the position, but in reality, near the beam line there is no ECAL. This indicates that some energy will be lost and unaccounted for in real data sets.

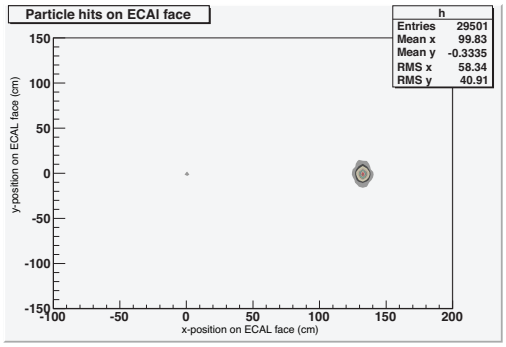
In the following plots, the contour coloring in the z direction represents the density of the number of events occurring in that region.



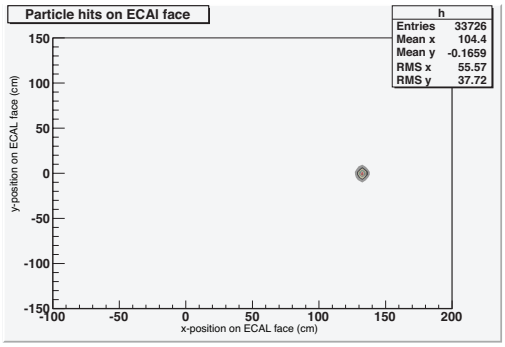
(a) $p = 50$ GeV



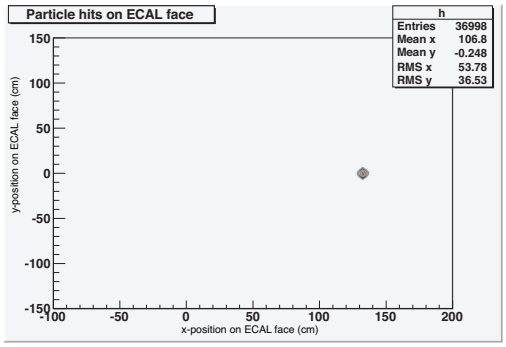
(b) $p = 100$ GeV



(c) $p = 200$ GeV



(d) $p = 500$ GeV

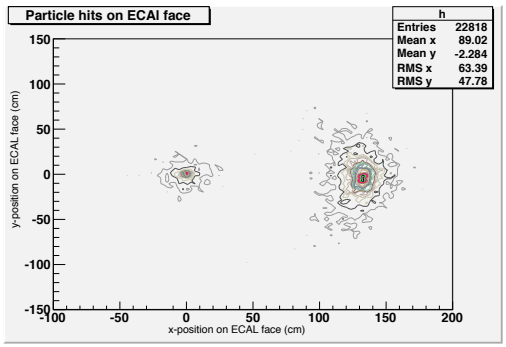


(e) $p = 1000$ GeV

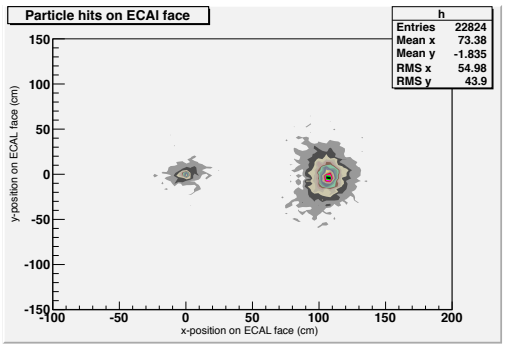
Figure 3.1: Contour plot of the position of the individual particles from jets at $\eta = 1.6$, with varying energies.

Figure 3.2 shows a 50 GeV jet for different η. We know the definition to be related with the angle above the beam line the particle hits the ECAL, with equation 1.1.

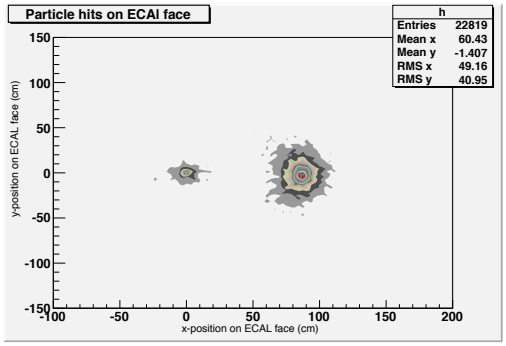
Because of this relation, as the η value gets smaller, the particle is oriented farther from the beam line. Thus, as η gets larger, the main collision area moves closer to the center where the beam line enters. This is shown in the figure well.



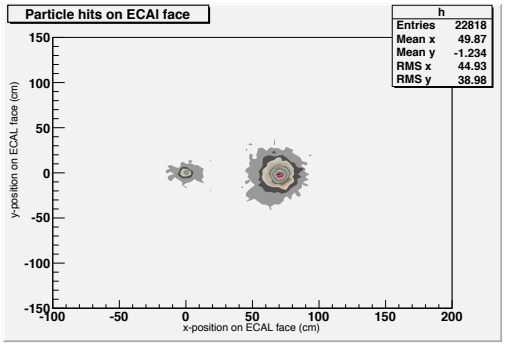
(a) $\eta = 1.6$



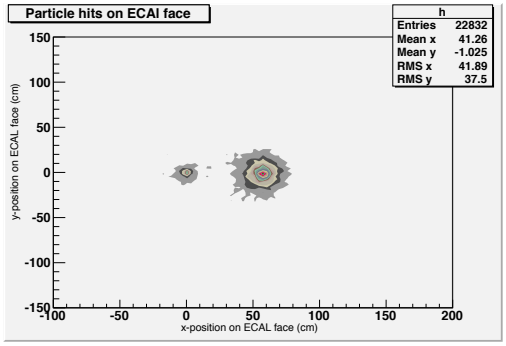
(b) $\eta = 1.8$



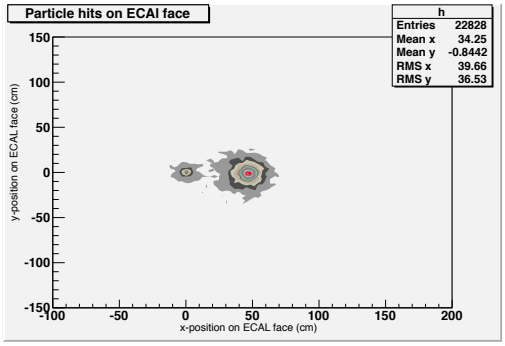
(c) $\eta = 2.0$



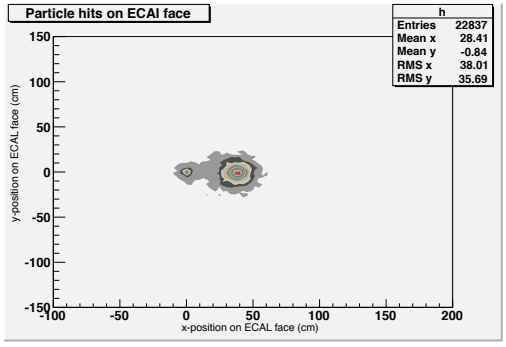
(d) $\eta = 2.2$



(e) $\eta = 2.4$



(f) $\eta = 2.6$



(g) $\eta = 2.8$

Figure 3.2: Contour plot of the particles position on the front face of the ECAL, with varying η at $p = 50$ GeV

CHAPTER 4

RESULTS

4.1 Comparisons of Clustering programs

Among the three different clustering methods, studies were done comparing the accuracy of matching to determine the best algorithm. Shown below is a good illustration of the differences for the electron case (Fig. 4.1). The Δr, which is a measure of how closely the tracks match with the clusters, varies greatly between all three clustering methods. The energy differences seem to be better for the Particle Flow clusters (the distribution is centered over zero).

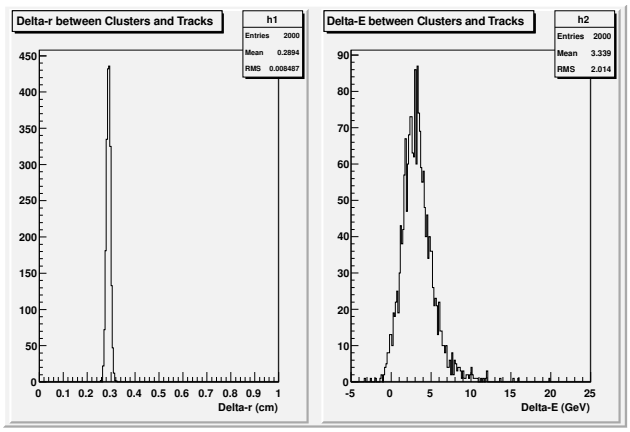
There is a trade off between the three clustering algorithms. On one hand, the linear weighted 5x5 has an average Δr of about 0.29 cm, as the log weighted 5x5 has an average Δr of 0.57 cm, both with a very narrow peak; conversely, the Particle Flow clusters have an average Δr of about 0.84 cm but with a more widespread distribution. To investigate, consider the cases when the same graphs vary with different energies, and η.

4.1.1 Linear Weighted 5x5:

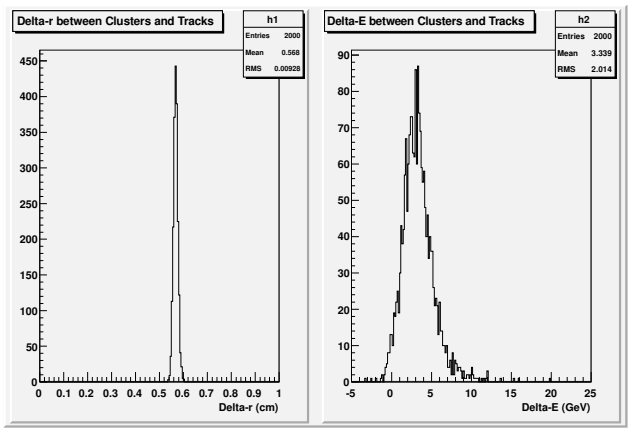
For increasing energy at constant η (Fig. 4.2), the energy differences and Δr seem to correlate inversely: with increasing energy, Δr narrows as ΔE widens.

For a really low energy, such as 10 GeV, the electron will curve a lot more than for a higher energy electron. This curving will produce more uncertainty in the Δr between the particle tracks and clusters because the clusters are based off of the energy depositions – an electron with a large radius of curvature that hits one cell, will leak into a neighboring cell because of the steep angle of incidence. On the other hand, since the electron is going so slow, essentially all of its energy will be deposited in the ECAL scintillating crystals. This is because there are enough interaction lengths for the particle to interact almost 100% of the time, thus almost all of its energy can be reconstructed.

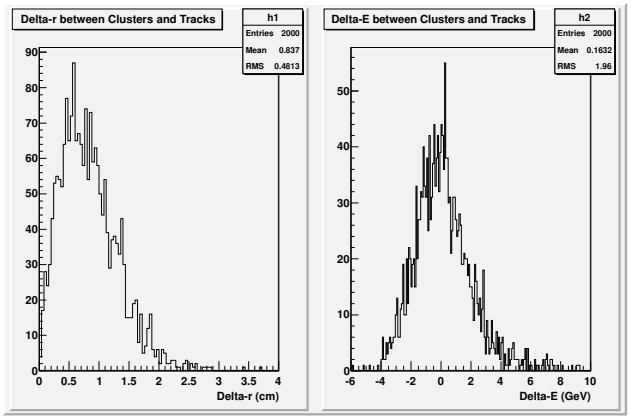
Conversely, for faster particles, like 500 and 1000 GeV, the opposite is true. The curving will be much less prevalent, resulting in better certainty in matching the clusters with the tracks; however,



(a) 5x5 Linear weight



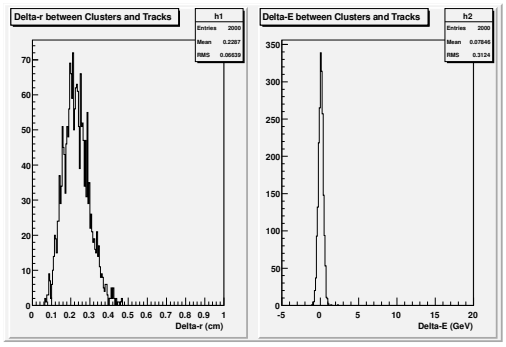
(b) 5x5 Log weight



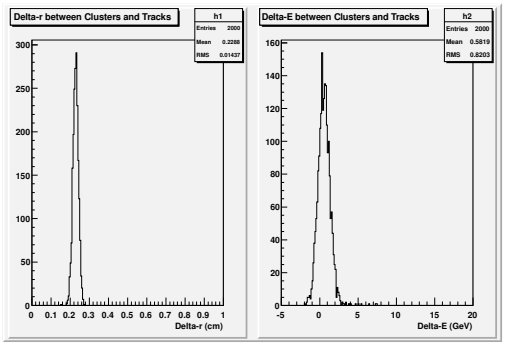
(c) Particle Flow

Figure 4.1: The three different clustering methods compared for a 200 GeV electron at an $\eta = 2.0$. For each algorithm, the left plot shows Δr while the right shows ΔE.

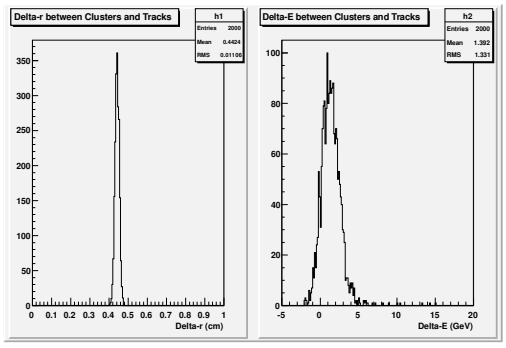
now that the particle is moving faster, it will travel through the scintillating and absorber parts of the detector much quicker. This explains the shape of the energy distribution. The fractional uncertainty of the energy differences show that it is relatively constant across all energies. For example, an electron with 1000 GeV of momentum will have a larger difference than one at 10 GeV. If the ΔE is around 3%, then it would be around 30 GeV, and 0.3 GeV respectively.



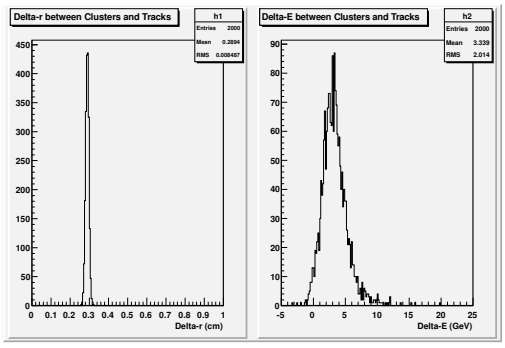
(a) $p = 10$ GeV



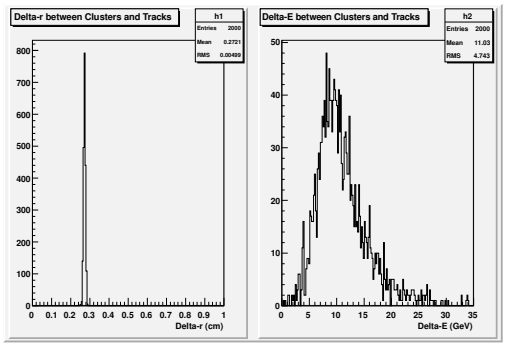
(b) $p = 50$ GeV



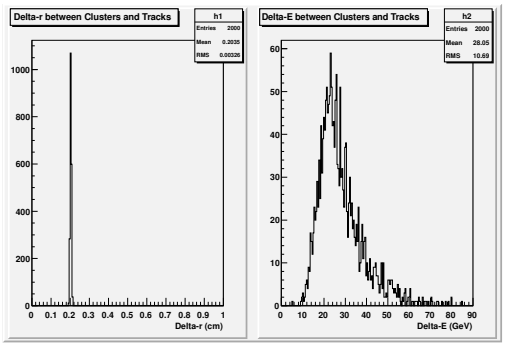
(c) $p = 100$ GeV



(d) $p = 200$ GeV



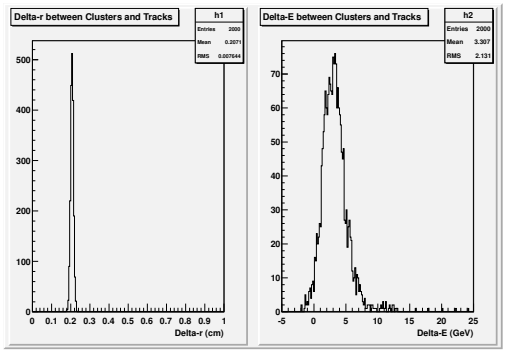
(e) $p = 500$ GeV



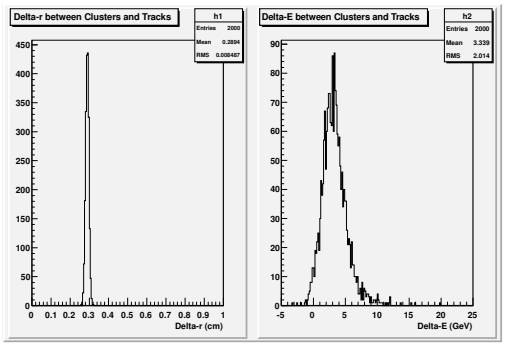
(f) $p = 1000$ GeV

Figure 4.2: 5x5 linear weighted clustering method for electrons at $\eta = 2.0$, with varying energy.

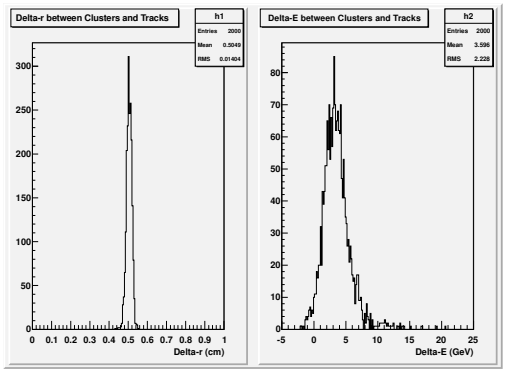
When the energy is held constant, and η is allowed to vary (Fig. 4.3), the Δr seems to vary with a constant term. Since the energy is constant, the width of the Δr will be the same throughout, so the variation, from 0.2 to 0.5 cm, has to deal with the angle the electron is hitting the crystal cells.



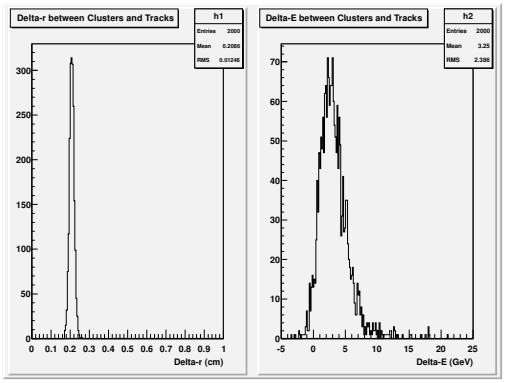
(a) $\eta = 1.8$



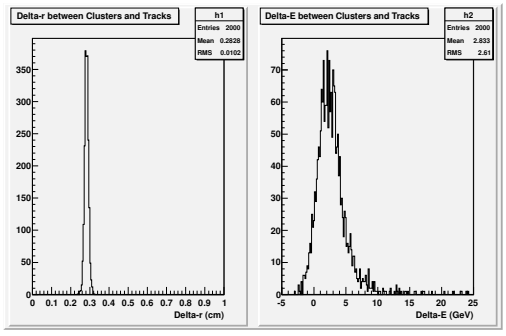
(b) $\eta = 2.0$



(c) $\eta = 2.2$



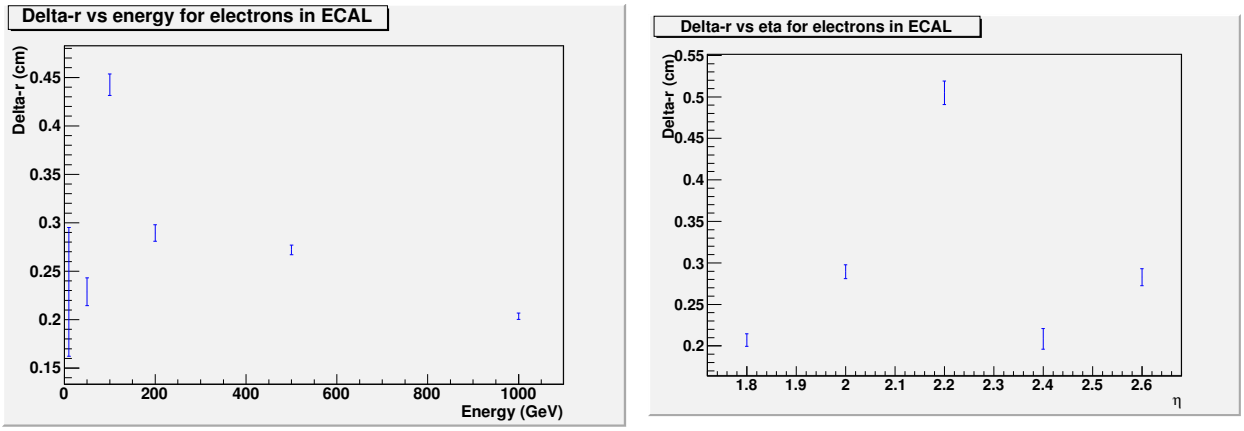
(d) $\eta = 2.4$



(e) $\eta = 2.6$

Figure 4.3: 5x5 linear weighted clustering method for electrons at $p = 200$ GeV, with varying η.

Figure 4.4 shows how the linear weighted 5x5 clustering method's Δr varies over the different energies and η. As seen in the figure, peaks occur at 100 GeV and $\eta = 2.2$. Since the uncertainties are relatively small for Δr, this implies a significant anomaly in this regions for the clustering method. However, since the electron is hitting in the same position each time, each event will result in a similar Δr, resulting in a smaller than expected spread.



(a) Δr as a function of energy.

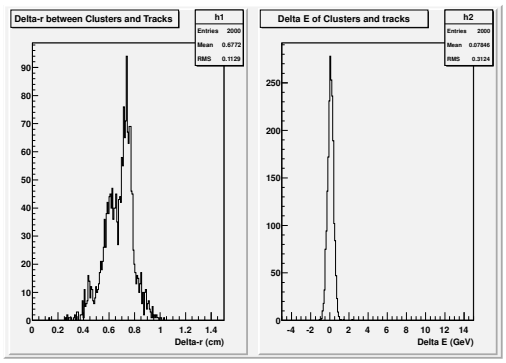
(b) Δr as a function of η.

Figure 4.4: Δr as a function of energy at $\eta = 2.0$ (left) and η at $E = 200$ GeV (right) for the linear weighted 5x5 clustering algorithm. The uncertainties are determined from the RMS, or the width of the distributions.

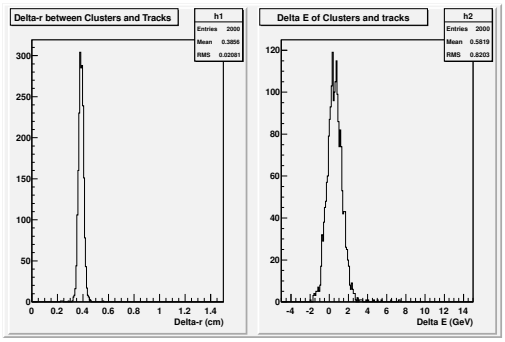
4.1.2 Log Weighted 5x5:

Now looking at the log weighted 5x5 clustering for varying energy at constant η (Fig. 4.5), there are similarities between this and the linear weighted clusters. The same patterns occur: for low energy, Δr is more widespread, with energy narrow, and as energy increases this trends move in reverse (Δr narrows as ΔE widens); and Δr reaches its highest value at 100 GeV. Another interesting observation is that the Δr average values seem to all be greater than the linear weighted clusters by a constant value – mostly greater by around 0.3 cm, with two cases (50 and 1000 GeV) where it is only 0.1 cm greater.

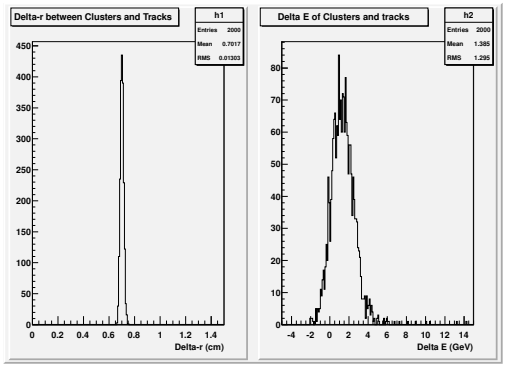
When energy is held constant, and η varies, the same patterns continue again (Fig. 4.6). The Δr shoots up to its highest value at $\eta = 2.2$. The average Δr value jumps around widely like above for the linear weighted clusters, with values ranging from 0.3 cm to 1.2 cm. This again is most likely due to the different angles of incidence.



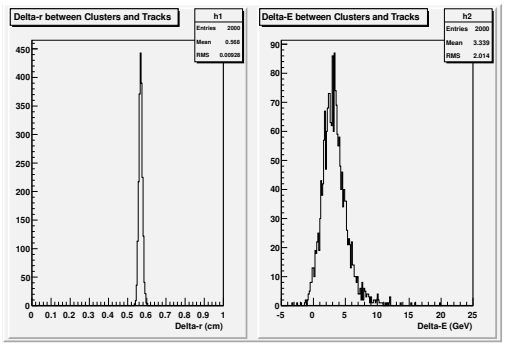
(a) $p = 10$ GeV



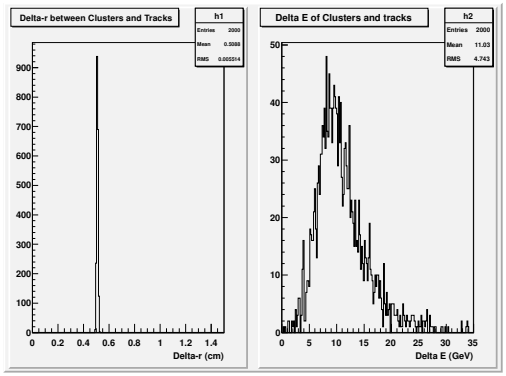
(b) $p = 50$ GeV



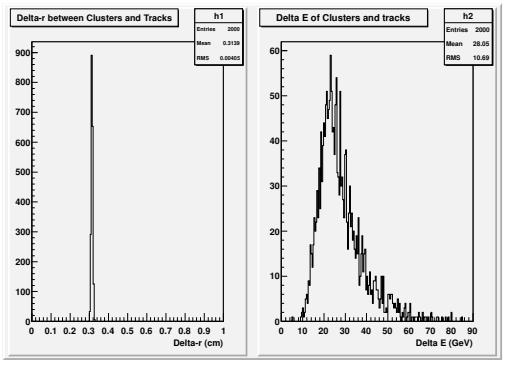
(c) $p = 100$ GeV



(d) $p = 200$ GeV

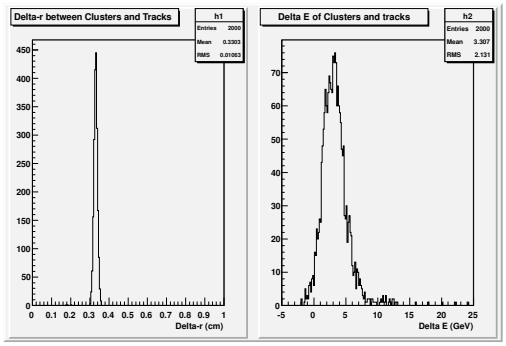


(e) $p = 500$ GeV

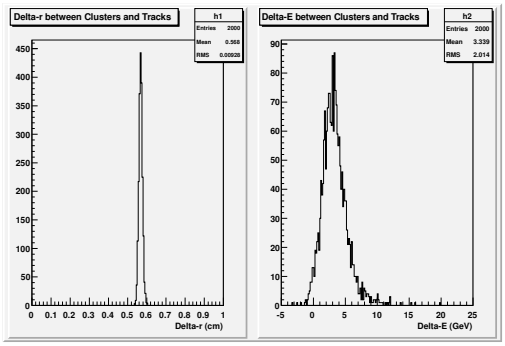


(f) $p = 1000$ GeV

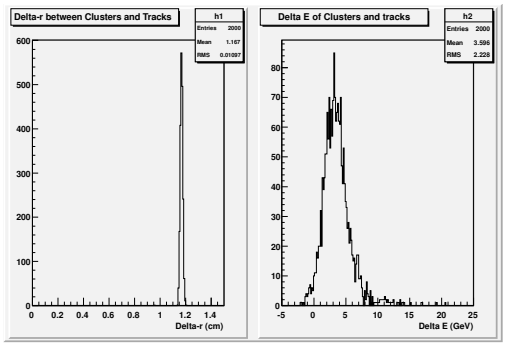
Figure 4.5: 5x5 log weighted clustering method for electrons at $\eta = 2.0$, with varying energy.



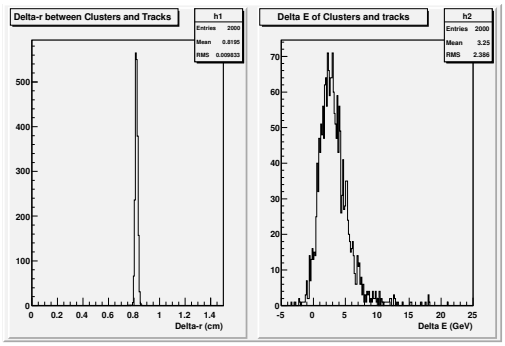
(a) $\eta = 1.8$



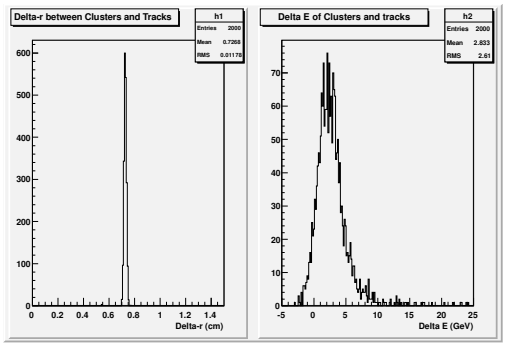
(b) $\eta = 2.0$



(c) $\eta = 2.2$



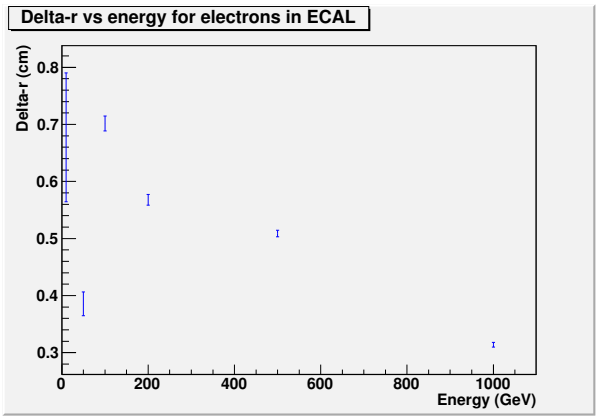
(d) $\eta = 2.4$



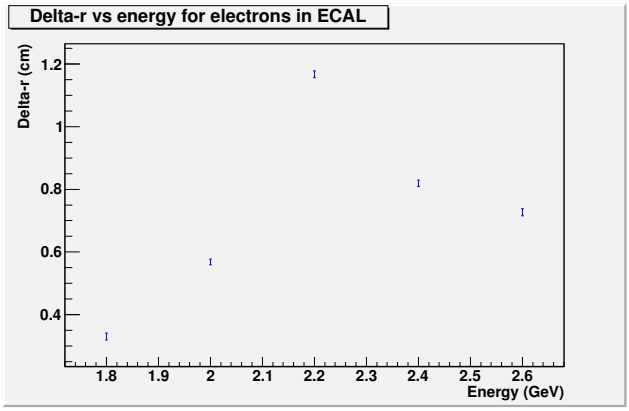
(e) $\eta = 2.6$

Figure 4.6: 5x5 log weighted clustering method for electrons at $p = 200$ GeV, with varying η.

Again, like the linear weighted clusters, there seem to be anomalies because of the small uncertainties in Δr (Fig. 4.7). This peculiarity will be discussed later on, after the case for the Particle Flow clusters is shown.



(a) Δr as a function of energy.



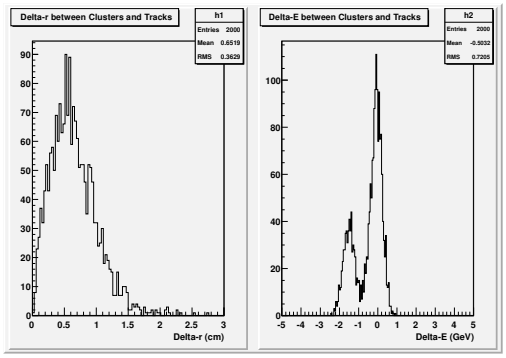
(b) Δr as a function of η.

Figure 4.7: Δr as a function of energy at $\eta = 2.0$ (left) and η at $E = 200$ GeV (right) for the log weighted 5x5 clustering algorithm. The uncertainties are determined from the RMS, or the width of the distributions.

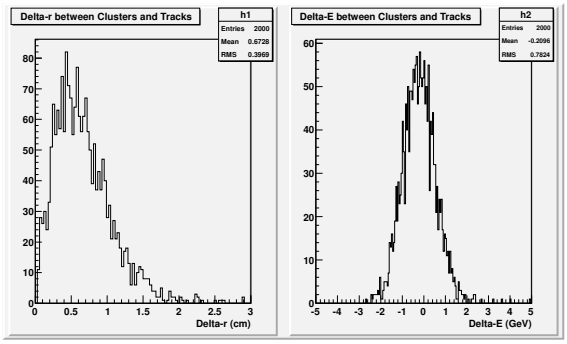
4.1.3 Particle Flow Clusters:

As opposed to the Linear weighted 5x5 clusters, the Particle Flow cluster distributions are more consistent across different energies at constant $\eta = 2.0$ (Fig. 4.8). Although the Δr distributions are more widespread than the linear weighted clusters, an interesting observation is that most of the peak is contained within the 1.4 cm that make up the crystal cell. There is a quick fall off outside of this region. The average Δr steadily increases with the energy, from about 0.65 cm to 1.5 cm. The actual peak of the Δr distributions seem to correlate with those of the log weighted 5x5 clusters.

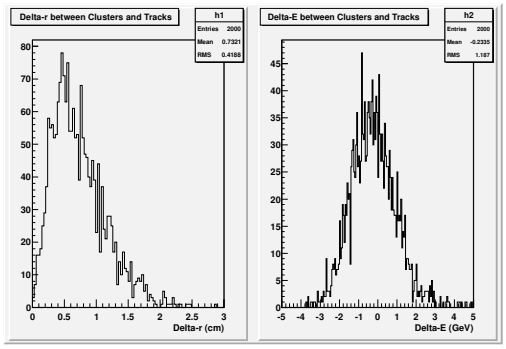
When varying the η with constant energy (200 GeV), the Δr is very consistent (Fig 4.9). The average Δr values vary between 0.83 and 0.85 cm.



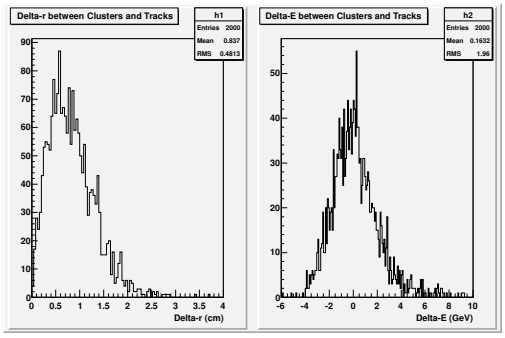
(a) $p = 10$ GeV



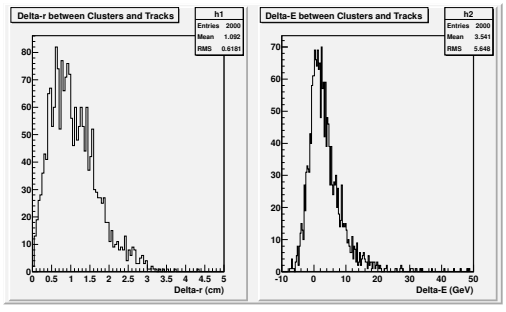
(b) $p = 50$ GeV



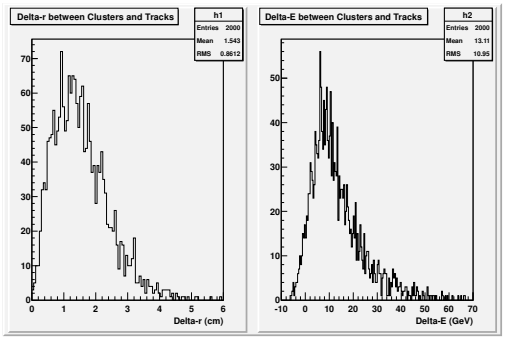
(c) $p = 100$ GeV



(d) $p = 200$ GeV

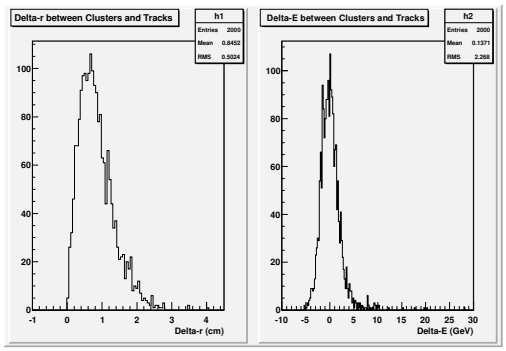


(e) $p = 500$ GeV

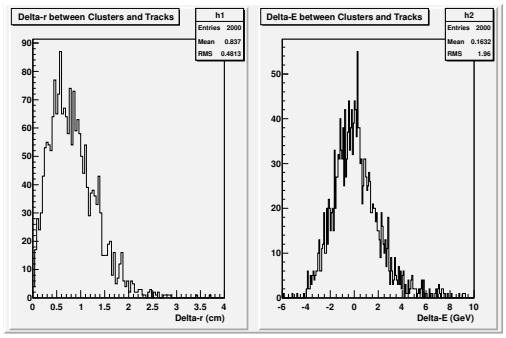


(f) $p = 1000$ GeV

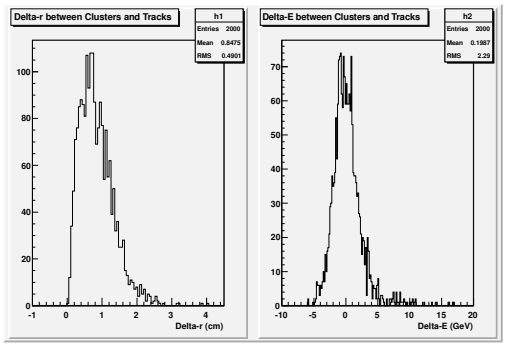
Figure 4.8: Particle Flow clustering method for electrons at $\eta = 2.0$, with varying energy.



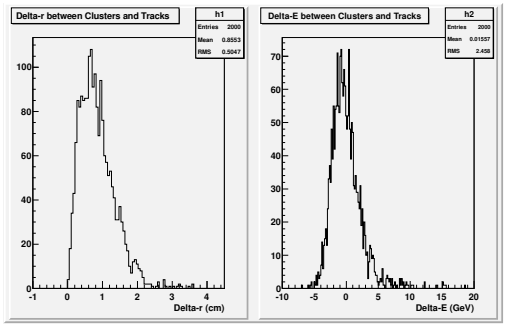
(a) $\eta = 1.8$



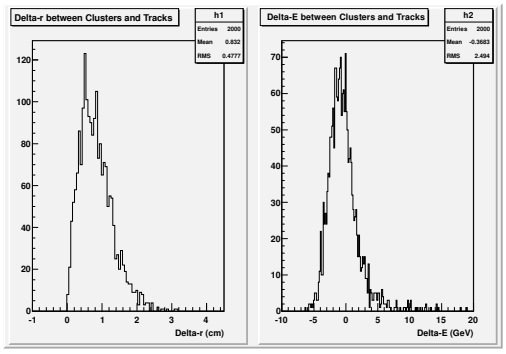
(b) $\eta = 2.0$



(c) $\eta = 2.2$



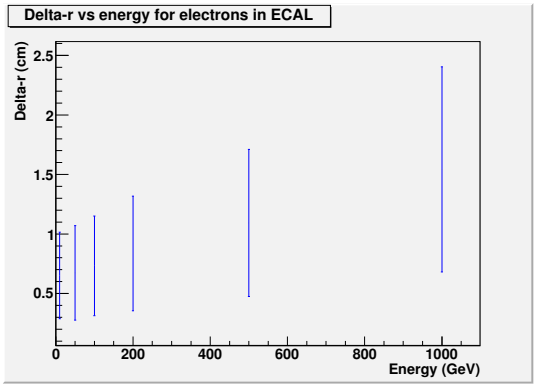
(d) $\eta = 2.4$



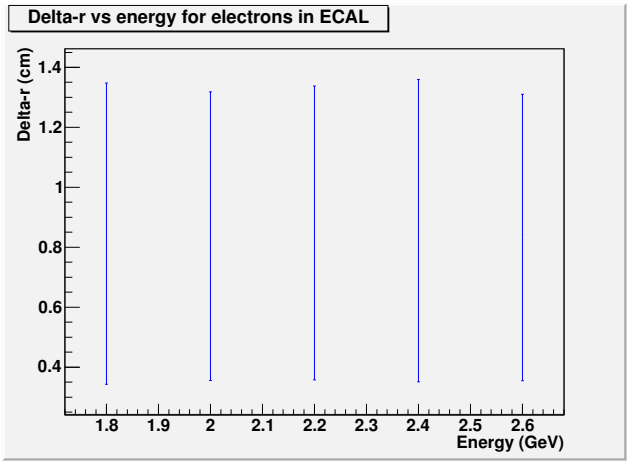
(e) $\eta = 2.6$

Figure 4.9: Particle Flow clustering method for electrons at $p = 200$ GeV, with varying η.

Figure 4.10 shows how the Particle Flow clustering method's Δr varies over the different energies and η. As the figure illustrates, the Particle Flow clustering method is much more consistent, showing no anomalies like the 5x5 clusters do.



(a) Δr as a function of energy.



(b) Δr as a function of η.

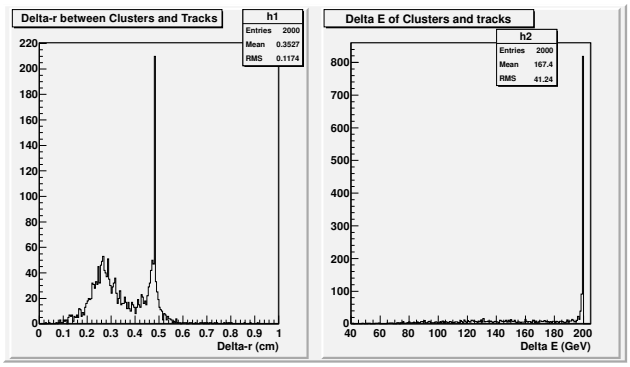
Figure 4.10: Δr as a function of energy at $\eta = 2.0$ (left) and η at $E = 200$ GeV (right) for the Particle Flow clustering algorithm. The uncertainties are determined from the RMS, or the width of the distributions.

4.1.4 Comparisons:

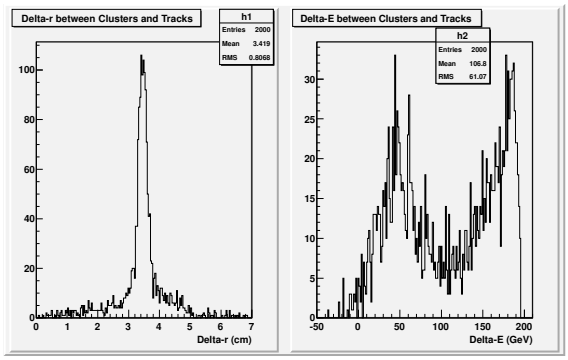
Although the linear weighted 5x5 clusters proved to have the smallest average Δr values among the three clustering methods, the shape and position of the peak raise some questions. This seems to persist within the log weighted clusters as well. There are several possibilities that can be investigated to explore these questions. There could be a bias due to systematic error when calculating the position of the cluster's centroid, possibly because the showering in the crystal has a varying depth in which the interactions begin. There could also be a bias from the positioning of the incident electron, as the simulations do not randomize the direction the electron travels and thus the position on the front face of the ECAL when the electron begins to shower.

4.1.5 Pions in ECAL and HCAL:

The figures above give good insight into the efficiency of the methods used for electrons. Figure 4.11 illustrates an example of a meson.



(a) π^- in the ECAL



(b) π^- in the HCAL

Figure 4.11: Δr and ΔE (left to right) for a π^- in both the ECAL and HCAL.

4.2 Resolution

The electrons can be used to estimate the resolution of the ECAL. Using the techniques described in the Methods section, a fit of the appropriate form was completed with the limited data available from simulations (Fig. 4.12).

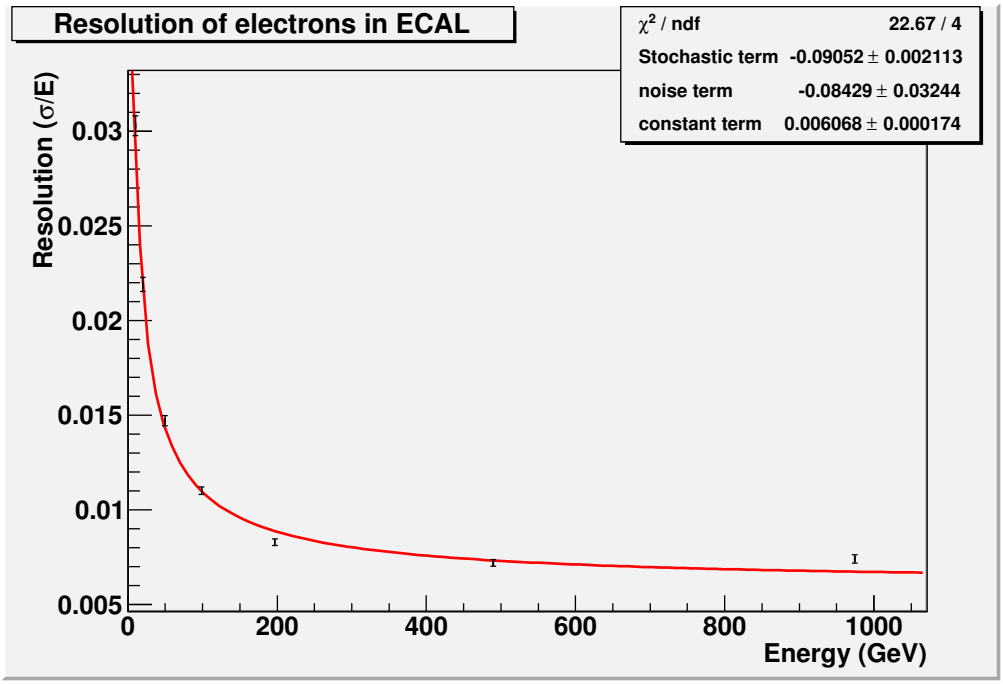


Figure 4.12: Resolution of the ECAL with electrons only

From this graph and the fitting, one can determine the coefficients for the resolution form of the detector, this is shown below:

$$\frac{\sigma}{E} = \frac{9.05\%}{\sqrt{E}} \oplus \frac{8.43\%}{E} \oplus 0.61\% \tag{4.1}$$

CHAPTER 5

CONCLUSION

The main purpose of this thesis was to investigate one of three possible electromagnetic calorimeter upgrades for the CMS detector, the Shashlik calorimeter. Simulations were run with the new ECAL in place at the end-cap of the detector. My contribution to the analysis and design studies of this new calorimeter were in the form of studying a Particle Flow algorithm, which involves forming energy deposition clusters and correlating each cluster to a specific particle in the event.

Three different clustering techniques were studied. All three clustering algorithms show merits and disadvantages in their methods when implemented; however, before a conclusion can be made on the better clustering technique, additional studies between the three clustering algorithms will be needed to determine the optimal one. Matching criteria for electrons were studied and the results presented. The implementation with pions and jets in both calorimeters will need to be refined to merge the ECAL and HCAL information for the more complicated hadronic showers.

The next step to be carried out for the Particle Flow algorithm would be to redesign the matching portion of it, replacing it with a more sophisticated matching technique and additional selection criteria. Further analysis would study the effectiveness of the Particle Flow algorithm for the case of large pileup.

Using electrons with no pileup, the ECAL resolution is measured to be,

$$\frac{\sigma}{E} = \frac{9.05\%}{\sqrt{E}} \oplus \frac{8.43\%}{E} \oplus 0.61\% \tag{5.1}$$

with the average energy resolution for electrons greater than 200 GeV around 6-8 %. It will be important to measure this for an effective Particle Flow algorithm in the environment of high pileup.

APPENDIX A

DERIVATION OF EQUATIONS OF MOTION

A.1 Charged Particles

For the Particle Flow Algorithm, you need a reliable way to calculate the trajectory of the multitude of particles that go through the detector. In our model, we approximate the magnetic field inside the solenoid to be constant at around 3.84 Teslas, in the $\hat{\mathbf{z}}$ direction. In addition, we assume that there is no internal electric field interfering with the moving particles. With these assumptions, we start with the Lorentz force law,

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$$

Which now becomes the simpler,

$$\mathbf{F} = q(\mathbf{v} \times \mathbf{B}) \tag{A.1}$$

where $\mathbf{F}$ is the magnetic force acting on a particle, q is the particle's charge, $\mathbf{v}$ is the particle's velocity, and $\mathbf{B}$ is the magnetic field strength. But Newton's second law says, $\mathbf{F} = \mathrm{m}\mathbf{a}$, and when generalizing to relativistic effects we then get the following

$$\gamma m \frac{d^2\mathbf{r}}{dt^2} = q\left(\frac{d\mathbf{r}}{dt} \times \mathbf{B}\right)$$

$$\frac{d^2\mathbf{r}}{dt^2} = \frac{q}{\gamma m}\begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ \frac{dx}{dt} & \frac{dy}{dt} & \frac{dz}{dt} \\ 0 & 0 & B_0 \end{vmatrix}$$

Which separates into three coupled differential equations:

$$\frac{d^2x}{dt^2} = \omega\frac{dy}{dt}; \qquad \frac{d^2y}{dt^2} = -\omega\frac{dx}{dt}; \qquad \frac{d^2z}{dt^2} = 0 \tag{A.2}$$

where the substitution $\omega \equiv \frac{qB_0}{m\gamma}$ has been made, with ω being the cyclotron frequency.

We know how to solve the equation of motion for z, it is trivial. The challenge is with the coupled differential equations. You can actually uncouple them easily with the following substitution: Take some complex number,

$$u = x + iy$$

then,

$$\frac{du}{dt} = \frac{dx}{dt} + i\frac{dy}{dt} \tag{A.3}$$

and,

$$\frac{d^2u}{dt^2} = \frac{d^2x}{dt^2} + i\frac{d^2y}{dt^2} \tag{A.4}$$

Now substitute A.2 into A.4:

$$\begin{aligned}
\frac{d^2u}{dt^2} &= \left(\omega\frac{dy}{dt}\right) + i\left(-\omega\frac{dx}{dt}\right) \\
&= \omega\left(\frac{dy}{dt} - i\frac{dx}{dt}\right) \\
&= -i\omega\left(\frac{dx}{dt} - \frac{1}{i}\frac{dy}{dt}\right) \\
&= -i\omega\underbrace{\left(\frac{dx}{dt} + i\frac{dy}{dt}\right)}_{\frac{du}{dt}} \\
\frac{d^2u}{dt^2} &= -i\omega\frac{du}{dt}
\end{aligned} \tag{A.5}$$

Take another substitution,

$$\text{Let} \quad \xi \equiv \frac{du}{dt}$$

then A.5 becomes,

$$\begin{aligned}
\frac{d\xi}{dt} &= -i\omega\xi \\
\xi(t) &= c_1 e^{-i\omega t} = \frac{du}{dt} \\
u(t) &= c_1\frac{i}{\omega}e^{-i\omega t} = x(t) + iy(t)
\end{aligned} \tag{A.6}$$

where c_1 is a constant independent of time. We can't quite solve it yet. However, one small manipulation makes it possible. Consider the original substitution for u and its complex conjugate.

$$u = x + iy \qquad \text{and} \qquad \overline{u} = x - iy$$

You can solve this system of equations for x and y, giving the following relations:

$$x = \frac{(u + \overline{u})}{2} \qquad \text{and} \qquad y = \frac{(u - \overline{u})}{2i} \tag{A.7}$$

we then get,

$$
\begin{aligned}
x(t) &= c_1 \frac{i}{2\omega}\left(e^{-i\omega t} - e^{i\omega t}\right) \\
&= \frac{c_1}{\omega}\frac{1}{2i}\left(e^{i\omega t} - e^{-i\omega t}\right) \\
x(t) &= \frac{c_1}{\omega} sin(\omega t)
\end{aligned} \tag{A.8}
$$

and similarly,

$$
y(t) = \frac{c_1}{\omega} cos(\omega t) \tag{A.9}
$$

We still need to determine the integration constant, c_1. If you look back at A.6, you see that the original coefficient is related to u', or the velocities in the x and y directions. So it should be proportional to the initial momentum of the particle.

In addition, we can make the equations A.8 and A.9 a little more general, accounting for possible phase shifts and differences. Adding in these considerations, as well as initial conditions regarding position and momentum, we can construct a complete solution [10]:

$$
\begin{aligned}
x(t) &= x_0 + \frac{v_{y0}}{\omega} - \frac{v_{y0}}{\omega} cos(\omega t) + \frac{v_{x0}}{\omega} sin(\omega t) \\
y(t) &= y_0 - \frac{v_{x0}}{\omega} + \frac{v_{x0}}{\omega} cos(\omega t) + \frac{v_{y0}}{\omega} sin(\omega t)
\end{aligned}
$$

Recall that,

$$
\omega = \frac{qB_0}{m\gamma}
$$

then we have,

$$
\begin{aligned}
x(t) &= x_0 + \frac{\gamma m v_{y0}}{qB_0} - \frac{\gamma m v_{y0}}{qB_0} cos(\omega t) + \frac{\gamma m v_{x0}}{qB_0} sin(\omega t) \\
y(t) &= y_0 - \frac{\gamma m v_{x0}}{qB_0} + \frac{\gamma m v_{x0}}{qB_0} cos(\omega t) + \frac{\gamma m v_{y0}}{qB_0} sin(\omega t)
\end{aligned}
$$

Notice the relativistic momentum $\gamma m v$ in each term. Therefore,

$$
\boxed{x(t) = x_0 + \frac{p_{y0}}{qB_0} - \frac{p_{y0}}{qB_0} cos\left(\frac{qB_0}{\gamma m} t\right) + \frac{p_{x0}}{qB_0} sin\left(\frac{qB_0}{\gamma m} t\right)} \tag{A.10}
$$

$$
\boxed{y(t) = y_0 - \frac{p_{x0}}{qB_0} + \frac{p_{x0}}{qB_0} cos\left(\frac{qB_0}{\gamma m} t\right) + \frac{p_{y0}}{qB_0} sin\left(\frac{qB_0}{\gamma m} t\right)} \tag{A.11}
$$

$$
\boxed{z(t) = z_0 + v_{z0} t} \tag{A.12}
$$

where $p_i \equiv \gamma m v_i$ is the relativistic momentum.

In all simulations used in this study, the vertex, or initial positions, of the particle is at the origin of our coordinate system. So x_0, y_0, and z_0 will all be zero for this simulation. This simplifies the coefficients we need to solve for. We can then solve for the time it takes along the z direction, since the distance is a fixed value (from vertex to front face of electromagnetic calorimeter).

$$z = \cancel{z_0} + v_{z0}t$$
$$d = \frac{p_{z0}}{m}t$$

$$\boxed{t = \frac{m}{p_{z0}}d} \tag{A.13}$$

where d is the distance from the vertex to the front face of the detector (315.4 cm), and we are provided with the mass and initial momentum values of each particle. Using this, we can then solve for the x and y positions. Equations A.10 and A.11 then become,

$$\boxed{x(t) = \frac{p_{y0}}{qB_0} - \frac{p_{y0}}{qB_0}cos\left(\frac{qB_0}{\gamma p_{z0}}d\right) + \frac{p_{x0}}{qB_0}sin\left(\frac{qB_0}{\gamma p_{z0}}d\right)} \tag{A.14}$$

$$\boxed{y(t) = -\frac{p_{x0}}{qB_0} + \frac{p_{x0}}{qB_0}cos\left(\frac{qB_0}{\gamma p_{z0}}d\right) + \frac{p_{y0}}{qB_0}sin\left(\frac{qB_0}{\gamma p_{z0}}d\right)} \tag{A.15}$$

A.2 Neutral Particles

The equations of motion for neutral particles is much simpler, as we do not have the Lorentz force affecting the trajectory. Using the same assumption of no electric field inside the solenoid, the particles follow a straight path.

$$\boxed{x(t) = x_0 + v_{x0}t} \tag{A.16}$$

$$\boxed{y(t) = y_0 + v_{y0}t} \tag{A.17}$$

$$\boxed{z(t) = z_0 + v_{z0}t} \tag{A.18}$$

Again, the vertex of the initial position of the particles are always zero in our simulations. As before, we can calculate the time it takes to travel to the front face.

$$\boxed{t = \frac{m}{p_{z0}}d} \tag{A.19}$$

With this, we can get the x and y positions for neutral particles.

$$x = v_{x0}\left(\frac{m}{p_{z0}}d\right)$$
$$y = v_{y0}\left(\frac{m}{p_{z0}}d\right)$$

which become,

$$\boxed{x = \left(\frac{p_{x0}}{p_{z0}}\right)d} \tag{A.20}$$

$$\boxed{y = \left(\frac{p_{y0}}{p_{z0}}\right)d} \tag{A.21}$$

BIBLIOGRAPHY

[1] The CMS Collaboration. *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*. Physics Letters B [online]. 2013. Available from: ≪http://arxiv.org/pdf/1207.7235v2.pdf≫.

[2] MolTech *GmbH*, Molecular Technology. *Scintillation Crystals and its General Characteristics* [online]. Available from: ≪http://www.mt-berlin.com/frames_cryst/descriptions/scintillators_gen%20.htm≫.

[3] Fabjan, C.W., Gianotti, F., 2003. *Calorimetry for Particle Physics*. Reviews of Modern Physics, Volume 75, Oct. 2003 [online]. Available from: ≪http://journals.aps.org/rmp/pdf/10.1103/RevModPhys.75.1243≫.

[4] Taylor, L., 2011. *Hadron Calorimeter* [online]. Available from: ≪http://cms.web.cern.ch/news/hadron-calorimeter≫ .

[5] Schopper, A., 2000. *LHCb ECAL Overview* [online]. Available from: ≪http://lhcb-calo.web.cern.ch/lhcb-calo/html/TDR/calo_tdr/node48.html≫.

[6] Omega Piezo Technologies. 2008. *BGO, LYSO and GSO Crystal Scintillators* [online]. Available from: ≪http://www.omegapiezo.com/crystal_scintillators.html≫.

[7] *GEANT4*. Available from: ≪http://geant4.cern.ch≫.

[8] Sjstrand, T., Christiansen, J., Desai, N., Ilten, P., Mrenna, S., Prestel, S., Skands, P., 2001. *Pythia*. Available from: ≪http://home.thep.lu.se/~torbjorn/Pythia.html≫.

[9] Garren, L., Knowles, I.G., Sjstrand, T., Trippe, T., 2002. *Monte Carlo Particle Numbering Scheme*. Available from: ≪http://pdg.lbl.gov/mc_particle_id_contents.html≫.

[10] Dilo, R., Alves-Pires, R., 1997. *Nonlinear Dynamics in Particle Detectors*. Singapore: World Scientific, 1997. Available from: ≪http://www.worldscientific.com/doi/pdf/10.1142/9789812798657_bmatter≫.

[11] The CMS Collaboration. *Energy Calibration and Resolution of the CMS Electromagnetic Calorimeter in pp collisions at* $\sqrt{s} = 7\ TeV$. Journal of Instrumentation [online]. 2013 [cited 2013 Oct 18]. Available from: ≪http://arxiv.org/abs/1306.2016≫.

BIOGRAPHICAL SKETCH

After graduation, I am signed up to go to Officer Candidate School (OCS) with the U.S. Navy in early June. I was accepted into a college program about a year ago with the Navy. It requires a five to six year service obligation, including about a year and a half of training. The training involves me becoming a naval officer, then learning skills to successfully operate a nuclear reactor onboard submarines.

Once I am finished with my military experience, I plan on going to graduate school for physics. Possible fields that interest me include: high-energy particle physics (experimental or theoretical), theoretical studies of new physics (such as supersymmetry, string theory, etc.), general relativity, and anything else that may strike my interest along the way.

After graduate school, I have several different career paths thought out, though which one I will choose is not yet decided. They include: become a quantitative analyst for a big investment firm/hedge fund on wall street or something similar, work at national and government research labs, go down the academia route (post-doc, associate professor, etc. until tenured), or work with government agencies (NASA, CIA, etc.).