GlueX–00–NNN
06. December 2005

The GlueX BCAL Reconstruction Code - Preliminary Studies

C. Xu, M. Barbi, G. Huber

University of Regina

3737 Wascana Parkway, Department of Physics,
Regina, S4S-0A2, Canada

Abstract

Preliminary efforts to develop the GlueX BCAL reconstruction code are presented. The code is based on the software package developed for the KLOE electromagnetic calorimeter which was consistently modified to fit into the BCAL requirements and overall GlueX detector environment. Results on energy and spatial position reconstruction are shown using single-γ, single-π^0 and $\gamma p \to p\pi^+\pi^-\pi^0$ data generated with the Hall-D Monte Carlo simulation software.

1 Introduction

The Barrel Calorimeter (BCAL) is one of the main components of the GlueX detector [1]. It will be positioned inside the solenoid which constrains its outer radius to be 90 cm. On the other hand, the inner radius is constrained to be 65 cm in order to allow the installation of the Cylindrical and Forward Drift Chambers, Start Counters and target. The BCAL is divided into 48 modules in ϕ direction. The module structure consists of a stack where alternating layers of scintillating fibers of 1 mm diameter are glued inside thin grooved lead layers of 0.5 mm thickness, and it is divided into two sections: the inner and outer. The inner (outer) section is further divided into 4 (3) sub-sections as depicted in fig. 1, representing the smallest subdivisions of the BCAL and therefore defing its segmentation. Each of these subdivisions is called a readout cell, and each cell is covered by two photon-detectors, one instrumenting the side A and the other the side B of the BCAL. Each BCAL module will have 64 readout channels, 40 in the inner and 24 in the outer layers, leading to a total of 3072 readout channels for the whole detector. Neutral and/or charged particles created in the photoproduction γp reaction will enter into the BCAL, producing showers inside the lead material and leaving some energy in the scintillating fibers. This energy will excite the scintillator producing light that will propagate along the fibers and will be read out on both BCAL sides, providing a signal for time and energy reconstruction.

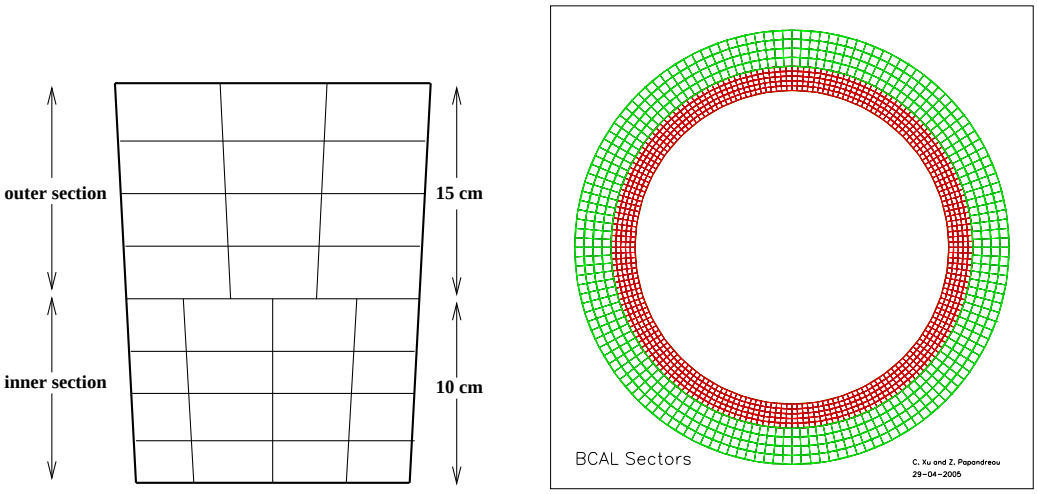

Figure 1: *The left figure shows the segmentation of each BCAL module. The right figure depicts the whole calorimeter.*

The main goal of the GlueX BCAL is to detect and measure the four-momentum of photons coming primarily from π^0's and η's decays. The purpose of the BCAL reconstruction code is, therefore, to find and associate the full collection of connected readout cells which

are related to a particle. This full collection of associated cells is called "cluster". The energy and angular direction of a particle is obtained from the cluster information.

This report is organized into four main sections. Section 2 presents the KLOE electromagnetic BCAL [2] reconstruction code. The GlueX BCAL geometrical definitions as used to interface the HallD MC simulation (HDGEANT) [3]) with the BCAL reconstruction code are discussed in section 3. Section 4 illustrutes the performance of the reconstruction code using MC simulation data. Finally, a brief discussion on future plans is presented in section 5.

2 The KLOE Reconstruction Code

Fig. 2 displays a flowchart for the KLOE calorimeter reconstruction code. It shows that the following items have to be fully understood:

1. Pre-clustering;
2. Cluster normalization;
3. Cluster unfolding;
4. Cluster merging;
5. Time order of the clusters.

2.1 Pre-cluster

Fig. 3B depicts the GlueX BCAL left-hand coordinate system, with the z axis pointing towards the photon beam momentum direction. The ϕ and r directions are shown in fig. 3A, where the open circles represent cells with energy deposition in the BCAL. The coordinate θ is defined with respect to the z axis, as depicted in fig. 3B. The KLOE coordinate system is defined in a similar way.

In the KLOE code, the raw data is first read in, decoded, and then arranged into arrays whose indices stand for the hit module, layer, and sector numbers. One can then know if the two readout cells are adjacent by judging the indices in r and ϕ directions, i.e., if the difference of the indices in both the r and ϕ directions is less than or equal to 1, one says that the two cells are adjacent. By searching all the adjacent cells, one can then define a cluster.

After finding all the cells belonging to a cluster, quantities such as the energy-weighted average positions, total energy and time as measured from a cell on both BCAL sides

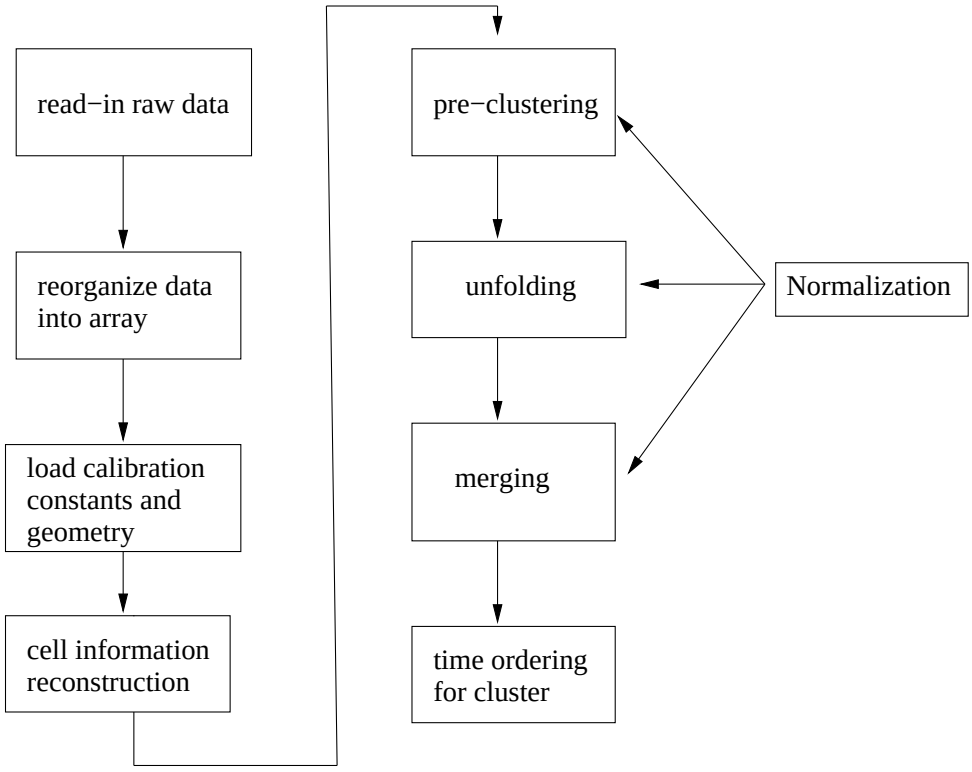

Figure 2: *The schematic for the KLOE reconstruction code.*

are calculated and used in the so called "normalization" and "unfolding" stages of the clustering algorithm, as described in sections 2.2 and 2.3, respectively.

2.2 Cluster Normalization

In the cluster normalization stage, the following quantities are calculated:

1. Energy-weighted average position and energy-weighted average time accounting for all cells within a cluster:

$$X = \sum_{1 \le i \le n_{cell}} (X_i E_i)/E \tag{1}$$

$$Y = \sum_{1 \le i \le n_{cell}} (Y_i E_i)/E \tag{2}$$

$$Z = \sum_{1 \le i \le n_{cell}} (Z_i E_i)/E \tag{3}$$

$$T = \sum_{1 \le i \le n_{cell}} (T_i E_i)/E \tag{4}$$

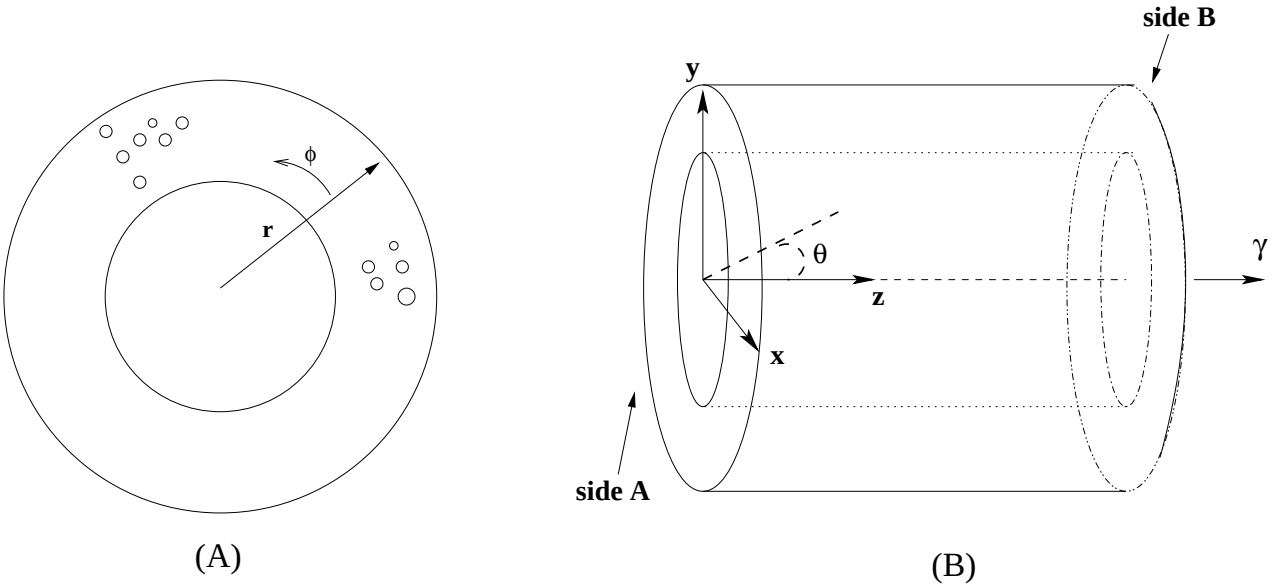

Figure 3: *The left figure (A) shows a transverse cross-section of the BCAL. It also shows two clusters in the pre-clustering stage. The right figure (B) depicts the BCAL xyz coordinate system.*

where n_{cell} is the total number of cells in a cluster; X_i, Y_i and Z_i are the spatial coordinates of cell i; T_i, E_i are the reconstructed time and energy information from cell i; and E is the total energy of the cluster:

$$E = \sum_{1 \leq i \leq n_{cell}} E_i \tag{5}$$

2. Cluster direction and apex point:

   The apex point of a cluster associated to a particle is defined by the spatial coordinate (AP_x,AP_y,AP_z) of the particle entrance point into the BCAL. The direction of the cluster is defined by the vector $\vec{C} = C_x\hat{i} + C_y\hat{j} + C_z\hat{k}$, as depicted in fig. 4, where $\hat{i}$, $\hat{j}$ and $\hat{k}$ are the unitary vectors in the x, y and z directions, respectively, and C_x, C_y and C_z are given by the following expressions:

$$C_x = P_x/P \tag{6}$$

$$C_y = P_y/P \tag{7}$$

$$C_z = P_z/P \tag{8}$$

   where P_x, P_y and P_z are the absolute values for the momentum in the x, y and z directions, respectively, and $P = \sqrt{P_x^2 + P_y^2 + P_z^2}$ is the total momentum of the particle associated to the cluster.

The dashed line in fig. 4 represents the result of a fit to the energy-weighted average positions using the KLOE fitting method, which provides the apex point position and particle direction. This method will be improved in the GlueX BCAL algorithm by using tracking information from the inner GlueX chambers. More details are given in section 4.2.

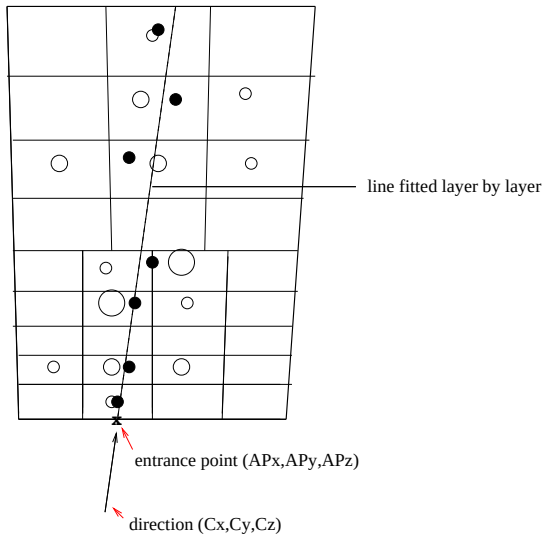

Figure 4: *A particle (full line) enters the BCAL at the apex point (AP_x,AP_y,AP_z) and with direction $\vec{C} = (C_x, C_y, C_z)$, leaving energy deposition in some of the BCAL cells (open circles). The filled circles represent the energy-weighted average positions of all cells with energy deposition in a given BCAL layer. The dashed line depicts the result of a fitting to these positions.*

3. Time fluctuations of the cluster from the BCAL sides A and B:

$$T^j = \sum_{1 \le i \le n_{cell}} (T_i^j E_i^j) / \sum_{1 \le i \le n_{cell}} E_i^j \tag{9}$$

$$T^j_{sqr} = \sum_{1 \le i \le n_{cell}} ((T_i^j)^2 E_i^j) / \sum_{1 \le i \le n_{cell}} E_i^j \tag{10}$$

$$T^j_{rms} = \sqrt{(T^j_{sqr} - (T^j)^2)} \tag{11}$$

where $j = a, b$ represents the sides A and B of detector respectively. E_i^j, T_i^j are, respectively, the energy and time information from cell i on side j of the detector.

2.3 Unfolding

An event has usually more than one track, with each track associated to a single particle. It may happen that two or more of such tracks are not resolved in the transverse $r\phi$

direction, leading to multiple clusters being misidentified as a single cluster in the pre-clustering stage. However, these tracks can still be well separated in the longitudinal z direction. In the unfolding stage, this extra information is used to assure that two different tracks are resolved in the BCAL reconstruction process.

The decision on whether the unfolding stage is necessary for a certain cluster obtained from preclustering is based on the variable Z_{rms}:

$$Z_{rms} = \sqrt{(Z_{sqr} - Z^2)} \tag{12}$$

where,

$$Z_{sqr} = \sum_{1 \le i \le n} (Z_i^2 E_i)/E \tag{13}$$

and Z_i, E_i, E and Z are defined in section 2.2.

Fig. 5 illustrates how a large Z_{rms} value would indicate that a cluster needs to be unfolded. In this figure, track (a) spans a smaller z range due to its smaller angle compared to track (b) which enters the BCAL at larger θ values (shaded areas). However, if both tracks are mixed together in the pre-clustering stage, the fluctuations in Z reconstructed of the mixed clusters will be large since the reconstructed position Z_i of a cell "i" in this cluster will be located at a distance far from the averaged Z position, as shown in fig. 5. Noting that the Z position is calculated from the measured time T, it is therefore more appropriate to use T_{rms} instead of Z_{rms} information in the unfolding stage. The following criterion is then used in the reconstruction algorithm to judge whether the unfolding stage is necessary:

$$\sqrt{(T^a_{rms})^2 + (T^b_{rms})^2} > 5.0\ ns \tag{14}$$

where T^a_{rms} and T^b_{rms} are defined in section 2.2 (eq. 11).

This criterion is supported by MC simulations performed by the KLOE Collaboration as a good way to resolve single and overlapped clusters. The factor 5 ns in the above expression is the one evaluated by KLOE. Very preliminary studies indicate that a value of 3 ns is more adequate for the GlueX BCAL, but more investigations are still needed. Therefore, the KLOE value is the one used throughout the rest of this report.

Once one has determined that the unfolding stage is necessary, the cells belonging to the original cluster are classified into four groups (a), (b), (c) and (d) according to the following rules:

$$(a)\ \ T^a_i > T^a,\ \ T^b_i > T^b \qquad (b)\ \ T^a_i > T^a,\ \ T^b_i < T^b$$

$$(c)\ \ T^a_i < T^a,\ \ T^b_i > T^b \qquad (d)\ \ T^a_i < T^a,\ \ T^b_i < T^b$$

where the above variables are defined in section 2.2 (eq. 9-10).

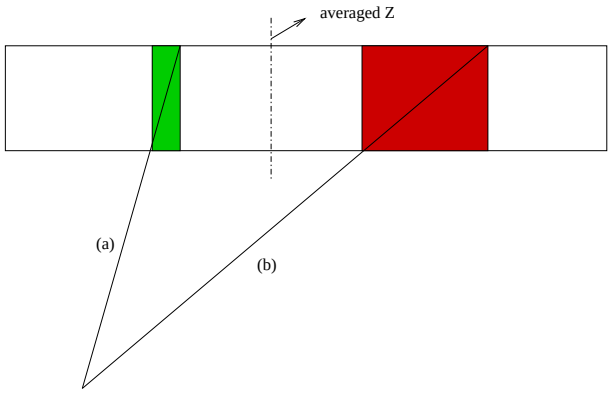

Figure 5: *Two tracks hit the BCAL transversely close to each other but at different longitudinal z direction component, causing a large fluctuation in the Z coordinate calculation. Track (a) has $\theta \simeq 90^o$, while track (b) is found at smaller θ angle with respect to the z axes.*

The reason for this classification is not clear yet, as it could not be provided by the KLOE group, since the person who introduced this logic has left the KLOE collaboration. However, it is known that this logic has been used in the KLOE reconstruction code, and therefore it is assumed that it leads to a good cluster reconstruction resolution. Further studies on this selection are still to be performed, and new options for the cells' classification will be investigated in the near future.

2.4 Merging

In the merging stage, the following considerations have to be taken into account:

1. In the unfolding stage, a cluster can be over-unfolded into as many as four clusters. Thus, these clusters have to be merged back into a unique cluster such that the particle to which it is associated can be unambigously identified.
2. It may happen that a track leaves energy in both the BCAL and FCAL. In this case, the clusters reconstruced in the BCAL have to be associated to those reconstructed in the FCAL, as depicted in fig. 6.
3. Dead and bad cells excluded from the readout after calibration procedures may introduce "holes" in the readout segmentation leading to some good cells, with energy deposition and adjacent to those dead/bad cell(s), not being associated to the cluster they belong to in the unfolding stage. These cells have to be merged back into the main cluster (see fig. 7).

The time and spatial distance between two different clusters are used in the merging stage. The spatial distance R, the transverse distance R^T_{diff} and the longitudinal distance Z_{diff}

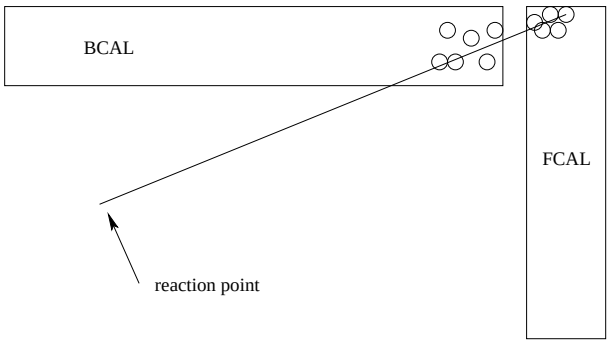

Figure 6: *A track hits both the barrel and forward calorimeters, resulting in showers in both detectors.*

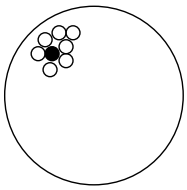

Figure 7: *If the filled circle in the figure is a dead channel, then one or more good cell (open circles) might not be considered as part of the group of adjacent cells belonging to a cluster, and some informations are therefore lost in the pre-clustering process.*

are depicted in fig. 8. The time difference T_{diff} is defined as:

$$T_{diff} = |T_1 - T_2| \tag{15}$$

where T_1 and T_2 are the time T for each of the two clusters as defined in section 2.2 (eq. 4.)

The following criteria, based on the above variables, are used in the BCAL reconstruction algorithm:

$$R < 40\ cm \tag{16}$$

$$T_{diff} < 2.5\ ns \tag{17}$$

$$Z_{diff} < 30\ cm \tag{18}$$

$$R^T_{diff} < 40\ cm \tag{19}$$

If one of the clusters is found in the BCAL and the other in the FCAL, only the criteria 16 and 17 are used in the merging stage.

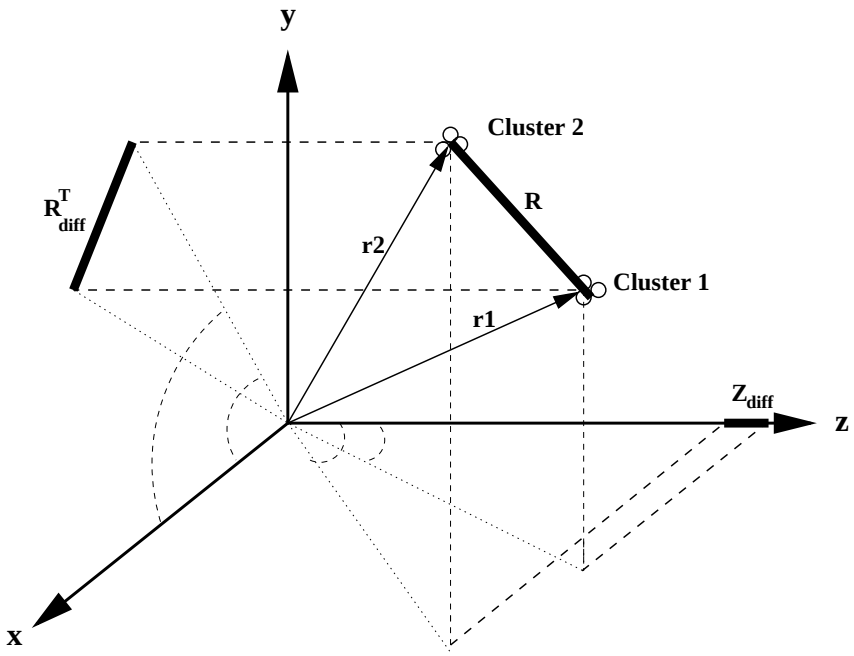

Figure 8: *R, R^T_{diff} and Z_{diff} are the tri-dimensional, transverse and longitudinal distance between two clusters "1" and "2", respectively.*

2.5 Time Ordering

After the merging stage, all the reconstructed "n" clusters in an event are ordered according to the following scheme:

cluster(1), cluster(2), ..., cluster(i), ..., cluster(j), ..., cluster(n)

where $i < j$ if $T_i < T_j$, and $i = [1, n]$.

This time ordering was established by the KLOE group and one has inherited it as a good way of organizing the different clusters in an event, though other methods, like energy ordering, could also be used.

3 BCAL Segmentation Interface for the HallD MC Simulation

The BCAL geometry is defined in section 1. However, in the HDGEANT simulation package the BCAL is divided into 48 modules without further segmentation, and a hit represents the particle information left in a specific position on the BCAL, either on side A or B. The total information from a module is then the sum of all hits in this module. In order to match the BCAL segmentation, a sub-routine was implemented in the simulation

package such that the information from a cell is the combination of all hits spanning the cell area as defined in section 1. Assigned to this combination of hits is a set of numbers representing the module, layer and sector numbers and the energy $E_i^{a,b}$ and time $T_i^{a,b}$, where "i" represents the corresponding cell associated to these hits, and (a, b) stands for sides A and B of the calorimeter.

With this small modification of the HDGEANT simulation package, one can test the reconstruction code, though further modifications are required in the future to account for the detector resolution.

4 Test of the BCAL Reconstruction Code

Simulation data, as provided by the HDGEANT package, are used to study the consistency of the reconstruction code. Three different tests are performed using single γ, single π^0 and $\gamma p \rightarrow p\pi^+\pi^-\pi^0$ events. Only reconstructed clusters with energy above 6 MeV are considered.

4.1 Single γ Events

A single γ event is supposed to have only one cluster to be reconstructed. However, fig. 9 shows that as the energy of the γ particle increases, the percentage of events having more than one reconstructed cluster also increases. The source of these additional clusters will be discussed in this section. Fig. 10 shows that the number of cells per cluster also increases with the γ energy, as it would be expected.

Fig. 11 shows the reconstructed γ energy for each MC data sample. The distributions are not Gaussians around the true γ energies, but instead have a tail on the left side and an enhancement near zero even for γ energies as low as 0.2 GeV. For this energy, it was found that 73%[1] of the events in the tail are due to photons that interact with materials before the BCAL, and that the other 34% are due to the mechanism depicted in Fig. 12[2]. In the latter case, some of the energy of the γ particle may leak from the BCAL, contributing to part of the tail in the energy distribution. Also, there is a chance that a low energy backward γ and/or electron is generated during the Bremsstrahlung process. This γ can escape the BCAL and be considered as energy leakage. On the other hand, the back-generated γ can also travel to another BCAL region, as shown in fig. 12, producing

[1] 66%, 46% and 21% for 0.5 GeV, 1.0 GeV and 2.0 GeV, respectively

[2] These events might also interact with materials before the BCAL and can, therefore, partially contribute to those percentages cited above.

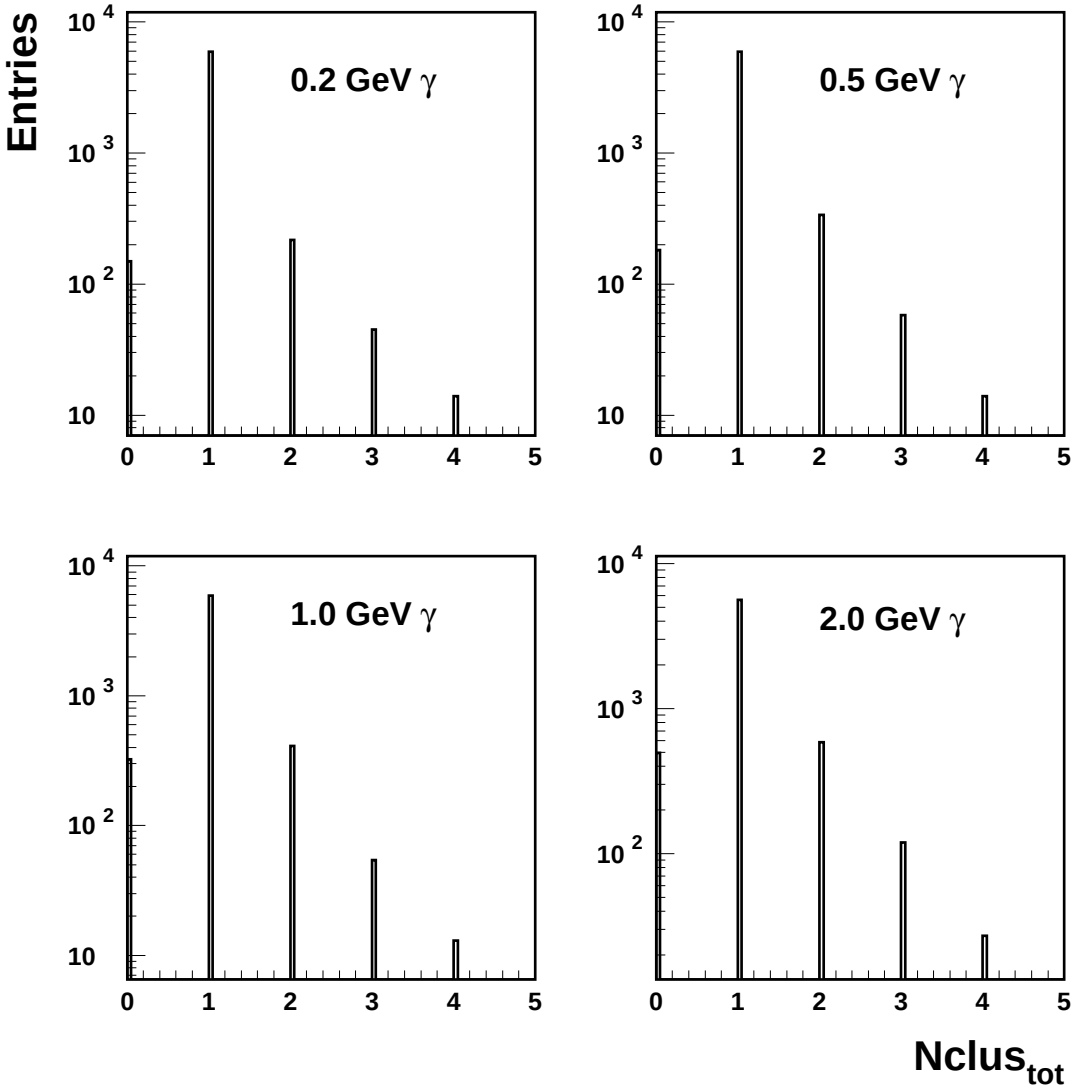

Figure 9: *Reconstructed total number of cluster $Nclus_{tot}$ distribution for single γ MC events. The events were generated with pz/p in [-0.4,0.9] so that the γ particle is at good angle to hit the BCAL. Some events have more than one cluster identified. The percentages of such events are 4.5%, 6.7%, 7.5% and 11.6% for 0.2 GeV, 0.5 GeV, 1.0 GeV and 2.0 GeV events respectively.*

an extra low energy cluster. This effect would partially explain the source for additional clusters in one-γ event as mentioned in section 4.

The reconstructed versus true γ C_x and C_z directions, as defined in section 2.2 (eq. 6-8) are shown in fig. 13. These reconstructed quantities were obtained by a linear layer-by-

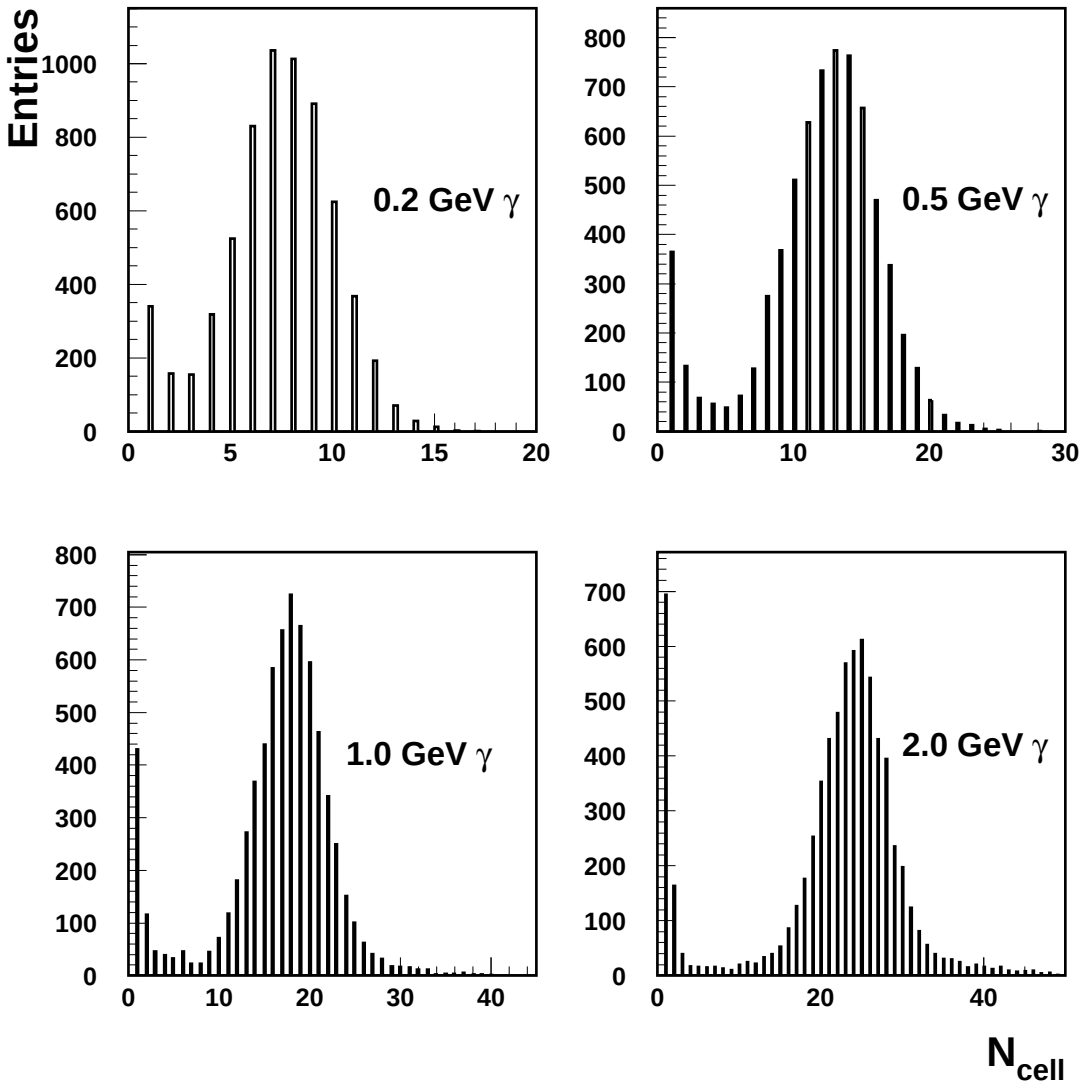

Figure 10: *Reconstructed total number of cells N_{cell} inside a cluster for single γ MC events with pz/p in [-0.4,0.9].*

layer fit to the energy-weighted positions. A poorer correlation can be observed between the true and reconstructed variables, but this is due to the simple fitting method used in this example. A more elaborate method will be developed to take into account extra fit constraints like information from other GlueX components as the inner tracking chambers.

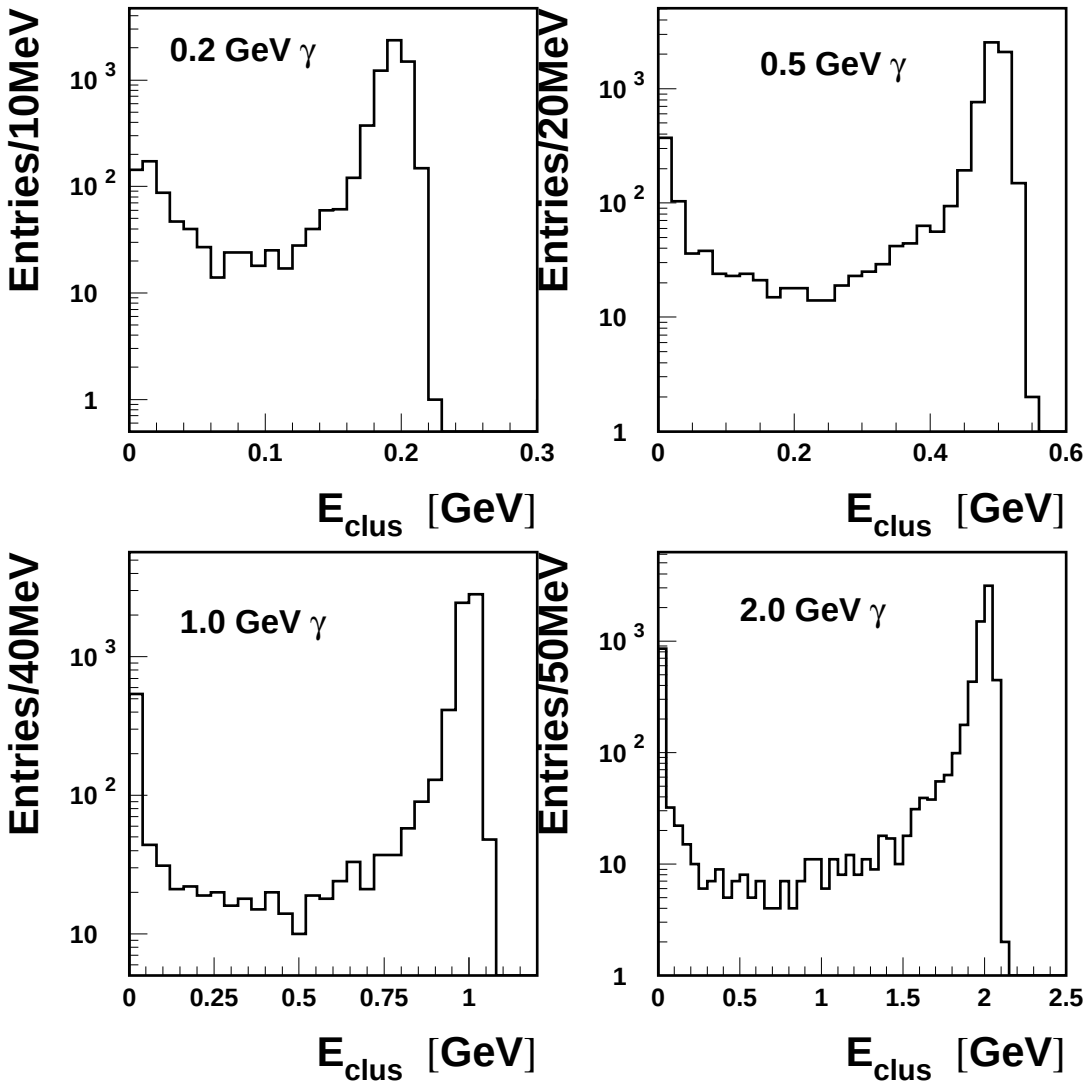

Figure 11: *Reconstructed energy spectrum for single γ MC events.*

4.2 π^0 Invariant Mass

In this section, the reconstructed π^0 invariant mass from single π^0 and $\gamma p \rightarrow p\pi^+\pi^-\pi^0$ events is analysed. The π^0 invariant mass is calculated considering any possible two-cluster combination of all clusters in an event, though care has been taken to remove double-counting pairs from the computation.

Fig. 14 depicts the π^0 distributions for .2 GeV, 0.5 GeV, 1.0 GeV and 2.0 GeV π^0 momentum. It shows that as the π^0 momentum increases, the possibility of resolving two

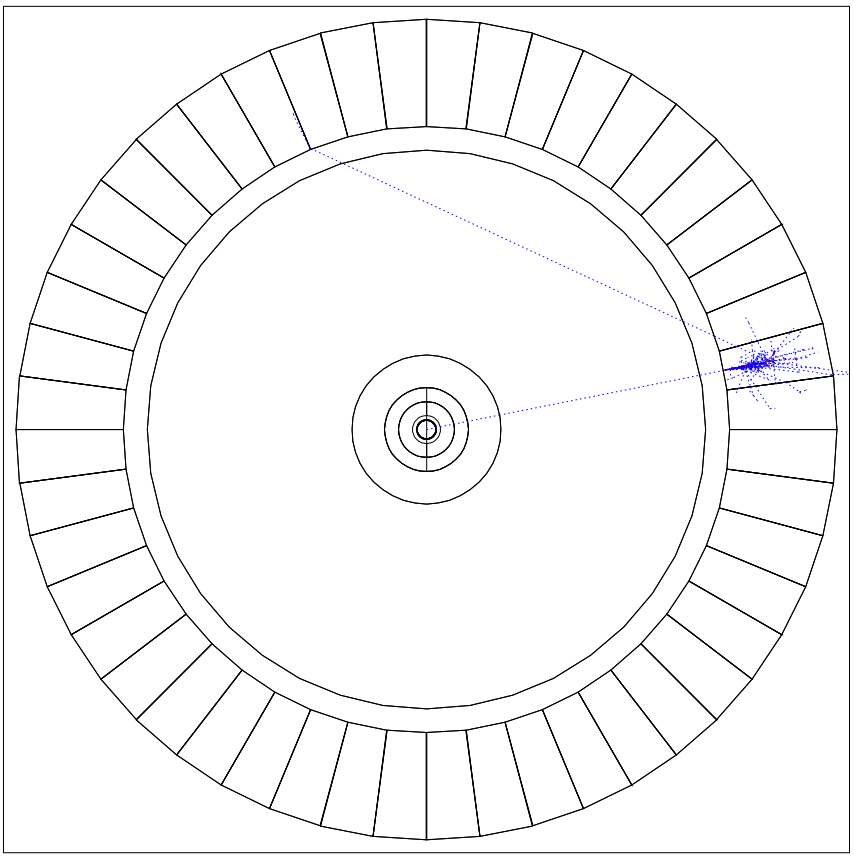

Figure 12: *Event display of a 0.5 GeV γ event. Processes like this one, in which a photon is ejected into a different part of the BCAL, helps to explain the long tail on the energy spectrum and also the additional clusters in fig. 9.*

γ's from a single π^0 decay is compromised by the fact that these two γ's get closer in angle resulting in poorer π^0 reconstruction. This will reflect on the reconstructed π^0 invariant mass distribution, as can be seen in fig. 15.

Fig. 16 shows the reconstructed π^0 invariant mass for $\gamma p \to \pi^+\pi^-\pi^0$ events. The top left figure shows the $Nclus_{tot}$ distribution (see 9 for its definition), and the top right figure is the π^0 invariant mass distribution considering all the reconstructed events. A large combinatoric background can be observed. This background can be reduced by cutting on $Nclus_{tot}$. The bottom figures show the improvement in the resolution of the same invariant mass distribution for a total number of cluster per event constrained to $Nclus_{tot} < 5$ and $Nclus_{tot} < 4$. Furthermore, it is expected that by adding information from the inner GlueX tracking chambers will significantly contribute to lower the combinatoric background in the invariant mass distribution.

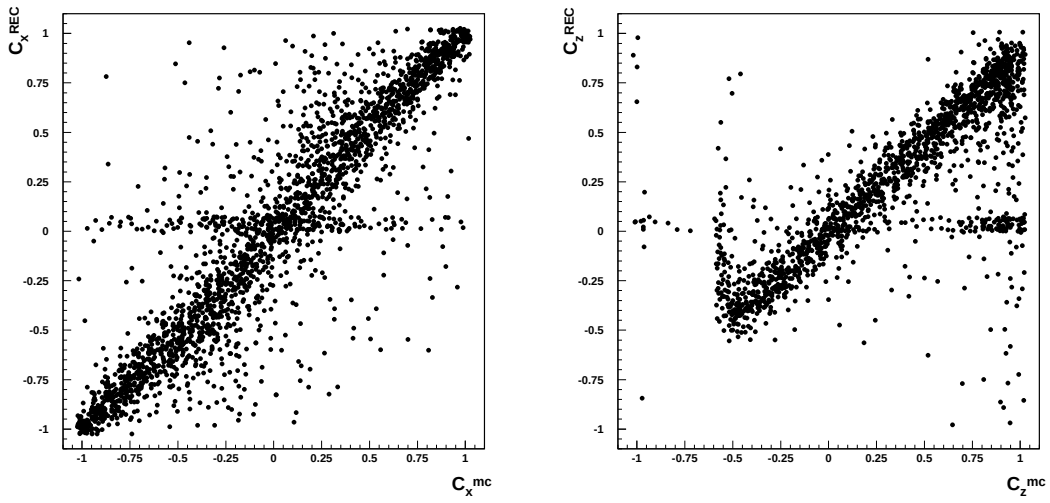

Figure 13: *MC true versus reconstructed particle direction information for 1.0 GeV incident γ events. The subscripts REC and mc in C_x and C_z stands for reconstructed and Monte Carlo, respectively.*

5 Future Plans

The future plans for the BCAL reconstruction code development include the following items:

1. Migration of the code from Fortran to C++ language.
2. At present, the material in the BCAL simulation as in HDGEAT is a mix of glue-lead-scintillator, leading to an overall effective interaction length for the BCAL. Also, HDGEANT provides only the energy deposit in a cell, whithout taking into account the effects of light transmission in the scintillator fibers. A more realistic approach would be to simulate the full BCAL geometry, however this would make the processing time to generate a data sample considerably slow, that, for many applications, is not desirable. An alternative solution would be to consider the average ratio between the energy deposit inside the fibers coupled to a readout cell and the whole energy deposit inside the same cell, and then parameterize the fluctuation of such ratio as a function of the whole deposit energy. This information is indeed available from standalone BCAL simulations [4]. With the total cell energy, such ratio and its fluctuation, the total energy deposit in a readout cell, as given by HDGEANT, can then be smeared leading to a more realistic approach. Furthermore, if the overall photo-detection efficiency is considered, an even more realistic approach can be accomplished and a better evaluation of the π^0 invariant mass reconstruction can be achieved.
3. Once the above item is completed, one may be able to make better optimization of

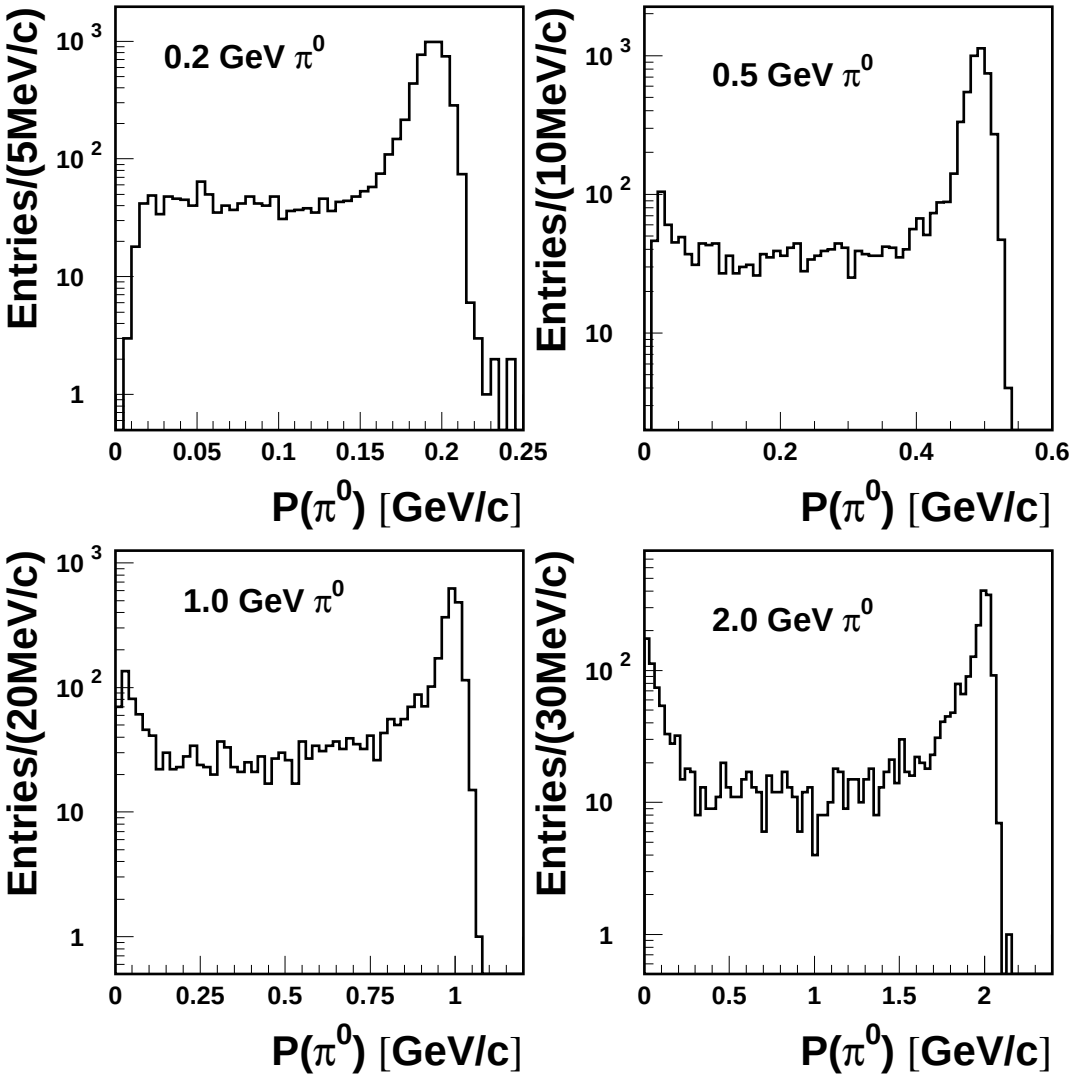

Figure 14: *Reconstructed single π^0 momentum distributions for π^0 with momentum of 0.2 GeV, 0.5 GeV, 1.0 GeV and 2.0 GeV. Only γ's found in the BCAL are considered for the π^0 reconstruction.*

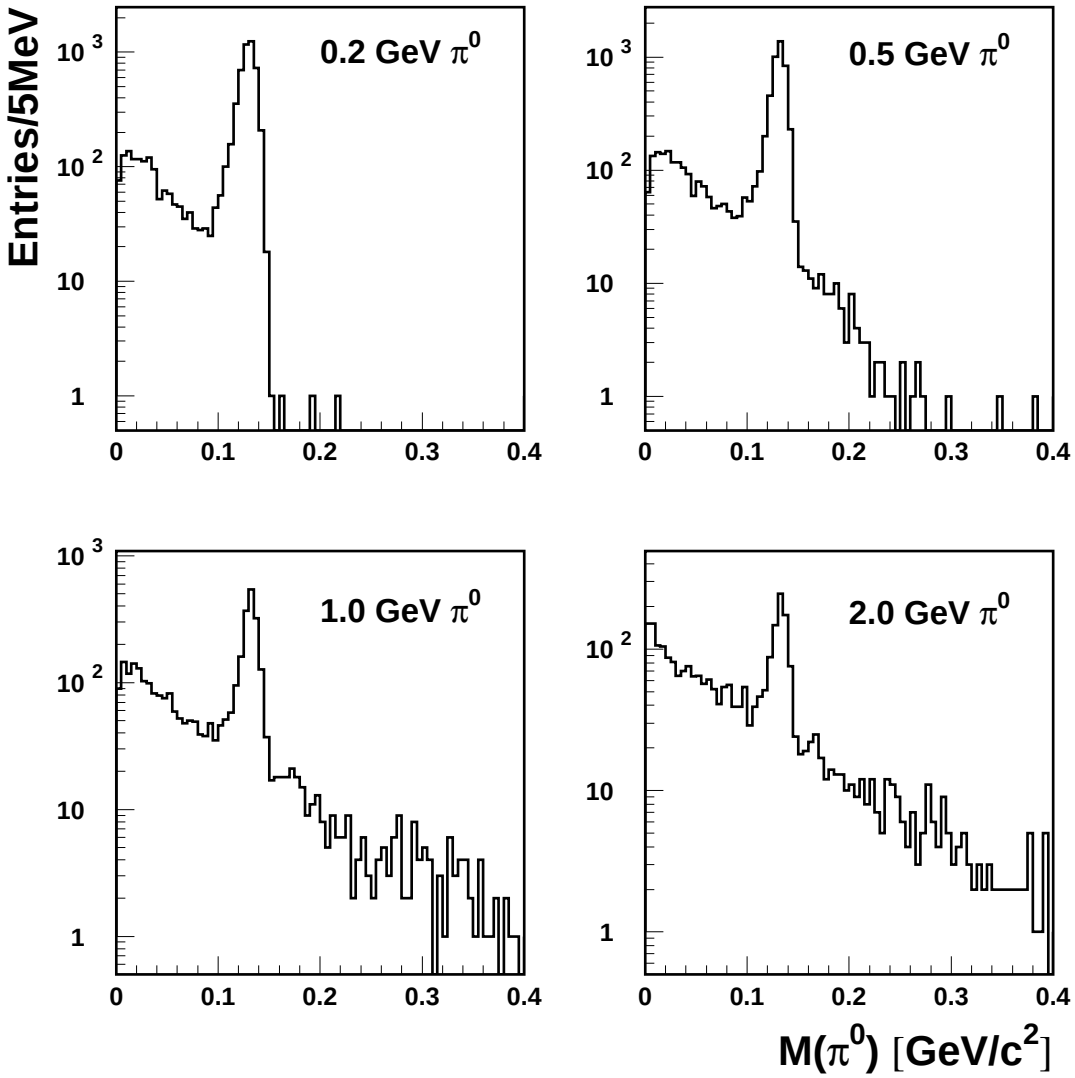

Figure 15: *Single π^0 invariant mass spectrum for different values of π^0 momentum.*

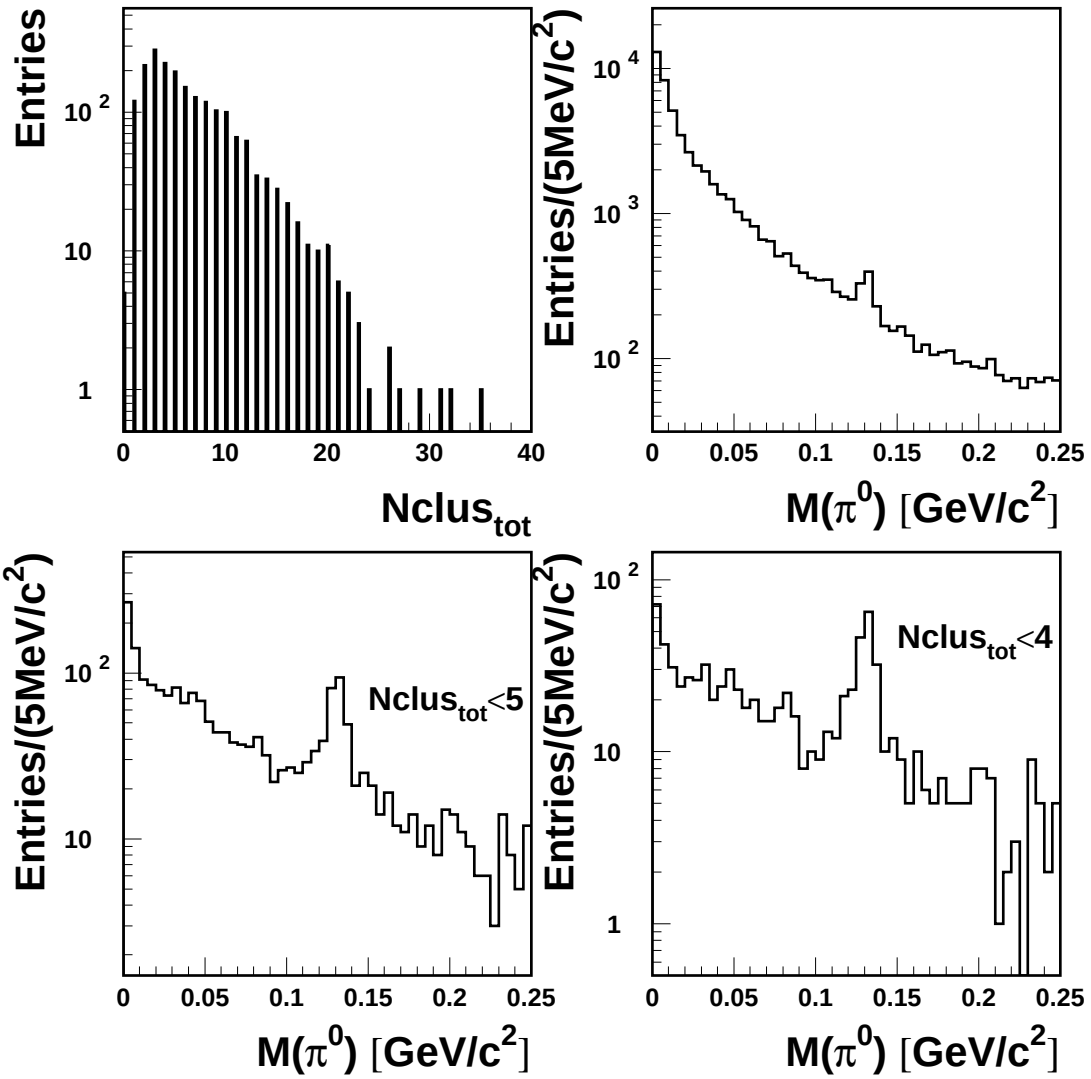

Figure 16: *The invariant mass distribution for π^0's from the $\gamma p \to \pi^+\pi^-\pi^0$ process. Only clusters in the BCAL are used. The top left plot is the total number of reconstructed clusters $Nclus_{tot}$. The top right plot is the π^0 invariant mass spectrum whithout applying the $Nclus_{tot}$ cut. The left and right bottom plots represent the same invariant mass distribution with $Nclus_{tot} < 5$ and $Nclus_{tot} < 4$, respectively.*

the BCAL segmentation, as well as provide the π^0 invariant mass resolution to a very good approximation.

4. Finally, the BCAL reconstruction code has yet to be fully improved with better fitting methods, and it has also to be independently tested by third parties to ultimately be

accepted as a full package.

6 Conclusions

The first studies using the BCAL reconstructed code are presented. Its consistency was tested using Monte Carlo simulation events generated with the HDGEANT package. Single γ and π^0 and $\gamma p \to \pi^+\pi^-\pi^0$ events were used to investigate the reconstruction efficiency and the π^0 invariant mass reconstruction. It was found that a large combinatoric background is presented in the latter case, while the addition of tracking information is expected to considerably improve the reconstruction efficiency and invariant mass resolution.

Finally, a number of items are listed as tasks to be implemented in order to improve the reconstruction algorithm.

7 Acknowledgments

The authors would like to thank Zisis Papandreou and Richard Jones for their help with the HDGEANT package, Rafael Hakobyan for his suggestions and for providing fig. 12, and to Stefano Miscetti, from KLOE, for providing the KLOE reconstruction algorithm and several other informations.

References

[1] A. Dzierba, Preprint hep-ex/0106010, 2001;
The GlueX Collaboration, *The GlueX Experiment. GlueX-doc-346*, Oct 2004, 2004;
The Gluex Collaboration, *Brief Summaries of Detector Systems. GlueX-doc-323-v6*, Apr 2005, 2005.

[2] M. Adinolf et al., NIM **A482**, 364 (2002).

[3] R. Jones, *The HDGeant Monte Carlo Program. University of Connecticut*, 2001, 2001.

[4] R. Hakobyan et al., *Simulation of the GlueX Barrel Calorimeter. GlueX-doc-529-v5*, Oct 2005, 2005.