GlueX-doc 849-v1
July 25, 2007

BCAL Simulation and Reconstruction in the DANA Framework

M. R. Shepherd
Department of Physics
Indiana University

This document overviews the current status and performance of the BCAL simulation and reconstruction algorithm. The algorithm is largely taken from the KLOE reconstruction code and initial development and performance characterization was done by Xu *et al.* and is documented extensively in Ref. [1]. The detector response and reconstruction algorithms are outlined below. Particular emphasis is placed on features of the algorithm that *have not* already been discussed in [1].

1 Detector Response

1.1 Geant Level

Showers the in BCAL are simulated using the GEANT-based Hall D Monte Carlo. This simulation models secondary particle development in a shower, but does not contain information about fiber geometry or the production and collection of scintillation light in the fibers and readout electronics. For each cell in the BCAL, the simulation produces three quantities:

$$E^{\text{cell}} = \sum_i dE_i, \tag{1}$$

$$z^{\text{cell}} = \frac{\sum_i z_i dE_i}{E^{\text{cell}}}, \text{ and} \tag{2}$$

$$t^{\text{cell}} = \frac{\sum_i t_i dE_i}{E^{\text{cell}}}. \tag{3}$$

Here the index i runs over the GEANT steps taken by charged particles in the shower and dE, t, and z are energy deposition, time, and z location at each step. Cells with energy, E^{cell}, greater than 1 MeV[1] are written to disk for further processing. At this level the deposited energy and energy weighted time and z position are stored for each cell. These stored values will then be turned to a electronics response for the photo-detecting elements at each end of the cell.

1.2 Sampling Fraction Fluctuations

Independent Monte Carlo studies [2] have shown that the scintillating fiber geometry of the BCAL samples approximately 12% of the charged particles in the shower roughly independent of energy. This sampling fraction however has statistical fluctuations that are energy dependent. This is logical since the energy deposited in a cell is directly proportional to the number of charged particles in a cell, and with more charged particles, the statistical fluctuations on the fraction of charged particles sampled by the scintillating fiber decreases. Monte Carlo studies [2] have shown that these fluctuations are Gaussian-distributed with a width as a fraction of the mean sampling fraction, $\frac{\sigma_f}{f}$, of

$$\frac{\sigma_f}{f} = \frac{A}{\sqrt{E^{\text{cell}}_{\text{dep}}}} + B, \tag{4}$$

[1] This threshold will be justified in the proceeding section.

where $E_{\rm dep}^{\rm cell}$ is the total energy deposited in the cell. The parameters A and B are 4.217% and 1.288%, respectively.

To properly simulate this effect in the Monte Carlo, for each cell we introduce a sampling fraction scale factor, $s^{\rm cell}$, where $s^{\rm cell}$ is drawn at random from a Gaussian distribution with width given by Equation 4. We then reset

$$E^{\rm cell} \rightarrow s^{\rm cell} E^{\rm cell}. \tag{5}$$

Note that the sampling fraction fluctuations on the readout of each end of the module are correlated. This procedure maintains this correlation since fluctuations are introduced before signals are propagated down to the ends of the module.

1.3 Time Delay and Attenuation

After sampling fraction fluctuations have been introduced, the propagation of the scintillation light down the module is modeled. If we take $\Delta z^{\rm up}$ and $\Delta z^{\rm down}$ as the distance from $z^{\rm cell}$ (defined above) to the upstream and downstream readout elements, then we can write the attenuated energy and delayed time as

$$E^{\rm up} = E^{\rm cell} e^{-\Delta z^{\rm up}/\lambda}, \tag{6}$$

$$E^{\rm down} = E^{\rm cell} e^{-\Delta z^{\rm up}/\lambda}, \tag{7}$$

$$t^{\rm up} = t^{\rm cell} + \Delta z^{\rm up}/c', \text{ and} \tag{8}$$

$$t^{\rm down} = t^{\rm cell} + \Delta z^{\rm down}/c', \tag{9}$$

where $\lambda = 300$ cm is the attenuation length and $c' = 16.75$ cm/ns is the effective speed of light in the fiber.

1.4 Electronics Threshold

To reduce data volume, the outputs of photo-sensing devices are zero-suppressed. Since this effectively removes cells with low energy around the borders of a shower, the same zero-suppression must be applied in the Monte Carlo in order to model the resolution of the detector. While this electronics threshold is a strong function of the choice of readout device, light output of the scintillating fibers, and desired zero-suppression factor, we attempt to make a reasonable estimate for the threshold as follows. We first make the following assumptions:

- the data readout limits the maximum occupancy in the BCAL for any given event to 5%
- the window of integration for BCAL pulses is 100 ns
- the dark rate in the readout device is purely single photoelectron noise at a rate of 60 MHz
- the width of the pedestal is dominated by dark rate and not electronics noise
- it takes 26 keV of unattenuated deposited energy in the fibers produce one photoelectron[2]

Given the 60 MHz dark rate the probability to have 11 or more dark pulses in a single 100 ns window is about 4%, so we take 11 photoelectrons as the threshold. To reach 11 photoelectrons we need 286 keV of deposited energy in the fibers, or, assuming a sampling fraction of 12%, 2.4 MeV of energy in the cell after attenuation. Note that this is significantly higher than the unattenuated 1 MeV cutoff used in GEANT generation. After sampling fraction smearing and attenuation procedures described above, all hits which are less than 2.4 MeV are removed.

[2]This number comes from discussion with G. Lolos. Studies to experimentally validate this are underway by A. Seminov.

1.5 Additional Response Effects

There are additional errors in digitization and response that are not currently modeled in the simulation. These errors include:

- errors due to statistical fluctuations in the number of photons detected by the photosensor
- errors in that contribute to the timing resolution of the pulses such as intrinsic timing resolution of scintillator and photosensor
- errors due to electronics digitization of the signals such as limited precision due to the size of the least count in both the ADC and TDC

As these effects are better understood they should be incorporated into the response algorithm.

2 Shower Reconstruction

The reconstruction algorithm implemented[3] is as described in [1]. In short, the algorithm converts upstream and downstream hits to single cell hits. These hits are characterized by x, y, z positions, E, and average t, where z is calculated by the end-to-end time difference and E has been corrected for attenuation based on the value of z. The algorithm then does a purely geometrical clustering of hits with no consideration of energy deposition. Cells that are simply neighboring other cells are collected together to form clusters. If $\vec{r_i}$ and t_i are the energy-weighted positions and average time of the cluster, clusters that meet the following criteria are merged.

$$|\vec{r_1} - \vec{r_2}| < 40 \text{ cm} \quad (10)$$

$$|t_1 - t_2| < 2.5 \text{ ns} \quad (11)$$

$$|r_{1z} - r_{2z}| < 30 \text{ ns} \quad (12)$$

$$\sqrt{(r_{1x} - r_{2x})^2 + (r_{1y} - r_{2y})^2} < 40 \text{ cm} \quad (13)$$

In addition, clusters are split in the z direction if their r.m.s. time is greater than 5.0 ns. These thresholds are implemented in the algorithm as parameters and should, at some point, be re-evaluated and adjusted if necessary. The values shown above are values determined by the authors of [1]. After the cluster merging and breaking routines, final energy-weighted spacial positions for the showers are determined.

As a final step in the reconstruction algorithm, the 3D energy-weighted position within a layer for each cluster is determined. These 3D points are projected onto the x, y, and z axis and fit to lines so that a parametric representation of the path of the photon can be inferred. One can write the location of photon as a function of some parameter, s, in a 3D "point-slope" form as

$$\vec{d}(s) = \vec{a} + s\vec{c}. \quad (14)$$

The vectors $\vec{a}$ and $\vec{c}$ (and their errors) are determined by the series of linear fits done by the reconstruction code. Ideally $\vec{c}$ should be parallel to the flight path of the photon.

To avoid creation of spurious clusters due to isolated low energy cells in the calorimeter, all showers with energy less than 20 MeV are removed. This cluster threshold can be adjusted with a parameter in the reconstruction code.

After examination of the reconstruction code, two items merit further study in the future:

- The energy deposited in each cell is a straight average of the attenuation-corrected energies on the upstream and downstream ends. The end with the largest energy detected presumably is the most precise, so perhaps a weighted average could be introduced to reduce the error on each cell's energy measurement.

[3] Currently this algorithm is utilized when data of the type `DBCALShower` is extracted with the tag `IU`. The functional differences between this algorithm and the default algorithm are the addition of a cluster threshold and removal of "corrected" energy. The most noticeable change is well over an order of magnitude speed increase.

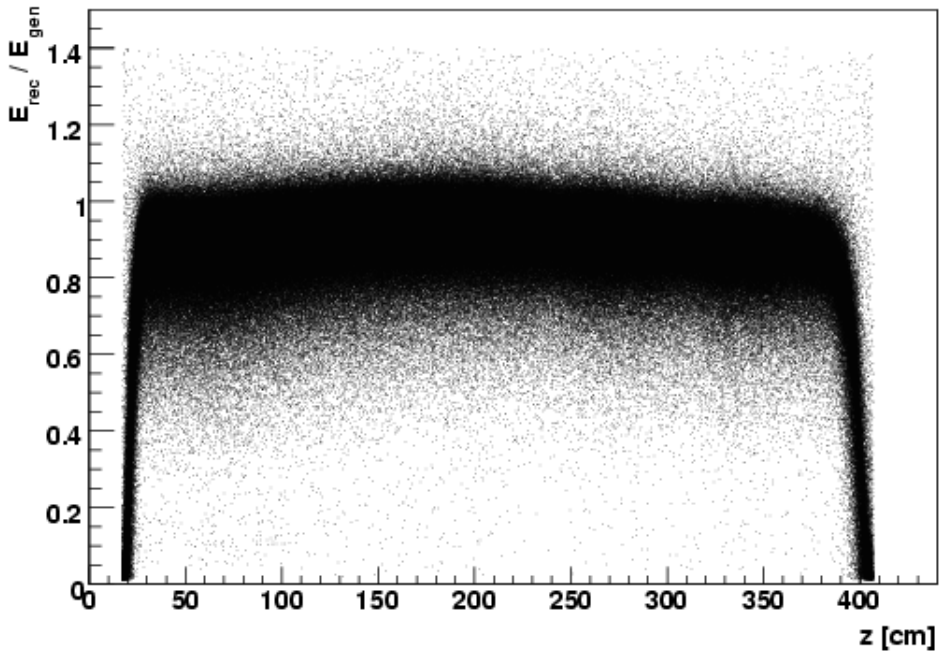


Figure 1: A plot of the ratio of the reconstructed energy to generated energy as a function of incident photon position along the BCAL. Photons are generated uniformly in energy from 0 - 2 GeV.

- The algorithm requires that both ends of a cell be hit for the cell to be included in reconstruction. Therefore no information is used from low energy cells where one hit is lost due to attenuation. The algorithm could probably be improved by adding back these single-sided hits after clusters have been formed.

At this stage of the reconstruction one has formed showers in the calorimeter, determined their position, and fit a photon trajectory using layer information. However, there is some energy loss due to leakage and cells that do not exceed the electronics threshold. The reconstructed shower energy, $E_{\rm rec}$ must be rescaled to account for these effects and accurately represent the incident photon energy.

3 Energy Calibration

Calibration of the BCAL is done by uniformly populating the detector in z and ϕ with photons that have an energy between 0 and 2 GeV. All photons are thrown from the center of the target. For the sake of energy calibration, we consider only events where the thrown photon did not convert before reaching the BCAL, and exactly one BCAL shower was reconstructed. For these events the ratio of reconstructed to generated energy as a function of incident photon position is show in Figure 1. As expected leakage at the ends of the BCAL causes an underestimate of reconstructed energy. One also notices a slight z dependence on this ratio – it is peaked in the near the middle of the BCAL.

To extract the z-dependent energy calibration, we divide the sample pictured in Figure 1 in slices along the x-axis (z position of the incident photon). We avoid the ends of the BCAL and take a total of fourteen 25 cm slices from 30 cm to 380 cm. The ratio of reconstructed to generated energy as a function of generated energy for each of the BCAL regions is shown in Figure 2. Following the parametrization suggested by

M. Kornicer, profile plots of the data are fit to the function

$$\frac{E_{\rm rec}}{E_{\rm gen}} = NE_{\rm gen}^{\epsilon}, \tag{15}$$

where N and ϵ are free parameters. The fit is performed over the range $E_{\rm gen} > 50$ MeV and models the data well. For $E_{\rm gen} < 50$ MeV there is an upward bias in the reconstructed energy because the modeled electronics threshold is more likely to select those those which have had large upward sampling fraction fluctuations, thus this point is excluded from the fit.

As one might expect, the parameters ϵ and A exhibit z-dependence show in Figure 3. The scale factor, N, shows a strong z-dependence that likely arises from several competing effects:

- Showers on the ends of the BCAL are likely to loose low energy cells because attenuation causes opposite side hits to be below the electronics threshold. This generates the peaking near the middle.
- Showers at higher z enter at a shallower angle. The increased path length in a cell boosts the energy deposition and offsets the first effect above.
- The leakage of energy outside the BCAL is has a z dependence.

The first effect is symmetric about the center of the BCAL while the latter two effects are symmetric about the target. To model this dependence, the scale factor, N, is fit with with a third-order polynomial for photons striking the BCAL downstream of the target ($z >= 65$ cm). For the region $z < 65$ cm, N is assumed constant at $N(z = 65$ cm). Thererefore, we have

$$N(z) = \begin{cases} 0.8015 + 1.740 \times 10^{-3} z - 6.804 \times 10^{-6} z^2 + 7.162 \times 10^{-9} z^3 & \text{if} \quad z \geq 65 \text{ cm}, \\ N(65 \text{ cm}) & \text{if} \quad z < 65 \text{ cm}. \end{cases} \tag{16}$$

For the nonlinearity term, ϵ,

$$\epsilon(z) = 4.441 \times 10^{-2} - 1.183 \times 10^{-4} z + 3.482 \times 10^{-7} z^2. \tag{17}$$

In both cases, z, in centimeters, is the position at which photon strikes the inner radius of the BCAL in the global coordinate system.

With N and ϵ given by the equations above we correct the reconstructed shower energy by setting

$$E_{\rm rec} \rightarrow \left(\frac{E_{\rm rec}}{N} \right)^{\frac{1}{1+\epsilon}}. \tag{18}$$

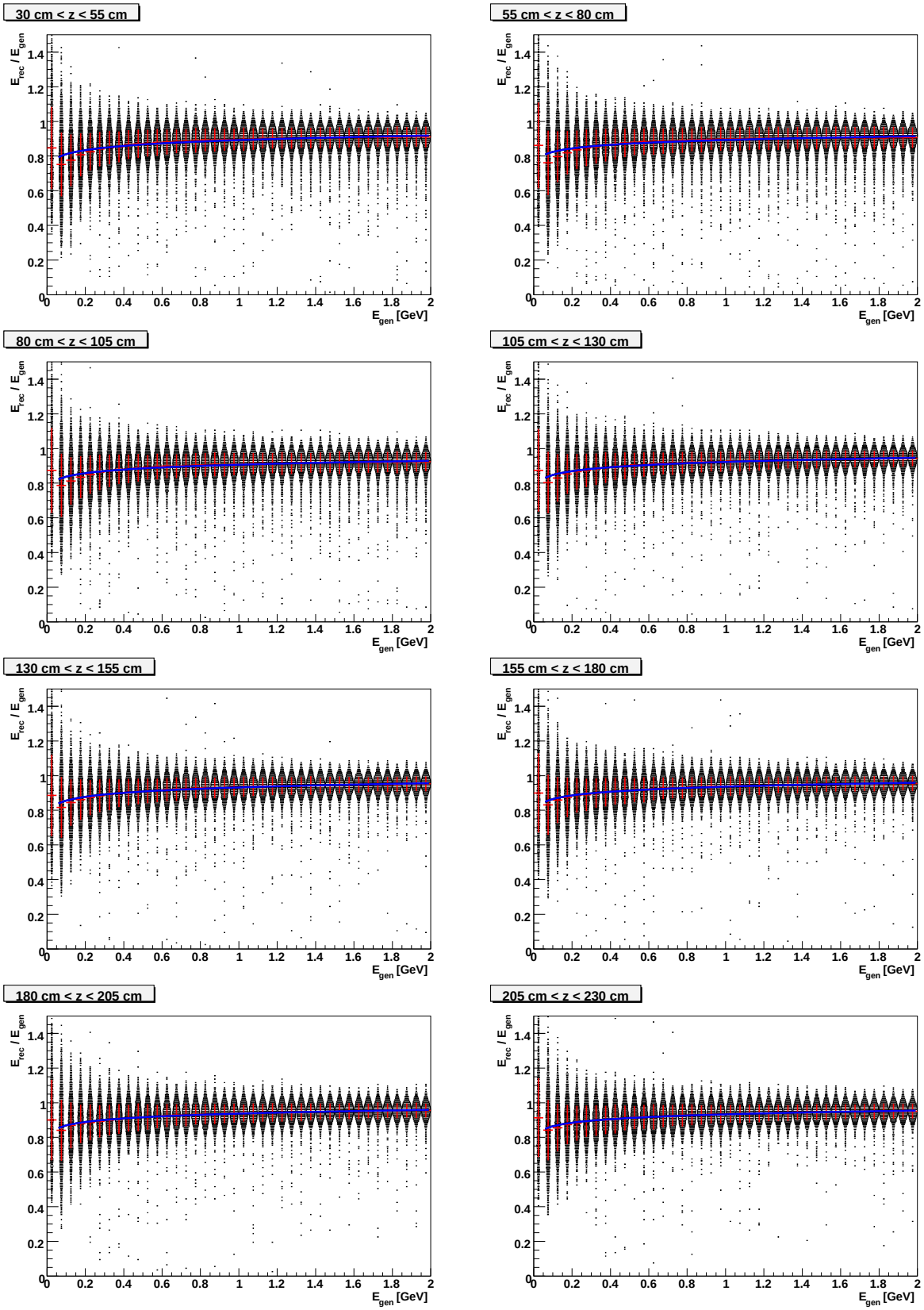
4 Reconstruction Performance

Users wishing to work with calibrated BCAL photons should extract `DBCALPhoton` objects, *not* the underlying `DBCALShower` objects from which they are built. The `DBCALPhoton` objects provide an interface to the location of the cluster in the global coordinate system, the lorentz momentum of the photon, and the results of position vs. layer fit done by the reconstruction code. Preliminary examination of the postion vs. layer fit indicates additional tuning is needed; therefore, we construct the flight direction of the photons by assuming they originate from the center of the target.

4.1 Energy Resolution

To examine the characteristic resolution of the BCAL, we consider photons distributed uniformly in energy between 0 and 2 GeV and uniformly in incident z position in the region 30 cm $< z <$ 380 cm. This avoids the ends of the BCAL where there is excessive leakage[4]. Furthermore we require that generated photons

[4]Additional studies of photon efficiency in the BCAL/FCAL overlap region are underway by M. Kornicer.


30 cm < z < 55 cm
55 cm < z < 80 cm
80 cm < z < 105 cm
105 cm < z < 130 cm
130 cm < z < 155 cm
155 cm < z < 180 cm
180 cm < z < 205 cm
205 cm < z < 230 cm
E_rec / E_gen
E_gen [GeV]

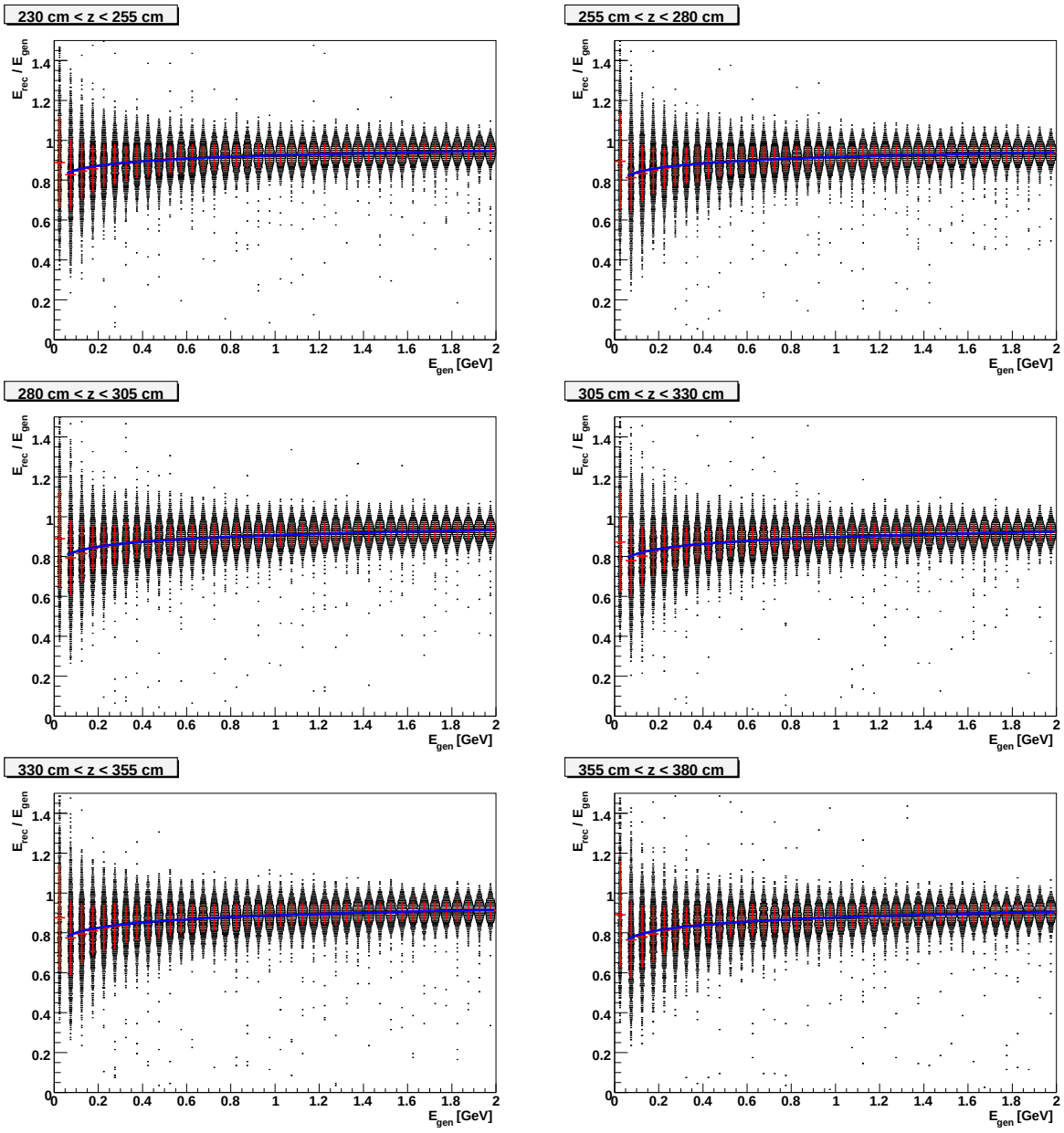


Figure 2: Energy dependence of the ratio of the reconstructed to generated energy for various ranges if z position of incident photons. Data are shown in boxes with an overlaid (red) profile plot. The fit for the > 50 MeV region is to the form $NE_{\rm gen}^{\epsilon}$, where N and ϵ are free parameters.

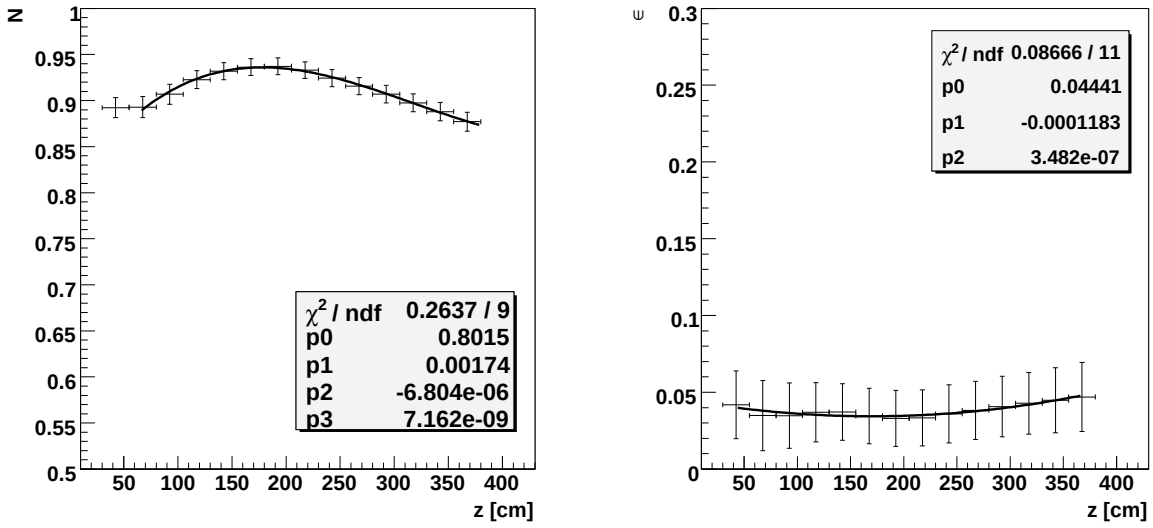


Figure 3: Plots of the z dependence of the energy scale, N, and nonlinearity, ϵ, parameters with fits to third-order (N) and second-order (ϵ) polynomials. *Note:* Error bars on these data points are artificially large due to the method of construction of the profile plots from which the parameters are extracted. Smooth z dependence on a scale much smaller than the indicated statistical error is visible.

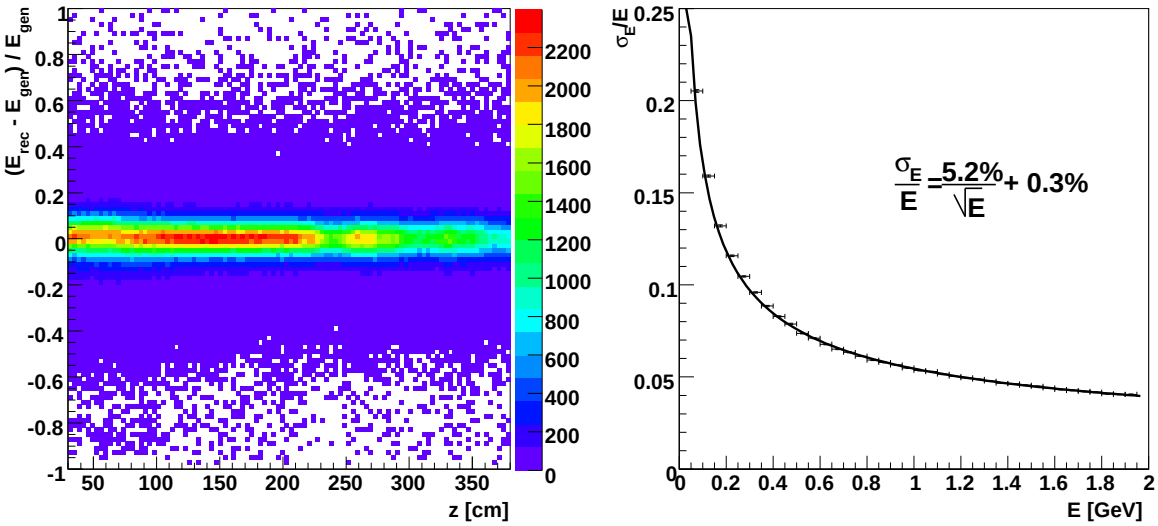


Figure 4: Left: A plot of the fractional energy error vs. z position. There is no noticeable z dependence in the corrected energy resolution. Note that the BCAL is populated uniformly in z; therefore, variations in efficiency introduce variations in the peak height as a function of z. Right: A fit to the standard deviation of the fractional energy error, σ_E/E as a function of energy.

did not convert before striking the BCAL and exactly one BCAL shower is found. Therefore, what follows is an accurate representation of the resolution of cleanly reconstructed photons in the BCAL, but does not reflect efficiency loss or combinatoric background that is produced by conversions. Figure 4 (left) shows the dependence of the fractional energy energy error on z position. The fact that no dependence is observable indicates that the previously mentioned z-dependent energy calibration is effective. The resolution appears to be relatively constant in z. This plot is derived from a sample that was generated uniformly in z; therefore, variations in the peak height are due to variations in the efficiency. Figure 4 (right) gives the characteristic energy resolution of the BCAL and is fit to a function of the form:

$$\frac{\sigma_E}{E} = \frac{A}{\sqrt{E}} + B \tag{19}$$

where the statistical (A) and floor (B) terms are 5.4% and 0.3%, to be compared to what was obtained in the BCAL test beam analysis: $A = 5.5\% - 6.1\%$ and $B = 2.4\% - 3.0\%$ depending on analysis method [3]. It is not surprising that the Monte Carlo resolution is better than beam tests indicate. The Monte Carlo is calibrated on a precisely known photon energy and also does not model all contributions to the resolution. (Some of the missing contributors are discussed in Section 1.5.)

4.2 Position Resolution

At this stage the Monte Carlo does not sufficiently account for effects that degrade the timing resolution of the detector. Therefore, current polar angle, θ, and longitudinal position, z, resolution estimates are optimistic. Figure 5 shows both the absolute error in polar angle and photon impact z position as a function of z position. Again photons are generated uniformly in z and in energy from 0 to 2 GeV. The observed structure in the peak height as a function of z is due to efficiency variation. Errors in polar angle can be written in terms of the (more fundamental) z position error as

$$\delta\theta = \frac{r}{z^2 + r^2}\delta z. \tag{20}$$

One notices from Figure 5 that while the z resolution degrades as one moves downstream due to dispersion of the shower along the BCAL axis the polar angle resolution still improves because of the z^2 factor in the denominator in RHS of the expression above. Give the speed of light in fibers is roughly 17 cm/ns, position resolutions on the order of a couple of millimeters (as indicated in the plots) are clearly optimistic. The simulation needs a more realistic model of the timing resolution of the detector response.

The resolution in azimuthal angle, ϕ, is derived from the shower position in the transverse plane, which is related to the energy-weighted average position of the cells the cluster. Unlike polar angle resolution, azimuthal angle resolution does not depend on timing information and therefore we expect the simulation to accurately model this. Figure 6 shows the absolute error in ϕ for photons evenly distributed in z and in energy between 0 and 2 GeV. The characteristic resolution obtained from this plot of 5.6 mrad is slightly better than the stated design goal of 8.5 mrad [4].

4.3 Efficiency, Conversions, and Shower Fragmentation

Up until now we have concerned ourselves primarily with the resolution for well reconstructed showers in the BCAL. Problems with reconstruction arise typically when photons convert before hitting the BCAL. Conversions can result in completely loss of efficiency (no photon detected in BCAL) or the resulting e^+/e^- pair can produce a pair of showers in the BCAL. Both efficiency loss and spurious showers produced by pair conversions are problematic for event reconstruction. Using the Monte Carlo sample which populates the BCAL uniformly in z with photons from 0 to 2 GeV, we can make a rough estimate of the size of these effects.

Figure 7 (left) shows the point of intersection in the z direction of generated photon trajectory with the BCAL for events with various numbers of reconstructed BCAL showers found. One notices that when

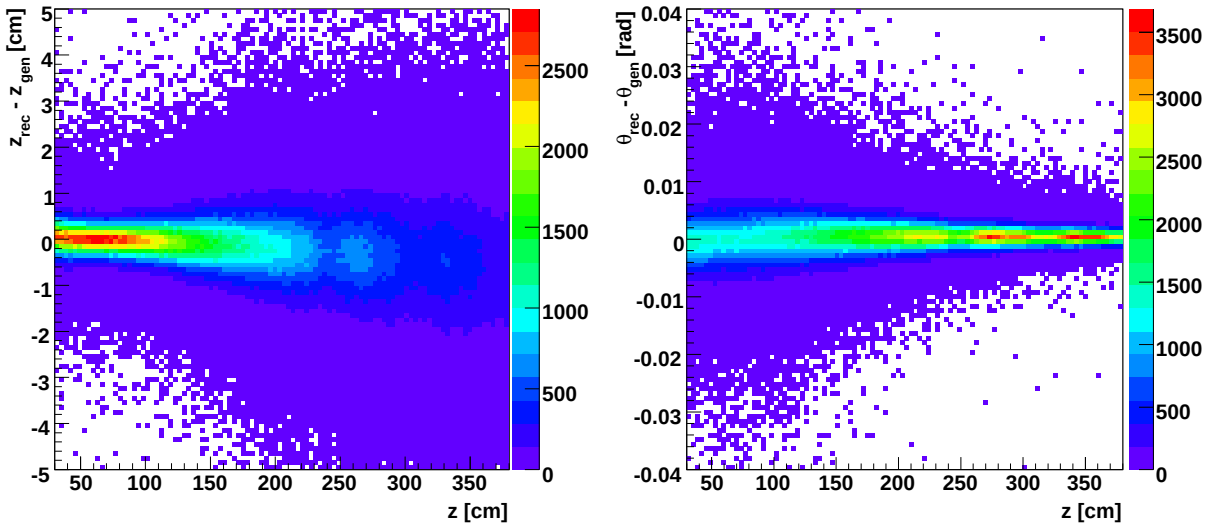


Figure 5: Absolute reconstruction error in z position of photon impact point (left) and polar angle (right) as a function z position of photon impact point.

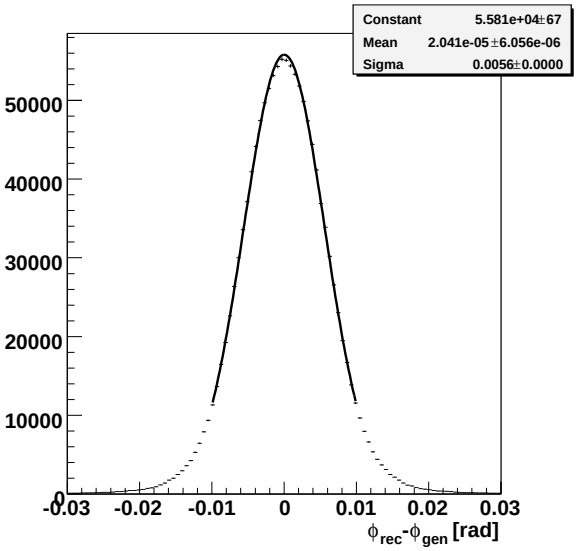


Figure 6: Distribution of absolute error in azimuthal angle reconstruction. The data are fit nicely to a Gaussian with $\sigma = 5.6$ mrad.

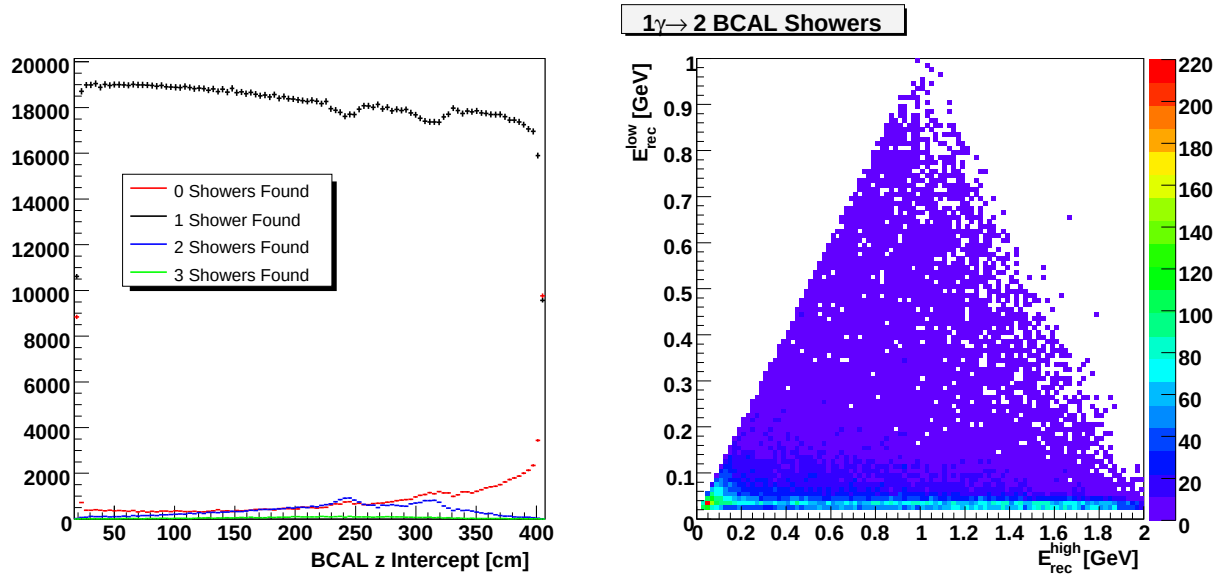


Figure 7: (left) z point of intersection of photon trajectory with BCAL for events in which various numbers of BCAL showers were found. (right) For events with two showers found in the BCAL, the energy of the lowest shower versus the energy of the highest shower.

photons are thrown near the ends of the BCAL they are frequently not reconstructed. In addition one can see structure corresponding to the FDC rings that suppresses the number of events in which one BCAL shower is found and enhances the number of events where zero or two BCAL showers are found[5]. Because the initial parent sample of events is flat in this variable, the efficiency for clean, single shower events is proportional to the black line.

Figure 7 (right) isolates those events in which two showers were detected in the BCAL and plots the energy of the low energy shower versus the energy of the high energy shower. One notices that for the full range of high energy showers, there is noticeable pileup of an additional low energy shower in the 50 - 100 MeV region. These low energy showers are a source of combinatoric background and could affect low energy photon reconstruction. Additional work should be done to examine the spacial separation between these showers and neighboring high energy showers since shower fragmentation could be a symptom of an sub-optimal reconstruction algorithm.

In summary, we have demonstrated that, while additional tuning is needed, the simulation and reconstruction algorithms are performing in a reasonable fashion. Further studies are best carried out in the context of full event simulations where the effect of backgrounds and efficiency can be better understood.

References

[1] C. Xu *et al.* The GlueX BCAL reconstruction code preliminary studies. *GlueX-doc 569*, 2005.

[2] B. Leverington. Regina BCAL standalone MC - sampling fraction fluctuations. *GlueX-doc 827*, 2007.

[3] B. Leverington. Analysis of the BCAL Hall-B beam test. *GlueX-doc 804*, 2007.

[4] B. Leverington *et al.* BCAL facts: what we know and how we know. *Gluex-doc 842*, 2007.

[5] These events were generated by using what was know at the time of writing as "FDC Option 2" for the FDC geometry and material.