On the Measurement of BCAL Photoelectron Statistics

R.T. Jones
University of Connecticut, Storrs, CT

September 28, 2007

Abstract

The most important factor that determines the time resolution of the GlueX barrel calorimeter is the number of detected photoelectrons per GeV deposited by a charged particle. Generally accepted values for scintillation yield and collection efficiency in the BCAL lead to an estimate of 120 detected photoelectrons per end for a 4 cm path length of a MIP at normal incidence. Analysis of cosmic ray data has recently produced a much lower estimate for this figure of about 25 pe/cm. The purpose of this note is to examine the assumptions that are entailed in the cosmic ray data analysis.

Table 1: Estimated photoelectron yield per end of a BCAL module sector that subtends 4 cm of a MIP track.

factor	value	
path length inside active scintillator	1.86	cm
specific energy deposition	2.0	MeV/cm
scintillation light yield	8000	/MeV
fiber capture fraction	9.6	%
attenuation inside fiber	0.43	
fiber-detector coupling factor	100	%
photoelectron detection efficiency	10	%
total expected yield	120	

1 Photoelectron Yield

The method that has been used until recently for estimating photoelectron statistics in the BCAL is illustrated in Table 1 for a readout segment that contains 4 cm of a MIP track. This number is for the geometric mean of the yield on the two ends of the calorimeter and so is independent of the position of the hit along the BCAL module. The estimate includes the photon detection efficiency, but does not take into account any loss at the coupling between the fiber and the end of the phototube.

2 Cosmic Ray Data

A check of this estimate with empirical data has been carried out by the Regina group [1]. They examine cosmic rays that pass at close to normal incidence through a sector of a BCAL module, while depositing negligible energy in the sectors on either side. The fluctuations in the magnitude of the pulse heights from event to event are attributed to the combination of sampling and Poisson statistics, while the differences between the two ends for a single event are attributed to Poisson statistics alone. The comparison between the two ends is done by observing the ratio of the pulse heights from the north and south ends. Attenuation effects on the ratio for cosmic tracks at different positions along the module are correctly taken into account. The result is 22 ± 2 photoelectrons.

One way to explain the apparent discrepancy between measured and expected photoelectron yields is to revise the figures in Table 1. The most obvious weakness is in the fiber - detector coupling factor of 100%. It would be easy to imagine that this factor is more like 75%. It has been suggested [3] that the scintillation yield of the fiber is only 2000/MeV instead of the nominal 8000/MeV. Then if the fiber - detector coupling efficiency is reduced to 75% the discrepancy is fully resolved. In view of the radical jump implied by this measurement for this broadly accepted property of scintillator, it is worth while considering critically the assumptions entailed in the cosmic ray data analysis.

3 Alternative Explanation

The goal in what follows is to examine carefully what is entailed in replacing the 100% fiber + light guide coupling efficiency with a more realistic model. Realistically one cannot assume that the average light from every fiber is uniformly attenuated by 75%. Some fibers will couple with 95% efficiency and some that lie near the edges of the light guides acceptance function may couple with 25% efficiency or even less in the corners. Let us assign a coupling efficiency factor $c_n(i)$ and $c_s(i)$ for the north and south ends of fiber i, respectively. There are correlations between the c factors for the two ends of a fiber but because of geometric imperfections they will not be identical. Accounting for this leads to the fact that sampling fluctuations to feed into differences in the north and south light yields that go beyond mere Poisson statistics. That is, the north/south ratio can no longer be interpreted in terms of Poisson statistics alone.

The proposed model builds on the analysis presented in Ref. [1]. The analysis begins with a sample of MIP tracks that leave an ionization trail 4 cm long in the volume subtended by a particular readout segment. Following the notation used by those authors, the symbol R represents the average value of the north/south photoelectron yields, taking into account the exponential attenuation factor. Let the symbol p_n [p_s] represent the expected number of photoelectrons seen at the north [south] ends, which will later be smeared with Poisson statistics to generate the pulse height spectrum predicted by the model.

$$\begin{aligned} p_n &= y\sqrt{R}\sum_i^n c_n(i)\,\varepsilon(i) \qquad (1) \\ p_s &= y\frac{1}{\sqrt{R}}\sum_i^n c_s(i)\,\varepsilon(i) \end{aligned}$$

The ideal yield factor y corresponds to the product of factors in Table 1 excluding the top two lines and assuming ideal coupling at the fiber-readout interface. The actual coupling efficiency is carried by the $c_{n,s}$ factors. The symbol $\varepsilon(i)$ represents the amount of energy deposited in fiber i in MeV. The sum is over the fibers that lie along the track in the cosmic ray event. There are roughly 30.

The dominant source of fluctuations in p_n and p_s from event to event under fixed trigger conditions is sampling fluctuations, ie. differences in how much total ionization energy is deposited and how that energy is allocated between lead, epoxy, and scintillator. In fact this is what drives the statistical term in the calorimeter energy resolution formula,

$$\sigma(\varepsilon) = a\sqrt{\varepsilon} \qquad (2)$$

where I let $a = 0.06\ \mathrm{GeV}^{\frac{1}{2}}$. Eq. 2 is justified by the following argument. It is known to hold for an entire shower, if ε represents the total shower energy [1]. All of the scintillation light in the calorimeter comes from ionization by electrons and positrons in the shower. Electrons deposit most of their ionization energy in the MIP region. Therefore the calorimeter response to a single passing high energy cosmic ray mimics its response to an electromagnetic shower that deposits the same amount of energy [2]. But the energy deposition fluctuations in different segments of the track are independent. Suppose that the total energy from a passing MIP seen in a readout segment were made of a sum of the energies from many smaller track segments as in $\varepsilon = \sum_i \varepsilon(i)$ and all of the $\varepsilon(i)$ are independent. Then it follows that Eq. 2 also holds for the individual $\varepsilon(i)$ with the same value

---

[1] Actually there is a small floor term that should also be included, but it has negligible effect at low energy and can be neglected.

[2] Cosmic ray events also have so-called Landau fluctuations in the total energy loss in the calorimeter that are not present in electromagnetic showers. This means that Eq. 2 underestimates the variances of p_n and p_s in a MIP sample. However, those fluctuations are completely correlated between north and south ends and do not affect the argument being made with this model.

for a. [3] Therefore, Eq. 2 holds equally for the energy deposited in all of the fibers in a readout sector and that seen by a single fiber.

Using Eq. 2, the variance of p_n due to sampling fluctuations is

$$V(p_n) = y^2 R a^2 \left\langle c_n^2 \right\rangle E + y^2 R n \left\langle \varepsilon^2 \right\rangle V(c_n) \tag{3}$$

where variations in the coupling coefficients c_n from one track to the next are explicitly taken into account. A similar expression applies to the south end, with R replaced with $1/R$ and c_n replaced with c_s. After several steps, this reduces to the following form.

$$V(p_n) = y^2 R\, a^2 \,\overline{c}^{\,2}\, E + 2y^2 R \left(a^2 + \frac{E}{2n} \right) E\, V(c_n) \tag{4}$$

In the last term $a^2 = 3.6$ MeV whereas $E = 4$ MeV and $n = 30$ so the final term proportional to E/n can be dropped. The first term in Eq. 4 takes into account the rescaling of the expected number of photoelectrons due to losses at the fiber-readout coupling, while the last term estimates how non-homogeneities in that coupling amplify the sampling fluctuations with respect to the ideal coupling case. This is very interesting because this exact effect can also be mocked up by modeling these fluctuations as due to Poisson photon statistics instead. That is, if one assumed that $V(c_n)$ were zero and introduced instead the extra Poisson fluctuations by artificially reducing the photoelectron statistics that model should fit the data equally well.

To see what magnitude of variation in the c_n would be required to explain the cosmic ray data in a way that is consistent with the numbers in Table 1, consider the following simple model for the frequency distribution of the c_n.

$$f(c_n) = 30\, c_n^4 \,(1 - c_n) \tag{5}$$

This distribution is peaked at 80% and has an average of 75%. Its variance is 2.6%. The first term in Eq. 4 is completely correlated between north and south ends and so does not contribute to the spread in the ratio. To get an idea of the magnitude of the contribution from the second term, I begin by assuming that the c_n and c_s are completely uncorrelated, just like they were in the Poisson model. This leads to an uncorrelated error $\sigma(p_n)/p_n = \sqrt{V(c_n)}/\overline{c_n} = 0.22$ which corresponds to the Poisson statistics of 22 photoelectrons. This is just about exactly the number of photoelectrons that were reported in Ref. [1] assuming perfectly uniform readout! This precision cannot be taken seriously, given the arbitrariness of the model. In fact, there are correlations between the c_n and the corresponding c_s and Eq. 5 is only a guess for their distribution. In addition, there actually are Poisson statistics to be added in the model presented here, on top of the sampling fluctuations in Eq. 4. The only significant result emerging from this study is that non-uniformities in the fiber-readout coupling do cause sampling fluctuations to feed into the spread of the north/south pulse height ratios, and that reasonable assumptions for the non-uniformities lead to corrections that are large enough to potentially explain the cosmic ray data without a dramatic reduction in the photon yield of the scintillator.

---

[3] The Poisson distribution also obeys this decomposition property, which has led some to assume that Eq. 2 is a consequence of some Poisson process underlying sampling fluctuations. This is not a sound argument. The form of Eq. 2 is not a unique signature of Poisson statistics, but holds for all statistical systems that are self-similar within some range of scales.

4 Conclusions

This author needs to be convince that the figure of 8000 scintillation photoelectrons/MeV in plastic scintillator is really wrong by a factor 4. This number refers to the part of the spectrum that is transmitted by the plastic, not the original VUV emission. Most of the photons from the original scintillation process are in the VUV region and do not go further than a few cm of the source, but the scintillator is doped with a waveshifter that converts most of these into the blue region where the scintillator is relatively transparent. It is these blue photons that are counted as 8000 per MeV deposited. Their absorption takes place on a length scale of order meters, and is described an exponential absorption length. If the yield is really 2000 instead of 8000 then this implies that 3/4 of the photons were absorbed within a few mm of the source, so close that they are not seen in the measured exponential attenuation curve that goes down to a few cm.

A detailed model of the observed width of the distribution of the ratio of the pulse heights at the two ends of a BCAL module exposed to cosmic rays is proposed that takes into account potential non-uniformities in the coupling between the fibers and the readout at the two ends. Making reasonable assumptions about the coupling in this model leads to fluctuations in the ratio that are large enough to explain the cosmic ray data before Poisson statistics are included. Including Poisson statistics at the level expected on the basis of 8000 photons/MeV does not appreciably affect this result. This does not vindicate any particular interpretation; either model can explain the data. However it does show that the cosmic ray data analysis does not reliably extract the number of photoelectrons in the cosmic ray data because it entails unwarranted assumptions about the uniformity of the readout on the two ends. The only thing that can be reliably said is that 22 photoelectrons is a *lower bound* on the yield in the cosmic ray data.

References

[1] A. Semenov, Z. Papandreou, "Analysis of Amplitude Information from 2006 BCAL Cosmics Runs", *gluex-doc-845* 2007.

[2] A. Semenov, Z. Papandreou, "Analysis of Amplitude Information from 2006 BCAL Cosmics Runs, version 070928a", *unpublished* 2007.

[3] G. Lolos, *private communication.*