Angular Reconstruction: Part III

G. Sinnis

March 30, 1998

Preface

The on/offline angular reconstruction algorithm has undergone several modifications. The purpose of this memo is to explain and document the changes and to show the performance of the current fitter. In two previous memos I discussed the angle fitter and modifications to the weighting scheme and sampling corrections. The algorithm for performing the angle fit has not changed. What has changed is the weighting scheme and the sampling corrections. This present round of changes was motivated by recent changes in the calibration constants (see memo of A. Smith), and the changing water depth in the pond. In this memo I will not discuss the different water depths. That result will follow. I will discuss some problems that cropped up while I was investigating the angle fitter and the resolution to these problems.

The Angle Fitter

The current angle fitter is a simple χ^2 minimization over the two direction cosines and the arrival time of the shower plane.

$$\chi^2 = \sum_{i=1}^{NHit} w_i (t_i - t_0 + \hat{i} x_i + \hat{j} y_i + \hat{k} z_i)^2 \tag{1}$$

Minimization with respect to t_0 leads to the following equation for t_0,

$$t_0 = \frac{1}{\mathcal{W}} [\mathcal{T} + \hat{i}\mathcal{X} + \hat{j}\mathcal{Y} + \hat{k}\mathcal{Z}] \tag{2}$$

where:

$$\mathcal{W} = \sum_{i=1}^{N_{Hit}} w_i \tag{3}$$

$$\mathcal{T} = \sum_{i=1}^{N_{Hit}} w_i t_i \tag{4}$$

$$\mathcal{X} = \sum_{i=1}^{N_{Hit}} w_i x_i \tag{5}$$

$$\mathcal{Y} = \sum_{i=1}^{N_{Hit}} w_i y_i \tag{6}$$

$$\mathcal{Z} = \sum_{i=1}^{N_{Hit}} w_i z_i \tag{7}$$

(8)

The Shower Front

It is well known that the shower front is not a plane wave. The wave-front itself is parabolic and the so called "sampling" corrections are dependent on the distance to the core. The width of the shower front is also a function of core distance. The weights (w_i) in the above formulae are functions of the intrinsic width of the shower front (which depends on core distance) and the "sampling" effect of a given timing measurement. therefore, the weights as well as the "Curve-O-Sampling" (COS) corrections are functions of core distance and measured pulse height. In Milagrito we have the added (so far ignored) complication that the weights are also a function of zenith angle. This is because the observed late tail in the timing distributions is larger for larger zenith angles.

The Fitting Method

Instead of performing a non-linear fit to a parabola, we attempt to correct the measured arrival times for "Curve-O-Sampling" and perform a fit to a plane wave. However, one must first perform a correct fit to know the correct corrections to apply. In the past this circular argument has been resolved by

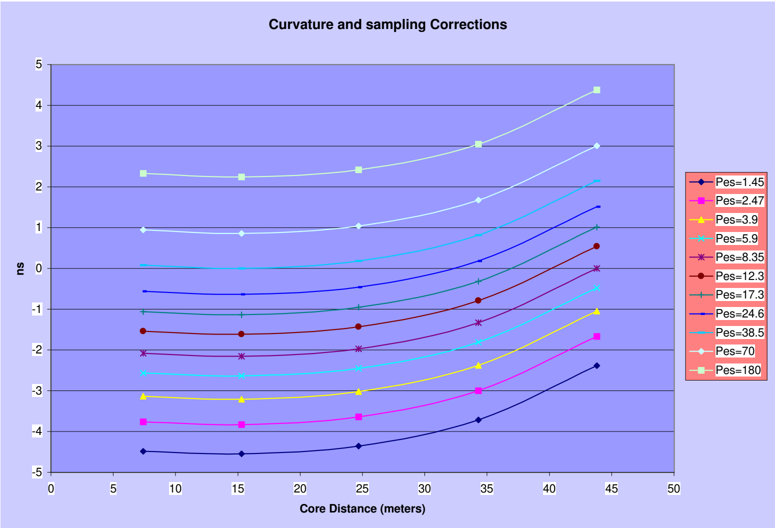

Figure 1: The observed curvature and sampling corrections using the new calibrations with the current weights and no timing corrections.

selecting "good" well-centered showers. Where we have timing measurements all around the shower, and the fitted angle is independent of the curvature correction. One then examines the timing residuals as a function of pulse height (PE) and core distance (R). The positions of the peaks of these distribution are taken to be the COS corrections, and the widths of these distributions are taken to be the inverse of the square root of the weights. When this technique was applied to the Milagrito data with the original calibration constants there was no evidence of curvature. The position of the peaks was independent of core distance.

When this technique is applied to the data with the current set of calibration constants there was clear dependence on the core distance. Figure shows the observed COS corrections. Note that larger times correspond to earlier arrival times. The figure clearly indicates negative curvature. I repeated this method using Monte Carlo data. The results were identical. When I used the fit angle to determine the COS corrections the result was negative curvature. I believe that this is an indication of a systematic error in the angular reconstruction.

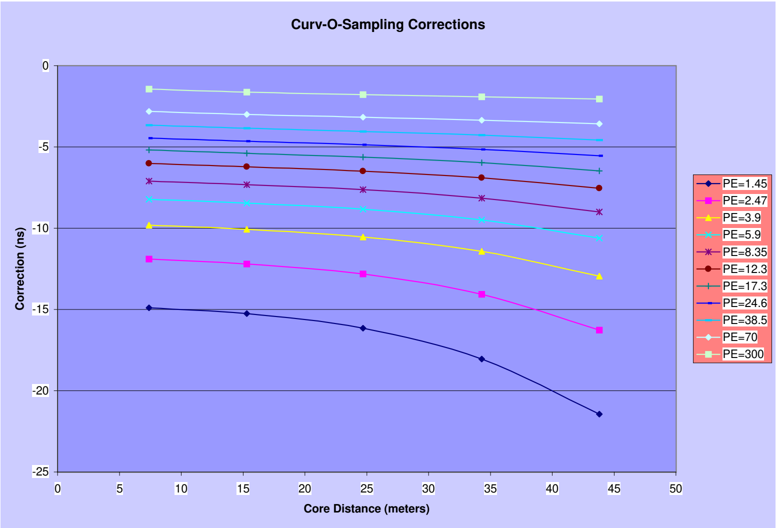

Figure 2: The curvature and sampling corrections using $t_0 = t_{HOT}$, and the true shower direction with Monte Carlo events.

The Problem and Solution

In our fitting algorithm the determination of the direction is correlated with the the determination of t_0. However, when we first calculate t_0 we get the wrong answer. Until we know the weights, and the COS corrections (because the t_i's are the corrected times), we can not correctly calculate t_0. I chose to redefine t_0 as the time that the hottest tube was hit. This removes the coupling between t_0 and the weights and the COS corrections. However, it also leads to a rather poor determination of t_0. I then looked at the timing residuals (t_χ) with respect to the true direction (using Monte Carlo generated showers). I made 130 histograms (10 radial bins $\times$ 13 pulse height bins). For each histogram I fit a Gaussian around the position of the peak. I then plotted the both the peak positions and the sigmas as a function of pulse height and core distance. The results are shown in Figure for the peaks and Figure for the sigmas. We clearer have true positive curvature. These data were then fit to the following function:

$$\mathcal{F} = a + bR^{\alpha} + cPE^{\beta} + dR^{\gamma}PE^{\delta} \tag{9}$$

But, the above widths are wider than they should be. This is due to the poor determination of t_0, in particular the fact that it is not corrected

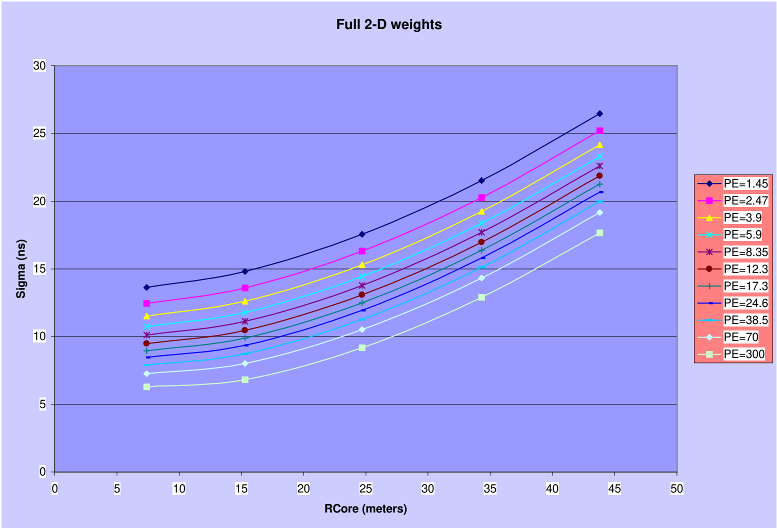

Figure 3: The sigmas of the timing distributions as a function of pulse height and core distance using $t_0 = t_{HOT}$, and the true shower direction with Monte Carlo events.

for arrival direction. Using the functions determined for the sigmas and the COS corrections I refit the showers with the original definition of t_0. The distributions were now quite a bit narrower and the peaks were not centered at zero. I performed one more iteration of the data using the same technique described above (fitting the 130 histograms to Gaussians, and again fitting the peaks and sigmas to the function given in equation 9. I show the net results in Figure for the peaks and Figure for the sigmas. The function given in the figure is a *net* result, and so represents the observed COS corrections. Note that there is still some evidence of negative curvature for the moderate to low pulse height regime near the core. But, the effect is quite small ($\sim$0.01 ns/m) and for large pulse heights and for small pulse heights far from the core we see a slight positive curvature.

Results

I applied the above corrections to both real data and Monte Carlo generated events. Over all of our data this fitter fits $\sim$92% of the events (the old fitter had yield of $\sim$88%). The results are shown in the following plots. Figure

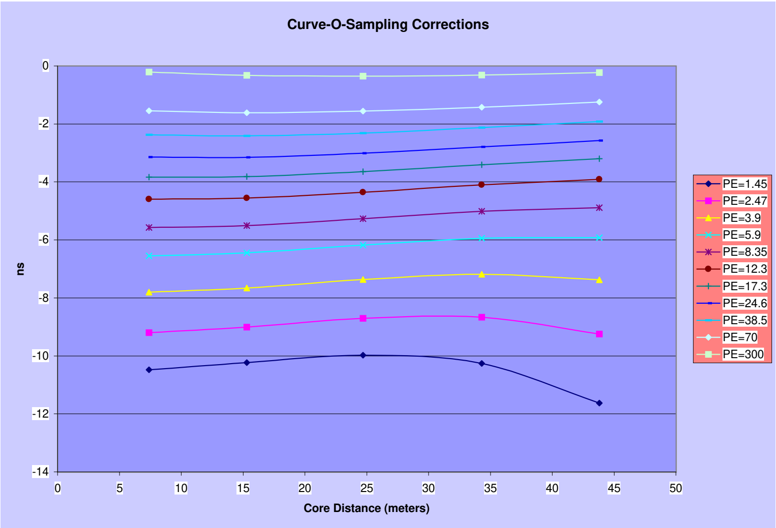

Figure 4: The curvature and sampling corrections using the original definition of t_0, and the fit shower direction with data.

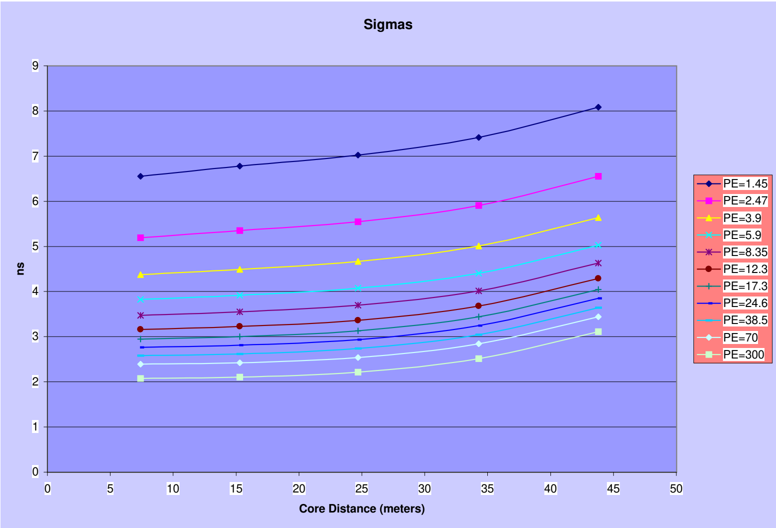

Figure 5: The sigmas of the timing distributions as a function of pulse height and core distance using the original definition of t_0, and the fit shower direction with data.

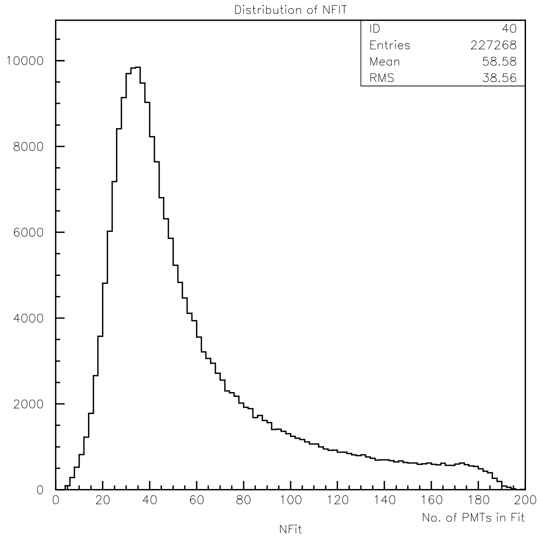

Figure 6: The distribution of NFit, the number of PMTS in the fit, from data.

shows the resulting distribution of NFIT the number of PMTs in the fit. The peak is now near 40 and the mean is near 60. This is in stark contrast to our previous result (see my first memo on angular reconstruction). Figure shows the same plot from the Monte Carlo. In this and all subsequent Monte Carlo plots I chose a trigger requirement of 80 PMTs. The peak position is slightly higher and the mean is significantly higher, probably due to the larger number of PMTs in the Monte Carlo (simulated PMTs never die).

The next plots show the distribution of DELEO for data, Figure and Monte Carlo . The peak is near 1.0 as it was in the past, however the mean is about 10% smaller (3.0 as opposed to 3.3). I have superimposed the previous result (dotted line) for DELEO on the data plot. The plots have the same total number of events (but for the old plot more of the events are in overflow). Both histograms use the new calibrations. The DELEO distribution derived from the Monte Carlo showers is markedly better than that derived from the data. However, some of the Monte Carlo runs used had their cores on the pond. So this should not yet be a cause for alarm.

The next two plots show the mean DELEO vs. NFIT (the number of PMTs in the fit). The numbers are higher then in the above plots because the events with DELEO larger than 10 degrees are not included in the averages

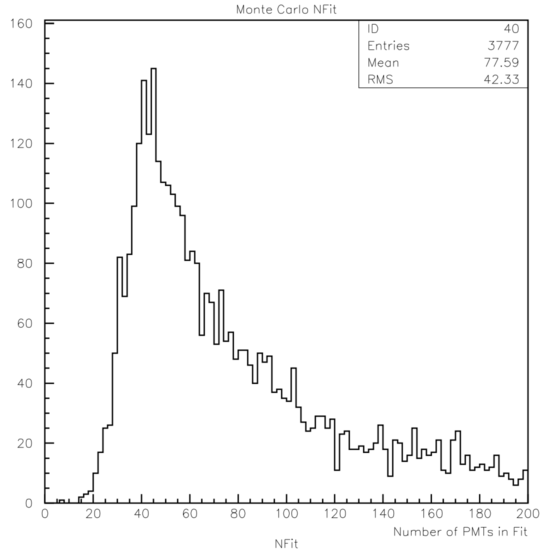

Figure 7: The distribution of NFit, the number of PMTS in the fit, from Monte Carlo.

calculated above, but they are included in these plots. There is a long tail to our angular resolution function.

For completeness I show DELEO vs. NHIT and then NFIT vs. NHIT (both for data and Monte Carlo). And DELEO vs. zenith angle.

Further Investigations

Using the Monte Carlo I have looked at ΔAngle, the space angle between the true and the fitted direction. And how this varies with the error in the core position and with additional curvature corrections. Figure shows the distribution of psace angle difference between the true and fitted directions. It is significant that this distribution has a similar peak and mean to the distribution of DELEO. Remember DELEO should be divided by 2, but not ΔAngle. Therefore our angular resolution is limited by our ability to locate the core, not by any other problems in our angle fitter (i.e. late tails, etc.). The other problem with poor core location is that it also limits our resolution even when we know where the core is. This is because our derived COS corrections and weights are averaged over all showers, those in which

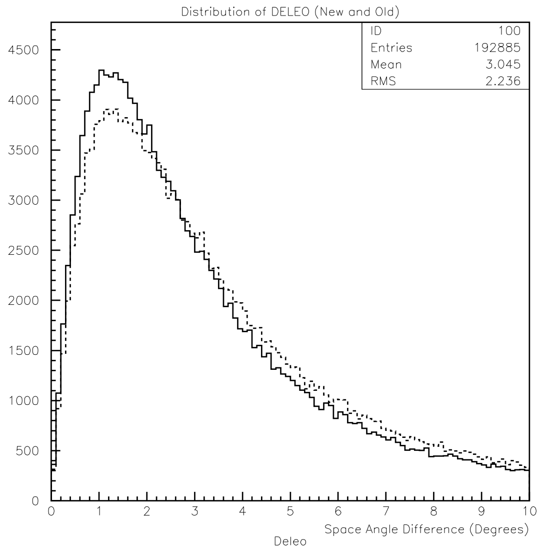

Figure 8: The distribution of DELEO. The solid histogram uses the new COS and weighting corrections, the dashed histogram uses the old sampling and weights. Both histograms were made using the new calibrations. From data.

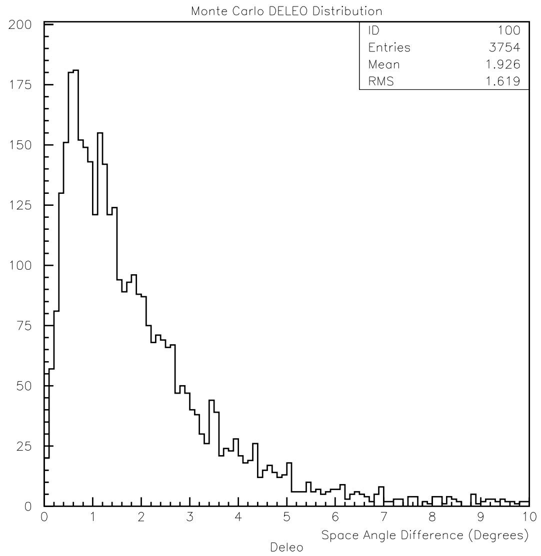

Figure 9: The distribution of DELEO. From Monte Carlo.

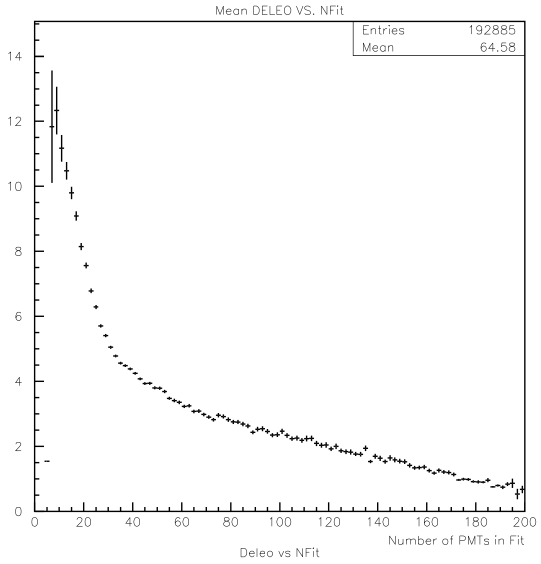

Figure 10: The mean value of DELEO as a function of NFit, data.

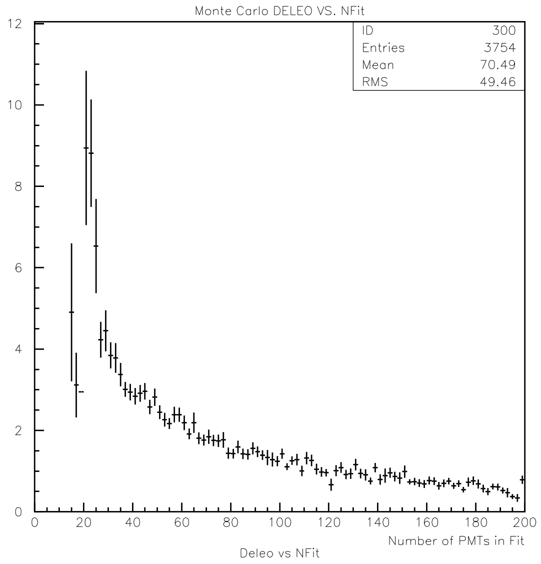

Figure 11: The mean value of DELEO as a function of NFit, Monte Carlo.

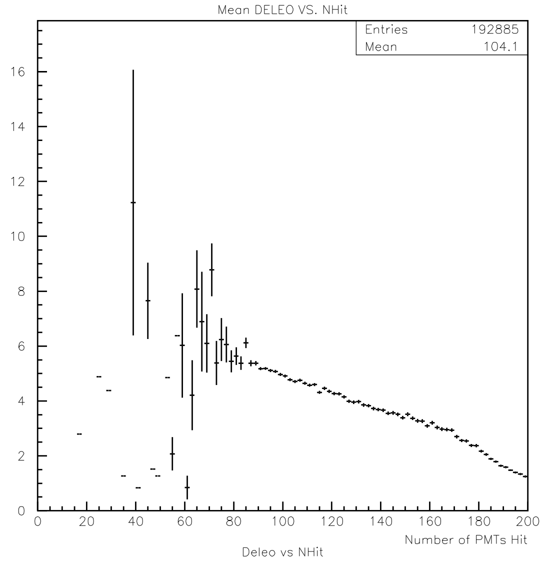

Figure 12: The mean value of DELEO as a function of NHit, data.

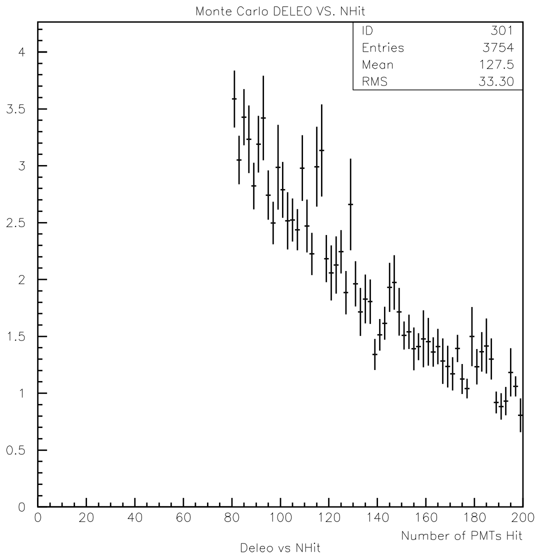

Figure 13: The mean value of DELEO as a function of NHit, Monte Carlo.

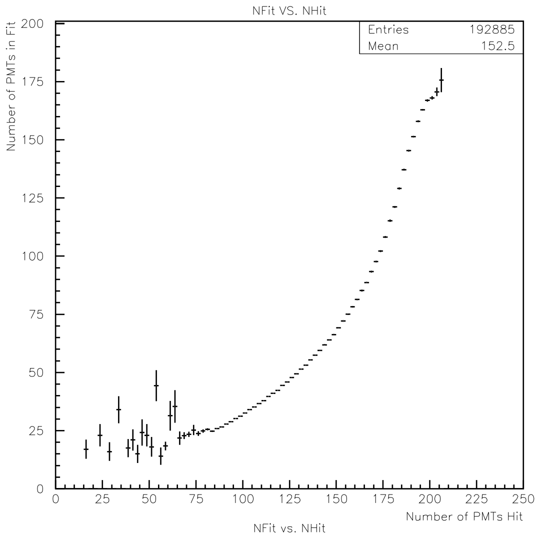

Figure 14: NFit vs. NHit, data.

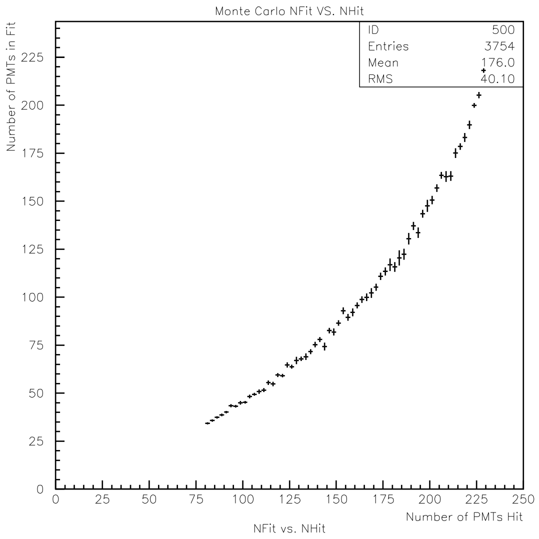

Figure 15: NFit vs. NHit, Monte Carlo.

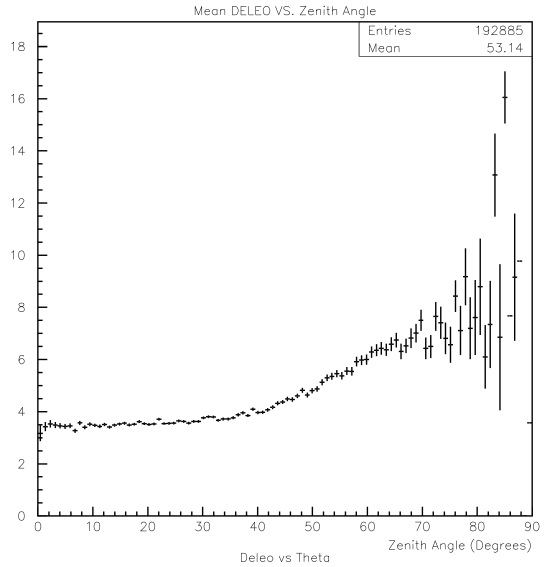

Figure 16: The mean value of DELEO vs. zenith angle, data.

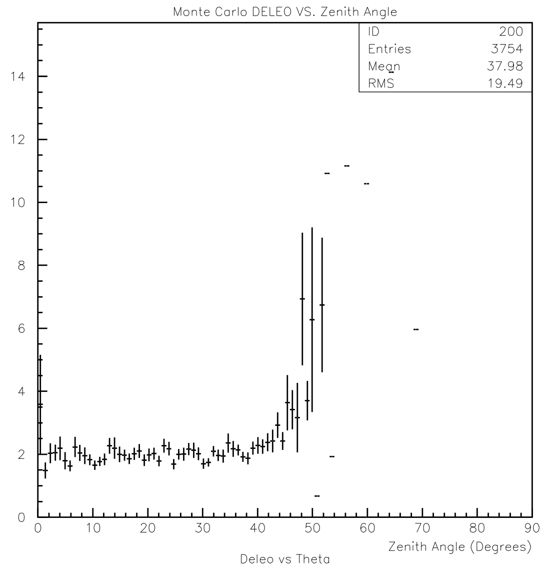

Figure 17: The mean value of DELEO vs. zenith angle, Monte Carlo.

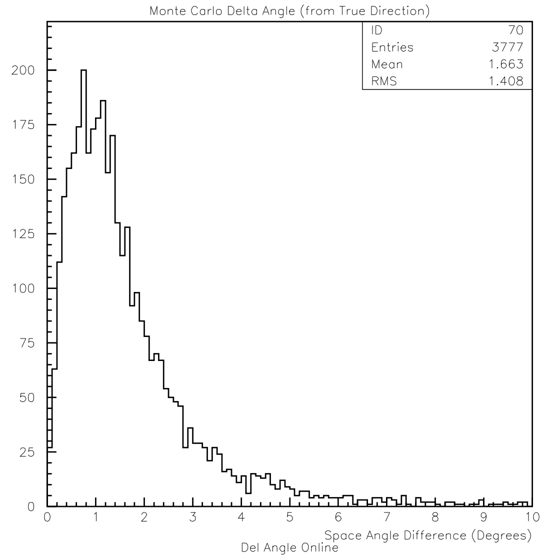

Figure 18: From Monte Carlo, the distribution of space angle differences between the fit and true shower direction (ΔAngle).

we know the core location and those where we do not. If (using the Monte Carlo) I derive COS corrections and weights using the true core location I get significantly different results. There is a *much* stronger dependence on core position. Thus, we are fitting our good events in a highly non-optimal fashion. What we need are 2 sets of COS corrections and weights - 1 set for events were we know the core position and another for those events were we do not know the core position. Our current core location algorithm does not supply a robust indicator of core knowledge.

Stefan discovered another systematic error in our angle fitter. The zenith angle of fitted events is systematically higher then the true zenith angle. I investigated this effect with the new corrections. Figure shows the average difference (θ_{fit} - θ_{true}) versus θ_{true}. We still observe a systematic of $\sim$0.2 degrees. Stefan has made a similar plot for baffled data and the effect disappears. Thus, it seems that the effect is due to the late tails in the timing distributions. It would be interesting to see if the effect disappears when one uses a maximum likelihood method for fitting directions.

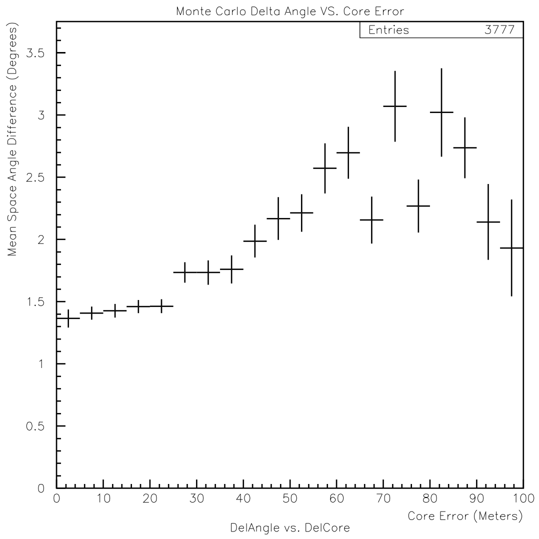

Figure 19: From Monte Carlo, the mean value of ΔAngle as a function of the core error.

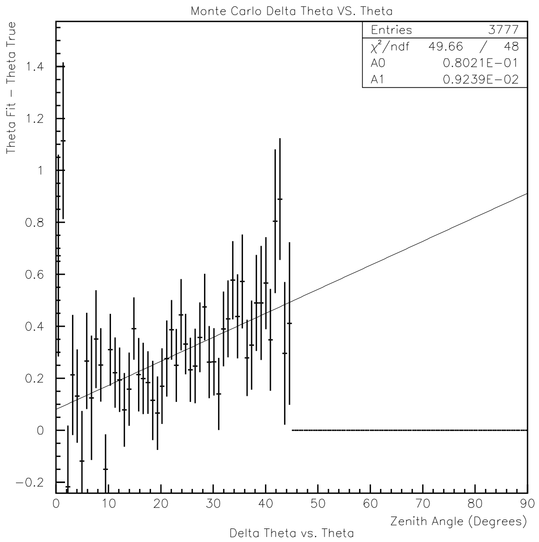

Figure 20: From Monte Carlo, the mean value of $\theta_{fit} - \theta_{true}$ vs. θ (zenith angle.)

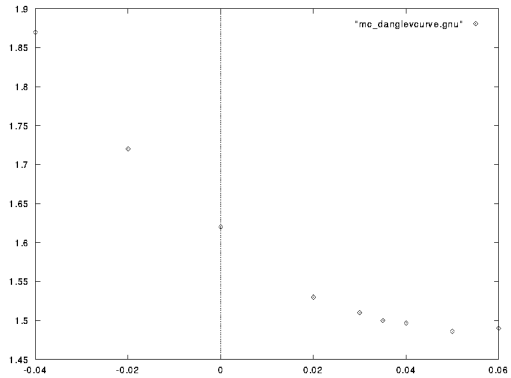

Figure 21: From Monte Carlo, the mean value of ΔAngle vs. the value of an additional curvature correction.

What Happened to the Curvature?

Using the Monte Carlo to generate the t_χ distributions we saw clear positive curvature. But, after using the data to re-generate the t_χ distributions essentially all of the curvature was washed out by the poorly fit showers. I went back to the Monte Carlo and looked at the distribution of ΔAngle (space angle difference between true and fit directions) as a function of an additional curvature correction (pulse height independent). The results are shown if Figure . As before ΔAngle is best for an additional curvature of 0.04 ns/m. This is not surprising since the Monte Carlo wanted this level of curvature to begin with. However, I think the data is telling us that we do not locate the core well enough to be using any "curvature correction".

Conclusions

Using the new calibration constants and redoing the COS corrections and weights I was able to make a $\sim$10% improvement in the mean of *deleo* and we now fit about 92% of the events. The Monte Carlo yields significantly better resolution than the data, perhaps because some of the Monte Carlo showers were thrown only over the pond. The resulting corrections and weights are an average over poorly and well fit showers (core fit), so using them we do not do the best job possible on good showers. Given the similarities between the ΔAngle and *deleo* distributions in the Monte Carlo it seems that further improvements in the fitting algorithm will not improve our angular resolution. We await outrigger counters and their knowledge of the core location. The previously observed θ systematic is still evident in our data. The current fitter systematically fits events at a larger zenith angle ($\sim 0.2^{\circ}$).