Milagro Memo
Gus Sinnis
2/9/00

Note on Removal of the Unwanted Hits in Milagro Events

Introduction:
The look-back time of the TDCs in Milagro is 1500 ns. With ~450 PMTs firing at ~17 kHz each, we expect to have ~10 PMTs hit in each event that have nothing to do with the air shower that triggered the detector. The current angle fitter does a good job of fitting the direction to the events if the "first guess" angle is near the true direction of the shower. However, the presence of 10 unassociated hits can lead to an incorrect first guess, and therefore a poor fit to the true direction of the primary particle. Recently several people have studied this effect and arrived at methods for removing these hits. See the memos by Guarang Yodh, Todd Haines, and Javier Bussons-Gordo. In this memo I demonstrate a trivial method (which actually speeds up our existing code) that gives essentially the same results as these methods. At David's request I have included a comparison of the new calibrations from Roman and Lazar (slewing with extension) and TOT-to-PE from Andy to the calibrations we are currently using online. There is a pretty dramatic improvement, in both cases I used the new edge finder.

The Problem:
In Figure 1 I show the start times of all of our data before the calibration routine is called. Figure 2 shows the start times after the calibration. Note that though the second distribution is much narrower there is a long tail associated with the random hits that come with every event. The first question is, "What are they doing there?". In Milagrito Scott Hugenburger had written a routine to find the good edges in every event. This routine required that the determined to be the starting edge be within 200ns of the trigger time. In the Milagro implementation of this routine that requirement was waived if the PMT had exactly 2 or exactly 4 edges. My trivial

method consists of simply removing the 2 lines of code in the existing edge finder that waive this requirement. Figure 3 shows the starting times after this code modification and also after the DOA cut has been applied to the data with the current edge finder. The effect is roughly identical.

Performance:

I will measure the performance of the method by comparing our sensitivity to a point source using the method described in my previous memo on determining the optimal bin size and NFIT cut for a binned analysis of the data. The one caution is that one must ensure that one is comparing apples to apples. For example if I use the current edge finder roughly 20% of the events do not pass deleo. Thus, deleo (which I use to parameterize our angular resolution) is biased because it does not include our worst events. Thus a naïve application of the method would lead to the conclusion that our current edge finder is better. If however, we compare the methods on the same events (by requiring that they pass deleo with the current edge finder) we have a fair estimate of the relative sensitivity of the different methods. However, to find the optimal bin size and NFIT cut we must use all the events. Thus a method, which passes more events through deleo, will lead to a better determination of the optimal bin size and NFIT cut.

The table and figures below compare 3 methods for the removal of bad hits: the current edge finder, the modified edge finder, and Javier's DOA cut applied to the current edge finder. The analysis was performed on a GRB file (all-sky / all data) taken when the detector was running at 1200 Hz.

Method	Fraction DELEO	Fraction Fit	Optimal bin size	Optimal NFIT cut	Significance of Excess
Current Edge	79.5%	91.0%	1.1	0	7.1
New Edge	90.0%	96.0%	1.2	0	7.03
DOA	94.9% (99%)	98.0%	1.3 (1.1)	0 (30)	6.9 (7.1)

In the table above the numbers in parenthesis are for the subset of events that pass deleo with the new edge finder. Thus Javier's cut is actually better than the modified edge finder, though not by much. However, the numbers returned by Javier's cut for the optimal bin size and NFIT cut are closer to the true best values, since his cut retains the largest fraction of the events in the histogram of deleo.

Figure 3 shows the significance versus bin size for the new edge finder and for the data with the DOA cut applied. Also shown is the DOA cut applied data but the deleo distribution used in the data consists of events that pass for only events that yield a deleo with the new edge finder. All histograms are made for the value of the NFIT cut that maximizes the significance of the signal for that particular analysis and are indicated in the caption. Figure 4 shows the significance of a signal versus the NFIT cut used for the 3 different analyses. In this plot the bin size is chosen to maximize the significance of the signal for each analysis. The dotted line deserves a little further explanation. As explained above if an event fails to provide a value of deleo it does not go into the calculation of our point-spread function. This creates a bias in that the (perhaps) worse analysis can seemingly give the best result, because only the best events yield a value of deleo. The dotted plot is meant to address this concern by selecting the same set of events (for the new edge finder and the DOA applied analysis) to determine deleo. This shows that the DOA cut gives slightly better performance then a modified edge finder.

Conclusion:

The veracity of the angle fitter is highly dependent on the quality of the initial guess of the direction. The "noise" hits in the pond make it difficult to guess the "correct" initial direction. Several memos have been written about how to remove the "noise" hits and therefore improve the angle fitter. Nearly all of the gain realized in these more sophisticated methods can be obtained with a trivial modification to the existing code.

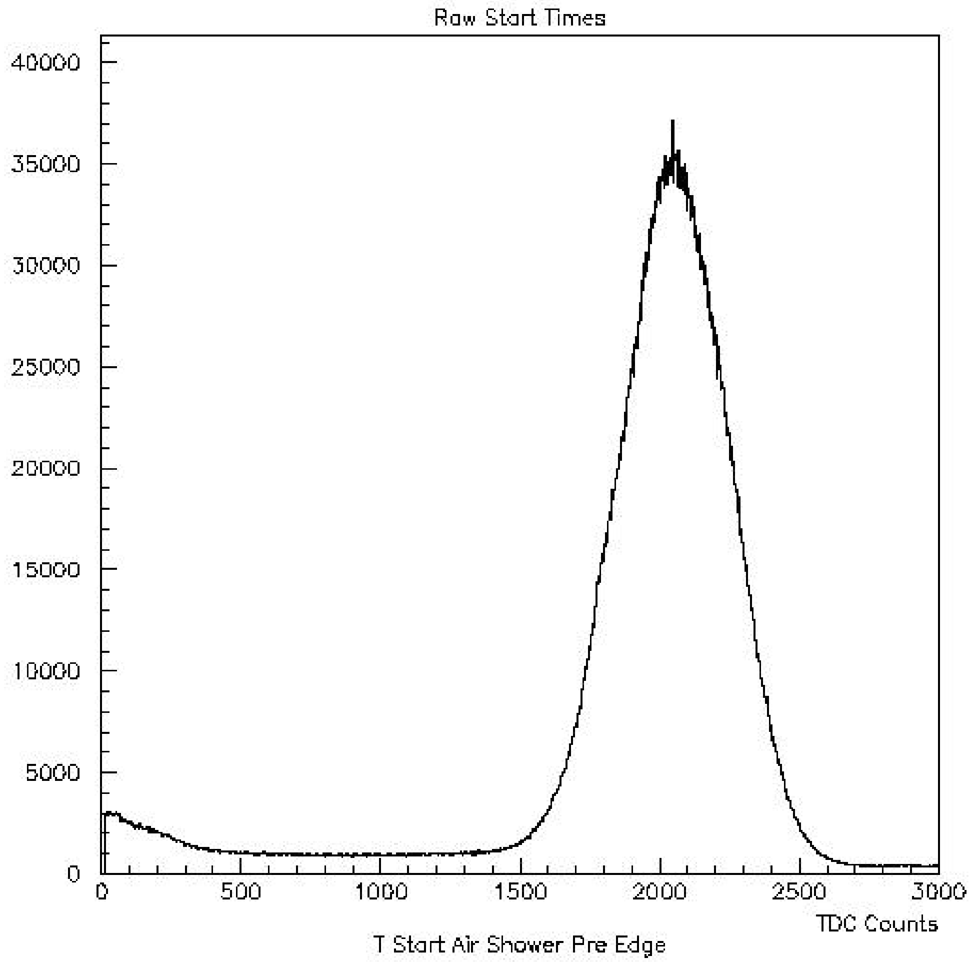

Figure 1 Distribution of start times for all hits (uncalibrated) in the pond.

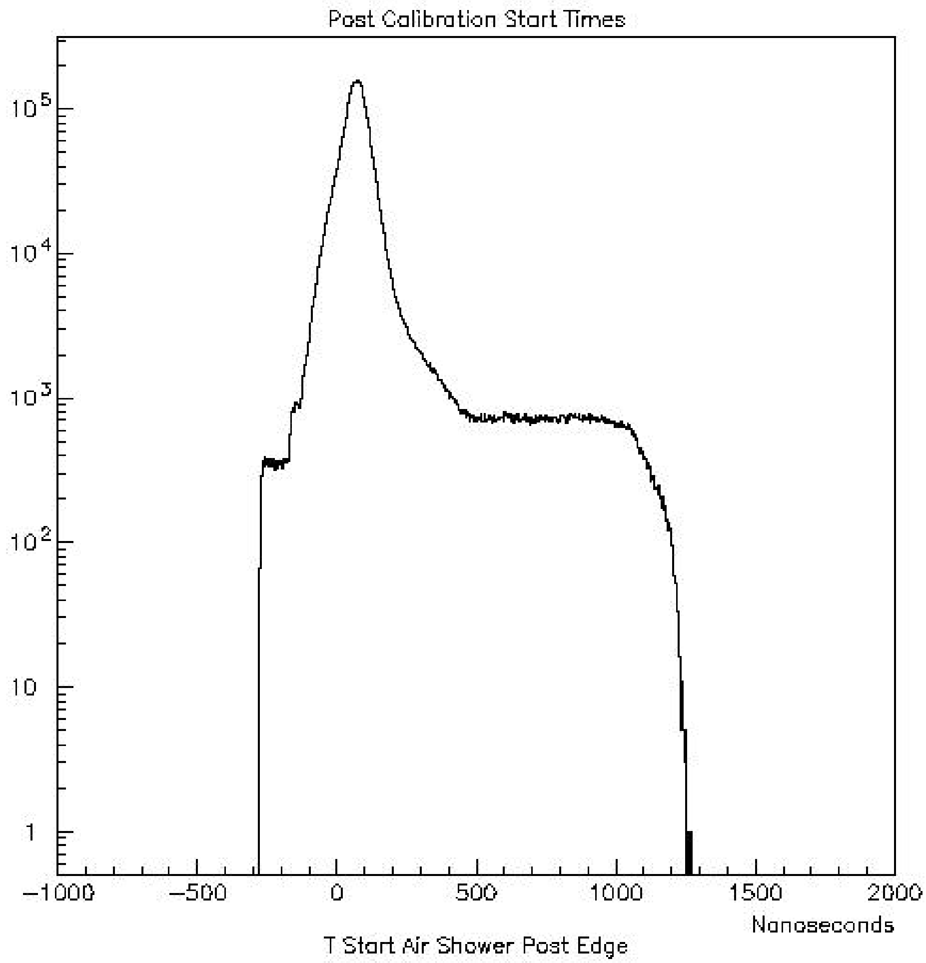

Figure 2 Distribution of start times for hits after calibration (and edge finder) for current version of the edge finder.

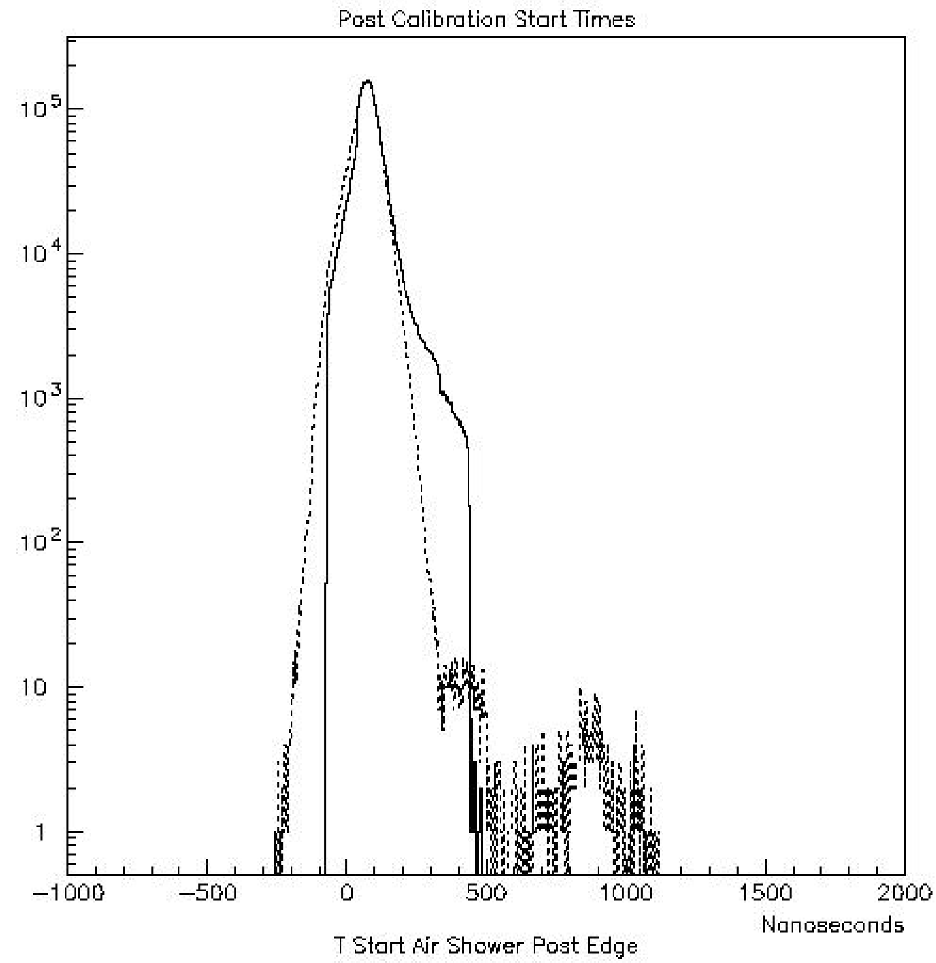

Figure 3 Distribution of start times after calibration with new edge finder (solid line) and old edge finder with DOA cut applied (dashed line).

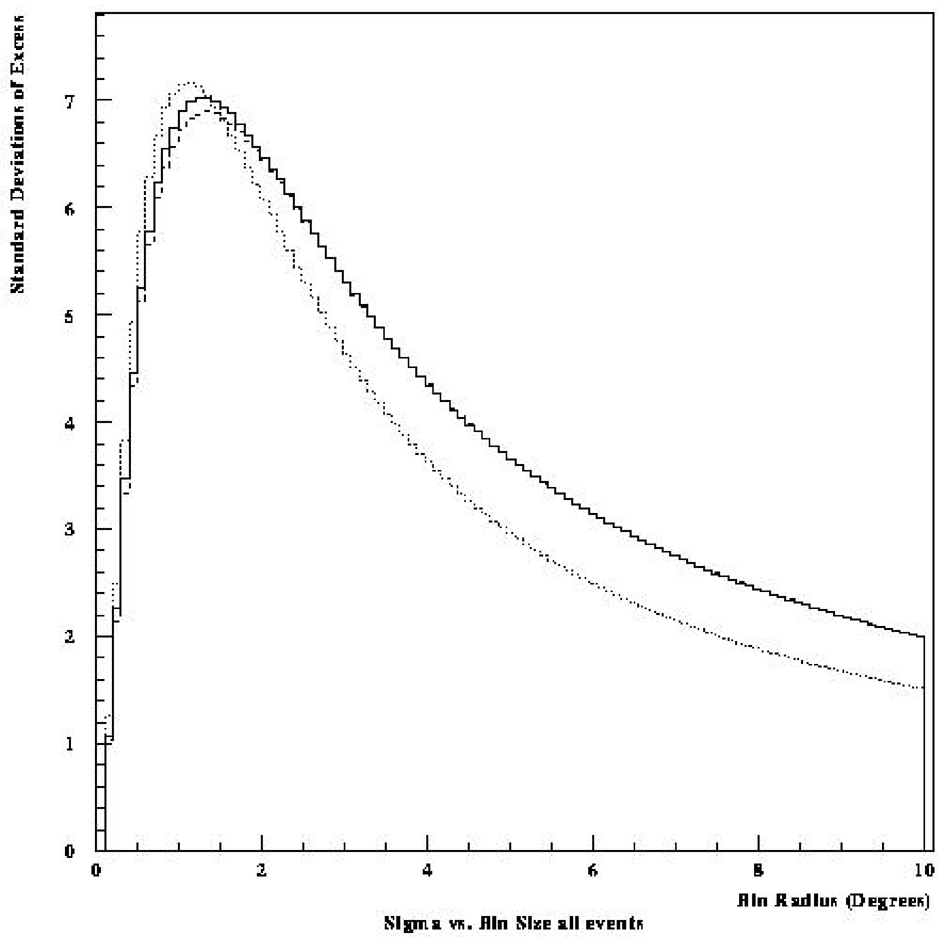

Figure 4 Significance of an excess with for three different analyses. Solid line is new edge finder for NFIT>=0. Dashed (Dotted) is old edge finder with DOA cut for NFIT>=0 (30 AND deleo based on event sthat pass deleo for solid line.

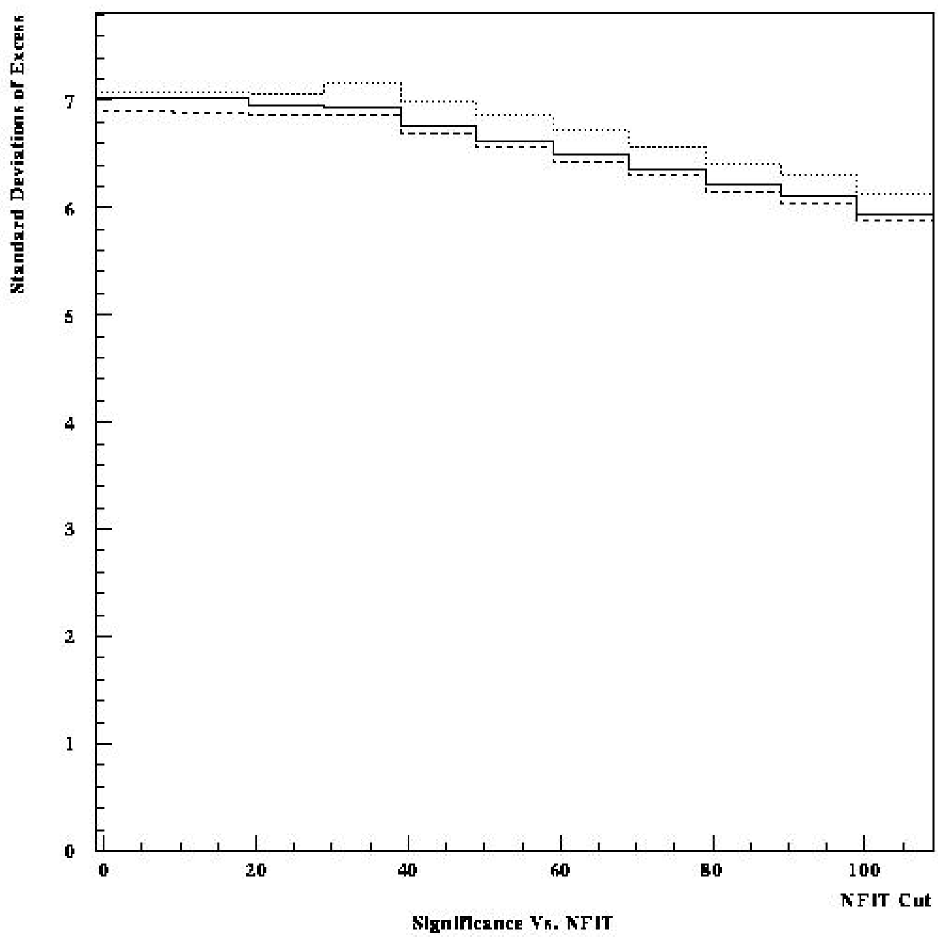

Figure 5 Significance of an excess vs. NFIT cut for same three analyses given in Figure 4. Same interpretation of lines.

The following four plots compare the new calibrations with the calibrations currently being run online. Figure 6 shows deleo for all events (no NFIT cut). This is the optimal cut in both cases. Figure 7 shows the radius of the optimal circular bin versus the NFIT cut for the two calibrations. Figure 8 shows the significance of a signal as a function of NFIT for the two calibrations (at each NFIT I selected the optimal bin radius), and finally Figure 9 shows the significance of an excess versus the bin radius for the optimal NFIT cut. There is roughly a 17% improvement with the new calibrations.

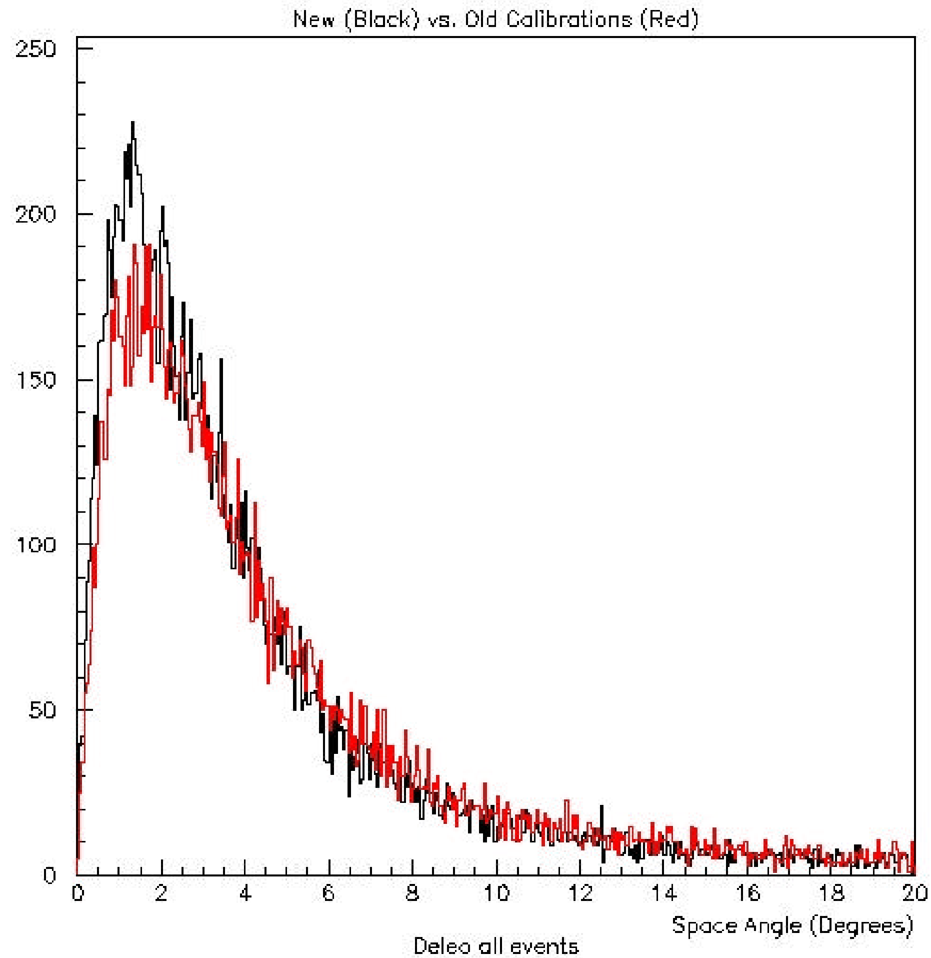

Figure 6 Deleo for all data. Black line is new calibrations and red line is current calibrations.

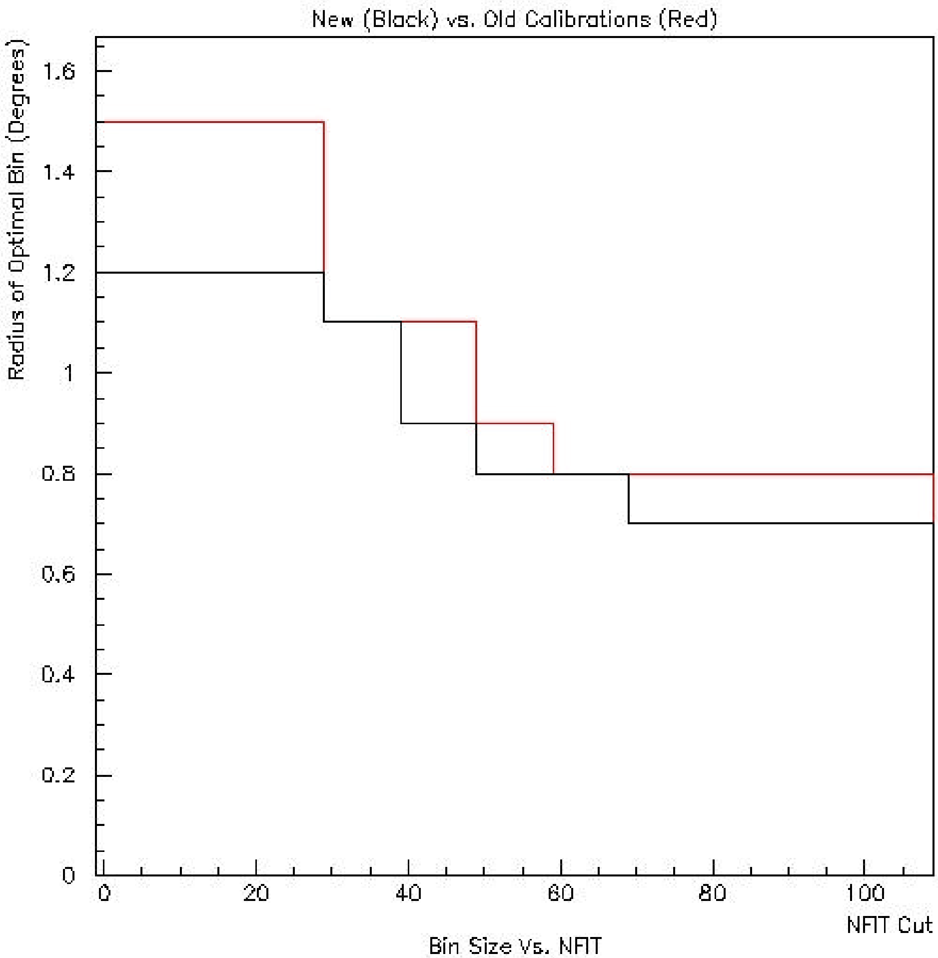

Figure 7 Optimal bin size as a function of NFIT for the new (black) and old (red) calibrations.

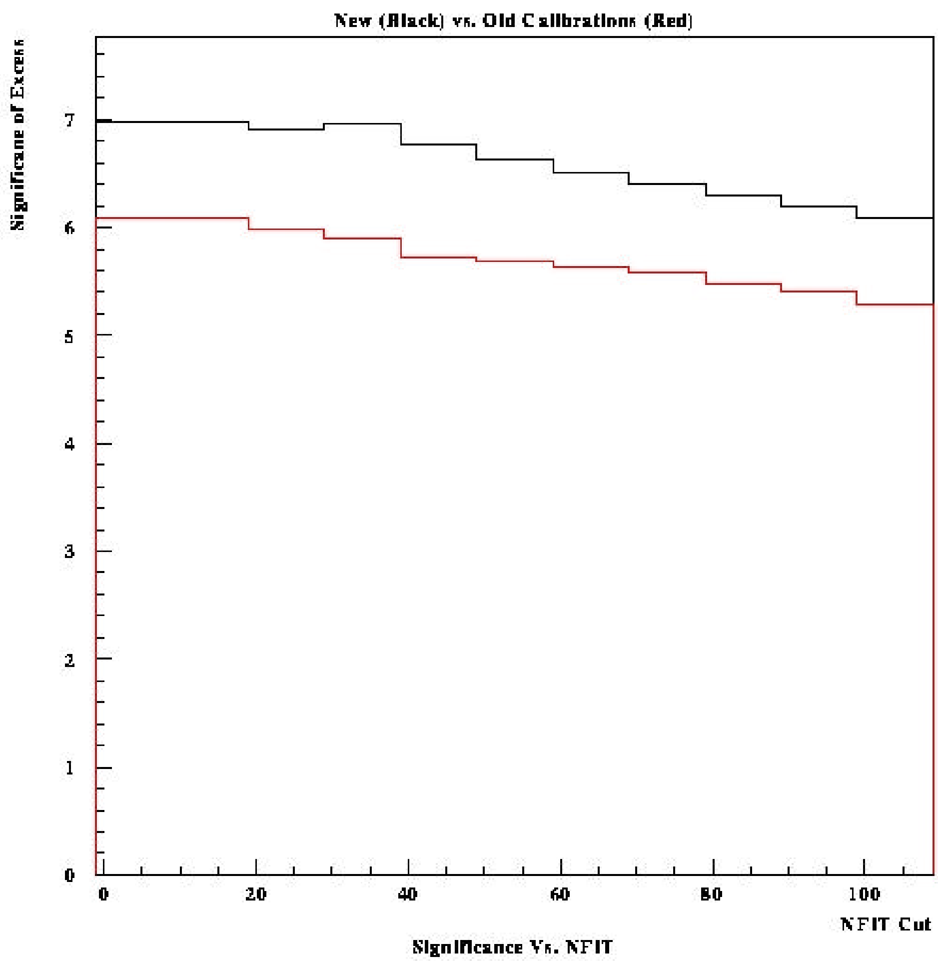

Figure 8 Significance of an excess as a function of NFIT cut for the new (black) and old (red) calibrations.

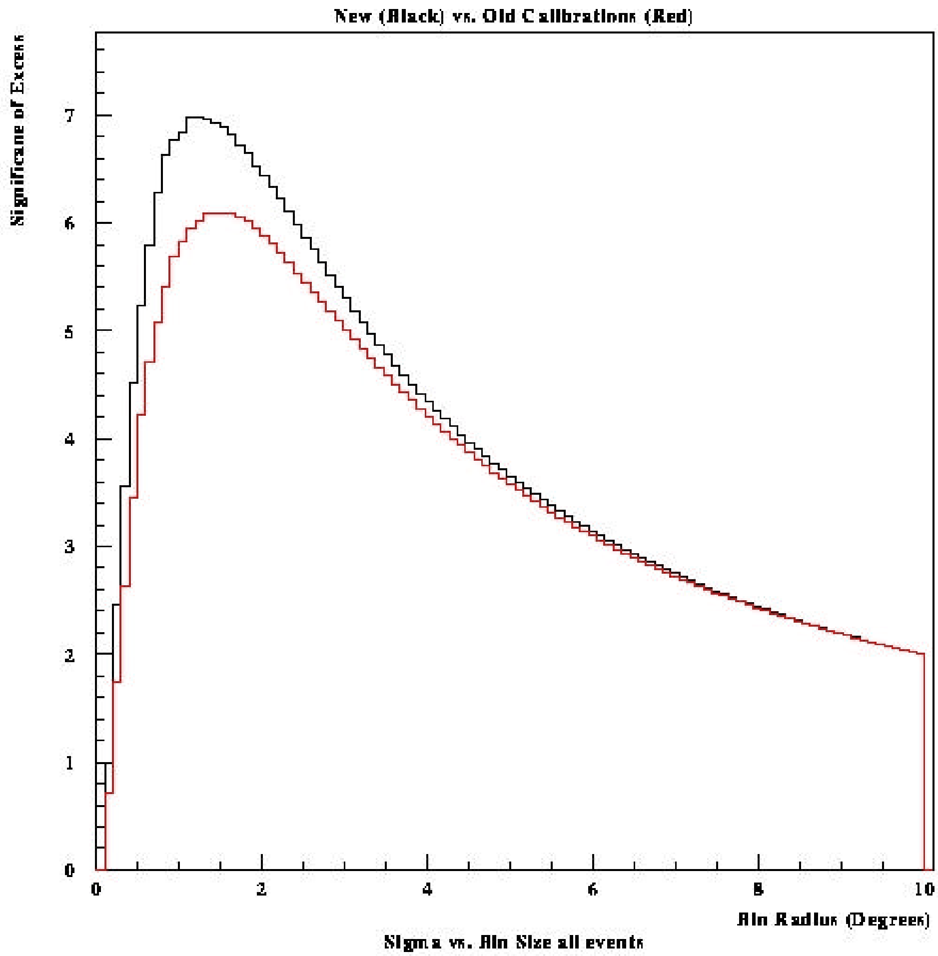

Figure 9 Significance of an excess as a function of bin radius for new (black) and old (red) calibrations.