Milagro Memo
Gus Sinnis
2/2/99

Study of Optimal Bin Size and NFIT Cut for a Binned Analysis

Introduction

In this note I determine the optimal bin size and NFIT cut to use when searching for a DC source at a known position. This result only applies to a binned analysis with a fixed bin size. There are inherent limitations in the technique that I used; primarily I assume that DELEO/2 is the point-spread function of Milagro. There are at least two problems with this assumption: 1. Event-by-event systematic errors caused by cores that land off the pond (80% of our data) and 2. Events that fail DELEO (either half of the pond could not be fit) yet are still fit by the angle fitter (this has a bigger effect for small values of NFIT). With these caveats my conclusions are that it is best to use no NFIT cut and a round bin of radius 1.1 degrees. Despite the steep fall-off in NFIT beyond an NFIT of 20 there is only a small decrease in sensitivity (15%) because of the improving angular resolution.

Technique

The significance of a signal is S/N where N is the square root of the background level. To find the effect of changing the NFIT cut and bin size I assume that the NFIT distribution is the same for signal and background, and also that our point spread function for signal (gamma rays) is the same as DELEO for protons (the bulk of our data). For the analysis here I used the Slewing calibrations from the laser data provided by Roman and Lazar, and the "Spectrum" based TOT-to-PE conversion provided by Andy. I use a recent run where the rate was 1200 Hz to determine the NFIT and DELEO distributions. For a given NFIT cut and bin size the significance of a signal is given by:

$$s = \frac{S}{\sqrt{B}} = \frac{S_0 f_{fit}(\geq NFIT) f_{DELEO}(\leq R)}{\sqrt{B_0 \left(R / R_0 \right)^2 f_{fit}(\geq NFIT)}}$$

Figure 1 shows the integral NFIT distribution, the fraction of events with nfit>= NFIT. Figure 2 shows the DELEO distribution for several values of the NFIT cut. Shown are curves corresponding to NFIT>=(0,30,50,70,100). There is no discernible difference between

NFIT=(0, 10, and 20). These distributions used with the above formula are used to derive the optimal bin size.

Result

Figures 3 and 4 show the result. I plot the number of standard deviations (based on an arbitrary S_0=316, B_0=1000, and R_0=2 degrees) versus the bin size for several values of the NFIT cut. Shown are NFIT>=(0,30,50,70, & 100). The best choice is no cut on NFIT and a bin with radius 1.1 degrees. However, as long as one modifies the bin size accordingly there is little loss in sensitivity for larger values of NFIT. The table below gives the optimal bin size as a function of NFIT, and the fraction of signal events retained (this includes the effect of the NFIT cut and the bin size).

NFIT Cut	Optimal Bin Radius	Fraction of Signal in Bin	Relative Significance of Signal
0	1.1	.39	7.1
10	1.1	.39	7.1
20	1.1	.38	7.0
30	1.1	.30	7.0
40	0.9	.20	6.8
50	0.9	.16	6.6
60	0.8	.12	6.4
70	0.7	.09	6.4
80	0.7	.08	6.2
90	0.7	.08	6.1
100	0.5	.05	6.0

For completeness I give Figures 5 (optimal bin size vs. NFIT), 6 (Significance vs. NFIT [assuming that the optimal bin size is used for the given value of NFIT]), and 7 (the fraction of signal events retained in the bin [including both the effect of the NFIT cut and the point-spread function of Milagro]).

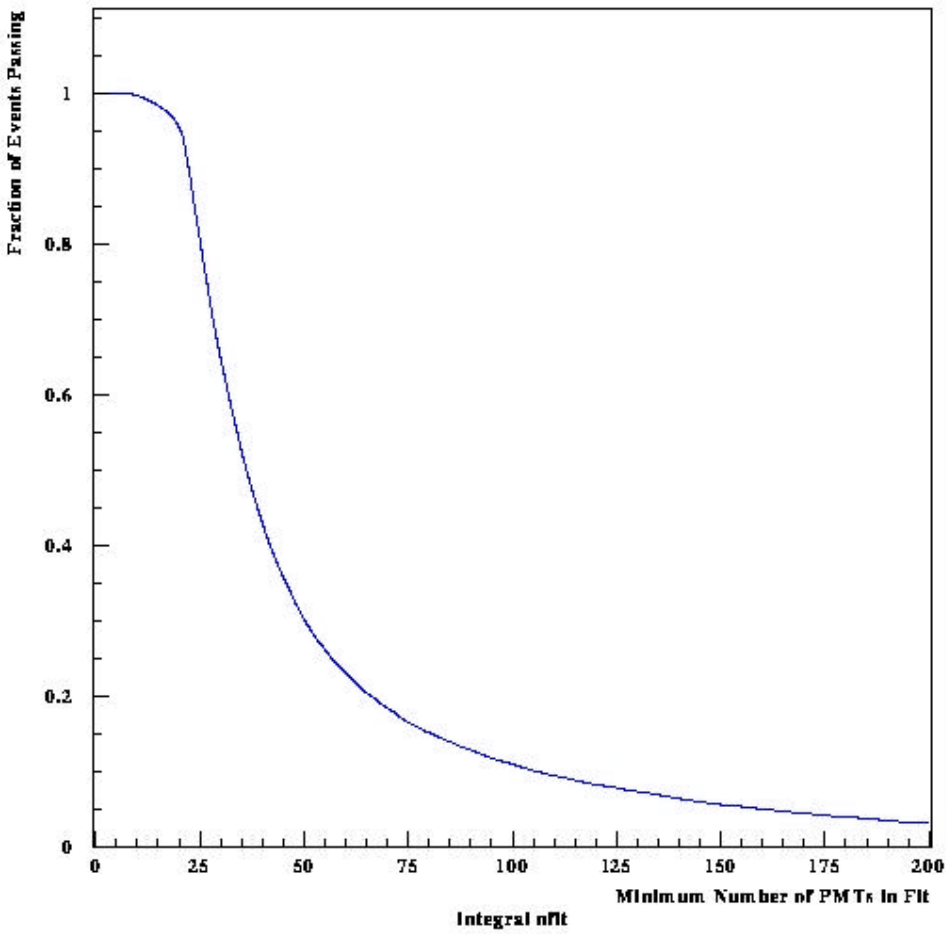

Figure 1. Integral NFIT distribution. Fraction of events with NFIT greater than or equal to x-axis value.

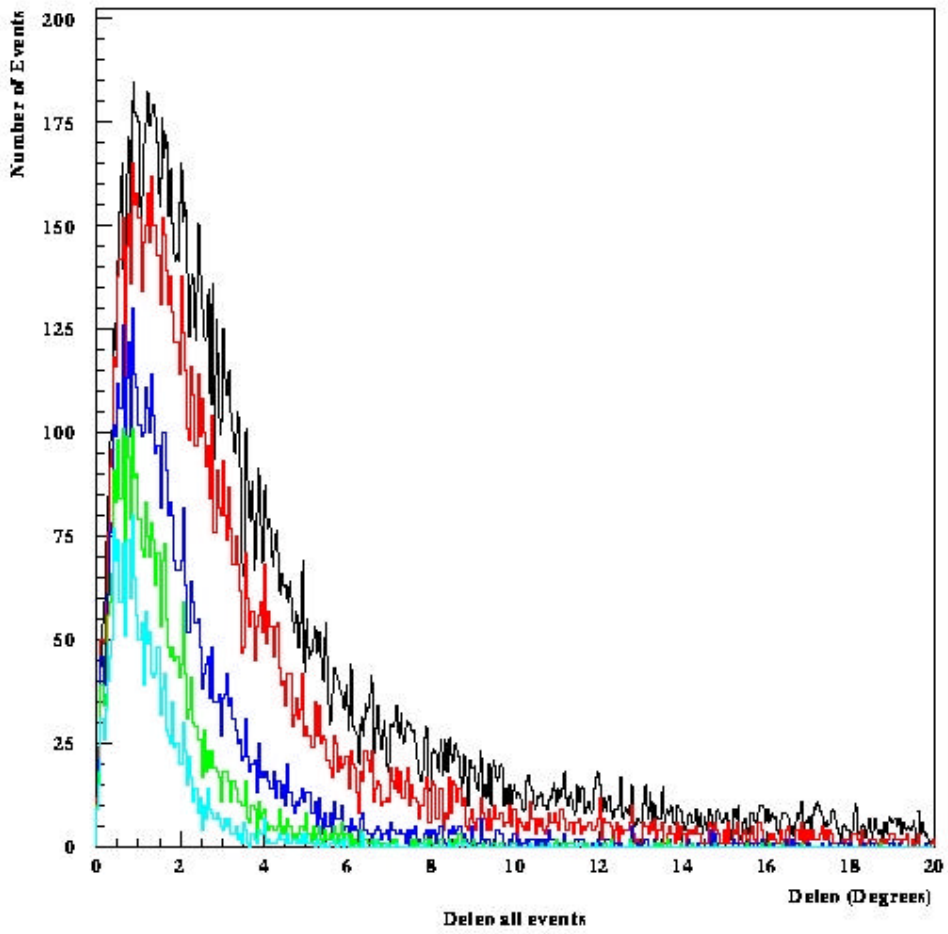

Figure 2 Deleo distribution for NFIT cut of (0=black, 30=red, 50=blue, 70=green, 100=sky blue). Or if your printer is color blind in order from the top down.

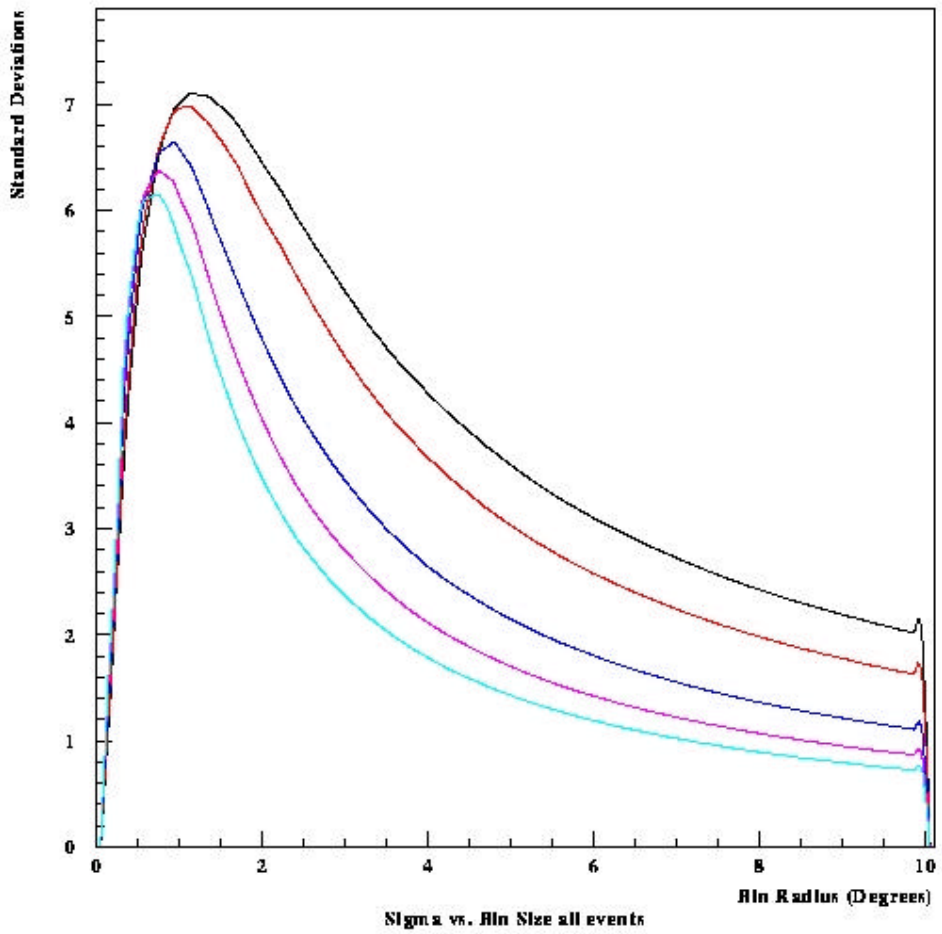

Figure 3. The significance of a signal as a function bin radius for several values of the NFIT cut. (0=black, 30=red, 50=blue, 70=pink, 100=light blue) or in order of decreasing height.

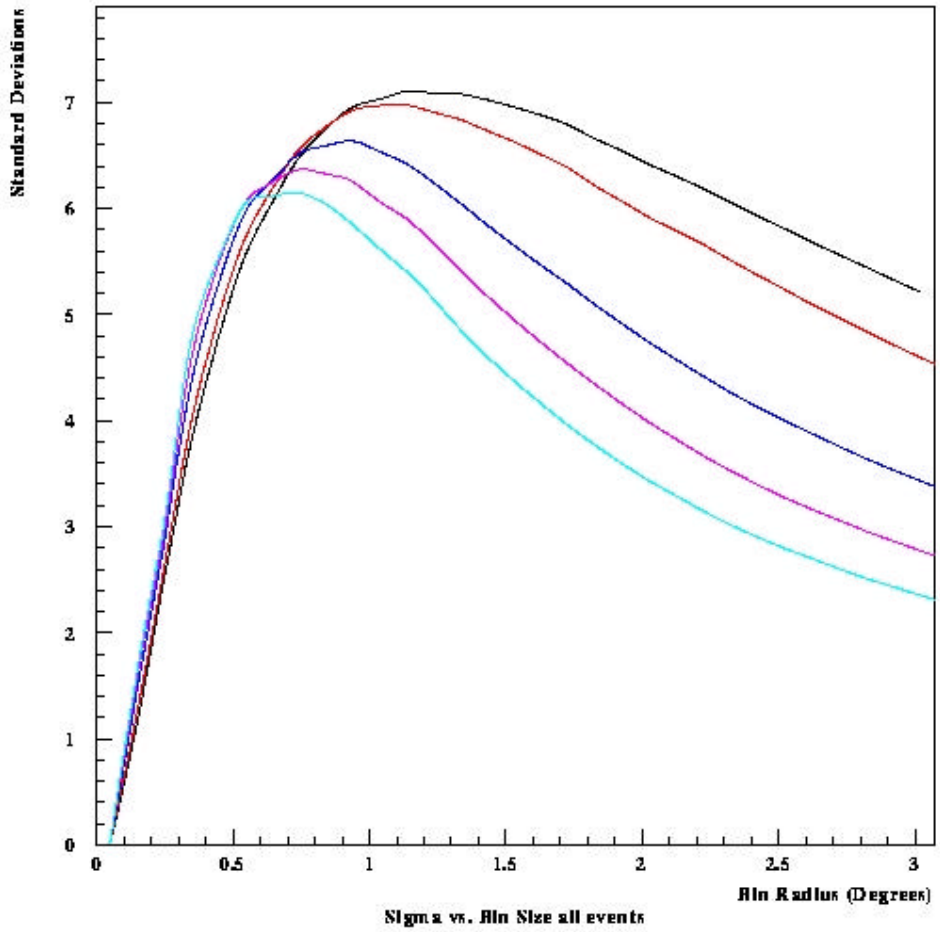

Figure 4. Same as Figure 3 expanded near the origin.

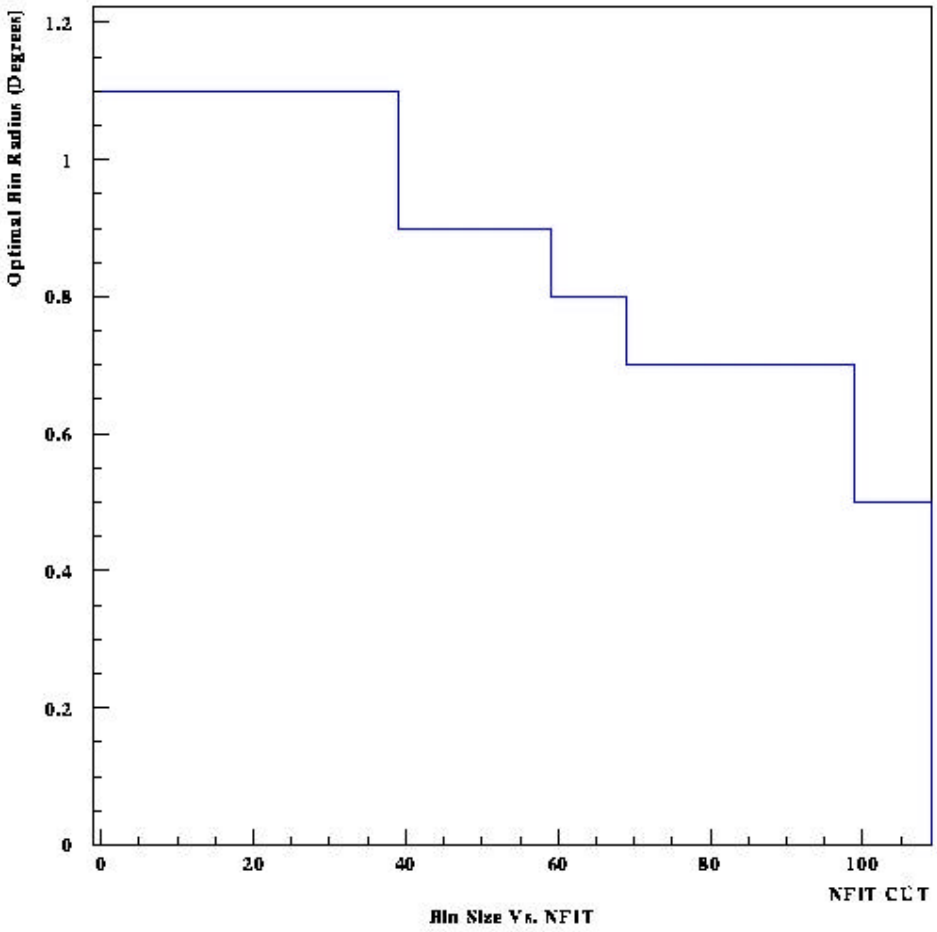

Figure 5. Optimal bin radius as a function of NFIT cut.

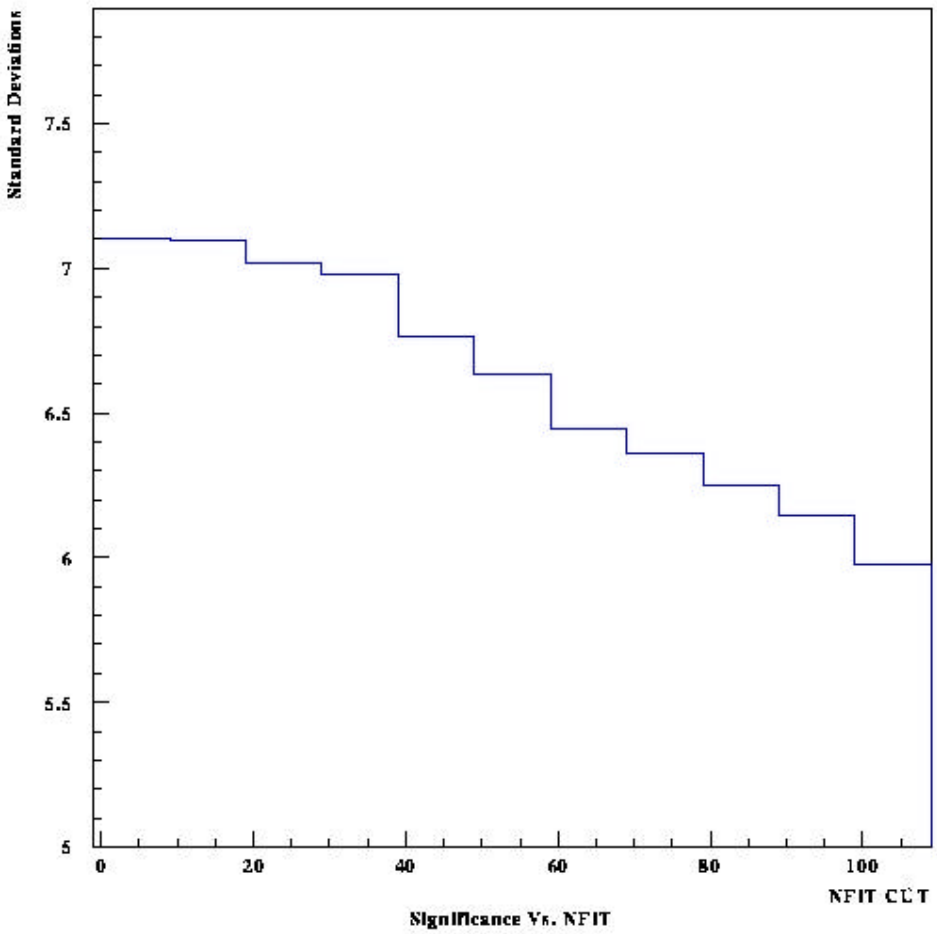

Figure 6. Relative significance of a signal as a function of NFIT cut.

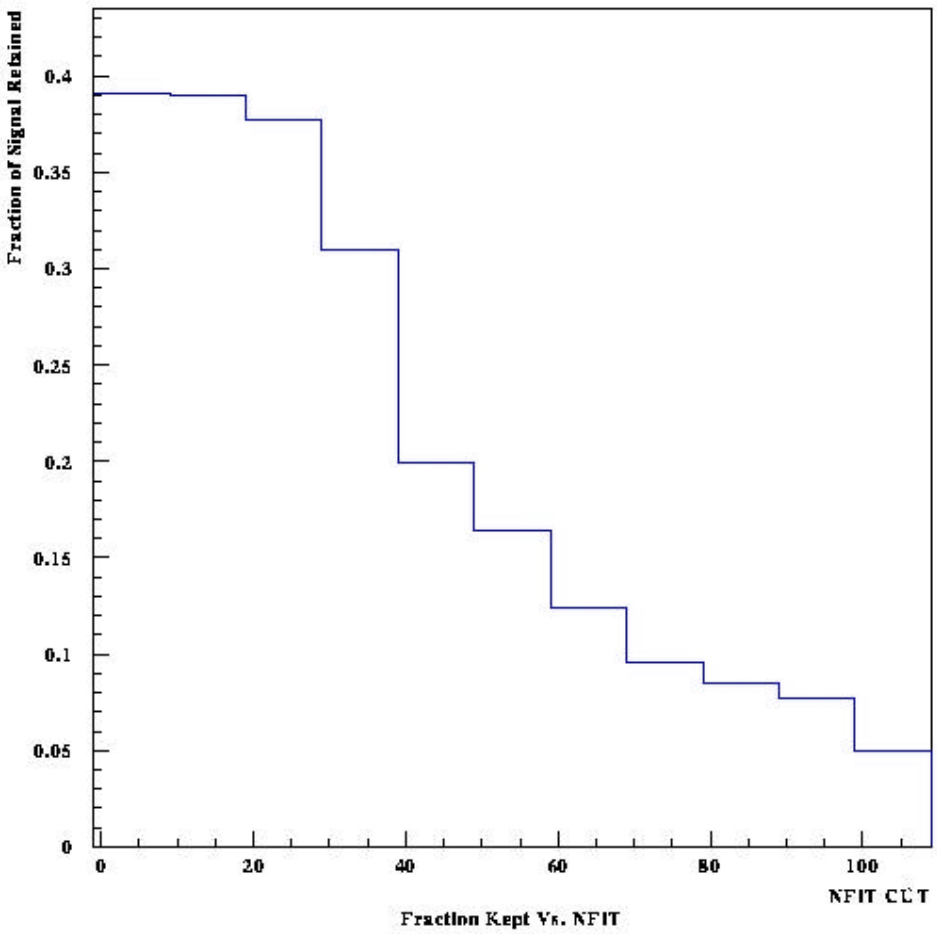

Figure 7. Fraction of signal retained in optimal bin after NFIT cut as a function of NFIT cut.