Gus Sinnis
LANL
Milagro memo
7/17/00

Hadron Rejection in Milagro

Introduction:

In this memo I study a generalization of the method of hadron rejection discussed by Gaurang in his earlier memo. The motivation behind this study is the fact that when the original cut was applied to data (as opposed to Monte Carlo) the Q factor was found to be less than 1. I have found two problems with the early work; one involves the reduction of our data (a problem with the use of the edgeFinder in the muon layer), the other in the comparison of data to Monte Carlo. At the collaboration meeting in Los Alamos of 7/2000 I outlined the results of this study using Version 43 and the spectrum based PE calibration. In this memo I use the new "Occupancy 4" PE calibration of Roman and Lazar, now the standard in V44 of the offline code.

I have optimized the cut over values of a PE threshold (p in the equation below) for each PMT and over the x parameter [x=Nbottom(>p PEs)/ MAXPE bottom]. While the individual PE, NHit, and x distributions produced by the Monte Carlo compare well with the data, at large values of x there is a significant discrepancy between the Monte Carlo and the data. This leads to the clearly incorrect conclusion that arbitrarily large values of Q are possible. Based solely on Monte Carlo Q values near 2 are attainable. We must tune our cut on a signal from the Crab to resolve this dilemma.

The x_p Cut:

The motivation for the x_p cut is that gamma ray induced air showers tend to yield relatively uniform pulse heights in the bottom layer of the pond, while hadronic air showers tend to be clumpy where a cluster of PMTs is characterized by 1-2 large pulse height PMTs. The exception to this rule is gamma showers whose core lands on the pond. The cut tends to remove these showers as well as proton showers. The definition of the x_p parameter is given here:

$$x_p \equiv \frac{NHit(\geq p_{PEs})}{PE_{Max}},$$

where all values are over the muon layer. The subscript p on x_p refers to the PE threshold value applied to the calculation of *NHit*. Figures 1-4 show the $NHit_p$ and PE_{Max} distributions for the data and for Monte Carlo proton showers. The Monte Carlo is version 22 and I have taken a 55 PMT threshold in both the simulations and the data. For the data this threshold keeps about 92 % of our data.

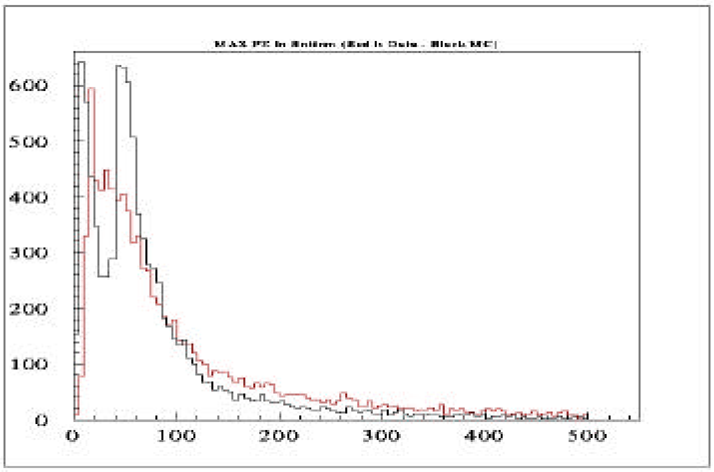

Figure 1 Distribution of PE MAX in bottom layer: red data, black MC protons.

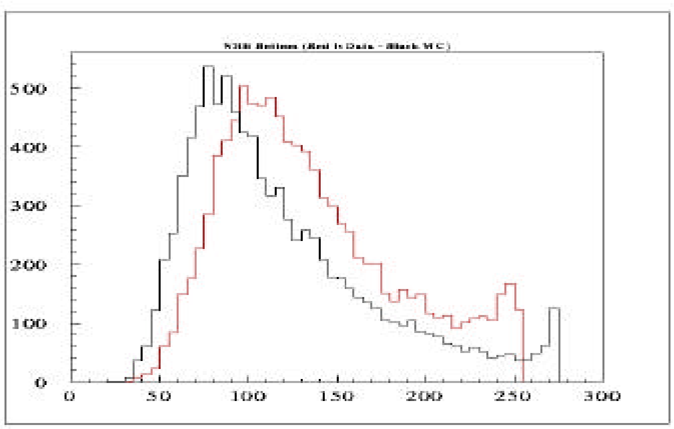

Figure 2 NHit distribution: red data, black MC protons.

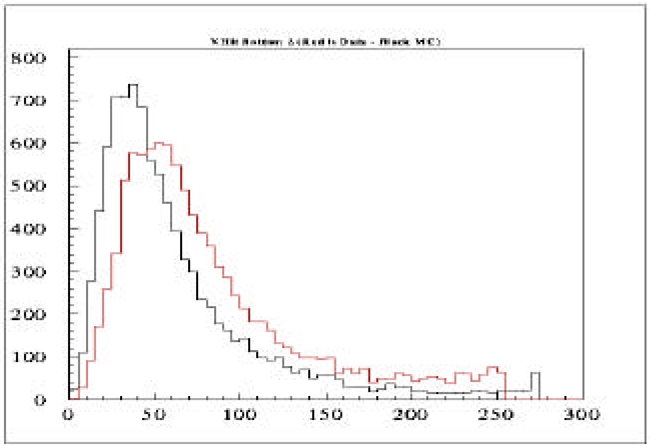

Figure 3 NHit (> 2 PE) distribution: red data - black MC protons.

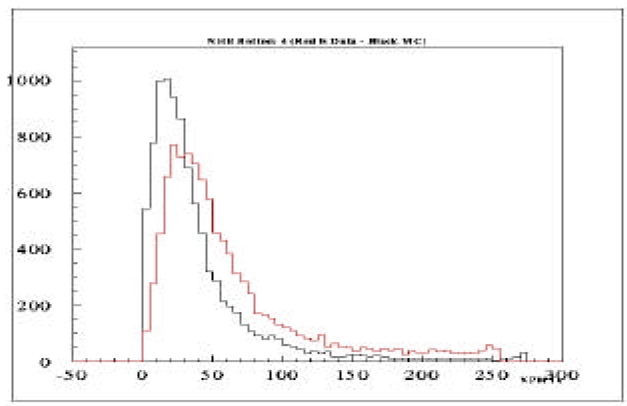

Figure 4 NHit (>4 PE) distribution.

There are several obvious features in the above plots. First the Monte Carlo does not have dead PMTs accounted for and in general there are more PMTs hit than predicted by the Monte Carlo. Though the agreement is better for larger values of the PE cut imposed. Note that I define a "Hit" in the data as a ***calibrated*** PMT, this ensures that the PMT passed the edgeFinder and should remove most of the random hits. Since all but 2 PMTs are calibrated in the bottom layer this has a minimal effect on the distributions. Gaurang has already discussed the problems with the PEMAX distribution in in an email message. Above ~50 PE the Monte Carlo and the data are in reasonable agreement.

In Figures 5-7 I show the distribution of x_p for various values of p. Figures 8-10 show the efficiency as a function of x for a cut on x_p. From these figures one can derive the Q factor for each value of p and each value of x.

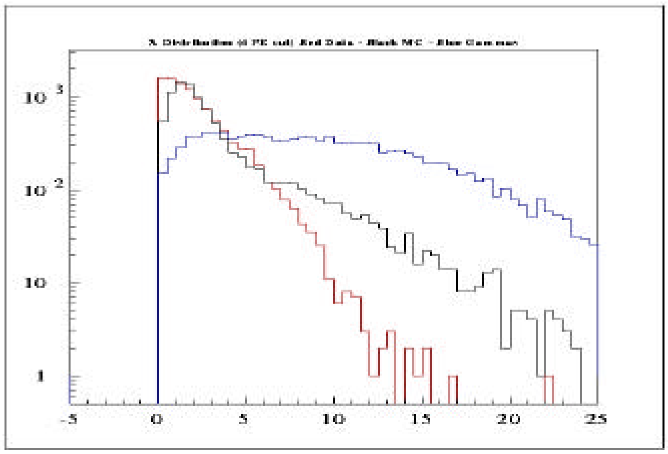

Figure 5 X distribution with 0 PE criteria imposed on bottom layer. Red data, black MC protons, blue MC gammas.

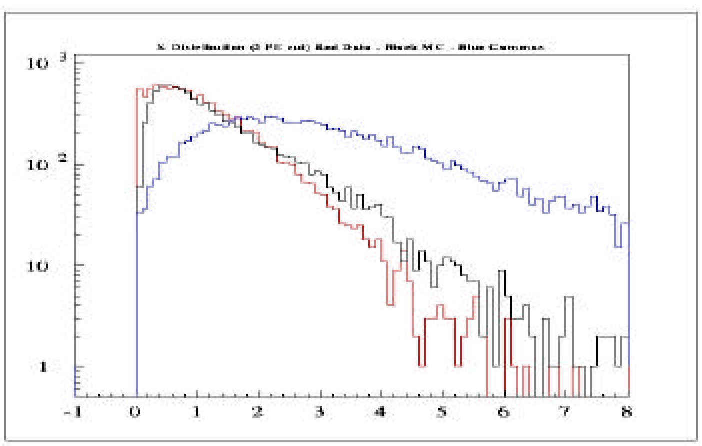

Figure 6 X distribution with 2 PE criteria imposed on bottom layer. Red data, black MC protons, blue MC gammas.

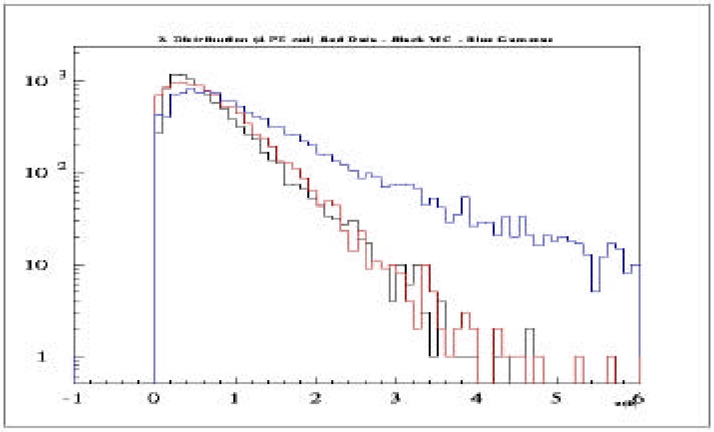

Figure 7 X distribution with 4 PE criteria imposed on bottom layer. Red data, black MC protons, blue MC gammas.

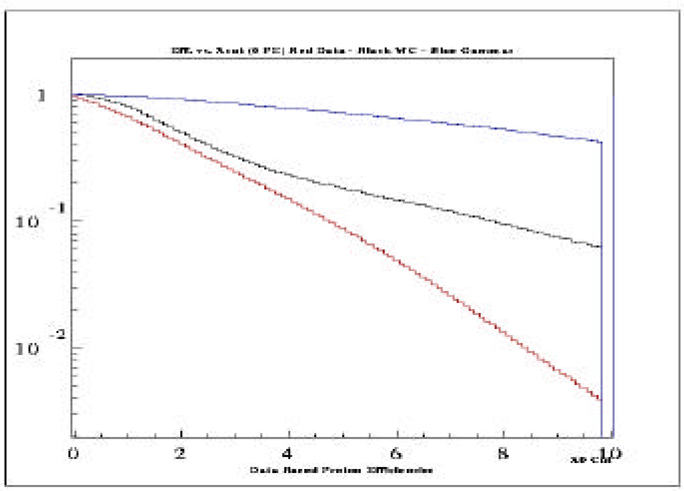

Figure 8 Efficiency as a function of X cut, 0 PE criteria imposed on bottom layer. Red data, black MC protons, blue MC gammas.

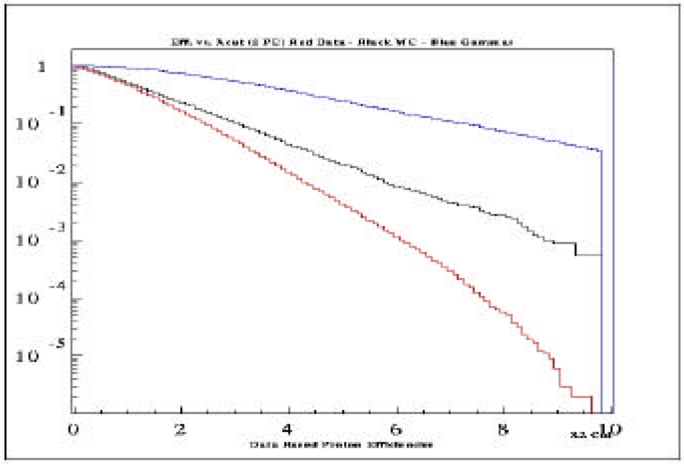

Figure 9 Efficiency as a function of X cut, 2 PE criteria imposed on bottom layer. Red data, black MC protons, blue MC gammas.

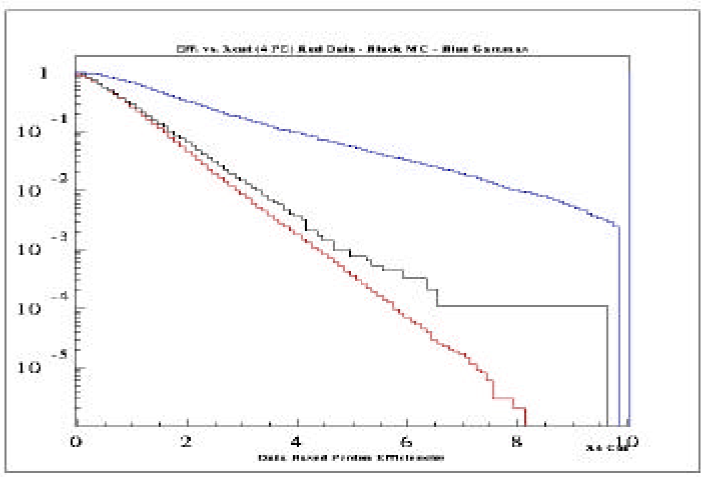

Figure 10 Efficiency as a function of X cut, 4 PE criteria imposed on bottom layer. Red data, black MC protons, blue MC gammas.

While the differences between the data and the Monte Carlo at and above the 2 PE cut appear small they are quite significant. This is illustrated by examining the *Q* factor as a function of the *x* cut for each of the above values of *p* (the PE cut imposed on the bottom layer). Figures 11-13 show this.

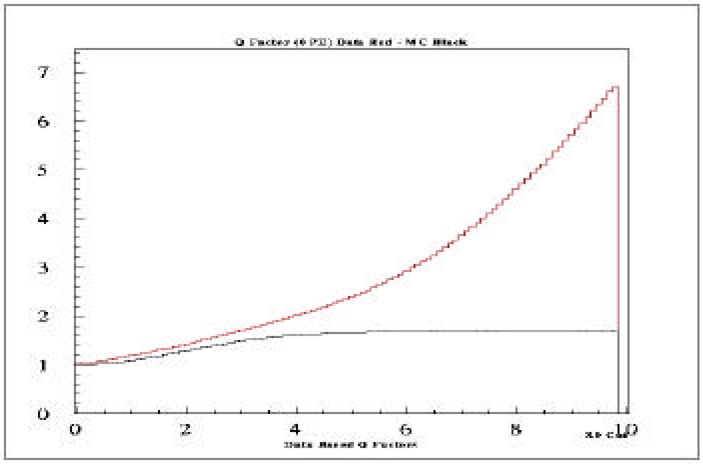

Figure 11 Q factor as a function of x cut, 0 PE criteria imposed on bottom layer. Red - Data and MC gammas, Black - MC protons and MC gammas.

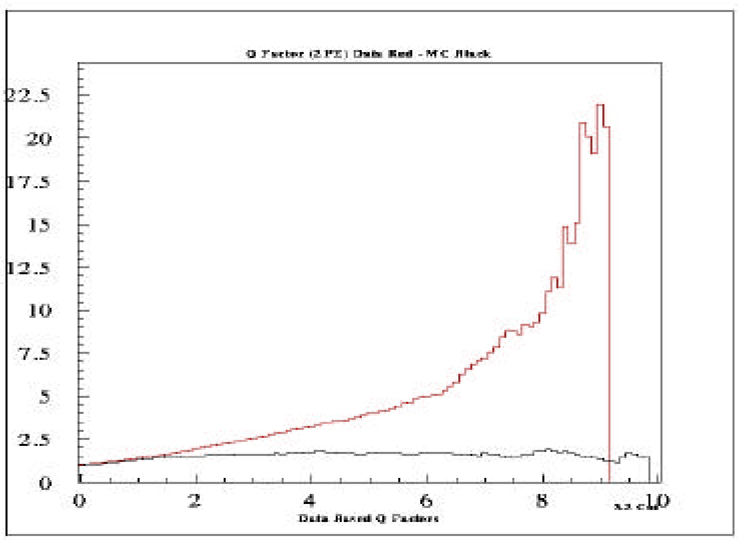

Figure 12 Q factor as a function of x cut, 2 PE criteria imposed on bottom layer. Red - Data and MC gammas, Black - MC protons and MC gammas.

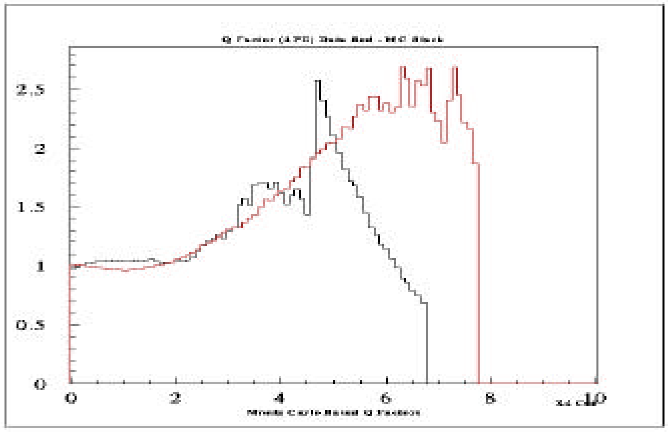

Figure 13 Q factor as a function of x cut, 4 PE criteria imposed on bottom layer. Red - Data and MC gammas, Black - MC protons and MC gammas.

An optimist might conclude from Figure 12 that Milagro is about to enter into the golden age of TeV astronomy. Given the large discrepancy between the simulations and the data at present the only solution would seem to be to tune the cut on any signal from the Crab, though beginning with a 2 PE criteria on the bottom layer and scanning in *x* seems like a reasonable starting point.

The more serious question is what is causing the large discrepancy between the data and the simulations. There are many places one may look. It is interesting that the discrepancy is at large values of *x*. This implies that large values of PEMAX are not the culprit and spending more time on the very high TOT extrapolation will not help us in gamma hadron separation. In Figure 14 I show the distribution of PEMAX for all events and superimposed for events with $3<X_2<5$. The two distributions are quite different and it appears that the problem is associated with the PE range near 15 PEs.

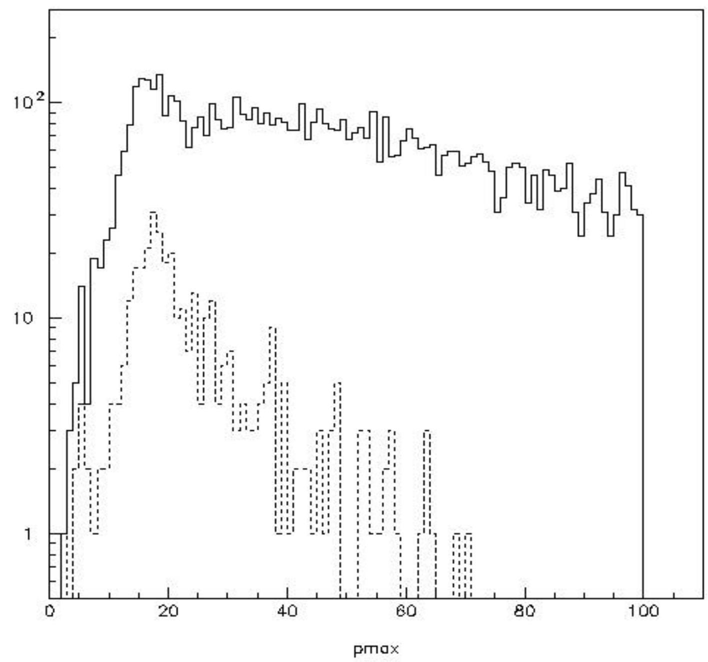

Figure 14 PEMax distribution. Solid line all events dashed line for events with 3<X_2<5.

This PE value is just above the high-low transition region (~12-13 PE in V44). We have traditionally had difficulties in this region and the laser calibration has some additional difficulties in the transition region, which require us to impose an artificially high value on HI TOT before it can be used to determine the PE level in the PMTs (the hardware transition occurs around 6 PE) and the resolution of the LO TOT becomes quite poor near 7 PEs. However, it appears that pulse heights in this region are quite important in gamma hadron separation.