Discovery Mode Search Techniques For Gamma-Ray Telescopes

Miguel F. Morales† for the Milagro Collaboration

†*Santa Cruz Institute for Particle Physics, University of California, Santa Cruz, CA 95064*

Abstract. Like many wide field gamma-ray telescopes, the Milagro observatory has a highly variable point spread function (PSF), dependent upon the characteristics of the incoming photon. Because of the large variations in the PSF, a binned transient search technique is not optimal, and maximum likelihood can be computationally unfeasible. I have expanded upon the Gaussian weighting technique [1] to develop a transient search method that is sensitive, model independent, and computationally tractable. This method is currently being used to look for TeV transients from 40 seconds – 2 hours duration with the Milagro detector.

THE PROBLEM

My goal in developing the weighted analysis technique was to perform a near optimal discovery mode search for point sources and transients in the Milagro data set. Milagro, like many wide field gamma-ray telescopes, has a highly variable resolution that depends on the characteristics of the individual photon. In an optimal bin analysis, one looks at the PSF of the detector and chooses a bin size that maximizes the expected signal to noise. These analyses are fast and work very well for detectors with a set PSF of near Gaussian shape. However, in wide field gamma-ray observatories the PSF can vary by more than an order of magnitude from one photon to the next. In binned analyses all photons are treated as equal, and the quality of an event is ignored. This can seriously degrade the significance of a signal and the flux limit of a detector.

The maximum likelihood technique is able to use the photon by photon PSF information, and is the standard tool of choice for dealing with variable PSF detectors. Unfortunately, maximum likelihood can be difficult to implement in a way that it is both computationally efficient and model independent. Because we don't know what the characteristics of a TeV transient signal will be, we must also make sure that our discovery mode search is not biased by an assumed signal type. The weighted analysis technique uses all of the available information, is computationally fast, and model independent.

THE METHOD

For a gamma-ray telescope every event has two associated probabilities, the probability that it was a gamma ray (as opposed to a cosmic ray - $P(\gamma)$), and the PSF, or probability that the initial photon came from the direction determined by the reconstruction algorithm. Both of these probabilities can vary considerably from one photon to the next, and the distributions are determined by the characteristics of the detector (for an example, see the next section where we characterize the Milagro detector).

We can make a map of the sky that exactly represents our knowledge by placing at each point the probability that a photon came from that location. This is equivalent to putting the product $(dPSF/d\Omega)P(\gamma)$ onto the sky map. Our sky map is then the probability density map and represents our complete knowledge of the dataset.

So at any given point, the cumulative photon probability density (or weight) is:

$$w = \sum_{i}^{all\,events} \frac{dPSF_i(r_i)}{d\Omega} P_i(\gamma)$$

Where i indexes all of the events, and r_i is the angular distance of the reconstructed event direction from the point in question. In a discovery mode search, we want to determine the probability that a given weight is a fluctuation of the background. There are two independent effects which contribute to variations in the weight. First there are fluctuations in the number of events observed (and thus used in the sum) compared to the expected number of events. Second, there are fluctuations in the weight due to unusually close events (PSF term), or from

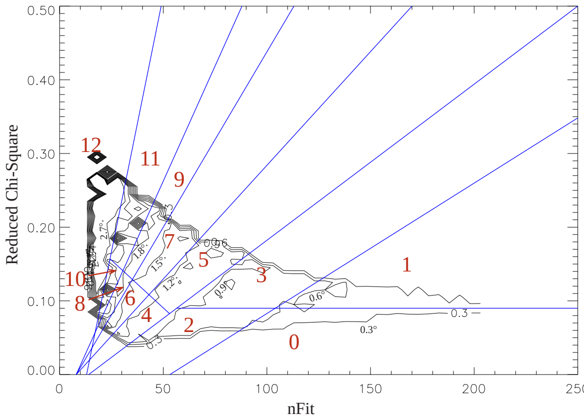

FIGURE 1. Contour plot of the PSF width as a function of the fit variables nFit and reduced chi-square. Overplotted are the 13 regions of similar PSF.

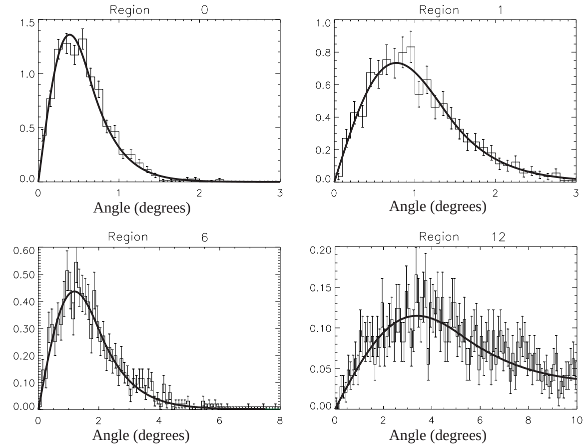

FIGURE 2. The radial PSFs for 4 of the 13 regions indicated in figure 1, with the associated fits.

unusually gamma-like events (gamma term). The effects of weight variations and number variations can be separated by looking independently at the fluctuations in the average weight and the number of events observed. The question then becomes, given the distribution of average weights and expected number of events, what is the probability of observing a particular average weight and number of events? Mathematically this can be written as:

$$P(w_{avg}, N_{obs} |\ P(w_{avg}|N_{obs}), N_{exp}) =$$

$$P(w_{avg}\ |\ P(w_{avg}|N_{obs}))\ P(N_{obs}|N_{exp})$$

The first term on the right is just the probability of getting a certain average weight (easily measured from our background), and the second term is straight Poisson statistics. Given the weight at a point, and the measured distribution of weights seen in the background, the probability that a given weight comes from the background distribution can be easily calculated.

For a full sky search, the sky map only needs to be sparsely sampled. The smallest PSF sets the spatial scale for variations, so sampling on a scale smaller than the smallest PSF retains all of the information. This spacing turns out to be only a little smaller than the bin size in an optimal binned analysis. Additionally, because we are sampling at points, tiling problems associated with binning a spherical sky never occur and separate sky maps can be directly summed. This can be a significant calculational advantage when searching over many time scales.

MILAGRO DETECTOR CHARACTERISTICS

In order to use the weighted analysis technique (or maximum likelihood), the characteristics of the detector must be understood. For Milagro, the PSF correlates well with the reconstruction variables nFit and reduced chi-square. The Milagro reconstruction algorithm performs an iterative chi-square fit where tubes with times far from the fit are removed or added in each iteration. nFit is the number of tubes used in the final fit, and the reduced chi-square is the true reduced chi-square times $\sim$0.1 (for historical reasons). Figure 1 is a contour plot of the PSF width as a function of nFit and reduced chi-square. In this work, this space is divided into the 13 regions shown in the figure, and the PSF of the detector is found for each of these regions. Events are categorized by which region they fall into and the corresponding PSF is used. Figure 2 shows the Monte Carlo distributions for gamma rays in 4 selected regions, and the corresponding fits. In order to avoid noise due to the low statistics of our Monte Carlo data set, these fits are used to determine the PSF used in the analysis.

The probability that an event was initiated by a gamma ray can be determined in a similar manner. The Milagro background rejection is currently based on the "compactness" parameter. Compactness characterizes the distribution of light in the lower layer of phototubes, and shows signficantly different behavior for gamma and proton initiated showers[2]. For the weighted analysis technique we want to determine the probability that a given shower was initiated by a gamma ray. For each region, we can use Monte Carlo simulations to determine the compactness distributions for both protons and gamma rays. Figure 3 shows the compactness distributions and gamma

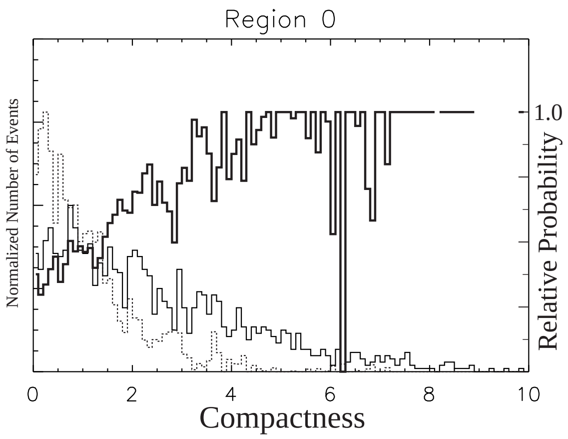

FIGURE 3. This figure shows the compactness histograms and gamma probability distributions for an example region. The compactness histograms for gamma and proton initiated showers are shown by the thin solid and dashed lines, respectively. The solid thick line shows the fraction of events which are gamma rays as a function of the compactness parameter

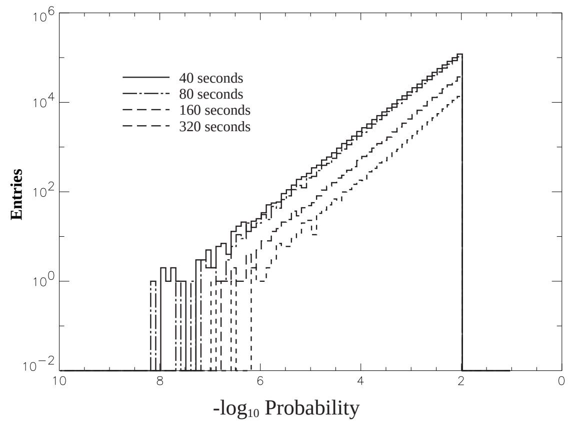

FIGURE 4. Histograms of the observed probability distributions for one day of TeV transient searches at 4 different time scales. These distributions are consistent with the expected background fluctuation distributions; any signals would appear as isolated points of very low probability. The weighted analysis method is currently being used for real time transient searches with the Milagro observatory.

probability for a representative region. The gamma probability distribution is fit to avoid noise from the low statistics of our Monte Carlo data set, and for computational reasons we currently use only showers with a gamma probability greater than 0.5.

IMPLEMENTATION

The weighted analysis technique is amenable to a number of approximations that can significantly speed up the analysis. First, the PSF can be truncated to some set distance; for Milagro 4 degrees is the largest radius used. The current implementation also does not calculate the distance from the photon to each point, but instead uses a table of approximate PSF values that are copied onto the sky map. Separate signal and background maps can be directly added to form new maps, greatly increasing the speed of analyzing additional time scales. On current hardware (dual $\sim$1Ghz Pentium 4, standard memory) the current algorithm can process $\sim$2000 events/second while analyzing 9 time scales between 40 seconds and 2 hours.

Recently, this technique has been used to search through the Milagro dataset in real time for transients of 40 seconds to 2 hours duration. Example background distributions for these searches are shown in figure 4. No significant burst has been detected to date, but rapid notification of any observed TeV transients will be distributed to the GRB community.

ACKNOWLEDGMENTS

This research was supported in part by the National Science Foundation (Grant Numbers PHY-9722617, -9901496, -0070927, -0070933, -0070968, -0096256), the U. S. Department of Energy Office of High Energy Physics, Los Alamos National Laboratory, the LDRD program at LANL, the University of California, the Institute of Geophysics and Planetary Physics, The Research Corporation, and the California Space Institute. The author would also like to acknowledge the support of the NASA Graduate Student Research Fellowship and the Mellon Foundation.

REFERENCES

1. Woodhams, M.D., *Scintillator Data Analysis for the JANZOS Array*, M.A. Thesis, University of Auckland, 1989, pp. 33-42.
2. Sinnis, C., for the Milagro Collaboration, "Background Rejection in the Milagro Gamma Ray Observatory," Proceedings 27th ICRC, Hamburg 2001, **OG187**