

A Radon Transform based Gamma-Hadron Discrimination technique for the ARGO-YBJ experiment

Milena Dattoli* §

on behalf of the ARGO-YBJ Collaboration

* *INAF-IFSI, c.so Fiume 4, 10133 Torino, Italy*

§ *INFN and University of Torino, via P.Giuria 1, 10125 Torino, Italy*

Abstract. The observation of γ -ray sources is deeply hampered by the charged cosmic rays (CRs) huge background, which overwhelms the very faint signal coming from these sources. The development of a background suppression method tailored to the detection technique is a key issue in Gamma Astronomy oriented experiments.

In this paper we discuss a Gamma Hadron Discrimination (GHD) method developed ad hoc for the ARGO-YBJ experiment (YangBaJing, P.R. China, 4300 m a.s.l.), an Extensive Air Shower (EAS) detector made by a single layer of Resistive Plate Counters (RPCs) covering a surface of about 6700 m². A very detailed image of the shower can be seen on the ARGO-YBJ central carpet thanks to the high granularity of the detector. The image undergoes a Radon Transform based analysis procedure, which highlights the differences in the particle distribution of gamma and hadron induced showers - especially in the region near the shower core.

The method efficiency has been valued by means of suitable Monte Carlo simulations, showing a good discrimination power (correct identification of about 80% of gamma and 60% of hadron events) for showers having the core within the detector central carpet and firing at least 500 pads, corresponding to a median energy of ~ 8 TeV for gamma showers. The description of the GHD technique and the results obtained applying the method to the Crab Nebula data are presented and discussed.

Keywords: Gamma-Hadron discrimination, Gamma Astronomy, Radon Transform

I. INTRODUCTION

A γ -ray source is identified as a statistically significant excess of events coming from a certain direction above the huge isotropic background of CRs, thus the sensitivity of ground based experiments is expressed in units of standard deviations of the CRs background ($S = \frac{N_\gamma}{\sqrt{N_{bkg}}}$, where N_γ is the excess of events coming from the direction of the source and N_{bkg} is the number of CRs events within a solid angle of observation $\Delta\Omega$). The reduction in the CRs background, which is fundamental to enhance the sensitivity of experiments aiming at studying γ -ray sources, can be achieved both with a

good angular resolution and with the application of some gamma/hadron discrimination (GHD) criterion: an accurate reconstruction of the arrival direction reduces the sky region contributing to the background level and produces a narrower spatial distribution of the signal, while a GHD method can enlarge the detector sensitivity of a factor $Q_f = \frac{\epsilon_\gamma}{\sqrt{1-\epsilon_{bkg}}}$, where ϵ_γ is the fraction of γ events correctly tagged as signal by the method and ϵ_{bkg} is the fraction of correctly discarded CRs events.

Even though efficient GHD methods have been developed by ground based experiments to reduce the hadronic background (e.g. [7],[1]), experiments based on different detection techniques can not use the same discrimination method: in this paper we present a GHD technique developed *ad hoc* for the ARGO-YBJ experiment.

II. THE ARGO-YBJ DETECTOR

ARGO-YBJ (YangBaJing, P.R. China, 4300 m a.s.l.) is an EAS detector devoted to the study of cosmic rays and γ radiation, optimized to work at TeV energies, with an energy threshold of few hundred GeV and a dynamical range extending to PeV. ARGO-YBJ is a modular detector, the basic module being the cluster (5.7×7.6 m²), divided into 12 Resistive Plate Counters (RPCs, 1.2×2.8 m²) [2]. Signals from each RPC are picked-up by 10 electrodes 55.6×61.8 cm² wide, called pads, which provide the space-time pattern of the shower front with a time resolution ~ 1.8 ns. Each pad is segmented into 8 strips which count the number of particles hitting the pad. The full detector is composed by 153 clusters, corresponding to a total active surface of ~ 6700 m²: a central full coverage carpet made up of 10×13 cluster is enclosed by a sampling guard ring in order to enlarge the active area and improve the capability to tag internal and external events.

III. THE TECHNIQUE: THE RADON TRANSFORM

ARGO-YBJ can perform γ -hadron discrimination thanks to its being a peculiar, full coverage detector with a very good spatial resolution. Unlike the traditional EAS arrays which sample the particle density at different distances from the shower core, ARGO-YBJ can detect the shower structure on a quite large continuous area in detail, thus allowing the study of the region near

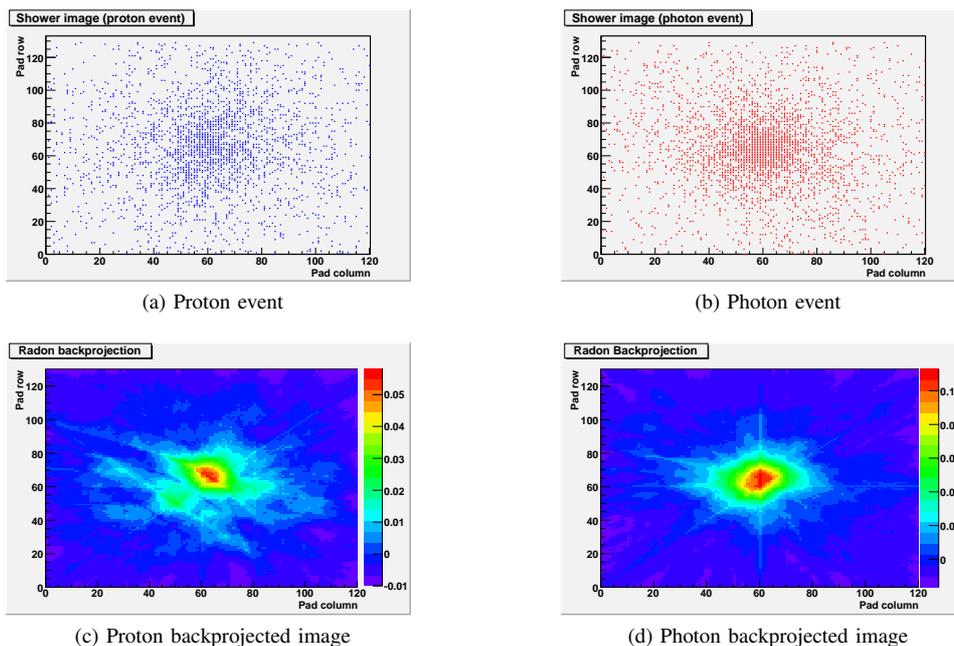


Fig. 1. Example of a photon and a proton shower firing almost the same number of pads (~ 2500) on ARGO-YBJ central carpet. Top: proton (left) and photon (right) shower images. Bottom: proton (left) and photon (right) backprojected images after the application of the Radon procedure. The colour scale is related to the local particle density.

the shower core, where showers induced by γ and by hadrons show the largest differences.

The GHD technique developed for ARGO-YBJ is based on the Radon Transform (RT) [8] of the shower image on the detector central carpet. The RT is an integral transform whose inverse is used to search for clusters in images, e.g. to reconstruct images from medical computed axial tomography (CAT) scans.

Given an image defined as a two-dimensional function $f(x, y)$, its RT $f(\rho, \theta)$ is found by integrating the image along all the straight lines across. If the line is described in its normal form $\rho = x \cos \theta + y \sin \theta$, ρ is the perpendicular distance from the line to the origin and θ is the angle formed by the distance vector.

In case of digital images, like the shower image on the ARGO-YBJ carpet, a discrete approximation of the RT is used, in which a linear sampling of the parameters ρ and θ is applied. With linear sampling the pixels of the starting image can also belong not to one of the sampling lines - thus a weight inversely proportional to the distance from the nearest line is assigned to the points not lying on the line.

IV. APPLICATION TO ARGO-YBJ DATA

The image that will be analyzed for GHD purposes is obtained from the ARGO-YBJ data - both Monte Carlo and raw data - projecting the shower image im_a onto the plane orthogonal to the arrival direction, reconstructed by using the procedure described in [3]; the resulting image undergoes the subsequent application of the direct and inverse RT¹, so that the final image im_B is again

¹In the following, the subsequent application of the direct and inverse RT will be referred as Radon Procedure (RP).

in the space (x, y) of the shower plane.

The original im_A and the backprojected im_B are not identical: the introduction both of the sampling and of the weighting of the pixels add some information to the particle density that was not showing on the starting image (see fig. 1). It is worthwhile to point out that the local density information added by the RP is different both from the strip multiplicity on each pad - because it takes into account also what happens in the neighbour pads - and from the *classical* density information that one can get for example studying the lateral distribution parameterized with the Nishimura-Kamata-Greisen function; by using the NKG function, in fact, the density is calculated on a fixed and symmetrical, growing shape - that is a circle having its centre in the shower core; with the RP, instead, the shape of the region changes and adapts itself to the actual shape of the shower image and the result is more sensitive to the shower features.

V. MONTE CARLO SIMULATION RESULTS

The image obtained via RP is the starting point for the actual GHD analysis: the discrimination efficiency of some parameters related to the features of γ and hadron images has been investigated by means of Monte Carlo simulations. A sample of events has been generated with the EAS simulation program CORSIKA 6.2 [6]: $7.8 \cdot 10^5$ γ and proton² events have been simulated within the energy range $[0.3, 100]$ TeV, with a primary energy spectrum generated according to power laws having a

²As a first approximation, we assume that the contribution of heavier than protons primary nuclei can be neglected.

spectral index of -2.7 for protons and of -2.5 for γ . The detector response has been simulated by using ARGOG, a GEANT3-based tool developed within the ARGO-YBJ collaboration. The arrival direction and the core position have been reconstructed with the standard reconstruction procedure; only events having $\theta < 40^\circ$ and the reconstructed core position within the central carpet undergo the RP. In order to select the region where γ and hadron showers show the most evident differences, a cut is applied at half the maximum intensity of the backprojected image. Some parameters characterizing the spot left by the cut are calculated:

- the equivalent radius, that is the radius of a circle having the same area of the spot;
- the variance of the spot border with respect to a circle;
- the ratio between the backprojection maximum intensity and the spot area;
- the ratio between the number of hits in the spot and its area;
- the opening angle at the vertex of the 3D backprojection;
- the equivalent value of π , that is the ratio between the border length and the spot diameter, defined as the maximum distance between two points belonging to the border.

If secondary clusters of particles exist in the shower image, more than one spot will be left by the cut. In this case, the analysis is done on the biggest one.

The events have been grouped in 6 multiplicity windows according to the number of fired pads on the detector, requiring at least 100 fired pads on the central carpet. The most efficient parameters have been combined by means of Fisher Linear Discriminant Analysis (LDA) [5], a simple and fast parametric method giving the linear combination of parameters best separating two classes by maximizing the between-classes variance with respect to the within-class variance.

The GHD method has been tuned on half of the simulated sample and its efficiency has been checked on the remaining sample.

TABLE I
VALUES OF THE QUALITY FACTOR AND OF THE EFFICIENCIES IN THE IDENTIFICATION OF γ AND PROTON SHOWERS IN THE DIFFERENT PAD MULTIPLICITY WINDOWS.

N_{pad}	ϵ_γ	ϵ_p	Q_f
100-200	0.34 ± 0.01	0.85 ± 0.01	0.88 ± 0.02
200-300	0.45 ± 0.01	0.78 ± 0.01	0.96 ± 0.03
300-500	0.87 ± 0.01	0.33 ± 0.01	1.06 ± 0.01
500-1000	0.84 ± 0.01	0.54 ± 0.01	1.24 ± 0.03
1000-2000	0.81 ± 0.02	0.62 ± 0.01	1.31 ± 0.05
2000-3000	0.80 ± 0.02	0.63 ± 0.02	1.32 ± 0.09
3000-10000	0.76 ± 0.02	0.67 ± 0.02	1.32 ± 0.12

The results listed in tab. I for the different pad multiplicity windows show that the method is efficient at quite high multiplicity ($N_{hit} > 500$): the following discussion is related to events firing at least 500 pads on the central carpet, corresponding to a median energy of $\sim 8TeV$

for γ showers. The efficiency in the identification of γ showers, decreasing with increasing hit multiplicity, is $\sim 80\%$, while the efficiency in the identification of proton showers, increasing with the hit multiplicity, is $\sim 60\%$. The worsening (improvement) in the efficiency of γ (proton) identification when increasing the number of fired pads on the detector is due to the fact that the higher the primary energy the more extended the region around the core in which the shower features have to be studied - so at very high energies this area is too wide to be entirely held within the ARGO-YBJ carpet. This means that the images of γ showers appear to be less sharp and smooth than they actually are, so they are more likely to be mistaken with proton showers, which on the contrary are correctly identified in a higher number of cases just because the features of γ showers are not accurately defined. In any case, the improvement in proton identification compensates for the worsening in γ identification and the value of Q_f is almost stable ($Q_f \sim 1.3$) for $N_{hit} > 1000$.

VI. THE CRAB NEBULA

The Crab Nebula data taken from day 1, 2008 to day 116, 2008 corresponding to ~ 515 hours of data taking (equivalent time: ~ 87 days) have been analyzed with the GHD technique. The analysis has been performed applying 5 multiplicity thresholds ($N_{hit} > 500, 600, 700, 800, 1000$); the background has been calculated with the time swapping method [4] and fixed smoothing windows have been used to build the maps.

Three different maps have been filled for each multiplicity threshold:

- (1) No cuts have been applied on the core position: all the events with their multiplicities above the corresponding threshold and having $\theta < 40^\circ$ have been used to build the map. These maps correspond to the *standard* analysis;
- (2) The events have been selected according to the criteria adopted for Monte Carlo simulations: out of the events in sample (a), only those having their cores located within the detector central carpet have been analyzed;
- (3) Only events of sample (b) tagged as γ by the GHD method have been used to fill the maps.

The results are listed in tab. II. For each multiplicity threshold the smoothing window and the significance of the detection are reported. The maps corresponding to $N_{hit} > 800$ are shown in fig. 2.

TABLE II
SIGNIFICANCE OF THE SIGNAL AT THE NOMINAL CRAB POSITION FOR THE THREE EVENT SELECTIONS (SEE TEXT).

$N_{hit, threshold}$	Smoothing window (deg)	σ_1	σ_2	σ_3
500	0.5	0.6	1.9	2.3
600	0.5	0.5	2.2	2.7
700	0.5	0.8	2.3	2.8
800	0.5	1.4	2.7	3.5
1000	0.4	0.3	2.2	2.2

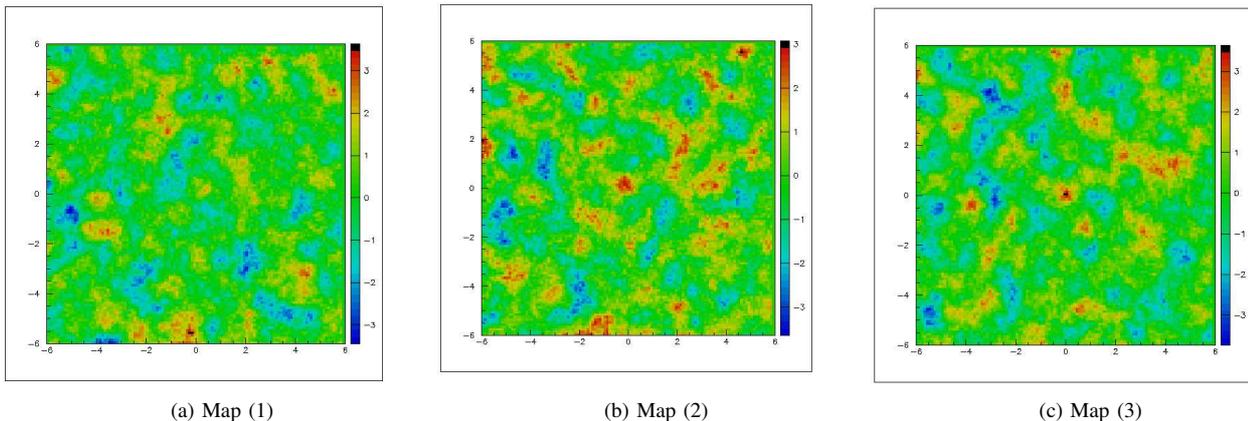


Fig. 2. Crab Nebula significance maps filled using events with $N_{hit} > 800$ and $\theta < 40^\circ$: (a) all the events; (b) reconstructed internal events; (c) events tagged as γ by the GHD method.

A clear improvement in the significance of the signal in all the multiplicity windows is achieved with the selection of internal events: this improvement is mainly due to the better angular resolution of internal events. The improvement in the significance when the GHD method is applied is in agreement with the expectations from Monte Carlo simulations (with the exception of the map with $N_{hit} > 1000$, where no improvement can be seen). A first check can be simply done by calculating the Q_f as the ratio between the significance of the signal when the GHD method is applied and the significance without discrimination: the Q_f values range from 1.2 to 1.3. A further confirmation of the expectations comes from a more detailed analysis on the behaviour of the GHD method on the signal and on the background events: for example, in the window centered in the Crab position, the number of events is $N_{signal} = 129.4$ in the map without discrimination, whereas it is $N_{signal} = 105$ in map (3). This means that $\sim 81\%$ of the signal is identified by the method used. As far as the background events are concerned, they are $N_{bkg} = 1926.6$ in map (2), while their number decreases to $N_{bkg} = 720.9$ in map (3). This means that $\sim 63\%$ of the background events are correctly recognized as noise by the GHD method. A rough estimation of the experimental value for the quality factor, based on the fraction of signal and background events identified by the method, is $Q_f = \frac{0.81}{\sqrt{1-0.63}} = 1.3 \pm 0.1$.

VII. CONCLUSIONS

An original gamma-hadron discrimination technique based on the analysis of the shower image on the ARGO-YBJ detector has been implemented: the image undergoes the Radon procedure - that is the subsequent application of a discrete Radon transform and of its backprojection - and then the features of the hottest spot left over after the application of a cut at half the maximum intensity of the backprojection are used to tag the shower as being a signal or a background event.

The application of the Radon procedure gives information on the local particle density calculated on a region whose shape depends on the real shower particle distribution and not - as usual - on a fixed shape of growing dimension.

A detailed Monte Carlo study has been done in order to tune the method; the resulting quality factor is $Q_f \sim 1.3$ for showers with $\theta < 40^\circ$, $N_{hit} > 500$ and core located on the whole central carpet.

If compared with the gamma-hadron discrimination techniques used by Čerenkov Telescopes or by Milagro, this method seems to be less efficient, but the information held in the lateral distribution of the electromagnetic component, even if studied in a such detailed way, does not suffice for a much better discrimination.

Moreover, the application of the method leads to an improvement of the sensitivity only at high energies, being $E \sim 8 \text{ TeV}$ the median energy of γ showers with $N_{hit} > 500$. Nevertheless, the method has been applied to the analysis of Crab Nebula data: 515 hours of data (corresponding to 87 days of Crab observation) have been analyzed and the significance maps have been built for different multiplicity thresholds. The improvement in the signal significance is consistent with the expectations from Monte Carlo simulations.

VIII. ACKNOWLEDGEMENTS

We would like to thank the National Institute for Astrophysics (INAF) for partially financing this work.

REFERENCES

- [1] A.A. Abdo *et al.*, *Astroph. J.*, 658 (2007) 33
- [2] G. Aielli *et al.*, *Nucl. Instr. and Meth. A*, 562 (2006) 92
- [3] G. Aielli *et al.*, *Astrop. Phys.* 30 (2009) 287
- [4] D. Alexandreas *et al.*, *Nucl. Instr. and Meth. A*, 328 (1992) 570
- [5] R.A. Fisher, *Annals of Eugenics*, 7 (1936) 179
- [6] D. Heck *et al.*, *Forschungszentrum Karlsruhe Report FZKA 6019* (1998)
- [7] A.M. Hillas, *Proc. 19th ICRC, La Jolla*, 3 (1985) 445
- [8] J. Radon, *Ber. Ver. Sächs. Akad. Wiss. Leipzig, Math-Phys. Kl.* 69 (1917) 262