Multiscale image analysis applied to γ/h discrimination for VHE gamma-ray astronomy with ARGO-YBJ

I. De Mitri * and F. Salamida

Dipartimento di Fisica dell'Università di Lecce and INFN Sezione di Lecce, I-73100 Lecce, Italy

On behalf of the ARGO-YBJ Collaboration

Abstract

Intrinsic differences in the processes involved in the electromagnetic and hadronic shower development in the atmosphere have been evidenced by means of a careful analysis of the event image given by the ARGO-YBJ detector. The images have been analyzed at different length scales and their multifractal nature has been studied. The use of the multiscale approach together with a properly designed and trained Artificial Neural Network, allowed us to obtain a good discrimination power. If confirmed by further studies on different event cathegories, this result would allow to nearly double the detector sensitivity to gamma ray sources.

Key words: Multiscale Analysis, Gamma Ray Astronomy
PACS: 98.70.Sa, 07.05.Kf, 07.05.Mh, 07.05.Pj

1. Introduction

Gamma ray astronomy at energies around 100 GeV-10 TeV is the main scientific goal of the ARGO-YBJ experiment [1]. The detector, which is now being assembled in Tibet (China) at 4300 m a.s.l., is a full coverage Extensive Air Shower array consisting of a Resistive Plate Chamber (RPC) carpet covering a total surface of more than $6000\,\mathrm{m}^2$ (see Fig.1). It is logically divided into 154 units called *clusters* ($7.64 \times 5.72\,\mathrm{m}^2$), made by 12 RPCs. Each RPC ($1.26 \times 2.85\,\mathrm{m}^2$) is read out by 10 pads ($62 \times 56\,\mathrm{cm}^2$), which are further divided into 8 different strips ($62 \times 7\,\mathrm{cm}^2$) providing the highest available space resolution. The signals coming from all the strips of a given pad are sent to the same channel of a multihit TDC. The whole system is designed in order to provide a single hit time resolution at the level of 1 ns, thus allowing a complete and detailed three-dimensional reconstruction of the shower front. The high altitude ($\sim 606\ \mathrm{g/cm}^2$) and the full coverage ensure a very low primary energy threshold ($E_\gamma\ \approx$100 GeV), while the detector time resolution gives a good pointing accuracy, thus allowing a high sensitivity to γ-ray sources.

Gamma/hadron discrimination is a key issue in Very High Energy (VHE) gamma ray astronomy since it allows, together with a good angular resolution, the rejection of the huge background

* Correspondig author (ivan.demitri@le.infn.it)

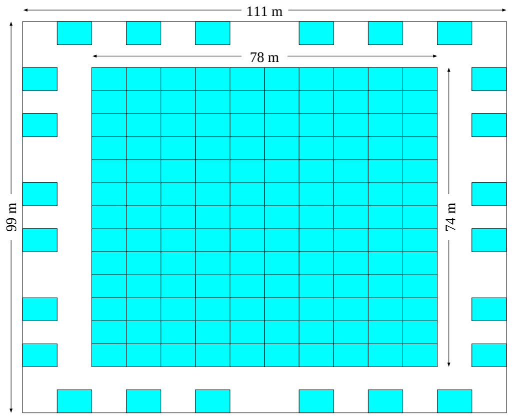

Fig. 1. *General layout of the ARGO-YBJ detector. Each indicated box represents a* $7.64 \times 5.72\ m^2$ *cluster of 12 RPCs (see text).*

due to charged primary hadrons. The use of a full coverage detector with a high space granularity - like ARGO-YBJ - can give detailed images of the shower front. Intrinsic differences in the processes involved in the electromagnetic and hadronic shower development in the atmosphere can then be evidenced by means of a careful analysis of the event[2,3]. Recently the use of multiscale behavior of event images has been showed to give good results in experiments exploiting the Imaging Atmospheric Cherenkov Techniques (IACT), together with the use of the so-called Hillas parameters [4,5]. In these experiments the image is integrated over the entire shower development, while in the case of ARGO-YBJ a section of the shower is provided at a given (slanted) depth only, thus giving a potential lack of information. However the typical disuniformities present in hadronic events might be better evidenced in the case of ARGO-YBJ , these being partially masked in the case of IACT detectors because of the smearing effect due to the integration of the information along the shower development.

In this work event images have been analyzed at different length scales and their multifractal nature has been studied. In particular the Discrete Wavelet Transforms have been applied since they allowed a differential approach to multifractality, that gave a higher discrimination power.

Since this is the first attempt of this kind of analysis in a EAS detector like ARGO-YBJ , we decided to restrict this study to events with the core at the detector center and zenith angles not larger than 15 degrees, while all the energies with the correct spectral dependencies have been simulated in a wide range. At this level we also neglected the contribution given by primary nuclei heavier than protons. This is a good first order approximation because of the proton-dominated cosmic ray composition in the considered energy region. Furthermore heavier-nuclei-induced showers would produce event patterns with characteristics even more different from gamma-initiated ones.

2. The multiscale event analysis

We have considered the shower image seen by ARGO-YBJ as a function defined on a two-dimensional space and corresponding to the amplitudes given by the measured strip multiplicity. As a first step we can calculate the multifractal (MF) moment $Z_q(\ell)$ of order q at the length scale ℓ following the definition given in [6]. As pointed out in [4–6], at each order q, the MF moment is expected to have a power law dependence on ℓ in the high resolution limit, namely: $Z_q(\ell) \sim \ell^{\tau(q)}$ for $\ell \to 1$. Therefore, by fitting the behavior of $Z_q(\ell)$ on ℓ (for each value of q) the MF scaling exponent $\tau(q)$ can be extracted. The dependence of $\tau(q)$ on q gives the main information on the MF properties of the event. As shown in [6], the MF approach might not sufficiently characterize the image. An approach to multifractality based on the discrete wavelet transformations (DWT) is more appropriate. DWT can be seen as an expansion of the event image on a discrete set of basis functions that are generated by scaling a so-called mother wavelet. Also in this case a DWT moment $W_q(\ell)$ of order q at lenght scale ℓ can be defined and easily calculated [6]. The scaling properties of the image can be evidenced at high resolutions: $W_q(\ell) \sim \ell^{\beta(q)}$ when $\ell \to 1$. The DWT scaling exponent $\beta(q)$ gives useful information on the analyzed event.

We have generated $\sim 2.8 \cdot 10^5$ gamma-initiated showers and $\sim 2.6 \cdot 10^5$ proton-initiated ones making use of the CORSIKA code [7]. The events have been taken within the energy range

30 GeV÷100 TeV with zenith angle between 0 and 15 degrees and core at the detector center. The primary energy spectra have been generated according to a power law with spectral index γ=2.5 for gammas (the CRAB nebula case) and γ=2.7 for hadrons (as measured for cosmic rays). The previously defined multiresolution quantities have been used to analyze each event image. As shown in Fig.1, the ARGO-YBJ detector is made by a central carpet and a guard ring. In order to preserve the same symmetry at different lenght scales ℓ, we decided to neglect, in this first analysis, the information coming from the external ring. We also decided to *mask* the central carpet with a square grid. In particular, since it is made of (120×130)pads, the first and the last row of pads were not considered, while four empty columns of pads were added (two on the left and two on the right), thus obtaining a (128×128) pad mask. In order to limit statistical fluctuations of the hit multiplicity in the smallest pixels, the minimum pixel size considered in the analysis was set at (2×2)pads - about 1.4 m^2 - which will then corresponds to the maximum resolution, i.e. ℓ=1. The analysis of an event goes then throught different steps, each correponding to different lenght scales. At the n-th step (with $n = 1, 2, ..., 6$), the image is divided into $2^n \times 2^n$ square pixels of size $\ell_n = \frac{64}{2^n}$, containing each $4\ell_n^2$ pads, and the total strip multiplicity is computed in each considered pixel.

The values of Log(Z$_q(\ell)$) and Log(W$_q(\ell)$) have been calculated and their dependences on Log(ℓ) have been fitted with a first order polynomial in the region where the scaling is expected (i.e. $\ell \to 1$), for different values of the moment q. The scaling exponents τ(q) and β(q) have then been obtained for each event. Their dependences on the moment q are shown in Fig.2 for a sample of gamma and proton initiated events. As can be seen there is a separation of the average values of the scaling exponents between electromagnetic and hadronic showers, which is however partly masked by the large fluctuations. Therefore the separations of the MF and DWT parameters are not sufficient in order to give a good discrimination between e.m. and hadronic showers, unless an Artificial Neural Network (ANN) is used as in ref.[4,5]. We then decided to use multifractal parameters as inputs to a properly designed and trained ANN.

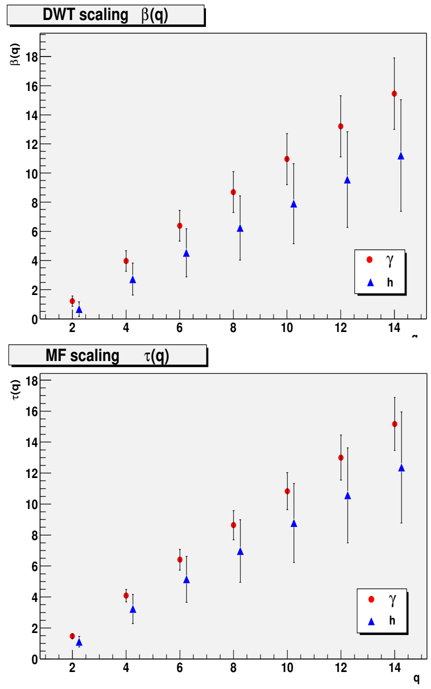

Fig. 2. *Dependences of the scaling exponents on the moment q for a sample of gamma and proton initiated-showers with energy from 3 TeV to 10 TeV . The error bars refer to the r.m.s values of the variuos distributions.*

In order to increase the γ/h separation, a study on the shape and the simmetry of the event image has also been made. In particular we studied the skewness of each event by means of the third moment of the distributions of the hit coordinates in the detector plane, namely x and y. The skewness has been found to be useful in γ/h separation and has been added to the list of the ANN input parameters.

The Neural Network we have chosen is of the *feed forward* type and it is made of 3 perceptrons layers. The ANN input is an eight-dimensional vector whose elements are: the event total hit multiplicity N_{hit}, the ratio of the skewness along x and y, the multifractal exponents τ and β for $q = 4, 6, 8$. The output vector is defined in a one dimensional space: it is trained to be 1 for gamma-initiated events and 0 for hadronic showers. Networks were implemented and optimized by using the **S**tuttgart **N**eural **N**etwork **S**imulator (SNNS) tool [8]. In designing the ANN, its characteristics have been deeply studied in order to reach a good compromise between the increase of recognition capability and the processing time. The network training was separately performed in 5 multiplicity windows by using several thousands events. The ANN were then tested by using an independent reduced sample of events and the γ recognition efficiency ε_γ, together with the proton contamination $(1 - \varepsilon_p)$, were measured. The ANN output for a couple of multiplicity windows is shown in Fig.3.

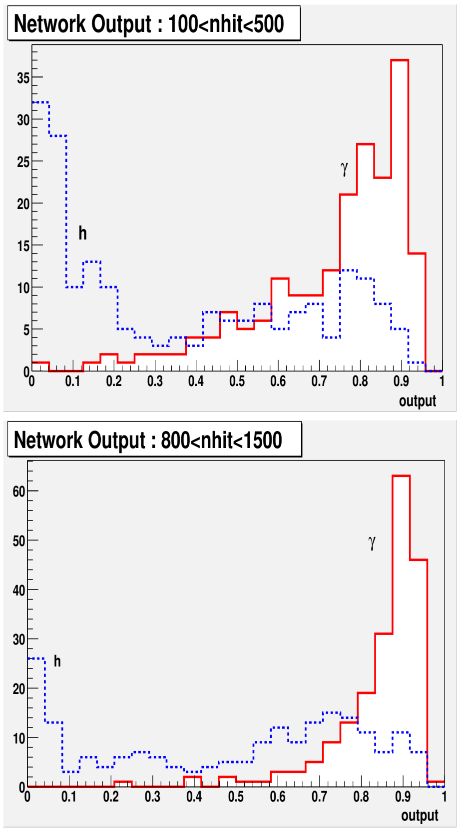

Fig. 3. *Outputs of the neural network in two of the five considered multiplicity regions.*

3. Results and Conclusions

The detector sensitivity to γ-ray sources is defined as $S = N_\gamma/\sqrt{N_h}$, where N_γ is the number of gamma-initiated events, while N_h is the hadron contamination of the considered sample. The use of

nhits	events (γ)	events (p)	$<E_p>$	$<E_\gamma>$	Q
50÷100	6657	3862	0.8 TeV	0.5 TeV	1.28±0.01
100÷500	11556	6862	1.8 TeV	1.1 TeV	1.42±0.02
500÷800	2571	1644	4.9 TeV	2.9 TeV	2.01±0.10
800÷1500	3087	1963	7.6 TeV	4.6 TeV	1.78±0.07
1500÷6000	4329	3053	18.4 TeV	11.3 TeV	1.78±0.06

Table 1
Main characteristics of the simulated data sample (no. of ANN training events, average primary energy, ...) together with the values of Q for γ/h discrimination that resulted from this work.

a γ/h discrimination tool, like the one we are considering here, makes the sensitivity S to be multiplied by the factor $Q = \varepsilon_\gamma/\sqrt{1 - \varepsilon_h}$. The value of Q depends on the event multiplicity, i.e. on the primary photon energy. In this work values of Q =2 have been reached (see Tab.1), which are among the largest obtained in the experiments working in the field [5]. In particular the best performances in γ/h discrimination have been obtained for photon primary energies in the few TeV range, while at higher energy this analysis might be well complemented by measuring the muon content of the shower. If the results obtained in this study will be confirmed by a further analysis on the whole event categories (now in progress), the detector sensitivity to a given source would nearly double or, equivalently, the time needed to observe it above the hadron background, with a given statistical significance, would be reduced by a factor four.

References

[1] C.Bacci et al., Astroparticle Phys. **17**, 151 (2002)

[2] R.S.Miller and S.Westerhoff, Astroparticle Phys. **11**, 379 (1999)

[3] S.Bussino and S.M.Mari, Astroparticle Phys. **15**, 65 (2001)

[4] A. Haungs et al., Astroparticle Phys. **12** 145 (1999)

[5] B. M. Schăfer et al., Nucl. Instr. & Meth. in Phys. Res., **A465** 342 (2001)

[6] Jan W. Kantelhardth et al., Physica **A220** 219 (1995)

[7] D.Heck et al., Report FZKA 6019, Forschungszentrum, Karlsruhe (1998)

[8] *http://www-ra.informatik.uni-tubingen.de/SNNS/*