# VHE $\gamma$-RAY ASTRONOMY WITH ARGO-YBJ: TOOLS FOR GAMMA/HADRON DISCRIMINATION

IVAN DE MITRI [a,b]

ON BEHALF OF THE ARGO-YBJ COLLABORATION

[a] *Dipartimento di Fisica, Università di Lecce, via per Arnesano, 73100 Lecce, Italy*

[b] *INFN, Sezione di Lecce, via per Arnesano, 73100 Lecce, Italy*

**Abstract**

The ARGO-YBJ experiment is devoted to the study of many issues in $\gamma$-ray astronomy and cosmic ray physics. The apparatus design is based on Resistive Plate Chambers (RPCs) operated in streamer mode and assembled in a single layer, covering a surface of more than $6000 m^2$. The detector installation is now going on in the YangBaJing Laboratory (Tibet, China) at 4300 $m$ above the sea level.

In this paper the present status of the experiment, the performances of the detector and the analysis of the first data will be presented, together with the results of a study devoted to $\gamma$/hadron discrimination.

In particular, 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 on the event image as given by a full simulation of the detector response. 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.

## 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 carpet of bakelite Resistive Plate Chambers (RPCs) covering a total surface of more than $6000\,\mathrm{m}^2$. The RPCs [2, 3] are operated in streamer mode, with a mixture of argon (15%), isobutane (10%) and tetrafluoroethane (75%).

The detector layout is shown in Fig. 1. It is logically divided into 154 units called *clusters* ($7.64 \times 5.72\,m^2$), made by 12 RPCs. Each RPC ($1.26 \times 2.85\,m^2$) is read out by 10 pads ($62 \times 56\,cm^2$), which are further divided into 8 different strips ($62 \times 7\,cm^2$) 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 (see an example in Fig. 2).

The high altitude ( the experimental site is located at an atmospheric depth of $\sim 606$ g/cm$^2$ ) and the full coverage ensure a very low primary energy threshold ($E_\gamma \approx$100 GeV, close to the limits of the satellite technology), while the detector time resolution gives a good pointing accuracy, thus allowing a high sensitivity to $\gamma$-ray sources. Moreover the large aperture and the high duty-cycle ensure the continuous monitoring of the sky in the declination band $-10° \div 70°$.

Data gathered with ARGO-YBJ will then allow to face with a wide range of fundamental issues:

- $\gamma$-ray astronomy, looking for point-like (galactic and extra-galactic) sources with an energy threshold of few hundreds GeV and for diffuse fluxes from the Galactic plane and/or SuperNova Remnants;
- Gamma Ray Burst (GRB) physics, extending the satellite measurements over the $GeV - TeV$ energy range [4];
- Cosmic Ray (hereafter CR) physics [5], that is measurements of antiproton/proton ratio at $TeV$ energy, studies of spectrum and composition around the knee ($E > 10\ TeV$);
- Sun and Heliosphere physics ($E > 10\ GeV$), looking for CR modulation, monitoring the interplanetary magnetic field and observing flares of high energy gammas and neutrons from the Sun.

In this work the present status of the experiment, the performances of the detector and the analysis of the first data will be presented, together with the results of a study devoted to $\gamma$/h discrimination. A different and complementary approach to $\gamma$/h discrimination is also presented in this workshop in a separate paper [6].

## 2 Detector performance and first measurements

Two main kinds of trigger have been designed for the data-acquisition: the *shower mode* and the *scaler mode*. In the first one, a minimum pad multiplicity is required on the central carpet, with a space/time pattern consistent with the one expected from a shower front. In the *scaler mode* the pad rate is measured from each cluster, with an integration time of 0.5 s. This last DAQ mode is devoted to the apparatus monitoring and the detection of unexpected increases in CR flux, as an effect of

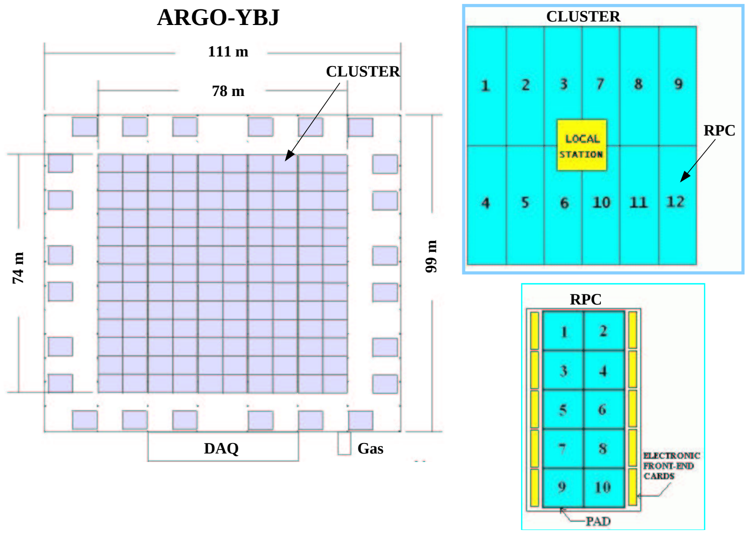


Figure 1: *General layout of the ARGO-YBJ detector (see text).*

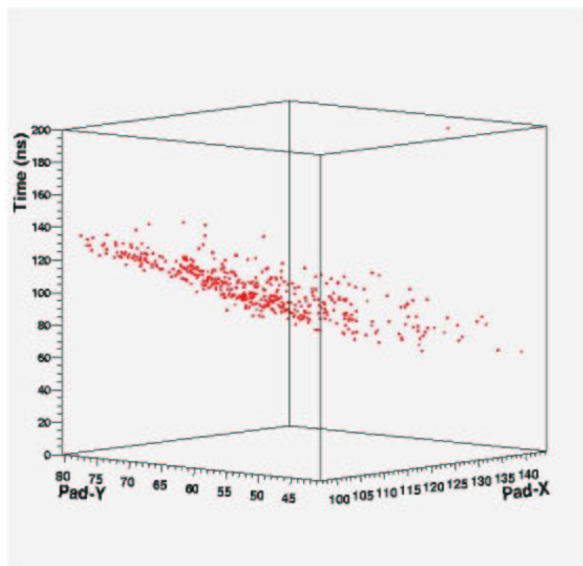


Figure 2: *Display of a real event in space-time view.*

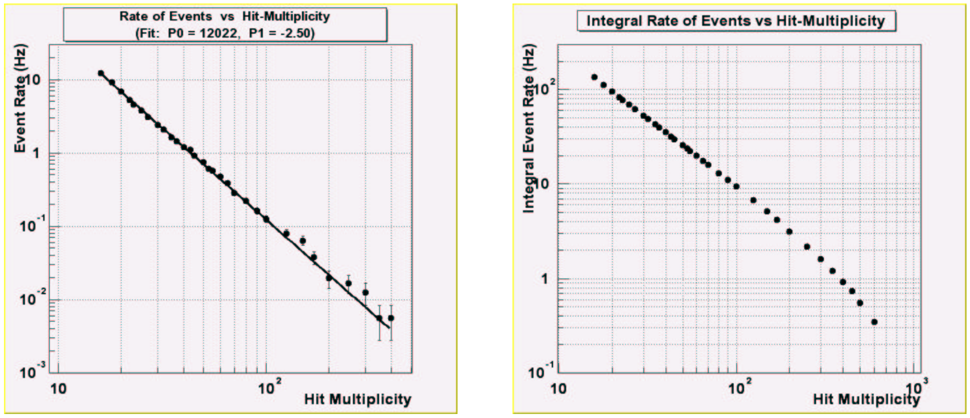


Figure 3: *Differential and integral rates versus hit multiplicity (trigger condition: more than 15 hits).*

GRB, solar flares and so on. The RPC operation is also successfully monitored by a Detector Control System (DCS), able to record HV, currents, temperature, humidity, pressure and gas flow.

After the collection of a first data sample, a modification to front-end electronics became mandatory in order to fix some unexpected damages occuring with few showers having huge particle density. Presently these problems have been fixed and 16 cluster have been in data-taking since many months with high stability and without any significant damages. The detector assembling is therefore going on and 48 clusters are expected to be operating within December 2004.

Here the first data collected in *shower mode* with 6 clusters (260 $m^2$), after the upgrading of the front-end electronics, are presented. The analysis results are in full agreement with those from data collected in the past with 16 clusters (before the aforementioned damages) and with what is expected from CR and EAS properties.

### Rate measurements

In Fig. 3 differential and integral rates are plotted versus the hit multiplicity. The differential rate follows a power law with slope $\eta \simeq -2.5$, in agreement with what estimated on the basis of the CR spectrum. Moreover the measured rates of the different triggers are consistent with the expected ones and the distribution of the time difference between consecutive events (Fig. 4) confirms the regular operation of the detector.

### Time calibration and angular resolution

Dedicated runs with high multiplicity trigger (more than 32 hits on each cluster) have been used for a preliminary time calibration. The goal of the procedure was

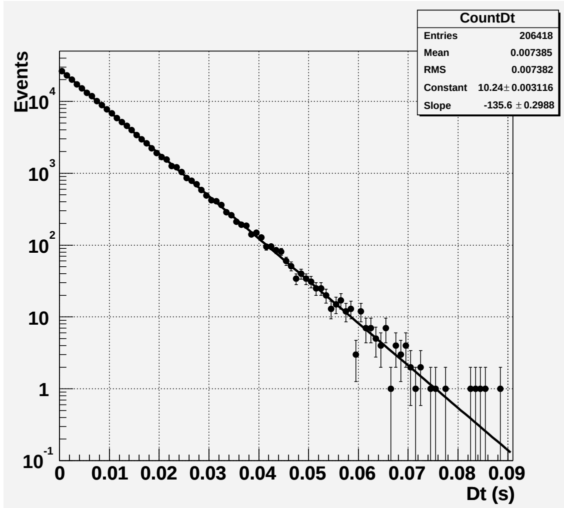


Figure 4: *Time difference between consecutive events (trigger condition: more than 32 hits on each cluster).*

to remove systematic time-offsets among the pads, due to different cable lenghts or any other electronic effects.

Corrections in the range $\pm 4$ $ns$ have been introduced to minimize the time residuals, that is the difference between the time measured by the pad and the time of the shower estimated by a planar fit of the shower front. After these corrections, the peak values of the residual distributions are smaller than 0.2 $ns$.

The uniformity of the azimuthal distribution is another parameter taken into account in the calibration. A systematic correction of the time measurement has been introduced according to the method suggested in [7]. The effect of this second correction on the azimuthal distribution is shown in Fig. 5. Also the symmetry of the direction cosine distributions improves as an affect of this systematics correction.

The angular resolution of the 6-cluster carpet has been estimated by dividing the detector into two independent sub-arrays ("odd" and "even" pads) and studying the angular difference $\Delta\varphi$ between the reconstructed shower directions. The angular resolution depends on the width of the $\Delta\varphi$-distribution [1] and decreases with the hit multiplicity, as shown in Fig. 6.

## Angular distributions of Extensive Air Showers

In the left side of Fig. 7 the distribution of the reconstructed zenith angle $\theta$ is shown, while the distribution of $sec\ \theta - 1$ is displayed on the right. By means of the fit with the function $e^{-\alpha(sec\theta-1)}$ we get $\alpha = 4.678 \pm 0.016$. Then we estimate $(128.3 \pm 0.4)$ $g/cm^2$ as attenuation length of showers, in excellent agreement with previous measurements [1].

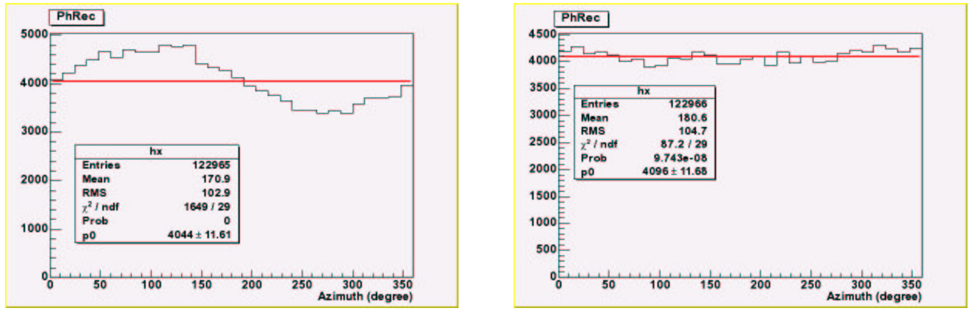


Figure 5: *Azimuthal distribution before and after the time calibration. The $\chi^2$ of the uniform fit is used as the estimator of the symmetry of the detector response.*

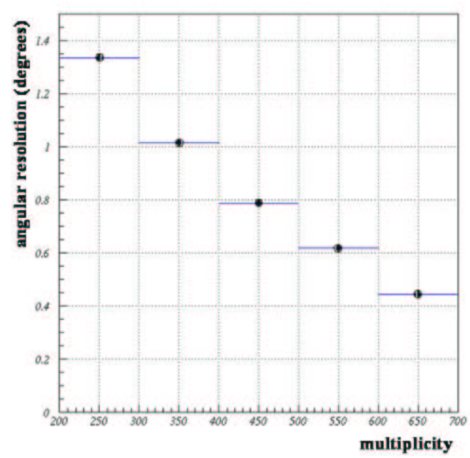


Figure 6: *Angular resolution versus hit multiplicity for the first data taken with six clusters (see text).*

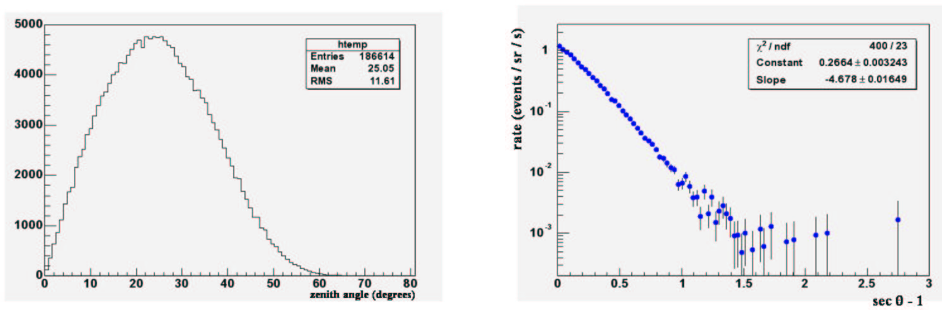


Figure 7: *Distributions of the zenith angle $\theta$ and of the quantity* ($sec\ \theta - 1$).

# 3 A tool for $\gamma/h$ discrimination

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 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[8, 9]. The use of multi-scale 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 [10, 11]. 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.

### 3.1 A multiscale event analysis and its results

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 $\ell$ following the definition given in [12]. As pointed out in [10, 11, 12], at each order $q$, the MF moment is expected to have a power law dependence on $\ell$ 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 $\ell$ (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 [12], 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 $\ell$ can be defined and easily calculated [12]. 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 [13]. 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 $\gamma$=2.5 for gammas (the CRAB nebula case) and $\gamma$=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 $\ell$, 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. $\ell$=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($\ell$) 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

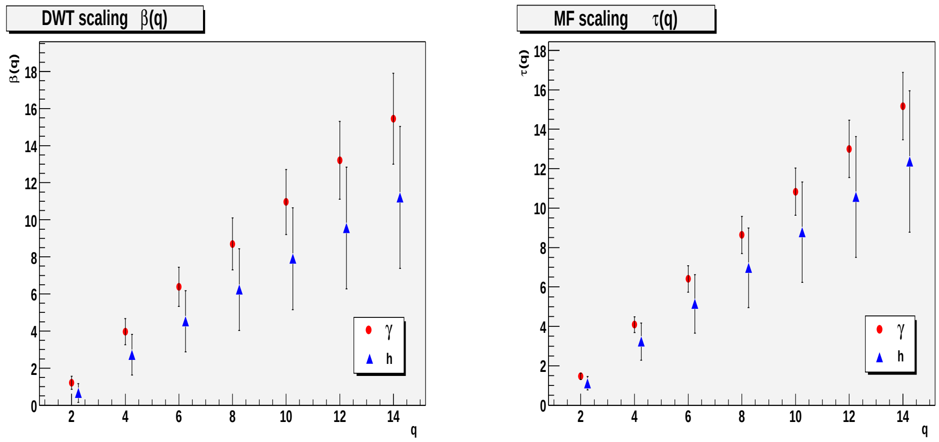


Figure 8: *Dependencies of the scaling exponents on the moment q for a sample of gamma and proton initiated-showers with energy from 3 TeV to 10 TeV . The error bars refer to the r.m.s values of the variuos distributions.*

$q$. The scaling exponents $\tau$(q) and $\beta$(q) have then been obtained for each event. Their dependences on the moment $q$ are shown in Fig.8 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.[10, 11]. We then decided to use multifractal parameters as inputs to a properly designed and trained ANN.

In order to increase the $\gamma$/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 $\gamma/h$ 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 $\tau$ and $\beta$ 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 [14]. 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

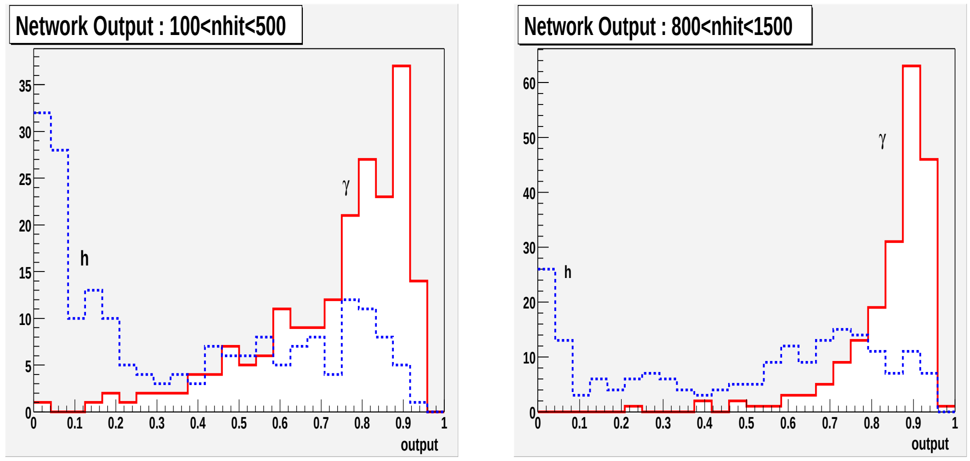


Figure 9: *Outputs of the neural network in two of the five considered multiplicity regions.*

reduced sample of events and the $\gamma$ recognition efficiency $\varepsilon_\gamma$, together with the proton contamination $(1-\varepsilon_p)$, were measured. The ANN output for a couple of multiplicity windows is shown in Fig.9.

The detector sensitivity to $\gamma$-ray sources is defined as $S = N_\gamma/\sqrt{N_h}$, where $N_\gamma$ is the number of gamma-initiated events, while $N_h$ is the hadron contamination of the considered sample. The use of a $\gamma$/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 [11]. As can be seen in Fig.10 the best performances in $\gamma$/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 [6]. If the results obtained in this Monte Carlo 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.

| nhits | events ($\gamma$) | 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 $\gamma/h$ discrimination that resulted from this work.*

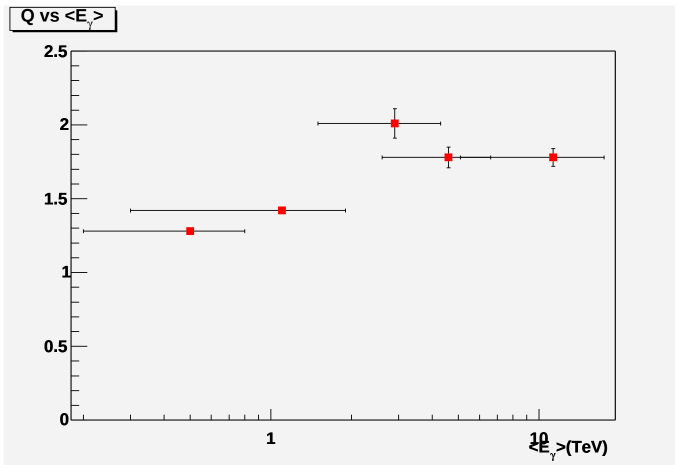


Figure 10: *Values of Q vs the average primary photon energy in the analyzed multiplicity regions. The bars on the horizontal axis refer to the r.m.s. of the energy distributions.*

## 4 Conclusions

The construction of the ARGO-YBJ detector is going on. The data collected with part of the apparatus have been used to check the detector performances and the analysis codes. The detector works as good as expected. Absolute trigger rates, shape of hit multiplicity distribution and preliminary shower reconstruction are consistent with the CR physics.

We foresee that 48 clusters ($\sim 2000\ m^2$) will be in data-taking at the end of December 2004 and that the central carpet will be completed within the begin of 2006. Stable data-taking and physics runs are expected already at the end of 2004. First physics results are close.

In the meantime tools for $\gamma/h$ discrimination are being studied on simulated events in order to make the detector sensitivity to $\gamma$-ray sources even higher. 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. Different approaches to $\gamma/h$ are viable at higher energies, mainly based on the measurement of the muon content of the showers. Also in this case promising results are being obtained [6].

## References


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

[2] ARGO-YBJ Collaboration, 6th International Conference on Advanced Technology and Particle Physics, Como (1998), Nucl. Phys. B (Proc. Suppl.) **78**, 38 (1999)

[3] C. Bacci et al.(ARGO-YBJ Coll.), Nuclear Instr. and Methods in Phys. Res. **A443**, 342 (2000)

[4] ARGO-YBJ Collaboration, Workshop "Gamma Ray Bursts in the Afterglow Era", Rome (1998), in Astronomy and Astrophysics, Suppl. Series **138** 597 (1999), also astro-ph/9904373

[5] L. Saggese et al. (ARGO-YBJ Collaboration), Proceedings of the 28th International Cosmic Ray Conference (Tsukuba), (2003) 263 [HE.1.2.9]

[6] K. Fratini et al. (ARGO-YBJ Coll.) *these proceedings*

[7] A.M. Elø, H. Arvela, Proceedings of 26th International Cosmic Ray Conference (Salt Lake City), **5** (1999) 320 [OG.4.4.07], 324 [OG.4.4.08] and 328 [OG.4.4.09]

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

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

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

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

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

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

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