



The ANN method to measure proton spectra below the ‘knee’ region with the ARGO-YBJ experiment

S.Z CHEN¹, M ZHA¹, AND H.H HE¹ ON BEHALF OF ARGO-YBJ COLLABORATION

¹Key Laboratory of Particle Astrophysics, Institute of High Energy Physics, Chinese Academy of Sciences, Beijing 100049, Beijing, China

chensz@ihep.ac.cn

Abstract: Based on Monte Carlo Simulation, the space-time information of the shower is used to study the difference between different primaries collected by ARGO experiment. With one month exposure, the proton spectrum from 20 TeV to 200 TeV energy region can be obtained with the data sample of digital read-out selected by an artificial neural network. With the analog read-out data, the energy region could be extrapolated to ‘knee’ region.

Introduction

The energy of cosmic ray (CR) particles can exceed many orders of magnitude from GeV energies up to at least 10^{20} eV. The all-particle energy spectrum has a steepening around several PeV, generally called the ‘knee’. To interpret this feature, many approaches have been discussed from the origin, acceleration and propagation of CR [1]. Even the mechanism is still obscure, but it is believed to be a corner stone in understanding the origin of CR now. The energy spectra of individual elements are believed to be the key point to understand the ‘knee’ mechanism. But different experiments give different results. The Tibet $AS\gamma$ experiment thinks that the ‘knee’ is caused by the steepening of heavy nuclei and proton should steepen at about 100 TeV [2]. While according to KASCADE experiment unfolding result, the ‘knee’ would be caused by light elements [3]. But the balloon-borne experiments above the atmosphere, RUNJOB [4] and JACEE [5], don’t see any feature up to 500 TeV.

Operated at high altitude and full coverage, the ARGO-YBJ is unique in better understanding the ‘knee’ problem. Its low energy threshold provides a “bridge” between direct and indirect measurement. While, the detailed space-time information of shower can be used to discriminate different primaries. In this paper by using the artificial neural

network(ANN) method we explore the feasibility of ARGO-YBJ to obtain the spectrum of different components.

The ARGO-YBJ detector

The ARGO-YBJ experiment [6], located at the Yangbajing Tibet of China (4300 m a.s.l), consists of a full coverage of detector of area $78 \times 74 \text{ m}^2$ realized with a single layer of RPCs, $280 \times 125 \text{ cm}^2$ each. The area surrounding this central carpet, up to $110 \times 100 \text{ m}^2$, is partially ($\approx 50\%$) instrumented with RPCs (the guard ring). A cluster consists of 12 contiguous RPCs while one RPC is divided into 10 basic detection units, which are called PADs and each has 8 digital readout strips. With the digital read-out ARGO reaches its energy upper limit at about 100 TeV, which can be extended to the ‘knee’ region by means of analog read-out (two BigPADs for each RPC).

The central carpet (10×13 clusters) has been collecting data with digital read-out since June, 2006, and the analog read-out is still under construction now. In the following, only the central carpet will be used and this central carpet is artificially divided into two parts: the internal detectors (the 6×9 clusters in the center) and the external detectors (the 76 clusters around).

Method

The general procedure for ANN method [7] is adopted here: with assuming of mixed composition of CR from some measurements, a large Monte Carlo simulation samples have been collected; we divide it to equally 2 parts, 50% of artificial samples are used to train the ANN, the other 50% are used to determine the energy spectrum.

Simulation

CORSIKA code [8] with QGSJET-II model is used to generate air showers induced by 5 primary chemical components: protons (P, 1), helium (He, 4), light nuclei(CNO, 7), median nuclei (Mg-Si, 13) and heavy nuclei (Fe, 56). Their flux in the low energy region is got by fitting the direct measurement result. The best fit in polygonato model(rigidity dependent on bending point $Z \times 4.49$ PeV, constant difference $\delta\gamma = 2.10$ between the spectral indices below and above the knee) [9] is used here to extrapolation to higher energy region. After normalization at 10 TeV, the component fractions of this hybrid model at 10 TeV, 100 TeV, 1 PeV and 10 PeV are listed in table 1.

The incident zenith angles of primary particles are isotropically sampled within 20° . The energy region for all five groups is from 10 TeV to 10 PeV. The detector response has been simulated via a GEANT3-based code, and an array with 10×13 clusters, detector efficiency of 95%, noise 380 Hz/PAD and sampling area $100 \times 100 m^2$ are taken into account. The sampling area is large enough for the following selection of internal events, and this has been verified with a much larger sampling area.

Data selection

In order to obtain a good data sample for the following data analysis, the internal events are selected by following conditions at first:

1. the number of fired PADs is more than 1000;
2. the number of fired strips N_{strip} is more than $10^{3.6}$;

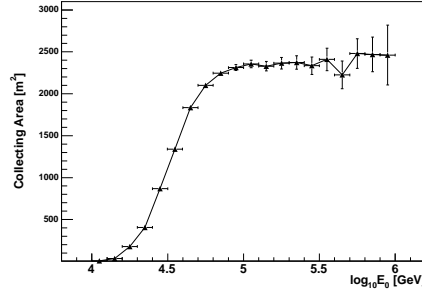


Figure 1: The effective collecting area for primary internal protons satisfying the selection conditions as a function of primary energy.

3. the strips density of the internal 54 Clusters is 1.1 times higher than that of the external 74 clusters;
4. the most fired Module (consists of 3 RPCs) is internal;
5. the reconstructed core is inside a fiducial area of 6×9 central clusters.
6. the reconstructed zenith angle is within 20° from the vertical;

After these cuts the efficiency to select internal P and Fe events are 99.3% and 95.7% respectively. In the end about 560 thousand simulation events are got, corresponding to a 24 days exposure of ARGO experiment. The fractions of different elements are P 48.09%, He 31.27%, CNO 10.23%, Mg-Si 4.74% and Fe 5.67%. The effective collecting area for the proton events satisfying the selection conditions as a function of energy is shown in figure 1.

Analysis using ANN

Each event obtained above can be characterized by four parameters:

- The number of total fired strips N_{strip} ;
- $\langle R \rangle (= \Sigma(N_i \times R_i) / \Sigma N_i)$: mean lateral spread radius of particle flow from the shower core position. N_i is the number of fired strips of the i th detector, and R_i is the distance to the shower axis;

Energy(GeV)	P	He	CNO	Mg-Si	Fe
10^4	29.1%	28.5%	16.1%	13.3%	13.0%
10^5	25.5%	31.5%	19.5%	10.4%	13.1%
10^6	21.1%	34.1%	23.5%	8.1%	13.2%
10^7	4.6%	25.0%	39.9%	9.7%	20.8%

Table 1: Composition fraction in hybrid model

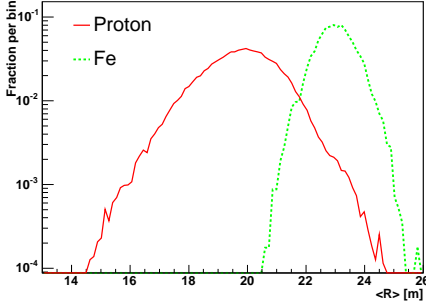


Figure 2: The difference between proton and Fe in parameter $\langle R \rangle$.

- N_{module} : the number of modules with fired strips more than 3;
- R_{slope} : the slope of the conical shower time front refer to the planar fit.

The proton-induced events can be characterized by small lateral spread and large conical slope, while those induced by heavy nuclei have the opposite character as their early development in the atmosphere. The difference between proton and Fe in parameters $\langle R \rangle$ and R_{slope} are shown in figure 2 and figure 3.

With these four input parameters, a 4 : 10 : 1 network is trained to discriminate the signal 0 (proton-induced events) from the background 1 with a half data sample. After 300 epochs the learning of the network becomes very stable and the network is able to select proton induced from the other half data sample. The figure 4 shows the output of the network within the range from 0 to 1. When the cut value of ANN output is set to 0.2, the fraction of correct classification is 90% and 46868 proton induced showers are selected.

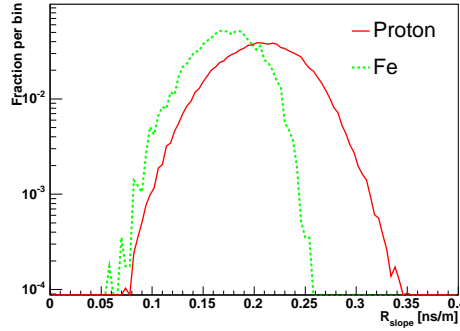


Figure 3: The difference between proton and Fe in parameter R_{slope} .

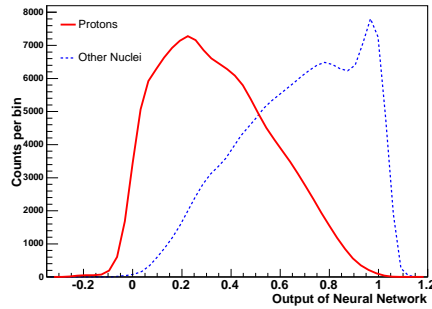


Figure 4: ANN output distribution of the proton-induced showers and other nuclei induced showers respectively.

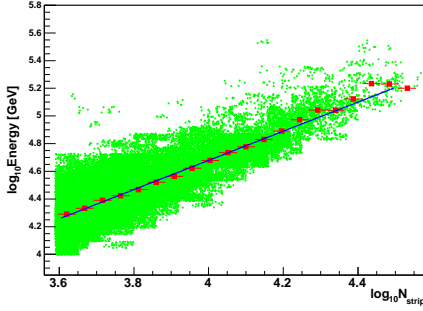


Figure 5: The scatter plots between the primary energy versus the number of total fired strip for the proton-induced events that selected by setting the ANN output value to 0.2. The solid circles denote the average values and solid curve is a fit using equation (1).

Result

The number of fired strips N_{strip} is in direct proportion to primary energy E_0 when the detectors are not saturated, thus the primary energy of each event can be estimated by the number of fired strips. The figure 5 shows the scatter plots between the E_0 and N_{strip} of proton events selected by ANN. The relationship between N_{strip} and E_0 is:

$$\log_{10}(E_0) = 0.448 + 1.058 \log_{10}(N_{strip}). \quad (1)$$

The uncertainty in this estimation is about 30%. With effective collecting area and ANN efficiency in selecting proton, the spectra of the proton can be determined and the result is shown in figure 6.

Discussion

Based on the result obtained above and assuming one month operation, a good agreement between the assumed and obtained spectra is seen in the energy region from 20 TeV to 200 TeV. This means that the ANN is effective to obtain the proton spectrum for ARGO experiment. The spectrum of other nuclei, like iron and helium, could be obtained with this method in principle. So, to get the spectrum of other nuclei and using the trained ANN to analyze

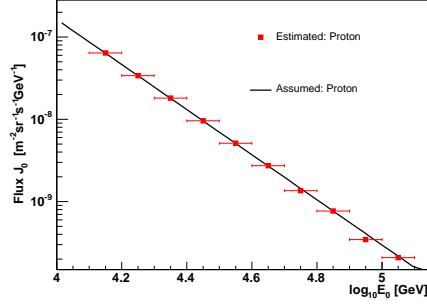


Figure 6: Comparison between the assumed and estimated spectrum of protons. The estimated spectrum is consistent with the assumed one in the energy region from 20 TeV to 200 TeV. Only statistic error is taken into account here.

the experiment data is our future plan. With the analog read-out data, the energy could be extrapolated to the 'knee' region.

References

- [1] J.R Hörandel, Astropart. Phys. 21(2004) 241-265
- [2] M. Amenomori, Phys Rev.D 62, 112002
- [3] T.Antoni, W.D.Apel, Astropart. Phys. 24(2005) 1-25
- [4] A.V. Apanasenko, Phys.Rev.D 62(2000) 112001
- [5] R.Asakimori, et al., APJ 502 (1998) 278
- [6] G.Aielli, Nucl. Phys.B (Proc. Suppl.) 166 (2007) 96-102
- [7] X.W Xu, J. Phys.G 29(2003)719-735
- [8] D. Heck et al., Report FZKA 6019, Forschungszentrum Karlsruhe, 1998.
- [9] J.R Hörandel, Astropart. Phys. 19(2003) 193