Cosmic Rays Composition Determination Method

A.D. Supanitsky and G. Medina-Tanco

Departamento de Física de Altas Energías, Instituto de Ciencias Nucleares, Universidad Nacional Autónoma de México, A. P. 70-543, 04510, México, D. F., México.

Abstract. In this work we present a multiparametric technique that attempts to tackle simultaneously the problems of composition determination and hadronic interaction uncertainty. This technique allows to test the compatibility between hadronic interaction models with real data. When the data results compatible with the hadronic interaction models considered, it also allows to determine the composition of a binary mixture of proton and iron nuclei very precisely and in a way that is independent of the assumed hadronic interaction models.

Keywords: Cosmic Rays, Chemical Composition
PACS: 96.50.S,96.50.sd,13.85.Tp

INTRODUCTION

The hadronic interactions at the highest energies are unknown. There are models that extrapolate the low energy data to higher energies in order to describe such interactions. The determination of the composition of the high energy cosmic rays is performed comparing the experimental data with simulations and, therefore, the main systematic uncertainty of such analyses comes from the hadronic interaction models assumed.

In this work we present a statistical method to test the compatibility of the high energy hadronic interaction models and real data [1]. We consider the hadronic interaction models QGSJET-II [2, 3] and Sibyll 2.1 [4], which are two of the most used for composition analyses. Our method allows to verify whether the experimental data is compatible with the hadronic models under consideration and, if so, to estimate the composition. We center our study in the region of the ankle, where composition carries critical astrophysical information [5], and use the two most important mass sensitive parameters: the number of muons at 600 m from the shower axis, $N_\mu(600)$, and the depth of the shower maximum, $X_{\max}$.

COMPOSITION TECHNIQUE

Let us consider two possible types of primaries, $A = a,b$, and samples of size $N = N_a + N_b$, where N_a and N_b are the number of events corresponding to type a and b, respectively. From each event of an individual sample it is possible to extract several observable parameters sensitive to the primary mass. Therefore, for a given mass sensitive parameter q we define,

$$\xi_q \equiv \frac{1}{N} \sum_{i=1}^{N} P_a(q_i) = \frac{1}{N} \left[\sum_{i=1}^{N_a} P_a(q_i^a) + \sum_{i=1}^{N_b} P_a(q_i^b) \right], \tag{1}$$

where q_i^A are N_A random variables distributed as $f_A(q)$ and $P_a(q) = f_a(q)/(f_a(q) + f_b(q))$, which is the probability that an event is of type a for a given q, assuming no prior knowledge of the primary type. The new statistic, so defined, is an estimator of the abundance of the primary of type a. Note that we restrict our analysis to the case in which the cosmic rays are the superposition of two components.

It can be shown that $\langle \xi_q \rangle (c_a) = m\ c_a + d$, where $c_a = N_a/N$ is the composition or the abundance of the primary a, $m = \int dq P_a(q)(f_a(q) - f_b(q))$, and $d = \int dq P_a(q) f_b(q)$. Therefore, $\langle \xi_q \rangle$ depends linearly on the composition.

Another important property of ξ_q is that the variance is also proportional to c_a and inversely proportional to the sample size: $Var[\xi_q](c_a) = [c_a \left(\sigma_a^2[P_a(q)] - \sigma_b^2[P_a(q)]\right) + \sigma_b^2[P_a(q)]]/N$, where $\sigma_A^2[P_a(q)] = \int dq P_a^2(q) f_A(q) - (\int dq P_a(q) f_A(q))^2$. Note that ξ_q is the sum of N random variables, therefore, for large enough values of N, it is distributed as a Gaussian variable.

The two parameters, $N_\mu(600)$ and X_{max}, will be available for the enhancements of the Pierre Auger Observatory AMIGA [6] and HEAT [7] as well as for Telescope Array and its low energy extensions [8]. However, we center here in the enhancements of the Pierre Auger Observatory. The development of the technique is performed from simulations. The simulation of the air showers is done by using the AIRES package version 2.8.2 [9] and the simulation of the detectors of AMIGA and HEAT is done following Ref. [10]. A power law energy spectrum, of spectral index $\gamma = -2.7$, is considered and all the analyses are performed for a primary energy of $E = 10^{18}$ eV assuming a 25% Gaussian uncertainty on its determination.

In order to calculate ξ_q (see Eq. (1)), we need the distribution functions, $f_A(q)$, of the different parameters sensitive to the primary mass considered, including the effects of the detectors and reconstruction methods. For that purpose, non-parametric method of kernel superposition [11] are used as an estimate of these probability density functions obtained from the simulated data (see Ref. [1] for details of the calculations).

Samples of $N = 100$ and $N = 1000$ events are considered, which are the number of hybrid events in the energy interval considered, expected for the 750 m-array in 2 and 20 years of data taking.

As mentioned above, the distribution functions of the variables ξ_{q1} and ξ_{q2} are Gaussian. The mean value and the covariance matrix depend on the proton abundance of the samples and therefore so will the ellipses that enclose regions of a given value of probability. Fig. 1 shows the ellipses corresponding to 68% and 95% probability for the parameters ξ_μ and $\xi_{X_{\text{max}}}$ for our case study: $\theta = 30°$, $N = 100$ and $N = 1000$ events, for samples built using QGSJET-II and Sibyll 2.1. As mentioned, we consider proton and iron primaries and then, c_p, refers to the proton abundance. The evolution of the abundance on the $\xi_\mu - \xi_{X_{\text{max}}}$ plane, and the shape and size of the associated ellipses, allow for a smooth estimation of the composition in a way that is reasonably independent of the assumed hadronic interaction model. Furthermore, given two possible interaction

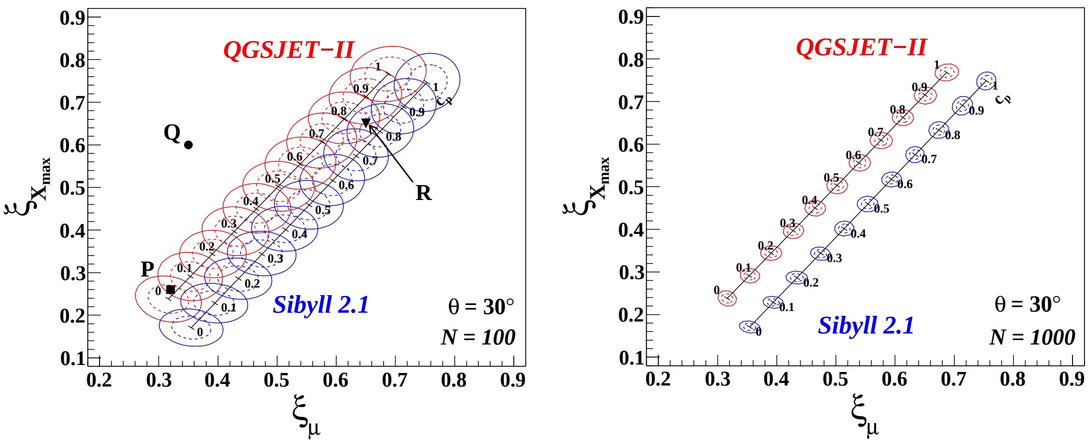

FIGURE 1. Ellipses corresponding to 68% and 95% probability for the Gaussian distributions of the parameters ξ_μ and $\xi_{X_{\max}}$ for $c_p \in [0,1]$, $\theta = 30°$, $N = 100$ (left panel) and $N = 1000$ (right panel) events and for samples corresponding to QGSJET-II and Sibyll 2.1.

models and an observed data set, a figure like the ones depicted in Fig. 1 can be used to assess the compatibility of these models and the experimental data. In order to clarify this point, the diagram in Fig. 2 shows a schematic view of the distributions corresponding to two hadronic interaction models, A and B. Since both ξ_1 and ξ_2 mean values are linearly dependent on c_p, the variation of the latter in the $\xi_1 - \xi_2$ plane is linear, as shown in the diagram. It is possible to approximate "composition isolines" originated by possible hadronic interaction models between A and B by straight lines that link the same values of composition corresponding to the two different models. These curves can be used to infer the real abundance, i.e. given a data sample we can attempt to obtain the real abundance independently of the hadronic interactions model by just finding the composition isoline which passes through the data point.

The statistical uncertainty on the determination of c_p is of the order of the size of the ellipses with dimensions according to the requested confidence level. Therefore, an estimation of the uncertainty over a composition isoline would be extracted from such an ellipse, namely at the crossing with the real abundance straight line.

Having an experimental sample, we can obtain a point on the $\xi_\mu - \xi_{X_{\max}}$ plane by using the density estimates obtained from simulations. As an example, let us consider the left panel of Fig. 1 and the experimental point $P = (0.32, 0.26)$. For this particular example, one can immediately tell form the position of the point with respect to the curves that the data is compatible with QGSJET-II and that the composition is, approximately, in the interval $[0, 0.05]$ at 68% confidence level. If, on the other hand, one considers a point like $R = (0.65, 0.65)$ it is not possible to discriminate between QGSJET-II and Sibyll 2.1; however one can still estimate the composition which is in the interval $[0.75, 0.85]$ at 68% confidence level. A point like $Q = (0.35, 0.6)$, that is located too far from the curves is inconclusive from the point of view of composition or hadronic interaction model, but is a strong indicative of large systematic errors in the detector. Note that if the experimental point results compatible with the hadronic interaction models under consideration, the composition and its uncertainty can be estimated in a more formal

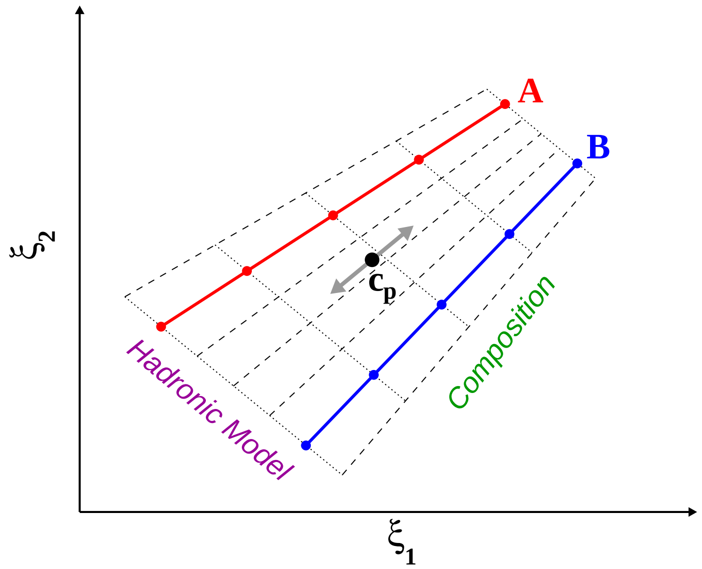

FIGURE 2. Diagram of the distributions of ξ_1 and ξ_2.

way, see Ref. [1] for details.

CONCLUSIONS

In this paper we present a new statistical method to perform composition studies in a two dimensional space. A main advantage of the method is that it minimizes the effects of the present uncertainty associated with the hadronic interaction models, used to simulate cosmic ray showers, on the inferred composition. Furthermore, besides the determination of the composition, it allows an independent verification of the compatibility between real shower data and hadronic interaction models.

REFERENCES

1. A. D. Supanitsky, G. Medina-Tanco, and A. Etchegoyen, *To be published in Astropart. Phys.* (2009), arXiv:0811.0545 [astro-ph].
2. S. Ostapchenko, arXiv:astro-ph/0412591 (2004).
3. S. Ostapchenko, arXiv:hep-ph/0501093 (2005).
4. R. Engel, T. Gaiser, P. Lipari, and T. Stanev, *Proc. 26th ICRC* **1**, 415 (2000).
5. G. Medina-Tanco for the Pierre Auger Collaboration, *Proc. 30th ICRC (Mérida-México)*, #991 (2007).
6. A. Etchegoyen for the Pierre Auger Collaboration, *Proc. 30th ICRC (Mérida-México)*, #1307 (2007).
7. H. Klages for the Pierre Auger Collaboration, *Proc. 30th ICRC (Mérida-México)*, #65 (2007).
8. K. Martens for the Telescope Array Collaboration, *Nucl. Phys. (Proc. Suppl.)* **B52**, 29 (2007).
9. S. Sciutto, AIRES package: http://www.fisica.unlp.edu.ar/auger/aires.
10. A. D. Supanitsky et al., *Astropart. Phys.* **29**, 461-470 (2008).
11. B. Silvermann, *Density Estimation for Statistics and Data Analysis*, ed. Chapman & Hall, New York, 1986.