

Anisotropy in the mass composition from the Telescope Array Surface Detector data

Grigory Rubtsov^a, Pierre Sokolsky^b, Sergey Troitsky^a, Yana Zhezher^{*a} for the Telescope Array collaboration[†]

^a*Institute for Nuclear Research of the Russian Academy of Sciences, 60th October Anniversary st. 7a, Moscow 117312, Russia*

^b*High Energy Astrophysics Institute and Department of Physics and Astronomy, University of Utah, Salt Lake City, Utah 84112, USA*

E-mail: zhezher.yana@physics.msu.ru

The mass composition of the Telescope Array surface detector data is analyzed with the use of the Boosted Decision Trees (BDT) technique, which assigns one parameter ξ for each event, which is then consequently used for the determination of the average atomic mass of the primary particle. Itself, ξ parameter may be distributed unevenly across the sky and indicate the mass composition anisotropy of the ultra-high-energy cosmic rays. The sensitivity of the BDT-based mass composition technique to the possible anisotropies is tested with the use of the ξ distribution with the use of the proton and iron Telescope Array surface detector Monte-Carlo sets.

*36th International Cosmic Ray Conference -ICRC2019-
July 24th - August 1st, 2019
Madison, WI, U.S.A.*

*Speaker.

†for collaboration list see PoS(ICRC2019)1177

1. Introduction

Ultra-high-energy cosmic rays (UHECRs) pose a number of long-standing questions despite a century of experimental observations along with theoretical developments of underlying physical mechanisms. The origin and acceleration mechanisms of UHECRs are the most tantalizing issues, which are still unknown at all energies.

The mass composition of the UHECRs at Earth is one of the measurable quantities directly connected to the cosmic-ray acceleration mechanisms in the source and source population is related to the propagation of the UHECR. Mass composition is usually derived from the observed extensive air showers (EAS) via the measurement of the longitudinal shape of EAS from fluorescence telescopes. In this case, by reconstructing the depth of the shower maximum X_{\max} , one may directly relate this observable to the mass composition of primary particles [1].

An alternative method was proposed by the Telescope Array collaboration from the analysis of solely the surface detector data, which has an important advantage of much higher statistics compared to the fluorescence measurements [2]. The former method is based on the multivariate analysis technique, namely the Boosted Decision Trees (BDT) which employs the set of 16 composition-sensitive observables reconstructed from the data. These observables are transformed into a single variable ξ which then determines the average atomic mass of a primary particle $\langle \ln A \rangle$.

The methods described above use a large number of observed extensive air showers to derive the average $\langle \ln A \rangle$ as a function of energy, and due to large fluctuations in signals registered from EAS, the current level of experimental analysis techniques doesn't allow to determine the mass of a primary particle separately for each observed shower. Yet, the study of mass distribution over the observed part of the sky is an important step forward in the study of UHECR sources and acceleration mechanisms.

In the present work, the ξ parameter derived from the BDT-based analysis is used as an UHECR mass indicator to study the mass composition anisotropy in the cosmic rays observed by the Telescope Array surface detector facility. The spatial distribution of the ξ parameter in the proton and iron Monte-Carlo data sets is used to determine the sensitivity of TA to the composition anisotropies, and thus the possibility to observe the composition anisotropy similar to the "hotspot" observed by the Telescope Array is investigated.

2. Telescope Array experiment

The Telescope Array (TA) experiment is the largest UHECR experiment in the Northern hemisphere, located near Delta, Utah, USA [3]. TA is designed to register the extensive air showers caused by the UHE cosmic rays entering the atmosphere. The experiment operates in hybrid mode and performs simultaneous measurements of the particle density and timing at the ground level with the surface detector array (SD) [4] and the fluorescence light with 38 fluorescence telescopes grouped into three fluorescence detector stations [5]. The SD array is comprised of 507 plastic scintillator detectors arranged on a square grid with 1.2 km spacing covering an area of approximately 700 km². Each detector is composed of two layers of 1.2 cm thick extruded scintillator of the 3 m² effective area.

3. Monte-Carlo simulations

For the Monte-Carlo simulations, CORSIKA software package [6] is used together with the QGSJETII-03 model for high-energy hadronic interactions [7], FLUKA [8] for low energy hadronic interaction and EGS4 [9] for electromagnetic processes. The detector response is simulated with the use of the GEANT4 package [10]. Real-time array status and detector calibration information for 9 years of observations are used for each simulated event [11]. Two separate Monte-Carlo data sets, for proton and iron primaries, are simulated and stored in the same data format as the SD data in the energy range $10^{17.5} - 10^{20.5}$ eV.

4. UHECR mass-composition analysis

The depth of shower maximum, X_{\max} , is the EAS observable, known to be the most sensitive to the composition of the UHECR primary particles. The measurement of X_{\max} , however, has an important drawback: it requires the EAS observations with fluorescence telescopes whose duty cycle is much smaller to that of surface detectors, since they operate only on clear moonless nights. On the contrary, there is no single observable, derived from the surface detector data, as sensitive to the primary particles composition, as X_{\max} .

The method, proposed in [2] is based on the use of one of the multivariate analysis techniques, the so-called Boosted Decision Trees [12, 13] which allows to address the advantages of SD data. The BDT classifier is trained with two Monte-Carlo sets, proton MC data set as a background and iron MC data set as signal events. It assigns the ξ value for each event, where $\xi \in [-1 : 1]$: 1 – pure signal event (Fe), -1 – pure background event (p).

The classifier is built upon 16 the following composition-sensitive variables reconstructed from the TA SD data:

1. Linsley front curvature parameter [14].
2. Area-over-peak (AoP) of the signal at 1200 m.
3. AoP slope parameter [15].
4. Number of detectors hit.
5. Number of detectors excluded from the fit of the shower front by the reconstruction procedure [16].
6. $\chi^2/d.o.f.$ of the joint geometry and LDF fit.
7. S_b parameter for $b = 3$ [17].
8. S_b parameter for $b = 4.5$.
9. The sum of the signals of all the detectors of the event.
10. Asymmetry of the signal at the upper and lower layers of detectors.
11. Total number of peaks within all FADC (flash analog-to-digital converter) traces.

12. Number of peaks for the detector with the largest signal.
13. Number of peaks present in the upper layer and not in the lower.
14. Number of peaks present in the lower layer and not in the upper.
15. Signal at 800 m from the shower core, S_{800} .
16. Zenith angle.

5. Mass composition anisotropy analysis framework

The present analysis is dedicated to the test of sensitivity of the BDT-based classification of events to the possible anisotropies in the distribution of the mass composition. For that purpose, the following sets are created from the available ξ distributions for the proton and iron-induced events:

1. Isotropic set with data composition, where the fractions of proton and iron-induced events are taken as in [2].
2. “Hotspot” composition set, where events with energies $\log E < 19.2$ are isotropically distributed according to the data composition, while on energies $\log E > 19.2$ the proton “hotspot” [18] is created and iron events are distributed isotropically.
3. The same as previous, but with iron “hotspot”.

Two energy bands are considered for the analysis: the full energy range $\log E > 18.0$ and the high-energy “hotspot” range $\log E > 19.2$. Derived ξ distributions are compared with the number of one- and two-dimensional tests independently in the described energy bands.

6. Results and discussion

6.1 One-dimensional tests

Boosted Decision Trees are trained with the number of composition-sensitive observables, one of which is the zenith angle. Thus ξ has a dependence on θ , and its comparison in declination bands is not legitimate. As a first one-dimensional test, mean ξ distributions are compared in right ascension bands of 20° for the MC sets described above.

The results are derived for χ^2 of the comparison of ξ distributions for the isotropic set with data composition with “hotspot” composition sets with iron (1) and proton (2) “hotspots” at energies $\log E > 19.2$:

$$\chi_{1,r.a.}^2/d.o.f. = 0.89, p_1 = 0.56,$$

$$\chi_{2,r.a.}^2/d.o.f. = 0.99, p_2 = 0.46,$$

where the subscript numbers correspond to the notions in the text. Number of degrees of freedom for all sets is $d.o.f. = 18$. The corresponding distributions are shown in Figure 1 and 2.

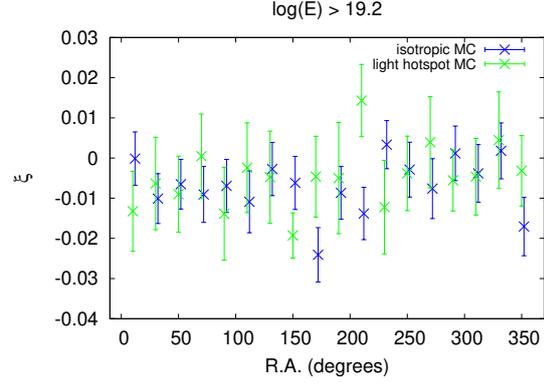
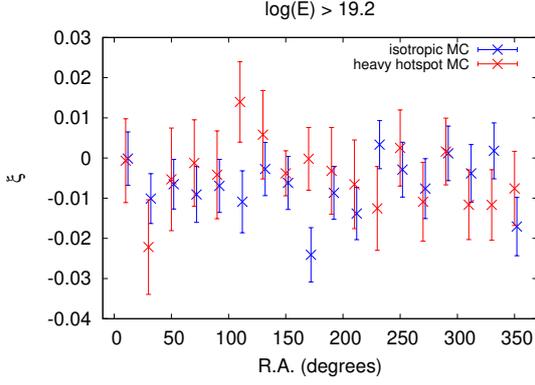


Figure 1: Average ξ distribution in the right ascension bands, the isotropic set with data composition (blue) compared with the set with data composition and the iron “hotspot” at energies $\log E > 19.2$ (red). **Figure 2:** Average ξ distribution in the right ascension bands, the isotropic set with data composition (blue) compared with the set with data composition and the proton “hotspot” at energies $\log E > 19.2$ (green).

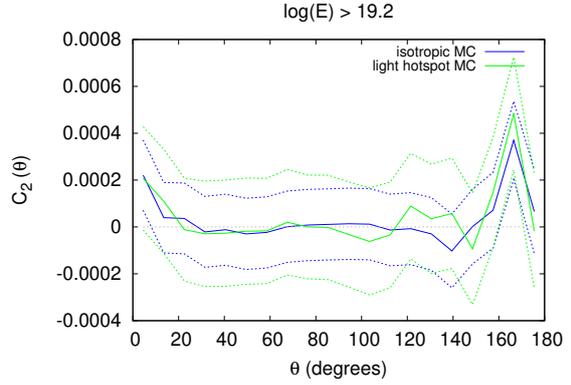
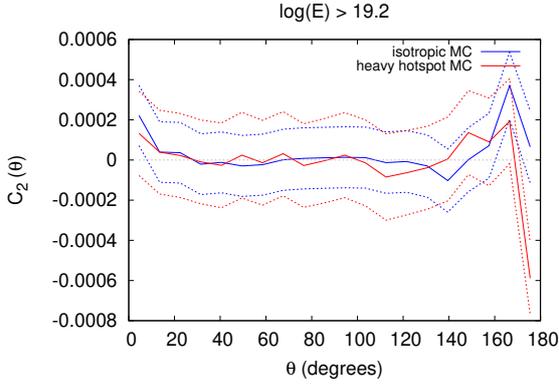


Figure 3: Two-point correlation functions for the isotropic set with data composition (blue) compared with the set with data composition and the iron “hotspot” at energies $\log E > 19.2$ (red). **Figure 4:** Two-point correlation functions for the isotropic set with data composition (blue) compared with the set with data composition and the proton “hotspot” at energies $\log E > 19.2$ (green).

6.2 Two-dimensional tests

The first straightforward two-dimensional test, which may take into account large-scale features of the ξ distribution across the part of the sky observed by the TA is done with the use of pixelized skymaps of ξ distributions. For each of the MC sets and energy bands, the HEALPIX pixelization is used [19] with the resolution parameter $n_{\text{side}} = 8$, which assures sufficient event statistics in each pixel. ξ distribution is averaged in each pixel, and the output full-sky maps are used in the HEALPIX storage format as one-dimensional arrays to calculate the χ^2 of the derived distributions.

For the two cases discussed in the previous section, the χ^2 values are derived as follows:

$$\chi_{1,HEALPIX}^2/d.o.f. = 1.63, p_1 = 6.3 \times 10^{-7},$$

$$\chi_{2,HEALPIX}^2/d.o.f. = 1.47, p_2 = 6.9 \times 10^{-5},$$

where only pixels in the TA field of view with $\delta > -10^\circ$ are used in the calculation. The derived results show 3.8σ difference between two maps for “light” hotspot case and 4.8σ difference for the “heavy” hotspot.

Finally, the Monte-Carlo data sets discussed above were compared with the use of two-point correlation functions [20, 21], which also takes into account large-scale distributions of the ξ parameter over the sky. Two-point correlation functions were calculated with the use of the HEALPIX pixelization with $n_{\text{side}} = 8$, introducing the optimization methods for the calculation of correlation functions in the pixelized maps [22]. The results are shown in Figures 3 and 4 for the same three cases as previous at the highest energies $\log E > 19.2$.

The obtained results indicate the potential capability to distinguish the mass composition anisotropy in the experimental data by using the ξ parameter distribution, if it is present in the events of the Telescope array surface detector. As for the future work, higher-order correlation functions are to be calculated to represent the ξ spatial distribution in precision. The present level of classification doesn’t allow to introduce the intermediate-mass primary nuclei in the mass composition anisotropy studies, such as helium and nitrogen. The latter may become possible with the development of machine learning techniques, which will separate primary particles with much better accuracy. On a long term scale, the introduced approach may become one of the crucial steps in the identification of the UHECR sources and acceleration mechanisms.

Acknowledgments

The development and application of the multivariate analysis method is supported by the Russian Science Foundation grant No. 17-72-20291 (INR).

References

- [1] T. K. Gaisser *et al.*, Phys. Rev. D **47**, 1919 (1993).
- [2] R. U. Abbasi *et al.* [Telescope Array Collaboration], Phys. Rev. D **99**, no. 2, 022002 (2019) [arXiv:1808.03680 [astro-ph.HE]].
- [3] H. Tokuno *et al.* [Telescope Array Collaboration], J. Phys. Conf. Ser. **293**, 012035 (2011).
- [4] T. Abu-Zayyad *et al.* [Telescope Array Collaboration], Nucl. Instrum. Meth. A **689**, 87 (2013) [arXiv:1201.4964 [astro-ph.IM]].
- [5] H. Tokuno *et al.* [Telescope Array Collaboration], Nucl. Instrum. Meth. A **676**, 54 (2012) [arXiv:1201.0002 [astro-ph.IM]].
- [6] D. Heck *et al.*, Report FZKA-6019 (1998), Forschungszentrum Karlsruhe.
- [7] S. Ostapchenko, Nucl. Phys. Proc. Suppl. **151**, 143 (2006) [hep-ph/0412332].
- [8] A. Ferrari, P. R. Sala, A. Fasso and J. Ranft, CERN-2005-010, SLAC-R-773, INFN-TC-05-11.
- [9] W. R. Nelson, H. Hirayama, D.W.O. Rogers, SLAC-0265 (permanently updated since 1985).
- [10] S. Agostinelli *et al.* [GEANT4 Collaboration], Nucl. Instrum. Meth. A **506**, 250 (2003).
- [11] T. Abu-Zayyad *et al.* [Telescope Array Collaboration], arXiv:1403.0644 [astro-ph.IM].

- [12] L. Breiman *et al.*, “Classification and Regression Trees” Wadsworth International Group (1984).
- [13] Y. Freund, R. E. Schapire, *Journal of JSAI*, **14(5)** (1999) 771-780.
- [14] M. Teshima *et al.*, *J. Phys. G* **12**, 1097 (1986).
- [15] J. Abraham *et al.* [Pierre Auger Collaboration], *Phys. Rev. Lett.* **100**, 211101 (2008) [arXiv:0712.1909 [astro-ph]].
- [16] T. Abu-Zayyad *et al.* [Telescope Array Collaboration], *ApJL* **768** (2013) L1.
- [17] G. Ros *et al.*, *Astropart. Phys.* **35**, 140 (2011) [arXiv:1104.3399 [astro-ph.HE]].
- [18] R. U. Abbasi *et al.* [Telescope Array Collaboration], *Astrophys. J.* **790**, L21 (2014) [arXiv:1404.5890 [astro-ph.HE]].
- [19] K. M. Gorski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke and M. Bartelman, *Astrophys. J.* **622**, 759 (2005) [astro-ph/0409513].
- [20] H. Totsuji, T. Kihara, *PASJ*, **21**, 221 (1969)
- [21] P. J. E. Peebles, *ApJ*, **185**, 413 (1973)
- [22] H. K. Eriksen, P. B. Lilje, A. J. Banday and K. M. Gorski, *Astrophys. J. Suppl.* **151**, 1 (2004) [astro-ph/0310831].