# Statistical methods for cosmic ray composition analysis at the Telescope Array Observatory

**Grigory I. Rubtsov, Sergey V. Troitsky for the Telescope Array collaboration**

Institute for Nuclear Research of the Russian Academy of Sciences, Moscow 117312, Russia

E-mail: `grisha@ms2.inr.ac.ru`

**Abstract.** The Telescope Array surface detector stations record the temporal development of the signal from the extensive air shower front which carries information about the type of the primary particle. We develop the method to study the primary mass composition of the ultra-high-energy cosmic rays based on multivariate analysis (MVA). We report the preliminary mass composition results based on the Telescope array 5 years data.

## 1. Introduction

The Telescope Array (TA) experiment [1, 2] is an ultra-high-energy cosmic ray (UHECR) detector operating in Millard County, Utah, USA since May 2008. The TA is a hybrid experiment with a surface detector array of 507 plastic scintillator stations with 1.2 km spacing covering $700$ km$^2$ area overlooked by 38 fluorescence telescopes located at the three sites. This *Talk* is dedicated to the study of the primary mass composition of UHECRs based on the time-resolved data of the TA surface detector (SD).

The primary composition of the UHECRs is one of the most intriguing questions. It is a key component in the studies of the potential sources of the cosmic rays and for predicting the flux of secondary photons and neutrino [3]. Moreover, one needs to know the primary composition in order to probe the interaction cross-section at the highest energies [4, 5] and to perform the precision tests of the Lorentz-invariance [6, 7]. The most accurate composition measurements up to date are based on the observation of the shower maximum $X_{\mathrm{MAX}}$ with the fluorescence telescopes (FD) working in stereo or in hybrid mode with the surface detectors (SD). A couple of alternative techniques are developed based on the data of the surface and muon detectors. The list of the available methods is given in Table 1.

It is well established that the photons and neutrino produce a minor fraction of the extensive air showers (EAS) in the atmosphere and therefore the most of the EAS are produced by hadrons [3]. Nevertheless, there is no agreement on the interpretation of the hadronic composition results between the experiments. The data of HiRes and TA are consistent with the pure protonic composition [8, 9], while the Pierre Auger observatory (PAO) indicate the presence of heavier nuclei at the energies higher than 10 EeV [10, 11]. The composition puzzle may be resolved by increasing the statistics of the Telescope Array experiment and increasing the control over systematic uncertainties in both TA and PAO.

In the present *Talk* the technique of the composition analysis based on the data of the TA surface detector is developed. The surface detector is capable of measuring chemical composition

**Table 1.** The techniques used for the composition measurements. $X_{\mathrm{MAX}}$ – depth of the shower maximum, $X_{\mathrm{MAX}}^{\mu}$ – muon production depth, $\rho_{\mu}$ – muon density, risetime – time interval between 10% and 50% of the total integral signal.

| Experiment | detector | observable | ref. |
|---|---|---|---|
| HiRes | FD stereo | $X_{\mathrm{MAX}}$ | [8] |
| Pierre Auger | FD+SD (hybrid) | $X_{\mathrm{MAX}}$ | [10] |
| Telescope Array | FD stereo | $X_{\mathrm{MAX}}$ | |
| Telescope Array | FD hybrid | $X_{\mathrm{MAX}}$ | [9] |
| Yakutsk | SD + muon detector | $\rho_{\mu}$ | [12] |
| Pierre Auger | SD | $X_{\mathrm{MAX}}^{\mu}$ | [11] |
| Pierre Auger | SD | risetime asymmetry | [11] |

independently of the fluorescence telescopes. It also provides better statistics at the highest energy due to the 95% duty cycle. The multivariate analysis (MVA) method is implemented using the two observables related to the area-over-peak of the signals at the surface detector stations. The data registered for the five years of observations are confronted with the proton and iron Monte-Carlo leading to the preliminary results on the mass-composition.

## 2. Method

Each of 507 TA surface detectors records time-resolved signals with a time step of 20 ns. For each detector we define Area-over-Peak ($AoP$) as the ratio of the integral of the FADC trace to it's peak value. The $AoP$ therefore is measured in time units (ns). The muons arrive close in time to the shower front making $AoP$ smaller, while electron-photon cascade shifts $AoP$ towards larger values. The $AoP$ observable was previously introduced by Pierre Auger collaboration for neutrino identification [13]. We fit $AoP$ as a linear function of the core distance:

$$AoP(r) = \alpha - \beta(r - 1200\,\mathrm{m})\,,$$

where $\alpha$ has a meaning of $AoP$ value at 1200 meters and $\beta$ is the $AoP$ slope.

Following [14, 15] we define percentile ranks of $\alpha$ and $\beta$ parameters relative to the proton showers $\mathcal{C}_{\alpha}, \mathcal{C}_{\beta}$:

$$\mathcal{C}_{\alpha}^{i} = \int\limits_{-\infty}^{\alpha^{i}} f_{MC,p}^{i}(\alpha)d\alpha\,,$$

$$\mathcal{C}_{\beta}^{i} = \int\limits_{-\infty}^{\beta^{i}} f_{MC,p}^{i}(\beta)d\beta\,,$$

where $f_{MC,p}^{i}(\alpha)$ is an $\alpha$ distribution function for proton QGSJETII-03 Monte-Carlo events compatible by zenith angle with the real event "i".

The composition analysis is based on the proton-iron classification procedure using the method of boosted decision trees (BDT). The *TMVA* package [16] for ROOT is used as an implementation of the method. The decision forest is constructed using the three observables: $C_{\alpha}$, $C_{\beta}$ and zenith angle. We construct independent BDT classifier for each energy bin spaced by 0.2 in $\log_{10}(E)$. The BDT is trained using the proton Monte-Carlo set as a background and the iron Monte-Carlo as a signal. The result of the BDT classifier is a single parameter $\xi^{i}$ for

each event "i" which has a limited range by definition $-1 < \xi^i < 1$. The $\xi$-parameter is finally used for traditional one-parametric composition analysis.

We note that the reduction of the composition problem to proton-iron classification is based on an approximation the all relevant shower properties scale linearly with the logarithm of the atomic mass - $\log A$. The latter is supported by the results of simulations with nitrogen, silicon and helium primaries.
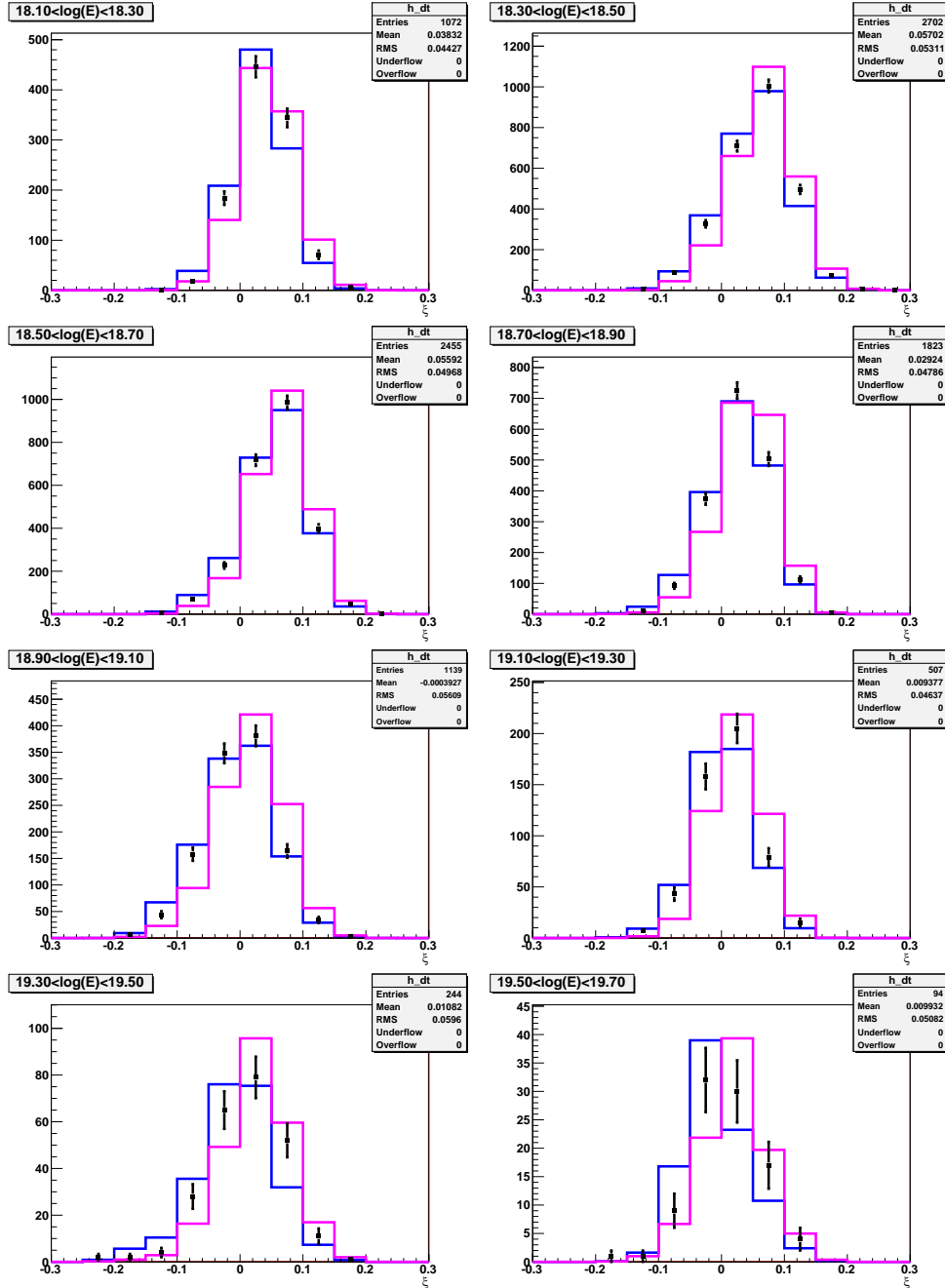


**Figure 1.** Distributions of the $\xi$ parameter for data and QGSJET-II Monte-Carlo (blue - protons, magenta - iron). /PRELIMINARY/

### 3. Data set, Monte-Carlo set and results

We use Telescope Array surface detector data set covering five years of observation from 2008-05-11 to 2013-07-13. Surface detector has been collecting data for more than 95% of time during that period [17]. The following cuts are applied to both data and MC events:

 (i) Quality cuts used for spectral analysis [17];

 (ii) Zenith angle cut: $0° < \theta < 45°$;

(iii) The number of detectors triggered is 7 or more;

(iv) The reconstructed energy is greater than $10^{18}$ eV.

The dataset contains 10242 events after the cuts. For $AoP(r)$ fit we require that the detector is not saturated and has a core distance larger than 600 m.

We use CORSIKA [18] with QGSJET II-03 [19] model to generate hadronic showers induced by primary protons and iron. The showers are simulated with thinning and the dethinning procedure is adopted [20] to simulate realistic shower fluctuations. The detector response is accounted for by using look-up tables generated by GEANT4 [21] simulations. Real-time array status and detector calibration information are used for each Monte Carlo (MC) simulated event. The Monte-Carlo events are produced in the same format as real events and the analysis procedures are applied in uniformly to both [22]. Each Monte-Carlo set is split into two halves - first is used for training the decision trees and the second for the final result.
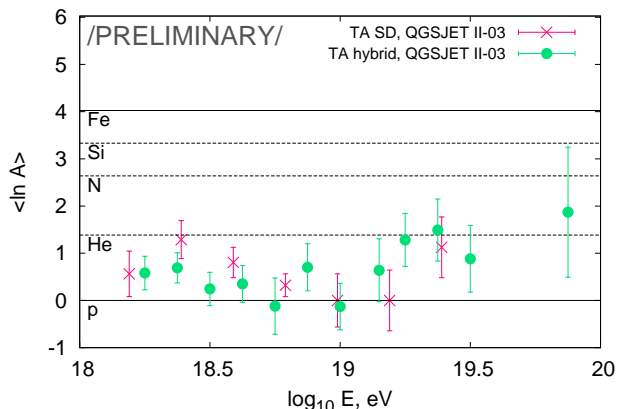


**Figure 2.** The preliminary result on the primary atomic mass based on the TA SD data compared to the TA hybrid result [9]

The histograms of $\xi$-parameter for data and Monte-Carlo are and QGSJET-II-03 and SIBYLL Monte-Carlo are shown in Figures 1. The data histograms are then compared to the mixture of proton and iron Monte-Carlo with the Kolmogorov-Smirnov probability. The preliminary result on the primary atomic mass from TA SD data is shown in Figure 2 in comparison with TA hybrid result. One may see the agreement of the results within the statistical errors. The advancemen of the method is required to reach sensitivity needed to resolve the composition puzzle. The systematic uncertainties and dependence on the hadronic model will be explained elsewhere.

### 4. Conclusion

We designed the new method of composition study based on the data of the array of scintillators. The preliminary composition based on the 5 years of TA SD is presented. Further development of the method involving more observable parameters is required to reach the sensitivity to discriminate between mono and mixed composition based on the TA surface detector data alone.

## Acknowledgments

## References

[1] Abu-Zayyad T et al. for the Telescope Array Collaboration 2012 *Nucl. Instrum. Meth. A* **689** 87
[2] Tokuno H et al. for the Telescope Array Collaboration 2012 *Nucl. Instrum. Meth. A* **676** 54
[3] Alvarez-Muniz J, Risse M, Rubtsov G and Stokes B 2013 *EPJ Web Conf.* **53** 01009
[4] Belov K et al. for the HiRes collaboration 2006 *Nucl. Phys. Proc. Suppl.* **151** 197
[5] Abreu P et al. for the Pierre Auger Collaboration 2012 *Phys.Rev.Lett.* **109** 062002
[6] Galaverni M and Sigl G 2008 *Phys. Rev. Lett.* **100** 021102
[7] Rubtsov G, Satunin P and Sibiryakov S 2014 *Phys. Rev. D* **89** 123011
[8] Abbasi R et al. for the HiRes collaboration 2010 *Phys.Rev.Lett.* **104** 161101
[9] Abbasi R et al. for the Telescope Array collaboration 2014 *Astropart.Phys.* **64** 49
[10] Abraham J et al. for the Pierre Auger collaboration 2010 *Phys.Rev.Lett.* **104** 091101
[11] Abreu P et al. for the Pierre Auger Collaboration 2011 *Studies of Cosmic Ray Composition and Hadronic Interaction models* (Proceeding of ICRC'11, Beijing) *preprint* arXiv:1107.4804
[12] Glushkov A, Makarov I, Pravdin M, Sleptsov I, Gorbunov D, Rubtsov G and Troitsky S 2008 *JETP Lett.* **87** 190
[13] Abreu P et al. for the Pierre Auger Collaboration 2013 *JCAP* **1305** 009
[14] Gorbunov D, Rubtsov G, Troitsky S 2007 *Astropart. Phys.* **28** 28
[15] Abu-Zayyad T et al. for the Telescope Array Collaboration 2013 *Phys. Rev. D* **88** 112005
[16] Hoecker A, Speckmayer P, Stelzer J, Therhaag J, von Toerne E and Voss H 2007 PoS ACAT 040,2007 *preprint* arXiv:physics/0703039
[17] Abu-Zayyad T et al. for the Telescope Array Collaboration 2013 *Astrophys. J.* **768** L1
[18] Heck D et al., 1998, *Forschungszentrum Karlsruhe* Report FZKA-6019.
[19] Ostapchenko S 2006 *Nucl. Phys. Proc. Suppl.* **151** 143
[20] Stokes B T et al. 2012 *Astropart. Phys.* **35**, 759
[21] Agostinelli S et al. for the GEANT4 Collaboration 2003 *Nucl. Instrum. Meth. A* **506** 250
[22] Abu-Zayyad T et al. for the Telescope Array Collaboration 2014 CORSIKA Simulation of the Telescope Array Surface Detector *Preprint* arXiv:1403.0644