

Mass composition of ultrahigh-energy cosmic rays with the Telescope Array Surface Detector data

R. U. Abbasi,¹ M. Abe,² T. Abu-Zayyad,¹ M. Allen,¹ R. Azuma,³ E. Barcikowski,¹ J. W. Belz,¹ D. R. Bergman,¹ S. A. Blake,¹ R. Cady,¹ B. G. Cheon,⁴ J. Chiba,⁵ M. Chikawa,⁶ A. di Matteo,⁷ T. Fujii,⁸ K. Fujita,⁹ M. Fukushima,^{8,10} G. Furlich,¹ T. Goto,⁹ W. Hanlon,¹ M. Hayashi,¹¹ Y. Hayashi,⁹ N. Hayashida,¹² K. Hibino,¹² K. Honda,¹³ D. Ikeda,⁸ N. Inoue,² T. Ishii,¹³ R. Ishimori,³ H. Ito,¹⁴ D. Ivanov,¹ H. M. Jeong,¹⁵ S. Jeong,¹⁵ C. C. H. Jui,¹ K. Kadota,¹⁶ F. Kakimoto,³ O. Kalashev,¹⁷ K. Kasahara,¹⁸ H. Kawai,¹⁹ S. Kawakami,⁹ S. Kawana,² K. Kawata,⁸ E. Kido,⁸ H. B. Kim,⁴ J. H. Kim,¹ J. H. Kim,²⁰ S. Kishigami,⁹ S. Kitamura,³ Y. Kitamura,³ V. Kuzmin,^{17,*} M. Kuznetsov,¹⁷ Y. J. Kwon,²¹ K. H. Lee,¹⁵ B. Lubsandorzhev,¹⁷ J. P. Lundquist,¹ K. Machida,¹³ K. Martens,¹⁰ T. Matsuyama,⁹ J. N. Matthews,¹ R. Mayta,⁹ M. Minamino,⁹ K. Mukai,¹³ I. Myers,¹ K. Nagasawa,² S. Nagataki,¹⁴ R. Nakamura,²² T. Nakamura,²³ T. Nonaka,⁸ H. Oda,⁹ S. Ogio,⁹ J. Ogura,³ M. Ohnishi,⁸ H. Ohoka,⁸ T. Okuda,²⁴ Y. Omura,⁹ M. Ono,¹⁴ R. Onogi,⁹ A. Oshima,⁹ S. Ozawa,¹⁸ I. H. Park,¹⁵ M. S. Piskunov,¹⁷ M. S. Pshirkov,^{17,25} J. Remington,¹ D. C. Rodriguez,¹ G. Rubtsov,¹⁷ D. Ryu,²⁰ H. Sagawa,⁸ R. Sahara,⁹ K. Saito,⁸ Y. Saito,²² N. Sakaki,⁸ N. Sakurai,⁹ L. M. Scott,²⁶ T. Seki,²² K. Sekino,⁸ P. D. Shah,¹ F. Shibata,¹³ T. Shibata,⁸ H. Shimodaira,⁸ B. K. Shin,⁹ H. S. Shin,⁸ J. D. Smith,¹ P. Sokolsky,¹ B. T. Stokes,¹ S. R. Stratton,^{1,26} T. A. Stroman,¹ T. Suzawa,² Y. Takagi,⁹ Y. Takahashi,⁹ M. Takamura,⁵ M. Takeda,⁸ R. Takeishi,¹⁵ A. Taketa,²⁷ M. Takita,⁸ Y. Tameda,²⁸ H. Tanaka,⁹ K. Tanaka,²⁹ M. Tanaka,³⁰ S. B. Thomas,¹ G. B. Thomson,¹ P. Tinyakov,^{7,17} I. Tkachev,¹⁷ H. Tokuno,³ T. Tomida,²² S. Troitsky,¹⁷ Y. Tsunesada,⁹ K. Tsutsumi,³ Y. Uchihori,³¹ S. Udo,¹² F. Urban,³² T. Wong,¹ K. Yada,⁸ M. Yamamoto,²² R. Yamane,⁹ H. Yamaoka,³⁰ K. Yamazaki,¹² J. Yang,³³ K. Yashiro,⁵ Y. Yoneda,⁹ S. Yoshida,¹⁹ H. Yoshii,³⁴ Y. Zhezher,^{17,35,†} and Z. Zundel¹

(Telescope Array Collaboration)

¹*High Energy Astrophysics Institute and Department of Physics and Astronomy, University of Utah, Salt Lake City, Utah 84112, USA*

²*The Graduate School of Science and Engineering, Saitama University, Saitama, Saitama 338-8570, Japan*

³*Graduate School of Science and Engineering, Tokyo Institute of Technology, Meguro, Tokyo 152-8550, Japan*

⁴*Department of Physics and The Research Institute of Natural Science, Hanyang University, Seongdong-gu, Seoul 04763, Korea*

⁵*Department of Physics, Tokyo University of Science, Noda, Chiba 278-0022, Japan*

⁶*Department of Physics, Kindai University, Higashi Osaka, Osaka 577-8502, Japan*

⁷*Service de Physique Theorique, Universit Libre de Bruxelles, Brussels 1160, Belgium*

⁸*Institute for Cosmic Ray Research, University of Tokyo, Kashiwa, Chiba 277-8582, Japan*

⁹*Graduate School of Science, Osaka City University, Osaka, Osaka 558-8585, Japan*

¹⁰*Kavli Institute for the Physics and Mathematics of the Universe (WPI), Todai Institutes for Advanced Study, University of Tokyo, Kashiwa, Chiba 277-8583, Japan*

¹¹*Information Engineering Graduate School of Science and Technology, Shinshu University, Nagano 380-0928, Japan*

¹²*Faculty of Engineering, Kanagawa University, Yokohama, Kanagawa 221-8686, Japan*

¹³*Interdisciplinary Graduate School of Medicine and Engineering, University of Yamanashi, Kofu, Yamanashi 400-8510, Japan*

¹⁴*Astrophysical Big Bang Laboratory, RIKEN, Wako, Saitama 351-0198, Japan*

¹⁵*Department of Physics, Sungkyunkwan University, Jang-an-gu, Suwon (16419) 2066, Korea*

¹⁶*Department of Physics, Tokyo City University, Setagaya-ku, Tokyo 158-8557, Japan*

¹⁷*Institute for Nuclear Research of the Russian Academy of Sciences, Moscow 117312, Russia*

¹⁸*Advanced Research Institute for Science and Engineering, Waseda University, Shinjuku-ku, Tokyo 169-8555, Japan*

¹⁹*Department of Physics, Chiba University, Chiba, Chiba 263-8522, Japan*

²⁰*Department of Physics, School of Natural Sciences, Ulsan National Institute of Science and Technology, UNIST-gil, Ulsan 44919, Korea*

²¹*Department of Physics, Yonsei University, Seodaemun-gu, Seoul 03722, Korea*

²²*Academic Assembly School of Science and Technology Institute of Engineering, Shinshu University, Nagano, Nagano 390-8621, Japan*

²³*Faculty of Science, Kochi University, Kochi, Kochi 780-8072, Japan*

- ²⁴*Department of Physical Sciences, Ritsumeikan University, Kusatsu, Shiga 603-8577, Japan*
- ²⁵*Sternberg Astronomical Institute, Moscow M.V. Lomonosov State University, Moscow 119992, Russia*
- ²⁶*Department of Physics and Astronomy, Rutgers University—The State University of New Jersey, Piscataway, New Jersey 08854, USA*
- ²⁷*Earthquake Research Institute, University of Tokyo, Bunkyo-ku, Tokyo 113-0032, Japan*
- ²⁸*Department of Engineering Science, Faculty of Engineering, Osaka Electro-Communication University, Neyagawa-shi, Osaka 572-8530, Japan*
- ²⁹*Graduate School of Information Sciences, Hiroshima City University, Hiroshima, Hiroshima 731-3166, Japan*
- ³⁰*Institute of Particle and Nuclear Studies, KEK, Tsukuba, Ibaraki 300-3256, Japan*
- ³¹*National Institute of Radiological Science, Chiba, Chiba 263-0024, Japan*
- ³²*CEICO, Institute of Physics, Czech Academy of Sciences, Prague 18200, Czech Republic*
- ³³*Department of Physics and Institute for the Early Universe, Ewha Womans University, Seodaemun-gu, Seoul 03760, Korea*
- ³⁴*Department of Physics, Ehime University, Matsuyama, Ehime 790-8577, Japan*
- ³⁵*Faculty of Physics, M.V. Lomonosov Moscow State University, Moscow 119991, Russia*

 (Received 14 August 2018; revised manuscript received 26 December 2018; published 24 January 2019)

The results on ultrahigh-energy cosmic rays (UHECR) mass composition obtained with the Telescope Array surface detector are presented. The analysis employs the Boosted Decision tree (BDT) multivariate analysis built upon 14 observables related to both the properties of the shower front and the lateral distribution function. The multivariate classifier is trained with Monte-Carlo sets of events induced by the primary protons and iron. An average atomic mass of UHECR is presented for energies $10^{18.0}$ – $10^{20.0}$ eV. The average atomic mass of primary particles shows no significant energy dependence and corresponds to $\langle \ln A \rangle = 2.0 \pm 0.1(\text{stat.}) \pm 0.44(\text{syst.})$. The result is compared to the mass composition obtained by the Telescope Array with X_{max} technique along with the results of other experiments. Possible systematic errors of the method are discussed.

DOI: [10.1103/PhysRevD.99.022002](https://doi.org/10.1103/PhysRevD.99.022002)

I. INTRODUCTION

The Telescope Array (TA) experiment is the largest ultrahigh-energy (UHE) cosmic-ray experiment in the Northern hemisphere, located near Delta, Utah, USA [1]. TA is designed to register the extensive air showers (EAS) 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) [2] and the fluorescence light with 38 fluorescence telescopes grouped into three fluorescence detector stations [3]. The SD is an array 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.

There is a continuous progress of the experimental techniques, which started since the discovery of the cosmic rays more than a century ago. Recently, the results of three independent experiments confirmed the cutoff in the

highest energy part [4–6] of the cosmic ray energy spectrum. The latter was predicted in 1966 by Greisen, Zatsepin and Kuzmin [7,8]. Still, the origin of the UHE cosmic rays remains unidentified. The mass composition of the UHE cosmic rays at Earth is one of the measurable quantities directly connected to the cosmic-ray acceleration mechanism in the source and source population as well as it is related to the propagation of the UHECR. Moreover, the mass composition is the main source of uncertainty in the expected cosmogenic photon and neutrino fluxes [9,10]. In the wider scope, one needs the mass composition for precision tests of the Lorentz-invariance [11] and to ensure the safety of the future 100 TeV colliders. The latter is based on the constraints on the black hole production derived from the stability of dense astrophysical objects, such as white dwarfs and neutron stars, which interact with the cosmic rays. Black hole production rate depends on the energy per nucleon and thus on the mass composition of the UHECR [12].

The most established method for the UHECR composition analysis is based on the measurements of the longitudinal shape of the EAS with the fluorescence telescope. This method uses the depth of the shower maximum X_{max} as a composition-sensitive observable [13]. There are UHE composition results available based on X_{max} measured by the three experiments: HiRes, Pierre Auger Observatory and

*Deceased.
†Corresponding author.
zhezher.yana@physics.msu.ru

Telescope Array [14–16]. The two latter results are compatible within the systematic errors in X_{\max} measurement which are of the order of 10–20 g/cm² in the energy range up to 10¹⁹ eV [17].

This paper is dedicated to an alternative approach to measure the mass composition. The method uses solely the data of the surface detector which has an undoubted advantage of the longer than 95% duty cycle [2]. Still, there is no single observable known that has a comparable to X_{\max} sensitivity to the mass composition, although measurements based on the risetime [18,19] have come close. In this paper we use the multivariate Boosted Decision Tree (BDT) [20,21] technique based on a number of composition-sensitive variables obtained during the reconstruction of the SD events. The BDT method has proved itself reliable with a number of successful applications for the astroparticle physics experiments, see, e.g., [22–24].

The general scheme of the analysis is the following. The proton-induced and iron-induced Monte-Carlo events are simulated using the real-time calibration of the Telescope Array. The Monte-Carlo events are stored in the same format as the SD data and are split into three parts used in the following stages. First, a BDT classifier is trained using the first part of the proton-induced Monte-Carlo (MC) events as a background and iron-induced events as signal. Second, the distribution of the classifier output ξ for data is compared to the second part of the proton and iron-induced MC events. The comparison results in the average atomic mass $\langle \ln A \rangle$ of the primary particle as a function of energy. Finally, the third part of the MC is used to estimate the bias of the method and to introduce a correction to $\langle \ln A \rangle$ in order to compensate it.

The paper is organized as follows: in Sec. II data and Monte-Carlo sets are described. Section III is dedicated to multivariate analysis method and its implementation to mass determination. Finally, results and discussion of the systematic uncertainties are provided in Sec. IV.

II. DATASET AND SIMULATIONS

A. Surface detector data

The data of the 9 years of the Telescope Array surface detector operation from May 11, 2008 to May 10, 2017 are used in this paper. Each event is a set of the time-dependent signals (waveforms) from both upper and lower layers of each triggered station. The waveforms are recorded by the 12-bit flash analog-to-digital converters (FADC) with the 50 MHz sampling rate and are converted to MIPs [2] at the calibration stage. The station is marked as saturated at this stage if the saturation effects are significant. In the case of saturated detectors only the signal incidence time is used in the analysis.

B. Event reconstruction and cuts

Surface detector array event reconstruction is done in two steps [6]. At the first step, event geometry is

reconstructed using the time of the arrival of the shower front particles measured by the triggered (>0.3 MIP) counters. Shower front is approximated with empirical functions proposed by Linsley [25] and later modified in AGASA experiment [26]. Secondly, pulse heights in the counters together with the event geometry information are used for determining the normalization of the shower lateral distribution profile S_{800} [27].

In order to determine the Linsley front curvature parameter an additional joint fit of shower front and lateral distribution function (LDF) is performed with 7 free parameters: x_{core} , y_{core} , θ , ϕ , S_{800} , t_0 , a [28]:

$$t_0(r) = t_0 + t_{\text{plane}} + a \times (1 + r/R_L)^{1.5} \text{LDF}(r)^{-0.5}, \quad (1)$$

$$S(r) = S_{800} \times \text{LDF}(r), \quad (2)$$

$$\text{LDF}(r) = f(r)/f(800 \text{ m}), \quad (3)$$

$$f(r) = \left(\frac{r}{R_m}\right)^{-1.2} \left(1 + \frac{r}{R_m}\right)^{-(\eta-1.2)} \left(1 + \frac{r^2}{R_1^2}\right)^{-0.6},$$

$$R_m = 90.0 \text{ m}, \quad R_1 = 1000 \text{ m}, \quad R_L = 30 \text{ m},$$

$$\eta = 3.97 - 1.79(\sec(\theta) - 1),$$

$$r = \sqrt{(x_{\text{core}} - x)^2 + (y_{\text{core}} - y)^2}, \quad (4)$$

where x_{core} , y_{core} , x and y are obtained from the predefined coordinate system of the array centered at the Central Laser Facility (CLF) [29], t_{plane} is the delay of the shower plane and a is the Linsley front curvature parameter. Including the Linsley front curvature, 14 composition-sensitive parameters are estimated for each event, see Appendix for details.

The parameters may be qualitatively split into three groups. The first group of parameters is related to the LDF which is known to be sensitive to X_{\max} . These are the S_b for $b = 3$ and $b = 4.5$ [30], the sum of the signals of all the detectors of the event, the number of the detectors hit and $\chi^2/\text{d.o.f.}$ of the LDF fit.

The second group is related to the shower front which is in turn sensitive to both X_{\max} and the muon content of the shower. The Linsley curvature parameter designates the shower front curvature, while the area-over-peak of the signal, its slope and the number of detectors excluded from the fit correlate with the shower front width.

The latter group indicates the muon content of the shower. Muons cause the single peaks in FADC traces as they propagate rectilinearly and have small dispersion of arrival time. Moreover, muons induce identical signals in the upper and in the lower layers of the detector. Hence, the total number of peaks within all FADC traces, number of peaks in the detector with the largest signal, number of peaks present in the upper layer and not in the lower and vice versa, and also the asymmetry of the signal at the upper

and at the lower layers of the detector are affected by the muonic component of the shower.

The following cuts are used to ensure the quality of reconstruction:

- (1) event includes 7 or more triggered stations;
- (2) zenith angle is below 45° ;
- (3) reconstructed core position inside the array with the distance of at least 1200 m from the edge of the array;
- (4) $\chi^2/\text{d.o.f.}$ does not exceed 4 for both the geometry and the LDF fits;
- (5) $\chi^2/\text{d.o.f.}$ does not exceed 5 for the joint geometry and LDF fit.
- (6) an arrival direction is reconstructed with accuracy less than 5° ;
- (7) fractional uncertainty of the S_{800} is less than 25%.

The same cuts are applied to both the data and the Monte-Carlo sets. The cuts listed above are tighter compared to the standard analysis cuts [6] due to the additional requirement of the curvature parameter reconstruction quality. Namely, 7 triggered stations is required instead of 5 and additional χ^2 condition for the joint fit is included [28].

After the cuts, the SD data set contains 18068 events with energy greater than 10^{18} eV and less than 10^{20} eV.

BDT parameters distribution histograms for energy bin $\log_{10}E = 18.8\text{--}19.0$ are denoted in Fig. 1, proton MC is shown with red lines, iron MC is shown with blue lines and black dots represent the data.

Let us discuss a contribution of individual parameters to overall BDT result. The TMVA package provides a relative importance value for each variable. The importance values are somewhat different in each energy range. Typically, the most discriminating variables are shower front curvature, χ^2 and zenith angle with importance about 8%. The least discriminative variables are number of detectors hit and number of detectors excluded from geometry fit with importance about 3% and 1% correspondingly. The remaining 11 parameters have importance value between 5% and 7%.

C. Simulations

For the Monte-Carlo simulations, CORSIKA software package [31] is used along with the QGSJETII-03 model for high-energy hadronic interactions [32], FLUKA [33,34] for low energy hadronic interaction and EGS4 [35] for electromagnetic processes.

Due to the large number of particles born in an extensive air shower, modern computer resources available make it impractical to track every single one in a simulation. Instead, a *thinning* procedure was proposed [36]. Within thinning, all particles with energies greater than a certain fraction of the primary energy ϵ_{th} are followed in detail, but below the threshold only one particle out of the secondaries produced in a certain interaction is randomly selected. This effective particle is assigned a weight to ensure energy conservation. The thinning level of $\epsilon_{\text{th}} = 10^{-6}$ with an

additional weight limitation according to [37] is used for simulations. The thinning allows to achieve CPU-time efficiency, but at the same time introduces artificial statistical fluctuations [38]. The *dethinning* procedure is developed and implemented [39] in order to restore the statistical properties of the shower. The detector response is simulated by the GEANT4 package [40]. Real-time array status and detector calibration information for 9 years of observations are used for each simulated event [41]. Two separate Monte-Carlo sets, for proton and iron primaries, are simulated and stored in the same data format as the SD data. In the energy range $10^{17.5}\text{--}10^{20.5}$ eV a set of 9800 CORSIKA showers was created. Using these showers, 200 million events were thrown on the detector for each MC set. The procedure of the Monte-Carlo set production for the Telescope Array is described in detail in [42].

For each of the fourteen variables, its data and MC distributions were verified to be in the reasonable agreement. Within errors, all distributions of variables of data events lie between the proton and iron distributions.

III. METHOD

A. BDT classifier

A number of composition-sensitive observables may be extracted from the data, and therefore one may benefit from using the multivariate analysis techniques. In this paper, Boosted Decision Trees (BDT) technique is implemented, available as a part of the ROOT toolkit for multivariate data analysis (TMVA) package [43]. The adaptive boosting (AdaBoost) algorithm is employed [21,44] with the number of trees $N_{\text{Trees}} = 1000$.

The proton and iron Monte-Carlo sets are split into 3 parts with equal statistics. The first part is used to build and train the BDT classifier based on 16 variables, including zenith angle, energy and 14 composition-sensitive parameters listed in Appendix. Proton-induced MC showers are used as a background and iron-induced ones as a signal events. A separate classifier is constructed for each energy bin with the width of $\log_{10}E = 0.2$: last two bins were merged together due to low number of data events. The classifier is applied to the data set as well as to the two remaining parts of the Monte-Carlo sets.

The result of the BDT classifier is a single value ξ for each data and Monte-Carlo event. ξ resides in the range $\xi \in [-1; 1]$, where $\xi = 1$ is a pure signal event, $\xi = -1$ —pure background event. The variable ξ is used in the following one-dimensional analysis. Figure 2 shows ξ parameter distribution histograms for all the energy bins, proton MC is shown with red lines, iron MC is shown with blue lines and black dots represent the data.

B. Estimation of an average atomic mass

Following the two-component approximation, the binned template fitting procedure is applied to p , Fe ,

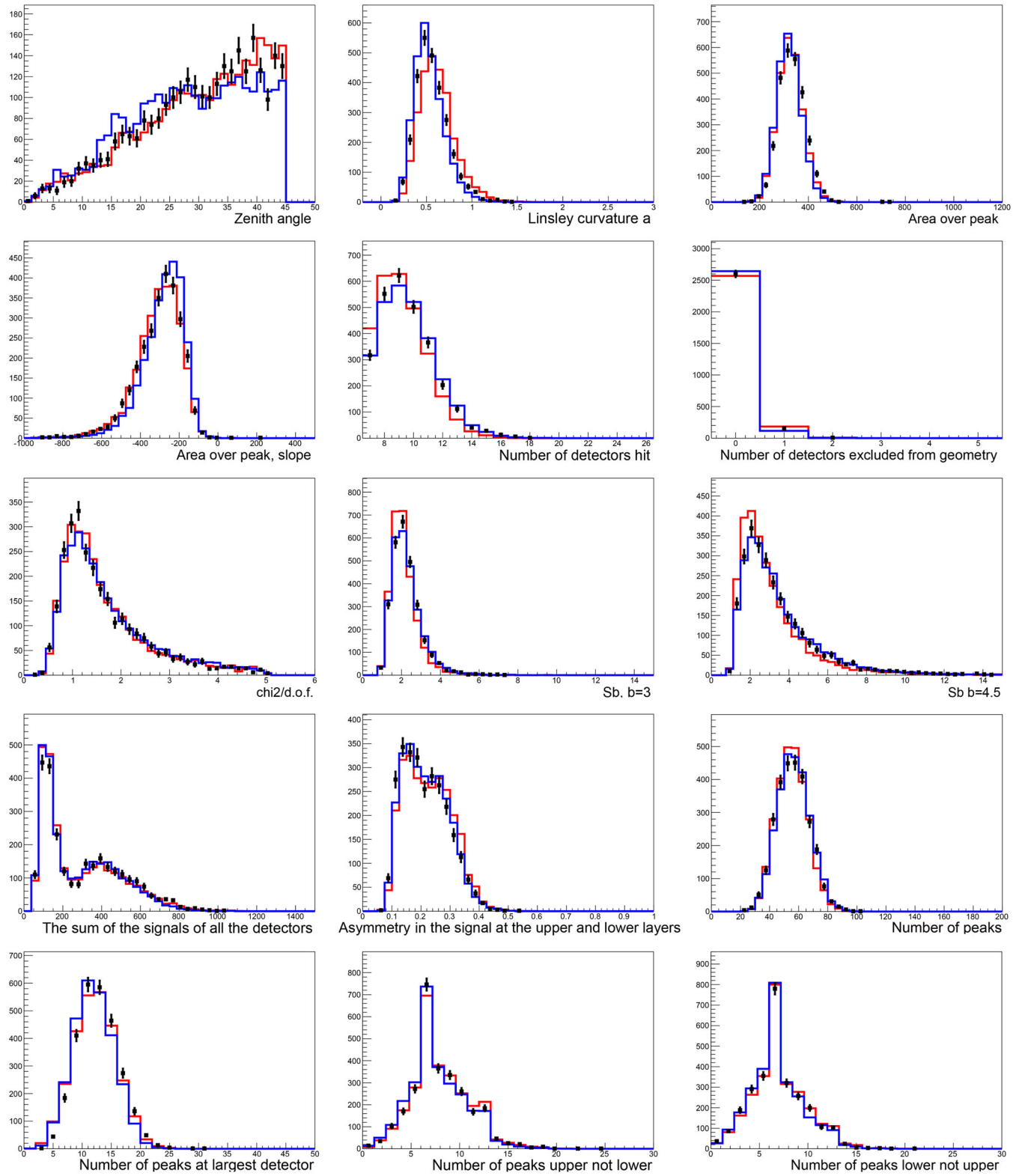


FIG. 1. Distributions of BDT parameters for energy bin $\log_{10}E = 18.8-19.0$. Proton MC is shown with red lines, iron MC is shown with blue lines and black dots represent the data.

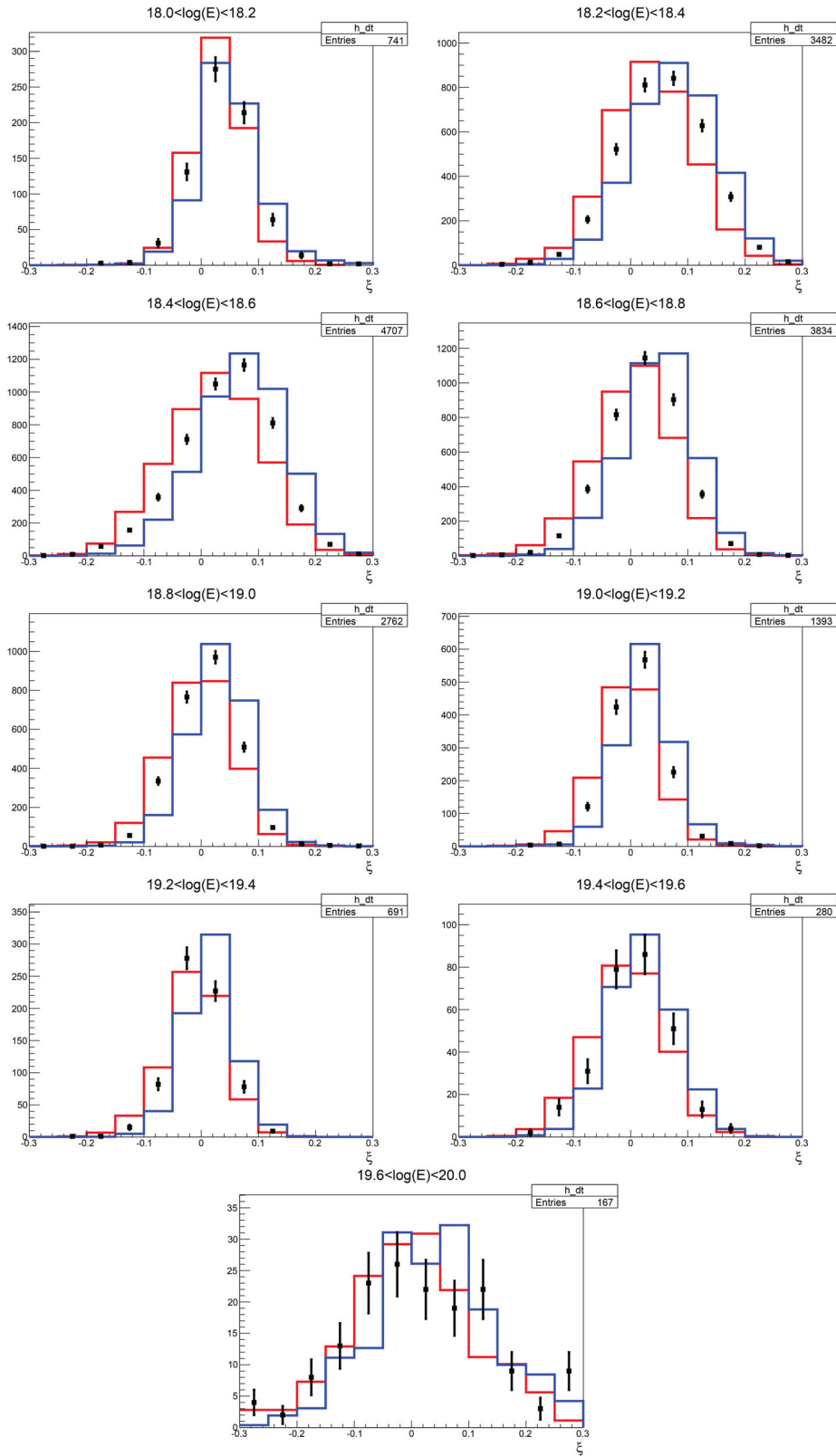


FIG. 2. ξ parameter distribution for different energy bins. Proton MC is shown with red lines, iron MC is shown with blue lines and black dots represent the data.

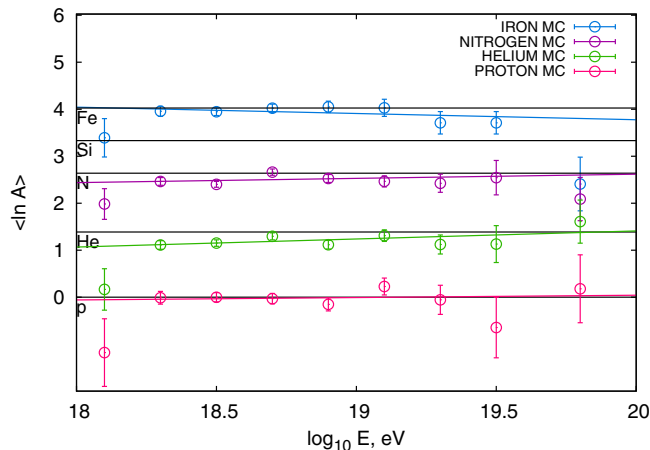


FIG. 3. $\langle \ln A \rangle$ approximated with a straight line for proton (red), helium (green), nitrogen (purple) and iron (blue) Monte-Carlo sets. Error bars for each $\langle \ln A \rangle$ point represent the statistical uncertainty of the method.

and data ξ distributions separately in each energy bin. The implemented method is `TFractionFitter` ROOT package [45,46]. The second part of the Monte-Carlo is used in this step to obtain the fraction of proton and iron in the data, ϵ_p and $\epsilon_{Fe} = 1 - \epsilon_p$, respectively.

The first estimate of an average atomic mass is based on the derived fraction of protons ϵ_p :

$$\langle \ln A \rangle^{(1)} = \epsilon_p \times \ln(M_p) + (1 - \epsilon_p) \times \ln(M_{Fe}), \quad (5)$$

where $M_p = 1.0$ and $M_{Fe} = 56.0$ are average atomic masses of proton and iron nuclei.

We note that the number of proton and iron-induced simulated showers is the same, while the trigger and reconstruction efficiency differ. The proton fraction ϵ_p is defined as the fraction of proton *simulated* events in the mixture which corresponds to the hypothesis that $\langle \ln A \rangle^{(1)}$ is the average atomic mass of the particles arriving to the atmosphere. It is assumed that the detector efficiency affects the statistics of the proton and iron MC showers in the same way it affects the proton and iron-induced events in the data.

C. Bias correction

One may go further and build the bias correction procedure based on the Fig. 3. Assuming that the cosmic ray flux is composed of particles of single type in each energy bin, it is possible to construct the quadratic polynomial function $\ln A_{\text{true}}(\langle \ln A \rangle)$ based on $\langle \ln A \rangle$ obtained for four MC sets, for which the $\ln A$ values are known.

In the Fig. 4 uncorrected $\langle \ln A \rangle^{(1)}$ and $\langle \ln A \rangle_{\text{non-linear}}$ obtained with nonlinear bias corrections are shown in comparison.

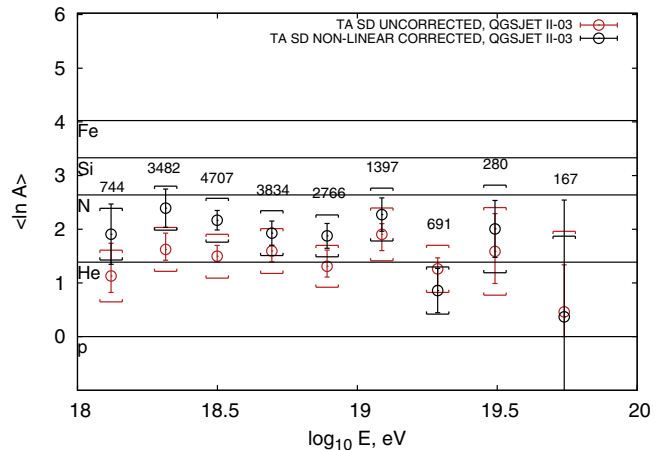


FIG. 4. Uncorrected $\langle \ln A \rangle^{(1)}$ in comparison with $\langle \ln A \rangle_{\text{non-linear}}$ corrected by non-linear function in each energy bin; statistical error is shown with error bars and systematic error as estimated in Sec. IV A is shown with brackets of the corresponding color. Numbers represent the number of data events in the corresponding energy bin.

IV. RESULTS AND DISCUSSION

A. Estimation of the systematic error

The nonlinear correction applied for the method is based on the assumption that the obtained composition is monotype. Thus the main source for the systematic error of the method is the inability to distinguish the mixture of a given elements and the single-type-particle composition.

To derive the systematic uncertainty, in each energy bin 100 mixtures of p , He , N , and Fe Monte-Carlo sets were created, among which 50 mixtures are random monotype, 25 are random two-component and 25 are random four-component. Its $\langle \ln A \rangle$ values were estimated with the use of `TFractionFitter` template fitting method and non-linear bias corrections applied and compared with the “true” values calculated from the known fractions. Mean systematic error is estimated as:

$$\delta \ln A_{\text{syst.}} = 0.44 \quad (6)$$

B. Hadronic models dependency

Composition results, both derived from surface detectors and in a hybrid mode, have a strong dependence on hadronic models used during Monte-Carlo simulations. Besides the one used in the above analysis, QGSJETII-04 [47], an improvement of QGSJETII-03 model, EPOS-LHC [48] and SYBILL [49] models are also widely used.

All of the hadronic interaction models are based on the collider data and extrapolated to the UHECR energies. The analysis by the Pierre Auger Observatory has shown the inconsistency between muon signal predicted by simulations and data [50]. The same conclusions were also made based on the Telescope Array SD data [51].

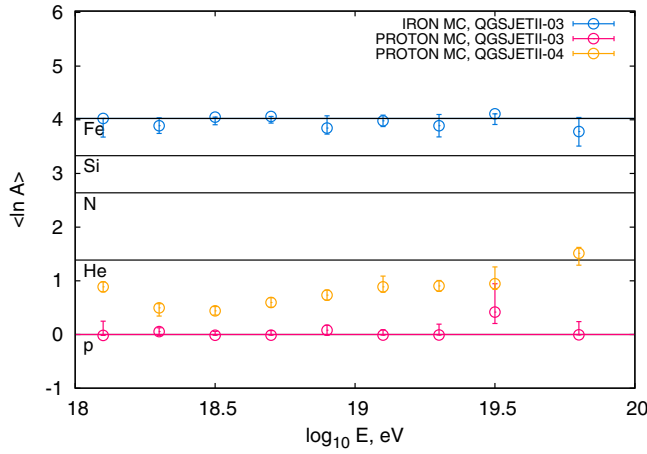


FIG. 5. $\langle \ln A \rangle$ approximated with a straight line for proton (red) and iron (blue) Monte-Carlo sets created with QGSJETII-03 hadronic interaction set and for proton MC set, created with QGSJETII-04 (orange line). Error bars for each $\langle \ln A \rangle$ point represent the statistical bias of the method.

This discrepancy may be the source of additional systematic bias which may affect the observables used for the composition study.

We study the systematic error introduced by the limited knowledge of the hadronic interaction models based on the comparison of the two models: QGSJETII-03 and QGSJETII-04 [47]. For the latter, an additional proton Monte-Carlo set with the use of QGSJETII-04 model is simulated. The set is subjected to the same multivariate analysis procedure trained with the original QGSJETII-03 Monte-Carlo. The result is shown in the Fig. 5, while the hadronic model uncertainty as a function of energy is shown in Fig. 6. The uncertainty from hadronic interaction models is minimal at $10^{18.5}$ eV with $\delta \ln A_{\text{hadr.}} = 0.23$ and maximal at $10^{19.75}$ eV with $\delta \ln A_{\text{hadr.}} = 0.74$.

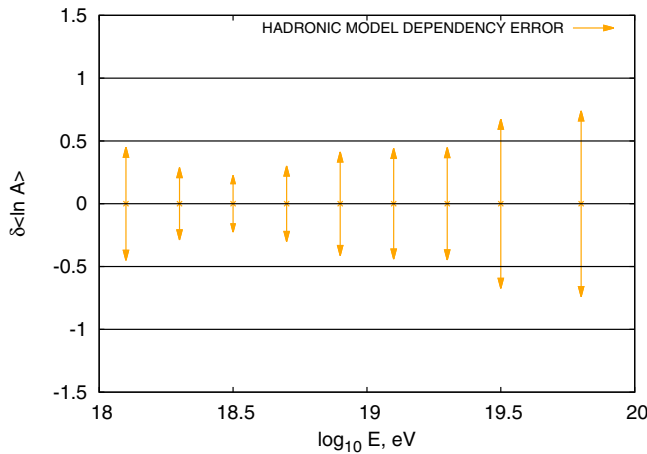


FIG. 6. Hadronic model dependency error of the method as a function of energy, based on a comparison with QGSJETII-04 hadronic interaction model.

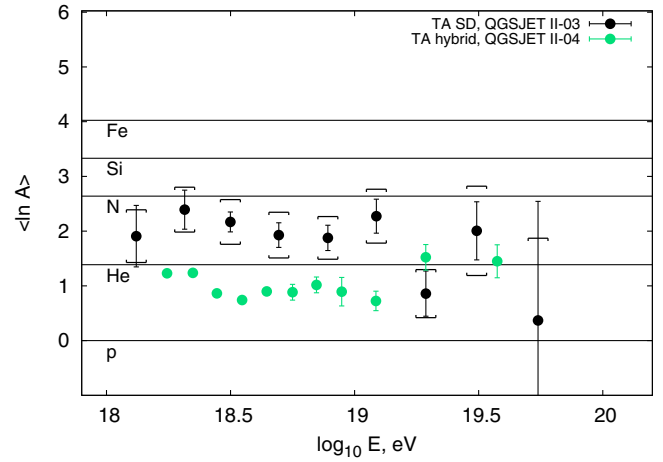


FIG. 7. Average atomic mass $\langle \ln A \rangle$ in comparison with the Telescope Array hybrid results [16]; statistical error is shown with error bars, systematic error is shown with brackets.

C. Composition

Mean logarithm of atomic mass as a function of energy without bias corrections and with the linear corrections applied is shown in Fig. 4. Within the errors, the average atomic mass of primary particles shows no significant energy dependence and corresponds to $\langle \ln A \rangle = 2.0 \pm 0.1(\text{stat.}) \pm 0.44(\text{syst.})$.

TA SD composition results in comparison with TA hybrid results are shown in Fig. 7. Comparisons with Pierre Auger Observatory SD X_{MAX}^{μ} based on muon density and muon arrival times and azimuthal risetime asymmetry, HiRes stereo X_{max} and Yakutsk muon detector results are shown in Fig. 8 and 9, respectively. We mention that while there exist composition results based on the Pierre Auger Observatory hybrid observations [54], we focus only on the comparison with the corresponding surface detector results.

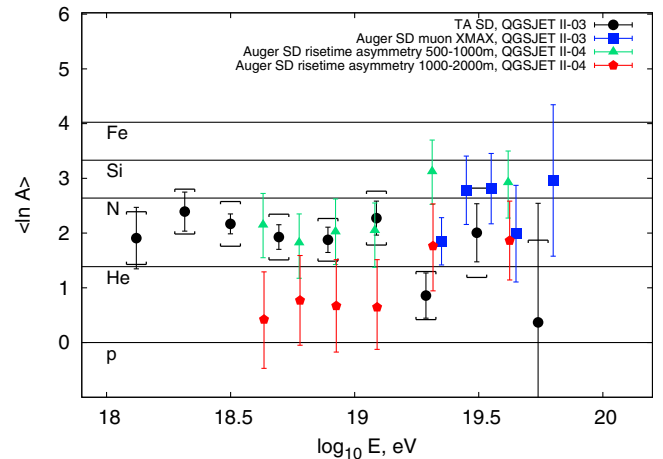


FIG. 8. Average atomic mass $\langle \ln A \rangle$ in comparison with the Pierre Auger Observatory X_{MAX}^{μ} and risetime asymmetry results [18,52]; statistical error is shown with error bars, systematic error is shown with brackets.

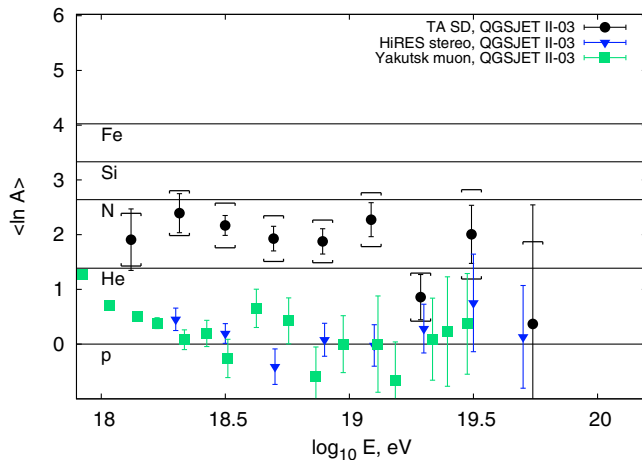


FIG. 9. Average atomic mass $\langle \ln A \rangle$ in comparison with the HiRes stereo results [14] and with the Yakutsk ρ_μ results [53]; statistical error is shown with error bars, systematic error is shown with brackets.

The obtained composition is qualitatively consistent with the TA hybrid and the Pierre Auger Observatory results, while all the points lie higher than the pure proton composition observed by HiRes and Yakutsk.

ACKNOWLEDGMENTS

The Telescope Array experiment is supported by the Japan Society for the Promotion of Science (JSPS) through Grants-in-Aid for Priority Area 431, for Specially Promoted Research JP21000002, for Scientific Research (S) JP19104006, for Specially Promote Research JP15H05693, for Scientific Research (S) JP15H05741 and for Young Scientists (A) JPH26707011; by the joint research program of the Institute for Cosmic Ray Research (ICRR), The University of Tokyo; by the U.S. National Science Foundation Grants No. PHY-0601915, No. PHY-1404495, No. PHY-1404502, and No. PHY-1607727; by the National Research Foundation of Korea (2017K1A4A3015188, 2016R1A2B4014967, 2017R1A2A1A05071429, 2016R1A5A1013277); by IISN Project No. 4.4502.13, and Belgian Science Policy under IUAP VII/37 (ULB). The development and application of the multivariate analysis method is supported by the Russian Science Foundation Grant No. 17-72-20291 (INR). The foundations of Dr. Ezekiel R. and Edna Wattis Dumke, Willard L. Eccles, and George S. and Dolores Dore Eccles all helped with generous donations. The State of Utah supported the project through its Economic Development Board, and the University of Utah through the Office of the Vice President for Research. The experimental site became available through the cooperation of the Utah School and Institutional Trust Lands Administration (SITLA), U.S. Bureau of Land Management (BLM), and the U.S. Air Force. We appreciate the assistance of the State of Utah and Fillmore offices of the BLM in crafting the Plan of

Development for the site. Patrick Shea assisted the collaboration with valuable advice on a variety of topics. The people and the officials of Millard County, Utah have been a source of steadfast and warm support for our work which we greatly appreciate. We are indebted to the Millard County Road Department for their efforts to maintain and clear the roads which get us to our sites. We gratefully acknowledge the contribution from the technical staffs of our home institutions. An allocation of computer time from the Center for High Performance Computing at the University of Utah is gratefully acknowledged. The cluster of the Theoretical Division of INR RAS was used for the numerical part of the work.

APPENDIX: COMPOSITION-SENSITIVE VARIABLES

In this work, a set of fourteen composition-sensitive variables is used:

- (1) Linsley front curvature parameter, as described in Sec. II B.
- (2)–(3) Area-over-peak (AoP) of the signal at 1200 m and AoP slope parameter [55]:
Given a time resolved signal from a surface station, one may calculate its peak value and area, which are both well-measured and not much affected by fluctuations.
 $AoP(r)$ is fitted with a linear fit:

$$AoP(r) = \alpha - \beta(r/r_0 - 1.0),$$

where $r_0 = 1200$ m, α is $AoP(r)$ value at 1200 m and β is its slope parameter.

- (4) Number of detectors hit.
- (5) Number of detectors excluded from the fit of the shower front by the reconstruction procedure [56].
- (6) $\chi^2/\text{d.o.f.}$ of the joint geometry and LDF fit.
- (7)–(8) S_b parameter for $b = 3$ and $b = 4.5$ [30]. The definition of the parameter is the following:

$$S_b = \sum_{i=1}^N \left[S_i \times \left(\frac{r_i}{r_0} \right)^b \right],$$

where S_i is the signal of i -th detector, r_i is the distance from the shower core to this station in meters and $r_0 = 1200$ m—reference distance. The value $b = 3$ and $b = 4.5$ are used as they provide the best separation.

- (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.

This value is summed over both upper and lower layers of all stations of the event. To suppress accidental peaks resulting from FADC noise, the peak is defined as a time bin with a signal exceeding 0.2 vertical equivalent muons (VEM) with the value higher than signals of the 3 preceding and 3 consequent time bins.

- (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.

-
- [1] H. Tokuno *et al.* (Telescope Array Collaboration), *J. Phys. Conf. Ser.* **293**, 012035 (2011).
 - [2] T. Abu-Zayyad *et al.* (Telescope Array Collaboration), *Nucl. Instrum. Methods Phys. Res., Sect. A* **689**, 87 (2012).
 - [3] H. Tokuno *et al.* (Telescope Array Collaboration), *Nucl. Instrum. Methods Phys. Res., Sect. A* **676**, 54 (2012).
 - [4] R. U. Abbasi *et al.* (HiRes Collaboration), *Phys. Rev. Lett.* **100**, 101101 (2008); *Astropart. Phys.* **32**, 53 (2009).
 - [5] J. Abraham *et al.* (Pierre Auger Collaboration), *Phys. Rev. Lett.* **101**, 061101 (2008); *Phys. Lett. B* **685**, 239 (2010).
 - [6] T. Abu-Zayyad *et al.* (Telescope Array Collaboration), *Astrophys. J.* **768**, L1 (2013).
 - [7] K. Greisen, *Phys. Rev. Lett.* **16**, 748 (1966).
 - [8] Z. T. Zatsepin and V. A. Kuz'min, *Zh. Eksp. Teor. Fiz. Pis'ma Red.* **4**, 144 (1966) [*JETP Lett.* **4**, 78 (1966)].
 - [9] G. Gelmini, O. E. Kalashev, and D. V. Semikoz, *J. Exp. Theor. Phys.* **106**, 1061 (2008).
 - [10] R. Aloisio, D. Boncioli, A. di Matteo, A. F. Grillo, S. Petrer, and F. Salamida, *J. Cosmol. Astropart. Phys.* **10** (2015) 006.
 - [11] A. Saveliev, L. Maccione, and G. Sigl, *J. Cosmol. Astropart. Phys.* **03** (2011) 046.
 - [12] A. V. Sokolov and M. S. Pshirkov, *Eur. Phys. J. C* **77**, 908 (2017).
 - [13] T. K. Gaisser *et al.*, *Phys. Rev. D* **47**, 1919 (1993).
 - [14] R. U. Abbasi *et al.* (HiRes Collaboration), *Phys. Rev. Lett.* **104**, 161101 (2010).
 - [15] A. Aab *et al.* (Pierre Auger Collaboration), *Phys. Rev. D* **90**, 122006 (2014).
 - [16] R. U. Abbasi *et al.* (Telescope Array Collaboration), *Astrophys. J.* **858**, 76 (2018).
 - [17] V. De Souza *et al.*, *Proc. Sci. (ICRC2017)* (2017) 522.
 - [18] A. Aab *et al.* (Pierre Auger Collaboration), *Phys. Rev. D* **93**, 072006 (2016).
 - [19] A. Aab *et al.* (Pierre Auger Collaboration), *Phys. Rev. D* **96**, 122003 (2017).
 - [20] L. Breiman *et al.*, *Classification and regression trees* (Wadsworth International Group, Belmont, California, 1984).
 - [21] R. E. Schapire, *Mach. Learn.* **5**, 197 (1990).
 - [22] M. Krause, E. Pueschel, and G. Maier, *Astropart. Phys.* **89**, 1 (2017).
 - [23] A. Aab *et al.* (Pierre Auger Collaboration), *J. Cosmol. Astropart. Phys.* **04** (2017) 009.
 - [24] R. Abbasi *et al.* (IceCube Collaboration), *Phys. Rev. D* **83**, 012001 (2011).
 - [25] J. Linsley and L. Scarsi, *Phys. Rev.* **128**, 2384 (1962).
 - [26] M. Teshima *et al.*, *J. Phys. G* **12**, 1097 (1986).
 - [27] M. Takeda *et al.*, *Astropart. Phys.* **19**, 447 (2003).
 - [28] T. Abu-Zayyad *et al.* (Telescope Array Collaboration), *Phys. Rev. D* **88**, 112005 (2013).
 - [29] Y. Takahashi *et al.* (Telescope Array Collaboration), *AIP Conf. Proc.* **1367**, 157 (2011).
 - [30] G. Ros, A. D. Supanitsky, G. A. Medina-Tanco, L. del Peral, J. C. D'Olivo, and M. D. Rodríguez Frías, *Astropart. Phys.* **35**, 140 (2011).
 - [31] D. Heck *et al.*, *Forschungszentrum Karlsruhe Report No. FZKA-6019*, 1998.
 - [32] S. Ostapchenko, *Nucl. Phys. B, Proc. Suppl.* **151**, 143 (2006).
 - [33] T. T. Böhlen, F. Cerutti, M. P. W. Chin, A. Fassò, A. Ferrari, P. G. Ortega, A. Mairani, P. R. Sala, G. Smirnov, and V. Vlachoudis, *Nucl. Data Sheets* **120**, 211 (2014).
 - [34] A. Ferrari, P. R. Sala, A. Fasso, and J. Ranft, Report Nos. CERN-2005-010, SLAC-R-773, INFN-TC-05-11, DOI: [10.2172/877507](https://doi.org/10.2172/877507).
 - [35] W. R. Nelson, H. Hirayama, and D. W. O. Rogers, Reports No. SLAC-0265, No. SLAC-265, SLAC-R-0265, SLAC-R-265.
 - [36] A. M. Hillas, *Nucl. Phys. B, Proc. Suppl.* **52**, 29 (1997).
 - [37] M. Kobal (Pierre Auger Collaboration), *Astropart. Phys.* **15**, 259 (2001).
 - [38] D. S. Gorbunov, G. I. Rubtsov, and S. V. Troitsky, *Phys. Rev. D* **76**, 043004 (2007).
 - [39] B. T. Stokes, R. Cady, D. Ivanov, J. N. Matthews, and G. B. Thomson, *Astropart. Phys.* **35**, 759 (2012).
 - [40] S. Agostinelli *et al.* (GEANT4 Collaboration), *Nucl. Instrum. Methods Phys. Res., Sect. A* **506**, 250 (2003).
 - [41] T. Abu-Zayyad *et al.* (Telescope Array Collaboration), [arXiv:1403.0644](https://arxiv.org/abs/1403.0644).
 - [42] T. Abu-Zayyad *et al.* (Telescope Array Collaboration), *Astrophys. J.* **768**, L1 (2013).
 - [43] A. Hocker *et al.*, *Proc. Sci. ACAT2007* (2007) 040.
 - [44] Y. Freund and R. E. Schapire, *ICML'96 Proceedings of the Thirteenth International Conference on International Conference on Machine Learning* (Morgan Kaufmann Publishers Inc. San Francisco, CA, USA, 1996), pp. 148.
 - [45] R. Brun and F. Rademakers, *Nucl. Instrum. Methods Phys. Res., Sect. A* **389**, 81 (1997).
 - [46] R. Barlow and C. Beeston, *Comput. Phys. Commun.* **77**, 219 (1993).
 - [47] S. Ostapchenko, *Phys. Rev. D* **83**, 014018 (2011).
 - [48] T. Pierog, I. Karpenko, J. M. Katzy, E. Yatsenko, and K. Werner, *Phys. Rev. C* **92**, 034906 (2015).

- [49] R. S. Fletcher, T. K. Gaisser, P. Lipari, and T. Stanev, *Phys. Rev. D* **50**, 5710 (1994).
- [50] A. Aab *et al.* (Pierre Auger Collaboration), *Phys. Rev. Lett.* **117**, 192001 (2016).
- [51] R. U. Abbasi *et al.* (Telescope Array Collaboration), *Phys. Rev. D* **98**, 022002 (2018).
- [52] P. Abreu *et al.* (Pierre Auger Collaboration), *Contributions to the 32nd International Cosmic Ray Conference, Beijing, China, 2011* (2011).
- [53] A. Sabourov *et al.*, Proc. Sci. ICRC2017 (**2017**) 553.
- [54] J. Bellido *et al.* (Pierre Auger Collaboration), Proc. Sci. ICRC 2017 (**2017**) 506.
- [55] J. Abraham *et al.* (Pierre Auger Collaboration), *Phys. Rev. Lett.* **100**, 211101 (2008).
- [56] T. Abu-Zayyad *et al.* (Telescope Array Collaboration), *Astrophys. J. Lett.* **768**, L1 (2013).