Mass composition of ultra-high-energy cosmic rays with the Telescope Array Surface Detector Data

R.U. Abbasi, M. Abe, T. Abu-Zayyad, M. Allen, R. Azuma, E. Barcikowski, J.W. Belz, D.R. Bergman, S.A. Blake, R. Cady, B.G. Cheon, J. Chiba, M. Chikawa, A. di Matteo, T. Fujii, K. Fujita, M. Fukushima, G. Furlich, T. Goto, W. Hanlon, M. Hayashi, Y. Hayashi, N. Hayashida, K. Hibino, K. Honda, D. Ikeda, N. Inoue, T. Ishii, R. Ishimori, H. Ito, D. Ivanov, H.M. Jeong, S. Jeong, C.C.H. Jui, K. Kadota, F. Kakimoto, O. Kalashev, K. Kasahara, H. Kawai, S. Kawakami, S. Kawana, K. Kawata, E. Kido, H.B. Kim, J.H. Kim, S. Kishigami, S. Kitamura, Y. Kitamura, V. Kuzmin, M. Kuznetsov, Y. J. Kwon, K.H. Lee, B. Lubsandorzhiev, J.P. Lundquist, K. Machida, K. Martens, T. Matsuyma, J.N. Matthews, R. Mayta, M. Minamino, K. Mukai, I. Myers, K. Nagasawa, S. Nagataki, R. Nakamura, T. Nakamura, T. Nonaka, H. Oda, S. Ogio, J. Ogura, M. Ohsawa, H. Ohoka, T. Okuda, Y. Omura, M. Ono, R. Onogi, A. Oshima, S. Ozawa, I.H. Park, M.S. Piskunov, M.S. Pshirkov, J. Remington, D.C. Rodriguez, G. Rubtsov, D. Ryu, H. Sagawa, R. Sahara, K. Saito, Y. Saito, N. Sakaki, N. Sakurai, L.M. Scott, T. Seki, K. Sekino, P.D. Shah, F. Shibata, T. Shibata, H. Shimodaira, B.K. Shin, H.S. Shin, J.D. Smith, P. Sokolsky, B.T. Stokes, S.R. Stratton, T.A. Stroman, T. Suzawa, Y. Takagi, Y. Takahashi, M. Takamura, M. Takei, T. Takeishi, A. Taketa, M. Takeda, Y. Takeda, H. Tanaka, K. Tanaka, S.B. Thomas, G.B. Thomson, P. Tinyakov, I. Tkachev, H. Tokuno, T. Tomida, S. Troitsky, Y. Tsubasada, K. Tsutsumi, Y. Uchihori, S. Udo, F. Urban, T. Weng, M. Yamamoto, R. Yamanouchi, H. Yamaoka, K. Yamazaki, J. Yang, K. Yashiro, Y. Yoneda, S. Yoshida, H. Yoshii, Y. Zhezher, Z. Zundel

1 High Energy Astrophysics Institute and Department of Physics and Astronomy, University of Utah, Salt Lake City, Utah, USA
2 The Graduate School of Science and Engineering, Saitama University, Saitama, Saitama, Japan
3 Graduate School of Science and Technology, University of Tsukuba, Tsukuba, Ibaraki, Japan
4 Department of Physics and Department of the Research Institute of Natural Science, Hanyang University, Seoul, Korea
5 Department of Physics, Tokyo University of Science, Noda, Chiba, Japan
6 Department of Physics, Kindai University, Osaka, Japan
7 Service de Physique Thorique, Universite Libre de Bruxelles, Belgium
8 Institute for Cosmic Ray Research, University of Tokyo, Kashiwa, Chiba, Japan
9 Graduate School of Science, Osaka City University, Osaka, Japan
10 Kavli Institute for the Physics and Mathematics of the Universe (WPI), Todai Institutes for Advanced Study, University of Tokyo, Kashiwa, Chiba, Japan
11 Information Engineering Graduate School of Science and Technology, Shinshu University, Nagano, Nagano, Japan
12 Faculty of Engineering, Kanagawa University, Yokohama, Kanagawa, Japan
13 Interdisciplinary Graduate School of Medicine and Engineering, University of Yamanashi, Kofu, Yamanashi, Japan
14 Astrophysical Big Bang Laboratory, RIKEN, Wako, Saitama, Japan
15 Department of Physics, Sungkyunkwan University, Jang-an-gu, Suwon, Korea
16 Department of Physics, Tokyo City University, Setagaya-ku, Tokyo, Japan
17 Institute for Nuclear Research of the Russian Academy of Sciences, Moscow, Russia
18 Advanced Research Institute for Science and Engineering, Waseda University, Shinjuku-ku, Tokyo, Japan
19 Department of Physics, Chiba University, Chiba, Chiba, Japan
20 Department of Physics, School of Natural Sciences, Ulsan National Institute of Science and Technology, UNIST-gil, Ulsan, Korea
21 Department of Physics, Yonsei University, Seodaemun-gu, Seoul, Korea
22 Academic Assembly School of Science and Technology Institute of Engineering, Shinshu University, Nagano, Nagano, Japan
23 Faculty of Science, Kochi University, Kochi, Kochi, Japan
24 Department of Physical Sciences, Ritsumeikan University, Kusatsu, Shiga, Japan
25 Sternberg Astronomical Institute, Moscow M.V. Lomonosov State University, Moscow, Russia
26 Department of Physics and Astronomy, Rutgers University - The State University of New Jersey, Piscataway, New Jersey, USA
27 Earthquake Research Institute, University of Tokyo, Bunkyo-ku, Tokyo, Japan
28 Department of Engineering Science, Faculty of Engineering, Osaka Electro-Communications University, Neyagawa-shi, Osaka, Japan
29 Graduate School of Information Sciences, Hiroshima City University, Hiroshima, Hiroshima, Japan
30 Institute of Particle and Nuclear Sciences, KEK, Tsukuba, Ibaraki, Japan
31 National Institute of Radiological Science, Chiba, Chiba, Japan
32 CEICO, Institute of Physics, Czech Academy of Sciences, Prague, Czech Republic
33 Department of Physics and Institute for the Early Universe, Ewha Womans University, Seodaemun-gu, Seoul, Korea
34 Department of Physics, Ehime University, Matsuyama, Ehime, Japan

* Deceased
** Corresponding author, zhezher.yana@physics.msusu.ru
The results on ultra-high-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 $\langle \ln A \rangle = 1.52 \pm 0.08$. The result is compared to the mass composition obtained by the Telescope Array with Xmax technique along with the results of other experiments. Possible systematic errors of the method are discussed.

Keywords: ultra-high-energy cosmic rays – Telescope Array – mass composition – boosted decision trees

I. INTRODUCTION

The Telescope Array (TA) experiment is the largest ultra-high-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 km$^2$. Each detector is composed of two layers of 1.2 cm thick extruded scintillator of the 3 m$^2$ 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 cut-off 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 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 Xmax as a composition-sensitive observable [13]. There are UHE composition results available based on Xmax measured by the three experiments: HiRes, Pierre Auger Observatory and Telescope Array [14–16]. The two latter results are compatible within the systematic errors in Xmax measurement which are of the order of $10 - 20$ g/cm$^2$ in the energy range up to $10^{19}$ 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 Xmax sensitivity to the mass composition, although approaches based on, for example, the risetime asymmetry of the signal have been proposed [18, 19]. 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 $\xi$ 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 the Section 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 Section IV.
II. DATA SET 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 detector. 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 \((> 0.3 \text{ MIP})\) 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_{\text{core}}, y_{\text{core}}, \theta, \phi, S_{800}, t_0, a \([28]\):

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

\[
S (r) = S_{800} \times \text{LDF} (r),
\]

\[
\text{LDF} (r) = f (r) / f (800 \text{ m}),
\]

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

\[
R_m = 90.0 \text{ m}, \ R_1 = 1000 \text{ m}, \ R_L = 30 \text{ m}, \]

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

\[
r = \sqrt{(x_{\text{core}} - x)^2 + (y_{\text{core}} - y)^2},
\]

where \(x_{\text{core}}, y_{\text{core}}, x \) and \(y \) are obtained from the predefined coordinate system of the array centered at the Central Laser Facility (CLF) \([29]\). \(t_{\text{plane}} \) 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 A 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 \(X_{\text{max}} \). These are the 3rd and 4.5th moments of the LDF \(S_3, \ S_6 \), the sum of the signals of all the detectors of the event, the number of the detectors hit and \(\chi^2/d.o.f. \) of the LDF fit.

The second group is related to the shower front which is in turn sensitive to both \(X_{\text{max}} \) 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\(^\circ\);
3. reconstructed core position inside the array with the distance of at least 1200 m from the edge of the array;
4. \(\chi^2/d.o.f. \) doesn’t exceed 4 for both the geometry and the LDF fits;
5. \(\chi^2/d.o.f. \) doesn’t exceed 5 for the joint geometry and LDF fit.
6. an arrival direction is reconstructed with accuracy less than 5\(^\circ\);
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 \(\chi^2 \) condition for the joint fit is included \([28]\).

After the cuts, the SD data set contains 18077 events with energy greater than \(10^{18} \text{ eV} \).
C. Simulations

For the Monte-Carlo simulations, CORSIKA software package [30] is used along with the QGSJETII-03 model for high-energy hadronic interactions [31]. FLUKA [32] [33] for low energy hadronic interaction and EGS4 [34] 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 [35]. Within thinning, all particles with energies greater than a certain fraction of the primary energy \( \epsilon \) 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 \( \epsilon = 10^{-6} \) with an additional weight limitation according to [30] is used for simulations. The thinning allows to achieve CPU-time efficiency, but at the same time introduces artificial statistical fluctuations [37]. The de-thinning procedure is developed and implemented [38] in order to restore the statistical properties of the shower.

The detector response is simulated by the GEANT4 package [39]. Real-time array status and detector calibration information for 9 years of observations are used for each simulated event [40]. 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 \( 10^{17.5} - 10^{20.5} \text{ eV} \) 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 creation for the Telescope Array is described in details in [41].

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 [42]. The adaptive boosting (AdaBoost) algorithm is employed [21] [43] with the number of trees NTrees=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 A. 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 \( \xi \) for each data and Monte-Carlo event. \( \xi \) resides in the range \( \xi \in [-1;1] \), where \( \xi = 1 \) is a pure signal event, \( \xi = -1 \) — pure background event. The variable \( \xi \) is used in the following one-dimensional analysis.

B. Estimation of an average atomic mass

Following the two-component approximation, the distribution of \( \xi \) in the data is compared to different mixtures of the p and Fe Monte-Carlo events with the Kolmogorov-Smirnov test. The second part of the Monte-Carlo is used in this step and the mixtures are made with the proton fraction \( \epsilon_p \) step of 2.5%. The first estimate of an average atomic mass is then given by the following equation:

\[
\langle \ln A \rangle^{(1)} = \epsilon_p \ln (M_p) + (1 - \epsilon_p) \ln (M_{Fe}) ,
\]

where \( \epsilon_p \) is the fraction of protons in the mixture with the smallest KS-distance, \( M_p = 1.0 \) 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 \( \epsilon_p \) 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. 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

The third part of the Monte-Carlo sets are used to correct possible bias of the estimator given by the Eq. 5. We estimate \( \langle \ln A \rangle \) for this remaining MC set and linearly approximate the result for proton and iron sets with \( y_p(x) \) and \( y_{Fe}(x) \), where \( x = \log_{10} E \). These lines slightly differ from the constants \( \ln A = \ln(M_p) \) and \( \ln A = \ln(M_{Fe}) \) due to limited statistics.

The result is then corrected with the following linear function:

\[
\langle \ln A \rangle = \ln (M_p) + \frac{\langle \ln A \rangle^{(1)} - y_p(x)}{y_{Fe}(x) - y_p(x)} \times (\ln (M_{Fe}) - \ln (M_p)) .
\]

where \( y_p(x) \) and \( y_{Fe}(x) \) are linear approximations for the MC \( \langle \ln A \rangle \) distributions.
IV. RESULTS AND DISCUSSION

A. Estimation of the systematic error

Figure 1 shows $\xi$ parameter distribution histograms for all the energy bins, BDT parameters distribution histograms for energy bin $\log_{10}E = 18.8 - 19.0$ are denoted in Fig. 2. On both figures proton MC is shown with red lines, iron MC is shown with blue lines and black dots represent the data.

Fractional difference for proton and iron $\xi$ distributions as an estimate for the method separation power for the one of the energy bins $10^{18.2} \text{eV} < E < 10^{18.4} \text{eV}$ is shown in Fig. 3.

To validate a method based on a mixture of two components, MC sets for events initiated by helium and nitrogen primaries were created. The same method was applied to it, and the results are shown in the Fig. 4.

Full systematic error is comprised of deviations of $p$, He, N and Fe $\langle \ln A \rangle$-distributions from “expected” values of $\langle \ln A \rangle$. One may see from the Fig. 4 that $p$ and Fe $\langle \ln A \rangle (E)$ points almost perfectly follow the lines $\langle \ln A \rangle = 0$ and $\langle \ln A \rangle = \ln (56{,}0)$ so their contribution to the systematic error is negligible. The lines for the helium and nitrogen Monte-Carlo sets do not perfectly follow the straight and horizontal lines $\ln (4{,}0)$ and $\ln (14{,}0)$ correspondingly. The systematic error of the method is estimated as the standard deviation:

$$\delta^2 \ln A_{\text{syst.}} = \sum_{X=p,He,N,Fe} \frac{w_X}{M} \sum_j \langle \ln A_X \rangle_j - \langle \ln A_X \rangle^2 \tag{7}$$

where $\langle \ln A_X \rangle_j$ is the result of the method in the j-th energy bin and $\ln A_X$ is the real atomic mass, which is equal to $1.0$, $4.0$, $14.0$, $56.0$ for protons, helium, nitrogen and iron correspondingly. Assuming a flat prior on composition $w_p = w_{He} = w_N = w_{Fe} = 0.25$, an estimate of the two-component method systematic error is $\delta \ln A_{\text{syst.}} = 0.36$.

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 [14], an improvement of QGSJETII-03 model, EPOS-LHC [15] and SYBILL [16] 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 [47]. The same conclusions were also made based on the Telescope Array SD data [48]. 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 [14]. 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 shown in the Fig. 5 indicates that the hadronic interaction model uncertainty estimated as the standard deviation is $\delta \ln A_{\text{hadr.}} = 0.4$.

C. Composition

Mean logarithm of atomic mass as a function of energy without bias corrections is shown in Fig. 6. The corrected $\langle \ln A \rangle$ is shown in Fig. 7. Within the errors, the average atomic mass of primary particles shows no significant energy dependence and corresponds to $\langle \ln A \rangle = 1.52 \pm 0.08 (\text{stat.}) \pm 0.36 (\text{syst.})$.

TA SD composition results in comparison with TA hybrid results are shown in Fig. 8. Comparisons with Pierre Auger Observatory SD $X_{\text{MAX}}$ and risetime asymmetry, HiRes stereo $X_{\text{MAX}}$ and Yakutsk results based on muon density all obtained with the use of the same QGSJETII-03 hadronic interaction model are shown in Fig. 9 and 10 respectively. The obtained composition is qualitatively consistent with the TA hybrid and the Pierre Auger Observatory results.

ACKNOWLEDGMENT

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) JP19H04006, 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 awards PHY-0601915, PHY-1404495, PHY-1404502, and PHY-1607727; by the National Research Foundation of Korea (2017R1A1A3015188 ; 2016R1A2B4011967; 2017R1A2A1A05071429, 2016R1A5A1013277); by IJSN 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. 17-72-05741 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 awards PHY-0601915, PHY-1404495, PHY-1404502, and PHY-1607727; by the National Research Foundation of Korea (2017R1A1A3015188 ; 2016R1A2B4011967; 2017R1A2A1A05071429, 2016R1A5A1013277); by IJSN 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 Dunke, 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
FIG. 1. $\xi$ 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.

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
FIG. 2. 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.

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.
FIG. 3. Fractional difference for proton and iron $\xi$ distributions for the energy bin $10^{18.2} \, \text{eV} < E < 10^{18.4} \, \text{eV}$.

FIG. 4. $\langle \ln A \rangle$ approximated with a straight line for proton (red), helium (green), nitrogen (purple) and iron (blue) Monte-Carlos. Error bars for each $\langle \ln A \rangle$ point represent the statistical bias of the method. The “shift” between He and N MCs and ln (4.0) and ln (14.0) is an estimate of the method’s systematic error.

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.

FIG. 6. Uncorrected $\langle \ln A \rangle$ as a function of energy; proton MC is shown with red line, iron MC is shown with blue line.

FIG. 7. Corrected $\langle \ln A \rangle$ as a function of energy; proton MC is shown with red line, iron MC is shown with blue line.

FIG. 8. Average atomic mass $\langle \ln A \rangle$ in comparison with the Telescope Array hybrid results [16].
FIG. 9. Average atomic mass \( \langle \ln A \rangle \) in comparison with the Pierre Auger Observatory \( X_{\mu}^{\text{MAX}} \) and risetime asymmetry results [18].

FIG. 10. Average atomic mass \( \langle \ln A \rangle \) in comparison with the HiRes stereo results [14] and with the Yakutsk \( \rho_{\mu} \) results [49].

APPENDIX A: COMPOSITION-SENSITIVE VARIABLES

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

1. Linsley front curvature parameter, as described in section II B.
2–3. Area-over-peak (AoP) of the signal at 1200 m and AoP slope parameter [50].

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.

\[
\text{AoP}(r) \text{ is fitted with a linear fit:}
\]

\[
\text{AoP}(r) = \alpha - \beta \left( \frac{r}{r_0} - 1 \right),
\]

where \( r_0 = 1200 \text{ m} \), \( \alpha \) is AoP \( (r) \) value at 1200 m and \( \beta \) 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 [51].
6. \( \chi^2/d.o.f. \) of the joint geometry and LDF fit.
7–8. \( S_b \) parameter for \( b = 3 \) and \( b = 4.5 \) [52]. 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 \( r_0 = 1200 \text{ m} \) – 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.
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. Meth. A 689, 87 (2013) [arXiv:1201.4964 [astro-ph.IM]].
[3] H. Tokuno et al. [Telescope Array Collaboration], Nucl. Instrum. Meth. A 676, 54 (2012) [arXiv:1201.0002 [astro-ph.IM]].
[4] PRL 100 (2008) & Astropart. Phys. 32 (2010)
[5] PRL 101 (2008) & Phys. Lett. B 685 (2010)
[6] T. Abu-Zayyad et al. [Telescope Array Collaboration], Astrophys. J. 768, L1 (2013) [arXiv:1205.5067 [astro-ph.IM]].
[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).
[9] G. Gelmini, O. E. Kalashev and D. V. Semikoz, J. Exp. Theor. Phys. 106, 1061 (2008) [astro-ph/0506128].
[10] R. Aloisio, D. Boncioli, A. di Matteo, A. F. Grillo, S. Petreka and F. Salamida, JCAP 1510, no. 10, 006 (2015) arXiv:1505.04020 [astro-ph.HE].
[11] A. Saveliev, L. Maccione and G. Sigl, JCAP 1103, 046 (2011) arXiv:1101.2903 [astro-ph.HE].
[12] A. V. Sokolov and M. S. Pshirkov, arXiv:1611.04949 [hep-ph].
[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) arXiv:0910.4184 [astro-ph.HE].
[15] A. Aab et al. [Pierre Auger Collaboration], Phys. Rev. D 90, no. 12, 122006 (2014) arXiv:1409.5803 [astro-ph.HE].
[16] W. Hanlon for the Telescope Array Collaboration, Contributions to the 2016 International Conference on Ultra-High Energy Cosmic Rays, Kyoto, Japan, October 2016.
[17] V. De Souza et al. Proceedings of the ICRC 2017, CR1167.
[18] P. Abreu et al. [Pierre Auger Collaboration], Contributions to the 32nd International Cosmic Ray Conference, Beijing, China, August 2011 arXiv:1107.4804 [astro-ph.HE].
[19] A. Aab et al. [Pierre Auger Collaboration], Phys. Rev. D 96, no. 12, 122003 (2017) arXiv:1710.07249 [astro-ph.HE].
[20] L. Breiman et al., Wadsworth International Group (1984).
[21] R.E. Schapire, Mach. Learn. 5 (1990) 197.
[22] M. Krause et al., Astropart. Phys. 89, 1 (2017) arXiv:1701.06928 [astro-ph.IM].
[23] A. Aab et al. [Pierre Auger Collaboration], JCAP 1704 (2017) no.04, 009 arXiv:1612.01517 [astro-ph.HE].
[24] R. Abbasi et al. [IceCube Collaboration], Phys. Rev. D 83 (2011) 012001 arXiv:1010.3980 [astro-ph.HE].
[25] J. Abraham et al. [IceCube Collaboration], Phys. Rev. D 88, no. 2, 022002 (2013) arXiv:1205.5067 [astro-ph.HE].
[26] T. Abu-Zayyad et al. [Telescope Array Collaboration], Phys. Rev. D 88, no. 11, 112005 (2013) arXiv:1304.5614 [astro-ph.HE].
[27] Y. Takahashi et al. [Telescope Array Collaboration], AIP Conf. Proc. 1367, 157 (2011).
[28] D. Heck et al., Report FZKA-6019 (1998), Forschungszentrum Karlsruhe.
[29] S. Ostapchenko, Nucl. Phys. Proc. Suppl. 151, 143 (2006) hep-ph/0412332.
[30] T. T. B¨ohlen et al., Nucl. Data Sheets 120, 211 (2014).
[31] A. Ferrari, P. R. Sala, A. Fasso and J. Ranft, CERN-2005-010, SLAC-R-773, INFN-TC-05-11.
[32] W. R. Nelson, H. Hirayama, D.W.O. Rogers, SLAC-0265 (permanently updated since 1985).
[33] A. M. Hillas, Nucl. Phys. Proc. Suppl. 52B, 29 (1997).
[34] M. Kobal [Pierre Auger Collaboration], Astropart. Phys. 15, 259 (2001).
[35] D. S. Gorbunov, G. I. Rubtsov and S. V. Troitsky, Phys. Rev. D 76, 043004 (2007).