



# Depth of Ultra High Energy Cosmic Ray Induced Air Shower Maxima Measured by the Telescope Array Black Rock and Long Ridge FADC Fluorescence Detectors and Surface Array in Hybrid Mode

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

(Telescope Array Collaboration)

<sup>1</sup> High Energy Astrophysics Institute and Department of Physics and Astronomy, University of Utah, Salt Lake City, UT, USA; [whanlon@cosmic.utah.edu](mailto:whanlon@cosmic.utah.edu)

<sup>2</sup> The Graduate School of Science and Engineering, Saitama University, Saitama, Saitama, Japan

<sup>3</sup> Graduate School of Science and Engineering, Tokyo Institute of Technology, Meguro, Tokyo, Japan

<sup>4</sup> Department of Physics and The Research Institute of Natural Science, Hanyang University, Seongdong-gu, Seoul, Republic of Korea

<sup>5</sup> Department of Physics, Tokyo University of Science, Noda, Chiba, Japan

<sup>6</sup> Department of Physics, Kindai University, Higashi Osaka, Osaka, Japan

<sup>7</sup> Service de Physique Théorique, Université Libre de Bruxelles, Brussels, Belgium

<sup>8</sup> Institute for Cosmic Ray Research, University of Tokyo, Kashiwa, Chiba, Japan; [iked@icrr.u-tokyo.ac.jp](mailto:iked@icrr.u-tokyo.ac.jp)

<sup>9</sup> Graduate School of Science, Osaka City University, Osaka, Osaka, Japan

<sup>10</sup> Kavli Institute for the Physics and Mathematics of the Universe (WPI), Todai Institutes for Advanced Study, University of Tokyo, Kashiwa, Chiba, Japan

<sup>11</sup> Information Engineering Graduate School of Science and Technology, Shinshu University, Nagano, Nagano, Japan

<sup>12</sup> Faculty of Engineering, Kanagawa University, Yokohama, Kanagawa, Japan

<sup>13</sup> Interdisciplinary Graduate School of Medicine and Engineering, University of Yamanashi, Kofu, Yamanashi, Japan

<sup>14</sup> Astrophysical Big Bang Laboratory, RIKEN, Wako, Saitama, Japan

<sup>15</sup> Department of Physics, Sungkyunkwan University, Jang-an-gu, Suwon, Republic of Korea

<sup>16</sup> Department of Physics, Tokyo City University, Setagaya-ku, Tokyo, Japan

<sup>17</sup> Institute for Nuclear Research of the Russian Academy of Sciences, Moscow, Russia

<sup>18</sup> Advanced Research Institute for Science and Engineering, Waseda University, Shinjuku-ku, Tokyo, Japan

<sup>19</sup> Department of Physics, Chiba University, Chiba, Chiba, Japan

<sup>20</sup> Department of Physics, School of Natural Sciences, Ulsan National Institute of Science and Technology, UNIST-gil, Ulsan, Republic of Korea

<sup>21</sup> Department of Physics, Yonsei University, Seodaemun-gu, Seoul, Republic of Korea

<sup>22</sup> Academic Assembly School of Science and Technology Institute of Engineering, Shinshu University, Nagano, Nagano, Japan

<sup>23</sup> Faculty of Science, Kochi University, Kochi, Kochi, Japan

<sup>24</sup> Department of Physical Sciences, Ritsumeikan University, Kusatsu, Shiga, Japan

<sup>25</sup> Sternberg Astronomical Institute, Moscow M.V. Lomonosov State University, Moscow, Russia

<sup>26</sup> Department of Physics and Astronomy, Rutgers University—The State University of New Jersey, Piscataway, NJ, USA

<sup>27</sup> Earthquake Research Institute, University of Tokyo, Bunkyo-ku, Tokyo, Japan

<sup>28</sup> Department of Engineering Science, Faculty of Engineering, Osaka Electro-Communication University, Neyagawa-shi, Osaka, Japan

<sup>29</sup> Graduate School of Information Sciences, Hiroshima City University, Hiroshima, Hiroshima, Japan

<sup>30</sup> Institute of Particle and Nuclear Studies, KEK, Tsukuba, Ibaraki, Japan

<sup>31</sup> National Institute of Radiological Science, Chiba, Chiba, Japan

<sup>32</sup> CEICO, Institute of Physics, Czech Academy of Sciences, Prague, Czech Republic

<sup>33</sup> Department of Physics and Institute for the Early Universe, Ewha Womans University, Seodaemun-gu, Seoul, Republic of Korea

<sup>34</sup> Department of Physics, Ehime University, Matsuyama, Ehime, Japan

Received 2018 January 16; revised 2018 March 27; accepted 2018 March 27; published 2018 May 9

## Abstract

The Telescope Array (TA) observatory utilizes fluorescence detectors and surface detectors (SDs) to observe air showers produced by ultra high energy cosmic rays in Earth's atmosphere. Cosmic-ray events observed in this way are termed hybrid data. The depth of air shower maximum is related to the mass of the primary particle that

<sup>35</sup> Deceased.

generates the shower. This paper reports on shower maxima data collected over 8.5 yr using the Black Rock Mesa and Long Ridge fluorescence detectors in conjunction with the array of SDs. We compare the means and standard deviations of the observed  $X_{\max}$  distributions with Monte Carlo  $X_{\max}$  distributions of unmixed protons, helium, nitrogen, and iron, all generated using the QGSJet II-04 hadronic model. We also perform an unbinned maximum likelihood test of the observed data, which is subjected to variable systematic shifting of the data  $X_{\max}$  distributions to allow us to test the full distributions, and compare them to the Monte Carlo to see which elements are not compatible with the observed data. For all energy bins, QGSJet II-04 protons are found to be compatible with TA hybrid data at the 95% confidence level after some systematic  $X_{\max}$  shifting of the data. Three other QGSJet II-04 elements are found to be compatible using the same test procedure in an energy range limited to the highest energies where data statistics are sparse.

*Key words:* acceleration of particles – astrochemistry – astroparticle physics – cosmic rays – elementary particles

## 1. Introduction

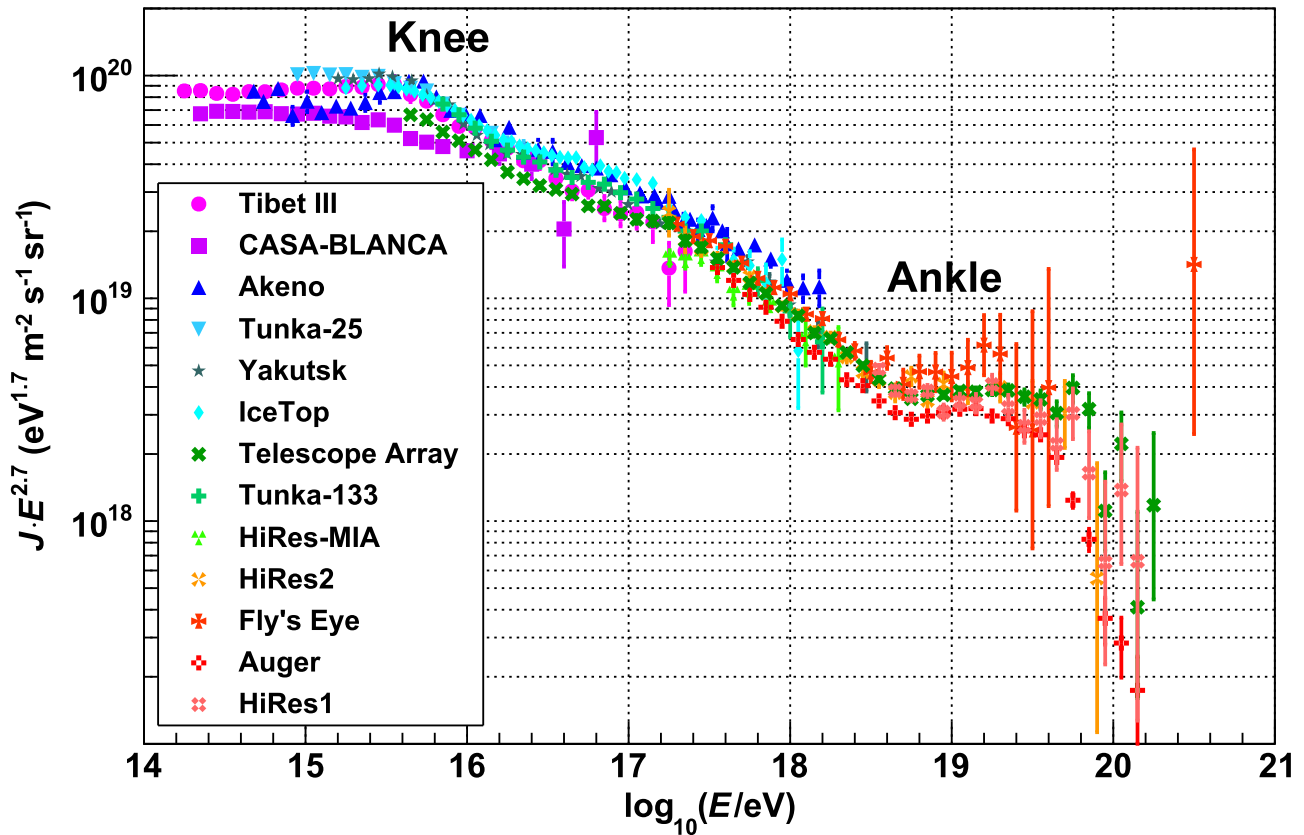
Ultra high energy cosmic-ray (UHECR) sources remain a mystery over a century since they were first observed by Hess in 1912 (Hess & Anderson 2013). Outstanding questions regarding the sources, acceleration mechanisms, propagation, and chemical composition of UHECRs have been studied now for over five decades, with the first of the large air shower arrays exceeding an area of 1 km<sup>2</sup> reporting results in 1961 (Linsley et al. 1961). The UHECR spectrum, as shown in Figure 1, exhibits structure that hints at correlated changes in chemical composition and energy that can help us resolve these long-standing questions. Of particular interest are the energy regions of UHECR flux dubbed the “knee” ( $E \approx 10^{15.5}$  eV), the “ankle” ( $E \approx 10^{18.7}$  eV), and the “GZK cutoff” (or suppression) ( $E \approx 10^{19.8}$  eV). Prior to the knee, the cosmic-ray flux is remarkably stable for six decades of energy, decreasing with energy as a power law,  $E^{-\gamma}$ , with  $\gamma \approx 2.7$ . The flux then steepens above the knee ( $\Delta\gamma \approx +0.4$ ), falling more rapidly. Around the energy of the ankle the flux begins to flatten ( $\Delta\gamma \approx -0.6$ ), until very rapidly dropping off nearly completely above the GZK cutoff. Models that wish to describe these changes in flux need to account for the maximum injection energy of astrophysical sources, acceleration either at the source or through other means such as shock waves, interactions with the interstellar medium, and chemical composition of cosmic rays observed in Earth’s atmosphere.

Composition provides a strong constraint on models describing UHECR sources and is therefore a fundamental parameter in modeling them. For instance, models that theorize about the origin of the knee come in two flavors: astrophysical and interaction models. Astrophysical models explain the knee as an intrinsic feature of the energy spectra of individual chemical species, resulting from magnetic rigidity dependence ( $E_{\max} \propto Z$ ). Different proposed acceleration mechanisms in magnetic field regions of galactic supernova remnants are theorized to boost the energy of particles to PeV and higher energies, limited by some maximum energy. The acceleration efficiency of higher- $Z$  nuclei allows those heavier elements to be boosted to higher energies than light elements. The spectra of individual nuclei show a series of cascading cutoffs as energy increases. The all-particle cosmic-ray spectrum around the knee under this model theorizes increasing particle mass with energy. Other versions of astrophysical models attempting to explain the origin of the knee use a leaky-box model in which light galactic nuclei cannot be contained in the galaxy as energy increases owing to their large gyroradii. Heavier nuclei cannot escape the galaxy as easily and contribute to the observed flux as a larger proportion of elements at higher energies. Interaction models of the knee propose that new

physics is in play as the air shower interacts in the atmosphere, producing, for example, undetected supersymmetric particles or other exotic particles not yet observed in nature. See Hoerandel (2004) for a review of many different models used to explain the UHECR knee feature.

In the energy region of the ankle, the cosmic-ray flux flattens, indicating a slight rise in the flux compared to energies below it. The ankle is traditionally thought of as the energy region where cosmic rays of extragalactic origin begin to dominate the spectrum. This is because there are few, if any, known sources in the galaxy able to accelerate nuclei to  $E \gtrsim 10^{19}$  eV while allowing the nuclei to remain contained in the galactic disk. A signature of galactic sources of UHECRs at the energy of the ankle would be anisotropy of arrival directions in the galactic plane, something that is not observed. Historically, the flattening of the spectrum at the ankle was described as the intersection of a steeply falling galactic spectrum ( $\gamma \approx 3.1$ ) and a flatter extragalactic spectrum ( $\gamma \approx 2-2.3$ ) at the energy of the ankle. This model is known simply as the ankle model. More recent models fitted to data from large cosmic-ray experiments use the signatures of propagation through the photon field of cosmic microwave background (CMB) radiation in intergalactic space. These propagation effects result in the suppression of flux above  $10^{20}$  eV due to photopion production with CMB photons (the GZK mechanism; Greisen 1966; Zatsepin & Kuzmin 1966), a bump due to pileup in the flux for primaries with energies below the photopion production energy, and a dip due to pair production during interaction with CMB photons. This is called the dip model. The dip model predicts that the galactic component of UHECRs disappears at a lower energy compared to predictions of the ankle model. The dip model is relatively insensitive to model parameters such as size and inhomogeneity of the source distributions, cosmological evolution of sources, and maximum acceleration energy. However, the dip model is sensitive to composition of heavy nuclei in the spectrum. Heavier nuclei interact and photospallate readily owing to larger cross section than protons, resulting in several primaries of lower total energy and a change in the shape of the dip. See, for example, Aloisio et al. (2007) for further discussion of the dip model and the transition from galactic to extragalactic cosmic rays. This model therefore makes testable predictions based on composition as well.

UHECR composition can be measured directly up to about  $10^{14}$  eV because the event rate here is about  $1 \text{ m}^{-2} \text{ hr}^{-1}$ , sufficient for balloon-borne or satellite experiments and their associated equipment to measure particle mass. Above this energy the particle cascades created by the cosmic-ray primary inelastically colliding with an air molecule must be observed if



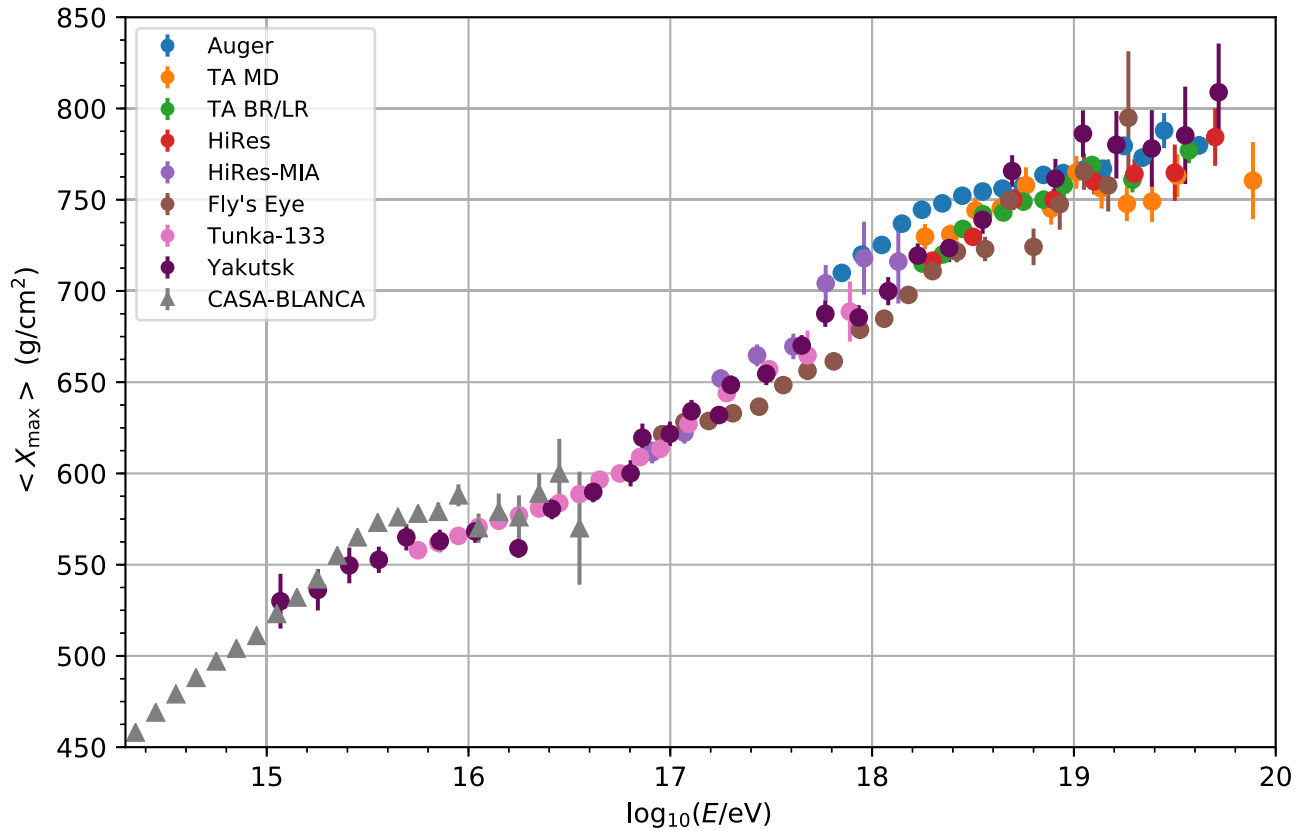
**Figure 1.** Cosmic-ray spectrum observed by recent experiments. In this figure the flux is scaled by  $E^{2.7}$ . Below the ankle energy this is approximately the power law followed by the flux and therefore appears flat in this figure. Near the energy of the ankle, the spectrum steepens, until flattening again near the ankle. Data are from Aartsen et al. (2013), Abbasi et al. (2008a), Abu-Zayyad et al. (2001), Amenomori (2008), Bird et al. (1994), Aab et al. (2015), Fowler et al. (2001), Ivanov (2016), Knurenko et al. (2015), Nagano et al. (1992), and Prosin et al. (2015).

one wishes to collect sufficient data. To do this, large ground-based experiments are required to observe the parts of the shower that survive to ground level or using fluorescence detectors to observe the light produced by the air shower. Neither method directly measures the mass of the primary, and a single observation of muons on the ground or light generated from an air shower cannot reveal the mass of an individual primary. Therefore, to measure the mass composition of UHECRs, we must resort to understanding the physics of extensive air showers, identifying those observables that can be related to the primary mass, and collecting large data sets to build statistical samples of sufficient size to reliably measure the average mass in some energy range. This method requires good hadronic models of high-energy interactions in matter to energy ranges not yet measured in the lab. Measurements from the Large Hadron Collider (LHC) reach up to about  $10^{17}$  eV in the lab frame, whereas UHECR primary particle energies above  $10^{20}$  eV have been measured. Hadronic models that predict particle elasticity, multiplicity, and interaction cross section currently require extrapolation over a few decades of energy for the energy region below the ankle and above.

UHECR composition measurements are performed best by fluorescence detectors that observe the depth in the atmosphere where the electromagnetic component of a cosmic-ray induced shower reaches a maximum. This atmospheric depth is called  $X_{\max}$  and is measured in  $\text{g cm}^{-2}$ . A toy model first developed by Heitler (Carlson & Oppenheimer 1937) demonstrates how  $X_{\max}$  is related to the primary particle energy and mass using a

simple branching model of electromagnetic (EM) showers in which a high-energy electron primary of energy  $E_0$  collides inelastically with a target particle. The EM shower is created and grows in size through the repeated processes of pair production and bremsstrahlung. In this model, making the simplistic assumption of a fixed interaction length,  $\lambda$ , between interactions, two new particles are generated and added to the shower for each existing particle. After  $n$  interactions, the total depth the shower has traveled is  $n\lambda = X$ , and the size of the shower is  $N(X) = 2^n = 2^{X/\lambda}$ . The average energy of each particle at depth  $X$  is  $E(X) = E_0/N(X) = E_0/2^{X/\lambda}$ . This process of particle generation at each interaction length continues until the average energy per particle decreases below some critical energy,  $E_c$ , defined as the energy at which particle energy lost as a result of collisions exceeds radiative energy losses. When the average particle energy is equal to  $E_c$ , the shower reaches its maximum size, called  $N_{\max}$ , and the depth is  $X_{\max}$ . Using the definition of  $N(X)$ , we find  $N_{\max} = E_0/E_c$  and  $X_{\max} \propto \ln(E_0/E_c)$ ; the number of particles generated at shower maximum is proportional to the primary particle energy, and  $X_{\max}$  is proportional to the log of that energy.

Showers initiated by a very high energy hadronic primary particle exhibit similar relationships (Matthews 2005). If the primary particle has mass  $A$  and energy  $E_0$ , we use the superposition principle to treat the particle as  $A$  independent nucleons each with an average initial energy of  $E_0/A$ . Using the Heitler model under the assumption of the superposition principle, we find  $X_{\max} \propto \ln(E_0/A)$ . In reality, hadronic



**Figure 2.**  $\langle X_{\max} \rangle$  observed by recent experiments. The dependence of mean depth with  $\log(E)$  over many decades of energy is readily apparent. Data are from Aab et al. (2014), Abbasi et al. (2010, 2014), Abu-Zayyad et al. (2001), Bird et al. (1994), Fowler et al. (2001), Knurenko & Petrov (2015), Prosin et al. (2015), and this work (TA BR/LR).

showers are more complicated, since for each hadronic interaction on average 2/3 of the particles produced are charged particles such as  $\pi^\pm$  and 1/3 are  $\pi^0$ . The  $\pi^0$  rapidly decay into two photons that contribute to the electromagnetic part of the shower. The relation between particle mass and  $X_{\max}$  predicted by the Heitler model is still valid. The property of shower universality tells us that for showers created by a hadronic primary particle of any mass the electromagnetic component evolves in the same way, parameterized by the shower age,  $s$  (Giller et al. 2003). Using this property, we can use the same method of observing  $X_{\max}$  of a shower to determine the mass of the primary particle, even if that particle is very light, such as a proton, or much heavier, such as an iron nucleus. For details about the treatment of hadronic showers as related to cosmic-ray composition, refer to Ulrich et al. (2011), Engel et al. (2011), and Kampert & Unger (2012).

Heavier primary particles are therefore expected to reach shower maximum at shallower depths in the atmosphere than light primaries. To experimentally measure UHECR composition, one can use fluorescence detectors to record energy and  $X_{\max}$  for many showers. For a given energy range  $\langle X_{\max} \rangle$  for lighter primaries will be larger than for heavy primaries. In addition, because of the superposition principle, the fluctuations in  $X_{\max}$  are expected to be smaller for heavy primaries. These data can be compared to models of individual primary species or mixtures of elements to determine the composition of UHECRs observed. Figure 2 shows  $\langle X_{\max} \rangle$  measured by many experiments over the past 30 yr.

Recent measurements of  $X_{\max}$  over the past decade have vastly improved statistics in the important region of the ankle

and above, with the three largest fluorescence-based measurements of HiRes, Telescope Array (TA), and Auger. HiRes reported on composition measured by events observed in stereo using simultaneous observation of two fluorescence detectors (Abbasi et al. 2010), and TA has presented results using hybrid reconstruction (Abbasi et al. 2014). Both results found indications of composition consisting of light primaries resembling mostly protons up to  $10^{19.8}$  eV, by comparing  $\langle X_{\max} \rangle$  and  $\sigma(X_{\max})$  of data to models. These results are consistent with the view (up until that time) that UHECRs with energies at the ankle and higher are most likely extragalactic protons. The existence of a suppression in the flux at the energy predicted by the GZK mechanism, first observed by HiRes (Abbasi et al. 2008a) and later confirmed by Auger (Abraham et al. 2008), fits in well with this scenario. However, Auger’s recent measurement of composition challenges this view. Auger has shown an energy evolution in  $\langle X_{\max} \rangle$  and  $\sigma(X_{\max})$  to heavier primaries above  $10^{18.3}$  eV (Aab et al. 2014). UHECR flux with increasing mass above the ankle leads to unexpected models that have been deemed “disappointing” for the field (Aloisio et al. 2012). Implications of such models are a lack of photopion production on CMB photons (and therefore very few ultra high energy cosmogenic neutrinos), no anisotropy of nearby sources due to strong deflection in magnetic fields, and no cutoff in the spectrum due to the GZK mechanism since the maximum energy of astrophysical accelerators is too low (Allard et al. 2008; Aloisio et al. 2012). The tension between these experimental results and the implications for particle astrophysics provide the impetus for further, high-precision studies of UHECR composition such as this one.

This work presents an analysis of  $X_{\max}$  data collected by the TA experiment over an 8.5 yr period. The  $X_{\max}$  distributions are collected in energy bins, and  $\langle X_{\max} \rangle$  and  $\sigma(X_{\max})$  are computed for each bin. Four sets of Monte Carlo, each representing a single chemical element, are generated and then reconstructed in the same manner as the data. The  $X_{\max}$  distributions,  $\langle X_{\max} \rangle$ , and  $\sigma(X_{\max})$  of the Monte Carlo are compared to the observed data. Statistical tests are used to compare the compatibility of the Monte Carlo to the data. Section 2 describes the design and operation of the TA experiment. Section 3 describes the hybrid method of observation and how data are reconstructed to measure  $X_{\max}$  for air showers. Section 4 examines the data collected, analysis cuts applied to the data, resolution, and bias of observables important to good  $X_{\max}$  reconstruction; compares data to Monte Carlo; and discusses the importance of understanding the different types of biases in measuring  $X_{\max}$ . Section 5 details the statistical tests used to measure compatibility of the different Monte Carlo sets to the data and the results of these tests. Conclusions of this analysis are presented in Section 6.

## 2. Apparatus

TA is one of the few detectors in the world able to shed light on the composition of UHECRs. TA is the successor experiment of the AGASA (Ohoka et al. 1997) and HiRes (Abu-Zayyad et al. 2000; Boyer et al. 2002) experiments. Expertise using surface arrays from the AGASA experiment and fluorescence detectors from the HiRes experiment is combined into a single cosmic-ray observatory able to observe UHECR flux over four decades of energy.

TA is located in Millard County, Utah (39°3N and 112°9W, 1400 m above sea level), consisting of 507 scintillation surface detectors (SDs) sensitive to muons and electrons and 48 fluorescence telescopes located in three fluorescence detector (FD) stations overlooking the counters. The spacing of the counters in the SD array is 1.2 km, and they are placed over an area of approximately 700 km<sup>2</sup>. Figure 3 shows the physical locations of the SDs and FDs.

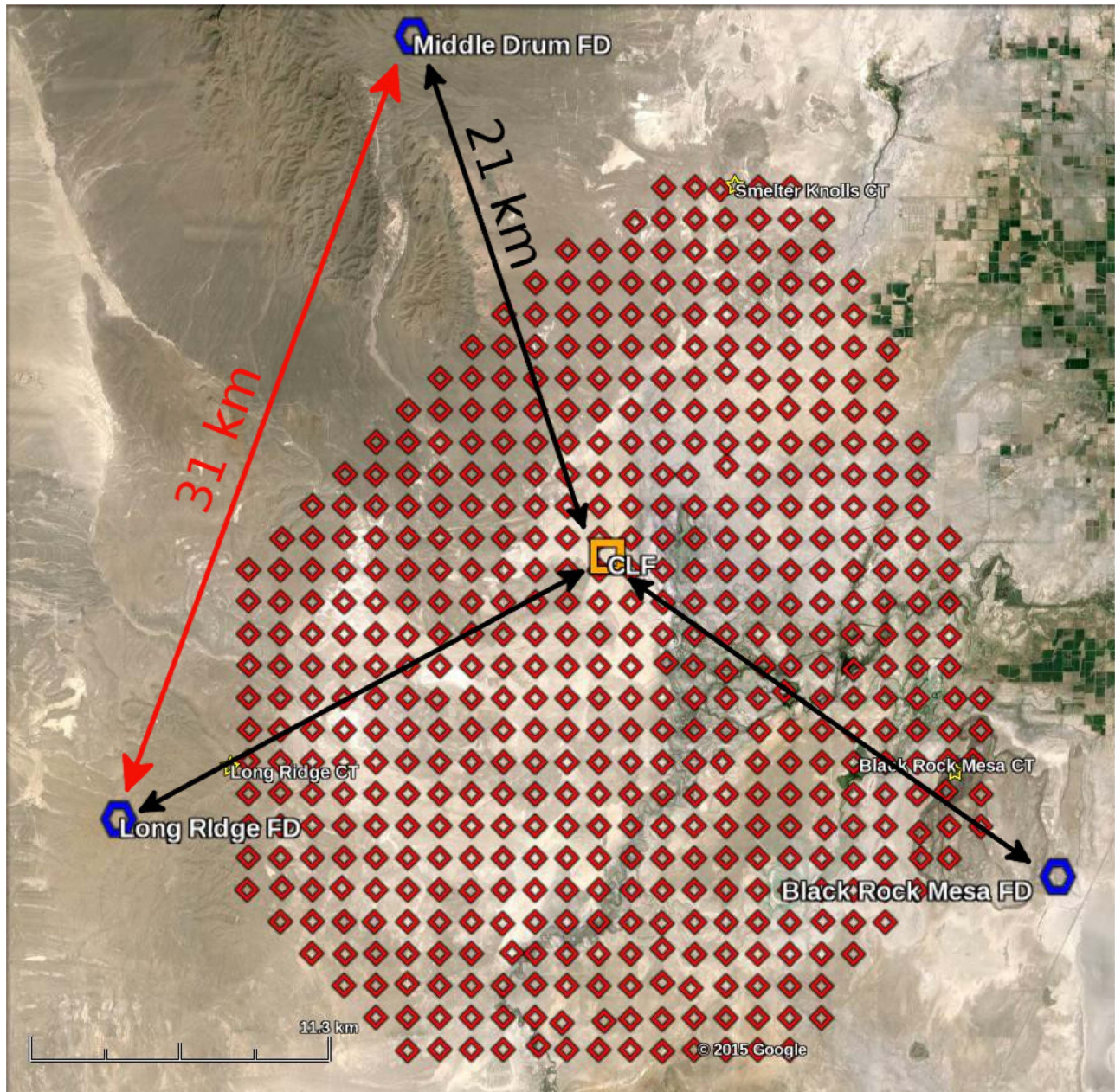
Each SD is made up of two layers of plastic scintillator 3 m<sup>2</sup> by 1.2 cm thick. Grooves running parallel along the length of each layer are etched into each scintillator layer, and 104 wavelength shifting fiber optic cables are embedded in them, for a total length of 5 m of fiber in each layer. When a charged particle passes through the scintillator and light is produced, the light is transmitted to a photomultiplier tube (PMT) via the fiber. Each scintillator layer has a dedicated PMT that is optically coupled to the fiber bundle to detect the passage of charged particles. The analog PMT signal is digitized via 12 bit Flash Analog-To-Digital Converter (FADC) electronics operating at 50 MHz sampling rate with signals stored in a local buffer. Each FADC bin is 20 ns wide, and a waveform consists of 128 FADC bins, providing a waveform buffer 2.56  $\mu$ s wide. Each SD electronics suite also has a Field-Programmable Gate Array (FPGA) that continuously monitors the FADC waveforms to monitor pedestals and to determine whether the event trigger condition is met. When an SD measures a signal above threshold, it can announce it to a remote Data Acquisition (DAQ) via radio communications. An SD can record two types of low-level triggers: a level 0 trigger, in which an integrated signal exceeding 15 FADC counts above pedestal is measured, and a level 1 trigger, in which an integrated signal exceeding 150 FADC (equivalent to three MIPs (minimum ionizing

particles)) counts above pedestal is measured. These remote DAQ locations are referred to as communication towers (CTs), as they monitor and receive data from many SDs and make the decision about high-level triggers based on the low-level trigger logic of all SDs that it communicates with. If three or more adjacent SDs announce level 1 triggers within an 8  $\mu$ s window, this constitutes a level 2 event trigger, and the CT directs all SDs that observed a level 0 trigger within  $\pm 32 \mu$ s of the event to send the waveform data to the CT for storage. Each SD has an onboard GPS unit to time-stamp event triggers, so the time of particle passage is also recorded by each SD and included as part of the event information.

SD event reconstruction is done by examining all level 2 triggers and finding SDs that have sufficient signal-to-noise ratio and also are connected in a small spacetime window. Using the positions of the SDs and their relative trigger times, the shower core and track direction can be determined. Shower energy is measured by relating the SD signal size 800 m from the shower axis (called S800) to a function that maps S800 and the shower zenith angle to primary particle energy. This mapping is determined by Monte Carlo simulation using CORSIKA (Heck et al. 1998) and is therefore model dependent. The final shower energy is found by scaling this energy mapping via a scale factor by using real events observed by both FD and SD and correcting the energy to that determined by the FD. Detailed information about the operation of the SD array can be found in Abu-Zayyad et al. (2013c) and Ivanov (2012).

The three FD stations are placed on the periphery of the SD array and look toward its center. Each is located the same distance from a central laser facility (CLF) about 21 km away, which contains a calibration laser that is fired throughout the night to monitor each FD's response and to monitor atmospheric quality. The Middle Drum (MD) FD station is located on the northern border of the SD array. It has 14 fluorescence telescopes that view 112° in azimuth arranged in two rings of elevation angle coverage. Ring 1 telescopes observe from 3° to 17° in elevation, and ring 2 telescopes observe from 17° to 31°. Seven telescopes are in each ring. Each telescope consists of a 5.1 m<sup>2</sup> mirror that reflects light from the sky onto a cluster of 256 PMTs arranged in a 16  $\times$  16 array, with each PMT monitoring approximately 1 milliradian solid angle. MD is built using the same sample and hold electronics and hardware that were used at the HiRes1 FD in the HiRes experiment (Abu-Zayyad et al. 2000). The MD site is also the location of the TALE FD station, which has 10 telescopes and is designed to observe low-energy cosmic rays near the energy of the ankle.

On the southeast and southwest borders of the SD array are the Black Rock Mesa (BR) and Long Ridge (LR) fluorescence detector stations. Each of these is made up of 12 telescopes in a two-ring configuration similar to MD. The electronics and hardware of these stations were newly built for the TA experiment and utilize FADC electronics. Each telescope has 256 PMTs focused onto a 6.8 m<sup>2</sup> mirror for light collection. Each PMT's analog signals are digitized by FADC electronics that employ 12 bit digitizers operating at 40 MHz. Before storage to the DAQ, four digital samples are summed to provide an equivalent 14 bit, 10 MHz sampling rate, providing a time resolution of 100 ns. Each telescope employs a track finder module that applies the trigger logic to incoming waveforms searching for spatial patterns that indicate a track caused by an extensive air shower. When an event trigger



**Figure 3.** Location of TA SDs and fluorescence detector stations. Each red diamond indicates the location of one of the 507 SDs. The blue hexagons show the locations of the fluorescence detectors that look inward over the SD array. Each FD is about 30 km distant from its neighboring FDs. All FDs are placed 21 km from a central laser facility.

occurs, a central computer orders the readout of telescope electronics to store the data for later offline analysis. More detailed information describing the construction and operation of the BR and LR fluorescence detectors can be found in Tameda et al. (2009) and Tokuno et al. (2012).

When a UHECR primary particle interacts in the atmosphere, an extensive air shower results, producing copious amounts of electrons and positrons among many other particle types. This electromagnetic component of the shower interacts with atmospheric  $N_2$ , producing fluorescence light, which is emitted isotropically and observed by FD telescopes on the ground. Because of the typical large distances to the showers and the relatively large size of each PMT's field of view, the shower appears as a downward-traveling line source. Each tube observes

the shower at different altitudes and therefore different atmospheric depth. Tube signals will vary depending on the amount of light produced in the sky at that depth, the atmospheric clarity, and the distance to the shower. The goal of FD shower reconstruction is to convert the signals and times measured by the passage of the shower, along with the fixed, known geometry of the PMTs, into the location and direction of a distant shower track, as well as the energy of the primary particle.

To accurately measure  $X_{\max}$ , fluorescence detectors must be used. For those events that also have their arrival time simultaneously measured by the SD array, though, the geometry of an individual shower can be very well measured. These *hybrid* events are very valuable for use in high-quality measurements of  $X_{\max}$ . TA began hybrid data collection in

2008 May and has now analyzed over 8.5 yr of data using this method. This paper is the first to report on  $X_{\max}$  measurements using the BR and LR stations.

The results of  $X_{\max}$  measurement using the MD FD and the SD array are reported in Abbasi et al. (2014). This analysis uses only the combined data of both BR and LR hybrid events. BR and LR employ identical electronics and hardware design, which are different from those found at MD. For example, BR and LR use larger mirrors than MD, which affects triggering and acceptance of tracks. Because hybrid reconstruction uses both the FDs and SDs, the locations of the FD stations relative to the SD array border are also an important consideration. BR and LR, located 3 and 4 km away from the SD border, respectively, are closer to the SD array than MD, which is 8 km away. This affects the acceptance of low-energy hybrid events. BR and LR are more efficient below  $10^{18.5}$  eV than MD is. Because of these differences, MD-reconstructed events are not considered for this analysis.

### 3. Measurement

#### 3.1. Hybrid Analysis Methodology

Hybrid reconstruction combines the separate SD and FD data streams by searching for time-coincident events. The kinematic properties of the shower, such as the charged particle depth profile, primary energy, and  $X_{\max}$ , are determined using the standard profile fitting procedure applied to FD data. The time-coincident SD data are used only to improve the shower track geometry, because it can very accurately measure the time of arrival and position of the shower core. Hybrid reconstruction strives to use the same routines as those used in the standard SD-only and FD-only analyses.

Hybrid reconstruction of data consists of the following steps:

1. Tagging events that trigger both the SD and FD.
2. FD and SD geometry measurement.
3. Hybrid fitting.
4. FD profile fitting.

##### 3.1.1. Tagging Hybrid Events

Tagging hybrid events is done by searching for hybrid-event candidates that independently trigger the SD array and either of the BR or LR FD stations. For real data the FD data stream is searched for downward-going events, and the SD data stream is searched for all level 2 triggers. Level 2 triggers are events in which three or more adjacent SD counters individually within 2400 m of each other detect a three MIP signal within  $8 \mu\text{s}$  (Abu-Zayyad et al. 2013c). FD downward-going events satisfy the requirements of the FD event trigger and are traveling from the atmosphere toward the ground (Tameda et al. 2009). Hybrid events are found by time matching SD and FD events using a  $500 \mu\text{s}$  window.

For Monte Carlo data a shower library of pregenerated CORSIKA events is used. The hybrid MC CORSIKA shower library is a collection of air showers generated at predetermined energies and zenith angles. For energies above  $10^{18}$  eV TA has 250 showers generated per 0.1 decade in energy. We generate showers of arbitrary energy by measuring the elongation rates of the Gaisser–Hillas profile parameters and correcting them from their generated values to the values expected for the energy being thrown for the MC. In this way we can throw a continuous distribution of shower profiles for any randomly

chosen energy. Each CORSIKA shower library element is thrown and set at a fixed zenith angle. Shower zenith angles are chosen according to a  $\sin \theta \cos \theta$  distribution for  $0^\circ \leq \theta \leq 60^\circ$ . To generate a random shower, the following procedure is performed:

1. A shower energy is randomly chosen according to the energy distribution of the combined HiRes1 and HiRes2 observed spectrum (Abbasi et al. 2008a).
2. A pregenerated shower element contained in the energy bin chosen from the previous step is randomly selected.
3. Shower azimuthal angle is chosen by a uniform distribution for  $0^\circ \leq \phi < 360^\circ$ .
4. Shower core is randomly selected within a circle of 25 km radius centered on the CLF.

Monte Carlo generation also assumes a given primary particle type and other choices that must be made when running CORSIKA, such as the high- and low-energy hadronic models.

At the conclusion of this stage of the reconstruction, each event will contain data from the SD counters and at least one of the FD stations. A hybrid-event candidate then has SD $\oplus$ BR data, or SD $\oplus$ LR data, or SD $\oplus$ BR $\oplus$ LR data, for events energetic enough to have been observed simultaneously by both FD stations.

This step of hybrid analysis is the only step that differs between reconstructing real data and Monte Carlo data. The real data and simulated data are packed into the same data format so that all analysis that follows uses the same software.

##### 3.1.2. FD and SD Geometry Measurement

This step of reconstruction determines the geometry of the shower as independently observed by the FD and by the SD. FD plane fitting is performed using the standard routines described in Abu-Zayyad et al. (2013b) and Stratton (2012). This procedure determines the geometry of the air shower track relative to the observing FD. This includes measuring the shower-detector plane (SDP) normal vector ( $\hat{n}$ ), shower zenith angle ( $\theta$ ), shower azimuth angle ( $\phi$ ), SDP angle ( $\psi$ ), shower impact parameter ( $R_p$ ), core location, and arrival time.

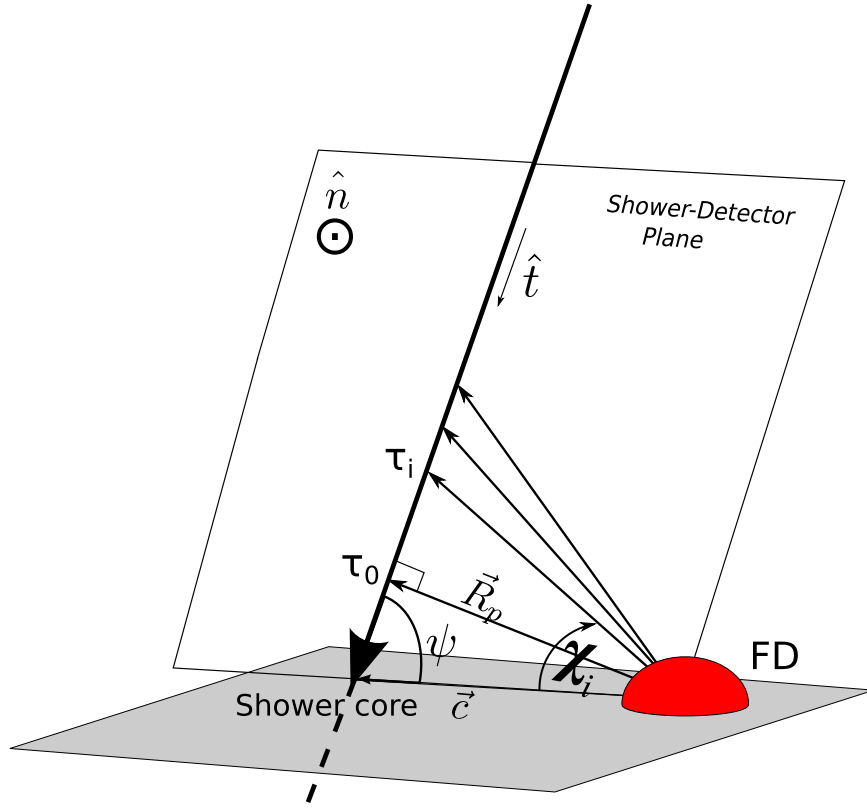
FD geometry reconstruction begins by selecting tubes that correlate closely in time and space on the PMT cluster face. For any event randomly triggered noise tubes are present as well. A filtering step is performed to reject tubes that are not part of the shower track. The SDP normal vector is found by finding those components of the vector that minimize the following sum:

$$\chi^2 = \sum_{i=1}^N (\hat{n} \cdot \hat{v}_i)^2 N_{pe,i}, \quad (1)$$

where  $N$  is the number of tubes along the shower track,  $\hat{v}_i$  is the pointing direction of tube  $i$  in the SDP, and  $N_{pe,i}$  is the tube signal measured as number of photoelectrons. FDs can accurately measure the SDP normal because many individual tubes are used to perform the fit. A typical shower contains 50 tubes along the shower track. Figure 4 illustrates the relationship among the different parts of the SDP.

Once the SDP is known, the shower impact parameter and the SDP angle are calculated by the time versus angle fit. The expected trigger time of PMT  $i$  is

$$\tau_i(\chi_i; R_p, \psi, t_0) = t_0 + \frac{R_p}{c} \tan\left(\frac{\pi - \psi - \chi_i}{2}\right). \quad (2)$$



**Figure 4.** Geometry of an air shower track as viewed by a fluorescence detector. The track vector  $\hat{t}$  is a unit vector that points along the shower in the direction in which the shower travels,  $\vec{c}$  points from the FD to the point on the ground where the shower impacts (the core),  $\vec{R}_p$  is the impact parameter, and  $\psi$  is the SDP angle. Each PMT that observes the shower has a viewing angle  $\chi_i$  and triggers at time  $t_i$ . The time  $\tau_i$  is the difference between the PMT trigger time and the light-travel time from the shower axis. The SDP normal vector is  $\hat{n}$ . All vectors and angles except  $\hat{n}$  are in the SDP.

$R_p$ , the shower impact parameter, also called the pseudo-distance, is the point of closest approach of the shower track and the origin of the observing FD.  $\psi$  is the SDP angle, or the angle on the ground between the shower track and the core vector, which points from the FD origin to the point of impact on the ground.  $\chi_i$  is the pointing direction of the PMT.  $t_0$  is the time when the shower front is located at the  $R_p$  point.  $R_p$ ,  $\psi$ , and  $t_0$  are parameters to be determined by fitting. A  $\chi^2$  minimization is performed on the function that measures the difference between the observed tube trigger times and the expected trigger times, described by Equation (2):

$$\chi^2 = \sum_{i=1}^N \frac{1}{\sigma_i^2} \left\{ t_i - \left[ t_0 + \frac{R_p}{c} \tan\left(\frac{\pi - \psi - \chi_i}{2}\right) \right] \right\}^2, \quad (3)$$

where  $t_i$  is the observed trigger time and  $\sigma_i$  is the timing uncertainty for tube  $i$ .

SD reconstruction is performed as well to determine the core arrival time and core location. A similar procedure for separating counters that trigger during an event and those that are noise triggers is employed by using time and space pattern recognition. Details about SD reconstruction can be found in Ivanov (2012).

At the end of this processing step, we have two independent determinations of the shower’s point of impact on the ground and its time of arrival.

### 3.1.3. Hybrid Fitting

Hybrid fitting is done by casting the individual “pixels” of the FD and SD that observe the passage of the shower into a

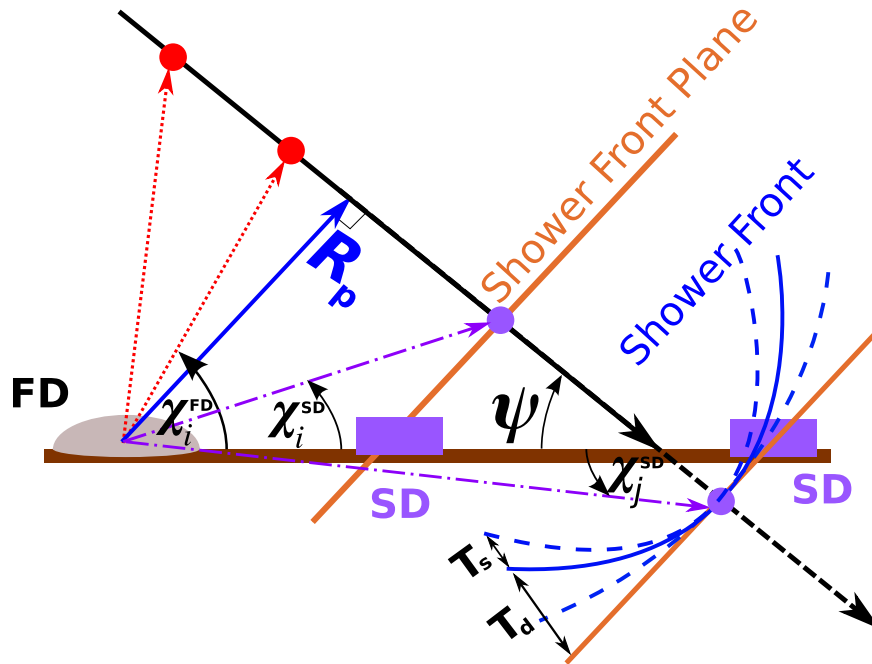
common frame of reference and then performing a four-component  $\chi^2$  minimization to redetermine some parameters of the shower geometry. By combining the independent FD and SD observations, the geometry fit is improved compared to FD-only or SD-only observations. The hybrid function minimized for this analysis is

$$\chi_{\text{hybrid}}^2 = \chi_{\text{FD}}^2 + \chi_{\text{SD}}^2 + \chi_{\text{SDP}}^2 + \chi_{\text{COC}}^2, \quad (4)$$

and the parameters being minimized are shower core position,  $C_x$  and  $C_y$ , zenith angle,  $\theta$ , azimuth angle,  $\phi$ , and shower core arrival time,  $t_c$ . Figure 5 shows how the common geometry of the FD and SDs is constructed. The FD pixels are treated as they typically are for normal FD reconstruction, each assigned a pointing direction and trigger time to determine the position and time of the shower point observed on the shower axis. The geometry and timing of the SDs that observe a common event are used to determine the pointing direction and times of those points on the shower axis in the reference frame of the FD.

$\chi_{\text{FD}}^2$  is the same as Equation (3) using the FD data described in Section 3.1.2. It is included in  $\chi_{\text{hybrid}}^2$  because in this fitting procedure the position of the shower core and the SDP are allowed to vary, potentially changing the previously calculated parameters of  $R_p$ ,  $\psi$ , and  $t_0$ .  $\chi_{\text{SDP}}^2$  (Equation (1)) is the same as described in Section 3.1.2 as well, allowing for varying shower track geometry. Because we cast the SD positions and times into a common reference frame as the FD tube data,  $\chi_{\text{SD}}^2$  uses Equation (1) to measure the best values for  $R_p$ ,  $\psi$ , and  $t_0$  of the shower track.  $\chi_{\text{COC}}^2$  is a term used to directly relate the





**Figure 5.** FD and SD components of shower geometry for a hybrid event. Each FD tube or SD counter can be considered a “pixel” with known geometry and timing that marked the passing of the shower. Timing of the individual SDs is corrected for shower front delay, where  $T_d$  is the mean delay from the shower front in particle arrival time at the SD, and  $T_s$  is the uncertainty on the delay time.

determination of the shower core as observed by the SDs (the center of charge) to the new shower core being fitted in Equation (4). This term is calculated by

$$\chi_{\text{COC}}^2 = \frac{(C_x - R_x)^2 + (C_y - R_y)^2}{\sigma_R^2}, \quad (5)$$

where  $R_x$  and  $R_y$  are the (fixed) components of the shower core as observed by the SD array, and  $\sigma_R$  is the SD uncertainty on the core location.

### 3.1.4. Profile Fitting

Once the shower geometry is improved by folding in the SD data, FD profile fitting is done as it is normally done and described in Abu-Zayyad et al. (2013b). This means that properties of the shower, such as the primary particle energy and the depth of shower maximum, are determined using only the light profile observed by the FD.

An inverse Monte Carlo method is used to determine the four parameters of the Gaisser–Hillas equation (Gaisser & Hillas 1977) as shown in Equation (6),  $N_{\text{max}}$ ,  $X_{\text{max}}$ ,  $X_0$ , and  $\lambda$ , which, after simulating the measured shower geometry, atmosphere, and detector acceptance, best mimics the tube-by-tube response measured in the FD:

$$N(x) = N_{\text{max}} \left( \frac{x - X_0}{X_{\text{max}} - X_0} \right)^{\frac{X_{\text{max}} - X_0}{\lambda}} \exp \left( \frac{X_{\text{max}} - x}{\lambda} \right). \quad (6)$$

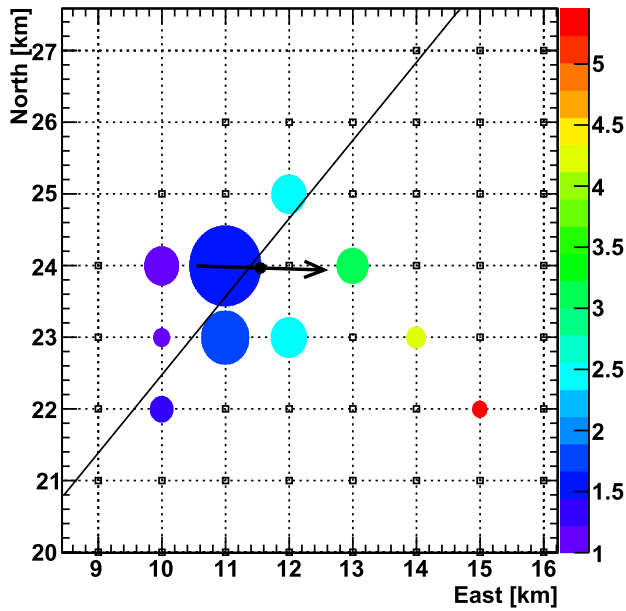
This analysis imposes constraints on the values of  $\lambda$  and  $X_0$ , fixing them to  $70 \text{ g cm}^{-2}$  and  $-60 \text{ g cm}^{-2}$ , respectively. It has been observed that  $\lambda$  and  $N_{\text{max}}$  are correlated, meaning that  $\lambda$  is strictly not an independent parameter. Also, ground-based fluorescence detectors cannot observe the depth of first interaction, which  $X_0$  is supposed to represent, but it is found to be physically meaningless since it often takes negative

values (Song 2004). We find, for this experiment, that by fixing these two parameters the bias in energy and  $X_{\text{max}}$  is reduced for simulated proton-induced showers over our energy range of interest. Both data and Monte Carlo utilize databases of detector PMT pedestals, PMT gains, and atmospheric conditions. These time-dependent databases are generated from the actual data and describe accurately the actual detector response recorded over each night’s observations.

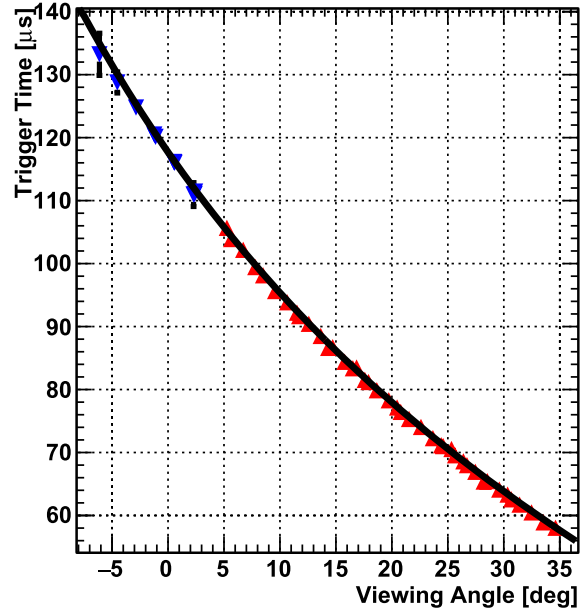
Once the best shower profile is determined, the energy of the primary particle is found by integrating the shower profile and applying corrections for energy transported by components of the shower not observed by direct fluorescence detection, i.e., neutrinos and muons. This correction is about 8%–5% for proton-induced showers at  $10^{18}$  and  $10^{20}$  eV, respectively. For iron-induced showers the missing energy correction is about 4% larger (Barcikowski 2011). This analysis applies the missing energy correction calculated from simulated proton-induced showers to all data and all Monte Carlo. We therefore expect a relative energy bias of around 4% in the reconstruction of proton and iron Monte Carlo data.

Figure 6 is an example of a typical hybrid event observed by TA. The SDs triggered in the event are shown in Figure 6(a). Marker size indicates signal measured by the SD, and color indicates trigger time. The arrow shows the azimuthal angle of the shower track, and the solid line is the projection of the SDP along the surface of Earth along the line between the observing FD and the shower core. Figure 6(b) shows the hybrid time versus angle fit. Red triangles show the trigger time and viewing angle of FD tubes that observed the passage of the shower. The blue triangles show the same information for SDs. The solid line is the fit to the combined SD and FD pixels using Equation (2). Figure 6(c) shows the FD tubes triggered in this event. Marker size indicates signal size, and color indicates trigger time. The solid line is the SDP found by fitting Equation (1). The flux and uncertainty in photons  $\text{deg}^{-1} \text{ m}^{-2}$  as a function of shower depth for each tube

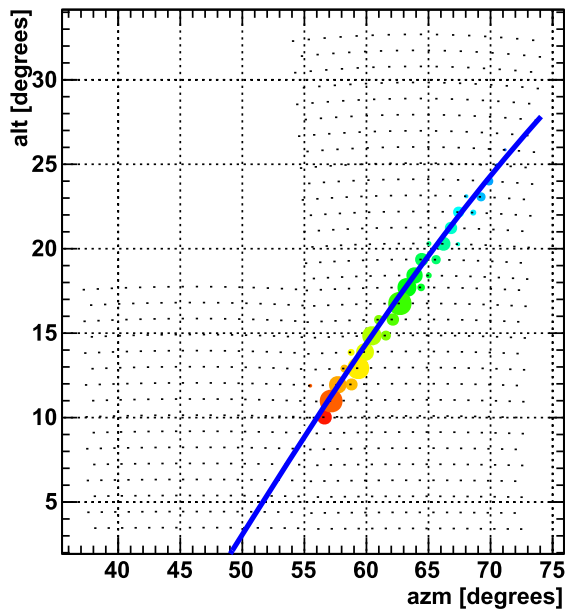
Event: 2014-01-22 04:42:34.391175



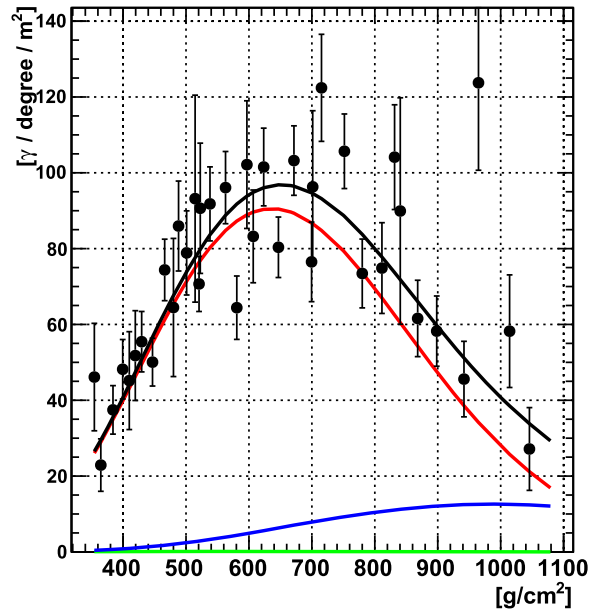
(a) Surface detector display.



(b) Time vs angle fit.



(c) Fluorescence detector display.



(d) Measured flux.

Figure 6. Typical hybrid event seen by TA in hybrid mode.

along the SDP are shown in Figure 6(d). The solid lines show the simulated flux generated for the best profile fit. Red is fluorescence flux, blue is Rayleigh-scattered flux, green is direct Cerenkov flux, and black is the sum of all these flux components.

#### 4. Analysis of the Data

The period of data collection covered in this work is from 2008 May 27 to 2016 November 29, about 8.5 yr. Because this analysis

is done using hybrid data, the collection period is limited by the approximately 10% duty cycle of the fluorescence detectors. This analysis examines 1500 nights of data collected over these 8.5 yr.

By performing time matching of SD and FD events as described in Section 3.1.1, 17,834 hybrid candidate events (hybrid events that have not gone through full reconstruction and analysis cuts) are found. The distribution of hybrid candidate events by the FD that observed them is as follows:

17,834	hybrid candidate events
10,381	BR events (mono and stereo)
8942	LR events (mono and stereo)
8892	BR events (mono)
7453	LR events (mono)
1489	BR + LR stereo events

Monocular hybrid candidate events are hybrid events observed by only one FD (either BR or LR for this analysis) and the SD array. Stereo hybrid candidate events are events observed by both FDs and the SD array. Even though an event is tagged as a stereo candidate event, there is no guarantee that the event is able to be fully reconstructed independently by each FD. For example, a medium-energy event located much closer to one FD than the other may not be of sufficient quality to pass all cuts for the FD located farther away.

Reconstruction of the data proceeds as described in Section 3.1. Once profile fitting is done, cuts are applied to the events to reject those that are poorly reconstructed, which may introduce  $X_{\max}$  bias and degrade  $X_{\max}$  resolution. These cuts are chosen to maximize the number of events collected without introducing large reconstruction biases.

1. *Boundary cut.* This cut ensures that the core of the event falls within the bounds of the SD array. Additionally, the core location must not be within 100 m of the boundary. This is to ensure that the charged particle distribution on the ground is fully contained within the array, allowing accurate reconstruction by the SDs.
2. *Track length cut.* This cut ensures that the shower track as observed by the FDs is sufficiently long enough to allow enough PMTs to observe it. Short shower tracks may be caused by distant showers or showers moving toward the detector, both instances of which indicate unfavorable geometry for accurate reconstruction. We require a shower track of  $10^\circ$  or greater to accept the track.
3. *Good tube cut.* Good tubes are tubes that have sufficient signal-to-noise ratio and are spatially and temporally part of the shower track as determined by the FD plane-fitting routines. We require 11 or more good tubes to accept a track.
4. *SDP angle cut.* This cut rejects events with SDP angle ( $\psi$ ) greater than  $130^\circ$ . Showers with  $\psi$  greater than  $90^\circ$  have some component of the track vector pointing toward the observing FD. As this angle grows, the shower is seen more and more head-on, increasing the contribution of direct Cerenkov light received. Showers with large direct Cerenkov signals are difficult to reconstruct accurately and are rejected.
5. *Time extent cut.* This cut rejects tracks with time profiles less than  $7 \mu\text{s}$ . This is another cut to ensure that short tracks, potentially approaching the observing FD directly, are removed from the data.
6. *Zenith angle cut.* Tracks with large zenith angle ( $\theta$ ) are difficult to reconstruct by both the SD and FD. FD-only reconstruction can accurately reconstruct events with relatively large zenith angles ( $\sim 75^\circ$ ; Abu-Zayyad et al. 2013b). The upper zenith angle limit for SD-only reconstruction is  $\sim 45^\circ$  (Abu-Zayyad et al. 2013a; Ivanov 2012). This cut removes tracks with zenith angles greater than  $55^\circ$ .
7. *Hybrid geometry  $\chi^2$  cut.* This cut requires that the reduced  $\chi^2$  of the hybrid fitting described in Section 3.1.3 must be less than 5 to accept the track. This ensures good

geometry reconstruction when combining the SD and FD geometry information.

8. *Profile  $\chi^2$  cut.* This cut requires that the reduced  $\chi^2$  of the profile fit described in Section 3.1.4 be less than 10. This ensures that the light profile of the track is well observed and the inverse Monte Carlo process sufficiently well simulated the shower as observed by the FD.
9.  *$X_{\max}$  bracketing cut.* Once the shower profile is reconstructed, the atmospheric depth versus angle for observed parts of the shower is known. Using this depth profile, the minimum depth observed,  $X_{\text{low}}$ , and the maximum depth observed,  $X_{\text{high}}$ , as well as  $X_{\max}$ , are calculated. The  $X_{\max}$  bracketing cut requires that the fitted value of  $X_{\max}$  be greater than  $X_{\text{low}}$  and less than  $X_{\text{high}}$  for the track to be accepted. This cut ensures that the turnover from rising shower size before  $X_{\max}$  and the falling size after it are in the field of view of the observing FD. This is required to get a good profile fit. If  $X_{\max}$  is not bracketed, the Gaisser–Hillas profile fit will often fail to accurately measure  $X_{\max}$  and  $N_{\max}$ , which also causes large uncertainty on the primary particle energy.
10. *Energy cut.* This cut is used to ensure that the energy of the shower is not less than  $10^{18.2}$  eV. Showers with energies below this cut are difficult to reconstruct by both the SD and FDs independently owing to low signal-to-noise ratio. The hybrid aperture falls steeply below  $E = 10^{18.5}$  eV. This cut represents the design limit of our detectors operating in hybrid mode.
11. *Weather cut.* Events that occur during bad weather nights are rejected. Weather is monitored by operators in the field and logged hourly. Operators record the state of cloud coverage in the four cardinal directions, as well as the amount of overhead coverage and cloud thickness as judged by eye. These weather codes are used to categorize nights into “excellent,” “good,” or “bad” weather nights. This analysis uses nights deemed “excellent” or “good.” Nights that are recorded as bad are rejected from the data. To ensure consistency among similar analyses, the same weather cut criteria as used by the FD monocular spectrum analysis (Abu-Zayyad et al. 2013b) are also used here.

Hybrid monocular events that pass all of the analysis cuts are accepted as part of the final event set. Stereo hybrid events must go through one more selection step before final acceptance. A stereo event is independently reconstructed using the data as observed by the BR FD and as observed by the LR FD. The cuts above are applied separately to the BR and LR reconstructed hybrid event information. If only one of the two site’s data pass the cuts, then the hybrid event data are accepted using that site’s reconstruction parameters. If both sites’ data pass the cuts, then the site’s data with the profile reduced  $\chi^2$  closest to 1 are accepted. The same set of cuts and the same selection procedures are used on data and Monte Carlo data.

After the cuts have been applied to the data, 3330 events remain. The distribution of fully reconstructed and accepted data events is as follows:

3330	hybrid accepted events
1743	BR events (mono and stereo)
1587	LR events (mono and stereo)
1676	BR events (mono)
1504	LR events (mono)
150	BR + LR stereo events

#### 4.1. Data/Monte Carlo Comparison

Monte Carlo is stored in the same format and analyzed in the same manner as data. This allows us to understand the acceptance of our detector by detailed simulations and to perform data/Monte Carlo comparisons to test the agreement of data with different composition models and mixtures. Figures 7 and 8 compare data and Monte Carlo of the parameters used by the analysis cuts for four different primary species: protons, helium, nitrogen, and iron. All of the Monte Carlo data used in this analysis are generated in CORSIKA using the QGSJet II-04 (Ostapchenko 2011) hadronic model. For all plots, the accepted energy range is  $E \geq 10^{18.2}$  eV, and all Monte Carlo simulated data histograms are normalized by area to the area of the data histogram. Each Monte Carlo data set represents 6 yr of TA operations with about 10 times the statistics collected in the data over that same period. The TA Monte Carlo mimics the real operating conditions of the FDs and SDs, by using time-dependent databases of the real operating conditions in the field, such as pedestals, atmosphere, and tube gains. This information is used in both reconstruction and event simulation. As the figures show, the simulation mimics the observed data well between the four chosen models.

To measure the bias and resolution of our detector for a given observable parameter  $X$ , for all reconstructed Monte Carlo events we histogram the difference  $X_{\text{recon}} - X_{\text{true}}$ , where  $X_{\text{recon}}$  is the reconstructed value of the parameter and  $X_{\text{true}}$  is the true value of the parameter. The parameter bias is the sample mean of this distribution, and the resolution is the sample standard deviation. Table 1 shows the measured bias and resolution of this analysis for four primary species for all reconstructed Monte Carlo events with  $E_{\text{true}} \geq 10^{18.2}$  eV. The table shows that the reconstruction biases for  $X_{\text{max}}$  are very small, about  $-1 \text{ g cm}^{-2}$  for protons and  $-4 \text{ g cm}^{-2}$  for iron, both of which are much smaller than the  $X_{\text{max}}$  resolutions of 17 and 13  $\text{g cm}^{-2}$ , respectively. Energy bias is less than 2% for protons and  $-6.5\%$  for iron. We expect a larger energy bias for iron because when the shower energy is computed the missing energy correction assumes a proton primary (see Section 3.1.4). In all cases the energy resolution is less than 6% for the four primary species shown. Angular resolution and bias for the geometric parameters are acceptably small in all cases, less than a degree, which is expected for hybrid reconstruction. Bias and resolution of the shower impact parameter are of order 0.1% of the average observed distance of  $R_p$ . The reconstruction accuracy of  $X_{\text{core}}$  and  $Y_{\text{core}}$  and the  $x$  and  $y$  components of the shower core location on the ground are also very good.

#### 4.2. $X_{\text{max}}$ Biases

$X_{\text{max}}$  bias in our simulation comes in two parts: bias due to detector acceptance and bias due to reconstruction. Reconstruction bias is bias that is affected by operating condition of the detector, selection of cuts, composition and hadronic model dependence, and proper modeling of the detector and air shower physics in the Monte Carlo. Acceptance bias is affected by physical detector design and detector response, such as choice of triggering algorithm.

Acceptance bias predominantly affects the deeply penetrating tail of the  $X_{\text{max}}$  distribution. This is because there is an upper bound to the maximum depth to which an air shower can be observed due to limited atmospheric mass overburden, which is part of the detector design (placement on Earth's surface). For very deeply penetrating primaries, the ability to reconstruct via

fluorescence observation is limited by the following scenarios: (1) the air shower track has small zenith angle and  $X_{\text{max}}$  occurs at the ground level or below, or (2) the air shower track achieves shower maximum in air but therefore has a very large zenith angle. The result is that as a function of energy,  $X_{\text{max}}$  acceptance bias in TA is seen as a systematic shift of the mean of the  $X_{\text{max}}$  distribution to smaller depths and narrowing in the width (rms) of the  $X_{\text{max}}$  distribution in each energy bin. This effect is shown explicitly in the Monte Carlo distributions in Figure 9.

This effect is dependent on the mass and energy of the primary particle; light particles penetrate more deeply and shower maximum occurs at deeper depths on average with increasing energy and decreasing primary mass. It is also dependent on the physics of UHECR hadronic interactions, which are not known for our energy range of interest. This appears as model dependence through our choice of hadronic generator in CORSIKA simulations. More recent hadronic models tuned to LHC results, such as QGSJet II-04 (Ostapchenko 2011) and EPOS LHC (Pierog et al. 2015), generate events that penetrate more deeply on average than older models, such as QGSJet II-03 (Ostapchenko 2007).

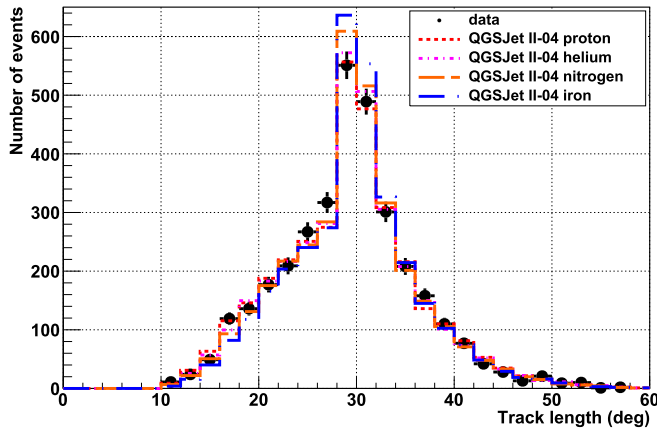
The sum of acceptance bias and reconstruction bias is called total bias, and it is important for us to understand because it appears as a systematic shift in the final reconstructed  $X_{\text{max}}$  distribution compared to the true generated  $X_{\text{max}}$ . Given that there is some combination of hadronic model and mixture of elements that represents the true distribution of  $X_{\text{max}}$  that is impinging on Earth's atmosphere, detectors with acceptance bias will never be able to fully reproduce the true distribution in nature simply by plotting the distribution of reconstructed events. Acceptance bias will distort the observed distribution, because information about the distribution is simply lost and it will typically appear as if it is the result of a heavier mixture of elements. To correct for this type of bias, one can attempt unfolding of the data, or resort to even more restrictive sets of cuts such as done by the Auger experiment (Aab et al. 2014).

An alternate method to understand measured composition is to simply use Monte Carlo to simulate biases incurred as a result of detector acceptance and compare the measured  $X_{\text{max}}$  distribution to the biased, simulated one. This is the method chosen by TA. It is important to understand how a measurement deals with this issue of acceptance bias before attempting to compare composition results between different experiments.

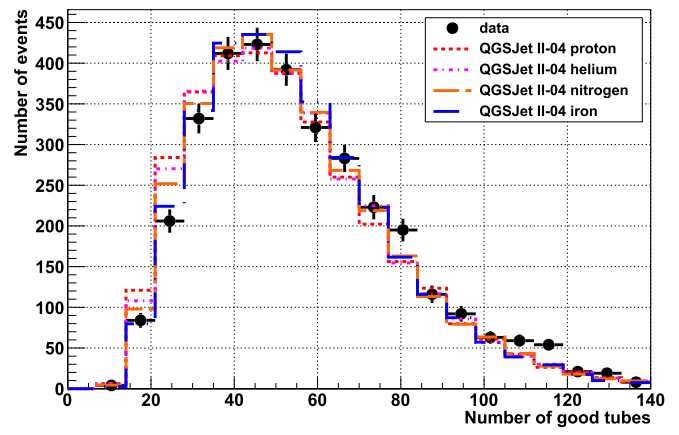
#### 4.3. TA $X_{\text{max}}$ Data

TA hybrid  $X_{\text{max}}$  data are binned by energy into 11 energy bins. Below  $10^{19}$  eV, there are sufficient statistics to use 0.1 decade wide energy bins, to provide  $\approx 100$  events per bin. Above  $10^{19}$  eV the bins are widened to try to capture more events. Figures 10 and 11 show the  $X_{\text{max}}$  distributions measured in this analysis. The distributions for reconstructed QGSJet II-04 Monte Carlo are shown as well. For each energy bin, the histogram of each individual species is normalized by area to the area of the data histogram. In a given energy bin, lighter elements have larger  $\langle X_{\text{max}} \rangle$  as expected because of the relationship  $X_{\text{max}} \propto \ln(E_0/A)$  discussed in Section 1.

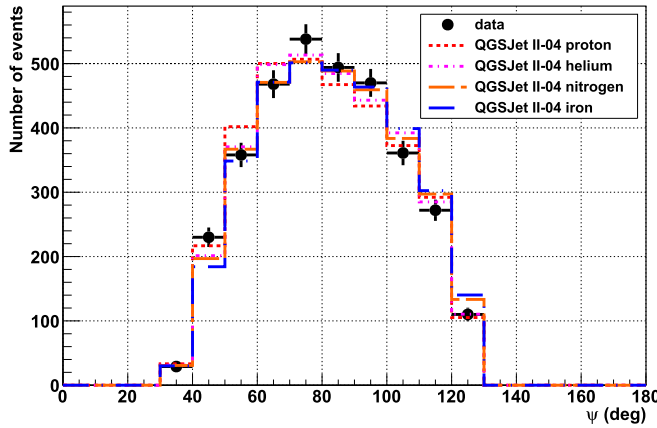
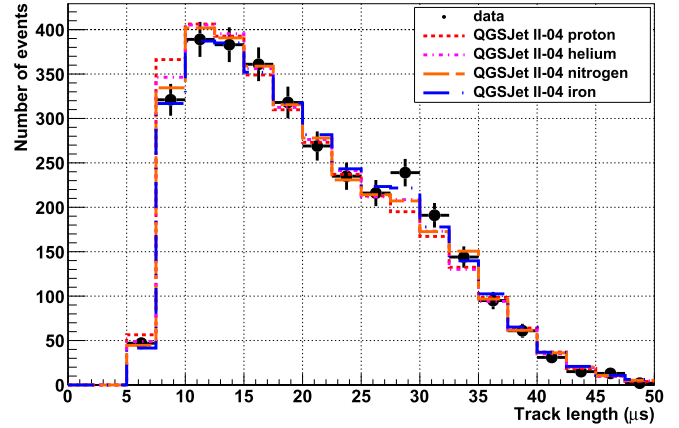
Each figure shows that the means and standard deviations of the distributions of the simulated elements decrease with increasing mass as we expect. We can use them to compare to the data to determine which pure element drawn from the QGSJet II-04 model most resembles the data and which



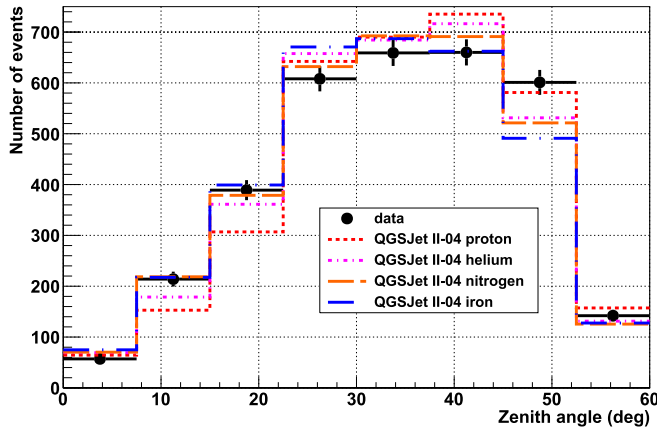
(a) Data/Monte Carlo track length comparison.



(b) Data/Monte Carlo number of good tubes comparison.

(c) Data/Monte Carlo  $\psi$  angle comparison.

(d) Data/Monte Carlo track time extent comparison.



(e) Data/Monte Carlo zenith angle comparison.

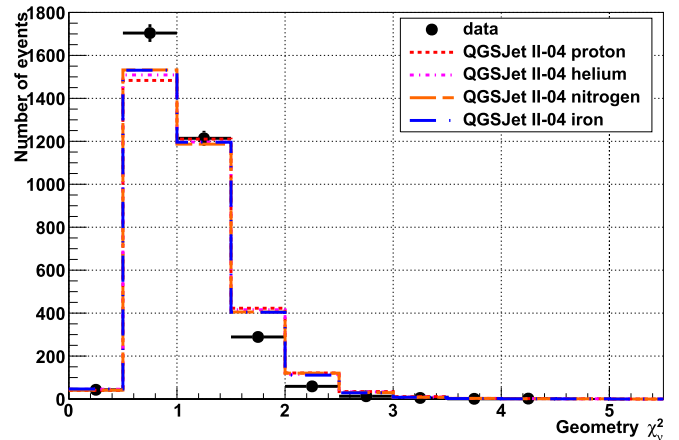
(f) Data/Monte Carlo geometry fit  $\chi^2_\nu$  comparison.

Figure 7. Data/Monte Carlo plots I.

elements may be excluded. Such a comparison does not imply that we believe cosmic rays in nature to be composed of a single chemical element in any given energy bin we have observed. In a future paper, we will investigate the compatibility of TA data with mixtures of elements. In this current work, we only compare TA data to pure CORSIKA elements.

UHECR composition measurements typically utilize the first and second moments of  $X_{\max}$  distributions of data and Monte

Carlo to compare observed results to those expected for the models under investigation. These individual quantities are too limited to fully understand the details of  $X_{\max}$  distributions, particularly for light elements that exhibit prominent non-Gaussian tails. While examining  $\langle X_{\max} \rangle$  and  $\sigma(X_{\max})$  as a function of energy is still useful, especially to place an experiment in historical context with older measurements, utilizing more powerful statistical techniques is more appropriate given that more powerful computers now exist to make

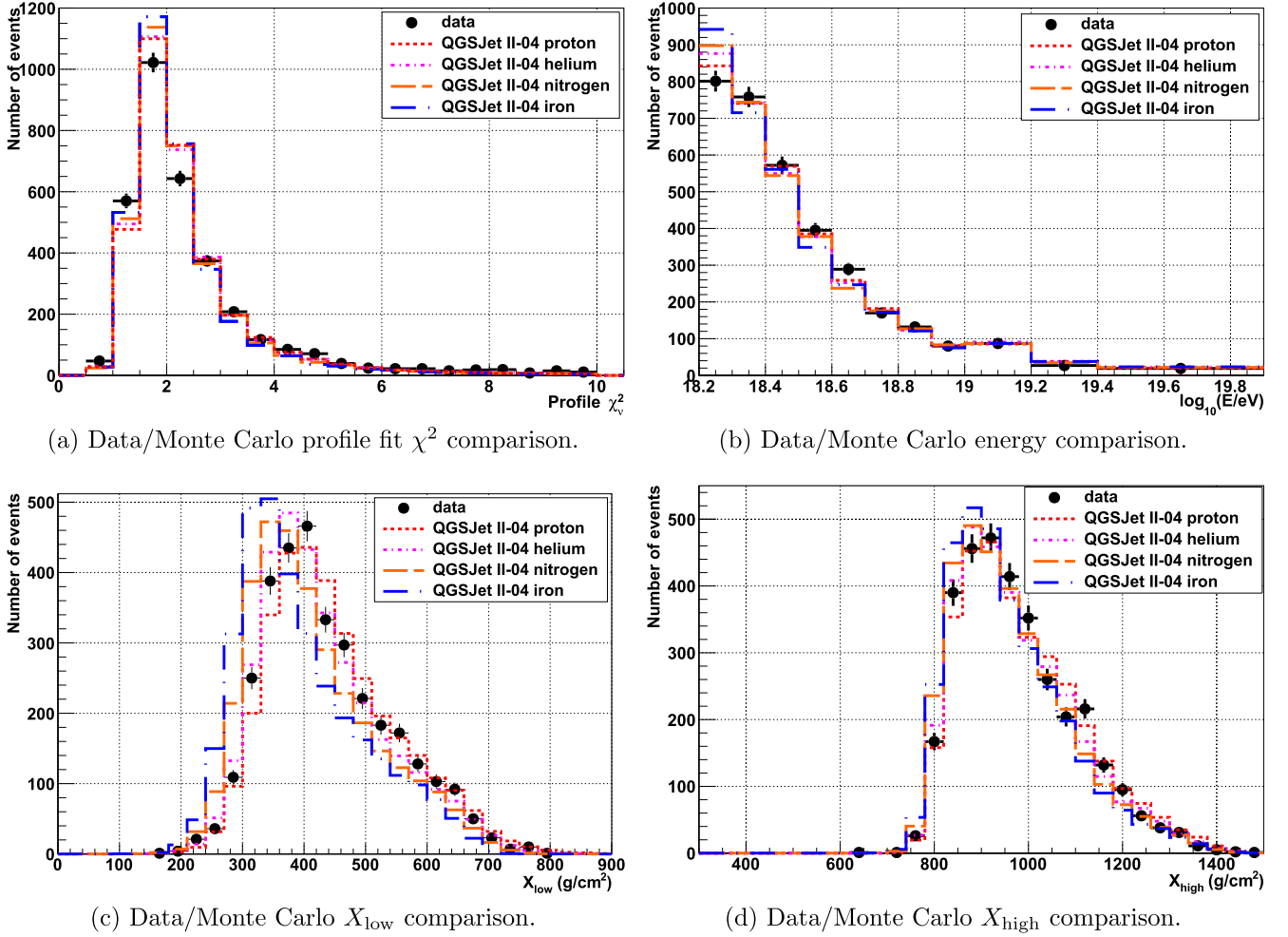


Figure 8. Data/Monte Carlo plots II.

**Table 1**  
Bias and Resolution of BR/LR Hybrid  $X_{\text{max}}$  Analysis Reconstruction

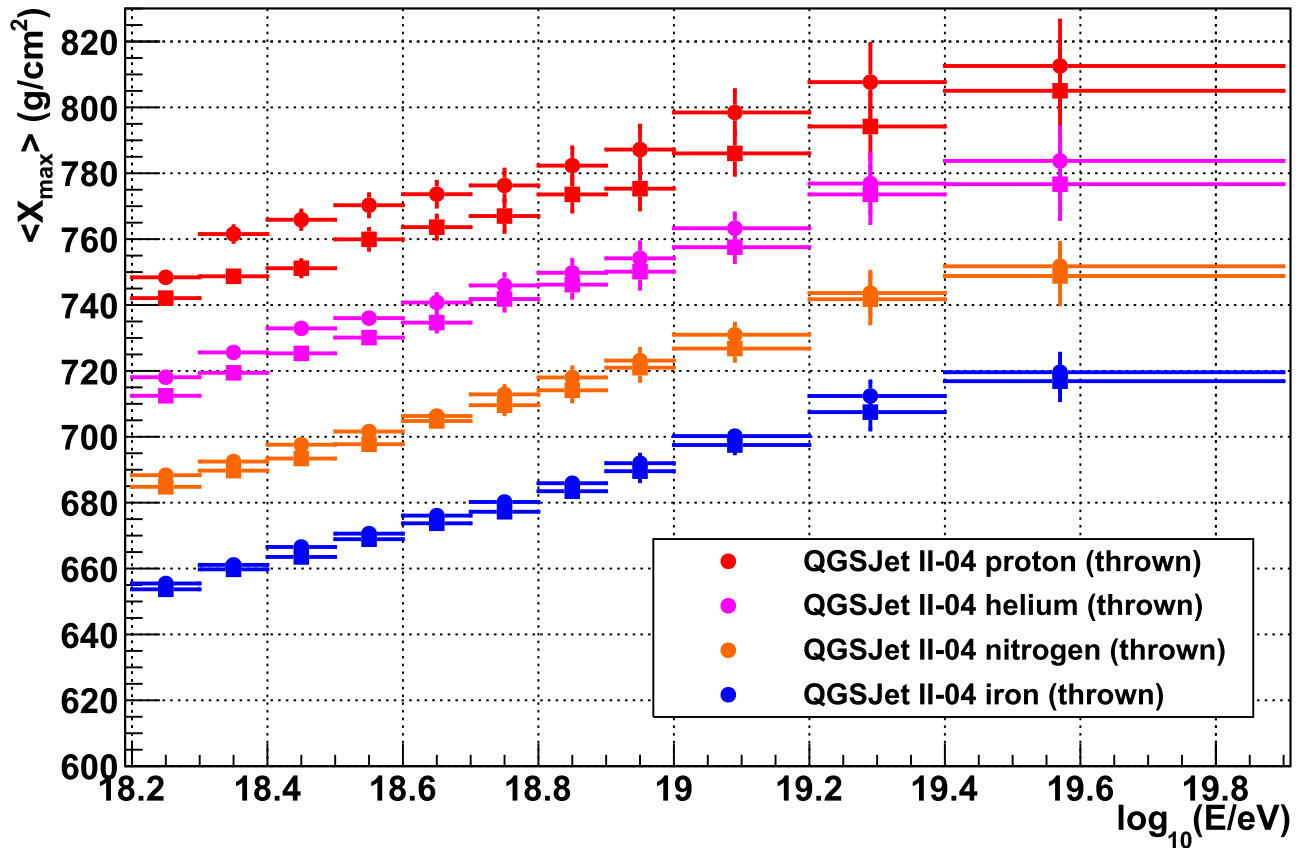
	Proton		Helium		Nitrogen		Iron	
	Bias	Res.	Bias	Res.	Bias	Res.	Bias	Res.
$X_{\text{max}} (\text{g cm}^{-2})$	-1.1	17.2	-3.3	15.7	-3.8	14.2	-3.8	13.2
$\sigma(X_{\text{max}}) (\text{g cm}^{-2})$	5.4	0.3	3.6	0.2	3.5	0.2	3.4	0.2
Energy (%)	1.7	5.7	-1.1	5.1	-3.5	4.4	-6.5	3.9
$\theta$ (deg)	0.014	0.377	0.006	0.364	0.0005	0.3553	-0.003	0.344
$\phi$ (deg)	-0.020	0.410	-0.017	0.399	-0.015	0.389	-0.015	0.374
$\psi$ (deg)	0.074	0.397	0.088	0.385	0.112	0.375	0.135	0.356
$R_p$ (m)	18.9	39.8	20.1	39.0	21.1	38.9	22.3	37.7
$X_{\text{core}}$ (m)	-3.6	49.8	-3.4	49.9	-3.3	50.6	-3.7	51.9
$Y_{\text{core}}$ (m)	8.7	42.9	8.1	42.3	8.3	42.9	8.4	43.3

**Note.** All primary species are generated using the QGSJet II-04 hadronic model.

these calculations much more practical. For this reason we will make our primary visual comparisons of data and Monte Carlo by simultaneously examining  $\langle X_{\text{max}} \rangle$  and  $\sigma(X_{\text{max}})$ , which is a more powerful way to understand the relationship of the data and Monte Carlo.

Recalling the discussion of the relationship of mass and energy to the mean and width of the  $X_{\text{max}}$  distribution from Section 1, we can examine the signature of a given element as observed by 8.5 yr of exposure in TA in hybrid mode by

simultaneously measuring the distributions of  $\langle X_{\text{max}} \rangle$  and  $\sigma(X_{\text{max}})$ . Light elements will have both larger  $\langle X_{\text{max}} \rangle$  and larger  $\sigma(X_{\text{max}})$  distributions, because of the larger fluctuations first in the interaction in the atmosphere and the subsequent shower development. We will also be able to see the effect of TA's acceptance on the reconstructed distributions and compare them to the observed data. To do this, the reconstructed  $X_{\text{max}}$  distribution for a single element, such as QGSJet II-04 protons shown in Figures 10 and 11, is sampled

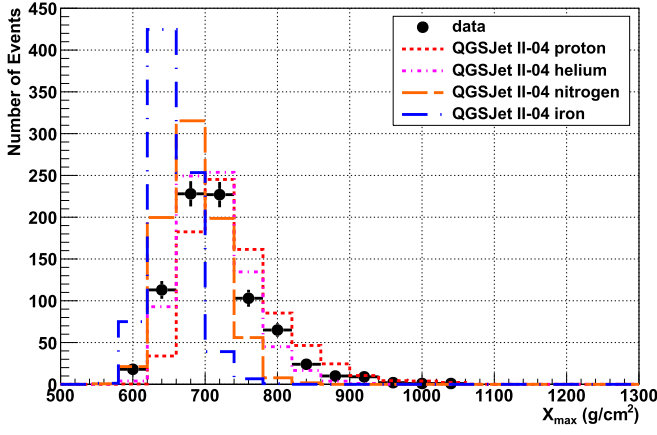
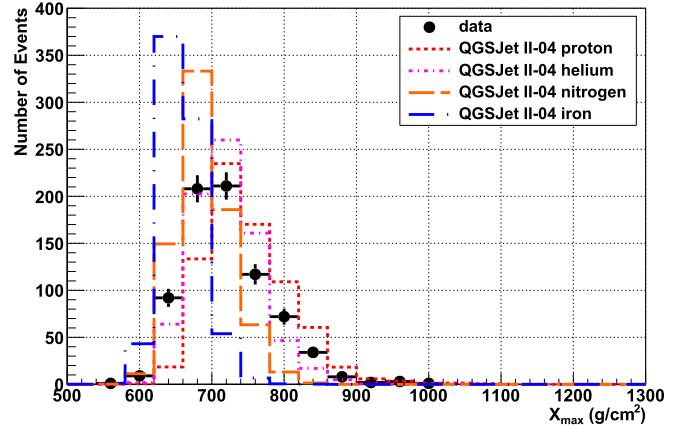
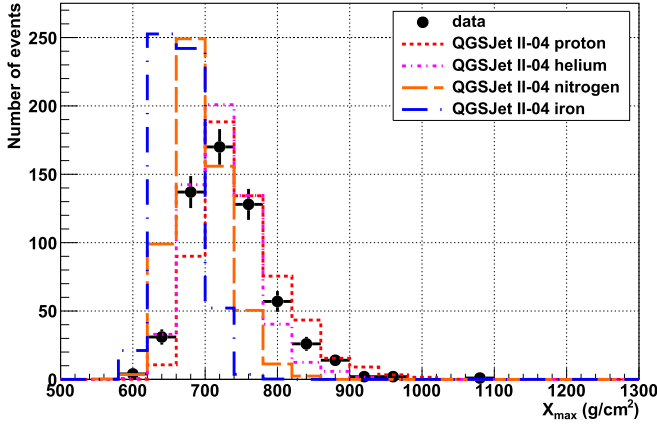
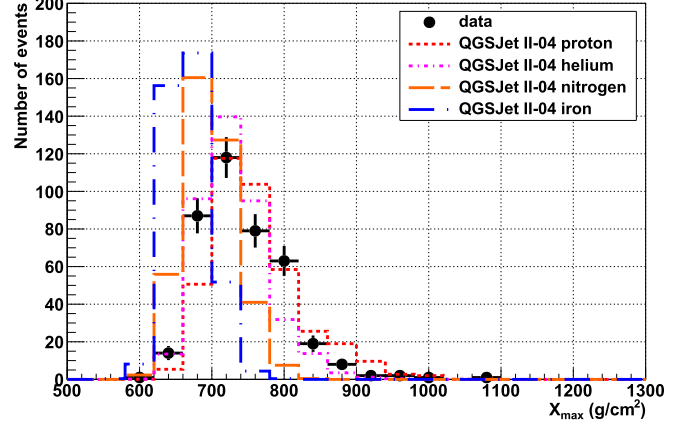
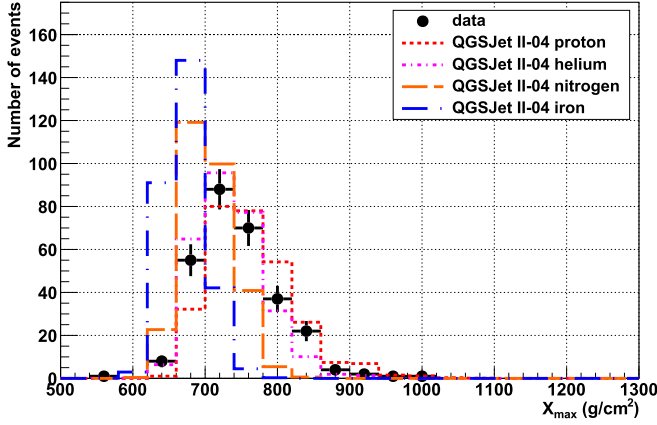
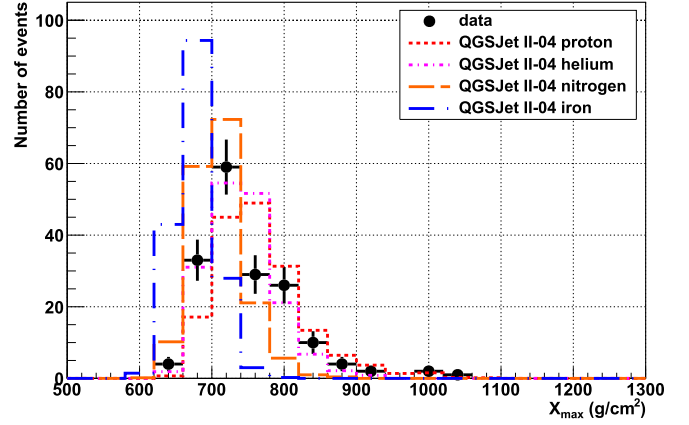


**Figure 9.** QGSJet II-04 Monte Carlo  $\langle X_{\max} \rangle$  used in this analysis. Circles represent  $\langle X_{\max} \rangle$  of the true distributions prior to any reconstruction. Squares represent the  $\langle X_{\max} \rangle$  after reconstruction. The difference in an energy bin between the thrown  $\langle X_{\max} \rangle$  and the reconstructed  $\langle X_{\max} \rangle$  (acceptance bias) is caused mainly by detector acceptance, with a small contribution from reconstruction biases as well. Light elements have larger acceptance bias because the primary effect of this type of bias is to cause the loss of very deeply penetrating events in the tails of the distributions. Uncertainties in the means are calculated using the equivalent exposure in the data.

according to the same number of events recorded in the data for a given energy bin.  $\langle X_{\max} \rangle$  and  $\sigma(X_{\max})$  of the energy are calculated and recorded for this sample. This procedure is then repeated 5000 times. The distribution of  $\langle X_{\max} \rangle$  and  $\sigma(X_{\max})$  is used to calculate the 68.3%, 90%, and 95% confidence intervals. The entire  $\langle X_{\max} \rangle$  and  $\sigma(X_{\max})$  calculated by this method are then plotted as a two-dimensional distribution along with the computed confidence intervals. This procedure is repeated for the other three chemical elements used in the analysis. Figures 12 and 13 show this measurement for all observed energy bins. The  $\langle X_{\max} \rangle$  and  $\sigma(X_{\max})$  of the data observed in each energy bin are also recorded as a single red star. Additionally, the statistical, systematic, and combined statistical and systematic error bounds are marked around the data.

Figures 12(a) and (b), corresponding to the energy range  $10^{18.2}-10^{18.4}$  eV, show that each of the four modeled chemical elements has clear separation and is individually resolvable by TA in those energy bins given our acceptance and statistics in the data (801 and 758 events, respectively). The  $\sigma(X_{\max})$  of the data resembles QGSJet II-04 protons, but the  $\langle X_{\max} \rangle$  of the data is lower by about our systematic uncertainty in those energy bins. Note that we do not account for systematic uncertainties in the QGSJet II-04 model, which will be discussed in Section 5. Figures 12(c)–(f), corresponding to the energy range  $10^{18.4}-10^{18.8}$  eV, show that  $\sigma(X_{\max})$  of the data continues to resemble QGSJet II-04 protons, and  $\langle X_{\max} \rangle$  of the data falls

within the 68.3% confidence interval of the proton distributions within the data’s systematic uncertainty. We also notice an effect of decreasing statistics in the data, by observing the increase in the size of the confidence intervals of the individual Monte Carlo elements. Figure 13(a), corresponding to the energy range  $10^{18.8}-10^{18.9}$  eV, shows a relatively large downward fluctuation in  $\sigma(X_{\max})$  of the data. In this energy bin, the 68.3% confidence intervals of QGSJet II-04 proton and helium both fall within the bounds of the systematic uncertainty of the data. In Figure 13(b), corresponding to the energy range  $10^{18.9}-10^{19.0}$  eV,  $\sigma(X_{\max})$  of the data fluctuates up from the previous energy bin and the systematic error bounds of the data fall within the 68.3% confidence interval of protons. In this energy, because of the small statistics in the data (80 events), the largest confidence intervals of proton and helium begin to overlap. This indicates that given TA’s current exposure in this energy bin, we are losing our ability to make precise statements about the signature of pure light chemical elements. The ability to distinguish nitrogen from iron or nitrogen from helium or protons remains. In Figure 13(c), corresponding to the energy range  $10^{19.0}-10^{19.2}$  eV, the 95% confidence intervals of proton and helium are once again separated, but here we have doubled the size of the energy bin in an attempt to enable us to still make a reasonably good measurement of  $\langle X_{\max} \rangle$  and  $\sigma(X_{\max})$  in the data. The 68.3% confidence intervals of both proton and helium fall within the bounds of the systematic uncertainty of the data. Figures 13(d) and (e), corresponding to the energy


 (a)  $18.2 \leq \log_{10}(E/eV) < 18.3$ 

 (b)  $18.3 \leq \log_{10}(E/eV) < 18.4$ 

 (c)  $18.4 \leq \log_{10}(E/eV) < 18.5$ 

 (d)  $18.5 \leq \log_{10}(E/eV) < 18.6$ 

 (e)  $18.6 \leq \log_{10}(E/eV) < 18.7$ 

 (f)  $18.7 \leq \log_{10}(E/eV) < 18.8$ 

**Figure 10.**  $X_{\max}$  distributions in energy bins for  $18.2 \leq \log_{10}(E/eV) < 18.8$ . The data are compared to Monte Carlo  $X_{\max}$  distributions generated using the QGSJet II-04 hadronic model for four primary elements.

range  $10^{19.2} - 10^{19.9}$  eV, show that TA's ability to resolve individual QGSJet II-04 elements is degraded owing to the overlap of the confidence intervals. According to these figures, when considering only the joint distributions of  $\langle X_{\max} \rangle$  and  $\sigma(X_{\max})$ , within the data's systematic uncertainty the data may resemble QGSJet II-04 proton, helium, or nitrogen.

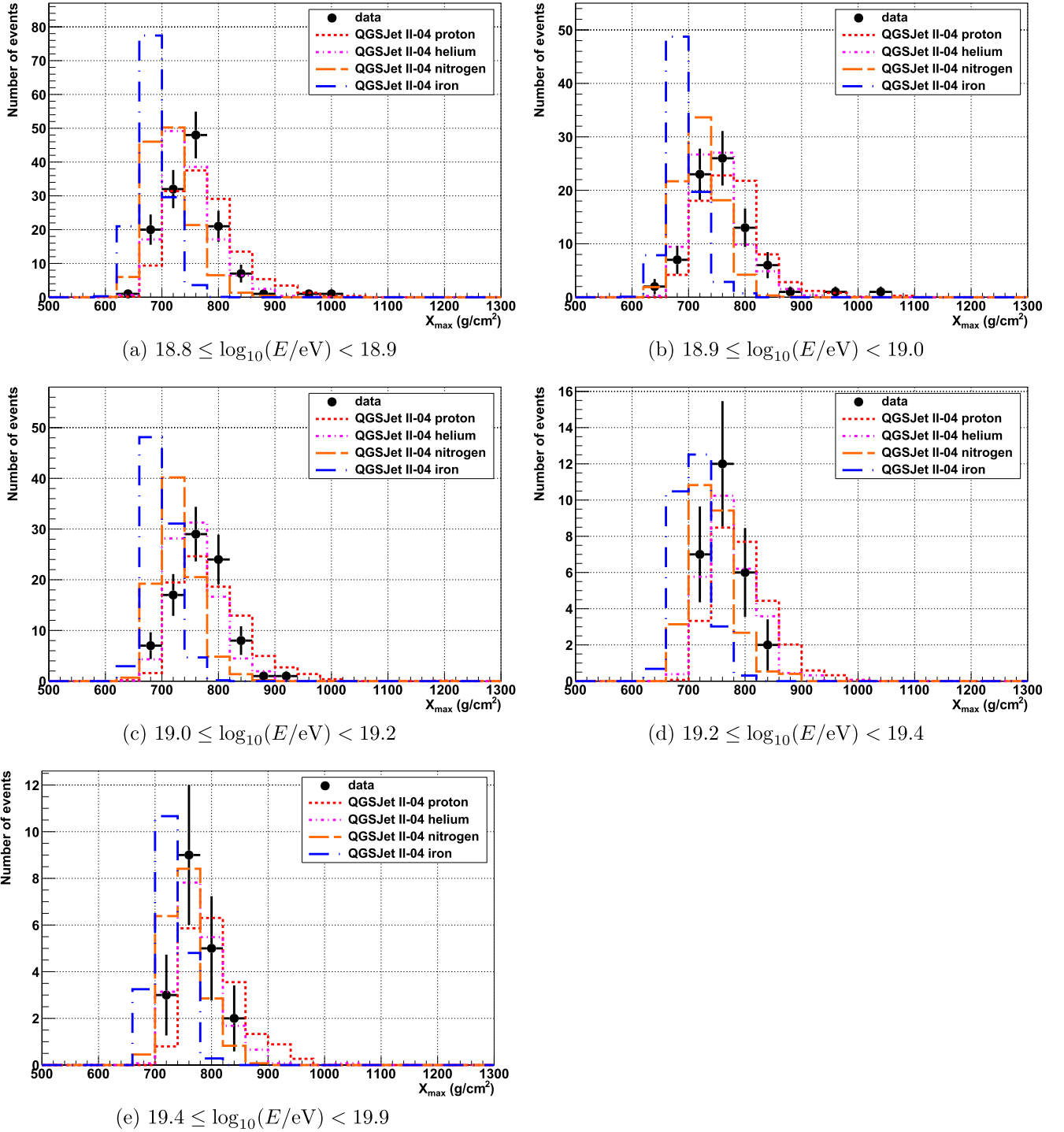
Figure 14 shows only the means of the  $X_{\max}$  distributions presented in Figures 10–13 of the observed data, as well as reconstructed Monte Carlo for four primary species. The gray

band around the data points indicates the systematic uncertainty in  $\langle X_{\max} \rangle$  of  $17.4 \text{ g cm}^{-2}$  estimated for this analysis.

#### 4.4. Systematic Uncertainties

Systematic uncertainties in the  $\langle X_{\max} \rangle$  and  $\sigma(X_{\max})$  are evaluated for four sources: detector modeling, atmosphere, fluorescence yield, and the reconstruction algorithm.

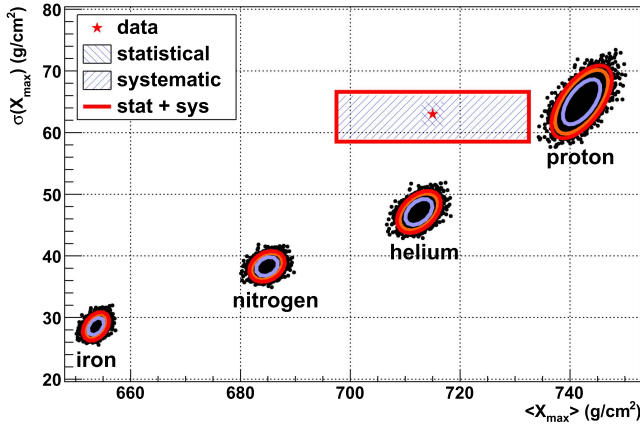
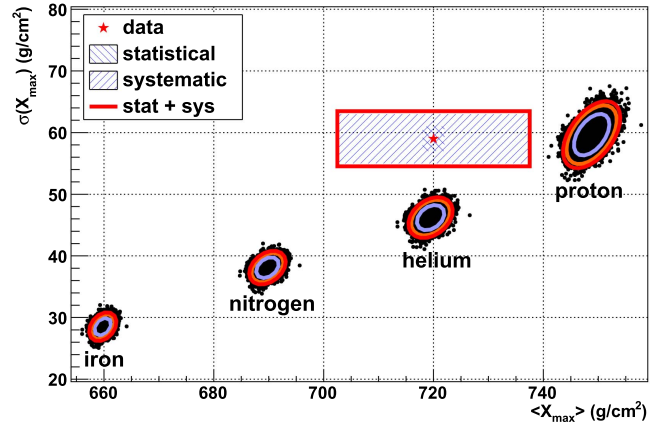
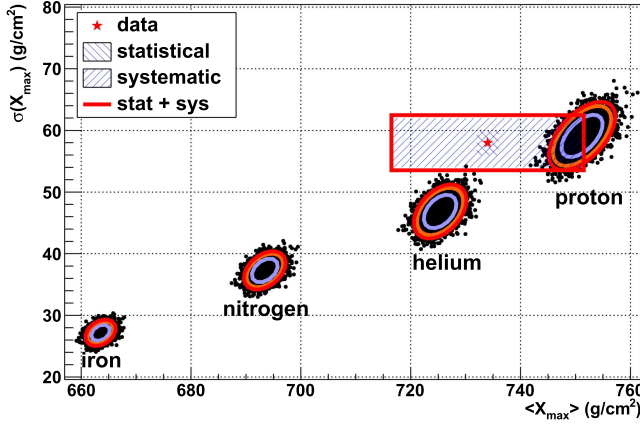
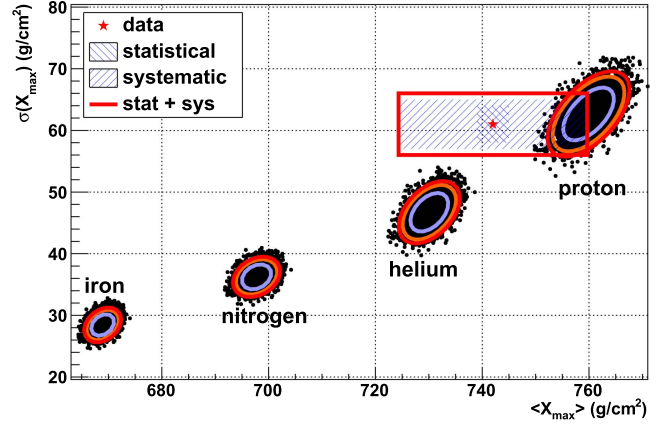
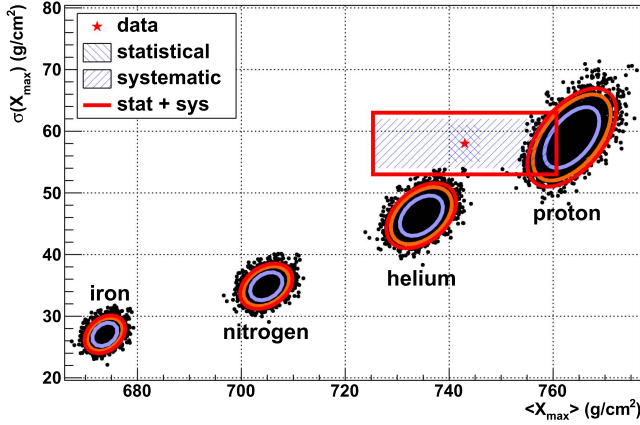
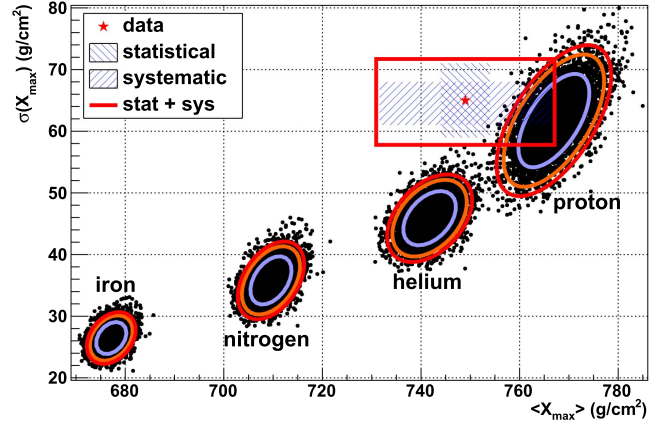




**Figure 11.**  $X_{\max}$  distributions in energy bins for  $18.8 \leq \log_{10}(E/\text{eV}) < 19.9$ . The data are compared to Monte Carlo  $X_{\max}$  distributions generated using the QGSJet II-04 hadronic model for four primary elements.

The pointing accuracies of the phototubes ( $\pm 0.05$  deg) and the relative timing between FD and SD (240ns) dominate the detector effects. These effects give  $\pm 3.3 \text{ g cm}^{-2}$  and  $\pm 3.8 \text{ g cm}^{-2}$  uncertainty in  $\langle X_{\max} \rangle$  and  $\pm 1.7 \text{ g cm}^{-2}$  and  $\pm 4.0 \text{ g cm}^{-2}$  in  $\sigma(X_{\max})$ , respectively. Some events are detected by both FD stations (BR and LR), and the  $X_{\max}$  differences of such stereo events can also be used to estimate the detector effect. We found that the BR-LR difference is smaller than  $10 \text{ g cm}^{-2}$  in  $\langle X_{\max} \rangle$  for events with energies greater than  $10^{18.2} \text{ eV}$ .

The atmospheric effect is dominated by the amount of aerosols. We have  $\sim 15\%$  uncertainty of aerosols in terms of vertical aerosol optical depth (VAOD), which gives a shift in  $\langle X_{\max} \rangle$  of  $3.4 \text{ g cm}^{-2}$ . Variations in atmospheric aerosols potentially have a large effect on  $\sigma(X_{\max})$ . Aerosols are measured every 30 minutes by the CLF (Tomida et al. 2013). If we compare how  $\sigma(X_{\max})$  varies using the VAOD data measured by the CLF, the effect on  $\sigma(X_{\max})$  is found to be  $18.9 \text{ g cm}^{-2}$ . Another effect comes from the atmospheric profile, i.e., the pressure and density of the atmosphere as

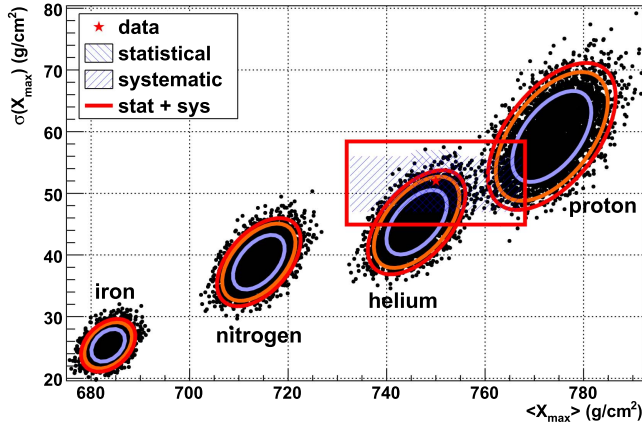
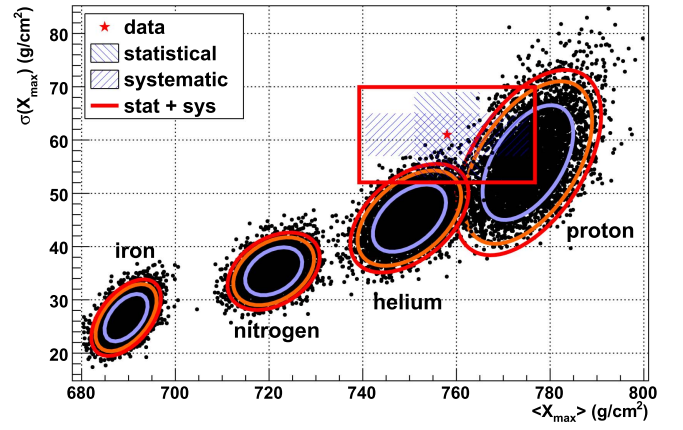
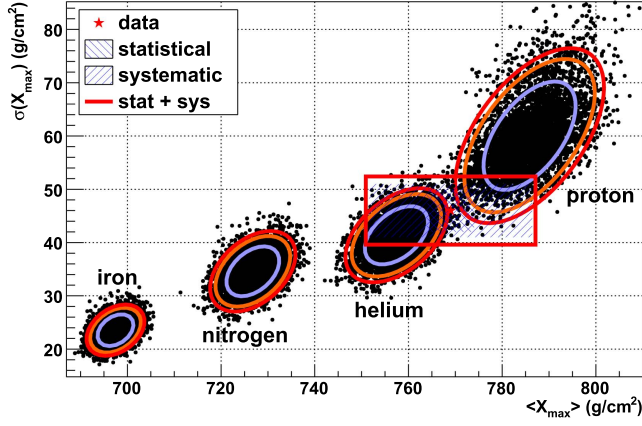
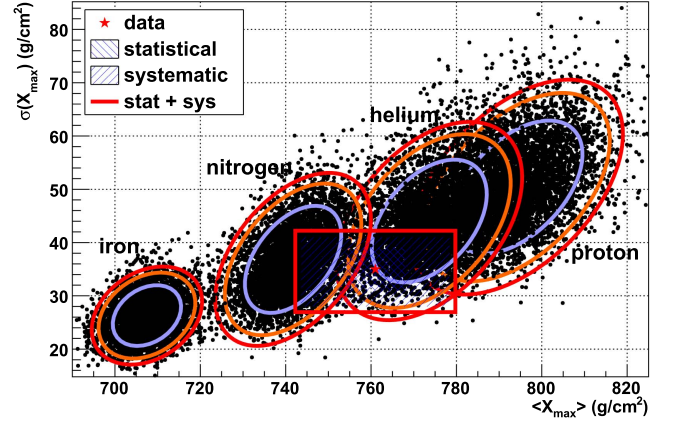
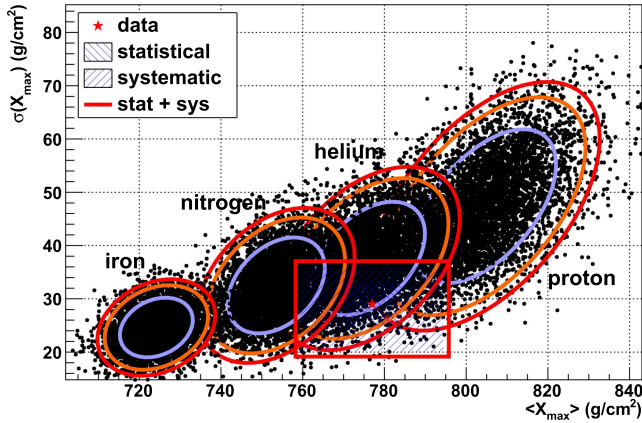
(a)  $18.2 \leq \log_{10}(E/eV) < 18.3$ (b)  $18.3 \leq \log_{10}(E/eV) < 18.4$ (c)  $18.4 \leq \log_{10}(E/eV) < 18.5$ (d)  $18.5 \leq \log_{10}(E/eV) < 18.6$ (e)  $18.6 \leq \log_{10}(E/eV) < 18.7$ (f)  $18.7 \leq \log_{10}(E/eV) < 18.8$ 

**Figure 12.** Measurements of data and QGSJet II-04 Monte Carlo  $\langle X_{\max} \rangle$  and  $\sigma(X_{\max})$  in energy bins for  $18.2 \leq \log_{10}(E/eV) < 18.8$ . Each Monte Carlo chemical element shows the 68.3% (blue ellipse), 90% (orange ellipse), and 95% (red ellipse) confidence intervals.

functions of height. When we reconstruct data using an atmospheric database that uses NOAA National Weather Service radiosonde data instead of the Global Data Assimilation System (GDAS; Laboratory 2004), the effects on  $\langle X_{\max} \rangle$  and  $\sigma(X_{\max})$  are found to be 5.9 and 7.4  $\text{g cm}^{-2}$ , respectively. The large difference between the systematic uncertainty of the  $\langle X_{\max} \rangle$  and  $\sigma(X_{\max})$  (3.4 and 18.9  $\text{g cm}^{-2}$ , respectively) is due to the large variation of VAOD relative to the mean VAOD value of 0.04. The systematic uncertainty of  $\langle X_{\max} \rangle$  is found by

reconstructing each event using the measured VAOD and  $\text{VAOD} = 0.04$  and then calculating the mean shift in  $X_{\max}$ ,  $\langle X_{\max}^{\text{VAOD}} - X_{\max}^{0.04} \rangle$ , where  $X_{\max}^{\text{VAOD}}$  is the  $X_{\max}$  found using the measured VAOD and  $X_{\max}^{0.04}$  is the  $X_{\max}$  found using an assumed VAOD of 0.04. The systematic uncertainty in  $\sigma(X_{\max})$  is calculated using the width of the distribution of those differences.

We use a fluorescence yield model that uses the absolute yield measurement by Kakimoto et al. (1996) and the

(a)  $18.8 \leq \log_{10}(E/eV) < 18.9$ (b)  $18.9 \leq \log_{10}(E/eV) < 19.0$ (c)  $19.0 \leq \log_{10}(E/eV) < 19.2$ (d)  $19.2 \leq \log_{10}(E/eV) < 19.4$ (e)  $19.4 \leq \log_{10}(E/eV) < 19.9$ 

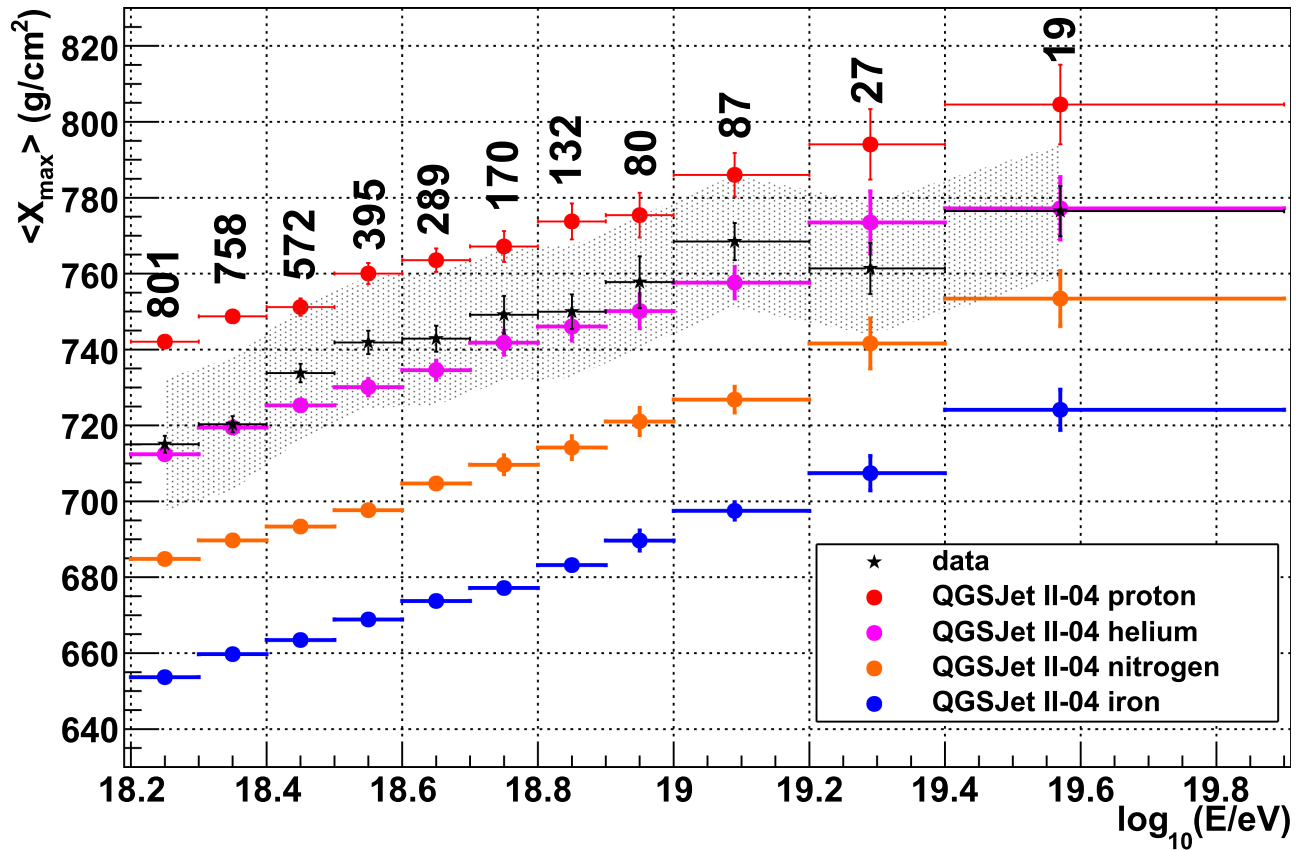
**Figure 13.** Measurements of data and QGSJet II-04 Monte Carlo  $\langle X_{\max} \rangle$  and  $\sigma(X_{\max})$  in energy bins for  $18.8 \leq \log_{10}(E/eV) < 19.9$ . Each Monte Carlo chemical element shows the 68.3% (blue ellipse), 90% (orange ellipse), and 95% (red ellipse) confidence intervals.

fluorescence spectral measurement by the FLASH experiment (Abbasi et al. 2008b). A  $\sim 5\text{--}6 \text{ g cm}^{-2}$  effect is expected if we use different fluorescence modeling. For example, we found a  $+5.6 \text{ g cm}^{-2}$  shift in  $\langle X_{\max} \rangle$  and a  $3.7 \text{ g cm}^{-2}$  effect in  $\sigma(X_{\max})$  when we use the model based on the measurements by the AirFly experiment (Ave et al. 2007, 2013) on the absolute yield, the spectrum, and the atmospheric parameter dependencies.

A systematic effect in  $X_{\max}$  also comes from the reconstruction program used in the analysis. We have two reconstruction

programs independently developed in TA for the same data. The reconstruction bias can be estimated by an event-by-event comparison of  $X_{\max}$  values calculated by these two separate reconstruction procedures, and this is smaller than  $4.1 \text{ g cm}^{-2}$  for events with  $E > 10^{18.2} \text{ eV}$ .

Some of these contributions are not fully independent. For example, the uncertainties evaluated from the BR–LR difference and the comparison of different analysis programs could be correlated. In the calculation of the total systematic uncertainty, we use a linear sum of these two sources of



**Figure 14.** Mean  $X_{\max}$  as a function of energy as observed by TA in BR/LR hybrid mode over 8.5 yr of data collection. The numbers above the data points indicate the number of events observed. The gray band is the systematic uncertainty of this analysis. Reconstructed Monte Carlo data of four different primary species generated using the QGSJet II-04 hadronic model are shown for comparison.

uncertainty ( $14.1 \text{ g cm}^{-2}$ ) as a conservative estimate. Other sources are added in quadrature, and we find the total systematic uncertainty in  $\langle X_{\max} \rangle$  to be  $17.4 \text{ g cm}^{-2}$ . The results are summarized in Table 2.

The systematic uncertainties of  $\sigma(X_{\max})$  from the sources discussed above are also evaluated and given in Table 3. Adding in quadrature, we obtain  $21.1 \text{ g cm}^{-2}$ .

As seen in Figure 14, within systematic uncertainties,  $\langle X_{\max} \rangle$  of the data is in agreement with QGSJet II-04 protons and helium for nearly all energy bins. There is clear separation between the region of systematic uncertainty and heavier elements such as nitrogen and iron. In the last two energy bins there is some overlap between the systematic uncertainty region of the data and the nitrogen, but statistics in the data there are very poor. Care must be taken in interpreting Figure 14, since  $\langle X_{\max} \rangle$  by itself is not a robust enough measure to fully draw conclusions about UHECR composition. When comparing  $\langle X_{\max} \rangle$  of data to Monte Carlo, in addition to detector resolution and systematic uncertainties in the data that may hinder resolving the different elements with relatively similar masses, the issue of systematic uncertainties in the hadronic model used to generate the Monte Carlo must also be recognized. This will be discussed in Section 5. Referring back to Figures 12 and 13, we can see that although the  $\langle X_{\max} \rangle$  of the data in Figure 14 lies close to QGSJet II-04 helium, the  $\sigma(X_{\max})$  of the data is larger than the helium model allows for energy bins with good data statistics. For this reason, we will test the agreement of data and Monte Carlo not just by comparing  $\langle X_{\max} \rangle$  and  $\sigma(X_{\max})$ , but by using the entire distributions. The elongation rate of the data in Figure 14 found by performing a  $\chi^2$  fit to the

**Table 2**  
The Systematic Uncertainties in  $\langle X_{\max} \rangle$  of TA Hybrid BR/LR Reconstruction

Items	$\Delta\langle X_{\max} \rangle$	Notes
Independent Sources		
Detector	$5.1 \text{ g cm}^{-2}$	Relative timing between FD and SD ( $3.8 \text{ g cm}^{-2}$ ), pointing direction of the telescope ( $3.3 \text{ g cm}^{-2}$ )
Atmosphere	$6.8 \text{ g cm}^{-2}$	Aerosol ( $3.4 \text{ g cm}^{-2}$ ), atmospheric depth ( $5.9 \text{ g cm}^{-2}$ )
Fluorescence yield	$5.6 \text{ g cm}^{-2}$	Difference in yield models
Quadratic sum	$10.2 \text{ g cm}^{-2}$	
Not Fully Independent Sources		
Detector Reconstruction	$10.0 \text{ g cm}^{-2}$	Difference in two FD stations
	$4.1 \text{ g cm}^{-2}$	Difference in reconstructions
Linear sum	$14.1 \text{ g cm}^{-2}$	
Total	$17.4 \text{ g cm}^{-2}$	

data is found to be  $56.8 \pm 5.3 \text{ g cm}^{-2} \text{ decade}^{-1}$ . The  $\chi^2/\text{dof}$  of this fit is 10.67/9. Table 4 summarizes the observed first and second moments of TA's observed  $X_{\max}$  for all energy bins.

## 5. Statistical Hypothesis Tests

### 5.1. Method

If one wishes to draw conclusions about agreement between the data and the models, we should employ a test that measures

**Table 3**  
The Systematic Uncertainties in  $\sigma(X_{\max})$  of TA Hybrid BR/LR Reconstruction

Items	$\Delta\sigma(X_{\max})$	Notes
Detector	4.3 g cm <sup>-2</sup>	Relative timing between FD and SD (1.7 g cm <sup>-2</sup> ), pointing direction of the telescope (4.0 g cm <sup>-2</sup> )
Atmosphere	20.3 g cm <sup>-2</sup>	Aerosol (18.9 g cm <sup>-2</sup> ), atmospheric depth (7.4 g cm <sup>-2</sup> )
Fluorescence yield	3.7 g cm <sup>-2</sup>	Difference in yield models
Quadratic sum	21.1 g cm <sup>-2</sup>	

the agreement of the entire distributions instead of relying on the first and second moments of the  $X_{\max}$  distributions. Comparisons of the means and standard deviations of  $X_{\max}$  distributions are difficult to fully characterize agreement or disagreement because these distributions are naturally skewed. The deep  $X_{\max}$  is problematic for energy bins with low exposure, which can lead to misinterpretation of the results if care is not taken and only the first and second moments of the distributions are considered. To test the compatibility of the data and the Monte Carlo, we use an unbinned maximum likelihood test.

To perform these tests, we fit the Monte Carlo  $X_{\max}$  distributions to a continuous function described by a convolution of a Gaussian with an exponential function (Toderò Peixoto et al. 2013) and then uniformly shift the  $X_{\max}$  distributions of the data within  $\pm 100$  g cm<sup>-2</sup> in 1 g cm<sup>-2</sup> steps, calculate the log likelihood, and record which  $\Delta X_{\max}$  shift gives the best likelihood between data and Monte Carlo. We allow for shifting of the data to account for possible systematic uncertainties in our reconstruction and for uncertainties in the models that we are testing against. An additional benefit of this method is that if the required shift is significantly larger than the combined experimental and theoretical uncertainties, pure elemental composition is strongly disallowed. However, in the current paper we focus on the shape comparisons exclusively.

For example, Figure 15 shows the log likelihood values measured for the chemical elements tested against the data in the  $18.2 \leq \log_{10}(E/\text{eV}) < 18.3$  energy bin. Figure 16 shows the data after shifting by the best  $\Delta X_{\max}$  and the Monte Carlo for each chemical element for the same energy bin. Data and protons appear to agree well over their entire distributions. Data and helium match well up until about 850 g cm<sup>-2</sup>, where the helium tail begins to fall off faster than the data. Nitrogen and iron show less agreement in the tails as well. Note that the same data are used for each panel shown in Figure 16, but they are shifted by a different amount in each one. The shifts applied to the data are +29, +7, -19, and -41 g cm<sup>-2</sup> in the proton, helium, nitrogen, and iron panels, respectively. The slight variation of the shape of the data histogram in each panel is due to the effect of systematic shifting of all data points and then binning in the plot. However, the maximum likelihood calculated for our tests uses an unbinned method.

We then calculate the probability ( $p$ -value) of measuring a log likelihood in the Monte Carlo equal to or more extreme than the one measured in the data shifted by the best  $\Delta X_{\max}$ . Figure 17 shows the distributions of log likelihood calculated for the  $18.2 \leq \log_{10}(E/\text{eV}) < 18.3$  energy bin, as well as the likelihood measured for the data when shifted by the best  $\Delta X_{\max}$ . The null hypothesis being tested is that the data after shifting and the Monte Carlo are drawn from the same continuous distribution. If the  $p$ -value we measure from this test statistic is 0.05 or less, we reject the null hypothesis at the

**Table 4**  
 $\langle X_{\max} \rangle$  and  $\sigma(X_{\max})$  Observed over 8.5 yr of Data by Telescope Array in BR/LR Hybrid Collection Mode

$E_{\text{low}}$	$\langle E \rangle$	$E_{\text{high}}$	$N_{\text{data}}$	$\langle X_{\max} \rangle$	$\sigma(X_{\max})$
18.20	18.25	18.30	801	$715 \pm 2_{-17.4}^{+17.4}$	$63 \pm 2_{-4}^{+3}$
18.30	18.35	18.40	758	$720 \pm 2_{-17.4}^{+17.4}$	$59 \pm 2_{-4}^{+4}$
18.40	18.45	18.50	572	$734 \pm 2_{-17.4}^{+17.4}$	$58 \pm 2_{-4}^{+4}$
18.50	18.55	18.60	395	$742 \pm 3_{-17.4}^{+17.4}$	$61 \pm 3_{-4}^{+4}$
18.60	18.65	18.70	289	$743 \pm 3_{-17.4}^{+17.4}$	$58 \pm 3_{-4}^{+4}$
18.70	18.75	18.80	170	$749 \pm 5_{-17.4}^{+17.4}$	$65 \pm 6_{-4}^{+3}$
18.80	18.85	18.90	132	$750 \pm 5_{-17.4}^{+17.4}$	$52 \pm 5_{-4}^{+4}$
18.90	18.95	19.00	80	$758 \pm 7_{-17.4}^{+17.4}$	$61 \pm 8_{-4}^{+4}$
19.00	19.09	19.20	87	$769 \pm 5_{-17.4}^{+17.4}$	$46 \pm 4_{-5}^{+5}$
19.20	19.29	19.40	27	$761 \pm 7_{-17.4}^{+17.4}$	$35 \pm 4_{-7}^{+6}$
19.40	19.57	19.90	19	$777 \pm 7_{-17.4}^{+17.4}$	$29 \pm 4_{-9}^{+7}$

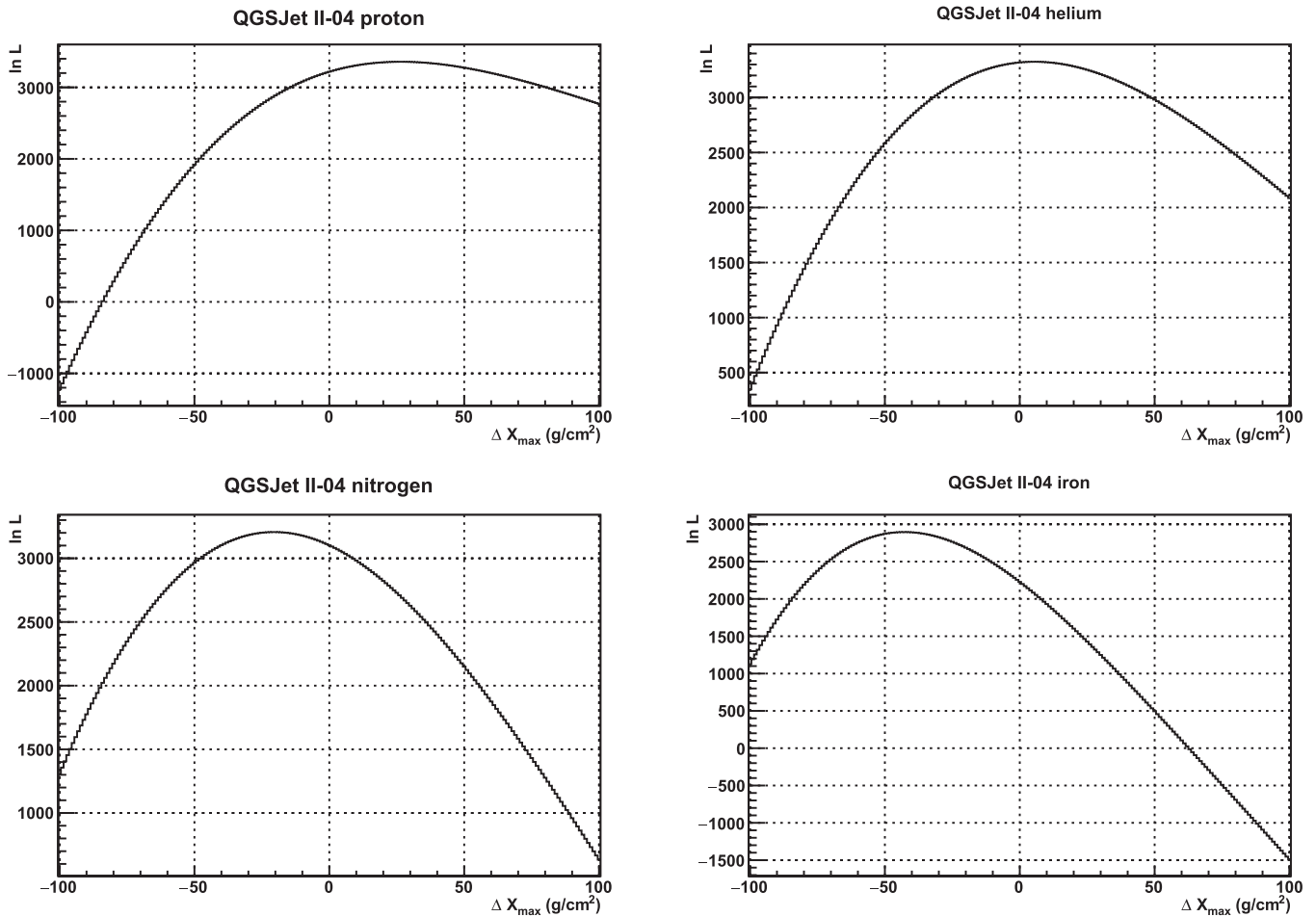
**Note.** Energy is in Units of  $\log_{10}(E/\text{eV})$ , and  $\langle X_{\max} \rangle$  and  $\sigma(X_{\max})$  are in g cm<sup>-2</sup>.

95% confidence level, and we say that the two distributions are not compatible. If the  $p$ -value is greater than 0.05, we fail to reject the systematically shifted data and Monte Carlo as being compatible.

Hadronic models in the UHECR energy regime are based on measurements made in accelerators. Cross section, multiplicity, and elasticity of the primary particle are fundamental parameters used by these models that are particularly sensitive for UHECR  $X_{\max}$ . These parameters are measured at relatively low energies ( $\sqrt{s} = 14$  TeV corresponds to about  $10^{17}$  eV in the lab frame), which need to be extrapolated up to  $10^{20}$  eV to fully describe the physics up to the highest-energy cosmic rays observed. Abbasi and Thomson have examined the uncertainty in  $\langle X_{\max} \rangle$  in several different popular hadronic models introduced by extrapolating these parameters. The estimated lower limits on the uncertainty in  $\langle X_{\max} \rangle$  from the extrapolation were found to  $\sim 6$  g cm<sup>-2</sup> at  $E_{\text{lab}} = 10^{17}$  eV and  $\sim 35$  g cm<sup>-2</sup> at  $E_{\text{lab}} = 10^{19.5}$  eV (Abbasi & Thomson 2016). This uncertainty in  $\langle X_{\max} \rangle$  at  $10^{19.5}$  eV is about the same as the difference in  $\langle X_{\max} \rangle$  predicted among the deepest model (EPOS LHC) and the shallowest model (QGSJet01c). The shapes of the  $X_{\max}$  distributions have a much smaller dependence on hadronic model assumptions. Because of these large uncertainties in the models that we compare our observed  $X_{\max}$  to, we simultaneously systematically shift the data and test the shapes of the distributions to measure compatibility between the data and model.

## 5.2. Results

Table 5 shows the results of these tests. For each QGSJet II-04 model tested against the data, the  $\Delta X_{\max}$  that gave the best log likelihood is shown, as well as the  $p$ -value for that shift. For



**Figure 15.** Log likelihood in the energy bin  $18.2 \leq \log_{10}(E/\text{eV}) < 18.3$  measured between data and Monte Carlo for the four chemical elements under test in this work. We systematically shift the data in each energy bin between  $-100 \text{ g cm}^{-2} \leq \Delta X_{\text{max}} \leq 100 \text{ g cm}^{-2}$  and calculate the unbinned maximum likelihood. The  $\Delta X_{\text{max}}$  that corresponds to the maximum measured likelihood is used to test the compatibility of the data and Monte Carlo  $X_{\text{max}}$  distributions.

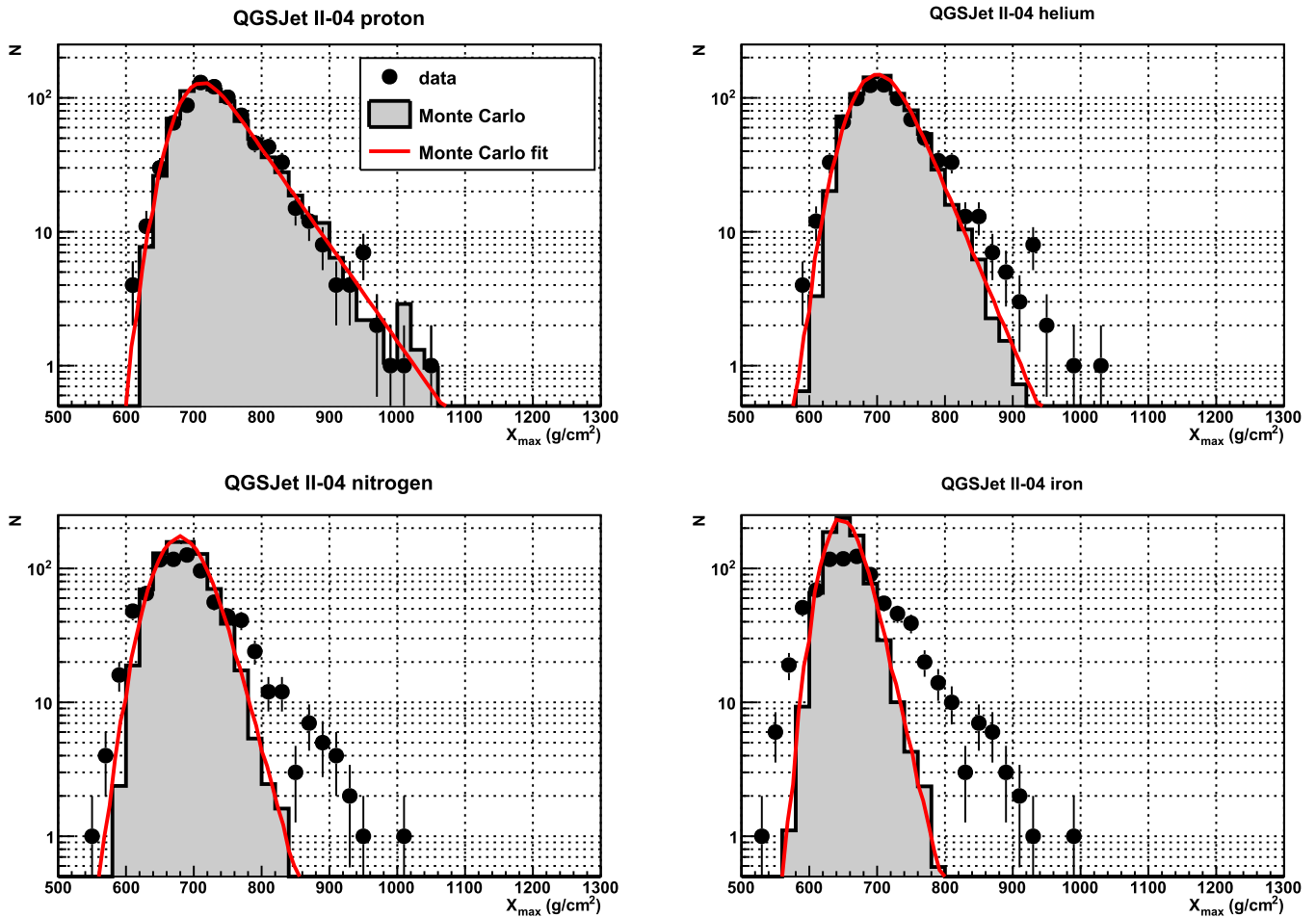
QGSJet II-04 protons, in most energy bins the  $\Delta X_{\text{max}}$  shifts are about the size of or slightly larger than the systematic uncertainty of  $\langle X_{\text{max}} \rangle$  for this analysis. The  $p$ -values, which measure the agreement of the shapes of the entire  $X_{\text{max}}$  distributions after shifting, all have values  $>0.05$ ; therefore, we fail to reject protons as being compatible with the data for all energy bins using this test. QGSJet II-04 helium has  $\Delta X_{\text{max}}$  shifts smaller than protons and within our quoted systematic uncertainties of  $\langle X_{\text{max}} \rangle$ , but the  $p$ -values indicate that once shifting is performed the shapes of the data and Monte Carlo do not agree for  $\log_{10}(E/\text{eV}) < 19.0$ . For those energy bins, the test rejects QGSJet II-04 helium as being compatible with data after systematic shifting. Above  $\log_{10}(E/\text{eV}) = 19.0$ , the  $p$ -values are  $>0.05$  and we fail to reject helium as being compatible with the data. QGSJet II-04 nitrogen requires  $\Delta X_{\text{max}}$  shifts slightly larger than our systematic uncertainty, but the  $p$ -values of the tests reject nitrogen for  $\log_{10}(E/\text{eV}) < 19.2$ . Iron requires  $\Delta X_{\text{max}}$  shifts larger than our systematic uncertainty and is rejected as being compatible with the data for  $\log_{10}(E/\text{eV}) < 19.4$ . Figure 18 visually summarizes the results of the tests and the data in Table 5.

We can understand why the likelihood test finds our data simultaneously compatible with elements with very different masses such as protons and nitrogen in the last two energy bins if we consider Figures 13(d) and (e). We see that because of poor detector exposure leading to very few events collected in

these energy bins, the confidence intervals of  $\langle X_{\text{max}} \rangle$  and  $\sigma(X_{\text{max}})$  of the Monte Carlo overlap among the different elements. Given our current exposure, we should not expect to be able to distinguish the difference between protons, helium, and nitrogen in these energy bins. Iron is also found compatible with the data in the last energy bin, but a shift of  $\Delta X_{\text{max}}$  larger than our systematic uncertainty is required. The agreement of the maximum likelihood test then comes from very few events in the data (19 events) and the lack of a tail in the data deep  $X_{\text{max}}$  distribution ( $X_{\text{max}} > 850 \text{ g cm}^{-2}$ ) as seen in Figure 11(e). This lack of deep  $X_{\text{max}}$  tail allows the shapes of the data and Monte Carlo to resemble each other after a sufficient amount of  $X_{\text{max}}$  shifting.

$X_{\text{max}}$  analysis using 5 yr of data analyzed by the MD FD also found compatibility of the data with protons and incompatibility with iron. In that analysis, the  $\chi^2$  test was applied to QGSJet II-03 protons and iron in three energy ranges:  $10^{18.2}\text{--}10^{18.4} \text{ eV}$ ,  $10^{18.4}\text{--}10^{18.6} \text{ eV}$ , and  $10^{18.6}\text{--}10^{18.8} \text{ eV}$ . For these energy bins, the  $p$ -values of the  $\chi^2$  tests rejected iron, but not protons (Abbasi et al. 2014).

As we have discussed previously, this analysis attempts to address the problem of potentially large, unknown systematic uncertainties between hadronic models and  $X_{\text{max}}$  distributions observed in nature by relying on a test of morphology. Systematic differences in  $X_{\text{max}}$  are removed by uniformly shifting the data distributions until the likelihood of observing



**Figure 16.** Data after shifting by the  $\Delta X_{\max}$  that provides the maximum likelihood with the Monte Carlo in the  $18.2 \leq \log_{10}(E/eV) < 18.3$  energy bin. The unbinned data are tested against the fit to the Monte Carlo (red line). For this energy bin the data and QGSJet II-04 protons were found to be compatible, with a  $p$ -value of 0.32. All other chemical elements had  $p$ -values  $< 0.05$  and were found to be incompatible with the data.

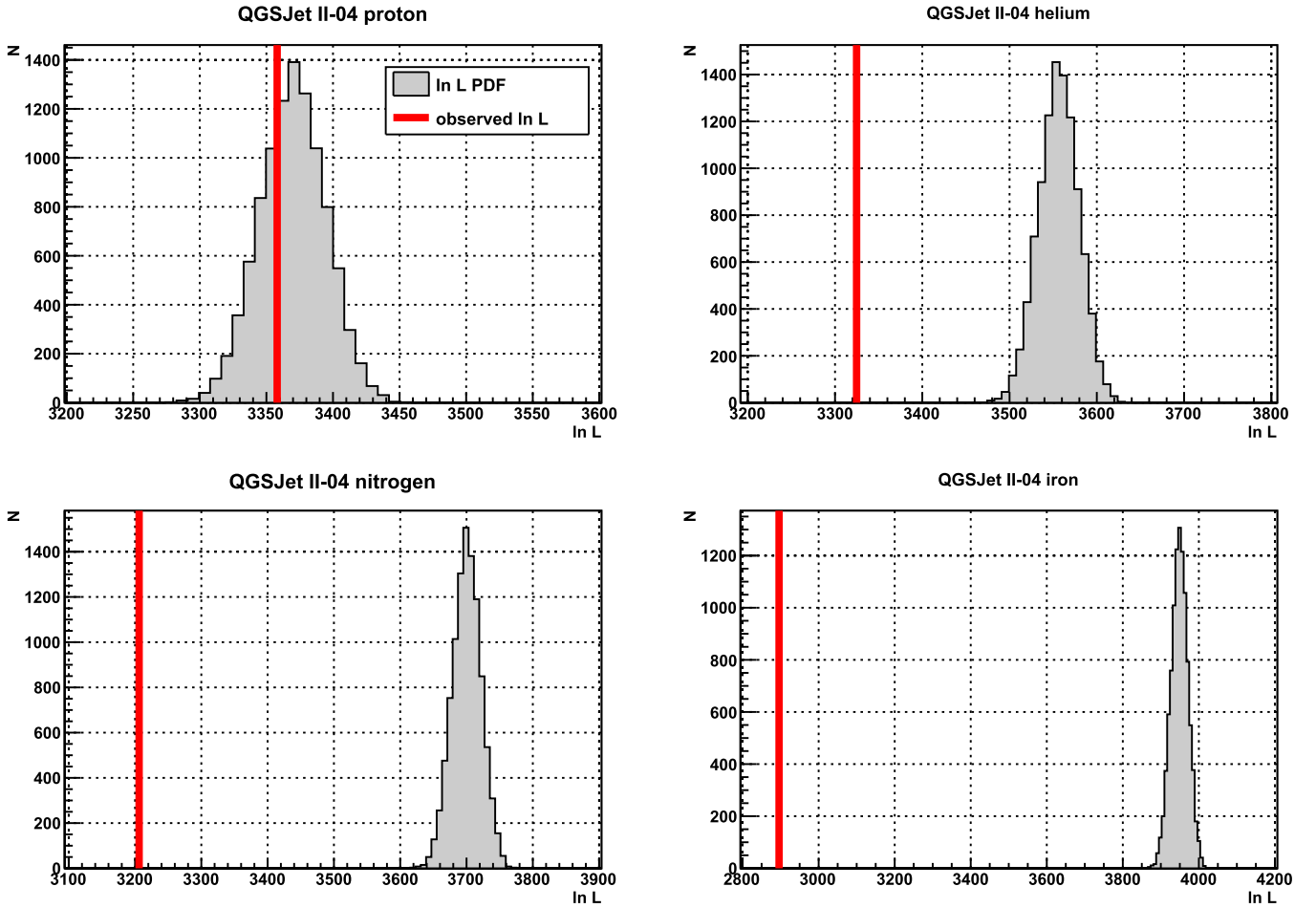
the data given each single element chemical species is maximized. The maximum likelihood in this way measures the agreement of the entire shapes of the simulated and observed distributions, which we have said is a more complete and powerful test of composition rather than relying merely on the first and second moments of the distributions. Since we have effectively “shifted” away any potential disagreement of the overall  $\langle X_{\max} \rangle$  in data and Monte Carlo, one may then ask about uncertainties in the shapes themselves, specifically the widths of the distributions characterized by the sample standard deviations.

In the case of simple, single-element composition, which we test against in this analysis, the deep tail of the  $X_{\max}$  distribution can be considered as a discriminator of light composition. By its very nature it consists of relatively few events compared to the rest of the distribution. The question then arises, what do we understand about effects based in composition theory, hadronic modeling, and our detector design and reconstruction that may introduce biases that can cause significant fluctuations of the few events in the tails? Given great enough distortions in the tails, two potential undesirable outcomes must be guarded against in detector design and reconstruction: poor acceptance or reconstruction bias causing a sparsely populated tail to be cut off, or poor resolution causing large upward fluctuations in the tail by asymmetrically moving events from shallow  $X_{\max}$  to deep  $X_{\max}$ . In the former case, these issues can make an  $X_{\max}$

distribution look as if it is heavier than it truly is; in the latter case, the distribution may be distorted to look lighter.

The Heitler model of electromagnetic showers introduced in Section 1 can be modified to investigate the expected properties of hadronic showers by considering a cascade of pions. This predicts  $X_{\max} \sim \lambda_1 + \lambda_e \ln(E_0/(n_{\text{mult}} E_c^{\text{e.m.}}))$  and  $V[X_{\max}] \propto V[X_1] + \lambda_e^2 V[\ln n_{\text{mult}}]$ , where  $\lambda_1$  is the pion interaction length in air,  $\lambda_e$  is the electron interaction length in air,  $E_0$  is the primary particle energy,  $n_{\text{mult}}$  is the number of secondary charged and neutral pions produced in each interaction,  $E_c^{\text{e.m.}}$  is the critical energy for electrons, and  $X_1$  is the depth of first interaction in the atmosphere. Because  $V[X_1] = \lambda_1^2 \propto 1/\sigma_1^2$ , where  $\sigma_1$  is the pion-air cross section, we find that fluctuations in  $X_{\max}$  are proportional to the fluctuations in first interaction depth, and therefore interaction cross section, and fluctuations in the multiplicity (Ulrich et al. 2011). If the cross section is very small,  $X_{\max}$  fluctuations are dominated by the cross section, and if the cross section is large, fluctuations in multiplicity dominate.

Ulrich et al. (2011) discuss the sensitivity of  $\sigma(X_{\max})$  to fundamental hadronic model parameters such as cross section, multiplicity, elasticity, and pion charge ratio. This is of interest to understanding UHECRs because of the large extrapolations of these parameters from energies that are achievable in the lab to UHECR energies. They find that  $\sigma(X_{\max})$  is very sensitive to choice of cross section for proton- and iron-initiated air showers at  $E_0 = 10^{19.5}$  eV. Protons also show sensitivity to



**Figure 17.** Calculation of  $p$ -value of the unbinned maximum likelihood test between data and Monte Carlo for the  $18.2 \leq \log_{10}(E/\text{eV}) < 18.3$  energy bin. The data are shifted by the  $\Delta X_{\text{max}}$  corresponding to the maximum likelihood, the value of which is indicated by the red line. The Monte Carlo is sampled according to the data statistics to measure the distribution of log likelihood expected. The  $p$ -value is the probability of measuring log likelihood less than or equal to that observed in the shifted data. The calculated  $p$ -values for this energy bin are 0.32 for proton,  $10^{-25}$  for helium,  $10^{-93}$  for nitrogen, and  $<10^{-324}$  for iron.

**Table 5**  
Results of Unbinned Maximum Likelihood Test of TA BR/LR Hybrid  $X_{\text{max}}$  Data against Four Pure QGSJet II-04 Chemical Models

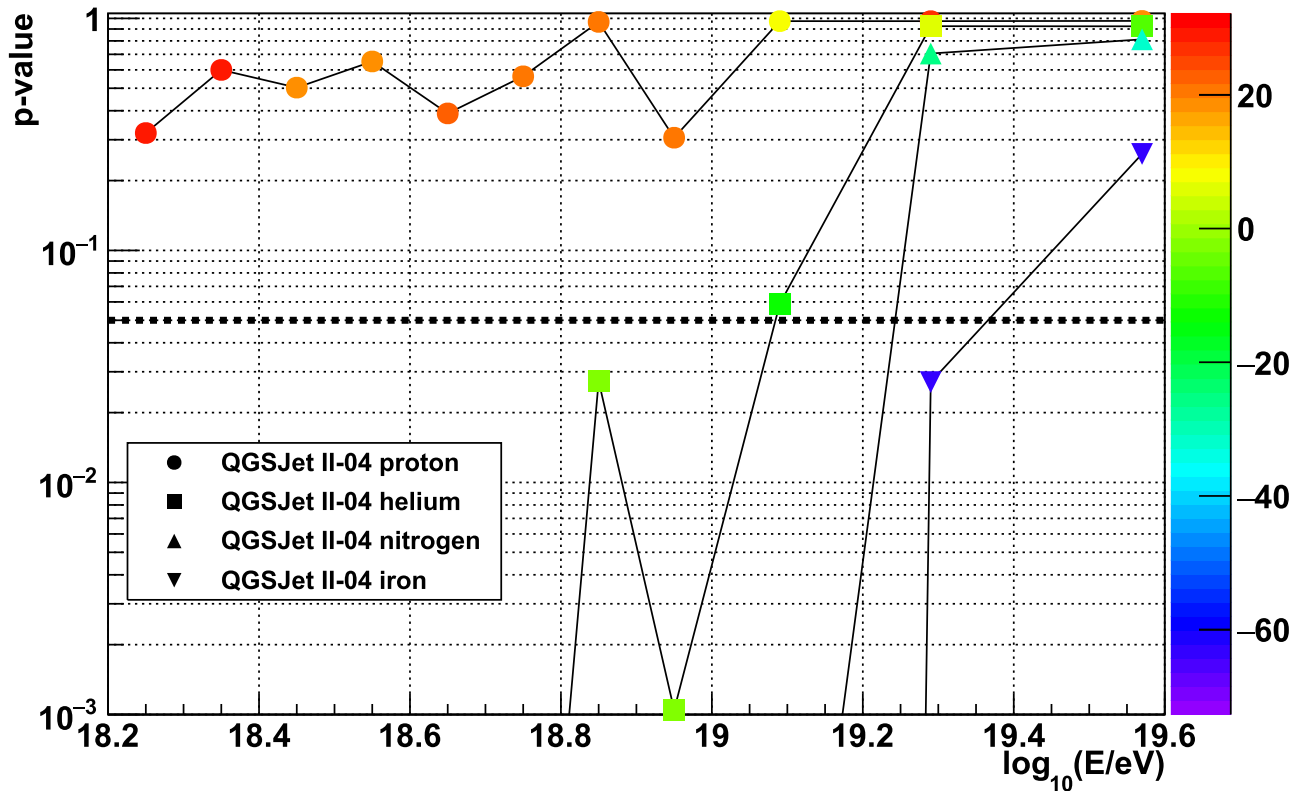
Energy	Proton		Helium		Nitrogen		Iron	
	$\Delta X_{\text{max}}$	$p$ -val	$\Delta X_{\text{max}}$	$p$ -val	$\Delta X_{\text{max}}$	$p$ -val	$\Delta X_{\text{max}}$	$p$ -val
18.2–18.3	$29 \pm 2$	0.32	$7 \pm 2$	—	$-19 \pm 1$	—	$-41 \pm 1$	—
18.3–18.4	$30 \pm 2$	0.59	$6 \pm 2$	$2 \times 10^{-18}$	$-19 \pm 1$	—	$-43 \pm 1$	—
18.4–18.5	$19 \pm 2$	0.50	$-2 \pm 2$	$9 \times 10^{-11}$	$-28 \pm 2$	—	$-53 \pm 1$	—
18.5–18.6	$19 \pm 2$	0.65	$-2 \pm 2$	$2 \times 10^{-11}$	$-33 \pm 2$	—	$-54 \pm 2$	—
18.6–18.7	$22 \pm 3$	0.38	$-1 \pm 3$	$3 \times 10^{-7}$	$-25 \pm 2$	—	$-52 \pm 2$	—
18.7–18.8	$20 \pm 4$	0.55	$2 \pm 3$	$6 \times 10^{-6}$	$-24 \pm 3$	—	$-53 \pm 2$	—
18.8–18.9	$20 \pm 4$	0.97	$2 \pm 3$	0.027	$-27 \pm 3$	$3 \times 10^{-6}$	$-51 \pm 2$	—
18.9–19.0	$21 \pm 5$	0.30	$1 \pm 5$	0.0010	$-25 \pm 4$	$1 \times 10^{-14}$	$-42 \pm 3$	—
19.0–19.2	$10 \pm 5$	0.98	$-7 \pm 4$	0.059	$-34 \pm 4$	$1 \times 10^{-5}$	$-57 \pm 3$	—
19.2–19.4	$26 \pm 8$	0.98	$9 \pm 8$	0.93	$-18 \pm 7$	0.71	$-50 \pm 5$	0.027
19.4–19.9	$19 \pm 8$	0.98	$-3 \pm 8$	0.93	$-23 \pm 7$	0.81	$-50 \pm 6$	0.26

**Note.** For each model the  $\Delta X_{\text{max}}$  shift required to find the maximum log likelihood is shown, as well as the  $p$ -value of the likelihood. After systematic shifting of the data, the maximum likelihood  $p$ -values reject all species except QGSJet II-04 protons at the 95% confidence level for energies below  $10^{19}$  eV. Above  $10^{19}$  eV, statistics are rapidly falling, and the likelihood test fails to reject at the 95% confidence level even very heavy elements. Entries shown as “—” have a  $p$ -value of  $<7.6 \times 10^{-24}$  (Significance  $>10\sigma$ ).  $X_{\text{max}}$  shifts are measured in  $\text{g cm}^{-2}$ .

elasticity, while iron much less so. They also note that  $\sigma(X_{\text{max}})$  observed by Auger at  $E = 10^{19.54}$  eV (Abraham et al. 2010) are compatible with a pure iron composition at the  $1\sigma$  level, but also compatible with a proton expectation with a large enough increase in the proton cross section. This study examined the

dependence of air shower observables such as  $\sigma(X_{\text{max}})$  on fundamental hadronic interaction model parameters at a single energy. To estimate the uncertainty in  $\sigma(X_{\text{max}})$  over a range of energies, large Monte Carlo sets generated using different hadronic models must be employed. The difference in pure





**Figure 18.** Unbinned maximum likelihood test on observed and simulated QGSJet II-04  $X_{\max}$  distributions after systematic shifting of the data to find the best log likelihood. Each point represents the probability of measuring a log likelihood more extreme than that observed in the data after it is shifted by the best  $\Delta X_{\max}$ . The color of the point indicates the  $\Delta X_{\max}$  measured in  $\text{g cm}^{-2}$  required to find the maximum log likelihood value. The dashed line at  $p\text{-value} = 0.05$  indicates the threshold below which the data are deemed incompatible with the Monte Carlo at the 95% confidence level.

proton  $\sigma(X_{\max})$  expected among the most recent versions of QGSJet II and EPOS is  $\lesssim 2 \text{ g cm}^{-2}$  between  $10^{18}$  and  $10^{19}$  eV and grows to about  $5 \text{ g cm}^{-2}$  at  $10^{19.7}$  eV (Abraham et al. 2010). The difference for QGSJet II and EPOS pure iron is consistently about  $3 \text{ g cm}^{-2}$  between  $10^{18}$  and  $10^{19.7}$  eV.

As shown in Table 4, below  $10^{19}$  eV, where TA has sufficient statistics and sensitivity to resolve the individual species presented in this work, the total systematic uncertainty on  $\sigma(X_{\max})$  is  $\sim \pm 4 \text{ g cm}^{-2}$ . Below  $10^{19}$  eV the average total  $\sigma(X_{\max})$  bias ( $\sigma(X_{\max}^{\text{recon}}) - \sigma(X_{\max}^{\text{thrown}})$ , where  $\sigma(X_{\max}^{\text{thrown}})$  is calculated for the Monte Carlo before any acceptance or trigger bias) caused by this reconstruction is  $-6$ ,  $+2$ ,  $+3$ , and  $+4 \text{ g cm}^{-2}$  for proton, helium, nitrogen, and iron, respectively, which is of order our calculated systematic uncertainty. The slight average narrowing of the proton distribution is due to the deeply penetrating nature of those primaries, which causes a decrease in acceptance. The slight increase in the other primaries' widths is due to shallower  $\langle X_{\max} \rangle$ , and therefore improved acceptance, and resolution smearing.

We have examined how large our  $X_{\max}$  resolution must be to make heavier distributions appear as a proton distribution. This is done by convolving different resolution functions with  $X_{\max}$  distributions that are accepted by our reconstruction but are not distorted by resolution. Again, below  $10^{19}$  eV, where TA has sufficient sensitivity to distinguish between individual chemical elements, we require resolutions at least 4–5 times worse than what our reconstruction provides to make helium appear as protons after reconstruction. For nitrogen and iron, resolutions worse than 5 times our calculated resolution are required to make those distributions appear to have a deep  $X_{\max}$  tail

resembling protons. Resolutions this large also have the effect of distorting the distributions in the shallow  $X_{\max}$  tail, causing large changes in the overall shapes.

Our reconstruction was also examined for potential changes in  $X_{\max}$  resolution as a function of depth. The concern is that for events with large  $X_{\max}$ , resolution is asymmetrically populating the deep  $X_{\max}$ , causing a heavy distribution to appear lighter. We find that  $|(X_{\max}^{\text{recon}} - X_{\max}^{\text{thrown}}) / \sigma(X_{\max}^{\text{thrown}})| < 1$  for all species at all depths for  $E < 10^{19}$  eV. This measure of  $X_{\max}$  resolution asymptotically approaches 0 with increasing energy and depth. This indicates that we have reasonable control over our resolution through event selection and detector acceptance over the energy range examined in this work.

## 6. Conclusions

TA has completed analyzing 8.5 yr of data collected in hybrid mode using events observed simultaneously by the SD array and the BR and LR fluorescence detectors. These data provided 3330 events after reconstruction and cuts are applied, and they were used to analyze the depth of shower maxima ( $X_{\max}$ ). Good operation of the detector was verified by using an extensive Monte Carlo suite with showers pregenerated using CORSIKA. This Monte Carlo allows us to verify that we understand the detector with a high degree of confidence and also to compare the observed  $X_{\max}$  distributions with CORSIKA models of four different single-element primaries: protons, helium, nitrogen, and iron, all generated using the QGSJet II-04 hadronic model. The data can be compared to the Monte Carlo by the traditional method, comparing the first and second moments ( $\langle X_{\max} \rangle$  and  $\sigma(X_{\max})$ ) of the observed  $X_{\max}$


















distributions to the Monte Carlo. This method may be overly simplistic and misleading, especially for energy bins with low exposure, which can change the shapes of the observed distributions. We have presented a new way to visualize  $\langle X_{\max} \rangle$  and  $\sigma(X_{\max})$  by plotting their joint distributions in the data, as well as the confidence intervals expected from Monte Carlo. We have extended the analysis of  $X_{\max}$  by using unbinned maximum likelihood, which allows us to measure the compatibility of the data and Monte Carlo using the entire distributions. This is especially important for statistical distributions that potentially exhibit a high degree of skew, such as those of light elements with  $X_{\max}$  distributions with deeply penetrating tails. Using this test, we can empirically reject certain chemical elements at a given confidence level as being compatible with our data.

After allowing for systematic shifting of the data  $X_{\max}$  distributions and performing the likelihood test on the data and Monte Carlo distributions of four pure chemical species, we find that we fail to reject QGSJet II-04 protons as being compatible with the data for all energy bins at the 95% confidence level. QGSJet II-04 helium is rejected as being compatible with the data for  $\log_{10}(E) < 19.0$ . QGSJet II-04 nitrogen is rejected for  $\log_{10}(E) < 19.2$ , and iron is rejected for  $\log_{10}(E) < 19.4$ . We have demonstrated that for  $\log_{10}(E) \geq 19.0$ , TA has insufficient exposure to accurately distinguish the difference between different individual elements. Energy bins in this energy range have poor statistics due to low exposure, and agreement among several of the models is found with the data. However, this agreement is physically unrealistic for the case of iron because of the large shifts required, in excess of our systematic uncertainty.

The Telescope Array experiment is supported by the Japan Society for the Promotion of Science (JSPS) through Grants-in-Aid for Specially Promoted Research JP21000002, for Scientific Research (S) JP19104006, for Specially Promoted 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 (2016R1A2B4014967, 2016R1A5A1013277, 2017K1A4A3015188, 2017R1A2A1A05071429); by the Russian Academy of Sciences, RFBR grant 16-02-00962a (INR), IISN project No. 4.4502.13, and Belgian Science Policy under IUAP VII/37 (ULB). The foundations of Dr. Ezekiel R. and Edna Wattis Dumke, Willard L. Eccles, and George S. and Dolores Doré 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 that

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.

### ORCID iDs

T. Abu-Zayyad  <https://orcid.org/0000-0001-5206-4223>  
 J. W. Belz  <https://orcid.org/0000-0001-9779-2750>  
 D. R. Bergman  <https://orcid.org/0000-0002-4450-7925>  
 W. Hanlon  <https://orcid.org/0000-0002-0109-4737>  
 S. Kawakami  <https://orcid.org/0000-0003-3820-7552>  
 K. Kawata  <https://orcid.org/0000-0001-6332-2005>  
 E. Kido  <https://orcid.org/0000-0001-7278-3049>  
 J. P. Lundquist  <https://orcid.org/0000-0002-4245-5092>  
 S. Nagataki  <https://orcid.org/0000-0002-7025-284X>  
 T. Nonaka  <https://orcid.org/0000-0003-4795-500X>  
 G. Rubtsov  <https://orcid.org/0000-0002-6106-2673>  
 D. Ryu  <https://orcid.org/0000-0002-5455-2957>  
 P. Sokolsky  <https://orcid.org/0000-0003-3391-1022>  
 S. B. Thomas  <https://orcid.org/0000-0002-8828-7856>  
 S. Troitsky  <https://orcid.org/0000-0001-6917-6600>  
 Y. Tsunesada  <https://orcid.org/0000-0001-9238-6817>  
 K. Yamazaki  <https://orcid.org/0000-0002-3771-2496>

### References

- Aab, A., Abreu, P., Aglietta, M., et al. 2014, *PhRvD*, **90**, 122005  
 Aab, A., Abreu, P., Aglietta, M., et al. 2015, arXiv:1708.06592  
 Aartsen, M. G., Abbasi, R., Adbou, Y., et al. 2013, *PhRvD*, **88**, 042004  
 Abbasi, R., Abu-Zayyad, T., Belov, K., et al. 2008b, *Aph*, **29**, 77  
 Abbasi, R. U., Abe, M., Abu-Zayyad, T., et al. 2014, *Aph*, **64**, 49  
 Abbasi, R. U., Abu-Zayyad, T., Al-Seady, M., et al. 2010, *PhRvL*, **104**, 161101  
 Abbasi, R. U., Abu-Zayyad, T., Amann, J. F., et al. 2008a, *PhRvL*, **100**, 101101  
 Abbasi, R. U., & Thomson, G. B. 2016, arXiv:1605.05241  
 Abraham, J., Abreu, P., Aglietta, M., et al. 2008, *PhRvL*, **101**, 061101  
 Abraham, J., Abreu, P., Aglietta, M., et al. 2013, *PhRvL*, **104**, 091101  
 Abu-Zayyad, T., Aida, R., Allen, M., et al. 2013a, *ApJL*, **768**, L1  
 Abu-Zayyad, T., Aida, R., Allen, M., et al. 2013b, *Aph*, **48**, 16  
 Abu-Zayyad, T., Aida, R., Allen, M., et al. 2013c, *NIMPA*, **689**, 87  
 Abu-Zayyad, T., Al-Seady, M., Belov, K., et al. 2000, *NIMPA*, **450**, 253  
 Abu-Zayyad, T., Belov, K., Bird, D. J., et al. 2001, *ApJ*, **557**, 686  
 Allard, D., Busca, N. G., Decerprit, G., Olinto, A. V., & Parizot, E. 2008, *JCAP*, **0810**, 033  
 Aloisio, R., Berezhinsky, V., Blasi, P., et al. 2007, *Aph*, **27**, 76  
 Aloisio, R., Berezhinsky, V., & Gazizov, A. 2012, *J. Phys. Conf. Ser.*, **337**, 012042  
 Amenomori, M. 2008, *AdSpR*, **42**, 467  
 Ave, M., Bohacova, M., Buonomo, B., et al. 2007, *Aph*, **28**, 41  
 Ave, M., Bohacova, M., Curry, E., et al. 2013, *Aph*, **42**, 90  
 Barcikowski, E. L. 2011, PhD thesis, Utah Univ., <http://cdmbuntu.lib.utah.edu/cdm/ref/collection/etd3/id/3514>  
 Bird, D. J., Corbato, S. C., Dai, H. Y., et al. 1994, *ApJ*, **424**, 491  
 Boyer, J. H., Knapp, B. C., Mannel, E. J., & Seman, M. 2002, *NIMPA*, **482**, 457  
 Carlson, J. F., & Oppenheimer, J. R. 1937, *PhRv*, **51**, 220  
 Engel, R., Heck, D., & Pierog, T. 2011, *ARNPS*, **61**, 467  
 Fowler, J. W., Fortson, L. F., Jui, C. C. H., et al. 2001, *Aph*, **15**, 49  
 Gaisser, T. K., & Hillas, A. M. 1977, *Proc. ICRC*, **8**, 353  
 Giller, M., Kacperczyk, A., Malinowski, J., et al. 2003, *Proc. ICRC*, **28**, 619  
 Greisen, K. 1966, *PhRvL*, **16**, 748  
 Heck, D., Knapp, J., Capdevielle, J. N., Schatz, G., & Thouw, T. 1998, CORSIKA: A Monte Carlo Code to Simulate Extensive Air Showers (Hannover: TIB)  
 Hess, V. F., & Anderson, C. D. 2013, *Physi*, **351**  
 Hoerandel, J. R. 2004, *Aph*, **21**, 241  
 Ivanov, D. 2012, PhD thesis, Rutgers Univ., <https://rucore.libraries.rutgers.edu/rutgers-lib/38839/>  
 Ivanov, D. 2016, *Proc. ICRC*, **34**, 349  
 Kakimoto, F., Loh, E. C., Nagano, M., et al. 1996, *NIMPA*, **372**, 527

- Kampert, K.-H., & Unger, M. 2012, *Aph*, **35**, 660
- Knurenko, S., & Petrov, I. 2015, *J. Phys. Conf. Ser.*, **632**, 012098
- Knurenko, S., Petrov, I., Petrov, Z., & Sleptsov, I. 2015, *EPJ Web Conf.*, **99**, 04001
- Laboratory, N.-A. R. 2004, Global Data Assimilation System (GDAS1) Archive Information, <https://ready.arl.noaa.gov/gdas1.php>
- Linsley, J., Scarsi, L., & Rossi, B. 1961, *PhRvL*, **6**, 485
- Matthews, J. 2005, *Aph*, **22**, 387
- Nagano, M., Teshima, M., Matsubara, Y., et al. 1992, *JPhG*, **18**, 423
- Ohoka, H., Yoshida, S., & Takeda, M. 1997, *NIMPA*, **385**, 268
- Ostapchenko, S. 2007, in AIP Conf. Proc. 928, C2CR07, ed. M. Tripathi & R. E. Breedon (Melville, NY: AIP), **118**
- Ostapchenko, S. 2011, *PhRvD*, **83**, 014018
- Pierog, T., Karpenko, I., Katzy, J. M., Yatsenko, E., & Werner, K. 2015, *PhRvC*, **92**, 034906
- Prosin, V. V., Berezhnev, S. F., Budnev, N. M., et al. 2015, *EPJ Web Conf.*, **99**, 04002
- Song, C. 2004, *Aph*, **22**, 151
- Stratton, S. R. 2012, PhD thesis, Rutgers Univ., <https://rucore.libraries.rutgers.edu/rutgers-lib/39045/>
- Tameda, Y., Taketa, A., Smith, J. D., et al. 2009, *NIMPA*, **609**, 227
- Todero Peixoto, C. J., de Souza, V., & Bellido, J. A. 2013, *Aph*, **47**, 18
- Tokuno, H., Tameda, Y., Takeda, A., et al. 2012, *NIMPA*, **676**, 54
- Tomida, T., Chikawa, M., Fukushima, M., et al. 2013, *EPJ Web Conf.*, **53**, 10003
- Ulrich, R., Engel, R., & Unger, M. 2011, *PhRvD*, **83**, 054026
- Zatsepin, G. T., & Kuzmin, V. A. 1966, *JETPL*, **4**, 78