PROCEEDINGS OF SCIENCE

Simulation study about event-by-event mass separation of cosmic rays using several mass sensitive air-shower observables

Benjamin Flaggs,[a,*] **Alan Coleman**[a,b] **and Frank G. Schröder**[a,c]

[a] *Bartol Research Institute, University of Delaware, Department of Physics and Astronomy, 104 The Green, Newark, DE 19716, USA*

[b] *Uppsala University, Department of Physics and Astronomy, SE-752 37, Uppsala, Sweden*

[c] *Institute for Astroparticle Physics, Karlsruhe Institute of Technology, D-76021, Karlsruhe, Germany*

E-mail: bflaggs@udel.edu, alan.coleman@physics.uu.se, fgs@udel.edu

Specific observables of extensive air showers, such as the depth of shower maximum and relative size of the air-shower muonic component, are sensitive to the mass of the primary cosmic-ray particle and therefore provide an avenue for cosmic-ray mass composition analyses. The separation power between proton, helium, oxygen, and iron cosmic rays, on an event-by-event basis, was studied using exact knowledge of mass sensitive air-shower observables determined from CORSIKA simulations at the sites of the IceCube Neutrino Observatory at the South Pole and the Pierre Auger Observatory in Argentina. The simulations cover shower energies from 10 PeV to a few EeV and a wide range of zenith angles, relevant to the next generation upgrades to the IceCube and Auger observatories. Combined knowledge of all studied mass sensitive observables yields promising mass separation power, even when accounting for typical reconstruction uncertainties of the observables. The combination of shower maximum and muon observables is of particular importance for event-by-event mass discrimination, while high-energy muons (> 500 GeV), measurable by IceCube, serve as an important mass sensitive observable on their own. This motivates equipping the next generation of air-shower arrays with multiple detection techniques for the simultaneous measurement of these shower observables.

38th International Cosmic Ray Conference (ICRC2023)
26 July - 3 August, 2023
Nagoya, Japan

ICRC2023
38th International Cosmic Ray Conference
The Astroparticle Physics Conference

*Speaker

PoS(ICRC2023)289

1. Introduction

Air-shower observables, such as the depth of shower maximum (X_{max}), a probe of the total muonic component of the shower, the shower electron-muon ratio, and the shape parameters of the longitudinal profile, are on average correlated with the primary cosmic-ray mass [1–3]. Fluctuations between showers constrain these relations to a statistical nature. As a result, single observables are limited in their event-by-event mass determination, yet the additional knowledge gained from simultaneously studying several observables increases the potential mass determination for single events [4]. The ability to distinguish ultra-high-energy cosmic rays by their mass would help to answer contemporary mysteries in cosmic-ray physics about their sources and acceleration mechanisms [5–7].

The next generation of air-shower arrays aims to provide answers to these unknowns. IceCube-Gen2, currently in planning, and AugerPrime, currently being deployed, are the respective next generation upgrades to the IceCube Neutrino Observatory (IceCube) and the Pierre Auger Observatory (Auger). Each upgrade will add both scintillator particle detectors and radio antennas within their existing surface arrays, where IceCube will also extend the surface footprint and in-ice volume [8, 9]. Motivated by the upgrades to these observatories, their locations were chosen to study the mass sensitivity of air-shower observables using a simulation basis.

In this article, information of several air-shower observables is extracted from CORSIKA [10] simulations at the locations of both IceCube and Auger for proton, helium, oxygen, and iron primaries. This observable information is then used within a Fisher linear discriminant analysis to determine the mass separation of individual observables, and their combinations, on an event-by-event basis. The event-by-event mass separation was investigated for various primary energies, zenith angles, observable reconstruction uncertainties, and high-energy hadronic interaction models, all of which are important for the next generation of air-shower arrays. In addition, the hadronic models are important for the interpretation of shower observables. This study serves to determine what is, in principle, possible for future air-shower arrays with regard to mass determination if multiple observables are simultaneously measured, and how various factors impact this mass separation. This is a summary of our more extensive paper [11], where this proceeding additionally includes a comparison between different hadronic interaction models.

2. Air-shower observables and their energy scale corrections

General-purpose CORSIKA libraries are used within this simulation study. Air-shower simulations at the location of IceCube range from $10^{16.0} - 10^{18.5}$ eV in energy and $0.0° - 71.6°$ in zenith angle, with approximately 45,000 air showers per primary. All showers were simulated with the Sibyll 2.3d [12] high-energy hadronic interaction model and with 10^{-6} thinning applied. The Auger simulation library is described in [13], yet only simulations with energies above 10 PeV are used in the analysis to remain consistent with the IceCube library. In addition, only the Auger library for hadronic interaction model Sibyll 2.3c [14] is used when directly comparing the IceCube and Auger locations. Sibyll 2.3c predicts slightly different values of certain mass sensitive observables than Sibyll 2.3d [12]; however, the results of this analysis will be unaffected, as no CORSIKA output is directly compared between observatory sites. Observables are determined from the CORSIKA

output for each shower, where observatory specific detector simulations are excluded to allow for a modular analysis that studies the theoretical limits of mass separation using knowledge of multiple observables. Studied observables include shape parameters of the electron (e^+/e^-) shower profile ($X_{\max}$, R, L) determined from a parameterized Gaisser-Hillas fit [3] to the CORSIKA longitudinal profile, along with the muon number (N_μ) and electron-muon ratio ($R_{e/\mu}$), both at observer level. Both N_μ and $R_{e/\mu}$ are studied within an annulus located 800 − 850 m from the shower axis. The annulus distance from the air-shower axis was motivated by the results of muon density reconstructions presented in [15]. High-energy muons (> 500 GeV), and the respective electron-muon ratio at ground for muons of this energy, were also studied, but only for simulations at the IceCube location as muons of such energies can trigger IceCube's array of in-ice detectors [16]. Simulated showers with anomalous profiles, i.e. double-bump showers, poorly constrained fits, etc., are excluded from the analysis with a frequency similar to the value found in [17].

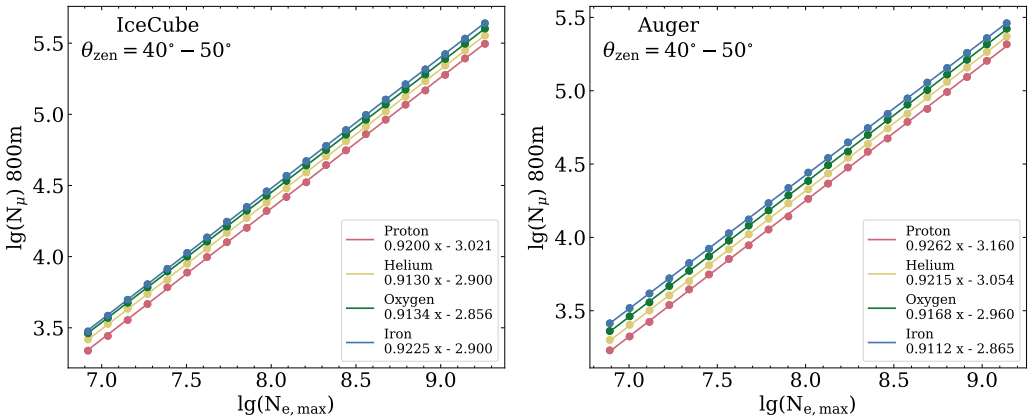

Figure 1: Scaling of the N_μ 800 m observable to the energy reference for all studied primaries at the IceCube (left) and Auger (right) locations. The choice of zenith range is described in the text.

An additional observable, the electron number at $X_{\max}$ ($N_{e,\,max}$) is used as a proxy for the primary energy. All observables are scaled with respect to this quantity to remove any energy dependence, allowing for events to be binned in energy without risk of smearing the observable distributions used in the mass-sensitivity analysis. Fig. 1 shows the scaling of N_μ 800 m with respect to the energy reference for simulations within a zenith range 40° − 50° at both IceCube and Auger. This range was chosen based on both the efficiency of simulated radio reconstructions at IceCube [18] and the distribution of radio events at Auger used for energy estimation in [19]. For the N_μ 800 m observable, a scaling of the form

$$\log\left(N_\mu\right) = \log\left(N_{\mu,\,\text{true}}\right) - 0.93 \times \log\left(\frac{N_{e,\,\max}}{N_{\text{EeV}}}\right) \tag{1}$$

was applied to each event, where N_{EeV} is the average electron number at $X_{\max}$ for all 1 EeV shower simulations. The scaling constant is determined by adding the average slope of the four primaries for this observable to the small difference from an exact linear relationship between the $N_{e,\,max}$ observable and true shower energy. The scalings are performed separately for each observatory location; however, in the case of the N_μ 800 m observable, the applied scaling for both locations follows Eq. 1 with slightly different N_{EeV} values for each location. The scaling of other studied observables follows this same method, but the results are excluded here for brevity.

3. Event-by-event mass separation

Knowledge of additional mass-sensitive observables can be used to soften the statistical limitations imposed by shower-to-shower fluctuations. Fig. 2 shows the two-dimensional contours of the N_μ 800 m and $X_{\max}$ distributions for all four primaries at both the IceCube and Auger locations. The energy range of the figure was chosen as an example and to provide smoother contours due to increased statistics. Clearly illustrated is the potential power for the combination of muon and $X_{\max}$ observables, as less overlap between contours indicates greater mass separation. The addition of other mass sensitive observables will further relax the statistical limitations for mass separation, although interpreting visuals of distributions across more than two dimensions can be perplexing and therefore are excluded for clarity.

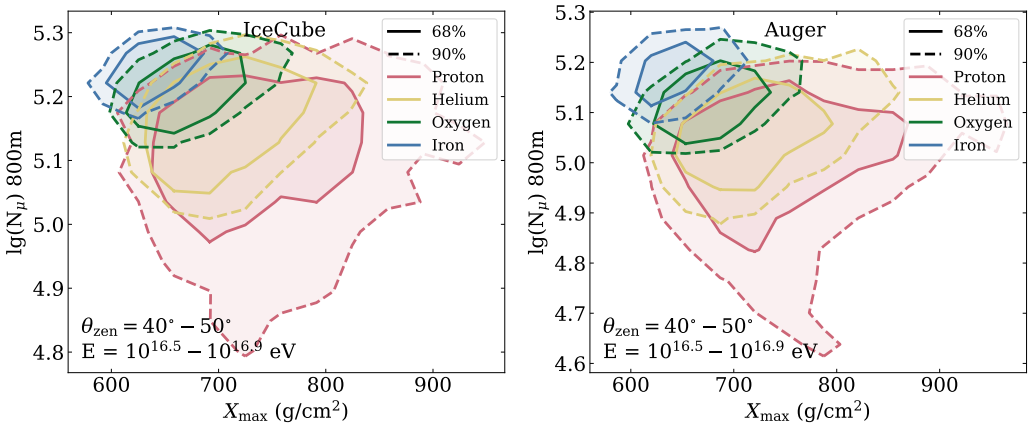

Figure 2: Contours of the scaled N_μ 800 m and $X_{\max}$ distributions for the IceCube (left) and Auger (right) locations. Distributions for all primaries used in this analysis are included. The choice of zenith and energy ranges are described in the text.

A Fisher analysis was used to reduce the dimensionality for all studied observables into a single dimension while preserving information about mass separation. A quantitative measure of the mass separation, known as the figure of merit (FOM), was calculated from the one-dimensional output of the linear discriminant analysis. For two populations, x and y, the FOM is defined as

$$\text{FOM} = \frac{|\mu_x - \mu_y|}{\sqrt{\sigma_x^2 + \sigma_y^2}} \tag{2}$$

where a larger FOM represents greater separation between the populations.

Fig. 3 shows the FOM metric for proton and iron populations calculated as a function of shower energy at both IceCube and Auger. Plots for each observatory show three types of FOM curves: 1) individual observables with an assumed uncertainty and combined observables both 2) with and 3) without included observable resolutions. The choice of resolutions are listed in the legend of Fig. 3 and influenced by current analysis techniques [20–23]. In addition, a 10% uncertainty is assumed for the energy reference observable, with motivation from contemporary radio and air-fluorescence reconstructions of this observable [19, 24]. Bootstrapping was used to calculate an uncertainty band for each FOM curve. Results are presented for large zenith ranges to improve statistics.

For the IceCube location, > 500 GeV muons are clearly an important observable for event-by-event mass separation, as this individual observable is contributing the most to the combined-

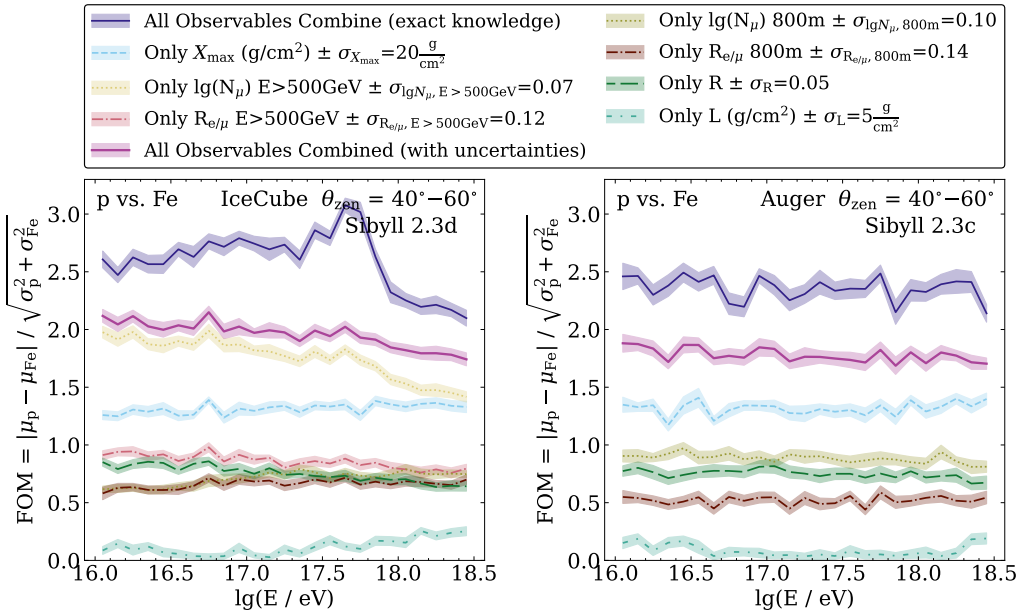

Figure 3: Proton-iron FOM as a function of air-shower energy for all studied observables at both the IceCube (left) and Auger (right) locations. FOM curves are calculated for combined knowledge of all observables and for knowledge of the individual observables within the assumed reconstruction uncertainties stated in the legend. Combined knowledge curves are calculated both when observables are known exactly and within the stated uncertainties. High-energy muon (> 500 GeV) observables are excluded from the Auger location.

knowledge FOM curves. The dip in separation power for high-energy muons towards higher primary energies has an unknown origin. Excluding the high-energy muons reduces the combined-knowledge proton-iron curves at IceCube to a level similar to the Auger location, indicating an overall consistency in the results between observatory locations. Once high-energy muons are excluded, the observable combination of N_μ 800 m and X_{max} serves as the primary influence on the event-by-event mass separation curves at the IceCube site. This same behavior is observed at the Auger site, highlighting the importance of combined muon and X_{max} measurements. In addition, R appears as an individual observable also important for mass separation, but to a lesser extent than the combination of muons and X_{max}. L shows little potential for event-by-event mass separation of protons and iron nuclei at these energies; however, we confirm the results of [25] that L may have importance for determining the proton-helium ratio when averaging over many events.

For less-inclined showers ($0° < \theta_{zen} < 20°$), the combined-knowledge FOM curves for both IceCube and Auger increase slightly (0.1 − 0.3 in FOM) compared to the curves shown in Fig. 3. Most of this increase comes from the heightened mass separation of surface muon observables at both locations, as curves for other observables remain approximately consistent between the different zenith ranges. Assuming larger uncertainties for observables will decrease event-by-event mass separation, where doubling uncertainties in all observables other than the energy reference will decrease the combined knowledge FOM by ≈ 0.5 at both IceCube and Auger. Doubling the uncertainty of the energy reference further decreases separation, but the effects are minimal (< 0.1 in FOM at both locations).

Moving beyond proton-iron mass separation, Fig. 4 shows the FOM metric calculated as a function of shower energy for helium and oxygen populations at both the IceCube and Auger locations. The zenith range and observable uncertainties are the same as those from Fig. 3. At both

observatory sites, helium-oxygen separation is about a factor of two less in FOM than the proton-iron separation. Hence, event-by-event mass separation of helium and oxygen nuclei proves to be difficult, even when all observables are known with exact precision. Although, the combination of muons and X_{max} is still the leading contribution to overall mass separation, where high-energy muons are still the individual observable with the most mass sensitivity at IceCube. Separation between proton and helium primaries was also studied, but shows even lesser separation than helium-oxygen due to larger shower-to-shower fluctuations for proton primaries.

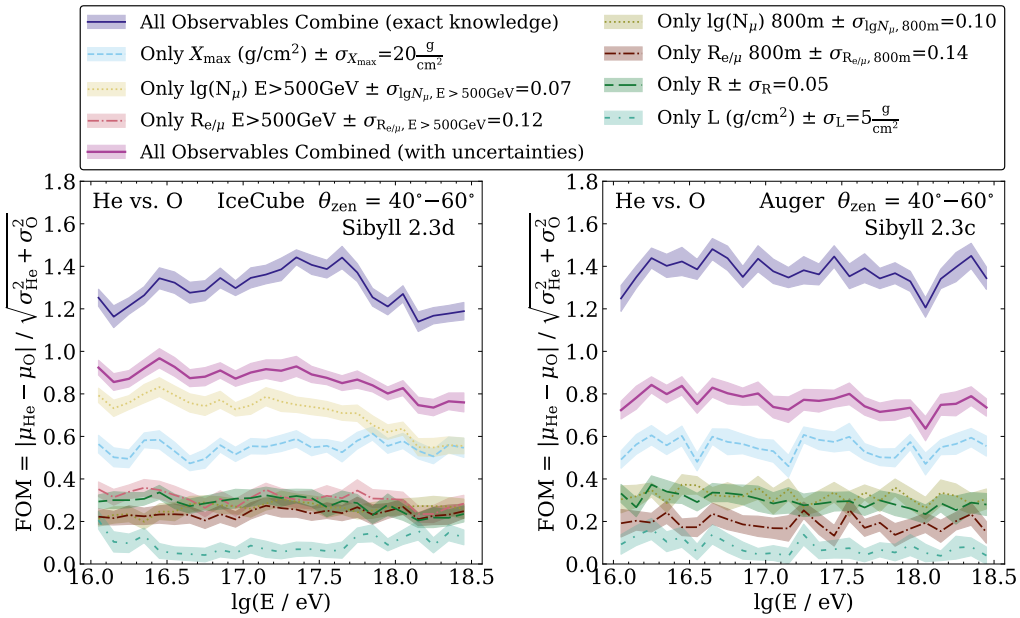

Figure 4: Helium-oxygen FOM as a function of air-shower energy for all studied observables at both the IceCube (left) and Auger (right) locations. See text and Fig. 3 for description.

4. Extension to higher energies and further hadronic interaction models at Auger

The simulation library for the Auger location extends beyond 100 EeV for showers simulated with Sibyll 2.3c and includes air-shower simulations up the 10 EeV for both EPOS-LHC [26] and QGSJET-II-04 [27]. Using the remainder of this library, we extended our simulation study at the Auger location to both higher energies and further hadronic models, as shown in Fig. 5. Observables studied and their uncertainties remain consistent; however, observable scalings have been performed separately for each hadronic model due to differences in both model predictions and the energy range of the simulations. While energy ranges are different for the models, the x- and y-axis scales for their plots are kept consistent to aid in direct comparison between their mass separation curves.

Comparing results between hadronic interaction models yields minuscule differences in proton-iron separation power, with little perceived change in the FOM values for the combined knowledge curves. For the individual observable FOM curves, only small changes appear as well, with the largest difference occurring for the QGSJET-II-04 N_μ 800 m observable. For QGSJET-II-04, the increase in proton-iron mass separation for this observable also slightly increases both combined knowledge curves. However, towards higher primary energies, the FOM curves of individual observables change sometimes drastically, with the largest differences apparent for the Sibyll 2.3c

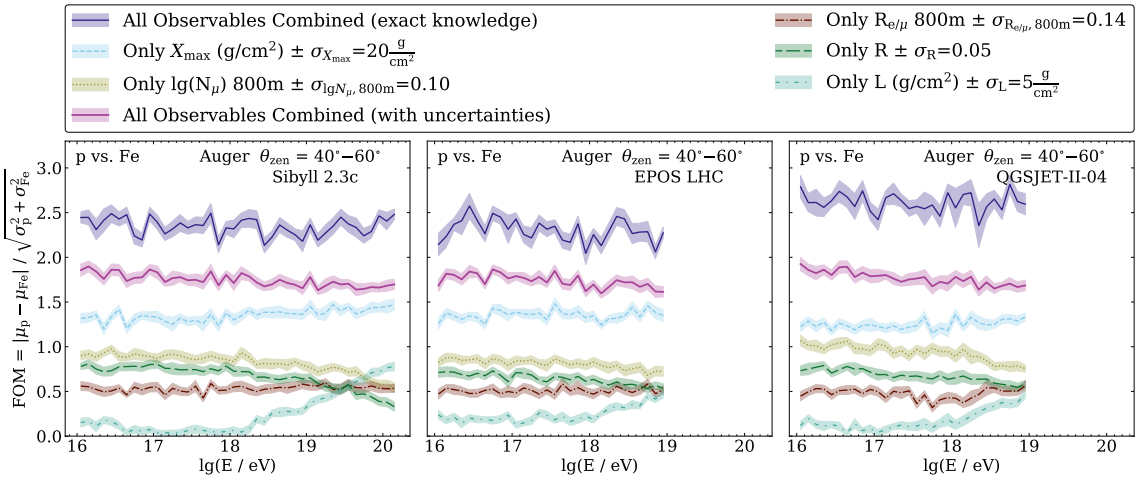

Figure 5: Proton-iron figure of merit of all studied Auger observables for the Sibyll 2.3c (left), EPOS LHC (middle) and QGSJET-II-04 (right) hadronic interaction models. The energy range has also been extended from the previous plots.

curves. The L observable exhibits a steadily increasing proton-iron separation power above ≈3 EeV, where slightly decreasing curves are seen for both the R and N_μ 800 m observables after this primary energy. Similar results are observed for the helium-oxygen and helium-proton separations, yet the L observable for helium-proton separation does not show this drastic increase in FOM. Hence, for the highest energy cosmic rays ever observed, mass composition studies may benefit from reconstructing L in tandem with X_{max} and muon measurements.

5. Conclusion

A study of event-by-event mass separation was performed using knowledge of several air-shower observables obtained from simulations of proton, helium, oxygen, and iron cosmic rays. The mass separation between primaries was studied with respect to multiple factors such as energy, zenith angle, observatory location, and hadronic interaction model. We find the combination of X_{max} and muon measurements provide the most promising avenue for separation of proton and iron primaries on an event-by-event basis, where the addition of L reconstructions at the highest observed cosmic ray energies will aid in this separation. Per-event separation between intermediate mass primaries is similarly influenced by these observables, although methods beyond this analysis are needed to determine if adequate separation between these populations is possible. Furthermore, we conclude that L, when averaged over many events, could prove as a measure of the proton-helium ratio for air showers. We find no major differences in separation power regarding the hadronic interaction models, hence the previous conclusions hold for all studied models. The results of this analysis demonstrate the importance of certain detection techniques with regard to the science goals of both planned upgrades to current observatories, such as IceCube-Gen2 [8] and AugerPrime [9], and for future air-shower arrays, such as the Global Cosmic Ray Observatory (GCOS) [28].

Acknowledgement: This research was supported in part through the use of Data Science Institute (DSI) computational resources at the University of Delaware. This research was supported in part through the use of DARWIN computing system: DARWIN – A Resource for Computational and Data-intensive

Research at the University of Delaware and in the Delaware Region, Rudolf Eigenmann, Benjamin E. Bagozzi, Arthi Jayaraman, William Totten, and Cathy H. Wu, University of Delaware, 2021, URL: https://udspace.udel.edu/handle/19716/29071. The authors are very grateful to the Prague Auger group for providing the simulations for the paper. The production of the simulations would not be possible without the use of the computing resources and the great support of the staff of the Computing Center of the Institute of Physics (CC IoP) of the Czech Academy of Sciences and of the DIRAC project. Funding for this research was provided by the United States National Science Foundation through NSF award #2046386.

References

[1] J. Matthews *Astropart. Phys.* **22** (2005) 387–397.
[2] **KASCADE** Collaboration, W. D. Apel *et al. Astropart. Phys.* **29** (2008) 412–419.
[3] S. Andringa, R. Conceicao, and M. Pimenta *Astropart. Phys.* **34** (2011) 360–367.
[4] E. M. Holt, F. G. Schröder, and A. Haungs *Eur. Phys. J. C* **79** no. 5, (2019) 371.
[5] A. Coleman *et al. Astropart. Phys.* **149** (2023) 102819.
[6] F. Sarazin *et al. Bull. Am. Astron. Soc.* **51** no. 3, (2019) 93.
[7] F. G. Schröder *et al. Bull. Am. Astron. Soc.* **51** (3, 2019) 131.
[8] **IceCube-Gen2** Collaboration, M. G. Aartsen *et al. J. Phys. G* **48** no. 6, (2021) 060501.
[9] **Pierre Auger** Collaboration, A. Aab *et al.* arXiv:1604.03637.
[10] D. Heck *et al.*, "CORSIKA: A Monte Carlo code to simulate extensive air showers," Tech. Rep. FZKA-6019, KIT, 2, 1998.
[11] B. Flaggs, A. Coleman, and F. G. Schröder. arXiv:2306.13246.
[12] F. Riehn *et al. Phys. Rev. D* **102** no. 6, (2020) 063002.
[13] **Pierre Auger** Collaboration, P. Abreu *et al. PoS* **ICRC2021** (2021) 232.
[14] F. Riehn *et al. PoS* **ICRC2017** (2018) 301.
[15] **IceCube** Collaboration, R. Abbasi *et al. Phys. Rev. D* **106** no. 3, (2022) 032010.
[16] **IceCube** Collaboration, M. G. Aartsen *et al. JINST* **9** (2014) P03009.
[17] C. Baus *et al.*, "Anomalous Longitudinal Shower Profiles and Hadronic Interactions," in *32nd International Cosmic Ray Conference*, vol. 2, pp. 206–209. 11, 2011. arXiv:1111.0504 [astro-ph.HE].
[18] **IceCube** Collaboration, A. Coleman *PoS* **ICRC2021** (2021) 317.
[19] **Pierre Auger** Collaboration, A. Aab *et al. Phys. Rev. D* **93** no. 12, (2016) 122005.
[20] **LOFAR** Collaboration, S. Buitink *et al. Phys. Rev. D* **90** no. 8, (2014) 082003.
[21] **Tunka-Rex** Collaboration, P. A. Bezyazeekov *et al. Phys. Rev. D* **97** no. 12, (2018) 122004.
[22] **IceCube** Collaboration, S. Verpoest *PoS* **ECRS** (2023) 074.
[23] **Pierre Auger** Collaboration, A. Aab *et al. JCAP* **03** (2019) 018.
[24] **Pierre Auger** Collaboration, A. Aab *et al. Astropart. Phys.* **95** (2017) 44–56.
[25] S. Buitink *et al. PoS* **ICRC2021** (2021) 415.
[26] T. Pierog *et al. Phys. Rev. C* **92** no. 3, (2015) 034906.
[27] S. Ostapchenko *EPJ Web Conf.* **52** (2013) 02001.
[28] **GCOS** Collaboration, R. Alves Batista *et al. PoS* **ICRC2023** (these proceedings) 281.