Eur. Phys. J. C (2025) 85:130
https://doi.org/10.1140/epjc/s10052-025-13861-3

THE EUROPEAN PHYSICAL JOURNAL C

Regular Article - Theoretical Physics

Ultra high energy cosmic rays versus models of high energy hadronic interactions

Blaž Bortolato[1,2,a], **Jernej F. Kamenik**[1,2,b], **Michele Tammaro**[3,c]

[1] Jožef Stefan Institute, Jamova 39, 1000 Ljubljana, Slovenia
[2] Faculty of Mathematics and Physics, University of Ljubljana, Jadranska 19, 1000 Ljubljana, Slovenia
[3] INFN Sezione di Firenze, Via G. Sansone 1, 50019 Sesto Fiorentino, Italy

Received: 6 December 2024 / Accepted: 19 January 2025
© The Author(s) 2025

Abstract We evaluate the consistency of hadronic interaction models in the CORSIKA simulation package with publicly available fluorescence telescope data from the Pierre Auger Observatory. By comparing the first few central moments of the extensive air shower depth maximum distributions, as extracted from measured events, to those predicted by the best-fit inferred compositions, we derive a statistical measure of the consistency of a given hadronic model with data. To mitigate possible systematic biases, we include all primaries up to iron, compensate for the differences between the measured and simulated energy spectra of cosmic rays and account for other known systematic effects. Additionally, we study the effects of including higher central moments in the fit and project our results to larger statistics.

1 Introduction

The study of ultra-high energy cosmic rays (UHECRs), particles with energies exceeding 10^{18} eV incident on Earth, is crucial for understanding the most energetic processes in the Universe. Key questions in UHECR research include their origin, propagation mechanisms, and composition. The origin of UHECRs is still debated, with potential sources ranging from active galactic nuclei to gamma-ray bursts [1,2]. Understanding their propagation involves considering interactions with cosmic microwave background photons and (inter)galactic magnetic fields, which crucially affect their trajectories and energy spectra. Additionally, the composition of UHECRs, whether primarily protons, or a mixture of heavier nuclei, has significant implications for both their sources, acceleration and propagation [3].

Detection of UHECRs currently relies on large ground based observatories such as the Pierre Auger Observatory [4,5] and the Telescope Array [6], comprised of ground detectors (water Cherenkov detectors and scintillators, respectively) and fluorescence telescopes. They measure various observables related to extensive air showers (EAS) produced by UHECRs, from which the energy, arrival direction, and the composition of primary cosmic rays can be inferred.

Interpreting the data from these observatories requires detailed simulations of EAS initiated by UHECRs in the Earth's atmosphere. These simulations depend crucially on models of hadronic interactions, which describe the behavior of particles produced in high-energy collisions. The accuracy of these models is essential for reliable interpretation of experimental data.

Historically, phenomenological hadronic interaction models suffered from theoretical inconsistencies leading to tensions with relevant accelerator data and biasing the model predictions for EAS properties [7]. Recent developments [8–11] however, have improved our understanding of particle interactions at extreme energies. Models are continually calibrated and refined using data from both cosmic ray experiments and particle accelerators.

Nonetheless, some discrepancies persist, particularly in the observed ground-level muon multiplicity, which appears inconsistent with predictions [12]. Conversely, measurements from fluorescence telescopes show better agreement with simulations, highlighting a possible venue for further refinement of hadronic interaction models.

Our recent work has focused on developing novel observables, such as the central moments of the shower depth maximum ($X_{\max}$) distributions recorded by the fluorescence

[a] e-mail: blaz.bortolato@ijs.si
[b] e-mail: jernej.kamenik@cern.ch
[c] e-mail: michele.tammaro@fi.infn.it (corresponding author)

Published online: 03 February 2025

detectors, to infer the mass composition of UHECRs [13–15]. These observables demonstrated differentiating sensitivity to hadronic models, providing a potential pathway to resolve current discrepancies and enhance our understanding of UHECR composition and interactions.

In this work, we carefully reexamine the consistency of different hadronic models with $X_{\rm max}$ measurements. Previously [13], we have shown that for the limited statistics of the Pierre Auger Open Data, the first three $X_{\rm max}$ central moments suffice to infer the full UHECR mass composition. Here we extend this study by projecting to larger statistics, relevant for the analysis of the full Pierre Auger dataset. In particular, we focus on the sensitivity of higher $X_{\rm max}$ moments to hadronic interaction models. By marginalizing over possible UHECR mass compositions we construct a robust statistical measure to differentiate between and assess the validity of different models.

The remainder of the article is structured as follows. In Sect. 2 we review the inference procedure for UHECR mass composition, with the respective likelihood based on $X_{\rm max}$ moments, and discuss the interplay of the CR energy spectrum, mass composition and hadronic interaction model effects. In Sect. 3 we present and discuss our results on hadronic model discrimination, while our main conclusions are summarized in Sect. 4.

2 Methodology

Following our previous work [13], we infer the full mass composition[1] of UHECR through the first n central moments of the $X_{\rm max}$ distribution. The statistical method based on bootstrapping allows to include in a natural way both systematic uncertainties from detector effects, and statistical errors from the finite size of data and simulated events available. Moreover, it provides a natural framework to make projections to larger statistics, in case the number of events is increased by a multiplicative factor f [14], as explained below.

2.1 Pierre Auger Open Data

We use publicly available data from the 2021 Pierre Auger Open Data (PAOD) release [16], consisting of ∼10% of the entire dataset collected by the Pierre Auger Observatory. In order to study $X_{\rm max}$, we are restricted to the use of hybrid showers, that is events recorded by both the Surface Detectors (SD) and the Fluorescent Detectors (FD); furthermore, we restrict our analysis to the energy bin $E \in [0.65, 1]$ EeV, containing 934 events. In this way we can select a statistical significant sample while limiting possible systematic errors from the energy dependence of the observables when comparing to simulations, as discussed in more details below. Systematic uncertainties on the measured $X_{\rm max}$ distributions have as main sources the detector calibration and the reconstruction of the data, as detailed in Ref. [17]. These amount to an error ~ 10 g/cm^2 for all energies. Larger uncertainties from the detector resolution are instead of statistical nature; fluctuations in the number of photo-electrons detected and in the shower arrival directions lead to a ~ 25 g/cm^2 systematic uncertainty on the observed $X_{\rm max}$ at low energies, while it decreases to ~ 15 g/cm^2 at higher energies. These values are expected to decrease in the future, as the larger dataset collected will improve the detector resolution spread.

2.2 Comparing with EAS simulations

In parallel, we simulate the $X_{\rm max}$ distributions of extensive air showers (EAS) generated by different primaries using CORSIKA v7.7550 [18]. For simplicity, we assume a uniform energy distribution within the bin $E \in [0.65, 1]$ EeV and consider all 26 primaries from proton ($Z = 1$) to iron ($Z = 26$), selecting the most abundant isotope for each atomic number Z. While current experimental data do not permit resolving the partial abundances of the considered primaries, including at least 16 primaries within this range is recommended to mitigate systematic biases associated with the choice of primary components [15]. We simulate 2000 showers per primary, focusing on their longitudinal electromagnetic energy profiles for each of the hadronic interaction models (EPOS [19], Sibyll [20], QGSJetII-04 [21], and QGSJet01 [22]).[2] Finally, we include systematic uncertainties in the simulated $X_{\rm max}$ distributions. The smearing from the detector resolution can be added following the procedure in Ref. [17]. Note that this effect is statistics limited and that the fit of the smearing parameters is based on the current PA dataset. We do not attempt to estimate the expected improvement on these parameters from the present and future data collection, thus projections to statistics beyond the present PA data can be considered conservative.

Another potentially important source of systematic uncertainty comes from the energy spectra of primaries, since it is well known that $X_{\rm max}$ distributions exhibit distinct primary energy dependence. Thus simulations based on a uniform energy distribution cannot be directly compared with PA data following steeply falling spectra. Furthermore, current EAS energy measurements only determine the energy spectrum of

[1] We define the full composition as $w = (w_p, \ldots, w_{Fe})$, that is with a total of 26 primaries, from proton (p) to iron (Fe). For a recent discussion on the choice of composition components see Ref. [15].

[2] Recently, new implementations of these models have been developed and will be included in CORSIKA: EPOS4 [23], Sibyll★ [20], and QGSJetIII [8]. Some of their predictions may be different with respect to the present versions.

all primaries combined. The indications of an energy dependent composition of UHECRs however point towards distinct energy spectra of light and heavier primaries, as one would expect if different mechanisms are responsible for the creation and/or acceleration of light and heavy primaries.

To illustrate this point, we plot in Fig. 1 the energy dependence of the first three central moments computed using the EPOS hadronic model for protons (solid lines) and iron (dashed lines). We observe the well known logarithmic dependence of $z_1 \equiv \langle X_{\rm max} \rangle$ while higher moments do not show any significant energy dependence on these scales. We also note that even for z_1, the energy dependence is small compared to the total range of values between light and heavy primaries. In absence of any energy dependence, the composition inferred from data in a finite energy interval would simply be the average composition. The energy dependence of z_1 thus introduces a systematic bias. In order to determine whether it significantly affects the inferred composition and its likelihood, we assume in the following that the spectra of all primaries are identical within the chosen narrow energy interval of 0.65 to 1 EeV. Even in this case, the uniform energy spectrum of the simulated data for each primary differs from the measured data. We can however account for this bias by reweighing the measured events so that their spectrum matches the one simulated (flat). We can then compare inferred compositions with and without reweighing to estimate the significance of this bias. The weights can be assigned to each measured event as the inverse of the energy spectrum density of the event itself, $u_i = 1/P(E_i)$. The energy spectrum density can be estimated in various ways, such as binning, template (i.e. power law) fitting, or using techniques like kernel density estimation. In this work, we estimate $P(E)$ by binning PAOD events. We split the relevant energy interval into 30 bins of equal length and compute the binned energy spectrum density by counting all events in a bin (N_b) and dividing by the total number of all events (N), $P(E)_b = N_b/N$.[3] For each event, we then determine the corresponding energy bin and assign it a weight equal to the inverse of the energy spectrum density of that bin, $1/P(E_i) \simeq 1/P(E)_b$.

2.3 Probing hadronic models

To compare how well a specific hadronic model describes the measured data, considering n central moments, we first compute the best composition for the case at hand. We start from the log-likelihood function

$$\ln \mathcal{L}(w) = \int \ln \Big[p\,(z|w) \Big]\; P\,(z)\; \mathrm{d}z, \tag{1}$$

where $p\,(z|w)$ is the probability distribution function (PDF) of moments $z = (z_1, \ldots, z_n)$, computed from simulated data for a given composition w; similarly, $P(z)$ is the PDF of moments z, computed from measured data (PAOD).

These distributions are not well-determined due to systematic and statistical uncertainties of the measured and simulated events, which in turn affect the $X_{\rm max}$ moments, the energy weights and ultimately the mass composition. We account for these uncertainties via the bootstrap method [24]. Namely, we use random uniform sampling, with repetitions allowed, to compute moments of the distributions, and repeat the procedure for $B = 10^4$ times. For a fixed energy bin, which includes N showers, the procedure is as follows:

1. Randomly choose N events, with uniform probability and allowing for repetitions. We denote the selected events as $\{(\tilde{E}_i, \tilde{X}_{{\rm max},i}, \delta\tilde{X}_{{\rm max},i})\}_\ell$, with $i = 1, \ldots, N$. Here ℓ is the current bootstrap step, while $\delta\tilde{X}_{{\rm max},i}$ is the total systematic uncertainty of $\tilde{X}_{{\rm max},i}$.
2. Estimate the energy density, $P(\tilde{E}_i)$, and compute the weight for each event, $\{u_i\} = 1/P(\tilde{E}_i)$.
3. Compute the moments:

$$z_1^\ell = \frac{\sum_{i=1}^N \left(\tilde{X}_{{\rm max},i} + \delta\tilde{X}_{{\rm max},i}\,\epsilon_i\right) u_i}{\sum_{i=1}^N u_i^l}, \tag{2}$$

$$z_n^\ell = \frac{\sum_{i=1}^N \left(\tilde{X}_{{\rm max},i} + \delta\tilde{X}_{{\rm max},i}\,\epsilon_i - z_1^l\right)^n u_i}{\sum_{i=1}^N u_i}, \tag{3}$$

   where ϵ_i are samples from a normal distribution with mean 0 and standard deviation 1.
4. Repeat the previous steps for $\ell = 1, \ldots, B$.

The procedure to compute central moments from data remains the same as described in Ref. [13].

For simulated events, this procedure needs to be carried out for each primary separately, with $N = N_{\rm sim} = 2000$ the total number of events in the simulated dataset; the moments of each primary then receive a different weight depending on the composition w [13]. When analyzing a real dataset instead, N corresponds to the actual number of observed events in the chosen energy bin; in our case, we have $N = N_{\rm PAOD} = 934$ for hybrid events with $E_i \in [0.65, 1]$ EeV in PAOD. The bootstrap then yields a distribution of observed moments, $P(z)$, to which we fit the distribution for simulated events to infer w. The variance of $P(z)$ strongly depends on the available statistics, that is $N_{\rm PAOD}$; we can make projections to larger statistics by sampling f-times the number of events in the actual PAOD dataset, where f is a statistical multiplicative factor, such that $N = f N_{\rm PAOD}$. Concretely, this factor represents an effective reduction in

[3] With the chosen binning, the number of events per energy bin, N_b, ranges between 20 and 60 for the interval $E \in [0.65, 1]$ EeV.

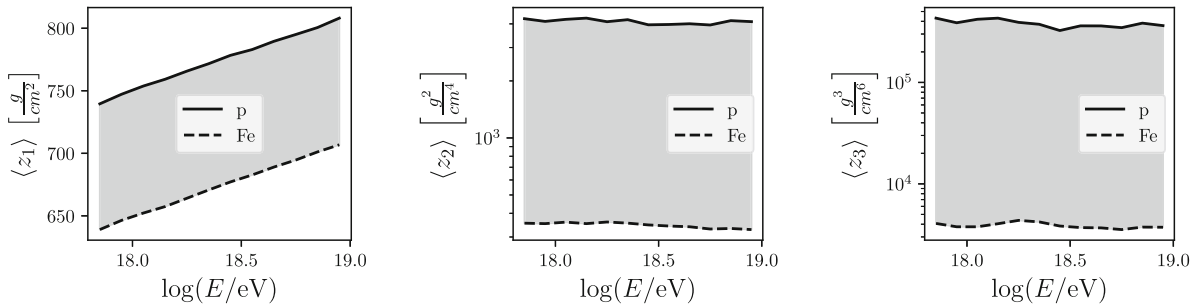

Fig. 1 The averaged value of the moments z_1 (top left), z_2 (top right) and z_3 (bottom), for proton (solid line) and Fe (dashed line), as a function of $\log_{10}(E/\text{eV})$. Simulated using the EPOS hadronic model

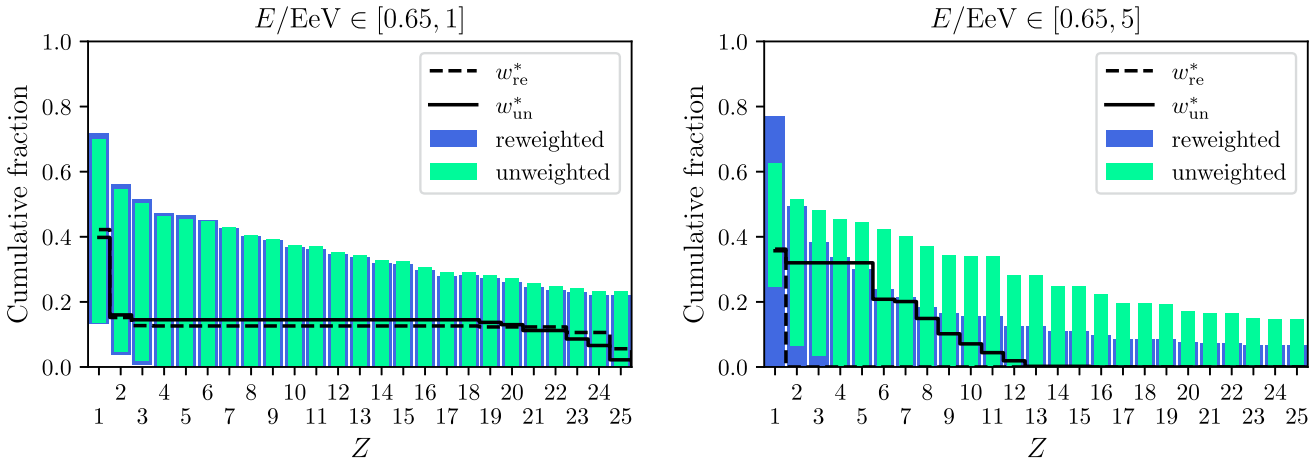

Fig. 2 Best fit of the cumulative composition and the respective confidence intervals, in two energy bins, [0.65, 1] EeV (Left) and [0.65, 5] EeV (Right), obtained with the EPOS hadronic model. The solid black line and green bars indicate the results obtained without introducing any weights for the energy distribution of measured events, while the dashed black line and blue bars show the results derived by including the reweighting procedure described in the text

the variance of $P(z)$ that would follow a larger statistical sample, assuming the central value remains unchanged.

Due to the bootstrap method, both $p(z|w)$ and $P(z)$ follow a multivariate normal distribution; that is $p(z|w) = \mathcal{N}_n(\mu_s, \Sigma_s)$ and $P(z) = \mathcal{N}_n(\mu_m, \Sigma_m)$, where μ_s and μ_m (Σ_s and Σ_m) indicate the vector of means (covariance matrix) obtained from simulations and measurements, respectively. The most probable composition, w^*, can then be computed by maximizing the log-likelihood in Eq. (1); given the properties of $p(z|w)$ and $P(z)$ discussed above, the integral can be computed analytically [13]. To estimate the confidence level of the resulting w^*, we use Nested Sampling techniques to efficiently sample the log-likelihood around its maximum [25–28]. Note that in general these results depend on both the number of moments included, n, and the statistical multiplier assumed, f.

Finally, we compute the confidence level of rejecting a given hadronic model, as function of n and f, by comparing the distribution of moments for the best fit composition, $p(z|w^*)$, to the measured one, $P(z)$. As the difference between random variables of two multivariate normal distributions also follows a multivariate normal distribution, $z^s - z^m \sim \mathcal{N}(\mu_s - \mu_m, \Sigma_s + \Sigma_m)$, we can compute the rejection power by evaluating the expression

$$\chi^2_n = (\mu_s - \mu_m)^T (\Sigma_s + \Sigma_m)^{-1} (\mu_s - \mu_m), \tag{4}$$

which follows a χ^2 distribution with n degrees of freedom.

3 Results

We first test the sensitivity of our approach to the size of the chosen energy interval by comparing the inferred most probable composition (and its likelihood) when accounting for, or not, the energy spectrum of measured events (i.e. using weights $u_i = 1/P(E_i)$ or $u_i = 1$, respectively). We find no significant difference between the inferred composition with original or energy reweighted events for the interval $E_i \in [0.65, 1]$ EeV, see left panel of Fig. 2. This justifies the choice of the energy interval size, indicating that the

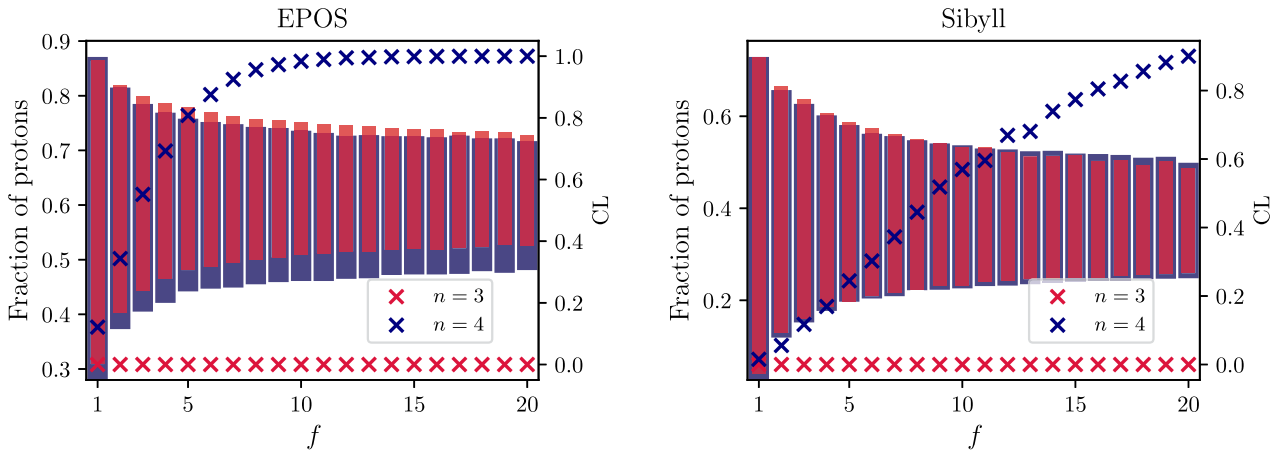

Fig. 3 Inferred fraction of protons (left axis) and model CL of rejection (right axis) as function of the statistics multiplier f, using EPOS (left plot) and Sibyll (right plot). Red and blue bars show the proton fraction results with $n = 3$ and $n = 4$ respectively, while crosses refer to the CL, with the same color code

described systematic bias does not significantly impact the results. On the other hand, selecting a larger energy bin would amplify this effect, as shown in the right panel of Fig. 2, where we consider all events in the energy bin [0.65, 5] EeV. To have a quantitative estimate of the associated systematic bias, we can compute the (energy reweighted) log-likelihood ($\log \mathcal{L}_{\rm re}$) ratio of the most probable compositions as inferred using reweighted ($w^*_{\rm re}$) and unweighted ($w^*_{\rm un}$) fits, assuming some statistical multiplier f. Namely, we compute

$$\lambda(f) = \frac{\log \mathcal{L}_{\rm re}(w^*_{\rm un}, f)}{\log \mathcal{L}_{\rm re}(w^*_{\rm re}, f)}. \tag{5}$$

In the small $E_i \in [0.65, 1]$ EeV bin, we obtain $\lambda(f = 1) \simeq 0.99$ with the PAOD statistics, and only decreases to $\lambda(f = 5) \simeq 0.97$. Conversely, for the choice of a larger energy bin, $E_i \in [0.65, 5]$ EeV, the log-likelihood ratio is already $\lambda(f = 1) \simeq 0.29$ with current statistics, and goes down to $\lambda(f = 5) \simeq 0.1$ assuming a dataset five times larger. We thus conclude, that while the effect is certainly important for large energy intervals, it is completely negligible for narrow enough bins, both in current PAOD as well as in anticipated larger statistics event samples.

The main result of this work is shown in Fig. 3, choosing the EPOS and Sibyll hadronic models as case studies.[4] The fraction of protons inferred with $n = 3$ (red) and $n = 4$ (blue) moments is shown as a function of f. As expected, the determination of w_p becomes more and more precise with increasing f, that is with increasing statistics, while the introduction of an additional moment in the inference procedure has a weak effect (see also Ref. [13]). At the same time, we show on the right y-axis the rejection confidence levels, with crosses that follow the same color code as for the bars, computed with Eq. (4). While the $n = 3$ case is fully consistent with the PAOD across the entire shown range of f, the inclusion of the fourth moment yields a rejection CL that is monotonically increasing. At $f = 1$, that is the actual PAOD, we have $CL \simeq 20\%$, which however grows rapidly, reaching $CL \gtrsim 90\%$ already at $f \lesssim 10$. In other words, already with the existing full PA dataset and by including additional information into the inference procedure, the fit obtained with the EPOS model could be statistically inconsistent with the data. A similar behaviour can be observed with the Sibyll hadronic model, see the right plot in Fig. 3, although the rejection level grows more slowly with f.

To detail further the effect of the number of moments and statistics on the rejection of hadronic models, we show in Fig. 4 the results for CL by adding an additional moment to the likelihood, that is $n = 5$, for both EPOS and Sibyll cases. Clearly, the rate of growth of model rejection confidence level increases significantly with the inclusion of an extra observational constraint, reaching $CL \gtrsim 95\%$ already at $f = 5$ for both models.

Finally, we can explicitly compute the deviations of the single best fit moments from the respective PAOD measured moments and thus identify the effects of including higher moments on the individual pulls in the likelihood fit. In Fig. 5 we show the first three moments of the best fit composition, z^*_i with $i = 1, 2, 3$, obtained with the EPOS model, with respect to the PAOD moments, $z^{\rm PA}_i$. We show the results for three different choices of n, $n = 3, 4, 5$, and two choices of f, $f = 1, 20$. The top plot, referring to the first moment z_1, indicates that the mean of the inferred $X_{\rm max}$ distribution is consistent with the data (the vertical black line) for all the cases considered. That is, in general the first moment is well reconstructed, even with several higher moments included in the fit and at much larger statistics than currently available. In contrast, the best fits for z_2 and z_3 start deviating from

[4] We present the same results for the two QGSJet models in Appendix A.

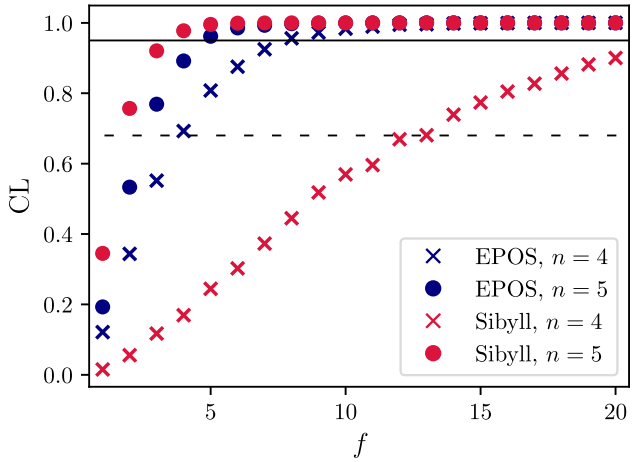

Fig. 4 Rejection CL of EPOS (blue) and Sibyll (red) as function of f, using $n = 4$ (cross) and $n = 5$ (dots) moments. The horizontal solid (dashed) line indicates the 68% (95%) confidence level

data once we include more moments, with a clear enhancement of the discrepancy when projecting to larger statistics. In the future, such analysis could thus point to particular tensions within fits to hadronic models failing to simultaneously account for all included $X_{\max}$ distribution moments.

4 Conclusions

This work examines the consistency of hadronic interaction models implemented in the CORSIKA simulation package (EPOS, Sibyll, QGSJetII-4 and QGSJet01), with the publicly

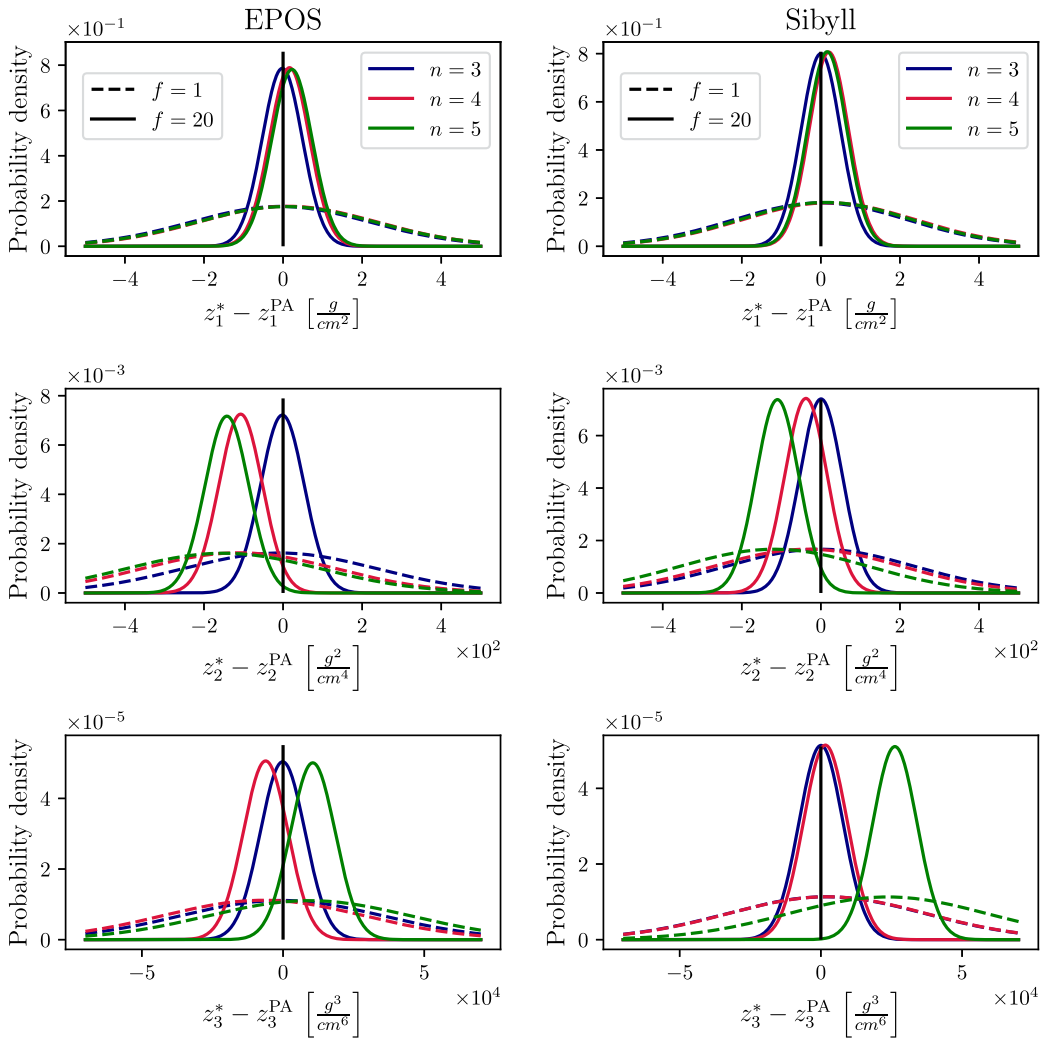

Fig. 5 Deviation of the first three moments of the best fit composition, z_i^*, from the PAOD moments, z_i^{PA}, using the EPOS (left column) and Sibyll (right column) models. Blue, red and green indicate the use of $n = 3, 4, 5$ moments respectively, while solid and dashed lines refer to $f = 1$ and $f = 20$

available fluorescence telescope data measured at the Pierre Auger Observatory.[5] For each model, the first n moments of the $X_{\max}$ distribution extracted from measured events are compared to the moments given by the best (most probable) inferred compositions, including all primaries up to iron and after accounting for all other known possible systematic effects. A statistical measure of consistency based on this comparison then allows us to compute confidence levels of rejecting a particular hadronic model. The results are furthermore projected to (future) larger datasets by assuming to have f - times more measured (and simulated) events.

As a byproduct of our analysis, the systematics associated with the unknown primary CR spectra and finite energy bin intervals are considered. We show that the effects of the overall CR energy spectrum can be accounted for by properly reweighing the measured events included in the fit to match the spectrum of simulated events. We also find that these effects are in practice negligible for narrow enough energy bins, for which the inferred composition can thus be interpreted as averaged over the relevant energy interval.

For EPOS and Sibyll models, we find that the best fit compositions are consistent with currently available public data. As long as we restrict the inference procedure to the first $n = 3$ moments the conclusion persists even when projecting to much larger datasets (up to $f = 20$). Once additional moments are introduced however, the rejection CL grows with projected increased statistics, highlighting a possible incompatibility of hadronic model predictions with measured data. Since our projections are based on fixed central values of measured moments, this does not mean that these models will in fact be incompatible once more data is analyzed. Instead it demonstrates the sensitivity of the existing full PA dataset (and its future upgrades) to discriminate between hadronic interaction models that currently best describe the available data.[6]

The method presented here can hopefully lay the foundations for more in-depth studies to test and improve hadronic interaction models employed in simulations of UHECR initiated atmospheric shower development. In turn this could shed light on existing discrepancies observed in ground detector data as well as improve our confidence in the primary composition spectra already inferred from measurements with fluorescence telescopes.

Acknowledgements B.B. and J.F.K. acknowledge the financial support from the Slovenian Research Agency (Grant no. J1-3013 and research core funding No. P1-0035).

Data Availability Statement Data will be made available on reasonable request. [Authors' comment: The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.]

Code Availability Statement Code/software will be made available on reasonable request. [Authors' comment: The code/software generated during and/or analysed during the current study is available from the corresponding author on reasonable request.]

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.
Funded by SCOAP[3].

Appendix A: Results for QGSJet models

See Figs. 6, 7 and 8.

[5] While in this work we employed the, at present, only publicly available dataset from the Pierre Auger Observatory, the same analysis can certainly be performed on $X_{\max}$ measurements from other observatories, such as Telescope Array [6].

[6] For the QGSJetII and QGSJet01 models, we show in Appendix A that already the first and the second moment, respectively, lead to poor fits to measured data.

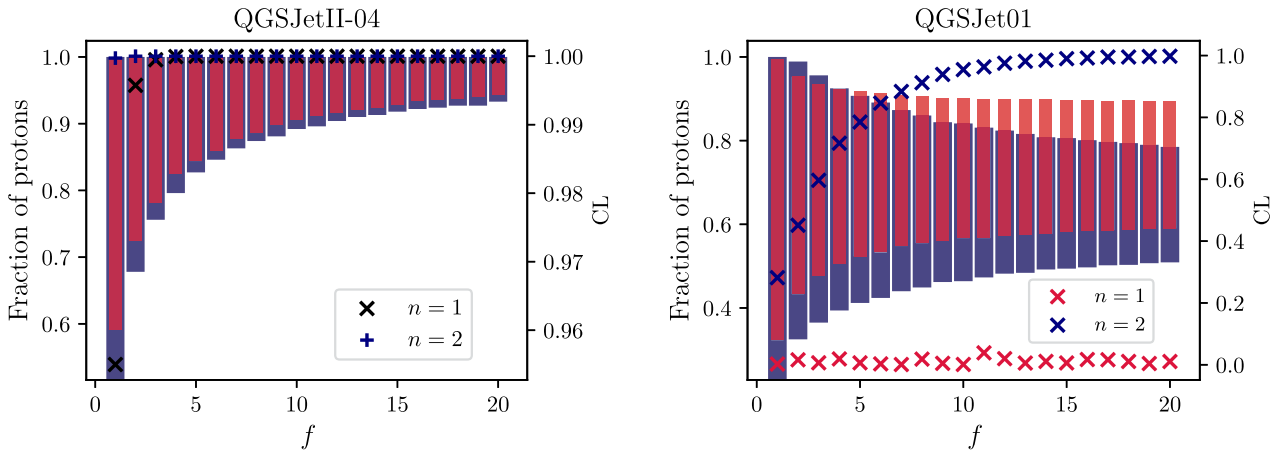

Fig. 6 Inferred fraction of protons (left axis) and model CL of rejection (right axis) as function of the statistics multiplier f, using QGSJetII-04 (left plot) and QGSjet01 (right plot). Red and blue bars show the proton fraction results with $n = 1$ and $n = 2$ respectively, while crosses refer to the CL, with the same color code except for QGSJetII-04

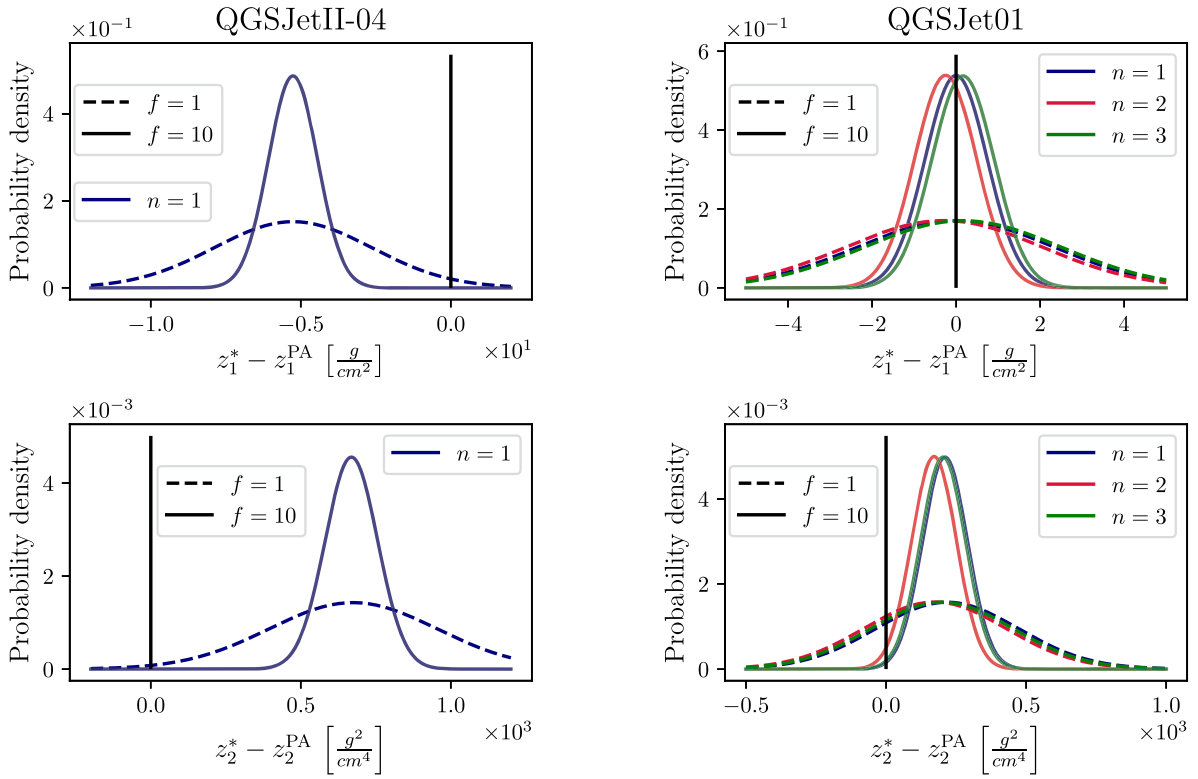

Fig. 7 Deviation of the first two moments of the best fit composition, z_i^*, from the PAOD moments, z_i^{PA}, using the QGSJetII-04 (left column) and QGSJet01 (right column) models. Blue, red and green indicate the use of $n = 1, 2, 3$ moments respectively, while solid and dashed lines refer to $f = 1$ and $f = 10$. For QGSJetII-04 the deviation of the first and second moment are the same for $n = 1, 2, 3$

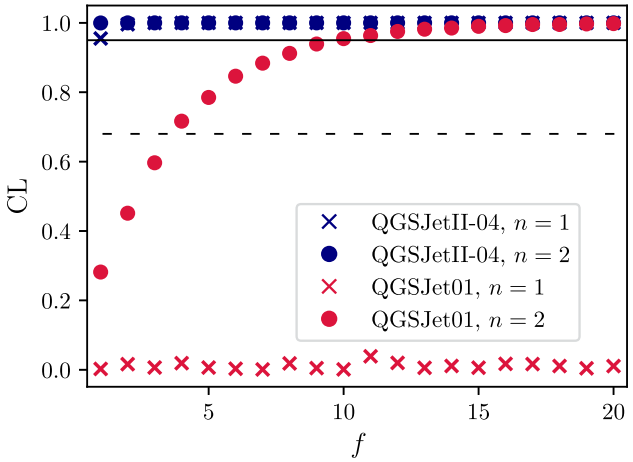

Fig. 8 Rejection CL of QGSJetII-04 (blue) and QGSJet01 (red) as function of f, using $n = 1$ (cross) and $n = 2$ (dots) moments. The horizontal solid (dashed) line indicates the 68% (95%) confidence level

References

1. G. Bertone, C. Isola, M. Lemoine, G. Sigl, Phys. Rev. D **66**, 103003 (2002). https://doi.org/10.1103/PhysRevD.66.103003. arXiv:astro-ph/0209192
2. J.R. Horandel, N.N. Kalmykov, A.V. Timokhin, Astropart. Phys. **27**, 119 (2007). https://doi.org/10.1016/j.astropartphys.2006.09.011. arXiv:astro-ph/0609490
3. S. Andringa, R. Conceicao, M. Pimenta, Astropart. Phys. **34**, 360 (2011). https://doi.org/10.1016/j.astropartphys.2010.10.002
4. A. Aab et al. (Pierre Auger), Nucl. Instrum. Methods A **798**, 172 (2015). https://doi.org/10.1016/j.nima.2015.06.058. arXiv:1502.01323 [astro-ph.IM]
5. J. Stasielak (Pierre Auger), in *9th International Conference on New Frontiers in Physics* (2021). arXiv:2110.09487 [astro-ph.HE]
6. H. Kawai, S. Yoshida, H. Yoshii, K. Tanaka, F. Cohen, M. Fukushima, N. Hayashida, K. Hiyama, D. Ikeda, E. Kido, Y. Kondo, T. Nonaka, M. Ohnishi, H. Ohoka, S. Ozawa, H. Sagawa, N. Sakurai, T. Shibata, H. Shimodaira, M. Takeda, A. Taketa, M. Takita, H. Tokuno, R. Torii, S. Udo, Y. Yamakawa, H. Fujii, T. Matsuda, M. Tanaka, H. Yamaoka, K. Hibino, T. Benno, K. Doura, M. Chikawa, T. Nakamura, M. Teshima, K. Kadota, Y. Uchihori, K. Hayashi, Y. Hayashi, S. Kawakami, T. Matsuyama, M. Minamino, S. Ogio, A. Ohshima, T. Okuda, N. Shimizu, H. Tanaka, D. Bergman, G. Hughes, S. Stratton, G. Thomson, A. Endo, N. Inoue, S. Kawana, Y. Wada, K. Kasahara, R. Azuma, T. Iguchi, F. Kakimoto, S. Machida, K. Misumi, Y. Murano, Y. Tameda, Y. Tsunesada, J. Chiba, K. Miyata, T. Abu-Zayyad, J. Belz, R. Cady, Z. Cao, P. Huentemeyer, C. Jui, K. Martens, J. Matthews, M. Mostofa, J. Smith, P. Sokolsky, R. Springer, J. Thomas, S. Thomas, L. Wiencke, T. Doyle, M. Taylor, V. Wickwar, T. Wilkerson, K. Hashimoto, K. Honda, K. Ikuta, T. Ishii, T. Kanbe, T. Tomida, Nucl. Phys. B Proc. Suppl. **175–176**, 221 (2008). https://doi.org/10.1016/j.nuclphysbps.2007.11.002. Proceedings of the XIV International Symposium on Very High Energy Cosmic Ray Interactions
7. S. Ostapchenko, in *25th European Cosmic Ray Symposium* (2016). arXiv:1612.09461 [astro-ph.HE]
8. S. Ostapchenko, Phys. Rev. D **109**, 094019 (2024). https://doi.org/10.1103/PhysRevD.109.094019. arXiv:2403.16106 [hep-ph]
9. A.A. Halim et al. (Pierre Auger), JCAP **01**, 022 (2024). https://doi.org/10.1088/1475-7516/2024/01/022. arXiv:2305.16693 [astro-ph.HE]
10. J. Blazek, J. Vicha, J. Ebr, R. Ulrich, T. Pierog, P. Travnicek, PoS **ICRC2021**, 441 (2021). https://doi.org/10.22323/1.395.0441
11. S. Ostapchenkoand, G. Sigl. (2024). arXiv:2409.05501 [hep-ph]
12. L. Cazon (EAS-MSU, IceCube, KASCADE Grande, NEVOD-DECOR, Pierre Auger, SUGAR, Telescope Array, Yakutsk EAS Array), PoS **ICRC2019**, 214 (2020). https://doi.org/10.22323/1.358.0214. arXiv:2001.07508 [astro-ph.HE
13. B. Bortolato, J.F. Kamenik, M. Tammaro, Phys. Rev. D **108**, 022004 (2023). https://doi.org/10.1103/PhysRevD.108.022004. arXiv:2212.04760 [astro-ph.HE]
14. B. Bortolato, J.F. Kamenik, M. Tammaro, Phys. Rev. D **109**, 043023 (2024). https://doi.org/10.1103/PhysRevD.109.043023. arXiv:2304.11197 [astro-ph.HE]
15. B. Bortolato, J.F. Kamenik, M. Tammaro. (2024b). arXiv:2409.06841 [astro-ph.HE]
16. T. P. A. Collaboration, Pierre Auger Observatory 2021 open data. (2021). https://doi.org/10.5281/zenodo.4487613
17. A. Aab et al. (Pierre Auger), Phys. Rev. D **90**, 122005 (2014). https://doi.org/10.1103/PhysRevD.90.122005. arXiv:1409.4809 [astro-ph.HE]
18. D. Heck, J. Knapp, J.N. Capdevielle, G. Schatz, T. Thouw, CORSIKA: a Monte Carlo code to simulate extensive air showers. (1998). https://doi.org/10.5445/IR/270043064
19. T. Pierog, I. Karpenko, J.M. Katzy, E. Yatsenko, K. Werner, Phys. Rev. C **92**, 034906 (2015). https://doi.org/10.1103/PhysRevC.92.034906. arXiv:1306.0121 [hep-ph]
20. F. Riehn, A. Fedynitch, R. Engel, Astropart. Phys. **160**, 102964 (2024). https://doi.org/10.1016/j.astropartphys.2024.102964. arXiv:2404.02636 [hep-ph]
21. S. Ostapchenko, Phys. Rev. D **83**, 014018 (2011). https://doi.org/10.1103/PhysRevD.83.014018. arXiv:1010.1869 [hep-ph]
22. N.N. Kalmykovand, S.S. Ostapchenko, Phys. Atom. Nucl. **56**, 346 (1993)
23. K. Werner. (2024). arXiv:2410.09955 [hep-ph]
24. B. Efronand, R.J. Tibshirani, *An Introduction to the Bootstrap, Monographs on Statistics and Applied Probability*, vol. 57 (Chapman & Hall/CRC, Boca Raton, 1993)
25. J. Skilling, AIP Conf. Proc. **735**, 395 (2004). https://doi.org/10.1063/1.1835238
26. J. Buchner, Stat. Comput. **26**, 383 (2014). https://doi.org/10.1007/s11222-014-9512-y
27. J. Buchner, Collaborative nested sampling: big data vs. complex physical models. (2017). https://doi.org/10.48550/ARXIV.1707.04476
28. J. Buchner, Ultranest – a robust, general purpose Bayesian inference engine. (2021). https://doi.org/10.48550/ARXIV.2101.09604