

# Testing Hadronic-Model Predictions of Depth of Maximum of Air-Shower Profiles and Ground-Particle Signals using Hybrid Data of the Pierre Auger Observatory

The Pierre Auger Collaboration  
The Pierre Auger Observatory, Av. San Martín Norte 306,  
5613 Malargüe, Mendoza, Argentina;  
<http://www.auger.org>\*  
(Dated: January 22, 2024)

We test the predictions of hadronic interaction models regarding the depth of maximum of air-shower profiles,  $X_{\max}$ , and ground-particle signals in water-Cherenkov detectors at 1000 m from the shower core,  $S(1000)$ , using the data from the fluorescence and surface detectors of the Pierre Auger Observatory. The test consists in fitting the measured two-dimensional  $(S(1000), X_{\max})$  distributions using templates for simulated air showers produced with hadronic interaction models EPOS-LHC, QGSJET-II-04, SIBYLL 2.3d and leaving the scales of predicted  $X_{\max}$  and the signals from hadronic component at ground as free fit parameters. The method relies on the assumption that the mass composition remains the same at all zenith angles, while the longitudinal shower development and attenuation of ground signal depend on the mass composition in a correlated way.

The analysis was applied to 2239 events detected by both the fluorescence and surface detectors of the Pierre Auger Observatory with energies between  $10^{18.5}$  to  $10^{19.0}$  eV and zenith angles below  $60^\circ$ . We found, that within the assumptions of the method, the best description of the data is achieved if the predictions of the hadronic interaction models are shifted to deeper  $X_{\max}$  values and larger hadronic signals at all zenith angles. Given the magnitude of the shifts and the data sample size, the statistical significance of the improvement of data description using the modifications considered in the paper is larger than  $5\sigma$  even for any linear combination of experimental systematic uncertainties.

## I. INTRODUCTION

The dominant contribution to uncertainties in the determination of the mass composition of ultra-high-energy cosmic rays (UHECR, energy  $E > 10^{18.0}$  eV) comes from the modelling of extensive air showers. Modern hadronic interaction models used for this purpose are based on extrapolation of interaction parameters like cross-sections, multiplicities, elasticities, etc. measured at accelerators at lower beam energies up to  $\sqrt{s} = 13$  TeV for proton-proton collisions at the LHC and pseudorapidities  $|\eta| \lesssim 5$  compared to energies  $\sqrt{s} \gtrsim 50$  TeV and pseudorapidities<sup>1</sup>  $\eta \approx 7$  to 11 driving the energy flow of the first interactions of UHECR in the atmosphere, where the target is different (mostly oxygen and nitrogen nuclei). Therefore, improvements in the description of the LHC data, implemented in the modern models do not necessarily lead to unambiguous, nearly hadronic-model independent, predictions for the mass-sensitive air-shower observables. For instance, at  $10^{18.7}$  eV the span in predictions for the mean depth of maximum of air-shower profiles,  $\langle X_{\max} \rangle$ , between models used by the UHECR community (EPOS-LHC [1], QGSJET-II-04 [2], SIBYLL 2.3d [3]) is  $\sim 25$  g/cm<sup>2</sup>, nearly independently of the primary particle mass and energy. Such a difference can be considered only a lower limit on the systematic uncertainty of the predicted  $X_{\max}$  scale. This is about one-quarter of the difference between  $\langle X_{\max} \rangle$  values of the two astrophysical extremes, protons and iron nuclei. As a consequence, the mass composition of cosmic rays can be referred only with respect to  $X_{\max}$  scale predicted by a par-

ticular model. The largest differences in the predicted  $\langle X_{\max} \rangle$  between the models, with a minimal impact on the elongation rate, come from the properties of the first hadronic interaction and production of nucleons-antinucleons in pion-air and kaon-air interactions that are not known well in the relevant kinematic region (for more detailed discussion, see e.g. Ref. [4]).

The difference in the standard deviation of  $X_{\max}$  distributions,  $\sigma(X_{\max})$ , between the models is within  $\sim 5$  g/cm<sup>2</sup>, whereas the difference between the  $X_{\max}$  fluctuations of protons and iron nuclei is  $\sim 40$  g/cm<sup>2</sup>. Therefore, the difference of  $\sigma(X_{\max})$  in model predictions has a smaller effect on the mass composition inferences compared to the difference in predictions of  $\langle X_{\max} \rangle$ . Note, that there is no direct correspondence of the model scales of  $X_{\max}$  and  $\sigma(X_{\max})$ , i.e. differences in  $\sigma(X_{\max})$  are not a mere consequence of the differences in  $\langle X_{\max} \rangle$ .

In general, the signal produced by air-shower particles reaching the ground shows much lower sensitivity to the mass composition than in the case of  $X_{\max}$ . The model differences in predictions of the ground-particle signal, for instance, at 1000 m from the impact point of an air shower at  $10^{18.7}$  eV detected at the Pierre Auger Observatory (Auger) [5] are at the level of  $\sim 3$  VEM<sup>2</sup>, whereas the difference between protons and iron nuclei is about twice this value. The fluctuations of the ground signal are dominated by the detector resolution, suppressing any significant sensitivity to the model or primary mass [6].

\* [auger\\_spokespersons@fnal.gov](mailto:auger_spokespersons@fnal.gov)

<sup>1</sup> The forward calorimeters at LHC can cover  $\eta = 8.4$  to 15, but of neutral particles only.

<sup>2</sup> This unit is the signal produced by a muon traversing the station on a vertical trajectory.

TABLE I: Indicative summary of the results of tests of models using Auger data. In the case of SIBYLL 2.3d, we also show estimations based on the previous version of the model Sibyll. The red shade of the boxes reflects the size of the tension between the models and measurements.

test	energy/EeV	$\theta/^\circ$	EPOS-LHC	QGSJET-II-04	SIBYLL 2.3d
$X_{\max}$ moments [7–10]	$\sim 3$ to 50	0 to 80	no tension <span style="color: green;">■</span>	tension <span style="color: red;">■</span>	no tension (2.3c) <span style="color: green;">■</span>
$X_{\max}:S(1000)$ correlation [10, 11]	3 to 10	0 to 60	no tension <span style="color: green;">■</span>	tension <span style="color: red;">■</span>	no tension (2.3c) <span style="color: green;">■</span>
mean muon number [12, 13]	$\sim 10$	$\sim 67$	tension <span style="color: red;">■</span>	tension <span style="color: red;">■</span>	tension <span style="color: red;">■</span>
mean muon number [14]	0.2 to 2	0 to 45	tension <span style="color: red;">■</span>	tension <span style="color: red;">■</span>	—
fluctuation of muon number [13]	4 to 40	$\sim 67$	no tension <span style="color: green;">■</span>	no tension <span style="color: green;">■</span>	no tension <span style="color: green;">■</span>
muon production depth [15]	20 to 70	$\sim 60$	tension <span style="color: red;">■</span>	no tension <span style="color: green;">■</span>	—
$S(1000)$ [16]	$\sim 10$	0 to 60	tension <span style="color: red;">■</span>	tension <span style="color: red;">■</span>	—

### A. Problems in the Description of Data with Hadronic Interaction Models

The correctness of the model predictions can be tested using data from air-shower experiments. At the Pierre Auger Observatory, for instance, negative variance of the logarithm of primary masses and poor description of the measured  $X_{\max}$  distributions are obtained when simulations with QGSJET-II-04 are used for the interpretation of the measured  $X_{\max}$  moments for  $E \gtrsim 10^{18.5}$  eV [7–10]. The problem is that, due to relatively shallow  $\langle X_{\max} \rangle$  predictions of QGSJET-II-04, the best possible description of the data is achieved with proton-helium mixes, but for these mixes the modeled  $X_{\max}$  distributions are broader than the observed ones. These findings are further supported by the observation of the negative correlation between  $X_{\max}$  and the signal in surface detector (SD) [10, 11] in the Auger data near the “ankle” feature ( $E \approx 10^{18.7}$  eV) of the UHECR spectrum. That can be achieved only for the mixed primary composition containing particles heavier than helium nuclei.

The deficiencies of the models are more evident for the SD observables, where, in some cases, the data are not even bracketed by the Monte Carlo (MC) predictions for protons and iron nuclei. The muon deficit in simulations, known as the “muon puzzle” [17], is the best-known example of this kind. In particular, this was observed by Auger for inclined showers (zenith angles  $\theta = 62^\circ$  to  $80^\circ$ ), dominated by the muon component, for EPOS-LHC and QGSJET-II-04 at  $\sim 10^{19.0}$  eV [12], and in direct measurements with underground muon detectors at  $E = 10^{17.3}$  to  $10^{18.3}$  eV and  $\theta < 45^\circ$  [14]. At the same time, the fluctuations of the muon signal measured by Auger [13] are consistent with the MC predictions, indicating that the muon deficit might originate from the accumulation of small deviations from the model predictions during the development of a shower, rather than be caused by a strong deviation in the first interaction. The range of predictions for the muon production depth of protons and iron nuclei is outside of the measured values for EPOS-LHC above  $10^{19.3}$  eV [15].

Less directly, the muon deficit in simulations was observed as a deficit of the total signal at 1000 m from the shower core,  $S(1000)$ , in the Auger SD stations [16] for vertical ( $\theta < 60^\circ$ ) showers with energies around  $10^{19.0}$  eV. In this analysis,

within the assumption that the electromagnetic (em) component,  $X_{\max}$ , and, correspondingly, the mass composition inferred from the  $X_{\max}$  measurements, are predicted correctly by a particular model, the deficit of  $S(1000)$  was interpreted as the underestimation of the hadronic signal (dominated by muons) in simulations by  $(33 \pm 16)\%$  for EPOS-LHC and by  $(61 \pm 21)\%$  for QGSJET-II-04, with a strong dependence on the energy scale. An indicative summary of all these tests for the three models used in this work is given in Table I.

The use of an incorrect MC  $X_{\max}$  scale would lead one to a biased inference on the mass composition and, through this, to a biased estimate of the muon deficit, since the muon content in a shower scales with the primary mass  $\propto A^{1-\beta}$ ,  $\beta \approx 0.9$  [18]. Therefore, in a more comprehensive approach a modification of the MC scales of both  $X_{\max}$  and SD signals, going along with the fitting of the primary mass composition accounting for these modifications, should be considered.

### B. Progressive Testing of Hadronic Interaction Models

In this work, progressive testing of the model predictions is performed. First, we allow for a rescaling of the signal on the ground produced by the hadronic shower component at 1000 m with a factor  $R_{\text{had}}$ . Then we add a zenith-angle dependence of  $R_{\text{had}}$ , and, finally, a shift in the predicted  $X_{\max}$  distributions ( $\Delta X_{\max}$ ) which is assumed to be independent of the primary mass and energy. In this way, considering freedom not only in the scale of the simulated hadronic part of the ground signal but also in the simulated  $X_{\max}$  scale, we reduce the main differences in  $X_{\max}$  and  $S(1000)$  predictions of the models and obtain similar mass composition inferences for the Auger data. Remarkably, the modification of only two scale parameters is sufficient to obtain convergence of predictions of the 3 models for the composition of the Auger data.

The analysis is performed for the energy region around the ankle ( $E \approx 10^{18.7}$  eV) in the energy spectrum where the UHECR mass composition is mixed [8, 10, 11]. Specifically, we find the values of  $\Delta X_{\max}$ ,  $R_{\text{had}}(\theta)$  and fractions of four primary particles (protons, and helium, oxygen, iron nuclei) for which the best fit of the measured two-dimensional distributions of  $(S(1000), X_{\max})$  is achieved. The remaining differences between the predictions of models with a smaller effect

on MC templates like the fluctuations of  $X_{\max}$  and hadronic signal, and the mass composition dependence of  $R_{\text{had}}(\theta)$  and  $\Delta X_{\max}$  are not considered in this work.

In Section II, we give a detailed description of the method. In Section III, the method is applied to the data of the Pierre Auger Observatory. The results of the data analysis are discussed in Section IV followed by a summary of our findings in Conclusions.

## II. METHOD

The method stems from Ref. [19] and it was first introduced in Ref. [20] followed by a slight modification in the approach to the ground-signal rescaling. A preliminary application of the method to the Auger data was presented in Ref. [21].

We perform a binned maximum-likelihood fit of the measured two-dimensional distributions  $(S, X)$  with MC templates for four primary species simultaneously in five zenith-angle bins. The observables are  $S(1000)$  and  $X_{\max}$  corrected for the energy evolution using the fluorescence detector (FD) energy  $E_{\text{FD}}$ ,

$$S = S(1000) \left( \frac{E^{\text{ref}}}{E_{\text{FD}}} \right)^{1/B}, \quad (1)$$

and

$$X = X_{\max} + D \lg \left( \frac{E^{\text{ref}}}{E_{\text{FD}}} \right), \quad (2)$$

where  $B = 1.031 \pm 0.004$  is the SD energy calibration parameter [22] and the elongation rate of a single primary  $D = 58 \text{ g/cm}^2/\text{decade}$  is taken as the average value over the four primary particles and three models used in this work. The value of the elongation rate varies only within  $\pm 2 \text{ g/cm}^2$  around the mean, in accordance with the universality with respect to the primary mass predicted within the phenomenological model in Ref. [23]. This is a consequence of energy dependencies of multiplicity, elasticity, and cross-section assumed in this phenomenological model. We chose the reference energy  $E^{\text{ref}} = 10^{18.7} \text{ eV}$  for the analyzed energy range of  $E_{\text{FD}} = 10^{18.5}$  to  $10^{19.0} \text{ eV}$ .

The signal  $S(1000)$  is assumed to be composed of the hadronic ( $S_{\text{had}}$ ) and electromagnetic ( $S_{\text{em}}$ ) components. The signal  $S_{\text{had}}$  is produced by muons, em particles from muon decays and low-energy neutral pions as in [16] according to the four-component shower universality model [24, 25]. The signal  $S_{\text{em}}$  is produced by em particles originating from high-energy neutral pions.

### A. Monte Carlo Templates

The MC templates were prepared using simulated showers from a library produced within the Auger collaboration [26]. Air showers were generated with CORSIKA 7.7400 [27–29] with a flat zenith-angle distribution in  $\sin^2 \theta$  (for  $\theta \in (0^\circ, 65^\circ)$ ) and the energy distribution  $\propto E^{-1}$ . Subsequently,

we re-weight the events to match the measured energy spectrum [22]. Four different atmospheric profiles are considered to represent the typical variations of atmospheric conditions at the location of the Pierre Auger Observatory. The simulation set includes three models (EPOS-LHC, QGSJET-II-04, SIBYLL 2.3d) and four primary particles (p, He, O, Fe). In productions with EPOS-LHC and SIBYLL 2.3d, the low-energy ( $E_{\text{kin}} < 80 \text{ GeV}$ ) interactions were simulated with URQMD [30], while QGSJET-II-04 was used in combination with FLUKA [31, 32]. No significant dependency of the results on the choice of the low-energy model was found.

The detector simulations and event reconstruction were performed with the Auger Offline Framework [33]. In the standard event processing chain, not all effects related to the detector calibration, atmospheric conditions, long-term performances, etc. are included in the simulations and reconstruction. To account for them, for the FD part an additional smearing of  $X_{\max}$  distributions by  $\approx 9 \text{ g/cm}^2$  is applied [8]. In the case of the SD part, the smearing of  $S(1000)$  by 9%, corresponding to the maximal expected contribution from realistic operational conditions, was tested without finding any statistically significant effect on the results. The event selection is the same as applied to the data (see Section III). After the selection, the MC templates contain  $\approx 15000$  showers per primary specie and model.

The analysis was carried out by splitting these simulated air-showers in five zenith-angle ranges containing nearly the same number of events, namely  $(0^\circ, 33^\circ)$ ,  $(33^\circ, 39^\circ)$ ,  $(39^\circ, 45^\circ)$ ,  $(45^\circ, 51^\circ)$ ,  $(51^\circ, 60^\circ)$ . Examples of the  $(S, X)$  distributions for protons and iron nuclei in the most vertical and most inclined angular ranges are shown in Fig. 1. Such two-dimensional distributions are then normalized and fitted with the ansatz function  $\Phi$  described in detail in Appendix A. This function is a convolution of the generalized Gumbel distribution of  $X$  and the Gaussian distribution of  $S$  with the mean value linearly changing with  $X$ , reflecting in this way their correlation indicated in Fig. 1. A set of these trial functions for each model, primary particle, and zenith-angle range is used as MC templates in the following fitting procedure.

### B. Fitting Procedure

For each model, we search for the most likely combination of the composition mix of the four primary species, the zenith-dependent rescaling parameter of the hadronic signal  $R_{\text{had}}(\theta)$ , and the constant shift of the depth of shower maximum  $\Delta X_{\max}$  in all the MC templates. The fitting method is a generalization of the fitting procedure used in Ref. [9] in the case of the  $X_{\max}$  distribution and applied here to the  $(X, S)$  distributions in five zenith-angle ranges simultaneously.

The negative log-likelihood-ratio expression that is minimized for a given model is of the form

$$\ln \mathcal{L} = \begin{cases} \sum_k \sum_j (C_{jk} - n_{jk} + n_{jk} \ln \frac{n_{jk}}{C_{jk}}), & n_{jk} > 0, \\ \sum_k \sum_j C_{jk}, & n_{jk} = 0, \end{cases} \quad (3)$$

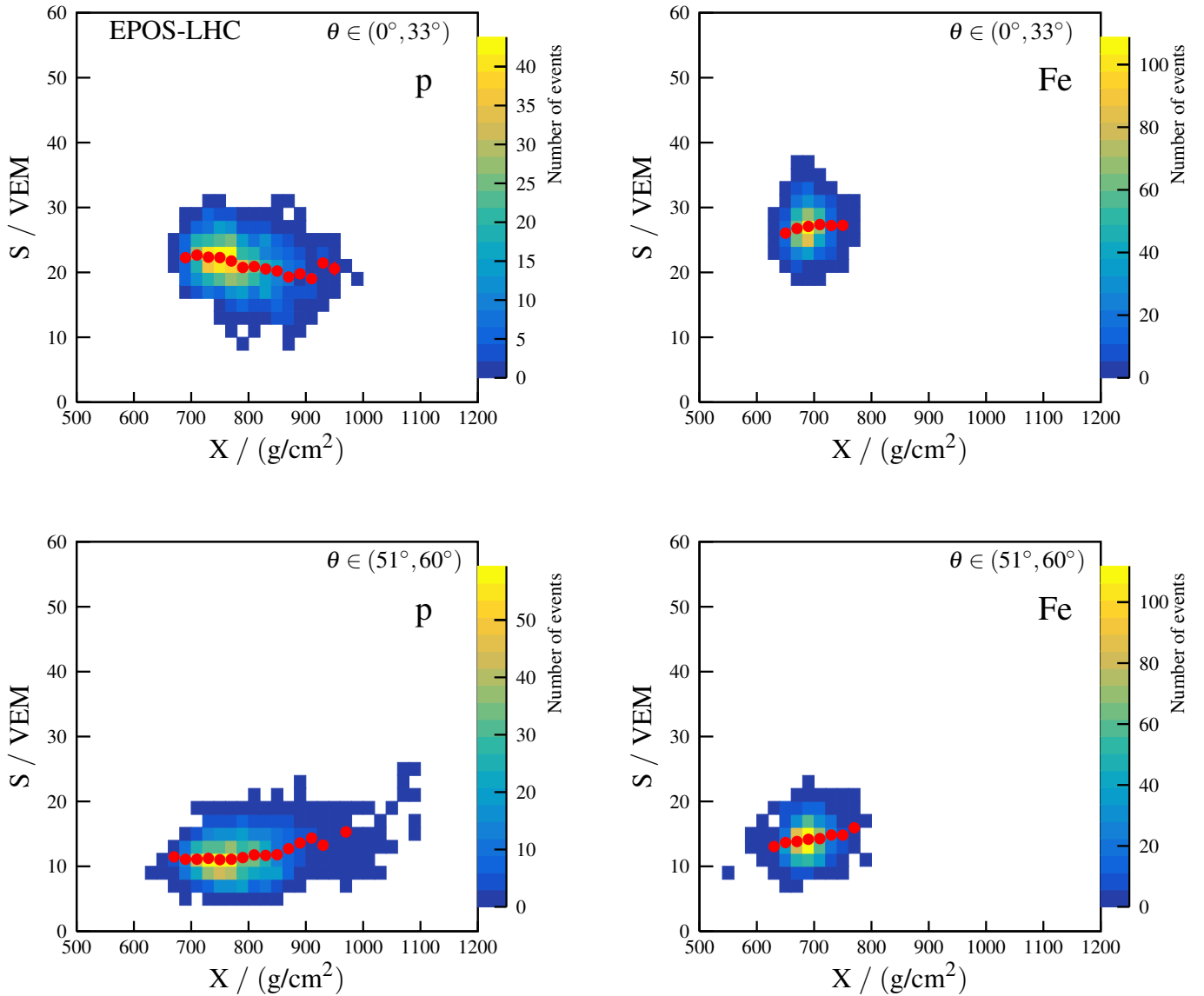


FIG. 1: Examples of two-dimensional distributions of  $S$  and  $X$  for protons (left) and iron nuclei (right) generated with EPOS-LHC for zenith angles between  $0^\circ$  and  $33^\circ$  (top) and  $51^\circ$  and  $60^\circ$  (bottom). The red points indicate the mean values of  $S$ .  $E = 10^{18.5}$  to  $10^{19.0}$  eV.

with the sums running over the two-dimensional bins  $j$  and five  $\theta$ -bins  $k$ . The corresponding number of showers measured in bins  $j, k$  is denoted by  $n_{jk}$  and the predicted number of MC showers by  $C_{jk}$ . The latter number is obtained using the total number of measured showers  $N_{\text{data}}^k$  as

$$C_{jk} = N_{\text{data}}^k \sum_i f_i \Phi_{i,k}(X'_{j,k}, S'_{j,k}), \quad (4)$$

where  $\Phi_{i,k}$  denotes the template function  $\Phi$  in  $\theta$ -bin  $k$  for a given model and primary particle  $i$  with relative fraction  $f_i$ . The modified  $X$  prediction is of the form

$$X'_{j,k} = X_{j,k} + \Delta X_{\text{max}}, \quad (5)$$

and the rescaled predicted ground signal

$$S'_{j,k} = S_{j,k} f_{\text{SD}}^k(R_{\text{had}}(\theta), \Delta X_{\text{max}}), \quad (6)$$

where  $X_{j,k}$  and  $S_{j,k}$  are the center bin values of  $X$  and  $S$ , respectively, of the original MC distribution ( $X, S$ ).

The rescaling parameter  $f_{\text{SD}}^k$  of all signals  $S_{j,k}$  is calculated as

$$f_{\text{SD}}^k = R_{\text{had}}(\theta) \frac{(E^{\text{ref}})^{\beta-1/B}}{\langle E_{\text{FD}}^{\beta-1/B} \rangle_k} g_{\text{had},k} f_{\text{had},k} + \frac{(E^{\text{ref}})^{1-1/B}}{\langle E_{\text{FD}}^{1-1/B} \rangle_k} g_{\text{em},k} (1 - f_{\text{had},k}), \quad (7)$$

where  $S(1000) \propto E_{\text{FD}}^B$  is assumed to be composed only of  $S_{\text{had}} \propto E_{\text{FD}}^\beta$  and  $S_{\text{em}} \propto E_{\text{FD}}$ , and  $f_{\text{had}} = S_{\text{had}}/S(1000)$ . The parameter  $\beta = 0.92$  is chosen following Ref. [23]. The mean energy factors  $\langle E_{\text{FD}}^{\beta-1/B} \rangle_k$  and  $\langle E_{\text{FD}}^{1-1/B} \rangle_k$  are calculated from

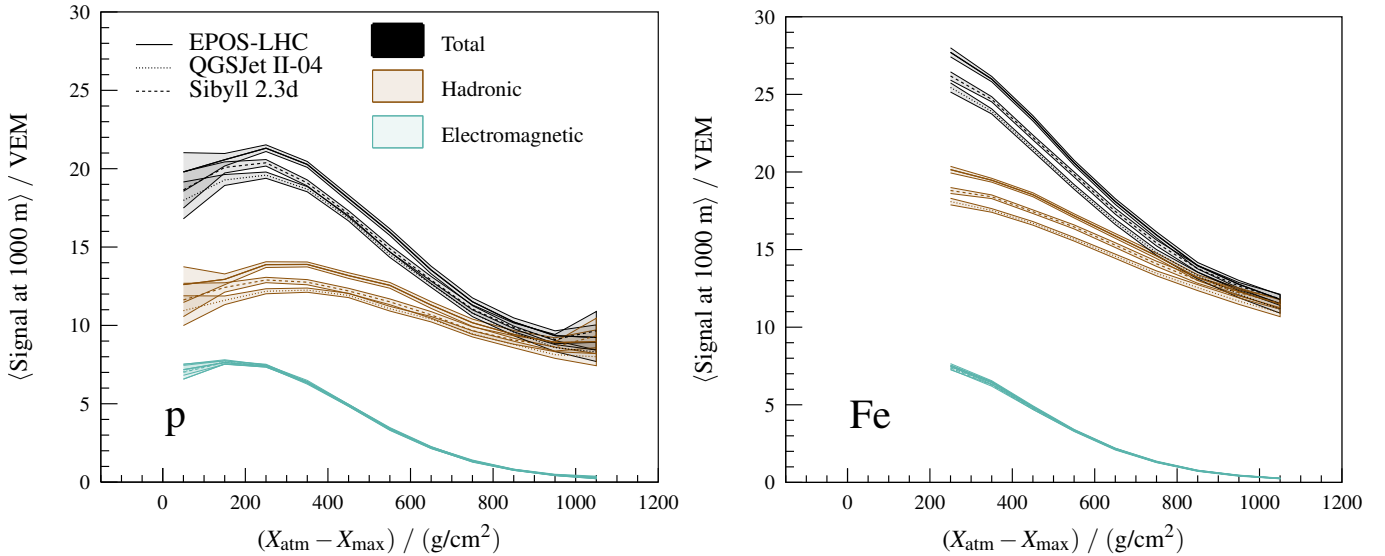


FIG. 2: The total ground signal at 1000 m from the shower core (black) and its hadronic (brown) and em (turquoise) components as a function of the distance from  $X_{\text{max}}$  to the ground in atmospheric depth units for protons (left) and iron nuclei (right) for different models. The bands contain the statistical uncertainty.  $E = 10^{18.5}$  to  $10^{19.0}$  eV,  $\theta < 60^\circ$ .

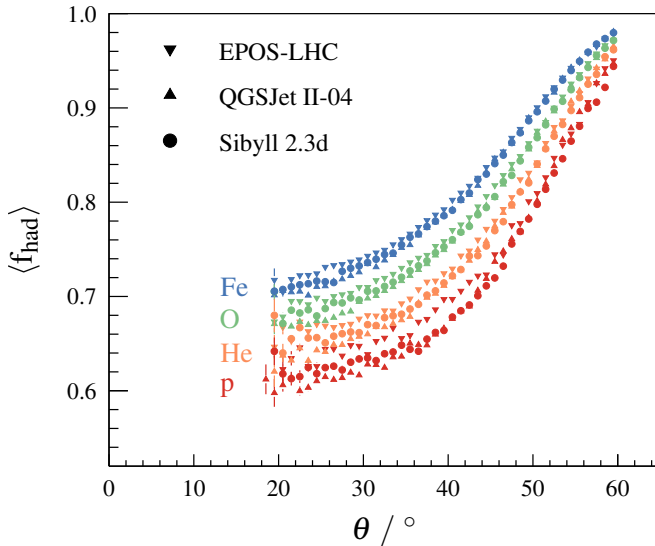


FIG. 3: The average fraction of hadronic signal at 1000 m from the shower core as a function of the reconstructed zenith angle for different models and primary masses.  $E = 10^{18.5}$  to  $10^{19.0}$  eV.

all measured showers in the energy range  $10^{18.5}$  to  $10^{19.0}$  eV and  $\theta$ -bin  $k$ .<sup>3</sup> The mean hadronic fraction, see also Fig. 3, in the  $\theta$ -bin  $k$ ,  $f_{\text{had},k}$ , is calculated using the average hadronic fractions  $f_{\text{had},k,i}$  for simulated showers induced by a primary  $i$

and weighted over the relative primary fractions  $f_i$  as

$$f_{\text{had},k} = \sum_i f_i f_{\text{had},k,i}. \quad (8)$$

The average effect of the  $X_{\text{max}}$  change on the ground signal is incorporated through the separate effects on the em  $g_{\text{em},k}$  and hadronic  $g_{\text{had},k}$  signals, see Appendix C for details. We parameterized the evolution of the mean ground signal parts with the distance of  $X_{\text{max}}$  to the ground in atmospheric depth units,  $X_{\text{atm}} - X_{\text{max}}$ , where  $X_{\text{atm}} = 880 \text{ g/cm}^2 / \cos \theta$ , see the examples in Fig. 2. In this way, the total ground signal is estimated to be modified (via  $g_{\text{em},k}$ ,  $g_{\text{had},k}$ ) at most by about 7% for a change of  $X_{\text{max}}$  by  $50 \text{ g/cm}^2$ . This is in accordance with the functional dependencies in Fig. 2 weighted over the relative contributions of hadronic and em signals (see Fig. 3) and over the primary fractions.

As a consequence of the four-component shower universality approach, the em signal is very similar in all three models, see Fig. 2. The differences in the total signal stem from the size of the hadronic signal at different zenith angles, corresponding to different  $X_{\text{atm}} - X_{\text{max}}$  values. Therefore, the freedom in  $R_{\text{had}}(\theta)$  and also  $\Delta X_{\text{max}}$  (see Section I) removes the main differences in predictions of  $S$  and  $X$  for the three models. We assumed a linear dependence of  $R_{\text{had}}(\theta)$  on  $X_{\text{atm}} - X_{\text{max}}$  and defined rescaling parameters of the hadronic signal at two extreme zenith angles,  $R_{\text{had}}(\theta_{\text{min}})$  and  $R_{\text{had}}(\theta_{\text{max}})$ , for  $\sim 28^\circ$  and  $\sim 55^\circ$ , respectively, see Appendix C for the definition.

We have verified using MC-MC tests, see Appendix B, that the method is performing well and the observed biases were taken as sources of systematic uncertainty of the results.

<sup>3</sup> Note that the choice of  $E^{\text{ref}} \approx \langle E_{\text{FD}} \rangle$  minimizes the effect of the energy factors:  $(E^{\text{ref}})^{\beta-1/B} / \langle E_{\text{FD}}^{\beta-1/B} \rangle_k \approx 1$  and  $(E^{\text{ref}})^{1-1/B} / \langle E_{\text{FD}}^{1-1/B} \rangle_k \approx 1$ .

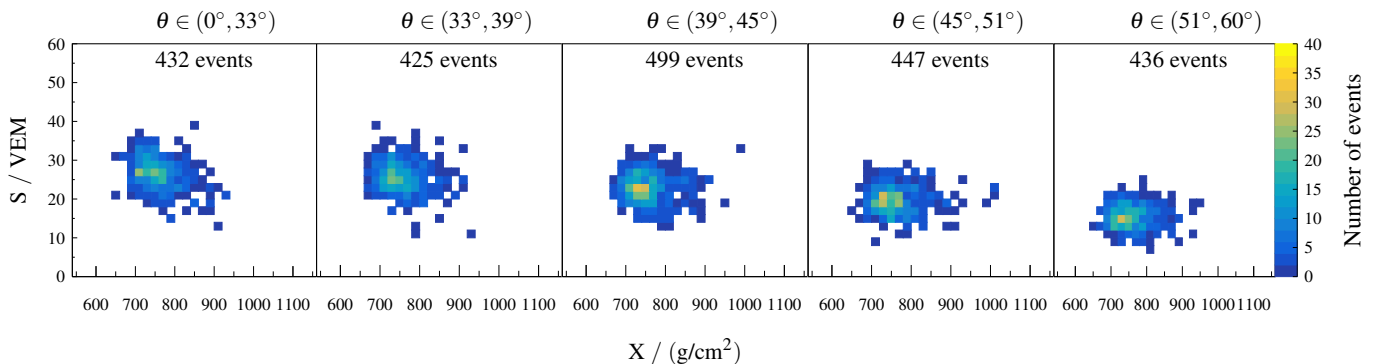


FIG. 4: Distributions of  $S$  and  $X$  for the data of the Pierre Auger Observatory in the energy range  $10^{18.5}$  to  $10^{19.0}$  eV in five zenith-angle bins.

TABLE II: Evolution of the minimum values of the log-likelihood expression (see Eq. (3)) fitting the data with different modifications of the model predictions. In all cases, except  $R_{\text{had}} = \text{const.}$  and  $\Delta X_{\text{max}}$ , the significance of improvement of data description with the  $R_{\text{had}}(\theta)$  and  $\Delta X_{\text{max}}$  fit is above  $5\sigma$  using the Likelihood-ratio test applying the Wilks' theorem [34] for nested model [35]. In the case of  $R_{\text{had}} = \text{const.}$  and  $\Delta X_{\text{max}}$ , the improvement of data description with  $R_{\text{had}}(\theta)$  and  $\Delta X_{\text{max}}$  fit is  $\sim 0.1\sigma$ ,  $\sim 4.4\sigma$  and  $\sim 2.0\sigma$  for EPOS-LHC, QGSJET-II-04 and SIBYLL 2.3d, respectively.

$\ln \mathcal{L}_{\text{min}}$	EPOS-LHC	QGSJET-II-04	SIBYLL 2.3d
none	2022.9	4508.0	2496.5
$\Delta X_{\text{max}}$	738.6	1674.8	1015.7
$R_{\text{had}} = \text{const.}$	489.2	684.4	521.6
$R_{\text{had}}(\theta)$	489.2	673.9	517.6
$R_{\text{had}} = \text{const. and } \Delta X_{\text{max}}$	452.2	486.7	454.2
$R_{\text{had}}(\theta)$ and $\Delta X_{\text{max}}$	451.9	476.3	451.6

### III. DATA ANALYSIS

#### A. Data Selection

The analysis is applied to hybrid events, i.e. events detected with both SD and FD of the Pierre Auger Observatory between 1 January 2004 and 31 December 2018. To select high-quality events, for the FD part we apply the selection criteria used for the  $X_{\text{max}}$  analysis [8, 10], for the SD part the selection is the same as in the measurements of the SD energy spectrum [22], additionally removing events with saturated SD stations. In this way, we ensure an accurate estimation of the observables  $X_{\text{max}}$ ,  $E_{\text{FD}}$ ,  $\theta_{\text{FD}}$ , and  $S(1000)$  used as the inputs to the method. The selection efficiency is similar for all four primary masses and all zenith-angle bins, therefore it does not introduce mass-composition biases. In total, 2239 hybrid events were selected in the FD energy range  $10^{18.5}$  to  $10^{19.0}$  eV ( $\langle E_{\text{FD}} \rangle \approx 10^{18.7}$  eV) and zenith angles between  $0^\circ$  and  $60^\circ$ . The data were divided into five zenith-angle bins containing nearly the same ( $N = 425 - 500$ ) and sufficiently large number of events, see Fig. 4.

To take into account long-term performance of the FD and SD [36], we applied a time-dependent correction to  $E_{\text{FD}}$  with a negligible systematic effect on the final results. The signal  $S(1000)$  was corrected for seasonal atmospheric effects [37].

#### B. Results

The minimized values of the log-likelihood expression (see Eq. (3)) are summarized in Table II, progressively applying different modifications to the models. At all stages, the fits of the mass composition are performed with (p, He, O, Fe) fractions as the free-fit parameters. First, the model predictions without any modifications are used for fitting the data. Then we perform fits adding freedom in only one of the modifications  $\Delta X_{\text{max}}$ ,  $R_{\text{had}}$  (independent on zenith angle),  $R_{\text{had}}(\theta)$  (zenith-angle dependent). Finally, the fits are performed using combinations of ( $R_{\text{had}}$ ,  $\Delta X_{\text{max}}$ ) and ( $R_{\text{had}}(\theta)$ ,  $\Delta X_{\text{max}}$ ). From the minimum values of the log-likelihood expression for these scenarios, and from the Likelihood-ratio test for the nested model, one can see that the major improvements in the data description are achieved due  $R_{\text{had}}$  and then  $\Delta X_{\text{max}}$  modifications. The improvement in the data description due to the introduction of the zenith-angle dependence in  $R_{\text{had}}$  is statistically significant for QGSJET-II-04, and less significant for SIBYLL 2.3d, with a negligible effect in the case of EPOS-LHC.

The measured two-dimensional ( $S$ ,  $X$ ) distributions are described acceptably by MC templates modified by  $R_{\text{had}}(\theta)$  and  $\Delta X_{\text{max}}$  in all five zenith-angle bins with  $p$ -values estimated to be higher than 10% for all three models. These probabilities are obtained from MC-MC tests (see Appendix B) us-

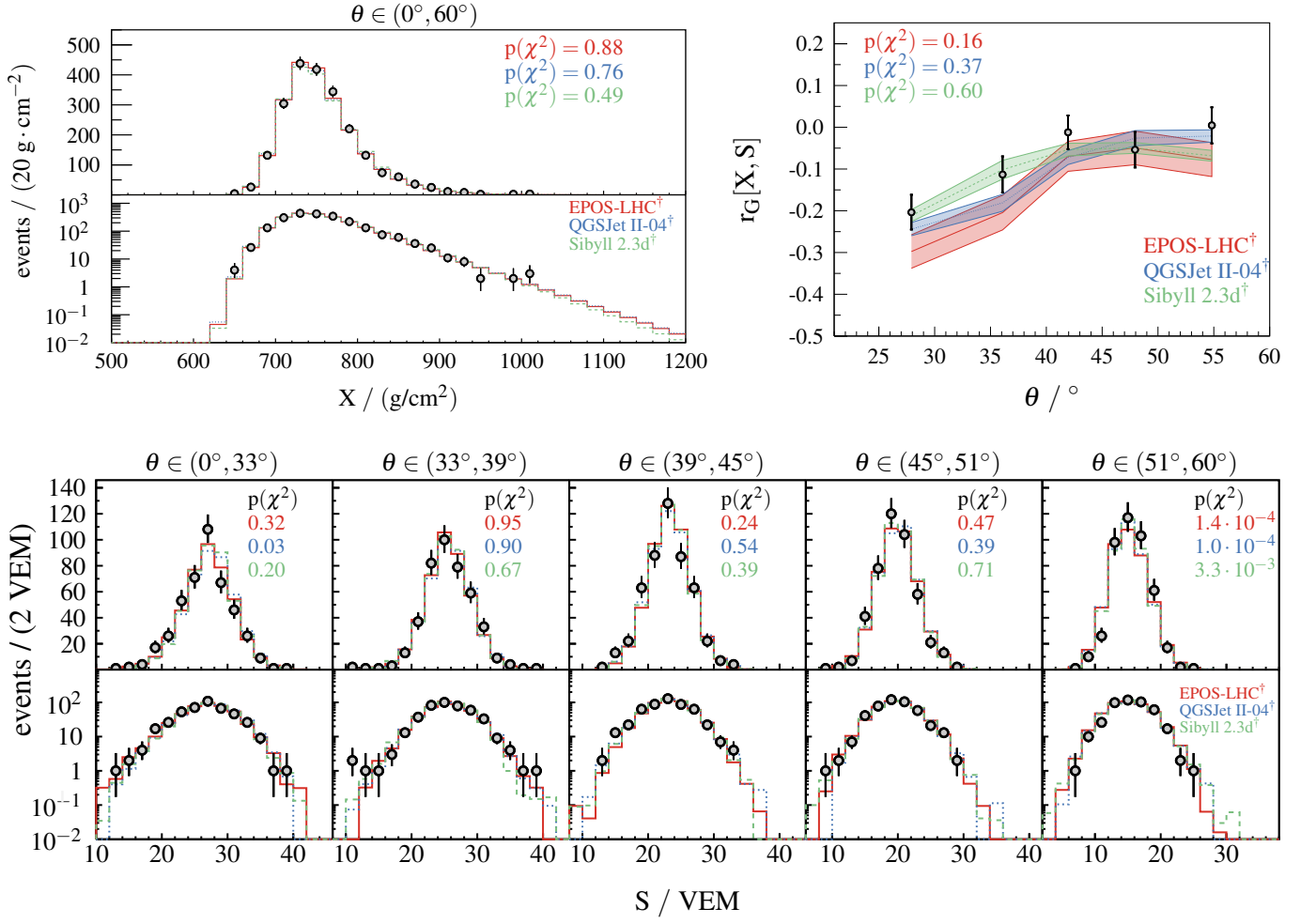


FIG. 5: Distributions of  $X$  (top-left) and  $S$  (bottom) for the Auger data (points) and for the best fits of  $(S, X)$  distributions with  $R_{\text{had}}(\theta)$  and  $\Delta X_{\text{max}}$  modifications in predictions of three models (denoted by <sup>†</sup> as templates were modified from original predictions). In the top-right panel, the dependence of the Gideon-Hollister coefficient of correlation [38] between  $S$  and  $X$  on the zenith angle is shown. The  $\chi^2$  probabilities characterize the compatibility between measurements and MC predictions in the individual plots.

TABLE III: Modifications of the model predictions and primary fractions in the energy range  $10^{18.5}$  to  $10^{19.0}$  eV with statistical and systematic uncertainties for the best data fits and the  $p$ -values obtained using MC-MC tests.

	$R_{\text{had}}(\theta_{\min})$	$R_{\text{had}}(\theta_{\max})$	$\Delta X_{\text{max}}(\text{g/cm}^2)$	$f_p$ (%)	$f_{\text{He}}$ (%)	$f_{\text{O}}$ (%)	$f_{\text{Fe}}$ (%)	$p$ -value (%)
EPOS-LHC	$1.15 \pm 0.01$ $^{+0.20}_{-0.16}$	$1.16 \pm 0.01$ $^{+0.14}_{-0.10}$	$22 \pm 3$ $^{+11}_{-14}$	$21 \pm 3$ $^{+14}_{-11}$	$20 \pm 4$ $^{+15}_{-6}$	$44 \pm 5$ $^{+15}_{-6}$	$15 \pm 4$ $^{+7}_{-15}$	10.6
QGSJET-II-04	$1.24 \pm 0.01$ $^{+0.22}_{-0.19}$	$1.18 \pm 0.01$ $^{+0.15}_{-0.12}$	$47^{+2}_{-1}$ $^{+9}_{-11}$	$16 \pm 2$ $^{+8}_{-10}$	$11 \pm 4$ $^{+20}_{-7}$	$36 \pm 5$ $^{+21}_{-5}$	$37 \pm 5$ $^{+6}_{-22}$	19.8
SIBYLL 2.3d	$1.18 \pm 0.01$ $^{+0.21}_{-0.17}$	$1.15 \pm 0.01$ $^{+0.15}_{-0.11}$	$29 \pm 2$ $^{+10}_{-13}$	$13 \pm 2$ $^{+18}_{-5}$	$15 \pm 4$ $^{+15}_{-12}$	$40 \pm 5$ $^{+22}_{-5}$	$32 \pm 5$ $^{+3}_{-25}$	32.6

ing distributions of values of log-likelihood expression from the fitting of 500 random samples consisting of 2239 simulated showers. Each MC sample had a mass composition and artificially modified  $X_{\text{max}}$  and  $S_{\text{had}}(\theta)$  following the values obtained from the best fits to data. We plot projected distributions of  $X$  and  $S$  and their zenith-angle dependent correlation for the data and MC templates with the modifications of  $(R_{\text{had}}(\theta), \Delta X_{\text{max}})$  in Fig. 5. The models without modifications or only with the zenith-angle independent modification of  $R_{\text{had}}$

do not describe the data equivalently well (see Appendix D).

The resulting parameters of the data fits with  $(R_{\text{had}}(\theta), \Delta X_{\text{max}})$  modifications are presented in Table III. For all three models, a deeper  $X_{\text{max}}$  scale is favored (see also left panels in Figs. 6 and 7) that would result in heavier primary mass composition derived from the  $X_{\text{max}}$  data compared to the inferences with non-modified predictions of the models [9, 39]. The modifications  $\Delta X_{\text{max}}$  are such that they reduce the difference in  $X_{\text{max}}$  scales between the models (see Fig. 10 in Sec-

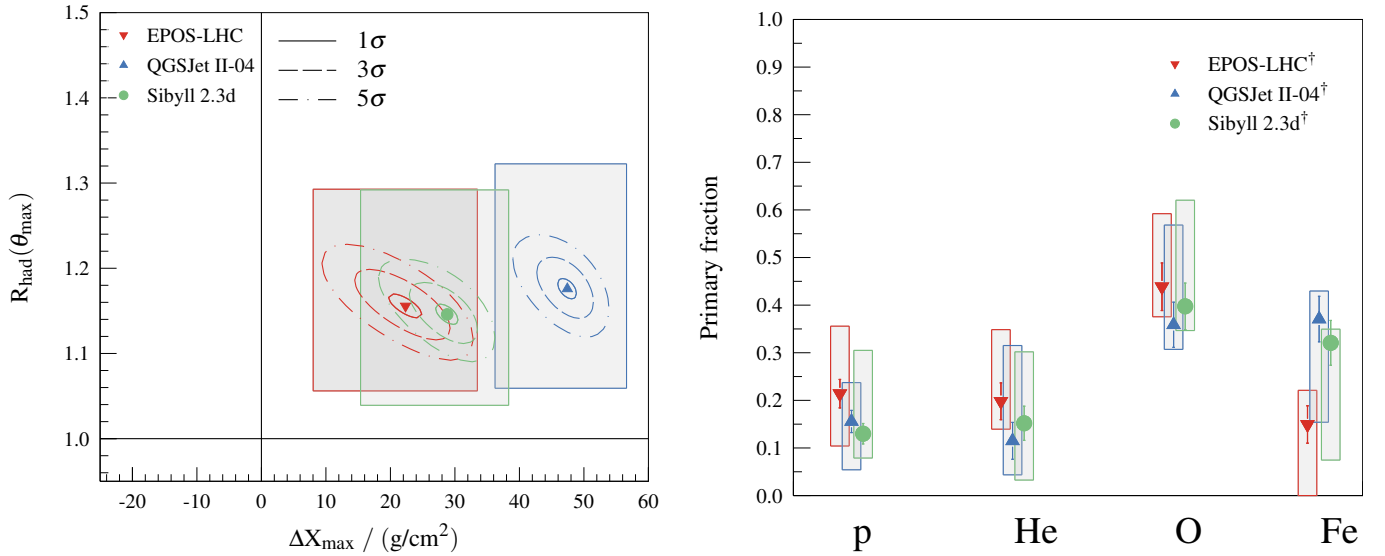


FIG. 6: *Left*: Correlations between  $\Delta X_{\max}$  and  $R_{\text{had}}(\theta_{\max} \approx 55^\circ)$  modifications of the model predictions obtained from the data fits. The contours correspond to  $1\sigma$ ,  $3\sigma$ , and  $5\sigma$  statistical uncertainties. The gray rectangles are the projections of the total systematic uncertainties. *Right*: The most likely primary fractions of the four components from the data fits using  $\Delta X_{\max}$  and  $R_{\text{had}}(\theta)$ . The height of the gray bands shows the size of projected total systematic uncertainties.

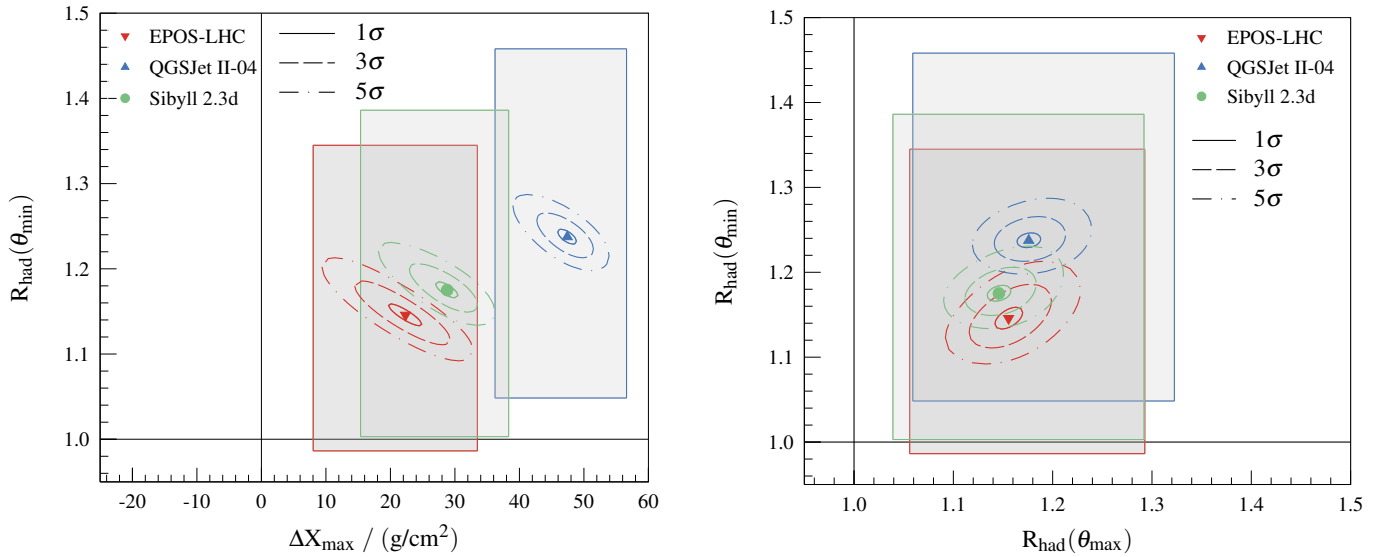


FIG. 7: *Left*: Correlations between  $\Delta X_{\max}$  and  $R_{\text{had}}(\theta_{\min} \approx 28^\circ)$  modifications of the model predictions obtained from the data fits. *Right*: Correlation between  $R_{\text{had}}(\theta_{\max} \approx 55^\circ)$  and  $R_{\text{had}}(\theta_{\min} \approx 28^\circ)$ . The contours correspond to  $1\sigma$ ,  $3\sigma$ , and  $5\sigma$  statistical uncertainties. The gray rectangles are the projections of the total systematic uncertainties.

tion IV) and, as a consequence, similar estimations of the fractions of the primary nuclei can be inferred from the data, see the right panel of Fig. 6.

Due to the shift of the  $X_{\max}$  predictions deeper in the atmosphere, more shower particles reach the ground producing a few percent larger SD signals compared to the non-modified models. Therefore, the total increase of  $S$  for the modified MC predictions consists of contributions from both  $\Delta X_{\max}$  ( $\sim 0$  to 7%) and  $R_{\text{had}}(\theta)$  ( $\sim 10$  to 16%) as shown in the left panel of Fig. 8. The two effects from the modification of the  $X_{\max}$  scale

— heavier primary mass composition and a larger number of shower particles reaching the ground — lead to the values of  $R_{\text{had}}(\theta)$  (Table III) that are smaller than the values previously found for EPOS-LHC and QGSJET-II-04 in Ref. [16]. The modified and original  $S_{\text{had}}$  scales are shown for the three models in the right panel of Fig. 8.



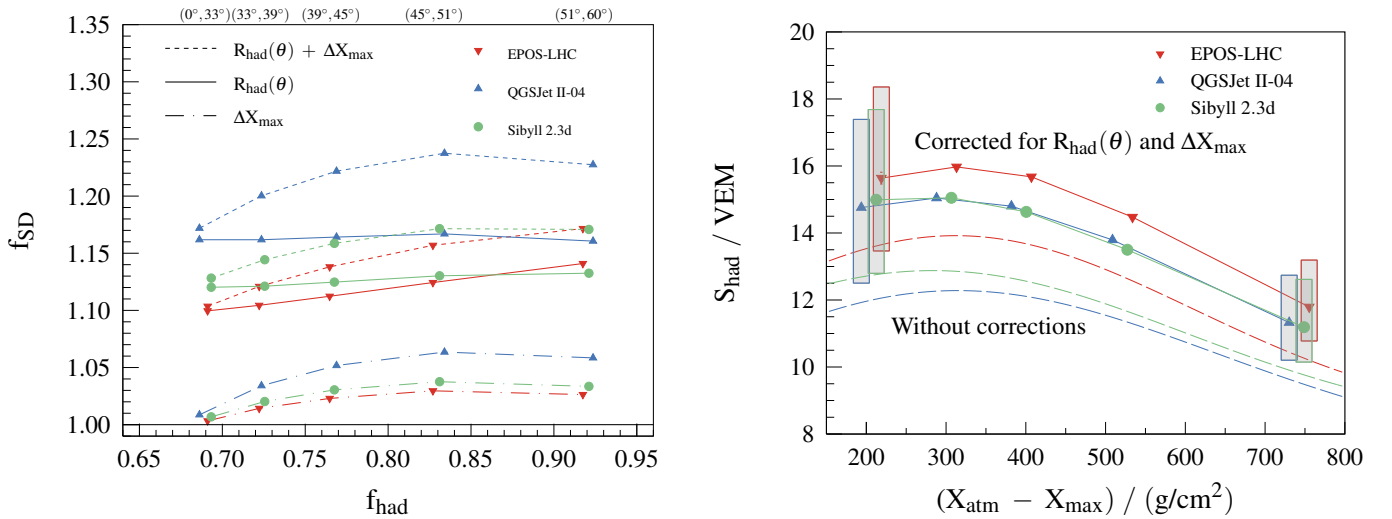


FIG. 8: *Left:* The total rescaling of  $S$  (dashed lines) broken down into the contributions from the rescaling of the hadronic signal  $R_{had}(\theta)$  (solid lines) and the change of the predicted  $X_{max}$  scale ( $\Delta X_{max}$ ) (dash-dotted lines) for different models. *Right:* The dependence of the hadronic signal at 1000 m on the distance of  $X_{max}$  to the ground in atmospheric depth units as predicted for proton showers generated using non-modified models (dashed lines) and accounting for the  $R_{had}(\theta)$ ,  $\Delta X_{max}$  modifications (solid lines). The height of the gray bands shows the size of projected total systematic uncertainties.

### C. Systematic Uncertainties

There are four dominant sources of systematic uncertainties in the fitted parameters:

- the uncertainty in the FD energy scale  $\pm 14\%$  [22],
- the uncertainty in the  $X_{max}$  measurements  $^{+8}_{-9} g/cm^2$  [8],
- the uncertainty in the  $S(1000)$  measurements  $\pm 5\%$  [6],
- the biases of the method estimated from the MC-MC tests (see Appendix C for the results of these tests).

Since the  $X_{max}$  systematic uncertainty is strongly correlated with the modification  $\Delta X_{max}$ , its effect on the nuclei fractions is nearly cancelled out by the corresponding change of  $\Delta X_{max}$ . In general, the nuclei fractions, and therefore the inferences on the mass composition, are weakly sensitive to all experimental systematic errors due to the simultaneous fitting of  $\Delta X_{max}$  and  $R_{had}$  in the method. To explore the effect of those systematics, and as a simplifying ansatz, the total systematic uncertainties on the fit parameters are obtained by summing all four contributions in quadrature (see Appendix E for the size of individual contributions).

One can also note, that within systematic errors no significant dependence of  $R_{had}$  on the zenith angle was found. The difference between  $R_{had}(\theta_{min})$  and  $R_{had}(\theta_{max})$  shows a tight correlation with the uncertainty on the energy scale. However, given all the experimental uncertainties in the case of QGSJET-II-04, the measured data prefers within the method rather flatter attenuation of the hadronic signal at 1000 m than predicted by the model, indicating too hard spectra of muons predicted by this model.

The systematic uncertainties on the parameters  $B$ ,  $D$ ,  $\beta$  used in Eqs. (1) and (2) for the energy correction of  $S$ ,  $X$ , as well as corrections of the long-term performance and other effects related to the operation of the SD and FD, have a negligible contribution to the systematic uncertainties. We could not identify any significant dependencies of the results on the zenith angle or energy in the studied ranges.

### D. Significance of Improvement in Data Description

In Fig. 9, the results of our method for  $\Delta X_{max}$  and  $R_{had}(\theta)$  applying also all possible combinations of the systematic uncertainties on  $E_{FD}$ ,  $X_{max}$  and  $S(1000)$  are shown with the full points. These points are located approximately in a plane, contour outlined with a dashed line, due to a correlation between the modification parameters through the mass composition describing the data (e.g. increase of  $\Delta X_{max}$  leads to a heavier fitted mass composition and consequently to a decrease of  $R_{had}$ ). The plane is tilted with respect to the  $[R_{had}(\theta_{min}), R_{had}(\theta_{max})]$  plane. It is a consequence of the effect of  $\Delta X_{max}$  on the ground signal  $S$  at different zenith angles, see Eqs. in Appendix C, and consequently on the fitted  $R_{had}(\theta_{min})$ ,  $R_{had}(\theta_{max})$ . The color of the points corresponds to the difference in fitted log-likelihood expressions ( $\Delta \ln \mathcal{L}$ ) in case of no modifications and in case of the assumed template modifications. The closest approach to the point of no modification is estimated through a dense scan of linear combinations of experimental systematic uncertainties for the lowest values of  $\Delta \ln \mathcal{L}$  that are, in some cases, even beyond the range of systematic uncertainties quoted by Auger, see Appendix F for more details. For all three models, the closest approach of  $\Delta \ln \mathcal{L}$  (indicated by a line in Fig. 9 connect-

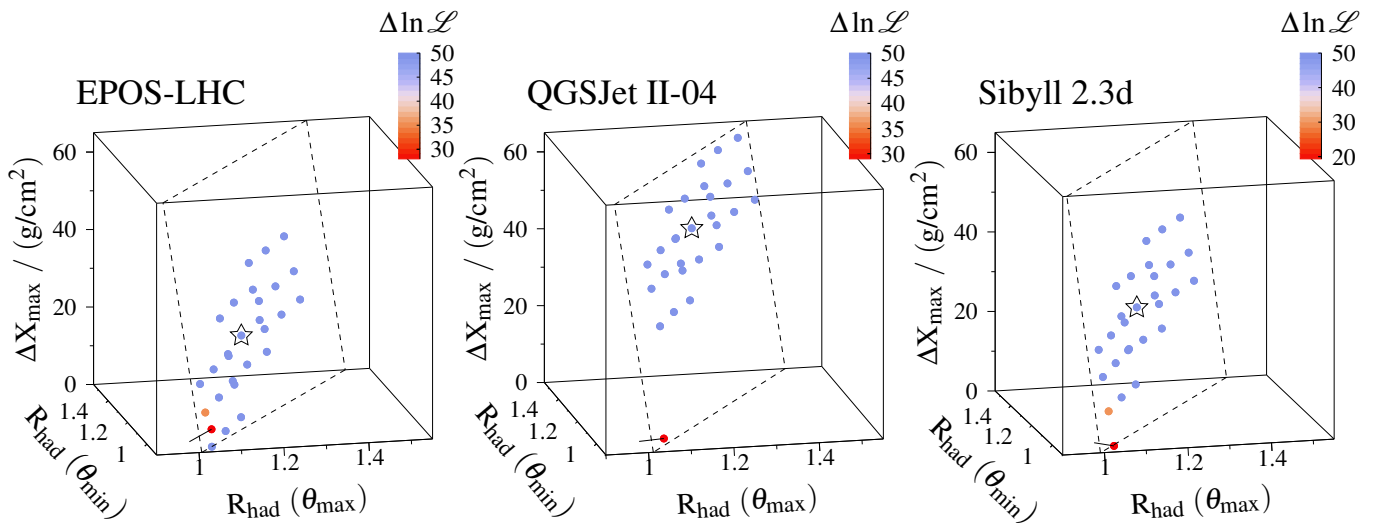


FIG. 9: Values of the modification parameters (points) for all possible combinations of experimental systematic uncertainties on the energy ( $\pm 14\%$ ),  $X_{\max}$  ( ${}_{-9}^{+8}$  g/cm $^2$ ) and  $S(1000)$  ( $\pm 5\%$ ). The color of the points shows the difference in log-likelihood expressions ( $\Delta \ln \mathcal{L}$ ) in the case of no modifications and in the case of the assumed template modifications, including the differences higher than 50 (note the slightly different scale between models). The results (see Table III) for no systematic shift of the data are highlighted by stars. Dashed lines outline the contour of the plane from the best fit to the points. The closest approach to the non-modified ( $R_{\text{had}}(\theta) = 1$ ,  $\Delta X_{\max} = 0$  g/cm $^2$ ) model predictions using a dense scan of linear combinations of experimental systematic uncertainties is connected with this point by a black line.

ing point [1,1,0 g/cm $^2$ ] with the plane) is  $> 19$  which is still higher than the value estimated using the Wilks' theorem in the Likelihood-ratio test for nested model at the level of  $5\sigma$  ( $\Delta \ln \mathcal{L} \approx 16.62$ ). This confirms that the modifications of  $X_{\max}$  and  $R_{\text{had}}$  scales are not an artifact of the systematic uncertainties of our measurements but a needed change in the model descriptions. A correction of the results for the biases seen in the MC-MC tests leads to even larger significance values.

## IV. DISCUSSION

### A. Implications for Inferences on Mass Composition

One straightforward consequence of the  $X_{\max}$  shift deeper in the atmosphere is the solution of the problem with the negative variance of the logarithmic mass  $\sigma^2(\ln A)$  derived with QGSJET-II-04 from the measured  $\langle X_{\max} \rangle$  and  $\sigma(X_{\max})$  (as discussed in Section I). After application of the corresponding  $\Delta X_{\max}$  shifts, one finds  $\sigma^2(\ln A) \approx 0.5$  to 2.5 in the energy range  $10^{18.5}$  to  $10^{19.0}$  eV for all models used in this work. These values are consistent with the degree of mixing of the primary composition found in the analysis of the correlation between  $X_{\max}$  and  $S(1000)$  with non-modified models, since the correlation analysis relies on the general phenomenology of air showers and this way is weakly sensitive to the uncertainties in the description of hadronic interactions [10, 11].

Another outcome of the method can be foreseen by taking into account the quasi-universal behavior of the  $X_{\max}$  elongation rate for all pure beams and models with values staying within 54 to 61 g/cm $^2$ /decade range. Changing the elongation

rate within these limits introduces an energy-dependent uncertainty on the MC  $X_{\max}$  scale of about 4 g/cm $^2$ /decade at most. Under the assumption that the difference  $\Delta X_{\max}$  between the models and data remains nearly independent of the primary energy, i.e. that there is no new physics that can significantly change the predictions for the  $X_{\max}$  elongation rate of single primary species, one could speculate that at the highest energies ( $E \gtrsim 10^{19.5}$  eV) the Auger  $X_{\max}$  measurements, see Fig. 10, can be described with a heavy mass composition having a low degree of mixing due to  $\langle X_{\max} \rangle$  and  $\sigma(X_{\max})$  staying between the extrapolated predictions of the modified models for oxygen and iron nuclei.

### B. Primary Species in the Cosmic-ray Beam

We checked if the shape of the data distribution of ground signal and  $X_{\max}$  and its zenith-angle evolution can be fitted better with an artificially reduced range of the masses in the MC templates and thus with different values of  $R_{\text{had}}(\theta)$  and  $\Delta X_{\max}$ . Due to the presence of deep events and  $\sigma(X_{\max})$  values close to the predictions for protons in  $10^{18.5}$  to  $10^{19.0}$  eV range, we kept protons in the MC templates and used (p, He, O) and (p, He) mixes for the data fits. In both cases, the quality of data fits was found to be inferior compared to the fits with (p, He, O, Fe) nuclei (see Table IV). This was confirmed by the MC-MC tests and was observed using the fits to the measured data for all three models that the obtained  $X_{\max}$  scale decreases by about 5 to 7, 10 to 17 and 30 to 40 g/cm $^2$  and the hadronic signal scale by about 2 to 5%, 4 to 9% and 15 to 20% when the heaviest primary Fe is replaced by Si, O and He in

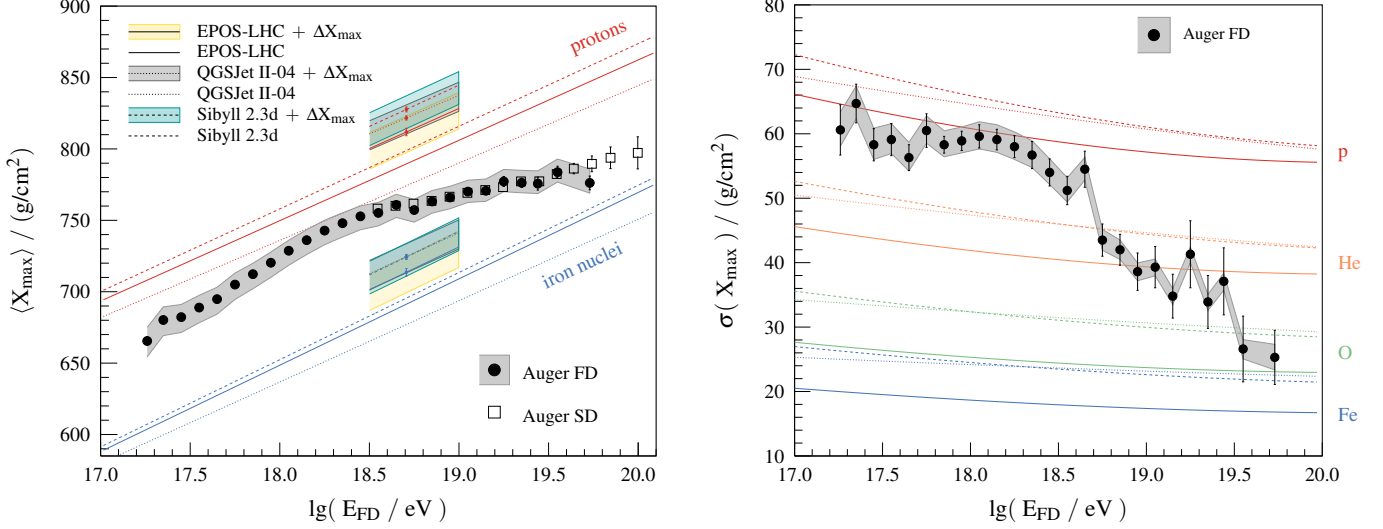


FIG. 10: The energy evolution of the mean (left) and the standard deviation (right) of the  $X_{\max}$  distributions measured by the Pierre Auger Observatory using FD [10] (solid circles) and SD [40] (open squares). The results of this paper for the modified  $X_{\max}$  scales (left panel) are shown with shaded bands with the heights corresponding to the systematic uncertainties. The original non-modified model predictions for different primary species are shown with lines for the entire energy range.

TABLE IV: Minimum values of the log-likelihood expression (see Eq. (3)) for different number of primaries assumed to be present in the model predictions. In all cases, the significance of improvement of data description with the  $R_{\text{had}}(\theta)$ ,  $\Delta X_{\max}$  and p, He, O, Fe fit using the Likelihood-ratio test applying the Wilks' theorem for nested model is above  $5\sigma$ .

$\ln \mathcal{L}_{\min}$	EPOS-LHC	QGSJET-II-04	SIBYLL 2.3d
p He	518.3	633.5	563.5
p He O	467.5	523.3	486.6
p He O Fe	451.9	476.3	451.6

the fit, respectively.

For the five-component (p, He, O, Si, Fe) fits, the values of  $R_{\text{had}}(\theta)$  and  $\Delta X_{\max}$  remain well within statistical errors from the values obtained from the (p, He, O, Fe) fits. Though the fractions of silicon and iron nuclei are strongly anticorrelated in such fits, the silicon fraction remains low,  $<5\%$ .

### C. Modifications of Hadronic Interactions

Assuming the same modifications of hadronic interaction features as in Ref. [41], the observed increase of the  $X_{\max}$  scale in MC predictions could be explained by an increase of the elasticity or a decrease of the multiplicity and cross-section. A detailed study of the combinations of such modifications at the energy equivalent to this study is ongoing [42].

The elasticity is a very good candidate for a potential source of  $X_{\max}$  modification because there are no precise data to constrain this in models at high energy. It could be precisely measured only in the region where no detector at the LHC

experiments exists. Consequently, the different models have relatively different predictions with large uncertainties.

A reduction of the multiplicity can also increase  $X_{\max}$  since the available energy is shared between fewer particles leading to higher  $\pi^0$  energy at the first interaction. But at the same time, the same effect will reduce the energy available for muon production effectively increasing the tension in  $R_{\text{had}}$ .

The superposition model [43] makes an ad-hoc modification of p-p cross-section that would explain a single change of  $X_{\max}$  scale for all primaries rather difficult. The change in the case of iron nuclei would correspond to a change of p-air cross-section at the energy 56-times smaller,  $\sim 9 \times 10^{16}$  eV, which starts to be in tension with the corresponding LHC measurements of p-p collisions [44]. Remarkably, the most recent cross-section measurements by the ALFA experiment are lower and more precise [45] than the one used to tune the models, possibly indicating an overestimation of the current cross-section of the air-shower simulations.

In the case of the multiplicity and elasticity, we can consider the model of shower development from Ref. [23] also assuming the energy increase of the total multiplicity as  $N \propto N_0 E^\alpha$  and a decrease of the elasticity with energy  $\kappa \propto \kappa_0 E^{-\omega}$ . We see that an ad-hoc change of the normalization of multiplicity ( $N_0$ ) or elasticity ( $\kappa_0$ ) would modify the  $X_{\max}$  independently of the primary and energy,

$$\begin{aligned}
 X_{\max}^A &= X_1^A + X_0 \ln \frac{\kappa E}{A 2N \xi_c^\pi} \\
 &= X_1^A + X_0 \left[ (1 - \alpha - \omega) \ln \frac{E}{A \xi_c^\pi} + \ln \frac{\kappa_0}{N_0} \right],
 \end{aligned} \tag{9}$$

where  $\xi_c^\pi$  is the critical energy of pions at which their decay and interaction lengths are equal.

The hadronic signal can be increased by increasing the multiplicity or decreasing the ratio of neutral pions according to Ref. [41]. But, as previously described, an increase in the multiplicity would decrease  $X_{\max}$  increasing the tension there. A change in the ratios of different hadrons, in particular with more strange particles, is a more likely possible explanation according to recent LHC data [46, 47].

The energy spectrum of muons naturally influences the dependence of  $S_{\text{had}}$  on  $X_{\text{atm}} - X_{\max}$ . Though the systematic uncertainty of the energy scale limits a significant conclusion deduced from the method (see  $R_{\text{had}}(\theta_{\min}) - R_{\text{had}}(\theta_{\max})$  in Appendix E), the QGSJET-II-04 model seems to predict too hard spectra of muons. For instance, a larger fraction of strange particles like kaons in the shower would lead to harder muon spectra because of the larger critical energy (smaller lifetime) of strange particles.

There are of course other possible modifications of hadronic interactions that could influence the observed differences between the predictions from the models and the measured data like cross-sections of low-mass mesons, energy spectra of pions etc.; see e.g. Ref. [17].

#### D. Limitations of the Method

There are remaining differences between the models and, correspondingly, with the data due to various limitations of the method. Though our approach leads to a reduction of the differences between models in  $\langle X_{\max} \rangle$ , the obtained modifications  $\Delta X_{\max}$  do not cancel out these differences completely. In particular, the modified  $X_{\max}$  scale for EPOS-LHC is shallower compared to QGSJET-II-04 and SIBYLL 2.3d (see the left panel of Fig. 10). We found, that a large part of this difference ( $\sim 10 \text{ g/cm}^2$ ) can be removed if an additional smearing of  $X_{\max}$  is applied to EPOS-LHC showers to compensate for the smaller  $X_{\max}$  fluctuations predicted by this model in comparison to the other two (possibly partly due to strong defragmentation of nuclei in EPOS-LHC [48], as recently confirmed by newer model EPOS-LHCR [49]). The remaining difference of about  $5 \text{ g/cm}^2$  between the models might be due to the statistical errors on  $\Delta X_{\max}$  and differences between the models in the separation of the primary species in  $\langle X_{\max} \rangle$ .

Possible dependencies of  $\Delta X_{\max}$ ,  $R_{\text{had}}(\theta)$ , and fluctuations of  $X_{\max}$  and  $S(1000)$  on the primary mass are out of the scope of this paper. We checked that adding a linear mass-dependence of  $\Delta X_{\max}$  into the method did not improve the fit significantly. The assumption of a mass-independent  $\Delta X_{\max}$  used in this paper was mainly motivated by the similar differences in predicted  $\langle X_{\max} \rangle$  for different primaries [49]. Such a quasi-universal difference is a consequence of the very similar energy dependencies of interaction features like multiplicity, elasticity, and cross-section assumed in the models [50]. It means that an offset in these features would lead to an approximately mass-independent difference in the  $X_{\max}$  predictions. An ad-hoc modification of the energy dependence of these features, like in [41, 42], would lead to mass-dependent  $X_{\max}$  shift.

## V. CONCLUSIONS

In this paper, we tested the predictions of the models QGSJET-II-04, SIBYLL 2.3d, EPOS-LHC regarding the depths of maximum of air-shower profiles  $X_{\max}$  and the signal produced by air-shower particles in water-Cherenkov stations at 1000 m from the shower core,  $S(1000)$ , composed of electromagnetic and hadronic ( $S_{\text{had}}$ ) parts. The test consisted of fitting with MC templates of two-dimensional distributions of  $(S(1000), X_{\max})$  measured at the Pierre Auger Observatory and obtaining the scales of  $X_{\max}$  and  $S_{\text{had}}(\theta)$  predicted by the models, as well as the fractions of four primary nuclei (p, He, O, Fe).

We found that for the best description of the data distributions in the energy range  $10^{18.5}$  to  $10^{19.0}$  eV the MC predictions of  $X_{\max}$  should be deeper in the atmosphere by about 20 to 50  $\text{g/cm}^2$ , and the hadronic signal should be increased by about 15 to 25% in all three models for  $\theta < 60^\circ$ . These modifications reduce the differences between the models in  $X_{\max}$  and  $S(1000)$ , and as a consequence, lead to smaller uncertainties on the estimated fractions of the primary nuclei. Due to the deeper MC  $X_{\max}$  scale and, correspondingly, a heavier mass composition inferred from the data compared with non-modified models, the scaling factors for the hadronic signal are found to be smaller than in previous estimations not considering any modifications to the MC  $X_{\max}$  scales. The statistical significance of the improvement in the data description using the assumed modifications to the MC templates is above  $5\sigma$  for all three models even accounting for all possible linear combinations of experimental systematic uncertainties.

The specific ways to produce the required changes in the models might consist in combinations of modifications of integral (cross-section, multiplicity, elasticity etc.) or differential (secondary particles energy spectra) characteristics of the hadronic interactions as discussed e.g. in Refs. [17, 41, 42]. Our method addresses only the first-order differences in the mean values of  $X_{\max}$  and hadronic signal without taking into account their possible dependencies on the primary mass or energy. These dependencies, as well as the investigation of a modification of fluctuations of air-shower observables, need to be studied further to corroborate the scale adjustments found in this paper, supplemented also by future data with increased mass-sensitivity from AugerPrime [51], the upgrade of the Pierre Auger Observatory.

## ACKNOWLEDGMENTS

The successful installation, commissioning, and operation of the Pierre Auger Observatory would not have been possible without the strong commitment and effort from the technical and administrative staff in Malargüe. We are very grateful to the following agencies and organizations for financial support:

Argentina – Comisión Nacional de Energía Atómica; Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT); Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET); Gobierno de la Provincia de Mendoza; Municipalidad de Malargüe; NDM Holdings and

Valle Las Leñas; in gratitude for their continuing cooperation over land access; Australia – the Australian Research Council; Belgium – Fonds de la Recherche Scientifique (FNRS); Research Foundation Flanders (FWO), Marie Curie Action of the European Union Grant No. 101107047; Brazil – Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq); Financiadora de Estudos e Projetos (FINEP); Fundação de Amparo à Pesquisa do Estado de Rio de Janeiro (FAPERJ); São Paulo Research Foundation (FAPESP) Grants No. 2019/10151-2, No. 2010/07359-6 and No. 1999/05404-3; Ministério da Ciência, Tecnologia, Inovações e Comunicações (MCTIC); Czech Republic – GACR 24-13049S, CAS LQ100102401, MEYS LM2023032, CZ.02.1.01/0.0/0.0/16.013/0001402, CZ.02.1.01/0.0/0.0/18.046/0016010 and CZ.02.1.01/0.0/0.0/17.049/0008422 and CZ.02.01.01/00/22.008/0004632; France – Centre de Calcul IN2P3/CNRS; Centre National de la Recherche Scientifique (CNRS); Conseil Régional Ile-de-France; Département Physique Nucléaire et Corpusculaire (PNC-IN2P3/CNRS); Département Sciences de l’Univers (SDU-INSU/CNRS); Institut Lagrange de Paris (ILP) Grant No. LABEX ANR-10-LABX-63 within the Investissements d’Avenir Programme Grant No. ANR-11-IDEX-0004-02; Germany – Bundesministerium für Bildung und Forschung (BMBF); Deutsche Forschungsgemeinschaft (DFG); Finanzministerium Baden-Württemberg; Helmholtz Alliance for Astroparticle Physics (HAP); Helmholtz-Gemeinschaft Deutscher Forschungszentren (HGF); Ministerium für Kultur und Wissenschaft des Landes Nordrhein-Westfalen; Ministerium für Wissenschaft, Forschung und Kunst des Landes Baden-Württemberg; Italy – Istituto Nazionale di Fisica Nucleare (INFN); Istituto Nazionale di Astrofisica (INAF); Ministero dell’Università e della Ricerca (MUR); CETEMPS Center of Excellence; Ministero degli Affari Esteri (MAE),

ICSC Centro Nazionale di Ricerca in High Performance Computing, Big Data and Quantum Computing, funded by European Union NextGenerationEU, reference code CN\_00000013; México – Consejo Nacional de Ciencia y Tecnología (CONACYT) No. 167733; Universidad Nacional Autónoma de México (UNAM); PAPIIT DGAPA-UNAM; The Netherlands – Ministry of Education, Culture and Science; Netherlands Organisation for Scientific Research (NWO); Dutch national e-infrastructure with the support of SURF Cooperative; Poland – Ministry of Education and Science, grants No. DIR/WK/2018/11 and 2022/WK/12; National Science Centre, grants No. 2016/22/M/ST9/00198, 2016/23/B/ST9/01635, 2020/39/B/ST9/01398, and 2022/45/B/ST9/02163; Portugal – Portuguese national funds and FEDER funds within Programa Operacional Factores de Competitividade through Fundação para a Ciência e a Tecnologia (COMPETE); Romania – Ministry of Research, Innovation and Digitization, CNCS-UEFISCDI, contract no. 30N/2023 under Romanian National Core Program LAPLAS VII, grant no. PN 23 21 01 02 and project number PN-III-P1-1.1-TE-2021-0924/TE57/2022, within PNCDI III; Slovenia – Slovenian Research Agency, grants P1-0031, P1-0385, I0-0033, N1-0111; Spain – Ministerio de Economía, Industria y Competitividad (FPA2017-85114-P and PID2019-104676GB-C32), Xunta de Galicia (ED431C 2017/07), Junta de Andalucía (SOMM17/6104/UGR, P18-FR-4314) Feder Funds, RENATA Red Nacional Temática de Astropartículas (FPA2015-68783-REDT) and María de Maeztu Unit of Excellence (MDM-2016-0692); USA – Department of Energy, Contracts No. DE-AC02-07CH11359, No. DE-FR02-04ER41300, No. DE-FG02-99ER41107 and No. DE-SC0011689; National Science Foundation, Grant No. 0450696; The Grainger Foundation; Marie Curie-IRSES/EPLANET; European Particle Physics Latin American Network; and UNESCO.

- 
- [1] T. Pierog, I. Karpenko, J. M. Katzy, E. Yatsenko, and K. Werner, EPOS LHC: Test of collective hadronization with data measured at the cern large hadron collider, *Phys. Rev. C* **92**, 034906 (2015).
- [2] S. Ostapchenko, Monte Carlo treatment of hadronic interactions in enhanced Pomeron scheme: QGSJET-II model, *Phys. Rev. D* **83**, 014018 (2011).
- [3] F. Riehn, R. Engel, A. Fedynitch, T. K. Gaisser, and T. Stanev, Hadronic interaction model Sibyll 2.3d and extensive air showers, *Phys. Rev. D* **102**, 063002 (2020).
- [4] S. Ostapchenko, LHC results and hadronic interaction models, *ArXiv e-prints* 1612.09461 [astro-ph.HE] (2016), arXiv:1612.09461 [astro-ph.HE].
- [5] Aab A. et al. (Pierre Auger Collaboration), The Pierre Auger Cosmic Ray Observatory, *Nucl. Instrum. Methods Phys. Res. A* **798**, 172 (2015).
- [6] A. Aab et al. (Pierre Auger Collaboration), Reconstruction of events recorded with the surface detector of the Pierre Auger Observatory, *Journal of Instrumentation* **15** (10), P10021.
- [7] P. Abreu et al. (Pierre Auger Collaboration), Interpretation of the depths of maximum of extensive air showers measured by the Pierre Auger Observatory, *Journal of Cosmology and Astroparticle Physics* **2013** (02), 026.
- [8] A. Aab et al. (Pierre Auger Collaboration), Depth of maximum of air-shower profiles at the Pierre Auger Observatory. I. Measurements at energies above  $10^{17.8}$  eV, *Phys. Rev. D* **90**, 122005 (2014).
- [9] A. Aab et al. (Pierre Auger Collaboration), Depth of maximum of air-shower profiles at the Pierre Auger Observatory. II. Composition implications, *Phys. Rev. D* **90**, 122006 (2014).
- [10] A. Yushkov (Pierre Auger Collaboration), Mass composition of cosmic rays with energies above  $10^{17.2}$  eV from the hybrid data of the Pierre Auger Observatory, *Proceedings, 36th International Cosmic Ray Conference (ICRC 2019): Madison, WI, U.S.A., July 24th - August 1st 2019*, PoS **ICRC2019**, 482 (2019).
- [11] A. Aab et al. (Pierre Auger Collaboration), Evidence for a mixed mass composition at the ‘ankle’ in the cosmic-ray spectrum, *Physics Letters B* **762**, 288 (2016).
- [12] A. Aab et al. (Pierre Auger Collaboration), Muons in air showers at the Pierre Auger Observatory: Mean number in highly inclined events, *Phys. Rev. D* **91**, 032003 (2015).

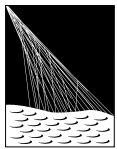
- [13] A. Aab *et al.* (Pierre Auger Collaboration), Measurement of the fluctuations in the number of muons in extensive air showers with the pierre auger observatory, *Phys. Rev. Lett.* **126**, 152002 (2021).
- [14] A. Aab *et al.* (Pierre Auger Collaboration), Direct measurement of the muonic content of extensive air showers between  $2 \cdot 10^{17}$  and  $2 \cdot 10^{18}$  eV at the Pierre Auger Observatory, *Eur. Phys. J. C* **210**, 751 (2020).
- [15] A. Aab *et al.* (Pierre Auger Collaboration), Muons in air showers at the pierre auger observatory: Measurement of atmospheric production depth, *Phys. Rev. D* **90**, 012012 (2014).
- [16] A. Aab *et al.* (Pierre Auger Collaboration), Testing Hadronic Interactions at Ultrahigh Energies with Air Showers Measured by the Pierre Auger Observatory, *Phys. Rev. Lett.* **117**, 192001 (2016).
- [17] J. Albrecht *et al.*, The Muon Puzzle in cosmic-ray induced air showers and its connection to the Large Hadron Collider, *Astrophysics and Space Science* **367**, 27 (2022).
- [18] J. Matthews, A Heitler model of extensive air showers, *Astroparticle Physics* **22**, 387 (2005).
- [19] J. Vicha, *Analysis of Air Showers with respect to Primary Composition of Cosmic Rays* (PhD thesis, Czech Technical University, Prague, 2016) pp. 83 – 114.
- [20] J. Vicha *et al.*, Testing Hadronic Interactions Using Hybrid Observables, *Proceedings, 36th International Cosmic Ray Conference (ICRC 2019): Madison, WI, U.S.A., July 24th - August 1st 2019*, PoS **ICRC2019**, 452 (2019).
- [21] J. Vicha (Pierre Auger Collaboration), Adjustments to Model Predictions of Depth of Shower Maximum and Signals at Ground Level using Hybrid Events of the Pierre Auger Observatory, *Proceedings, 37th International Cosmic Ray Conference (ICRC 2021): Berlin, Germany, July 12th - 23rd 2021*, PoS **ICRC2021**, 310 (2021).
- [22] A. Aab *et al.* (Pierre Auger Collaboration), Measurement of the cosmic-ray energy spectrum above  $2.5 \times 10^{18}$  eV using the pierre auger observatory, *Phys. Rev. D* **102**, 062005 (2020).
- [23] K.-H. Kampert and M. Unger, Measurements of the cosmic ray composition with air shower experiments, *Astroparticle Physics* **35**, 660 (2012).
- [24] M. Ave, R. Engel, M. Roth, and A. Schulz, A generalized description of the signal size in extensive air shower detectors and its applications, *Astroparticle Physics* **87**, 23 (2017).
- [25] M. Ave, M. Roth, and A. Schulz, A generalized description of the time dependent signals in extensive air shower detectors and its applications, *Astropart. Phys.* **88**, 46 (2017).
- [26] E. dos Santos (Pierre Auger Collaboration), Monte Carlo simulations for the Pierre Auger Observatory using the VO auger grid resources, *Proceedings, 37th International Cosmic Ray Conference (ICRC 2021): Berlin, Germany, July 12th - 23rd 2021*, PoS **ICRC2021**, 232 (2021).
- [27] D. Heck *et al.*, CORSIKA: A Monte Carlo code to simulate extensive air showers, *Report FZKA Forschungszentrum Karlsruhe* **6019** (1998).
- [28] D. Heck and J. Knapp, Upgrade of the Monte Carlo code CORSIKA to simulate extensive air showers with energies  $> 10^{20}$  eV, *Report FZKA Forschungszentrum Karlsruhe* **6097** (1998).
- [29] D. Heck, Air shower simulation with CORSIKA at arbitrary direction of incidence, *Report FZKA Forschungszentrum Karlsruhe* **7254** (2006).
- [30] M. Bleicher, E. Zabrodin, C. Spieles, S. A. Bass, C. Ernst, S. Soff, L. Bravina, M. Belkacem, H. Weber, H. Stöcker, and W. Greiner, Relativistic hadron-hadron collisions in the ultra-relativistic quantum molecular dynamics model, *Journal of Physics G: Nuclear and Particle Physics* **25**, 1859 (1999).
- [31] G. Battistoni *et al.*, Overview of the FLUKA code, *Annals Nucl. Energy* **82**, 10 (2015).
- [32] T. Böhlen, F. Cerutti, M. Chin, A. Fassò, A. Ferrari, P. Ortega, A. Mairani, P. Sala, G. Smirnov, and V. Vlachoudis, The fluka code: Developments and challenges for high energy and medical applications, *Nuclear Data Sheets* **120**, 211 (2014).
- [33] S. Argirò, S. Barroso, J. Gonzalez, L. Nellen, T. Paul, T. Porter, L. Prado Jr., M. Roth, R. Ulrich, and D. Veberič, The offline software framework of the pierre auger observatory, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **580**, 1485 (2007).
- [34] S. S. Wilks, The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses, *Annals Math. Statist.* **9**, 60 (1938).
- [35] J. Fan, C. Zhang, and J. Zhang, Generalized Likelihood Ratio Statistics and Wilks Phenomenon, *The Annals of Statistics* **29**, 153 (2001).
- [36] K. Choi (Pierre Auger Collaboration), Long Term Performance of the Pierre Auger Observatory, *Proceedings, 36th International Cosmic Ray Conference (ICRC 2019): Madison, WI, U.S.A., July 24th - August 1st 2019*, PoS **ICRC2019**, 222 (2021).
- [37] A. Aab *et al.* (Pierre Auger Collaboration), Impact of Atmospheric Effects on the Energy Reconstruction of Air Showers Observed by the Surface Detectors of the Pierre Auger Observatory, *JINST* **12** (02), P02006, arXiv:1702.02835 [astro-ph.IM].
- [38] R. A. Gideon and R. A. Hollister, A rank correlation coefficient resistant to outliers, *Journal of the American Statistical Association* **82**, 656 (1987).
- [39] J. Bellido for the Pierre Auger Collaboration, Depth of maximum of air-shower profiles at the Pierre Auger Observatory: Measurements above  $10^{17.2}$  eV and Composition Implications, in *Proceedings of the 35th International Cosmic Ray Conference ICRC 2017, PoS(ICRC2017)506* (2018).
- [40] C. J. T. Peixoto (Pierre Auger Collaboration), Estimating the Depth of Shower Maximum using the Surface Detectors of the Pierre Auger Observatory, *Proceedings, 36th International Cosmic Ray Conference (ICRC 2019): Madison, WI, U.S.A., July 24th - August 1st 2019*, PoS **ICRC2019**, 440 (2019).
- [41] R. Ulrich, R. Engel, and M. Unger, Hadronic multiparticle production at ultrahigh energies and extensive air showers, *Phys. Rev. D* **83**, 054026 (2011).
- [42] J. Ebr *et al.*, Impact of modified characteristics of hadronic interactions on cosmic-ray observables for proton and nuclear primaries, *Proceedings, 38th International Cosmic Ray Conference (ICRC 2023): Nagoya, Japan, July 26th - August 3rd 2023*, PoS **ICRC2023**, 245 (2023).
- [43] J. Engel, T. K. Gaisser, P. Lipari, and T. Stanev, Nucleus-nucleus collisions and interpretation of cosmic-ray cascades, *Phys. Rev. D* **46**, 5013 (1992).
- [44] G. Antchev *et al.* (TOTEM Collaboration), First measurement of elastic, inelastic and total cross-section at  $\sqrt{s} = 13$  tev by totem and overview of cross-section data at lhc energies, *European Physical Journal C* **79** (2019).
- [45] P. Grafstrom (ATLAS Collaboration), Measurements of the elastic, inelastic and total cross sections in pp collisions with ATLAS sub-detectors, PoS **EPS-HEP2015**, 486 (2015).
- [46] S. Baur, H. Dembinski, M. Perlin, T. Pierog, R. Ulrich, and K. Werner, Core-corona effect in hadron collisions and muon production in air showers, *Phys. Rev. D* **107**, 094031 (2023).
- [47] L. A. Anchordoqui, C. G. Canal, F. Kling, S. J. Sciutto, and J. F. Soriano, An explanation of the muon puzzle of ultrahigh-energy cosmic rays and the role of the Forward Physics Facility for

model improvement, *JHEAp* **34**, 19 (2022), arXiv:2202.03095 [hep-ph].

- [48] S. Ostapchenko, *High Energy Cosmic Ray Interactions and UHECR Composition Problem, talk at the UHECR 2018, Paris, October 08-12, 2018 (Paris)* (2018 (accessed Oct, 2018)).
- [49] T. P. K. Werner, EPOS LHC-R : up-to-date hadronic model for EAS simulations, *Proceedings, 38th International Cosmic Ray Conference (ICRC 2023): Nagoya, Japan, July 26th - August 3rd 2023, PoS ICRC2023*, 230 (2023).
- [50] J. Alvarez-Muñiz, R. Engel, T. K. Gaisser, J. A. Ortiz, and T. Stanev, Hybrid simulations of extensive air showers, *Phys. Rev. D* **66**, 033011 (2002).
- [51] A. Castellina (Pierre Auger Collaboration), AugerPrime: the Pierre Auger Observatory Upgrade, *EPJ Web Conf.* **210**, 06002

(2019), arXiv:1905.04472 [astro-ph.HE].

- [52] M. D. Domenico, M. Settimo, S. Riggi, and E. Bertin, Reinterpreting the development of extensive air showers initiated by nuclei and photons, *Journal of Cosmology and Astroparticle Physics* **2013** (07), 050.
- [53] A. Justel, D. Pena, and R. Zamar, A multivariate Kolmogorov-Smirnov test of goodness of fit, *Statistics & Probability Letters* **35**, 251 (1997).
- [54] T. K. Gaisser and A. M. Hillas, Reliability of the Method of Constant Intensity Cuts for Reconstructing the Average Development of Vertical Showers, in *International Cosmic Ray Conference*, International Cosmic Ray Conference, Vol. 8 (1977) p. 353.



PIERRE  
AUGER  
OBSERVATORY

A. Abdul Halim<sup>13</sup>, P. Abreu<sup>73</sup>, M. Aglietta<sup>55,53</sup>, I. Allekotte<sup>1</sup>, K. Almeida Cheminant<sup>71</sup>, A. Almela<sup>7,12</sup>, R. Aloisio<sup>46,47</sup>, J. Alvarez-Muñiz<sup>79</sup>, J. Ammerman Yebra<sup>79</sup>, G.A. Anastasi<sup>59,48</sup>, L. Anchordoqui<sup>86</sup>, B. Andrada<sup>7</sup>, S. Andringa<sup>73</sup>, L. Apollonio<sup>60,50</sup>, C. Aramo<sup>51</sup>, P.R. Araújo Ferreira<sup>43</sup>, E. Arnone<sup>64,53</sup>, J.C. Arteaga Velázquez<sup>68</sup>, P. Assis<sup>73</sup>, G. Avila<sup>11</sup>, E. Avocone<sup>58,47</sup>, A. Bakalova<sup>33</sup>, F. Barbato<sup>46,47</sup>, A. Bartz Mocellin<sup>85</sup>, J.A. Bellido<sup>13,70</sup>, C. Berat<sup>37</sup>, M.E. Bertaina<sup>64,53</sup>, G. Bhatta<sup>71</sup>, M. Bianciotto<sup>64,53</sup>, P.L. Biermann<sup>i</sup>, V. Binet<sup>5</sup>, K. Bismark<sup>40,7</sup>, T. Bister<sup>80,81</sup>, J. Biteau<sup>38,b</sup>, J. Blazek<sup>33</sup>, C. Bleve<sup>37</sup>, J. Blümer<sup>42</sup>, M. Boháčová<sup>33</sup>, D. Boncioli<sup>58,47</sup>, C. Bonifazi<sup>8,27</sup>, L. Bonneau Arbetche<sup>22</sup>, N. Borodai<sup>71</sup>, J. Brack<sup>k</sup>, P.G. Brichetto Orcherá<sup>7</sup>, F.L. Briechle<sup>43</sup>, A. Bueno<sup>78</sup>, S. Buitink<sup>15</sup>, M. Buscemi<sup>48,62</sup>, M. Büsken<sup>40,7</sup>, A. Bwembya<sup>80,81</sup>, K.S. Caballero-Mora<sup>67</sup>, S. Cabana-Freire<sup>79</sup>, L. Caccianiga<sup>60,50</sup>, F. Campuzano<sup>6</sup>, R. Caruso<sup>59,48</sup>, A. Castellina<sup>55,53</sup>, F. Catalani<sup>19</sup>, G. Cataldi<sup>49</sup>, L. Cazon<sup>79</sup>, M. Cerda<sup>10</sup>, A. Cermenati<sup>46,47</sup>, J.A. Chinellato<sup>22</sup>, J. Chudoba<sup>33</sup>, L. Chytka<sup>34</sup>, R.W. Clay<sup>13</sup>, A.C. Cobos Cerutti<sup>6</sup>, R. Colalillo<sup>61,51</sup>, M.R. Coluccia<sup>49</sup>, R. Conceição<sup>73</sup>, A. Condorelli<sup>38</sup>, G. Consolati<sup>50,56</sup>, M. Conte<sup>57,49</sup>, F. Convenga<sup>58,47</sup>, D. Correia dos Santos<sup>29</sup>, P.J. Costa<sup>73</sup>, C.E. Covault<sup>84</sup>, M. Cristinziani<sup>45</sup>, C.S. Cruz Sanchez<sup>3</sup>, S. Dasso<sup>4,2</sup>, K. Daumiller<sup>42</sup>, B.R. Dawson<sup>13</sup>, R.M. de Almeida<sup>29</sup>, J. de Jesús<sup>7,42</sup>, S.J. de Jong<sup>80,81</sup>, J.R.T. de Mello Neto<sup>27,28</sup>, I. De Mitri<sup>46,47</sup>, J. de Oliveira<sup>18</sup>, D. de Oliveira Franco<sup>49</sup>, F. de Palma<sup>57,49</sup>, V. de Souza<sup>20</sup>, B.P. de Souza de Errico<sup>27</sup>, E. De Vito<sup>57,49</sup>, A. Del Popolo<sup>59,48</sup>, O. Deligny<sup>35</sup>, N. Denner<sup>33</sup>, L. Deval<sup>42,7</sup>, A. di Matteo<sup>53</sup>, M. Dobre<sup>74</sup>, C. Dobrigkeit<sup>22</sup>, J.C. D'Olivo<sup>69</sup>, L.M. Domingues Mendes<sup>73,16</sup>, Q. Dorosti<sup>45</sup>, J.C. dos Anjos<sup>16</sup>, R.C. dos Anjos<sup>26</sup>, J. Ebr<sup>33</sup>, F. Ellwanger<sup>42</sup>, M. Emam<sup>80,81</sup>, R. Engel<sup>40,42</sup>, I. Epicoco<sup>57,49</sup>, M. Erdmann<sup>43</sup>, A. Etchegoyen<sup>7,12</sup>, C. Evoli<sup>46,47</sup>, H. Falcke<sup>80,82,81</sup>, G. Farrar<sup>88</sup>, A.C. Fauth<sup>22</sup>, N. Fazzini<sup>8</sup>, F. Feldbusch<sup>41</sup>, F. Fenu<sup>42,f</sup>, A. Fernandes<sup>73</sup>, B. Fick<sup>87</sup>, J.M. Figueira<sup>7</sup>, A. Filipčić<sup>77,76</sup>, T. Fitoussi<sup>42</sup>, B. Flaggs<sup>90</sup>, T. Fodran<sup>80</sup>, T. Fujii<sup>89,h</sup>, A. Fuster<sup>7,12</sup>, C. Galea<sup>80</sup>, C. Galelli<sup>60,50</sup>, B. García<sup>6</sup>, C. Gaudu<sup>39</sup>, H. Gemmeke<sup>41</sup>, F. Gesualdi<sup>7,42</sup>, A. Gherghel-Lascu<sup>74</sup>, P.L. Ghia<sup>35</sup>, U. Giaccari<sup>49</sup>, J. Glombitza<sup>43,d</sup>, F. Gobbi<sup>10</sup>, F. Gollan<sup>7</sup>, G. Golup<sup>1</sup>, M. Gómez Berisso<sup>1</sup>, P.F. Gómez Vitale<sup>11</sup>, J.P. Gongora<sup>11</sup>, J.M. González<sup>1</sup>, N. González<sup>7</sup>, D. Góra<sup>71</sup>, A. Gorgi<sup>55,53</sup>, M. Gottowik<sup>79</sup>, T.D. Grubb<sup>13</sup>, F. Guarino<sup>61,51</sup>, G.P. Guedes<sup>23</sup>, E. Guido<sup>45</sup>, L. Güllow<sup>42</sup>, S. Hahn<sup>40</sup>, P. Hamal<sup>33</sup>, M.R. Hampel<sup>7</sup>, P. Hansen<sup>3</sup>, D. Harari<sup>1</sup>, V.M. Harvey<sup>13</sup>, A. Haungs<sup>42</sup>, T. Hebbeker<sup>43</sup>, C. Hojvat<sup>8</sup>, J.R. Hörandel<sup>80,81</sup>, P. Horvath<sup>34</sup>, M. Hrabovský<sup>34</sup>, T. Huege<sup>42,15</sup>, A. Insolia<sup>59,48</sup>, P.G. Isar<sup>75</sup>, P. Janecek<sup>33</sup>, V. Jilek<sup>33</sup>, J.A. Johnsen<sup>85</sup>, J. Jurysek<sup>33</sup>, K.-H. Kampert<sup>39</sup>, B. Keilhauer<sup>42</sup>, A. Khakurdikar<sup>80</sup>, V.V. Kizakke Covilakam<sup>7,42</sup>, H.O. Klages<sup>42</sup>, M. Kleifges<sup>41</sup>, F. Knapp<sup>40</sup>, J. Köhler<sup>42</sup>, N. Kunka<sup>41</sup>, B.L. Lago<sup>17</sup>, N. Langner<sup>43</sup>, M.A. Leigui de Oliveira<sup>25</sup>, Y. Lema-Capeans<sup>79</sup>, A. Letessier-Selvon<sup>36</sup>, I. Lhenry-Yvon<sup>35</sup>, L. Lopes<sup>73</sup>, L. Lu<sup>91</sup>, Q. Luce<sup>40</sup>, J.P. Lundquist<sup>76</sup>, A. Machado Payeras<sup>22</sup>, M. Majercakova<sup>33</sup>, D. Mandat<sup>33</sup>, B.C. Manning<sup>13</sup>, P. Mantsch<sup>8</sup>, F.M. Mariani<sup>60,50</sup>, A.G. Mariuzzi<sup>3</sup>, I.C. Mariş<sup>14</sup>, G. Marsella<sup>62,48</sup>, D. Martello<sup>57,49</sup>, S. Martinelli<sup>42,7</sup>, O. Martínez Bravo<sup>65</sup>, M.A. Martins<sup>79</sup>, H.-J. Mathes<sup>42</sup>, J. Matthews<sup>a</sup>, G. Matthiae<sup>63,52</sup>, E. Mayotte<sup>85,39</sup>, S. Mayotte<sup>85</sup>, P.O. Mazur<sup>8</sup>, G. Medina-Tanco<sup>69</sup>, J. Meinert<sup>39</sup>, D. Melo<sup>7</sup>, A. Menshikov<sup>41</sup>, C. Merx<sup>42</sup>, S. Michal<sup>33</sup>, M.I. Micheletti<sup>5</sup>, L. Miramonti<sup>60,50</sup>, S. Mollerach<sup>1</sup>, F. Montanet<sup>37</sup>, L. Morejon<sup>39</sup>, C. Morello<sup>55,53</sup>, K. Mulrey<sup>80,81</sup>, R. Mussa<sup>53</sup>, W.M. Namasaka<sup>39</sup>, S. Negi<sup>33</sup>, L. Nellen<sup>69</sup>, K. Nguyen<sup>87</sup>, G. Nicora<sup>9</sup>, M. Niechciol<sup>45</sup>, D. Nitz<sup>87</sup>, D. Nosek<sup>32</sup>, V. Novotny<sup>32</sup>, L. Nožka<sup>34</sup>, A. Nucita<sup>57,49</sup>, L.A. Núñez<sup>31</sup>, C. Oliveira<sup>20</sup>, M. Palatka<sup>33</sup>, J. Pallotta<sup>9</sup>, S. Panja<sup>33</sup>, G. Parente<sup>79</sup>, T. Paulsen<sup>39</sup>, J. Pawlowsky<sup>39</sup>, M. Pech<sup>33</sup>, J. Pękala<sup>71</sup>, R. Pelayo<sup>66</sup>, L.A.S. Pereira<sup>24</sup>, E.E. Pereira Martins<sup>40,7</sup>, J. Perez Armand<sup>21</sup>, C. Pérez Bertolli<sup>7,42</sup>, L. Perrone<sup>57,49</sup>, S. Pettera<sup>46,47</sup>, C. Petrucci<sup>58,47</sup>, T. Pierog<sup>42</sup>, M. Pimenta<sup>73</sup>, M. Platino<sup>7</sup>, B. Pont<sup>80</sup>, M. Pothast<sup>81,80</sup>, M. Pourmohammad Shahvar<sup>62,48</sup>, P. Privitera<sup>89</sup>, M. Prouza<sup>33</sup>, S. Querschfeld<sup>39</sup>, J. Rautenberg<sup>39</sup>, D. Ravnani<sup>7</sup>, J.V. Reginatto Akim<sup>22</sup>, M. Reininghaus<sup>40</sup>, J. Ridky<sup>33</sup>, F. Riehn<sup>79</sup>, M. Risse<sup>45</sup>, V. Rizi<sup>58,47</sup>, W. Rodrigues de Carvalho<sup>80</sup>, E. Rodriguez<sup>7,42</sup>, J. Rodriguez Rojo<sup>11</sup>, M.J. Roncoroni<sup>7</sup>, S. Rossoni<sup>44</sup>, M. Roth<sup>42</sup>, E. Roulet<sup>1</sup>, A.C. Rovero<sup>4</sup>, P. Ruehl<sup>45</sup>, A. Saftoiu<sup>74</sup>, M. Saharan<sup>80</sup>, F. Salamida<sup>58,47</sup>, H. Salazar<sup>65</sup>, G. Salina<sup>52</sup>, J.D. Sanabria Gomez<sup>31</sup>, F. Sánchez<sup>7</sup>, E.M. Santos<sup>21</sup>, E. Santos<sup>33</sup>, F. Sarazin<sup>85</sup>, R. Sarmento<sup>73</sup>, R. Sato<sup>11</sup>, P. Savina<sup>91</sup>, C.M. Schäfer<sup>40</sup>, V. Scherini<sup>57,49</sup>, H. Schieler<sup>42</sup>, M. Schimassek<sup>35</sup>, M. Schimp<sup>39</sup>, D. Schmidt<sup>42</sup>, O. Scholten<sup>15,j</sup>, H. Schoorlemmer<sup>80,81</sup>, P. Schovánek<sup>33</sup>, F.G. Schröder<sup>90,42</sup>, J. Schulte<sup>43</sup>, T. Schulz<sup>42</sup>, S.J. Sciutto<sup>3</sup>, M. Scornavacche<sup>7,42</sup>, A. Sedoski<sup>7</sup>, A. Segreto<sup>54,48</sup>, S. Sehgal<sup>39</sup>, S.U. Shivashankara<sup>76</sup>, G. Sigl<sup>44</sup>, G. Silli<sup>7</sup>, O. Sima<sup>74,c</sup>, K. Simkova<sup>15</sup>, F. Simon<sup>41</sup>, R. Smau<sup>74</sup>, R. Šmída<sup>89</sup>

P. Sommers<sup>1</sup>, J.F. Soriano<sup>86</sup>, R. Squartini<sup>10</sup>, M. Stadelmaier<sup>50,60,42</sup>, S. Stanič<sup>76</sup>, J. Stasielak<sup>71</sup>, P. Stassi<sup>37</sup>, S. Strähnz<sup>40</sup>, M. Straub<sup>43</sup>, T. Suomijärvi<sup>38</sup>, A.D. Supanitsky<sup>7</sup>, Z. Svozilikova<sup>33</sup>, Z. Szadkowski<sup>72</sup>, F. Tairli<sup>13</sup>, A. Tapia<sup>30</sup>, C. Taricco<sup>64,53</sup>, C. Timmermans<sup>81,80</sup>, O. Tkachenko<sup>42</sup>, P. Tobiska<sup>33</sup>, C.J. Todero Peixoto<sup>19</sup>, B. Tomé<sup>73</sup>, Z. Torrès<sup>37</sup>, A. Travaini<sup>10</sup>, P. Travnicek<sup>33</sup>, C. Trimarelli<sup>58,47</sup>, M. Tueros<sup>3</sup>, M. Unger<sup>42</sup>, L. Vaclavěk<sup>34</sup>, M. Vacula<sup>34</sup>, J.F. Valdés Galicia<sup>69</sup>, L. Valore<sup>61,51</sup>, E. Varela<sup>65</sup>, A. Vásquez-Ramírez<sup>31</sup>, D. Veberič<sup>42</sup>, C. Ventura<sup>28</sup>, I.D. Vergara Quispe<sup>3</sup>, V. Verzi<sup>52</sup>, J. Vicha<sup>33</sup>, J. Vink<sup>83</sup>, S. Vorobiov<sup>76</sup>, C. Watanabe<sup>27</sup>, A.A. Watson<sup>e</sup>, A. Weindl<sup>42</sup>, L. Wiencke<sup>85</sup>, H. Wilczyński<sup>71</sup>, D. Wittkowski<sup>39</sup>, B. Wundheiler<sup>7</sup>, B. Yue<sup>39</sup>, A. Yushkov<sup>33</sup>, O. Zapparrata<sup>14</sup>, E. Zas<sup>79</sup>, D. Zavrtnik<sup>76,77</sup>, M. Zavrtnik<sup>77,76</sup>

### The Pierre Auger Collaboration

- <sup>1</sup> Centro Atómico Bariloche and Instituto Balseiro (CNEA-UNCuyo-CONICET), San Carlos de Bariloche, Argentina
- <sup>2</sup> Departamento de Física and Departamento de Ciencias de la Atmósfera y los Océanos, FCEyN, Universidad de Buenos Aires and CONICET, Buenos Aires, Argentina
- <sup>3</sup> IFLP, Universidad Nacional de La Plata and CONICET, La Plata, Argentina
- <sup>4</sup> Instituto de Astronomía y Física del Espacio (IAFE, CONICET-UBA), Buenos Aires, Argentina
- <sup>5</sup> Instituto de Física de Rosario (IFIR) – CONICET/U.N.R. and Facultad de Ciencias Bioquímicas y Farmacéuticas U.N.R., Rosario, Argentina
- <sup>6</sup> Instituto de Tecnologías en Detección y Astropartículas (CNEA, CONICET, UNSAM), and Universidad Tecnológica Nacional – Facultad Regional Mendoza (CONICET/CNEA), Mendoza, Argentina
- <sup>7</sup> Instituto de Tecnologías en Detección y Astropartículas (CNEA, CONICET, UNSAM), Buenos Aires, Argentina
- <sup>8</sup> International Center of Advanced Studies and Instituto de Ciencias Físicas, ECyT-UNSAM and CONICET, Campus Miguelete – San Martín, Buenos Aires, Argentina
- <sup>9</sup> Laboratorio Atmósfera – Departamento de Investigaciones en Láseres y sus Aplicaciones – UNIDEF (CITEDEF-CONICET), Argentina
- <sup>10</sup> Observatorio Pierre Auger, Malargüe, Argentina
- <sup>11</sup> Observatorio Pierre Auger and Comisión Nacional de Energía Atómica, Malargüe, Argentina
- <sup>12</sup> Universidad Tecnológica Nacional – Facultad Regional Buenos Aires, Buenos Aires, Argentina
- <sup>13</sup> University of Adelaide, Adelaide, S.A., Australia
- <sup>14</sup> Université Libre de Bruxelles (ULB), Brussels, Belgium
- <sup>15</sup> Vrije Universiteit Brussels, Brussels, Belgium
- <sup>16</sup> Centro Brasileiro de Pesquisas Físicas, Rio de Janeiro, RJ, Brazil
- <sup>17</sup> Centro Federal de Educação Tecnológica Celso Suckow da Fonseca, Petropolis, Brazil
- <sup>18</sup> Instituto Federal de Educação, Ciência e Tecnologia do Rio de Janeiro (IFRJ), Brazil
- <sup>19</sup> Universidade de São Paulo, Escola de Engenharia de Lorena, Lorena, SP, Brazil
- <sup>20</sup> Universidade de São Paulo, Instituto de Física de São Carlos, São Carlos, SP, Brazil
- <sup>21</sup> Universidade de São Paulo, Instituto de Física, São Paulo, SP, Brazil
- <sup>22</sup> Universidade Estadual de Campinas (UNICAMP), IFGW, Campinas, SP, Brazil
- <sup>23</sup> Universidade Estadual de Feira de Santana, Feira de Santana, Brazil
- <sup>24</sup> Universidade Federal de Campina Grande, Centro de Ciências e Tecnologia, Campina Grande, Brazil
- <sup>25</sup> Universidade Federal do ABC, Santo André, SP, Brazil
- <sup>26</sup> Universidade Federal do Paraná, Setor Palotina, Palotina, Brazil
- <sup>27</sup> Universidade Federal do Rio de Janeiro, Instituto de Física, Rio de Janeiro, RJ, Brazil
- <sup>28</sup> Universidade Federal do Rio de Janeiro (UFRJ), Observatório do Valongo, Rio de Janeiro, RJ, Brazil
- <sup>29</sup> Universidade Federal Fluminense, EEIMVR, Volta Redonda, RJ, Brazil
- <sup>30</sup> Universidad de Medellín, Medellín, Colombia
- <sup>31</sup> Universidad Industrial de Santander, Bucaramanga, Colombia
- <sup>32</sup> Charles University, Faculty of Mathematics and Physics, Institute of Particle and Nuclear Physics, Prague, Czech Republic
- <sup>33</sup> Institute of Physics of the Czech Academy of Sciences, Prague, Czech Republic
- <sup>34</sup> Palacky University, Olomouc, Czech Republic
- <sup>35</sup> CNRS/IN2P3, IJCLab, Université Paris-Saclay, Orsay, France
- <sup>36</sup> Laboratoire de Physique Nucléaire et de Hautes Energies (LPNHE), Sorbonne Université, Université de Paris, CNRS-IN2P3, Paris, France
- <sup>37</sup> Univ. Grenoble Alpes, CNRS, Grenoble Institute of Engineering Univ. Grenoble Alpes, LPSC-IN2P3, 38000 Grenoble, France
- <sup>38</sup> Université Paris-Saclay, CNRS/IN2P3, IJCLab, Orsay, France
- <sup>39</sup> Bergische Universität Wuppertal, Department of Physics, Wuppertal, Germany
- <sup>40</sup> Karlsruhe Institute of Technology (KIT), Institute for Experimental Particle Physics, Karlsruhe, Germany
- <sup>41</sup> Karlsruhe Institute of Technology (KIT), Institut für Prozessdatenverarbeitung und Elektronik, Karlsruhe, Germany



- 42 Karlsruhe Institute of Technology (KIT), Institute for Astroparticle Physics, Karlsruhe, Germany
- 43 RWTH Aachen University, III. Physikalisches Institut A, Aachen, Germany
- 44 Universität Hamburg, II. Institut für Theoretische Physik, Hamburg, Germany
- 45 Universität Siegen, Department Physik – Experimentelle Teilchenphysik, Siegen, Germany
- 46 Gran Sasso Science Institute, L’Aquila, Italy
- 47 INFN Laboratori Nazionali del Gran Sasso, Assergi (L’Aquila), Italy
- 48 INFN, Sezione di Catania, Catania, Italy
- 49 INFN, Sezione di Lecce, Lecce, Italy
- 50 INFN, Sezione di Milano, Milano, Italy
- 51 INFN, Sezione di Napoli, Napoli, Italy
- 52 INFN, Sezione di Roma “Tor Vergata”, Roma, Italy
- 53 INFN, Sezione di Torino, Torino, Italy
- 54 Istituto di Astrofisica Spaziale e Fisica Cosmica di Palermo (INAF), Palermo, Italy
- 55 Osservatorio Astrofisico di Torino (INAF), Torino, Italy
- 56 Politecnico di Milano, Dipartimento di Scienze e Tecnologie Aerospaziali, Milano, Italy
- 57 Università del Salento, Dipartimento di Matematica e Fisica “E. De Giorgi”, Lecce, Italy
- 58 Università dell’Aquila, Dipartimento di Scienze Fisiche e Chimiche, L’Aquila, Italy
- 59 Università di Catania, Dipartimento di Fisica e Astronomia “Ettore Majorana“, Catania, Italy
- 60 Università di Milano, Dipartimento di Fisica, Milano, Italy
- 61 Università di Napoli “Federico II”, Dipartimento di Fisica “Ettore Pancini”, Napoli, Italy
- 62 Università di Palermo, Dipartimento di Fisica e Chimica “E. Segrè”, Palermo, Italy
- 63 Università di Roma “Tor Vergata”, Dipartimento di Fisica, Roma, Italy
- 64 Università Torino, Dipartimento di Fisica, Torino, Italy
- 65 Benemérita Universidad Autónoma de Puebla, Puebla, México
- 66 Unidad Profesional Interdisciplinaria en Ingeniería y Tecnologías Avanzadas del Instituto Politécnico Nacional (UPIITA-IPN), México, D.F., México
- 67 Universidad Autónoma de Chiapas, Tuxtla Gutiérrez, Chiapas, México
- 68 Universidad Michoacana de San Nicolás de Hidalgo, Morelia, Michoacán, México
- 69 Universidad Nacional Autónoma de México, México, D.F., México
- 70 Universidad Nacional de San Agustín de Arequipa, Facultad de Ciencias Naturales y Formales, Arequipa, Peru
- 71 Institute of Nuclear Physics PAN, Krakow, Poland
- 72 University of Łódź, Faculty of High-Energy Astrophysics, Łódź, Poland
- 73 Laboratório de Instrumentação e Física Experimental de Partículas – LIP and Instituto Superior Técnico – IST, Universidade de Lisboa – UL, Lisboa, Portugal
- 74 “Horia Hulubei” National Institute for Physics and Nuclear Engineering, Bucharest-Magurele, Romania
- 75 Institute of Space Science, Bucharest-Magurele, Romania
- 76 Center for Astrophysics and Cosmology (CAC), University of Nova Gorica, Nova Gorica, Slovenia
- 77 Experimental Particle Physics Department, J. Stefan Institute, Ljubljana, Slovenia
- 78 Universidad de Granada and C.A.F.P.E., Granada, Spain
- 79 Instituto Galego de Física de Altas Enerxías (IGFAE), Universidade de Santiago de Compostela, Santiago de Compostela, Spain
- 80 IMAPP, Radboud University Nijmegen, Nijmegen, The Netherlands
- 81 Nationaal Instituut voor Kernfysica en Hoge Energie Fysica (NIKHEF), Science Park, Amsterdam, The Netherlands
- 82 Stichting Astronomisch Onderzoek in Nederland (ASTRON), Dwingeloo, The Netherlands
- 83 Universiteit van Amsterdam, Faculty of Science, Amsterdam, The Netherlands
- 84 Case Western Reserve University, Cleveland, OH, USA
- 85 Colorado School of Mines, Golden, CO, USA
- 86 Department of Physics and Astronomy, Lehman College, City University of New York, Bronx, NY, USA
- 87 Michigan Technological University, Houghton, MI, USA
- 88 New York University, New York, NY, USA
- 89 University of Chicago, Enrico Fermi Institute, Chicago, IL, USA
- 90 University of Delaware, Department of Physics and Astronomy, Bartol Research Institute, Newark, DE, USA
- 91 University of Wisconsin-Madison, Department of Physics and WIPAC, Madison, WI, USA

<sup>a</sup> Louisiana State University, Baton Rouge, LA, USA

<sup>b</sup> Institut universitaire de France (IUF), France

<sup>c</sup> also at University of Bucharest, Physics Department, Bucharest, Romania

<sup>d</sup> now at ECAP, Erlangen, Germany

- e* School of Physics and Astronomy, University of Leeds, Leeds, United Kingdom
- f* now at Agenzia Spaziale Italiana (ASI). Via del Politecnico 00133, Roma, Italy
- g* Fermi National Accelerator Laboratory, Fermilab, Batavia, IL, USA
- h* now at Graduate School of Science, Osaka Metropolitan University, Osaka, Japan
- i* Max-Planck-Institut für Radioastronomie, Bonn, Germany
- j* also at Kapteyn Institute, University of Groningen, Groningen, The Netherlands
- k* Colorado State University, Fort Collins, CO, USA
- l* Pennsylvania State University, University Park, PA, USA

### Appendix A: Parameterization of MC templates

The MC templates of the two-dimensional distributions of  $(S, X)$  normalized by the total number of simulated showers  $N_{\text{MC}}$  and weighted to correspond to the measured energy spectrum are fitted using the following function:

$$\begin{aligned}\Phi &= \frac{dN}{N_{\text{MC}} dX dS} \\ &= A_{\text{Gauss}} A_{\text{Gumbel}} f_{\text{Gumbel}}(X) f_{\text{Gauss}}(X, S).\end{aligned}\quad (\text{A1})$$

The  $X_{\text{max}}$  part is described by the generalized Gumbel distribution [52]

$$f_{\text{Gumbel}}(X) = \exp[-\lambda(x - \exp x)], \quad (\text{A2})$$

where  $x = (X - m)/s$  and the ground-signal part is assumed to follow the Gaussian distribution with the mean value linearly dependent on  $X$

$$f_{\text{Gauss}}(X, S) = \exp\left[-\frac{y^2}{2r^2}\right], \quad (\text{A3})$$

where  $y = S - pX - q$ . The normalization terms are given by

$$A_{\text{Gauss}} = \frac{1}{\sqrt{2\pi}r} \quad \text{and} \quad A_{\text{Gumbel}} = \frac{\lambda^\lambda}{s\Gamma(\lambda)}, \quad (\text{A4})$$

where  $\Gamma$  is the gamma function. The six free parameters in each of the MC template fits are  $m$ ,  $s$ ,  $\lambda$  of the generalized Gumbel distribution, and  $p$ ,  $q$ ,  $r$  of the Gaussian part. These fitted parameters for the three models used in this work are listed in Tables V to VII.

In Fig. 11, we show the mean and the standard deviation of the pull distribution for the description of each bin of the two-dimensional distribution  $(S, X)$ , see examples in Fig. 1, by the fits with a function Eq. (A1). The goodness of the description of the MC templates with these fits is also tested with the two-dimensional Kolmogorov-Smirnov test [53] demonstrating very good consistency between the MC templates and MC distributions of  $(S, X)$ .

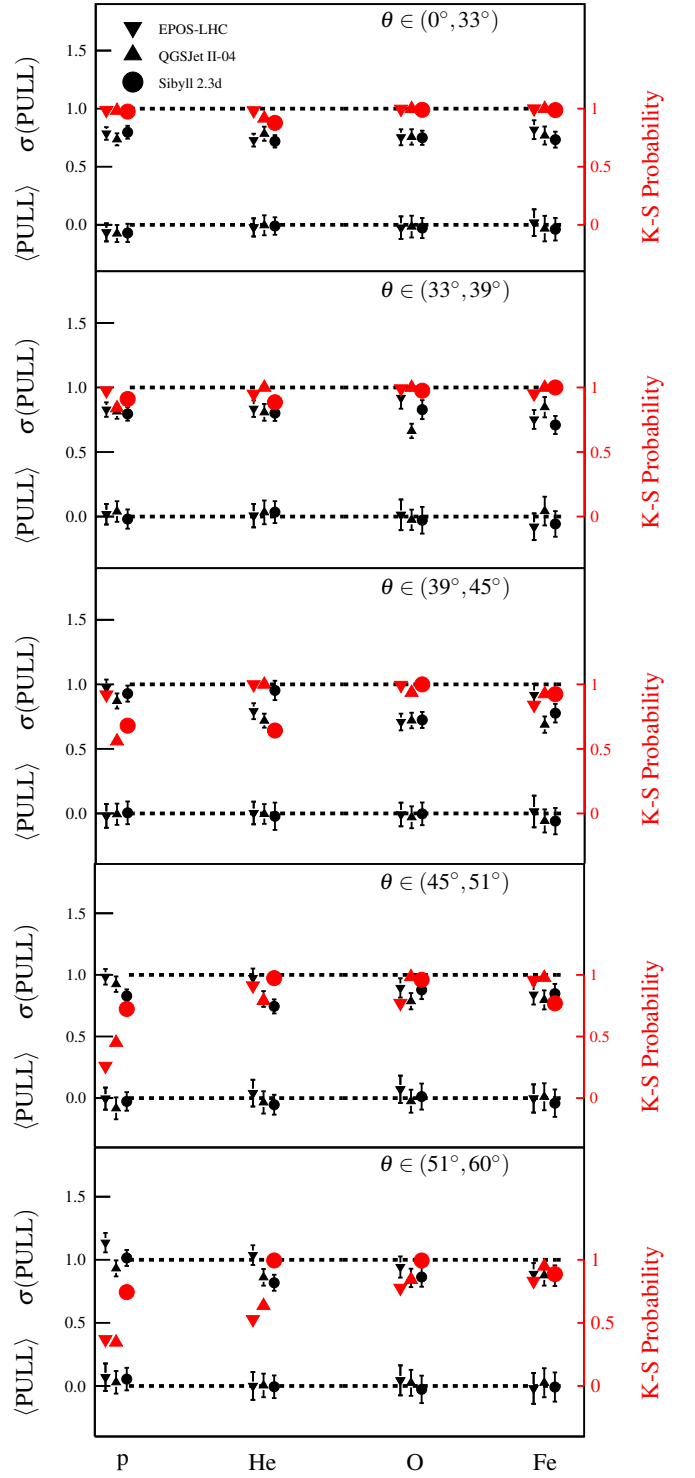


FIG. 11: The mean and the standard deviation of the MC distribution of  $\text{PULL} = (N_{\text{fit}} - N_{\text{mc}})/\sqrt{N_{\text{mc}}}$  where  $N_{\text{fit}}$  is the value of the fitted function Eq. (A1) and  $N_{\text{mc}}$  is the value for each bin of the two-dimensional MC distribution  $(S, X)$ . The probability of consistency between the parameterized function and MC distribution (red) is tested with the two-dimensional Kolmogorov-Smirnov (K-S) test.  $E = 10^{18.5}$  to  $10^{19.0}$  eV.

TABLE V: Parameters of MC templates for air showers generated with model EPOS-LHC and initiated by a primary particle  $i$ , see Eq. (A1).

$i$	$\theta$	$f_{\text{had}}$	$m/(\text{g}/\text{cm}^2)$	$s/(\text{g}/\text{cm}^2)$	$\lambda$	$p/(\text{VEM}/(\text{g}/\text{cm}^2))$	$q/\text{VEM}$	$r/\text{VEM}$
p	(0°, 33°)	0.64	757	49	1.2	$-1.6 \times 10^{-2}$	3.10	34.3
p	(33°, 39°)	0.67	755	44	1.1	$-3 \times 10^{-3}$	2.85	22.9
p	(39°, 45°)	0.71	755	43	$8.9 \times 10^{-1}$	$6.7 \times 10^{-3}$	2.88	13.8
p	(45°, 51°)	0.77	754	47	$9.8 \times 10^{-1}$	$1.2 \times 10^{-2}$	2.83	6.3
p	(51°, 60°)	0.88	757	53	1.3	$1.1 \times 10^{-2}$	2.68	3.0
He	(0°, 33°)	0.67	741	51	2.1	$-2.7 \times 10^{-3}$	2.78	25.3
He	(33°, 39°)	0.70	739	49	2.0	$7.7 \times 10^{-3}$	2.60	16.1
He	(39°, 45°)	0.75	742	61	2.6	$1.7 \times 10^{-2}$	2.47	7.2
He	(45°, 51°)	0.81	741	57	2.5	$1.7 \times 10^{-2}$	2.55	3.5
He	(51°, 60°)	0.91	739	64	2.7	$1.3 \times 10^{-2}$	2.50	2.6
O	(0°, 33°)	0.71	714	63	5.7	$-1.6 \times 10^{-3}$	2.61	26.2
O	(33°, 39°)	0.74	714	65	5.4	$8.8 \times 10^{-3}$	2.45	16.9
O	(39°, 45°)	0.78	714	66	5.5	$1.9 \times 10^{-2}$	2.41	6.8
O	(45°, 51°)	0.84	711	48	2.9	$2.2 \times 10^{-2}$	2.37	1.8
O	(51°, 60°)	0.93	714	82	7.5	$1.7 \times 10^{-2}$	2.33	1.0
Fe	(0°, 33°)	0.74	688	76	13.2	$1.4 \times 10^{-2}$	2.65	17.4
Fe	(33°, 39°)	0.77	688	100	19.4	$1.3 \times 10^{-2}$	2.62	15.7
Fe	(39°, 45°)	0.81	687	81	13.4	$1.4 \times 10^{-2}$	2.51	11.9
Fe	(45°, 51°)	0.87	688	134	32.7	$1.8 \times 10^{-2}$	2.41	6.2
Fe	(51°, 60°)	0.94	687	176	48.1	$1.3 \times 10^{-2}$	2.34	4.9

TABLE VI: Same as in Table V, but for model QGSJET-II-04.

$i$	$\theta$	$f_{\text{had}}$	$m/(\text{g}/\text{cm}^2)$	$s/(\text{g}/\text{cm}^2)$	$\lambda$	$p/(\text{VEM}/(\text{g}/\text{cm}^2))$	$q/\text{VEM}$	$r/\text{VEM}$
p	(0°, 33°)	0.62	741	49	1.1	$-1 \times 10^{-2}$	2.67	27.6
p	(33°, 39°)	0.65	748	70	1.9	$1.1 \times 10^{-3}$	2.53	18.0
p	(39°, 45°)	0.69	743	56	1.3	$1.2 \times 10^{-2}$	2.54	8.0
p	(45°, 51°)	0.76	739	47	$9.5 \times 10^{-1}$	$1.6 \times 10^{-2}$	2.51	2.2
p	(51°, 60°)	0.88	737	50	1.0	$1 \times 10^{-2}$	2.35	2.7
He	(0°, 33°)	0.65	727	68	2.7	$3.3 \times 10^{-3}$	2.49	19.0
He	(33°, 39°)	0.69	729	88	4.7	$1.2 \times 10^{-2}$	2.35	11.0
He	(39°, 45°)	0.74	727	74	3.1	$1.8 \times 10^{-2}$	2.38	4.5
He	(45°, 51°)	0.80	725	60	2.3	$2 \times 10^{-2}$	2.20	$4.6 \times 10^{-1}$
He	(51°, 60°)	0.91	724	74	3.0	$1.1 \times 10^{-2}$	2.21	3.2
O	(0°, 33°)	0.69	700	99	9.5	$1.1 \times 10^{-2}$	2.40	15.1
O	(33°, 39°)	0.73	702	87	6.6	$1.5 \times 10^{-2}$	2.38	10.9
O	(39°, 45°)	0.77	702	91	7.5	$2.1 \times 10^{-2}$	2.27	4.0
O	(45°, 51°)	0.84	704	102	8.0	$2.1 \times 10^{-2}$	2.22	$8.4 \times 10^{-1}$
O	(51°, 60°)	0.93	701	84	5.6	$1.5 \times 10^{-2}$	2.15	1.8
Fe	(0°, 33°)	0.72	675	68	6.7	$1.6 \times 10^{-2}$	2.41	13.9
Fe	(33°, 39°)	0.76	676	77	8.2	$2.1 \times 10^{-2}$	2.37	8.2
Fe	(39°, 45°)	0.81	676	91	11.9	$2.2 \times 10^{-2}$	2.37	4.7
Fe	(45°, 51°)	0.87	675	87	10.7	$1.6 \times 10^{-2}$	2.24	6.4
Fe	(51°, 60°)	0.94	675	162	31.6	$1.1 \times 10^{-2}$	2.19	5.6

TABLE VII: Same as in Table V, but for model SIBYLL 2.3d.

$i$	$\theta$	$f_{\text{had}}$	$m/(\text{g}/\text{cm}^2)$	$s/(\text{g}/\text{cm}^2)$	$\lambda$	$p/(\text{VEM}/(\text{g}/\text{cm}^2))$	$q/\text{VEM}$	$r/\text{VEM}$
p	$0^\circ - 33^\circ$	0.63	769	70	2.1	$-9 \times 10^{-3}$	2.97	27.7
p	$33^\circ - 39^\circ$	0.65	771	68	1.8	$7.8 \times 10^{-4}$	2.96	19.0
p	$39^\circ - 45^\circ$	0.69	772	71	2.0	$1.1 \times 10^{-2}$	2.84	9.0
p	$45^\circ - 51^\circ$	0.76	766	62	1.5	$1.4 \times 10^{-2}$	2.73	4.4
p	$51^\circ - 60^\circ$	0.87	766	52	1.1	$9.2 \times 10^{-3}$	2.77	4.0
He	$0^\circ - 33^\circ$	0.66	747	56	2.2	$-1.3 \times 10^{-3}$	2.68	23.3
He	$33^\circ - 39^\circ$	0.69	745	50	1.8	$1 \times 10^{-2}$	2.55	13.1
He	$39^\circ - 45^\circ$	0.73	742	54	1.8	$1.6 \times 10^{-2}$	2.51	6.8
He	$45^\circ - 51^\circ$	0.80	744	58	2.1	$2 \times 10^{-2}$	2.41	$7.7 \times 10^{-1}$
He	$51^\circ - 60^\circ$	0.90	743	58	2.1	$1.5 \times 10^{-2}$	2.41	$6 \times 10^{-1}$
O	$0^\circ - 33^\circ$	0.70	715	48	2.3	$8.2 \times 10^{-3}$	2.44	18.1
O	$33^\circ - 39^\circ$	0.73	716	65	4.3	$1.6 \times 10^{-2}$	2.41	10.8
O	$39^\circ - 45^\circ$	0.77	716	62	3.6	$2.1 \times 10^{-2}$	2.36	4.9
O	$45^\circ - 51^\circ$	0.84	717	57	3.0	$1.8 \times 10^{-2}$	2.26	3.8
O	$51^\circ - 60^\circ$	0.93	717	67	3.9	$1.5 \times 10^{-2}$	2.25	1.8
Fe	$0^\circ - 33^\circ$	0.73	688	60	5.6	$1.4 \times 10^{-2}$	2.61	16.2
Fe	$33^\circ - 39^\circ$	0.77	690	62	5.8	$2.2 \times 10^{-2}$	2.40	8.6
Fe	$39^\circ - 45^\circ$	0.81	690	60	5.5	$2.3 \times 10^{-2}$	2.40	5.5
Fe	$45^\circ - 51^\circ$	0.87	690	97	12.1	$2 \times 10^{-2}$	2.35	3.8
Fe	$51^\circ - 60^\circ$	0.94	690	96	11.8	$1.9 \times 10^{-2}$	2.24	$8.7 \times 10^{-1}$

## Appendix B: MC-MC Tests of the Method

To evaluate the precision and reliability of the method, we performed tests on the MC simulations with artificially modified predictions on the event-by-event basis using the shift  $\Delta X_{\max}$  and the factors  $R_{\text{had}}(\theta_{\min})$ ,  $R_{\text{had}}(\theta_{\max})$  obtained from the fits to the measured data, see Table III.

For each model, all possible mixes of (p, He, O, Fe) with relative fractions changing in 0.1 steps are used to estimate the biases on modification parameters as a function of the mean logarithmic mass of the primary beam  $\langle \ln A \rangle$ . In the tests, we use only the composition mixes containing primary iron nuclei, as there can be additional biases stemming from the assumption on the presence of different species in the primary beam, see discussion in Section IV B. For each composition mix, five different samples of simulated showers are randomly selected with the same number of events as in the data ( $N = 2239$ ) and following the shape of the measured energy spectrum. The biases of the method for each composition mix are calculated as an average over the biases for these five random sets. The results of the MC-MC tests are summarized in Fig. 12. We find that for  $2 < \langle \ln A \rangle < 3$ , the range of the masses inferred from the data analysis (Section III B), the maximum overestimation of the fitted  $X_{\max}$  scale with the method is  $5 \text{ g/cm}^2$ , and biases on the fitted  $\langle S_{\text{had}} \rangle$  scale are within 3%. These resulting values of these MC-MC tests are considered as systematic uncertainties of the method on the modification parameters, see Fig. 16. The systematic uncertainties of the method on the primary fractions are within  $\pm 20\%$ , which corresponds to the maximum bias of the method on the fractions inferred from  $\sim 90\%$  of the studied MC samples, see Fig. 13.

The performance of the method [20] is compatible with these estimations of systematic uncertainties also for a test when one of the models is used for the analysis of the MC samples prepared with other models. This is in agreement with the expectations that the main differences in MC templates between the models are due to the differences in  $X_{\max}$  and  $S_{\text{had}}$  scales. Finally, we checked that the number of zenith-angle bins does not affect the results unless it is too small (2 bins are not enough to disentangle the hadronic and em parts of the ground signal) or too large ( $> 10$ ) when the event statistics per zenith-angle bin becomes too low.

### Appendix C: Parameterization of Attenuation of Ground Signals

The electromagnetic and hadronic signals at 1000 m were corrected for the energy evolution as  $S_{\text{em}}^{\text{ref}} = S_{\text{em}}(E^{\text{ref}}/E_{\text{FD}})$  and  $S_{\text{had}}^{\text{ref}} = S_{\text{had}}(E^{\text{ref}}/E_{\text{FD}})^{1/\beta}$ . The dependence of these average signals on the distance of  $X_{\max}$  to the ground in atmospheric depth units,  $t = X_{\text{atm}}(\theta) - X_{\max}$ , see Fig. 2, was parameterized with the Gaisser-Hillas function [54] allowing its vertical offset,

$$\langle S_{\alpha}^{\text{ref}} \rangle(t) = S_{\alpha}^0 \left( \frac{t - t_{\alpha}}{u_{\alpha} - t_{\alpha}} \right)^{Z(t_{\alpha})} \exp(Z(t)) + w_{\alpha}, \quad (\text{C1})$$

where  $\alpha = \text{had}$  or  $\text{em}$ , and normalization scaling  $Z(x) = (u_{\alpha} - x)/v_{\alpha}$ .  $t_{\alpha}$  is the value of  $X_{\text{atm}}(\theta) - X_{\max}$  where the function reaches its maximum,  $u_{\alpha}$  and  $v_{\alpha}$  are parameters without a straightforward physics interpretation, and  $S_{\alpha}^0$ ,  $w_{\alpha}$  are the rescale and offset parameters, respectively. The fitted parameters are listed in Table VIII and Table IX for em and muon signal, respectively.

The factor  $g_{\text{em},k}$  reflecting the average change of em signal due to the change of  $X_{\max}$  scale for a mix of primary species  $i$  with relative fractions  $f_i$  in zenith-angle bin  $k$  is calculated as

$$g_{\text{em},k} = \sum_i f_i \frac{S_{\text{em}}(\langle t \rangle_k - \Delta X_{\max})}{S_{\text{em}}(\langle t \rangle_k)}, \quad (\text{C2})$$

where  $\langle t \rangle_k$  is the average measured  $X_{\text{atm}}(\theta) - X_{\max}$  in a zenith-angle bin  $k$ :  $\sim 241.2 \text{ g/cm}^2$ ,  $\sim 335.5 \text{ g/cm}^2$ ,  $\sim 429.4 \text{ g/cm}^2$ ,  $\sim 556.1 \text{ g/cm}^2$ ,  $\sim 777.6 \text{ g/cm}^2$ , for respective increasing average values of the zenith angle. The factor  $g_{\text{had}}$  for the hadronic signal is obtained as

$$g_{\text{had},k} = \sum_i f_i \frac{S_{\text{had}}(\langle t \rangle_k - \Delta X_{\max})}{S_{\text{had}}(\langle t \rangle_k)} \frac{R_{\text{had}}(\langle t \rangle_k - \Delta X_{\max})}{R_{\text{had}}(\langle t \rangle_k)}, \quad (\text{C3})$$

where

$$R_{\text{had}}(\theta) \equiv R_{\text{had}}(\langle t \rangle_k) = R_{\text{had}}(\theta_{\min}) + \quad (\text{C4})$$

$$+ (R_{\text{had}}(\theta_{\max}) - R_{\text{had}}(\theta_{\min})) \frac{\langle t \rangle_k - \langle t \rangle_{\min}}{\langle t \rangle_{\max} - \langle t \rangle_{\min}}, \quad (\text{C5})$$

for  $\langle t \rangle_{\min}$  and  $\langle t \rangle_{\max}$  corresponding to minimum and maximum zenith-angle bins, respectively.

### Appendix D: Data description using fits with less freedom in MC templates

The description of Auger data using fits without any modification to MC templates and with zenith-independent  $R_{\text{had}}$  are shown in Figs. 14 and 15, respectively.

### Appendix E: Systematic Uncertainties

The individual contributions to the total systematic uncertainties are plotted in Fig. 16.

### Appendix F: Scan for the linear combinations of experimental systematic uncertainties most favorable for the models

In Fig. 17, the histogram of difference in log-likelihood expressions ( $\Delta \ln \mathcal{L}$ ) for fits using non-modified and  $\Delta X_{\max}$ ,  $R_{\text{had}}(\theta)$  modifications is shown for dense scans in linear combinations of experimental systematic uncertainties. The ranges of these uncertainties were selected in a way to estimate the closest approach to the non-modified value ( $R_{\text{had}}(\theta) = 1$ ,  $\Delta X_{\max} = 0 \text{ g/cm}^2$ ) even for cases when the uncertainties were out of the range of uncertainties quoted by

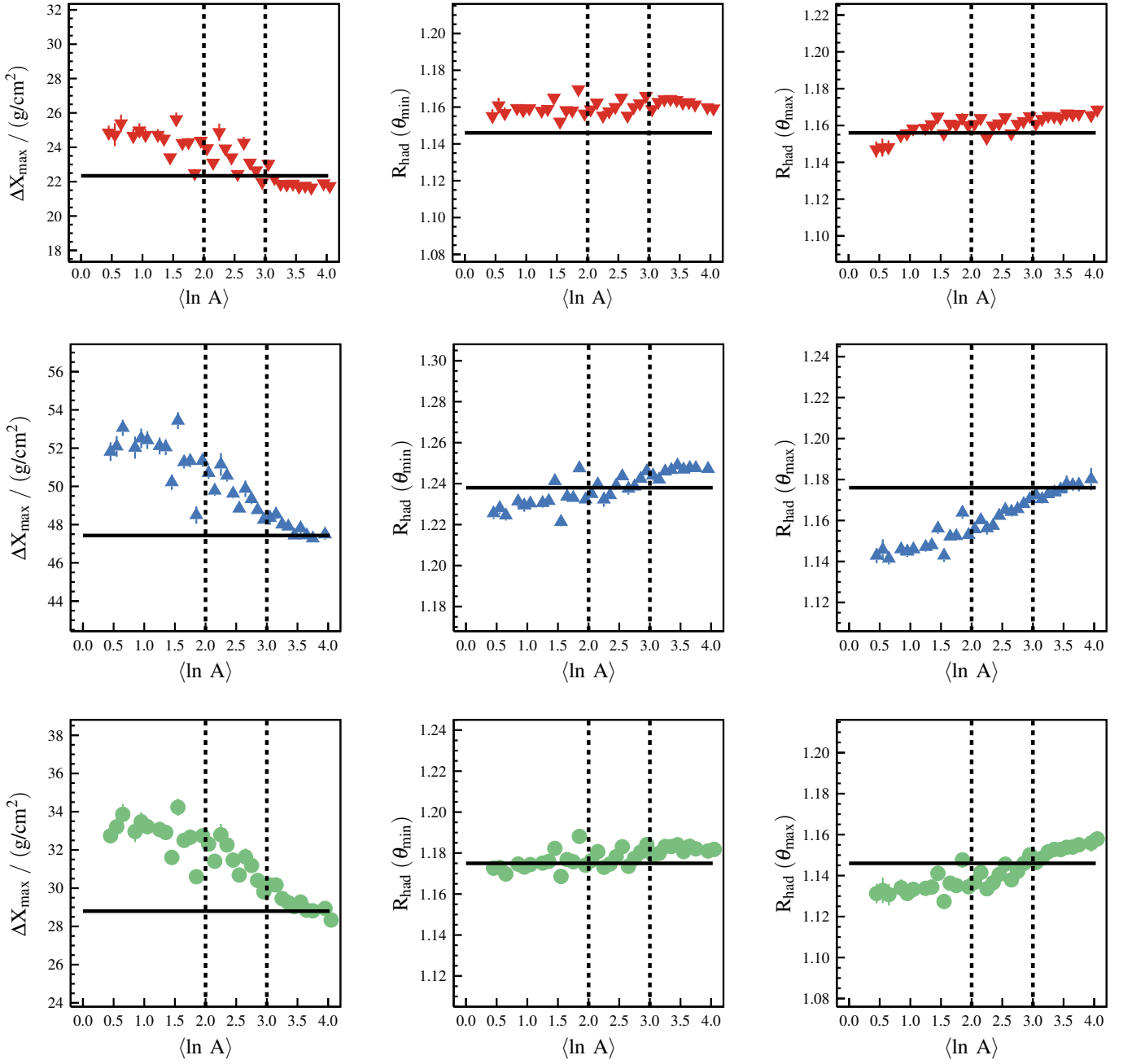


FIG. 12: The values of modification parameters found by the method (points) in the MC samples with artificially modified  $\Delta X_{\max}$  and  $R_{\text{had}}(\theta)$ . The values of artificial modifications, shown with solid horizontal lines, are applied to each simulated shower individually. The MC samples contain all possible combinations of primary nuclei with 0.1 steps in relative fractions, iron nuclei are present in all samples (relative fraction  $\geq 0.1$ ), and the mean logarithmic masses of the samples are marked on horizontal axis. Hadronic interaction models used in the tests: EPOS-LHC (top row, red), QGSJET-II-04 (middle row, blue), Sibyll 2.3d (bottom row, green). The systematic uncertainties are taken as the maximum biases within the  $\langle \ln A \rangle = 2$  to 3 range indicated with dashed lines.

Auger. It was not possible to find such a linear combination of experimental systematic uncertainties that would decrease the significance of improvement in data description below  $5\sigma$ .



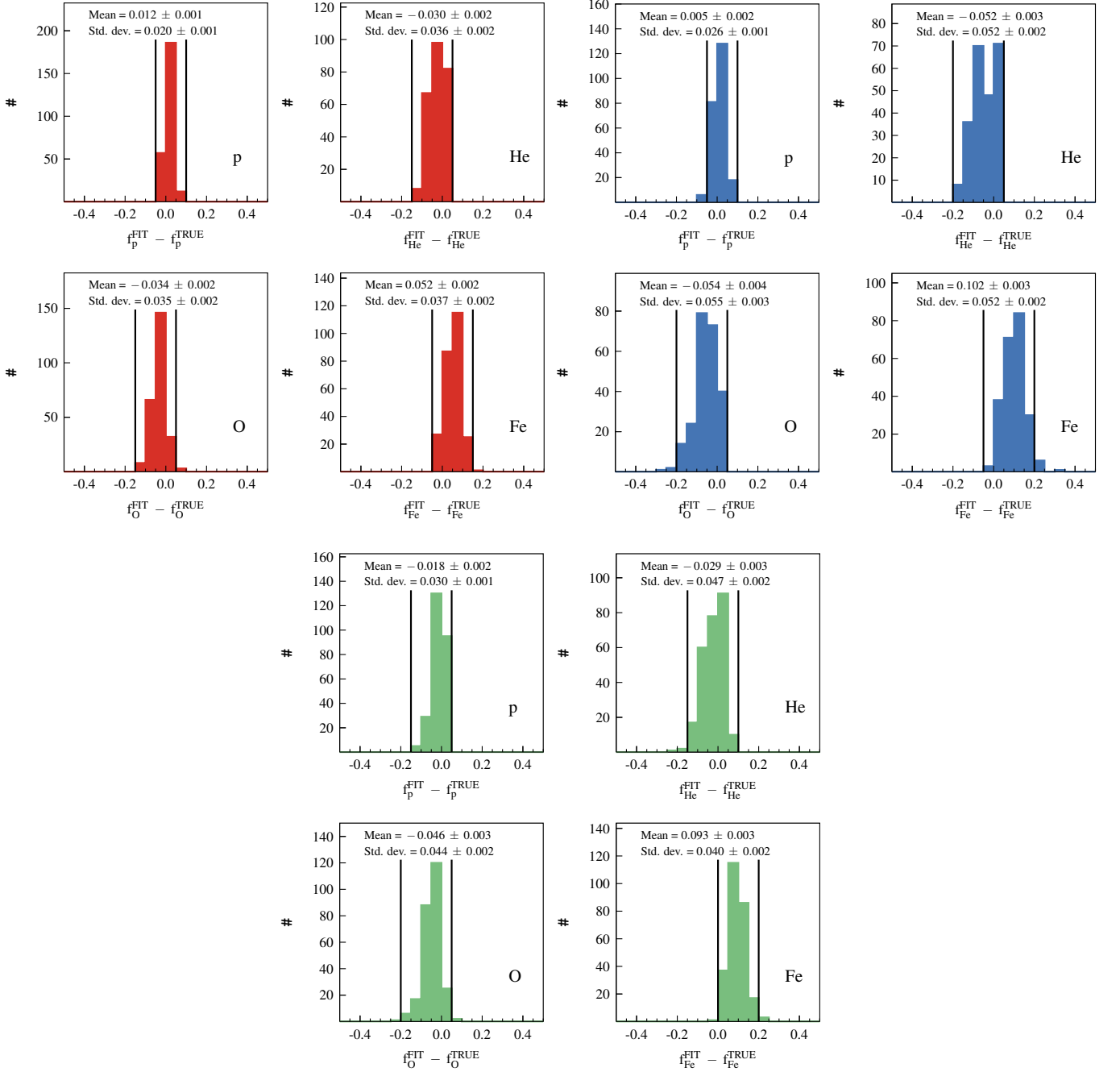


FIG. 13: Biases of the method on the (p, He, O, Fe) fractions for the MC samples with artificially modified  $\Delta X_{\max}$  and  $R_{\text{had}}(\theta)$  as described in Fig. 12. Additionally, we select the fitted relative fractions with  $0.0 \leq f_p^{\text{FIT}} \leq 0.4$ ,  $0.0 \leq f_{\text{He}}^{\text{FIT}} \leq 0.3$ ,  $0.2 \leq f_{\text{O}}^{\text{FIT}} \leq 0.6$  and  $0.0 \leq f_{\text{Fe}}^{\text{FIT}} \leq 0.5$  which correspond to the ranges obtained by the data fits. The range considered as a contribution to the systematic uncertainty is indicated by vertical lines. Hadronic interaction models used in the tests: EPOS-LHC (top left, red), QGSJET-II-04 (top-right, blue), SIBYLL 2.3d (bottom, green).

TABLE VIII: Fitted parameters of functional dependence of the average em signal part of  $S(1000)$  vs. distance of  $X_{\max}$  to the ground in atmospheric depth units for air showers initiated by a primary particle  $i$ , see Eq. (C1).

	$i$	$S_{\text{em}}^0/\text{VEM}$	$t_{\text{em}}/(\text{g}/\text{cm}^2)$	$u_{\text{em}}/(\text{g}/\text{cm}^2)$	$v_{\text{em}}/(\text{g}/\text{cm}^2)$	$w_{\text{em}}/\text{VEM}$
EPOS-LHC	p	7.58	-1,638	172	42	$1 \times 10^{-1}$
EPOS-LHC	He	7.76	-580	179	81	$6 \times 10^{-2}$
EPOS-LHC	O	7.83	-1,412	161	48	$9 \times 10^{-2}$
EPOS-LHC	Fe	7.64	-407	204	86	$7 \times 10^{-2}$
QGSJET-II-04	p	7.56	-2,006	175	35	$1 \times 10^{-1}$
QGSJET-II-04	He	7.71	-495	183	87	$3 \times 10^{-2}$
QGSJET-II-04	O	7.77	-763	167	72	$7 \times 10^{-2}$
QGSJET-II-04	Fe	7.52	-278	206	103	$2 \times 10^{-3}$
SIBYLL 2.3d	p	7.57	-1,152	176	53	$1 \times 10^{-1}$
SIBYLL 2.3d	He	7.62	-490	185	88	$3 \times 10^{-2}$
SIBYLL 2.3d	O	7.54	-665	187	73	$6 \times 10^{-2}$
SIBYLL 2.3d	Fe	7.74	-1,892	152	39	$9 \times 10^{-2}$

TABLE IX: Same as in Table VIII, but for the average hadronic-signal part of  $S(1000)$ .

	$i$	$S_{\text{had}}^0/\text{VEM}$	$t_{\text{had}}/(\text{g}/\text{cm}^2)$	$u_{\text{had}}/(\text{g}/\text{cm}^2)$	$v_{\text{had}}/(\text{g}/\text{cm}^2)$	$w_{\text{had}}/\text{VEM}$
EPOS-LHC	p	5.6	-9,999	312	9	8.3
EPOS-LHC	He	7.2	-9,419	232	16	8.4
EPOS-LHC	O	9.0	-1,298	228	98	8.7
EPOS-LHC	Fe	10.7	-8,304	171	24	9.6
QGSJET-II-04	p	5.1	-1,880	314	47	7.2
QGSJET-II-04	He	7.2	-300	255	222	6.8
QGSJET-II-04	O	7.2	-4,934	205	34	8.7
QGSJET-II-04	Fe	9.6	-1,614	166	115	8.7
SIBYLL 2.3d	p	4.8	-8,519	283	11	8.1
SIBYLL 2.3d	He	6.7	-3,557	225	44	8.0
SIBYLL 2.3d	O	8.4	-1,521	207	100	8.4
SIBYLL 2.3d	Fe	8.9	-5,078	194	35	10.1

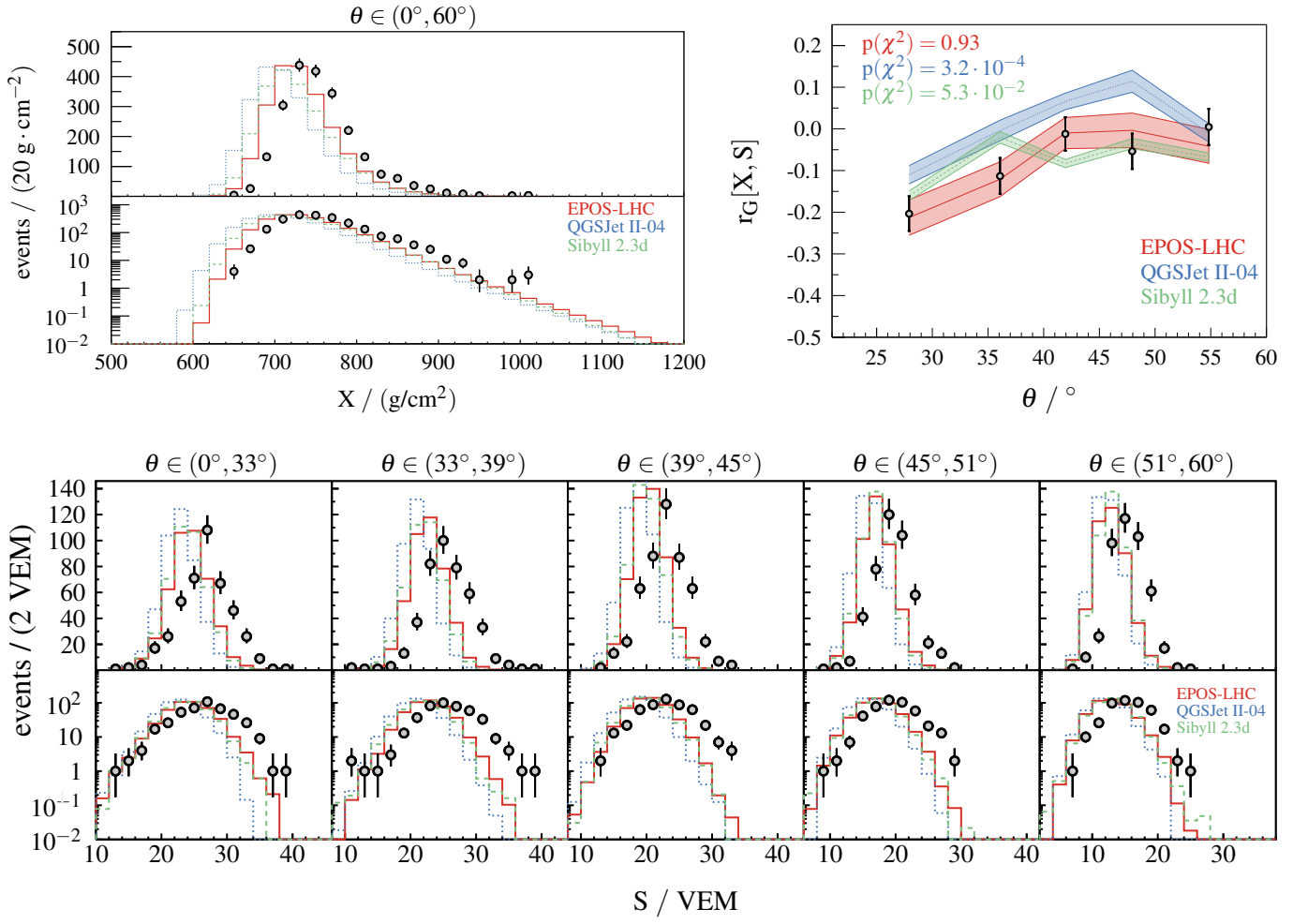


FIG. 14: Same as in Fig. 5, but for the data fits performed without any modification of the predictions of models.

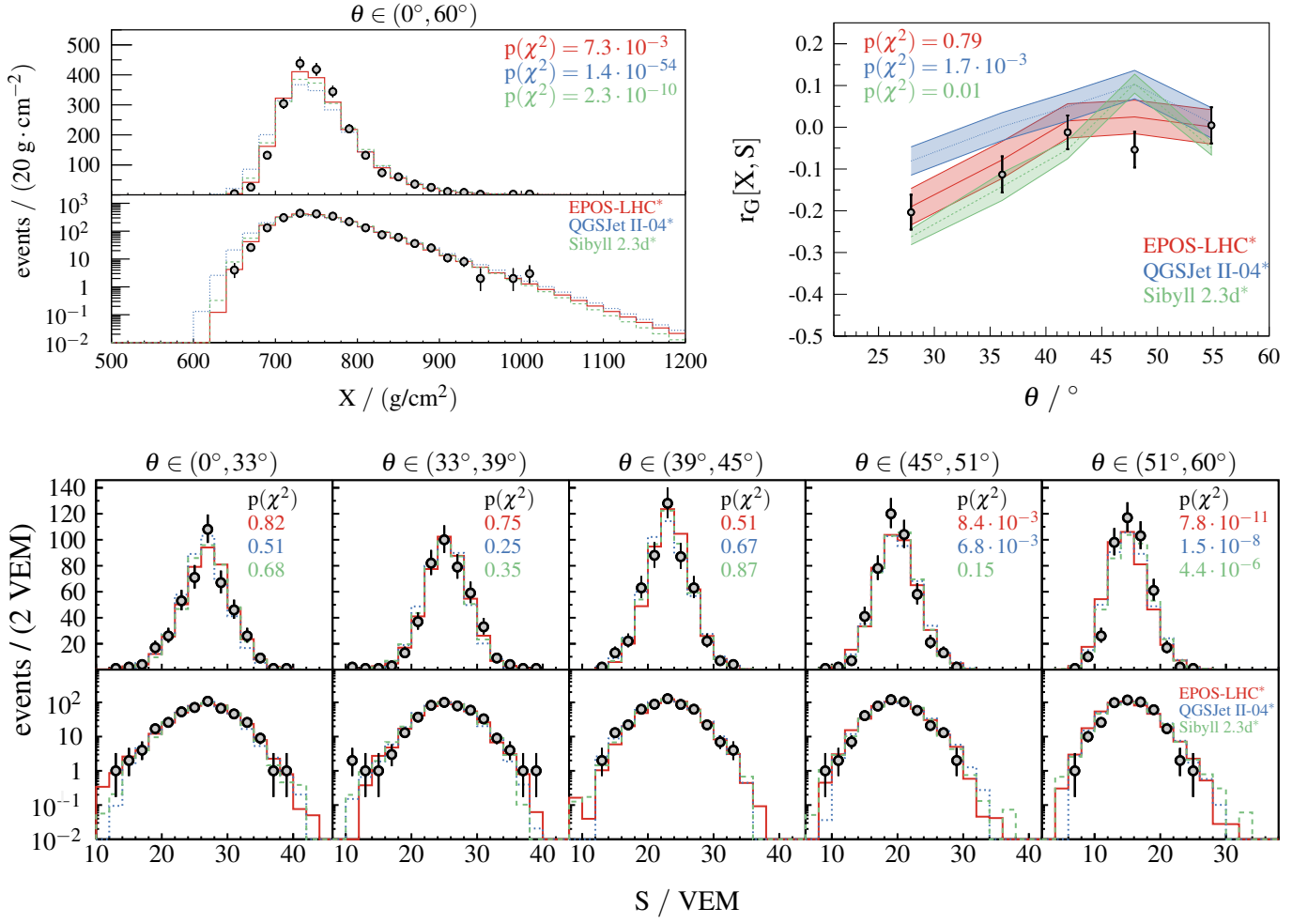


FIG. 15: Same as in Fig. 5, but for the data fits performed with only zenith-angle independent modification  $R_{\text{had}}$  of the predictions of models (denoted by \* as templates were modified from the original predictions).

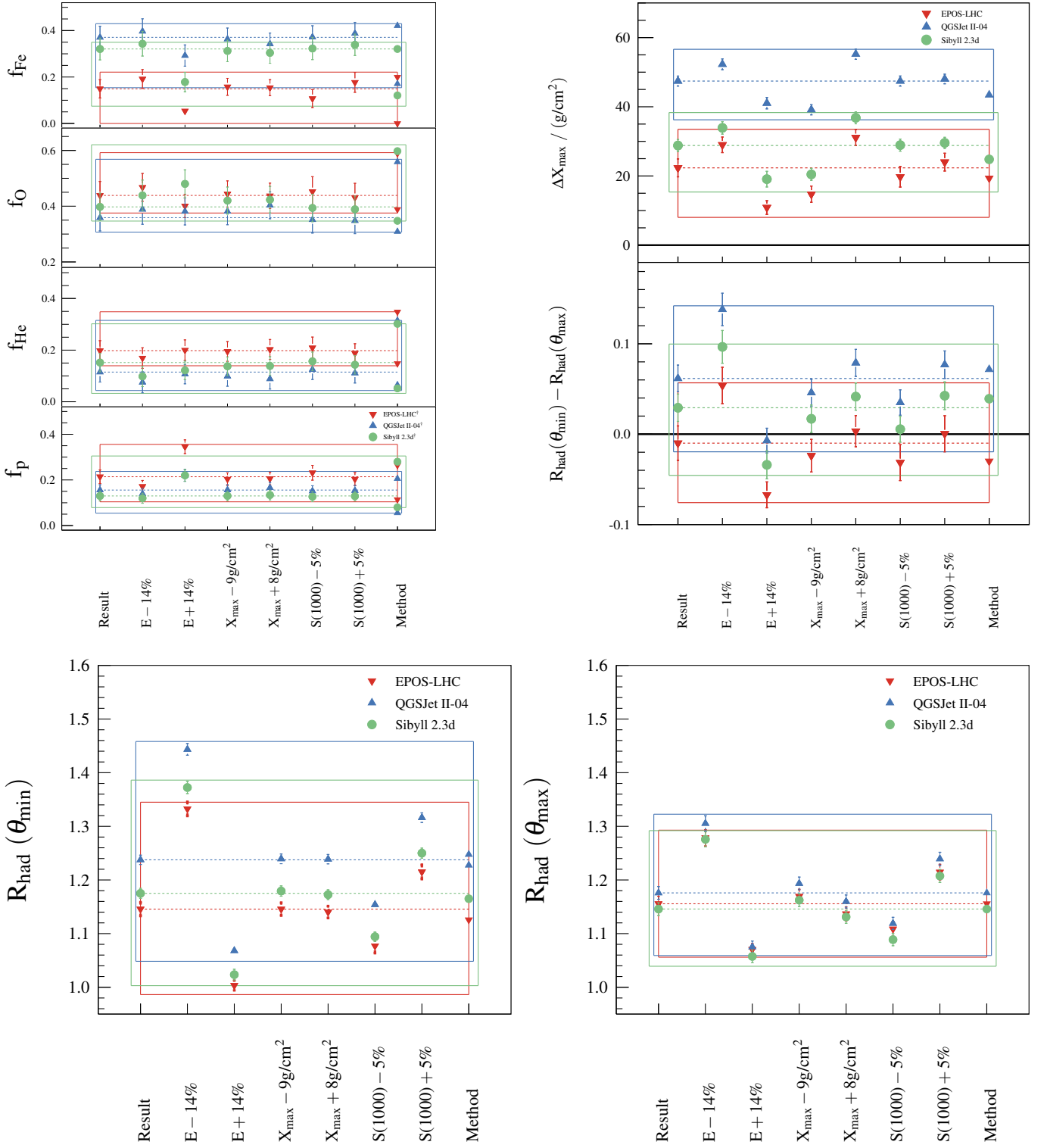


FIG. 16: Contributions of individual experimental systematic uncertainties on energy,  $X_{\text{max}}$ ,  $S(1000)$  and of the method biases inferred from the MC-MC tests (see Appendix C) to the total systematic uncertainties on fractions of primary nuclei (top left),  $\Delta X_{\text{max}}$  (top right),  $R_{\text{had}}(\theta)$  (bottom) and zenith-angle difference of  $R_{\text{had}}$  (top right). Coloured bands are the total systematic uncertainties obtained by summing individual contributions in quadrature with the best-fit results indicated by dashed horizontal lines.

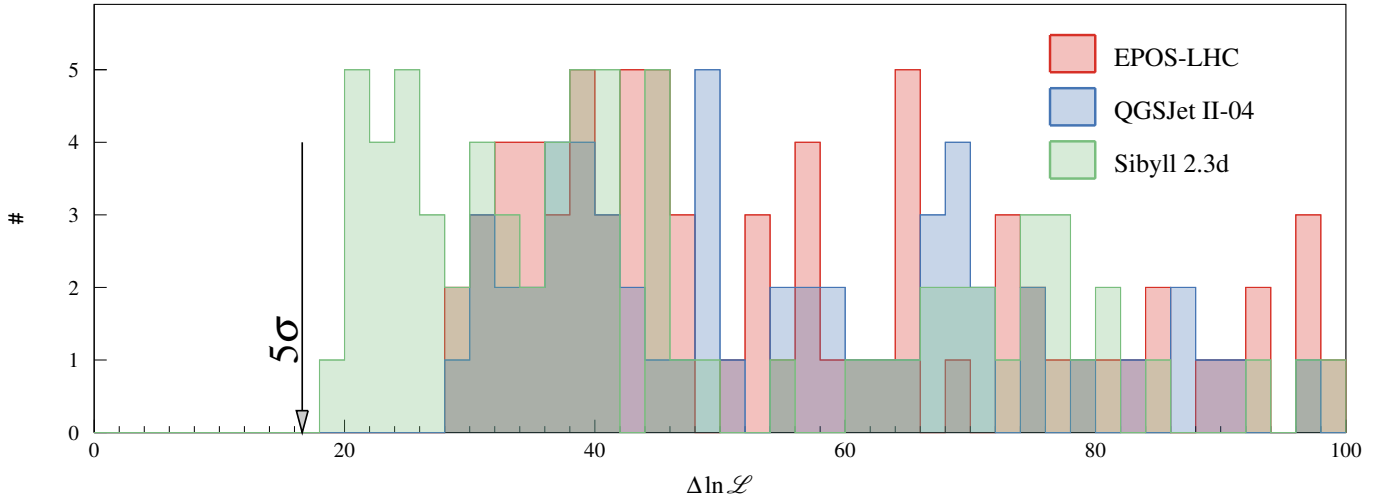


FIG. 17: Number of scans in linear combinations of experimental systematic uncertainties for fits using modifications  $\Delta X_{\max}$  and  $R_{\text{had}}(\theta)$  to estimate the closest approach to the point ( $R_{\text{had}}(\theta) = 1$ ,  $\Delta X_{\max} = 0 \text{ g/cm}^2$ ) from fits with non-modified MC templates in the difference of log-likelihood  $\Delta \ln \mathcal{L}$ . The value estimated using Wilks' theorem in the Likelihood-ratio test for the nested model at the level of  $5\sigma$  ( $\Delta \ln \mathcal{L} \approx 16.62$ ) is indicated by the arrow.