

Testing Hadronic Interactions Using Hybrid Observables

Jakub Vicha*, Alexey Yushkov, Petr Trávníček, Eva dos Santos

FZU, Institute of Physics of the Czech Academy of Sciences, Prague, Czech Republic

E-mail: vicha@fzu.cz

Dalibor Nosek

Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic

We present a new method to infer the mass composition of ultra-high energy cosmic rays performing simultaneous modifications of the simulated depth of shower maximum (X_{\max}) and the ground signal. We introduce a likelihood fit to two-dimensional distributions of "hybrid observables", X_{\max} and the ground signal induced by muons and electromagnetic particles at different zenith angles. With this method, the composition fractions of primary particles are simultaneously estimated together with a shift of X_{\max} and multiplicative rescaling parameters of the muon and electromagnetic signals in the ground detector. The performance of the method is presented using Monte Carlo simulations showing the importance of a global description of the hybrid data from ground and optical detectors resulting in reduced systematic uncertainties stemming from the modelling of hadronic interactions.

*36th International Cosmic Ray Conference -ICRC2019-
July 24th - August 1st, 2019
Madison, WI, U.S.A.*

*Speaker.

1. Introduction

The mass composition of ultra-high energy cosmic rays (UHECR, $> 10^{18}$ eV) can be inferred from measurements of the depth of shower maximum (X_{\max}), which is currently the most reliable observable that is sensitive to the mass composition of UHECR. The interpretation of the X_{\max} distributions measured by fluorescence detectors (FDs) is usually based on the comparison with combinations of simulated X_{\max} distributions of showers induced by different primary particles for a given model of hadronic interactions (HI model).

These HI models extrapolate properties of particle interactions (cross-sections, multiplicities, elasticities etc.) measured at accelerators with lower beam energies (up to $\sqrt{s} = 13$ TeV for proton-proton collisions at LHC) and lower pseudorapidity regions ($|\eta| < \sim 5$) than those covering most of the energy flow of the first interactions of UHECR in the atmosphere (\sqrt{s} above 50 TeV and $\eta \approx 7 - 11$)¹. The differences in the predicted X_{\max} distributions of the HI models tuned to the LHC data (EPOS-LHC [1], QGSJet II-04 [2] and Sibyll 2.3c [3]) remain at the level of ~ 30 g/cm² at 5 EeV for the mean of the X_{\max} distribution ($\langle X_{\max} \rangle$). Extrapolating the uncertainties of accelerator measurements, the lower limit on the uncertainty of $\langle X_{\max} \rangle$ was estimated to be at least ~ 35 g/cm² at the energy $10^{19.5}$ eV [4]. Although this uncertainty decreases with decreasing energy, the current uncertainties of $\langle X_{\max} \rangle$ at the energy 5 EeV can be estimated to be at least about 1/3 of the difference between the two considered astrophysical extremes - protons and iron nuclei. In case of the square root of the variance of the X_{\max} distribution, the HI models predict similar values with differences up to a level of ~ 4 g/cm².

Furthermore, we know that these HI models have problems to describe different aspects of measured data at ultra-high energies well, see e.g. the negative variance of the mean logarithmic mass number for QGSJet II-04 above ~ 2 EeV [5] that was derived from the lowest two central moments of X_{\max} distributions. For EPOS-LHC, the range of predictions for the muon production depth of protons and iron nuclei is not bracketing the measured values above ~ 20 EeV [6]. The incompatibilities between the FD measurements and ground measurements were observed for EPOS-LHC [7] and QGSJet II-04, Sibyll 2.1 [8].

At the Pierre Auger Observatory [9], the largest hybrid cosmic-ray observatory equipped with fluorescence and surface detectors, a deficit of the simulated ground signal in stations at 1000 m from the shower core, $S(1000)$, was observed with respect to the measured $S(1000)$ at the energy $\sim 10^{19}$ eV for zenith angles (θ) within 60° [10]. This deficit was interpreted as a necessity of the rescaling of the hadronic part (consisting mostly of muons) of the ground signal for EPOS-LHC and QGSJet II-04. The rescaling factor was strongly correlated with the energy scale. The electromagnetic (EM) component was assumed to be correct for a given HI model. This assumption included also the position of the maximum of the EM component - X_{\max} and, consequently, the mass composition that was best describing the observed X_{\max} distribution at the energy 10^{19} eV. The size of the ground signal depends, for a given θ and primary energy, on the mass number of the primary nucleus mainly due to the fact that more nucleons give rise to more hadronic interactions and finally, to more muons produced by the decays of charged pions. However, X_{\max} is influenced mainly by the properties of the first hadronic interactions that are far from being known, as dis-

¹The forward calorimeters at LHC can cover $\eta = 8.4 - 15$, but of neutral particles only.

cussed above, or e.g. in [11]. A similar deficit of muons in simulations was observed in inclined showers ($62^\circ \leq \theta \leq 80^\circ$) for EPOS-LHC and QGSJet II-04 at energy $\sim 10^{19}$ eV [12].

In this work, we propose a more complex approach to test the imperfections of HI models (in principle depending strongly on the mass composition of UHECR) using a hybrid detector like the Pierre Auger Observatory. We apply a likelihood fit to two-dimensional distributions of the ground signal and X_{\max} using Monte Carlo (MC) templates. The fit implements a correction for the X_{\max} distribution (single shift ΔX_{\max} for all primaries) and rescaling parameters of the simulated muon (R_μ) and EM signals (R_{em}) on the ground to find the best combination of MC templates of different primaries. The performance of the method is tested using various mixes of four primary particles and different HI models demonstrating low sensitivity of the fitted mass composition to the details of hadronic interactions contrary to the fits of X_{\max} distributions. The promising results show that the method can be used to infer data-driven scales of X_{\max} and ground signal.

2. Simulations and Observables

We use simulated showers of mean energy ~ 5 EeV and zenith angle within 60° that successfully pass the reconstruction and the high-quality requirements of the Pierre Auger Observatory using Auger Offline software [13]. For each primary particle (proton p, nucleus of helium He, oxygen O or iron Fe) and each HI model (EPOS-LHC, QGSJet II-04, Sibyll 2.3c) around 5000 showers were generated with CORSIKA 7.64 [14] and reconstructed ten times (ten random core positions) with the Auger Offline software to create Monte Carlo (MC) templates at the reconstruction level. We adopted standard high-quality FD and SD selection cuts of the Observatory [8]. In our simulation set, we select the showers with reconstructed FD energy $E_{\text{FD}} = 10^{18.5-19.0}$ eV and reconstructed zenith angles up to 60° . In a final step, the selected showers (around 13000 per primary and HI model) were re-weighted in order to describe the energy spectrum measured by the Pierre Auger Observatory [15].

The observable of ground signal, $S^{\text{Ref}}(1000)$, is defined as the signal $S(1000)$ corrected for the energy evolution event-by-event via $S^{\text{Ref}}(1000) = S(1000) \cdot \left(\frac{E_{\text{Ref}}}{E_{\text{FD}}}\right)^{1/B}$ assuming $S(1000) \propto E_{\text{FD}}^{1/B}$ where the parameter $B = 1.042$ is the slope of the energy-calibration curve [15]. The reference energy E_{Ref} was chosen to be $10^{18.7}$ eV which is a value very close to the average FD energy in the energy interval $10^{18.5-19}$ eV.

The observable of the depth of shower maximum is defined as X_{\max}^{Ref} . It is X_{\max} corrected for the energy evolution event-by-event via $X_{\max}^{\text{Ref}} = X_{\max} + D \cdot \log_{10} \left(\frac{E_{\text{Ref}}}{E_{\text{FD}}}\right)$, where $D = \frac{dX_{\max}}{d \log_{10} E} = 58$ g/cm².

3. Method

We apply the binned maximum likelihood method to find simultaneously the best description of four 2D distributions of $S^{\text{Ref}}(1000)$ and X_{\max}^{Ref} with MC templates in four θ bins. For a given HI model, we find the most likely combination of four primary species considering the rescaling of the muon signal R_μ , the rescaling of the EM signal R_{em} , and the shift of the depth of shower maximum ΔX_{\max} in all the MC templates of given HI model.

3.1 MC Templates

The four examples of 2D distributions of $S^{\text{Ref}}(1000)$ and $X_{\text{max}}^{\text{Ref}}$ out of 48 in total (3 HI models \times 4 primaries \times 4 zenith-angle bins) are shown in Fig. 1a. It is worth noting that the larger the distance of X_{max} from the ground (larger θ or heavier primary) the more positive correlation between $S^{\text{Ref}}(1000)$ and $X_{\text{max}}^{\text{Ref}}$ is present. These effects are the consequences of the evolution of the ground signal with mass distance of X_{max} to the ground (DX) (see Fig. 1b).

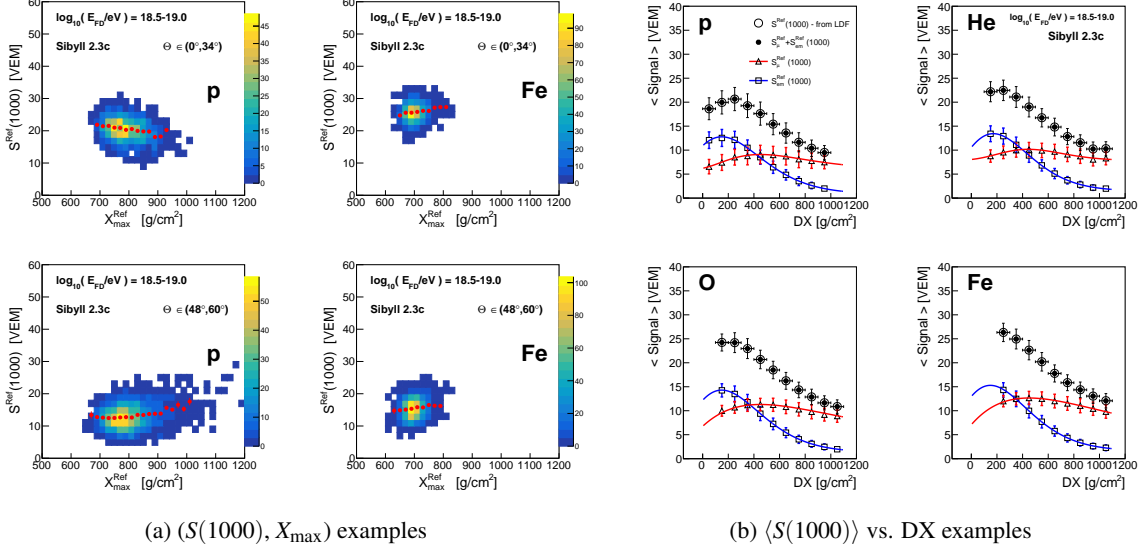


Figure 1: **Four left panels (a):** Four examples of two-dimensional distributions of the ground signal and depth of shower maximum for protons (left) and iron nuclei (right) simulated for HI model Sibyll 2.3c. Top and bottom panels correspond to showers of low and high θ values, respectively. The red markers denote the mean values of the ground signal. **Four right panels (b):** Examples of evolution of the mean ground signal with mass distance of X_{max} to the ground for four primaries in case of Sibyll 2.3c. The part of the signal induced only by muons (red) and EM particles (blue) is depicted as well together with their parametrized curves.

For each HI model, each primary particle and each θ bin, the 2D distribution was normalized to unity and then fitted with the following Ansatz function:

$$\Phi = \frac{dN}{dX_{\text{max}}^{\text{Ref}} \cdot dS^{\text{Ref}}(1000) \cdot N_{\text{MC}}} = A_{\text{Gauss}} \cdot A_{\text{Gumbel}} \cdot f_{\text{Gumbel}}(X_{\text{max}}^{\text{Ref}}) \cdot f_{\text{Gauss}}(X_{\text{max}}^{\text{Ref}}, S^{\text{Ref}}(1000)). \quad (3.1)$$

The two normalization terms $A_{\text{Gauss}} = \frac{1}{\sqrt{2 \cdot \pi \cdot \sigma^2}}$, $A_{\text{Gumbel}} = \frac{\lambda^\lambda}{s \cdot \Gamma(\lambda)}$ where Γ is the gamma function, are followed by the Generalized Gumbel [16] part of the $X_{\text{max}}^{\text{Ref}}$ distribution

$$f_{\text{Gumbel}}(X_{\text{max}}^{\text{Ref}}) = e^{-\left(\frac{X_{\text{max}}^{\text{Ref}} - m}{s} - e^{-\frac{X_{\text{max}}^{\text{Ref}} - m}{s}}\right) \cdot \lambda} \quad (3.2)$$

that is suited to describe the distribution of maxima. The Gaussian part describes the $S^{\text{Ref}}(1000)$ distribution with the mean value being linearly dependent on $X_{\text{max}}^{\text{Ref}}$

$$f_{\text{Gauss}}(X_{\text{max}}^{\text{Ref}}, S^{\text{Ref}}(1000)) = e^{-\frac{(S^{\text{Ref}}(1000) - a \cdot X_{\text{max}}^{\text{Ref}} - b)^2}{2 \cdot \sigma^2}}. \quad (3.3)$$

The six fitted parameters for each MC template are m , s , λ of the Generalized Gumbel part, and σ , a , b of the Gaussian part of Φ .

3.2 Fitting Procedure

The fitting method is a generalization of the fitting procedure used in [17] to derive primary fractions comparing simulated and measured X_{\max} distributions. The minus log-likelihood expression that is in our analysis minimised for each HI model is of the form

$$\mathcal{L} = \sum_z \sum_j \left(C_{jz} - n_{jz} + n_{jz} \cdot \ln \frac{n_{jz}}{C_{jz}} \right) \quad \text{for } n_{jz} > 0, \quad \mathcal{L} = \sum_z \sum_j C_{jz} \quad \text{for } n_{jz} = 0 \quad (3.4)$$

with the sums running over all the 2D bins j and all four zenith-angle bins z . For the zenith-angle bin z and the 2D bin j , the number of showers selected in the studied set is denoted by n_{jz} and the predicted number of MC showers by C_{jz} . The latter number is $C_{jz} = N_{\text{data}}^z \sum_i f_i \cdot \Phi_{i,z}(\widehat{X_{\max}^{\text{Ref}}(j)}, S^{\text{Ref}}(1000)^{(j,z)})$, where N_{data}^z is the number of showers in the studied MC sample in the zenith-angle bin z , $\Phi_{i,z}$ denotes the MC template Φ from Eq. (3.1) for given HI model, the primary i and the zenith-angle bin z , and $\widehat{X_{\max}^{\text{Ref}}(j)} = X_{\max}^{\text{Ref}(j)} - \Delta X_{\max} - D \cdot \log_{10} R_E$, and $S^{\text{Ref}}(1000)^{(j,z)} = S(1000)^{\text{Ref}(j)} / f_{\text{SD}}^z(R_\mu, R_{\text{em}}, R_E) / f_{\Delta X}^{(j,z)}(\Delta X_{\max})$. $X_{\max}^{\text{Ref}(j)}$ and $S(1000)^{\text{Ref}(j)}$ denote the center values of X_{\max}^{Ref} and $S^{\text{Ref}}(1000)$ for the 2D bin j of the studied distribution in the zenith-angle bin z , respectively.² The rescaling parameter f_{SD}^z of all signals $S(1000)^{\text{Ref}(j)}$ in a given zenith angle bin z is derived from the assumed rescaling parameter f_{SD} of the ground signal $S^{\text{Ref}}(1000)$ of single shower using the possible rescaling parameters R_μ , R_{em} and R_E of the muon signal S_μ^{Ref} , of the EM signal $S_{\text{em}}^{\text{Ref}}$, both at E_{Ref} ($S_\mu \propto E_{\text{FD}}^\beta$, $S_{\text{em}} \propto E_{\text{FD}}$), and of the energy scale, respectively, as $S(1000)(\theta) \cdot \left(\frac{E_{\text{Ref}}}{E_{\text{FD}}} \right)^{1/B} \cdot f_{\text{SD}}(\theta) = R_\mu \cdot S_\mu(\theta) \cdot \left(\frac{R_E \cdot E_{\text{Ref}}}{E_{\text{FD}}} \right)^\beta + R_{\text{em}} \cdot S_{\text{em}}(\theta) \cdot \left(\frac{R_E \cdot E_{\text{Ref}}}{E_{\text{FD}}} \right)$. The signal $S(1000)$ was assumed to be composed only of S_μ and S_{em} , which is a fair approximation at the distance 1000 m from the shower core. Then, using the definition of muon fraction of the ground signal as $f_\mu = S_\mu / S(1000)$ and considering the average value of f_{SD} over showers in a given bin z :

$$f_{\text{SD}}^z(R_\mu, R_{\text{em}}, R_E) = R_\mu \cdot R_E^\beta \cdot \frac{(E_{\text{Ref}})^{\beta-1/B}}{\langle E_{\text{FD}}^{\beta-1/B} \rangle_z} \cdot f_\mu^z + R_E \cdot \frac{(E_{\text{Ref}})^{1-1/B}}{\langle E_{\text{FD}}^{1-1/B} \rangle_z} \cdot R_{\text{em}} \cdot (1 - f_\mu^z) \quad (3.5)$$

where $\langle E_{\text{FD}}^{\beta-1/B} \rangle_z$ and $\langle E_{\text{FD}}^{1-1/B} \rangle_z$ are calculated from all selected showers in the zenith-angle bin z . The mean muon fraction in the zenith-angle bin z (f_μ^z) is calculated as the average value over average muon fractions $f_\mu^{z,i}$ for simulated showers induced by a primary i and weighted over the relative primary fractions as $f_\mu^z = \sum_i f_i \cdot f_\mu^{z,i}$.

The modification of the ground signal due to the change of X_{\max} by ΔX_{\max} is estimated by $f_{\Delta X}^{(j,z)}(\Delta X_{\max}) = f_{\Delta X}(\Delta X_{\max}, DX^{(j,z)})$, where $DX^{(j,z)} = X_{\text{vert}} / \cos(\langle \theta \rangle_z) - X_{\max}^{\text{Ref}(j)}$ for average zenith angle $\langle \theta \rangle_z$ of the bin z , $X_{\text{vert}} = 880 \text{ g/cm}^2$. The size of this effect is calculated from parametrizations of EM and muon signals shown as examples in right panels of Fig. 1 and it is about 5% per ΔX_{\max} of 40 g/cm^2 for $\theta > \sim 34^\circ$ and about half of it for $\theta < \sim 34^\circ$.

²We implement here also the influence of the rescaling of the energy scale (R_E) on X_{\max}^{Ref} for completeness, although this effect should be small ($\sim 3.3 \text{ g/cm}^2$ for a shift of the energy scale by the systematic uncertainty claimed by the Pierre Auger Observatory 14% [9]).

The rescaling parameters R_{em} and R_E are almost indistinguishable within our method (see Eq. (3.5)). Therefore, in the following tests, we decided to fix the parameter R_E to 1 and release only the parameter R_{em} and use only simulations from the same energy range that was used for the parametrizations of the MC templates. However, applying this method to the measured data, the precision of R_{em} would be limited by the systematic uncertainty on the energy scale.

Due to the fact that $\sum_i f_i = 1$, we apply the transformation according to [18] from f_i where $i = p, He, O, Fe$ to ξ_k where $k = 1, 2, 3$ to obtain only three free parameters related to the mass composition. Within this approach, six parameters in total ($\xi_1, \xi_2, \xi_3, R_\mu, R_{em}$ and ΔX_{max}) are derived in the following to find the best description of the four 2D distributions modifying the MC templates of a given HI model.

4. Method Performance

We studied the performance of the presented method on sets of showers simulated by the HI model Sibyll 2.3c and adopting MC templates of all the HI models. In total, we produced 286 combinations of the four primaries in steps of 10% covering all the possible combinations. Each combination contained 2500 showers. To compare with the standard X_{max} fit approach, the size of the studied sample was doubled.

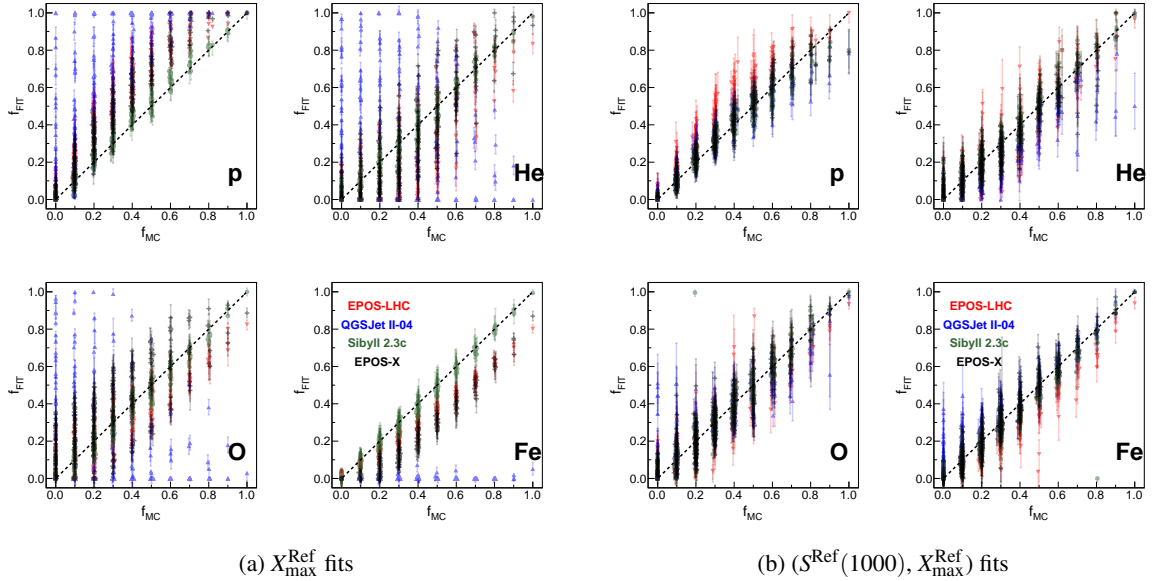


Figure 2: Fitted (f_{FIT}) vs. true (f_{MC}) primary fractions in the case of X_{max}^{Ref} distribution fits (**four left panels (a)**) and in the case of our method using 2D fits (**four right panels (b)**).

4.1 Mass Composition

The panels of Fig. 2a illustrate the precision of the fitted primary fractions when only the distribution of X_{max}^{Ref} was fitted with X_{max}^{Ref} templates for the four primaries. The panels of Fig. 2b were obtained with the presented method fitting the 2D distributions in four zenith-angle bins.

Generally, our method provides much less biased primary fractions wrt. simple fits of X_{\max}^{Ref} . The bias in the case of X_{\max}^{Ref} fits increases with the difference of $\langle X_{\max} \rangle$ between Sibyll 2.3c and the HI model used for MC templates ($\sim -6 \text{ g/cm}^2$ for EPOS-LHC and $\sim -26 \text{ g/cm}^2$ for QGSjet II-04) in consequence of a lighter mass composition to be fitted wrt. true mass composition. In our method, the largest bias, which is at the level of $\sim 8\%$, comes from the different widths of X_{\max} distributions between the HI models for all primaries, that is about 4 g/cm^2 between EPOS-LHC and Sibyll 2.3c. To illustrate this effect, we additionally smeared the X_{\max} distributions of EPOS-LHC to make it $\sim 4 \text{ g/cm}^2$ wider for all primaries, denoted as EPOS-X. In such a case, the shapes of X_{\max} distributions for the four primaries are fairly universal wrt. HI models minimising the bias of the fitted primary fractions.

4.2 MC Correction Factors

Fig. 3 illustrates the fitted values of MC corrections for all the 286 combinations of the four primaries compared with the true differences between the HI models (lines). The $\sim 10 \text{ g/cm}^2$ bias of ΔX_{\max} in case of EPOS-LHC is minimised removing the difference in the width of X_{\max} distribution (EPOS-X). The $\sim 5\text{-}6\%$ bias of the rescaling parameters R_{μ} and R_{em} however persists even in the case of EPOS-X and it is probably a consequence of different values of R_{μ} and R_{em} for different zenith angles that are in the method assumed to be independent on the zenith angle.

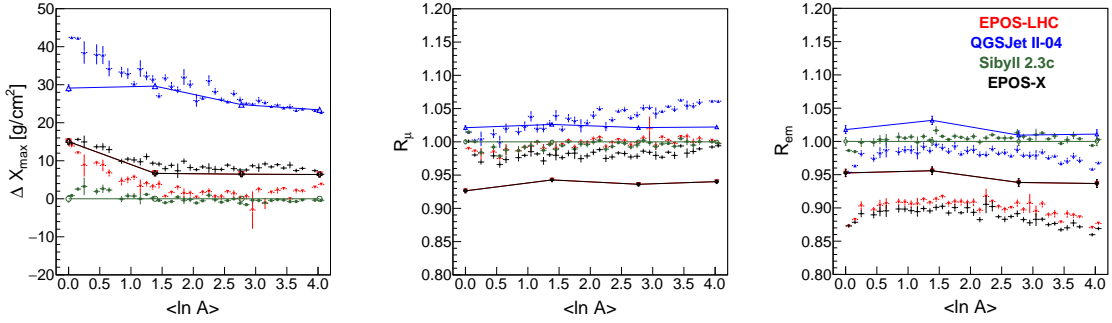


Figure 3: Corrections to MC predictions on X_{\max} (**left**), muon (**middle**) and EM (**right**) signal on the ground. The lines indicate the true differences in corrections for the HI models from the Sibyll 2.3c.

5. Towards Less Model-Dependent Mass Composition

The presented method indicates the way towards the less model-dependent estimation of the mass composition of cosmic rays at ultra-high energies using a hybrid observatory. Considering simple MC corrections of X_{\max} and ground signal, the mass composition is driven by the shape of X_{\max} distributions and the zenith-angle dependence of the correlation between the ground signal and X_{\max} . The main differences between the predictions of different HI models on primary fractions are minimised by fitting the absolute scales of X_{\max} and ground signal. On the other hand, the method can also test the HI models revealing the differences in absolute scales of X_{\max} and ground signal.

Acknowledgements

This work is funded by Ministry of Education, Youth and Sports of the Czech Republic under the projects LTT18004, LM2015038 and CZ.02.1.01/0.0/0.0/16_013/0001402. The authors are very grateful to the Pierre Auger Collaboration for providing the simulations for this contribution.

References

- [1] T. Pierog *et al.*, *EPOS LHC: Test of collective hadronization with data measured at the CERN Large Hadron Collider*, Phys. Rev. C **92** (2015) 034906.
- [2] S. Ostapchenko, *Monte Carlo treatment of hadronic interactions in enhanced Pomeron scheme: QGSJET-II model*, PRD **83** (2011) 014018.
- [3] F. Riehn *et al.*, *A new version of the event generator Sibyll*, in proceedings of the ICRC 2015, PoS(ICRC2015)558.
- [4] R. U. Abbasi and G. B. Thomson, *$\langle X_{max} \rangle$ Uncertainty from Extrapolation of Cosmic Ray Air Shower Parameters*, in proceedings of UHECR 2016, arXiv:1605.05241 [hep-ex].
- [5] A. Aab *et al.*, *Depth of maximum of air-shower profiles at the Pierre Auger Observatory. I. Measurements at energies above $10^{17.8}$ eV*, PRD **90** (2014) 122005.
- [6] A. Aab *et al.*, *Erratum: Muons in air showers at the Pierre Auger Observatory: Measurement of atmospheric production depth*, PRD **90** (2014) 012012, Erratum: PRD **92** (2015) 019903.
- [7] A. Aab *et al.*, *Inferences on mass composition and tests of hadronic interactions from 0.3 to 100 EeV using the water-Cherenkov detectors of the Pierre Auger Observatory*, PRD **96** (2017) 122003.
- [8] A. Aab *et al.*, *Evidence for a mixed mass composition at the "ankle" in the cosmic-ray spectrum*, Phys. Lett. B **762** (2016) 288.
- [9] A. Aab *et al.*, *The Pierre Auger Cosmic Ray Observatory*, NIM A **798** (2015) 172.
- [10] A. Aab *et al.*, *Testing Hadronic Interactions at Ultrahigh Energies with Air Showers Measured by the Pierre Auger Observatory*, Phys. Rev. Lett. **117** (2016) 192001.
- [11] S. Ostapchenko, *LHC results and hadronic interaction models*, in proceedings of XXV ECRS 2016, arXiv:1612.09461 [astro-ph.HE].
- [12] A. Aab *et al.*, *Muons in air showers at the Pierre Auger Observatory: Mean number in highly inclined events*, Phys. Rev. D **91** (2015) 032003.
- [13] S. Argiro *et al.*, *The Offline Software Framework of the Pierre Auger Observatory*, NIM A **580** (2007) 1485.
- [14] D. Heck *et al.*, *Upgrade of the Monte Carlo code CORSIKA to simulate extensive air showers with energies $> 10^{20}$ eV*, Report FZKA Forschungszentrum Karlsruhe **6019** (1998).
- [15] F. Fenu *for the Pierre Auger Collaboration*, *The cosmic ray energy spectrum measured using the Pierre Auger Observatory*, in proceedings of the ICRC 2017, PoS(ICRC2017)486.
- [16] M. De Domenico *et al.*, *Reinterpreting the development of extensive air showers initiated by nuclei and photons*, JCAP **07** (2013) 050.
- [17] A. Aab *et al.*, *Depth of maximum of air-shower profiles at the Pierre Auger Observatory. II. Composition implications*, Phys. Rev. D **90** (2014) 122006.
- [18] W. J. Metzger, *Statistical methods in data analysis*, Nijmegen Univ. Fys. Lab. 2002.