

Study on the Combined Estimate of the Cosmic-Ray Composition and Particle Cross Sections at Ultrahigh Energies

O. Tkachenko,* R. Engel, R. Ulrich and M. Unger

Institute for Astroparticle Physics, Karlsruhe Institute of Technology, Germany

E-mail: olena.tkachenko@kit.edu

The mass composition is one of the key observables to understand the nature and origin of ultra-high energy cosmic rays. The study of hadronic interactions at energies well beyond human-made accelerators is a fundamental probe of elementary particle physics. In previous analyses, the properties of the hadronic interactions were estimated under the assumption of a certain mass composition, typically proton-dominated, and the cross sections were calculated by fitting the tail of the X_{\max} distribution. In such an analysis, the impact of a possible He-contamination on the cross section measurement is quoted as a systematic uncertainty. Vice versa, the cosmic-ray mass composition is typically determined using air shower simulations by assuming the validity of the considered hadronic interaction models.

In this contribution, we present a fully self-consistent approach of varying the proton-proton cross sections, with the nucleus-nucleus cross sections being predicted via the Glauber theory, and making a full X_{\max} distribution fit to get an independent and simultaneous estimation of the interaction cross sections and cosmic-ray primary composition. We will discuss the degeneracy between mass composition and hadronic interactions and compare the sensitivity of the proposed method to one of the previous approaches.

*37th International Cosmic Ray Conference (ICRC 2021)
July 12th – 23rd, 2021
Online – Berlin, Germany*

*Presenter

1. Introduction

The existence of cosmic rays with energies beyond 10^{20} eV has been known for more than half a century, but the understanding of their astrophysical origin remains one of the biggest unsolved problems of modern astrophysics. Studying the mass composition and hadronic interaction properties helps to improve our knowledge about the nature of ultra-high energy cosmic rays. However, this understanding is complicated by a strong dependence of extensive air shower (EAS) variables on the underlying hadronic interactions. Since both, mass composition and hadronic interaction properties, are not yet well determined at ultrahigh energies, the inference of either of these quantities usually requires certain assumptions about the other.

In this work we present and discuss the combined approach allowing the independent and simultaneous estimation of both mass composition and interaction cross sections without underlying assumptions present in the separate analyses. The main essence of the proposed method is to perform a standard composition fit using the X_{\max} distributions in which the nucleus-air cross sections (or more precisely, the fundamental hadron-nucleon cross sections) are allowed to vary leading to best-fit estimates of the combination of the composition fractions and interaction cross sections.

2. Standard mass composition and cross section measurements

One of the most robust techniques for estimation of the mass composition is based on the depth of atmospheric shower maximum, X_{\max} , as can be measured by fluorescence detectors. Comparing the measured X_{\max} distributions to the predictions obtained from air shower simulations with varying nuclear fractions, the best-fit composition can be estimated, see e.g. [1, 2]. Given the sensitivity of the longitudinal air shower development to the hadronic interaction properties, a model for hadronic interactions is needed which extrapolates the existing accelerator. The hadronic interaction properties at ultra-high energies are not well known and different models predict different characteristics of hadronic interactions. For instance the interaction cross sections and particle multiplicities can differ substantially depending on the adopted model and therefore these theoretical uncertainties dominate the uncertainty on the mass composition of cosmic rays at ultra-high energies.

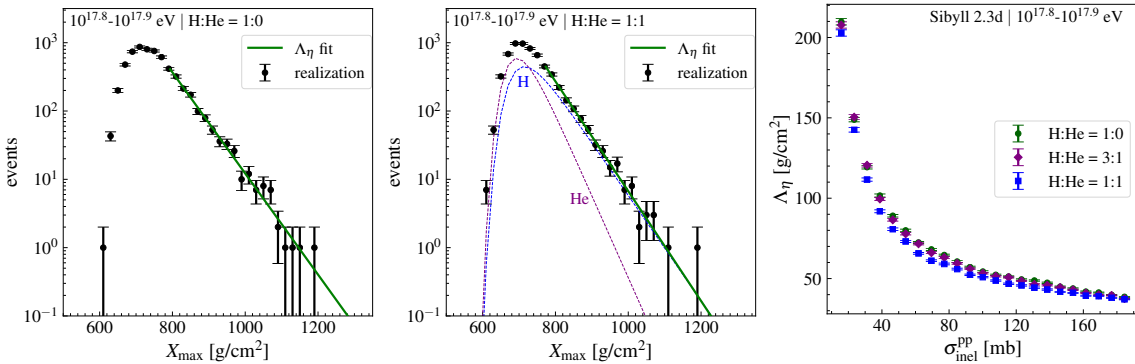


Figure 1: Examples of the exponential fit to the tail of the X_{\max} distribution with fully proton-dominated composition (left) and with 50% He contribution (center), and the conversion from the attenuation length Λ_η into the proton-proton inelastic cross section (right).

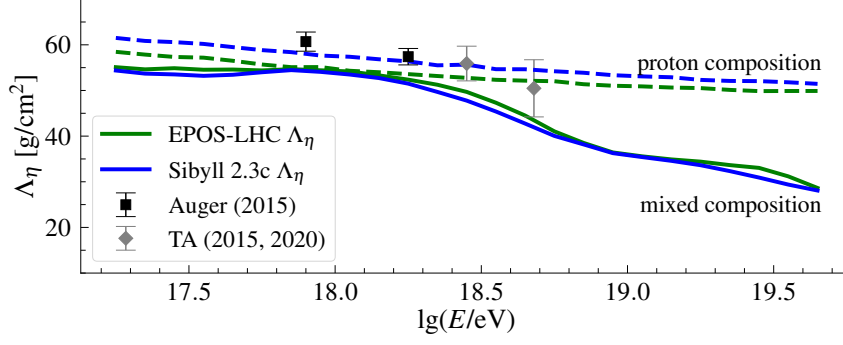


Figure 2: Measured attenuation lengths (data points) and expectations for a pure proton composition (dashed lines) and a mixed composition (solid lines). Different colors denote two hadronic interaction models: EPOS-LHC (green) and Sibyll 2.3c (blue). Simulations and Auger data describe the exponential tail of the 20% deepest air showers ($\Lambda_{\eta=0.2}$). See text for a discussion of the TA data points.

The standard techniques to estimate the interaction cross sections are based on the measurements of the attenuation length of cosmic-ray particles in the atmosphere Λ_{η} by fitting the tail of the X_{\max} distribution with the exponential profile, where η corresponds to the fraction of deep- X_{\max} events used for the fit. In the analysis of the Pierre Auger Collaboration, η is set to 0.2 to achieve a bias originating from a possible helium contamination of up to 25% that is smaller than the statistical uncertainty of the Λ_{η} measurement [3, 4]. Another analysis approach is followed by the Telescope Array (TA) Collaboration, which determined the exponential slope of the X_{\max} distribution in a fixed depth range [5, 6]. In each case, the measured attenuation is converted into an estimate of the proton-air cross section using a look-up table $\sigma_{p+\text{air}} = f(\Lambda)$ obtained from simulations of the X_{\max} distribution of proton primaries. Both analyses therefore implicitly assume that the tail of the X_{\max} distribution is dominated by protons. The impact of a possible He-contamination is included in the systematic uncertainty of the result.

Two examples of an unbinned likelihood Λ_{η} fit to data realizations with different percentages of He in the proton-dominated X_{\max} distributions are shown in the left and middle panels of Fig. 1. These were obtained by simulating air showers with the CONEX [7] program using the Sibyll 2.3d [8] hadronic interactions model. The relation between the attenuation length and proton-proton inelastic cross section is shown in the right panel of Fig. 1. Since here and in what follows we assume an ideal detector response, the detector resolution and acceptance effects are not considered.

In Fig. 2 we compare the expected values for $\Lambda_{\eta=0.2}$ to measurements of the attenuation length from Auger [4] and TA [5, 6]¹. The simulations labeled as “mixed composition” are based on the composition fractions derived by the Pierre Auger Collaboration in [9], smoothed with a Gaussian kernel with a bandwidth of half a decade [10]. The dashed lines labeled as “proton composition” are for a pure beam of cosmic protons. For both cases we show the results derived for air shower simulations using the EPOS-LHC [11] and Sibyll 2.3c [8] hadronic interaction models.

¹Note that TA reports the exponential slope fitted within an X_{\max} -range of 790 to 1000 g/cm². We have verified with simulations that the difference to $\Lambda_{\eta=0.2}$ at the energies of the two TA data points is ≤ 2 g/cm². The analysis in [6] was performed at full efficiency, but the highest energy TA data point from [5] is possibly affected by detector effects. We estimate the effect of the telescope acceptance, in particular the loss of events below ground-level, assuming that the TA acceptance is similar to the one of Auger published in [12], as the both observatories are located at a similar altitude and thus ground-level depth. This results in an estimate of the TA acceptance bias of ≤ 3 g/cm².

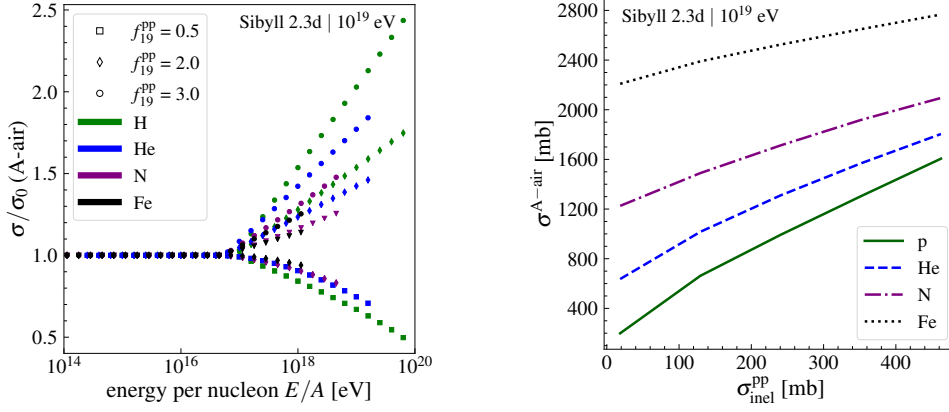


Figure 3: The effect of the scaling factor on the nucleus-air interaction cross sections at 10^{19} eV. *Left:* The normalized by the non-modified values σ_0 nucleus-air cross sections as a function of energy per nucleon E/A . *Right:* The relation between the nucleus-air and proton-proton interaction cross sections.

As can be seen, the difference between Λ_η derived from the X_{\max} distribution of a mixed composition and pure proton composition is small in the energy range probed by Auger. However at the somewhat higher energies probed by TA, the two scenarios diverge considerably. It is therefore unexpected to find the TA data points close to the attenuation values for proton-initiated air showers [13], but given the large statistical uncertainties of these measurements, these data points are also in reasonable agreement with the predictions for a mixed composition.

3. Modified hadronic interactions

To be able to alter hadronic particle production in air shower simulations to affect the EAS properties in a combined fit of composition and cross section, a self-consistent scheme of calculating nucleus-nucleus interaction cross sections is needed. Here we directly modify the proton-proton interactions to obtain self-consistent rescaling of any nucleus-nucleus cross sections via Glauber [15] theory given the input proton-proton interaction cross section. Technically this is achieved by a modified implementation of Sibyll within the CONEX air shower simulation program in which the original proton-proton cross section can be rescaled by an energy-dependent factor similar to the approach introduced in [14],

$$\sigma^{\text{mod,pp}} = \sigma^{\text{orig,pp}} f(E), \quad (1)$$

with a linear scaling factor $f(E)$ between the energy threshold E_0 and energy of interest E

$$f(E) = 1 + H(E - E_0)(f_{lg E_1} - 1) \frac{\lg(E/E_0)}{\lg(E_1/E_0)}, \quad (2)$$

where $f_{lg E_1}$ is the rescaling factor at $E = E_1$ and $H(x)$ denotes the Heaviside step function. The energy threshold of $E_0 \approx 10^{16.95}$ eV is set by the laboratory energy equivalent to the LHC center of mass energy, below which interaction models are constructed from the ground-based accelerator measurements.

The effect of the proton-proton rescaling on the nucleus-air interaction cross sections and the relation between the nucleus-air and proton-proton cross sections are shown in Fig. 3, where on the

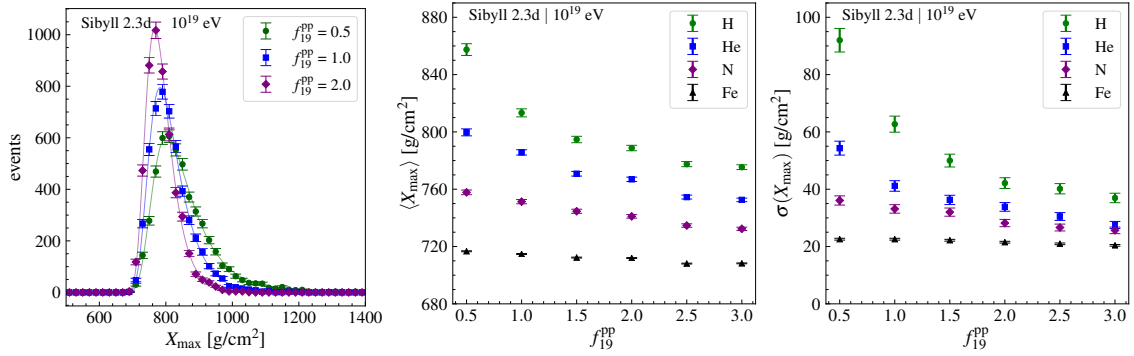


Figure 4: X_{\max} distributions for the modified cross sections: proton X_{\max} distribution for three different scaling factors (left), mean (center) and standard deviation, (right) of the X_{\max} distribution (simulations at 10^{19} eV).

left we plot the normalized cross sections as a function of energy per nucleon to bring the start of the cross section modifications for different elements at one point since the heavier the element the later in energy it starts to be affected by the introduced modifications.

In Fig.4 the effect of the cross section modifications on the X_{\max} distributions is shown, in which we are most interested as it is the observable widely used for the mass composition estimation. As expected, the increase in the scaling factor (and thus cross section) causes the decrease in the moments of the X_{\max} distribution, i.e. the mean, $\langle X_{\max} \rangle$, and standard deviation, $\sigma(X_{\max})$. The distribution becomes narrower and shifts towards smaller X_{\max} values. The heavier the element, the less it is affected.

To avoid random fluctuations originating from finite statistics in the simulations, we use the generalized Gumbel distribution [16] for the description of the X_{\max} distributions. The functional dependencies of the generalized Gumbel distribution shape parameters on the scaling factor are obtained by fitting it to the simulated X_{\max} distributions for a range of input proton-proton cross sections and interpolate the results. In Fig.4 (right) the generalized Gumbel distributions for three different scaling factor values are superimposed on top of the corresponding X_{\max} distributions from air shower simulations. With this setup in place, we can quickly generate the X_{\max} distributions with large statistics, since we describe the distribution with a continuous function, for the arbitrary scaling factor value and use them e.g. to generate many realizations of simulated data sets.

4. Combined mass composition and cross section fit

To estimate the best-fit mass composition and proton-proton interaction cross sections we perform a 4-component binned maximum likelihood fit with the python-based *iminuit* [17] interface for the MINUIT2 C++ library. We perform a discrete scan over a range of scaling factors $f(E)$ and fit the best composition fractions for each scaling factor. This leads to an estimate of the best-fit combination of nuclear species fractions and the modified proton-proton interaction cross sections. To measure the goodness of fit we use the fact that the deviance defined by the logarithm of the Poissonian likelihood adopted in the maximum-likelihood fit is approximately χ^2 -distributed [18]. Since we have discrete points in the scan, the best-fit scaling factor value corresponding to the

minimum χ^2 is found by approximating the region around the smallest obtained χ^2 with a parabolic function. As statistical uncertainty limits we then can take values corresponding to the χ^2+1 .

The performance of the combined fit was tested on simulated data sets with the modified proton-proton cross sections and different mass compositions. The underlying hadronic interaction model used here is Sibyll 2.3d. Since we expect the ratio between H and He to have the defining effect on the fit results, we focus mainly on the two-component composition with varied He contamination, in particular we looked at the compositions with 0%, 5%, 25%, 40%, 50%, 60%, 75%, 95% and 100% of He. The total number of H and He events is always 6000, similar to the number of events available to Auger in the energy bins around at 10^{18} eV. The data as well as the X_{\max} distribution templates were generated using the parameterization of the generalized Gumbel distribution as a function of scaling factor and a binning of 1 g/cm^2 was used. The fits were performed to data sets simulated for three different scaling factor values of 0.75, 1.0 and 1.4. The selected energy range is $10^{17.8}-10^{17.9}$ and it can be noted that given such low energies, the variations in the proton-proton interaction cross sections do not affect the EAS variables for heavier elements, so the X_{\max} distributions for N and Fe remain unchanged regardless of the scaling factor value. For each composition and cross section scenario we generated 100 data realizations and we evaluate the fit performance with the average obtained results.

An example of one realization of a 4-component mass composition fit with varied hadronic interactions applied to the proton-dominated mix of H and He is shown in Fig.5. The four upper panels display the estimated fractions as a function of the scaling factor and in the bottom panel the χ^2 goodness of fit is shown. In the case when the mass composition is mostly dominated by one element, the χ^2 as a function of the scaling factor has a parabolic shape close to the minimum with a prominent minimum, and correspondingly the best-fit interaction cross sections are well-defined. However, if the composition is more mixed, then the fit can have two minima or a very flat χ^2 (see Fig.6). While the latter

case does not affect the results on average (but just causes a poor resolution of f), the former case of having two minima, may result in incorrect cross section and mass composition values. An extreme example of degeneracy is shown in Fig.6 (center), where in the second minimum the “true” p-He mix can be well described by a pure helium composition with a smaller cross section value. The X_{\max} distribution fits corresponding to the estimated from these two minima values describe the data equally well and are almost indistinguishable (7). The only possibility to distinguish these from minima is on grounds of theoretical prejudice in that the proton-proton cross section does not drop to a quarter of its value at LHC energies within a decade in beam energy (factor of 3 in center of mass energy). In the following, we will present result in which the second minimum was rejected if it was found to be very close to boundaries of the considered scaling factor range.

In Fig.8 (left) we compare the estimates of the inelastic proton-proton cross sections obtained by fitting the tail of the X_{\max} distribution with the ones determined with the combined analysis discussed

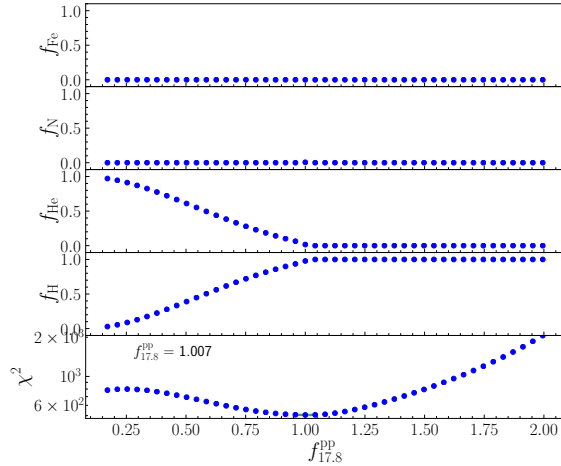


Figure 5: Combined mass-composition and cross section fit to the data set simulated with a ratio between composition fractions of H:He: 0.95:0.05.

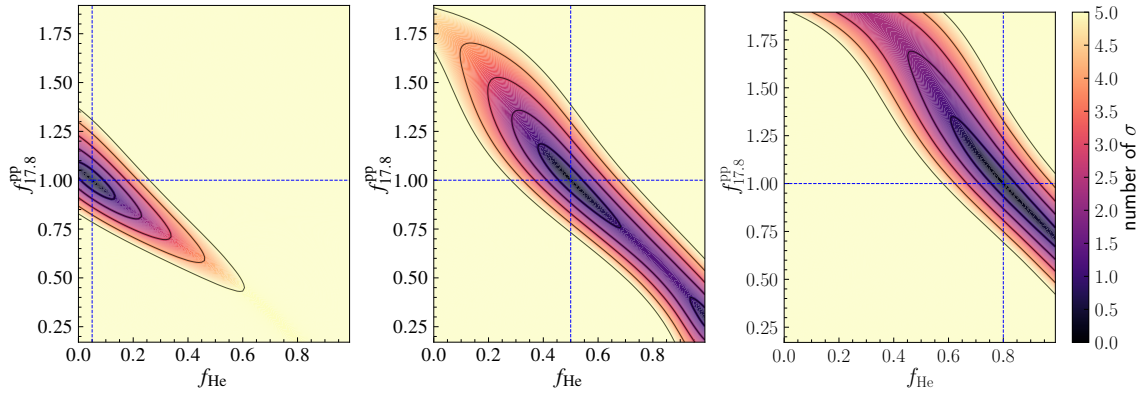


Figure 6: The difference between the X_{\max} distribution with fixed and with a scanned over the whole considered range parameters for three different composition realizations: proton-dominated (left), equal H-He mix, and He-dominated with 20% of H. The color bar shows how much the χ^2 deviates in units of sigma. Blue dashed line denotes the true simulated parameters.

in the previous section. The points are the mean of 100 realizations with 6000 events each and the bands denote the standard deviation of the results. As can be seen, while for the Λ_{η} method the relative bias $(\sigma_{\text{est}}^{\text{PP}} - \sigma_{\text{sim}}^{\text{PP}})/\sigma_{\text{sim}}^{\text{PP}}$ increases gradually as the He contribution grows depending on the simulated cross section, the combined mass composition and cross section fit recovers the simulated values quite well for any He contribution and does not depend much on the input scaling factor. The 1-sigma standard deviation of the combined fit the interval becomes broader as He fraction increases and the largest spread in the individual fit results is present when there is about three times more He than H in the data. As expected this is also where χ^2 is very flat over a broad range in scaling factors. Regarding the mass composition estimation, the fitted He fractions are compared to the simulated ones, and as can be seen in Fig.8 (right), the difference is indeed very small, although the fitted He fractions are slightly biased towards smaller values for non-zero He contribution. First studies with the inclusion of other nuclear species than p and He in the simulated data show that the presence of heavier nuclei leads to a broader dispersion in the fitted cross section values and a slight increase in the corresponding bias, but as long as H and He remain the main components, the results are qualitatively similar to those displayed in Fig. 7.

5. Summary

We have presented a method for the combined mass composition and cross section analysis. Such a combined approach is possible if the proton-proton cross section is used as a fit parameter to self-consistently predict the corresponding nucleus-nucleus cross sections and air shower properties.

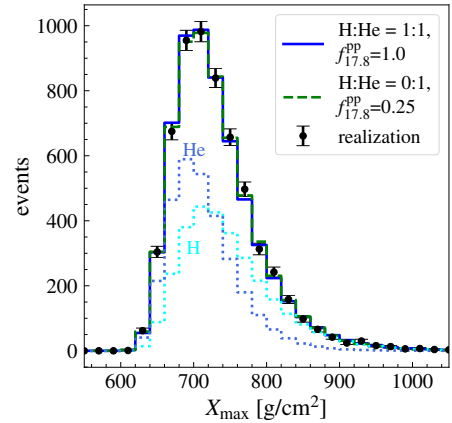


Figure 7: X_{\max} distributions corresponding to two minima in the fit: H:He=1:1 with $f_{17.8}^{\text{PP}}=1.0$ and H:He=0:1 with $f_{17.8}^{\text{PP}} \approx 0.25$. In addition, the separate H and He contributions into the best-fit with a mixed composition are shown.

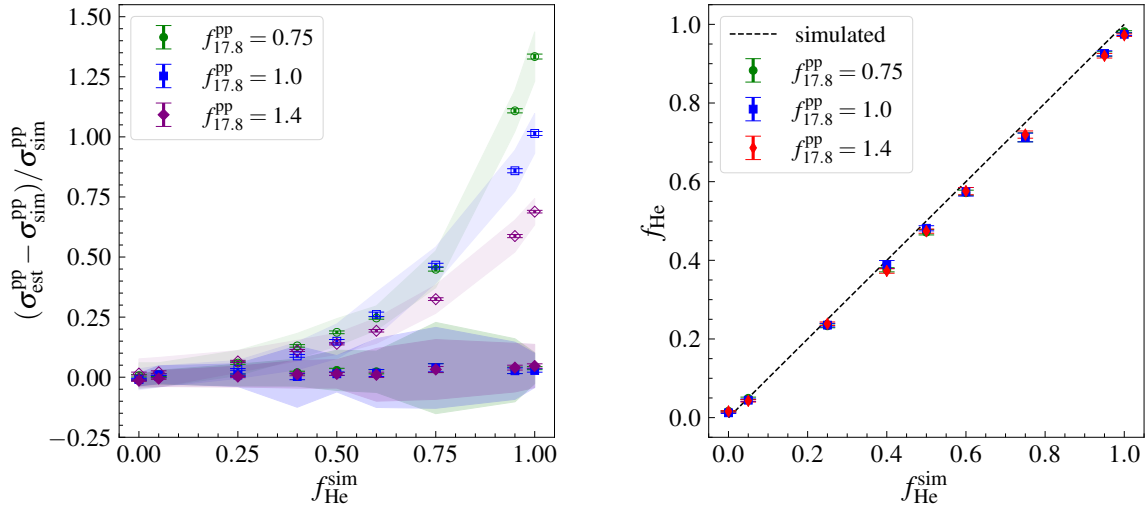


Figure 8: Combined cross section and mass composition fit results. *Left:* Comparison of the inelastic proton-proton cross sections estimated with combined fit (filled symbols) and with standard Λ_n approach (open symbols). *Right:* Comparison of the simulated and reconstructed He fractions.

The combined fit was applied to the simulated data with varied H-He ratio and proton-proton interaction cross sections. By comparing the obtained results to the cross sections estimated from the attenuation length fit to the X_{max} distribution tail we can see that the performance of our method is compatible with the standard approach for the case of a proton-dominated composition but can be used to obtain near-unbiased results at arbitrarily high helium fractions.

Further studies are needed before this promising method can be applied to data. In particular, the performance of the method in case of the presence of a large fraction of intermediate-mass nuclei needs to be investigated as well as the stability of the method with respect to experimental uncertainties of the X_{max} scale, other properties of hadronic interactions (multiplicity, elasticity,...) and parameters of the Glauber calculation. Given that the method is designed to be able to work beyond the narrow energy range in which the ultra-high energy mass composition is believed to be proton-dominated, the large lever-arm of modern X_{max} data sets from near-LHC center-of-mass energies at 10^{17} eV to beyond 10^{19} eV will help to constrain the energy evolution of both, hadronic interactions and mass composition.

References

- [1] T. K. Gaisser *et al.* [Fly’s Eye Collaboration], *Comments Astrophys.* **17** (1993) 103.
- [2] A. Aab *et al.* [The Pierre Auger Collaboration], *Phys. Rev. D* **90** (2014) 122006.
- [3] P. Abreu *et al.* [The Pierre Auger Collaboration], *Phys. Rev. Lett.* **109** (2012) 062002.
- [4] R. Ulrich for the Pierre Auger Collaboration, *PoS ICRC2015* (2015) 401.
- [5] R.U. Abbasi *et al.* [Telescope Array Collaboration], *Phys. Rev. D* **92** (2015) 032007.
- [6] R.U. Abbasi *et al.* [Telescope Array Collaboration], *Phys. Rev. D* **102** (2020) 062004.
- [7] T. Bergmann *et al.*, *Astropart. Phys.* **26** (2007) 420.
- [8] F. Riehn *et al.*, *Phys. Rev. D* **102** (2020) 063002.
- [9] J. Bellido for the Pierre Auger Collaboration, *PoS ICRC2017* (2017), 506.
- [10] R. Alves Batista *et al.*, *Front. Astron. Space Sci.* **6** (2019), 23.
- [11] T. Pierog *et al.*, *Phys. Rev. C* **92** (2015) 034906.
- [12] A. Aab *et al.* [Pierre Auger Collaboration], *Phys. Rev. D* **90** (2014) 122005.
- [13] P. Lipari, *Phys. Rev. D* **103** (2021) no.10, 103009
- [14] R. Ulrich, R. Engel, and M. Unger, *Phys. Rev. D* **83** (2011) 054026.
- [15] R.J. Glauber and G. Matthiae, *Nuclear Physics B* **21** (1970) 135.
- [16] M. De Domenico, M. Settimo, S. Riggi and E. Bertin, *JCAP* **07** (2013), 050
- [17] H.Dembinski *et al.*, *scikit-hep/iminuit: v2.0.0 Zenodo* 2020.
- [18] S. Baker and R. D. Cousins, *NIM* **221** (1984) 437.