Constraining Models of the Extragalactic Cosmic-Ray Origin with the Pierre Auger Observatory

Von der Fakultät für Mathematik, Informatik und Naturwissenschaften
der RWTH Aachen University zur Erlangung des akademischen Grades
eines Doktors der Naturwissenschaften genehmigte Dissertation

vorgelegt von

Diplom-Physiker
David Simon Walz
aus Mainz

Berichter:
Universitätsprofessor Herr Prof. Dr.rer.nat Martin Erdmann
Universitätsprofessor Herr Prof. Dr.rer.nat Günter Sigl

Tag der mündlichen Prüfung: 7.11.2016

Contents

Declaration of Pre-released Extracts

Acknowledgements

Chapter 1

Introduction

Our planet is constantly hit by a stream of energetic subatomic particles, which are known as ‘cosmic rays’. These particles are mostly protons, but also heavier nuclei, as well as small fractions of leptons, photons and neutrinos. Cosmic rays are observed with energies ranging from below 1 GeV to up more than 10^{20} eV, making them the most energetic particles known to man.

Upon entering the atmosphere a highly energetic cosmic ray interacts with air nuclei, thereby producing a cascade of secondary particles. To illustrate this, a cosmic ray of 10^{20} eV energy creates an ‘extensive air shower’ of $\sim 10^{11}$ secondary particles spread over a $10\,\mathrm{km}^2$ area on the ground. These extensive air showers are essential to the detection of cosmic rays. Since the flux of cosmic rays is strongly decreasing with energy, direct detection via small balloon-borne or space-based experiments is only feasible up to about 10^{15} eV energy. The decreasing flux at higher energies requires vast detection volumes that can be realized by using the Earth’s atmosphere for an indirect measurement of cosmic rays via their extensive air showers.

Since their discovery more than 100 years ago, cosmic rays have been intensively studied and numerous advancements in their understanding have been made (see [1] for a review). Still, many aspects of the phenomenon remain unclear. For instance, while there is a general belief that the vast majority of cosmic rays above $10^{18.7}$ eV are charged hadrons accelerated at discrete extragalactic sources, neither the positions and type of these sources, nor their acceleration mechanism is established.

A prominent feature in the observed cosmic ray spectrum is the flattening at $10^{18.7}$ eV, the so called ‘ankle’. The ankle could mark the transition between galactic and extragalactic origin of cosmic rays. Alternatively, this transition could already occur at lower energies, with the ankle being a signature from interactions of cosmic ray protons with the cosmic microwave background (CMB).

Another distinct feature of the spectrum is a cutoff at around $10^{19.7}$ eV. This cutoff was first predicted by Greisen, Zatsepin and Kuz’min (GZK) for cosmic ray protons due to energy losses in photoproduction interactions with the CMB, and a similar effect is expected for heavier nuclei due to photodisintegration. However, it is still unclear whether the observed cutoff is due to these interactions, or due to cosmic ray sources reaching their maximum acceleration energy, or a combination of both.

Recent measurements of the cosmic ray mass composition can help to shed light on these questions. These measurements suggest an increasingly heavy and pure composition above $10^{18.2}$ eV. This observation could be explained by a rigidity-dependent maximum acceleration energy in combination with propagation effects.

In this thesis we aim to assess the aforementioned questions through a combined investigation of the spectrum and composition measured with the Pierre Auger Observatory. Starting from a simple model of the extragalactic sources, detailed simulations of cosmic ray propagation are performed. The resulting model predictions are compared to the measurements in a statistical analysis and constraints on the source model are derived. Complementary to similar recent studies special attention will be given to identifying all relevant theoretical and experimental uncertainties and to estimating their impact.

The thesis is structured as follows: Chapter 2 starts with a general description of cosmic ray air showers and its features that can be used for inferring the properties of the primary cosmic ray. Then, the observational results on cosmic rays are reviewed. Finally, an astrophysical interpretation of the measurements is presented, and the main open questions are specified. In chapter 3 a theoretical overview is given of the physics of cosmic-ray acceleration and propagation through extragalactic and galactic space. Chapter 4 describes the implementation of the propagation physics in CRPropa 3, which was developed as part of this thesis. The Pierre Auger Observatory along with its key measurements, the spectrum and composition of cosmic rays above $10^{17.8}$ eV, are described in chapter 5. Then, in chapter 6 we define a set of astrophysical models and formulate a statistical method for testing these models for agreement with the measurements. Using the statistical analysis we constrain the model parameters in chapter 7 and interpret the results in comparison with recent studies. We also study the impact of the major systematic uncertainties of the experiment and the simulations, and estimate the effect of several simplifying assumptions in the source model. Finally, we summarize the main results and discuss the implications.

Chapter 2

Cosmic Rays

At energies above $\sim 10^{15}\,\mathrm{eV}$ the cosmic-ray flux becomes too small for direct detection with balloon-borne or space-based experiments. Instead, by observing extensive air showers, the atmosphere can serve as a giant detector volume to indirectly measure the primary cosmic rays. In the following we first describe the air shower phenomenon and how the main cosmic-ray properties — energy, composition and direction — can be reconstructed from the measurements of the shower properties. Then, using this understanding we review the experimental results on the energy spectrum (section 2.2), the composition (section 2.3) and the arrival directions (section 2.4), all with a focus on ultra-high energy cosmic rays (UHECR), i.e. cosmic rays with an energy $E > 10^{18}\,\mathrm{eV}$. Finally, on the basis of these observations we will motivate a generally accepted astrophysical interpretation of the cosmic-ray phenomenon in section 2.5 and formulate the open questions.

2.1 Extensive air showers

Upon entering the atmosphere a cosmic ray interacts with an air nucleus at a typical height of $15 - 35\,\mathrm{km}$ above ground [2] and starts a massive cascade of secondary particles. The front of such an extensive air shower forms a curved disk of only a few meters thickness, moving with the speed of light in the cosmic ray's incident direction, and extending over square-kilometers on the ground.

The particle cascade, as seen in figure 2.1, is generally separated into three components: hadronic, electromagnetic and muonic. If the cosmic ray is a nucleon or nucleus, the shower development typically starts with a hadronic interaction with an air nucleus. Multiple hadrons, mostly pions, kaons and nucleons, are produced in this and each subsequent interaction, thereby forming the hadronic component. With a lifetime of $c\tau \approx 3 \times 10^{-9}\,\mathrm{m}$ neutral pions effectively immediately decay into photons ($\pi^0 \longrightarrow \gamma\gamma$) and are the main source of the electromagnetic component. The electromagnetic cascade continues by $e^{\pm}$ pair production of photons and bremsstrahlung of electrons and positrons. Charged pions and kaons have a relatively long lifetime of $c\tau \approx 8\,\mathrm{m}$ and $4\,\mathrm{m}$, respectively, which allows them to interact before decaying. However, at lower Lorentz factors they decay into muons and neutrinos ($K^{\pm}, \pi^{\pm} \longrightarrow \mu^{\pm} + \nu_\mu/\bar{\nu}_\mu$) before interacting, thereby feeding the muonic component. Muons reach ground level mostly unattenuated, due to their long life-time ($c\tau \approx 660\,\mathrm{m}$) and small energy loss in the atmosphere. Since deflections by elastic scatterings off air nuclei are also of low importance, muons travel on approximately straight lines and keep information about their production point along the shower axis. This is a useful property, because the point of maximum particle production is related to the mass of the primary cosmic ray. In very inclined showers, i.e.

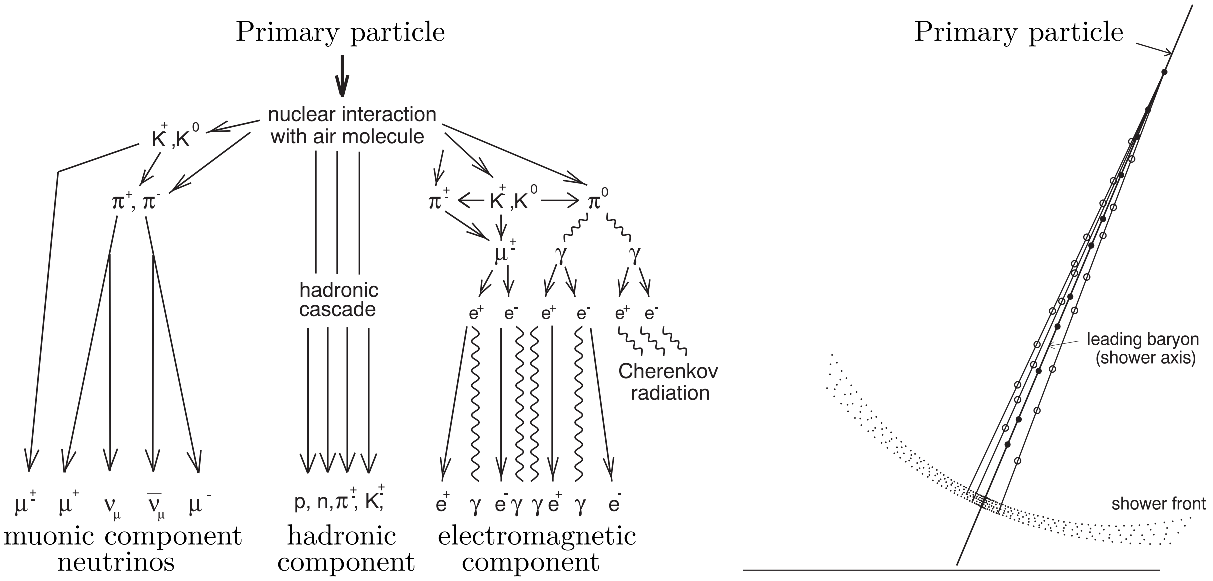


Figure 2.1: Sketch of an extensive air shower (modified from [3]). Left: schematic development of the main components of the particle cascade. Right: geometry of shower front and axis .

with zenith angle $\theta > 65^\circ$, muons, and the electrons from their decay, are essentially the only detectable particles on the ground. The full shower development, however, is always dominated by the photons and electrons of the electromagnetic component. This is exemplified in figure 2.2, showing the average lateral distribution and longitudinal development of particles for the case of a vertical shower induced by a 10^{19} eV proton. The number of photons and electrons is seen to be higher by several orders of magnitude than of that hadrons and muons. Concretely, for a 10^{19} eV proton shower the electromagnetic component makes up 99% of the shower particles and carries around 85% of the energy [4].

The lateral distribution (figure 2.2 left) describes the density of shower particles as a function of distance r to the shower axis. It can be modeled by a Nishimura-Kamata-Greisen function [5, 6]

$$S(r) \propto \left(\frac{r}{r_M}\right)^{s-2} \left(1 + \frac{r}{r_M}\right)^{s-4.5} \tag{2.1}$$

where r_M is the Molière radius, and the shower age s describes the stage of the longitudinal shower development corresponding to the height of observation. The longitudinal development (figure 2.2 right) describes the number of shower particles along the shower trajectory through the atmosphere. It is best expressed as a function of 'slant depth'

$$X(l) = \int_l^\infty \rho(\vec{r}(l'))\, dl' \ , \tag{2.2}$$

which is the integrated atmospheric density ρ traversed by the shower front on its path $\vec{r}(l)$ with the longitudinal coordinate l along the shower axis. The longitudinal development is often

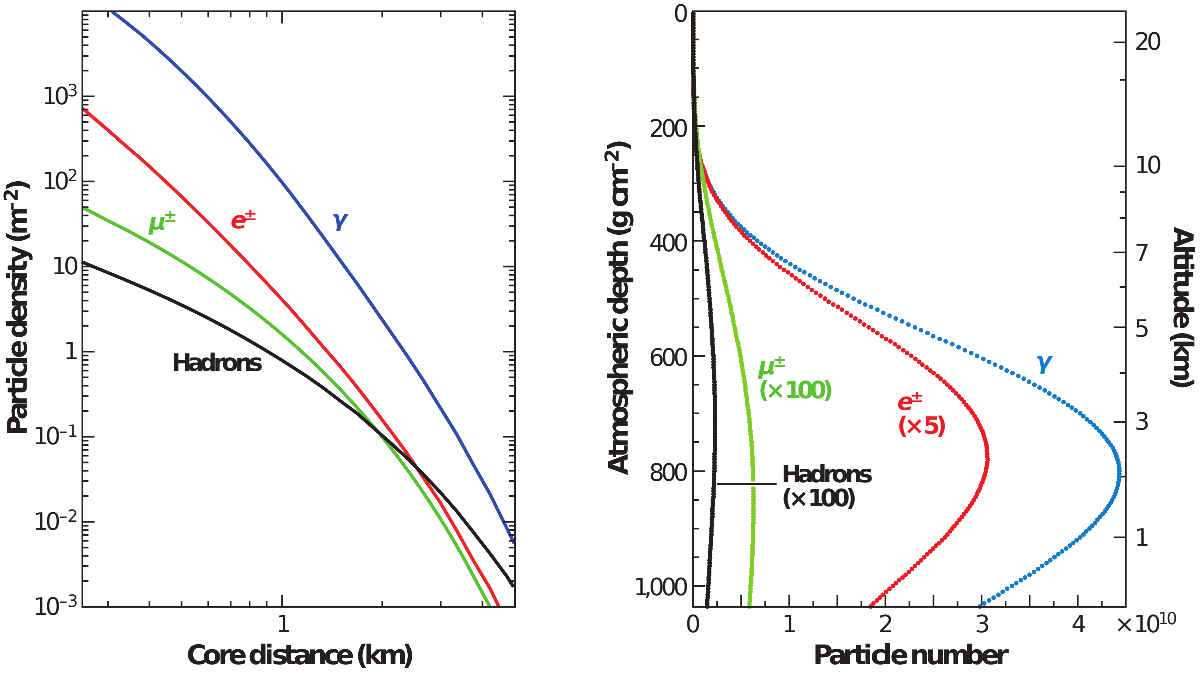


Figure 2.2: Average profiles for vertical showers of 10^{19} eV protons (from [2]). Left: Lateral distribution of the particle density at a slant depth of $X = 870\,\mathrm{g/cm^2}$ corresponding the atmospheric overburden of the Pierre Auger Observatory. Right: Longitudinal distribution of the particle number as function of traversed atmospheric depth. The individual components are scaled for better visibility. For reference the altitude above sea level is shown as well.

modeled by a ‘Gaisser-Hillas’ function [7]

$$N(X) = N_{\max} \left(\frac{X - X_0}{X_{\max} - X_0} \right)^{\frac{X_{\max} - X_0}{\lambda}} \exp \left(\frac{X_{\max} - X_0}{\lambda} \right) \tag{2.3}$$

where X_0 is the first interaction point, $X_{\max}$ is the point of maximum particle number $N_{\max}$, and λ is a shape parameter.

2.1.1 Main Shower Features

The properties of the primary cosmic ray, i.e. energy E, mass number A and arrival direction, need to be inferred from the macroscopic features of the shower, e.g. $N_{\max}$ and $X_{\max}$. The main features of the relation between shower features and cosmic-ray properties are easily derived from two simple models, namely the Heitler [8] and Heitler-Matthews model [9] of electromagnetic and hadronic cascades. Additionally the superposition model is used, in which a cosmic-ray nucleus of mass number A is approximated as A independent nucleons of E/A energy each. This approximation is valid since binding energies are small compared to collision energies early in the shower. Using these, the principal relations are as follows (cf. [2, 4, 9]):

1. The number of shower particles scales with energy $N_{\max} \propto E$.

2. The mean depth of shower maximum scales logarithmically with energy $\langle X_{\max} \rangle \propto \ln E$.
3. Nuclear showers develop earlier than proton showers $\langle X_{\max}(p) \rangle - \langle X_{\max}(A) \rangle \propto \ln A$.
4. Fluctuations of $X_{\max}$ from shower to shower decrease with mass number.
5. Nuclear showers have more muons compared to proton showers $N_\mu^A \approx N_\mu^p A^{0.15}$.

The aforementioned relations are generally reproduced and more precisely predicted by detailed shower simulations, the CORSIKA code [10] being the most widely used. Due to the indirect nature of the air shower measurements, these simulations play a crucial role in cosmic-ray physics. While the electroweak interactions in air showers are well understood, there are considerable uncertainties in modeling the hadronic interactions [2, 11, 12].

The highest center-of-mass energy that can be investigated at a collider experiment is currently $\sqrt{s_{pp}} = 13\,\text{TeV}$ in proton-proton interactions at the LHC. This corresponds to a fixed target interaction of cosmic ray at $E = s_{pp}/(2m_p) \approx 10^{17}\,\text{eV}$, assuming a proton target. Therefore, the measured cross sections need to be significantly extrapolated in order to model the first interactions in UHECR showers. Also, the majority of hadronic interactions inside a shower are soft processes for which pertubative computation methods are not feasible and effective theories need to be applied instead. Furthermore, the hadronic cross sections for most nuclei are not measured so far. As a substitute, nucleus-nucleus interactions are described in terms of elementary nucleon-nucleon interactions via the Glauber approach, see e.g. [13] for a review. While this approach describes well the measured total and elastic cross sections, its validity in very high energy multiparticle production is not experimentally verified [2]. Also, cross sections are unknown for a number of shower hadrons, e.g. strange baryons and charmed or bottom mesons. As a consequence of these challenges there are significant differences among available hadronic interaction models, namely EPOS [14], QGSJet [15] and SIBYLL [16, 17]. These differences are manifested, among others, in different predictions of $X_{\max}$ and muon numbers. For a review of these models and constraints derived from accelerator data refer to [18]. In this thesis we make use of shower simulations with each of the aforementioned interaction models.

2.1.2 Measurement Techniques

There are two well established methods for observing cosmic-ray air showers: particle detector arrays and fluorescence telescopes. Particle detector arrays sample the shower particles at multiple positions on the ground. The incident cosmic-ray direction is then reconstructed from the arrival times. The primary energy is measured by estimating the total number of particles from a model of the lateral distribution function, see equation 2.1. Information on the cosmic-ray mass can be obtained by separating electromagnetic and muonic component in order to measure the muon production depth $X^\mu_{\max}$ [19], the electron-muon ratio [20], or certain shape parameters of the shower (cf. [21]).

Fluorescence telescopes observe the faint fluorescence light from excited air molecules as the shower front passes through the atmosphere. Since the amount of fluorescence light is proportional to the number of shower particles, fluorescence telescopes are able to observe the longitudinal shower development. Integrating the longitudinal development, cf. equation 2.3, provides a calorimetric measurement of the cosmic-ray energy, except for a small fraction of 'invisible energy' that is carried away by neutrinos and high energy muons. The depth of shower maximum X_{max} is sensitive to the cosmic-ray mass number. The arrival direction can be reconstructed by observing the shower from at least two locations, or from one location in combination with a measurement of the shower impact point on the ground.

The shower front contains a large number of particles that collectively move at relativistic velocities. The charged particles in the shower front cause several types of radiation including bremsstrahlung, Cherenkov and synchrotron radiation. Due to the Lorentz-boost these radiation processes are strongly collimated in the forward direction and can be observed by suitable detectors that are located close to the shower axis. The Cherenkov radiation can be observed by imaging telescopes, such as HESS [22], or arrays of non-imaging detectors (e.g. HAWC [23] and Tunka-HiScore [24]). The shower front emits a radio signal, which is in first order caused by shower electrons and positrons radiating in the geomagnetic field, and in second order due to the Askaryan effect [25] caused by a negative net charge of the shower front. The radio signal propagates largely unattenuated through the atmosphere and, therefore, contains information on the shower development. At ground level the beamed radio signal results in a short $\sim 10\,\mathrm{ns}$ pulse at MHz-GHz frequencies, which can be observed relatively close to the shower axis with suitable radio detector arrays, such as AERA [26].

Currently the two largest cosmic-ray experiments are the Pierre Auger Observatory (hereafter Auger), described in chapter 5, and the Telescope Array experiment [27] (hereafter TA). Both observatories are 'hybrid' detectors, consisting of surface and fluorescence detectors as well as several enhancements, and are focusing on detecting UHECRs. The main results of Auger, TA and other experiments are reviewed in the following.

2.2 Energy spectrum

The cosmic-ray flux strongly depends on energy but shows little directional dependency, as reviewed in section 2.4. Therefore, one of the most basic ways to describe cosmic rays is the differential flux that quantifies the number of particles N crossing an area A per time t, energy E and solid angle Ω. The 'energy spectrum' describes this differential flux as a function of energy.

$$J(E) = -\frac{d^4N}{dE\,dA\,d\Omega\,dt}(E) \tag{2.4}$$

A non-exhaustive collection of spectrum measurements is seen figure 2.3. The selected measurements extend from 10^8 eV to beyond 10^{20} eV, thus covering more than 12 orders of magnitude in energy. The cosmic flux decreases roughly by a factor of 500 in each energy decade. At $E = 10^9$ eV the flux measures more than 1000 particles per m^2 and s, decreasing to about one per m^2 and year at 10^{15} eV, and to less than one particle per km^2 and century above 10^{20} eV.

Below $\sim 10^9$ eV per nucleon the cosmic radiation is modulated by magnetized solar winds, which decelerate charged particles and partially shield them off the inner solar system. On

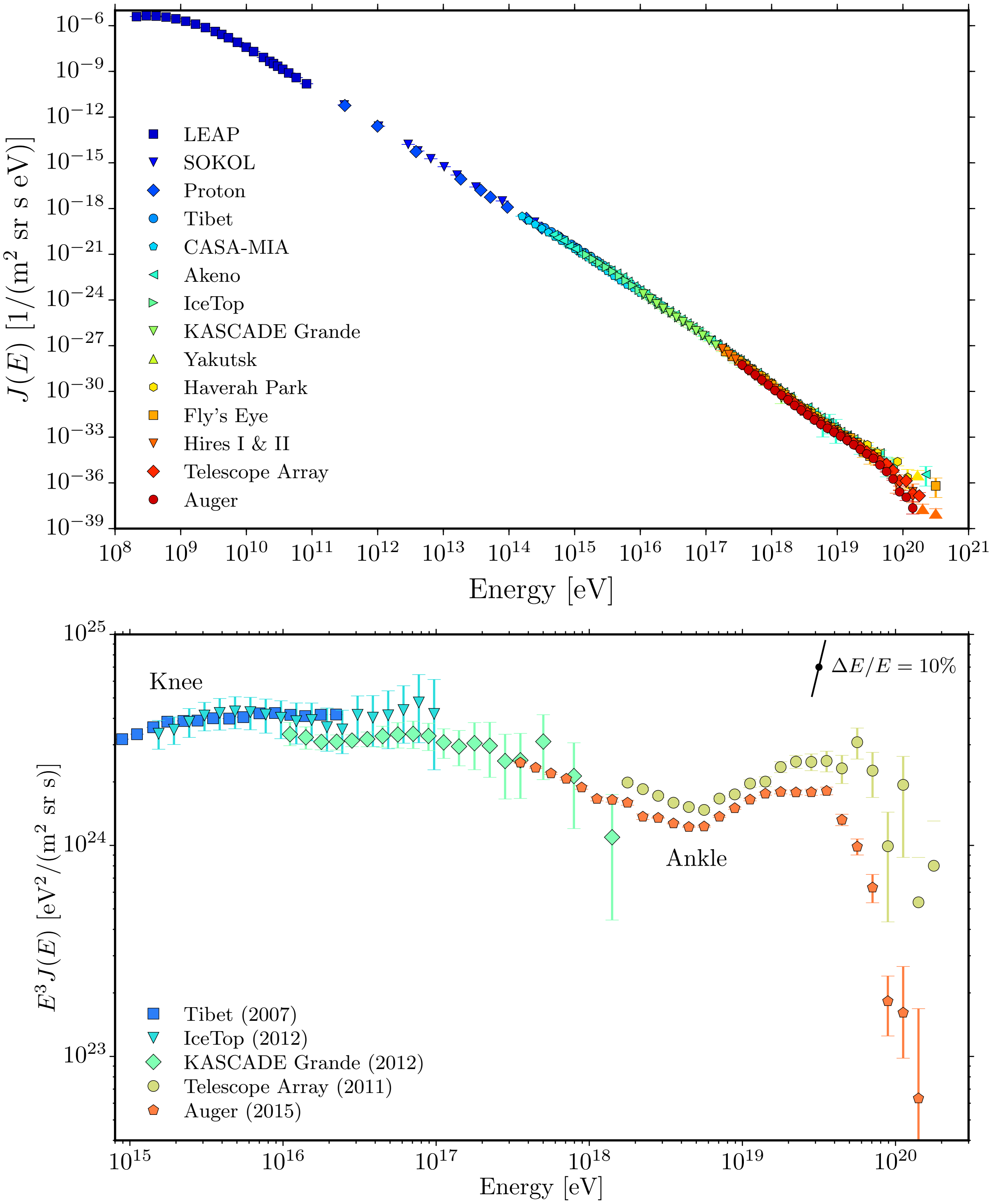


Figure 2.3: Top: All-particle cosmic-ray spectrum measured by various experiments [29–31, 36–45]. Bottom: Spectrum above 10^{15} eV from air shower experiments, multiplied with E^3 to emphasize signature features: knee, ankle and cutoff. The effect of 10% energy scale shift is shown exemplarily in the upper right.

Earth the geomagnetic field causes an additional shielding against charged particles [28]. This causes the flattening in the spectrum at the lowest energies.

Above $\sim 10^9$ eV the spectrum in its double-logarithmic representation is seen to follow a smooth power-law $J(E) \propto E^{-\alpha}$. The bottom of figure 2.3 shows the spectrum scaled with a factor E^3, which cancels out the general slope in order to highlight distinct features of the spectral shape. These features are named according to the spectrum's resemblance to a human leg. At the 'knee' around 3×10^{15} eV the spectral slope steepens from $\alpha \sim 2.6$ to 3.1. At the 'ankle' at $\sim 5 \times 10^{18}$ eV the slope flattens again to $\alpha \sim 2.6$. Finally, at the highest energies, above 4×10^{19} eV, a flux suppression has been established with high significance [29, 30]. This cutoff indicates a limit to the possible energies of cosmic rays reaching Earth.

Beside these main features, additional spectral changes between knee and ankle have been recently observed by several experiments [31–33], showing a flattening at $\sim 2 \times 10^{16}$ eV and a steepening at $\sim 1.3 \times 10^{17}$ eV ('second knee'). These features in the cosmic-ray spectrum are of great interest since they are indicative of the origin and propagation of cosmic rays, which will be discussed in section 2.5.

In the scaled spectrum in the bottom of figure 2.3, differences in the energy scale between experiments are emphasized. For instance, the positions of the ankle measured by Auger and TA are different, but can be brought to agreement through a constant shift of the energy scale. This does not hold for the energy range above the ankle, where the spectrum measured by TA is relatively higher, even after accounting for a shift of energy scale. A possible explanation is a difference in spectrum on the northern and southern hemispheres. However, a comparison of the spectrum measured by Auger in different bands of declination does not show a declination dependency [30] despite a significant overlap with the TA field of view.

2.3 Composition

At energies below 10^{15} eV the composition of cosmic rays is known from direct measurements above the atmosphere where cosmic rays did not get to interact. Effectively all elements of the periodic table have been found in cosmic rays and their abundance at GeV per nucleon energies largely corresponds to the abundance in the solar system [21]. The light elements, $Z = 3 - 5$, as well as the elements closely below iron and lead, are overabundant in cosmic rays, which can be interpreted in terms of propagation processes. The spectra for the relatively abundant elements up to iron have been measured individually, see figure 2.4. All elements are seen to follow the same power-law behavior.

Above 10^{15} eV cosmic rays can only be measured by air shower experiments, and information on the nature of the primary cosmic ray needs to be inferred from the shower observables as outlined in section 2.1. The most sensitive air shower observable is the depth of shower maximum $X_{\max}$. On average, a shower induced by an iron nucleus develops faster than a proton shower at the same energy, leading to a difference in $X_{\max}$ of about $100\,\mathrm{g/cm^2}$. However, since average shower-to-shower fluctuations are of the order $\sigma_{\mathrm{sh}}(X_{\max}) \sim 20 - 60\,\mathrm{g/cm^2}$, and since detector resolution is typically not better than $20\,\mathrm{g/cm^2}$, a mass determination on an event basis is not possible. Instead, the composition is inferred statistically from the $X_{\max}$

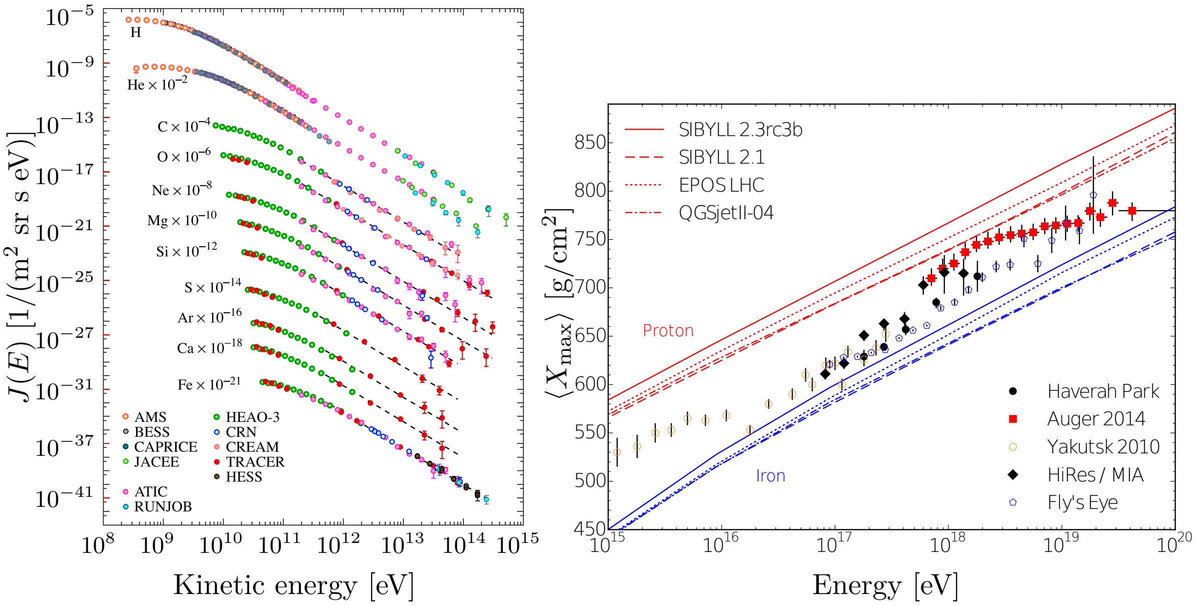


Figure 2.4: Composition measurements. Left: Fluxes of individual elements from direct measurements (modified from [28]). The absolute fluxes are scaled as indicated for better visibility. Right: Average depth of air shower maximum from several experiments, and comparison with shower simulations using current hadronic interaction models (modified from [17]).

distribution of an ensemble of air showers measured at similar energies [34]. The means of these distributions as measured by several experiments are shown in figure 2.4. A comparison of the measurements with the expectation from shower simulations indicates the evolution of mass composition with energy. Above the knee at 2×10^{15} eV to 10^{17} eV the composition moves from light to heavy, then reversing and becoming increasingly proton-like at around 10^{18} eV before the ankle, and finally becoming heavier again towards the highest energies. The measured energy range does not extend beyond 10^{20} eV, because fluorescence $X_{\max}$ measurements require dark nights, which implies a significant reduction of statistics compared to surface detectors. One should note that measurements by the HiRes and TA collaborations are not shown in figure 2.4 for lack of comparability, because the detector response has not been unfolded from these measurements. Although TA measures a dominantly light composition above the ankle in contrast to Auger, a recent analysis showed that applying the TA detector response on Auger data gives consistent results [35].

2.4 Arrival directions

The measurement of arrival directions is complementary to that of the energy spectrum and composition for understanding the origin and propagation of cosmic rays. At all energies the observed arrival directions are found to be largely isotropic, with no clear correlations with source candidates [46]. However, small anisotropic features have been found at different energies and on a wide range of angular scales.

The angular distribution of cosmic-ray directions $\phi(\vec{n})$ is often described in terms of spherical harmonics $Y_{lm}(\vec{n})$,

$$\phi(\vec{n}) = \sum_{l \geq 0} \sum_{m=-l}^{l} a_{lm} Y_{lm}(\vec{n}) \,. \tag{2.5}$$

Over- and under-densities in the cosmic-ray flux show up as non-zero multipole amplitudes a_{lm}. The power spectrum is defined as the average squared amplitude

$$C_l = \frac{1}{2l+1} \sum_{m=-l}^{l} |a_{lm}|^2 \tag{2.6}$$

and quantifies the anisotropic signal power that is present on an angular scale of $180°/l$.

At TeV to PeV energies, multiple experiments have discovered anisotropies at the level of 10^{-4}–10^{-3} in the directional flux. On the northern hemisphere several regions of relative excess and deficit of have been measured by the Tibet ASγ [47], Milagro [48], ARGO-YBJ [49] and HAWC [50] experiments, among others. On the southern hemisphere the only measurements come from the IceCube experiment [51] showing qualitatively similar results.

On both hemispheres, a decomposition of the arrival directions into spherical harmonics shows that most of the anisotropy is contained in the low-multipole moments $l < 5$, but higher multipoles down to angular scales of 10 degrees are also found to contribute [50, 51]. The amplitude and phase of the dominating dipole moment is found to vary with energy in the measured range of $10^{13} - 10^{16}$ eV. A limitation arises in the spherical harmonics analysis from the incomplete sky-coverage of each experiment, which generally causes correlations between multipole moments. For this reason, work is being done towards combining data-sets in order to achieve a full-sky coverage [52].

At energies above 10^{18} eV the main experimental data is provided by the Pierre Auger Observatory and the Telescope Array experiment. At energies $E > 10^{19}$ eV the cosmic-ray data recorded at Auger and TA has been combined for a full sky harmonic analysis. The initial analysis [54] showed no excess from the isotropic expectation at the 99% confidence level at any angular scale. However, a repeated analysis considering an additional year of data and extending the Auger field of view to 85% of the sky by including highly inclined events, found evidence for a dipole moment [53]. The measured dipole, see figure 2.5, has an amplitude of 6.5(19) % and points to 93(24)° in right ascension and −46(18)° in declination. This reinforces the evidence for an equatorial dipole found in the Auger dataset alone [55], as well as hints in earlier analyses at lower energies [56, 57], where the dipole phase was found to be consistent across energy bins and showing a smooth transition from $\sim -270°$ to $\sim 90°$ in right ascension at around 10^{18} eV. It is interesting to note that these phases correspond to a transition from the azimuthal direction of the galactic center to the galactic anti-center at higher energies. The declination of the dipole reconstructed at $E > 10^{19}$ eV, however, does not match the direction of the anti-center.

At the highest energies $E > 5 \times 10^{19}$ eV the Pierre Auger collaboration previously reported evidence of a correlation of cosmic rays above ~ 60 EeV with the extragalactic matter distribution in the nearby universe [58, 59]. An update of this analysis found a reduced correlation

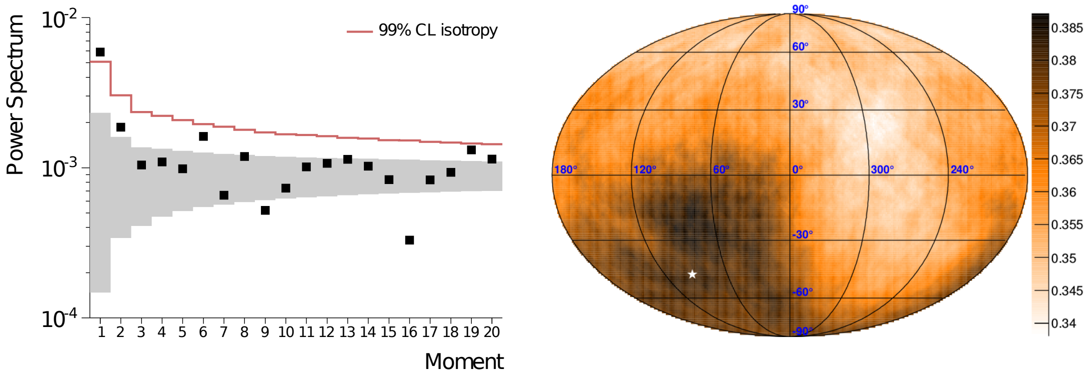


Figure 2.5: Harmonic analysis of cosmic rays $E > 10^{19}$ eV recorded by the Pierre Auger Observatory and the Telescope Array experiment [53]. Left: Measured power spectrum (black dots), and expectation from an isotropic distribution with gray bands showing the RMS around the mean values and the red solid line indicating the upper limit at 99% confidence level. Right: Sky map in equatorial coordinates of the flux above 10^{19} eV in units of $(\mathrm{km}^2\,\mathrm{yr}\,\mathrm{sr})^{-1}$. The flux is smoothed over 60° to emphasize the dipole, whose direction is denoted by a white star.

strength compared to the initial results [60]. Also, no significant correlation was found with the galactic and super-galactic plane, and with several catalogs of source candidates. The autocorrelation, counting the number of event pairs within a given angular separation, is found to exhibit a modest excess from the isotropic expectation. The largest contribution comes from a clustering of events within ∼15° around the radiogalaxy Centaurus A [60, 61].

Further searches using variants of the autocorrelation observable showed no significant excess [64]. A search for a magnetically-induced alignment in the arrival directions of cosmic rays with $E > 20$ EeV has been performed without finding ‘multiplet’ structures beyond the isotropic expectation [64]. Finally, the local regions of 15° around events of $E > 60$ EeV have been investigated for patterns in the distribution of cosmic rays with $E > 5$ EeV using two observables [65]: The ‘energy-energy correlation’ quantifies the radial energy ordering within each region and is sensitive to the energy ordered smearing arising from a turbulent magnetic field. The ‘principal axes’ search for a linear collimation of energy inside the regions and are sensitive to a spectrometer-like ordering induced by regular magnetic fields. Neither observable showed a significant deviation from isotropy. Additional targeted and untargeted searches for point sources have been performed for neutrons [66, 67], neutrinos [68] and photons [69].

In the northern sky the Telescope Array has found evidence for a localized excess in the arrival directions of cosmic rays above 57 EeV, with 19 out of 72 events correlating in a region of 20° radius [63]. The hotspot is centered at about 19° off the supergalactic plane and does not directly coincide with a prominent cosmic-ray source candidate. The chance probability of the excess is 3.7×10^{-4} An unchanged significance was found, when considering two additional years of data [62]. A sky map of the statistical significance of clustering for cosmic rays with $E > 57$ EeV is shown in figure 2.6, indicating the hotspots observed by both experiments.

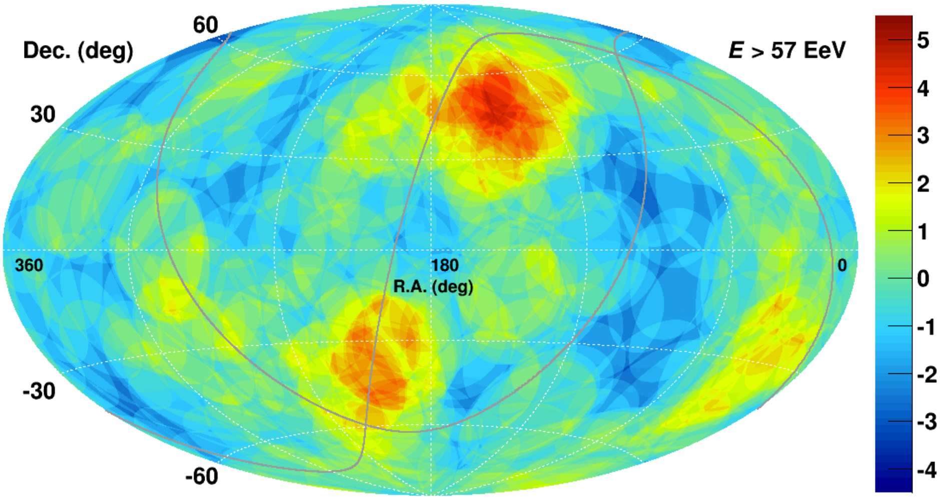


Figure 2.6: Sky map of events above $E > 57\,\mathrm{EeV}$ measured by TA and Auger in equatorial coordinates [62]. The (109 + 157) events are oversampled with 20°-radius circle to emphasize the hotspots seen in both hemispheres at that angular scale. The color code shows the statistical significance as described in [63] without accounting for random clustering. The southern hotspot in the direction of Centaurus A is not significant. Note, that no corrections are applied to account for the different energy scales of the experiments.

2.5 Astrophysical Interpretation

In the following we briefly review the prevailing astrophysical interpretation of the cosmic-ray phenomenon, leaving the theoretical details to the next chapter. More in-depth astrophysical reviews can be found in [4, 21, 70].

The power-law shape and the extreme range of the energy spectrum (figure 2.3) strongly indicate a non-thermal origin of cosmic rays. One set of theories explains cosmic rays as decay products of super-heavy relic particles or topological defects (cf. [71]). However, these hypotheses are disfavored from the non-observation of photons at ultra-high energies from such decays [72]. Instead, it is now commonly believed that cosmic rays are protons and nuclei that are accelerated by powerful astrophysical objects. cosmic-ray photons and neutrinos are then products of cosmic-ray hadrons when interacting with ambient radiation fields during acceleration or propagation.

The cosmic-ray flux up to some $10^{15}\,\mathrm{eV}$ is generally believed to be caused by supernova remnants (SNR) in our Galaxy. The changing spectral shape at this energy, the knee, could then be an effect of reaching the maximum energy of galactic magnetic confinement or acceleration capability of SNR. In both cases the maximum energy is proportional to the cosmic-ray charge number Z, so that heavier elements can potentially reach Z times higher energies. A natural assumption is therefore that protons first reach their maximum energy, then give way to helium and subsequently to increasingly heavy elements. Since heavier elements are less abundant, this transition would be observed as a steepening of the cosmic-ray spectrum, as sketched in figure 2.7. Evidence for this so called ‘Peters cycle’ [73] is in fact observed in

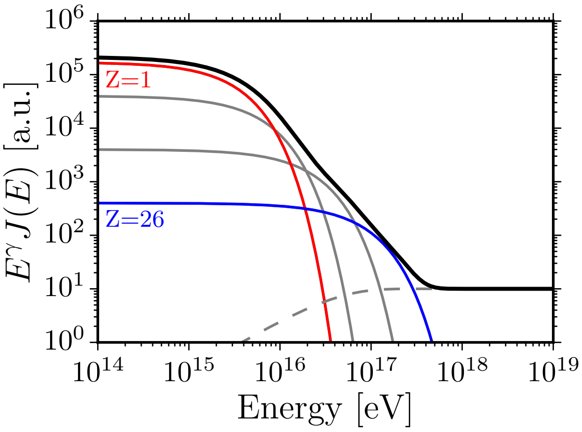


Figure 2.7: Sketch of a Peters cycle. Exemplarily shown are the energy spectra of four nuclei: protons (red), helium and nitrogen (grey) and iron (blue). Due to a maximum rigidity the dominant proton component gives way to increasingly heavy nuclei, which results in a spectral steepening in the total spectrum (black). A separate component (dashed line) takes over somewhere above the maximum energy of iron, causing a spectral flattening.

both the energy spectrum and the composition measurements: In the spectrum, if the knee at $(3-4)\times 10^{15}$ eV marks the maximum energy for protons, then for iron the maximum energy should be seen at $\sim 10^{17}$ eV, which roughly coincides with the second knee. In the direct composition measurements below 10^{15} eV (left of figure 2.4), the relative fluxes of individual elements are seen to be independent of energy, showing that the maximum energy for protons has not been reached yet. At higher energies, the $X_{\max}$ measurements (right of figure 2.4) show that the composition becomes increasingly heavy towards the second knee, as expected from a Peters cycle.

Above the ankle at 5×10^{18} eV a galactic origin is unlikely, firstly, because there are no accelerators capable of reaching these energies. Secondly, the galactic magnetic field is not strong enough to contain UHECRs inside the Galaxy in order to generate the observed level of isotropy. Thus, cosmic rays above the ankle are believed to be of extragalactic origin. They could be accelerated by gamma ray bursts (GRB) or active galactic nuclei (AGN), among other candidates [74]. Since the sources of UHECR are not known, the location and even the number density of the sources are unknown as well. On average, the number of sources in a spherical shell increases with the distance squared r^2 and the flux from each source decreases with r^{-2}, so that in principle every distance should contribute equally to the observed flux. However, above energies of ~ 50 EeV the distance that cosmic rays can propagate is strongly limited by interactions with background photons, which is called the GZK effect [75, 76]. The large-scale structure of the matter distribution in the universe is not isotropic within the distance limit that is set by the GZK effect. Consequently, the distribution of potential sources, which follow the large-scale structure, should in principle give rise to anisotropies in the cosmic-ray arrival directions. The absence of strong anisotropies therefore indicates the presence of considerable deflections in galactic and extragalactic magnetic fields. Nevertheless, there is evidence for a dipole structure in the angular distribution of cosmic rays at energies above 10 EeV, and for a small scale clustering at $E > 60$ EeV. The dipole moment could result from a local cosmic-ray density gradient due to the distribution of nearby sources, and the small scale clustering may indicate the long-awaited directional excess from individual source.

While cosmic rays below $E \sim 10^{17}$ eV and above $E \approx 5 \times 10^{18}$ eV are generally considered to be of galactic and extragalactic origin, respectively, it is not clear at which energy the extragalactic component becomes dominant. The composition between second knee and ankle evolves from heavy to light, and no anisotropy is observed in this energy range. Therefore, it has been argued that the cosmic rays above 10^{18} eV should be of extragalactic origin, because galactic protons at these energies would give rise to anisotropies in excess of the observational limits [77]. Previous to the Auger X_{max} measurements, UHECRs were commonly believed to be predominantly protons, with the transition between galactic and extragalactic component occurring at $\sim 10^{17}$ eV. Both, ankle and cutoff could then be explained in the so called 'dip model' [78, 79] as signatures of interactions with the cosmic microwave background. Since the dip model requires a high proton fraction above the ankle, it is disfavored by the current composition measurements. Moreover, the predicted neutrino flux resulting from a pure proton composition is in excess of the limits set by IceCube [80]. Alternatively, the ankle could simply mark the transition to the extragalactic component. A concave feature such as the ankle naturally forms when one component gives way to a flatter component, $E^{-\gamma_1} \longrightarrow E^{-\gamma_2}$ with $\gamma_1 > \gamma_2$, without requiring a fine-tuning of the individual flux normalizations that would be necessary for a convex feature such as the knee. On the other hand, if the extragalactic component falls off to quickly towards lower energies, there may be a gap between the end of galactic and onset of extragalactic cosmic rays, thus requiring an additional sub-ankle component, see e.g. [81]. A third option is that the ankle forms due to photodisintegration, either inside the sources [82] or in a magnetized region around the sources [83]. Cosmic rays below the ankle would be predominantly protons knocked off nuclei of higher energies.

The focus of this thesis lies on the cosmic rays beyond the ankle. The observed cutoff in the energy spectrum around 50 EeV can be explained either by energy losses during propagation or by another Peters cycle, when reaching the maximum acceleration energies of the extragalactic sources. Since the increasingly heavy composition is hinting at the latter, much attention has recently been given to modeling this scenario [81, 83–87]. A surprising result of these studies is that a hard injection spectrum $dN/dE \propto E^{-1}$ is required to simultaneously fit the spectrum and composition. In this thesis we are going to investigate this scenario with respect to the main theoretical uncertainties in the simulation from source to observation, and to the dominant experimental uncertainties. This will allow us to test the robustness of the published results, and to identify the main sources of uncertainties.

Chapter 3

UHECRs: Origin and Propagation

This chapter gives a brief theoretical overview of the origin and propagation of UHECRs. First, the possible methods of acceleration and the implications for the source spectrum and composition are reviewed. Then, the effects are discussed that can modify the cosmic-ray flux during extragalactic and galactic propagation, most importantly the interactions with the extragalactic photon background.

3.1 Acceleration and Possible Sources

It is an open question how exactly UHECRs are accelerated to such extreme energies that reach 10^{20} eV, which in macroscopic units corresponds to about 16 J. The two main classes, briefly reviewed in the following, are direct and stochastic acceleration. A more extensive review of these mechanisms and the possible sources is given in e.g. [74].

Direct (one-shot) acceleration can be achieved either by electric or by time-varying magnetic fields. Whereas electric fields in the universe are generally quickly neutralized by astrophysical plasmas, the magnetic fields of rotating objects, such as neutron stars or black holes with magnetized disks, have been considered for acceleration [74]. Rapidly rotating neutron stars generally create outflows of plasmas with relativistic velocities $\vec{\beta}$, in which the combination of the rotational energy and the strong magnetic field $\vec{B}$ induces an electric field $\vec{E} = \vec{\beta} \times \vec{B}$. In young and fast rotating neutron stars of very high magnetic field strength (magnetars) particles can potentially be accelerated to the highest observed energies [88–91]. The energy of accelerated particles depends on the rotational speed and as the star spins down, this energy decreases. The energy spectrum then follows from the spin down rate is predicted as $dN/dE \propto E^{-1}$.

Stochastic acceleration (cf. [92, 93] for reviews on the subject) is based on the idea that relativistic particles can gain energy by elastically scattering off magnetic regions moving with a characteristic velocity β. The particle gains or loses energy depending on whether a head-on or tail-on collision takes place.

In the original theory [94] magnetized gas clouds in the interstellar medium were considered, which move with $\beta c \sim 15\,\mathrm{km/s}$ in random directions. Both, energy gains and losses take place, but due to the relative motion of particle and cloud, head-on collision with an energy gain occur more frequently. The resulting average energy gain is $\langle \Delta E/E \rangle \propto \beta^2$, which is why the mechanism is generally called 2^{nd} order Fermi acceleration.

The 1^{st} order Fermi acceleration is thought to take place in shocks, i.e. wave fronts propagating with a velocity greater than the sound velocity of the medium. The shock front separates the unshocked upstream and shocked down-stream regions. The crucial aspect is that in the rest frame of either one region the other region is approaching. Thus, a particle that scatters off magnetic irregularities inside one region and is eventually deflected across the shock front always experiences a head-on collision on the other side. The resulting energy gain is of first order, $\langle \Delta E/E \rangle \propto \beta$, which makes for a more efficient acceleration. Shock fronts potentially capable of accelerating cosmic rays occur in the expanding remnants of supernovae (SNR), in the jets of active galactic (AGN) [93], in the collimated outflows of gamma ray bursts (GRB) [95], in starburst galaxies [96], as well as in the accretion shocks around massive clusters of galaxies [97].

Stochastic acceleration naturally leads a power-law spectrum as a result of the competition between the fractional energy gain $\langle \Delta E/E \rangle$ and the chance p of escaping the accelerator region within one acceleration step. The probability that a particle is accelerated from E_0 to $E_n = E_0(1 + \langle \Delta E/E \rangle)^n$ within n steps is $p_n = (1-p)^n$. Hence, the number of particles with energy E_n or higher is proportional to

$$N(\geq E_n) \propto \sum_{i=n}^{\infty} (1-p)^i = \frac{(1-p)^n}{p} \propto \left(\frac{E_n}{E_0}\right)^{1-\gamma} \tag{3.1}$$

which is the integral form of a power-law spectrum with $\gamma = 1 - \ln(1-p)/\ln(1 + \langle \Delta E/E \rangle)$. In non-relativistic shocks a $dN/dE \propto E^{-2}$ spectrum is expected, whereas the expectation for relativistic shocks is $dN/dE \propto E^{-(2.2-2.3)}$ [98].

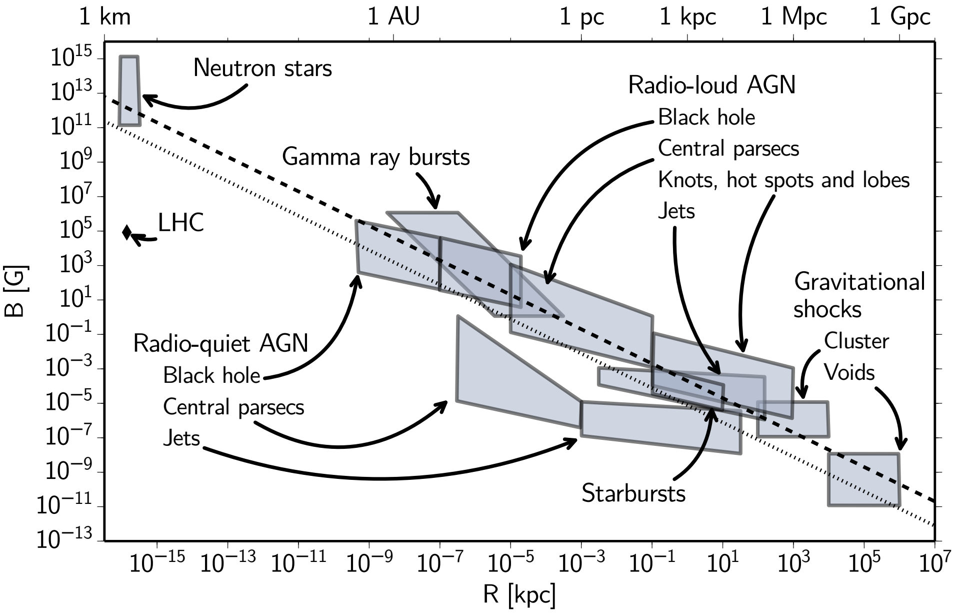


Figure 3.1: Hillas diagram, showing the size R and magnetic field strength B of possible UHECR accelerator sites. The dashed (dotted) line corresponds to the minimum requirements for accelerating protons (iron nuclei) to $E = 10^{20}$ eV. The parameters of the LHC, the largest artificial accelerator, are shown for comparison. Figure originally from [99], modified in [100].

The necessity of containing a particle within the source site during acceleration puts a limit on the maximum energy

$$E_{\max} \sim ZeBR\,, \tag{3.2}$$

where Ze is the charge of the particle, B is the magnetic field strength of the accelerator region and R its size. This necessary condition is often called the 'Hillas criterion' [101] and, with possible modifications, holds for all proposed acceleration sites that are summarized in figure 3.1. Additional necessary conditions for UHECR accelerators are [99]

- Energy losses: the energy lost by a particle through synchrotron radiation in the accelerating field, and through photonuclear and hadronic interactions cannot exceed the rate of energy gain during acceleration.
- Energy budget: the source needs to provide the energy to transfer it to accelerated particles, and the source population is required to sustain the observed UHECR flux.

Few source candidates satisfy the Hillas criterion for accelerating protons to $E = 10^{20}\,\mathrm{eV}$, as seen in figure 3.1. However, the requirements are less demanding for accelerating nuclei of higher charge number, as they can be better contained within the accelerator region. A prominent effect arises when the Hillas criterion is the limiting factor for the maximum energy, instead of energy losses. Then, nuclei can be accelerated to Z times higher energies compared to protons, which leads to a Peters cycle [73], where a dominating proton component gives way to higher elements with increasing energy. The source spectrum in this scenario can be described by

$$\frac{dN_Z}{dE} \propto (E/E_0)^{-\gamma}\, a_Z\, f_{\mathrm{cut}}(R/R_{\max}) \tag{3.3}$$

where $(E/E_0)^{-\gamma}$ is the power law shape of the acceleration mechanism, a_Z the relative abundance of element Z, $R = E/Z$ is the rigidity with a maximum $R_{\max}$ above which particles cannot be contained, and f_{cut} is a dimensionless function describing the cutoff shape. Such a source spectrum can be the reason for the observed cutoff at the highest energies, which would simultaneously explain the observed increase of cosmic-ray mass. The rate of mass-increase with energy depends on the power law and cutoff shapes, and on the relative abundances. In the source scenarios of AGN and the gravitational shocks, the accelerated particles originate from the interstellar and intergalactic medium, in which the fraction of heavier elements is small. Alternatively, the environment supplying the injected particles could already be enriched in heavy elements. This is the case for shock acceleration in GRBs [95], and in the neutron star scenario, where the accelerated particles are extracted from the iron-rich surface [91].

3.2 Cosmological Effects

Cosmic rays can, depending on their energy and particle type, travel over cosmological distances $d \gg 10\,\mathrm{Mpc}$ ($1\,\mathrm{parsec} \approx 3 \times 10^{16}\,\mathrm{m}$), where several effects of the expanding universe need to be considered. Using the redshift z to indicate the cosmic time, the expansion of the universe can be expressed by a scale factor $a = (1+z)^{-1}$, which relates the proper distance between two comoving objects at present time $z = 0$ to that at an earlier time $z > 0$. The scale factor can be used to define comoving coordinates $\vec{\chi} = \vec{r}(z)/a$ which factor out the cosmological expansion from proper coordinates $\vec{r}$.

In the standard ΛCDM cosmology the evolution of the scale factor, i.e. the rate of expansion, is given by the Hubble parameter

$$H(z) = \frac{\dot{a}}{a} = H_0\sqrt{\Omega_m(1+z)^3 + \Omega_\Lambda}\ . \tag{3.4}$$

where $H_0 \approx 63.7\,\mathrm{km/s/Mpc}$ is the Hubble constant, $\Omega_m \approx 0.315$ is the present density of baryonic and dark matter, $\Omega_\Lambda \approx 0.685$ is the dark energy density, and radiation density and curvature are neglected [28].

All relativistic particles, i.e. photons and high-energy cosmic rays, are observed with an energy $E(z = 0)$ that is redshifted compared to the initial energy $E(z) = E(z = 0)(1 + z)$. This 'adiabatic energy loss' or redshift loss is given as a function of traveled proper distance by

$$\frac{1}{E}\frac{dE}{dx} = -\frac{1}{c}H(z)\,. \tag{3.5}$$

For a cosmic ray traveling over 100 Mpc distance ($z \approx 0.024$) this results in a 2% energy loss. Nevertheless, redshift losses are the dominant energy loss for cosmic rays hadrons below $E \sim 2\,\mathrm{EeV}$ where interactions with the extragalactic photon background become negligible, see section 3.4. For cosmic rays of higher energies a more significant effect arises from the cosmological evolution of the photon backgrounds. In addition to the formation history of these backgrounds, the photon number per proper volume is diluted with time as $(1 + z)^{-3}$, decreasing the overall interaction rate of cosmic rays, and the photon energies are redshifted by $(1+z)$, which increases the energy threshold for interactions by the same amount. Likewise, the density of cosmic-ray sources decreases as $\rho(z) = \rho(z = 0)(1 + z)^3$, and magnetic fields are diluted in a flux conserving manner $B(z) = B(z = 0)(1 + z)^2$, both in addition to their respective formation histories.

3.3 Cosmic Photon Background

The universe is filled, even in the largest voids, with a diffuse radiation. This radiation spans a wide range of the electromagnetic spectrum. Its components most relevant to cosmic-ray propagation are shown in figure 3.2 and outlined in the following in order of increasing photon energy.

The cosmic radio background (CRB) at photon energies $\epsilon < 10^{-6}\,\mathrm{eV}$ is formed by the synchrotron emission of relativistic electrons in both normal and radio-loud galaxies [102]. Measuring the CRB is particularly difficult due to the dominating radio foreground of our Galaxy. Recent measurements [104] indicate that the CRB intensity could be significantly larger than estimated in earlier models [102, 105].

The cosmic microwave background (CMB) is the most intense extragalactic radiation field with 400 photons per cm^3. The CMB is a relic of the epoch of recombination 400,000 years

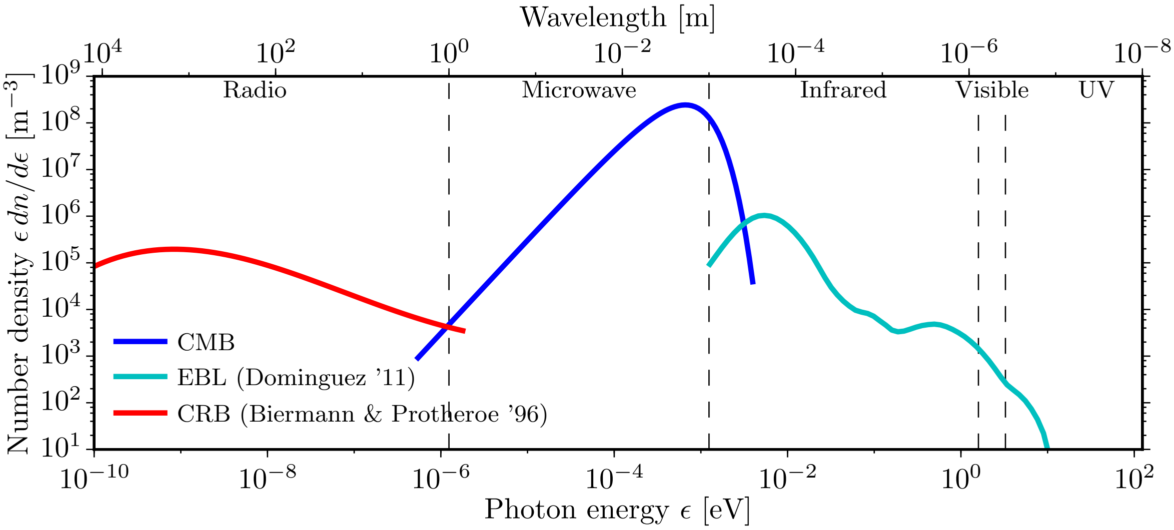


Figure 3.2: Spectrum of the diffuse extragalactic background radiation at $z = 0$: cosmic microwave background (CMB), cosmic radio background (CRB, model from [102]) and extragalactic background light (EBL, model from [103]). Dashed lines show the typical classification of the electromagnetic spectrum.

after the Big Bang and shows a near perfect blackbody spectrum of temperature $T_0 = 2.726\,\mathrm{K}$ [28], thus peaking at meV photon energies.

The extragalactic background light (EBL) is composed of the light emitted by stars and galaxies since the epoch of re-ionization. Although its density of about two photons per cm^3 is 200 times smaller than that of the CMB, the EBL plays an important role in limiting the propagation of TeV photons [106] and EeV nuclei [107]. The spectrum of the EBL has a double peak structure. The contribution of direct starlight peaks around eV energies and dominates in the near-infrared, visible and ultraviolet spectrum. A significant part of starlight is absorbed by dust in the interstellar medium and re-radiated thermally in the far-infrared, causing the second peak at 10^{-2} eV. Also, line emission of polycyclic aromatic hydrocarbons in the interstellar matter is visible in the EBL spectrum [108]. The EBL evolves with star formation rate in the universe and thus exhibits a strong cosmological evolution. An important challenge in modeling the EBL is therefore to predict its evolution based on measurements of galactic populations and their luminosities at different redshift. Multiple EBL models have been formulated using different techniques, including forward evolution of star and galaxy formation models [108–111], backward evolution models of present day observations [112, 113], and purely observational methods [103, 114]. A direct comparison of these models in the context of simulating cosmic-ray propagation is given in the next chapter.

3.4 Interaction with Photons

Interactions of cosmic rays with photons are important to consider during both acceleration and propagation, as they generally result in energy losses. During acceleration, the energy losses compete with the energy gain, and during propagation they limit the maximum propagation distance. The interactions can also modify the particle type, and they give rise to a flux of secondary particles. The interaction rate or inverse mean free path λ^{-1} of a cosmic ray of energy E, mass m and velocity $\beta \approx 1$ is given by [92]

$$\lambda^{-1}(E) = \frac{1}{8\beta E^2} \int_{\epsilon_{\text{min}}}^{\infty} \frac{n(\epsilon)}{\epsilon^2} \int_{s_{\text{th}}}^{s_{\text{max}}} \sigma(s)(s - m^2c^4) ds \, d\epsilon \tag{3.6}$$

where $n(\epsilon) = dn/d\epsilon$ is the differential photon number density per energy interval $d\epsilon$ and $\sigma(s)$ is the cross section of the considered process. The squared center of momentum frame energy

$$s = m^2c^4 + 2\epsilon E(1 - \beta \cos\theta) \tag{3.7}$$

is determined by the angle θ between the directions of cosmic rays and photons in the observer frame, for given E and ϵ. s is maximal for head-on collisions ($\theta = 180°$) and minimal for tail-on collisions ($\theta = 0°$)

$$s_{\text{max}} - m^2c^4 = 2\epsilon E(1 + \beta) \approx 4\epsilon E \tag{3.8}$$

$$s_{\text{min}} - m^2c^4 = 2\epsilon E(1 - \beta) \approx 0 \; . \tag{3.9}$$

In general, a threshold energy $s_{\text{th}} = (mc^2 + \sum_i m_i c^2)^2$ is required for the production of new particles of mass m_i. For a given photon background with an effective maximum photon energy ϵ_{max} this translates into a threshold energy for the cosmic ray of

$$E_{\text{th}} \approx \frac{s_{\text{th}} - m^2c^4}{4\epsilon_{\text{max}}} \; . \tag{3.10}$$

Once the cosmic ray exceeds this threshold $E > E_{\text{th}}$ it starts interacting with the photon field. The expression for the interaction rate in equation 3.6 holds for both massive and massless particles. For massive particles it is more convenient to consider — instead of s — the photon energy in the cosmic-ray rest frame given by the Lorentz boost

$$\epsilon' = \Gamma\epsilon(1 - \beta \cos\theta) \tag{3.11}$$

with $\Gamma = E/mc^2$. Here the threshold photon energy is given by $\epsilon'_{\text{th}} = (s_{\text{th}} - m^2c^4)/2mc^2$. In the following a brief overview of the relevant interactions is given, while a more detailed treatment including the spectra secondary particles is presented in chapter 4.

3.4.1 Protons and Nuclei

For cosmic protons and nuclei the relevant interactions with background photons γ_b are:

- electron pair production $p + \gamma_b \longrightarrow p + e^+ + e^-$

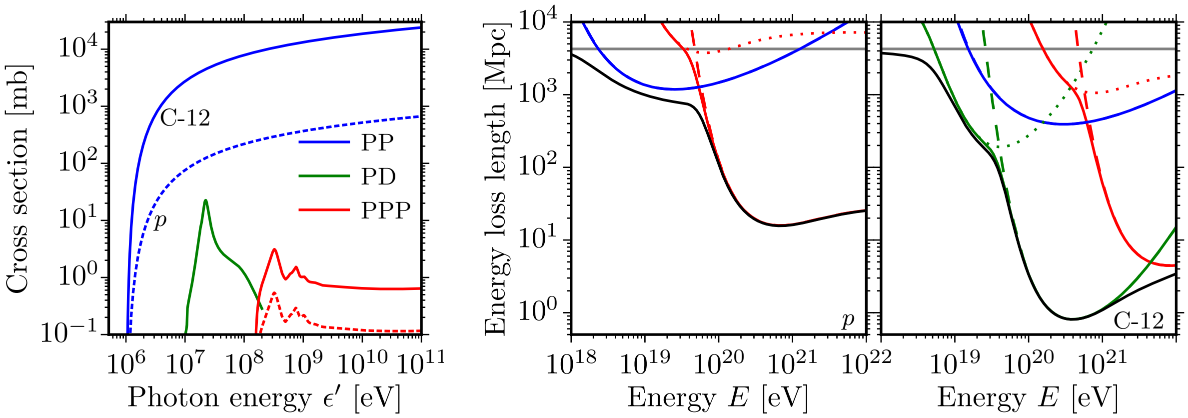


Figure 3.3: Left: Cross sections for pair production (PP), disintegration (PD) and pion production (PPP) for protons and carbon-12 as function of the photon energy in the nuclear rest frame. Right: Corresponding energy loss lengths for interactions with the extragalactic photon background at $z = 0$. Individual contributions are shown for CMB (dashed line) and EBL (dotted line). The horizontal line shows the redshift loss from the expanding universe, and the black line gives the total energy loss length. For details refer to chapter 4.

- photodisintegration, e.g. $X_A^Z + \gamma_b \longrightarrow X_{A-1}^{Z-1} + p$
- photoproduction of hadrons, e.g. $p + \gamma_b \longrightarrow p/n + \pi^0/\pi^+$

The corresponding cross sections are presented in figure 3.3 for the cases of protons and carbon-12 nuclei. Pair production has the lowest energy threshold $\epsilon'_{\rm th} \approx 2m_e c^2$ and largest cross section. Photodisintegration of nuclei is possible, once the photon energy is in the range of nuclear binding energies $\epsilon' \sim 10\,\mathrm{MeV}$ and peaks at the giant dipole resonance around 20 MeV. Pion production starts at $\epsilon'_{\rm th} = m_\pi c^2(1 + m_\pi/(2m)) \approx 145\,\mathrm{MeV}$, peaks at the excitation energy of the Δ-resonance and then slowly increases again as more mesons and baryons become available for production.

To account for the different inelasticities $\eta(E) = 1 - \Delta E/E$ of each process, the impact for cosmic-ray propagation is often assessed through the 'energy loss length' [115]

$$\lambda_{\rm loss}^{-1} \equiv \frac{1}{E}\frac{dE}{dx} = \frac{\eta(E)}{\lambda(E)}\,, \tag{3.12}$$

expressing the distance over which the energy would drop to $1/e$ of its initial value, if the attenuation length were to remain constant. For protons, pair production on the CMB is the most important interaction for energies $E = 2-50\,\mathrm{EeV}$, as seen in the right of figure 3.3. Below 2 EeV only redshift losses are relevant. Above 50 EeV pion production on the CMB renders the universe opaque to protons, which is the famous GZK effect [75, 76]. The EBL does not give rise to significant contributions, and the CRB becomes relevant only at extremely high energies of 10^{24} eV. For cosmic-ray nuclei the situation is different, with photodisintegration dominating the energy loss throughout the relevant energy range of $E/A \sim 0.3 - 300\,\mathrm{EeV}$, where A is the nuclear mass number. Below and above $E/A \sim 3 - 4\,\mathrm{EeV}$ photodisintegration

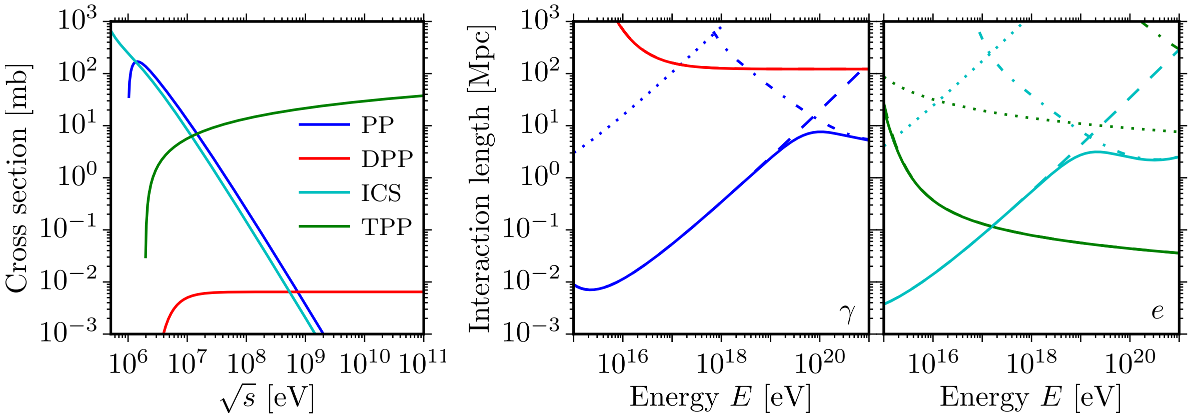


Figure 3.4: Left: Cross sections for pair production (PP) and double pair production (DPP) of photons, and inverse Compton scattering (ICS) and triplet pair production (TPP) of electrons. Right: Corresponding interaction lengths for interactions with the extragalactic photon background at $z = 0$. Individual contributions are shown for CMB (dashed line), EBL (dotted line) and CRB (dash-dotted line).

occurs mainly on the EBL and CMB, respectively. Pion production on the CMB takes over only at extremely high energies of $E/A \sim 200 - 500\,\mathrm{EeV}$. Note that for this general description it is sufficient to consider only one element, here carbon-12, because the differences between individual elements are mostly due to the different Lorentz factor $\Gamma \propto E/A$ at equal energies, which translate into energy thresholds shifted to $A \times E_{\mathrm{th}}$.

As a result of these energy losses, cosmic-ray protons and nuclei that are observed with $E > 10^{20}\,\mathrm{eV}$ can only originate from within $\sim 100\,\mathrm{Mpc}$ distance, a concept which is called the GZK horizon. All source candidates presented in section 3.1 are generally distributed along the large-scale structure of the universe, which is not homogeneous within the horizon. Therefore, anisotropies in the cosmic-ray arrival directions are expected and individual nearby sources should be identifiable, if the magnetic deflections of cosmic rays are not too large.

3.4.2 Photons and Electrons

For comparison we also consider the photon interactions of cosmic-ray photons and electrons. Electrons suffer energy losses through inverse Compton scattering $e + \gamma_b \longrightarrow e + \gamma$ and triplet pair production $e + \gamma_b \longrightarrow e + e^+ + e^-$, whereas photons undergo pair production $\gamma + \gamma_b \longrightarrow e^+ + e^-$ through the Breit-Wheeler process [116, 117] as well as double pair production. The cross sections and resulting interaction lengths are shown in figure 3.4. Compared to nuclei, electrons and photons have lower threshold energies E_{th} for interactions with the photon backgrounds. Thus, the importance of these backgrounds is shifted towards lower energies. For photons the dominating pair production occurs mainly on the CMB up to $E \approx 2 \times 10^{18}\,\mathrm{eV}$, whereafter the CRB takes over. The EBL becomes relevant at energies below $10^{15}\,\mathrm{eV}$, which is the energy range of γ-ray astronomy. For electrons the CMB is the most important photon background at all relevant energies, because at $E > 10^{17}\,\mathrm{eV}$ triplet pair production on the CMB contributes more strongly than inverse Compton scattering on the CRB.

3.5 Interactions with Matter

Apart from photon fields, interactions with the matter distribution in the universe can be relevant for cosmic-ray propagation. Hydrogen makes up more than 90% of the interstellar gas and of baryonic matter in general, thus it is most important to consider interactions between cosmic rays and thermal protons. For a rough estimate of the impact we consider a total proton-proton cross section of $\sigma_{pp}(\sqrt{s} = 5 \times 10^{14}\,\mathrm{eV}) \sim 100\,\mathrm{mb}$ [28], which corresponds to a $1\,\mathrm{EeV}$ cosmic ray hitting a proton of the interstellar medium. Cross sections for spallation of nuclei by protons are of similar magnitude (cf. [95]). The interaction probability is $p = n_H^{\mathrm{col}} \sigma_{pp}$, where n_H^{col} is the column density of hydrogen, i.e. the integrated number density along a given line of sight. Values for the hydrogen column density in the galaxy range from $n_H^{\mathrm{col}} \approx 10^{22}\,\mathrm{cm}^{-2}$ within the disk in direction of the galactic center, to $10^{20}\,\mathrm{cm}^{-2}$ towards of the galactic poles [118][1]. Thus, the probability for a hadronic interaction of a cosmic-ray proton propagating non-diffusively inside the galaxy can be estimated as $p = 10^{-5}$ to 10^{-3}, which is negligible as energy loss process, but may be relevant for the production of secondary particles. For energies $E < 10^{17}\,\mathrm{eV}$ magnetic diffusion inside the Galaxy significantly increases the propagation distance, and hadronic interactions need to be considered, as presented e.g. in [119].

Regarding the impact on extragalactic propagation the mean free path as function of hydrogen number density n_H reads

$$d_{pp} = (n_H \sigma_{pp})^{-1} \approx 3\,\mathrm{Mpc} \left(\frac{\mathrm{cm}^{-3}}{n} \right) . \tag{3.13}$$

The average baryonic density in the universe corresponds to $n_H \sim 10^{-7}\,\mathrm{cm}^{-3}$ [28], and is furthermore strongly concentrated in the galaxy clusters that form the cosmic 'large-scale structure'. The largest volume fraction of the universe consists of voids with far lower density. Hence, hadronic interactions can be neglected for extragalactic propagation. However, they have been considered for the propagation in cores of galaxy clusters (see [120]).

3.6 Magnetic Fields

Cosmic rays of charge q are scattered by the pervasive magnetic fields in the universe according to the relativistic Lorentz force $\dot{\vec{p}} = q\vec{v} \times \vec{B}$. The gyroradius of a cosmic ray is given by

$$r_g = \frac{p}{|q| B_\perp} \simeq \frac{E/c}{ZeB_\perp} \tag{3.14}$$

where $B_\perp$ denotes the magnetic field component perpendicular to the cosmic-ray heading. Note that in the ultra-relativistic limit the gyroradius does not depend on the particle mass. At equal energy electrons and protons have the same gyroradius, whereas nuclei experience Z times stronger deflections. More specifically, all magnetic field effects only depend on the

---

[1] https://heasarc.gsfc.nasa.gov/cgi-bin/Tools/w3nh/w3nh.pl

cosmic-ray rigidity $R = E/Z$. The only exception is the energy loss through synchrotron emission [121]

$$-\frac{dE}{dx} = \frac{\sigma_T}{6\pi c} B^2 \left(\frac{q m_e}{m}\right)^4 \left(\frac{E}{m_e}\right)^2 \tag{3.15}$$

where σ_T is the Thompson cross section and m_e the electron mass. However, due to the m^{-4} dependency synchrotron radiation of protons and nuclei is negligible for magnetic fields encountered during propagation and only needs to be considered for cosmic-ray electrons.

For the following discussion it is useful to first consider the propagation inside an isotropic turbulent field (cf. [98, 122]). Such a field can be characterized by its root-mean-square value $B_{\text{rms}}^2 \equiv \langle B(x)^2 \rangle$ and correlation length $l_c \equiv B_{\text{rms}}^{-2} \int_0^\infty \langle \vec{B}(0) \cdot \vec{B}(\vec{x}(l)) \rangle dl$, which describes the average distance over which the magnetic field vectors are coherent. For a turbulent field, $B_\perp$ in equation 3.14 is approximatively replaced by B_{rms}. A critical rigidity $R_c = ecB_{\text{rms}} l_c$ is given by the rigidity at which the gyroradius equals the correlation length. Three propagation regimes can now be defined:

- For $E/Z < R_c$ particles experience large deflections by the magnetic field modes of scales comparable to the gyroradius, and a spatial *resonant diffusion* occurs.
- In contrast, a particle with $E/Z > R_c$ scatters only by a small amount on each correlation length, and thus propagates in the *quasi-rectilinear regime.*
- Only after passing many coherence lengths the average deflection exceeds $\sim 60°$, which marks the onset of the *non-resonant diffusion* regime.

In the rectilinear regime after propagating over a distance $d \gg l_c$, i.e. covering multiple coherence length, the distribution of scattering angles takes the form of a normal distribution on the sphere, with zero mean and root-mean-square [123] (the reference uses $l_c' = \frac{3}{8} l_c$) of

$$\sqrt{\langle \alpha^2 \rangle} \simeq \frac{\sqrt{d\, l_c}}{2 r_g} \approx 37.5° \left(\frac{E/Z}{\text{EeV}}\right)^{-1} \left(\frac{B_{\text{rms}}}{\text{nG}}\right) \left(\frac{\sqrt{d\, l_c}}{\text{Mpc}}\right) . \tag{3.16}$$

Note that this formula describes the angular separation between initial and final heading, whereas the angular separation between initial (source) position and final heading is $1/\sqrt{3}$ of this value. Due to the deflections the propagation distance is increased compared to the linear distance d by [123]

$$\langle \Delta d \rangle \simeq \frac{l_c d^2}{24 r_g^2} \approx 116\,\text{kpc} \left(\frac{E/Z}{\text{EeV}}\right)^{-2} \left(\frac{B_{\text{rms}}}{\text{nG}}\right)^2 \left(\frac{l_c}{\text{Mpc}}\right)^2 \left(\frac{d}{\text{Mpc}}\right) . \tag{3.17}$$

At distances where this effect becomes significant, cosmic rays enter the (non-resonant) diffusive regime. As a result the cosmic-ray density is locally increased, and decreased at greater distances.

Two domains of magnetic field need to be considered for cosmic-ray propagation: the extragalactic magnetic field (EGMF) filling the space in between galaxies, and the galactic magnetic field (GMF) in the interstellar medium of the Milky Way. In the following we outline the properties of EGMF and GMF and discuss the possible implications for cosmic rays.

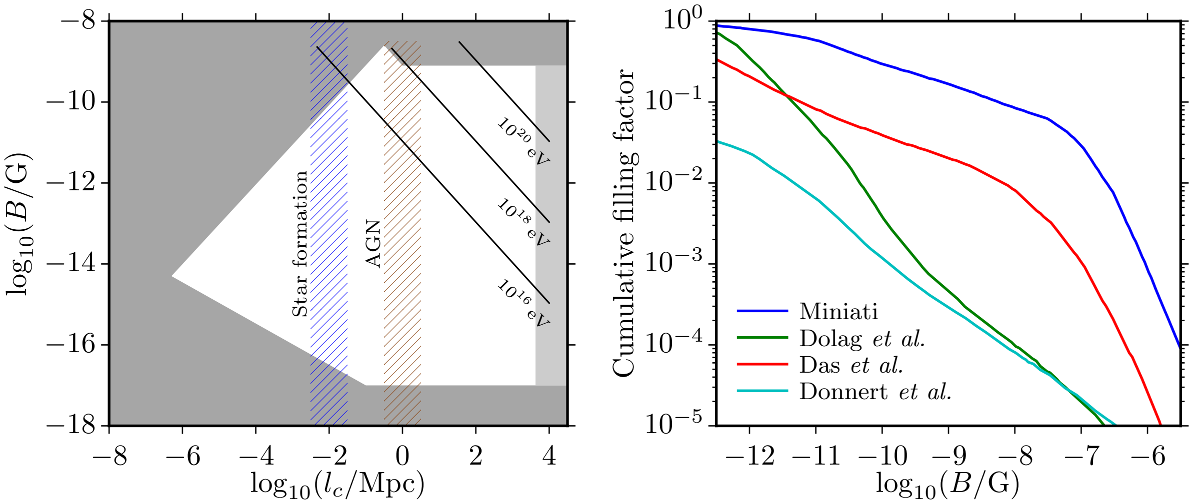


Figure 3.5: Left: Bounds on the strength B and coherence length l_c of magnetic fields in the universe (modified from [124]). The grey shaded area is excluded by observational and theoretical bounds. Hatched areas show the expected properties for a seed field of galactic origin (starbursts and AGN). Diagonal lines indicate different critical rigidities; the parameter space below each line allows for rectilinear propagation, whereas diffusion takes place above. Right: Cumulative filling factors of the EGMF strength from large-scale structure simulations, showing the volume that is filled with B equal or higher to a given value (modified from [125]).

3.6.1 Extragalactic Magnetic Field

Magnetic fields in the universe are generally believed to be produced by amplification of preexisting magnetic seed fields, see [124] for an extensive review. The amplification takes place through different types of dynamo mechanisms, and through adiabatic compression during the gravitationally driven structure formation. Both mechanisms require a seed field, the origin of which is unknown. They are either generated in the early epochs of the universe, or at the onset of structure formation. The two hypotheses lead to different predictions of the unobserved magnetic fields in the voids of the large-scale structure, where no amplification takes place. While in the former hypothesis the voids would be filled with the mostly unaltered primordial field, in the latter the only field would be due to magnetic outflows of adjacent galaxies.

Measurements of the magnetic field strength in galaxies and galaxy clusters are of the order of $B \sim (1 - 10)\mu$G. In the filaments connecting the clusters the situation is less clear and in the voids no observation has been possible so far [126]. However, the overall strength and coherence length of the EGMF are constrained by a number of observational and theoretical limits [124], as shown in figure 3.5.

The ignorance of the EGMF properties translates into considerable uncertainties on expected deflections of UHECRs. Several authors have studied the expected deflection using EGMF models from magnetohydrodynamical large-scale structure simulations ([127–129] among others). While these EGMF models all reproduce the observed magnetic field strength in galaxy clusters, the predicted properties of filaments and voids are largely different. Dolag *et al.* [130] found that for 40 EeV protons coming from 110 Mpc distance the expected deflections are small, exceeding 1° for less then 20% of the sky. Miniati *et al.* [127] predict much larger deflections of at least 45° for 60% of protons with $E > 40$ EeV coming from a structured source distribution. The prediction of Das *et al.* [129] lies in between with 75% of trans-GZK protons arriving within 15° of their source direction. These models agree that at least towards the highest energies the propagation of cosmic-ray protons should become ballistic. However, considering the observation of an increasingly heavy composition, this can be easily counteracted by the Z times larger deflection that a cosmic-ray nucleus experiences. It is therefore an open question if the extragalactic cosmic-ray flux is already at the edge of the galaxy to a large degree isotropic.

Another potentially important effect stems from the time-delay through the deflections in magnetic fields. The maximum source distance is limited by the time it takes to reach the observer, which ultimately cannot exceed the age of the universe. Thus, in a regime of diffusive propagation a magnetic horizon effect causes a flux suppression of low rigidity cosmic rays from distant sources. However, according to the 'propagation theorem' [78, 131] the cosmic-ray spectrum takes universal shape, irrespective of the mode of propagation, as long as the average source distance is small compared all other length scales, i.e. energy loss length and diffusion length. This can explained when considering that a flux suppression from distant sources is compensated by an increased flux of nearby sources, as long as the observer lies within the diffusion sphere of the nearby sources. The appearance of magnetic horizon effects thus depends on the distances to these closest sources, as well the parameters of the EGMF, and is controversially discussed. For a homogeneous magnetic field, with parameters towards the upper allowed limit, see figure 3.5, it was shown that a flux suppression can appear at $E/Z \sim 10^{18}$ eV [132], thereby affecting the energy spectrum and the composition. In contrast, when considering the structured EGMF models mentioned above, a flux suppression was found to only become relevant at rigidities below 10^{16} eV [125].

3.6.2 Galactic Magnetic Field

The galactic magnetic field is known much better than the EGMF, see [134] for a review of observational results. Radio observations allow to constrain not only its statistical properties, but also the concrete spatial structure. The GMF is typically described to be composed of a large-scale regular field whose strength and orientation in the galactic plane follows the spiral arms, as well as small-scale random components. While the regular component is subdominant in strength, it is expected to give a dominant contribution to deflections of UHECRs, as compared to the incoherent deflections in the random field [135].

Currently the most sophisticated GMF model is the JF12 model [133, 136] with further refinements in [137]. It was obtained by simultaneously fitting maps of the total and polarized synchrotron emission by galactic relativistic electrons, together with a large number of Faraday rotation measures of mostly extragalactic sources. The two measurements are complementary, because synchrotron emission is sensitive to the magnetic field perpendicular to the line of sight, while Faraday rotation probes the parallel component.

The expected cosmic-ray deflections are large, see figure 3.6. The average deflection

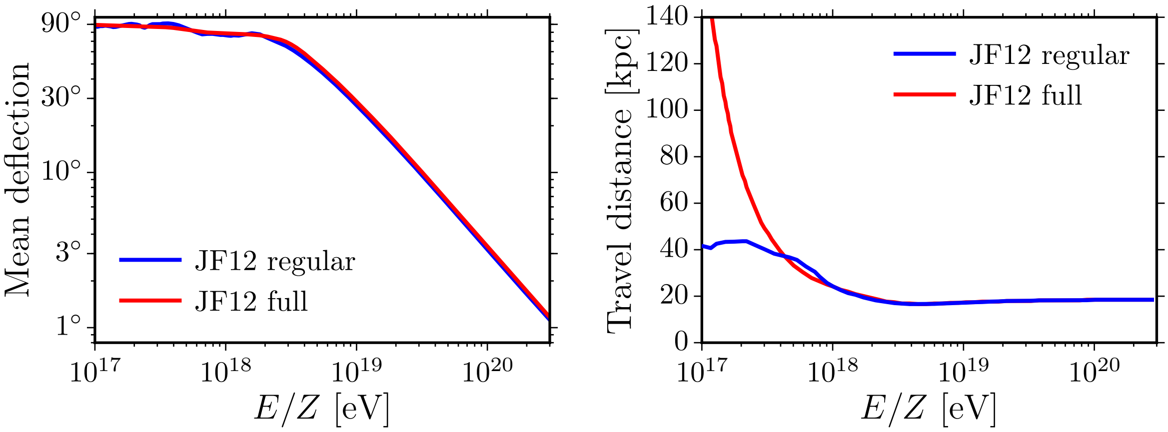


Figure 3.6: Impact of the JF12 galactic magnetic field model [133] for extragalactic cosmic rays arriving at Earth. The blue line shows the impact of the regular component and the red line both regular and random components assuming a 60 pc turbulent coherence length. Left: Mean angular deflection showing a $(E/Z)^{-1}$ dependency for the rectilinear regime and saturating at 90° for $E/Z < 2 \times 10^{18}\,\text{eV}$. Right: Mean travel distance inside the Galaxy. When considering the random field components, the travel distance below 10^{18} eV shows the expected $(E/Z)^{-2}$ dependency from equation 3.17.

is $\langle\theta(E, Z)\rangle \approx 5^\circ Z(60\,\text{EeV}/E)$, which significantly impedes the identification of cosmic-ray sources in case of a heavy composition. Still, the cosmic-ray trajectories inside the Galaxy can be reconstructed by using an adequate GMF model, and provided that the cosmic-ray charge is known. It is therefore of great interest to obtain information on the charge number of the highest energy cosmic rays.

In principle the GMF can modify not only the arrival directions, but also the energy spectrum and composition of cosmic rays by suppressing or enhancing the flux from certain directions [138]. However, a fundamental limitation to such effects derives from Liouville's theorem, which states that the phase space density is constant along particle trajectories. In general the galactic propagation of UHECRs is short compared to all energy loss lengths, as seen in figure 3.6. Also, constant magnetic fields do not alter particle energies, and modifications by the non-constant magnetic field component due to the galactic rotation are small [139]. Therefore, cosmic-ray energies can be considered as conserved during galactic propagation. In combination with Liouville's theorem it then follows that an isotropic cosmic-ray flux outside of the Galaxy is isotropically observed on Earth as well. This implies that the GMF cannot generate anisotropies and does neither alter the spectrum nor the composition, assuming an isotropic flux of extragalactic cosmic rays. Consequently, any modification of the UHECR flux during galactic propagation is limited to its anisotropic component.

Chapter 4

CRPropa 3: A Simulation Toolkit for Astrophysical Scenarios of Cosmic Rays

With the advent of the Pierre Auger Observatory and Telescope Array experiments, the past years have seen huge advancements in the quality and quantity of UHECR data. Interpreting these data in terms of concrete astrophysical scenarios requires detailed simulations of cosmic-ray propagation from the source to the observer. In these simulations the magnetic deflections of UHECRs need to be computed over several orders of magnitude in energy and length scales, ranging from hundreds of megaparsecs down to galactic kiloparsec scales. Furthermore, all relevant interactions as outlined in the previous chapter need to be considered. Secondary particles from these interactions such as photons and neutrinos should also be included in order to best exploit the available measurements in a multi-messenger approach. A comparison of predictions from these simulations with measured data then allows to reject invalid scenarios, and to identify scenarios that are compatible with all available measurements.

Fully exploiting the experimental advancements requires that the accuracy of simulations matches the increasing experimental accuracies. Recently, more effort is being invested into comparing simulations and identifying the main sources of uncertainty [140, 141]. Such comparisons are strongly facilitated with the simulation software publicly available, its source code open and well documented, and ideally openly developed. In the following we attempt an overview of the current status of UHECR simulations, noting that most of the referenced simulation tools are private codes, which are not publicly available.

The most general simulation approach is the forward propagation of individual particles using the Monte Carlo method, see e.g. [142–150]. Propagation in one spatial dimension enables simulations of the diffuse cosmic-ray flux, whereas in three dimensions the angular distribution can be simulated as well. A limitation in current three-dimensional forward simulations is that effects of the expanding universe cannot be considered, because the cosmological time of emission is not known beforehand due to deflections of charged particles in magnetic fields.

In backward propagation simulations, e.g. [151–154], the treatment is inverted. Starting from their observed properties, individual cosmic rays are backtracked to the possible sources. In three dimensions this approach significantly increases the simulation efficiency, because the chance of connecting particles to one of many sources is generally larger than the chance of hitting a single observer in a forward simulation. Backtracking is especially useful for galactic propagation where energy loss processes are negligible. Over extragalactic distances backtracking relies on suitable approximations, e.g. [151], since stochastic interaction processes in general cannot be inverted [155]. For nuclei, which fragmentate during propagation, and for secondary photons and neutrinos the approach has not yet been adopted.

Alternative to the Monte Carlo method, the diffuse cosmic-ray flux can be calculated for protons by solving the one-dimensional transport equation [156] or with a kinetic equation approach for the case of nuclei, see [157] and references therein. These solutions are analytical by definition but still require dedicated simulation codes for numerically solving a large number of integrals. The main advantage of these analytic solutions is the greatly reduced computational effort compared to the Monte Carlo method, cf. [158]. Limitations of analytic calculations are that magnetic field effects can only be approximatively treated [125, 131] and that arrival directions are obviously not simulated. The latter limitation is tackled in a related approach [159], where parameterizations for the energy losses and the magnetic deflections are combined to model the angular distributions of the cosmic-ray flux. However, the approach has not been fully adapted to nuclei, and is limited to homogeneous magnetic fields.

Similar to the backtracking method, the transport equations can also be inverted [160], which in principle allows to drop assumptions on the source spectrum and to derive it from the data. However, this inverse problem is mathematically ill defined and leads to instabilities of the derived solutions [160].

Although all of the aforementioned techniques have certain advantages, only the forward Monte Carlo method allows for comparisons with all experimental data including the arrival directions and secondary messengers, and without loss of generality. In an effort to provide an open framework for detailed simulations of UHECR scenarios the Monte Carlo code CRPropa 3 [161] was co-developed in the course of this thesis. While the implementation of the physical processes is to a large extent based on CRPropa 2, the code was completely rewritten to enable multiple new use cases and to serve as a long-term framework for future developments. For instance, while CRPropa 3 is currently focused on the propagation of extragalactic cosmic rays above 10^{17} eV, the framework allows to reach far into the energy range of galactic cosmic rays, see e.g. [162]. CRPropa 3 is publicly developed[1] and attention was given to documenting every aspect of the implementation such as the preprocessing of experimental cross sections, in order to make the simulation as transparent and reproducible as possible. In the following, the main simulation method is described (section 4.1) as well as the treatment of magnetic fields (section 4.2) and cosmological effects (section 4.3). Then follows a detailed description of all relevant physical processes for cosmic-ray nuclei (sections 4.4, 4.5 ,4.6 and 4.7), and finally the propagation of non-hadronic secondary particles is outlined (section 4.8).

4.1 Simulation Method and Software Design

CRPropa uses the Monte Carlo method to compute trajectories of individual cosmic rays in the forward tracking approach. Each particle is injected with specified initial properties at a source position and propagated in discrete steps of variable size Δs. In each step, stochastic processes such as photonuclear interactions or nuclear decay are modeled to occur with a probability $p = 1 - e^{-\Delta s/\lambda}$ according to the respective interaction length λ. The propagation proceeds until a user defined break conditions is met, e.g. when the particle energy falls short of the energy range of interest. The particles that reach the observer position represent the simulated

---

[1] https://github.com/CRPropa

cosmic-ray flux to be compared to measurements.

Technically, all simulation aspects in CRPropa 3 are separated into independent modules that correspond to physical processes, boundary conditions, observers, output etc. These modules simply provide a method to update the cosmic-ray properties according to the propagation step and the module's purpose. The simulation itself is a user-defined sequence of modules, where each iteration through the entire sequence corresponds to one propagation step. Since there are no direct dependencies between modules, any combination of modules can in principle be selected. With this flexible setup simulations can be easily adapted to different use cases such as the extragalactic or galactic propagation. Furthermore, the framework can be easily extended with new processes or alternative implementations without modifying any part of the existing code.

To enable this modular structure the cosmic-ray particle class serves as the single interface between the modules. It carries information on e.g. its initial and current particle state — particle type, energy, position and direction — as well as the current propagation step and cosmological time point. In the observer frame all cosmic rays are modeled in the ultra-relativistic limit $\beta = 1,\ E = pc$. This is allowed since even for heavy cosmic-ray nuclei the Lorentz factors of $\Gamma \simeq 10^8 (E/10^{17}\,\mathrm{eV})/A$ are large. For the particle type we make use of the PDG Monte Carlo numbering scheme [163] which provides a standardized way to represent elementary particles and hadrons. Nuclear codes are given in this scheme as ten digit numbers $\pm 100ZZZAAAI$, where Z and A are the charge and mass number and the isomer level I indicates nuclear excitations. Nuclear masses are approximated from the atomic weights m_atomic as present in the NIST database [164] by subtracting the electron masses, $m(A, Z) = m_\mathrm{atomic}(A, Z) - Z \cdot m_e$, thereby neglecting electron binding energies ($\sim$ keV) compared to electron masses and nuclear binding energies ($\sim$ MeV). For completeness the nuclear masses of extremely unstable isotopes which are experimentally not available are approximated as $m = A \cdot u - Z \cdot m_e$. The masses of non-hadronic cosmic rays, i.e. electrons, positrons and neutrinos, are currently not needed for any calculation during the propagation and are therefore neglected.

Efficient Monte Carlo propagation depends on dynamically adjusting the step size to accommodate for varying simulation conditions, e.g. making smaller steps in regions of strong magnetic deflections or at energies where interactions rates are large. Therefore, a central aspect of the modular structure is a bidding system that enables all modules to bid for the maximum allowed size of the next step. The lowest bid is then selected as the step size for the next iteration of the module sequence. This way the propagation always proceeds in the largest possible steps that still ensure the accuracy of the individual processes. Concretely, the following constraints exist:

- When propagating in magnetic fields the estimated directional error has to be smaller than a user defined value, as described in section 4.2.
- Observers and other spatial boundary conditions limit the step according to the closest distance to the respective surface in order to prevent the particle from overshooting.
- Continuous energy loss processes limit the step to 10% of the energy loss length. The default limit can be changed by the user.

- Stochastic interactions limit the step to 10% of the current interaction length, which corresponds to an interaction probability of $p = e^{-0.1} < 10\%$. Again, the default limit can be changed by the user.

The step size constraint of stochastic interaction modules also ensures that the order of modules in the simulation sequence is unimportant. In the worst case of two modules with the same interaction length winning the bid for the maximum step, the first interaction process preempts the second one with a probability of merely $p^2 < 1\%$. Likewise, since the modules do not take into account an eventual previous interaction of another module in the same step, the interaction frequency of two concurrent processes is overestimated in this worst case by $\sim p^2/2 < 0.5\%$, where the factor 2 accounts for the average distance until the first interaction. Cross-testing with CRPropa 2 showed that in practice these effects do not lead to noticeable differences.

The interaction lengths of the photonuclear processes — electron pair production, pion production and photodisintegration — are numerically calculated[2] using equation 3.6 for the CMB and the extragalactic background light models of [103, 108, 109, 111–113] shown in figure 4.2. Differences between these models were found to be significant for the photodisintegration of nuclei [141]. Since the computation of photonuclear interaction rates are too expensive to be performed in each propagation step, the interaction rates are tabulated for a range of nuclear Lorentz factors and interpolated during runtime. Table 4.1 lists the implemented processes together with their tabulation range, considered interaction model and possible secondary particles. Secondary particles can be created with arbitrary direction in the cosmic-ray rest frame. In the observer frame the direction is strongly beamed towards that of the primary cosmic ray, with angular separations of generally less than $\theta \sim 1/\Gamma$. Therefore all secondary particles are modeled to have the same velocity vector as the primary cosmic ray.

Process	$\log_{10}\Gamma$	Model	Secondaries
Pair production	6-14	analytic	$e^\pm$
Disintegration	6-14	TALYS/PSB/Kossov	$p, n, d, t, {}^3\mathrm{He}, \alpha\ (\gamma)$
Pion production	6-16	SOPHIA	$p, n, \bar{p}, \bar{n}, e^\pm, \nu_e, \bar{\nu}_e, \nu_\mu, \bar{\nu}_\mu, \gamma$
Nuclear decay	—	experimental data	$p, n, \alpha, e^\pm, \nu_e, \bar{\nu}_e\ (\gamma)$

Table 4.1: Implemented processes for cosmic-ray nuclei. Tabulated interaction rates are available for the indicated range of nuclear Lorentz factors. References to the considered interaction models are given in the corresponding sections together with a detailed description of the implementation. The last columns lists the possible secondary particles. Secondary photons in brackets are subdominant at energies $> 10^{17}$ eV and therefore currently neglected.

CRPropa 3 is written in C++ and interfaced to Python using SWIG [3]. This allows to set up and steer simulations in a high level scripting language while all computations are performed with the underlying C++ code. The SWIG interface enables cross-language polymorphism,

---

[2] https://github.com/CRPropa/CRPropa3-data

[3] `www.swig.org`

which can be used to extend a CRPropa simulation directly from the Python script that runs it. The user can for example write a custom simulation module in Python to be used in combination with the existing C++ modules. While the Python usage is the advised steering mode, backwards-compatibility to the XML steering of CRPropa 2 is provided as well.

Cosmic ray propagation is a perfectly parallel task as interactions between cosmic rays are negligible. Current multicore processors can therefore be adequately utilized by simply running multiple simulation instances in parallel. However, for better memory utilization CRPropa 3 implements shared-memory multiprocessing using OpenMP [4]. This allows to use higher resolution magnetic fields and matter distributions in the simulation. The parallelization occurs on the level of the module sequence with a dynamic distribution of cosmic rays among the available threads. As the simulation modules are stateless, only a single instance of each module is needed. The resulting speedup is limited by the number of critical sections that are not thread safe and can only be executed by one thread at a time. The critical section with the largest impact is the call to the external library SOPHIA [165] when simulating photopion interactions. Thus, the speedup depends on the frequency of these interactions. In a typical simulation the speedup is close to optimal for up to eight threads. Thus, on a standard computing cluster with 2 GB RAM per core, a CRPropa 3 simulation can efficiently run on eight cores in parallel, providing 16 GB RAM for simulation data.

4.2 Propagation in Magnetic Fields

The equation of motion of a particle with charge q, position $\vec{x}$ and direction $\vec{\beta}$ in a magnetic field $\vec{B}$ can be written as

$$\frac{d}{dt}\begin{pmatrix}\vec{x}\\ \vec{\beta}\end{pmatrix} = \begin{pmatrix}c\vec{\beta}\\ qc^2(\vec{\beta}\times\vec{B})/E\end{pmatrix} . \tag{4.1}$$

In CRPropa this ordinary differential equation is solved using the Cash-Karp method [166], which is an implicit Runge-Kutta method of 4th order accuracy, embedded in a 5th order accurate computation to estimate the error. We require that the direction error satisfies $|\Delta\vec{\beta}| < r$, where the default target error $r = 10^{-4}$ can be user defined. If the estimated error is larger than the target error, the step is rejected and the calculation repeated with a smaller step size. If it is less than the target error, the step is accepted and the next step size is increased proportional to the ratio of target and estimated error. This way the computation dynamically adapts to the strongly varying gyroradii in structured magnetic fields. An additional maximum step size can be set to prevent overshooting sharply localized regions of high magnetic field strength that are much smaller than the optimal step size in the surrounding volume. However, magnetic fields from large scale structure simulations such as [130] exhibit relatively smooth transitions towards the highest field strengths in the cores of galaxy clusters, so that a low maximum step size is typically not required. For uncharged particles, i.e neutrons, photons and neutrinos, a simple rectilinear propagation is performed.

---

[4] `www.openmp.org`

Multiple ways to represent magnetic fields are implemented. The most generic option is the representation by a three-dimensional grid lattice, where the field is evaluated through a trilinear interpolation of the eight neighboring grid points. For the often needed case of homogeneous turbulent fields, cf. section 3.6, we provide a method to instantiate random realizations of a specified power spectrum. The field is first constructed in wave vector space $\vec{B}(\vec{k})$ by drawing for each grid point a random amplitude from its power spectrum, a random complex phase, and a random orientation with $\vec{k} \cdot \vec{B} = 0$ to satisfy $\vec{\nabla} \cdot \vec{B} = 0$. The field is then transformed to real space with a Fast Fourier Transform method [167]. Since only the real components are of interest here, we employ an in-place complex to real Fourier transform, thereby maximizing the achievable resolution of the grid for given computing resources. By construction the method results in periodic boundary conditions. Thus, the field can be periodically repeated for covering larger spaces without producing unphysical discontinuities at the boundaries. Since the extragalactic magnetic field is not homogeneous, we devised and implemented a method to modulate the aforementioned homogeneous turbulent fields with a scalar field grid that describes the spatial distribution of the field strength, e.g. from LSS simulations. Here, using a high-resolution turbulent field in combination with a large modulation field of lower resolution can provide a suitable approximation to simultaneously resolve the small-scale structure within galaxy clusters, as well to represent the cosmic web of filaments and voids on large scales.

Additional magnetic field implementations allow to access the highly compressed magnetic field representations that are commonly used by LSS simulations, such as the adaptive mesh refinement [168] and smooth particle [169] schemes. However, the computational cost of these magnetic field lookups is greatly increased compared to that on a (modulated) grid, which increases the computing time of propagation simulations substantially.

For the galactic field, the models by Pshirkov *et al.* [170] and by Jansson & Farrar [133, 136] (JF12) are implemented, the latter including the regular and both the anisotropic random ('striated') and isotropic random ('turbulent') field components. For the turbulent component we use the above described homogeneous turbulent field on a periodically repeated, high-resolution grid lattice, and modulate it according to the spatial dependency of the B_{rms} strength as specified in [133]. In addition to the Pshirkov and JF12 models, any combination of axisymmetric or bisymmetric spiral fields in the galactic disc, toroidal halo fields, and (modulated) turbulent field components can be considered.

In the Galaxy, where energy losses are negligible, backtracking simulations can be realized simply by inverting the equation of motion 4.1 with a charge reversal $q \rightarrow -q$. Hence, the forward propagation of extragalactic protons in the Galaxy can be simulated in CRPropa 3 by propagating anti-protons from Earth to the edge of galactic space, as shown in figure 4.1. Large sets of these trajectories can serve as lookup tables to efficiently simulate the effect of the GMF on extragalactic cosmic rays. The information on the start and end positions of the trajectories can be neglected here, because Earth (the Milkyway) can be considered point-like compared to galactic (extragalactic) distances. The effect of the GMF on charged cosmic rays then represents a transformation $(E/Z, \vec{\beta}_{\text{extragalactic}}) \rightarrow (E/Z, \vec{\beta}_{\text{Earth}})$ which for binned rigidities and directions takes the form of a matrix vector multiplication. This so

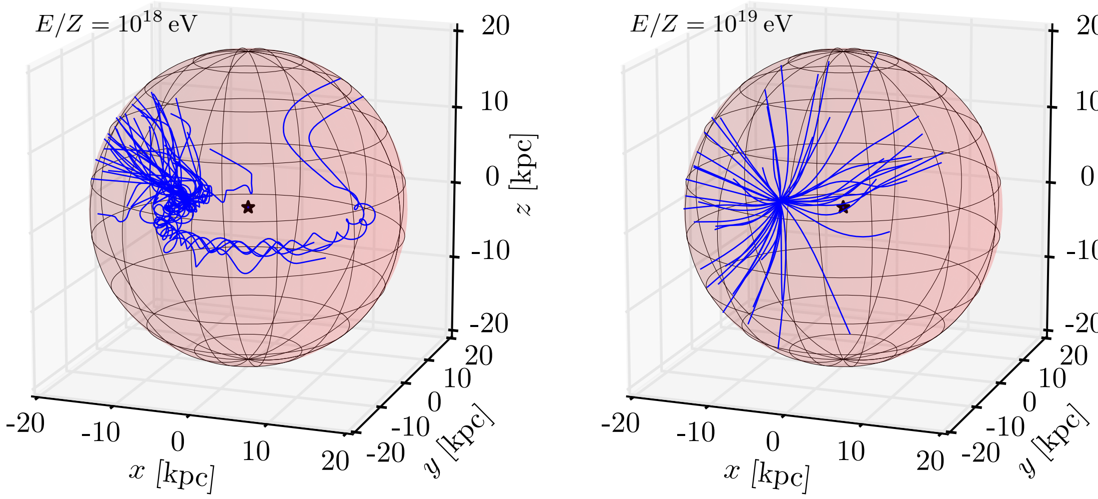


Figure 4.1: Backtracking simulations of cosmic rays of rigidity $E/Z = 10^{18}$ eV (left) and 10^{19} eV (right) in the regular component of the JF12 GMF model. Each blue line shows the trajectory of one cosmic ray propagated from Earth to the galactic border (red sphere) which is modeled at a distance of 20 kpc from the galactic center (black star).

called 'lensing' technique [159] is an efficient approach to simulating the galactic deflection of extragalactic cosmic rays, and is now implemented in the CRPropa framework. GMF lenses produced with CRPropa 3 were already used in multiple analyses, cf. [65, 100, 171–174].

4.3 Implementation of Cosmology

CRPropa 3 uses a comoving coordinate system for all spatial coordinates, such as the source, the cosmic ray or the observer position. In this system the proper vector between two comoving points $\vec{r}(z) = a(z)\vec{\chi}$ factors into a constant comoving vector $\vec{\chi}$ and a scale factor $a = (1+z)^{-1}$, expressing the cosmological expansion. This facilitates defining the general case of a comoving source distribution, i.e. sources without a proper motion. Moreover, solving the equations of motion 4.1 in comoving coordinates implicitly takes into account the adiabatic dilution of magnetic fields $\vec{B}(z) = (1+z)^2 \vec{B}_{\mathrm{com}}(z)$ due to the expanding universe, where the comoving magnetic field $\vec{B}_{\mathrm{com}}$ is constant in the absence of magneto-hydrodynamical amplifications.

Cosmic rays are injected at a certain cosmological time, corresponding to a redshift z. This redshift is stored as an attribute of the cosmic ray. In each propagation step a redshift module decrements this redshift $\Delta z = -H(z)/c\,\Delta s$ according to the comoving step size Δs and using a small step approximation. The adiabatic energy loss of the particle along one step reads $\Delta E/E = \Delta z/(1+z)$. Here, the resulting energy loss over the entire propagation is exact, independent of the performed step sizes. In all interaction modules the redshift information can be used to adapt to the expanding universe, e.g. to translate the comoving step Δs into a proper step $\Delta d = \Delta s/(1+z)$ on which physical interactions take place. Not setting the initial redshift of the cosmic ray (default is $z = 0$) implicitly turns off all cosmological effects.

For cosmic rays above a few EeV the most important cosmological effect is the evolution of the photon backgrounds. In the case of the CMB, the photon number density per unit energy and per unit proper volume evolves passively as [175]

$$n(\epsilon,\, z) = (1+z)^2\, n\left(\epsilon(1+z)^{-1},\, z=0\right)\ . \tag{4.2}$$

Note that the differential $d\epsilon(z) = (1+z)\, d\epsilon(z=0)$ reduces the exponent of the volume factor $(1+z)^3$ by one. Substituting into equation 3.6 leads to a convenient scaling of the interaction rates

$$\lambda^{-1}(E,z) = (1+z)^3 \lambda^{-1}\left(E(1+z)\,,\, z=0\right) \tag{4.3}$$

which allows to use the (precomputed) interaction rates at $z=0$ for any redshift. The EBL on the other hand is determined by the sum of galactic luminosities integrated over the entire age of the universe, hence its redshift dependence is non-trivial. However, by assuming the spectral *shape* of the EBL to be constant in time, the redshift dependency can be absorbed in an evolution factor $e(z)$ yielding the approximate scaling relation

$$n(\epsilon,z) \approx (1+z)^2\, e(z)\, n\left(\epsilon(1+z)^{-1},\, z=0\right) \tag{4.4}$$

$$\lambda^{-1}(E,z) \approx (1+z)^3\, e(z)\, \lambda^{-1}\left(E(1+z),\, z=0\right) \tag{4.5}$$

$$e(z) = \begin{cases} 1 & \text{CMB} \\ \frac{1}{(1+z)^3} \frac{\int n(\epsilon,z) d\epsilon}{\int n(\epsilon,0) d\epsilon} & \text{EBL} \end{cases} \tag{4.6}$$

which is exact for the case of the CMB. The spectrum at $z=0$ and scaling factor $e(z)$ of the implemented EBL models is shown figure 4.2. Compared to the true redshift dependency of the interaction rates this introduces an error on the simulated cosmic-ray spectrum of less than 1% in typical simulations [161].

In the absence of magnetic deflections the propagation distance and source distance are identical. In that case, which is best simulated in one dimension, the initial redshift of the cosmic ray is known *a priori*. However, once magnetic fields are considered the propagation distance of charged particles is known only after the simulation. First neglecting the cosmological effects during propagation, the resulting propagation distance can be used for *a posteriori* corrections, such as the redshift loss. However due to the redshift dependency of the photonuclear interactions, as well as the interplay between distance dependent energy losses and energy dependent magnetic time delay, any *a posteriori* correction can only be of approximative nature. A general solution to consider both magnetic and cosmological effects is a four-dimensional simulation in three spatial dimensions and in time. To this end cosmic rays can be propagated starting at a certain time given by e.g. a source luminosity evolution, and only registered if they arrive at the observer within a small redshift window $\pm\Delta z_{\mathrm{obs}}$ around $z=0$. The symmetrical window enables a faster convergence with decreasing window size, compared to a non-symmetrical window between 0 and Δz_{obs}. Here, negative redshifts corresponding to future time points are explicitly allowed by extrapolating the corresponding quantities. Although the computational effort significantly increases through the decreased detection efficiency, this direct approach allows to develop and validate suitable approximative methods, e.g. [176].

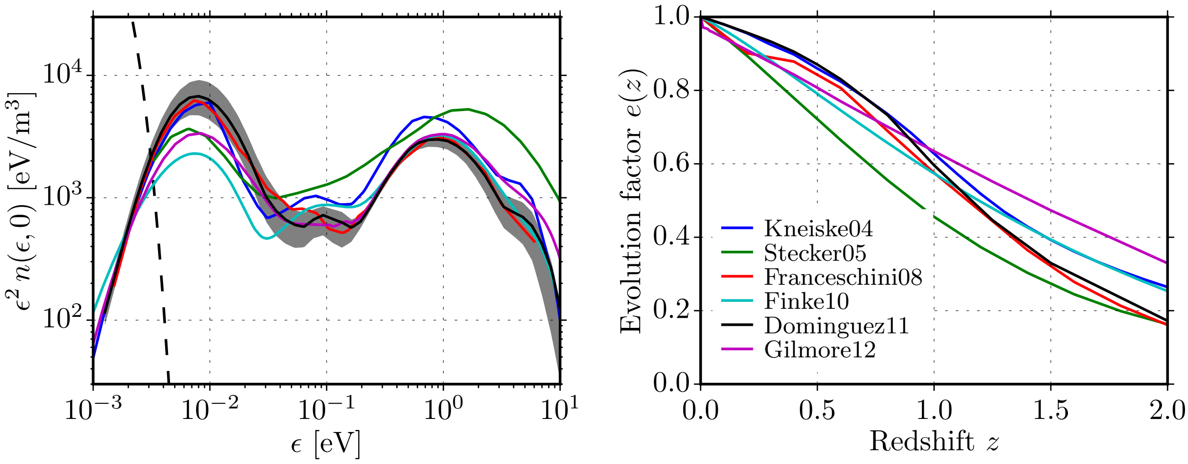


Figure 4.2: Left: Spectral number density at $z = 0$ (scaled with ϵ^2) of the implemented EBL models [103, 108, 109, 111–113] with the legend given on the right. The Dominguez *et al.* model is shown together with its 68% statistical uncertainty interval. The dashed line shows the CMB for comparison. Right: EBL evolution factor indicating the evolution of the total photon number density with redshift.

4.4 Electron Pair-Production by Hadrons

Electron pair-production through the Bethe-Heitler process [177] is the photonuclear interaction with lowest energy threshold and the highest cross section. However, the inelasticity of the process is small, $\eta < 10^{-3}$ [121], so that electron pair-production is generally treated in the continuous energy loss approximation. Here, instead of considering discrete stochastic interactions, the cosmic-ray energy is reduced by $\Delta E = dE/dx \Delta s$ in each propagation step according to the step size Δs and the energy loss rate dE/dx. For an isotropic photon field this energy loss rate is given by [178]

$$-\frac{dE}{dx} = \alpha r_0^2 Z^2 \frac{m_e}{m} \int_2^\infty n\left(\frac{\kappa}{2\Gamma}\right) \frac{\phi(\kappa)}{\kappa^2} d\kappa \tag{4.7}$$

where α is the fine-structure constant, r_0 the classical electron radius, $\kappa = 2\Gamma\epsilon/(m_e c^2)$ the maximum photon energy in the collision frame in units of electron mass m_e, and $n(\kappa)$ the spectral number density of the photon field. The dimensionless function $\phi(\kappa)$ describes the photon energy dependence of the energy loss rate. We use its parametrization given in [179]

$$\phi(\kappa) = \begin{cases} \frac{\pi}{12} \frac{(\kappa-2)^4}{\sum_{i=1}^4 c_i (\kappa-2)^i} & \text{for } \kappa < 25 \\ \frac{\kappa \sum_{i=0}^3 d_i \ln(\kappa)^i}{1-\sum_{i=1}^3 f_i/\kappa^i} & \text{for } \kappa > 25 \end{cases} \tag{4.8}$$

with coefficients $c_i = (0.8048, 0.1459, 0.00137, -3.879 \times 10^{-6})$, $d_i = (-86.07, 50.96, -14.45, 8/3)$ and $f_i = (2.910, 78.35, 1837)$. From equation 4.7 one can see that for nuclei of equal Γ the loss rate scales only weakly with $Z^2/A \approx Z/2$ (since $A \approx 2Z$ for $Z = 2-26$) compared to protons.

Contrarily, at equal energy the Lorentz factor is smaller by $1/A$ so that the threshold energy increases $E_{\text{th}} = A \times E_{\text{th},p}$.

For the produced e^+e^- pairs we directly sample from their energy distribution in the observer frame which was tabulated in CRPropa 2 following the description in [180]. In order to connect the continuous energy loss of the primary cosmic ray with the discrete energies of the secondary particles we apply the following algorithm: In each propagation step a number of e^+e^- pairs is created with energies $E^{\text{pair}} = E^{e^+} + E^{e^-}$ until the total energy $\sum_i E_i^{\text{pair}}$ exceeds the energy loss ΔE of the primary cosmic ray. The last pair $i = n$ is randomly accepted with a probability $p = (\Delta E - \sum_i^{n-1} E_i^{\text{pair}})/E_n^{\text{pair}}$. This method ensures that the number and spectrum of electron pairs is independent of the step size, and that energy is on average conserved over many propagation steps. Finally, due to the strong Lorentz boost the electron and positron directions in the observer frame are assumed to be parallel to the primary cosmic ray.

4.5 Photodisintegration of Nuclei

Photodisintegration is the process, in which a nucleus enters an excited state by absorbing a high energy photon and deexcites through the evaporation of mostly single protons, neutrons or α-particles. Each channel is described by an energy dependent *exclusive* cross section, e.g. $\sigma_{\gamma,n}(\epsilon')$ for the emission of a single neutron $X_A^Z + \gamma \longrightarrow X'^Z_{A-1} + n$. The sum of all exclusive cross section equals the total photoabsorption cross section

$$\sigma_{\text{abs}}(\epsilon') = \sigma_{\gamma,n} + \sigma_{\gamma,p} + \sigma_{\gamma,pn} + \sigma_{\gamma,\alpha} + ... + \sigma_{\gamma,\gamma} + \sigma_{\gamma,F} \tag{4.9}$$

where (γ, γ) denotes (in-)elastic scattering and (γ, F) photofission, which only becomes important for very heavy elements $A \gg 56$ [155] and is therefore not considered in the following. Energy dependent branching ratios can be defined with respect to the photoabsorption cross section, e.g. $BR_{\gamma,n}(\epsilon') = \sigma_{\gamma,n}/\sigma_{\text{abs}}$. The main contribution to the photoabsorption cross section is the giant dipole resonance (GDR), in which protons and neutrons oscillate collectively and with opposite phase. The giant dipole resonance is seen as a broad peak in the photoabsorption cross section at energies $\epsilon' = 10 - 30\,\text{MeV}$. The energy range between 30 and 150 MeV is attributed to the 'quasi-deuteron' effect which can be phenomenologically thought of as an interaction between the photon and a proton-neutron pair inside the nucleus. At higher energies $\epsilon' > 145\,\text{MeV}$ photoproduction of hadrons becomes dominant, which is modeled in CRPropa through a separate process as described in the next section.

Despite large collections of photonuclear experimental data, e.g. [181], nuclides with measured cross sections and branching ratios make up for only a small fraction of the potentially relevant cosmic-ray nuclides. Therefore, a general modeling is required to simulate photodisintegration during cosmic-ray propagation. For this purpose we use the state-of-the-art nuclear reaction code TALYS [182, 183], which allows to simulate any type of light particle-induced nuclear reaction for nuclides with $A \geq 12$. The photodisintegration cross sections derived with TALYS were found to accurately describe the available photonuclear data [184]. Concretely, we used TALYS version 1.6 to compute cross sections for all 169 isotopes in the range $A = 12 - 56$, $Z \leq 26$ with a lifetime of $\tau > 2\,\text{s}$, based on the evaluated lifetimes described in

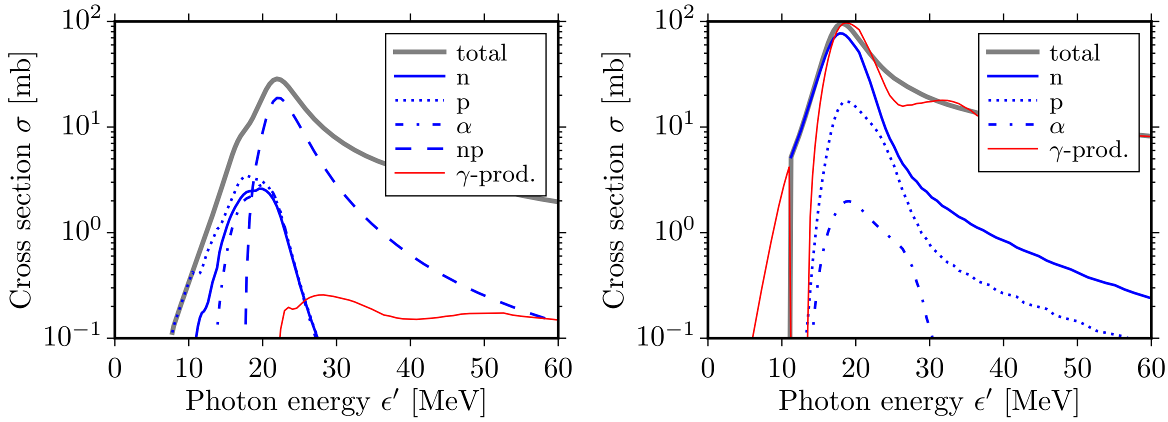


Figure 4.3: Photodisintegration cross sections for ^{14}N (left) and ^{56}Fe (right) as computed by TALYS. The total cross section is shown in grey, blue lines give the main exclusive cross sections. For ^{14}N the main channel is the (γ, np) disintegration into a ^{12}C nucleus, whereas for 56 $^{\mathrm{F}}$e the (γ, n) channel is dominant. The total photo-production cross section (red line) summarizes the elastic scattering mainly below the disintegration threshold, and the photon emission that can accompany disintegration.

section 4.7. For shorter lifetimes photodisintegration can be safely neglected, since the decay length $\lambda_{\mathrm{decay}} < \Gamma 5 \times 10^{-9}$ pc is much smaller than their disintegration length at all energies. We tested that the current version 1.8 computes identical photodisintegration cross sections. We set TALYS to use a generalized Lorentzian (Kopecky-Uhl) to parametrize the so called E_1 transition, which accounts for most of the giant dipole resonance [183]. As done in [184, 185] we use the GDR parameters given in the 'Atlas of Giant Dipole Resonances' [181], if available, instead of the default values in TALYS. The complete list of GDR parameters, as well as an exemplary steering card is given in appendix A.1. The cross sections are computed for incident photon energies of $\epsilon' = 2 - 200\,\mathrm{MeV}$ in logarithmic steps of $\Delta \log_{10}(\epsilon'/\mathrm{MeV}) = 0.01$. This ensures that the individual threshold energies which have significant impact on the interaction rates are well captured by the tabulation. Branching ratios are taken into account for every channel that can be separately output by TALYS, namely with a simultaneous evaporation of up to eight nucleons in form of protons, neutrons, deuterons, tritons, He-3 nuclei and α particles. In practice, a large fraction of the resulting more than 25000 channels is of negligible impact. Thus, to increase performance, channels are removed that have a branching ratio of less than 5% at every photon energy in the tabulated range, and the branching ratios of the remaining channels are scaled up accordingly. For the total cross section, however, all channels are considered, in order to prevent a systematic underestimation of the cross section.

TALYS should not be used for very light nuclides $A < 12$ [183]. For these nuclides we use the following parameterized cross sections instead, as described in [143, 186]:

- ^{2}H, ^{3}H, ^{3}He, ^{4}He, ^{9}Be are taken from [155]. Here, the cross sections for ^{3}H and ^{3}He are scaled by 1.7 and 0.66 and for ^{9}Be the parametrization is refitted to better reproduce the experimental data (cf. [186]).

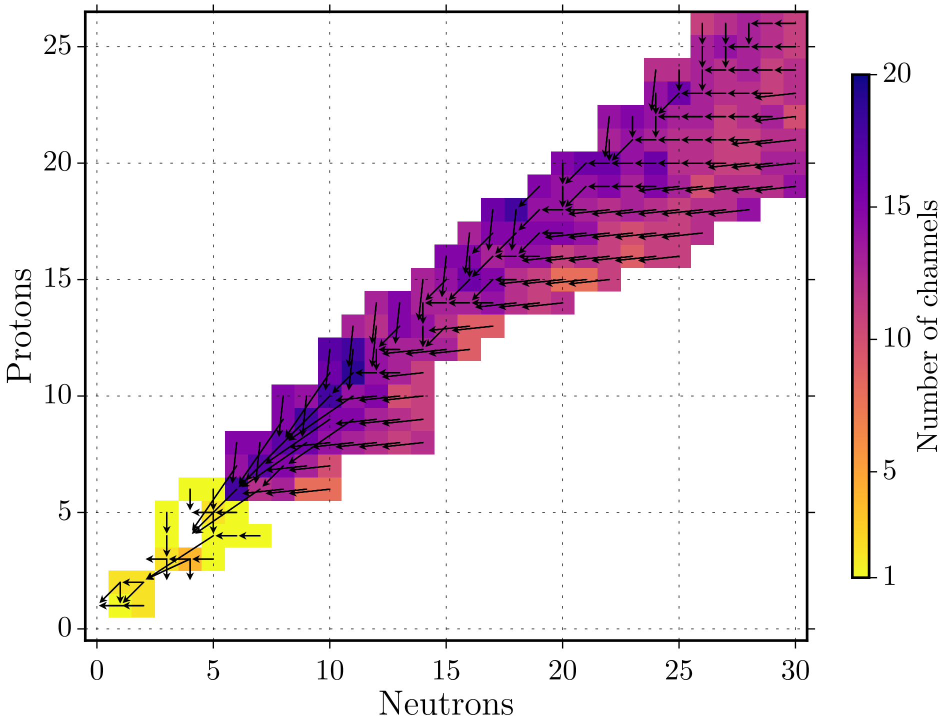


Figure 4.4: Photodisintegration: Overview of the adopted TALYS model (see text for details). The number of disintegration channels after the thinning procedure is color coded. Arrows indicate the main disintegration channel in the dominating GDR range. For light nuclides with $A < 12$ all disintegration channels are shown. Unstable nuclides with a lifetime $\tau < 2\,\mathrm{s}$ are not considered for disintegration since $\lambda_{\mathrm{decay}} \ll \lambda_{\mathrm{disintegration}}$. Several disintegration paths end at ^{8}Be, which decays into two α particles.

- ^{6}Li, ^{8}Li, ^{7}Be, ^{10}Be, ^{11}Be, ^{8}B, ^{10}B, ^{11}B, ^{10}C, ^{11}C are taken from [187]. Since branching ratios are not modeled in this reference, we assume the loss of one proton (neutron) if the neutron number is $N < Z$ ($N > Z$). For $N = Z$ the loss of one neutron or proton is modeled with equal probability. This is motivated by the higher nuclear binding energies in the valley of stability $Z \approx N$, and by the low relevance of multi-nucleon evaporation for light elements.

- For ^{7}Li an interpolation of experimental data is used (cf. [186]).

The resulting photodisintegration model, called TALYS model hereafter, comprises cross sections for 183 nuclides and 2200 channels. An overview of the large number of possible disintegration paths is given in figure 4.4, showing for each nuclide the number of channels and the main channel in the GDR region.

Alternatively to the above described TALYS model, we implemented the option of modeling photodisintegration for all nuclides $A = 1 - 56$ using the parametrization from Kossov [187], which is also used in the general particle-matter simulation toolkit GEANT4 [188]. While the Kossov model describes both photodisintegration and photoproduction of hadrons, we consider only the former part up to the pion production threshold here. Also, since the parametrization only models the total cross section, branching ratios for the individual evaporation channels are taken from TALYS. This mixed procedure is possible, because the photoabsorption and the consequent evaporation of nucleons are phenomenologically separate processes [183].

As a third option, we implemented the photodisintegration model of Puget, Stecker & Bredekamp [107] including its modifications described in [189] (hereafter called PSB model). In the PSB model a single element Z is considered for every mass number A, i.e. all nuclides of equal mass number are considered the same. The total cross section in the GDR region for $2 < \epsilon'/\mathrm{MeV} < 30$ is approximated as a sum of two Gaussians fitted to an early set of experimental data. Branching ratios are estimated for single and double-nucleon emission, thereby not differentiating between protons, neutrons and deuterons, and are assumed to be energy-independent. At energies above 30 MeV the total cross section is assumed constant and with constant branching ratios for the emission of 1-15 nucleons. Reference [189] describes a modification, where the constant threshold energy $\epsilon'_{\mathrm{th}} = 2\,\mathrm{MeV}$ is replaced by the threshold given by the mass difference between residual and initial nuclei for each possible single and double-nucleon channel, $\epsilon'_{\mathrm{th}}/c^2 = n_p m_p + n_n m_n + m(Z - n_p, N - n_n) - m(Z, N)$, where n_n and n_p denote the number of neutron and protons, respectively. Despite its simplicity the PSB model it is still often used in cosmic-ray propagation, e.g. in [146, 149]. In CRPropa 3, we implemented the PSB model in the following way: Since it cannot be guaranteed that the particle stays on the single photodisintegration path, cross sections and disintegration channels are evaluated for same set of nuclides as for the TALYS and Kossov models. The evaporation of nucleons is assumed to take place in form of individual nucleons, where protons or neutrons are selected based on the stability of the residual nucleus, and in form of α-particles. The nuclides with mass numbers $A = 5 - 8$ are not considered in the PSB model. For completeness, we assume instant photodisintegration by single nucleon emission for these nuclides.

A comparison of the aforementioned models to the available photodisintegration measurements is given in appendix A.2. We find that all three models give a reasonable description of the photoabsorption cross sections, with TALYS performing best in predicting the total interaction rate. Differences between the TALYS and PSB model are most pronounced in the individual channels. While in the PSB model the single nucleon evaporation dominates for all elements, the TALYS model predicts a large contribution of the α channel, which leads to a more efficient disintegration of cosmic-ray nuclei. This difference was found to be a main uncertainty in simulating the propagation of cosmic-ray nuclei [141].

For each model the interaction rates on the CMB and each EBL model are calculated using equation 3.6 and tabulated for cosmic-ray Lorentz factors $\log_{10}(\Gamma) = 6 - 14$. The same is done for the exclusive interaction rates in order to obtain the branching ratios, e.g. $BR_{\gamma,n}(\Gamma) = \lambda_{\gamma,n}(\Gamma)/\lambda(\Gamma)$. During a propagation step Δs, a cosmic-ray nuclei disintegrates with a probability $p = 1 - \exp(-\Delta s/\lambda)$. If an interaction occurs, a channel is selected according to the branching ratios at the given Lorentz factor.

Nucleons are evaporated with kinetic energies of 1-2 MeV in the mainly contributing GDR range [190]. Since these energies are small compared to the nuclear masses, the evaporated particles and the residual nucleus are well approximated to be produced at rest in the nuclear rest frame. In the observer frame primary and secondary particles thus carry the energies $E_i = \Gamma m_i$ and propagate in direction as the incident cosmic ray.

There are two sources of secondary photons in the context of photoabsorption which are currently neglected: (in-)elastic photon scattering on the nuclear structure, and γ-decay of the residual nucleus, which is generally left in an excited state after having disintegrated [183]. Based on TALYS computations, which are exemplarily shown in figure 4.3 for the cases of nitrogen-14 and iron-56, elastic scattering is only relevant below the disintegration threshold energy. The γ-decays following a photodisintegration are more relevant and have been studied for cosmic rays of lower energy as a source of TeV γ-rays [191]. For UHECR nuclei the typical $E'_\gamma \sim$ MeV decay photons contribute to the photon flux at $E_\gamma \sim (\Gamma/10^9)(E'_\gamma/\mathrm{MeV})10^{15}$ eV and thus may be relevant, depending on the considered UHECR and photon energies.

4.6 Pion Production

The photohadronic cross sections for protons and neutrons are experimentally well known [28], which allows to calculate the corresponding interaction rates on the CMB and EBL using equation 3.6.

Pion production starts at a threshold photon energy of $\epsilon'_{\mathrm{th}} \approx 145\,\mathrm{MeV}$. At these energies the background photon electromagnetically scatters with a virtual charged meson, giving it enough energy to materialize [165],

$$p + \gamma \longrightarrow n + \pi^+ .$$

Scattering on virtual uncharged mesons is suppressed, hence an isospin change $p \leftrightarrow n$ of the cosmic ray is strongly favored. Resonance production becomes dominant at the nucleon excitation energies, most importantly the Δ resonance. The resulting strong decay conserves the isospin leading to the channels

$$\begin{aligned} p + \gamma &\longrightarrow \Delta^+ \longrightarrow p + \pi^0 \\ p + \gamma &\longrightarrow \Delta^+ \longrightarrow n + \pi^+ \end{aligned}$$

with branching ratios of 2/3 and 1/3, respectively. At higher energies inelastic multiparticle production, also of heavier particles, becomes important. Charged and uncharged pions are an important source of UHE neutrinos $\pi^+ \longrightarrow e^+\nu_e\bar{\nu}_\mu\nu_\mu$ and UHE photons $\pi^0 \longrightarrow \gamma\gamma$, respectively.

Accurately modeling the individual photohadronic processes and the energy distribution of the produced particles is a substantial task. For this purpose we use the dedicated event generator SOPHIA [165]. SOPHIA models the energy loss and potential isospin change of the primary particle and calculates the energy of secondary particles in the observer frame. Since the lifetime of pions and muons is negligibly short for cosmic-ray propagation, their instant decay is also handled by SOPHIA. All secondary particles from SOPHIA can be considered in CRPropa: photons, electrons $e^\pm$, neutrinos $\nu_e, \bar{\nu}_e, \nu_\mu, \bar{\nu}_\mu$, and occasionally baryon pairs $p/n, \bar{p}/\bar{n}$ at very high photon energies $\epsilon' > 2\,\mathrm{GeV}$.

For nuclei the experimental situation is similar to that of photodisintegration (cf. [187, 192]) thus requiring a general modeling of the cross sections for all potentially relevant cosmic-ray nuclides. Here we employ the same effective scaling as in CRPropa 2, i.e. approximating nuclei by a superposition of the proton and neutron cross sections

$$\sigma_{A,Z} = 0.85 \times \left(Z^{\zeta} \sigma_p + N^{\zeta} \sigma_n \right) \tag{4.10}$$

with $\zeta = 2/3$ for $A \leq 8$ and $\zeta = 1$ for heavier nuclei [143]. The collision energy is large compared to nuclear binding energies, thus the photon is assumed to interact with a single free nucleon carrying a fraction E/A of the cosmic-ray energy. Therefore the interaction rate of a nucleus is given by

$$\lambda_{A,Z}^{-1}(E) = 0.85 \times \left(Z^{\zeta} \lambda_p^{-1}(E/A) + N^{\zeta} \lambda_n^{-1}(E/A) \right) , \tag{4.11}$$

where λ_p^{-1} and λ_n^{-1} are the interaction rates for protons and neutrons, respectively. In case of an interaction, a proton or neutron of the nucleus is selected with a probability according to its contribution to the interaction rate. The residual nucleus loses the interacting nucleon and the corresponding fraction of its energy while keeping a constant Lorentz factor. SOPHIA is then called to simulate the photon-nucleon interaction for calculating the final nucleon energy and for generating secondary particles.

From equation 4.11 it can be seen that for nuclei the threshold energy increases by a factor A compared to protons, $E_{\mathrm{th}}^{A} = A E_{\mathrm{th}}^{p}$. Consequently, pion production is subdominant compared to photodisintegration up to extremely high energies of about $A \times (200 - 500)$ EeV. Therefore, pion production of nuclei is a second order effect for which the above described approximative treatment is adequate, unless the cosmic-ray sources have much stronger acceleration capabilities than generally assumed. Note that for the same reason most other simulation codes [144–147, 150] neglect pion production of nuclei altogether.

4.7 Nuclear Decay

Photo-nuclear interactions often result in unstable cosmic-ray nuclei. Therefore, nuclear decay is a relevant process for the cosmic-ray nucleus, changing its type and energy, and for the creation of secondary cosmic-ray hadrons, electrons, photons and neutrinos. In CRPropa we consider $\beta^{\pm}$ and α decays, as well as proton and neutron dripping, including their branching ratios in the presence of multiple decay channels. Isomeric transitions leading to UHE photons are currently not considered. Decay channels and lifetimes are obtained from the ENSDF database through the NuDat2 web interface [193] for all unstable nuclides in the range $Z = 1 - 26$. The database is cleansed of isomeric transitions, entries with missing lifetime or decay channels, and of duplicate channels.

The decay length of a nuclide is given by $\lambda_{\mathrm{decay}} = \Gamma c \tau$ according to its lifetime τ and Lorentz factor Γ. In the case of β^{+} decay the measured lifetimes in the ENSDF database include the effect of electron-capture (EC), in which an electron of the atomic shell is caught by a proton in the nucleus ($e^{-} + p \longrightarrow n + \nu_e$). Cosmic rays, however, are fully ionized and cannot decay through electron-capture. Therefore the decay rates of β^{+} emitters need to be

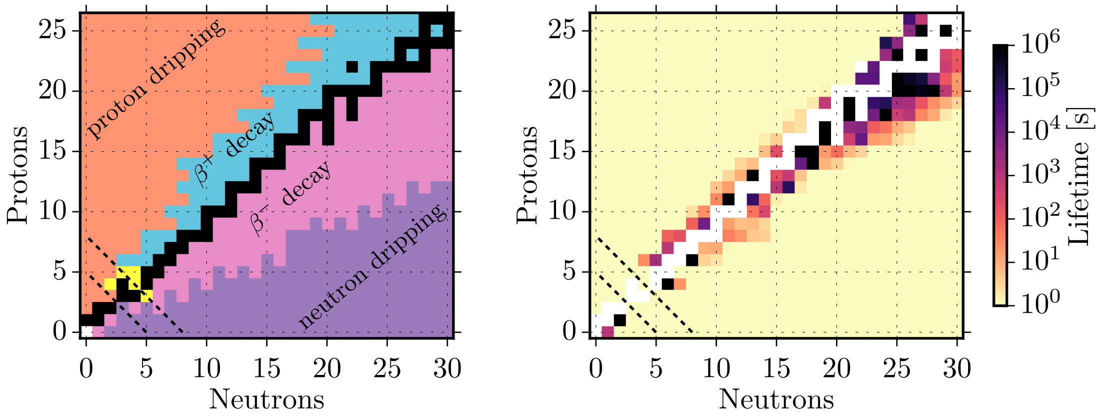


Figure 4.5: Nuclear decay as implemented in CRPropa 3. Left: Chart of the main decay modes. Stable nuclides are shown in black, revealing the stability gaps at $A = 5$ and 8 (dashed lines). Nuclides above (below) the valley of stability undergo β^+ (β^-) decay and proton (neutron) dripping further away from the valley. Nuclides with dominant α-emission are shown in yellow. Right: Mean lifetime. For reference, a 100 s lifetime corresponds to a $\sim$ 10 kpc mean travel distance for a typical UHECR Lorentz factor of $\Gamma = 10^{10}$.

corrected for the missing electron-capture contribution, as described in the following. The Q-factors of the two reactions are

$$Q_{\beta^+} = (\Delta m - m_e)c^2$$
$$Q_{\mathrm{EC}} = (\Delta m + m_e)c^2$$

with $\Delta m = m_{A,Z} - m_{A',Z'}$ the mass difference of initial and residual nucleus. Up to the squared matrix elements the decay rates are proportional to the available phase space of final states, cf. [194]. Hence, the β^+ decay rate is proportional to the integral over the β-spectrum, which is given by the three-body decay spectrum (neglecting the small Coulomb correction)

$$f(E_e) = E_e \sqrt{E_e^2 - m_e^2 c^4} (\Delta m c^2 - E_e)^2 \ , \tag{4.12}$$

whereas in electron-captures the neutrino carries the entire energy $E_\nu = Q_{\mathrm{EC}}$. Explicit expressions for the rates can be found in [194], allowing to formulate the ratio

$$\frac{\tau_{\beta^+}}{\tau_{\mathrm{EC}}} = \frac{\pi^2}{2} \left(\frac{Z c \hbar}{a_0} \right)^3 \frac{(\Delta m + m_e)^2 c^4}{\int_{m_e c^2}^{\Delta m c^2} f(E_{e^+}) dE_{e^+}} \tag{4.13}$$

where a_0 is the Bohr radius and Δm is the only variable. Since decay rates are additive, τ_{β^+} can be obtained from the experimentally measured lifetime τ of unionized nuclei using

$$\tau_{\beta^+} = \left(1 + \frac{\tau_{\beta^+}}{\tau_{\mathrm{EC}}} \right) \tau \ . \tag{4.14}$$

This is done for all β^+/EC emitters in the ENSDF database. If the β^+ decay of a nuclide is energetically not allowed ($\Delta m < m_e$), the channel is removed altogether. The resulting

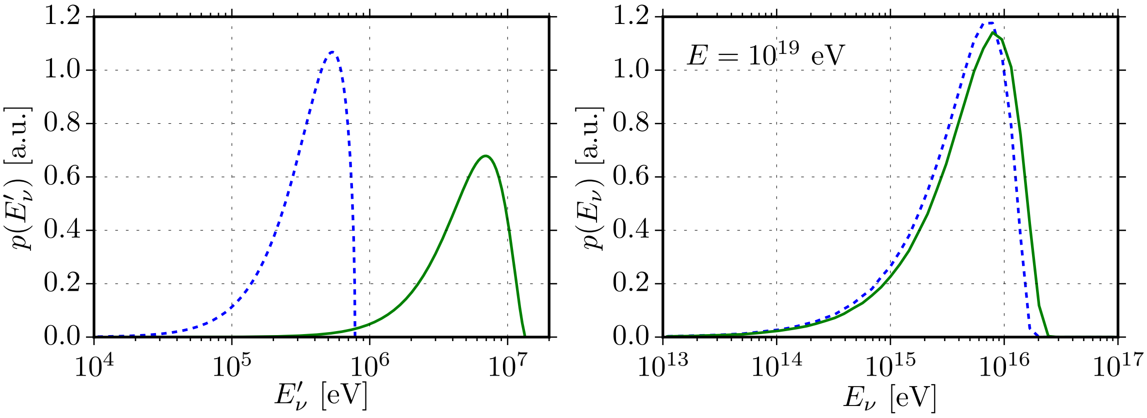


Figure 4.6: Neutrino energy spectrum from the β^- decay of free neutrons (blue dashed line) and of $^{12}\mathrm{B} \longrightarrow {}^{12}\mathrm{C}$ (green). Left: In the decay rest frame. Right: In the observer frame for a primary cosmic-ray energy of 10^{19} eV. The decay neutrinos carry about 10^{-3} of this energy.

tables of lifetimes and main decay channels are shown in fig. 4.5. All nuclides not listed in the ENSDF database (as stable or unstable) are modeled to instantly undergo neutron or proton dripping, depending on whether the nuclide features an excess of protons or neutrons. This is done in order to prevent a situation, where repeated photopion interactions result in a highly unstable nuclide that, for the lack of experimental data, would be otherwise considered stable.

Since the kinetic energy released in nuclear decays is small compared to nucleon masses $Q/m_p c^2 < 10^{-3}$ we assume that hadronic decay products α, p, n and the residual nucleus are produced at rest in the decay frame. The energy E_i of hadronic decay products in the observer frame is then simply proportional to their mass m_i, $E_i = \Gamma m_i c^2$.

In $\beta^\pm$ decays the electron/positron energy E'_e in the decay rest frame is sampled from the three-body decay spectrum in equation 4.12 and the neutrino is attributed the remaining energy $E'_\nu = \Delta m c^2 - E'_e$, as seen in figure 4.6. Electron/positron and neutrino decay back-to-back in an arbitrary direction $\pm\cos\theta$ with respect to the cosmic-ray heading. The energies in the observer frame are then given by the Lorentz boost

$$E_e = \Gamma(E'_e - p'_e \cos\theta) \tag{4.15}$$

$$E_\nu = \Gamma E'_\nu (1 + \cos\theta) \ , \tag{4.16}$$

where the neutrino mass is neglected. In the observer frame, the energies of leptonic secondaries from nuclear decays typically do not exceed those from neutron decays due to the smaller Lorentz boost. This is seen in figure 4.6 for the case of $^{12}\mathrm{B} \longrightarrow {}^{12}\mathrm{C}$. In both cases the secondary leptons carry a fraction $\sim 10^{-3}$ of the cosmic-ray energy.

4.8 Secondary Messengers

Neutrinos and photons produced during UHCER propagation can provide important information on the cosmic-ray universe. In general the CRPropa 3 framework allows to propagate any kind of subatomic particle in the ultra-relativistic limit. Redshift losses, magnetic deflections and simulation constraints, such as the detection by an observer object, work for any type of cosmic-ray particle. Thus, a physically adequate simulation of electron, photon and neutrino propagation simply requires the inclusion of the corresponding interaction processes.

For neutrinos, being only weakly interacting particles, this is trivial: Their interactions with matter are only relevant during the passage through compact objects, which is outside of the current simulation scope of CRPropa. Neutrino interactions with the cosmic photon and and neutrino backgrounds during extragalactic propagation can be neglected as well [195]. Photonuclear processes that lead to cosmic-ray neutrinos are the decay of charged pions in pion production and the β-decay of nuclei and neutrons. This covers all relevant production channels in the context of cosmic-ray propagation.

In contrast to neutrinos, photons and electrons are subject to frequent interactions with background photons γ_b (cf. section 3.4), resulting in a kind of electromagnetic cascade: Photons create electrons mainly via pair production $\gamma\gamma_b \to e^+e^-$ and in second order via double pair production $\gamma\gamma_b \to e^+e^-e^+e^-$. Electrons produce further electrons via triplet pair production $e^\pm\gamma_b \to e^\pm e^+e^-$, and photons via inverse Compton scattering $e^\pm\gamma_b \to e^\pm\gamma$ and synchrotron radiation in magnetic fields.

To calculate the electromagnetic cascade CRPropa 3 uses the specialized codes DINT [121] and EleCa [196]. DINT calculates the observed energy spectra of photons and electrons based on the transport equation approach which is particularly suited for calculating the cascades down to lower energies $E < 10^{17}$ eV where the particle number strongly increases. In contrast, EleCa provides a Monte Carlo tracking of the individual particles in one dimension with a stochastic interaction treatment. An implementation of the electron and photon interaction processes from EleCa directly in CRPropa is currently under way [197].

The highest energy photons result from the decay of neutral pions, whose production is in CRPropa taken into account through the SOPHIA code. Additional production channels for photons that are currently not considered are the inverse Compton scattering of background photons by UHECR nuclei, and the radiative decay of excited nuclei following photodisintegration or nuclear decays. Production channels for electrons are fully considered with the creation in pair production, β-decay and the decay of charged pions in pion production.

Chapter 5

The Pierre Auger Observatory

The Pierre Auger Observatory [198] is the largest experiment for the investigation of UHECRs to date. It is located in Argentina, in the Province of Mendoza between $35.0° - 35.3°$ latitude and $69.0° - 69.4°$ longitude on a generally flat area close to the Andes. The average altitude of $\sim 1400\,\mathrm{m}$ corresponds to an atmospheric depth of $875\,\mathrm{g/cm^2}$ for vertical cosmic-ray showers, which is ideal for ground observation of UHECR showers shortly after their maximum shower development. Further advantages of the location are the dry weather conditions and a small amount of light pollution from civilization.

The layout of the observatory in figure 5.1 shows the two main detector components. The 'surface detector' (SD) comprises 1660 water Cherenkov tanks spread over an area of $\sim 3000\,\mathrm{km}^2$. The 'fluorescence detector' (FD) consists of 24 fluorescence telescopes observing the night sky from the perimeter of the surface detector array. These detectors have different advantages and different systematic uncertainties. The surface detector measures the electromagnetic shower particles on ground level with a 100% duty cycle, while the fluorescence detector observes the longitudinal shower development in clear moonless nights, thus providing a calorimetric energy measurement. Combining both detectors in a hybrid design allows for cross-calibration, cross-checking and measurement redundancy. In addition to the SD and FD baseline detectors a number of extensions are operated at the Pierre Auger Observatory. The 'High-Elevation Auger Telescopes' (HEAT) and a more densely instrument SD array (750 m spacing) have extended the minimum sensitive energy of the baseline detectors from $10^{18}\,\mathrm{eV}$ down to $10^{17}\,\mathrm{eV}$.

In the following the two baseline detectors are described with their main characteristics. The Pierre Auger Observatory recently provided key measurements of the cosmic-ray energy spectrum and X_{max} distributions up to the highest observed energies. In the present analysis these measurements are compared to simulations down to the detector level in order to constrain astrophysical models of the UHECR origin. To this end the spectrum and X_{max} measurements are reviewed together with descriptions of the detector response.

5.1 Fluorescence Detector

The baseline fluorescence detector [198, 200] consists of 24 fluorescence telescopes, overlooking the array from four sites. Each telescope has a $30° \times 30°$ field of view, so that the six telescopes in one FD building cover 180° in azimuth. The minimum and maximum elevation are 1.5° and 31.5° above the horizon. Figure 5.2a depicts a FD building, while the schematics of an individual FD telescope are shown in figure 5.2b.

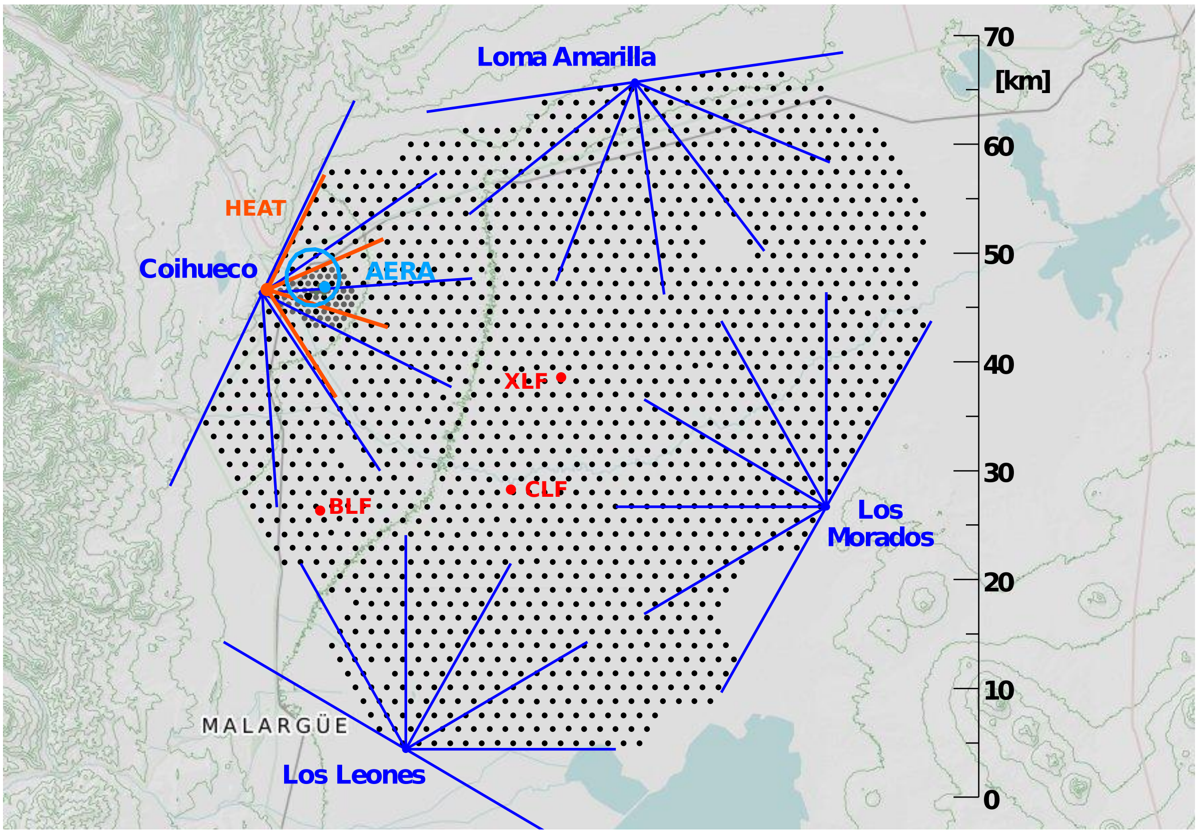
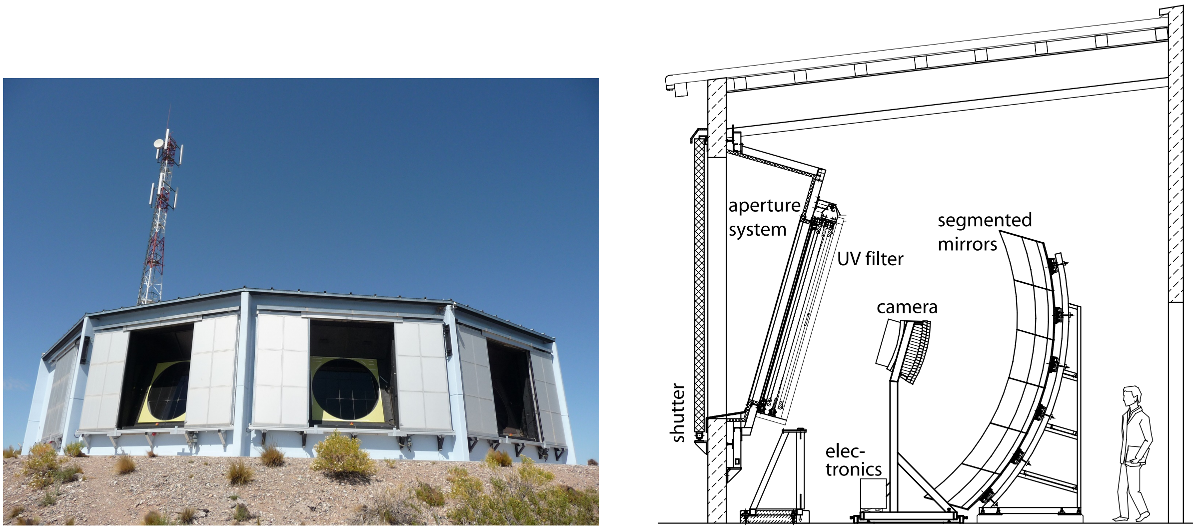


Figure 5.1: Layout of the Pierre Auger Observatory near the city of Malargüe, Argentina [199]. Black dots mark the surface detector stations separated by 1500 m. The fluorescence detector telescopes are located at four sites around the surface detector with their viewing angles indicated as blue lines. Three additional telescopes of the HEAT extension are located at the Coihueco site (orange lines), overlooking the more densely instrumented 750 m array of the surface detector, as well as the AERA radio detector and other extensions.

The telescope design is based on a Schmidt telescope, where the light passes through a corrector lens and is reflected by a large spherical mirror onto a curved camera inside the optical system. Schmidt optical systems allow for a large field of view without coma and with low spherical aberration. The aperture system consists of a 3.8 m^2 large circular opening covered by a UV pass filter glass window. The filter improves the signal-to-noise ratio by blocking light outside of the fluorescence light range and closes off the climate controlled FD building to the outside. A simplified annular lens mounted to the inside of the aperture serves as the corrector lens to reduce spherical aberration. The camera is located at the focal plane of the 13 m^2 large spherical mirror. It is composed of 22 × 20 pixels of 1.5° angular size. Each pixel consists of a 4.56 cm wide hexagonal photomultiplier tube (PMT) with a Winston cone attached for collecting the light from the dead space between PMTs.

The PMT signals are digitized with 10 MHz frequency (100 ns binning), which is fast enough to resolve air showers. Three local triggers are used. The first is a threshold trigger for each



(a) FD building (b) FD telescope

Figure 5.2: (a) FD building at Los Leones with a communications tower behind the building. Shutters to three of the six telescopes are opened for maintenance. (b) Schematic view of a fluorescence telescope with its main components.

pixel, where the threshold level is continuously adjusted to keep the pixel trigger rate at $\sim 100\,\mathrm{Hz}$. The second level trigger searches for tracks of at least five pixels in length within a camera. The third level trigger filters out high multiplicity signals that are caused by lightning and cosmic-ray muons, and it rejects randomly triggered pixels far off a shower track. This trigger is sent to the central data acquisition system together with an estimate of the region of impact in order to query the corresponding SD stations for their signal traces.

The observation of fluorescence light from air showers requires dark background conditions. Thus the FD operates only at nights with little direct or scattered moonlight and when weather conditions are suitable for operation. On average the FD operates 17 nights per month with a total duty cycle of $\sim 15\%$.

Event reconstruction: The longitudinal shower development is observed as a track of triggered PMT pixels in one or several fluorescence telescopes of a FD site. This track, in combination with the impact position and time from at least one SD station, determines the shower geometry. In order to measure the shower development the signal of all contributing FD pixels is converted to the photon flux for each time bin. With help of the reconstructed geometry, every time bin is projected to piece of the path length on the shower axis at slant depth X. The energy deposit on this path length is obtained from the measured photon flux by considering the pixel aperture, the light attenuation in the atmosphere, the fluorescence yield, and the contributions from direct and scattered Cherenkov light. The resulting profile of deposited energy dE/dX is then fitted with a Gaisser-Hillas function, see equation 2.3. The depth of the maximum energy deposit X_{max} is an observable that is sensitive to the primary cosmic-ray mass. The total energy deposit provides a near calorimetric measurement of the cosmic-ray energy, except for the small fraction of 'invisible energy' carried away by neutrinos and high energy muons. This fraction is estimated from shower simulation and ranges from 8.5 to 17% at 1 EeV and from 7 to 13.5% at 10 EeV [201].

The fluorescence detector has an energy resolution of $\leq$ 10% [202]. The energy scale uncertainty is 14%, and consists of contributions from fluorescence yield (4%), atmospheric attenuation (5%), FD calibration (10%), longitudinal profile reconstruction (6%), invisible energy (1%) and the energy scale stability (5%) [203].

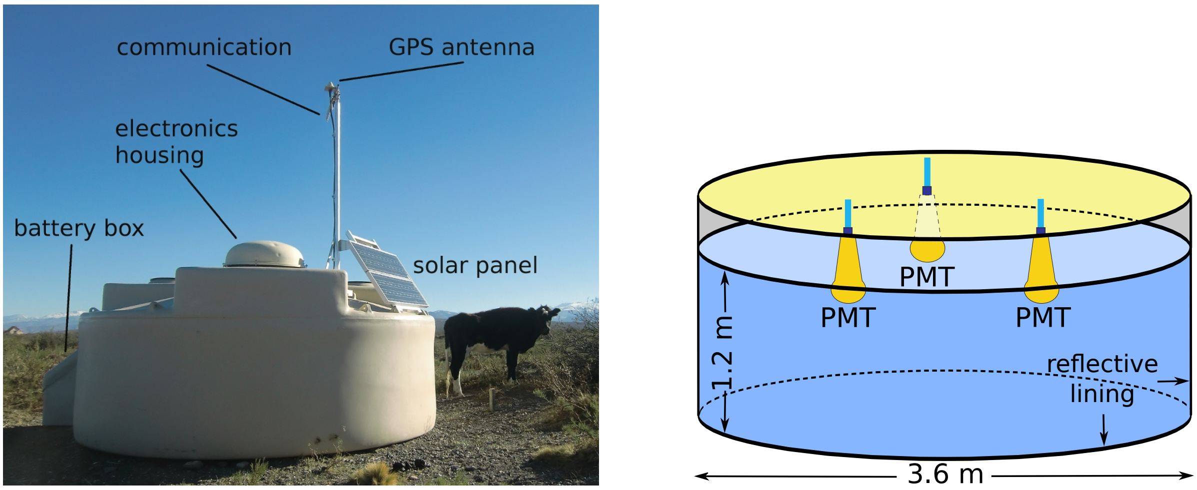
5.2 Surface Detector

The surface detector [198, 204] comprises more than 1600 detector stations arranged on a triangular grid with 1500 m spacing and covers an area of 3000 km^2. An additional 49 stations make up the more densely instrumented 750 m array covering 24 km^2. Due to the large extent of the array the stations are designed to operate autonomously.

The station layout is depicted in figure 5.3. A cylindrical water tank of 10 m^2 surface area holds 12,000 liters of ultra-pure water in a sealed liner with a reflective inner surface. Three 9 inch photo-multiplier tubes (PMTs) are installed at the top of the tank and look downwards through transparent windows in the liner to detect the Cherenkov light when relativistic charged particles pass through the water. Due to the 1.2 m height of the water volume the stations are also sensitive to shower photons initiating electron-pair cascades inside the water and to very inclined showers. A radio transceiver and antenna mounted on top of the SD tank are used for communication with the central data acquisition systems at the observatory campus via one of the four FD sites. The PMTs, station electronics and radio transceiver require an average of 10 W electrical power which is provided by two solar panels of 55 W nominal power and two 105 Ah lead-acid car batteries. The stations are designed to withstand the severe environmental conditions for the lifetime of the observatory, estimated at 20 years. This includes temperature variations from $-15\,^\circ\mathrm{C}$ to $50\,^\circ\mathrm{C}$ as well as strong UV radiation, dust and wind speeds up to 160 km/h. The polyethylene tanks are colored beige to blend in the landscape and do not need fencing against livestock, cf. figure 5.3a. As a result of the robust design, the efficiency of data collection is better than 98%.

The PMT signals are digitized with 40 MHz sampling rate and converted into units of 'vertical-equivalent-muons' (VEM). A VEM corresponds to the signal in one PMT that is produced by a single muon vertically passing the detector center. The maximum signal before saturation corresponds to a $\sim$ 100 EeV cosmic ray at 500 m distance to the shower core. With the GPS receiver installed at each station the signal data is synchronized to GPS time within 10 ns RMS [198].

The PMT signal traces are monitored for cosmic-ray events using three on-line (T1-T3), as well as two off-line trigger levels (T4, T5). The first trigger level (T1) either requires a signal above a threshold of 1.75 VEM in all three PMTs, or a 'time-over-threshold' signal above 0.2 VEM for a minimum duration of 325 ns in at least two PMTs. The former condition corresponds to the large but short signal from the dominant muonic component in very inclined showers, while the latter applies to the numerous signals from the electromagnetic component



(a) SD station (b) SD detector

Figure 5.3: (a) SD station with description of its main components and a cow for size comparison (from [205]). (b) Schematic view of the water Cherenkov detector inside a SD station (modified from [3]).

that is dominant in vertical showers. Two additional T1 modes are active since June 2013, which both reduce the influence of muons in the time-over-threshold trigger. The second trigger level (T2) applies only to the T1 1.75 VEM threshold trigger and requires a coincident signal of all three PMTs above 3.2 VEM. All T2s send their timestamps to the central data acquisition system, which forms the global trigger level (T3) by searching for coincidently triggered neighboring stations. In case of a T3, the system requests the PMT signal traces of the T2 triggered stations and all stations up to six stations away. The trigger efficiency of the SD increases with energy and reaches 100% at 3 EeV. Two additional off-line triggers are used to ensure a good data quality. The T4 trigger rejects stations that are randomly triggered by low energy showers based on the time coincidence between adjacent detector stations within the propagation time of the shower front. Finally, the T5 trigger ensures that the shower footprint is well contained. For the highest quality events the station with the highest signal is required to be surrounded by six operating stations (6T5), i.e. a full hexagon.

Event reconstruction: The arrival times and deposited energy in the SD stations allow to reconstruct the shower geometry and energy. First, to reconstruct the shower geometry, a spherical wave model is fitted to the SD signal times, which gives a virtual origin and start time of the shower. The shower impact point on the ground is estimated by fitting a model of the lateral distribution to the station signals, cf. equation 2.1.

$$S(r) = S(r_{\mathrm{opt}}) \left(\frac{r}{r_{\mathrm{opt}}}\right)^{\beta} \left(\frac{r + r_1}{r_{\mathrm{opt}} + r_1}\right)^{\beta+\gamma} . \tag{5.1}$$

Here, r is the distance to the shower axis, $r_1 = 700\,\mathrm{m}$ is the Moliere radius, r_{opt} is the distance for which the model dependence on lateral distribution is minimal and $S(r_{\mathrm{opt}})$ is an

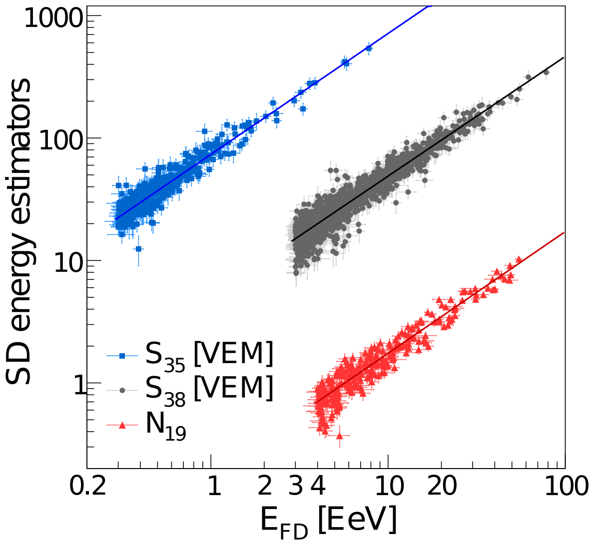


Figure 5.4: Calibration of the SD energy estimators S_{35}, S_{38} and N_{19} using events that are measured in combination with the fluorescence detector.

estimator for the shower size. r_{opt} depends on the detector spacing and is 1000 m (450 m) for the 1500 m (750 m) array. The shower axis (arrival direction) is then given by the line connecting the virtual origin with the ground impact point. For events with a 6T5 trigger, the angular resolution is better than 0.9° [206]. For the energy reconstruction of the 1500 m array, 'vertical' events with $\theta < 60°$ are distinguished from 'inclined' events with $60° < \theta < 80°$ zenith angle. For the 750 m array, only vertical events with $\theta < 55°$ are used. In vertical showers (initiated by protons or nuclei) the electromagnetic signal contribution is comparable or larger than the muon signal. In inclined events muons dominate the SD signal, as the electromagnetic component is largely absorbed in the atmosphere, and the signal response to muons increases with zenith angle due to the longer track length inside the water tank.

For vertical events, the shower size $S(r_{\text{opt}})$ is used as energy estimator. Due to the attenuation of shower particles in the atmosphere, $S(r_{\text{opt}})$ contains a zenith angle dependency. The 'Constant Intensity Cut' (CIC) method [207] is used to infer this attenuation from data, in order to estimate for a given shower size $S(1000)$ ($S(450)$) the signal S_{38} (S_{35}) it would have had at the median angle of 38° (35°).

Inclined events are reconstructed with a different procedure [208], as they exhibit a large asymmetry in the shower footprint due to deflections in the geomagnetic field. Here, the muon number density at the ground is modeled with

$$\rho_\mu(\vec{r}) = N_{19}\, \rho_{\mu,19}(\vec{r}|\phi, \theta)\,, \tag{5.2}$$

where N_{19} is the relative normalization with respect to the reference muon distribution $\rho_{\mu,19}$ for a primary proton with $E = 10^{19}$ eV and arrival direction (ϕ, θ) as obtained from simulations. After subtracting the estimated remaining electromagnetic component the model is fitted to the SD signal in order to determine N_{19} as a measure of the shower size.

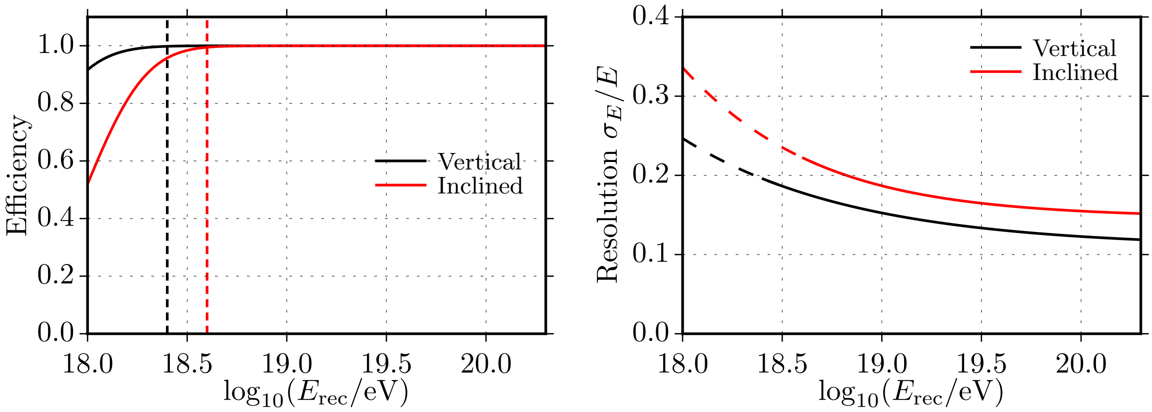


Figure 5.5: Parameterizations of the detector response for the vertical and inclined event samples used in the spectrum analysis [30, 209]. Left: The efficiency is seen to saturate above $10^{18.4}$ eV and $10^{18.6}$ eV for the vertical and inclined reconstructions, respectively. Right: Relative resolution of the reconstructions. The vertical resolution is seen to decrease towards the highest energies due to saturation of SD stations close to the shower core.

The absolute calibration of these energy estimators is done using 'hybrid' events, i.e. events that are measured in coincidence with the FD. The relation between shower size $\hat{S} = S_{38}, S_{35}, N_{19}$ and FD energy E_{FD} is well described by a simple power-law

$$E_{\mathrm{FD}} = A(\hat{S})^B \tag{5.3}$$

By fitting this relation to the hybrid events, as shown in figure 5.4, the individual SD data sets are cross-calibrated to the near-calorimetric energy measurement of the FD. The energy resolution is 15.3%, 19% and 13% for the vertical and inclined events of the 1500 m array and the 750 m array, respectively [30]. The SD energy scale uncertainty is composed of the statistical uncertainty of the calibration fit and the uncertainty of the FD energy scale. Since the uncertainties on the calibration parameters are less than 2%, the SD effectively shares the FD energy scale uncertainty of 14% [30].

5.3 Key Measurements

5.3.1 Energy Spectrum

The latest result on the cosmic-ray energy spectrum measured at the Pierre Auger Observatory was presented in [30]. The spectrum was measured individually by the 750 m and 1500 m SD arrays, the latter distinguished by vertical and inclined events, and by the FD in combination with at least one SD station (hybrid data).

The main steps in measuring the energy spectrum are: counting the cosmic ray events in bins of reconstructed energy, dividing by the experimental exposure, and correcting for effects of the finite energy resolution. As described in the previous two sections, the energy is reconstructed using energy estimators that are cross-calibrated with the fluorescence detector. The

Reconstruction	Vertical	Inclined	750 m	Hybrid
Exposure [$\mathrm{km}^2\,\mathrm{sr}\,\mathrm{yr}$]	42500 ± 1300	10900 ± 300	150 ± 5	1500 ± 20 at 10^{19} eV
Threshold energy	3×10^{18} eV	4×10^{18} eV	3×10^{17} eV	10^{18} eV

Table 5.1: Exposures and threshold energies of the four spectrum measurements [30].

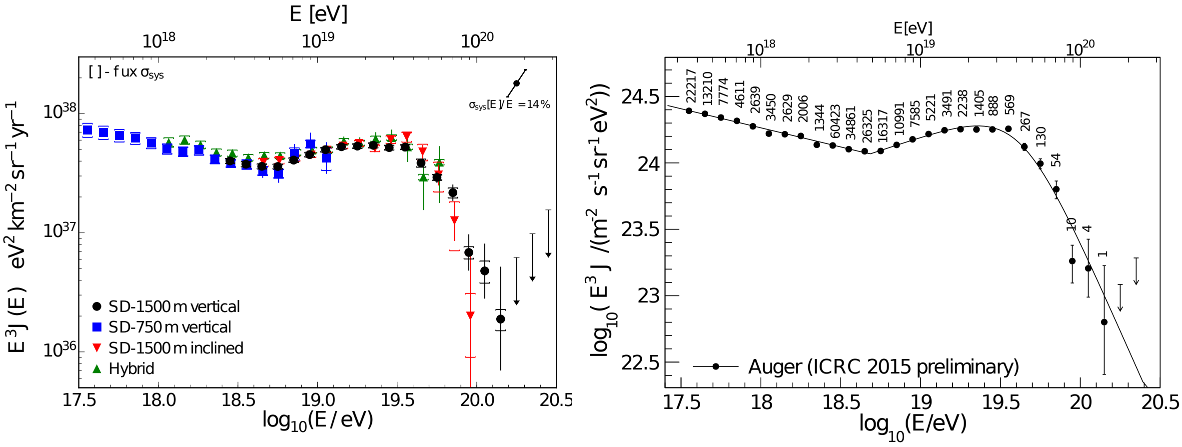


Figure 5.6: Energy spectrum as measured at the Pierre Auger Observatory [30]. Left: Individual spectra using SD and FD data. Statistical and systematical uncertainties (excluding the energy scale) are shown as error bars and brackets, respectively. Right: Combined spectrum fitted with by an empirical flux model to guide the eye. Only statistical uncertainties are shown. Total numbers of events are given for each bin. Upper limits above $E = 10^{20.2}$ eV correspond to the 84% C.L.

exposures for the surface detector measurements above the energy thresholds for full detection efficiency are obtained by integrating the number of active hexagons over time and multiplying with the respective aperture of one hexagon. The hybrid (FD) exposure is calculated from time-dependent MC simulations which reproduce the data taking conditions and the hybrid detector response. The threshold energies and exposures are given in table 5.1. In the steeply falling energy spectrum the finite detector resolution results in more upwards than downwards fluctuations of the reconstructed energies, which causes a net increase of the measured flux at a given energy. In order to estimate the true spectrum this effect is accounted for in a forward folding approach: An empirical model of the true flux is folded with the detector response and fitted to the measurements. The energy dependent ratio of the true and the folded flux models then serves as a correction factor for the measured spectrum. The resulting estimate of the true spectrum has a considerable dependence on the flux model that is used in the forward folding procedure. Therefore, comparisons between simulated and measured cosmic-ray spectra should be performed on the reconstruction level for a self-consistent treatment of the detector response.

All four measured spectra agree within uncertainties, as seen on the left side of figure 5.6. Since the SD energy estimators are calibrated to the FD, all measurements share the

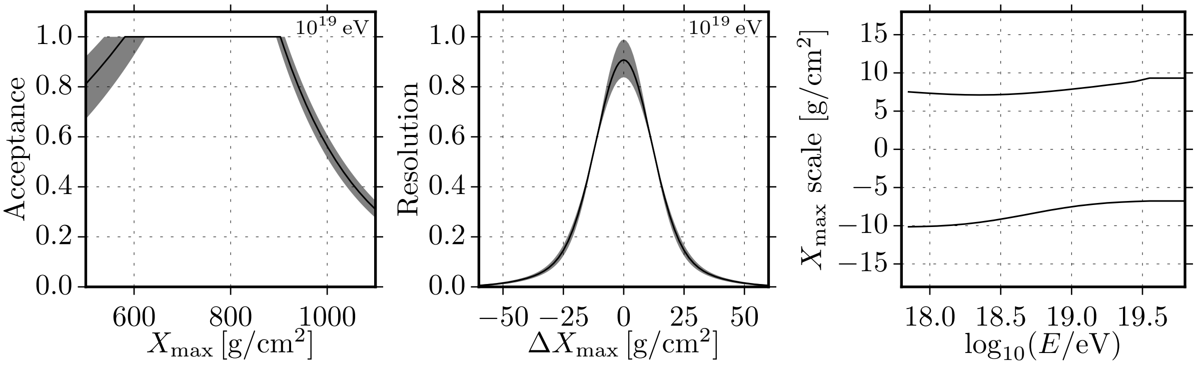


Figure 5.7: Parameterizations of the detector response for the X_{max} data selection in [210]. Left and middle: Acceptance and resolution at $E = 10^{19}$ eV with the shaded area indicating the systematical uncertainty. Right: Systematic uncertainty on the X_{max} scale as function of energy.

same energy scale uncertainty of 14%. Additional systematical uncertainties arise from the energy calibration fit (for the SD spectra), the exposure calculation and the unfolding. The independent measurements are then combined with a maximum likelihood method, where the systematical uncertainties are taken into account and the flux normalizations are allowed to vary within the respective exposure uncertainties. The combined spectrum is shown on the right side of figure 5.6 along with the number of observed events for each energy bin. This constitutes the most precise measurement of the cosmic-ray spectrum above $10^{17.5}$ eV. The ankle is observed at $10^{17.68}$ eV and the observed flux suppression at the highest energies is established with more than 20σ significance.

5.3.2 Depth of Shower Maximum

The Pierre Auger Collaboration recently reported the first measurement of the full X_{max} distributions of cosmic-ray air showers with energies above $10^{17.8}$ eV [210]. Whereas previous Xmax measurements have concentrated on determining the mean and standard deviation of the distributions, only the distributions themselves contain the full information on the cosmic-ray composition that can be obtained from measuring Xmax.

From 2.6×10^6 shower candidates 1.1×10^5 showers with energies above $10^{17.8}$ eV are reconstructed under good data taking conditions. From these, 2×10^4 high quality events, shown in figure 5.8, are selected for which the distortions of the measured X_{max} distributions are minimal. In particular, one quality cut requires that X_{max} is observed within the field of view. The baseline fluorescence telescopes observe the sky within elevation angles of 1.5° to 30°, which sets an upper and lower limit on the observable X_{max} range. Additionally, the observable range is affected by the limited distance up to which showers are detected and by other quality cuts. A non uniform acceptance of X_{max} values can bias the inferred mass composition. Therefore, a fiducial field of view cut is applied that ensures a uniform acceptance of X_{max} values in the relevant range.

Still, the measured distributions $f(X_{\mathrm{max}}^{\mathrm{obs}})$ are different from the true distributions $f(X_{\mathrm{max}})$

on account of the detector response

$$f(X_{\max}^{\text{obs}}) = \int_0^\infty f(X_{\max})\epsilon(X_{\max})R(X_{\max}^{\text{obs}} - X_{\max})dX_{\max} \ . \tag{5.4}$$

Here ϵ is the detector acceptance and R the detector resolution. Parameterizations of the acceptance and resolution for the high quality selection have been obtained from simulated events and detailed detector simulations. They are shown in figure 5.7 and described in the following. The relative acceptance after the field of view cut is given by

$$\epsilon(X_{\max}) = \begin{cases} \exp(+(X_{\max} - x_1)/\lambda_1) & X < x_1 \\ 1 & x_1 < X < x_2 \\ \exp(-(X_{\max} - x_2)/\lambda_2) & X > x_2 \end{cases} \ , \tag{5.5}$$

where $x_1, x_2, \lambda_1, \lambda_2$ are energy dependent parameters listed in [210]. The resolution determines the broadening of the distribution due to the statistical uncertainty in measuring $X_{\max}$ and can be parameterized by a sum of two Gaussians

$$R(X_{\max}^{\text{obs}} - X_{\max}) = fG(\sigma_1) + (1 - f)G(\sigma_2) \tag{5.6}$$

with an energy dependent relative contribution f and standard deviations σ_1, σ_2 listed in [210]. Finally, a systematic uncertainty on the $X_{\max}$ scale arises from the detector calibration and reconstruction as well as the knowledge of fluorescence yield, Cherenkov light and atmospheric attenuation. The total systematic uncertainty, shown on the right of figure 5.7, is $\leq 10\,\text{g/cm}^2$ at all energies.

From the energy binned $X_{\max}$ distributions the unbiased values of the first two moments, mean and spread, are obtained by accounting for the detector response, see figure 5.8. The energy dependence of the moments is seen to be different from the expectation for either a pure-proton or pure-iron composition. A comparison of the moments with shower simulations indicates that the average mass decreases towards $10^{18.25}$ eV and increases for higher energies.

A more detailed analysis was performed in [34] by considering the full $X_{\max}$ distributions. While the large shower-to-shower fluctuations of $X_{\max}$ prevent a determination of the contribution of every mass number A, it is possible to constrain the contributions of a limited number of representative mass numbers. In [34] four representative mass numbers $A = 1, 4, 14, 56$ were selected, which correspond to stable isotopes of the elements hydrogen (protons), helium, nitrogen and iron. The motivation for this choice is that these values are approximately equidistant in $\ln(A)$, which is the main dependence of the $X_{\max}$ observable, cf. section 2.1.1. Templates for the $X_{\max}$ distribution for each of these elements were generated from shower simulation and compared to the measured distributions in a binned likelihood fit. As a general feature it was found that the data is inconsistent with a composition dominated by either protons or iron nuclei and that intermediate nuclei are required for acceptable fit qualities. In the present analysis we make use of a similar statistical analysis in order to compare the prediction of astrophysical scenarios to the measured $X_{\max}$ distributions.

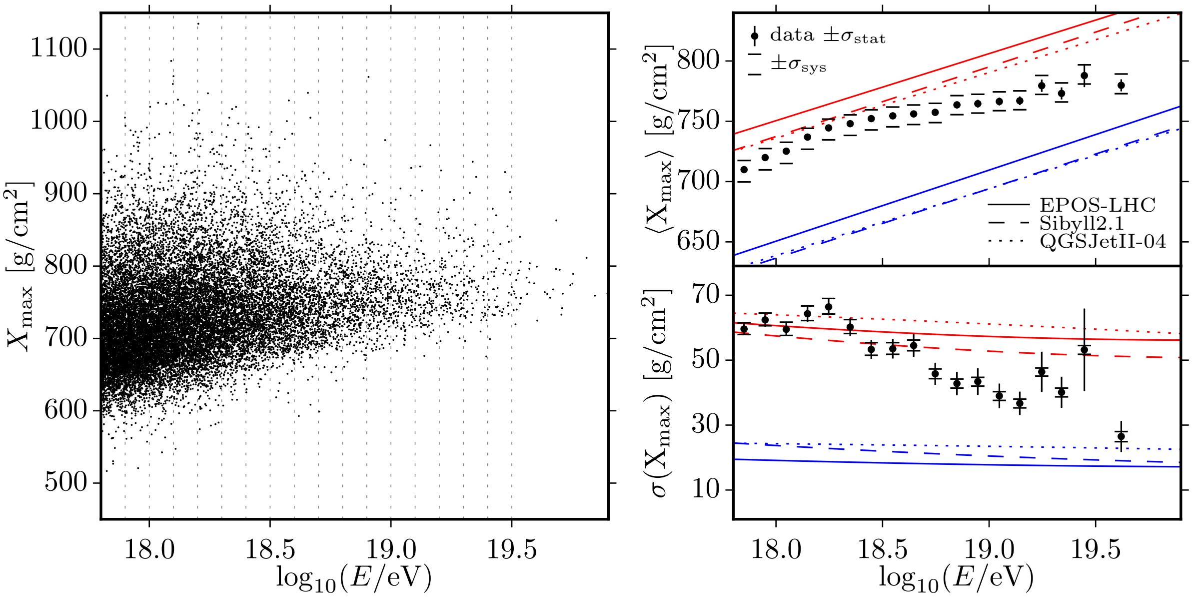


Figure 5.8: Left: Distribution of energy and $X_{\max}$ of the selected air showers [210]. Dashed lines indicate the energy bins for the analysis of moments on the right. Right: Mean and spread of the $X_{\max}$ distribution as function of energy. For comparison the expectation for proton (red) and iron (blue) are shown for different hadronic interaction models. Note that for a mixed composition the spread can easily exceed that of the indicated pure proton composition.

Chapter 6

Method of Constraining Astrophysical Scenarios of UHECRs

This chapter describes a method of a combined fit to the energy dependent flux and composition of ultra-high energy cosmic rays measured by the Pierre Auger Observatory, using suitable astrophysical scenarios. The method is applied to energies $E > 5\,\mathrm{EeV}$, i.e. above the ankle feature in the cosmic-ray all-particle spectrum. At these energies a galactic origin is unlikely, mainly because the gyroradius of cosmic rays is too large to be contained in the Galaxy, and because of the lack of adequately powerful accelerators within the Milky Way. Hence, an extragalactic origin is assumed when adopting an astrophysical scenario for this energy range. Here we consider a simple model consisting of identical, uniformly distributed sources, which emit cosmic-ray nuclei of a few representative elemental species with a rigidity dependent mechanism. This is in line with recent analyses [81, 83–87]. Complementary to these analyses we will formulate a more rigorous treatment of the statistical and systematic experimental uncertainties in order to better exploit the information in the available cosmic-ray measurements. We will also study the impact of the main simulation uncertainties and consider several variations of the simplified source model.

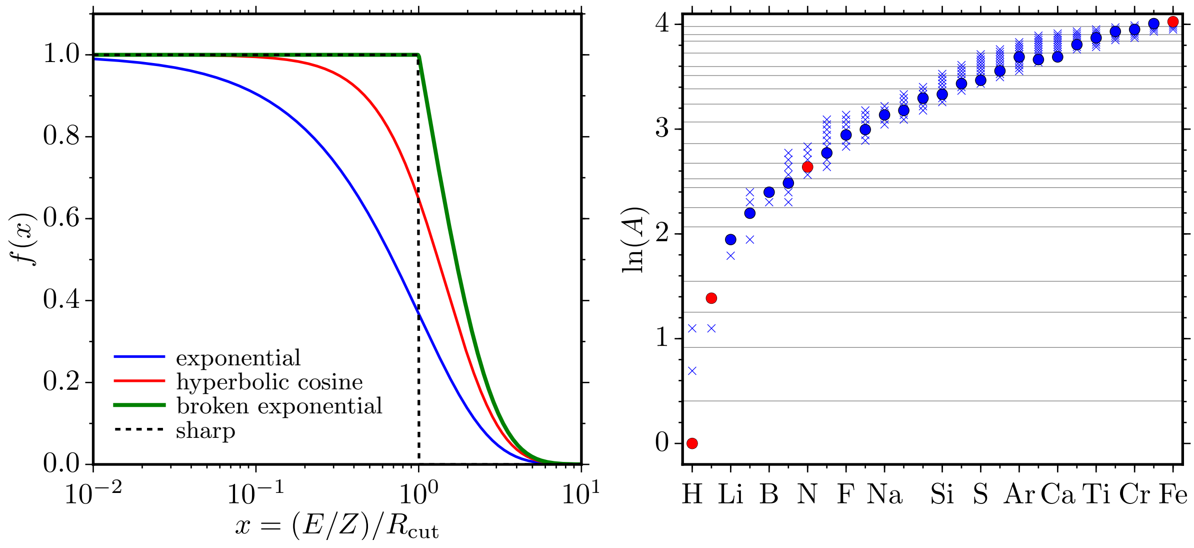
The chapter is structured as follows: First, the astrophysical scenario is presented, consisting of the source model and the modification of the cosmic-ray flux during extragalactic propagation (section 6.1). Then, the calculation of the predicted observables from the simulated cosmic-ray flux is described (section 6.2). Finally, the statistical method for evaluating the astrophysical scenario and constraining the model parameters is explained (section 6.3). While the method of calculating the cosmic-ray flux is to some extent specific to the considered scenario, the procedure of calculating the predicted observables and the method of statistical inference are more general and can be applied to a wide range of cosmic-ray scenarios.

6.1 Astrophysical Scenario

6.1.1 Source Model

As a simple source model we assume a uniform distribution of identical sources, each accelerating cosmic rays with a power-law spectrum. Following the arguments outlined in section 3.1 we consider a rigidity dependent maximum acceleration energy $E_{\mathrm{max}} = Z \cdot R_{\mathrm{cut}}$. The source spectrum thus reads

$$J_0(Z_0, E_0) = \phi_0\, a(Z_0)\, f_{\mathrm{cut}}\left(\frac{E_0/Z_0}{R_{\mathrm{cut}}}\right)\left(\frac{E_0}{1\,\mathrm{EeV}}\right)^{-\gamma}. \tag{6.1}$$



(a) Cutoff functions

(b) $\ln(A)$ for different nuclides

Figure 6.1: (a) Comparison of functional cutoff shapes in equation 6.2. (b) $\ln(A)$ for nuclides in the range $Z = 1-26$, indicating the composition sensitivity of the X_{max} observable. Circles show the most abundant isotopes for each element, blue crosses all other isotopes with a lifetime $\tau > 2\,\text{s}$. The nuclides ^{1}H, ^{4}He, ^{14}N and ^{56}Fe (red circles) are approximately equidistant in $\ln(A)$. Horizontal lines show the mass bins in which the simulated events are stored, see section 6.2 for details.

Here, ϕ_0 is the flux level of the cosmic ray sources, $a(Z_0)$ denotes the fraction of cosmic rays of element Z_0 at a given energy in absence of a cutoff, f_{cut} describes the cutoff shape around the maximum rigidity, and γ the spectral index. The shape of the cutoff depends on the details of the acceleration mechanism and the environment in which it takes places, see e.g. [92] and references therein. As done in recent analyses [81, 87] we will primarily consider the case of a broken exponential cutoff. However, we will also study the impact of different functional cutoff forms, namely

$$f_{\text{cut}}\left(x = \frac{E_0/Z_0}{R_{\text{cut}}}\right) = \begin{cases} e^{-x} & \text{exponential} \\ \cosh(x)^{-1} & \text{hyperbolic cosine} \\ \Theta(1-x) + \Theta(x-1)e^{1-x} & \text{broken exponential} \\ \Theta(1-x) & \text{sharp} \end{cases}, \tag{6.2}$$

where Θ is the Heaviside step function. As seen in figure 6.1a, these functions mainly vary in the abruptness of the predicted cutoffs. For instance, the broken exponential and sharp cutoffs don't modify the spectrum for $R < R_{\text{cut}}$, whereas with an exponential cutoff the flux is already suppressed by 40% at half of the cutoff rigidity.

Regarding the source composition, in principle any element that is present in the universe can be accelerated at the cosmic-ray sources. Recent astrophysical analysis have focused on considering hydrogen (protons), helium, nitrogen and iron as representative elements at the source. This follows the same motivation as the empirical analysis of the observed $X_{\max}$ distributions in [34]: If the cosmic rays reach Earth largely intact, their mass numbers are approximately equidistant in the principal $\ln(A)$ dependency of the $X_{\max}$ observable, as seen in figure 6.1b. Hence, their contributions to the source spectrum is much better constrained than the individual contributions of nearby mass number, such as ^{12}C, ^{14}N and ^{16}O. For comparability we will follow the choice of ^{1}H, ^{4}He, ^{14}N and ^{56}Fe as representative elements. We will also consider variations of these representative elements in order to assess the impact on the astrophysical fit and to investigate possible differences of individual nuclides in terms of cosmic ray propagation.

Due to the aforementioned assumption of continuously distributed sources, the propagation theorem described in section 3.6.1 applies and the resulting cosmic-ray flux takes a universal shape. In particular, the spectrum and composition do not depend on magnetic field effects. For a discrete source distribution, deviations from this universal shape are expected at energies where the interaction lengths are smaller than the distance to the closest sources [211]. In that case the universal spectrum can be further modified by magnetic field effects [212]. Investigating these effects, sometimes called ‘ensemble fluctuations’ or ‘cosmic variance’, is beyond the scope of the present analysis.

6.1.2 Simulated Cosmic-Ray Flux

The cosmic-ray flux injected by the sources is modified during its extragalactic propagation in a non-trivial way, as explained in chapters 3 and 4. In order to calculate this modification we perform Monte Carlo simulations of the propagation with CRPropa 3. Since the analysis focuses on the energy spectrum and composition, the simulation can be efficiently performed in one spatial dimension including the effects of the expanding universe. An exemplary steering card can be found in appendix A.4.

In the simulation all relevant processes for cosmic-ray nuclei are considered, namely redshift losses, electron pair production, photodisintegration, photoproduction of hadrons and nuclear decay. The main uncertainties in modeling these processes lie in the photodisintegration cross sections and in the model of the extragalactic background light [141]. For the photodisintegration models (see section 4.5) we use TAYLS as the baseline model, but will also consider the Kossov and PSB models. Regarding the extragalactic background light (see section 4.3) we use the model of Gilmore *et al.* [108] as the baseline, and consider the models by Dominguez *et al.* [103] and Franceschini *et al.* [113] as variations. As cosmological parameters the current standard values $H_0 = 63.7\,\mathrm{km/s/Mpc}$, $\Omega_m = 0.315$ and $\Omega_\Lambda = 0.685$ are used [28].

For each simulation setup, we inject a large number of cosmic rays at random source positions and record all particles that arrive at the observer. The simulation is done separately for each injected element. In CRPropa 3, 183 nuclides are considered up to iron, leading to a choice of the specific isotope for a given element. Here, we always select the most abundant isotope (on Earth) from [193], as shown in figure 6.1b. The number of injected and recorded particles are listed in table 6.1. We verified that in the analysis the statistical uncertainty of the simulated observables due to the finite number of simulated events is sufficiently small compared to the experimental statistical uncertainty.

Element	Z_0	A_0	N_{injected}	N_{recorded}
Hydrogen	1	1	4×10^6	4×10^6
Helium	2	4	4×10^6	14×10^6
Nitrogen	7	14	8×10^6	83×10^6
Iron	26	56	4×10^6	140×10^6

Table 6.1: Type and number of simulated cosmic rays for each considered scenario. Note that for nuclei $A > 1$ the number of recorded particles exceeds the injected numbers due fragmentation in up to A nucleons.

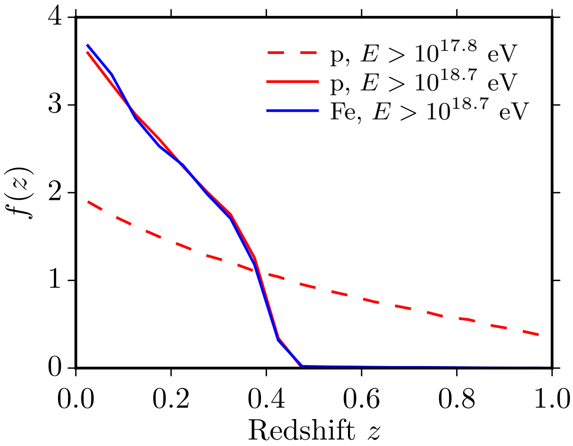


Figure 6.2: Redshift distribution for events from initial ^{1}H and 56 $^{\mathrm{F}}$e recorded with $E > 10^{17.8}$ eV and $E > 10^{17.8}$ eV. The simulated range of $z < 2$ is seen to be sufficient for covering all contributions from all distances. Note that the width of the distribution decreases for lower maximum injected energies.

The cosmic rays are simulated with an initial spectrum of $dN/dE_0 \propto E_0^{-1}$ between a minimum and maximum energy of $10^{17.5}$ eV and $Z_0 \times 21.5$ eV, respectively. This allows to consider the cosmic-ray flux in the range of reconstructed energies $E^{\text{rec}} = 10^{17.8}\,\text{eV} - 10^{20.5}\,\text{eV}$ including systematic shifts of the energy scale of $\Delta \log_{10}(E/\text{eV}) = \pm 0.3$, which corresponds to variations of more than $\pm 3\sigma$. In the analysis the cutoff rigidity of the sources can be varied up to maximum value of $R_{\text{cut}} = 10^{20.5}$ eV while still ensuring at least a decade in energy to model the shape of the source spectrum above the cutoff. This is potentiallRedshift disy important, because for hard source spectra even a minor flux of heavy high-energy nuclei can cause a significant pile-up of light secondary particles at lower energies.

The sources are uniformly distributed with a minimum distance of $d = 3$ Mpc and a maximum distance corresponding to a redshift $z = 2$. As seen in figure 6.2 this maximum distance safely encompasses all contributions for observed energies above $10^{17.8}$ eV.

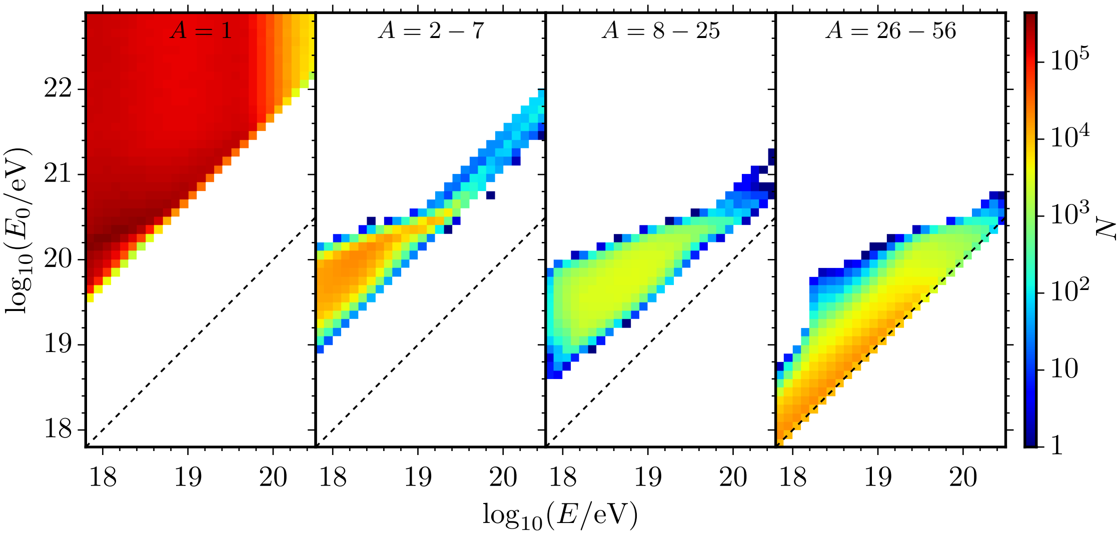


Figure 6.3: Histogram of simulated events from initial iron nuclei (see text for details), separated by groups of observed mass. The observed and initial energy are given by the x and y-axis, respectively. The dashed line indicates the border case of no energy loss. The gap between this line and the observed events corresponds to the energy fraction E/A that the secondary particles can carry away. For heavier observed cosmic rays the ratio of initial and observed energy is closer to one.

6.1.3 Event Weighting

For a fast calculation of the cosmic-ray flux for different source parameters the propagation simulation is performed only once and the simulated events are weighted according to the modeled source spectrum in equation 6.1. To this end the simulated cosmic rays are first filled into a histogram $N(Z_0^i, E_0^j, A^k, E^l)$ by bins of initial element Z_0, initial energy E_0, final mass number A and final energy E. Here, the bins of initial and final energy have a width of $\log_{10}(E/\mathrm{eV}) = 0.1$, which corresponds to the energy bin width of the observables.

Due to photonuclear interactions the final mass number of an initial iron nucleus can in general take any value between 1 and 56. In practice certain mass numbers are observed with much higher frequencies than others. Moreover, the difference between neighboring mass numbers in the $\ln(A)$ dependency of the X_{max} observable decreases with increasing mass, as seen in figure 6.1b. Therefore, in order to reduce the size of the histogram we select the following 20 bins of final mass number: $A = 1, 2, 3, 4, 5-7, 8-9, 10-11, 12, 13-14, 15-17, 18-21, 22-25, 26-29, 30-33, 34-36, 37-41, 42-46, 47-49, 50-53$ and $54-56$ as shown in figure 6.1b. Compared to considering each mass number individually, this binning is found to have a negligible impact on the simulated observables, while allowing for a significant speed-up in the following calculations. The histogram of simulated events is visualized in figure 6.3 for the example of initial iron, showing the number and energy distribution of secondary particles as resulting from the unweighted simulated energy spectrum. Here and in the following, the mass numbers are combined for illustrative purposes into groups which are roughly centered at the four customary representatives, hydrogen, helium, nitrogen and iron.

The number of events N_w that are predicted to arrive according to a specific set of source parameters (R_{cut}, γ, $a(Z)$) can now be written in form of a matrix multiplication and summation:

$$N_w(A^k, E^l) = \sum_{i=1} \sum_{j=1} w(Z_0^i, E_0^j)\, N(Z_0^i, E_0^j, A^k, E^l) \tag{6.3}$$

$$w(Z_0^i, E_0^j) = a(Z_0^i) f_{\text{cut}} \left(\frac{E_0^j / Z_0^i}{R_{\text{cut}}} \right) \left(\frac{E_0^j}{1\,\text{EeV}} \right)^{-\gamma+1} \quad . \tag{6.4}$$

Note that the weight $w(Z_0, E_0 | R_{\text{cut}}, \gamma, a_Z)$ is equal to the source spectrum in equation 6.1, except for the power law index $-\gamma$ being incremented by one to account for the injected energy spectrum $\propto (E_0/1\,\text{EeV})^{-1}$ in the simulation. The flux normalization is not yet considered here, because it is only needed for the comparison with the measured energy spectrum. This matrix concept enables the usage of optimized implementations of matrix products and summations. The resulting fast computation is essential for the Bayesian analysis described below in section 6.3.

6.2 Simulated Observables

The histogram of weighted events $N_w(A^k, E^l)$ contains the information on the predicted energy spectrum and mass composition. In order to compare this information to the measurements, the histograms need to be translated into the simulated energy spectrum and X_{max} distributions. Specifically, we are considering the energy spectra measured by the Pierre Auger Observatory with the vertical and inclined reconstructions surface detector with 1500 m spacing. Both reconstructions are fully efficient in the considered energy range and provide the bulk of the total experimental exposure with 42 500 km^2 sr yr and 10 900 km^2 sr yr, respectively. For the X_{max} observable we consider the full distributions as measured by the fluorescence detector.

As described in the previous chapter (section 5.3) the measurements are distorted by the detector response: Both the reconstructed energy and X_{max} values are modified due to the finite resolution and the limited acceptance. While it is possible to unfold the detector response from the measured distributions, in doing so the statistical uncertainties may propagate into large uncertainties and negative correlations in the resulting estimators of the true distributions. Furthermore, when unfolding the measured energy spectrum with a forward folding method as described in section 5.3.1, the resulting estimate has a dependency on the assumed model of the spectral shape. In this sense, an astrophysical fit has the advantage of providing the ‘true’ cosmic-ray flux according to the simulation. Therefore, a self-consistent treatment of the detector and smaller statistical uncertainties can be achieved by folding the detector response into the predicted observables. The comparison of simulation and measurement is thus performed on the detector level. The present analysis is the first to use this approach for both the energy spectrum and X_{max} distributions.

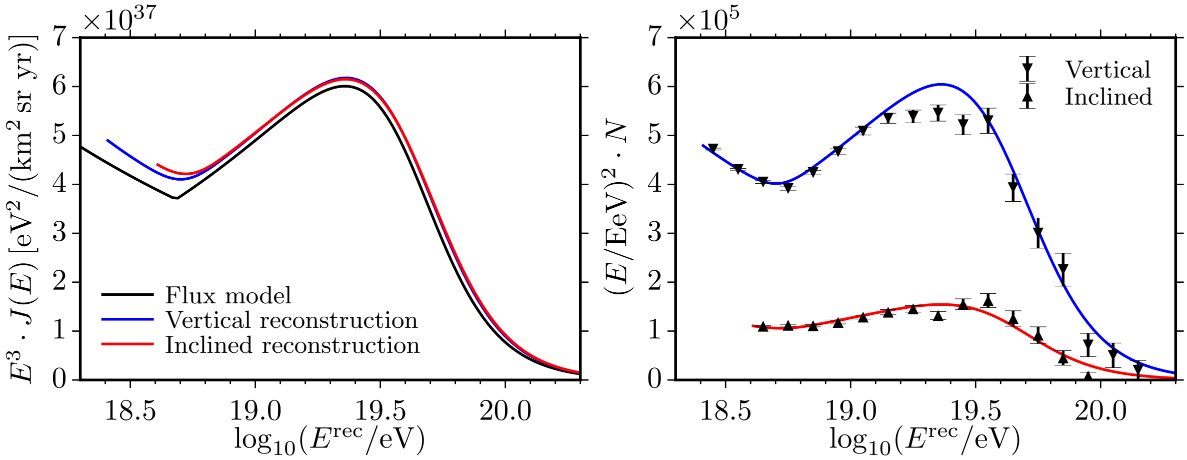


Figure 6.4: Impact of the detector response on the empirical flux model from [30]. Left: Comparison between unfolded (black) and folded energy spectra for the vertical (blue) and inclined (red) reconstruction. The folding causes a net upwards fluctuation, which is seen as an increased flux level. Right: Comparison between the folded flux models and the observed events (black dots) with error bars showing the statistical Poisson uncertainties of the measurement.

6.2.1 Energy Spectrum

The energy spectrum follows from the weighted number of simulated events summed over all mass numbers, $N_w(E^l) = \sum_k N_w(A^k, E^l)$. The next step is to fold in the detector response, consisting of the energy resolution, a possible energy bias and the detector efficiency. Here, the finite energy resolution $\sigma_E(E)$ causes a Gaussian smearing of the reconstructed energies about the true energies. The energy bias $b(E)$ shifts the reconstructed energies and the detector efficiency $\varepsilon(E)$ decreases the contribution from a given true energy. The impact of the efficiency is small, since the surface detector is fully efficient above the threshold energies of 3×10^{18} eV and 4×10^{18} eV for the vertical and inclined measurements, respectively. For the considered energy range $E_{\mathrm{rec}} > 5 \times 10^{18}$ eV the impact is thus limited to events that migrate to higher energies. The entire detector response can be expressed through a response matrix, containing the probabilities of an event of energy E^l to be reconstructed with energy E^m_{rec}. The matrix reads

$$M(E^m_{\mathrm{rec}}, E^l) = \varepsilon(E^l) \exp\left(-\frac{(E^m_{\mathrm{rec}} - b(E^l)E^l)^2}{2\sigma_E^2(E^l)}\right). \tag{6.5}$$

Corresponding to the event histogram N_w, the matrix is calculated with an energy binning of $\Delta \log_{10}(E/\mathrm{eV}) = 0.1$ in both the true and reconstructed energies. To take into account the energy distribution within each energy bin, the matrix is first calculated with a ten-fold finer binning and weighted according the empirical model of the true energy spectrum in [30] before down-sampling to the final bin width. We verified that this procedure is compatible with the unfolding reported in [30]. Using the response matrix the predicted energy spectra for the

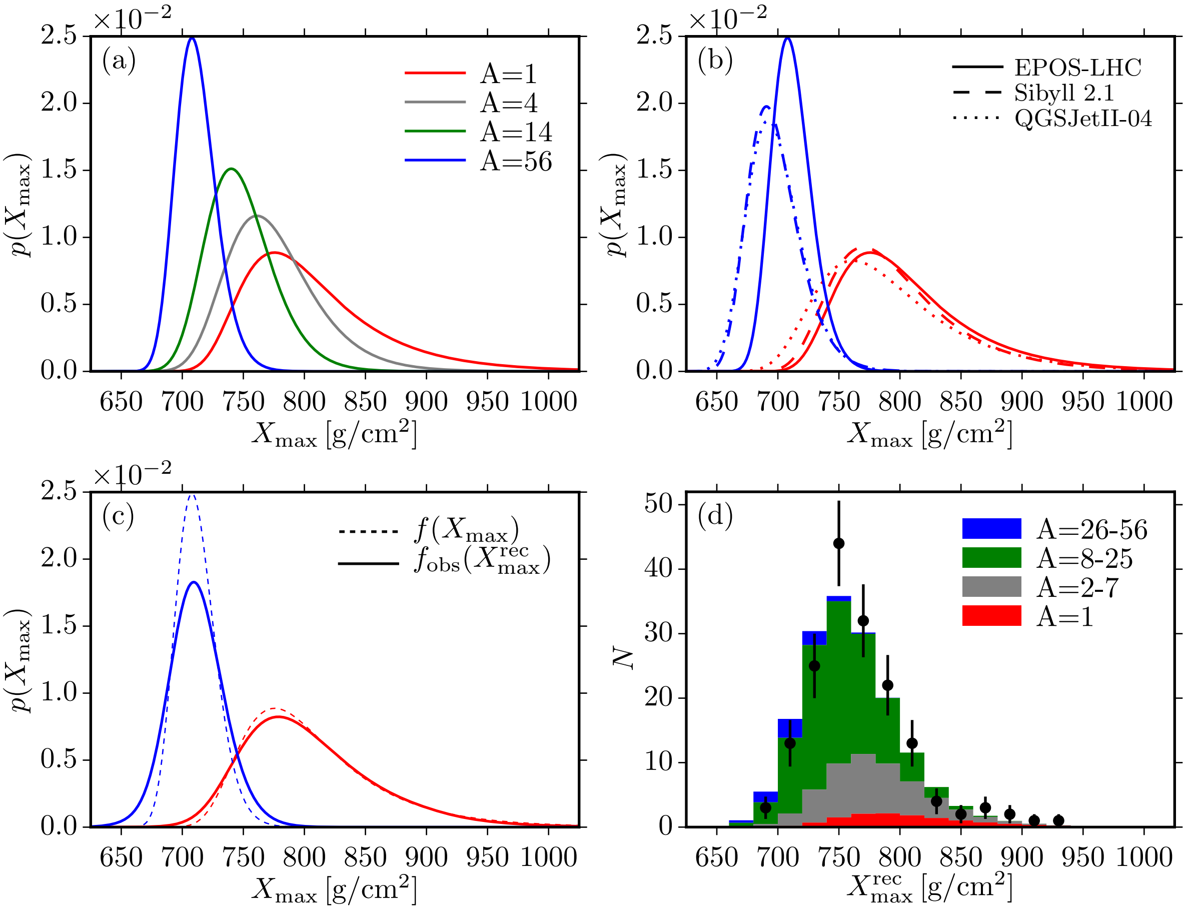


Figure 6.5: Simulated X_max probability distributions at $E = 10^{19-19.1}$ eV. (a) Distributions for different mass numbers. (b) Differences between hadronic interaction models for proton and iron showers. (c) Modification of the simulated distribution by the detector response for proton and iron showers. (d) Comparison between the measured X_max distribution (black dots) and the prediction for a generic mixed composition using EPOS-LHC.

vertical and inclined reconstructions, denoted by the superscripts V and I, are given by

$$N^V_\mathrm{spectrum}(E^m_\mathrm{rec}) = \phi_0\,\omega^V \sum_l M^V(E^m_\mathrm{rec}, E^l) N_w(E^l) \tag{6.6}$$

$$N^I_\mathrm{spectrum}(E^m_\mathrm{rec}) = \phi_0\,\omega^I \sum_l M^I(E^m_\mathrm{rec}, E^l) N_w(E^l)\,. \tag{6.7}$$

Here ϕ_0 is the flux normalization of the sources and ω is the experimental exposure. The effect of the folding is illustrated in figure 6.4 for the example of the empirical flux model reported in [30]. The figure on the right shows the comparison with the measured energy spectra.

6.2.2 $X_{\max}$ Distributions

To model the X_{max} distribution we use the parametrization developed in [213], which takes the form of a generalized Gumbel distribution

$$p(X_{\text{max}}|E,A) = \frac{\lambda^{\lambda}}{\sigma\Gamma(\lambda)}\left(e^{-\lambda z-\lambda e^{-z}}\right), \quad z = (X_{\text{max}}-\mu)/\sigma \ . \tag{6.8}$$

The energy and mass dependent parameters μ, σ, λ characterize the location, scale and shape of the distribution, respectively. The parameters were empirically obtained by fitting to distributions of simulated showers for several hadronic interaction models. In this analysis we consider the cases of QGSJetII-04, Sibyll 2.1 and EPOS-LHC with the latter serving as the baseline model. Differences between the distributions for individual masses are seen in figure 6.5a and for different hadronic interaction models figure 6.5b.

The Gumbel parameterization has been used in [214, 215] to infer the fractions of the representative mass numbers $A = 1, 4, 14, 56$ in the X_{max} distributions measured by the Pierre Auger Observatory. In this analysis we consider the distribution of all mass numbers $A = 1-56$ from the simulation of extragalactic propagation, represented by the 20 mass bins previously described (section 6.1.3).

We construct templates of the X_{max} distribution for each energy and mass bin by evaluating the Gumbel parametrization and folding with the X_{max} detector response in equation 5.4. The effect of the folding is shown in figure 6.5c for the cases of proton and iron showers. The resulting templates $G(A^k, E^l, X_{\text{max}}^{\text{rec},x})$ give the probability of a cosmic ray of mass number A^k and energy E^l to be reconstructed as $X_{\text{max}}^{\text{rec},x}$. Note that the acceptance is explicitly not used for normalizing the templates, so that $\sum_x G(A^k, E^l, X_{\text{max}}^{\text{rec},x}) \leq 1$. This is done to include the overall acceptance for certain elements in the predicted X_{max} distributions. This overall acceptance ranges from 0.977 for protons at $E = 10^{19.7}$ eV and EPOS-LHC, up to 1 for iron nuclei at all energies.

For a given simulated cosmic-ray flux the X_{max} probability distribution reads

$$p(E^l, X_{\text{max}}^{\text{rec},x}) = c\sum_k N_w(A^k, E^l)\, G(A^k, E^l, X_{\text{max}}^{\text{rec},x}) \tag{6.9}$$

where c normalizes the right hand side, summed over all X_{max} bins. Due to the low number of events, the last energy bin of the *measured* X_{max} distribution combines the energies $\log_{10}(E/\text{eV}) > 19.5$. The mean energy in this bin is $\langle\log_{10}(E/\text{eV})\rangle = 19.62$ and the highest energy is about 19.9. We therefore combine the energy bins $\log_{10}(E/\text{eV}) = 19.5 - 19.9$ in the simulated X_{max} distribution.

An example for the conducted fits is given in figure 6.5d, where the measured X_{max} distribution at $E = 10^{19-19.1}$ eV is compared to the simulated distribution according to a generic mixed composition. The presented method is tested in appendix A.3, where the empirical fit to the X_{max}distributions reported in [34] is repeated with compatible results regarding the observed composition.

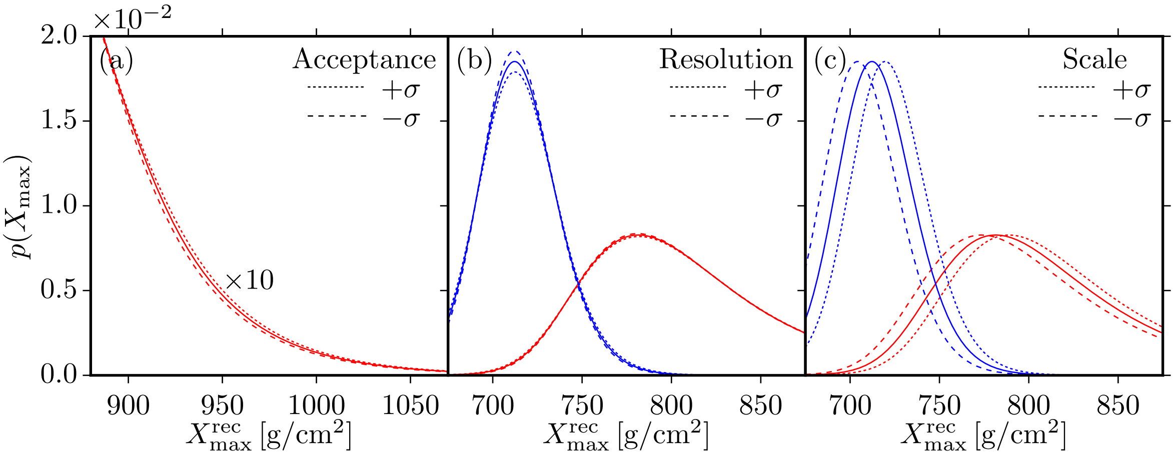


Figure 6.6: Impact of the $X_{\max}$ detector response uncertainties on the probability distributions $p(X_{\max})$ for $A = 1$ (red) and $A = 56$ (blue) at $E = 10^{19-19.1}$ eV. (a) Variations of the acceptance have a small impact at high $X_{\max}$ values. The distribution is scaled by 10 for emphasis. (b) Variations of the resolution are most pronounced at the narrow peaks at high mass numbers. (c) The $X_{\max}$ scale uncertainty allows for considerable shifts towards lower or higher values.

6.2.3 Experimental Uncertainties

Uncertainties in the detector response cause systematic uncertainties in the measured energy and $X_{\max}$ observables. The statistical evaluation described below requires that these uncertainties can be continuously varied in form of 'nuisance parameters' θ. A nuisance is any parameter of the physical model that is not a parameter of interest. Here, the nuisance parameters are expressed in terms of standard scores, i.e. $\theta = 1$ corresponds to a variation by $+1\sigma$. The considered uncertainties and the methods for their continuous variations are described in the following.

The most important systematic uncertainty is that of the energy scale. As described in chapter 5 the surface detector is cross-calibrated to the fluorescence detector. Since the uncertainty of the calibration is small, SD and FD have the same energy scale uncertainty, which is estimated at $\sigma_E/E = \pm 14\%$. The energy scale uncertainty affects both observables. A shift of the energy scale is included in the analysis by shifting the simulated spectrum and $X_{\max}$ distributions through a linear interpolation in energy. Applying the shift on the level of the observables is possible, because all parameterizations that are used in their calculation are functions of true energy.

In the energy spectrum measurements an additional systematic uncertainty stems from the exposures ω^V and ω^I, which have relative uncertainties of $\approx 3\%$, each. Together with the flux normalization ϕ_0, these exposures represent three multiplicative factors for two observables, and are consequently not independent. Therefore, we consider one of the two exposures to be fixed and only vary the other exposure and the flux normalization. Concretely, the vertical exposure is kept at its nominal value and the *relative* inclined exposure $r = \omega^I/\omega^V$ is varied within a relative uncertainty that corresponds to the relative uncertainties of the individual exposures added in quadrature

$$\frac{\sigma_r}{r} = \sqrt{\left(\frac{1300\,\mathrm{km}^2\,\mathrm{sr\,yr}}{42\,500\,\mathrm{km}^2\,\mathrm{sr\,yr}}\right)^2 + \left(\frac{300\,\mathrm{km}^2\,\mathrm{sr\,yr}}{10\,900\,\mathrm{km}^2\,\mathrm{sr\,yr}}\right)^2} = 4.1\% \,. \tag{6.10}$$

Formally, this relative exposure, being the ratio of two normal distributed quantities, follows a ratio distribution. However, the relative uncertainties are small enough to consider the relative exposure as normal distributed as well.

The effect of variations of the X_{max} detector response are seen in figure 6.6 for the cases of $A = 1$ and 56 at $E = 10^{19-19.1}\,\mathrm{eV}$. The fiducial field of view cut, described in section 5.3.2, ensures a uniform acceptance in the range where the bulk of the X_{max} distribution is measured. Therefore, a simulated distribution that reproduces the measurement is little affected by variations of the acceptance. Specifically, the variations are limited to the tails of the predicted X_{max} i.e. for $A = 1$ at high energies. The finite X_{max} resolution causes a smearing of the distribution, which is most pronounced at the peaks of the X_{max} distributions for high mass numbers A. Finally, the X_{max} scale uncertainty implies a possible shift of the distributions towards higher or lower values. To allow for continuous variations of the detector acceptance and resolution, we make use of the 'template morphing' method. In high energy particle physics this method is employed to consider the effect of shape-changing uncertainties in the simulated distributions that are compared to the measurements, see e.g. [216, 217]. Concretely, we construct six additional templates for $G(A, E, X_{\mathrm{max}}^{\mathrm{rec}}|\theta_{\mathrm{accept.}}, \theta_{\mathrm{resol.}}, \theta_{\mathrm{scale}})$ that correspond to the separate variations of resolution, acceptance and X_{max} scale by $\pm 1\sigma$. This is is done in every bin of A, E and X_{max}. Using a cubic spline to interpolate between the templates and a linear extrapolation beyond, the modification according to the variation θ of one systematic uncertainty reads

$$\Delta G(\theta) = \begin{cases} (G_+ - G_0)\theta & \theta > 1 \\ (G_+ - G_-)\theta + (G_+ + G_- - 2G_0)\theta^2 \pm (2G_0 - G_+ - G_-)\theta^3 & |\theta| < |\pm 1| \\ (G_0 - G_-)\theta & \theta < -1 \end{cases} , \tag{6.11}$$

where G_0 denotes the nominal template and $G_\pm(\theta)$ the variation by $\pm 1\sigma$. Since the effects of the uncertainties are approximately independent, we can separately account for them, using

$$G = G_0 + \Delta G(\theta_{\mathrm{acceptance}}) + \Delta G(\theta_{\mathrm{resolution}}) + \Delta G(\theta_{\mathrm{scale}})\,. \tag{6.12}$$

While clearly of approximative nature, the template morphing technique represents an effective method to calculate the required continuous variations of the detector response with acceptable precision.

6.3 Statistical Evaluation

The present analysis aims to estimate the free parameters of the astrophysical model in section 6.1 with a combined fit of the energy spectrum and the X_{max} distributions. In the context of parameter estimation, Bayesian inference provides an explicit formalism to translate the uncertainties about model and measurement into uncertainties of the parameter estimates. To this end the likelihood $P(\boldsymbol{d}|\boldsymbol{\theta}, H) \equiv \mathscr{L}(\boldsymbol{\theta})$ is formulated, which is the conditional probability to observe the data $\boldsymbol{d}$, assuming the model H with a set of parameters $\boldsymbol{\theta}$ is true. Additionally, the prior distribution $P(\boldsymbol{\theta}|H) \equiv \pi(\boldsymbol{\theta})$ of the model parameters is stated, containing the prior knowledge or a subjective degree of belief about the parameter values. The posterior probability of the model parameters given the data is then obtained from Bayes' theorem:

$$P(\boldsymbol{\theta}|\boldsymbol{d}, H) = \frac{P(\boldsymbol{d}|\boldsymbol{\theta}, H)\, P(\boldsymbol{\theta})}{P(\boldsymbol{d}|H)} \propto \mathscr{L}(\boldsymbol{\theta})\, \pi(\boldsymbol{\theta}) \ . \tag{6.13}$$

The denominator normalizes the posterior probability over the model parameters

$$P(\boldsymbol{d}|H) = \int_{\Omega_\theta} \mathscr{L}(\boldsymbol{\theta})\, \pi(\boldsymbol{\theta}) d\boldsymbol{\theta} \ . \tag{6.14}$$

and is called the Bayesian evidence, because it quantifies how much the data favor the model, which can be used for model selection, see section 6.3.5. Since $P(\boldsymbol{d}|H)$ is independent of $\boldsymbol{\theta}$, it can be ignored for purpose of parameter estimation.

The posterior probability $P(\boldsymbol{\theta}|\boldsymbol{d}, H)$ contains the full information about the model parameters that can be inferred from the measurement. For a given parameter θ_i the 'marginalized' posterior probability $P(\theta_i|\boldsymbol{d}, H)$ can be calculated by integrating out all other parameters. As previously described, systematic uncertainties can be included in form of nuisance parameters, which are not of prime interest. To arrive at a posterior probability that only depends on the parameters of interest, all nuisance parameters are integrated out. This treatment exploits the full information about the effects of the nuisance parameters that is encoded in the likelihood function. Thereby, the influence of the systematic uncertainties is reduced by letting the measurement constrain their values.

The main features of a posterior distribution can be summarized using a point estimate and an interval estimate. As point estimate, the maximum of the posterior distribution

$$\hat{\boldsymbol{\theta}} = \arg\max_{\theta} P(\boldsymbol{\theta}|\boldsymbol{d}, H) \tag{6.15}$$

represents the set of parameters with the highest posterior density. For a uniform prior distribution $\pi(\boldsymbol{\theta}) = \text{const}$, it coincides with the maximum likelihood estimate that is used in frequentist methods. Since $\hat{\boldsymbol{\theta}}$ is not necessarily representative of the posterior distribution a more often used point estimate is the posterior mean

$$\langle\boldsymbol{\theta}\rangle = \int_{\Omega_\theta} \boldsymbol{\theta} P(\boldsymbol{\theta}|\boldsymbol{d}, H) d\boldsymbol{\theta} \ , \tag{6.16}$$

indicating the average parameter values in the posterior distribution.

Uncertainty intervals, which are called 'credibility intervals' in the context of Bayesian inference, reflect the range of posterior parameter values. The most often considered representation is the 'highest posterior density' interval, which is the smallest interval containing the true parameter value with a given probability. Consequently, every point in this interval has a higher probability than any point outside.

In the following sections we first formulate the likelihood function by identifying the statistical distributions that the observables follow. As a second step, the ‘deviance’ is introduced which provides a measure for the goodness-of-fit of the astrophysical scenario. Then, the prior distributions of the model parameters are stated. Finally, we describe the used computational method for evaluating the posterior distribution.

6.3.1 Likelihood

The first step in the statistical inference is the formulation of the likelihood function. In the combined fit the likelihood needs to quantify the agreement with both the energy spectrum and $X_{\max}$ distributions.

In the energy spectrum the number of events in each energy bin represents a counting experiment and thus follows a Poisson distribution. The likelihood reads

$$\mathscr{L} = \mathrm{Poisson}(\boldsymbol{k}|\boldsymbol{\lambda}) = \prod_m \frac{(\lambda_m)^{k_m}}{k_m!} e^{-\lambda_m} \ , \tag{6.17}$$

where $\lambda_m = N_{\mathrm{spectrum}}(E^m_{\mathrm{rec}})$ is the predicted number of events according to equation 6.7, k_m is the measured number of events, and m denotes the energy bin.

The vertical and inclined event reconstructions represent two independent measurements of the energy spectrum, $\boldsymbol{k}_V$ and $\boldsymbol{k}_I$. In principle this leaves the choice of fitting both spectra individually, $\mathscr{L} = \mathrm{Poisson}(\boldsymbol{k}_V|\boldsymbol{\lambda}_V) \cdot \mathrm{Poisson}(\boldsymbol{k}_I|\boldsymbol{\lambda}_I)$, or fitting to the summed spectrum, $\mathscr{L} = \mathrm{Poisson}(\boldsymbol{k}_V + \boldsymbol{k}_I|\boldsymbol{\lambda}_V + \boldsymbol{\lambda}_I)$. Since the shape of the model prediction for the vertical and inclined reconstructions are not identical due to the different detector response, the correct approach is to fit these spectra separately. Otherwise, when fitting to the summed spectrum, the exposures of the two analyses could get varied in such a way that the agreement with the individual measurements deteriorate, while the agreement with summed measurement improves. Hence, the spectrum likelihood is given by

$$\mathscr{L}_{\mathrm{spectrum}} = \mathrm{Poisson}(\boldsymbol{k}^V|\boldsymbol{\lambda}^V) \cdot \mathrm{Poisson}(\boldsymbol{k}^I|\boldsymbol{\lambda}^I) \ . \tag{6.18}$$

Note that a sufficiently large number of events in most energy bins would also allow for using the normal approximation of the Poisson distribution. Under the normal approximation the likelihood approach then corresponds to a χ^2 minimization as used in the astrophysical fits in e.g. [81, 84]. However, the approximation does not hold for the low number of events at $E > 19.9\,\mathrm{eV}$, where the Gaussian probability deviates from the correct Poisson probability. Specifically, the fitted energy range spans from $E = 10^{18.7}\,\mathrm{eV}$ to $10^{20.2}\,\mathrm{eV}$ for both measurements. For the inclined energy spectrum this includes two energy bins with zero observed events.

In the $X_{\max}$ distributions $\boldsymbol{k}_m$, the total number of events $n_m = \sum_x k_{mx}$ per energy E_m is also a measure of the spectral flux. Since the information on the spectral flux is already captured by the spectrum likelihood, we consider the total number of events in each distribution as fixed, i.e. $n_m = \mathrm{const} \ \forall \ m$. With a fixed number of events these binned distributions then each follow a multinomial distribution and the likelihood is given by the product of all energy bins, reading

$$\mathscr{L}_{X_\mathrm{max}} = \prod_m \mathrm{Multinomial}\left(\boldsymbol{k}_m|n_m, \boldsymbol{p}_m\right) = \prod_m n_m! \prod_x \frac{1}{k_{mx}!}(p_{mx})^{k_{mx}} \ . \tag{6.19}$$

Here $p_{mx} = p(E^m, X_\mathrm{max}^{\mathrm{rec};x})$ is the predicted probability to observe an event in the X_max bin x given by equation 6.9.

A similar binned likelihood was also used in the empirical analysis of the present X_max distributions in [34]. An alternative to this method would be an unbinned likelihood approach, which generally has a higher statistical power. However, the increased computational complexity of evaluating the model probability for up to 20.000 events in each fitting step is impractical for the present statistical evaluation. Moreover, the binned likelihood has the additional advantage of providing deviance as a direct goodness-of-fit measure, see below. This is not possible for an unbinned likelihood, where one has to resort to p-values.

In the aforementioned formulation, the energy spectrum and the X_max distributions are independent observables. Therefore, the total likelihood is given by the product

$$\mathscr{L} = \mathscr{L}_\mathrm{spectrum} \cdot \mathscr{L}_{X_\mathrm{max}} \ . \tag{6.20}$$

6.3.2 Deviance

To assess the goodness-of-fit we make use of a generalized χ^2-method [28], which is often referred to as deviance [218]. The deviance is defined as the log-likelihood ratio between the considered model and the 'saturated model' that perfectly describes the data

$$D = -2\ln\frac{\mathscr{L}}{\mathscr{L}^\mathrm{sat}} \ . \tag{6.21}$$

The likelihood of the saturated model $\mathscr{L}^\mathrm{sat}$ is obtained by replacing the modeled number of events with the observed number of events, and is constant. Concretely, the saturated likelihood values for the energy spectrum and X_max distributions read

$$\mathscr{L}^\mathrm{sat}_\mathrm{spectrum} = \mathrm{Poisson}(\boldsymbol{k}^V|\boldsymbol{\lambda}^V = \boldsymbol{k}^V) \cdot \mathrm{Poisson}(\boldsymbol{k}^I|\boldsymbol{\lambda}^I = \boldsymbol{k}^I) \tag{6.22}$$

$$\mathscr{L}^\mathrm{sat}_{X_\mathrm{max}} = \prod_m \mathrm{Multinomial}(\boldsymbol{k}_m|n_m, \boldsymbol{p}_m = \boldsymbol{k}_m/n_m)) \ . \tag{6.23}$$

Since the total likelihood is given by the product of the energy spectrum and X_max likelihood functions, the total deviance can be expressed as the sum of the corresponding deviances,

$$D = -2\ln\frac{\mathscr{L}_\mathrm{spectrum}}{\mathscr{L}^\mathrm{sat}_\mathrm{spectrum}} - 2\ln\frac{\mathscr{L}_{X_\mathrm{max}}}{\mathscr{L}^\mathrm{sat}_{X_\mathrm{max}}} = D_\mathrm{spectrum} + D_{X_\mathrm{max}} \ . \tag{6.24}$$

Using equations 6.18, 6.22 for the energy spectrum and equations 6.19, 6.23 for the $X_{\max}$ distributions the deviances simplify to

$$
\begin{aligned}
D_{\text{spectrum}} &= -2 \ln \left(\frac{\prod_m (\lambda_m)^{k_m} e^{-\lambda_m} / k_m!}{\prod_m (k_m)^{k_m} e^{-k_m} / k_m!} \right) \\
&= 2 \sum_m \underbrace{k_m \ln(k_m/\lambda_m)}_{\overset{!}{=}\, 0 \text{ for } k_m = 0} + \lambda_m - k_m
\end{aligned}
\tag{6.25}
$$

$$
\begin{aligned}
D_{X_{\max}} &= -2 \ln \left(\frac{\prod_m n_m! \prod_x (p_{mx})^{k_{mx}} / k_{mx}!}{\prod_m n_m! \prod_x (k_{mx}/n_m)^{k_{mx}} / k_{mx}!} \right) \\
&= 2 \sum_m \sum_x \underbrace{k_{mx} \ln(k_{mx}/n_m p_{mx})}_{\overset{!}{=}\, 0 \text{ for } k_{mx} = 0} \; .
\end{aligned}
\tag{6.26}
$$

Note that equation 6.25 shows the contribution of only one of the two measurements, vertical or inclined, for better readability. The cases for bins with zero observed entries are specifically noted. As can be seen, empty bins do not contribute to the multinomial deviance, whereas the Poisson deviance still sees a contribution. If the number of entries is fixed $\sum \boldsymbol{k} = \sum \boldsymbol{\lambda}$, the Poisson and multinomial deviances coincide, which is expected as this corresponds to the additional requirement of a multinomial distribution.

According to Wilks' theorem, for sufficiently large samples $\boldsymbol{k}$ the minimum deviance follows a χ^2 distribution [28]. In the analysis performed in the next chapter, the minimum deviances of the individual observables are indicated together with the corresponding number of data points, D/n. Above $E > 10^{18.7}$ eV the vertical and inclined energy spectra have $n = 15$ and 13 non-zero bins, respectively. In this energy range, the $X_{\max}$distributions consist of $n = 110$ non-zero bins. The total goodness-of-fit of the astrophysical model can be assessed with the deviance per degrees of freedom $D/ndf = (D_{\text{spectrum}} + D_{X_{\max}})/(15 + 13 + 110 - n_p)$, where n_p is the number of model parameters. To obtain reference values for these goodness-of-fits, we fit the measurements with suitable empirical models in appendix A.3.

6.3.3 Prior Distributions

The second step for the Bayesian inference is to assign a joint prior probability $\pi(\boldsymbol{\theta})$ to both the model parameters and the nuisance parameters, which represent the systematic uncertainties. This probability can be based on prior knowledge or, in the absence of such knowledge, on a degree of belief about different parameter values. In scientific analyses the aim is to obtain objective results that do not depend on a subjective belief of the experimenter. Therefore, one tries to identify uninformative or objective priors, which only contain objective information, e.g. that the elemental fractions at the source cannot be negative. To this end we assign the following prior probabilities to the parameters of the astrophysical model in section 6.1.1, see table 6.2, which we consider objective in the context of the model.

For the cutoff rigidity R_{cut} we assign a logarithmically flat prior, motivated by the logarithmic energy binning of the observables. Previous analyses and tests showed that the most likely parameter is contained with the chosen range $R_{\text{cut}} = 10^{17.8} - 10^{20.5}$ eV. A requirement

Source parameter	Prior distribution
Maximum rigidity	$\pi(\log_{10}(R_{\rm cut}/{\rm eV}))$ = Uniform(17.8, 20.5)
Power law index	$\pi(\gamma)$ = Uniform(-3, 3)
Elemental fractions	$\pi(a(\boldsymbol{Z_0})) = {\rm const}$
Flux constant	uninformative
Nuisance parameter	**Relative uncertainty**
Relative exposure	4.1%
Energy scale	14%
Xmax acceptance	parametrized, see figure 6.6
Xmax resolution	parametrized, see figure 6.6
Xmax scale	parametrized, see figure 6.6

Table 6.2: Overview of the astrophysical model parameters (section 6.1.1) with their priors distributions, and of the nuisance parameters representing the experimental systematic uncertainties (section 6.2.3) with their relative uncertainties.

for the minimum $R_{\rm cut}$value is that the sources have to be able to accelerate cosmic rays to the highest observed energies, i.e. $E \sim 10^{20.15}\,{\rm eV}$. In case of the smooth cutoff functions in equation 6.2 this is possible for all $R_{\rm cut}$values, whereas the hard cutoff requires a minimum value of $R_{\rm cut} \approx 10^{20.15}/26\,{\rm eV} \approx 18.75\,{\rm eV}$. With the chosen upper bound of $R_{\rm cut} = 20.5\,{\rm eV}$ at least one order of magnitude is left for modeling a smooth cutoff beyond the maximum rigidity, considering that maximum rigidity of the simulated events is $R_{\rm cut} = 10^{21.5}\,{\rm eV}$.

For the power law index of the source spectrum $dN/dE \propto E^{-\gamma}$ we assign a uniform prior in the range of $\gamma = (-3) - (3)$. This range covers the predictions for one shot acceleration by pulsars with $\gamma \sim 1$, Fermi acceleration with $\gamma \sim 2 - 2.3$, as well as the effective index $\gamma > 2$ for a population of Fermi accelerators with exponentially distributed maximum rigidities. We explicitly extend the range down to negative γ values, which corresponds to a source emissivity that increases with energy. While such a behavior does not correspond to any of the predictions described in section 3.1, we found that these extreme spectral indices maximize the likelihood in some of the tested scenarios.

The elemental fractions at the source need to satisfy $\sum_i a(Z_0^i) = 1$, hence they are not mutually independent. To achieve a uniform sampling under this side condition we apply a method described in [219] and used in [214]. The method consist of sampling a corresponding number of independent unit-exponential random variables, which are then normalized to their sum:

$$a(Z_0^i) = \zeta_i / \sum_i \zeta_i \quad \text{with } f(\zeta_i) = \text{Exponential}(\zeta_i | \beta = 1). \tag{6.27}$$

This realizes a uniform prior, $\pi(a(\boldsymbol{Z_0})) = {\rm const}$, with each set of elemental fractions having the same probability.

The flux normalization ϕ_0 only contributes to the spectrum observable, because the number of events in the $X_{\rm max}$ distributions is considered fix. Thus, in contrast to the other model parameters, an analytic expression can be found for the optimum flux normalization. This is done by drawing ϕ_0 out of the model expectation for the vertical and inclined spectra, $\boldsymbol{\lambda} = \phi_0 \tilde{\boldsymbol{\lambda}}$, and setting the derivate of the spectrum deviance (equation 6.25) to zero:

$$D = 2 \sum_m k_m^V \ln \left(\frac{k_m^V}{\phi_0 \tilde{\lambda}_m^V} \right) + k_m^I \ln \left(\frac{k_m^I}{\phi_0 \tilde{\lambda}_m^I} \right) + \phi_0 \left(\tilde{\lambda}_m^V + \tilde{\lambda}_m^I \right) - k_m^V - k_m^I \tag{6.28}$$

$$\frac{dD}{d\phi_0} = 2 \sum_m -\frac{1}{\phi_0} (k_m^V + k_m^I) + \tilde{\lambda}_m^V + \tilde{\lambda}_m^I \overset{!}{=} 0 \tag{6.29}$$

$$\phi_0 = \sum_m (k_m^V + k_m^I) / \sum_m (\tilde{\lambda}_m^V + \tilde{\lambda}_m^I) \ . \tag{6.30}$$

Using this expression the flux normalization is set to its maximum likelihood value in each fitting step. This procedure corresponds to assigning an uninformative prior distribution. Uninformative priors are equivalent to uniform prior, except for a missing normalization to the size of the allowed parameter range.

In contrast to the parameters of the astrophysical model, a well defined prior knowledge is available for the systematic uncertainties of the detector response. Corresponding to the implementation described in section 6.2.3, each systematic uncertainty is assigned a standard normal distributed nuisance parameter

$$\pi(\theta_i^{\mathrm{sys}}) = \mathrm{Normal}(\theta_i^{sys} | \mu = 0, \sigma = 1) \ . \tag{6.31}$$

Hence, the value of θ^{sys} represents the deviation of the nominal detector response in units of standard deviations.

All model and nuisance parameters are assumed to be independent. Therefore, the joint prior probability is given by the product

$$\pi(\boldsymbol{\theta}) = \pi(R_{\mathrm{cut}}) \cdot \pi(\gamma) \cdot \pi(a(\boldsymbol{Z_0})) \cdot \prod_i \pi(\theta_i^{\mathrm{sys}}) \ . \tag{6.32}$$

As a consequence of the uniform prior distributions for the astrophysical model parameters, the maximum of the posterior distribution $\hat{\boldsymbol{\theta}}$ corresponds to the maximum likelihood estimate, as long as the nuisance parameters are not considered, i.e. for $\theta_i^{\mathrm{sys}} = 0$. In this case $\hat{\boldsymbol{\theta}}$ is comparable to the point estimates reported in frequentist analyses, e.g. [87].

6.3.4 Markov-Chain Monte Carlo

The posterior distribution in equation 6.13 quickly becomes difficult to evaluate analytically for a dimensionality of the parameter space > 3. For higher dimensionalities, Bayesian inference is enabled by computational techniques that allow to efficiently sample from the posterior distribution. The most widely used general technique is the Markov chain Monte Carlo (MCMC) method. A Markov chain is a series of samples, randomly drawn from a varying distribution that approaches the desired target distribution as the number of samples becomes sufficiently large. MCMC methods produce such Markov chains of the model parameters $\boldsymbol{\theta}_s$, $s = 1...n$, which have the posterior as target distribution. Point and interval estimates are readily obtained from these chains, e.g. $\langle \boldsymbol{\theta} \rangle = 1/n \sum_s \boldsymbol{\theta}_s$.

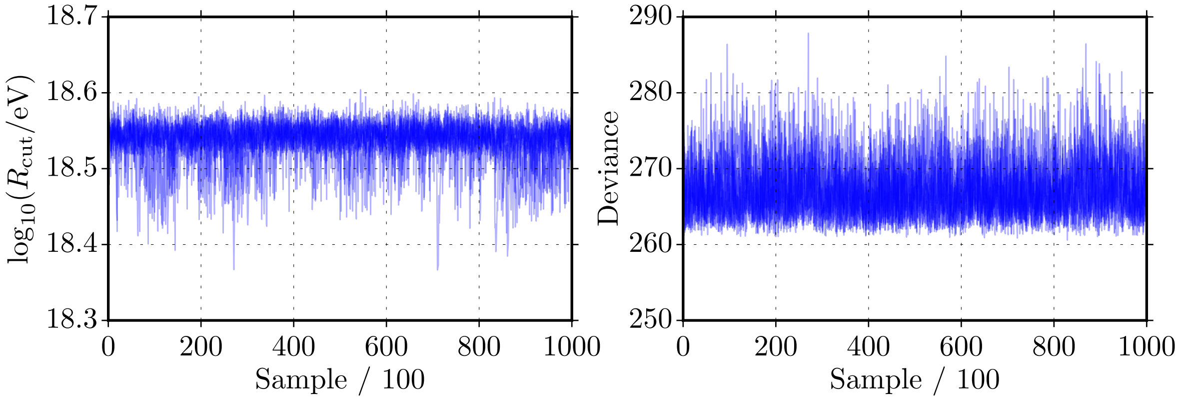


Figure 6.7: Overlay of ten Markov chains showing the traces every hundredth value sampled from the posterior distribution. The left figure shows the cutoff rigidity and the right figure the corresponding deviances. The Markov chains are seen to have reached a stationary behavior.

The process of approaching the target distribution loosely corresponds to a standard model optimization, i.e. parameters are identified which optimize the agreement with the measurement. However, instead only searching for the optimum parameters, MCMC methods explore the probability density around the maximum and thereby sample the full information about the parameters that is contained in the posterior distribution.

In this analysis we use the software package **PyMC**, version 2.3 [220], which provides the statistical framework and several MCMC algorithms. In the present analysis we employ a standard Metropolis-Hastings algorithm. For each tested scenario, we generate a minimum of ten Markov chains starting from suitable random initial conditions Each chain passes an initial tuning phase until the target distribution is approximately reached, before sampling 100 000 values from the posterior distribution.

We verify that the traces of the deviance show a stationary behavior and that each chain has converged to the optimum parameter range, see figure 6.7. Additionally, we test the posterior distribution of each model parameter for convergence using the Gelman-Rubin $\hat{R}$ statistic [221]. This statistic uses the variances within the chains and between chains to formulate an upper limit to the level of convergence. We require that the generated chains satisfy $\hat{R} \leq 1.05$ for each parameter, indicating a convergence to within a relative uncertainty of 5%. All chains are then combined for the following evaluation. With this procedure the computational uncertainties on the resulting point and interval estimates of the parameters are small compared to the stated precision.

6.3.5 Model Selection

Bayesian model selection between two competing models H_1 and H_0 can be achieved by comparing their respective posterior probabilities given the data [222], reading

$$\frac{P(H_1|\boldsymbol{d})}{P(H_0|\boldsymbol{d})} = \frac{P(\boldsymbol{d}|H_1)}{P(\boldsymbol{d}|H_0)} \frac{P(H_1)}{P(H_2)} \,. \tag{6.33}$$

$P(\boldsymbol{d}|H_i)$ is the Bayesian evidence (see equation 6.14) for the model H_i and $P(H_i)$ its prior probability. When both models are a priori equally probable, $P(H_1) = P(H_0)$, the ratio of the posterior probabilities simplifies to the ratio of the Bayesian evidences, which is called the Bayes factor

$$B_{10} = \frac{P(\boldsymbol{d}|H_1)}{P(\boldsymbol{d}|H_0)} = \frac{\int \mathscr{L}(\boldsymbol{\theta}|H_1)\,\pi(\boldsymbol{\theta}|H_1)d\boldsymbol{\theta}}{\int \mathscr{L}(\boldsymbol{\theta}|H_0)\,\pi(\boldsymbol{\theta}|H_0)d\boldsymbol{\theta}}\,. \tag{6.34}$$

Thus, compared to a likelihood ratio test, the likelihood is averaged over the model parameters instead of maximized. The evidence is thus generally larger for models with high likelihood values over a large range parameters values, than for models with sharply peaked likelihood functions. This can be seen as penalizing fine-tuning a model to the observed data [223]. Table 6.3 lists a categorization for model selection in terms of Bayes factors.

$2\ln(B_{10})$	B_{10}	Evidence against H_0
$0-2$	$1-3$	Not worth more than a bare mention
$2-6$	$3-20$	Positive
$6-10$	$20-150$	Strong
>10	>150	Very strong

Table 6.3: Suggested categories for Bayes factors from [222].

Unfortunately, the integral in the Bayesian evidence is difficult to compute and is not directly accessible from the traces of posterior samples that are produced by MCMC methods. In this case approximative methods such as the ‘deviance information criterion’ (DIC, [224]) or the ‘Bayes predictive information criterion’ (BPIC, [225]) can be employed to estimate the Bayesian evidence from the trace of deviance values obtained in the MCMC optimization. In the following we will use the latter and indicate

$$\Delta S_{10} = 2\ln(\mathrm{BPIC}(H_1)) - 2\ln(\mathrm{BPIC}(H_0)) \approx 2\ln(B_{10}) \tag{6.35}$$

when performing variations of the astrophysical model.

Chapter 7

Tests of Astrophysical Cosmic-Ray Scenarios

In this chapter we perform a combined fit to energy the energy spectrum and $X_{\max}$ distributions of the Pierre Auger Observatory in the order to constrain the parameters of the cosmic ray source model as described in the previous chapter. These constraints can help us to assess key open questions in the astrophysical interpretation of ultra-high energy cosmic rays. This includes the source properties and the nature of the observed suppression of the cosmic-ray flux above 40 EeV. Answering these question would be a major step towards identifying the elusive sources of the highest energy cosmic rays: Firstly, a comparison of the fitted source parameters with the predictions from accelerator theories puts constraints on the type of astrophysical objects responsible for the acceleration. Secondly, providing information on the cosmic-ray rigidity *during* propagation would help in disentangling the magnetic deflections from the arrival directions in order to locate the sources.

To this end, we first perform the analysis for the specified baseline scenario (section 7.1) and describe the features of the found solutions. Then the impact of main simulation uncertainties are assessed (section 7.2). In order to test the validity of the derived conclusions with respect to the simplifications in the source model we consider several scenario variations (section 7.3). Finally, we describe implications of the obtained scenario (section 7.4) and summarize our findings in (section 7.5).

7.1 Baseline Scenario

In the baseline scenario we consider a uniform distributions of sources injecting hydrogen, helium, nitrogen and iron nuclei according to a power-law energy spectrum with a broken-exponential cutoff. The cosmic-ray propagation from the sources to the observation is simulated including all relevant interactions with the photon background. Here we consider the extragalactic background light (EBL) model by Gilmore *et al.* and use the photodisintegration cross sections computed by the TALYS software. For comparing with measured $X_{\max}$ distributions the cosmic-ray air shower properties are simulated using EPOS-LHC as hadronic interaction model. Details on the source and propagation model are found in sections 6.1.1 and 6.1.2, respectively.

7.1.1 Likelihood scan

For a qualitative understanding of the constraints that are imposed on the model parameters through the measurements it is useful to first assess the main properties of the model likelihood function. To this end we scan the likelihood in the two-dimensional parameter plane of the

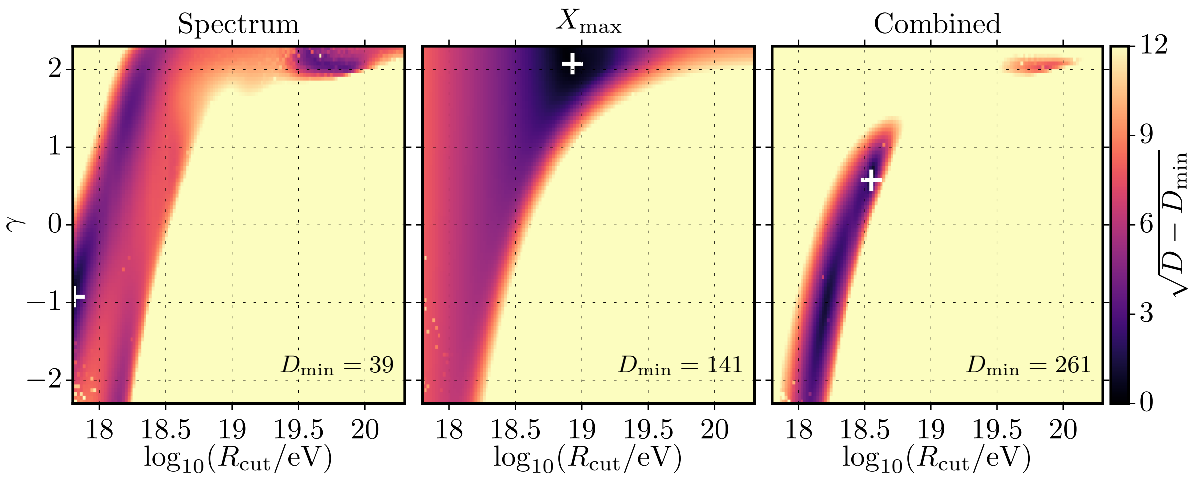


Figure 7.1: Scan of the deviance in the (R_{cut}, γ) plane, considering only the energy spectrum (left), the $X_{\max}$ distributions (middle), and both observables simultaneously (right). The color code gives the number of pseudo standard deviations $n_\sigma = \sqrt{D - D_{\min}}$ with respect to the optimum value $D_{\min}$ found at the white cross. The combined fit results in an extended valley at low cutoff rigidities and second disjunct region around $R_{\mathrm{cut}} \approx 10^{19.9}$ eV.

cutoff rigidity R_{cut} and power-law index γ, which to a large extent define the shape of the total source spectrum. The scan is performed in steps of $\Delta \log_{10}(R_{\mathrm{cut}}/\mathrm{eV}) = 0.02$ and $\Delta\gamma = 0.05$. For each (R_{cut}, γ)-pair the likelihood is maximized with respect to the elemental fractions $a(Z_0)$ and the flux normalization ϕ_0. Systematic uncertainties are not considered in this scan, however, we vary the relative exposure of the vertical and inclined energy spectrum measurements to a value that better describes the ratio of the measurements, see section 7.1.3.

Since the deviance is a negative log-likelihood ratio, maximizing the likelihood corresponds to minimizing the deviance. The total deviance is the sum of the deviances of the energy spectrum and of the $X_{\max}$ distributions, $D = D_{\mathrm{spectrum}} + D_{X_{\max}}$. In order to assess the individual contributions by each observable we separately scan the corresponding deviances in addition to the total deviance. The results are shown in figure 7.1 in units of $n_\sigma = \sqrt{D - D_{\min}}$, which, according to Wilk's theorem, can be interpreted as the number of (pseudo) standard deviations in the large sample limit. Both the energy spectrum and $X_{\max}$ distributions are seen to separately exclude the lower right part of the parameter space, i.e. high cutoff rigidities and negative γ values. However, the shape of the deviance contours largely differs between the energy spectrum and the $X_{\max}$ distributions, and the minimum deviances are found at different positions (marked by a white cross). The energy spectrum favors a cutoff rigidity at the lower boundary of the scanned parameter range and a spectral index $\gamma \approx -1$, but also finds a solution at $R_{\mathrm{cut}} \approx 10^{19.7}$ eV and $\gamma \approx 2$, whereas the $X_{\max}$ deviance exhibits a minimum at $R_{\mathrm{cut}} \approx 10^{19}$ eV and $\gamma \approx 2$. Overall, the $X_{\max}$ observable is seen to less strongly constrain the (R_{cut}, γ) plane, which is understandable since in the considered formulation the $X_{\max}$ likelihood is not sensitive to the relative number of events in the distributions at different energies, and thus to the energy spectrum.

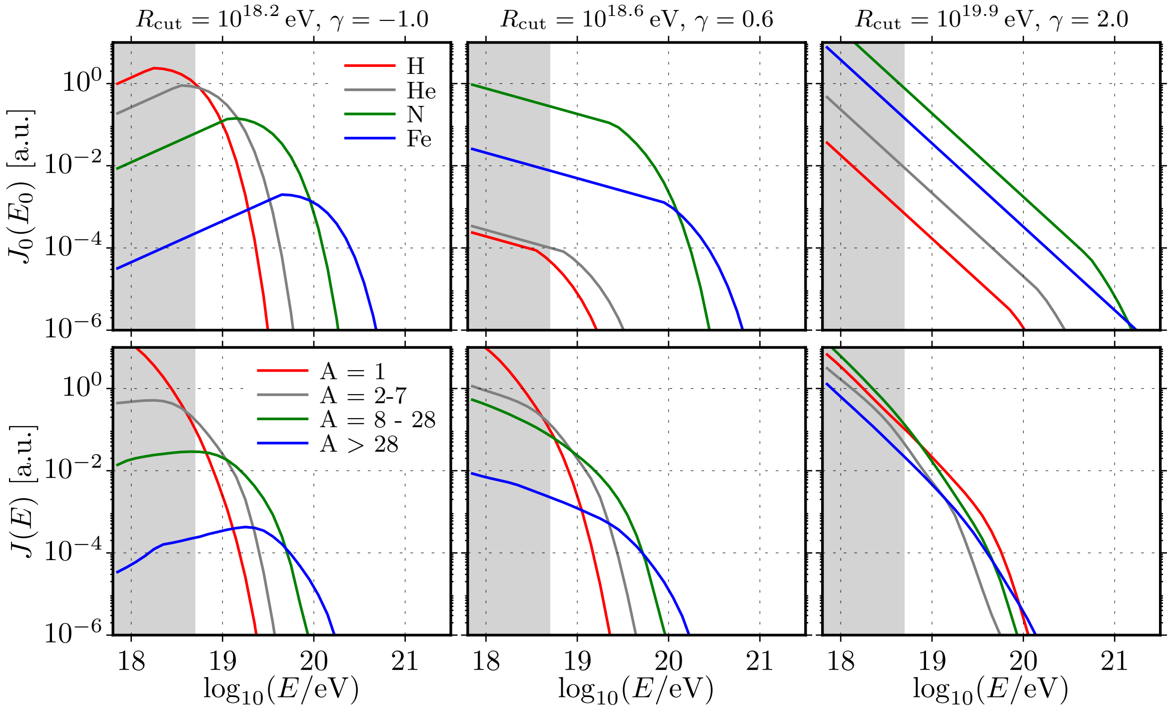


Figure 7.2: Comparison of the injected (top row, by injected element) and observed energy spectrum (bottom row, by observed mass groups) for different points in the $(R_{\rm cut}, \gamma)$ plane: the global minimum (middle), a second parameter point in the valley (left) and the disjunct second minimum (right). The gray area indicates the energy range $E < 10^{18.7}$ eV not considered in the fit.

Obviously, the strongest constraints are set by considering the total deviance of both observables, as seen in the right of figure 7.1. Simultaneously fitting both observables necessarily limits the parameter space to regions where each individual observable can be well fitted. However, since for a given ($R_{\rm cut}$, γ)-pair the observables may require different elemental fractions $a(Z)$ to be well described, the constraints in the combined fit are stronger than the superposition of the individual deviance contours. In particular, both the individual minima are excluded and the total minimum deviance is significantly larger than the sum of the individual deviances. The global minimum is found at $R_{\rm cut} \approx 10^{18.6}$ eV and $\gamma \approx 0.6$ within an extended valley that spans over a spectral index from $\gamma \approx 1$ down to the lower boundary of the scanned parameter range. A second disjunct region is found at a high cutoff rigidity of $R_{\rm cut} \approx 10^{19.9}$ eV and $\gamma \approx 2$, although with a significantly higher deviance corresponding to $\sqrt{D - D_{\rm min}} > 7$ pseudo standard deviations. This region, called the second minimum hereafter, will be separately investigated in the following, due to the accordance of the spectral index with the prediction from Fermi acceleration.

The implications of different parameter points in the $(R_{\rm cut}, \gamma)$ plane are visualized in figure 7.2, where the global minimum (middle column) is compared to an arbitrary second parameter point within the extended valley (left column) and to the disjunct parameter region at $R_{\rm cut} \approx$

$10^{19.9}$ eV and $\gamma \approx 2$(right column). In each case the injected and predicted observed spectra are shown, which allows to assess the importance of interactions during propagation and the amount of secondary particles.

The shown solution at $R_{\text{cut}} = 10^{18.2}$ eV (left of figure 7.2) corresponds to a typical Peters cycle, with a dominating proton component giving way to successively heavier elements of decreasing abundance. In this case the spectral shape of each component before its respective cutoff is hidden by the previously dominating lighter component and the power-law index of the total source spectrum is only visible below the cutoff energy of the lightest component. For cutoff rigidities $R_{\text{cut}} < 10^{18.7}$ eV this occurs below the fitted energy range, and thus the power-law index γ is not directly visible in the observed energy spectrum. Instead, the extreme values of $\gamma < 0$, which denotes that the injected flux increases with energy, cause a delay of the exponential source cutoff in order to still reproduce the observed spectrum, despite the very early cutoff. The power-law index and cutoff rigidity are thus positively correlated under the measurements, which leads to the elongated valley in the likelihood scan. Since the valley is almost flat for $\gamma < 0$, one may say that the fit loses sensitivity to the spectral index for cutoff rigidities that are far below the considered energy range.

In contrast, in the global minimum at $R_{\text{cut}} \approx 10^{18.55}$ eV (middle of figure 7.2) the power-law index is visible in the total source spectrum within the fitted energy range, because the proton and helium components are subdominant to the nitrogen component. Comparing the injected and observed energy spectrum, the observed spectral shape around the ankle is seen to be largely determined by a flux of light particles that is not present in the source spectrum. These are secondary particles from photodisintegration, which are able to dominate the flux of primary particles at lower energies due to the hard spectral index. The observed energy spectrum is thus determined by a combination of interactions and the maximum acceleration capability of the sources. Towards lower R_{cut} parameter values the level of photodisintegration decreases and the importance of the source cutoff increases accordingly.

Conversely, at the second minimum at $R_{\text{cut}} \approx 10^{19.9}$ eV the observed flux suppression can be mostly attributed to energy losses during propagation, because the cutoff energies of all injected species are in excess of the observed flux suppression energy at around $10^{19.6}$ eV. Here, the soft spectral index $\gamma \approx 2$ implies that the observed spectrum keeps a better memory of the source spectrum, because the secondary particles from interactions do not overwhelm the primary particles at lower energies. For this reason the spectral index is well constrained around $\gamma \sim 2$, with the difference to the observed spectral shape above the ankle $dN/dE \propto E^{-2.7}$ stemming from energy losses. Note that a similar difference between injected and observed spectral index is predicted in the dip model of a pure proton composition [78, 79].

In general while the amount of interactions with the extragalactic photon background, especially of photodisintegration, increases with higher cutoff rigidities, the *importance* of photodisintegration for the astrophysical fit is highest in the global minimum, because here the light secondary particles give rise to dominant contributions to the cosmic-ray flux within the fitted energy range.

7.1.2 Statistical Analysis

In the following we perform the statistical inference of the baseline scenario with the method described in the previous chapter. Systematic uncertainties are not yet considered, again with the exception of the relative exposure in the vertical and inclined energy spectrum measurements. Thus, the astrophysical fit is first performed with the nominal detector response, in particular regarding the systematic uncertainties on the energy scale and the $X_{\max}$ scale of the experiment. Due to the use of uniform prior distributions for the model parameters, the results can be compared to those obtained with frequentist analyses, e.g. [87, 226, 227].

The best fit (maximum likelihood) to the energy spectrum and $X_{\max}$ distributions is shown in figure 7.3. The comparison with the measurement shows a qualitatively good fit. Quantitatively, the minimum deviance $D_{\min} = 260.4$ considerably exceeds the number of degrees of freedom $ndf = 132$: 138 data points minus 6 model parameters, namely R_{cut}, γ, the flux normalization and 3 independent fractions due to the side condition $\sum_i a(Z_i) = 1$. This is to large part due to the vertical and inclined energy spectrum, for which the model achieves a deviance of $D_{\mathrm{spectrum}}/n = (51 + 44)/(15 + 13)$ per number of data points. For reference, simultaneously fitting both spectra with an empirical flux model results in a deviance of $D^{\mathrm{ref}}_{\mathrm{spectrum}} = 56$, see appendix A.3. The $X_{\max}$ distributions are relatively better described with a deviance of $D_{X_{\max}} = 165.5$ for 110 data points. The largest individual contributions arise in the energy bin $10^{18.7-18.8}$ eV, where the predicted distribution is slightly shifted towards higher $X_{\max}$ values, and in the bin $10^{19.4-19.5}$ eV, where the predicted distribution is shifted towards lower $X_{\max}$ values and is narrower than measured. The reference value from fitting the $X_{\max}$ distributions with individual mass fractions in each energy bin is $D^{\mathrm{ref}}_{X_{\max}} = 122$ at the cost of introducing 27 parameters, see appendix A.3.

	First minimum		Second minimum	
Parameter	$\hat{\theta}$	$\langle\theta\rangle\pm\sigma$	$\hat{\theta}$	$\langle\theta\rangle\pm\sigma$
$\log_{10}(R_{\mathrm{cut}}/\mathrm{eV})$	18.56	$18.54^{+0.02}_{-0.01}$	19.88	$19.87^{+0.02}_{-0.02}$
γ	0.62	$0.59^{+0.10}_{-0.07}$	2.03	$2.04^{+0.01}_{-0.01}$
H [%]	0.1	12^{+2}_{-12}	0.0	2^{+0}_{-2}
He [%]	0.1	7^{+1}_{-7}	0.3	3^{+1}_{-3}
N [%]	98.5	79^{+16}_{-5}	92.0	86^{+5}_{-2}
Fe [%]	1.2	$1.0^{+0.3}_{-0.3}$	7.7	8^{+1}_{-1}
$D\,(D_{\mathrm{spectrum}} + D_{X_{\max}})$	260.4 (95.0 + 165.5)		312.5 (56 + 256)	

Table 7.1: Model parameter results in the baseline scenario: best fit value $\hat{\theta}$, posterior mean $\langle\theta\rangle$ and 68% highest posterior density interval denoted by $\pm\sigma$. The corresponding posterior distributions are shown in figure 7.4.

Table 7.1 summarizes the model parameters by their best fit value $\hat{\theta}$, posterior mean $\langle\theta\rangle$ and 68% highest posterior density interval, cf. section 6.3 for definitions. The posterior distributions of the model parameters are displayed in figure 7.4.

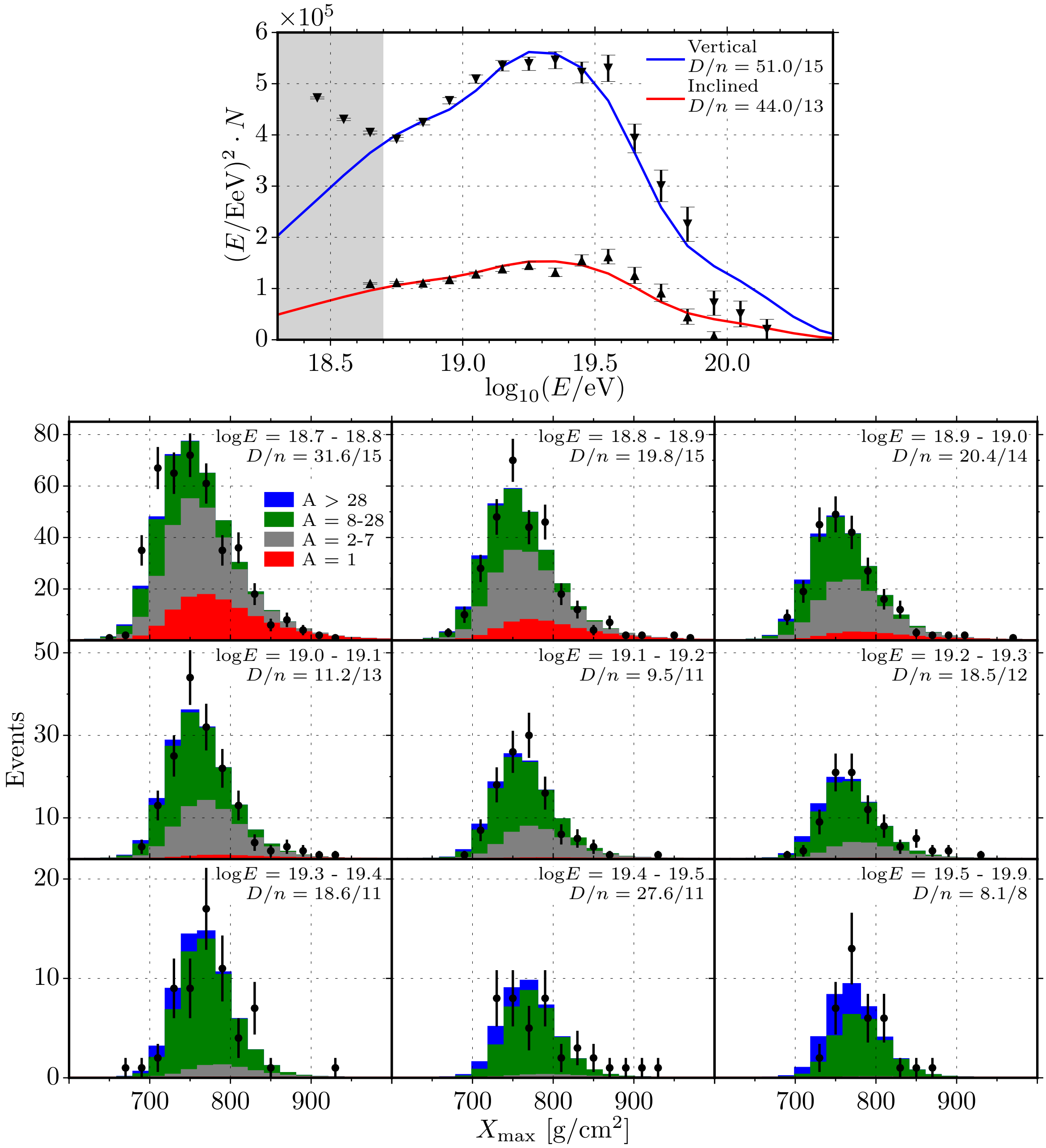


Figure 7.3: Best fit observables in the baseline scenario. Black markers show the measured number of events with their Poisson uncertainties. The deviance per number of data points D/n indicates the goodness-of-fit for each observable. Top: Raw energy spectrum from the vertical and inclined reconstructions, scaled with E^2 for better visibility. Bottom: $X_{\max}$ distributions in the indicated energy bins. The stacked histograms show the model prediction separated by groups of observed mass number. For a comparison of the measured and predicted first two moments of $X_{\max}$ see the left of figure 7.5.

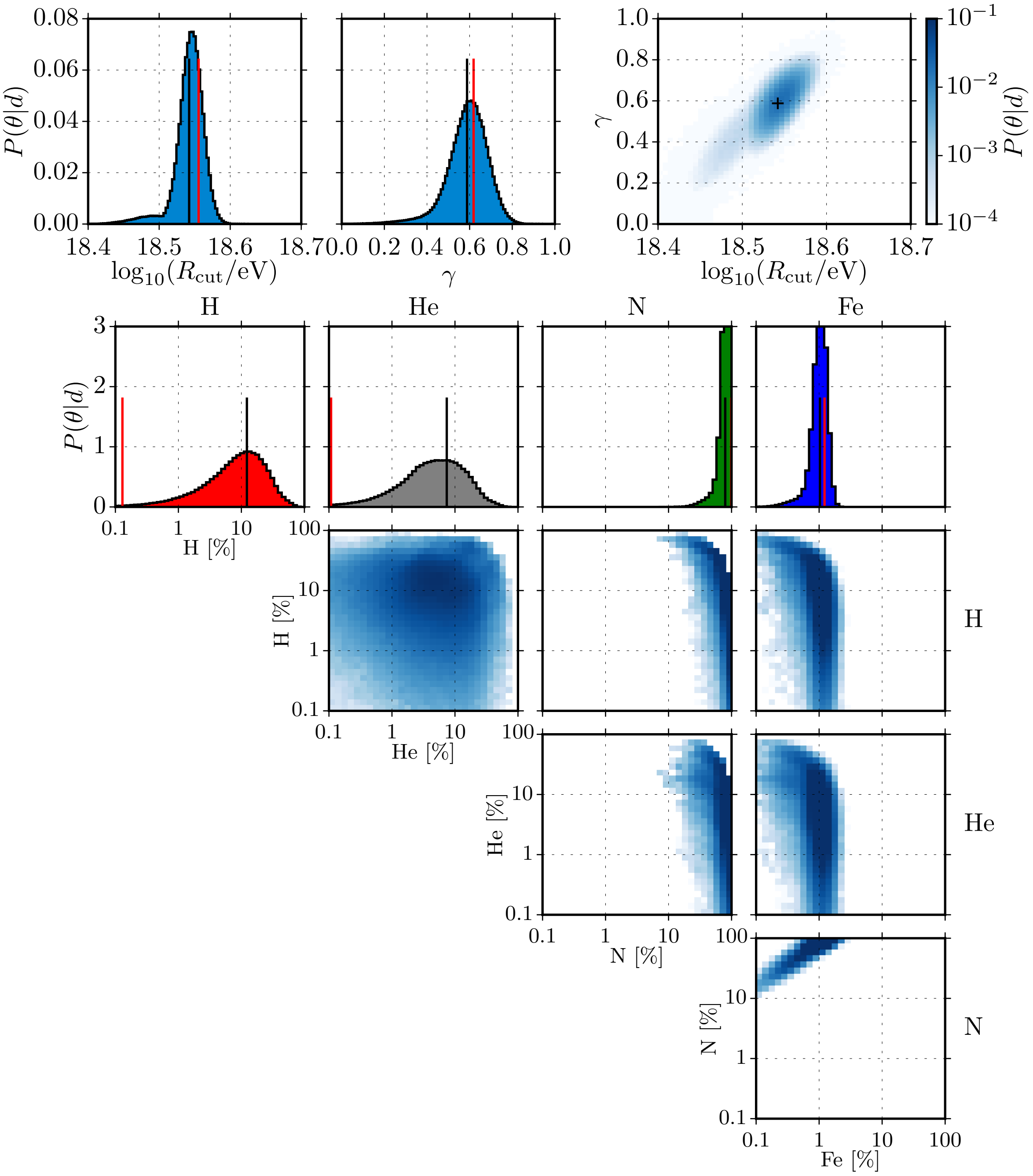


Figure 7.4: Posterior distribution of the model parameters in the baseline scenario: Cutoff rigidity and spectral index (top) and elemental fractions (bottom). One-dimensional histograms show the marginalized distributions together with the posterior mean $\langle\theta\rangle$(black line) and the maximum likelihood value (red line). Correlations among the parameters are visible from the pair-wise joint distributions with the probability indicated on a logarithmic color scale. The parameters are summarized in table 7.1.

The global minimum of the fitted scenario is characterized by a low cutoff rigidity of $\log_{10}(R_{\mathrm{cut}}/\mathrm{eV}) = 18.54^{+0.02}_{-0.01}$ in order to reproduce the early transition towards heavier elements as implied by the X_{max} data. At this cutoff rigidity, a hard power-law index $\gamma = 0.6^{+0.1}_{-0.1}$ is required for fitting the energy spectrum and composition. The power-law index of sources is surprisingly hard, conflicting with both the expectations for Fermi acceleration ($\gamma > 2$), as also noted by previous astrophysical analyses [81, 84], as well as for one-shot acceleration by fast-rotating neutron stars (see section 3.1, $\gamma \approx 1$).

The source composition is dominated by nitrogen, in order to result in a large contribution by cosmic rays of intermediate mass as shown in the best fit X_{max} distributions in figure 7.3. As explained in the previous section the light composition around the ankle then results from secondary particles from photodisintegration mainly of the nitrogen nuclei. Correspondingly, the source fractions of proton and helium are small. With the source cutoff of the nitrogen component occurring at $7 \times R_{\mathrm{cut}} \approx 10^{19.4}$ eV a small fraction of iron is required for reproducing the cosmic-ray flux at higher energies.

In the posterior distributions in figure 7.4 we can see that for hydrogen and helium fractions the maximum likelihood value is significantly lower than the posterior mean. This shows that while hydrogen or helium are not required, significant admixtures of these elements are still allowed in the fit.

Correlations among the model parameters can be assessed from the corresponding joint distributions in figure 7.4. The cutoff rigidity and spectral index have a small positive correlation resulting from the valley in the likelihood scan. The four elemental source fractions are already a priori correlated due to the side condition $\sum_i a(Z_i) = 1$, which is seen by the anti-correlation towards large parameter values. Apart from this feature the only significant correlation is found between the fractions of nitrogen and iron. The correlation is positive because a certain ratio of nitrogen to iron is required for reproducing the energy spectrum towards the highest energies. If both fractions are simultaneously reduced, a similar source spectrum can be recovered by accordingly increasing the flux normalization.

Due to the important role of photodisintegration in this scenario, we expect the model parameters to be sensitive to the photodisintegration cross sections used in the propagation simulation. Since at energies below $E/A \sim 3 - 4$ disintegration interactions are mostly induced by EBL photons, we also expect a sensitivity to the considered EBL model. This is further investigated in section 7.2, together with a variation in the air shower modeling, which determines the composition that is implied by the X_{max} measurements. Regarding the source model, the role of the cutoff shape, of the representative source elements, and of the source distribution is investigated in the section 7.3.

In addition to the aforementioned global minimum, we also list in table 7.1 the results for the second minimum at high cutoff rigidities $\log_{10}(R_{\mathrm{cut}}/\mathrm{eV}) = 19.87 \pm 0.02$. Here, the spectral index $\gamma = 2.04 \pm 0.01$ is compatible with Fermi acceleration, but the heavy source composition again does not correspond to the expectation for accelerating particles from the interstellar or intergalactic medium. Despite a theoretically appealing spectral index, the second minimum exhibits a poor goodness-of-fit fit with a deviance of $D_{\mathrm{min}} = 312.6$ compared to $D_{\mathrm{min}} = 260.4$ in the first minimum. The reason for this poor description is obvious in a

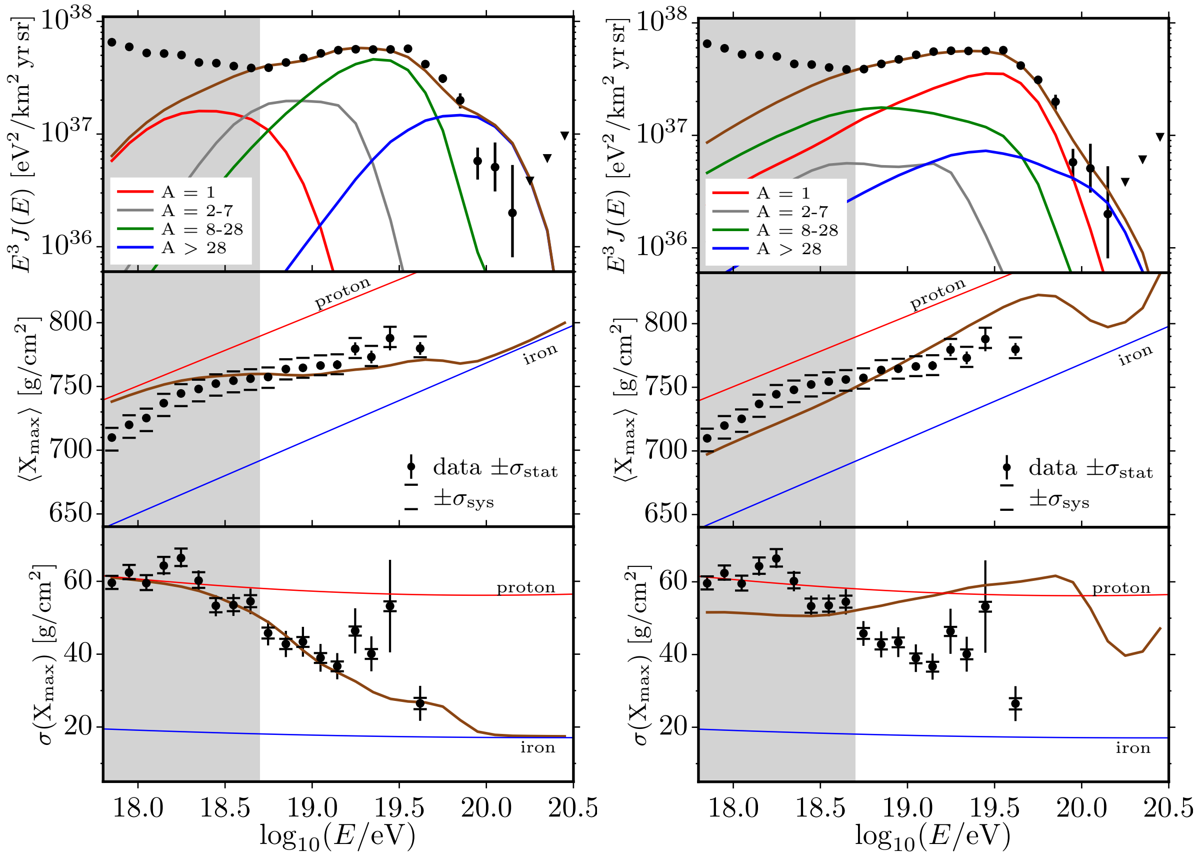


Figure 7.5: Comparison of the unfolded energy spectrum and moments of $X_{\max}$ for the global (left) and the second minimum (right). The top row shows the unfolded energy spectrum [30] together with the model prediction (brown line) and the contribution by mass groups. The middle and bottom rows show the unfolded mean and spread of the $X_{\max}$ distributions [210] with the model prediction and the expectation for the case of a pure proton (iron) composition.

comparison of the predicted energy spectrum and X_{max} moments (mean and spread of the distributions) for the first and second minimum, shown in figure 7.5. In the first minimum the low cutoff rigidity results in a quick transition towards heavier elements with small spread in the X_{max} distributions. With the high cutoff rigidity in the second minimum a large fraction of primary and secondary protons is present up to the highest observed energies. The large proton component prevents the average X_{max} value to follow the apparent trend of an increasingly heavy and pure composition. A characteristic feature of this second minimum is a dip in both X_{max} moments at around $10^{20.2}$ eV revealing the cutoff in the flux of secondary protons from nitrogen and iron. Beyond this energy the composition again becomes lighter and more spread due a emergence of alpha particles (helium) from photodisintegration. Extending the composition measurements to higher energies through a significantly larger exposure would allow to search for such a feature.

7.1.3 Experimental Uncertainties

Up to now the experimental systematic uncertainties have not been taken into account, with the exception of discrete shifts of the energy scale and the X_{max} scale by $\pm\sigma$ considered in [87, 227]. In the following we consider continuous variations of the energy scale, the relative exposure of the vertical and inclined flux measurements, and on the acceptance, resolution and scale of the X_{max} measurements. As described in section 6.2.3 this is done by representing each of these uncertainties with a nuisance parameter, which is then constrained through the measurement along with the model parameters. The number of parameters thus increases to $n_p = 6 + 5 = 11$.

The resulting comparison of the predicted observables with the measurements are shown in figure 7.6. The posterior distributions are shown in figure 7.7 and summarized in the first row of table 7.2. The relative number of events in the vertical and inclined energy spectrum measurements imply a small rescaling of the relative exposure of these measurements. The resulting posterior mean $\langle\theta_{\mathrm{expo}}\rangle = +1.6\,\sigma$ corresponds to a shift $\omega^{\mathrm{Inclined}}/\omega^{\mathrm{Vertical}} - 1 \approx +6.7\%$. This is in good agreement with the value obtained in the combined energy spectrum analysis [30, 209]. Due to the similar detector response in the vertical and inclined reconstructions, this nuisance parameter in first order does not depend on the predicted cosmic ray flux, but strongly affects the absolute likelihood value. Hence, for the likelihood scan and the nominal fit in the previous sections we considered a fixed variation by $\theta_{\mathrm{expo}} = +1.6\,\sigma$.

In the baseline scenario the energy scale is fitted to $\langle\theta_E\rangle \approx -2\,\sigma$. Thus, the true energies are lower by $\Delta E/E \approx -28\%$ with respect to the reconstructed energies. This energy shift is preferred by the model in order to better reproduce both the energy spectrum and the X_{max} distributions as shown in figure 7.6. In the energy spectrum the shift allows to better match the change in spectral slope above $E = 10^{19.55}$ eV with the cutoff of the dominating nitrogen component. The lower true energies also imply less photodisintegration during propagation which can better reproduce the narrow X_{max} distributions. A negative energy shift also has consequences for interpreting the X_{max} distribution. The average depth of shower maximum increases with energy and decreases with mass number. Thus, if the true energy scale is lower than reconstructed the measured composition becomes lighter.

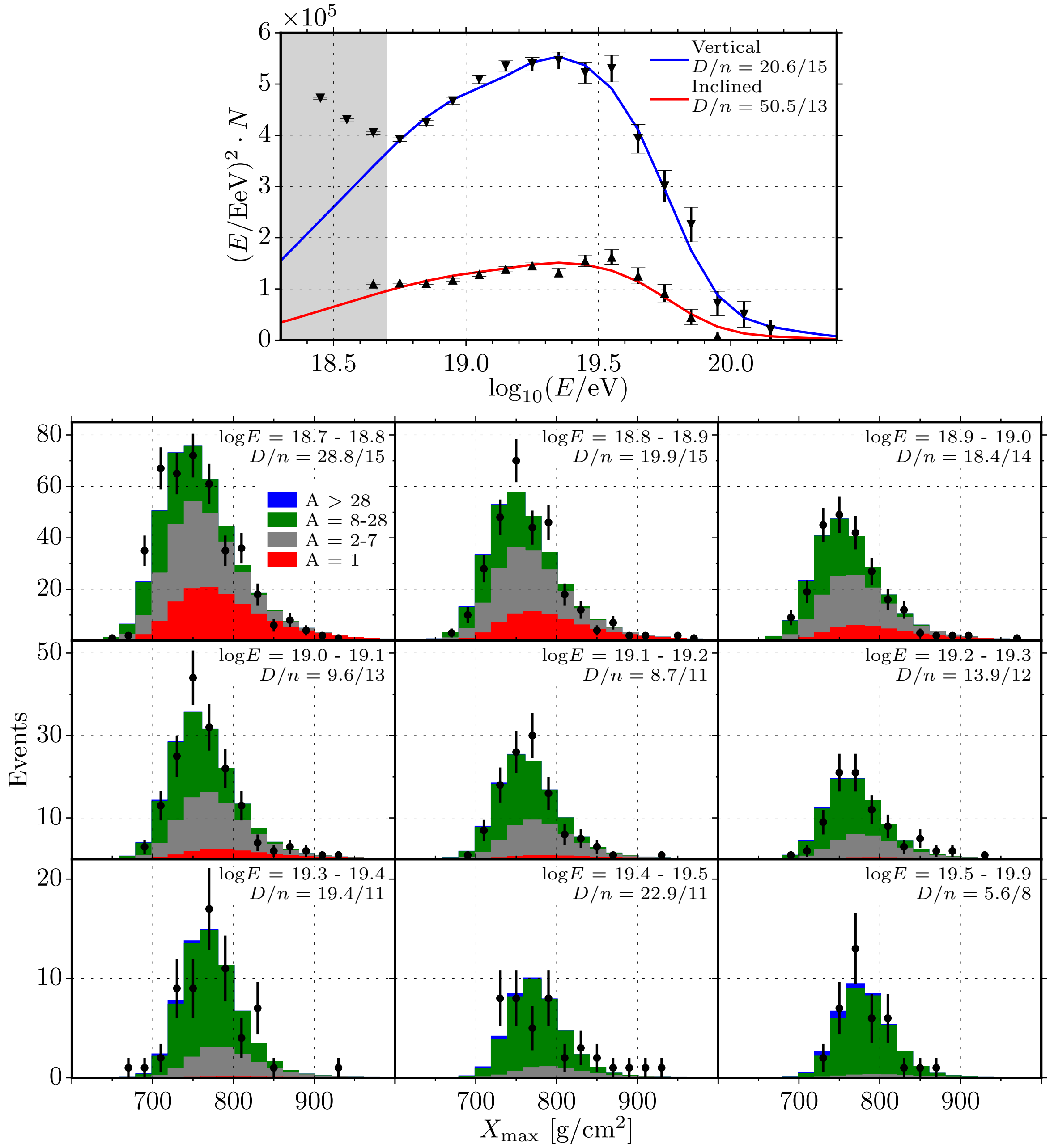


Figure 7.6: Best fit observables in the baseline scenario including the experimental systematic uncertainties. Black markers show the measured number of events with their Poisson uncertainties. The deviance per number of data points D/n indicates the goodness-of-fit for each observable. Top: Raw energy spectrum from the vertical and inclined reconstructions, scaled with E^2 for better visibility. Bottom: X_{max} distributions in the indicated energy bins. The stacked histograms show the model prediction separated by groups of observed mass number. Both observables are seen to be better reproduced compared to the fit without systematic uncertainties, shown in figure 7.3.

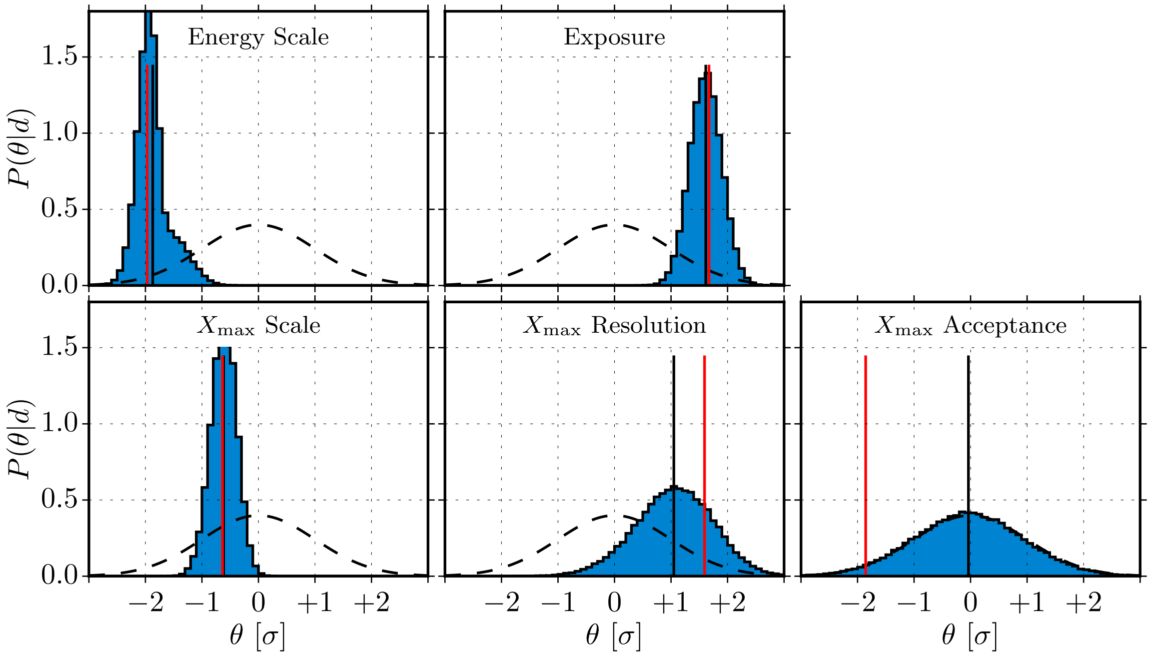


Figure 7.7: Posterior distribution of the nuisance parameters representing the experimental systematic uncertainties in units of σ. The Gaussian prior is shown as a dashed line. Vertical lines show the posterior mean (black) and the best fit value (red).

For the $X_{\max}$ related systematic uncertainties the following characteristics are found. The detector acceptance is uniform over a large range of possible $X_{\max}$ values due the employed fiducial field of view cut as described in section 5.3.2. In the considered energy range, the relative acceptance decreases only for high $X_{\max}$ values that correspond to deeply penetrating proton showers at the highest energies. Since the model predicts a negligible proton flux above 10^{19} eV, the impact of the acceptance and thus its uncertainty are negligible. As a result the posterior in figure 7.7 perfectly traces the Gaussian prior distribution.

The true $X_{\max}$ distributions are smeared out by the finite resolution of the fluorescence detector. Varying the parametrization of the detector resolution within its uncertainties, the fit shows a preference for a higher resolution corresponding to $\langle\theta_{X,\text{reso}}\rangle \approx +1\,\sigma$. This leads to slightly narrower predicted $X_{\max}$ distributions which improve the agreement with the measurement.

Of the $X_{\max}$ related uncertainties, the uncertainty on the scale has the largest impact, which can be inferred from the strongly localized posterior distribution in figure 7.7. A mean parameter value of $\langle\theta_{X,\text{scale}}\rangle \approx -0.6\,\sigma$ is found, which represents a shift of $\Delta X_{\max} \approx -5.4\,\text{g/cm}^2$ ($-4.2\,\text{g/cm}^2$) at a nominal energy of $10^{18.7}$ eV ($10^{19.7}$ eV). Thus, according to the fit the true $X_{\max}$ values are lower than reconstructed, and the true composition correspondingly heavier. This counteracts the opposite effect of the negative energy shift and prevents the composition of becoming too light compared to the model prediction. The shift of $X_{\max}$ scale may therefore be seen as consequence of the energy scale shift.

All posterior distributions are found to be well contained within the allowed parameter range of $\theta = \pm 3\,\sigma$. We tested all pairs of nuisance parameters for correlations. The largest correlation was found between energy and X_{max} scale parameters, with a Pearson's correlation coefficient of $\rho = 0.31$. This positive correlation is expected, because a stronger negative energy scale shift needs to be accompanied a stronger negative X_{max} shift.

In the following when probing the impact of simulation uncertainties we indicate the fit results both with and without the experimental systematic uncertainties.

7.2 Simulation Uncertainties

In this section the impact of the main simulation uncertainties on the fit results are investigated: the extragalactic background light model, the photodisintegration model, and the hadronic interaction model in the air-shower simulation.

Due to the relation $Z \approx A/2$ for nuclei $A > 1$, the cutoff in the source spectrum for the global minimum of the baseline scenario corresponds to a maximum energy per nucleon of $E/A \approx R_{\mathrm{cut}}/2 \approx 2 \times 10^{18}\,\mathrm{eV}$ or to a Lorentz factor of $\Gamma \approx 2 \times 10^9$. At these Lorentz factors a background photon of a few meV energy can be boosted to nuclear rest frame energies $\epsilon' > \epsilon'_{th} \sim 10\,\mathrm{MeV}$ above the threshold for photodisintegration. As can be seen in the top left and right of figure 7.8 this photon energy corresponds to the transition between the CMB and EBL as the main background: For higher E/A the photodisintegration on the CMB becomes possible, leading to large interaction rates, whereas for lower E/A photodisintegration occurs mostly on the far-infrared peak in the EBL spectrum. This behavior is similar for all nuclides $A \leq 56$, see the bottom right of figure 7.8, due to similar shapes of the photodisintegration cross sections. In fact, the emerging photodisintegration on the CMB for $R_{\mathrm{cut}} > 18.6\,\mathrm{eV}$ is seen in the combined likelihood scan in figure 7.1 by a strongly decreasing deviance. Hence, uncertainties in the photodisintegration and EBL model are expected to translate to large uncertainties in the astrophysical fit.

The third major source of uncertainty in the astrophysical interpretation is due to the hadronic interaction modeling in the air shower simulation. In order to test the impact of these uncertainties on the astrophysical fit, we consider the variations of the respective models. Regarding the EBL model and the photodisintegration model, separate propagation simulation were performed. To vary the air shower modeling it is sufficient to construct different sets of the X_{max} templates described in section 6.2.2 of the previous chapter. The fit results for these variations are summarized in table 7.2, both with and without experimental systematic uncertainties, and described in the following.

7.2.1 Photodisintegration Model

In the baseline scenario the photodisintegration model is based on the cross sections and branching ratios obtained from the TALYS software, see chapter 4.5 for a detailed description. This model was found to overall well describe the available experimental photodisintegration data [184]. Still, considerable uncertainties in modeling the photodisintegration remain, as shown in appendix A.2. The model uncertainties are especially large regarding the branching ratios for the emission of specific particles from the nucleus. Here, the α-channel is of particular

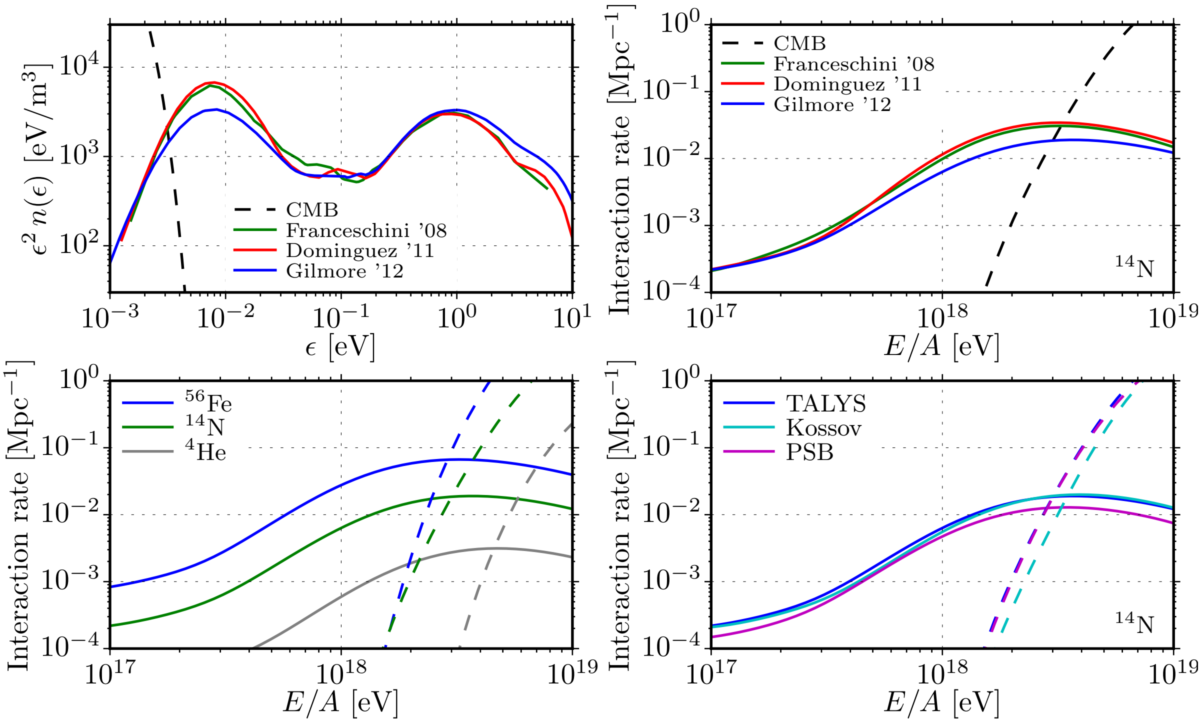


Figure 7.8: Uncertainties in the photodisintegration modeling. Top left: Spectral number densities $n(\epsilon)$ of the considered EBL models (Franceschini, Dominguez and Gilmore) and the CMB. Top right: Photodisintegration rates of ^{14}N as function of energy per nucleon E/A for the different EBL models. The transition from the EBL to the CMB as dominating background field occurs at a few EeV. Bottom left: photodisintegration rates for the three injected nuclides on the Gilmore *et al.* EBL (solid line) and the CMB (dashed line). Bottom right: Photodisintegration rates of ^{14}N for the different considered disintegration models.

importance because it enables an efficient disintegration of cosmic rays of intermediate mass, which are found to be dominating the considered scenario, cf. also [141, 227].

An alternative photodisintegration model commonly used in cosmic-ray propagation is the PSB model [107, 189], in which the emission of more than two nucleons is neglected altogether in the dominating giant dipole resonance. Therefore, despite similar total interaction rates, as seen in the bottom right of figure 7.8, the level of photodisintegration for intermediate nuclei such as nitrogen is effectively reduced.

As a third option we consider the cross section model by Kossov [187], which is used in CRPropa in combination with the branching ratios from TALYS. Here, the α channel is less important compared to the TALYS model because of the smaller total cross section at low photon energies, where the branching ratio for α emission is large, see appendix A.2.

Model	$\log_{10}(R_\mathrm{cut})$	γ	H [%]	He [%]	N [%]	Fe [%]	θ_E	$\theta_{X,\mathrm{scale}}$	$\theta_{X,\mathrm{reso}}$	Min. Deviance
Baseline	$18.56^{+0.02}_{-0.01}$	$0.6^{+0.1}_{-0.1}$	12^{+2}_{-12}	7^{+1}_{-7}	79^{+16}_{-5}	$1.0^{+0.3}_{-0.3}$				260.4 (95 + 166)
	$18.57^{+0.02}_{-0.02}$	$0.9^{+0.1}_{-0.1}$	11^{+2}_{-11}	8^{+2}_{-8}	80^{+14}_{-7}	$0.4^{+0.1}_{-0.3}$	$-1.9^{+0.2}_{-0.3}$	$-0.6^{+0.4}_{-0.5}$	$1.0^{+1.3}_{-1.4}$	218.4 (71 + 147)
Kossov	$18.52^{+0.02}_{-0.01}$	$0.5^{+0.1}_{-0.1}$	7^{+0}_{-7}	1^{+0}_{-1}	91^{+8}_{-1}	$0.8^{+0.2}_{-0.2}$				233.4 (70 + 163)
	$18.51^{+0.04}_{-0.02}$	$0.5^{+0.2}_{-0.1}$	18^{+4}_{-18}	15^{+2}_{-15}	67^{+26}_{-11}	$0.3^{+0.1}_{-0.2}$	$-1.1^{+0.3}_{-0.3}$	$-0.2^{+0.4}_{-0.4}$	$0.8^{+1.4}_{-1.5}$	215.7 (63 + 153)
PSB	$18.51^{+0.03}_{-0.01}$	$0.5^{+0.1}_{-0.1}$	13^{+1}_{-13}	6^{+0}_{-6}	81^{+16}_{-2}	$0.7^{+0.2}_{-0.2}$				236.9 (74 + 163)
	$18.52^{+0.04}_{-0.01}$	$0.6^{+0.2}_{-0.1}$	15^{+3}_{-15}	10^{+2}_{-10}	74^{+23}_{-5}	$0.3^{+0.1}_{-0.3}$	$-1.0^{+0.3}_{-0.3}$	$-0.3^{+0.4}_{-0.4}$	$0.8^{+1.4}_{-1.4}$	217.1 (65 + 152)
Franceschini	$18.29^{+0.04}_{-0.03}$	$-0.6^{+0.3}_{-0.2}$	40^{+14}_{-40}	28^{+8}_{-28}	32^{+8}_{-28}	$0.1^{+0}_{-0.1}$				245.5 (71 + 175)
	$18.34^{+0.03}_{-0.01}$	$0.1^{+0.2}_{-0.1}$	19^{+2}_{-19}	4^{+1}_{-4}	77^{+22}_{-4}	$0.1^{+0}_{-0.1}$	$-2.2^{+0.2}_{-0.3}$	$-0.8^{+0.4}_{-0.4}$	$0.8^{+1.3}_{-1.3}$	219.3 (61 + 158)
Dominguez	$18.25^{+0.03}_{-0.02}$	$-0.8^{+0.2}_{-0.2}$	35^{+11}_{-35}	27^{+8}_{-27}	38^{+12}_{-30}	$0.1^{+0}_{-0.1}$				242.6 (69 + 174)
	$18.20^{+0.07}_{-0.04}$	$-0.8^{+0.6}_{-0.3}$	31^{+10}_{-31}	41^{+14}_{-41}	28^{+8}_{-27}	0^{+0}_{-0}	$-2.0^{+0.2}_{-0.5}$	$-0.4^{+0.5}_{-0.5}$	$0.8^{+1.4}_{-1.4}$	205.1 (47 + 158)
QGSJetII-04	$18.14^{+0.01}_{-0.02}$	$-2.8^{+0.1}_{-0.2}$	95^{+2}_{-1}	5^{+1}_{-2}	0.1^{+0}_{-0}	0^{+0}_{-0}				355.6 (65 + 290)
	$18.44^{+0.15}_{-0.08}$	$0.1^{+0.8}_{-0.3}$	30^{+5}_{-30}	26^{+5}_{-26}	44^{+25}_{-43}	$0.2^{+0.1}_{-0.2}$	$-2.0^{+0.2}_{-0.5}$	$-0.4^{+0.5}_{-0.5}$	$0.8^{+1.4}_{-1.4}$	234.6 (54 + 181)
Sibyll 2.1	$18.15^{+0.01}_{-0.03}$	$-2.7^{+0.1}_{-0.3}$	95^{+2}_{-1}	5^{+1}_{-2}	$0.1^{+0}_{-0.1}$	0^{+0}_{-0}				270.8 (63 + 208)
	$18.55^{+0.03}_{-0.02}$	$0.8^{+0.1}_{-0.1}$	12^{+2}_{-12}	13^{+3}_{-13}	75^{+18}_{-9}	$0.3^{+0.1}_{-0.3}$	$-1.7^{+0.1}_{-0.5}$	$-2.1^{+0.4}_{-0.4}$	$-0.9^{+2.0}_{-1.5}$	219.6 (72 + 148)

Table 7.2: Fit results of the baseline scenario (Gilmore '12, TALYS model and EPOS-LHC) compared to variations of the EBL model (Franceschini '08 and Dominguez '11), the photodisintegration model (PSB and Kossov) and the hadronic interaction model in the air shower simulation (QGSJet II-04 and Sibyll 2.1). In each case, the second row gives the fit results including the experimental systematic uncertainties. Model and nuisance parameters are summarized by the posterior mean $\langle\theta\rangle$ and the 68% HPD interval. The nuisance parameters of the relative exposure and the X_max acceptance are similar in all cases and are omitted. The minimum deviance, D_min ($D^\mathrm{spectrum}_\mathrm{min} + D^{X_\mathrm{max}}_\mathrm{min}$), indicates the goodness-of-fit.

The fit results using these alternative photodisintegration models are listed in table 7.2. The measurements constrain the model parameters to qualitatively similar values, i.e. the scenario is still described by a low cutoff rigidity $R_{\mathrm{cut}} \approx 10^{18.5}\,\mathrm{eV}$, a hard spectral index $\gamma \approx 0.5$ and a dominating nitrogen component. The goodness-of-fit significantly improves to $D_{\mathrm{min}} \approx 233$ and $D_{\mathrm{min}} \approx 237$ for the Kossov and PSB models, respectively. This is seen to be mostly due the energy spectrum. The improvement is not surprising, because, as mentioned in the previous section, the goodness-of-fit profits from a reduced level of photodisintegration. Assuming the source model is true, the minimum deviance of the fits with the nominal detector response thus indicates a preference for the Kossov and PSB models compared to the TALYS model.

Since the measurements are already better described with the nominal detector response, the impact of the detector systematics decreases compared to the baseline fit, see table 7.2. We conclude that the modeling of photodisintegration during cosmic-ray propagation poses a significant source of uncertainty, but the general picture of the fit results in the presented scenarios does not change.

7.2.2 Extragalactic Background Light Model

Recent predictions for the EBL spectrum include the models by Gilmore *et al.* (our baseline model), Dominguez *et al.* and Franceschini *et al.*. In this section we evaluate their influence on the fit results. As shown in the top left of figure 7.8 the differences between the models are particularly pronounced in the far-infrared peak around $\epsilon \sim 10^{-2}\,\mathrm{eV}$. For kinematical reasons photons at these energies contribute most to the photodisintegration rate at $E/A \sim 10^{18}\,\mathrm{eV}$, which leads to considerable differences in the predicted interaction rates (top right of figure 7.8). In particular, both the Franceschini and Dominguez models predict higher interaction rates than the Gilmore model. Here we note that, while there are also differences in the redshift evolution of the models, see chapter 3.3, these are less relevant than the differences in the spectrum at $z_0 = 0$, because in the present scenario the bulk of cosmic rays originates from low redshifts.

The fit results listed in table 7.2 show that with the alternative EBL models the posterior distribution significantly moves in the $(R_{\mathrm{cut}}, \gamma)$ parameter plane. Compared to the baseline fit the best fit is found at lower cutoff rigidities $R_{\mathrm{cut}} \sim 10^{18.25-18.29}\,\mathrm{eV}$ and an extremely hard spectral index $\gamma \sim -(0.8 - 0.6)$, implying an increasing emissivity with increasing energy. This change is due to the increased level of photodisintegration that, using the alternative EBL models, is predicted at the parameter values of the baseline fit. Thus, the most probable region moves to lower cutoff rigidities in order to suppress photodisintegration through lower cosmic-ray energies. The minimum deviance decreases, indicating a preference for the alternative EBL models under the considered source hypothesis.

7.2.3 UHECR-Air Interaction Model

The cosmic-ray air showers represent the second main simulation aspect in modeling cosmic rays from the source to the observation. Air shower simulations are needed to translate the cosmic-ray composition arriving at Earth into the observable X_{max} distributions, as described in section 6.9. Here, the main uncertainties reside in the hadronic interactions in the shower, for which we consider several models: EPOS-LHC (our baseline model), as well as QGSJetII-04 and Sibyll 2.1. The differences between these models are large in terms of interpreting the measured $X_{\max}$ distributions with an empirical fit using four representative mass numbers $A = 1, 4, 14, 56$, cf. [34] and appendix A.3. Specifically, QGSJetII-04 and Sibyll 2.1 indicate a dominating component of the helium group $A = 4$ above the ankle, whereas EPOS-LHC requires a large fraction of the nitrogen group $A = 14$, see figure A.8.

Using either QGSJet or Sibyll in the fit of the astrophysical scenario, the resulting model parameters in table 7.2 are seen to shift to lower cutoff rigidities $\langle R_{\mathrm{cut}}\rangle \approx 10^{18.14}$ eV and harder spectral indices $\langle\gamma\rangle \approx -2.8$. Here, the posterior distribution of the spectral index is located at the boundary of the considered parameter range, hence the minimum is likely located at even more extreme values $\gamma < -3$. Using QGSJet the goodness-of-fit strongly decreases compared to the baseline model, even accounting for the decreased expectation from the empirical fit shown in appendix A.3. This is a bit surprising, because the more helium-like interpretation of the measured $X_{\max}$ distributions using QGSJet or Sibyll could in principle be accounted for by a late cutoff $R_{\mathrm{cut}} > 10^{19}$ eV of a large helium component at the source. However, the corresponding large energies per nucleon would imply the presence of efficient photodisintegration off the CMB, see figure 7.8, leading to an admixture of secondary protons in the observed flux above the ankle. Since the measured $X_{\max}$ distributions imply an overall increasingly pure composition above the ankle, this would be at variance with the measurements. Therefore, we note that EPOS-LHC is strongly preferred compared to QGSJetII-04 and slightly preferred compared to Sibyll 2.1 within the present astrophysical scenario at the nominal detector response.

Allowing for variations of the detector response, a systematic shift of the energy scale by $\langle\theta_E\rangle \sim -2\,\sigma$ again leads to a much improved goodness-of-fit. The model parameters then move towards those found in the baseline fit with a cutoff rigidity around $10^{18.5}$ eV and a spectral index somewhat harder than one.

7.3 Scenario Variations

In this section we test several modifications of the astrophysical source model, namely the cutoff shape, the source distribution and the source elements. We thereby test the validity of the derived conclusion with respect to deviations of the simplified source model. For each tested scenario variation H_1 we give the approximative Bayes factor $\Delta S \approx 2\ln(B_{10})$ (see equation 6.35) with respect to the baseline scenario H_0. Here, positive values indicate evidence against the baseline, negative values against the scenario variation. A categorization of the typical interpretation is given in table 6.3. Considering the simple nature of the source model, the large simulation uncertainties and the approximative calculation of the Bayes factors, we will not draw firm conclusions from these values, but only use to them to quantify the preference of the astrophysical fit.

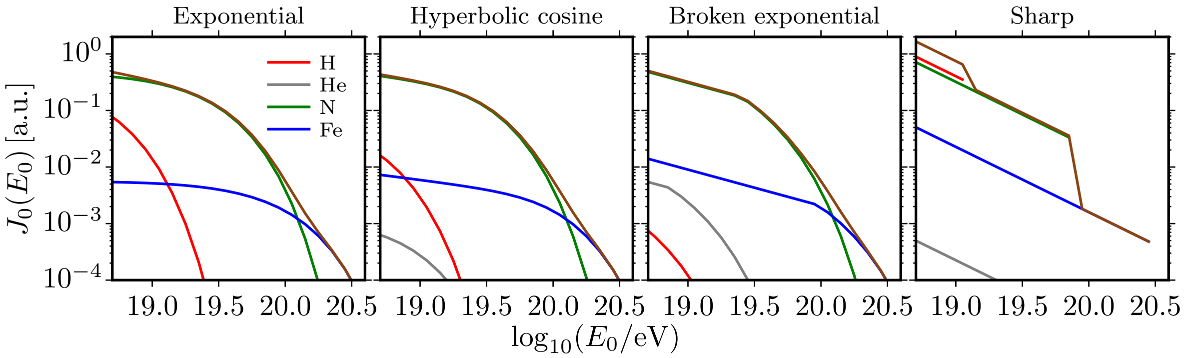


Figure 7.9: Variation of the cutoff shape in the source spectrum (best fit parameters). The brown line shows the total flux and colored lines the individual contributions by injected element. The broken exponential cutoff corresponds to the baseline scenario.

7.3.1 Source Cutoff shape

The cutoff shape describes the end of the source spectrum near the maximum acceleration capability of the sources. In the baseline scenario we assumed a broken exponential cutoff that leaves the source spectrum below the cutoff rigidity unmodified and models an exponential suppression beyond. As variations we consider three additional cutoff functions given in equation 6.2, which are chosen to cover a range of different abruptnesses to the end of the source spectrum. Here, the sharp cutoff corresponds to the limiting case in terms of abruptness, whereas in the other direction the cutoff could, in principle, proceed slower than modeled by the exponential function. All of the tested cutoffs are functions of rigidity, implying that the acceleration is limited by the capability of the source to magnetically confine the cosmic rays during acceleration. Testing different functional dependencies that would arise if the maximum energy is limited by synchrotron losses or interactions with radiation in the accelerator region, is beyond the scope of this work.

Cutoff shape	$\log_{10}(R_\mathrm{cut})$	γ	H [%]	He [%]	N [%]	Fe [%]	Min. Deviance	ΔS
Exponential	$18.46^{+0.02}_{-0.02}$	$-0.1^{+0.1}_{-0.1}$	21^{+6}_{-21}	17^{+5}_{-17}	63^{+17}_{-16}	$0.3^{+0.1}_{-0.2}$	247.1 (82 + 165)	+14
Hyperb. cos.	$18.49^{+0.03}_{-0.02}$	$0.2^{+0.1}_{-0.1}$	22^{+6}_{-22}	14^{+3}_{-14}	64^{+22}_{-15}	$0.5^{+0.2}_{-0.2}$	247.0 (82 + 165)	+15
Broken exp.	$18.54^{+0.02}_{-0.01}$	$0.6^{+0.1}_{-0.1}$	12^{+2}_{-12}	7^{+1}_{-7}	79^{+16}_{-5}	$1.0^{+0.3}_{-0.3}$	260.4 (95 + 166)	0
Sharp	$19.08^{+0.02}_{-0.02}$	$1.2^{+0.1}_{-0.1}$	41^{+4}_{-4}	1^{+0}_{-1}	56^{+4}_{-4}	$1.9^{+0.3}_{-0.3}$	305.7 (142 + 164)	-46

Table 7.3: Variation of the cutoff shape in the source spectrum, cf. figure 6.1a. For each parameter the posterior mean and 68% HPD interval are listed. D_min indicates the goodness-of-fit and ΔS the evidence against the baseline scenario (broken exponential cutoff). The table is ordered by increasing abruptness of the cutoff: smoother cutoffs require lower cutoff rigidities and harder spectral indices, and give a better fit to the energy spectrum.

The results for the considered variations, leaving all other settings as in the baseline scenario, are shown in table 7.3. Here, the cutoff rigidity and power-law index are seen to be correlated with the cutoff shape. As a general feature, smoother cutoff shapes require a lower cutoff rigidity and a harder spectral index compared to more abrupt cutoffs. This is because a stronger softening of the injected spectrum due to the cutoff shape is compensated by a harder spectral index, and vice versa, as shown in figure 7.9. Additionally, a smoother cutoff signifies a slower transition to heavier mass numbers. Since the observed X_{max} distributions imply a swift transition with increasing energy to higher cosmic-ray masses, the modeled transition is sped up by lowering the cutoff rigidity, and simultaneously hardening the spectral index. In all cases the nitrogen fraction dominates, which hides the transition (Peters cycle) from hydrogen to helium and helium to nitrogen in the source spectrum.

In the case of the sharp cutoff the cutoff rigidity is limited downwards by the requirement that the maximum acceleration energy $E_{\mathrm{max}} = 26 \times R_{\mathrm{cut}}$ has to be equal or larger than the highest observed energies. The maximum energy $\langle E_{\mathrm{max}} \rangle = 26 \times \langle R_{\mathrm{cut}} \rangle \approx 10^{20.5}$ eV is close to this minimum requirement. A sharp cutoff means less cosmic rays at the highest energies where CMB photons can disintegrate nuclei. Thus, less secondary protons are produced, which is compensated by a larger fraction of primary protons in order to reproduce the light composition around the ankle.

Qualitatively, a variation of the cutoff shape does not change the general picture of a low cutoff rigidity, hard spectral index and intermediate-mass source composition. However, as demonstrated the spectral index and to some extent the cutoff rigidity depend on the cutoff shape, which has to be kept in mind when comparing fit results from analyses using different cutoff models.

Inspecting the X_{max} deviance, the predicted composition is comparable in all cases. In contrast, the energy spectrum deviances in the exponential and hyperbolic cosine cutoff cases are better compared to the broken exponential cutoff in the baseline scenario, and much better than for the sharp cutoff case. Note however, that with the sharp cutoff the limited number of injected elements lead to a few discrete jumps in the source spectrum, which is likely the reason for the poorer fit to the observed energy spectrum. The decreased goodness-of-fit should therefore not be taken as argument against the sharp cutoff case, without testing scenarios with larger numbers of injected elements.

7.3.2 Source Evolution

In the baseline scenario the sources are assumed to be homogeneously distributed, i.e. sources with a constant comoving density injecting a constant flux of cosmic rays. Since cosmic rays can travel over cosmological distances the formation history of the sources becomes relevant. This formation history depends on the source type and can be expressed as function of redshift z, e.g. in form of the comoving source density $\rho_{\mathrm{com}}(z)$. A frequently used assumption for small enough redshifts ($z < 1$) is to consider a power-law $\rho_{\mathrm{com}}(z) = \rho(0)(1+z)^m$ with an evolution parameter m. The case $m = 0$ then corresponds to homogeneous distribution in the baseline fit. Typically, the evolution parameter is positive, indicating a higher comoving source density at earlier cosmological times, e.g. $m = 3.4$ for the star formation rate, $m = 4.8$ for gamma ray bursts and $m = 5$ for active galactic nuclei [228]. Conversely, nearby sources are more abundant in case of a negative evolution parameter.

In the following we test the impact of different evolution parameters on the fit results. In principle the evolution can be fitted as well within the present framework. However, this requires an additional distance binning in the simulated event matrix $N(Z_0, E_0, A, E)$ (see section 6.1.3), which increases the computation time by a factor equal to the number of distance bins. Instead, we construct a separate event matrix for each tested evolution parameter and perform the analysis as before. The results are listed in table 7.4.

Evolution	$\log_{10}(R_{\mathrm{cut}})$	γ	H [%]	He [%]	N [%]	Fe [%]	Min. Deviance	ΔS
$m = +5$	$18.08^{+0.01}_{-0.01}$	$-2.9^{+0.1}_{-0.1}$	25^{+7}_{-25}	24^{+7}_{-24}	51^{+22}_{-23}	$0.0^{+0.0}_{-0.0}$	256.5 (74 + 182)	+6
$m = +4$	$18.16^{+0.02}_{-0.02}$	$-2.0^{+0.3}_{-0.2}$	27^{+7}_{-27}	33^{+10}_{-33}	40^{+16}_{-29}	$0.0^{+0.0}_{-0.0}$	250.2 (68 + 182)	+12
$m = +3$	$18.23^{+0.03}_{-0.03}$	$-1.3^{+0.3}_{-0.3}$	35^{+11}_{-35}	49^{+29}_{-21}	16^{+2}_{-14}	$0.0^{+0.0}_{-0.0}$	252.2 (72 + 180)	+10
$m = +2$	$18.29^{+0.03}_{-0.03}$	$-0.8^{+0.4}_{-0.3}$	50^{+41}_{-20}	38^{+15}_{-32}	13^{+3}_{-12}	$0.0^{+0.0}_{-0.0}$	254.8 (76 + 178)	+8
$m = +1$	$18.40^{+0.02}_{-0.02}$	$-0.1^{+0.4}_{-0.2}$	38^{+12}_{-38}	29^{+10}_{-26}	33^{+12}_{-30}	$0.2^{+0.1}_{-0.2}$	258.6 (79 + 180)	+3
$m = 0$	$18.54^{+0.02}_{-0.01}$	$0.6^{+0.1}_{-0.1}$	12^{+2}_{-12}	7^{+1}_{-7}	79^{+16}_{-5}	$1.0^{+0.3}_{-0.3}$	260.4 (95 + 166)	–
$m = -1$	$18.59^{+0.02}_{-0.02}$	$0.8^{+0.1}_{-0.1}$	10^{+2}_{-10}	8^{+2}_{-8}	81^{+14}_{-5}	$1.2^{+0.3}_{-0.4}$	250.7 (88 + 163)	+9
$m = -2$	$18.65^{+0.02}_{-0.02}$	$0.9^{+0.1}_{-0.1}$	9^{+2}_{-9}	5^{+1}_{-5}	84^{+11}_{-4}	$1.5^{+0.3}_{-0.4}$	243.2 (82 + 162)	+16
$m = -3$	$18.68^{+0.02}_{-0.02}$	$1.1^{+0.1}_{-0.1}$	11^{+3}_{-11}	5^{+1}_{-5}	82^{+13}_{-6}	$1.6^{+0.3}_{-0.4}$	245.1 (85 + 161)	+14

Table 7.4: Comparison of the fit results for different source evolutions $(1+z)^m$. For each parameter the posterior mean and 68% HPD interval are listed. $D_{\min}$ indicates the goodness-of-fit and ΔS the evidence against the baseline scenario ($m = 0$).

The most apparent feature is the strong anti-correlation between the source evolution parameter and the spectral index. For the typical positive evolution parameters, a lower cutoff rigidity and harder spectral index are required, whereas for a negative evolution the fitted cutoff rigidity increases and the source spectrum turns softer. This is understandable, considering that a positive evolution increases the average source distance and thus enhances the importance of photodisintegration during propagation. To compensate, lower cutoff rigidities are again preferred within the scenario in order to reduce the level of photodisintegration. The case of an evolution parameter $m = 5$, corresponding to active galactic nuclei or similarly to gamma ray bursts, requires extremely hard spectral indices $\gamma \approx -3$. Since this value is at the boundary of the allowed prior parameter range, the optimum value is likely found at even lower values. In contrast, a negative source evolution parameter indicating an overabundance of nearby sources can help to bring the power-law index into agreement with theoretical expectations. This theoretical bias for a negative source evolution was first reported in [86]. The present results are in qualitative agreement, however as mentioned before, we find somewhat harder spectral indices.

7.3.3 Source Elements

Up to now we considered ^{1}H, ^{4}He, ^{14}N and ^{56}Fe nuclei as representatives of the source composition, based on the argument that their mass numbers are approximately equidistant in the the $X_{\max}$ observable, and can thus be best differentiated if observed undisintegrated. Two questions arise regarding this choice: First, are four elements sufficient to adequately model the smooth Peters cycle resulting from a full elemental composition? Specifically, the rigidity dependent cutoff energies of the four elements leave gaps of more than half a decade in the source spectrum, $\log_{10}(E^{\mathrm{N}}_{\mathrm{cut}}/E^{\mathrm{He}}_{\mathrm{cut}}) \approx 0.54$ and $\log_{10}(E^{\mathrm{Fe}}_{\mathrm{cut}}/E^{\mathrm{N}}_{\mathrm{cut}}) \approx 0.57$, before the onset of the cosmic-ray flux of the next higher charge number. A higher number of injected elements decreases theses gaps, which could improve the fit to the energy spectrum and thereby enable different solutions in the model parameter space.

Second, since photodisintegration rates and the disintegration paths depend on the specific nuclide, how representative are the selected elements in terms of their properties during propagation? In order to assess these questions we tested multiple variations to the number and type of the source elements, keeping all other settings as in the baseline scenario. As before we consider for each element the most abundant isotope on Earth. Apart from the theoretical bias that the injected cosmic rays are unlikely to be exceptional in terms of their isotopic composition, less abundant stable or unstable isotopes would also increase the level of photodisintegration and nuclear decay, which in the previous section was seen to be disfavored. An overview of the result for a selection of the tested variations is given in table 7.5.

As a first step we tested the relevance of the four elements for the scenario by injecting only subsets of these elements. The iron fraction, although small, is required at a low cutoff rigidity in order to produce the observed cosmic-ray flux at the highest energies. Thus, removing iron as the highest source element only the minimum at $R_{\mathrm{cut}} \approx 10^{19.9}$ eV is viable, albeit with the usual poor fit to the $X_{\max}$ distributions. The nitrogen component is dominating in the baseline fit. Eliminating this source element, the observed intermediate composition has to be produced by a partial disintegration of the iron component at correspondingly higher energies. Therefore, again only the high R_{cut} solution is viable.

The source fractions of hydrogen and helium are subdominant at the global minimum. Removing either element does not significantly impact the parameter results, and the deviance is very close to the baseline fit. The Bayes factor indicates a slight preference towards these reduced models due to the lower model complexity. In fact, the global minimum at $R_{\mathrm{cut}} \approx 10^{18.6}$ eV and $\gamma \approx 0.6$ is seen to work with only nitrogen and iron as source elements. This is not the case for the parameter space at extremely low cutoff rigidities, which requires a Peters cycle of all four elements, see the left of figure 7.2. Therefore, we continue to consider hydrogen and helium for the following variation, in order to not miss the possibility of this solution.

Next, we considered variations of four source elements. Since the nitrogen component is dominating in the baseline fit, we tested whether its properties in terms of initial charge number and photodisintegration are indeed exceptional. We therefore replaced $^{14}_{7}$N with other elements of intermediate mass, namely $^{7}_{3}$Li, $^{9}_{4}$Be $^{9}_{5}$B, $^{12}_{6}$C, $^{16}_{8}$O, $^{19}_{9}$F, $^{20}_{10}$Ne, $^{24}_{12}$Mg, $^{28}_{14}$Si, $^{32}_{16}$S, $^{40}_{18}$Ar and $^{40}_{20}$Ca. The parameter results are comparable for intermediate elements in the range $Z_0 = 5 - 14$ with the exception of ^{12}C. If no intermediate source element is present in this range, and for

3 Elements	$\log_{10}(R_\mathrm{cut})$	γ	H [%]	He [%]	N [%]	Fe [%]		Min. Deviance	ΔS
1,2,7	$19.90^{+0.02}_{-0.02}$	$2.0^{+0.0}_{-0.0}$	2^{+0}_{-2}	2^{+0}_{-2}	97^{+3}_{-1}	–		367.1 (58 + 309)	-106
1,2,26	$19.89^{+0.02}_{-0.03}$	$2.2^{+0.0}_{-0.0}$	2^{+0}_{-2}	86^{+2}_{-1}	–	13^{+1}_{-1}		423.1 (53 + 371)	-163
1,7,26	$18.56^{+0.01}_{-0.01}$	$0.6^{+0.1}_{-0.1}$	14^{+5}_{-14}	–	85^{+14}_{-5}	$1.1^{+0.2}_{-0.2}$		260.4 (95 + 166)	+1
2,7,26	$18.56^{+0.02}_{-0.01}$	$0.6^{+0.1}_{-0.1}$	–	9^{+2}_{-9}	90^{+9}_{-2}	$1.2^{+0.2}_{-0.3}$		260.4 (95 + 166)	+1
7,26	$18.56^{+0.01}_{-0.01}$	$0.6^{+0.1}_{-0.1}$	–	–	99^{+0}_{-0}	$1.3^{+0.2}_{-0.3}$		260.4 (95 + 166)	+2
4 Elements	$\log_{10}(R_\mathrm{cut})$	γ	H [%]	He [%]	X [%]	Fe [%]		Min. Deviance	ΔS
1,2,5,26	$18.72^{+0.01}_{-0.02}$	$0.9^{+0.1}_{-0.1}$	6^{+1}_{-6}	4^{+1}_{-4}	87^{+8}_{-3}	$2.1^{+0.3}_{-0.5}$		293.8 (129 + 165)	-35
1,2,6,26	$19.91^{+0.03}_{-0.02}$	$2.1^{+0.0}_{-0.0}$	2^{+0}_{-2}	3^{+1}_{-3}	86^{+4}_{-2}	9^{+1}_{-1}		314.7 (54 + 261)	-56
1,2,7,26	$18.54^{+0.02}_{-0.01}$	$0.6^{+0.1}_{-0.1}$	12^{+2}_{-12}	7^{+1}_{-7}	79^{+16}_{-5}	$1.0^{+0.3}_{-0.3}$		260.4 (95 + 166)	–
1,2,8,26	$18.64^{+0.02}_{-0.02}$	$0.8^{+0.1}_{-0.1}$	6^{+1}_{-6}	7^{+1}_{-7}	85^{+10}_{-3}	$1.1^{+0.2}_{-0.3}$		225.9 (64 + 162)	+34
1,2,9,26	$18.66^{+0.04}_{-0.04}$	$0.9^{+0.1}_{-0.1}$	14^{+4}_{-14}	20^{+7}_{-19}	66^{+11}_{-16}	$0.2^{+0.1}_{-0.1}$		214.6 (49 + 166)	+45
1,2,10,26	$18.63^{+0.04}_{-0.02}$	$0.8^{+0.1}_{-0.1}$	19^{+5}_{-19}	23^{+7}_{-23}	60^{+13}_{-20}	$0.2^{+0.1}_{-0.0}$		213.5 (50 + 164)	+48
1,2,12,26	$18.56^{+0.05}_{-0.05}$	$0.9^{+0.1}_{-0.1}$	13^{+4}_{-13}	71^{+10}_{-10}	16^{+3}_{-5}	$0.0^{+0.0}_{-0.0}$		256.9 (62 + 195)	+7
5 Elements	$\log_{10}(R_\mathrm{cut})$	γ	H [%]	He [%]	Ne [%]	X [%]	Fe [%]	Min. Deviance	ΔS
1,2,10,5,26	$18.64^{+0.03}_{-0.02}$	$0.8^{+0.1}_{-0.1}$	11^{+2}_{-11}	17^{+5}_{-17}	58^{+13}_{-14}	14^{+5}_{-12}	$0.2^{+0.0}_{-0.2}$	212.7 (52 + 161)	+48
1,2,10,12,26	$18.63^{+0.04}_{-0.03}$	$0.9^{+0.1}_{-0.1}$	17^{+5}_{-17}	24^{+7}_{-24}	56^{+15}_{-20}	3^{+1}_{-3}	$0.2^{+0.0}_{-0.2}$	213.5 (50 + 164)	+47
1,2,10,14,26	$18.64^{+0.04}_{-0.02}$	$0.9^{+0.1}_{-0.1}$	10^{+1}_{-10}	27^{+8}_{-27}	62^{+18}_{-19}	1^{+0}_{-1}	$0.2^{+0.0}_{-0.2}$	213.5 (50 + 164)	+47
1,2,10,16,26	$18.63^{+0.04}_{-0.02}$	$0.9^{+0.1}_{-0.1}$	18^{+5}_{-18}	23^{+7}_{-23}	59^{+15}_{-19}	1^{+0}_{-1}	$0.1^{+0.0}_{-0.1}$	213.5 (50 + 164)	+47
1,2,10,18,26	$18.63^{+0.04}_{-0.02}$	$0.9^{+0.1}_{-0.1}$	17^{+4}_{-17}	23^{+6}_{-23}	60^{+15}_{-18}	$0.4^{+0.1}_{-0.4}$	$0.1^{+0.0}_{-0.1}$	213.5 (50 + 164)	+47

Table 7.5: Variations of the source elements, denoted by their charge number Z. For each parameter the posterior mean and 68% HPD interval are listed. the parameter X represents the varied source element. D_min indicates the goodness-of-fit and ΔS the evidence against the baseline scenario (H, He, N, Fe).

^{12}C, only the second minimum is found, in which the observed intermediate composition is produced through photodisintegration.

The reason for the differences of the ^{12}C case is not immediately evident. Secondary particles from disintegrations retain the Lorentz factor of the parent nucleus. For nuclei around the cutoff rigidity $R_\mathrm{cut} \sim 18.7\,\mathrm{eV}$ the secondary α-particles contribute to the cosmic ray flux in the fitted energy range above the ankle α-particles $E_\alpha = 4\frac{Z}{A}R_\mathrm{cut} \approx 2R_\mathrm{cut} > 10^{18.7}\,\mathrm{eV}$, whereas secondary nucleons fall short of the ankle. This argument is the same for all nuclei, since $Z/A \approx 2$. However, a particular feature of ^{12}C is the availability of a short disintegration path into three α particles, $\gamma + {}^{12}\mathrm{C} \longrightarrow {}^{8}\mathrm{Be} + \alpha$, with the subsequent nuclear decay into two α-particles. In the case of ^{12}C the large branching ratio for this α-channel in the TALYS model may prevent an acceptable fit to the energy spectrum.

The baseline case of nitrogen as intermediate source element requires a slightly lower cutoff rigidity and harder spectral index ($\gamma = 0.6$) compared to the other elements, e.g. oxygen, fluorine or neon ($\gamma = 0.8 - 0.9$). Also, in these latter cases the goodness-of-fit to the energy spectrum is significantly higher and the Bayes factors indicate a strong preference against the nitrogen case. This could be either again due to the photodisintegration features of ^{14}N, or due to the cutoff energies of higher elements better reproducing the observed cutoff in the energy spectrum. Since similar improvements are seen for different charge numbers $Z_0 = 8 - 10$ without a corresponding change of the cutoff rigidity, the different initial charge numbers are likely not the main reason. However, a similar reasoning as for ^{12}C can be applied. In the TALYS model among the main disintegration channels of ^{14}N are the emissions of a proton, neutron or deuteron. The deuteron channel directly ends on ^{12}C ($\gamma + {}^{14}\mathrm{N} \longrightarrow {}^{12}\mathrm{C} + d$), the proton and neutron channels can indirectly end there: ^{13}N decays into ^{13}C, which has approximately equal branching ratios for the neutron channel ending in ^{12}C, and for the α channel ending in ^{9}Be, which disintegrates into two α-particles and a neutron. Thus ^{14}N can efficiently disintegrate into two to three α-particles.

Inspecting the goodness-of-fits of the variations of four elements, the minimum is found with neon as intermediate element. The minimum deviance of $D_{\mathrm{min}}/ndf = 213.5/137$ per degrees of freedom, with a contribution of $D_{\mathrm{spectrum}} \approx 50$ and $D_{X_{\mathrm{max}}} \approx 164$. Thus, the fit to the energy spectrum is much better than in the baseline scenario, and is at the level of the expectation from the empirical fit performed in appendix A.3.

Considering five or more source elements does not lead to further significant improvements. Here, additional elements between helium and neon only slightly improve the deviance, and elements above neon are found to be negligible for the fit and the parameter results.

We conclude that the selection of possible source elements does play a role, which is related to their somewhat uncertain properties regarding photodisintegration. However, a stable solution can be found with few elements, which can therefore be considered as representatives of a full elemental composition. In particular, an acceptable fit of the scenario to the measurements requires a dominant fraction of intermediate elements and small fraction of heavy elements.

7.4 Scenario Implications

In this section we investigate the implications of the scenario fits with respect to cosmic-ray flux below the ankle, the expected magnetic deflections and the cosmogenic flux of photons and neutrinos. In particular we assess whether these aspects can provide further constraints to the presented scenarios.

7.4.1 Cosmic-Ray Flux below the Ankle

As a consequence of the spectral index $\gamma \ll 2.7$ that is found with the presented scenarios, the predicted flux below the ankle quickly falls short of the observed flux. This is shown on the left in figure 7.10, where the missing sub-ankle flux is inferred from the difference to the unfolded measurement.

As described in sections 5.3.1 and 6.2.1, the finite energy resolution of the detector causes a migration of cosmic-ray events towards higher energies. Thus, since the model does not

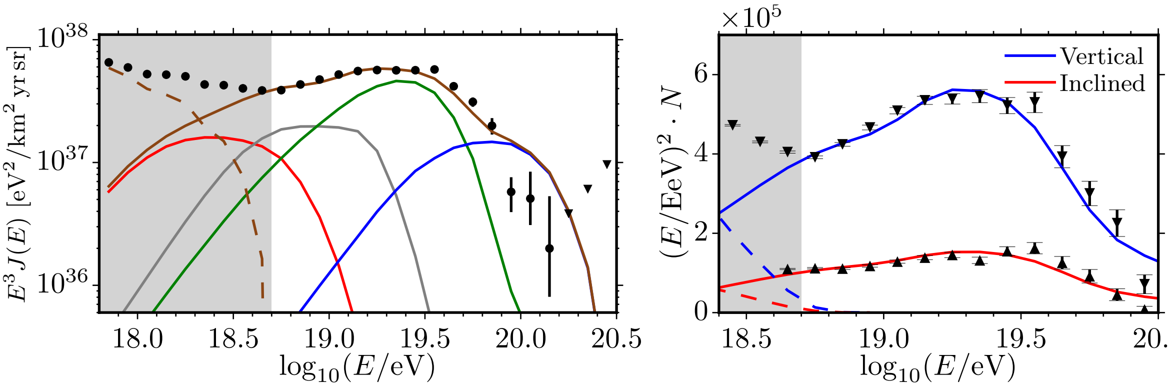


Figure 7.10: Estimate of the cosmic-ray flux below the ankle for the baseline scenario (best fit). Black markers and error bars show the measurements with their statistical uncertainty. The gray area indicates the energy range not considered in the fit. Left: Unfolded energy spectrum (brown solid line) and the inferred sub-ankle flux (brown dashed line). Right: Predicted raw energy spectra for the vertical and inclined reconstructions (solid lines). The finite detector resolution causes a migration of the sub-ankle flux to the fitted energy range (dashed lines).

account for the full cosmic-ray flux below the ankle, the additional contribution of the sub-ankle component to the fitted energy range above the ankle is currently neglected. Estimating the sub-ankle flux from the difference to the unfolded measurement, the additionally migrated flux in the energy bin $E = 10^{18.7-18.8}$ eV amounts to about 3% and 4% for the vertical and inclined spectrum measurements, respectively, and to the less than 0.5% in the second fitted energy bin $E = 10^{18.8-18.9}$, see the right of figure 7.10. While this contribution is significant within the measured number of events, we find that including the estimated additional migration in the fitting procedure only slightly modifies the fit results. Fully taking this effect into account requires a suitable modeling of the cosmic-ray flux below the ankle and is left for future analyses.

In the baseline scenario the sub-ankle component is predicted to start dominating below $E \approx 10^{18.4}$ eV. This turnover energy only weakly varies with the parameters of the extragalactic source model. The $X_{\max}$ measurements in this energy range, see the left of figure 7.5, indicate that the sub-ankle component has a dominantly light composition. As first discussed in [81] the sub-ankle component could either originate from a separate class of extragalactic sources, or from the galactic sources reaching very high energies. Both hypotheses are challenging. Introducing a separate class of extragalactic sources of different properties is not well motivated, and the emissivity and spectral index of the sources need to be fine-tuned in order to provide a smooth transition to the galactic cosmic rays at lower energies without producing a sharp spectral feature.

Attributing the sub-ankle component to the galactic cosmic-ray domain is difficult as well, due to the lack of source candidates that are suitable for reaching theses energies. This is especially challenging because of the measured light cosmic-ray composition at EeV energies, whereas the maximum acceleration energies are expected to be accompanied by a heavy composition. Therefore alternative explanations for the ankle are currently discussed [83, 85, 229], such as regions of strong magnetic fields surrounding the extragalactic sources. These magnetized regions could trap cosmic rays of lower rigidity long enough inside to facilitate their disintegration, or to delay their arrival beyond the present time, while letting cosmic rays of higher rigidity escape, thus acting as a *high-pass filter*. As single class of extragalactic sources could then account for the cosmic-ray flux above and below the ankle, with the ankle feature itself being a product of the source environment. This scenario could be further investigated with the present analysis, to better test the agreement with the Pierre Auger Observatory measurements.

7.4.2 Magnetic Deflections

Based on the measured level of isotropy in the cosmic-ray arrival directions, it is possible to set lower limits on combinations of the source density and the strength of magnetic deflections. These deflections depend on the properties of the intervening extragalactic and galactic magnetic fields, and on the cosmic-ray rigidity during propagation. Since the source composition is unknown, past analyses have focused on interpreting the measured level of isotropy in terms of a pure proton composition [230, 231]. The derived constraints on the minimum source density and strength of deflections are expected to be significantly relaxed by the higher charge numbers in mixed composition scenarios such as the presented ones. The predicted low maximum rigidities further increase the average extragalactic deflections. The lack of strong anisotropic features in the cosmic-ray arrival directions thus allows for more sparsely distributed source candidates and weaker magnetic fields compared to scenarios with a lighter composition or higher maximum rigidity.

For an estimate of the impact of the predicted mixed composition in the baseline scenario on magnetic deflections, we show in figure 7.11 the average charge number (top left figure) and rigidity (top right figure) of injected and observed cosmic rays. Since the deflections in a given magnetic field only depend on the rigidity $R = E/Z$, the mean charge number or rigidity as function of energy indicate how much stronger cosmic rays are on average deflected than protons at the same energy. For instance, at $E = 10^{19.5}$ eV the mean observed charge number is $\langle Z \rangle \approx 10$, meaning that at this energy cosmic rays behave in average as protons at a ten times lower energy.

Both the injected and observed charge numbers are seen to increase towards higher energies. In the observed flux the rise towards $\langle Z \rangle \sim 26$ from the emerging iron fraction occurs earlier than in the injected flux. This is due to non-destructive energy losses that shift the energy dependent composition to lower energies. Additionally, at equal energies lighter nuclei have higher Lorentz factors and thus higher interaction rates, which causes the iron fraction to emerge earlier from the dominant nitrogen flux. Around the ankle the situation is inverse with the pile-up of light secondary particles from disintegration decreasing the observed charge numbers.

In disintegration interactions the energy per nucleon of all disintegration products is similar to that of the mother nucleus, $E/A \approx E'/A'$. Therefore, since $Z/A \approx 2$ for $A \gg 1$, the rigidity is conserved during disintegration. The exception are secondary protons ($Z/A = 1$), which have half the rigidity of their mother nuclei and thus experience stronger deflections than

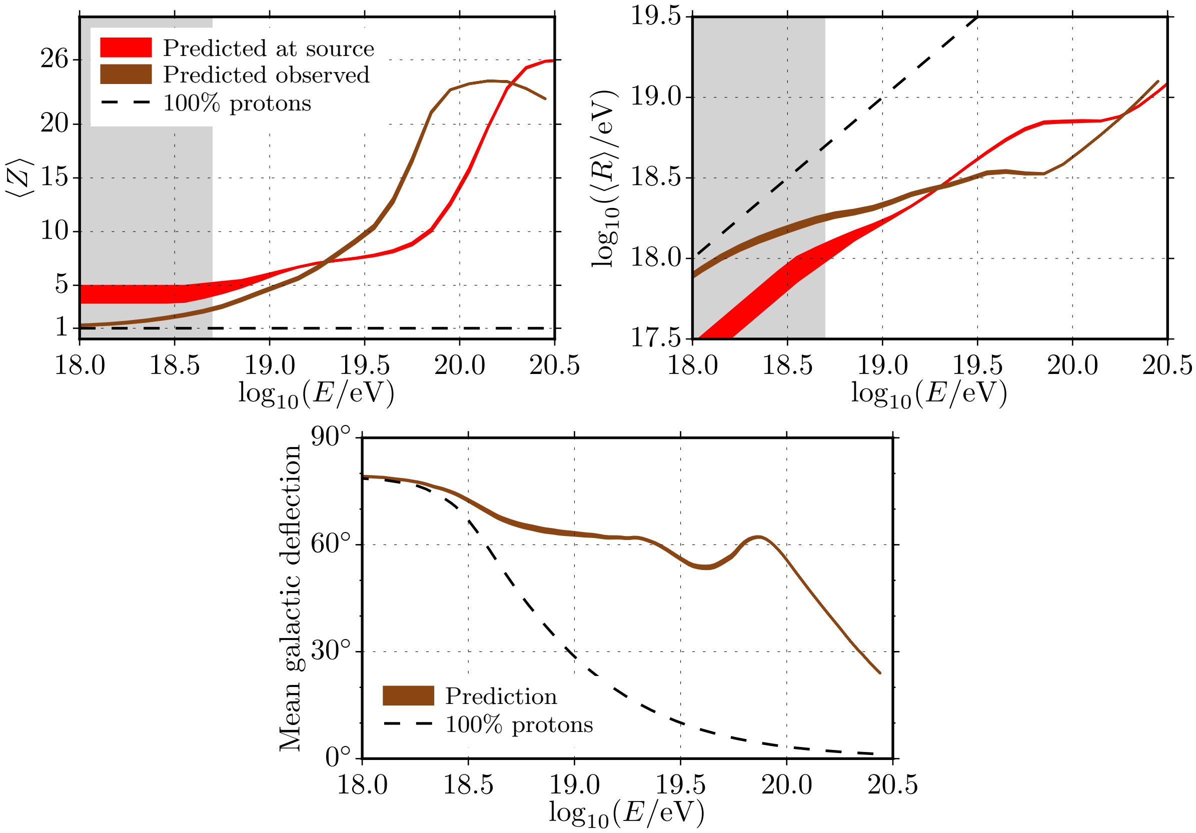


Figure 7.11: Average charge number (top left), rigidity (top right) and galactic angular deflection (bottom) that are predicted for the baseline scenario. The properties of the injected and observed cosmic-ray flux are shown in red and brown, respectively, with the bands indicating the 68% central posterior probability interval. For comparison the black dashed line shows the case of a pure proton composition. Bottom: Mean angular deflection in the JF12 galactic magnetic field model including the random field components, averaged over arrival directions at the galactic border.

other disintegration products. In total the average charge number during the extragalactic propagation is bound by the curves of the injected and observed cosmic rays, which allows to estimate the amplification of deflections in extragalactic magnetic fields compared to a pure proton composition.

A general feature is that the increasing average charge number towards higher energies significantly slows down the increase of magnetic rigidity. Both the injected and observed rigidities exceed $E/\langle Z\rangle > 10^{18.8}\,\mathrm{eV} \approx 6\,\mathrm{EeV}$ only for energies of more than 200 EeV. In particular the observed rigidity is almost flat in energy, implying that magnetic deflections remain at a rather constant level.

We can use the observed rigidity for a concrete prediction of the angular deflections in the galactic magnetic field. This is shown at the bottom of figure 7.11 for the case of the full JF12 model, see chapter 3.6.2 for details. Here, the average angular deflections are seen to remain at around 60° for energies $E < 100\,\mathrm{EeV}$, and larger than 30° higher energies. We note that this estimate is rather independent of the source model, since the predicted observed composition mainly depends on the interpretation of the measured $X_{\max}$ distributions. Uncertainties in this estimate thus mostly arise from the air shower modeling (here EPOS-LHC) when translating the $X_{\max}$ distributions into mass or charge distributions.

The predicted large deflections in the galactic field alone illustrate the challenges when searching for correlations between the arrival directions of charged cosmic rays and neutral messengers such as neutrinos, or the directions of astrophysical objects that are source candidates. Even at the highest energies, identifying such a correlation will likely require information on both the individual charges of the cosmic rays, and on the intervening magnetic fields.

7.4.3 Secondary Messengers

Neutrinos and photons are necessarily produced in interactions of cosmic-ray protons and nuclei during propagation with the extragalactic photon background, as outlined in chapter 4.8. Additional constraints for cosmic-ray scenarios are therefore provided by the (non-)observation of such cosmogenic neutrinos [80, 228, 232, 233] and photons [234–236].

In order to assess the flux of cosmogenic photons and neutrinos that are expected in the baseline scenario we simulated the production of neutrinos, photons and electrons/positrons during the propagation of the hadronic cosmic rays with CRPropa, cf. chapter 4.8. To that end we performed separate simulations in which the energies of injected hadronic cosmic-ray were extended down to $E_0 > 10^{15}\,\mathrm{eV}$. Since the energy loss length of neutrinos is effectively equal to the Hubble radius, even very distant sources can contribute to the observed neutrino flux. Hence, we simulated the injection of 3.5×10^6 cosmic-rays for each nuclear species up to a redshift of $z = 4$, compared to $z = 2$ in the main simulations. For the same reason the neutrino flux is much more sensitive to the source evolution than the hadronic flux as tested in section 7.3.2. Therefore, in addition to the baseline scenario of non-evolving sources, we consider the case of a strong evolution as $(1 + z)^m$ with $m = 6$, thereby emphasizing the cosmogenic neutrino flux.

For the cosmogenic photon flux it is sufficient to simulate a smaller number of hadronic cosmic rays, here 7×10^5 injected particles per nuclear species, due to the larger number of produced photons and electrons. Since the horizon of these particles is much smaller than that of hadronic cosmic rays, it is also sufficient to consider a smaller maximum distance. Here we set a maximum comoving distance of 1 Gpc, corresponding to a redshift of $z \approx 0.24$. We then used the transport code DINT within CRPropa to calculate the ensuing electromagnetic cascade down to energies of $10^7\,\mathrm{eV}$. Note that for the cosmogenic photons several possibly relevant production channels are not yet considered in CRPropa, cf. chapter 4.8, hence the simulated photon flux represents a lower limit.

The predicted photon and neutrino fluxes are shown in figure 7.12 for the global and second minimum in the baseline scenario. For the neutrinos the low maximum energies per nucleon $Z/AR_{\mathrm{cut}} \approx 1/210^{18.6}\,\mathrm{eV}$ in the global minimum results in a low predicted flux thats peaks at $E_\nu \sim 10^{17}\,\mathrm{eV}$. This flux is mainly composed of neutrinos from β-decays of unstable residual nuclei and neutrons in photodisintegration interactions. In the second minimum the cosmic

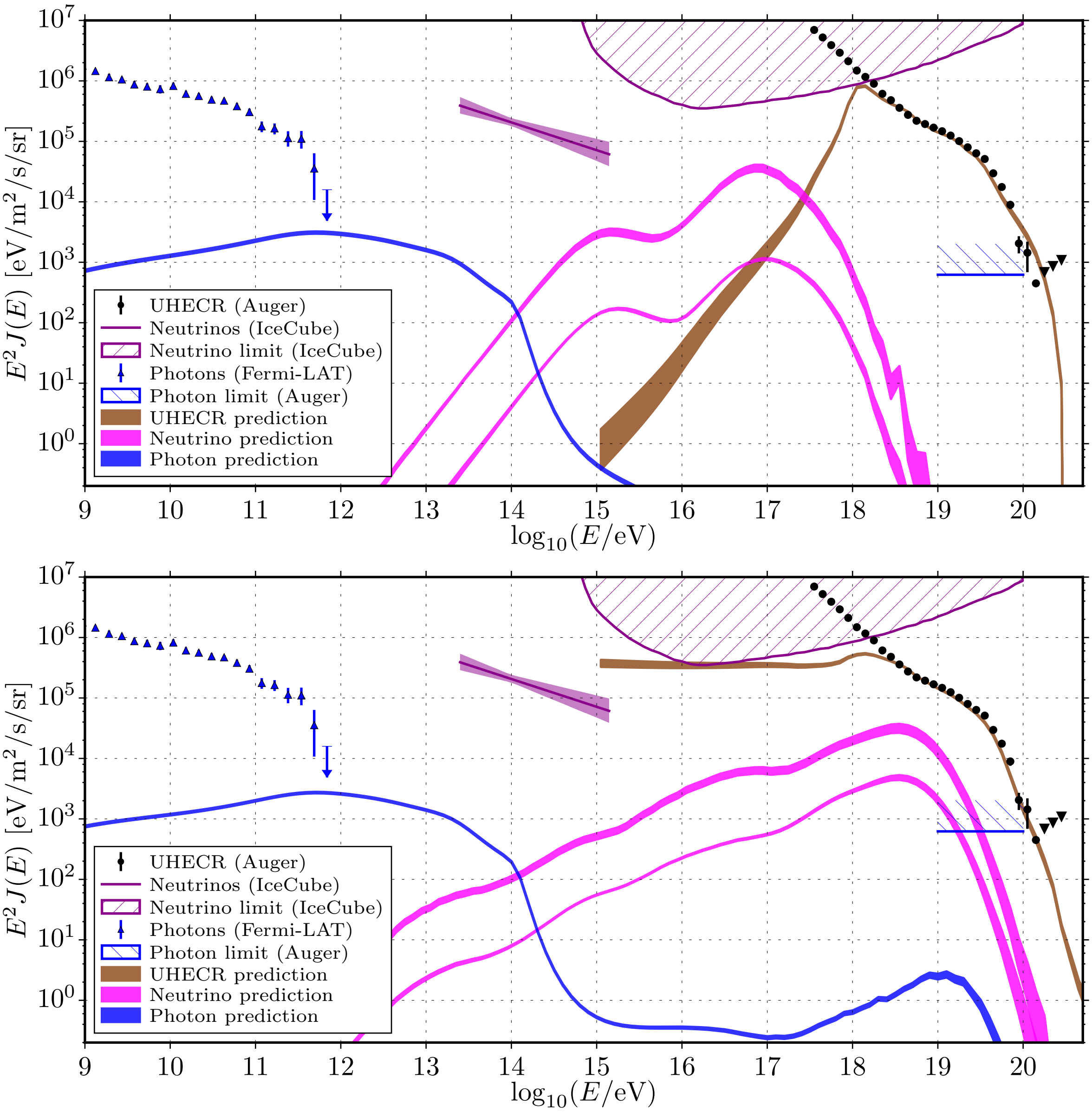


Figure 7.12: Simulated flux of cosmogenic photons (blue) and neutrinos (magenta) that are produced by the hadronic cosmic rays (brown) with bands indicating the central 68% posterior probability interval. For the neutrinos the upper band shows the cases of a strong source evolution with $m = 6$. The top and bottom figures show the prediction for the global and the second minimum, respectively. In both cases the predicted fluxes are well below the measured fluxes and limits: the diffuse isotropic photon flux by Fermi-LAT [237] and integral photon flux limits by Auger [238] (95% C.L.), the measured neutrino flux level [239] and differential all-flavor neutrino limit [240] (90% C.L.) by IceCube, and the all-particle flux by Auger [30].

rays are accelerated to much higher energies per nucleon of $E/A \approx 1/2 10^{19.8}$ eV. Here the neutrinos from nuclear decays may receive larger Lorentz boosts, and there is a contribution of high-energy neutrinos from the charged pion decay in photopion interactions. This leads to a higher energy neutrino flux peaking at $E_\nu \sim 10^{18.5}\,\mathrm{eV}$. For comparison the measured neutrino flux at PeV energies [239] and the upper limit [240] by IceCube are shown. The predicted cosmogenic neutrino fluxes in either minimum of the baseline scenario are orders of magnitude smaller than the measured flux and the upper flux limit. This remains valid even when emphasizing the cosmogenic neutrino flux through a strong source evolution with $m = 6$, the predicted flux is significantly lower the measurements. Hence, the measured neutrinos at PeV energies are not cosmogenic neutrinos from the considered extragalactic cosmic rays, and the upper flux limit does not constrain this scenario based on the cosmogenic neutrinos alone.

In contrast to neutrinos, the flux of cosmogenic photons is strongly modified in the ensuing electromagnetic cascade. Especially for energies $E_\gamma \sim (10^{14} - 10^{19})\,\mathrm{eV}$ the very high interaction rates on the CMB suppress the predicted photon flux and cause a cascade to lower energies. At sub-PeV energies the predicted photon flux is similar in both minima of baseline scenario. At higher energies, $E_\gamma \sim (10^{17} - 10^{19})\,\mathrm{eV}$, the flux in the second minimum is orders of magnitude larger than in the global minimum. These are photons from neutral pion decays in photopion interactions of cosmic-ray hadrons at the highest energies. Still the predicted flux is considerably smaller than the upper limit to the integral photon flux for energies $E_\gamma > 10^{19}\,\mathrm{eV}$ reported by the Pierre Auger Collaboration [30]. At energies $E_\gamma = 100\,\mathrm{MeV} - 820\,\mathrm{GeV}$ the Fermi Large Area Telescope (Fermi-LAT) has measured the diffuse γ-ray flux [237] at high galactic latitudes. The two main components of this flux are the extragalactic point sources that are too faint or too diffuse to be individually resolved, and the truly diffuse flux by processes such as the electromagnetic cascade induced by UHECRs. Therefore, the Fermi-LAT measurement provides an upper limit to the cosmogenic photon flux. In a recent analysis, the Fermi-LAT collaboration estimated that unresolved point source account for at least $86^{+16}_{-14}\%$ of the measured diffuse γ-ray flux [241] above $50\,\mathrm{GeV}$. Considering this it is very interesting to see that the predicted UHECR induced photon flux approaches the diffuse flux to within one order of magnitude and is thus close to the limit. Additional contributions arise from the sub-ankle cosmic rays and are alone expected to reach or exceed this limit [236].

We conclude that the presented astrophysical scenario is presently compatible with photon and neutrino measurements. Specifically regarding neutrinos and EeV photons, no constraints for the scenario are currently expected. In contrast, it is possible that the predicted electromagnetic cascade induced by cosmic rays below and above the ankle overproduces photons at $E_\gamma < 820\,\mathrm{GeV}$.

7.5 Summary of the Analysis Results

In the following we summarize the main findings of this analysis. A key step in advancing our knowledge about the UHECR universe is to identify scenarios that reproduce all available measurements. We find that a simple model of sources injecting a mixed composition with a power-law spectrum and a rigidity dependent cutoff can reasonably well fit the measured energy spectrum and X_max distributions of cosmic rays above the ankle, $E > 10^{18.7}\,\mathrm{eV}$. In this scenario the expected angular deflections in the extragalactic and galactic magnetic fields are especially large, implying that it is relatively easy to accommodate for the observed high level of isotropy in the cosmic-ray arrival directions. Also, the predicted cosmogenic flux of neutrinos and high energy photons is compatible with the corresponding measurements. Hence, the present scenario is compatible with all current observations and can serve as a working point for further investigations.

Complementary to recent similar analyses the treatment of the experimental uncertainties was improved in the present analysis. Specifically, the predicted and measured observables were compared on the detector level, thereby avoiding unfolding methods that generally increase the statistical and systematic uncertainties. For the energy spectrum the implemented forward folding approach allows for using the correct Poisson statistics. Regarding the mass sensitive X_{max} observable we compared the model prediction to the full X_{max} distributions, which contain more information than its first two moments, considered previous analyses. Additionally we considered the main systematic uncertainties as nuisance parameters in a Bayesian treatment.

We find that, while the energy spectrum and the X_{max} distributions alone leave a large allowed parameter space, the combination of both observables strongly constrains the model parameters. In the parameter plane $(R_{\mathrm{cut}}, \gamma)$ the global minimum is found inside an elongated valley at low cutoff rigidities and hard spectral indices $\gamma < 1$. Considering the uncertainties in the source model and the simulation, the spectral index can be compatible with the expectation for fast-rotating neutron stars, $\gamma \sim 1$, but is incompatible with Fermi acceleration, $\gamma \approx 2-2.3$. A second disjunct minimum is found at a Fermi compatible $\gamma \approx 2$, but is strongly disfavored, mainly because the observation of an increasingly heavy and pure mass composition with increasing energy is not reproduced. These observations require that lighter cosmic-ray components vanish early and rapidly within the fitted energy range above the ankle. This is realized by the low optimum cutoff rigidity of $10^{18.6}\,\mathrm{eV}$ in the baseline scenario. Here, the injected composition is found to be dominated by nitrogen as a representative for other elements of intermediate mass. The light observed composition around the ankle is then due to light secondary particles from photodisintegration during propagation. The extended valley in $(R_{\mathrm{cut}}, \gamma)$ shows that even lower cutoff rigidities are possible, if the sources are able to provide extremely hard spectral indices, e.g. $\gamma \approx -1$ for $R_{\mathrm{cut}} \approx 10^{18.2}\,\mathrm{eV}$. In that case the cosmic-ray energies are too low for the efficient photodisintegration on the CMB, and thus the light observed composition around the ankle is due to a dominating light injected composition at the onset of a Peters cycle.

The experimental systematic uncertainties are found to have a large impact on the goodness-of-fit, but their variation does not strongly change the parameter results in the baseline scenario and most of the considered scenario variations. For the fit results the most important systematic uncertainties are that of the energy scale and the X_{max} scale, which are positively correlated in the fit due to the energy dependence of the the X_{max} observable. In all cases a negative shift of the energy scale is preferred, because, if the true cosmic-ray energies are lower than observed, the level of photodisintegration during the extragalactic propagation decreases and the X_{max} distributions are better reproduced.

As the main simulation uncertainties we assessed the role of the photodisintegration model,

the EBL model and the hadronic air-shower model. Using the photodisintegration models by Kossov or the PSB, the goodness-of-fit increases significantly compared the baseline TALYS model, while the parameter results remain relatively unchanged. This can be attributed to the lower cross sections for the emission of α particles, which are specifically relevant because of the low threshold energy, the large loss of the primary cosmic ray's mass and the resulting flux of α particles of relatively high energy, $E_\alpha = 4/AE$.
As variations for the EBL we considered the Franceschini *et al.* and Dominguez *et al.* model. The most important difference for the present scenario is the higher infra-red peak compared to the baseline model by Gilmore *et al.*, which leads to higher photodisintegration rates around $E/A \sim 10^{18}\,\mathrm{eV}$. This causes the parameter results to shift to lower cutoff rigidities inside the valley in $(R_{\mathrm{cut}}, \gamma)$, thereby reducing the importance of photodisintegration.
Regarding the hadronic interaction modeling in the air shower simulation, we find that QGSJetII-04 and Sibyll 2.1 require extremely low cutoff rigidities and power-law indices that are at the boundary of the allowed prior parameter range $\gamma > -3$. The goodness-of-fit decreases compared to the baseline EPOS-LHC model, specifically for QGSJetII-04. When including the systematic uncertainties the parameter results are more reasonable, i.e. $\gamma > 0$, and the goodness-of-fit is comparable to the baseline fit.

In order to test the impact of simplifying assumptions in the source model we considered variations of the cutoff shape in the source spectrum, the source evolution and the representative source elements. The cutoff shape depends on details of the accelerator region and is theoretically not well constrained. Therefore, we tested several cutoff functions of varying abruptness, among which the baseline broken exponential function represents an intermediate choice. We find that the smoother cutoff shapes, such as an exponential cutoff, require slightly lower cutoff rigidities for the rapid transition towards heavier cosmic-ray mass, and harder spectral indices in order to reproduce the observed energy spectrum. Likewise, with a sharp cutoff the cutoff rigidity the cutoff rigidity is higher, necessarily allowing for reaching the observed highest energies, and the source spectrum softer, $\gamma \approx 1.2$. The injected elemental composition is similar in all cases with a dominating nitrogen component.
In the baseline scenario the sources are assumed to be non-evolving. Typical source candidates show a positive evolution with redshift, $(1+z)^m$ with evolution parameters $m > 0$, meaning that the injection rate was higher in the past. In contrast a negative source evolution $m < 0$ represents an overabundance of nearby sources. We find that in the baseline scenario the goodness-of-fit increases for either positive of negative m. For positive m the level of photodisintegration increases due to the higher average source distance, and hence increasingly lower cutoff rigidities and harder spectral indices are required to again suppress these interactions. For negative m, cutoff rigidity and spectral index likewise increase, while the elemental composition remains unchanged compared to the baseline scenario with $m = 0$.
The validity of considering a specific set of injected elements depends on whether the results are representative for a more full elemental composition that is likely to be accelerated inside a source region. In the global minimum of the baseline scenario a dominant nitrogen fraction and a necessary small fraction of iron are found. Significant admixtures of hydrogen and helium are allowed but not necessary because they can be substituted with the corresponding

secondary particles from photodisintegration. Hence, the intermediate element is most the important component of the scenario, for which we considered a range of alternative representatives. With the exception of ^{12}C, all intermediate elements give similar parameter results due to their similar properties regarding photodisintegration. However, the minimum deviance decreases significantly compared to baseline scenario when slightly higher intermediate elements such as $^{16}_{8}$O or $^{20}_{10}$Ne are used. This should be taken into account when assessing the absolute value of the goodness-of-fit. No further improvements where found when considering five or more injected elements, thus we confirm that a small set of elements can indeed serve to represent a more complex elemental composition at the sources.

We conclude that there are multiple uncertainties in modeling the injection, propagation and observation of cosmic rays, which translate into considerable uncertainties in the parameter results of the present scenario. Nevertheless, across all scenario variations two main features are essential for achieving a reasonable fit to the measured energy spectrum and X_{max} distribution: First, a mixed source composition and an early cutoff of the source spectrum, occurring for protons around or below the ankle, are required for the observed transition towards a heavier cosmic-ray mass with increasing energy. Second, the required hard spectral index $\gamma < 1.2$ is incompatible with Fermi acceleration. Thus, the general results obtained in this analysis significantly constrain the parameter space for scenarios with a rigidity dependent maximum acceleration energy.

Chapter 8

Conclusions

The past years have seen major advancements in the quality and quantity of cosmic-ray measurements. Testing theories about the still unknown origin of cosmic rays against these measurements is a challenging task due to the large open parameter space and many interconnected aspects of cosmic-ray acceleration and propagation. Here, detailed simulations have proven to be an indispensable tool for connecting theories and observations. These simulations combine our knowledge in many fields of physics, including astrophysics and cosmology as well as plasma, nuclear and elementary particle physics. The existing simulation codes, however, are either limited in scope or not publicly available. In order to provide the astroparticle community with an open and flexible simulation tool the software CRPropa 3 was co-developed in the course of this thesis. CRPropa is a general simulation framework for the galactic and extragalactic Monte Carlo propagation of ultra-high energy cosmic-ray nuclei, photons and neutrinos. It enables the calculation of the cosmic-ray arrival directions, energy spectrum and composition for arbitrary source configurations, as well as the fluxes of secondary particles in a multi-messenger approach. Due to its general applicability, modular structure and well documented open development, CRPropa 3 can serve as a standard tool for cosmic-ray physics.

Making use of CRPropa 3, we simulated the predicted cosmic-ray flux for a scenario of uniformly distributed extragalactic sources, emitting a mixed composition with the nuclear contributions, the power of the power-law spectrum, and the rigidity dependent cutoff as the model parameters. The predicted cosmic-ray flux was then compared to the Pierre Auger Observatory measurements of the energy spectrum and the composition-sensitive depth of the shower maximum X_{max}. To this end we formulated the likelihood of simulated and measured observables on the detector level, thus allowing for a better account of the detector response and a higher statistical power compared to previous analyses. A Bayesian analysis was then performed to optimize the model parameters, and to calculate the probability distributions of the model parameters given the measurements. We also quantified the impact of the main systematic uncertainties in the detector description and the simulation.

We find that in general the considered source model can well describe the measurements, but only for source spectra that are significantly harder than expected for Fermi-type accelerators and for low maximum rigidities. On the simulation side the largest uncertainties stem from the branching ratios in the photodisintegration model, the extragalactic background light density in the infrared, and the hadronic air shower modeling. On the detector side, the largest uncertainties are due to the energy scale and the X_{max} scale.

An important implication of the considered scenario is that it does not account for the cosmic-ray flux below the ankle, thus requiring a different origin of cosmic rays at these energies. In the scenario the expected cosmic-ray deflections in galactic and extragalactic magnetic fields are large up to the highest observed energies, and the predicted fluxes of high energy cosmogenic photons and neutrinos are below the current experimental sensitivities, hence no constraints are expected from these observables. The presented method can be used in future analyses for investing different source scenarios and for modeling cosmic rays over a larger energy range in order to further expand our understanding of the cosmic-ray universe.

Appendix A

Appendix

A.1 TALYS Settings

Photodisintegration is the most important interaction for cosmic-ray nuclei with energies $E > 10^{19}$ eV. Cross sections for this interaction are dominated by the giant dipole resonance (GDR) for photons with energies $\epsilon' < 30$ MeV in the nucleus rest frame. In Ref. [184] an "accurate description" of the available experimental data was found using a preliminary version of TALYS. TALYS was used in this comparison [185] with the GDR parameters from the IAEA atlas [181]. In contrast, the publicly available versions of TALYS by default uses the GDR parameters from the RIPL-2 database [242], which leads to predicted cross sections at variance with the experimental data on the total cross sections, see figure A.1.

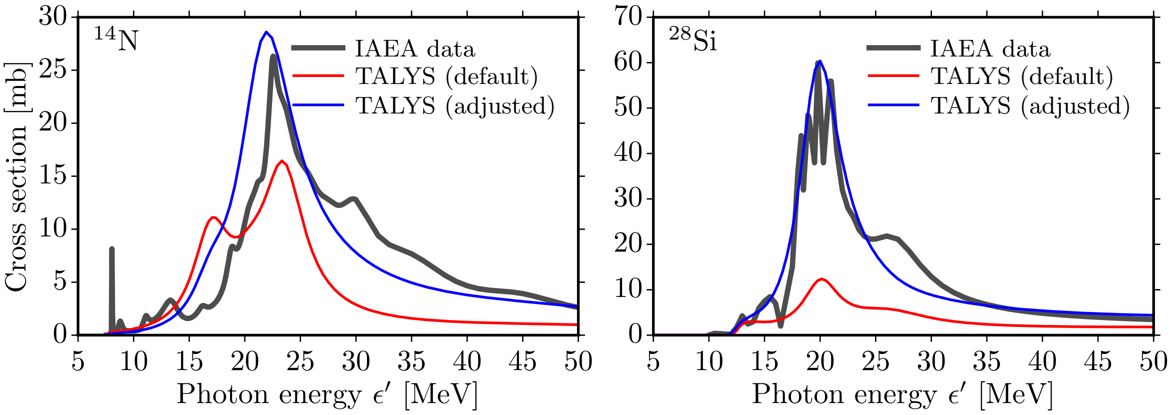


Figure A.1: Comparison of total photodisintegration cross sections for ^{14}N (left) and ^{28}Si (right) with the evaluated experimental data compiled in the IAEA atlas [181], see section A.2. TALYS (default) and TALYS (adjusted) correspond to the models implemented in CRPropa 2 and CRPropa 3, respectively.

Therefore, for CRPropa 3 we use TALYS with the GDR parameters of the IAEA atlas, if available, as the resulting cross sections are in much better agreement with the available measurements. The complete list of used GDR parameters is given in table A.1. Note that the current version of CRPropa 2 still uses TALYS with the default GDR parameters, and should therefore not be used for propagating cosmic-ray nuclei.

Isotope	E_0 [MeV]	σ_0 [mb]	Γ_0 [MeV]	E_1 [MeV]	σ_1 [mb]	Γ_1 [MeV]	Source
C-12	22.70	21.36	6.00				Atlas
N-14	22.50	27.00	7.00				Atlas
O-16	22.35	30.91	6.00				Atlas
Na-23	23.00	15.00	16.00				Atlas
Mg-24	20.80	41.60	9.00				Atlas
Al-27	21.10	12.50	6.10	29.50	6.70	8.70	RIPL-2
Si-28	20.24	58.73	5.00				Atlas
Ar-40	20.90	50.00	10.00				Atlas
Ca-40	19.77	97.06	5.00				Atlas
V-51	17.93	53.30	3.62	20.95	40.70	7.15	RIPL-2
Mn-55	16.82	51.40	4.33	20.09	45.20	4.09	RIPL-2

Table A.1: Giant dipole resonance parameters used with TALYS (as parameters for the Kopecky-Uhl generalized Lorentzian model of the E1-strength function): peak energy E_i, peak cross section σ_i and width Γ_i for resonances with a single ($i = 0$) or a split peak ($i = 0, 1$). Default values from the RIPL-2 database are replaced, if available, with the total cross section parameters from the atlas of GDR parameters. Note that for isotopes not listed, as well as for higher order contributions, TALYS uses a compilation of formulas listed in [183].

A.2 Comparison of Photodisintegration Models

Photo-induced reactions are relevant in the context of e.g. radiation shielding, radiotherapy, nuclear transmutation and as second order effects for fission and fusion. Consequently, a large body of photonuclear cross sections has been measured for nuclides that are typical to these use cases. The mainly used experimental techniques are the detection of emitted particles from a target material when irradiated with a continuous bremsstrahlung spectrum, or with quasi-monoenergetic photons from positron annihilation. There are significant systematic uncertainties inherent to these techniques (cf. [181] for a discussion) and efforts have been made by several national nuclear agencies to compile databases of *evaluated* cross sections, that reevaluate and combine the various measurements. A combination of these databases is provided in form of the IAEA Photonuclear Data Library [181], which contains a set of recommended evaluated photonuclear data for 164 nuclides and photon energies up to 140 MeV. Unfortunately, only 49 out of these nuclides lie in the range $Z \leq 26$, which represents only a minor fraction of the nuclides that are potentially relevant for cosmic-ray propagation as shown in figure A.2.

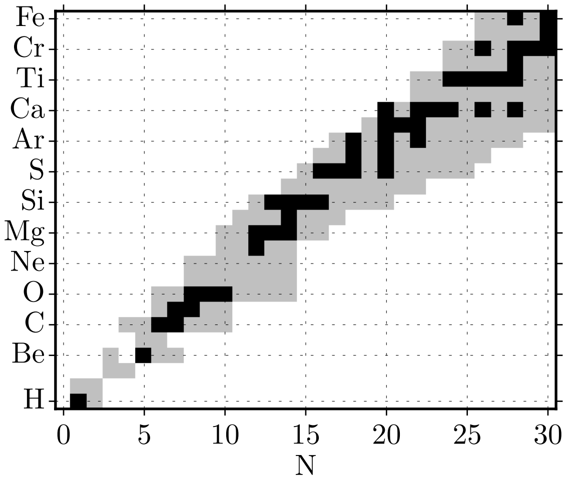


Figure A.2: Overview of isotopes in the IAEA Photonuclear Data Library [181] (black) and all isotopes (gray) with a lifetime of $\tau > 2\,\mathrm{s}$, see chapter 4.7.

In order estimate the global agreement between the photodisintegration models and the data, we perform a comparison of the resulting interactions rates with the CMB and EBL for all 49 isotopes in the IAEA database. Here, we consider the EBL model by Gilmore *et al.* 2012, noting that the choice of the EBL model has little impact for the following comparison. The cross sections in the IAEA library are tabulated for varying photon energy ranges. For comparability the photon energy range for each nuclide is chosen identically for model and data, spanning from 100 keV to the maximum energy in the IAEA library but not larger than 50 MeV. The interaction rates are calculated for cosmic ray energies between $E = 10^{18.5} - 10^{20.5}\,\mathrm{eV}$, since the region around and above the ankle is most relevant for the current astrophysical interpretation.

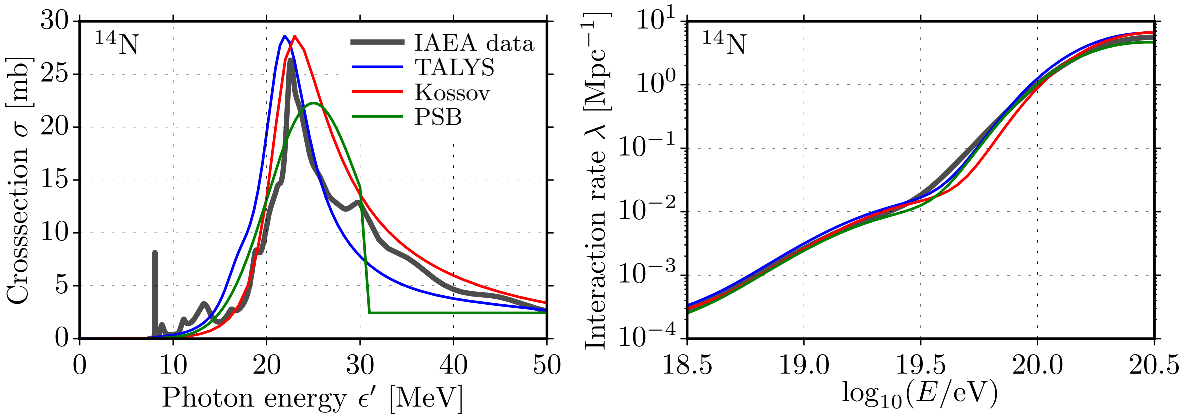


Figure A.3: Photodisintegration cross section (left) for ^{14}N from the IAEA database and resulting interaction rate (right), in comparison with the TALYS, Kossov and PSB models.

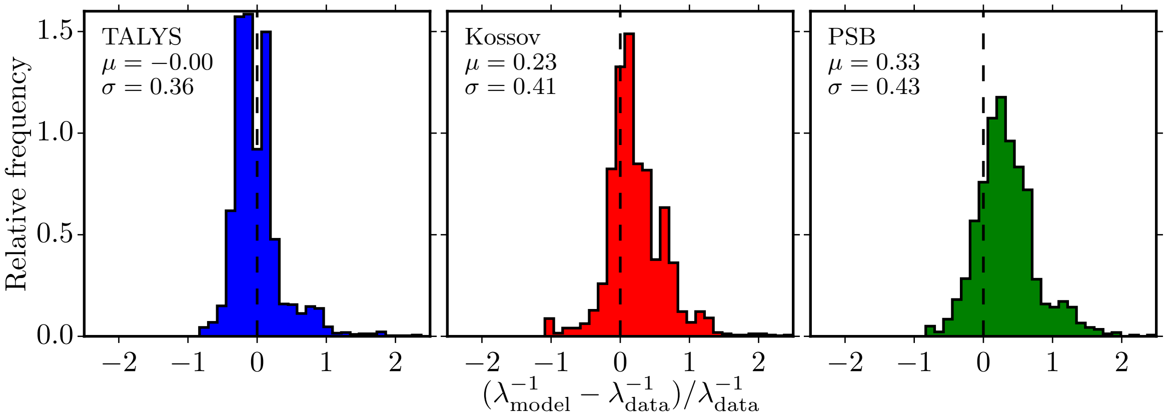


Figure A.4: Distributions for the measure of agreement between the models and the IAEA for all isotopes in A.2 and log-spaced energies in the range $\log_{10}(E/\mathrm{eV}) = 18.5 - 20.5$.

Figure A.3 shows the comparison of the cross sections and interaction rates for the example of ^{14}N. While all model reproduce the general shape of the giant dipole resonance, there are considerable differences in between models and data and among the models. This translates into uncertainties in modeling the total interaction rates. In order to derive a global measure of agreement the relative difference

$$\Delta(E) = \left(\lambda^{-1}_{\mathrm{model}}(E) - \lambda^{-1}_{\mathrm{data}}(E)\right) / \lambda^{-1}_{\mathrm{data}}(E) \tag{A.1}$$

is computed for each nuclide. Then, for each model the distribution $f(\Delta)$ of the relative differences for all nuclides and all energies $\log_{10}(E/\mathrm{eV}) = 18.5 - 20.5$ in logarithmic steps is calculated, as shown in figure A.4.

We find that the Kossov and PSB model in average overestimate the total interaction rate with respect to the IAEA data. All models show a similar spread compared to the data, with the TALYS model performing slightly better. The central 68% uncertainty intervals of the distributions of relative difference read $(-0.28, 0.23)$ for the TALYS model, $(-0.08, 0.66)$ for

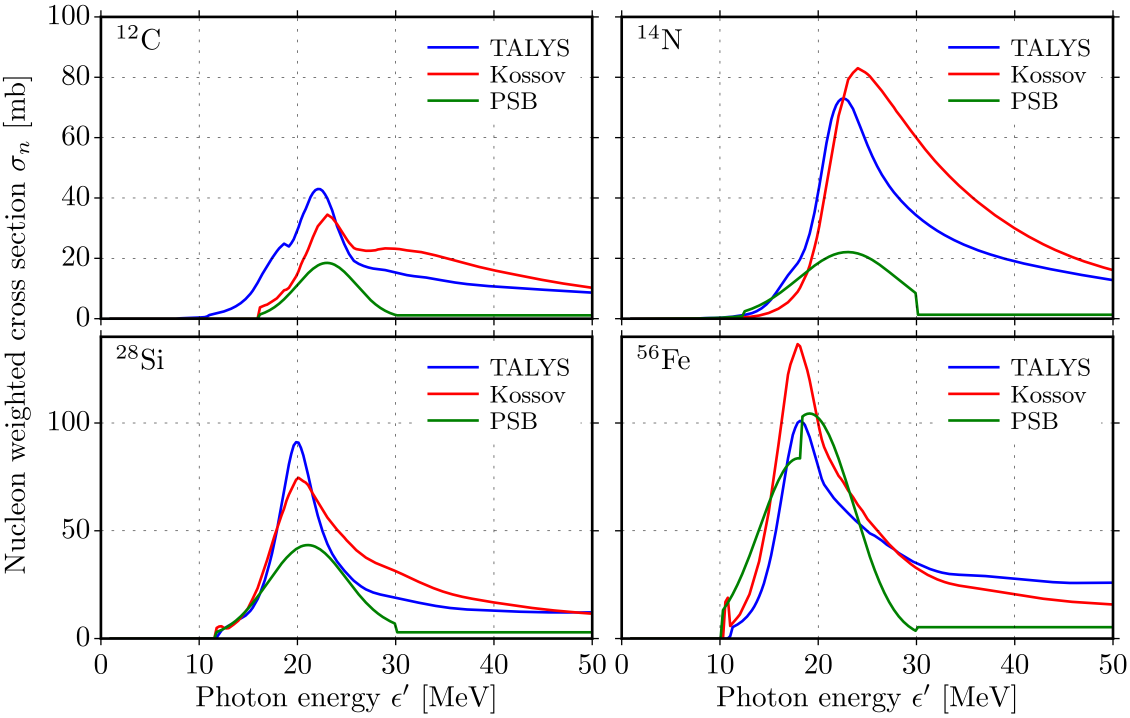


Figure A.5: Comparison of the nucleon weighted cross sections σ_n, indicating the different impact of photodisintegration for the TALYS, Kossov and PSB models. Since the PSB model only considers one and two nucleon emission, the weighted cross section is systematically lower.

the Kossov model, and (−0.05, 0.70) for the PSB model. We note that this overall measure of agreement depends on the choice of isotopes and energy range, and that depending on a specific cosmic ray scenario, certain isotopes have a larger impact than others.

Apart from the total interaction rate, the impact of photodisintegration also depends on the branching ratios of the individual evaporation channels. For instance, the emission of an α-particle represents a four times more efficient disintegration than the emission of a single nucleon. Here, the differences between the models are particularly large, with the TALYS model predicting considerable branching ratios for the α-channel, whereas the PSB model at low collision energies only considers the emission of or two nucleons. Unfortunately, the *evaluated* experimental data regarding the branching ratios is very sparse and does not allow for an overall comparison. A comparison with *unevaluated* experimental data for a few important nuclides, such as ^{12}C, indicates that TALYS significantly overpredicts the cross section of the α-channel [141].

In order to highlight the impact of the different branching ratios in the models, we compare the nucleon weighted cross sections $\sigma_n(\epsilon') = \sum_i n_i \sigma_i(\epsilon')$, where i denotes the individual channels with the emission of n_i nucleons. This is shown in figure A.5 for the cases of ^{12}C, ^{14}N, ^{28}Si and ^{56}Fe. As expected the nucleon weighted cross sections are systematically higher in the TALYS and Kossov models compared to the PSB model. The differences are most pronounced for nuclei of low and intermediate mass, where the branching ratios for the α channel are comparatively larger. Note that while the Kossov model is used in combination with the branching ratios from TALYS, the nucleon weighted cross sections are different, because of a different total cross section. The differences between the TALYS and Kossov models are most pronounced for low mass nuclei, in particular ^{12}C, where Kossov model features a higher threshold energy.

A.3 Empirical Fit to the Spectrum and $X_{\max}$ Distributions

In this section we perform separate fits to the energy spectrum and $X_{\max}$ distributions above $E = 10^{18.7}$ eV using suitable empirical models to describe these measurements. We argue that the obtained goodness-of-fits can serve as a reference for the astrophysical modeling.

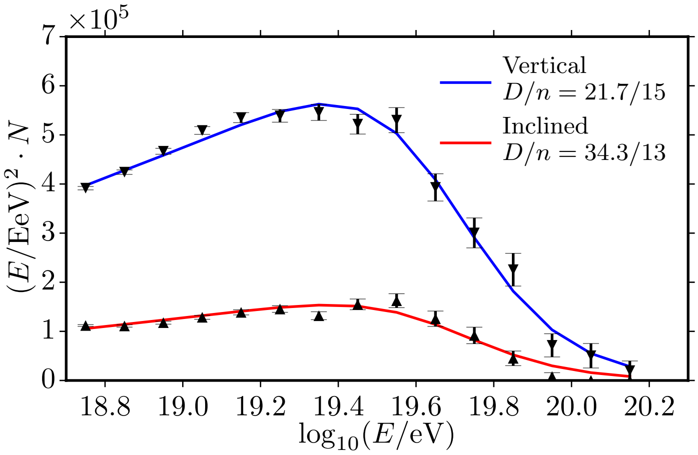


Figure A.6: Simultaneous fit with an empirical flux model to the vertical and inclined measurements of the raw energy spectrum.

For the energy spectrum we consider the empirical flux model presented in [30], describing the spectrum by a power-law below the ankle, $J(E) = J_0(E/E_{\mathrm{ankle}})^{\gamma_1}$, and a power-law with a smooth suppression beyond:

$$J(E) = J_0 \left(\frac{E}{E_{\mathrm{ankle}}}\right)^{\gamma_2} \left[1 + \left(\frac{E_{\mathrm{ankle}}}{E_{\mathrm{s}}}\right)^{\Delta\gamma}\right] \left[1 + \left(\frac{E}{E_{\mathrm{s}}}\right)^{\Delta\gamma}\right]^{-1} . \tag{A.2}$$

Since we are only interested in the trans-ankle region, we can omit the former term. To compare with vertical and inclined measurements, the flux model is folded with the detector response as in equation 6.7, reading

$$N^V(E^m_{\mathrm{rec}}) = \omega^V \sum_l M^V(E^m_{\mathrm{rec}}, E^l)\, J(E^l)\, \Delta E^m_{\mathrm{rec}} \tag{A.3}$$

$$N^I(E^m_{\mathrm{rec}}) = \omega^I \sum_l M^I(E^m_{\mathrm{rec}}, E^l)\, J(E^l)\, \Delta E^m_{\mathrm{rec}} . \tag{A.4}$$

As in the astrophysical fit, the flux normalization J_0 and the individual exposures represent three parameters for two independent quantities regarding the energy spectrum. We therefore fit the flux normalization J_0 and the relative exposure parameter $r = \omega^I/\omega^V$. The model is thus described by six parameters, yielding $ndf = 15 + 13 - 6 = 22$ degrees of freedom. We can employ the same deviance as in the astrophysical fit, see equation 6.25. The best fit in figure A.6 shows a qualitatively good description of the measurements. The χ^2 per degrees of freedom is $D/ndf = 56.0/22$. Apparently the measured spectra exhibit a slightly more complex structure than the smooth flux model can reproduce. In the considered astrophysical scenario, see section 6.1.1, the source spectrum is described by a similarly simple model and the appearance of more complex features in the observed spectrum is limited to propagation effects. Therefore, the goodness-of-fit from the empirical flux model can reasonably serve as a reference value for the astrophysical fit.

For the $X_{\max}$ distributions we follow the Auger analysis [34] by fitting the contributions of four representative mass numbers $A = 1, 4, 14$ and 56. separately in each energy bin. Analogously to the model prediction 6.9 in the astrophysical fit, in each energy bin the $X_{\max}$ distribution according to a set of mass number fractions $f(A^k)$ is given by

$$p(X_{\max}^{\mathrm{rec},x}) = c \sum_k f(A^k)\, G(A^k, X_{\max}^{\mathrm{rec},x})\,. \tag{A.5}$$

Here, $G(A^k, X_{\max}^{\mathrm{rec},x})$ is the $X_{\max}$ template including the detector response as described in section 6.2.2, and c normalizes the right hand side. Due to the side condition $\sum_k f(A^k) = 1$ there are three independent mass fractions in each energy bin, hence, 27 parameters in the energy range above $10^{18.7}$ eV.

The best fit results are shown in figure A.7 for the case of EPOS-LHC as hadronic interaction model in the underlying air shower simulations. The sum of minimum deviances above $10^{18.7}$ eV is $D = 122.4$ compared to $ndf = 116 - 27 = 99$ degrees of freedom. When using QGS-JetII-04 and Sibyll 2.1 the minimum deviances are $D = 202.8$ and $D = 129.4$, respectively.

As a cross-check for the present analysis we compare in figure A.8 the fitted mass fractions to the results reported in [34], in which templates from shower simulations were used instead of the $X_{\max}$ parametrization from [213]. The results are seen to be in fair agreement, showing the validity of the considered approach. In the case of EPOS-LHC, the mass fractions are seen to have rather complex energy dependence, for instance showing a re-emergence of a proton ($A = 1$) fraction in the energy bin $10^{19.4-19.5}$ eV, which is necessary to describe the tails in the measured $X_{\max}$ distributions. This feature is also seen for the cases of QGSJetII-04 and Sibyll 2.1. However instead of a large nitrogen fraction ($A = 14$) these models predict a dominant Helium ($A = 4$) fraction in the energy range above the ankle. As for the energy spectrum we argue that the our simple astrophysical model cannot outperform this empirical fit, which uses independent mass fractions in each energy bin. The obtained deviances thus serve as a reference for the maximum achievable goodness-of-fit in the astrophysical fit.

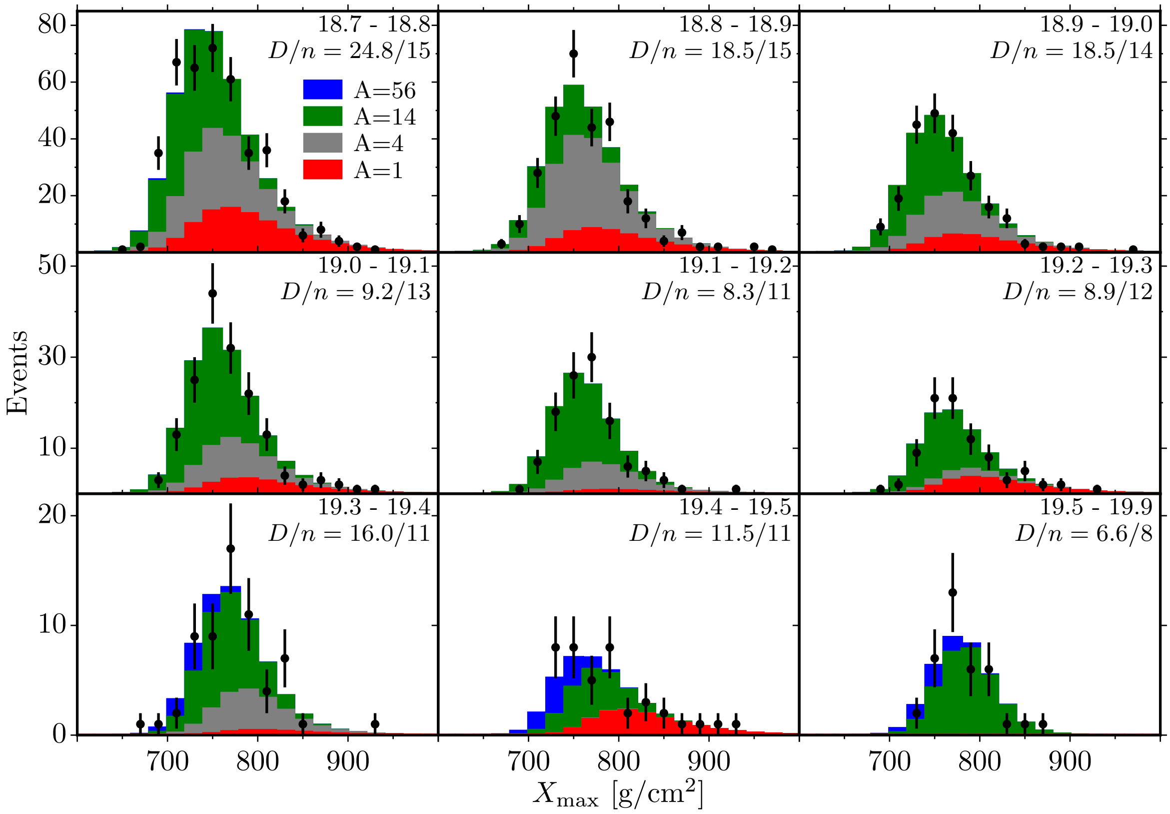


Figure A.7: Best fit result of fitting four mass fractions $A = 1, 4, 14$ and 56 to the measured X_{max} distributions (black circles). Indicated in the upper right of each plot are energy bin extent (in $\log_{10}(E/\mathrm{eV})$) and the deviance divided by the number of data points.

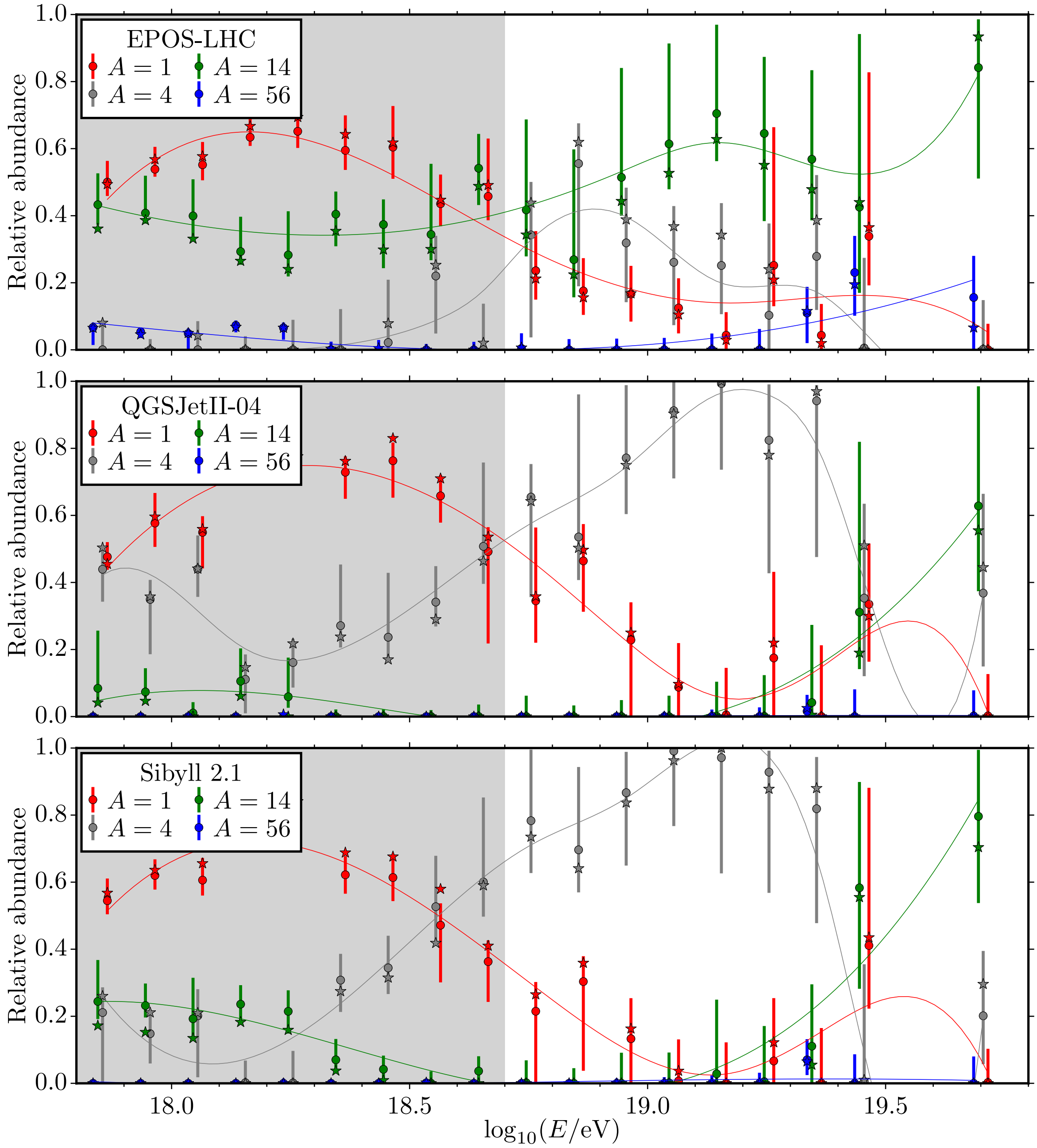


Figure A.8: Comparison of the most likely mass fractions obtained in the present analysis (circles and error bars showing the symmetrical 68% interval) with the results from [34] (stars) as function of energy. Solid lines highlight the general trend for each mass number. The energy range below $10^{18.7}$ eV is marked in gray. The plots from top to bottom show the analyses assuming EPOS-LHC, QGSJetII-04 and Sibyll 2.1 as hadronic interactions models in the underlying shower simulations.

A.4 Simulation Settings

The following python script describes the simulation setup for the analysis presented in 6. The simulations were run on RWTH compute cluster using CRPropa 3, revision 99b78ec. Note that the constraint on the minimum source distance is applied after the simulation.

```
from crpropa import *

EBL = IRB_Gilmore12
A, Z, n = ...
Emin, Emax = 10**17.5 * eV, Z*10**21.5 * eV

# module setup
m = ModuleList()
m.add(SimplePropagation())
m.add(Redshift())
m.add(PhotoPionProduction(CMB))
m.add(PhotoPionProduction(EBL))
m.add(PhotoDisintegration(CMB))
m.add(PhotoDisintegration(EBL))
m.add(NuclearDecay())
m.add(ElectronPairProduction(CMB))
m.add(ElectronPairProduction(EBL))

# observer
obs = Observer()
obs.add(ObserverPoint())
obs.onDetection(TextOutput(filename, Output.Event1D))
m.add(obs)

# source
source = Source()
source.add(SourceUniform1D(0, redshift2ComovingDistance(2)))
source.add(SourceRedshift1D())
source.add(SourcePowerLawSpectrum(Emin, Emax, -1))
source.add(SourceParticleType(nucleusId(A, Z)))

m.run(source, n, True)
```

Bibliography

[1] K.-H. Kampert and A. A. Watson. “Extensive Air Showers and Ultra High-Energy Cosmic Rays: A Historical Review”. In: *Eur. Phys. J.* H37 (2012), pp. 359–412. DOI: `10.1140/epjh/e2012-30013-x`. arXiv: `1207.4827 [physics.hist-ph]`.

[2] R. Engel, D. Heck, and T. Pierog. “Extensive air showers and hadronic interactions at high energy”. In: *Ann. Rev. Nucl. Part. Sci.* 61 (2011), pp. 467–489. DOI: `10.1146/annurev.nucl.012809.104544`.

[3] B. G. Keilhauer. “Investigation of Atmospheric Effects on the Development of Extensive Air Showers and their Detection with the Pierre Auger Observatory”. PhD thesis. Karlsruhe University, 2003.

[4] A. Letessier-Selvon and T. Stanev. “Ultrahigh energy cosmic rays”. In: *Rev. Mod. Phys.* 83 (July 2011), pp. 907–942. DOI: `10.1103/RevModPhys.83.907`. arXiv: `1103.0031 [astro-ph.HE]`.

[5] K. Kamata and J. Nishimura. “The lateral and the angular structure functions of electron showers”. In: *Progress of Theoretical Physics Supplement* 6 (1958), p. 93.

[6] K. Greisen. “Cosmic ray showers”. In: *Annu. Rev. Nucl. Sci.* 10 (1960), pp. 63–108.

[7] T. K. Gaisser and A. M. Hillas. “Reliability of the method of constant intensity cuts for reconstructing the average development of vertical showers”. In: *Proceedings of the 15th ICRC, Plovdiv.* Vol. 8. 1977, pp. 353–357.

[8] W. Heitler. “The Quantum Theory of Radiation”. In: Oxford University Press, 1954. Chap. Section 38, p. 386.

[9] J. Matthews. “A Heitler model of extensive air showers”. In: *Astropart. Phys.* 22 (2005), pp. 387–397. DOI: `10.1016/j.astropartphys.2004.09.003`.

[10] D. Heck, G. Schatz, T. Thouw, J. Knapp, and J. N. Capdevielle. *CORSIKA: A Monte Carlo code to simulate extensive air showers.* FZKA 6019 6019. Forschungszentrum Karlsruhe, 1998.

[11] P. Lipari. “Cosmic rays and hadronic interactions”. In: *EPJ Web Conf.* 99 (2015), p. 14001. DOI: `10.1051/epjconf/20159914001`.

[12] R. Ulrich, C. Baus, and R. Engel. “Relation between hadronic interactions and ultra-high energy extensive air showers”. In: *EPJ Web Conf.* 99 (2015), p. 11001. DOI: 10.1051/epjconf/20159911001.

[13] M. L. Miller, K. Reygers, S. J. Sanders, and P. Steinberg. “Glauber modeling in high energy nuclear collisions”. In: *Ann. Rev. Nucl. Part. Sci.* 57 (2007), pp. 205–243. DOI: 10.1146/annurev.nucl.57.090506.123020. arXiv: nucl-ex/0701025 [nucl-ex].

[14] K. Werner, I. Karpenko, and T. Pierog. “The ’Ridge’ in Proton-Proton Scattering at 7 TeV”. In: *Phys. Rev. Lett.* 106 (2011), p. 122004. DOI: 10.1103/PhysRevLett.106.122004. arXiv: 1011.0375 [hep-ph].

[15] S. Ostapchenko. “Monte Carlo treatment of hadronic interactions in enhanced Pomeron scheme: QGSJET-II model”. In: *Phys. Rev.* D83 (2011), p. 014018. DOI: 10.1103/PhysRevD.83.014018. arXiv: 1010.1869 [hep-ph].

[16] E. J. Ahn, R. Engel, T. K. Gaisser, P. Lipari, and T. Stanev. “Cosmic ray interaction event generator SIBYLL 2.1”. In: *Phys. Rev.* D80 (2009), p. 094003. DOI: 10.1103/PhysRevD.80.094003. arXiv: 0906.4113 [hep-ph].

[17] F. Riehn, R. Engel, A. Fedynitch, T. K. Gaisser, and T. Stanev. “A new version of the event generator Sibyll”. In: *Proceedings of the 34th International Cosmic Ray Conference (ICRC 2015)*. 2015. arXiv: 1510.00568 [hep-ph].

[18] D. d’Enterria, R. Engel, T. Pierog, S. Ostapchenko, and K. Werner. “Constraints from the first LHC data on hadronic event generators for ultra-high energy cosmic-ray physics”. In: *Astropart. Phys.* 35 (2011), pp. 98–113. DOI: 10.1016/j.astropartphys.2011.05.002. arXiv: 1101.5596 [astro-ph.HE].

[19] A. Aab et al. “Muons in air showers at the Pierre Auger Observatory: Measurement of atmospheric production depth”. In: *Phys. Rev.* D90.1 (2014). [Erratum: Phys. Rev.D92,no.1,019903(2015)], p. 012012. DOI: 10.1103/PhysRevD.92.019903, 10.1103/PhysRevD.90.012012, 10.1103/PhysRevD.90.039904. arXiv: 1407.5919 [hep-ex].

[20] W. D. Apel et al. “Kneelike structure in the spectrum of the heavy component of cosmic rays observed with KASCADE-Grande”. In: *Phys. Rev. Lett.* 107 (2011), p. 171104. DOI: 10.1103/PhysRevLett.107.171104. arXiv: 1107.5885 [astro-ph.HE].

[21] J. Blümer, R. Engel, and J. R. Hörandel. “Cosmic rays from the knee to the highest energies”. In: *Progress in Particle and Nuclear Physics* 63 (Oct. 2009), pp. 293–338. DOI: 10.1016/j.ppnp.2009.05.002. arXiv: 0904.0725 [astro-ph.HE].

[22] F. Aharonian et al. “First ground based measurement of atmospheric Cherenkov light from cosmic rays”. In: *Phys. Rev.* D75 (2007), p. 042004. DOI: 10.1103/PhysRevD.75.042004. arXiv: astro-ph/0701766 [ASTRO-PH].

[23] A. U. Abeysekara et al. “HAWC Contributions to the 34th International Cosmic Ray Conference (ICRC2015)”. In: *Proceedings, 34th International Cosmic Ray Conference (ICRC 2015)*. 2015. arXiv: 1508.03327 [astro-ph.HE].

[24] M. Tluczykont et al. “The HiSCORE concept for gamma-ray and cosmic-ray astrophysics beyond 10 TeV”. In: *Astropart. Phys.* 56 (2014), pp. 42–53. DOI: `10.1016/j.astropartphys.2014.03.004`. arXiv: `1403.5688 [astro-ph.IM]`.

[25] G. A. Askar’yan. “Excess negative charge of an electron-photon shower and its coherent radio emission”. In: *Sov. Phys. JETP* 14.2 (1962). [Zh. Eksp. Teor. Fiz.41,616(1961)], pp. 441–443.

[26] J. Kelley for the Pierre Auger Collaboration. “AERA: the Auger Engineering Radio Array”. In: *Pierre Auger Observatory: Contributions to the 32nd International Cosmic Ray Conference (ICRC 2011)*. 2011.

[27] H. Kawai et al. “Telescope Array Experiment”. In: *Nucl. Phys. B (Proc. Suppl.)* 175–176 (2008). Proceedings of the XIV International Symposium on Very High Energy Cosmic Ray Interactions, pp. 221–226. ISSN: 0920-5632. DOI: `10.1016/j.nuclphysbps.2007.11.002`.

[28] K. A. Olive et al. “Review of Particle Physics”. In: *Chin. Phys.* C38 (2014), p. 090001. DOI: `10.1088/1674-1137/38/9/090001`.

[29] T. Abu-Zayyad et al. “The Cosmic Ray Energy Spectrum Observed with the Surface Detector of the Telescope Array Experiment”. In: *Astrophys. J.* 768 (2013), p. L1. DOI: `10.1088/2041-8205/768/1/L1`. arXiv: `1205.5067 [astro-ph.HE]`.

[30] I. Valiño for the Pierre Auger Collaboration. “The flux of ultra-high energy cosmic rays after ten years of operation of the Pierre Auger Observatory”. In: *Pierre Auger Observatory: Contributions to the 34th International Cosmic Ray Conference (ICRC 2015)*. 2015. arXiv: `1509.03732 [astro-ph.HE]`.

[31] W. D. Apel et al. “The spectrum of high-energy cosmic rays measured with KASCADE-Grande”. In: *Astropart. Phys.* 36 (2012), pp. 183–194. DOI: `10.1016/j.astropartphys.2012.05.023`.

[32] B. Ruzybayev. “Measuring the Cosmic Ray Energy Spectrum and Composition with IceCube”. In: *Phys. Procedia* 61 (2015), pp. 443–449. DOI: `10.1016/j.phpro.2014.12.106`.

[33] V. V. Prosin et al. “Primary CR energy spectrum and mass composition by the data of Tunka-133 array”. In: *EPJ Web Conf.* 99 (2015), p. 04002. DOI: `10.1051/epjconf/20159904002`.

[34] A. Aab et al. “Depth of Maximum of Air-Shower Profiles at the Pierre Auger Observatory: Composition Implications”. In: *Phys.Rev.D* (2014). arXiv: `1409.5083 [astro-ph.HE]`.

[35] R. Abbasi et al. “Report of the Working Group on the Composition of Ultra High Energy Cosmic Rays”. In: *2014 Conference on Ultrahigh Energy Cosmic Rays (UHECR2014) Springdale, USA, October 12-15, 2014*. 2015. arXiv: `1503.07540 [astro-ph.HE]`.

[36] E. S. Seo et al. “Measurement of cosmic-ray proton and helium spectra during the 1987 solar minimum”. In: *Astrophys. J.* 378 (Sept. 1991), pp. 763–772. DOI: `10.1086/170477`.

[37] I. P. Ivanenko et al. “Energy Spectra of Cosmic Rays above 2 TeV as Measured by the ’SOKOL’ Apparatus”. In: *International Cosmic Ray Conference* 2 (1993), p. 17.

[38] N. Grigorov et al. “Energy Spectrum of Primary Cosmic Rays in the 10^{11} - 10^{15} eV According to the Data of Proton-4 Measurements”. In: *Proceedings of the 12th ICRC (1970)*. 1970.

[39] M. Amenomori et al. “The Cosmic ray energy spectrum between $10^{14.5}$ eV and $10^{16.3}$ eV covering the ’knee’ region”. In: *Astrophys. J.* 461 (1995), pp. 408–414. DOI: `10.1086/177069`.

[40] M. Glasmacher et al. “The cosmic ray composition between 10^{14} and 10^{16} eV”. In: *Astropart. Phys.* 12.1-2 (1999), pp. 1–17. ISSN: 0927-6505.

[41] M. G. Aartsen et al. “Measurement of the cosmic ray energy spectrum with IceTop-73”. In: *Phys. Rev.* D88.4 (2013), p. 042004. DOI: `10.1103/PhysRevD.88.042004`. arXiv: `1307.3795 [astro-ph.HE]`.

[42] M. Nagano et al. “Energy spectrum of primary cosmic rays above 10^{17} eV determined from the extensive air shower experiment at Akeno”. In: *J. Phys.* G18 (1992), pp. 423–442. DOI: `10.1088/0954-3899/18/2/022`.

[43] M. A. Lawrence, R. J. O. Reid, and A. A. Watson. “The cosmic ray energy spectrum above 4×10^{17} eV as measured by the Haverah Park array”. In: *J. Phys. G* 17 (1991), p. 733. DOI: `10.1088/0954-3899/17/5/019`.

[44] D. J. Bird et al. “The Cosmic ray energy spectrum observed by the Fly’s Eye”. In: *Astrophys. J.* 424 (1994), pp. 491–502. DOI: `10.1086/173906`.

[45] R. U. Abbasi et al. “First observation of the Greisen-Zatsepin-Kuzmin suppression”. In: *Phys. Rev. Lett.* 100 (2008), p. 101101. DOI: `10.1103/PhysRevLett.100.101101`. arXiv: `astro-ph/0703099 [astro-ph]`.

[46] K.-H. Kampert and P. Tinyakov. “Cosmic rays from the ankle to the cutoff”. In: *Comptes Rendus Physique* 15 (2014), pp. 318–328. DOI: `10.1016/j.crhy.2014.04.006`. arXiv: `1405.0575 [astro-ph.HE]`.

[47] M. Amenomori et al. “Large-scale sidereal anisotropy of Galactic cosmic-ray intensity observed by the Tibet air shower array”. In: *Astrophys. J.* 626 (2005), pp. 29–32. DOI: `10.1086/431582`. arXiv: `astro-ph/0505114 [astro-ph]`.

[48] A. Abdo, B. Allen, T. Aune, D. Berley, S. Casanova, et al. “The Large Scale Cosmic-Ray Anisotropy as Observed with Milagro”. In: *Astrophys. J.* 698 (2009), pp. 2121–2130. DOI: `10.1088/0004-637X/698/2/2121`. arXiv: `0806.2293 [astro-ph]`.

[49] B. Bartoli et al. “Medium scale anisotropy in the TeV cosmic ray flux observed by ARGO-YBJ”. In: *Phys. Rev.* D88.8 (2013), p. 082001. DOI: `10.1103/PhysRevD.88.082001`. arXiv: `1309.6182 [astro-ph.HE]`.

[50] A. U. Abeysekara et al. "Observation of Small-scale Anisotropy in the Arrival Direction Distribution of TeV Cosmic Rays with HAWC". In: *Astrophys. J.* 796.2 (2014), p. 108. DOI: `10.1088/0004-637X/796/2/108`. arXiv: `1408.4805 [astro-ph.HE]`.

[51] S. Westerhoff. "Anisotropy in Cosmic-Ray Arrival Directions Using IceCube and Ice-Top". In: *The IceCube Neutrino Observatory Contributions to ICRC 2015 Part III: Cosmic Rays*. 2015. arXiv: `1510.05225 [astro-ph.HE]`.

[52] J. C. for the HAWC Collaboration and he IceCube Collaboration. "Full-Sky Analysis of Cosmic-Ray Anisotropy with IceCube and HAWC". In: *Proceedings, 34th International Cosmic Ray Conference (ICRC 2015)*. 2015. arXiv: `1510.04134 [astro-ph.HE]`.

[53] O. Deligny for the Pierre Auger and Telescope Array Collaborations. "Large-Scale Distribution of Arrival Directions of Cosmic Rays Detected at the Pierre Auger Observatory and the Telescope Array above 10^{19} eV". In: *Pierre Auger Observatory and Telescope Array: Joint Contributions to the 34th International Cosmic Ray Conference (ICRC 2015)*. 2015. arXiv: `1511.02103 [astro-ph.HE]`.

[54] A. Aab et al. "Searches for Large-Scale Anisotropy in the Arrival Directions of Cosmic Rays Detected above Energy of 10^{19} eV at the Pierre Auger Observatory and the Telescope Array". In: *Astrophys. J.* 794.2 (2014), p. 172. DOI: `10.1088/0004-637X/794/2/172`. arXiv: `1409.3128 [astro-ph.HE]`.

[55] A. Aab et al. "Large Scale Distribution of Ultra High Energy Cosmic Rays Detected at the Pierre Auger Observatory With Zenith Angles up to 80°". In: *Astrophys. J.* 802.2 (2015), p. 111. DOI: `10.1088/0004-637X/802/2/111`. arXiv: `1411.6953 [astro-ph.HE]`.

[56] P. Abreu et al. "Search for First Harmonic Modulation in the Right Ascension Distribution of Cosmic Rays Detected at the Pierre Auger Observatory". In: *Astropart.Phys.* 34 (2011), pp. 627–639. DOI: `10.1016/j.astropartphys.2010.12.007`. arXiv: `1103.2721 [astro-ph.HE]`.

[57] P. Abreu et al. "Large scale distribution of arrival directions of cosmic rays detected above 10^{18} eV at the Pierre Auger Observatory". In: *Astrophys.J.Suppl.* 203 (2012), p. 34. DOI: `10.1088/0067-0049/203/2/34`. arXiv: `1210.3736 [astro-ph.HE]`.

[58] J. Abraham et al. "Correlation of the highest-energy cosmic rays with the positions of nearby active galactic nuclei". In: *Astropart.Phys.* 29 (2008), pp. 188–204. DOI: `10.1016/j.astropartphys.2008.06.004,10.1016/j.astropartphys.2008.01.002`. arXiv: `0712.2843 [astro-ph]`.

[59] J. Abraham et al. "Correlation of the highest energy cosmic rays with nearby extragalactic objects". In: *Science* 318 (2007), pp. 938–943. DOI: `10.1126/science.1151124`. arXiv: `0711.2256 [astro-ph]`.

[60] P. Abreu et al. "Update on the correlation of the highest energy cosmic rays with nearby extragalactic matter". In: *Astropart.Phys.* 34 (2010), pp. 314–326. DOI: `10.1016/j.astropartphys.2010.08.010`. arXiv: `1009.1855 [astro-ph.HE]`.

[61] A. Aab et al. “Searches for Anisotropies in the Arrival Directions of the Highest Energy Cosmic Rays Detected by the Pierre Auger Observatory”. In: *Astrophys. J.* 804.1 (2015), p. 15. DOI: 10.1088/0004-637X/804/1/15. arXiv: 1411.6111 [astro-ph.HE].

[62] K. Kawata, for the Telescope Array Collaboration. “Ultra-High-Energy Cosmic-Ray Hotspot Observed with the Telescope Array Surface Detectors”. In: *Telescope Array: Contributions to the 34th International Cosmic Ray Conference (ICRC 2015).* 2015.

[63] R. U. Abbasi et al. “Indications of Intermediate-Scale Anisotropy of Cosmic Rays with Energy Greater Than 57 EeV in the Northern Sky Measured with the Surface Detector of the Telescope Array Experiment”. In: *Astrophys. J.* 790 (2014), p. L21. DOI: 10.1088/2041-8205/790/2/L21. arXiv: 1404.5890 [astro-ph.HE].

[64] P. Abreu et al. “Search for signatures of magnetically-induced alignment in the arrival directions measured by the Pierre Auger Observatory”. In: *Astropart.Phys.* 35 (2012), pp. 354–361. DOI: 10.1016/j.astropartphys.2011.10.004. arXiv: 1111.2472 [astro-ph.HE].

[65] A. Aab et al. “Search for patterns by combining cosmic-ray energy and arrival directions at the Pierre Auger Observatory”. In: *Eur.Phys.J.C* (2014). arXiv: 1410.0515 [astro-ph.HE].

[66] P. Abreu et al. “A Search for Point Sources of EeV Neutrons”. In: *Astrophys.J.* 760 (2012), p. 148. DOI: 10.1088/0004-637X/760/2/148. arXiv: 1211.4901 [astro-ph.HE].

[67] A. Aab et al. “A Targeted Search for Point Sources of EeV Neutrons”. In: (2014). arXiv: 1406.4038 [astro-ph.HE].

[68] P. Abreu et al. “A search for anisotropy in the arrival directions of ultra high energy cosmic rays recorded at the Pierre Auger Observatory”. In: *JCAP* 1204 (2012), p. 040. DOI: 10.1088/1475-7516/2012/04/040. arXiv: 1210.3602 [astro-ph.HE].

[69] A. Aab et al. “A search for point sources of EeV photons”. In: (2014). arXiv: 1406.2912 [astro-ph.HE].

[70] P. Blasi. “Origin of very high- and ultra-high-energy cosmic rays”. In: *Comptes Rendus Physique* 15 (2014), pp. 329–338. DOI: 10.1016/j.crhy.2014.02.008. arXiv: 1403.2967 [astro-ph.HE].

[71] V. Berezinsky. “Ultrahigh-energy cosmic rays from cosmological relics”. In: *Nucl. Phys. B (Proc. Suppl.)* 87 (2000), pp. 387–396. DOI: 10.1016/S0920-5632(00)00701-5. arXiv: hep-ph/0001163 [hep-ph].

[72] J. Abraham et al. “Upper limit on the cosmic-ray photon fraction at EeV energies from the Pierre Auger Observatory”. In: *Astropart. Phys.* 31 (2009), pp. 399–406. DOI: 10.1016/j.astropartphys.2009.04.003. arXiv: 0903.1127 [astro-ph.HE].

[73] B. Peters. “Primary cosmic radiation and extensive air showers”. English. In: *Il Nuovo Cimento (1955-1965)* 22.4 (1961), pp. 800–819. DOI: 10.1007/BF02783106.

[74] K. Kotera and A. V. Olinto. “The Astrophysics of Ultrahigh-Energy Cosmic Rays”. In: *Annu. Rev. Astron. Astrophys.* 49 (Sept. 2011), pp. 119–153. DOI: `10.1146/annurev-astro-081710-102620`. arXiv: `1101.4256 [astro-ph.HE]`.

[75] K. Greisen. “End to the Cosmic-Ray Spectrum?” In: *Phys. Rev. Lett.* 16 (1966), pp. 748–750. DOI: `10.1103/PhysRevLett.16.748`.

[76] G. T. Zatsepin and V. A. Kuz’min. “Upper Limit of the Spectrum of Cosmic Rays”. In: *JETP Lett.* 4 (1966), p. 78. URL: `http://adsabs.harvard.edu/abs/1966JETPL...4...78Z`.

[77] G. Giacinti, M. Kachelrieß, D. V. Semikoz, and G. Sigl. “Cosmic ray anisotropy as signature for the transition from galactic to extragalactic cosmic rays”. In: *JCAP* 7, 031 (July 2012), p. 031. DOI: `10.1088/1475-7516/2012/07/031`. arXiv: `1112.5599 [astro-ph.HE]`.

[78] V. Berezinsky, A. Gazizov, and S. Grigorieva. “On astrophysical solution to ultrahigh-energy cosmic rays”. In: *Phys. Rev.* D74 (2006), p. 043005. DOI: `10.1103/PhysRevD.74.043005`. arXiv: `hep-ph/0204357 [hep-ph]`.

[79] R. Aloisio, V. Berezinsky, P. Blasi, A. Gazizov, S. Grigorieva, and B. Hnatyk. “A dip in the UHECR spectrum and the transition from galactic to extragalactic cosmic rays”. In: *Astropart. Phys.* 27 (2007), pp. 76–91. DOI: `10.1016/j.astropartphys.2006.09.004`. arXiv: `astro-ph/0608219 [astro-ph]`.

[80] J. Heinze, D. Boncioli, M. Bustamante, and W. Winter. “Cosmogenic Neutrinos Challenge the Cosmic Ray Proton Dip Model”. In: *Astrophys. J.* 825.2 (2016), p. 122. DOI: `10.3847/0004-637X/825/2/122`. arXiv: `1512.05988 [astro-ph.HE]`.

[81] R. Aloisio, V. Berezinsky, and P. Blasi. “Ultra high energy cosmic rays: implications of Auger data for source spectra and chemical composition”. In: *JCAP* 1410.10 (2014), p. 020. DOI: `10.1088/1475-7516/2014/10/020`. arXiv: `1312.7459 [astro-ph.HE]`.

[82] N. Globus, D. Allard, and E. Parizot. “A complete model of the cosmic ray spectrum and composition across the Galactic to extragalactic transition”. In: *Phys. Rev.* D92.2 (2015), p. 021302. DOI: `10.1103/PhysRevD.92.021302`. arXiv: `1505.01377 [astro-ph.HE]`.

[83] M. Unger, G. R. Farrar, and L. A. Anchordoqui. “Origin of the ankle in the ultrahigh energy cosmic ray spectrum, and of the extragalactic protons below it”. In: *Phys. Rev.* D92.12 (2015), p. 123001. DOI: `10.1103/PhysRevD.92.123001`. arXiv: `1505.02153 [astro-ph.HE]`.

[84] A. M. Taylor. “UHECR Composition Models”. In: *Astropart. Phys.* 54 (2014), pp. 48–53. DOI: `10.1016/j.astropartphys.2013.11.006`. arXiv: `1401.0199 [astro-ph.HE]`.

[85] N. Globus, D. Allard, R. Mochkovitch, and E. Parizot. “UHECR acceleration at GRB internal shocks”. In: *Mon. Not. Roy. Astron. Soc.* 451.1 (2015), pp. 751–790. DOI: `10.1093/mnras/stv893`. arXiv: `1409.1271 [astro-ph.HE]`.

[86] A. M. Taylor, M. Ahlers, and D. Hooper. "Indications of Negative Evolution for the Sources of the Highest Energy Cosmic Rays". In: *Phys. Rev.* D92.6 (2015), p. 063011. DOI: `10.1103/PhysRevD.92.063011`. arXiv: `1505.06090 [astro-ph.HE]`.

[87] A. di Matteo for the Pierre Auger Collaboration. "Combined fit of spectrum and composition data as measured by the Pierre Auger Observatory". In: *Pierre Auger Observatory: Contributions to the 34th International Cosmic Ray Conference (ICRC 2015).* 2015. arXiv: `1509.03732 [astro-ph.HE]`.

[88] P. Blasi, R. I. Epstein, and A. V. Olinto. "Ultrahigh-energy cosmic rays from young neutron star winds". In: *Astrophys. J.* 533 (2000), p. L123. DOI: `10.1086/312626`. arXiv: `astro-ph/9912240 [astro-ph]`.

[89] J. Arons. "Magnetars in the metagalaxy: an origin for ultrahigh-energy cosmic rays in the nearby universe". In: *Astrophys. J.* 589 (2003), pp. 871–892. DOI: `10.1086/374776`. arXiv: `astro-ph/0208444 [astro-ph]`.

[90] K. Fang, K. Kotera, K. Murase, and A. V. Olinto. "Testing the Newborn Pulsar Origin of Ultrahigh Energy Cosmic Rays with EeV Neutrinos". In: *Phys. Rev.* D90.10 (2014). [Phys. Rev.D90,103005(2014)], p. 103005. DOI: `10.1103/PhysRevD.90.103005,10.1103/PhysRevD.90.10300510.1103/PhysRevD.92.129901,10.1103/PhysRevD.92.129901`. arXiv: `1311.2044 [astro-ph.HE]`.

[91] K. Kotera, E. Amato, and P. Blasi. "The fate of ultrahigh energy nuclei in the immediate environment of young fast-rotating pulsars". In: *JCAP* 1508.08 (2015), p. 026. DOI: `10.1088/1475-7516/2015/08/026`. arXiv: `1503.07907 [astro-ph.HE]`.

[92] R. Protheroe. "Acceleration and Interaction of Ultra High Energy Cosmic RaysTopics in cosmic-ray astrophysics". In: ed. by M. A. DuVernois. Nova Science Publishing.

[93] F. M. Rieger, V. Bosch-Ramon, and P. Duffy. "Fermi acceleration in astrophysical jets". In: *Astrophys. Space Sci.* 309 (2007), pp. 119–125. DOI: `10.1007/s10509-007-9466-z`. arXiv: `astro-ph/0610141 [astro-ph]`.

[94] E. Fermi. "On the Origin of the Cosmic Radiation". In: *Phys. Rev.* 75 (1949), pp. 1169–1174. DOI: `10.1103/PhysRev.75.1169`.

[95] S. Horiuchi, K. Murase, K. Ioka, and P. Mészáros. "The Survival of Nuclei in Jets Associated with Core-collapse Supernovae and Gamma-Ray Bursts". In: *Astrophys. J.* 753, 69 (July 2012), p. 69. DOI: `10.1088/0004-637X/753/1/69`. arXiv: `1203.0296 [astro-ph.HE]`.

[96] L. A. Anchordoqui, G. E. Romero, and J. A. Combi. "Heavy nuclei at the end of the cosmic ray spectrum?" In: *Phys. Rev.* D60 (1999), p. 103001. DOI: `10.1103/PhysRevD.60.103001`. arXiv: `astro-ph/9903145 [astro-ph]`.

[97] S. Inoue, G. Sigl, F. Miniati, and E. Armengaud. "Ultrahigh energy cosmic rays as heavy nuclei from cluster accretion shocks". In: (2007). arXiv: `astro-ph/0701167 [astro-ph]`.

[98] A. Achterberg, Y. A. Gallant, J. G. Kirk, and A. W. Guthmann. “Particle acceleration by ultrarelativistic shocks: Theory and simulations”. In: *Mon. Not. Roy. Astron. Soc.* 328 (2001), p. 393. DOI: `10.1046/j.1365-8711.2001.04851.x`. arXiv: `astro-ph/0107530 [astro-ph]`.

[99] K. V. Ptitsyna and S. V. Troitsky. “Physical conditions in potential sources of ultra-high energy cosmic rays. I. Updated Hillas plot and radiation-loss constraints”. In: *Phys. Usp.* 53 (2010), pp. 691–701. DOI: `10.3367/UFNe.0180.201007c.0723`. arXiv: `0808.0367 [astro-ph]`.

[100] T. Winchen. “The Principal Axes of the Directional Energy Distribution of Cosmic Rays Measured with the Pierre Auger Observatory”. PhD thesis. RWTH, 2013. URL: `http://wwwo.physik.rwth-aachen.de/fileadmin/user_upload/www_physik/Personen/erdmann/Dissertation-Winchen.pdf`.

[101] A. M. Hillas. “The origin of ultra-high-energy cosmic rays”. In: *Annu. Rev. Astron. Astrophys.* 22 (1984), pp. 425–444.

[102] R. Protheroe and P. Biermann. “A New estimate of the extragalactic radio background and implications for ultrahigh-energy gamma-ray propagation”. In: *Astropart. Phys.* 6 (1996), pp. 45–54. DOI: `10.1016/S0927-6505(96)00041-2`. arXiv: `astro-ph/9605119 [astro-ph]`.

[103] A. Dominguez, J. Primack, D. Rosario, F. Prada, R. Gilmore, et al. “Extragalactic Background Light Inferred from AEGIS Galaxy SED-type Fractions”. In: *Mon.Not.Roy.Astron.Soc.* 410 (2011), p. 2556. DOI: `10.1111/j.1365-2966.2010.17631.x`. arXiv: `1007.1459 [astro-ph.CO]`.

[104] D. J. Fixsen et al. “ARCADE 2 Measurement of the Absolute Sky Brightness at 3-90 GHz”. In: *Astrophys. J.* 734.1 (2011), p. 5.

[105] M. Gervasi, A. Tartari, M. Zannoni, G. Boella, and G. Sironi. “The contribution of the Unresolved Extragalactic Radio Sources to the Brightness Temperature of the sky”. In: *Astrophys. J.* 682 (2008), p. 223. DOI: `10.1086/588628`. arXiv: `0803.4138 [astro-ph]`.

[106] T. M. Kneiske, T. Bretz, K. Mannheim, and D. Hartmann. “Implications of cosmological gamma-ray absorption. 2. Modification of gamma-ray spectra”. In: *Astron. Astrophys.* 413 (2004), pp. 807–815. DOI: `10.1051/0004-6361:20031542`. arXiv: `astro-ph/0309141 [astro-ph]`.

[107] J. Puget, F. Stecker, and J. Bredekamp. “Photonuclear Interactions of Ultrahigh-Energy Cosmic Rays and their Astrophysical Consequences”. In: *Astrophys. J.* 205 (1976), pp. 638–654. DOI: `10.1086/154321`.

[108] R. Gilmore, R. Somerville, J. Primack, and A. Dominguez. “Semi-analytic modeling of the EBL and consequences for extragalactic gamma-ray spectra”. In: *Mon.Not.Roy.Astron.Soc.* 422 (2012), p. 3189. DOI: `10.1111/j.1365-2966.2012.20841.x`. arXiv: `1104.0671 [astro-ph.CO]`.

[109] T. M. Kneiske, K. Mannheim, and D. H. Hartmann. “Implications of cosmological gamma-ray absorption - I.evolution of the metagalactic radiation field”. In: (2002). arXiv: `astro-ph/0202104 [astro-ph]`.

[110] T. M. Kneiske and H. Dole. “A Lower-Limit Flux for the Extragalactic Background Light”. In: (2010). arXiv: `1001.2132 [astro-ph.CO]`.

[111] J. D. Finke, S. Razzaque, and C. D. Dermer. “Modeling the Extragalactic Background Light from Stars and Dust”. In: *Astrophys. J.* 712 (2010), pp. 238–249. DOI: `10.1088/0004-637X/712/1/238`. arXiv: `0905.1115 [astro-ph.HE]`.

[112] F. W. Stecker, M. Malkan, and S. Scully. “Intergalactic photon spectra from the far ir to the uv lyman limit for 0 \$¡Z\$ $<$ 6 and the optical depth of the universe to high energy gamma-rays”. In: *Astrophys. J.* 648 (2006), pp. 774–783. DOI: `10.1086/506188`. arXiv: `astro-ph/0510449 [astro-ph]`.

[113] A. Franceschini, G. Rodighiero, and M. Vaccari. “The extragalactic optical-infrared background radiations, their time evolution and the cosmic photon-photon opacity”. In: *Astron. Astrophys.* 487 (2008), p. 837. DOI: `10.1051/0004-6361:200809691`. arXiv: `0805.1841 [astro-ph]`.

[114] F. W. Stecker, M. A. Malkan, and S. T. Scully. “A Determination of the Intergalactic Redshift Dependent UV-Optical-NIR Photon Density Using Deep Galaxy Survey Data and the Gamma-ray Opacity of the Universe”. In: *Astrophys. J.* 761 (2012), p. 128. DOI: `10.1088/0004-637X/761/2/128`. arXiv: `1205.5168 [astro-ph.HE]`.

[115] R. J. Protheroe and P. A. Johnson. “Propagation of Ultra High Energy Protons over Cosmological Distances and Implications for Topological Defect Models”. In: *Astropart. Phys.* 4 (1996), pp. 253–269. DOI: `10.1016/0927-6505(95)00039-9`.

[116] G. Breit and J. A. Wheeler. “Collision of Two Light Quanta”. In: *Phys. Rev.* 46 (1934), pp. 1087–1091. DOI: `10.1103/PhysRev.46.1087`.

[117] R. Gould and G. Schréder. “Opacity of the Universe to High-Energy Photons”. In: *Phys. Rev. Lett.* 16.6 (1966), pp. 252–254. DOI: `10.1103/PhysRevLett.16.252`.

[118] J. M. Dickey and F. J. Lockman. “HI in the galaxy”. In: *Ann. Rev. Astron. Astrophys.* 28 (1990), pp. 215–261. DOI: `10.1146/annurev.aa.28.090190.001243`.

[119] G. Giacinti, M. Kachelrieß, and D. V. Semikoz. “The escape model for Galactic cosmic rays”. In: *J. Phys. Conf. Ser.* 632.1 (2015), p. 012094. DOI: `10.1088/1742-6596/632/1/012094`.

[120] K. Kotera et al. “Propagation of Ultrahigh Energy Nuclei in Clusters of Galaxies: Resulting Composition and Secondary Emissions”. In: *Astrophys. J.* 707 (Dec. 2009), pp. 370–386. DOI: `10.1088/0004-637X/707/1/370`. arXiv: `0907.2433 [astro-ph.HE]`.

[121] S. Lee. “Propagation of extragalactic high energy cosmic and γ rays”. In: *Phys. Rev. D* 58 (4 July 1998), p. 043004. DOI: `10.1103/PhysRevD.58.043004`.

[122] D. Harari, S. Mollerach, E. Roulet, and F. Sanchez. "Lensing of ultra-high energy cosmic rays in turbulent magnetic fields". In: *Journal of High Energy Physics* 3 (2002), p. 45. DOI: `10.1088/1126-6708/2002/03/045`.

[123] A. Achterberg, Y. A. Gallant, C. A. Norman, and D. B. Melrose. "Intergalactic propagation of uhe cosmic rays". In: (1999). arXiv: `astro-ph/9907060 [astro-ph]`.

[124] R. Durrer and A. Neronov. "Cosmological Magnetic Fields: Their Generation, Evolution and Observation". In: *Astron. Astrophys. Rev.* 21 (2013), p. 62. DOI: `10.1007/s00159-013-0062-7`. arXiv: `1303.7121 [astro-ph.CO]`.

[125] R. A. Batista and G. Sigl. "Diffusion of cosmic rays at EeV energies in inhomogeneous extragalactic magnetic fields". In: (2014). arXiv: `1407.6150 [astro-ph.HE]`.

[126] D. Ryu, D. R. G. Schleicher, R. A. Treumann, C. G. Tsagas, and L. M. Widrow. "Magnetic Fields in the Large-Scale Structure of the Universe". In: *Space Science Review* 166 (May 2012), pp. 1–35. DOI: `10.1007/s11214-011-9839-z`. arXiv: `1109.4055`.

[127] G. Sigl, F. Miniati, and T. A. Ensslin. "Ultrahigh-energy cosmic rays in a structured and magnetized universe". In: *Phys. Rev.* D68 (2003), p. 043002. DOI: `10.1103/PhysRevD.68.043002`. arXiv: `astro-ph/0302388 [astro-ph]`.

[128] K. Dolag, D. Grasso, V. Springel, and I. Tkachev. "Mapping deflections of ultrahigh energy cosmic rays in constrained simulations of extragalactic magnetic fields". In: *JETP Lett.* 79 (2004), pp. 583–587. DOI: `10.1134/1.1790011`. arXiv: `astro-ph/0310902 [astro-ph]`.

[129] S. Das, H. Kang, D. Ryu, and J. Cho. "Propagation of Ultra-High-Energy Protons through the Magnetized Cosmic Web". In: *Astrophys. J.* 682 (2008), p. 29. DOI: `10.1086/588278`.

[130] K. Dolag, D. Grasso, V. Springel, and I. Tkachev. "Constrained simulations of the magnetic field in the local Universe and the propagation of UHECRs". In: *JCAP* 0501 (2005), p. 009. DOI: `10.1088/1475-7516/2005/01/009`. arXiv: `astro-ph/0410419 [astro-ph]`.

[131] R. Aloisio and V. Berezinsky. "Diffusive Propagation of Ultra-High-Energy Cosmic Rays and the Propagation Theorem". In: *Astrophys. J.* 612 (Sept. 2004), pp. 900–913. DOI: `10.1086/421869`. eprint: `arXiv:astro-ph/0403095`.

[132] S. Mollerach and E. Roulet. "Magnetic diffusion effects on the ultra-high energy cosmic ray spectrum and composition". In: *JCAP* 1310 (2013), p. 013. DOI: `10.1088/1475-7516/2013/10/013`. arXiv: `1305.6519 [astro-ph.HE]`.

[133] R. Jansson and G. R. Farrar. "The Galactic Magnetic Field". In: *Astrophys. J.* 761 (2012), p. L11. DOI: `10.1088/2041-8205/761/1/L11`. arXiv: `1210.7820 [astro-ph.GA]`.

[134] R. Beck and R. Wielebinski. "Magnetic fields in galaxies". In: *Planets* Populations (2013). DOI: `10.1007/978-94-007-5612-0_13`. arXiv: `1302.5663 [astro-ph.GA]`.

[135] P. Tinyakov and I. Tkachev. “Deflections of cosmic rays in a random component of the Galactic magnetic field”. In: *Astropart. Phys.* 24.1-2 (2005), pp. 32–43. ISSN: 0927-6505. DOI: 10.1016/j.astropartphys.2005.05.003.

[136] R. Jansson and G. R. Farrar. “A New Model of the Galactic Magnetic Field”. In: *Astrophys. J.* 757 (2012), p. 14. DOI: 10.1088/0004-637X/757/1/14. arXiv: 1204.3662 [astro-ph.GA].

[137] M. C. Beck, A. M. Beck, R. Beck, K. Dolag, A. W. Strong, and P. Nielaba. “New constraints on modelling the random magnetic field of the Milky Way”. In: (2014). arXiv: 1409.5120 [astro-ph.GA].

[138] D. Harari, S. Mollerach, and E. Roulet. “Signatures of galactic magnetic lensing upon ultra high energy cosmic rays”. In: *Journal of High Energy Physics* 2000.02 (2000), p. 035. DOI: 10.1088/1126-6708/2000/02/035.

[139] D. Harari, S. Mollerach, and E. Roulet. “Effects of the galactic magnetic field upon large scale anisotropies of extragalactic Cosmic Rays”. In: *JCAP* 1011 (2010), p. 033. DOI: 10.1088/1475-7516/2010/11/033. arXiv: 1009.5891 [astro-ph.HE].

[140] E. Kido and O. Kalashev. “Constraining UHECR source models by the TA SD energy spectrum”. In: (2013). arXiv: 1310.6093 [astro-ph.HE].

[141] R. Alves Batista, D. Boncioli, A. di Matteo, A. van Vliet, and D. Walz. “Effects of uncertainties in simulations of extragalactic UHECR propagation, using CRPropa and SimProp”. In: *JCAP* 1510.10 (2015), p. 063. DOI: 10.1088/1475-7516/2015/10/063. arXiv: 1508.01824 [astro-ph.HE].

[142] E. Armengaud, G. Sigl, T. Beau, and F. Miniati. “CRPropa: A numerical tool for the propagation of UHE cosmic rays, gamma-rays and neutrinos”. In: *Astropart. Phys.* 28 (2007), pp. 463–471. DOI: 10.1016/j.astropartphys.2007.09.004. arXiv: astro-ph/0603675 [astro-ph].

[143] K.-H. Kampert, J. Kulbartz, L. Maccione, N. Nierstenhoefer, P. Schiffer, et al. “CRPropa 2.0 – a Public Framework for Propagating High Energy Nuclei, Secondary Gamma Rays and Neutrinos”. In: *Astropart. Phys.* 42 (2013), pp. 41–51. DOI: 10.1016/j.astropartphys.2012.12.001. arXiv: 1206.3132 [astro-ph.IM].

[144] D. Allard, E. Parizot, E. Khan, S. Goriely, and A. V. Olinto. “UHE nuclei propagation and the interpretation of the ankle in the cosmic-ray spectrum”. In: *Astron. Astrophys.* 443 (2005), pp. 29–32. DOI: 10.1051/0004-6361:200500199. arXiv: astro-ph/0505566.

[145] N. Globus, D. Allard, and E. Parizot. “Propagation of high-energy cosmic rays in extragalactic turbulent magnetic fields: resulting energy spectrum and composition”. In: *Astron. Astrophys.* 479.1 (2008), pp. 97–110. DOI: 10.1051/0004-6361:20078653.

[146] H. Takami, S. Inoue, and T. Yamamoto. “Propagation of ultra-high-energy cosmic ray nuclei in cosmic magnetic fields and implications for anisotropy measurements”. In: *Astroparticle Physics* 35 (July 2012), pp. 767–780. DOI: 10.1016/j.astropartphys.2012.03.008. arXiv: 1202.2874 [astro-ph.HE].

[147] M. De Domenico. “HERMES: Simulating the propagation of ultra-high energy cosmic rays”. In: *Eur. Phys. J. Plus* 128 (2013), p. 99. DOI: `10.1140/epjp/i2013-13099-7`.

[148] R. Aloisio, D. Boncioli, A. Grillo, S. Petrera, and F. Salamida. “SimProp: a Simulation Code for Ultra High Energy Cosmic Ray Propagation”. In: *JCAP* 1210 (2012), p. 007. DOI: `10.1088/1475-7516/2012/10/007`. arXiv: `1204.2970 [astro-ph.HE]`.

[149] R. Aloisio, D. Boncioli, A. di Matteo, A. F. Grillo, S. Petrera, and F. Salamida. “SimProp v2r2: a Monte Carlo simulation to compute cosmogenic neutrino fluxes”. In: (2015). arXiv: `1505.01347 [astro-ph.HE]`.

[150] R. Aloisio, D. Boncioli, A. di Matteo, A. Grillo, S. Petrera, and F. Salamida. “SimProp v2r3”. PhD thesis. 2016. arXiv: `1602.01239 [astro-ph.HE]`.

[151] H. Yoshiguchi, S. Nagataki, and K. Sato. “A New Method for Calculating Arrival Distribution of Ultra-High-Energy Cosmic Rays above 10^{19} eV with Modifications by the Galactic Magnetic Field”. In: *Astrophys. J.* 596.2 (2003), p. 1044. DOI: `10.1086/378201`.

[152] H. Yoshiguchi, S. Nagataki, and K. Sato. “Numerical Study on the Propagation of Ultra-High-Energy Cosmic Rays in the Galactic Magnetic Field”. In: *Astrophys. J.* 607.2 (2004), p. 840. DOI: `10.1086/386277`.

[153] H. Takami, H. Yoshiguchi, and K. Sato. “Propagation of ultra-high energy cosmic rays above 10**19 eV in a structured extragalactic magnetic field and Galactic magnetic field”. In: *Astrophys. J.* 639 (2006). [Erratum: Astrophys. J.653,1584(2006)], pp. 803–815. DOI: `10.1086/499420`. arXiv: `astro-ph/0506203 [astro-ph]`.

[154] M. Sutherland, B. Baughman, and J. Beatty. “CRT: A numerical tool for propagating ultra-high energy cosmic rays through Galactic magnetic field models”. In: *Astropart. Phys.* 34 (2010), pp. 198–204. DOI: `10.1016/j.astropartphys.2010.07.002`. arXiv: `1010.3172 [astro-ph.IM]`.

[155] J. P. Rachen. “Interaction Processes and Statistical Properties of the Propagation of Cosmic Rays in Photon Backgrounds”. PhD thesis. Universität Bonn, 1996.

[156] C. T. Hill and D. N. Schramm. “The Ultrahigh-Energy Cosmic Ray Spectrum”. In: *Phys. Rev.* D31 (1985), p. 564. DOI: `10.1103/PhysRevD.31.564`.

[157] R. Aloisio, V. Berezinsky, and S. Grigorieva. “Analytic calculations of the spectra of ultra-high energy cosmic ray nuclei. I. The case of CMB radiation”. In: *Astropart. Phys.* 41 (2013), pp. 73–93. DOI: `10.1016/j.astropartphys.2012.07.010`. arXiv: `0802.4452 [astro-ph]`.

[158] O. E. Kalashev and E. Kido. “Simulations of Ultra High Energy Cosmic Rays propagation”. In: *J. Exp. Theor. Phys.* 120.5 (2015), pp. 790–797. DOI: `10.1134/S1063776115040056`. arXiv: `1406.0735 [astro-ph.HE]`.

[159] H.-P. Bretz, M. Erdmann, P. Schiffer, D. Walz, and T. Winchen. “PARSEC: A Parametrized Simulation Engine for Ultra-High Energy Cosmic Ray Protons”. In: *Astropart. Phys.* 54 (2014), pp. 110–117. DOI: `10.1016/j.astropartphys.2013.12.002`. arXiv: `1302.3761 [astro-ph.HE]`.

[160] V. S. Ptuskin, S. I. Rogovaya, and V. N. Zirakashvili. “Inverse problem for extragalactic transport of ultra-high energy cosmic rays”. In: *JCAP* 1503.03 (2015), p. 054. DOI: 10.1088/1475-7516/2015/03/054. arXiv: 1409.4654 [astro-ph.HE].

[161] R. A. Batista et al. “CRPropa 3—a public astrophysical simulation framework for propagating extraterrestrial ultra-high energy particles”. In: *Journal of Cosmology and Astroparticle Physics* 2016.05 (2016), p. 038. DOI: 10.1088/1475-7516/2016/05/038.

[162] L. Merten. “Monte Carlo Simulation des Transports der Galaktischen kosmischen Strahlung”. MA thesis. Ruhr-Universität Bochum, 2015.

[163] P. D. Group. *Monte Carlo Number Scheme Text and Tables*. 2014. URL: http://pdg.lbl.gov/2015/mcdata/mc_particle_id_contents.html.

[164] National Institute of Standards and Technology. *NIST Standard Reference Database 144*.

[165] A. Mucke, R. Engel, J. Rachen, R. Protheroe, and T. Stanev. “SOPHIA: Monte Carlo simulations of photohadronic processes in astrophysics”. In: *Comput.Phys.Commun.* 124 (2000), pp. 290–314. DOI: 10.1016/S0010-4655(99)00446-4. arXiv: astro-ph/9903478 [astro-ph].

[166] J. R. Cash and A. H. Karp. “A Variable Order Runge-Kutta Method for Initial Value Problems with Rapidly Varying Right-hand Sides”. In: *ACM Trans. Math. Softw.* 16.3 (Sept. 1990), pp. 201–222. ISSN: 0098-3500. DOI: 10.1145/79505.79507.

[167] M. Frigo and S. Johnson. “The Design and Implementation of FFTW3”. In: *Proc. IEEE* 93.2 (Feb. 2005), pp. 216–231. ISSN: 0018-9219. DOI: 10.1109/JPROC.2004.840301.

[168] R. Teyssier. “Cosmological hydrodynamics with adaptive mesh refinement: a new high resolution code called ramses”. In: *Astron. Astrophys.* 385 (2002), pp. 337–364. DOI: 10.1051/0004-6361:20011817. arXiv: astro-ph/0111367 [astro-ph].

[169] K. Dolag and F. Stasyszyn. “An MHD Gadget for cosmological simulations”. Submitted to Mon. Not. R. Astron. Soc. 2008.

[170] M. Pshirkov, P. Tinyakov, P. Kronberg, and K. Newton-McGee. “Deriving global structure of the Galactic Magnetic Field from Faraday Rotation Measures of extragalactic sources”. In: *Astrophys. J.* 738 (2011), p. 192. DOI: 10.1088/0004-637X/738/2/192. arXiv: 1103.0814 [astro-ph.GA].

[171] P. Bigalke. “A method for searching for Anisotropy in the arrival directions of Cosmic Nuclei traversing the Galactic Magnetic Field”. Bachelors thesis. RWTH Aachen, 2014. URL: http://wwwo.physik.rwth-aachen.de/fileadmin/user_upload/www_physik/Personen/erdmann/Bachelorthesis-Bigalke.pdf.

[172] D. Reinert. “Autocorrelations of Ultra-High Energy Cosmic Rays with Galactic Magnetic Field Corrections at the Pierre Auger Observatory”. MA thesis. RWTH Aachen, 2015. URL: http://wwwo.physik.rwth-aachen.de/fileadmin/user_upload/www_physik/Personen/erdmann/Masterthesis_Reinert.pdf.

[173] M. Erdmann, G. Müller, M. Urban, and M. Wirtz. "The Nuclear Window to the Extragalactic Universe". In: *(submitted)* (2016).

[174] M. Lauscher. "Search for the sources of ultra-high energy cosmic rays with the Pierre Auger Observatory". PhD thesis. RWTH Aachen, 2016.

[175] E. V. Bugaev, A. Misaki, and K. Mitsui. "Neutrinos from extragalactic cosmic ray interactions in the far infrared background". In: *Astropart. Phys.* 24 (2005), pp. 345–354. DOI: `10.1016/j.astropartphys.2005.08.001`. arXiv: `astro-ph/0405109 [astro-ph]`.

[176] R. A. Batista and G. Sigl. *Cosmological Effects in Three-Dimensional Simulations of Propagation of UHE Protons.* GAP Note. 2013.

[177] H. Bethe and W. Heitler. "On the Stopping of fast particles and on the creation of positive electrons". In: *Proc. Roy. Soc. Lond.* A146 (1934), pp. 83–112. DOI: `10.1098/rspa.1934.0140`.

[178] G. R. Blumenthal. "Energy Loss of High-Energy Cosmic Rays in Pair-Producing Collisions with Ambient Photons". In: *Phys. Rev. D* 1 (6 Mar. 1970), pp. 1596–1602. DOI: `10.1103/PhysRevD.1.1596`.

[179] M. J. Chodorowski, A. A. Zdziarski, and M. Sikora. "Reaction rate and energy-loss rate for photopair production by relativistic nuclei". In: *Astrophys. J.* 400 (Nov. 1992), pp. 181–185. DOI: `10.1086/171984`.

[180] S. R. Kelner and F. A. Aharonian. "Energy spectra of gamma rays, electrons, and neutrinos produced at interactions of relativistic protons with low energy radiation". In: *Phys. Rev. D* 78 (3 Aug. 2008), p. 034013. DOI: `10.1103/PhysRevD.78.034013`.

[181] M. Chadwick et al. *Handbook on photonuclear data for applications: cross-sections and spectra.* Tech. rep. 1178. International Atomic Energy Agency, 2000. URL: `https://www-nds.iaea.org/publications/tecdocs/iaea-tecdoc-1178/`.

[182] A. J. Koning, S. Hilaire, and M. C. Duijvestijn. "TALYS: Comprehensive Nuclear Reaction Modeling". In: *International Conference on Nuclear Data for Science and Technology.* Ed. by R. C. Haight, M. B. Chadwick, T. Kawano, and P. Talou. Vol. 769. American Institute of Physics Conference Series. May 2005, pp. 1154–1159. DOI: `10.1063/1.1945212`.

[183] A. Koning, S. Hilaire, and S. Goriely. *TALYS 1.6 User Manual.*

[184] E. Khan et al. "Photodisintegration of ultra-high-energy cosmic rays revisited". In: *Astropart. Phys.* 23 (2005), pp. 191–201. DOI: `10.1016/j.astropartphys.2004.12.007`. arXiv: `astro-ph/0412109 [astro-ph]`.

[185] S. Goriely. *private communication.*

[186] N. Nierstenhoefer. "On the Origin and Propagation of Ultra-High Energy Cosmic Rays". PhD thesis. Bergische Universität Wuppertal, 2011.

[187] M. Kossov. "Approximation of photonuclear interaction cross-sections". In: *EPJ A* 14.3 (2002), pp. 377–392. ISSN: 1434-6001. DOI: `10.1140/epja/i2002-10008-x`.

[188] S. Agostinelli et al. "GEANT4: A Simulation toolkit". In: *Nucl. Instrum. Meth.* A506 (2003), pp. 250–303. DOI: `10.1016/S0168-9002(03)01368-8`.

[189] F. Stecker and M. Salamon. "Photodisintegration of ultrahigh-energy cosmic rays: A new determination". In: *Astrophys. J.* 512 (1999), pp. 521–526. DOI: `10.1086/306816`. arXiv: `astro-ph/9808110 [astro-ph]`.

[190] J. S. Levinger. *Nuclear Photo-Disintegration.* Oxford University Press, 1960.

[191] L. A. Anchordoqui, J. F. Beacom, H. Goldberg, S. Palomares-Ruiz, and T. J. Weiler. "TeV gamma-rays from photo-disintegration/de-excitation of cosmic-ray nuclei". In: *Phys. Rev. Lett.* 98 (2007), p. 121101. DOI: `10.1103/PhysRevLett.98.121101`. arXiv: `astro-ph/0611580 [astro-ph]`.

[192] N. Bianchi et al. "Absolute total photoabsorption cross-sections on nuclei in the nucleon resonance region". In: *Phys. Lett. B* B325 (1994), pp. 333–336. DOI: `10.1016/0370-2693(94)90021-3`.

[193] B. N. L. National Nuclear Data Center. *Nuclear Structure and Decay Data.* 2012. URL: `http://www.nndc.bnl.gov/nudat2/`.

[194] J.-L. Basdevant, J. Rich, and M. Spiro. *Fundamentals In Nuclear Physics: From Nuclear Structure to Cosmology.* Springer, 2005.

[195] R. Ruffini, G. V. Vereshchagin, and S. .-.-S. Xue. "Cosmic absorption of ultra high energy particles". In: (2015). arXiv: `1503.07749 [astro-ph.HE]`.

[196] M. Settimo and M. De Domenico. "Propagation of extragalactic photons at ultra-high energy with the *EleCa* code". In: *Astropart. Phys.* 62 (2015), pp. 92–99. DOI: `10.1016/j.astropartphys.2014.07.011`. arXiv: `1311.6140 [astro-ph.HE]`.

[197] C. Heiter. "in progress". MA thesis. RWTH Aachen, 2016.

[198] A. Aab et al. "The Pierre Auger Cosmic Ray Observatory". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 798 (2015), pp. 172–213. ISSN: 0168-9002. DOI: `http://dx.doi.org/10.1016/j.nima.2015.06.058`. arXiv: `1502.01323`.

[199] D. Veberic. *Layout of Pierre Auger Observatory.* 2013. URL: `https://commons.wikimedia.org/wiki/File:Layout_of_Pierre_Auger_Observatory.svg`.

[200] The Pierre Auger Collaboration. "The Fluorescence Detector of the Pierre Auger Observatory". In: *Nucl. Instrum. Meth. A* A620 (2010), pp. 227–251. DOI: `10.1016/j.nima.2010.04.023`.

[201] M. J. Tueros for the Pierre Auger Collaboration. "Estimate of the non-calorimetric energy of showers observed with the fluorescence and surface detectors of the Pierre Auger Observatory". In: *Pierre Auger Observatory: Contributions to the 33rd International Cosmic Ray Conference (ICRC 2013).* 2013. arXiv: `1307.5059 [astro-ph.HE]`.

[202] B. R. Dawson for the Pierre Auger Collaboration. "Hybrid Performance of the Pierre Auger Observatory". In: *The Pierre Auger Observatory: Contributions to the 30th International Cosmic Ray Conference (ICRC 2007).* 2007. arXiv: `0706.1105 [astro-ph]`.

[203] V. Verzi for the Pierre Auger Collaboration. “The Energy Scale of the Pierre Auger Observatory”. In: *Pierre Auger Observatory: Contributions to the 33rd International Cosmic Ray Conference (ICRC 2013)*. 2013. arXiv: `1307.5059 [astro-ph.HE]`.

[204] The Pierre Auger Collaboration. “The Surface Detector System of the Pierre Auger Observatory”. In: *Nucl. Instrum. Meth. A* 586 (2008), pp. 409–420. DOI: `10.1016/j.nima.2007.12.016`.

[205] K. Weidenhaupt. “Antenna Calibration and Energy Measurement of Ultra-High Energy Cosmic Rays with the Auger Engineering Radio Array”. PhD thesis. RWTH Aachen, 2014. URL: `https://publications.rwth-aachen.de/record/465387/files/5264.pdf`.

[206] C. Bonifazi for the Pierre Auger Collaboration. “The angular resolution of the Pierre Auger Observatory”. In: *Nucl. Phys. Proc. Suppl.* 190 (2009), pp. 20–25. DOI: `10.1016/j.nuclphysbps.2009.03.063`. arXiv: `0901.3138 [astro-ph.HE]`.

[207] J. Hersil, I. Escobar, D. Scott, G. Clark, and S. Olbert. “Observations of Extensive Air Showers near the Maximum of Their Longitudinal Development”. In: *Phys. Rev. Lett.* 6 (1961), pp. 22–23. DOI: `10.1103/PhysRevLett.6.22`.

[208] A. Aab et al. “Reconstruction of inclined air showers detected with the Pierre Auger Observatory”. In: *JCAP* 1408.08 (2014), p. 019. DOI: `10.1088/1475-7516/2014/08/019`. arXiv: `1407.3214 [astro-ph.HE]`.

[209] A. E. S. Task. *private communication.*

[210] A. Aab et al. “Depth of Maximum of Air-Shower Profiles at the Pierre Auger Observatory: Measurements at Energies above $10^{17.8}$ eV”. In: *Phys.Rev.D* (2014). arXiv: `1409.4809 [astro-ph.HE]`.

[211] M. Ahlers, L. A. Anchordoqui, and A. M. Taylor. “Ensemble fluctuations of the flux and nuclear composition of ultrahigh energy cosmic ray nuclei”. In: *Phys. Rev.* D87.2 (2013), p. 023004. DOI: `10.1103/PhysRevD.87.023004`. arXiv: `1209.5427 [astro-ph.HE]`.

[212] A. Supanitsky and G. Medina-Tanco. “Ensemble fluctuations of the cosmic ray energy spectrum and the intergalactic magnetic field”. In: *Phys. Rev.* D91.12 (2015), p. 123006. DOI: `10.1103/PhysRevD.91.123006`. arXiv: `1412.7475 [astro-ph.HE]`.

[213] M. De Domenico, M. Settimo, S. Riggi, and E. Bertin. “Reinterpreting the development of extensive air showers initiated by nuclei and photons”. In: *JCAP* 1307 (2013), p. 050. DOI: `10.1088/1475-7516/2013/07/050`. arXiv: `1305.2331 [hep-ph]`.

[214] M. Urban. “Investigation of the Mass Composition of Ultra High Energy Cosmic Rays with the Pierre Auger Observatory”. MA thesis. RWTH Aachen University, 2013. URL: `http://wwwo.physik.rwth-aachen.de/fileadmin/user_upload/www_physik/Personen/erdmann/Masterthesis-Urban.pdf`.

[215] G. T. Elipe and R. Vazquez. *The Analysis of Xmax composition using Bayesian methods.* GAP Note. 2013.

[216] J. Steggemann. “Search for New Particles Decaying to a Top Quark Pair with the CMS Experiment”. PhD thesis. RWTH Aachen, 2012. URL: http://wwwo.physik.rwth-aachen.de/fileadmin/user_upload/www_physik/Personen/erdmann/PhDthesis-Steggemann.pdf.

[217] D. Klingebiel. “Measurement of the t-Channel Single-Top-Quark-Production Cross Section and the CKM-Matrix Element Vtb with the CMS Experiment”. PhD thesis. RWTH Aachen, 2014. URL: http://wwwo.physik.rwth-aachen.de/fileadmin/user_upload/www_physik/Personen/erdmann/PhDthesis-Klingebiel.pdf.

[218] J. K. Lindsey. *Parametric Statistical Inference*. New York: Oxford University Press, 1996.

[219] S. Onn and I. Weissman. “Generating uniform random vectors over a simplex with implications to the volume of a certain polytope and to multivariate extremes”. English. In: *Annals of Operations Research* 189.1 (2011), pp. 331–342. ISSN: 0254-5330. DOI: 10.1007/s10479-009-0567-7.

[220] A. Patil, D. Huard, and C. J. Fonnesbeck. “PyMC: Bayesian Stochastic Modelling in Python”. In: *Journal of Statistical Software* 35.4 (July 2010), pp. 1–81. ISSN: 1548-7660.

[221] A. Gelman and D. B. Rubin. “Inference from Iterative Simulation Using Multiple Sequences”. In: *Statistical Science* 7.4 (1992), pp. 457–472. DOI: 10.1214/ss/1177011136.

[222] R. E. Kass and A. E. Raftery. “Bayes Factors”. In: *Journal of the American Statistical Association* 90.430 (1995), pp. 773–795. DOI: 10.1080/01621459.1995.10476572. eprint: http://amstat.tandfonline.com/doi/pdf/10.1080/01621459.1995.10476572.

[223] F. Feroz, M. P. Hobson, E. Cameron, and A. N. Pettitt. “Importance Nested Sampling and the MultiNest Algorithm”. In: (2013). arXiv: 1306.2144 [astro-ph.IM].

[224] D. J. Spiegelhalter, N. G. Best, B. P. Carlin, and A. Van Der Linde. “Bayesian measures of model complexity and fit”. In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 64.4 (2002), pp. 583–639. ISSN: 1467-9868. DOI: 10.1111/1467-9868.00353.

[225] T. Ando. “Predictive Bayesian Model Selection”. In: *American Journal of Mathematical and Management Sciences* 31.1-2 (2011), pp. 13–38. DOI: 10.1080/01966324.2011.10737798. eprint: http://dx.doi.org/10.1080/01966324.2011.10737798.

[226] B. Sarkar. “Constraints to UHE cosmic ray source scenarios with all the major observables of the Pierre Auger Observatory”. PhD thesis. Bergische Universität Wuppertal, 2015.

[227] A. di Matteo. “Ultra-high-energy cosmic ray phenomenology: Monte Carlo simulations and experimental data”. PhD thesis. University of L’Aquila, 2015.

[228] G. B. Gelmini, O. Kalashev, and D. V. Semikoz. “Gamma-Ray Constraints on Maximum Cosmogenic Neutrino Fluxes and UHECR Source Evolution Models”. In: *JCAP* 1201 (2012), p. 044. DOI: `10.1088/1475-7516/2012/01/044`. arXiv: `1107.1672 [astro-ph.CO]`.

[229] D. Harari, S. Mollerach, and E. Roulet. “The shape of the extragalactic cosmic ray spectrum from Galaxy Clusters”. In: (2016). arXiv: `1605.08113 [astro-ph.HE]`.

[230] P. Abreu et al. “Bounds on the density of sources of ultra-high energy cosmic rays from the Pierre Auger Observatory”. In: *JCAP* 1305 (2013), p. 009. DOI: `10.1088/1475-7516/2013/05/009`. arXiv: `1305.1576 [astro-ph.HE]`.

[231] A. Aab et al. “Energy Estimation of Cosmic Rays with the Engineering Radio Array of the Pierre Auger Observatory”. In: *Submitted to: Phys. Rev. D* (2015). arXiv: `1508.04267 [astro-ph.HE]`.

[232] D. Allard et al. “Cosmogenic Neutrinos from the propagation of Ultra High Energy Nuclei”. In: *JCAP* 609 (2006), p. 5. DOI: `10.1088/1475-7516/2006/09/005`.

[233] M. Ahlers and F. Halzen. “Minimal Cosmogenic Neutrinos”. In: *Phys. Rev.* D86 (2012), p. 083010. DOI: `10.1103/PhysRevD.86.083010`. arXiv: `1208.4181 [astro-ph.HE]`.

[234] V. Berezinsky, A. Gazizov, M. Kachelriess, and S. Ostapchenko. “Restricting UHECRs and cosmogenic neutrinos with Fermi-LAT”. In: *Phys. Lett. B* B695 (2011), pp. 13–18. DOI: `10.1016/j.physletb.2010.11.019`. arXiv: `1003.1496 [astro-ph.HE]`.

[235] D. Hooper, A. M. Taylor, and S. Sarkar. “Cosmogenic photons as a test of ultra-high energy cosmic ray composition”. In: *Astroparticle Physics* 34 (Jan. 2011), pp. 340–343. DOI: `10.1016/j.astropartphys.2010.09.002`. arXiv: `1007.1306 [astro-ph.HE]`.

[236] R.-Y. Liu, A. M. Taylor, X.-Y. Wang, and F. A. Aharonian. “Indication of a local fog of subankle ultrahigh energy cosmic rays”. In: *Phys. Rev.* D94.4 (2016), p. 043008. DOI: `10.1103/PhysRevD.94.043008`. arXiv: `1603.03223 [astro-ph.HE]`.

[237] M. Ackermann et al. “The spectrum of isotropic diffuse gamma-ray emission between 100 MeV and 820 GeV”. In: *Astrophys. J.* 799.1 (2015), p. 86. DOI: `10.1088/0004-637X/799/1/86`. arXiv: `1410.3696 [astro-ph.HE]`.

[238] C. Bleve for the Pierre Auger Collaboration. “Updates on the neutrino and photon limits from the Pierre Auger Observatory”. In: *Pierre Auger Observatory: Contributions to the 34th International Cosmic Ray Conference (ICRC 2015)*. 2015. arXiv: `1509.03732 [astro-ph.HE]`.

[239] M. Aartsen et al. “Observation of High-Energy Astrophysical Neutrinos in Three Years of IceCube Data”. In: *Phys. Rev. Lett.* 113 (2014), p. 101101. DOI: `10.1103/PhysRevLett.113.101101`. arXiv: `1405.5303 [astro-ph.HE]`.

[240] M. Aartsen et al. “Probing the origin of cosmic rays with extremely high energy neutrinos using the IceCube Observatory”. In: *Phys. Rev.* D88 (2013), p. 112008. DOI: `10.1103/PhysRevD.88.112008`. arXiv: `1310.5477 [astro-ph.HE]`.

[241] M. Ackermann et al. “Resolving the Extragalactic γ-Ray Background above 50 GeV with the Fermi Large Area Telescope”. In: *Phys. Rev. Lett.* 116.15 (2016), p. 151105. DOI: `10.1103/PhysRevLett.116.151105`. arXiv: `1511.00693 [astro-ph.CO]`.

[242] T. Belgya et al. *Handbook for calculations of nuclear reaction data, RIPL-2*. Tech. rep. 1506. IAEA, 2006. URL: `https://www-nds.iaea.org/publications/tecdocs/iaea-tecdoc-1506/`.

[243] *The Pierre Auger Observatory: Contributions to the 34th International Cosmic Ray Conference (ICRC 2015)*. 2015. arXiv: `1509.03732 [astro-ph.HE]`.

[244] *The Pierre Auger Observatory: Contributions to the 33rd International Cosmic Ray Conference (ICRC 2013)*. 2013. arXiv: `1307.5059 [astro-ph.HE]`.

Declaration of Pre-released Extracts

The following publications and internal notes of the Pierre Auger Collaboration contain pre-released extracts of this work. The corresponding sections of the publications have been written by the author of this thesis and were revised and copy-edited by the co-authors of the publications.

- M. Erdmann *et al.*, A Benchmark Scenario for UHECR Propagation, Internal Note of the Pierre Auger Collaboration, GAP-2012-138 (2012)
  The internal note of the Pierre Auger Observatory contains parts of chapter 4.

- R. Alves Batista *et al.*, CRPropa 3.0 – a Public Framework for Propagating UHE Cosmic Rays through Galactic and Extragalactic Space, Proceedings, 33rd International Cosmic Ray Conference (ICRC 2013)
  This publication contains parts of chapter 4.

- D. Walz and M. Erdmann, Constraining UHECR source scenarios with spectrum and X_{max} distributions, Internal Note of the Pierre Auger Collaboration, GAP-2014-034 (2014)
  The internal note of the Pierre Auger Observatory contains parts of chapters 6 and 7.

- R. Alves Batista *et al.*, Cosmic ray propagation with CRPropa 3, Proceedings, 16th International workshop on Advanced Computing and Analysis Techniques in physics (ACAT 2014), JPCS 608 (2015) 012076
  This publication contains parts of chapter 4.

- R. Alves Batista *et al.*, Effects of uncertainties in simulations of extragalactic UHECR propagation, using CRPropa and SimProp, JCAP (2015) 063 This publication contains parts of appendix A.1 and A.2.

- R. Alves Batista *et al.*, CRPropa 3 – a public astrophysical simulation framework for propagating extraterrestrial ultra-high energy particles, JCAP (2016) 038
  This publication contains parts of chapter 4 and of appendix A.1 and A.2.

The following publication based on work presented in this thesis is in preparation.

- Aab *et al.* (The Pierre Auger Collaboration) Combined fit of spectrum and composition data as measured by the Pierre Auger Observatory, in preparation (2016) The publication will contain parts of 6 and 7.

Acknowledgments

Firstly, I would like to express my sincere gratitude to my advisor Prof. Dr. Martin Erdmann for the continuous support of my Ph.D study and related research, for his constant motivation, and for his unique way to put things into a new perspectives. I would also like to thank Prof. Dr. Günter Sigl for co-refereeing my thesis, as well as Prof. Thomas Hebbeker and Prof. Achim Stahl for forming my thesis committee.

I thank Martin Urban, Raphael Krause, Daniel Kümpel, Olga Walz and Martin Erdmann for critical comments to the manuscript of this thesis.

My analysis would not have been possible without the enthusiastic work of the members and technical staff of the Pierre Auger Collaboration. I particularly appreciate the valuable discussions on UHECR propagation and measurements with my colleagues in L'Aquila and Hamburg.

My fellow colleagues in the Auger and CMS groups at my institute I would like to thank for countless stimulating discussions on experiments, statistics, computing and life in general.

Last but not the least, I would like to thank my family: my parents and my sisters for supporting me throughout my life. Most of all I would like to thank my wife Olga for her love, support and sacrifices for keeping my back free whenever I needed it.