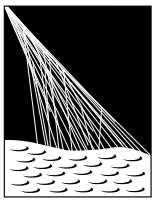

Measurement of the Mass Composition of Cosmic Rays Using Cherenkov-Light Dominated Air Showers Detected at the Pierre Auger Observatory

Master's Thesis of

Marc L. Schumann

at the Department of Physics
Institute for Astroparticle Physics (IAP)

Reviewer:	Prof. Dr. Ralph Engel
Second reviewer:	Prof. Dr. Guido Drexlin
Advisor:	Dr. Michael Unger

10th March 2024 – 10th March 2025

Karlsruher Institut für Technologie
Fakultät für Physik
76128 Karlsruhe

Measurement of the Mass Composition of Cosmic Rays Using Cherenkov-Light Dominated Air Showers Detected at the Pierre Auger Observatory (Master's Thesis)

I declare that I have developed and written the enclosed thesis completely by myself. I have not used any other than the aids that I have mentioned. I have marked all parts of the thesis that I have included from referenced literature, either in their original wording or paraphrasing their contents. I have followed the by-laws to implement scientific integrity at KIT.

Karlsruhe, 10th March 2025

. . . . . .

(Marc L. Schumann)

Abstract

Cosmic rays are charged particles that can be accelerated by Galactic and extragalactic sources to energy levels far beyond what is possible with accelerators on Earth. High-energy cosmic rays are measured indirectly by ground-based detectors like the Pierre Auger Observatory through extensive air showers. Fluorescence light and Cherenkov light are produced by air showers as they traverse the atmosphere. This light can be detected with sensitive telescopes in moonless nights to infer the development of the air showers and to determine the energy of the primary particles. The energy range between 10^{16} eV and 10^{17} eV is generally linked to the transition from Galactic to extragalactic cosmic rays. In this range, measurements are only feasible by detecting the intense Cherenkov light.

The mass of individual cosmic rays cannot be measured directly from air showers. Instead, the mass composition can be inferred from average differences in the depth of shower maximum for different cosmic ray masses in comparison to detailed Monte Carlo simulations. Such an analysis is presented in this work for Cherenkov-light dominated events of the Pierre Auger Observatory. In preparation for the analysis, the reconstruction procedure of Cherenkov-dominated events is assessed for possible improvements. The agreement between Monte Carlo simulations and corresponding data is confirmed. A mass composition analysis for calorimetric energies between 10^{16} eV and 10^{17} eV, including an examination of systematic effects and the stability of the result, is performed.

This work demonstrates that the absolute value of the mean mass measured with Cherenkov-light dominated showers still has large systematic uncertainties ($\approx$ 20 % in units of proton-iron difference). However, this analysis found evidence for a feature in the evolution of the mass composition at an energy around 4×10^{16} eV, close to the low-energy ankle in the cosmic ray flux. Below this energy, the average mass of cosmic rays increases steadily, whereas above it the mass decreases, which could signify the transition to a different component in the cosmic ray flux.

Zusammenfassung

Kosmische Strahlen sind geladene Teilchen, die durch galaktische und extragalaktische Quellen auf Energieniveaus beschleunigt werden können, die außerhalb der Möglichkeiten von Beschleunigern auf der Erde liegen. Hochenergetische kosmische Strahlen werden auf indirekte Art und Weise durch Luftschauer von Detektoren auf der Erdoberfläche wie dem Pierre-Auger-Observatorium gemessen. Von Luftschauern, die die Atmosphäre durchqueren, wird Fluoreszenzlicht und Tscherenkov-Licht emittiert. Dieses Licht kann in mondlosen Nächten durch sensitive Teleskope gemessen werden, um auf die Entwicklung der Luftschauer zu schließen und die Energie der Primärteilchen zu bestimmen. Der Energiebereich zwischen 10^{16} eV und 10^{17} eV wird generell mit dem Übergang von galaktischen zu extragalaktischen kosmischen Strahlen in Verbindung gebracht. In diesem Bereich sind Messungen nur durchführbar, indem das intensive Tscherenkov-Licht detektiert wird.

Die Masse individueller kosmischer Strahlen kann nicht direkt durch Luftschauer gemessen werden. Stattdessen wird die Massenkomposition von mittleren Unterschieden in der Tiefe des Schauermaximums für verschiedene Massen kosmischer Strahlen durch einen Vergleich mit detaillierten Monte-Carlo-Simulationen abgeleitet. Eine solche Analyse wird in dieser Arbeit mit Tscherenkov-Licht-dominierten Schauern des Pierre-Auger-Observatoriums vorgestellt. Zur Vorbereitung der Analyse wird die Rekonstruktionsprozedur Tscherenkov-dominierter Ereignisse auf mögliche Verbesserungen hin überprüft. Die Übereinstimmung zwischen Monte-Carlo-Simulationen und entsprechenden Daten wird bestätigt. Eine Massenkompositionsanalyse für kalorimetrische Energien zwischen 10^{16} eV und 10^{17} eV wird durchgeführt, einschließlich einer Betrachtung systematischer Effekte sowie der Stabilität der Resultate.

Diese Arbeit zeigt, dass die absoluten Werte der mittleren Masse Tscherenkov-Licht-dominierter Luftschauer über große systematische Unsicherheiten verfügen ($\approx 20\,\%$ in Einheiten des Proton-Eisen-Unterschieds). Jedoch werden Hinweise auf ein Merkmal der Entwicklung der Massenkomposition bei Energien um 4×10^{16} eV gefunden, nahe des niederenergetischen Knöchels des Flusses kosmischer Strahlen. Unterhalb dieser Energie erhöht sich die mittlere Masse der kosmischen Strahlen stetig, während sich die Masse oberhalb verringert. Dies könnte einen Übergang zu einer separaten Komponente des Flusses kosmischer Strahlen bedeuten.

Contents

1. Introduction

Since the discovery of Cosmic Rays (CRs) in 1912 by Victor Hess during balloon flights [1], many advances have been made in understanding the nature of these particles through measurements conducted by large-scale detectors across the globe. However, many open questions about CRs still remain, related for example to their Galactic and extragalactic origins, as well as processes of acceleration and propagation [2]. Obtaining knowledge about the mass composition, and subsequently the charge composition of CRs across the measured energy range, plays a crucial role in furthering the understanding about all of these aspects. This applies especially to the energy range related to the transition between Galactic and extragalactic CRs, which is examined in this work.

To gather information about high-energy CRs, ground-based detectors such as the Pierre Auger Observatory use the atmosphere as a calorimeter. When a CR arrives at Earth's atmosphere, a cascade of secondary particles is initiated, known as an Extensive Air Shower (EAS). EASs can cause light emission through fluorescence of excited nitrogen molecules or through Cherenkov light from shower particles which move faster than the speed of light in the air. Measurements of this emission along the longitudinal development of an EAS can be used to extrapolate information about the primary CR particle, such as its arrival direction and its calorimetric energy deposited in the atmosphere. One of the characteristic air shower parameters is the depth into the atmosphere at which the shower reaches its maximum size. The average depth of shower maximum is mass dependent, with showers induced by heavier CR particles reaching their maximum at higher altitudes for a given energy. This mass dependency of the depth of shower maximum can be quantified with detailed Monte Carlo (MC) simulations for EASs with underlying hadronic interaction models. A subsequent comparison of the simulations with measured data can provide insights into the mass composition of the measured CR flux.

Addressing this, the goal of this work is to perform a mass composition analysis with measurements by the Pierre Auger Observatory in the energy range related to the end of the Galactic cosmic ray spectrum, between calorimetric energies of 10^{16} eV and 10^{17} eV. In this energy range, EAS measurements are dominated by contributions from Cherenkov light. The reconstruction of such air showers at Auger is investigated for possible improvements. A focus is placed on the reconstruction of the depth of shower maximum in comparison to generated values in simulations. The alignment of simulations and data on key shower parameters is investigated and a mass composition analysis is performed based on the reconstruction of the depth of shower maximum.

The structure of this work is as follows: in Chapter 2, an overview of current understandings about the CR spectrum and mass composition is provided, along with the light emission and a simplified model of EASs. Chapter 3 introduces the Pierre Auger Observatory with its main detectors and the reconstruction procedure of EASs with a dominant contribution of Cherenkov light. Subsequently, Chapter 4 provides a detailed description

of the Monte-Carlo simulations used in this work and an analysis of different methods for the reconstruction of Cherenkov-dominated events. Following this investigation, a mass composition analysis takes place in Chapter 5, which includes comparisons between simulations and data, the measurement of the evolution of the mean mass between 10^{16} eV and 10^{17} eV in calorimetric energy, as well as stability checks of the results.

2. Cosmic Rays and Extensive Air Showers

CRs are the most energetic particles known in the universe. In this chapter, an introduction on measurements of the CR energy spectrum and models for its description is given. The Heitler model for the development of EASs is described, as well as an explanation about the emission of fluorescence and Cherenkov light.

2.1. Energy spectrum and mass composition

Several experiments measure the CR flux directly with satellites [3, 4] or balloons [5, 6]. However, at energies above 10^{15} eV, the CR flux becomes so small that direct measurements are not feasible with the surface area and exposure time of those instruments [7]. At these energies, CRs are studied through air showers in the atmosphere, which are explained in detail in Section 2.2, by Earth-based detectors. Fig. 2.1 shows the CR energy spectrum measured by several such experiments. The spectrum roughly follows a power law for many decades in energy for the CR flux:

$$I = \frac{\mathrm{d}N}{\mathrm{d}t\mathrm{d}E\mathrm{d}\Omega\mathrm{d}A} \sim E^{-\gamma}, \tag{2.1}$$

where γ is the spectral index. Several distinct breaks are visible, which indicate changes in the spectral index. The most prominent softenings in the CR energy spectrum at about 4×10^6 GeV and 10^8 GeV are known as the first and second knee, respectively [8]. Between the two knees at 3×10^7 GeV lies a hardening, the first ankle or low-energy ankle [9, 10]. An additional hardening of the spectrum at 5×10^9 GeV is historically known as *the* ankle. The spectral index lies between the values of 2.5 and 3.3 between the first knee and the ankle. However, after the ankle, at about 5×10^{10} GeV, the CR spectrum drops rapidly with $\gamma \approx 5$ [8, 10]. Greisen, Zatsepin and Kuzmin predicted a suppression of CRs with energies in this range due to interactions with the cosmic microwave background [11, 12]. The drop in CR flux at the highest measured energies could be an indication of this so-called GZK cut-off or it could signify the maximum possible energy that CRs can be accelerated to, see e.g. Ref. [13].

While the ankle is presumably a feature related to extragalactic CRs [8], the energy range between the first and second knees is generally linked to the end of the Galactic cosmic ray spectrum. According to Peters [14], particles accelerated in a magnetic field can reach a certain critical magnetic rigidity R_c, where rigidity is defined as $R = E/Z$ with the energy E and charge Z of the particle. This leads to a maximum energy proportional to the nucleus' charge:

$$E_\mathrm{max} = R_\mathrm{c} Z. \tag{2.2}$$

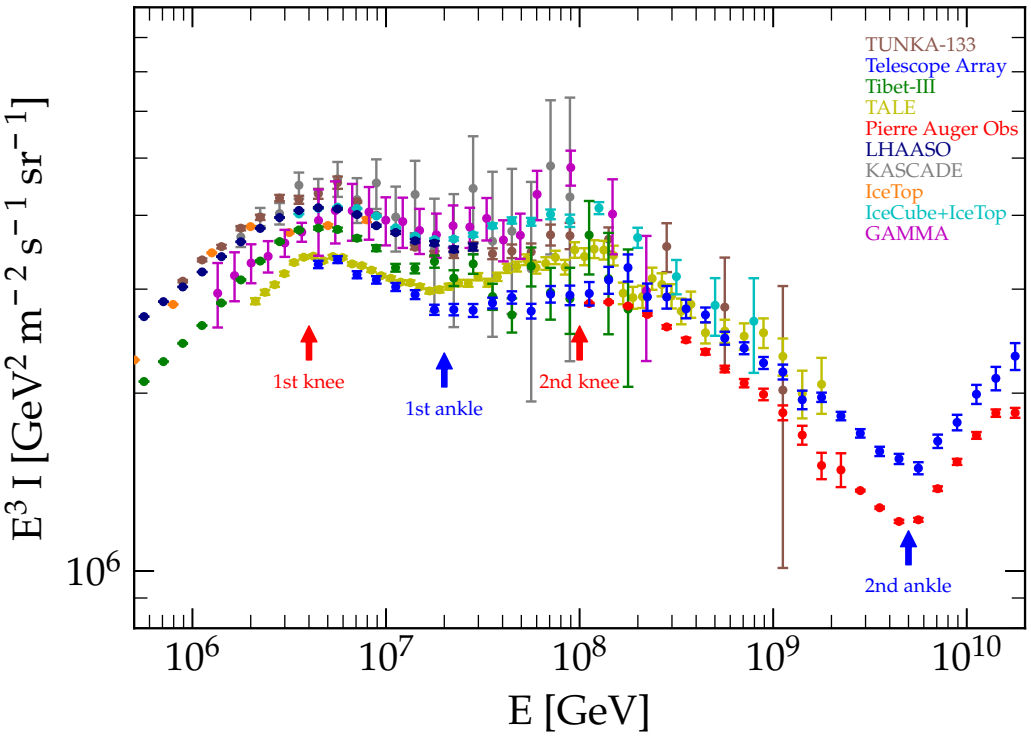

Figure 2.1.: Cosmic ray energy spectrum measured by various ground-based experiments. The flux is multiplied by E^3 for visibility. It can be described with a power law $I \sim E^{-\gamma}$, where the spectral index γ changes at the knee and ankle energies marked in the spectrum. Taken from [9].

Above this energy, the particles are no longer confined in the accelerator region. Hillas proposed that the knee could be the result of the maximum energy of diffusive shock acceleration in supernova remnants [15]. Following Eq. (2.2), the cut-off energies for proton and iron would be separated by a factor of 26, which fits the positions of the first and second knees [9]. With a transition to an extragalactic component above 10^{19} eV, this leaves a gap in the composition between the knees and the ankle. To address this, Hillas introduced a component of unknown Galactic origin [16]. This model with three major components was implemented by Gaisser [17] and modified in the Gaisser-Stanev-Tilav (GST) model [18] with an optional fourth component. Each component has a characteristic critical rigidity and a specific composition of several elements with respective spectral indices. The two Galactic components have critical rigidities of 120 TV and 4 PV, respectively, and are comprised of five elements between proton and iron and small contributions from heavier elements in the second component. The extragalactic flux is modelled with either one component consisting of protons and iron with a critical rigidity of 1.3 EV or two components with 1.5 EV for the first one consisting of protons and iron and 40 EV for the second with purely protons. A comparison of the models with measurements of the mass composition at the time is shown in Fig. 2.2, with the GST model showing a better agreement with data than the older model by Gaisser. The behaviour of the logarithmic mass $\ln A$ between 10^7 GeV and 10^8 GeV is explained in the GST model by the drop-off

of the second Galactic CR component according to the Peters cycle, leading to a heavier composition, and then the rise of an extragalactic component, resulting in the composition becoming lighter again.

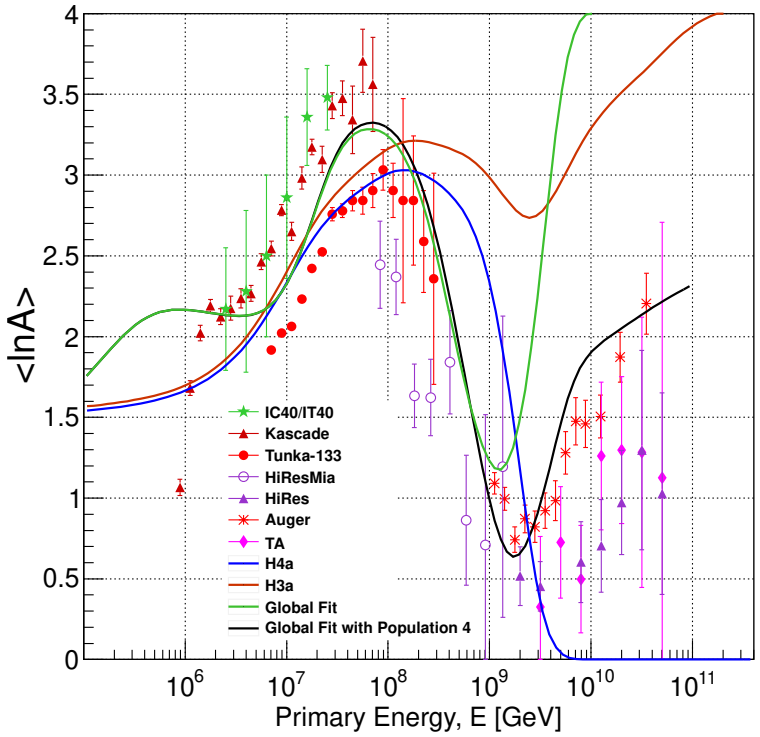

Figure 2.2.: Fit of the Hillas model with a mixed (H3a) and proton-only (H4a) extragalactic component as well as the GST model (Global Fit) with three and four components to data. Taken from [18].

2.2. Development of extensive air showers

After travelling through the vacuum of space, when a high-energy CR arrives at Earth, it will interact with its atmosphere. The first interaction sparks a cascade of secondary particles and interactions known as an Extensive Air Shower (EAS). EASs can then be detected at Earth's surface. To derive information about the primary particles such as their energy and mass composition from EAS measurements, a sufficient model of EASs is needed. Monte Carlo simulations are usually used for that purpose in praxis. Highly detailed simulations in three-dimensional space with tracing of individual shower particles can be performed with CORSIKA [19]. The high runtime of CORSIKA prevents the creation of large samples of simulated air showers. Instead, the one-dimensional MC simulation software CONEX can be used, which relies on numerical solutions for cascade equations to calculate secondary particle distributions [20]. To give an overview about air shower development in the atmosphere, a basic shower model is explained in the following section.

2.2.1. Heitler model

2.2.1.1. Electromagnetic shower

A basic model of the electromagnetic shower component consisting of photons, electrons and positrons was developed by Heitler [21]. The step size of the cascade depends on the radiation length λ_r of photons and electrons in the medium. The radiation length in air expressed in slant depth is $\approx 37\,\mathrm{g/cm^2}$, where slant depth is a measure of the penetrated material at a certain distance l into the atmosphere:

$$X(l) = \int_0^l \mathrm{d}l' \rho(l'), \tag{2.3}$$

with the height-dependent density $\rho(l)$ of air. After travelling a slant depth of $X = \lambda_r \ln 2$, a photon creates an electron and a positron through pair production, while an electron or positron emits a photon via Bremsstrahlung. In this way, the number of particles is doubled in each step: after n steps, $N_n = 2^n$. Consequently, individual particles possess an energy of

$$E_n = \frac{E_0}{N_n} = \frac{E_0}{2^n}, \tag{2.4}$$

where E_0 is the energy of the primary particle or photon initiating the shower. The cascade continues developing in this way, illustrated in Fig. 2.3 (left), until the individual secondary products approach the critical energy $E_c^{em} \approx 85\,\mathrm{MeV}$ in air. At this point the energy loss from radiation becomes smaller than the loss from ionisation and the number of particles has reached its maximum N_{max}. The latter is directly proportional to the primary energy following Eq. 2.4:

$$N_{max} = \frac{E_0}{E_c^{em}}. \tag{2.5}$$

The number of steps n_{max} to reach the shower maximum according to Eq. 2.4 is then

$$n_{max} = \frac{\ln\left(E_0/E_c^{em}\right)}{\ln 2}, \tag{2.6}$$

leading to a depth of shower maximum of

$$X_{max} = n_{max}\,\lambda_r \ln 2 = \lambda_r \ln\left(E_0/E_c^{em}\right). \tag{2.7}$$

Although the Heitler model does not describe real showers as accurately as detailed simulations, important characteristics and proportionalities still apply, especially for the maximum shower size $N_{max} \sim E_0$ and its penetration depth $X_{max} \sim \ln E_0$ [22].

2.2.1.2. Hadronic shower

The Heitler model was adapted by Matthews for the hadronic and muonic component of an EAS [22]. A schematic view of the shower development can be seen in Fig. 2.3 (right). The Heitler-Matthews model assumes a constant interaction length $\lambda_I \approx 120\,\mathrm{g/cm^2}$ of strongly interacting particles in air. In each step corresponding to an atmosphere thickness of

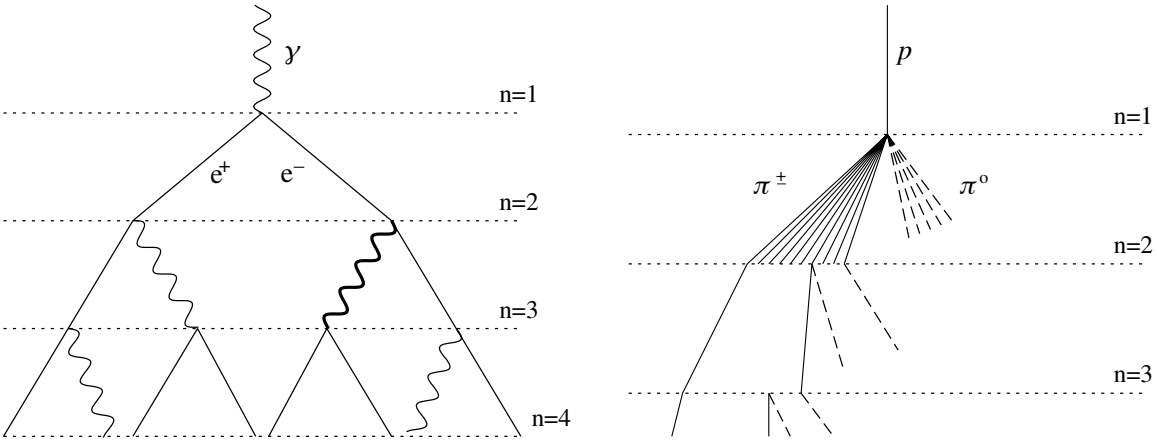

Figure 2.3.: Qualitative shower development for an electromagnetic shower according to the Heitler model (left) and a hadronic shower according to the Heitler-Matthews model (right). Taken from [22].

$\lambda_I \ln 2$, charged pions and neutral pions are created at a ratio of $2:1$. Neutral pions decay into two photons, prompting electromagnetic showers according to the Heitler model. For that reason, one-third of the energy of the hadronic component is transferred into electromagnetic showers in each step of the cascade. The remaining two thirds consisting of charged pions continue to interact until their decay length becomes smaller than the distance between two interactions, marked by a critical energy $E_c^\pi \approx 20\,\text{GeV}$. Below this energy, the pions are assumed to decay into muons.

After n steps, the number of charged pions is $N_n^\pi = (N_{ch})^n$ with the multiplicity $N_{ch} \approx 10$. The energy of each charged pion is then

$$E_n^\pi = \frac{\left(\frac{2}{3}\right)^n E_0}{N_n^\pi} = \frac{E_0}{\left(\frac{3}{2}N_{ch}\right)^n}. \tag{2.8}$$

Following Eq. (2.8), the number of steps to reach the critical energy $E_c^\pi = E_{n_{max}}^\pi$ is

$$n_{max} = \frac{\ln\left(E_0/E_c^\pi\right)}{\ln\left(\frac{3}{2}N_{ch}\right)}. \tag{2.9}$$

Based on the assumption that charged pions decay into muons after falling below the critical energy, the total number of muons is given as

$$N_\mu = N_{n_{max}}^\pi = (N_{ch})^{n_{max}}, \tag{2.10}$$

with each muon inheriting the critical energy E_c^π. The total number of electrons is related to the maximum number of particles of the electromagnetic part of the cascade by $N_e = N_{max}/g$ with the correction factor $g \approx 10$. With Eq. (2.5), this leads to an energy of the primary particle of

$$E_0 = gE_c^{em}N_e + E_c^\pi N_\mu. \tag{2.11}$$

In the context of air shower measurements, the shower energy is obtained by extrapolating the energy deposit of the shower particles in the atmosphere. This calorimetric energy E_{cal}

does not account for the so-called *invisible* energy carried by muons and neutrinos beyond the ground level. The full primary energy is calculated from the calorimetric energy with a correction factor f_{cal}:

$$E_0 = \frac{1}{f_{\mathrm{cal}}} E_{\mathrm{cal}}. \tag{2.12}$$

The correction factor is energy dependent and can be determined from detailed simulations [23] or using data-driven methods [24]. In the Heitler-Matthews model, it can be approximated by equating the calorimetric energy with the energy going into the electromagnetic shower component. Using this approximation as well as Eqs. (2.8), (2.10) and (2.11), the correction factor is

$$f_{\mathrm{cal}} = \frac{E_0 - E_{\mathrm{c}}^{\pi} N_\mu}{E_0} = 1 - \frac{N_\mu}{\left(\frac{3}{2} N_{\mathrm{ch}}\right)^{n_{\max}}} = 1 - \left(\frac{2}{3}\right)^{n_{\max}}. \tag{2.13}$$

Thus, for $n_{\max} \approx 5.7$ at a primary energy of 10^{17} eV, an energy fraction of $f_{\mathrm{cal}} \approx 90\,\%$ is retained in the electromagnetic shower. This is in reasonable agreement with a value of 86 % obtained from air shower simulations for protons [24].

For a precise calculation of the depth of shower maximum $X_{\max}$ with regard to electrons, positrons and photons in the Matthews-Heitler model, each individual electromagnetic subshower would have to be evaluated independently. Instead, $X_{\max}$ can be approximated using only the first interaction and the resulting first wave of electromagnetic subshowers. This underestimates the depth of shower maximum itself but gives an accurate elongation rate Λ according to Matthews [22], with Λ defined as[1]

$$\Lambda = \frac{\mathrm{d}X_{\max}}{\mathrm{d}\log E_0}. \tag{2.14}$$

The first interaction at a depth of $X_1 = \lambda_{\mathrm{I}} \ln 2$ produces $N_{\mathrm{ch}}/2$ neutral pions that carry a total energy of $E_0/3$. The decay into photons leads to a number of N_{ch} electromagnetic subshowers, each with an energy of $(E_0/3)/N_{\mathrm{ch}}$. With the starting point X_1, they reach their maximum according to Eq. 2.7 at a depth of

$$X_{\max}^{\mathrm{p}} = \lambda_{\mathrm{I}} \ln 2 + \lambda_{\mathrm{r}} \ln\left(\frac{E_0}{3 N_{\mathrm{ch}} E_{\mathrm{c}}^{\mathrm{em}}}\right). \tag{2.15}$$

This can be calculated for protons, leading to an elongation rate of $\Lambda_p = 58\,\mathrm{g/cm^2}$ in accordance with simulations [22]. For this work, it is of particular interest to consider the shower maximum for different primary masses. In the Matthews-Heitler model it is suitable to use the superposition model, in which a nucleus of mass A is treated as A individual nucleons with respective energy E_0/A, thus producing A simultaneous hadronic showers. This results in a lower correction factor for the calorimetric energy, for example $f_{\mathrm{cal}} \approx 82\,\%$ for iron at 10^{17} eV according to Eqs. (2.9) and (2.13) in accordance with a value of 80 % from simulations [24]. To account for the superposition model regarding the depth of shower maximum, one only needs to substitute E_0/A for E_0 in Eq. 2.15, resulting in

$$X_{\max}(A) = X_{\max}^{\mathrm{p}} - \lambda_{\mathrm{r}} \ln A. \tag{2.16}$$

[1] Throughout this thesis the convention $\log x := \log_{10} x$ is used.

The proportionality $X_{\max} \sim \ln A$ makes the measurement of the depth of shower maximum a basis for mass composition analysis of CRs. For example, the difference in shower maximum between protons and iron according to the Heitler-Matthews model is $\approx 150\,\mathrm{g/cm^2}$ in reasonable agreement with simulations which predict a difference of $\approx 100\,\mathrm{g/cm^2}$ at 10^{17} eV [22].

2.2.2. Fluorescence light and Cherenkov light

In the development of an EAS in the atmosphere, nitrogen molecules are excited. These then de-excite after about 10 ns by isotropically emitting fluorescence photons. The measured nitrogen fluorescence spectrum has multiple peaks in the near-ultraviolet range between 300 nm and 400 nm [25], shown in Fig. 2.4. Fluorescence efficiency can be measured in the laboratory and is $\approx 10\,\%$ per ionizing particle [26]. The number of fluorescence photons is proportional to the fluorescence yield and energy deposit in the atmosphere and therefore directly related to the calorimetric energy [27].

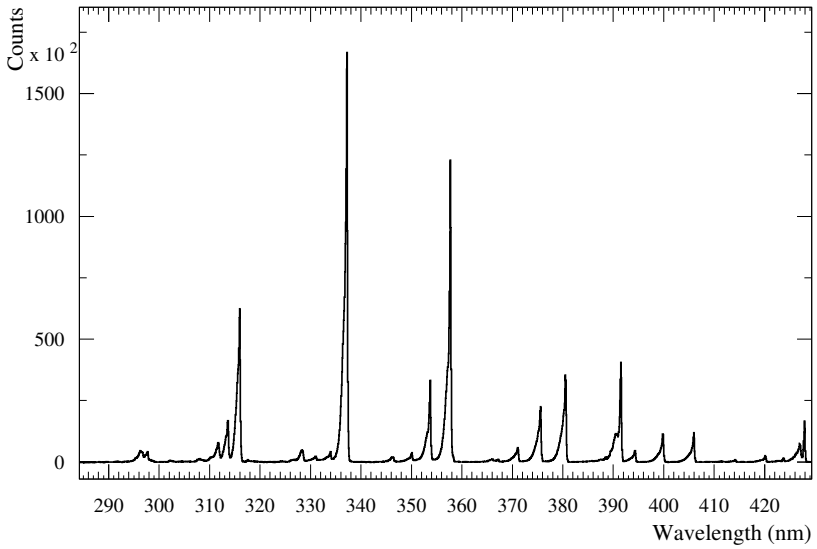

Figure 2.4.: Nitrogen fluorescence spectrum in the near-ultraviolet range. Taken from [28].

Shower particles that move at a speed greater than the speed of light in air emit Cherenkov radiation. For a particle moving at relativistic speed β, Cherenkov photons are emitted at an angle of

$$\cos\theta = \frac{1}{n\beta} \tag{2.17}$$

to the trajectory of the particle. At a height of 10 km in the atmosphere with a refractive index of $n(10\,\mathrm{km}) = 1.00007$ [25], this results in a narrow cone of Cherenkov light with an opening angle of $2\theta = 1.4°$, though the cone widens slightly with decreasing height. Because of the forward-peaked nature of Cherenkov radiation, it can only be measured from an EAS if the shower is roughly pointing towards the detector. Nevertheless, the emission angle relative to the shower axis can be much larger than the Cherenkov emission angle from Eq. (2.17), as seen in Fig. 2.5. This is because of the shower particles that do not move in parallel to the shower axis due to Coulomb scattering [25].

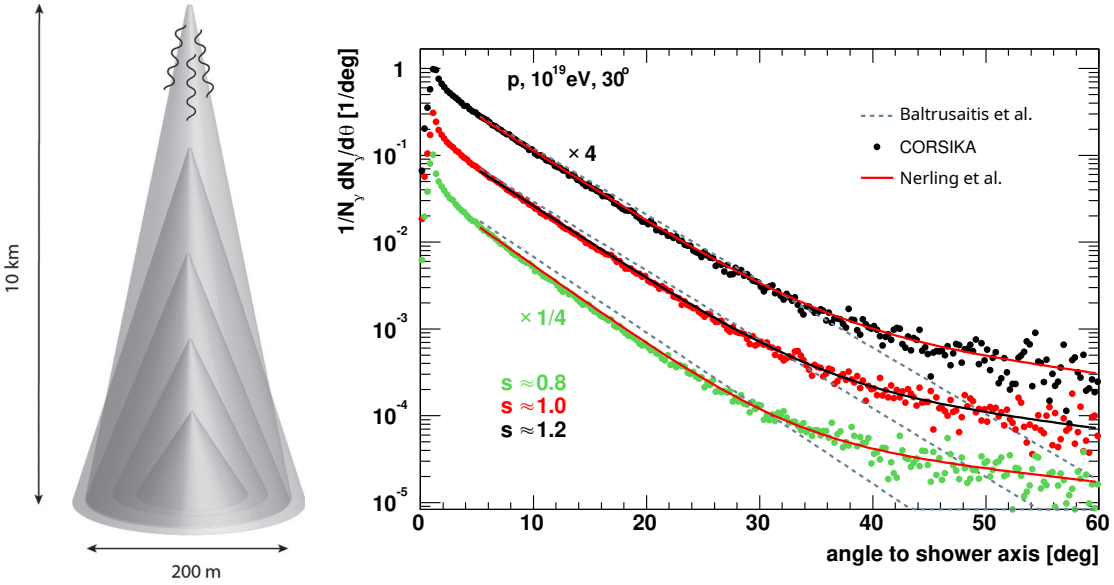

Figure 2.5.: Left: Wider Cherenkov opening angles at lower altitudes together with Coulomb scattering result in a characteristic Cherenkov distribution at ground level. Taken from [25]. Right: Angular distributions of Cherenkov photons at different shower ages s for one CORSIKA shower. Adapted from [29].

In contrast to the peaked fluorescence emission, the spectrum of Cherenkov radiation is more continuous. Taking scattering and absorption into account, it can be measured at wavelengths between 300 nm to 450 nm, placing it in the same range as fluorescence photons [25]. It can be used as an additional source of information for the EAS, as the number of Cherenkov photons is proportional to the number of particles above the Cherenkov threshold energy [27], the latter given by

$$E = \frac{n(h)}{\sqrt{n(h)^2 - 1}} m \tag{2.18}$$

with the height-dependent spectral index $n(h)$ in the air and a particle of mass m [25].

Light emission in air from EASs induced by ultra-high energy cosmic rays with energies greater than 10^{18} eV is mostly measured as fluorescence light. Because fluorescence light is emitted isotropically, it can also be detected from showers that are not directed at the detector. However, the fluorescence yield per charged particle is only $\approx$ 5 photons/m [26]. At energies below the ultra-high energy range, the total intensity of fluorescence light becomes too small to be efficiently measured. Instead, the intense forward-peaked Cherenkov light can be used for the air shower measurement at these energies, even though it is limited to showers that are directed approximately at the detector.

3. Pierre Auger Observatory

As mentioned in Section 2.1, high-energy CRs can only be measured through EASs by ground-based experiments. The largest of these experiments is conducted in the Pierre Auger Observatory. It is based near the city of Malargüe in Argentina and has been collecting data since 2004 with multiple detection techniques [30].

3.1. Surface detector

The Surface Detector (SD) of the Pierre Auger Observatory consists of approximately 1600 water Cherenkov tanks covering an area of 3000 km^2. Each station contains 12 000 l of ultra-pure water and three Photomultiplier Tubes (PMTs) to measure Cherenkov light from shower particles passing through the water. The electromagnetic part of a shower induces a signal proportional to its energy, while the signal from the muonic component is proportional to the geometrical path length in the water tank [31]. The SD duty cycle is almost 100 % and the spacing of the stations ensures the detection of all EASs with primary energies greater than 3×10^{18} eV [32].

The main SD detector based on a hexagonal grid with a spacing of 1500 m between the stations is supplemented by two low-energy extensions. The infill array with 61 stations and a spacing of 750 m is part of the Auger Muon and Infill for the Ground Array (AMIGA), which has been gathering data since 2008 [33]. A map of all the stations of the main and 750 m array can be seen in Fig. 3.1. Another extension with an even smaller spacing of 433 m, consisting of 19 tanks in a triangular grid, is installed in the area of the infill array and demonstrates full efficiency for measurements of primary energies above 10^{17} eV [34].

An EAS triggering a number of SD stations gives a lateral view of the shower front. Individual station trigger times can be used in combination to determine the direction of the shower, as well as the impact point of the shower on the ground. The signal densities, on the other hand, are compared with an expected lateral distribution function to obtain the shower size [35].

3.2. Fluorescence detector

3.2.1. Setup

While the SD measures the lateral distribution of an EAS, the Fluorescence Detector (FD) observes the longitudinal development using fluorescence and Cherenkov light. The shower reconstruction using FD is described in detail in Section 3.2.2. FD data acquisition is performed with 24 telescopes deployed in four sites around the main SD array called Los Leones, Los Morados, Loma Amarilla, and Coihueco [36]. Three additional telescopes were

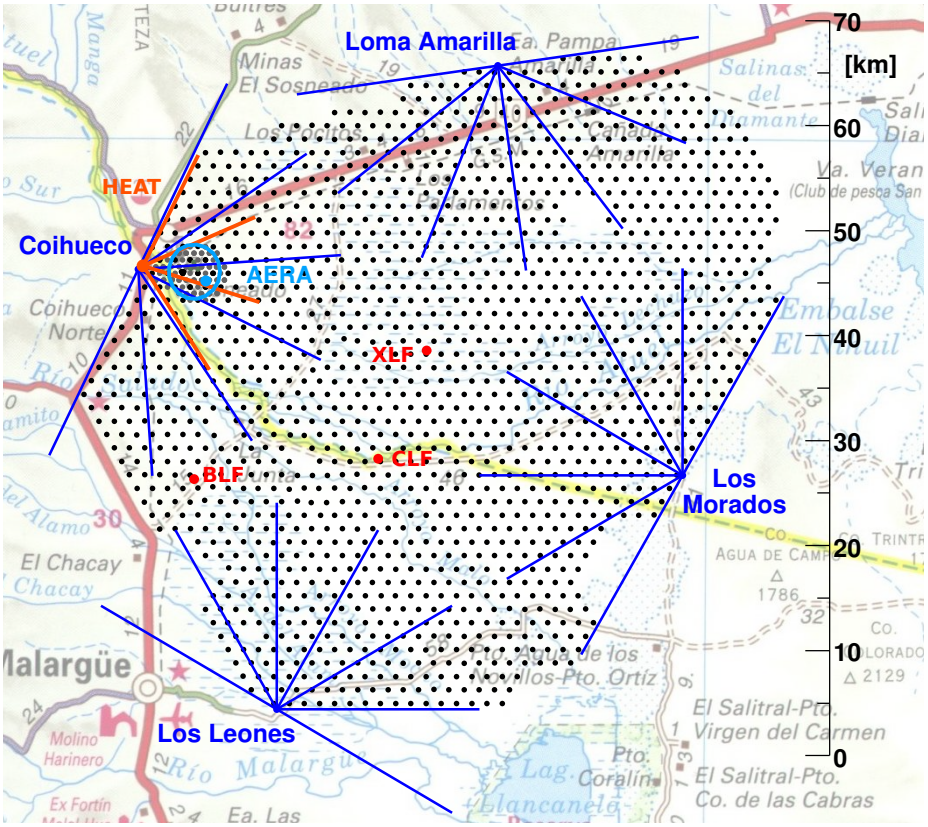

Figure 3.1.: Map of the Pierre Auger Observatory. Black dots are SD stations. The four main FD sites and HEAT with the field of view of their telescopes are shown in blue and red, respectively. Also visible are the radio array AERA and laser facilities for atmosperic measurements. Taken from [2].

deployed at the Coihueco site and can be tilted upwards by 29°. These High-Elevation Auger Telescopes (HEAT) are better suited for the detection of EAS with a shower maximum at larger altitudes. HEAT overlooks the infill array to enable hybrid measurements with it. Events from Coihueco and HEAT can be combined into a virtual FD detector called HeCo. The optics of HEAT telescopes resemble those of standard FD telescopes, with a field of view of 1.5° per camera pixel and an effective aperture of 3 m^2 [37]. Each individual FD telescope has a full field of view spanning 30 by 30 degrees. Light is focused by a curved mirror onto the camera, consisting of 440 PMTs with light collectors acting as pixels. Data can only be collected on clear nights without the moon in the field of view of a telescope. Weather conditions or scattered moonlight can also cause dead time, leading to a duty cycle of about 13 % for the FD telescopes [36]. The location of the FD sites can be seen in Fig. 3.1.

The signals that come from the PMT electronics are digitised through an analog board. For HEAT telescopes, the signal is digitised at a rate of 20 MHz instead of 10 MHz as employed for standard telescopes [37]. The effect of aliasing occurs for signal with a frequency larger than the so-called Nyquist frequency, which is equal to half of the sampling frequency. Above this frequency, part of the signal is not sampled fast enough

and incorrectly measured at a lower frequency. To avoid this, antialiasing filters are applied.

The FD trigger system is composed of four stages, starting with the First Level Trigger (FLT). If the sum of Analogue-Digital Converter (ADC) samples over multiple time bins exceeds a threshold, a pixel trigger is generated. A constant pixel trigger rate of 100 Hz is kept by adjusting the threshold depending on the background light.

The FLT pixel triggers are then evaluated by the Second Level Trigger (SLT), which selects events that have multiple triggered pixels in a continuous track. This track must be at least five pixels long with a basic pattern as illustrated in Fig. 3.2. It should be noted that tracks with one missing pixel also pass the SLT, which accounts for defective or below-threshold PMTs.

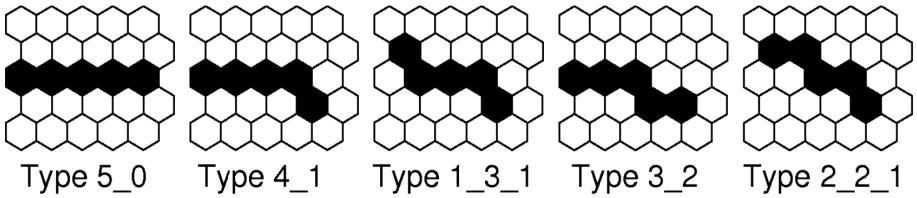

Figure 3.2.: Pixel trigger patterns regarded as valid traces by SLT if four out of five pixels are triggered. Taken from [36].

Events that pass the FLT and SLT hardware triggers are read out and further processed by two software triggers. The first of these, the Third Level Trigger (TLT), is designed to reject events caused by lightning, muons, or noise. In order to reduce the processing time, lightning events are classified solely with the number of triggered pixels and their multiplicity from FLT, without reading the complete ADC trace.

The last trigger step is a hybrid trigger known as T3. A first basic reconstruction of an event is performed to obtain a rough estimate of the shower direction and ground impact time. Events are classified by T3 as different types of showers or are rejected as muons, noise, or by other filters. For events accepted by T3, SD stations close to the FD telescope are checked for signals at the calculated impact time. This is because events do not always create SD triggers for three or more stations, especially at energies below 3×10^{18} eV [36].

3.2.2. Reconstruction of Cherenkov-light dominated showers

The first large-scale experiment that used fluorescence light to measure EASs was the Fly's Eye detector in Utah [38]. In this section, the FD reconstruction at the Pierre Auger Observatory will be described instead, where a combination of information from both Cherenkov light and fluorescence light is used in the procedure. A focus is placed on the reconstruction of Cherenkov-light dominated events.

3.2.2.1. Standard geometry fit

The first part of reconstructing Cherenkov-dominated showers is identical to the reconstruction of pure fluorescence events. Fig. 3.3 shows the triggered pixels for a typical

shower observed by a HEAT telescope. The directional data of the PMTs is used to identify the Shower Detector Plane (SDP), the plane that includes the shower axis and the detector. In the SDP, the shower axis is defined by its angle to the ground χ_0 and the smallest distance between the shower axis and the detector, the impact parameter R_p, as seen in Fig. 3.4. After finding the SDP, the geometry of the shower in the SDP has to be determined using the timing information provided by the PMTs. The trigger times t_i are related to the pixel angles χ_i in the SDP in the following way:

$$t_i = t_0 + \frac{R_\mathrm{p}}{c} \tan\left(\frac{\chi_0 - \chi_i}{2}\right). \tag{3.1}$$

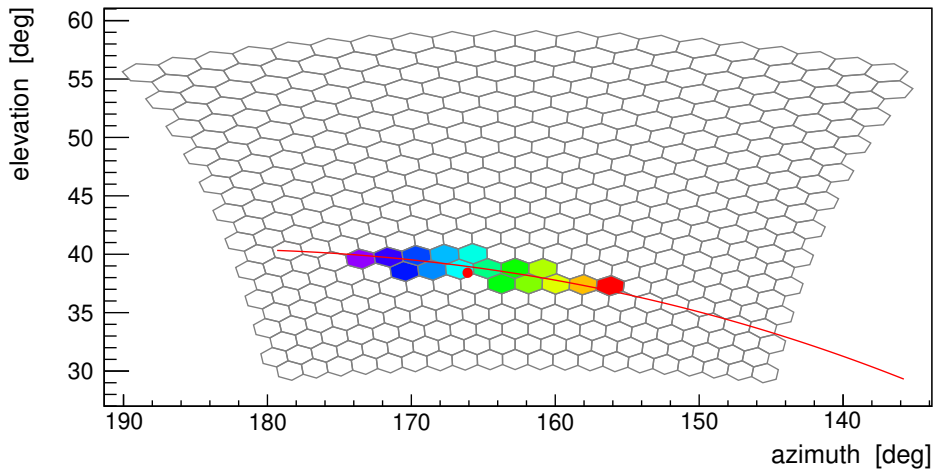

Figure 3.3.: A typical telescope image, recorded for Auger Event 170027052700 in HEAT telescope 3. Pixel colors correspond to trigger times from violet (early) to red (late).

For the short measured time intervals of several 100 ns, the relation between trigger times and angles appears to be nearly linear, as seen in Fig. 3.5. For that reason, it is difficult to perform a fit with Eq. (3.5) with three free parameters. One possibility of reconstructing such showers is to use SD timing information. However, at low energies between 10^{16} eV and 10^{17} eV, as examined in this work, showers are typically not detected by SD, making this hybrid approach unfeasible. As an alternative, the so-called Profile Constrained Geometry Fit (PCGF) is employed. This approach utilises fits to the energy deposit profile for a given geometry, a concept that will be introduced next.

3.2.2.2. Profile fit

The energy deposit profile of a shower is obtained by relating the measured light flux $\boldsymbol{y}$ to the energy deposit $\boldsymbol{w}$. $\boldsymbol{y}$ and $\boldsymbol{w}$ are vectors with each component of either vector belonging to a specific time bin or corresponding slant depth interval. They are related by the Cherenkov Fluorescence Matrix (CFM) $\boldsymbol{C}$ through

$$\boldsymbol{y} = \boldsymbol{C}\boldsymbol{w}. \tag{3.2}$$

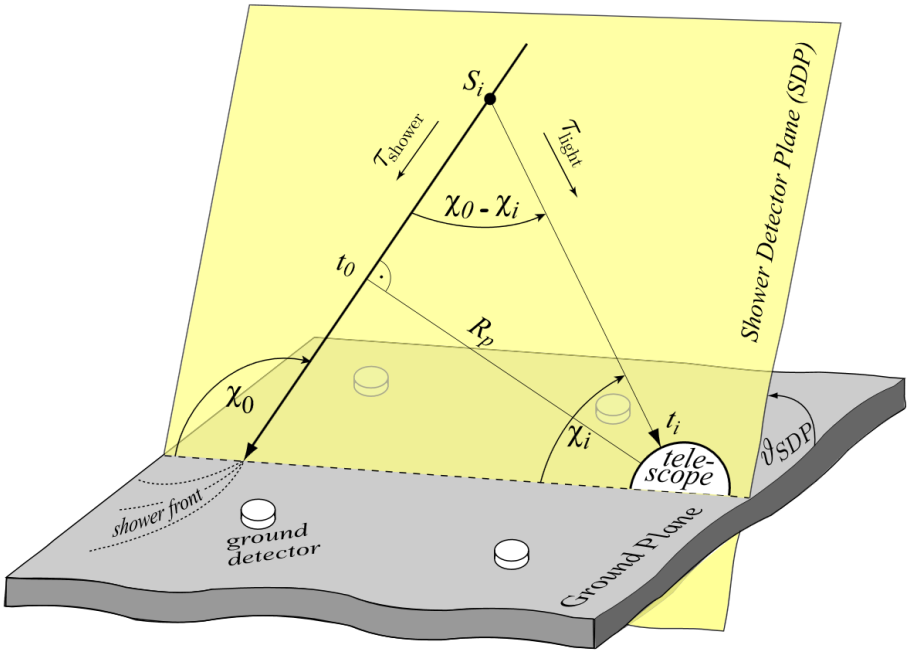

Figure 3.4.: Schematic depiction of the shower detector plane, defined by the shower axis and location of the FD telescope. Taken from [39].

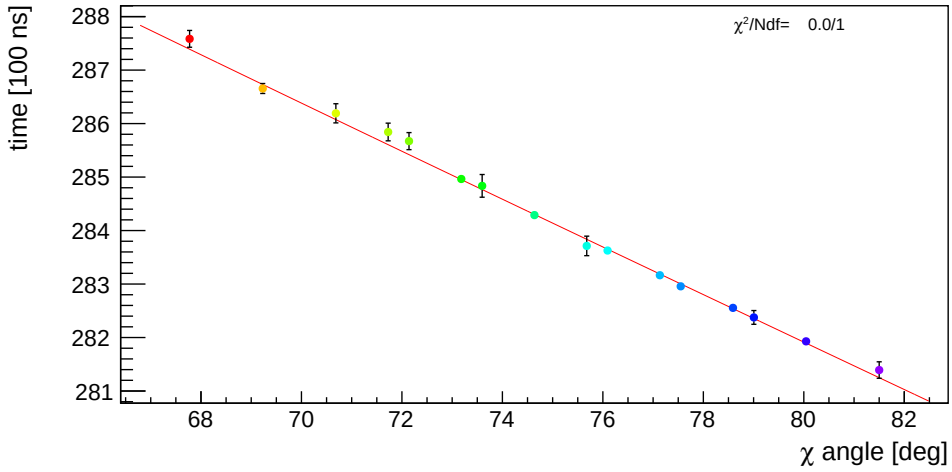

Figure 3.5.: Trigger times and SDP angles for Auger event 170027052700. Each data point corresponds to one pixel in Fig. 3.3. Shown in red is the timing fit of Eq. (3.1), which is approximately linear.

The triangular CFM can be easily inverted to give an estimate of the energy deposit: $\hat{\boldsymbol{w}} = \boldsymbol{C}^{-1}\boldsymbol{y}$ [27]. A common fit function for the energy deposit profile is the Gaisser-Hillas (GH) function [40]:

$$N(X) = N_{\max}\left(\frac{X - X_0}{X_{\max} - X_0}\right)^{(X_{\max}-X_0)/\lambda} \exp\left(\frac{X_{\max} - X}{\lambda}\right). \tag{3.3}$$

It has four free parameters: the depth of shower maximum $X_{\max}$ and the maximum number of particles $N_{\max}$ as well as the two shape parameters λ and X_0.

3.2.2.3. Profile constrained geometry fit

The PCGF method was first used for the HiRes detector [41]. It was implemented into the Auger analysis software Offline [42] as a reconstruction module [43]. Instead of fitting the complete timing equation Eq. (3.1), the angle of the shower axis in the SDP χ_0 is set to specific scan values, removing one of the free parameters. For each χ_0, a linear regression, the so-called timing fit, provides values for the parameters t_0 and R_p in Eq. (3.1). With this combination of the geometry parameters, a GH fit is performed for the energy deposit profile. Subsequently, the light flux at the telescope is predicted for the particular geometry and profile.

Following the completion of the χ_0 scan and the corresponding fits for each value, the most suitable candidate for χ_0 must be determined. This is conducted using a minimum χ^2 analysis with multiple χ^2 contributions. Usually, after a first scan in 1° steps, an additional scan in 0.1° steps is performed around the initial χ^2 minimum. An example for a χ_0 scan with corresponding χ^2 values is shown in Fig. 3.6. The total χ^2 is formed with the following contributions:

$$\chi^2_{\text{tot}} = \chi^2_{\text{time}} + \chi^2_{\text{prof}} + \sum_Y \chi^2_Y. \tag{3.4}$$

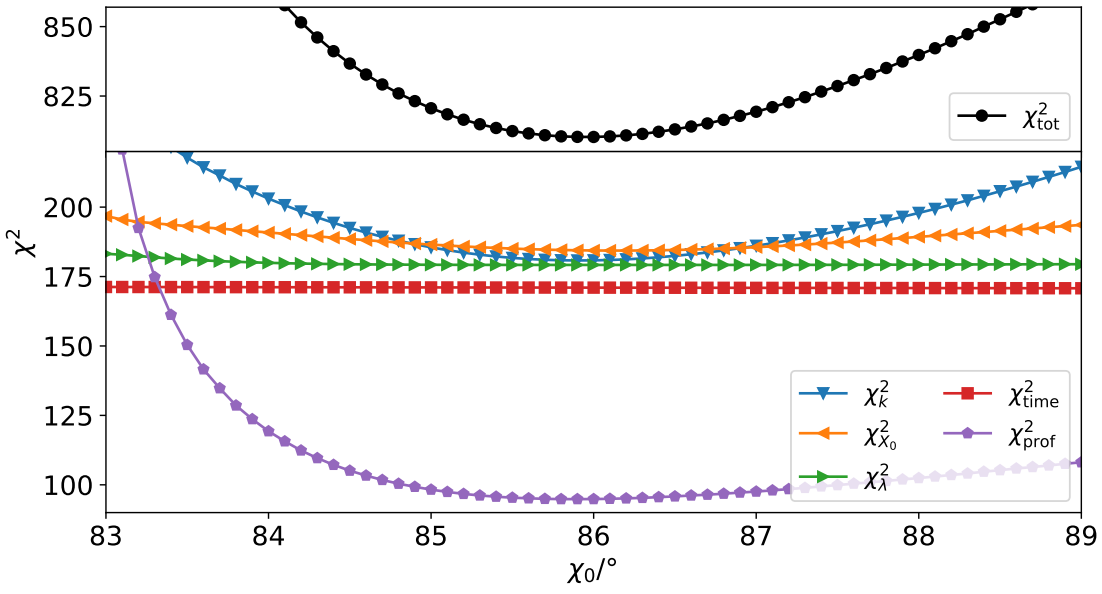

Figure 3.6.: Second χ_0 scan in 0.1° steps as part of the PCGF reconstruction for Auger event 170027052700. The total value for χ^2 (top) is formed by adding individual contributions from the profile fit, timing fit as well as profile constraints based on the shape parameters X_0 and λ and the area over peak k of the profile (bottom). The result of the scan is $\chi_0 = 86.0°$.

The contribution χ^2_{time} originates from the timing fit, a linear regression to Eq. (3.1) with the parameters t_0 and R_p. Similarly, the profile contribution χ^2_{prof} originates from the fit of the GH function to the energy deposit profile. χ^2_{prof} includes the agreement between the predicted and measured light flux for each χ_0 value. Additional contributions are constraints placed on the shape parameters λ and X_0 of the GH function. Shape constraints are based on a Gaussian distribution of the shape parameters, with mean and variance of the distribution derived from measured data [44]. The constraint for a given shape parameter $Y \in \{\lambda, X_0\}$ yields

$$\chi^2_Y = \frac{Y - \langle Y \rangle}{\sigma(Y)^2}. \tag{3.5}$$

An additional third constraint was introduced based on the area-over-peak k of the profile, which was found to be approximately universal:

$$k = \frac{E_\text{cal}}{N_\text{max}}. \tag{3.6}$$

The constraint on k is also based on a Gaussian distribution with mean and variance values from simulations [45]. The main purpose of the shape constraints is to ensure a successful GH fit if only a small slant depth around X_max was observed or the shower maximum even lies outside the field of view [27]. Usually, shape constraints are included as part of a combined χ^2_{prof}, but they are considered separate contributions in this thesis for convenience.

Different parametrisations of the GH function with resulting constraints can be used in the profile fit. Eq. (3.3) is also called the *Classic* parametrisation in that context. A different yet mathematically equivalent way to express the GH function is the so-called Universal Shower Profile (USP) parametrisation [46]. Instead of the parameters λ and X_0, the USP parametrisation uses parameters L and R given by

$$R = \sqrt{\frac{\lambda}{|X_0 - X_\text{max}|}} \quad \text{and} \quad L = \sqrt{\lambda |X_0 - X_\text{max}|},$$

resulting in the GH function

$$N(X) = N_\text{max} \left(1 + \frac{R(X - X_\text{max})}{L}\right)^{R^{-2}} \exp\left(\frac{X_\text{max} - X}{RL}\right). \tag{3.7}$$

Higher correlation between λ and X_0 leads to a given shower profile having multiple possible pairs of shape parameters in the classic parametrisation, while L and R are less correlated and thus well-defined for a given shape [46]. A third constraint is not necessary for the USP parametrisation, as the parameter L already approximates the area-over-peak k of the distribution: $L \simeq k/\sqrt{2\pi}$ [45].

After selecting the best χ_0 candidate using the χ^2 minimisation, the light at the aperture is recalculated with the result geometry and a final energy deposit fit is performed, shown as examples for a Cherenkov-light dominated shower in Figs. 3.7 and 3.8, respectively. The calorimetric energy as introduced in Section 2.2 is then obtained from the integral of the energy deposit profile.

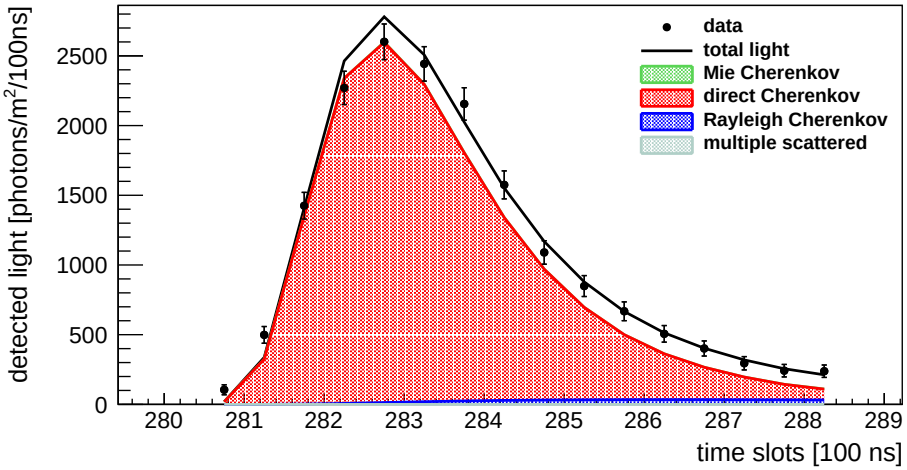

Figure 3.7.: Calculated light at the aperture for Auger event 170027052700. The Cherenkov light fraction was found to be 88 %.

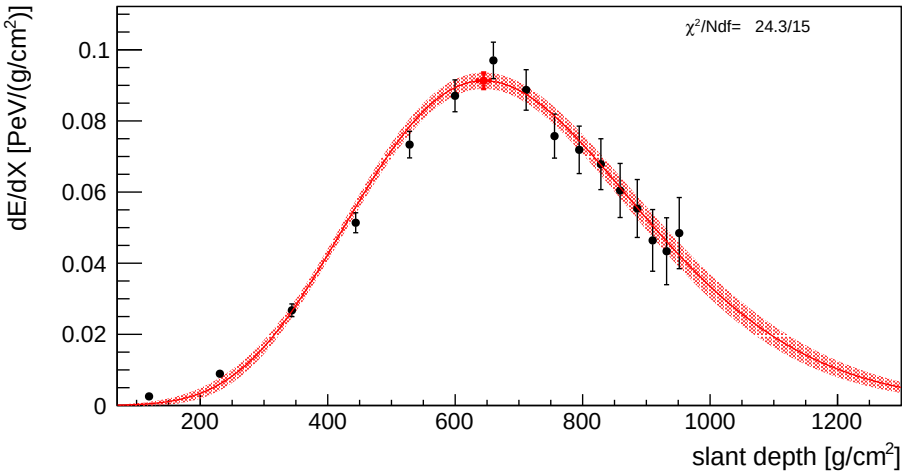

Figure 3.8.: Profile fit using a GH function (shown in red) for Auger event 170027052700 with black dots representing the energy deposit profile.

4. Performance of the Profile-Constrained Geometry Reconstruction

In this chapter, the performance of the event reconstruction using the PCGF is studied. The main focus will be the bias and resolution of the measurement of the shower maximum, which is the observable used in this work to study the composition of CRs. For that purpose, detailed realistic Monte Carlo (RealMC) simulations including the detector response are used. The generated and reconstructed values $X_{\mathrm{max}}^{\mathrm{gen}}$ and $X_{\mathrm{max}}^{\mathrm{rec}}$ of the simulations are compared and the reconstruction technique of Cherenkov-light dominated showers is investigated in the following sections.

4.1. Event simulation with RealMC and quality selection

The existing simulations used in the main analysis of this work were created for exposure calculations for the International Cosmic Ray Conference (ICRC) in 2021 [10]. The energy range is $10^{15.3}$ eV to $10^{18.2}$ eV, with showers generated following a power law with spectral index $\gamma = 2.2$. Events are simulated exclusively for HEAT and Coihueco telescopes combined into the virtual system HeCo with a mixed composition consisting of equal contributions from protons, helium, nitrogen, and iron. The underlying hadronic interaction model is Sibyll 2.3d [47] and the simulation code CONEX, introduced in Section 2.2, was used. The events are processed with the Auger reconstruction software Offline using PCGF to determine the geometry of the shower as explained in Section 3.2.2. In this work, several basic selection cuts are applied to the Advanced Data Summary Trees (ADSTs), which are output of the Offline software. These cuts exclude

- events caused by atmospheric monitoring laser facilities,
- bad FD periods or periods with HEAT telescopes tilted downwards,
- events with saturated or otherwise unusable pixels,
- events without atmospheric database data, with a large cloud cover, or a high vertical aerosol optical depth (VAOD),
- events without a successfully reconstructed energy.

The exact cut values and efficiencies are shown in Tab. 4.1. After applying basic selection cuts, 674013 events with a successful reconstruction remain for further analysis.

In a first step, the ICRC21 simulations are compared with an analysis performed by the Telescope Array (TA) collaboration [48], in which one of its fluorescence detectors, the

Table 4.1.: ADST selection cuts with efficiencies ε for simulations as well as for data reconstructed in the HeCo system.

ADST cut	ε (sim.)	ε (data)
!isCLF	100.0 %	100.0 %
!isXLF	100.0 %	100.0 %
HEATdownCampaignFile	99.7 %	99.5 %
heatOrientationUp	100.0 %	98.9 %
badFDPeriodRejection	100.0 %	100.0 %
skipSaturated	93.8 %	93.8 %
noBadPixelsInPulse	100.0 %	100.0 %
good10MHzCorrection	100.0 %	100.0 %
hasMieDatabase	84.6 %	84.1 %
hasGDASDatabase	100.0 %	100.0 %
maxVAOD 0.1	90.6 %	88.7 %
LidarCloudRemovalPCGF 25	95.2 %	96.0 %
minLgEnergyFD 1e-20	41.0 %	23.0 %

Telescope Array Low-energy Extension (TALE) [49], was used. In a mass composition analysis by TA [50], MC simulations based on the interaction models QGSJetII-03 [51] and EPOS-LHC [52] were reconstructed using the PCGF. To compare the two sets, the same quality cuts as in the TA analysis are applied to the HeCo events, as given in Tab. 4.2. Efficiencies are given based on the complete ICRC21 sample with basic selection cuts and the generated mixed composition with equal parts of proton, helium, nitrogen and iron showers. It is worth noting that in the case of HeCo the cuts on δX_{max} and $\delta E/E$ do not include the propagated geometrical uncertainties, and instead only the statistical uncertainty of the profile. Further details are explained in Appendix A.2.

Table 4.2.: Quality cuts used in a TA mass composition analysis [50] and applied to ICRC21 simulations reconstructed in HeCo. Efficiencies ε are given in relation to events with basic selection cuts from Tab. 4.1.

cut variable	cut value	ε (Auger sim.)
Cherenkov light fraction	$f_{\mathrm{Ch}} > 35\,\%$	65.3 %
inverse angular speed	$0.014\,\mu\mathrm{s}/^\circ < 1/\omega < 0.1\,\mu\mathrm{s}/^\circ$	65.6 %
angular tracklength	$\alpha_{\mathrm{tr}} > 6.1^\circ$	56.4 %
zenith angle	$28^\circ < \theta < 65^\circ$	92.2 %
impact parameter	$0.4\,\mathrm{km} < R_{\mathrm{p}} < 5\,\mathrm{km}$	100.0 %
depth of shower maximum	$435\,\frac{\mathrm{g}}{\mathrm{cm}^2} < X_{\mathrm{max}} < 920\,\frac{\mathrm{g}}{\mathrm{cm}^2}$	92.7 %
estimated X_{max} error	$\delta X_{\mathrm{max}} < 200$	94.5 %
estimated relative energy error	$\delta E_{\mathrm{cal}}/E_{\mathrm{cal}} < 0.6$	98.9 %
profile χ^2/N_{dof}	$\chi^2_{\mathrm{prof}}/N_{\mathrm{dof}} < 12$	88.3 %
timing χ^2/N_{dof}	$\chi^2_{\mathrm{time}}/N_{\mathrm{dof}} < 4.5$	100.0 %

As seen in Fig. 4.1, at lower energies between $10^{15.5}$ eV and 10^{16} eV, the ICRC21 events are almost exclusively recorded in HEAT telescopes, due to the upward tilted nature of the latter. At $10^{16.75}$ eV, events from HEAT and Coihueco are approximately equal. Above this energy, Coihueco events are more common. Showers recorded as a mirror crossing between HEAT and Coihueco telescopes also become more common around 10^{17} eV, likely due to a larger number of triggered pixels at higher energies. Compared to basic selection cuts, quality cuts suppress pure Coihueco events at all energies to a higher extent than HEAT events.

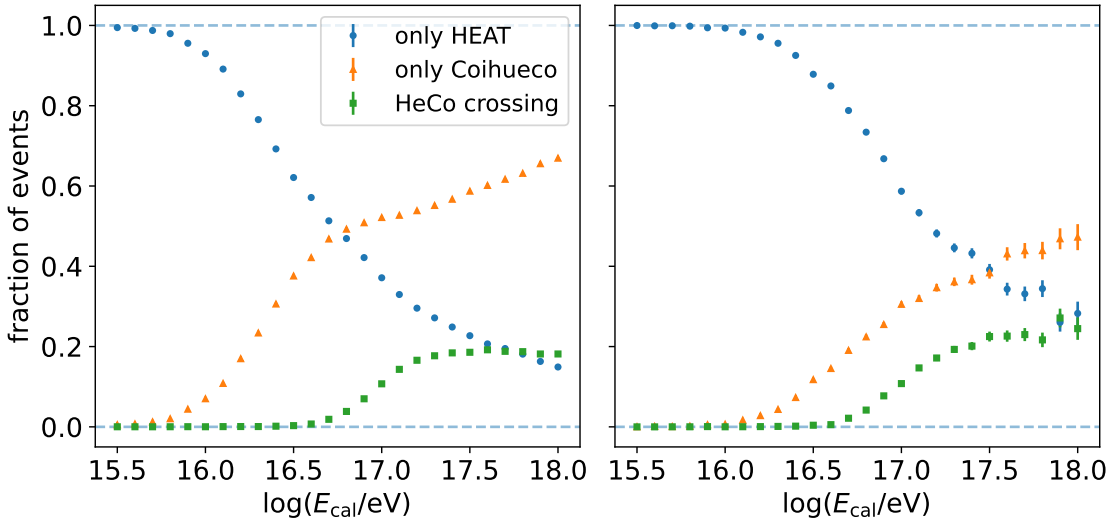

Figure 4.1.: Fraction of RealMC HeCo events recorded in HEAT or Coihueco telescopes or as a mirror-crossing event between the two with basic selection cuts (left) and all quality cuts (right) applied.

After the quality cuts are applied, the Auger and TA simulations are compared based on the difference between the reconstructed and the generated depth of shower maximum:

$$\Delta X_{\max} = X_{\max}^{\mathrm{rec}} - X_{\max}^{\mathrm{gen}}. \tag{4.1}$$

The mean $\langle \Delta X_{\max} \rangle$ and standard deviation $\sigma(\Delta X_{\max})$ of this difference are also called the $X_{\max}$ bias and $X_{\max}$ resolution, respectively. The bias and the resolution of the ICRC21 simulations compared to the values given by TA are shown as an example for protons in Fig. 4.2. There is a significant difference in bias and resolution. As can be seen, the bias is much larger for the Auger simulations, especially at lower energies, with $\langle X_{\max} \rangle = 44.5\,\mathrm{g/cm^2}$ at $10^{15.75}$ eV compared to $-15.3\,\mathrm{g/cm^2}$ for the TA simulations. Moreover, the resolution is larger for Auger simulations in a range of $25\,\mathrm{g/cm^2}$ to $45\,\mathrm{g/cm^2}$ between $10^{15.5}$ eV and 10^{18} eV. Similar findings apply to other primaries, shown in Appendix A.1. It is unlikely that these differences originate from hardware differences, as TALE has similar characteristics to the ones in Auger FD telescopes as introduced in Section 3.2. More specifically, TALE has a field of view per pixel of 1° [53], a viewing area of $3.7\,\mathrm{m^2}$ per mirror, and time binning of 10 MHz [50].

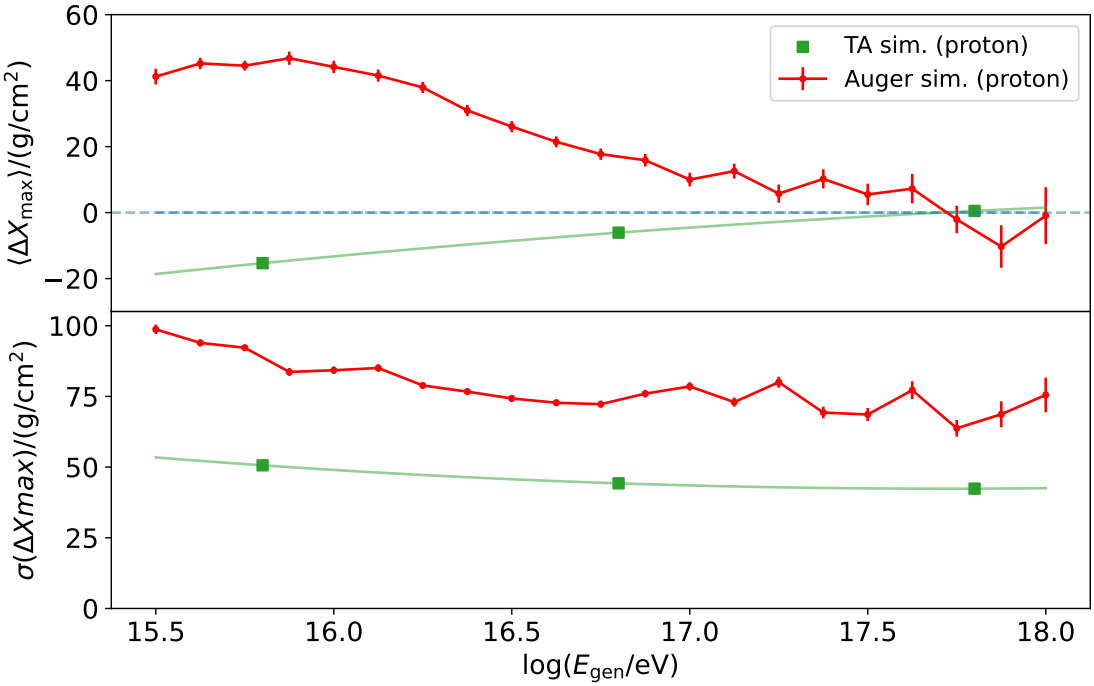

Figure 4.2.: X_{max} bias and resolution of Auger proton simulations reconstructed in the HeCo system using the PCGF, in comparison to values given by TA in Ref. [50]. The same quality cuts given in Tab. 4.2 are applied in both cases.

In conclusion, the comparison of the X_{max} bias and resolution serves as a motivation to examine the reconstruction of Cherenkov-dominated showers at Auger in-depth and to seek possible improvements before performing a mass composition analysis with real HeCo data.

4.2. Testing changes in the reconstruction

The RealMC showers from ICRC21 are only available in reconstructed forms and not as pure simulations. To investigate changes in the reconstruction, a more suitable way would be to have a stable set of simulations to which different reconstruction scenarios can be applied. In this way, fluctuations from repeated simulations can be avoided. Specifically, eight energy settings are chosen in logarithmic steps of 0.25 between $10^{15.75}$ eV and $10^{17.5}$ eV. At each energy setting, 250 showers are generated for protons and for iron using CONEX with the hadronic interaction model Sibyll 2.3d. In this process, the shower geometries are not randomly simulated. Instead, to improve selection efficiency, geometries of showers with similar energies are taken from the existing ICRC21 simulations with quality cuts. In the next step, the newly generated CONEX showers are then simulated with these geometries and reconstructed with the same Offline configuration as the ICRC21 set. All ADST and TA quality selection cuts introduced in Section 4.1 are also applied to the reconstructed showers.

To test if the smaller sample of newly simulated geometries is representative of the complete ICRC21 simulations, the two sets are compared based on the resulting $X_{\max}$ bias and resolution, shown in Fig. 4.3 for proton events. No significant difference in the bias of the two sets can be observed. However, the resolution has improved for the sample generated in this Section. This could be a result of the pre-selection of shower geometries or it could come from differences in the time frame used for the simulations. For the ICRC21 set, the showers were generated for different data taking conditions between 2012 and 2017, while newly generated showers were simulated only for a 10 minute interval (2021-07-09, 23:19:59 - 23:29:59). For iron, slight differences in bias and resolution are also observed (Fig. A.4, Appendix A.1). However, the relative response to changes in the reconstruction is expected to be comparable between the new set and the ICRC21 set. Effects of different changes are discussed in the following sections. Results are only shown for the proton part of the sample, but select cases were also applied to the sample of iron showers and similar results were observed.

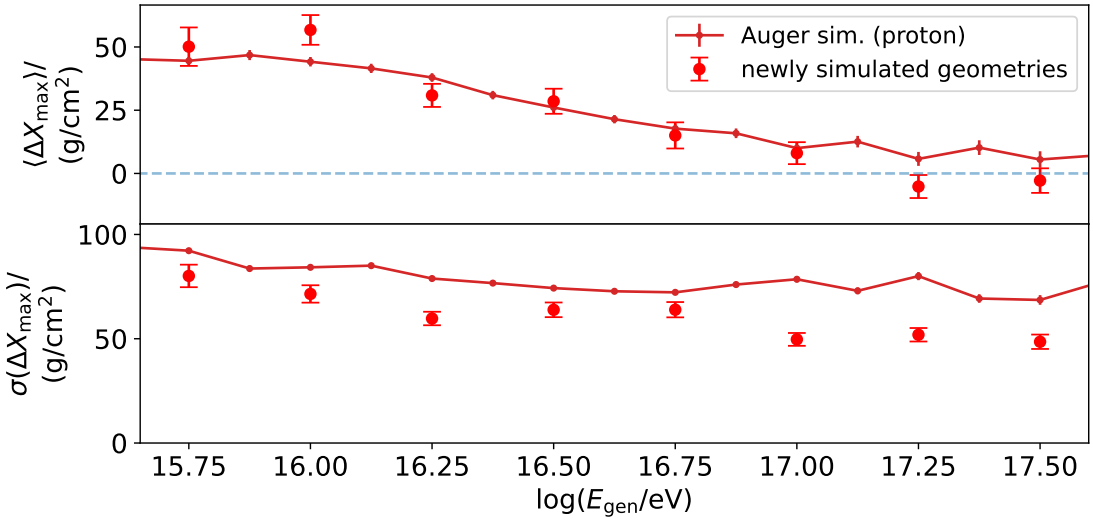

Figure 4.3.: $X_{\max}$ bias and resolution of newly simulated sample geometries from the ICRC21 sample, with the same reconstruction configuration as well as basic selection cuts and quality selection cuts applied.

4.2.1. True generated geometry

True generated shower parameters can be inserted at different points of the reconstruction to assess the implication of a correct measurement of these parameters on the reconstruction result. Specifically, the goal is to see which parameters result in improvements in the $X_{\max}$ bias and resolution when their true values are inserted. The reconstruction with true geometry parameters is performed separately in Coihueco and HEAT for an easier implementation of the true parameters. Events are combined afterwards, with showers recorded in both Coihueco and HEAT only being taken from one of the telescopes. The same is done for the standard PCGF reconstruction to ensure comparability.

Several implementations of true shower parameters are shown in Fig. 4.4. In a first step, the angles θ_{SDP} and ϕ_{SDP} defining the SDP are set to their true values. A standard PCGF is then performed with a χ_0 scan and corresponding timing fits of t_0 and R_{p} as well as profile fits. The result is very similar to the standard reconstruction without true SDP values. This implies that the SDP is usually well reconstructed and has only a minor contribution to the large bias and resolution.

The use of true shower parameters is then expanded into the PCGF by limiting the χ_0 scan to a single value, the true generated angle. A clear improvement in both bias and resolution can be observed, indicating that the PCGF scan result plays an important role in the accuracy of the X_{max} reconstruction result. The two remaining geometry parameters, R_{p} and t_0, which are calculated in the timing fit, can also be set to their true values. In this case, the PCGF is omitted and only the final energy deposit fit is performed. This leads to an additional slight improvement in resolution. Since the greatest improvement in bias and resolution is seen for a correct PCGF scan result, in the following sections an in-depth investigation into the PCGF implementation will be taken.

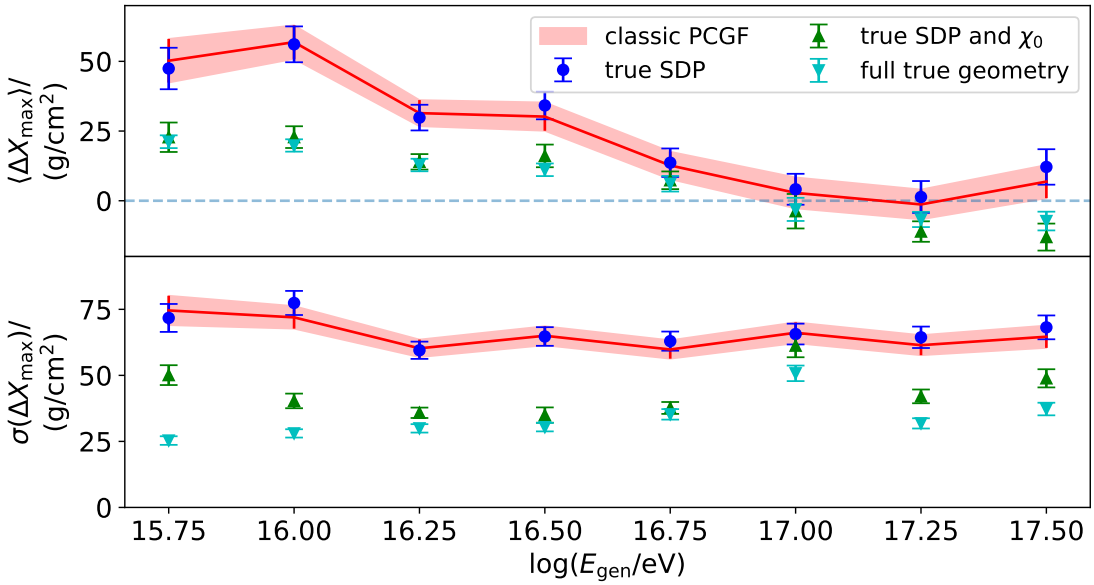

Figure 4.4.: Usage of true generated parameters in comparison to the classic PCGF reconstruction. For true SDP parameters with a standard PCGF, no change in the X_{max} bias and resolution can be observed. An improvement is only seen if true geometry parameters are used as a result of the PCGF and timing fit. The showers are reconstructed separately in Coihueco and HEAT, and combined afterwards.

4.2.2. Changes to Gaisser-Hillas shape constraints

In the PCGF the geometry parameters are obtained from the minimum combined χ^2_{tot} given by Eq. (3.4). Several possibilities are explored to improve this technique. The contribution

χ^2_{time} from the timing fit to Eq. (3.1) changes very little during the course of the χ_0 scan, previously shown for an example in Fig. 3.6. No alteration in the average result is observed in the case of its complete exclusion from the scan (Fig. 4.5). The shape constraints defined by Eq. (3.5) for the parameters λ, X_0 and k in the classic parametrisation, as well as L and R in the USP parametrisation, have a much bigger impact on χ^2_{tot}. By greatly increasing the standard deviation $\sigma(Y)$ of a given shape parameter, the shape constraints in the profile fit are effectively removed. A reconstruction without any shape constraints results in a worse resolution by $\approx 50\,\text{g/cm}^2$ and a bias with a higher energy dependence, as seen in Fig. 4.5. For the classic PCGF, either only using the universality constraint on k or only the two shape constraints for λ and X_0 yields similar results to the standard reconstruction with all three shape constraints.

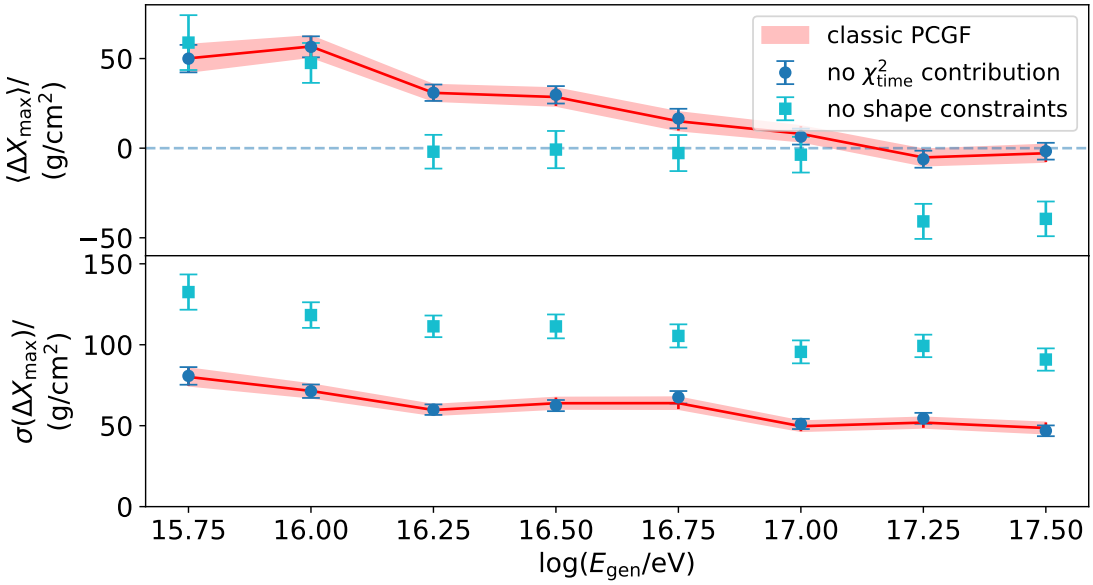

Figure 4.5.: A χ_0 scan without the χ^2_{time} contribution exhibits no variation in the results when compared to the standard PCGF configuration. A reconstruction without shape constraints in the profile fit results in a worse resolution and shifted bias.

By gradually adjusting the standard deviation of the shape parameters in the PCGF configuration, the strength of individual or multiple constraints was varied according to Eq. (3.5). However, neither of these changes resulted in a difference in the bias and resolution. By changing the mean values of the constraints, the bias can be slightly shifted up or down, but its energy dependence remains. Additionally, the mean true values of the shape parameters derived later in this section were not found to differ significantly from the ones implemented in the PCGF configuration. In the PCGF technique used by TA, the shape parameters are fixed to values of $\lambda = 70\,\text{g/cm}^2$ and $X_0 = 40\,\text{g/cm}^2$ [41], instead of $\langle\lambda\rangle = 61\,\text{g/cm}^2$ and $\langle X_0\rangle = -121\,\text{g/cm}^2$ used for the shape constraints in Auger. The TA configuration with fixed shape parameters and consequently no shape constraints was tested as well and did not lead to any significant change of the result.

The USP parametrisation performs similarly to the classic PCGF, as seen in Fig. 4.6, and is equally robust to changes in the shape constraints. Another approach of using true generated values is also applied for the USP parametrisation. For the showers generated with CONEX that were used to re-simulate geometries from the ICRC2021 simulations as described in Section 4.1, the true USP shape parameters are derived. This is done by performing a GH fit to the generated energy deposit profile of each CONEX shower. The fitted "true" values for L and R can then be used in the reconstruction with the PCGF. For each shower, the USP shape parameters were fixed, removing the shape constraints and providing the correct shape in the GH fit, thereby leaving only X_{max} and N_{max} as free parameters. Even in this best-case scenario for the shape parameters, the bias and resolution are not improved as seen in Fig. 4.6.

To conclude, the current PCGF implementation shows no potential improvement regarding X_{max} for any of the described changes of the shape parameters. As a further check, the precision of the PCGF module and implementations of important shower characteristics are investigated in the following sections.

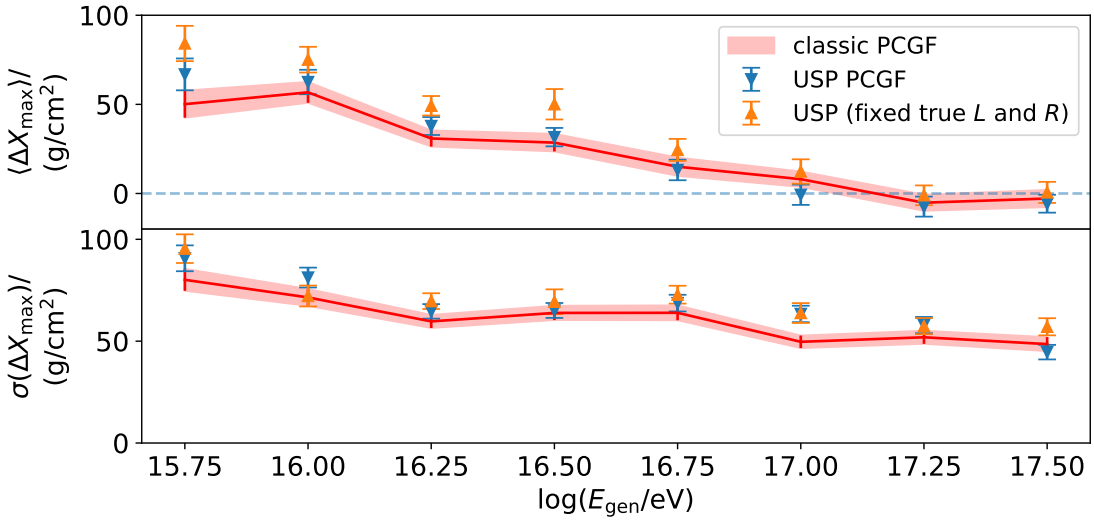

Figure 4.6.: Using the USP parametrisation leads to similar results compared to the classic parametrisation. Fixing the USP shape parameters to their true values for each shower yields no improvement in the reconstruction performance.

4.2.3. χ_0 scan with "brute-force" calculations

For the standard PCGF configuration, several shortcuts are taken to optimise the runtime. The light at the aperture is calculated before a χ_0 scan and used for all χ_0 values and corresponding fits. For the energy deposit profile fits, the CFM is calculated with a smaller number of elements and without doing pre-fits, resulting in a less detailed calculation compared to fits performed at the end of a monocular reconstruction for the calorimetric energy. These shortcuts that are normally used to speed up the calculation are removed for

potentially more precise results. Technically, this means that the PCGF module is removed from the Offline configuration. Instead, a complete monocular HEAT reconstruction is performed for each value of χ_0 by fixing it to the current scan value in the reconstruction of the shower axis. After the shower axis reconstruction for each scan value is completed, a detailed energy deposit fit including pre-fits is performed. A combined χ^2_{tot} with the typical contributions from the profile and the timing fit is produced. The best resulting reconstruction is then chosen by the minimum χ^2_{tot} as in the PCGF. The manual χ_0 scan is performed in 0.1° steps in an interval of ±5° around the generated value of χ_0 to reduce runtime. If the minimum χ^2_{tot} value lies in close proximity to the boundaries of this interval, the shower is discarded. However, since the average χ_0 bias is below 0.5° even at low energies with a comparatively X_{max} bias, this only affects a small fraction of events.

The result of this "brute-force" reconstruction for the sample of simulations introduced in Section 4.1, in comparison to a standard PCGF reconstruction in HEAT, is shown in Fig. 4.7. Due to the separate aperture light calculation for each χ_0 value, the number of degrees of freedom in the profile fit can vary across the scan, leading to jumps in the χ^2_{tot} profile and fewer showers with a smooth χ^2_{tot} minimum. This results in larger uncertainties for the X_{max} bias and resolution despite the more detailed reconstruction. No improvement in the bias and resolution is observed, indicating that the simplified calculations for the PCGF described at the beginning of this section are valid and do not negatively impact the reconstruction outcomes.

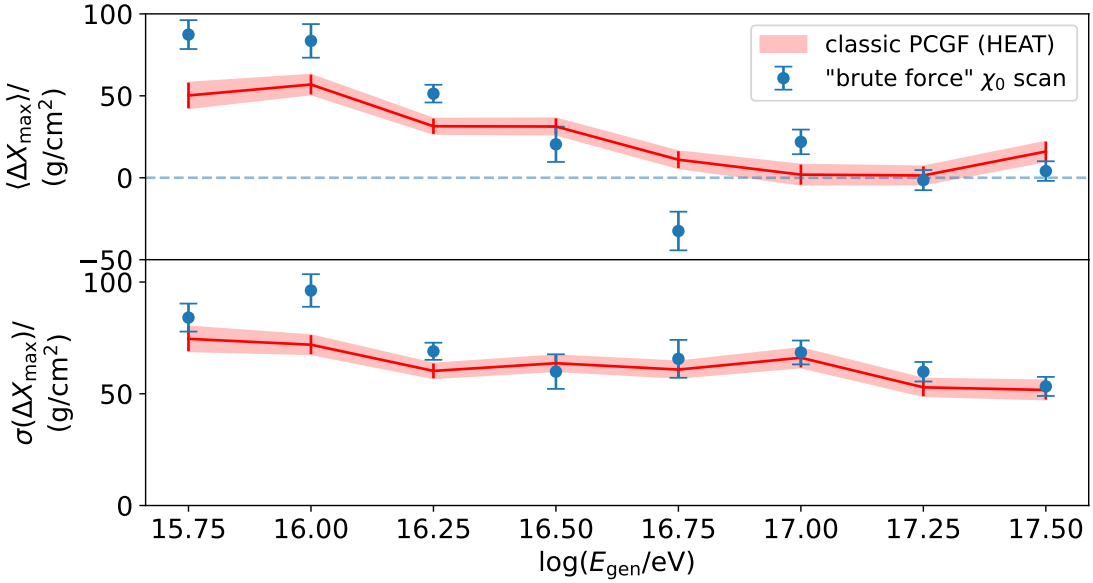

Figure 4.7.: Reconstruction with a manual χ_0 scan and detailed profile fit calculations compared to the standard PCGF reconstruction. The results are comparable

4.2.4. Changes to lateral distribution of Cherenkov light and antialiasing filter correction

A possible source of bias could be the implementation of shower characteristics in the simulation and reconstruction. One of these is the parametrisation of the Cherenkov light Lateral Distribution Function (LDF). When simulating with Offline using CONEX showers, the generation of Cherenkov photons in respect to the shower axis is based on parametrisations by Nerling et al. [29]. The lateral distribution of photons is implemented with a Gora function, which was originally derived for fluorescence light [54]. This implementation of the LDF can be deactivated in the simulation as well as the reconstruction, treating all Cherenkov photons as if they come directly from the shower axis. This leads to an improvement mainly in the $X_{\max}$ bias, as seen in Fig. 4.8. The effect mostly comes from the LDF for direct Cherenkov light, as it is not observed when only the scattered Cherenkov LDF is switched off.

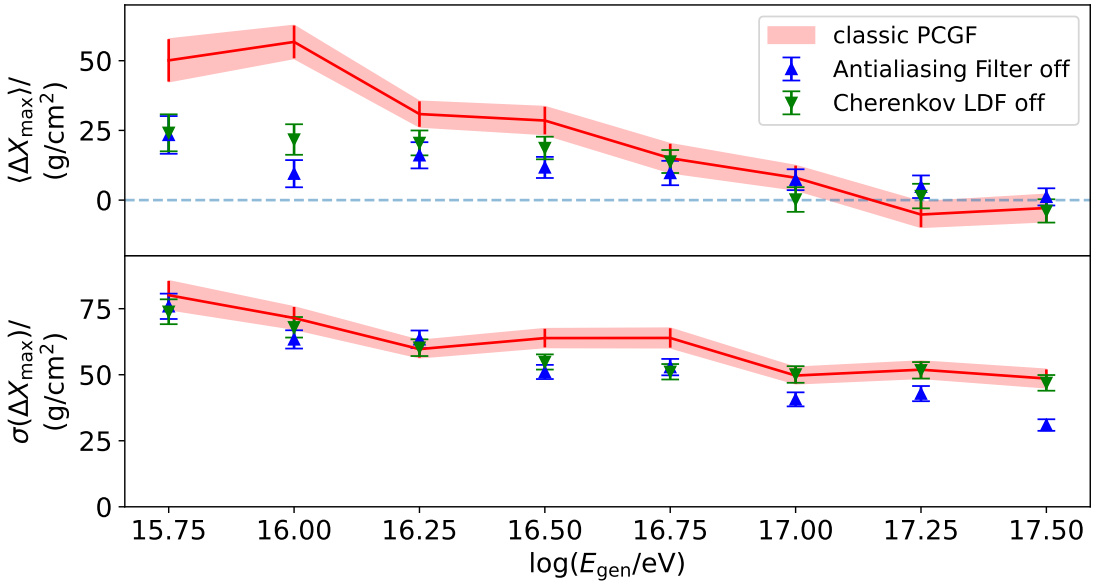

Figure 4.8.: Reconstruction of showers simulated without a Cherenkov LDF and without the antialiasing filter correction compared to the standard procedure.

Another change in the configuration that has a similar effect, also shown in Fig. 4.8, concerns the antialiasing filter. Disabling the application of the antialiasing filter for the simulations and the corresponding antialiasing filter correction in the reconstruction also leads to a reduced bias of $\langle \Delta X_{\max} \rangle \leq 25\,\mathrm{g/cm^2}$ in the observed energy range, while the resolution is similar to the standard reconstruction. Comparable improvement is observed if the antialiasing filter application is only disabled in the simulation, but not if the correction is only disabled in the reconstruction.

For both the reconstruction without a Cherenkov LDF as well as without the antialiasing filter, the use of the complete true generated geometry leads to an improved bias of $\langle \Delta X_{\max} \rangle \leq 10\,\mathrm{g/cm^2}$ compared to the standard reconstruction with true geometry

parameters with $\langle \Delta X_{\mathrm{max}} \rangle \leq 25\,\mathrm{g/cm^2}$ (previously shown in Fig. 4.4). If both the LDF and the filter are switched off simultaneously, the result is similar to only one of the effects being disabled. This indicates that turning the effects off does not cause an additive shift in bias and instead a tangible improvement. Only one time bin on average is used additionally if the effects are disabled, implying that the improvement is not simply the result of more light being collected for the reconstruction. Instead, the reason could be the implementation of the Cherenkov LDF and the antialiasing filter in the Offline framework. However, it is beyond the scope of this work to analyse the implementation in detail.

All the different tests described so far in this Chapter are summarised in Tab. 4.3. In conclusion, most of the tested configurations do not change the outcome of the X_{max} bias and resolution significantly. Using the generated geometry parameters, the largest improvement is seen if the angle χ_0 is set to its true value. However, changes in the profile shape constraints as well as expanded calculations do not lead to improvements of the PCGF result. The bias and resolution reported by TA using their detector TALE could not be reproduced for simulations in Auger with reconstruction in HeCo. An improvement in bias is observed for disabling the Cherenkov light LDF and the antialiasing filter application in the simulations and reconstruction, however, the exact reason for this improvement is unclear.

Table 4.3.: Overview of changes of the reconstruction tested in this chapter. Approximate values at $10^{16.25}$ eV are given if bias or resolution do not agree with the standard reconstruction within the statistical uncertainties.

description of changes	effect on bias [g/cm^2]	effect on resolution [g/cm^2]
true SDP	none	none
true SDP, χ_0	-18	-24
true SDP, χ_0, R_{p}, T_0	-19	-30
no χ^2_{time}	none	none
no shape constraints	-33	+52
TA shape parameters	none	none
USP parametrisation	none	none
true USP shape	+18	+10
"brute force" calculations	+20	+9
Antialiasing filter off	-15	none
Cherenkov LDF off	-10	none

4.3. Comparison with CORSIKA simulations

An existing library of Cherenkov-light dominated showers (see Ref. [55]) generated with CORSIKA is used for a comparison with CONEX events in this section. The CORSIKA showers have generated zenith angles below 60° in an energy range of $10^{15.2}$ eV to 10^{17} eV. Generated shower geometries were selected from CONEX showers with a high Cherenkov fraction and adjusted to ensure a sufficient detection efficiency in HEAT for the CORSIKA events [55]. Only proton showers are available. Due to the preselected nature of the

CORSIKA showers, simply placing cuts for the used energy and zenith angle range on the ICRC21 CONEX showers is not sufficient to align the two samples on key shower parameters. Therefore, another method is chosen for the comparison. The CORSIKA showers are simulated in two ways using Offline. In the first case, they are simulated with the regular CORSIKA configuration. In the second case, the CONEX configuration is used, simulating CORSIKA showers like CONEX showers with the corresponding parametrisation for Cherenkov light. The reconstruction is identical for both cases. The X_{max} bias and resolution after application of TA cuts for the two simulation methods are shown in Fig. 4.9. Between 10^{16} eV and 10^{17} eV the bias and resolution of the CORSIKA showers are on average 10.3 g/cm^2 and 7.1 g/cm^2 higher than for the CONEX-like showers, respectively. The difference could be a result of the different parametrisation of Cherenkov light in the two simulation methods. In CORSIKA, individual Cherenkov photons are generated, while CONEX relies on the parametrisation of the longitudinal profile of charged particles. CORSIKA simulations are presumed to describe real data more accurately than CONEX simulations. However, as described above, the existing CORSIKA showers are only available for protons and in the preselected form. Therefore, the difference in bias is included as a systematic error in the following mass composition analysis using the ICRC21 CONEX showers in Chapter 5.

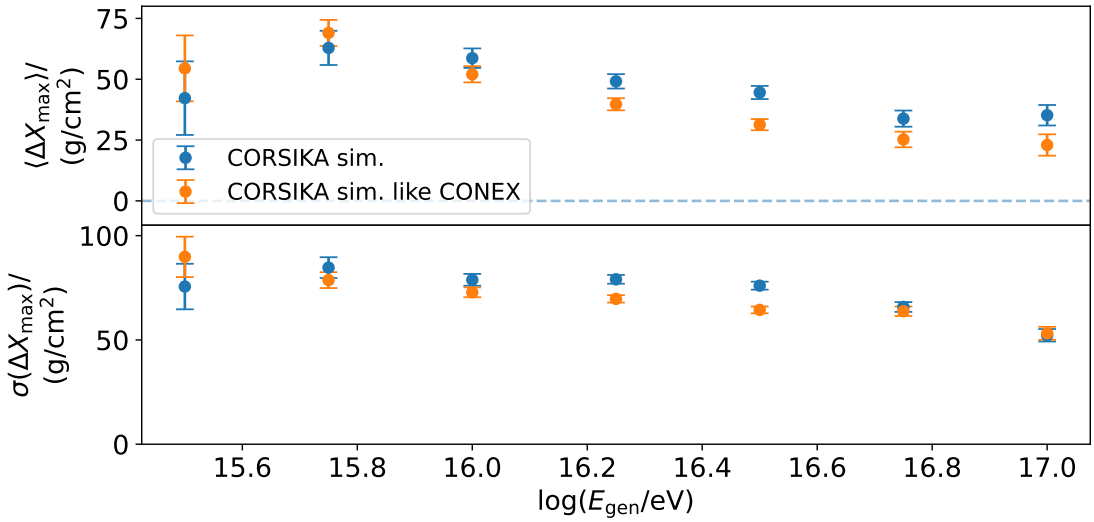

Figure 4.9.: X_{max} bias and resolution of CORSIKA showers compared to the same showers simulated and reconstructed with a CONEX configuration. A difference in bias and resolution between 10^{16} eV and 10^{17} eV is observed.

5. Mass Composition Analysis

In this chapter the mass composition analysis with data recorded by the HEAT and Coihueco telescopes is peformed. The agreement between simulations and data as well as the stability of the results are checked. Finally, the results are compared to the GST model of the cosmic ray composition, which was introduced in Section 2.1. The calorimetric energy range for the mass composition analysis is chosen between 10^{16} eV and 10^{17} eV. At energies below this range, the X_{max} resolution is worsening for all primary elements. This is because fewer light is measured per shower, making a mass separation more difficult. At higher energies, Cherenkov-light dominated events play a less important role, as the information from fluorescence light becomes dominant.

5.1. Weighting of simulated events

The data used for the mass composition analysis were measured by HEAT and Coihueco and reconstructed within the virtual HeCo system between 2012 and 2017. Events are subjected to the FD trigger system described in Section 3.2.1. All T3 accepted HEAT and Coihueco events combined into HeCo are included in the given time range, with 80.4 % of successfully reconstructed events classified by T3 as close showers, 18.6 % as shower candidates and only 1.0 % as horizontal showers. In addition, 10 % of T3 rejected events are used because they often include Cherenkov-dominated events [43], but only T3 rejected showers reconstructed in HEAT are included in the used dataset. They are mostly classified by T3 as muons, noise, or a combination thereof. Lastly, the sample also contains 1 % of TLT rejected HEAT events.

The RealMC simulation reconstructed in HeCo that were introduced in Section 4.1 are also part of this analysis. They include all TLT accepted events. For that reason, real events that are TLT rejected are not included in this analysis. They are identified by their TLT label that is applied before passing Offline. TLT rejected events constitute a mere 0.5 % of the entire dataset. Additionally, the simulations are weighted to reflect the usage of T3 accepted and rejected events in the data. Specifically, T3 rejected pure Coihueco events are excluded and T3 rejected events that have the telescope with maximum triggered pixels in HEAT receive a weight of 0.1. In this way, T3 rejected mirror crossings between HEAT and Coihueco are only kept in the simulations if they are mainly recorded in HEAT, while purely HEAT reconstructed T3 rejected events are stored in the data. This discrepancy is expected to have a negligible effect as the amount of mirror crossings is small at all energies, as will be shown in Fig. 5.3.

An additional weight is applied to the simulations based on their generation in regards to energy. The ICRC21 events were generated with a spectral index of 2.2. An energy weight is introduced to adjust the distribution to realistic spectral indices derived from

measurements. The energy value 2.8×10^{16} eV for the low energy ankle as well as the spectral indices γ below and above it are taken from Ref. [10]. The energy weight is then calculated as

$$w(E_{\text{gen}}) = \left(\frac{E_{\text{gen}}}{2.8 \times 10^{16}\,\text{eV}} \right)^{-(\gamma-2.2)}, \gamma = \begin{cases} 3.09, & E_{\text{gen}} < 2.8 \times 10^{16}\,\text{eV} \\ 2.85, & \text{otherwise} \end{cases} \tag{5.1}$$

5.2. Quality cuts

For the basic ADST selection, real HeCo events undergo the same cuts as the simulations in Chapter 4, with cut values and efficiencies shown in Tab. 4.1. In total, there are 1044770 data events after the application of these cuts, excluding TLT rejected events. While TA quality cuts were used to enable a comparison to TALE simulations in earlier analysis, new quality cuts were developed for this analysis that are more suited for the particular sample. The TA cuts were used as a basis, with individual cuts being examined for their effects specifically on the X_{max} bias and resolution in the simulations. This was done either by applying only single quality cuts on top of the basic selection or by applying them in addition to the other TA cuts, with several different cut values being tested in each case. Additional cut parameters were also looked into, including the number of profile fit degrees of freedom N_{dof} and a requirement for the depth of shower maximum to be measured in the Field Of View (FOV) or in the expected FOV. The cuts were modified from the TA cuts to a new configuration by iterative tests of single cuts or additions to existing cuts. TA cuts with little effect on the sample were discarded, such as cuts on the profile and timing χ^2 contributions. The zenith angle cut was avoided to enable a stability analysis with the whole zenith angle range at a later stage. The X_{max} and E_{cal} error cuts were disregarded, not only due to the incomplete calculation laid out in Appendix A.2, but also because other cuts could easily be substituted for them. The resulting quality cut variables are

- the Cherenkov light fraction to maintain Cherenkov-light dominated events especially at higher energies,
- a requirement for X_{max} to be measured in the FOV,
- the angular track length to have sufficient directional information about events,
- the number of degrees of freedom N_{dof} which includes used time bins and corresponding slant depth bins to ensure that enough timing information about showers is available especially at lower energies,
- X_{max} to remove outliers at the lower and upper ends of the measured range. The cut is less stringent compared to the TA cuts.

At lower energies the most important cut is the N_{dof} cut. Its effect on the X_{max} bias and resolution between 10^{16} eV and $10^{16.25}$ eV for different cut values is shown in Fig. 5.1. The cut value $N_{\text{dof}}^{\text{cut}} = 6$ was chosen after taking the effect on bias, resolution and efficiency

in combination with other cuts into account. Further cut value analyses can be seen in Appendix A.3. Chosen cut values and efficiencies are given in Tab. 5.1. The efficiency in relation to basic ADST selections is overall smaller than for the TA cuts by about 15 to 20 percentage points across the energy range. The efficiencies of the data lie between those of the simulations for protons and iron, as illustrated in Fig. 5.2.

Without quality cuts, the respective fractions of HEAT events, Coihueco events and mirror crossings are similar in the data to the values in the simulations described in Section 4.1. With quality cuts, HEAT showers are almost the only remaining events for energies lower than about $10^{16.5}$ eV (Fig. 5.3). In contrast to the application of TA quality cuts, HEAT events are suppressed to a larger extent at higher energies, leading to a higher fraction of Coihueco events at 10^{17} eV. Overall, the fraction of mirror crossings is smaller in data compared to RealMC events. This effect could be attributed to the inclusion of some T3 rejected HeCo mirror crossings in the simulations, as previously noted in Section 5.1.

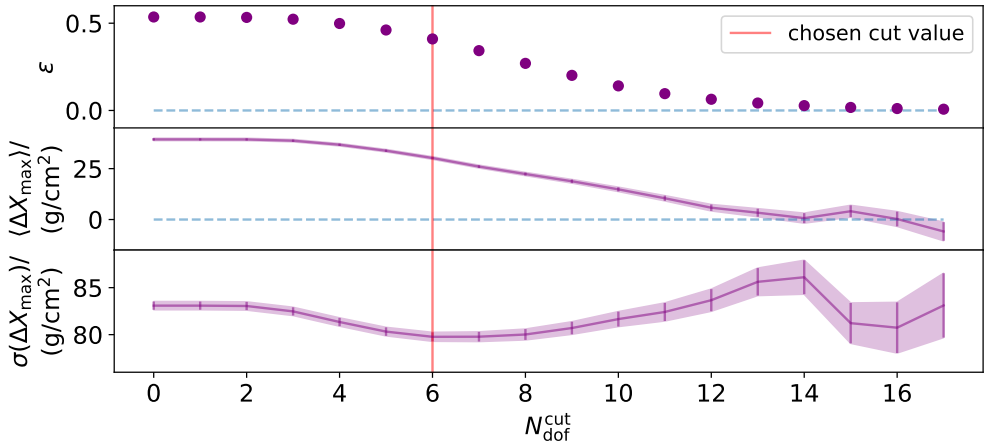

Figure 5.1.: Example of cut optimisation: efficiency ε compared to basic selection cuts as well as X_{max} bias and resolution for different cut values of the profile fit degrees of freedom when applied in addition to the other quality cuts given in Tab. 5.1 in the calorimetric energy bin between 10^{16} eV and $10^{16.25}$ eV. The red vertical line denotes the selected cut value $N_{dof}^{cut} = 6$.

Table 5.1.: Quality cut variables, respective cut values and efficiencies ε when applied to simulations and data with basic selection cuts. The simulations have equal parts of protons, helium, nitrogen and iron.

cut variable	cut value	ε (sim.)	ε (data)
Cherenkov light fraction	$f_{Ch} > 60\,\%$	88.5 %	90.5 %
X_{max} in FOV	-	72.6 %	68.9 %
ang. tracklength	$\alpha_{tr} > 6^\circ$	73.3 %	80.6 %
profile fit deg. of freedom	$N_{dof} > 6$	83.3 %	59.9 %
X_{max}	$300\,\frac{g}{cm^2} < X_{max} < 1000\,\frac{g}{cm^2}$	98.8 %	99.0 %

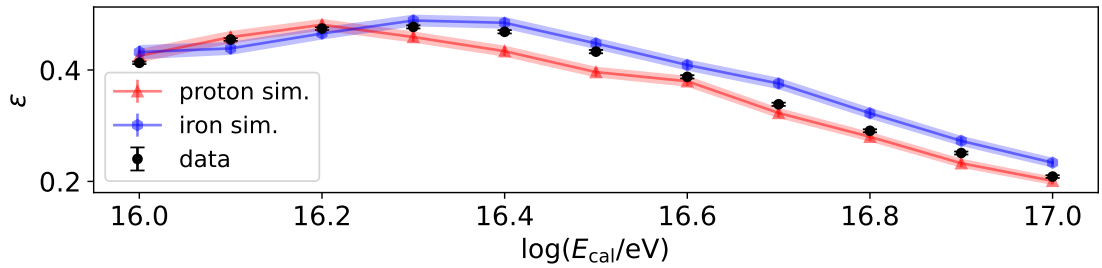

Figure 5.2.: Efficiency of quality cuts compared to events with basic ADST selection. Values for data fall between the proton and iron simulations.

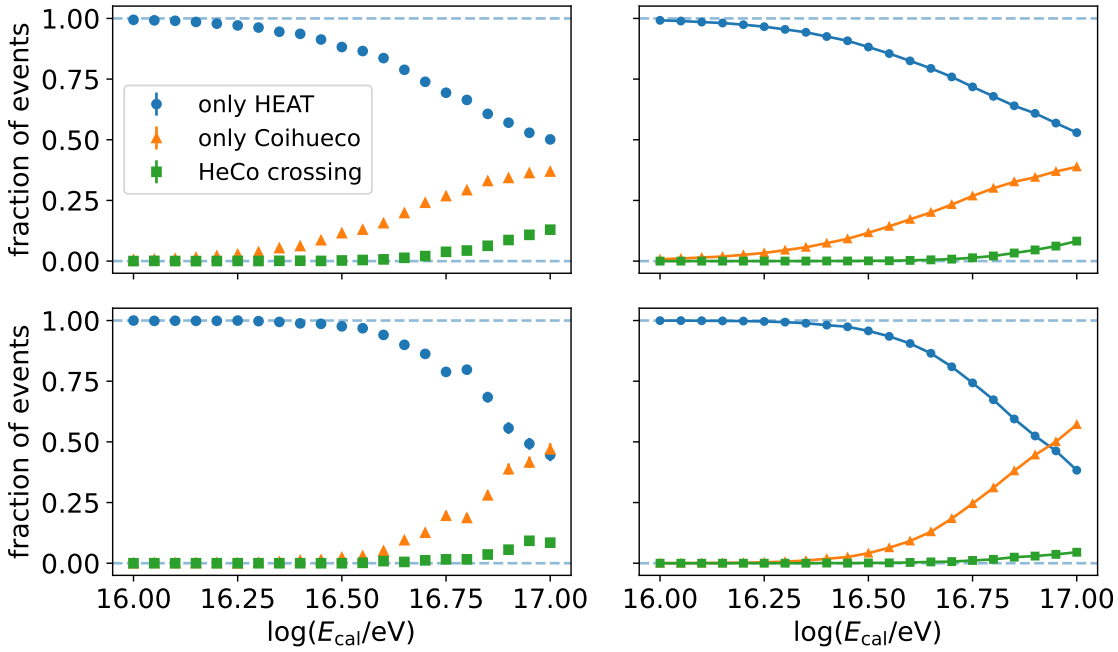

Figure 5.3.: Fractions of events recorded exclusively in HEAT or Coihueco telescopes or as a HeCo mirror-crossing for simulations (left) and data (right) without (top) and with (bottom) quality cuts. Simulations are weighted as described in Section 5.1.

The shower core distribution with applied quality cuts for the lowest calorimetric energy bin is illustrated in Fig. 5.4. Most showers have an impact point in front of the telescopes. For increasing energies, showers with cores farther away from the telescopes are detected. Visible lines in the core distributions come from showers that only have triggered pixels on a horizontal line as variations of the trace type 5_0 in Fig. 3.2. This leads to the events being reconstructed with the same angle ϕ_{SDP}. They are not completely excluded by the quality cuts.

Fig. 5.5 shows the X_{max} bias and resolution in the simulations with the new quality cuts. Compared to TA quality cuts, the energy dependence of the X_{max} bias was reduced, with it taking values between 6 g/cm^2 to 25 g/cm^2 for protons and 25 g/cm^2 to 42 g/cm^2 for iron. The X_{max} resolution remains similar to the one previously shown in Fig. 4.2 with TA cuts, with values between 65 g/cm^2 to 80 g/cm^2 for protons and iron. The new quality cuts as well as all simulation weights are applied in the following sections, unless stated otherwise.

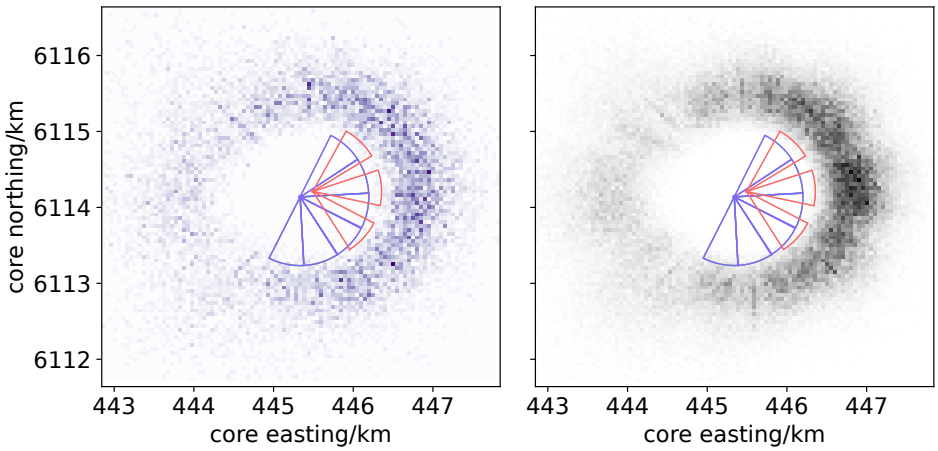

Figure 5.4.: Distributions of the reconstructed shower core for simulations (left) and data (right) with quality cuts applied for calorimetric energies between 10^{16} eV and $10^{16.25}$ eV. The field of view of Coihueco and HEAT telescopes is marked in blue and red, respectively. Lines in the distribution are the result of showers with a horizontal pixel trace sharing the angle ϕ_{SDP}.

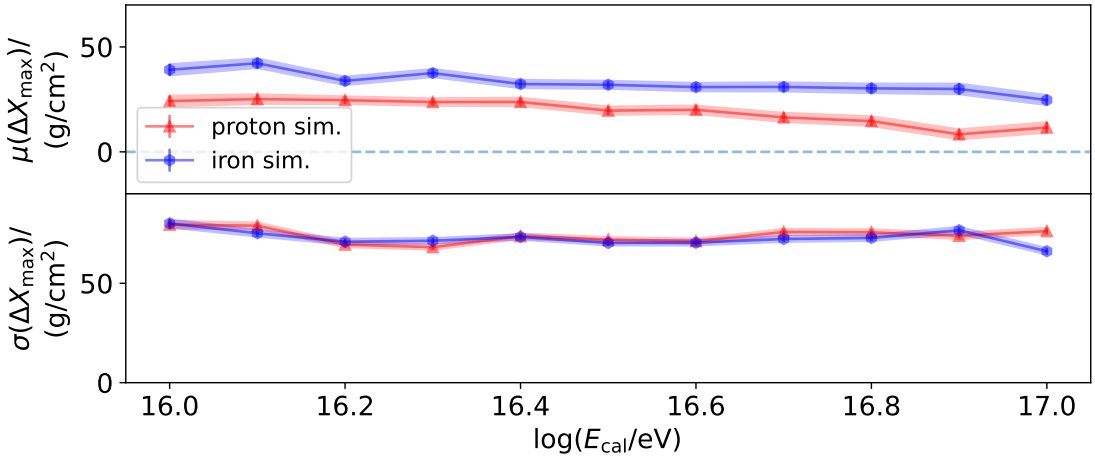

Figure 5.5.: X_{max} bias and resolution of the HeCo simulations with applied quality cuts. Biases and resolutions for helium and nitrogen lie between those for proton and iron and are omitted for better visibility.

5.3. Comparison of data with simulations

Before performing the mass composition analysis, it has to be confirmed that data and simulations align on key shower parameters. If there is good agreement, the bias for the X_{max} measurement can also be assumed to be similar. The first shower parameter that is compared is the reconstructed calorimetric energy E_{cal} itself. The energy spectra with basic selection cuts and quality cuts are shown in Fig. 5.6. The data and simulation energy spectra agree very well in both cases, with the data falling within the uncertainties of the proton and iron MC samples in the complete range of the calorimetric energy observed here.

Contributions of proton and iron are presented in all plots in this Section, as results for simulations of helium and nitrogen invariably fall between them. The statistic uncertainties of the simulations are considerably larger than the ones in the data because of the limited

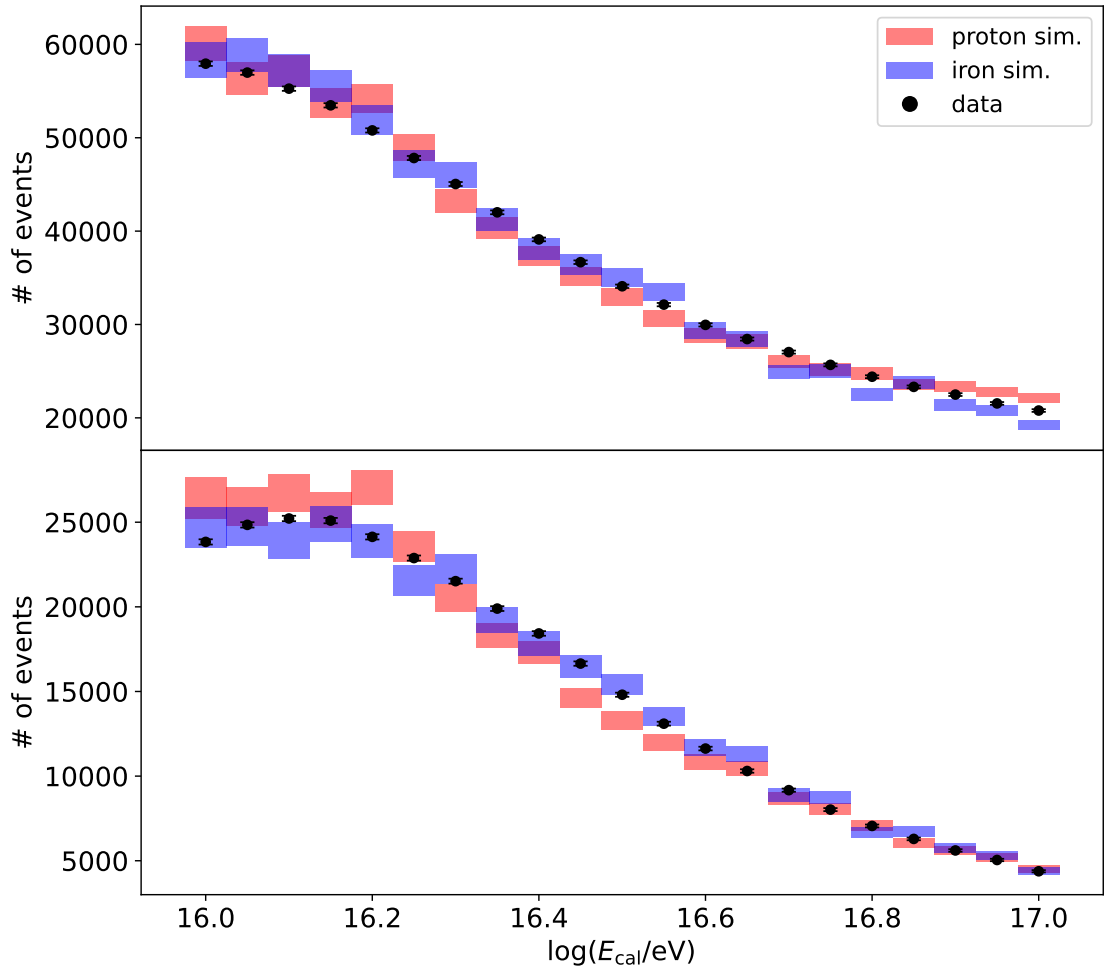

Figure 5.6.: Energy spectrum of data and simulations without (top) and with (bottom) quality cuts show a very good agreement. Simulations are scaled to the number of data events.

size of the sample. All further plots in this Section show the lowest energy bin between $E_{\mathrm{cal}} = 10^{16.0}$ eV and $10^{16.25}$ eV. In addition, a comparison was also conducted in three additional bins up to 10^{17} eV, which yielded similar results. The comparison in the bin between $E_{\mathrm{cal}} = 10^{16.75}$ eV and $10^{17.0}$ eV is shown in Appendix A.4.

Important angles in the shower geometry of the data compared to proton and iron simulations are shown in Fig. 5.7. Similarly to the energy spectrum, most parameters show a very good agreement among data and simulations. Firstly, the zenith angles θ of reconstructed and selected events mainly fall between 30° to 60°. At higher energies, more inclined showers around $\theta = 60°$ are detected. Secondly, most showers are measured from the front of the telescopes with azimuth angles ϕ between ±90°. Thirdly, the X_{max} viewing angle $\alpha_{X_{\mathrm{max}}}$ has a narrow distribution around 10° at lower energies and slightly larger angles around 12° in the higher bins. It is defined as $\alpha_{X_{\mathrm{max}}} = \chi_0 - \chi_{\mathrm{max}}$, where χ_{max} is the angle between the ground and the line of sight between the telescope and the shower maximum in the SDP, analogous to χ_i in Fig. 3.4. Lastly, the distribution of χ_0 has two local maxima above and below 90° respectively in all energy bins.

Further comparisons between data and simulation between $10^{16.0}$ eV and $10^{16.25}$ eV are shown in Fig. 5.8. Firstly, the Cherenkov light fraction f_{Ch} is barely affected by the cut at 60 % in the lowest energy bin, as more than 80 % of events have a fraction of $f_{\mathrm{Ch}} \geq 85\,\%$. At higher energies, events are more evenly distributed. Between $10^{16.75}$ eV and 10^{17} eV, the number of events with a fraction of $f_{\mathrm{Ch}} = 80\,\%$ is only approximately 1.5 times the number of events with $f_{\mathrm{Ch}} = 60\,\%$. Secondly, the number of profile fit degrees of freedom N_{dof} is the only investigated parameter for which the data does not lie between the simulations in the histogram. This is evident in the two lower energy bins. However, there is an overall good agreement in the shape of the N_{dof} distributions, with data being slightly overrepresented at $N_{\mathrm{dof}} \leq 10$ and slighty underrepresented at $N_{\mathrm{dof}} > 12$. Data with N_{dof} below the cut value of $N_{\mathrm{dof}}^{\mathrm{cut}} = 6$ also seem to be overrepresented in comparison to simulations, as the efficiency for the cut (given previously in Tab. 5.1) is more than 20 percentage points lower for the data. However, the overall efficiency with all quality cuts has a good agreement between data and simulations, as previously shown in Fig. 5.2. Lastly, for the shower impact parameter R_{p}, closer showers around 1.5 km are part of the analysis at lower energies and a broader distribution around 1.8 km is observed for higher energies.

With a very good agreement between the data and simulations in almost all investigated shower parameters, the X_{max} distributions are compared in the following section to assess the mass composition of the data.

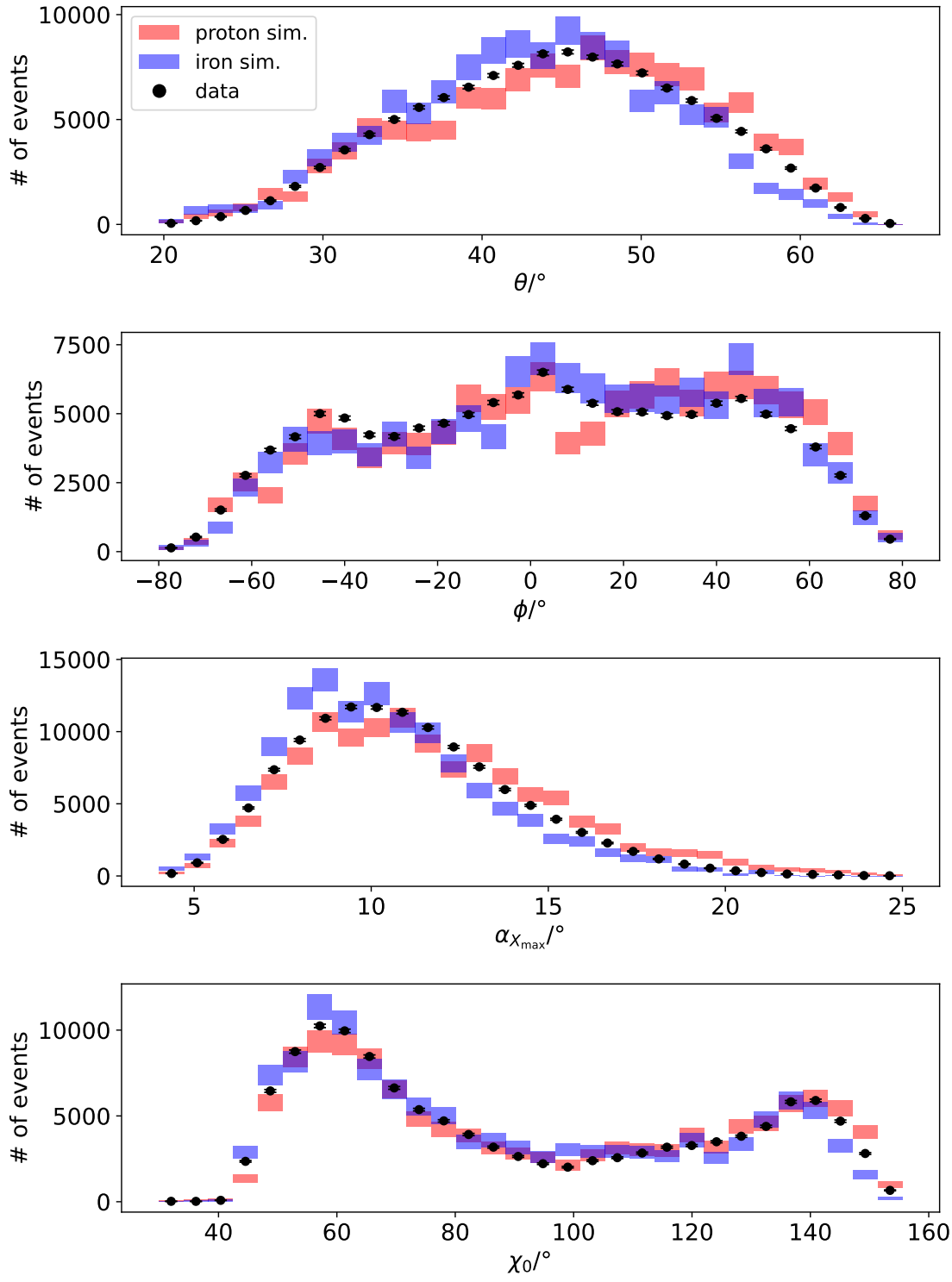

Figure 5.7.: Distributions of events across various parameters: zenith angle θ, azimuth angle ϕ, X_{max} viewing angle $\alpha_{X_{max}}$ and χ_0, for data and simulation with calorimetric energies between $10^{16.0}$ eV and $10^{16.25}$ eV. Simulations are scaled to the number of data events.

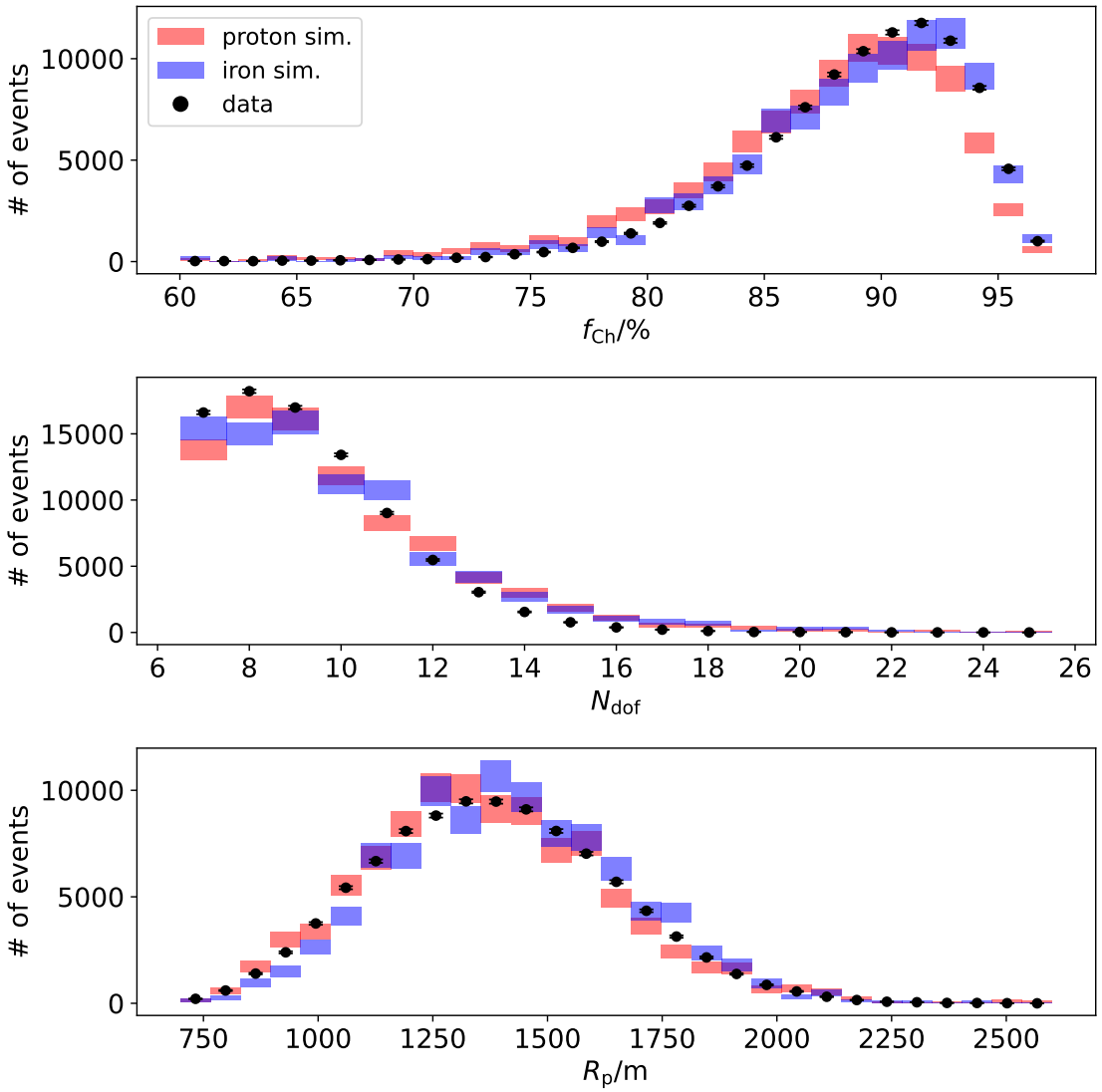

Figure 5.8.: Distributions of events across various parameters: Cherenkov fraction f_{Ch}, profile fit degrees of freedom N_{dof}, and impact parameter R_p for data and simulations with calorimetric energies between $10^{16.0}$ eV and $10^{16.25}$ eV. Simulations are scaled to the number of data events.

5.4. X_{max} Analysis

5.4.1. Mean and standard deviation of X_{max}

The X_{max} difference ΔX_{max} introduced in Chapter 4 compares the reconstructed and generated values of the simulations with quality cuts. The mean generated X_{max}^{gen} of these selected events can differ from the value in the pure simulations. The latter can be derived directly from the complete ICRC21 set of simulated CONEX showers, with fits to the mean generated X_{max} depending on the generated energy. Only a small difference of $\leq 10\,\mathrm{g/cm^2}$

between generated values of the CONEX showers and generated values of the selected events is observed. Thus, the trigger and selection of showers do not introduce a large bias on the mean of the X_{max} distribution. However, as can be seen in Fig. 5.9, the *reconstructed* $\langle X_{\mathrm{max}} \rangle$ differs considerably from the generated one, due to the reconstruction bias discussed in Chapter 4. The bias is larger for the heavier elements, placing the reconstructed values closer together than the generated ones. However, there is still a clear separation of at least $20\,\mathrm{g/cm^2}$ between all elements and of $90\,\mathrm{g/cm^2}$ between protons and iron.

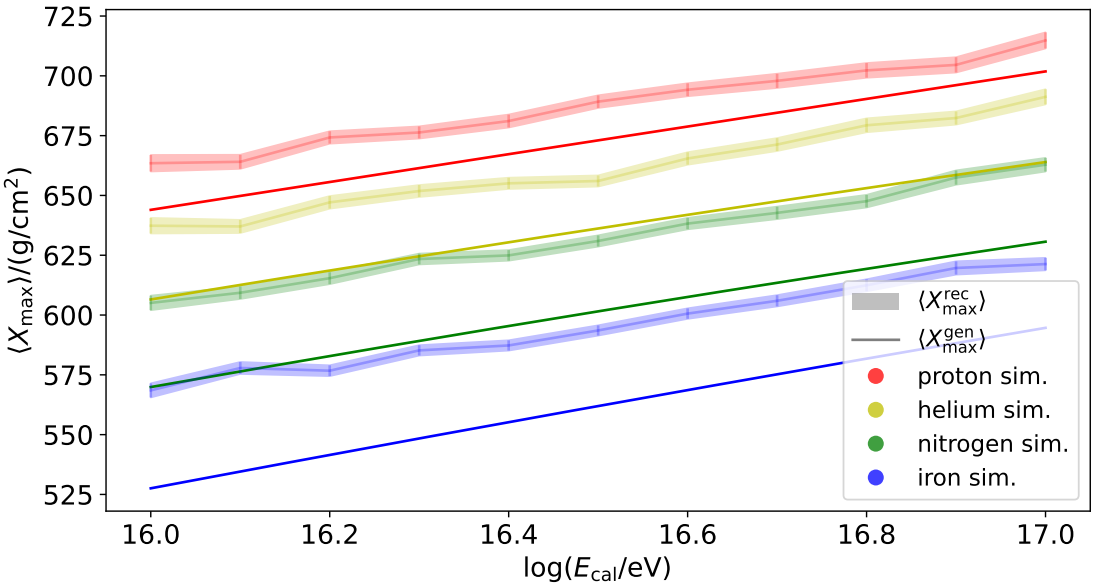

Figure 5.9.: Reconstructed X_{max} of events with quality cuts compared to the generated values obtained from pure simulations. The bias is larger for bigger primary masses. The curves from top to bottom are those of proton, helium, nitrogen, and iron, respectively.

The results of the mean reconstructed depth of shower maximum of data and simulations is visualised in Fig. 5.10. At 10^{16} eV, the mean X_{max} of the data starts at $\approx 620\,\mathrm{g/cm^2}$ and only increases slightly to $\approx 630\,\mathrm{g/cm^2}$ at $10^{16.5}$ eV. In relation to $\langle X_{\mathrm{max}} \rangle$ found for the different primary particles in the simulations, this signifies a transition from a helium-like composition to a nitrogen-like composition. Above this point, from about $10^{16.6}$ eV to 10^{17} eV, the slope of $\langle X_{\mathrm{max}} \rangle$ for data is slightly larger than that of the nitrogen simulations, i.e., the composition becomes slightly lighter again.

Fig. 5.11 shows the corresponding standard deviation of the depth of shower maximum for data and simulations. In contrast to $\langle X_{\mathrm{max}} \rangle$, the results for the individual elements are much closer, partially even with overlap, making it more difficult to draw conclusions about the mass composition. Specifically, the distributions of proton and iron, which have the highest and the lowest values, are only separated by $\approx 20\,\mathrm{g/cm^2}$ across the energy range. The standard deviation of the data is situated in the range of the proton standard deviation, slightly below and above it at energies lower and higher than $10^{16.5}$ eV, respectively. It is important to note that this does not strictly mean that the composition is proton-like -

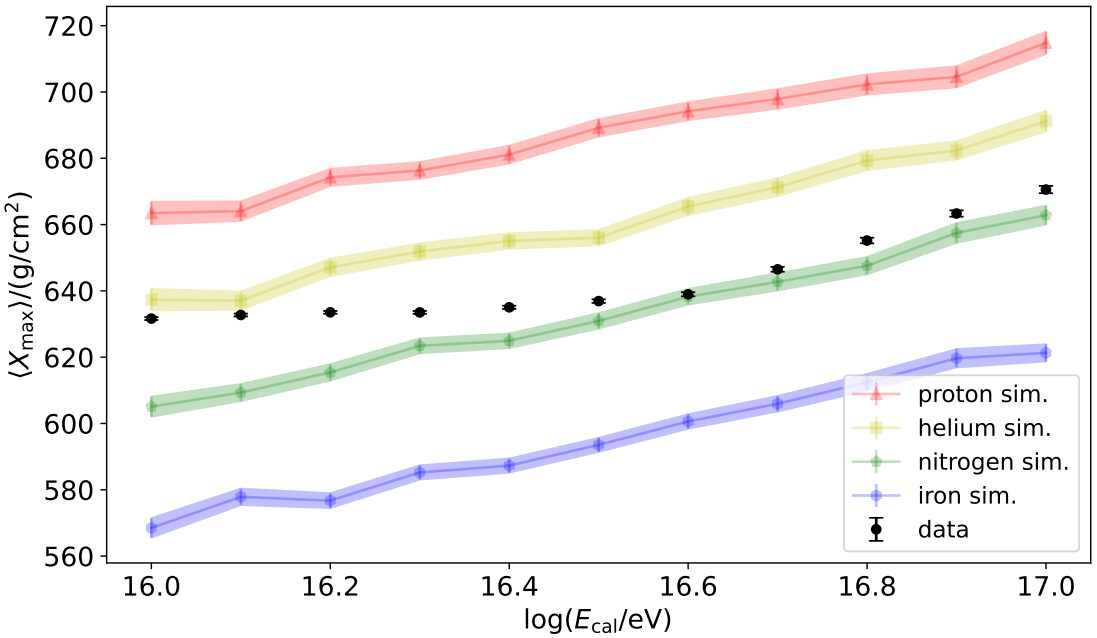

Figure 5.10.: Mean X_{max} of the data compared to simulations. The composition changes from approximately helium-like to approximately nitrogen-like.

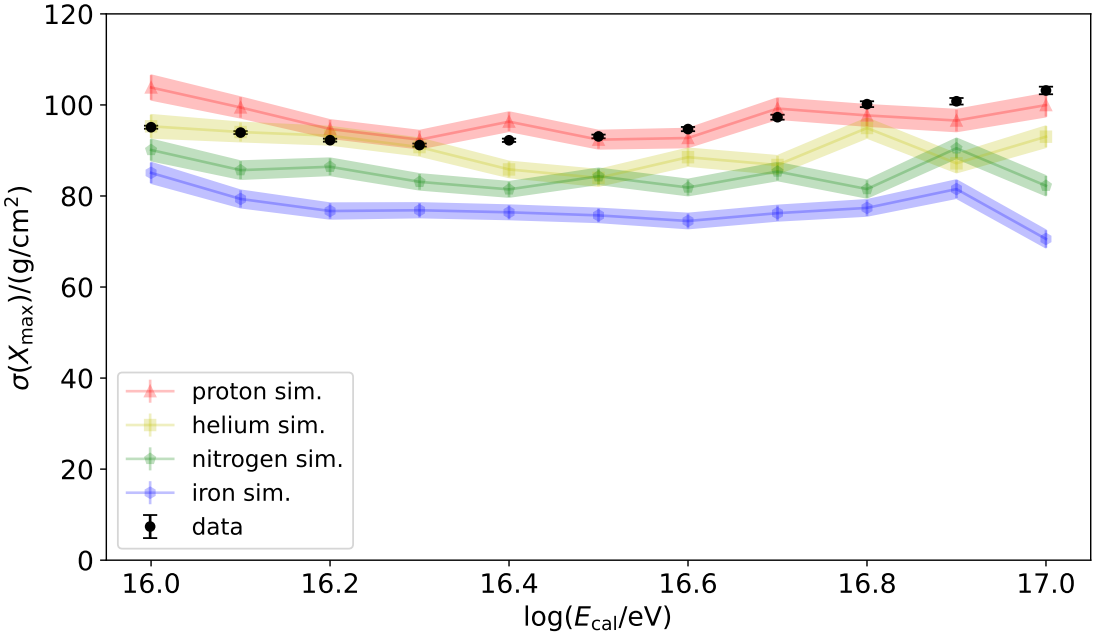

Figure 5.11.: Standard deviation of X_{max}. Small differences and partial overlaps between mass groups make it difficult to interpret the result.

rather, the superposition of elements has an X_{max} standard deviation similar to protons. The small variation and partial overlap in $\sigma(X_{\mathrm{max}})$ between different primary masses is likely a consequence of the poor resolution of the X_{max} reconstruction that was found in Chapter 4. Specifically, the difference between the generated and the reconstructed $\sigma(X_{\mathrm{max}})$ for showers with quality cuts is on average 28 g/cm^2 for protons and 50 g/cm^2 for iron. For the reasons laid out here, the standard deviation of the reconstructed X_{max} is not used further for the mass composition analysis. Instead, the focus lies on $\langle X_{\mathrm{max}} \rangle$.

A way to use the result for the mean depth of shower maximum to quantify the mass of the composition more directly is to calculate the relative location of $\langle X_{\mathrm{max}} \rangle$ compared to the proton and iron simulations. This is conducted with the mass estimation parameter

$$z = \frac{\langle X^{\mathrm{p}}_{\mathrm{max}} \rangle - \langle X_{\mathrm{max}} \rangle}{\langle X^{\mathrm{p}}_{\mathrm{max}} \rangle - \langle X^{\mathrm{Fe}}_{\mathrm{max}} \rangle}. \tag{5.2}$$

It ranges between 0 for a pure proton composition and 1 for a pure iron composition. In the case of equal selection efficiency for all primaries, $z \ln 56$ is equal to the logarithmic mass $\ln A$. However, this does not apply for this sample, as seen in Fig. 5.12. As mentioned in Section 4.1, the simulations used in this analysis are generated with equal parts of the four used primary particle types. With all basic selection cuts and quality selection cuts, the primary fractions are different and dependent on the calorimetric energy. For medium energies between $10^{16.4}$ eV and $10^{16.6}$ eV, all four primary fractions are within two percentage points of each other. At the low and high end of the used energy range, the lighter elements are overrepresented by 1 to 5 percentage points and the heavier elements are underrepresented by the same amount.

Fig. 5.13 shows the calculated parameter z for the dataset according to Eq. (5.2). As described for the $\langle X_{\mathrm{max}} \rangle$ result, the composition becomes heavier and then slightly lighter again, with the slope $D_z := \mathrm{d}z/\mathrm{d}\log E_{\mathrm{cal}}$ changing at a break energy E^*_{cal}. A fit for the

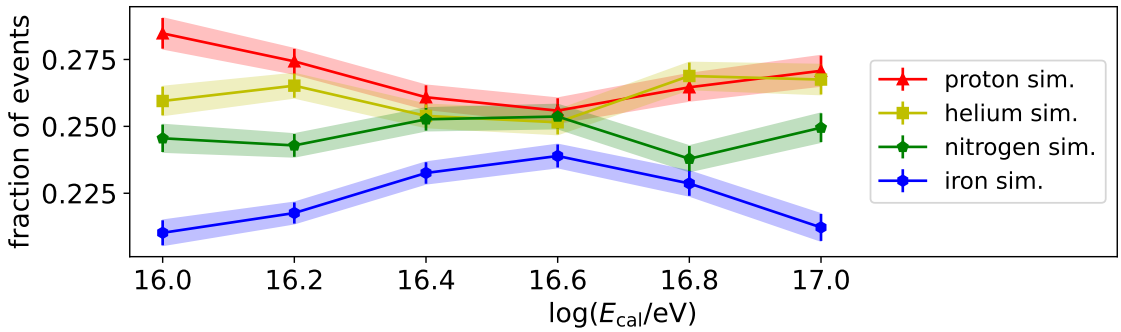

Figure 5.12.: Primary fraction of the four simulated elements with all weights and quality cuts applied. The four primaries are initially simulated with equal fractions of 0.25 (note the zero-suppressed y axis).

calculated z parameter is performed using the function

$$z(E_{cal}) = \begin{cases} z^* + D_z^a \log\left(E_{cal}/E_{cal}^*\right), & E_{cal} < E_{cal}^* \\ z^* + D_z^b \log\left(E_{cal}/E_{cal}^*\right), & \text{otherwise} \end{cases} \tag{5.3}$$

with fit parameters being the break energy E_{cal}^* and the corresponding value $z^* = z(E_{cal}^*)$ as well as the slopes D_z^a and D_z^b below and above the break energy, respectively. The fit results are given in Tab. 5.2 in the following Section, where systematic effects for the result will be derived.

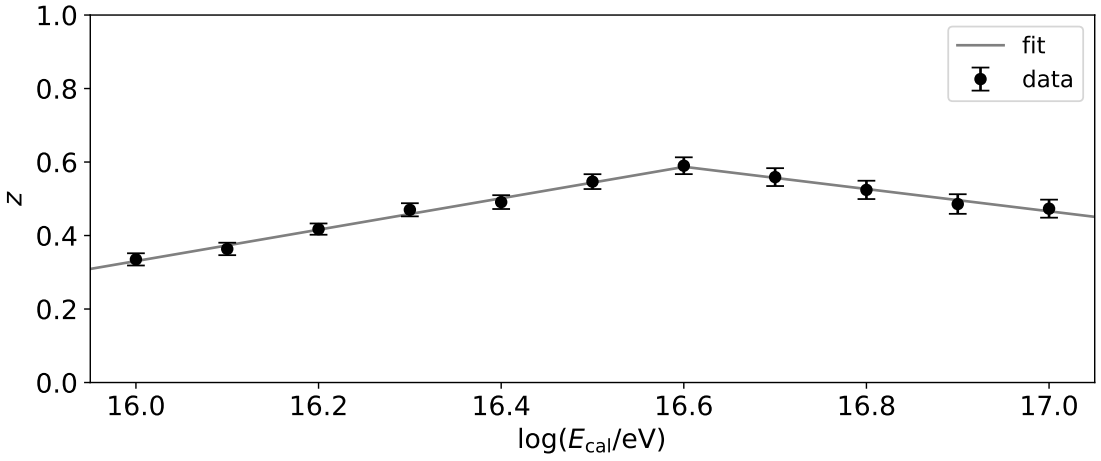

Figure 5.13.: Mass estimator z of the data calculated according to Eq. (5.2). A change in slope of the parameter is marked by a fit following Eq. (5.3).

5.4.2. Subsample analysis

In this section, the stability of the mass composition results is examined using the mass estimator z. For several shower parameters, two subsamples are created for the simulations and data based on a median cut: In each energy bin, the median value of a given parameter is calculated for the data. The data and simulations are then split into two samples, one below this median value and one above it.

The mean depth of shower maximum can experience a substantial shift in both the simulations and the data between the subsamples, as seen in Fig. 5.14 for a split based on the median zenith angle $\tilde{\theta}$. However, for a stable mass composition that is not dependent on the given shower parameter, the relative position of $\langle X_{max} \rangle$ of the data compared to the proton and iron simulations is expected to remain the same in the two samples. This means that the parameter z is expected to be stable when compared between the complete dataset and the two subsamples if the simulation describes the data well.

Fig. 5.15 shows the results of the subsample analysis for the shower geometry angles whose distributions were previously shown in the histograms in Fig. 5.7. A separation of z is seen for the zenith angle across the energy range, more prominent at higher energies.

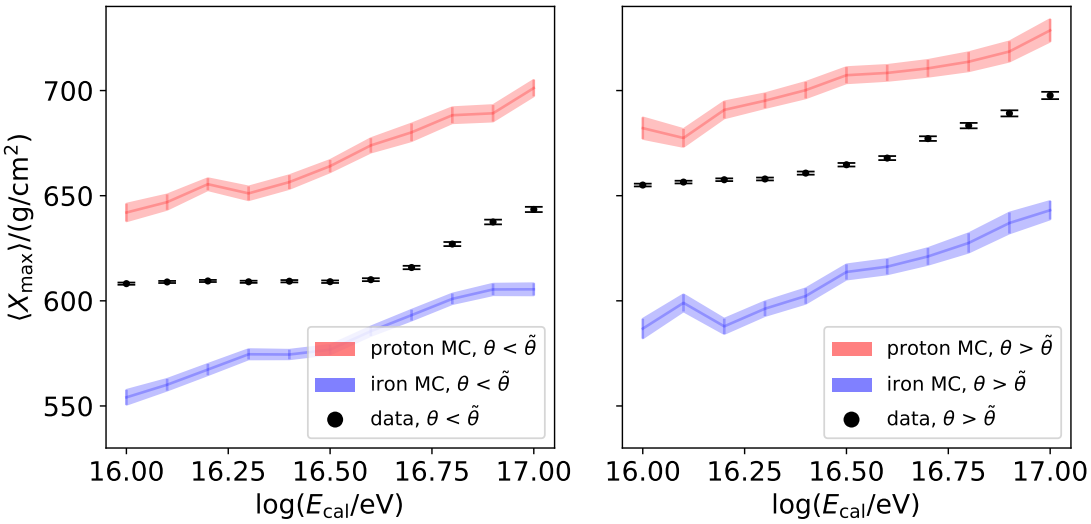

Figure 5.14.: Mean X_{max} in the subsamples below (left) and above (right) the zenith angle median. The whole distribution is shifted upwards above the median cut.

Above the median value $\tilde{\theta}$, which lies between 44° and 57° depending on the energy bin, the composition is lighter, and below the median, the composition is heavier. The separation of z is in the range of $\Delta z = 0.06$ to 0.21.

Other parameters are more stable than the zenith angle. No separation is identified for the azimuth angle. For the X_{max} viewing angle, a small separation is observed at the lower and upper end of the energy range. Similarly, for χ_0, a slight separation occurs at the medium energies between $10^{16.2}$ eV and $10^{16.8}$ eV. The latter two separations are assumed to be correlated with the zenith angle separation. Since they are much smaller with $\Delta z < 0.1$, the zenith angle will be the main focus for further analysis.

A subsample analysis for additional shower parameters is shown in Fig. 5.16. A small separation of $\Delta z < 0.1$ is seen for the Cherenkov light fraction f_{Ch}. For the profile fit degrees of freedom N_{dof} and the shower impact parameter R_p, separations in the range of $\Delta z \approx 0.1$ occur at high energies. Subsamples for the year showers were recorded in are shown and have no separation. Samples based on the season (summer/winter) and time of night (early/late) were also examined and no separation was found.

The largest separation is still observed for the subsample analysis based on the zenith angle. Solely based on the result shown in Fig. 5.15, the separation could be caused by a zenith angle dependent selection, where heavier elements have a lower cut efficiency at higher zenith angles, while lighter elements have a higher efficiency, and vice versa for lower zenith angles. This dependency does exist in the simulations, shown in Fig. 5.17. However, it is not the sole cause for the separation. Instead, it will be shown in Section 5.5 that the separation could be an effect of a difference between simulations and data. For that reason, the energy spectrum is examined above and below the median values, shown in Fig. 5.18. However, neither with nor without quality cuts can a disagreement between data and simulations be seen in the subsamples.

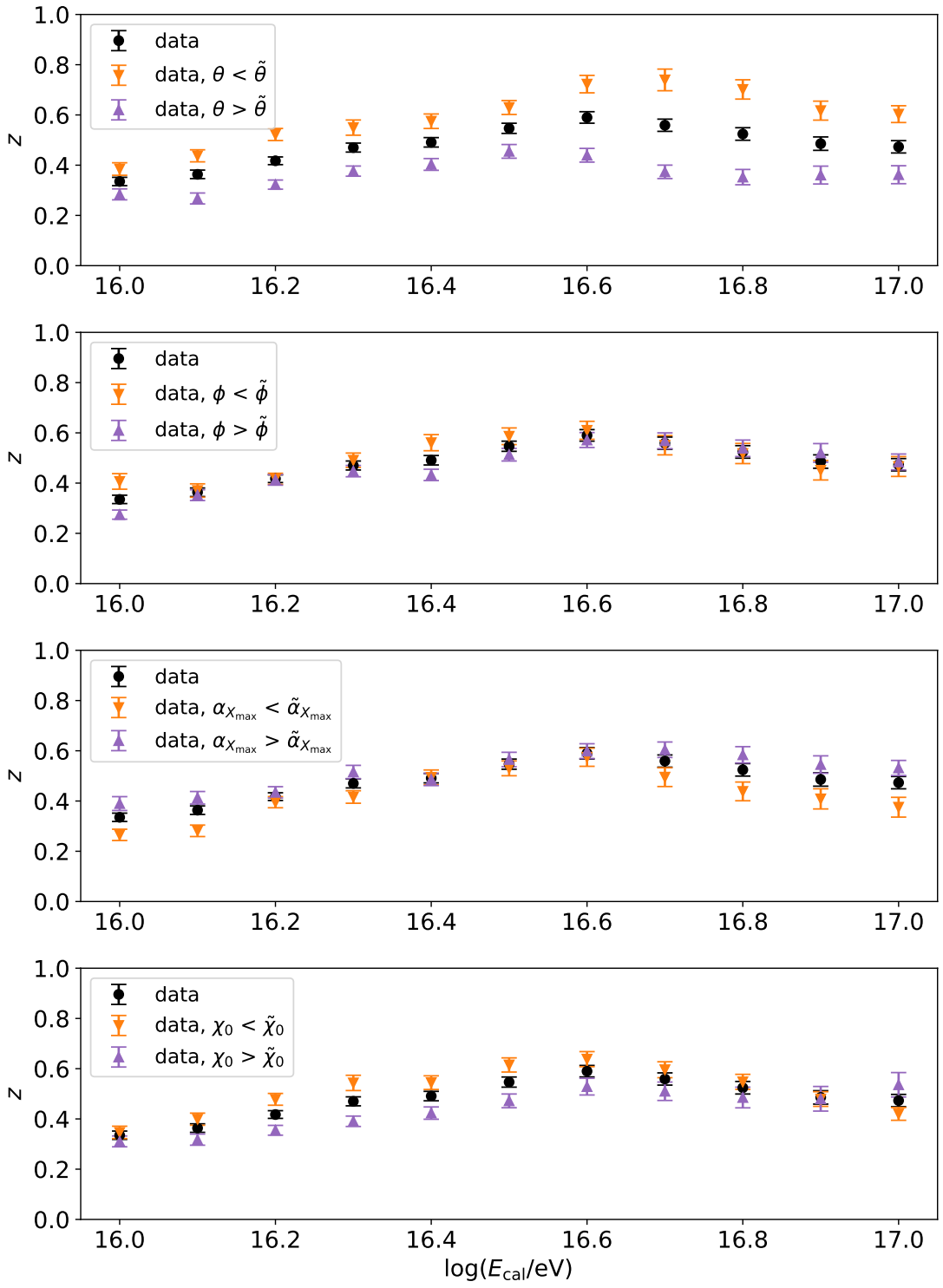

Figure 5.15.: z dependence on median cuts, with median values in data for the zenith angle $\tilde{\theta}$, azimuth angle $\tilde{\phi}$, X_{max} viewing angle $\tilde{\alpha}_{X_{max}}$ and $\tilde{\chi}_0$. A separation is mainly observed for the zenith angle.

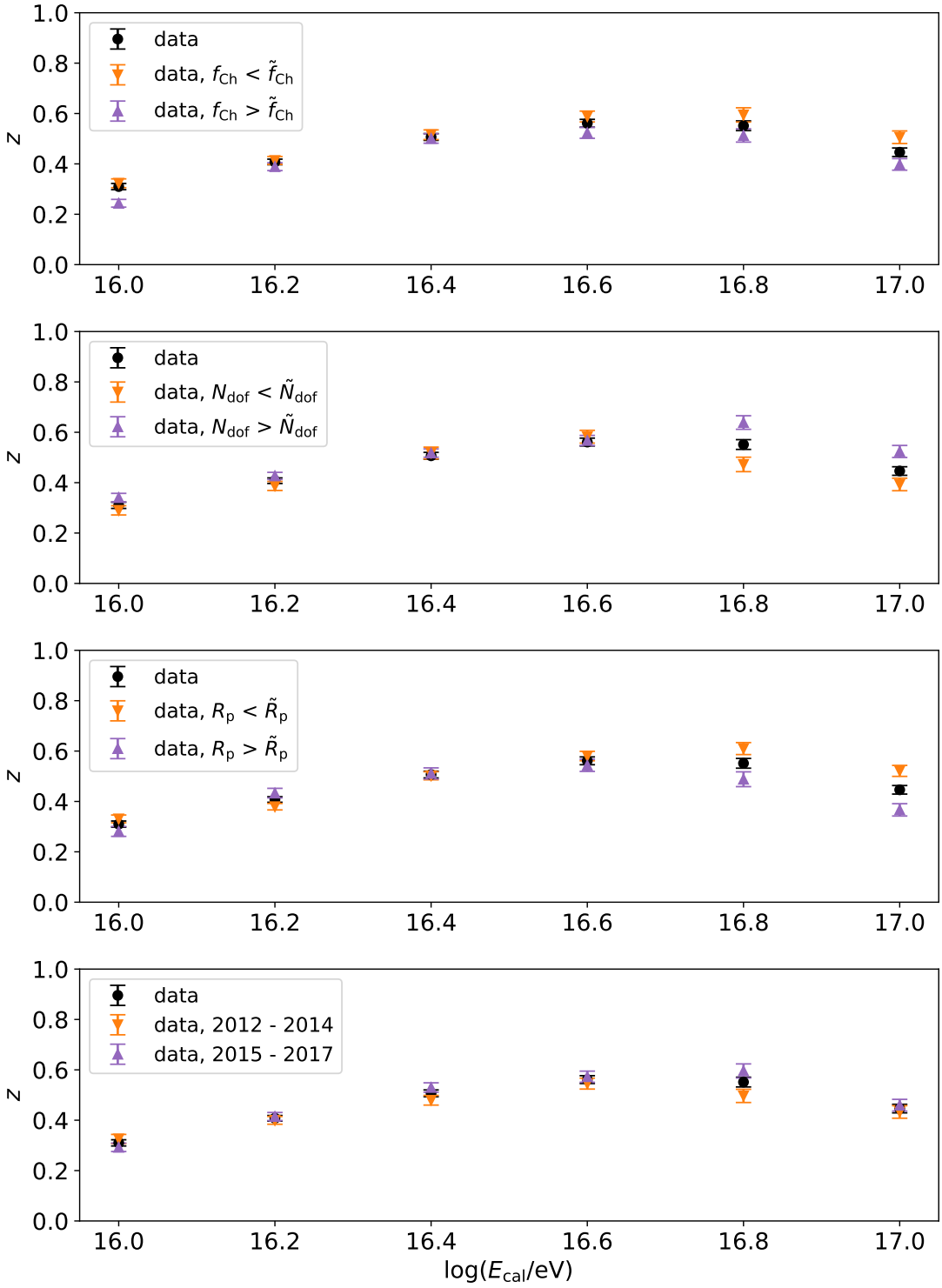

Figure 5.16.: z dependence on median cuts, with median values in data for the Cherenkov fraction $\tilde{f}_{Ch}$, profile fit degrees of freedom $\tilde{N}_{dof}$, impact parameter $\tilde{R}_p$ and year that showers were recorded.

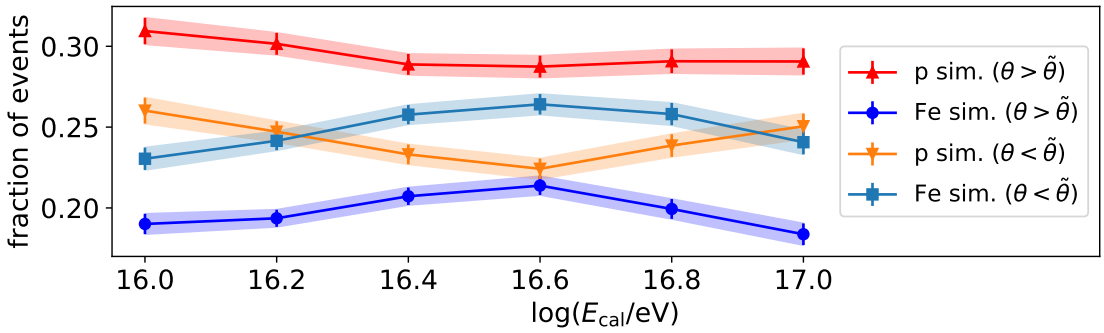

Figure 5.17.: Different selection efficiencies for proton and iron above and below the median zenith angle $\tilde{\theta}$. Above the median, proton showers have a higher fraction and iron showers have a lower fraction compared to the values below the median.

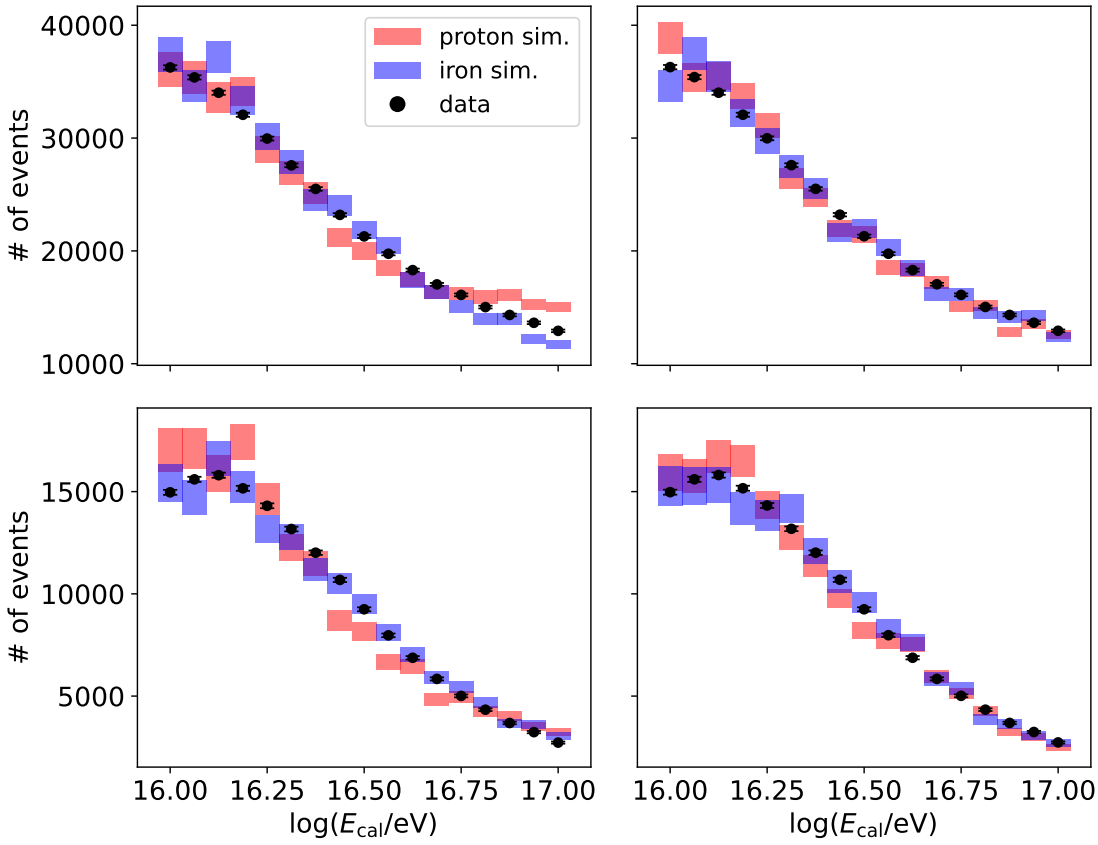

Figure 5.18.: Energy spectrum below (left) and above (right) the median value $\tilde{\theta}$ without (top) and with (bottom) quality cuts applied. The agreement between data and simulations is very good, similar to the full dataset.

Table 5.2.: Fit parameters and statistical uncertainties from the fit of Eq. (5.3) to z in the complete dataset and the subsamples based on the median zenith angle $\tilde{\theta}$.

fit parameter	full dataset	data ($\theta < \tilde{\theta}$)	data ($\theta > \tilde{\theta}$)
$\log(E^*_{\mathrm{cal}}/\mathrm{eV})$	16.60 ± 0.02	16.69 ± 0.03	16.50 ± 0.06
z^*	0.587 ± 0.007	0.74 ± 0.02	0.44 ± 0.02
D_z^{a}	0.42 ± 0.02	0.51 ± 0.04	0.36 ± 0.06
D_z^{b}	-0.30 ± 0.04	-0.5 ± 0.1	-0.21 ± 0.09

Despite the separation of z, a break energy is observed in the zenith angle subsamples as well. Fits are performed with Eq. (5.3) in the subsamples. The results compared to the parameters derived from the full dataset are shown in Tab. 5.2. The qualitative behaviour of the data and its subsamples is similar, although the numeric values of the fit parameters differ. The maximum difference for each parameter in the subsamples compared to the data is included as a systematic uncertainty. This leads to a result for the parameters in the full dataset of

$$\begin{aligned}
\log(E^*_{\mathrm{cal}}/\mathrm{eV}) &= 16.60 \pm 0.02 \pm 0.10,\\
z^* &= 0.587 \pm 0.007 \pm 0.2,\\
D_z^{\mathrm{a}} &= 0.42 \pm 0.02 \pm 0.08,\\
D_z^{\mathrm{b}} &= -0.30 \pm 0.04 \pm 0.2,
\end{aligned}$$

with the first uncertainty being statistical and the second one systematic.

5.5. Comparison with Gaisser-Stanev-Tilav model

Results of the HeCo mass composition analysis are compared with the GST model in this section. The GST model is chosen for comparison, rather than a measurement of mass composition, because the parametrisation of the flux for different primary masses can be adapted in the HeCo simulations as a function of the calorimetric energy E_{cal}. This is done by weighting the equally mass distributed ICRC21 events according to the primary fractions in the model. These fractions are derived using the python module *crflux* which was used as research code in Ref. [56]. It compiles different cosmic ray flux models including the GST model. In the model parameters, carbon and oxygen are used as primaries with masses between helium and iron [18]. The carbon and oxygen fluxes are added to calculate the nitrogen fraction weight for the ICRC21 simulations, as the three elements have very similar masses. In the second Galactic component of the GST model, elements heavier than iron are also included, which are not used in this analysis. The fraction weights are applied to the respective simulations based on their generated primary particle element and their generated energy.

The parameter z is calculated in the same manner as for the data, using Eq. (5.2) and with $\langle X_{\mathrm{max}} \rangle$ of the GST-weighted ICRC21 simulations. The parameter is once more examined for stability regarding median cuts. For almost all the parameters described in Section 5.4.2, no separation of z is observed. It is only possible to discern a separation for the zenith angle θ and the X_{max} viewing angle $\alpha_{X_{\mathrm{max}}}$. The two parameters have a similar separation

$\Delta z < 0.1$ in the whole energy range. As an example, the separation for the zenith angle is visualised in Fig. 5.19. It is much smaller than the separation for z calculated in the data described in Section 5.4.2. Because strong separation for the zenith angle analysis occurs for the data but not for the weighted simulations, the separation can be excluded as an effect purely of the selection. It could instead be caused by a difference in the data and simulations.

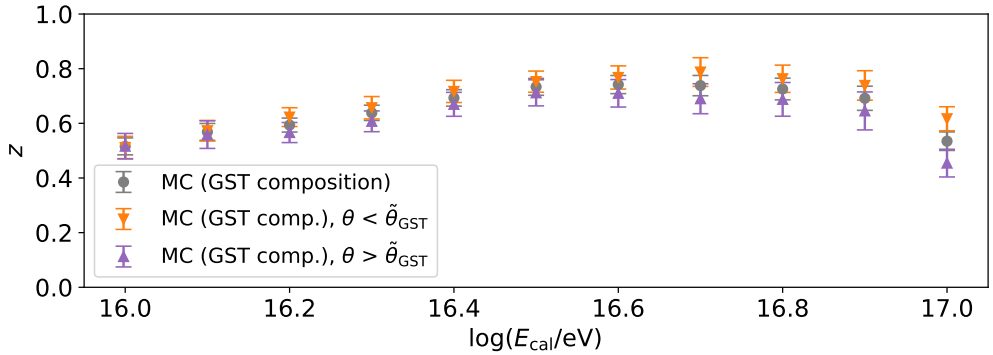

Figure 5.19.: Mass estimator z of simulations weighted according to the GST model. After a zenith angle median cut, only a small separation is seen, in contrast to the HeCo data in Fig. 5.15.

To compare with the GST-weighted simulations, systematic uncertainties are included for the z parameter calculated in the data. The z separation observed for the zenith angle in Fig. 5.15 is taken as the first contribution to the uncertainty. Another systematic effect introduced in Section 4.3 is the difference in X_{max} bias when comparing CORSIKA showers simulated with the standard CORSIKA configuration and with the CONEX configuration. Similarly to the result of the analysis with TA cuts, the bias of the CORSIKA sample is on average 9.6 g/cm^2 higher compared to the CONEX configuration when applying the quality cuts introduced in Section 5.2. This applies to the energy range between 10^{16} eV and 10^{17} eV for protons. The difference in bias is also found to be zenith angle dependent. For events with $\theta < 47°$, the average bias difference is only 4.3 g/cm^2, while above 47° it increases to 17.3 g/cm^2. If a similar effect also applies to heavier elements, it would suggest that $\langle X_{\mathrm{max}} \rangle$ of the simulations would have to be corrected upward relative to the data at higher zenith angles. This would lead to a heavier composition result for larger zenith angles and, to a lesser extent, for the complete dataset. In this way, the separation seen for the z parameter in data could be reduced. However, since Cherenkov-light dominated CORSIKA showers are only available for protons, no detailed analysis can be performed to examine the difference in X_{max} bias with respect to heavier primaries. Instead, in a simplified way, the X_{max} bias is assumed to be 9.6 g/cm^2 higher for the iron simulations as well, and the resulting higher value

$$z = \frac{\langle X_{\mathrm{max}}^{\mathrm{p}} \rangle - \langle X_{\mathrm{max}} \rangle + 9.6\,\mathrm{g/cm^2}}{\langle X_{\mathrm{max}}^{\mathrm{p}} \rangle - \langle X_{\mathrm{max}}^{\mathrm{Fe}} \rangle}. \tag{5.4}$$

is included as the second contribution to the systematic uncertainty for z.

As a result, the mass estimator is compared between the GST-weighted simulations and the data. With the large systematic uncertainty, the GST model is still compatible with the data. The general trend that the composition becomes heavier up to around $10^{16.5}$ eV in the GST model and $10^{16.6}$ eV in the data, followed by a stagnant mass, is similar. However, the negative slope in the GST model is very small up to $10^{16.9}$ eV with a more sudden drop in mass at 10^{17} eV. In contrast, this transition is more continuous in the data, and the composition is slightly heavier at 10^{17} eV than at 10^{16} in calorimetric energy.

Fig. 5.20 illustrates the result of the data and GST-weighted simulations (left) and provides a comparison with measurements of $\ln A$ by different experiments. These experiments include Auger starting at energies above 10^{17} eV, TALE, and the non-imaging Cherenkov light detector Tunka-133 [57], which is part of the Tunka Advanced Instrument for cosmic ray physics (TAIGA) [58]. A comparison to these measurements is only possible on a qualitative basis, as they derive $\ln A$ dependent on the primary energy E_0, in contrast to the calorimetric energy E_{cal} used in this work. Nonetheless, the result by Tunka-133 shows a break in the mass composition in the energy range between 10^{16} eV and 10^{17} eV, while such a break is not observed in the measurements of the TA collaboration using TALE. Because a similar break is also observed for the measurement of the mass composition in this work, it can be concluded that this exhibits more similarity with the Tunka-133 result than with the TALE result.

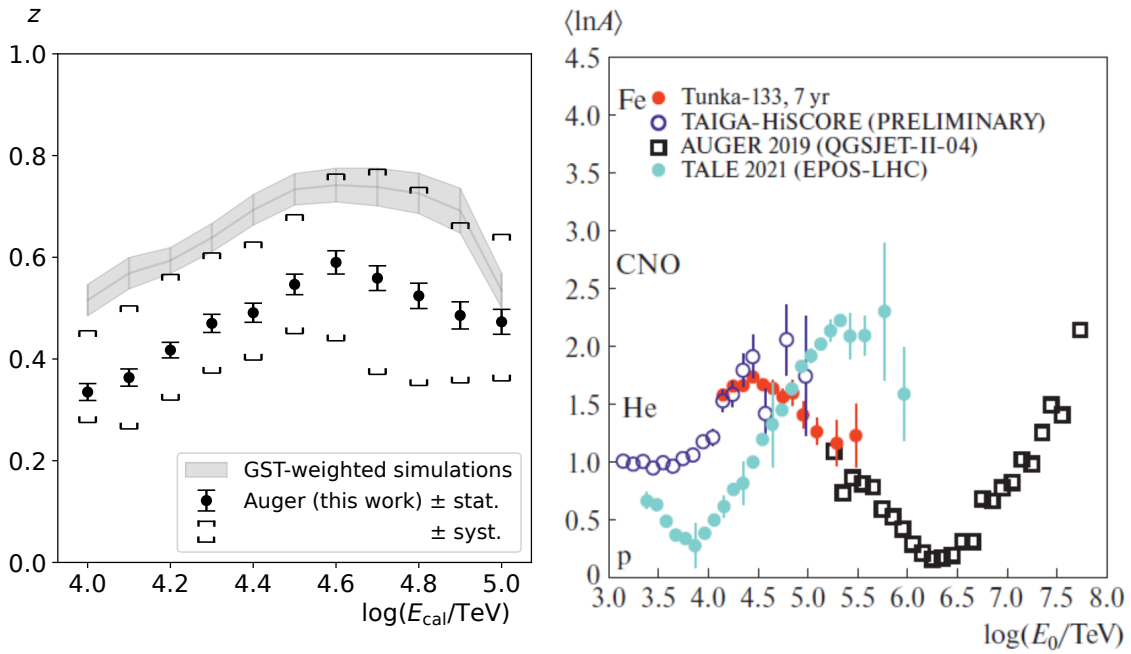

Figure 5.20.: Left: mass estimator z of simulations weighted according to the GST model and of HeCo data. The systematic error of the data originates from the separation observed for the zenith angle and an additional one-sided contribution from the difference to CORSIKA showers. Right: measurements of $\ln A$ by the detectors Tunka, TAIGA, TALE, and Auger. A break in the measurements by Tunka is evident between 4 TeV and 5 TeV. Concurrently, a steady increase in mass is measured by TALE in the same range. Taken from [59].

6. Summary and Conclusions

The goal of this work was to perform a mass composition analysis in the energy range that marks the transition from Galactic to extragalactic cosmic rays. The shower parameter that was used to differentiate between mass groups is the reconstructed depth of shower maximum X_{max}. In preparation for the mass composition analysis, the reconstruction procedure of Cherenkov-light dominated events at the Pierre Auger Observatory was examined.

An existing library of simulated Cherenkov-dominated showers was used in the analysis. The difference between reconstructed and simulated depth of shower maximum ΔX_{max} was compared between Auger simulations and results presented by the TA Collaboration in the same energy range. It was found that the reconstruction of the depth of shower maximum performed better in the TA analysis regarding the mean and standard deviation of ΔX_{max}. Motivated by this discrepancy, the Auger reconstruction of Cherenkov-light dominated events was examined for possible improvements in multiple ways. By using the true geometry for the reconstruction of simulated shower, it was found that the main source of bias and resolution in the reconstruction of X_{max} is the Profile Constrained Geometry Fit (PCGF) through which the geometry of an extensive air shower in the shower detector plane is reconstructed. The PCGF reconstruction includes constraints placed on the shape of the Gaisser-Hillas fit function for the energy deposit profile fit. Although the shape constraints were modified in multiple ways, including a configuration used by TA and the use of the true energy deposit profile shape, no improvements were found in the reconstruction of X_{max}. The precision of the PCGF module implementation was further investigated and confirmed.

Improvements in the measurement of the depth of shower maximum can be achieved by disabling the Cherenkov light lateral distribution function and the antialiasing filter correction in the simulation and reconstruction. Further investigation of the implementation of these two shower characteristics would be necessary to assess if they can be modified to improve the reconstruction.

The agreement between Cherenkov-dominated data and simulations was found to be very good for the energy spectrum and the distribution of other shower parameters, in particular regarding the geometry of triggered and selected showers. This good agreement is a necessary condition to use the simulations for a mass composition analysis. Based on the differences in the average reconstructed depth of shower maximum of four mass groups including protons, helium, nitrogen and iron, the simulations were compared to data to obtain the mass composition result. A composition that gets heavier up to the calorimetric break energy E^{*}_{cal} and then slightly lighter again was measured. A stability analysis showed that the derived composition is very robust with respect to most shower parameters. However, a zenith angle dependence of the mass composition result was found. Differences between CONEX and CORSIKA simulations indicate that the simplified

generation of Cherenkov light in the former case might be a source of the zenith angle dependence. For a more detailed analysis, CORSIKA showers for heavier components than protons would be required. The zenith angle dependence and differences between CONEX and CORSIKA were incorporated as systematic uncertainties, leading to a result for the break energy of

$$E^*_{\mathrm{cal}} = (4.0 \pm 0.2 \pm 0.9) \times 10^{16}\,\mathrm{eV},$$

which is near the low-energy ankle at $E_0 = 2.8 \times 10^{16}$ eV.

The mass composition result is comparable to the Gaisser-Stanev-Tilav model within the systematic uncertainties. In that model, the composition changes at the end of the Galactic cosmic ray spectrum with the transition to an extragalactic component. A similar effect could be the cause of the mass composition behaviour found in this work around the energy E^*_{cal}. The composition getting heavier below this energy would correspond to a rigidity-dependent cut-off of a galactic cosmic ray component according to the Peters cycle. The change to a lighter composition above E^*_{cal} would represent the transition to a new Galactic or extragalactic component in this scenario. Beyond this qualitative interpretation of the results, the interpolation to specific mass groups is difficult. With improved uncertainties of future analyses using Cherenkov-light dominated events, it will be possible to more accurately describe the mass composition at the end of the Galactic cosmic ray spectrum.

Bibliography

[1] V. F. Hess. "Über Beobachtungen der durchdringenden Strahlung bei sieben Freiballonfahrten." *Z. Phys.* 13 (1912), p. 1084.

[2] A. Coleman et al. "Ultra high energy cosmic rays: The intersection of the Cosmic and Energy Frontiers". *Astroparticle Physics* 149, 102819 (2023).

[3] M. Circella et al. "The space mission PAMELA". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 518.1–2 (2004), p. 153.

[4] F. Barao et al. "AMS—Alpha Magnetic Spectrometer on the International Space Station". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 535.1–2 (2004), p. 134.

[5] E. S. Seo et al. "Cosmic-ray energetics and mass (CREAM) balloon project". *Advances in Space Research* 33.10 (2004), p. 1777.

[6] M. Ave et al. "The TRACER instrument: A balloon-borne cosmic-ray detector". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 654.1 (2011), p. 140.

[7] J. R. Hörandel. "Overview on direct and indirect measurements of Cosmic Rays". *International Journal of Modern Physics A* 20.29 (2005), p. 6753.

[8] L. A. Anchordoqui. "Ultra-high-energy cosmic rays". *Physics Reports* 801 (2019), p. 1.

[9] C. Evoli. *The Cosmic-ray Spectrum across the Knees*. 2024. URL: https://zenodo.org/records/14508816.

[10] V. Novotný et al. "Energy spectrum of cosmic rays measured using the Pierre Auger Observatory". *Proceedings of 37th International Cosmic Ray Conference — PoS(ICRC2021)*. Sissa Medialab, 2021, p. 324.

[11] K. Greisen. "End to the Cosmic-Ray Spectrum?" *Physical Review Letters* 16.17 (1966), p. 748.

[12] G. T. Zatsepin and V. A. Kuz'min. "Upper Limit of the Spectrum of Cosmic Rays". *Soviet Journal of Experimental and Theoretical Physics Letters* 4 (1966), p. 78.

[13] R. Aloisio, V. Berezinsky, and A. Gazizov. "Ultra high energy cosmic rays: The disappointing model". *Astroparticle Physics* 34.8 (2011), p. 620.

[14] B. Peters. "Primary cosmic radiation and extensive air showers". *Il Nuovo Cimento* 22.4 (1961), p. 800.

[15] A. M. Hillas. "Can diffusive shock acceleration in supernova remnants account for high-energy galactic cosmic rays?" *Journal of Physics G: Nuclear and Particle Physics* 31.5 (2005), R95.

[16] A. M. Hillas. *Cosmic Rays: Recent Progress and some Current Questions*. 2006. URL: https://arxiv.org/abs/astro-ph/0607109.

[17] T. K. Gaisser. "Spectrum of cosmic-ray nucleons, kaon production, and the atmospheric muon charge ratio". *Astroparticle Physics* 35.12 (2012), p. 801.

[18] T. K. Gaisser, T. Stanev, and S. Tilav. "Cosmic ray energy spectrum from measurements of air showers". *Frontiers of Physics* 8.6 (2013), p. 748.

[19] D. Heck et al. *CORSIKA: A Monte Carlo Code to Simulate Extensive Air Showers*. FZKA-6019. Forschungszentrum Karlsruhe, 1998.

[20] T. Bergmann et al. "One-dimensional hybrid approach to extensive air shower simulation". *Astroparticle Physics* 26.6 (2007), p. 420.

[21] W. Heitler. *The quantum theory of radiation*. Second edition. Courier Corporation, 1944.

[22] J. Matthews. "A Heitler model of extensive air showers". *Astroparticle Physics* 22.5–6 (2005), p. 387.

[23] H. M. J. Barbosa et al. "Determination of the calorimetric energy in extensive air showers". *Astroparticle Physics* 22.2 (2004), p. 159.

[24] A. Aab et al. "Data-driven estimation of the invisible energy of cosmic ray showers with the Pierre Auger Observatory". *Physical Review D* 100.8, 082003 (2019).

[25] T. K. Gaisser, R. Engel, and E. Resconi. *Cosmic Rays and Particle Physics*. Second edition. Cambridge University Press, 2016. ISBN: 9781139192194.

[26] P. Sokolsky. *Introduction to ultrahigh energy cosmic ray physics*. Ed. by Gordon Thomson. Second edition. CRC Press, Taylor & Francis Group, 2020. ISBN: 9780429055157.

[27] M. Unger et al. "Reconstruction of longitudinal profiles of ultra-high energy cosmic ray showers from fluorescence and Cherenkov light measurements". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 588.3 (2008), p. 433.

[28] M. Ave et al. "Measurement of the pressure dependence of air fluorescence emission induced by electrons". *Astroparticle Physics* 28.1 (2007), p. 41.

[29] F. Nerling et al. "Universality of electron distributions in high-energy air showers - description of Cherenkov light production". *Astroparticle Physics* 24.6 (2005), p. 421.

[30] A. Aab et al. "The Pierre Auger Cosmic Ray Observatory". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 798 (2015), p. 172.

[31] I. Allekotte et al. "The Surface Detector System of the Pierre Auger Observatory". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 586.3 (2007), p. 409.

[32] D. Allard et al. "Aperture calculation of the Pierre Auger Observatory surface detector". *Proc. of the 29th ICRC (Pune, India)*. 2005, p. 71.

[33] E. Varela et al. "The low-energy extensions of the Pierre Auger Observatory". *Journal of Physics: Conference Series* 468, 012013 (2013).

[34] G. Silli et al. "Performance of the 433 m surface array of the Pierre Auger Observatory". *Proceedings of 37th International Cosmic Ray Conference — PoS(ICRC2021)*. Sissa Medialab, 2021, p. 224.

[35] A. Aab et al. "Reconstruction of events recorded with the surface detector of the Pierre Auger Observatory". *Journal of Instrumentation* 15.10, P10021 (2020).

[36] J. Abraham et al. "The Fluorescence Detector of the Pierre Auger Observatory". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 620.2–3 (2009), p. 227.

[37] T. H. J. Mathes et al. "The HEAT Telescopes of the Pierre Auger Observatory: Status and First Data". *Proc. of the 32nd ICRC, Beijing, China*. 2011, p. 153.

[38] R. M. Baltrusaitis et al. "The Utah Fly's Eye detector". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 240.2 (1985), p. 410.

[39] D. Kuempel, K. Kampert, and M. Risse. "Geometry reconstruction of fluorescence detectors revisited". *Astroparticle Physics* 30.4 (2008), p. 167.

[40] T. K. Gaisser and A. M. Hillas. "Reliability of the Method of Constant Intensity Cuts for Reconstructing the Average Development of Vertical Showers". *International Cosmic Ray Conference, 15th, Plovdiv, Bulgaria, August 13-26, 1977, Conference Papers*, p. 353.

[41] T. Abu-Zayyad. "The energy spectrum of ultra high energy cosmic rays". PhD thesis. University of Utah, 2000.

[42] S. Argirò et al. "The offline software framework of the Pierre Auger Observatory". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 580.3 (2007), p. 1485.

[43] V. Novotný. "Measurement of the energy spectrum of cosmic rays using Cherenkov–dominated data at the Pierre Auger Observatory". PhD thesis. Charles University Prague, 2020.

[44] M. Unger. *Shower Profile Reconstruction from Fluorescence and Cherenkov light*. Auger-internal document GAP-2006-010. 2006.

[45] J. Bellido et al. *Shower Profile Fitting*. Auger-internal document GAP-2022-050. 2022.

[46] S. Andringa, R. Conceição, and M. Pimenta. "Mass composition and cross-section from the shape of cosmic ray shower longitudinal profiles". *Astroparticle Physics* 34.6 (2011), p. 360.

[47] F. Riehn et al. "Hadronic interaction model sibyll 2.3d and extensive air showers". *Physical Review D* 102.6, 063002 (2020).

[48] H. Kawai et al. "Telescope Array Experiment". *Nuclear Physics B - Proceedings Supplements* 175–176 (2008), p. 221.

[49] G. B. Thomson et al. "The Telescope Array Low Energy Extension (TALE)". *Proc. of the 32nd ICRC, Beijing, China*. 2011, p. 338.

[50] R. U. Abbasi et al. "The Cosmic-Ray Composition between 2 PeV and 2 EeV Observed with the TALE Detector in Monocular Mode". *The Astrophysical Journal* 909.2, 178 (2021).

[51] S. Ostapchenko. "Status of QGSJET". *AIP Conference Proceedings*. Vol. 928. 2007, p. 118.

[52] T. Pierog et al. "EPOS LHC: Test of collective hadronization with data measured at the CERN Large Hadron Collider". *Physical Review C* 92.3, 034906 (2015).

[53] Z. J. Zundel. "Spectrum measurement with the Telescope Array Low Energy Extension (TALE) fluorescence detector". PhD thesis. University of Utah, 2016.

[54] D. Góra et al. "Universal lateral distribution of energy deposit in air showers and its application to shower reconstruction". *Astroparticle Physics* 24.6 (2006), p. 484.

[55] A. Kuotb. "Measurements of the maximum depth of air shower profiles at LHC energies with the High-Elevation Auger Telescopes". PhD thesis. Karlsruhe Institute of Technology, 2018.

[56] A. Fedynitch, J. Becker Tjus, and P. Desiati. "Influence of hadronic interaction models and the cosmic ray spectrum on the high energy atmospheric muon and neutrino flux". *Physical Review D* 86.11, 114024 (2012).

[57] S. F. Berezhnev et al. "The Tunka-133 EAS Cherenkov light array: Status of 2011". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 692 (2012), p. 98.

[58] N. M. Budnev et al. "The TAIGA experiment: From cosmic-ray to gamma-ray astronomy in the Tunka valley". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 845 (2017), p. 330.

[59] N. M. Budnev et al. *TAIGA - an advanced hybrid detector complex for astroparticle physics and high energy gamma-ray astronomy*. 2022. URL: https://arxiv.org/abs/2208.13757v1.

A. Appendix

A.1. Comparison between Auger and TA simulations

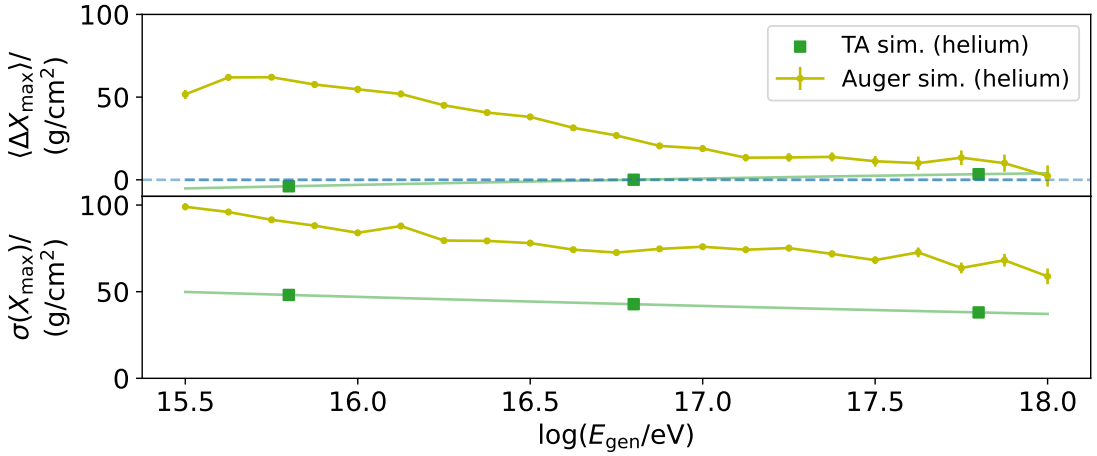

Figure A.1.: Comparison of the X_{max} bias and resolution of Auger simulations to values of TALE simulations published by TA in Ref. [50] for helium showers.

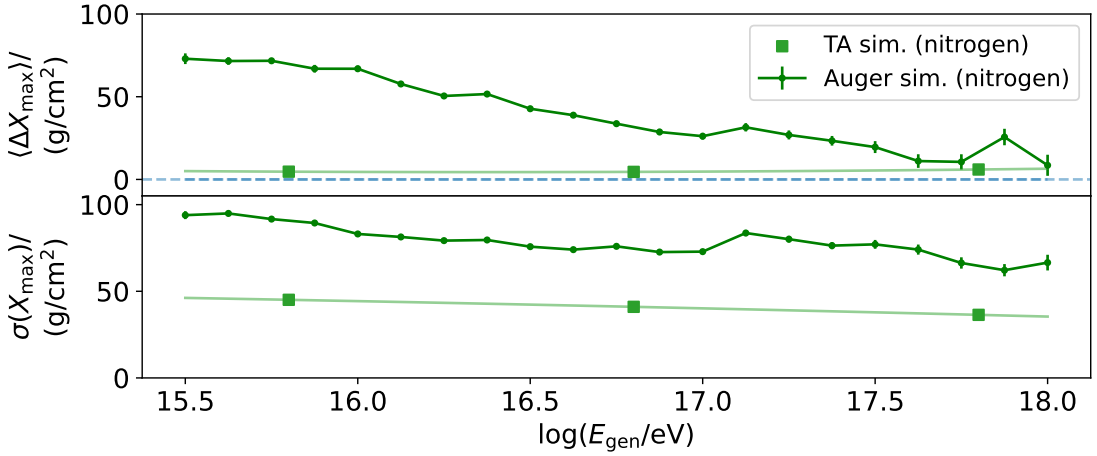

Figure A.2.: Comparison of the X_{max} bias and resolution of Auger simulations to values of TALE simulations published by TA in Ref. [50] for nitrogen showers.

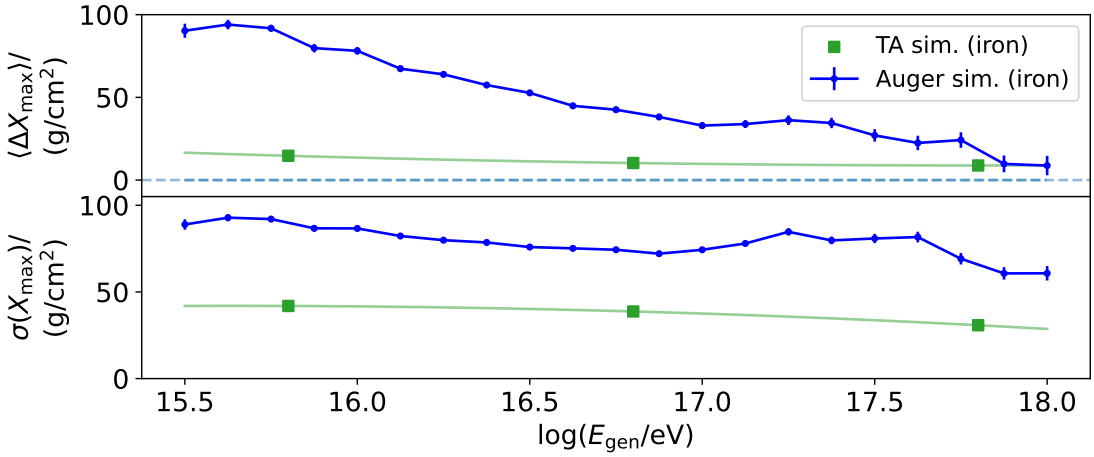

Figure A.3.: Comparison of the X_{max} bias and resolution of Auger simulations to values of TALE simulations published by TA in Ref. [50] for iron showers.

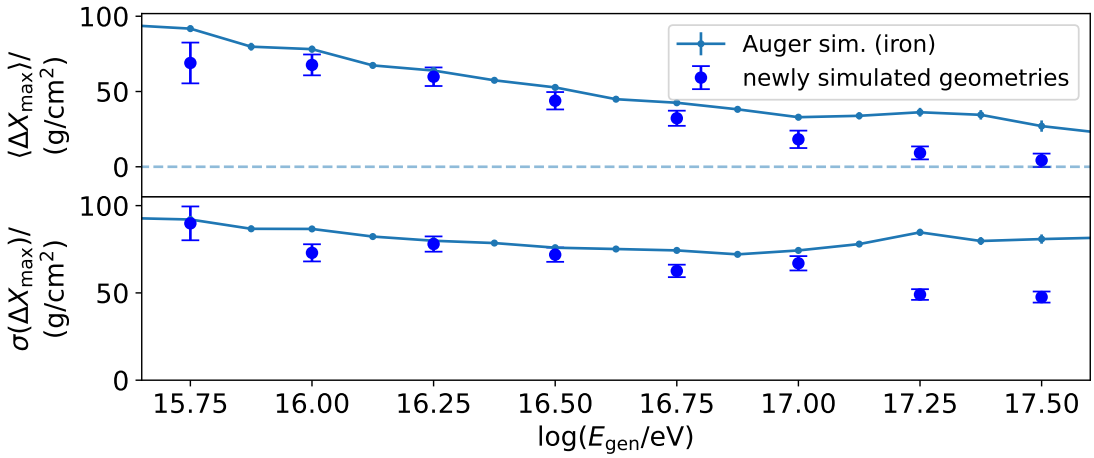

Figure A.4.: Comparison of the X_{max} bias and resolution of the complete ICRC21 simulations to the smaller sample of newly simulated geometries introduced in Section 4.1 for iron.

A.2. Incomplete error propagation in HeCo

The quality cut in the estimated error of δX_{max} as well as the relative energy error $\delta E_{\mathrm{cal}}/E_{\mathrm{cal}}$ used in Chapter 4 is incomplete for HeCo events, as the geometric error is not propagated for the virtual detector and is therefore missing from the estimated error. The reconstruction with PCGF was performed only for HeCo in the ICRC21 simulations and not for other telescopes, which makes a correctly calculated error unavailable. Part of the geometric X_{max} error can be calculated from the estimated error $\delta\chi_0$. However, this contribution was found to be small compared to the full geometric error calculated in other eyes. The effect of incomplete error estimation was investigated with the smaller sample of simulations introduced in Section 4.1 by performing two reconstructions, one in HEAT and one in HeCo, with the same reconstruction configuration. In the case of HeCo, only events where the telescope with the maximum number of triggered pixels belongs to HEAT are chosen to compare with pure HEAT events.

The difference in the estimated errors δX_{max} and $\delta E_{\mathrm{cal}}/E_{\mathrm{cal}}$ between the two eyes is illustrated in Fig. A.5, with generally higher estimated errors in HEAT. However, the application of the respective cuts (as documented in Tab. 4.2) demonstrated only a minor effect on cut efficiencies, especially when combined with the other TA quality cuts. When all other cuts are applied first, the efficiency for the δX_{max} cut for HeCo and for HEAT is 98.3 % and 95.5 %, respectively. Accordingly, the efficiency for the $\delta E_{\mathrm{cal}}/E_{\mathrm{cal}}$ cut for HeCo and for HEAT is 98.0 % and 97.7 %, respectively. Changes of the X_{max} bias and resolution are not observed between the reconstruction in the two eyes. Nevertheless, the affected cuts were disregarded in the quality selections for the mass composition analysis in Chapter 5.

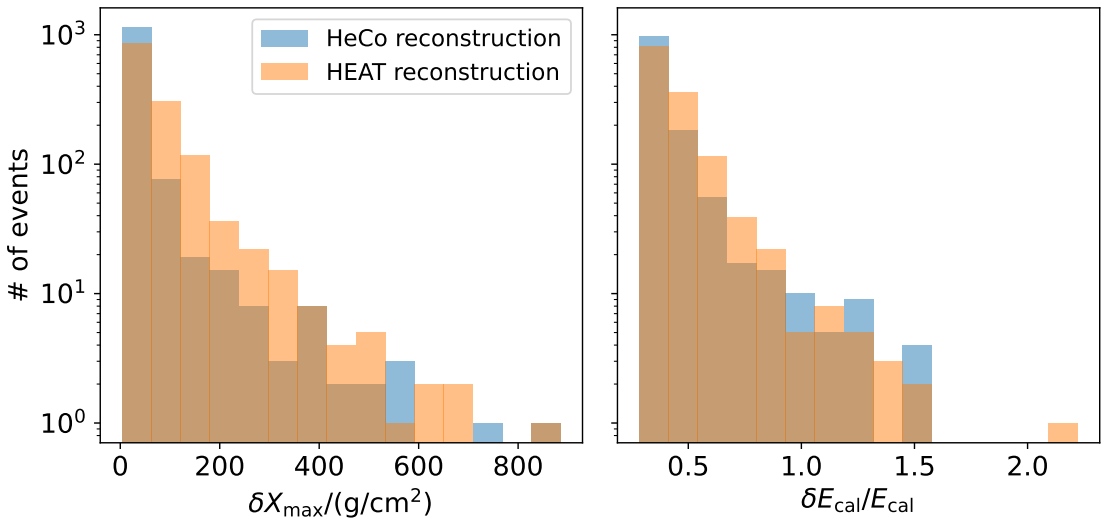

Figure A.5.: Differences between the estimated X_{max} error and estimated relative E_{cal} error for the same simulations reconstructed in the HEAT system and HeCo system. The geometric error propagation is missing in the HeCo reconstruction.

A.3. Quality selection cut analysis

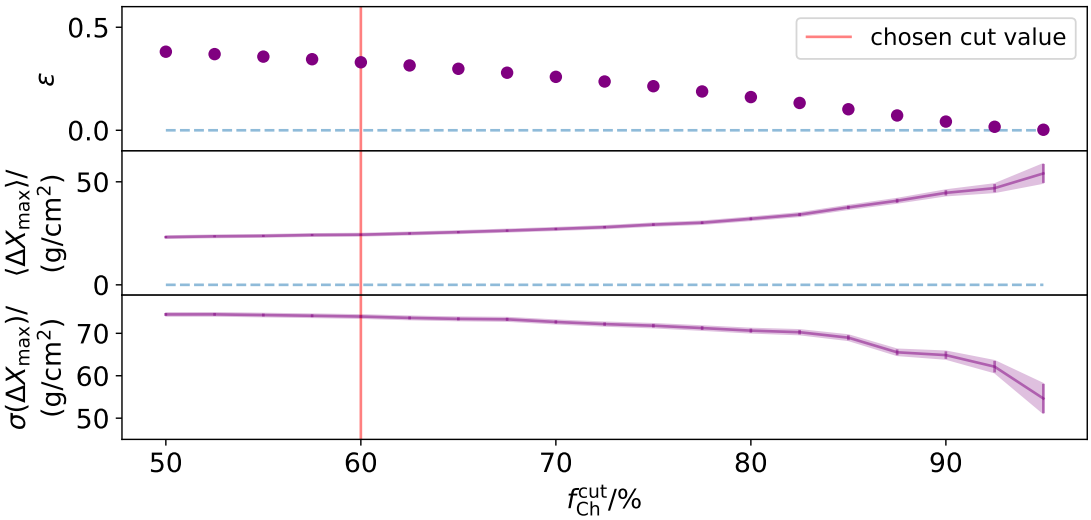

Figure A.6.: Cut analysis for the Cherenkov light fraction, applying it in addition to the other quality cuts from Tab. 5.1 in the highest energy bin between $10^{16.75}$ eV and 10^{17} eV. The chosen cut value is $f_{\mathrm{Ch}}^{\mathrm{cut}} = 60\,\%$. At lower energies the cut only has a minor effect.

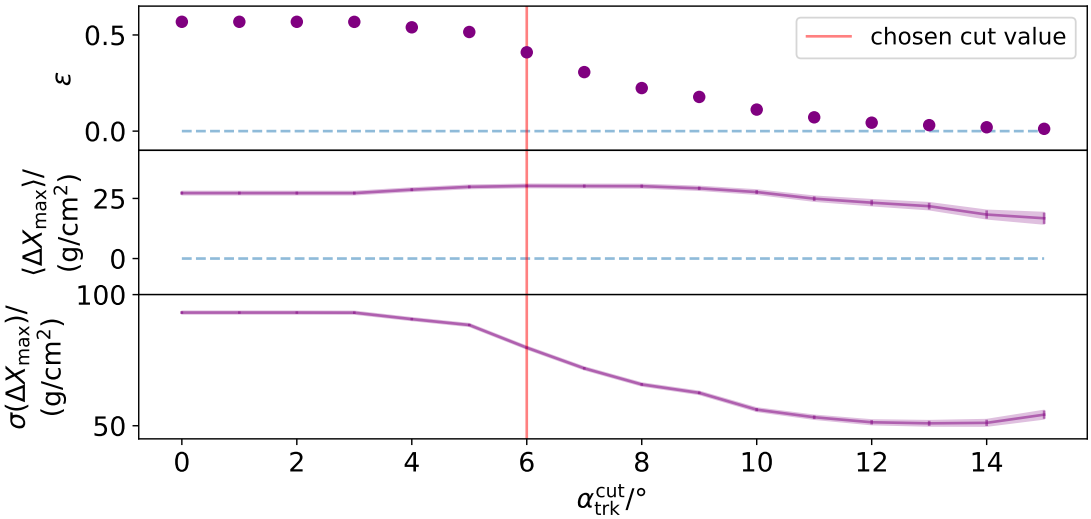

Figure A.7.: Cut analysis for the angular track length, applying it in addition to the other quality cuts from Tab. 5.1 in the lowest energy bin between $10^{16.25}$ eV and $10^{16.5}$ eV. The chosen cut value is $\alpha_{\mathrm{cut}}^{\mathrm{cut}} = 6°$.

A.4. Comparison between Auger data and simulations

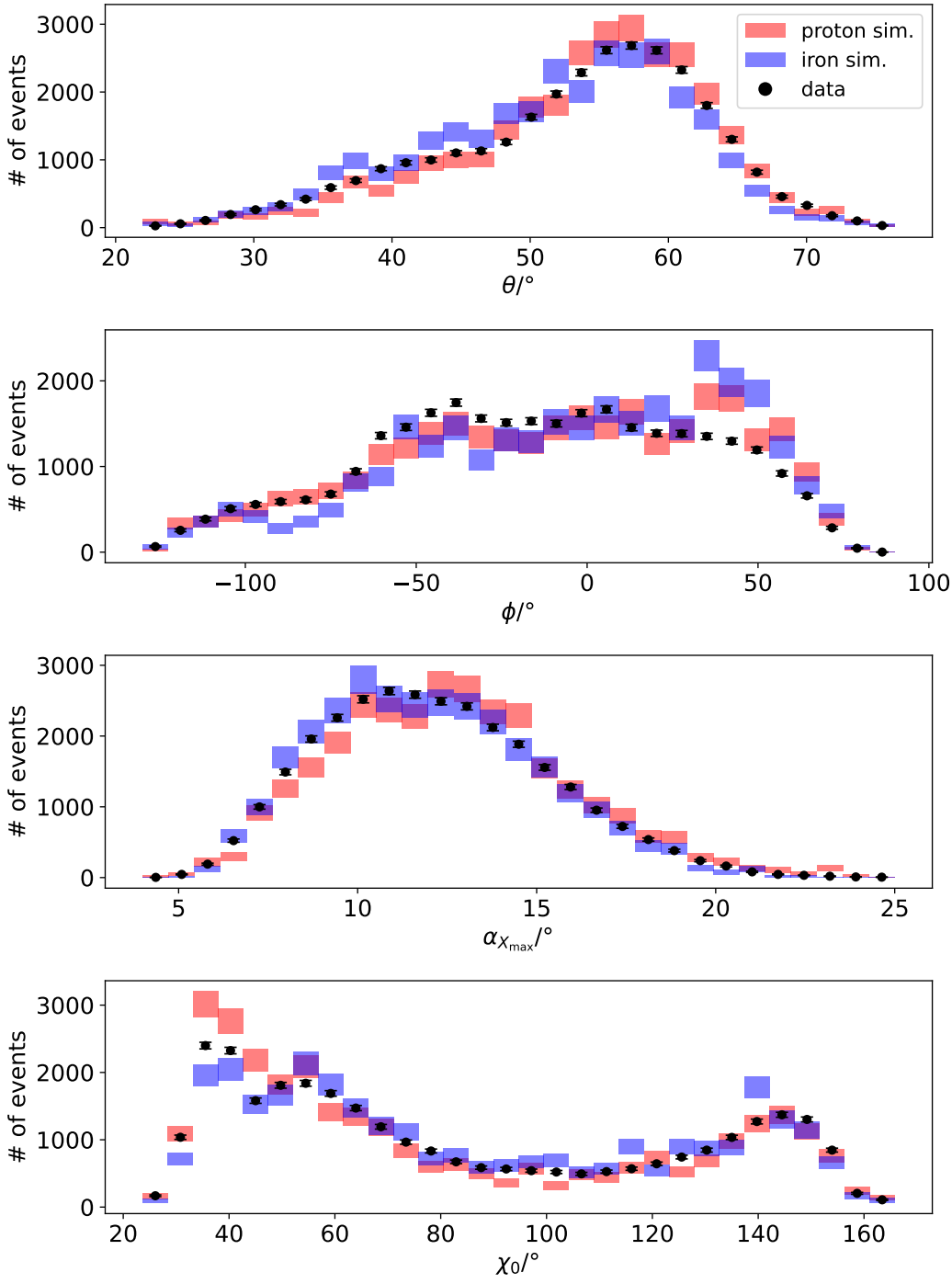

Figure A.8.: Distributions of events across various parameters: zenith angle θ, azimuth angle ϕ, X_{max} viewing angle $\alpha_{X_{max}}$ and χ_0, for data and simulation with calorimetric energies between $\log E = 16.75$ and $\log E = 17.0$. Simulations are scaled to the data.

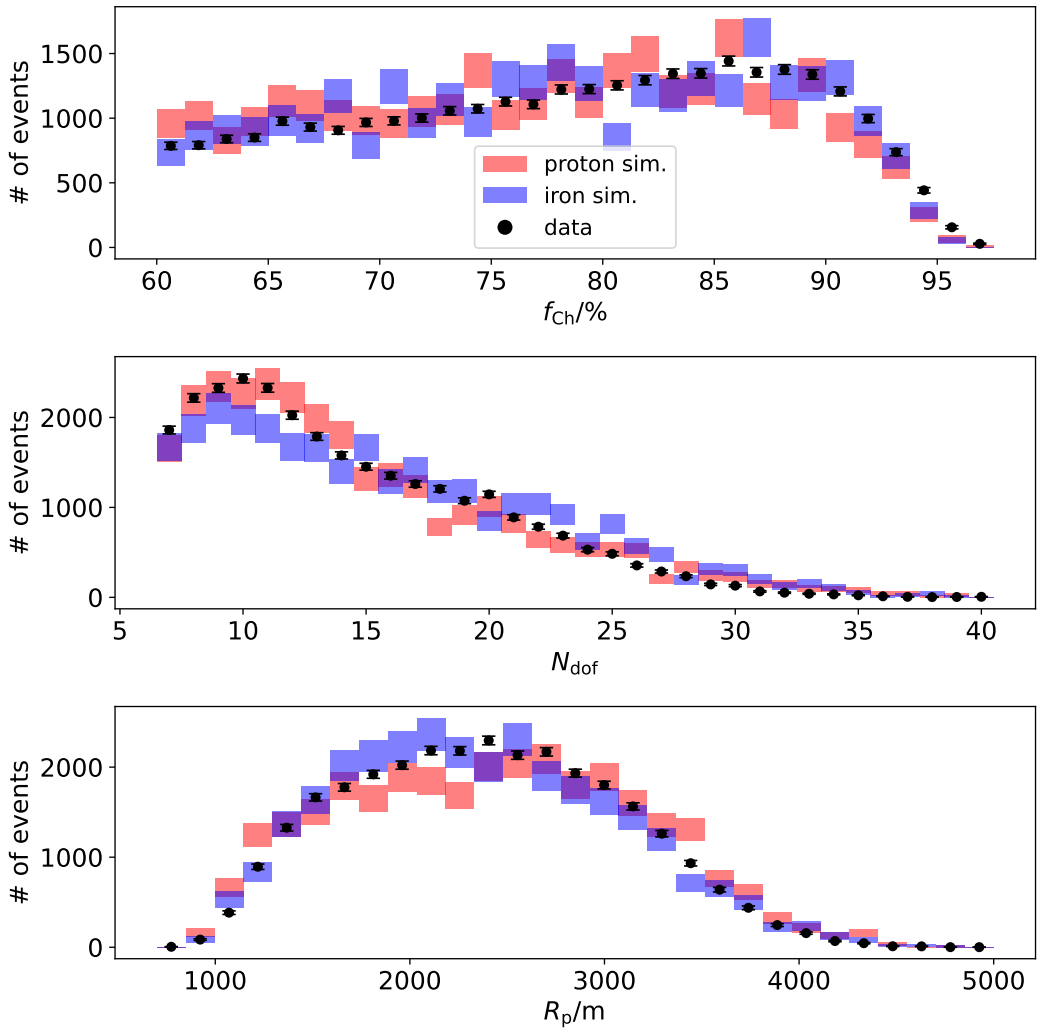

Figure A.9.: Distributions of Cherenkov fraction f_{Ch}, profile fit degrees of freedom N_{dof} and impact parameter R_p for data and simulations with calorimetric energies between $\log E = 16.75$ and $\log E = 17.0$. Simulations are scaled to the data.

Acknowledgements

First, I would like to extend my sincere gratitude to my supervisor, Dr. Michael Unger, for providing insight, support and crucial ideas for the realisation of the work presented in this thesis.

I would like to thank Prof. Dr. Ralph Engel for his constructive feedback on the status of my work at various points during my year-long stay in the Auger group at KIT. Furthermore, I thank Prof. Dr. Guido Drexlin for agreeing to act as the second reviewer for this thesis.

I extend my gratitude to all colleagues at the IAP who provided the welcoming and pleasant working environment that made this work so much easier. Of course, other members of the Pierre Auger Collaboration also provided valuable insight and assistance. I especially want to thank Vladimír Novotný for sharing his vast knowledge about Cherenkov-light dominated showers at Auger with me.

Finally, my heartfelt thanks go to my family and friends for providing mental support and constant encouragement. Specifically, I would like to thank Siling Chen for the fruitful exchange regarding academic writing, and of course I want to thank Adrian for always being there for me.