On Air-Shower Universality and the Mass Composition of Ultra-High-Energy Cosmic Rays

Zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

an der KIT-Fakultät für Physik des
Karlsruher Institut für Technologie (KIT)
und des Instituto de Tecnología "Prof. Jorge A. Sábato" der
Universidad Nacional de San Martín (UNSAM)

angenommene

Dissertation

von

M. Sc. Maximilian Klaus Stadelmaier

aus Obereisesheim

Tag der mündlichen Prüfung: 04. 02. 2022
Referent: Prof. Dr. Ralph Engel
Korreferent: Prof. Dr. Federico Sanchez
Betreuer: Dr. Markus Roth and Dr. Darko Veberič

On Air-Shower Universality and the Mass Composition of Ultra-High-Energy Cosmic Rays

For the attaiment of the academic degree of

Doctorate in Science

from the
Karlsruher Institut für Technologie (KIT)
and the
Universidad Nacional de San Martín (UNSAM)

accepted

Dissertation

of

M. Sc. Maximilian Klaus Stadelmaier

of Obereisesheim

Day of the oral examination: 04. 02. 2022
Referee: Prof. Dr. Ralph Engel
Co-referee: Prof. Dr. Federico Sanchez
Supervisors: Dr. Markus Roth and Dr. Darko Veberič

On Air-Shower Universality and the Mass Composition of Ultra-High-Energy Cosmic Rays

Tesis aceptada para optar en el título de

Doctor en Astrofísica

del Karlsruher Institut für Technologie (KIT)
y del Instituto de Tecnología "Prof. Jorge A. Sábato" de la
Universidad Nacional de San Martín (UNSAM)

por

M. Sc. Maximilian Klaus Stadelmaier

de Obereisesheim

Fecha de la defensa oral: 04. 02. 2022
Director: Prof. Dr. Ralph Engel
Co-director: Prof. Dr. Federico Sanchez
Colaboradores: Dr. Markus Roth and Dr. Darko Veberič

Erklärung zur Selbstständigkeit

Ich versichere, dass ich diese Arbeit selbstständig verfasst habe und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe, die wörtlich oder inhaltlich übernommenen Stellen als solche kenntlich gemacht und die Satzung des KIT zur Sicherung guter wissenschaftlicher Praxis in der aktuellen Fassung beachtet habe.

Karlsruhe, den 19.01.2022, _____

Maximilian Stadelmaier

Del Rigor en la Ciencia

... en aquel Imperio, el Arte de la Cartografía logró tal Perfección que el mapa de una sola Provincia ocupaba toda una Ciudad, y el mapa del Imperio, toda una Provincia. Con el tiempo, estos Mapas Desmesurados no satisficieron y los Colegios de Cartógrafos levantaron un Mapa del Imperio, que tenía el tamaño del Imperio y coincidía puntualmente con él. Menos Adictas al Estudio de la Cartografía, las Generaciones Siguientes entendieron que ese dilatado Mapa era Inútil y no sin Impiedad lo entregaron a las Inclemencias del Sol y los Inviernos. En los desiertos del Oeste perduran despedazadas Ruinas del Mapa, habitadas por Animales y por Mendigos; en todo el País no hay otra reliquia de las Disciplinas Geográficas.

- J. L. Borges

Abstract

The question of the origin of ultra-high-energy cosmic rays is one of the unsolved problems in physics today. An important step towards the answer is to be able to identify minimally deflected particles in the sky, whose arrival directions potentially point towards their sources. This can be achieved only by an event-level estimate of the mass – and thus the charge – of cosmic rays. In this work I revisit and extend the work on air-shower universality done in the last decades to identify light and heavy cosmic rays with modern surface-detector experiments.

This thesis comprises a detailed review on analytical approaches to describe extensive air showers, as well as a step-by-step guide of how to develop a model of the spatial and temporal distribution of particles in extensive air showers. I present considerations on which observables are dependent and sensitive on the depth of the shower maximum as well as on the relative muon content. The model is parametrized and tested using Monte Carlo simulations. Furthermore, I propose a method to calculate the atomic mass number of primary cosmic rays directly from the combined information of the depth of the shower maximum and the relative muon content, along with a way to calibrate the method using data from fluorescence detectors. At last, I present results on the mass composition of the highest-energy cosmic rays as well as a selection of high-rigidity events, which are identified in the data of the Pierre Auger Observatory.

ZUSAMMENFASSUNG

Die Frage nach dem Urpsrung der ultrahochenergetischen kosmischen Strahlung ist eines der bislang ungelösten Probleme der Physik. Ein unabdingbarer Schritt auf dem Weg zur Lösung ist die Möglichkeit, kosmische Teilchen zu identifizieren, die nur minimal durch Magnetfelder abgelenkt wurden, da deren Ursprungsrichtungen potenziell auf ihre Quellen deuten. Dies kann nur mittels eines ereignisbasierten Schätzers der atomaren Masse – und damit der Kernladungszahl – der kosmischen Strahlung geschehen. In dieser Dissertation überarbeite und erweitere ich die Arbeit zur Luftschaueruniversalität der letzten Jahrzehnte, um leichte von schweren kosmischen Teilchen mit modernen Oberflächendetektoren unterscheiden zu können.

Diese Arbeit umfasst eine detaillierte Zusammenfassung der analytischen Beschreibung extensiver Luftschauer, sowie eine schrittweise Beschreibung wie ein darauf basierendes Modell der Teilchendichten von Luftschauern aufgebaut werden kann. Darauf aufbauend beschreibe ich eine Abschätzung des Verhaltens möglicher Observablen in Abhängigkeit der Tiefe der Schauermaxima und der relativen Menge an Myonen in Luftschauern und der damit verbundenen Sensitivität. Das Modell der Teilchendichten wird durch Monte-Carlo-Simulationen parametrisiert und getestet. Des Weiteren schlage ich eine Methode zur Bestimmung der logarithmischen atomaren Masse von kosmischen Strahlen anhand der Tiefe der Schauermaxima und der relativen Zahl der Myonen vor, sowie eine Möglichkeit die dargelegte Methode dank der Daten von Fluoreszenzteleskopen zu kalibrieren. Zum Schluss zeige ich Resultate zur Bestimmung der mittleren Kernmasse höchstenergetischer kosmischer Strahlung, sowie eine Auswahl höchstrigider Teilchen, die ich in den Daten des Pierre Auger Observatoriums identifiziert habe.

RESUMEN

El interrogante respecto al origen de los rayos cósmicos de ultra alta energía es uno de los problemas no resueltos de la física actual. Poder identificar partículas mínimamente desviadas en el cielo nocturno, cuyas direcciones de arribo potencialmente apuntan hacia sus fuentes, constituye un paso importante hacia la respuesta. Esto sólo puede alcanzarse estimando, evento a evento, la masa – y por lo tanto la carga – de los rayos cósmicos.

En este trabajo se revisita y extiende la labor de las últimas décadas sobre universalidad de lluvias atmosféricas para identificar rayos cósmicos livianos y pesados con experimentos modernos de detectores de superficie. Esta tesis comprende una revisión detallada de los enfoques analíticos que describen las lluvias atmosféricas así como el desarrollo de un modelo para la distribución espacio-temporal de las partículas en una lluvia atmosférica. Se presentan consideraciones respecto a qué observables son sensibles a la profundidad del máximo de la lluvia, así como a la cantidad relativa de muones. El modelo es parametrizado y probado usando simulaciones de Monte Carlo. Además, se propone un método para calcular el logaritmo del número másico del rayo cósmico combinando la información de la profundidad del máximo de la lluvia y el contenido relativo de muones, junto con una manera de calibrar el método usando datos tomados por detectores de fluorescencia. Por último, se presentan resultados de la composición de masa utilizando eventos de alta energía, así como eventos de alta rigidez magnética tomados por el Observatorio Pierre Auger.

Contents

CHAPTER I

INTRODUCTION

Particles of extraterrestrial and extragalactic origin arrive at the Earth as *Cosmic Rays* (CRs) with energies reaching far beyond the limitations of man-made accelerators. Not only is the exact origin of the highest-energetic CRs still unclear, but also their energies challenge the theorized limits for their sheer existence. The key to understanding the origins and acceleration mechanisms that are accountable for these particles is to unravel their exact physical properties – especially their masses and charges – which will be the main focus of this work. While the highest-energy CRs are predominantly ionized nuclei, their chemical composition and exact mass number is unknown.

At the highest energies, where the flux of particles is extremely low, it is statistically impossible to acquire enough data through direct detection to determine the mass composition of the *Ultra-High-Energy Cosmic Rays* (UHECRs) within the possibilities of today's technology. Only a profound understanding of the air-shower phenomenon, which occurs when UHECRs interact with the Earth's atmosphere, makes it possible to reconstruct the mass of this particles from secondary observables using ground-based detectors on an event-by-event basis. Ground-detector arrays can detect the air showers of such cosmic rays with a frequency of approximately one event per century per square kilometer of area covered. Reasonably large detector arrays such as the surface detector of the Pierre Auger Observatory, which is currently the largest detector of this kind, are already taking data and continue to do so even with upgraded detector elements to boost the sensitivity on the observables used to determine the mass of the primary particle.

The data of surface-detector arrays can be interpreted in many ways, for example using empirical and data driven approaches, such as machine learning. In this thesis I develop a purely physics-driven approach. For this approach I revisit and extend the work on air-shower universality of the last decades. Utilizing the implications of air-shower universality, extensive air showers can be described from a perspective in which both the shower-to-shower fluctuations and the systematic differences occurring from different types of primary particles manifest in only two observables besides the energy of the primary particle. These observables are the depth of the shower maximum and the relative number of muons. Using these observables, a model of the expected particle densities in air showers can be formulated. Additionally, I propose a new model of the temporal distribution of particles and signal. The model is parametrized on simulated responses of the water-Cherenkov detectors (WCDs) of the Pierre Auger Observatory, as well as on the scintillator surface detectors (SSDs), which are currently deployed for the upgraded detector *AugerPrime*.

Furthermore, using this model of air showers, I describe a reconstruction algorithm to estimate the depth of the shower maximum and the relative number of muons of air showers from data collected only by a surface detector array. Both of these observables are primary-mass dependent individually, but both separately yield only a finite sensitivity on the mass of the primary cosmic ray. Using the correlated information of both the depth of

the shower maximum and the relative number of muons I propose a novel method to estimate the atomic mass number with improved precision. This method minimizes the impact of shower-to-shower fluctuations on the reconstructed atomic mass and can be calibrated using the measurements of the fluorescence detector of the Pierre Auger.

I will present results on the reconstructed primary-mass sensitive observables as well as on the atomic mass number as a function of the primary energy. The results are compared against the results from the fluorescence detector in the region of energy where these measurements are available.

From the event-level reconstruction of the atomic mass number I will estimate the rigidity of the primary cosmic ray. In this way, we can identify minimally deflected particles in galactic and extragalactic magnetic fields, opening a new window into the Cosmic Ray Astronomy.

CHAPTER II

COSMIC RAYS

Cosmic Rays (CRs) are high-energy particles of extraterrestrial origin. Their sources include Supernovae, *Active Galactic Nuclei* (AGNs), the Sun and multiple other celestial objects and phenomena, from which some are still unknown or not yet understood [1]. Most CRs are relativistic particles with an energy that is much larger than their rest mass. The spectrum of CRs can be well described by a steeply falling power law (see Section 2.2), rapidly thinning at the highest energies.

The highest-energy CRs, however, being accelerated up to 10^{20} eV and beyond, are particularly interesting. First of all, their existence alone challenges our current understanding of the propagation of ultra-high energy particles through the intergalactic medium (see Section 2.2.2). Secondly, their Larmor radius is estimated to be larger than the size of our own galaxy, not only implying that they are of extragalactic origin, but also throwing up questions about the process of their acceleration [2].

In this chapter, we outline the most important properties of CRs and put recent findings into historical context. An overview over selected types of celestial objects is presented, discussing the features of possible source candidates.

2.1 EARLY HISTORY

After the discovery of radioactivity by Henri Becquerel, it became clear very quickly that the atmosphere is not a perfect isolator, but it can be and is constantly ionized [3]. As a consequence, any electrometer exposed to the atmosphere is discharged over time, even in the absence of nearby radioactive sources [4]. At first, the general scientific consensus was that the cause of the ionization of the atmosphere is solely radiation originating from radioactive elements that are abundant in the Earth's crust.

In the early 20^{th} century Theodor Wulf created the most precise electrometer of that time and did a series of experiments to try and quantify the possible origins of this radiation not only in soil, but also in the atmosphere. To test his theory that the atmosphere is ionized not only by radiation originating from the ground, he took an electrometer to the top of the Eiffel tower and measured the degree of ionization about 300 m above the ground [5]. His results, however, were inconclusive. The logical next step was to move to higher altitudes, which was carried out by multiple physicists using hot-air balloons, however again with inconclusive results. In 1912 Victor Hess, whose experiment used sealed electrometers and thus maintained a constant number of particles inside his apparatus, carried out a successful series of daily and nightly balloon flights to test the hypothesis of an extraterrestrial source of radiation. It was shown that the ionization of the atmosphere increased drastically at an altitude larger than approximately 1 km above sea level. He concluded that his findings can be very well explained by a powerful and omnipresent radiation penetrating

the atmosphere from above [6]. For the discovery of the cosmic radiation Victor Hess was awarded the Nobel prize in 1936 [4].

For a long time, gamma rays were favored over charged particles as the nature of the cosmic radiation. One of the reasons for this was that the exponential absorption of gamma rays fitted the experimental data quite well, identifying a hard and soft component of the cosmic radiation, with a shorter and longer wavelength, respectively [3]. Robert Milikan theorized that fusion of atoms in the interstellar medium produces the gamma rays accountable for the ionization of the atmosphere and thus employed the name *Cosmic Ray*. Consequently, to this day, the deceptive name *ray* is given also to massive particles. Testing the hypothesis that CRs can in fact be massive, charged particles, the possible deflection by the Earth's magnetic field was examined. John Clay and Arthur Compton confirmed an increased flux of CRs in moderate European latitudes with respect to tropical regions. Furthermore, an East-West asymmetry, as predicted by Bruno Rossi [7], was confirmed shortly thereafter. This excluded light, negatively charged particles such as electrons and left protons as natural candidates for the primary particles of the cosmic radiation. Later balloon experiments then also identified alpha particles and heavier nuclei as primary particles of the cosmic radiation [8]. Still, an inexplicably strong penetration power was observed from the overall cosmic radiation. This penetration power could have not been achieved by protons, which because of their relatively large mass ($m_p \simeq 938\,\mathrm{MeV}$) lose much of their energy through ionization, and also not by electrons, which lose a significant amount of their energy due to bremsstrahlung because of their relatively small mass ($m_e \simeq 0.511\,\mathrm{MeV}$) [9]. Rossi extended previous measurements of CRs using a triangular grid of scintillators which were covered by absorbers of lead and iron of different thickness [10]. An increase of coincident signals in the detectors was observed whilst the scintillators were shielded. He correctly concluded that a CR produces secondary particles of which some have an extremely strong penetration power and whose abundance in the overall particle cascade increases with the zenith angle of the primary particle[1]. Recent findings in elementary particle physics of that time then further completed the picture of the cosmic radiation by the discovery of the muon, μ. The mass of the muon ($m_\mu \simeq 1.06\,\mathrm{MeV}$) protects it from excessive energy loss due to both radiation and ionization at the observed energies. Identifying the muon as part of the secondary cosmic radiation, i.e. being produced in collisions of CRs with the Earth's atmosphere, easily explains the penetrative power of the cosmic radiation.

In 1937 Pierre Auger and his collaborators investigated the phenomenon of long-distance coincidences of muons and electrons produced in particle cascades by CRs, named air showers [11]. Not only did they confirm the existence of large extensive air showers (see Chapter 3), but they concluded that the primary particles accountable for these events must carry an energy above $10^{15}\,\mathrm{eV}$. Furthermore, the differential spectrum of the CRs was measured to be a power law of the approximate form $\mathrm{d}N/\mathrm{d}E \propto E^{-2}$. Already then the question of the origin of the highest energy CRs was thrown up also by Pierre Auger, stating: "[...] it is actually impossible to imagine a single process able to give to a particle such an energy".

2.2 THE COSMIC-RAY SPECTRUM

Using modern detectors that are much more sophisticated than Pierre Auger's, the CR spectrum can be measured with remarkable precision. The differential energy spectrum

[1] It became obvious shortly after that Rossi had discovered the air shower phenomenon for which Pierre Auger is often credited for. Rossi, however, assumed that the detected particle cascades are mainly produced upon impact of the CRs with the walls of houses and similarly dense matter, because he conducted his experiment indoors.

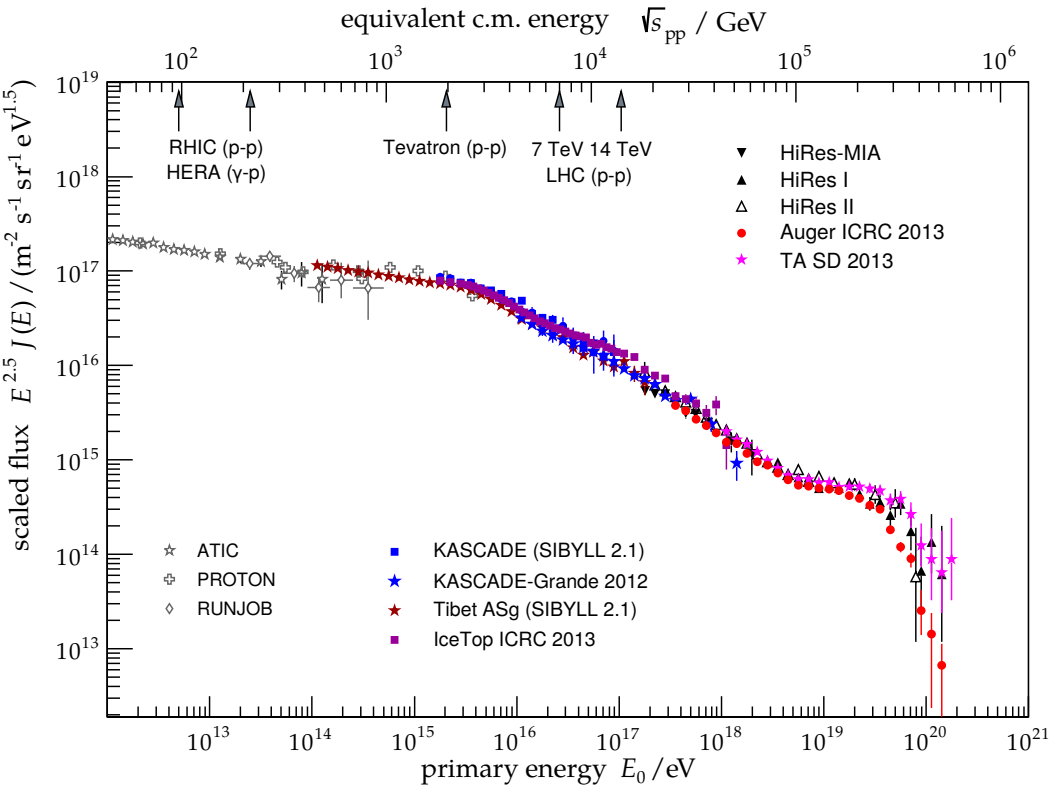


Figure 2.1: The differential energy spectrum of cosmic rays as measured by several experiments, reproduced from Ref. [2]. The equivalent energy for proton-proton collisions in the laboratory system is depicted on the upper x-axis, including the center-of-mass (c.m.) energy from several colliders.

of CRs can be well described by a steeply falling piece-wise power law, $\mathrm{d}N/\mathrm{d}E \propto E^{-\gamma}$, for which the spectral index γ changes several times with increasing energy [12]. In Fig. 2.1 the all-particle flux of high-energy CRs is depicted as measured by several experiments. To better visualize the behavior, especially at the highest energies, the flux is scaled by $E^{2.5}$. In this way, the spectrum can be illustrated over several orders of magnitude in energy. The energies at which the behavior of the flux deviates from a global power law appear as breaking points. The two most obvious characteristic energies of the depicted spectrum are known as the *knee* at $E_{\mathrm{knee}} \simeq 3{\times}10^{15}\,\mathrm{eV}$, where the spectrum softens, and the *ankle* at $E_{\mathrm{ankle}} \simeq 5{\times}10^{18}\,\mathrm{eV}$, where the spectrum hardens again. A more subtle second knee is emerging at energies around $E_{\mathrm{2nd\,knee}} \simeq 1{\times}10^{17}\,\mathrm{eV}$, which is associated with a change in the composition of the primary particles [13]. The flux is suppressed at energies above $E_{\mathrm{GZK}} \simeq 4{\times}10^{19}\,\mathrm{eV}$ (see Section 2.2.2). CRs in the energy region of the suppression are of particular interest, firstly, because on present theoretical grounds it remains challenging to find suitable accelerators to produce particles of these energies and secondly, because the details of the cutoff remain unclear even in pure GZK scenarios (see Section 2.2.2), if the composition of the UHECRs is not known [1]. The composition at these highest energies needs to be determined to complete the picture of ultra-high energy cosmic rays (UHECRs). The flux of UHECRs, however, is so low that on average only one particle is expected to be detected by a detector of 1 km^2 size every century.

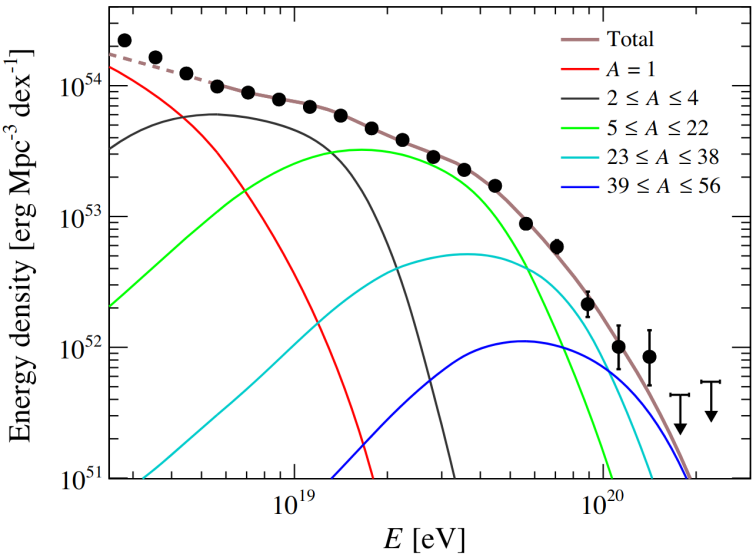


Figure 2.2: Best fit results from the combined fit of the composition and spectrum of UHECRs, see Ref. [16]. The dashed curve shows the region where data is not considered in the fit and an additional (possibly galactic) component is needed to explain the all-particle flux. Black dots mark the all-particle flux of CRs as measured by the Pierre Auger Observatory.

2.2.1 THE HIGHEST-ENERGY COSMIC RAYS

The highest-energy CRs are of particular interest for several reasons. Firstly, in the region of E_{ankle}, a transition from galactic to extragalactic sources occurs, which was only recently confirmed by the detection of a large-scale dipole anisotropy of the arrival direction of CRs [14]. Secondly, at the ankle the spectrum hardens and then softens significantly [15], implying a possible change of the acceleration mechanisms, injection spectra, or source distributions for UHECRs whilst a single local source scenario is disfavoured by data [16]. A scenario of universally distributed sources accelerating protons with the ankle as a consequence of the energy loss due to pair production is disfavoured as well, since the transition region would then be proton dominated. A proton only or proton and helium composition in the ankle region, however, is excluded at the $6.4\,\sigma$ level. The benchmark scenario obtained from the best fit to both the spectrum and the mass-sensitive data, depicted in Fig. 2.2, implies the presence of a subdominant light component at the highest energies of the all-particle spectrum of CRs. It is these light particles that are of special interest, since they are less charged and thus not as deflected by the extragalactic magnetic fields as the heavier, strongly charged ones and therefore might point directly towards their sources. An illustration of simplified trajectories of light and heavy CRs is given in Fig. 2.3. These sources, however, should be within our galactic neighborhood due to the GZK cutoff that is explained in the following.

2.2.2 THE GZK CUTOFF

Soon after the discovery of the cosmic microwave background (CMB) in 1965 [17], it was suggested that the Universe is opaque to the highest-energy cosmic rays [18, 19]. The process of UHECRs interacting with photons of the CMB is named after its discoverers, Greisen, and Zatsepin and Kuzmin and therefore known as the GZK effect or GZK cutoff.

Photons of the CMB appear blue-shifted in the center-of-mass system with the ultra-high energy protons so that they interact producing an intermediate Δ^+-baryon, which de-

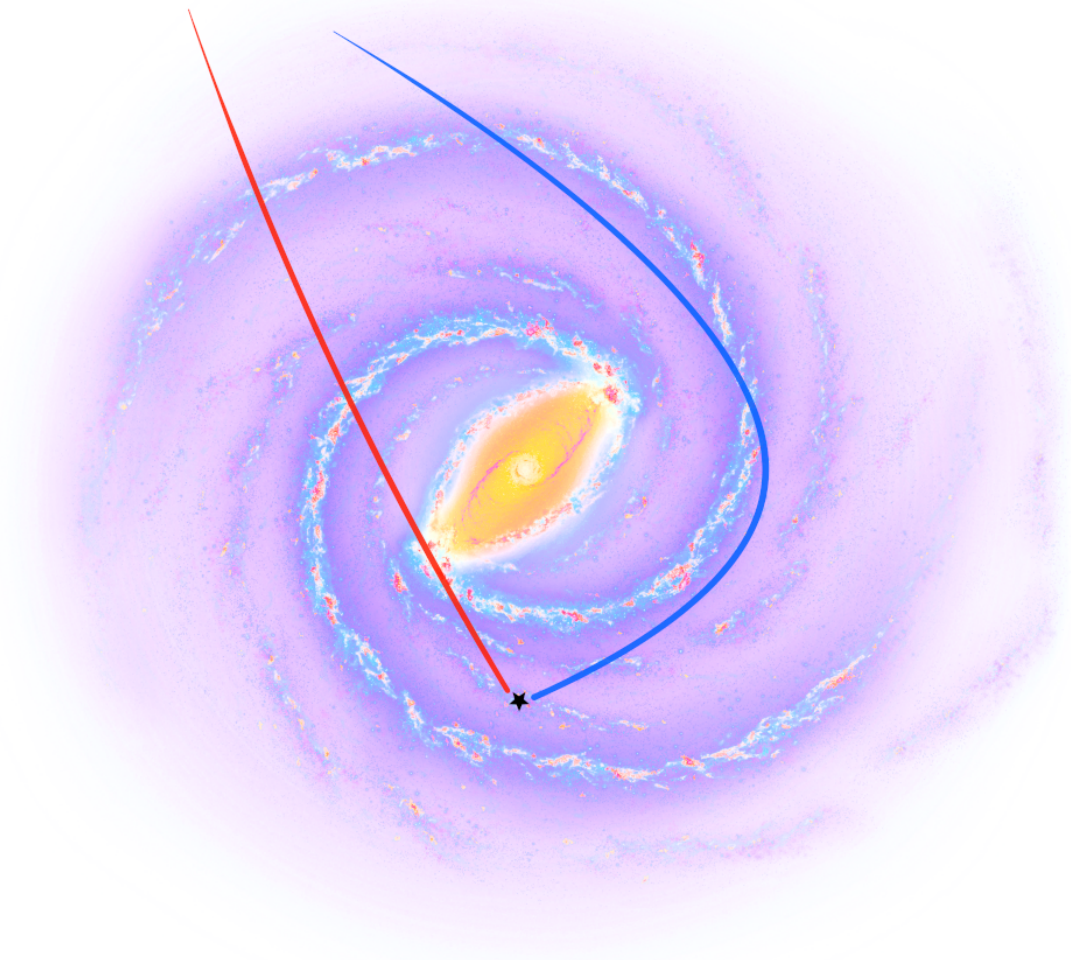


Figure 2.3: Artistic impression of extragalactic CRs arriving at the Solar System. The trajectory of a light CR is depicted in red, whereas a strongly curved trajectory, corresponding to a heavier, more charged CR, is depicted in blue. The approximate position of the Solar System in the Milky Way is given as black star. Background image: NASA/JPL-Caltech (colors modified).

cays into a neutral pion and a proton,

$$\mathrm{p} + \gamma_{\mathrm{CMB}} \longrightarrow (\Delta^{+})^{*} \longrightarrow \mathrm{p} + \pi^{0}. \tag{2.1}$$

In this process the neutral pion in the final state carries approximately 20 % of the energy of the UHECR in the initial state, thus softening it significantly. The power spectrum of the CMB resembles an almost perfect black body spectrum, see Fig. 2.4 (a), resulting in an energy-dependent number density of CMB-photons that can be given by [20]

$$n_{\gamma_{\mathrm{CMB}}}(E_{\gamma_{\mathrm{CMB}}}) = n_0 \frac{E_{\gamma_{\mathrm{CMB}}}^2}{\mathrm{e}^{\frac{E_{\gamma_{\mathrm{CMB}}}}{k_{\mathrm{B}}T}} - 1}, \tag{2.2}$$

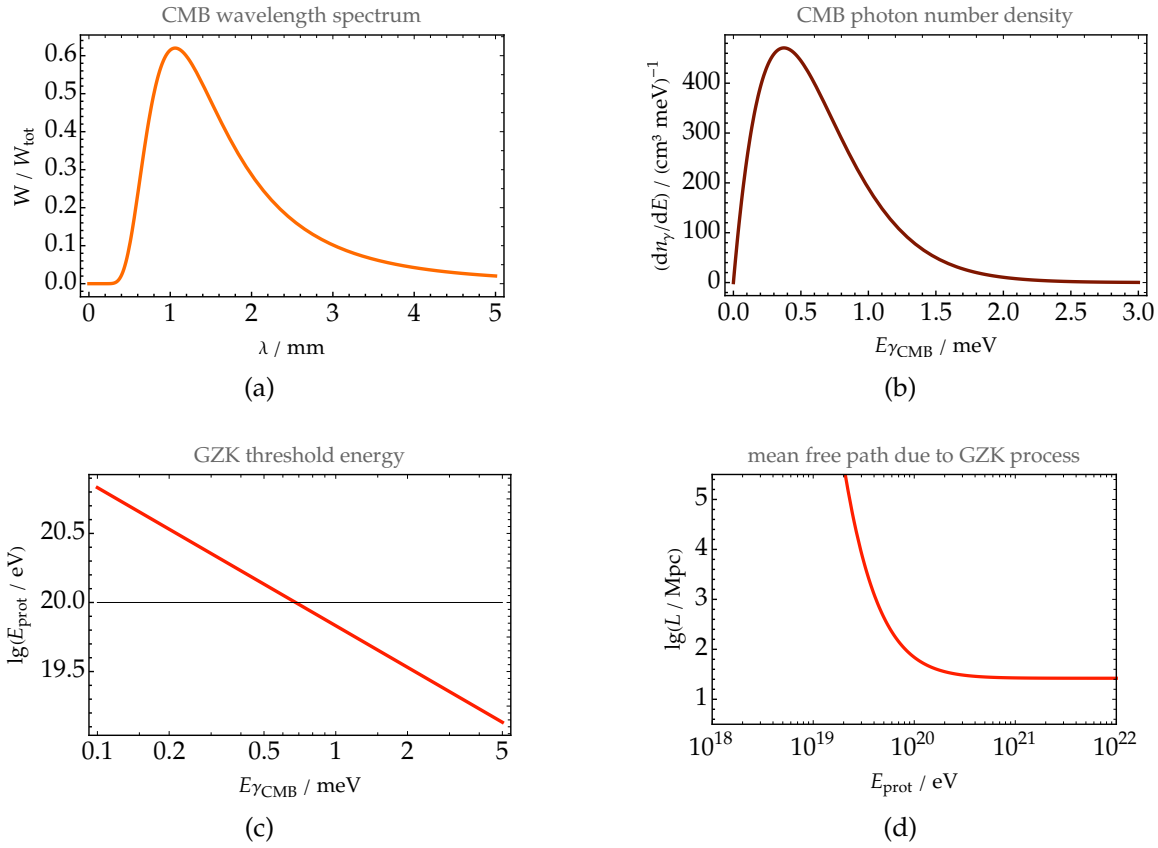


Figure 2.4: CMB spectra and characteristic limits of the GZK cutoff for protons, showing (a) the normalized wavelength spectrum of the CMB-photons, (b) the differential number spectrum of the CMB-photons with respect to the energy, (c) the threshold energy for ultra-high energy protons to produce a neutral pion in the final state by interacting with CMB-photons, and (d) the approximated mean free path L as a function of the proton energy, E_{prot}.

see Fig. 2.4 (b), with a normalization constant n_0 that depends on the power density of the CMB. With increasing energy of the UHECR, even the lowest energy CMB-photons allow for the process of neutral pions, π^0, to be produced in the laboratory system. The threshold energy of UHECRs to interact with a CMB-photon of the energy $E_{\gamma_{\text{CMB}}}$ is given by

$$E_{\text{CR}} = \frac{\frac{1}{4}m_{\pi^0}^2\,(2m_{\text{CR}} + m_{\pi^0})^2 - E_{\gamma_{\text{CMB}}}^2 m_{\text{CR}}^2}{E_{\gamma_{\text{CMB}}} m_{\pi^0}(2m_{\text{CR}} + m_{\pi^0})} \approx \frac{1}{2}\frac{m_{\text{CR}}\, m_{\pi_0}}{E_{\gamma_{\text{CMB}}}}, \tag{2.3}$$

using natural units, $c \equiv 1$. The approximation on the right-hand side is valid for nuclei CRs, i.e. $m_{\text{CR}} \gg m_{\pi_0} \gg E_{\gamma_{\text{CMB}}}$. The threshold energy for protons as a function of $E_{\gamma_{\text{CMB}}}$ is depicted in Fig. 2.4 (c). The number of CMB-photons which are sufficiently blue shifted to partake in the process of Eq. (2.1) thus increases with E_{CR}. Consequently, the mean free path of an UHECR decreases rapidly with energy. The highest-energy CR protons should therefore originate from our galactic neighborhood within a sphere on the radius on the order of the mean free path[2] $L \simeq 60\,\text{Mpc}$, see Fig. 2.4 (d). For heavier particles the threshold energy to interact with CMB photons of a certain energy is higher, therefore the mean free path due to this process becomes significantly larger. Other processes such as photo-disintegration of heavy nuclei, however, then become important [21].

[2] $1\,\text{pc} \simeq 3.1\times10^{16}\,\text{m} \simeq 3.3\,\text{ly}$

Similarly to the process described in Eq. (2.1), a slightly heavier but charged pion can be produced, leaving an ultra-high energy neutron instead of a proton in the final state,

$$p + \gamma_{CMB} \longrightarrow (\Delta^+)^* \longrightarrow n + \pi^+. \tag{2.4}$$

Being uncharged, however, neutrons would propagate on rectilinear trajectories away from their source region. A source emitting ultra-high energy protons that convert into neutrons by the GZK effect would produce a simple signature of many proton-like CRs from a single celestial direction. Since free neutrons have a limited lifetime due to the β decay, neutrons in the uppermost region of the CR spectrum have a mean free path of $\sim$10 kpc to 100 kpc only. The radius of our Galaxy is approximately 15 kpc, meaning respective sources from within our galaxy should be visible in their neutron signature. No such source, however, has yet been identified [22, 23].

2.3 EXPERIMENTS

In the past there have been numerous experiments to detect CRs, which are historically reviewed in detail for example in Ref. [24]. CR experiments can be classified into two main categories. The first type, indirectly measuring CRs by observing the particle cascades that occur when CRs interact with the atmosphere (see Chapter 3), and the second type directly detecting the CR. Indirect measurements usually make use of ground based detector arrays or telescopes, whereas the direct measurements typically make use of balloon-borne probes or satellites that operate either high up or even outside of the atmosphere.

Instruments for direct detection usually employ a calorimeter that gets ionized on impact of a CR. Because of the steeply falling flux of the CRs and the limited size of such devices, these kind of experiments are not suited to detect the highest-energy, extragalactic CRs and thus are not directly relevant for the context of this work. Still, they provide precise information on the nature and origins of CRs up to an energy of $\mathcal{O}(10\,\text{TeV})$. The balloon-borne ATIC experiment, for example, was equipped with a thin layered calorimeter and a segmented tracking region which allowed for the redunant idendification of the mass of the particles and their charges. ATIC gained significant exposure to CRs during multiple flights in Antarctica of one month of duration, each. Among hadronic CRs, ATIC observed an excess of positrons of supposedly cosmic origin at energies between 300 GeV to 800 GeV that could be the decay products of dark-matter particles [25]. A more advanced experiment operating outside of the Earth's atmosphere is the Alpha Magnetic Spectrometer (AMS), which is mounted on the International Space Station (ISS). AMS also combines a tracker as well as a calorimeter to detect the momentum and the charge of CRs. Furthermore, AMS uses a ring-imaging Cherenkov detector to additionally and independently determine the charge and the velocity of CRs. Being constantly in orbit, the experiment gained much more exposure than balloon-borne instruments, however at lower energies. Therewith AMS confirms the measurements of ATIC of an excess of positrons [26]. Recent discoveries of antihelium-3 candidates in AMS data further imply that some CRs are candidates for the decay products of dark matter [27].

Measurements of the particle cascades, or air showers, of CRs that interact with the atmosphere allowed for the detection of UHECRs already in the early days of CR physics. The Volcano Ranch experiment used a hexagonal array of scintillators to detect the footprints of air showers that result in temporal and spatial coincidences of the signal in the detector array. With this technique, the first particle with an energy above 10^{20} eV was detected at the Volcano Ranch in 1962 [28]. A method of directly observing the air shower as it develops in the atmosphere was employed by the Fly's Eye experiment [29]. If the primary particle

exceeds an energy of $\sim 10^{17}$ eV, a detectable amount of fluorescence light in the upper region of the visible spectrum[3] is emitted by excited nitrogen molecules in the atmosphere. The Fly's Eye experiment and its direct successors used fluorescence telescopes made up of several independent pixels to directly detect the brightness and the development the fluorescence light emitted by air showers. With this setup, a CR with a reconstructed energy of $(3.2 \pm 0.9) \times 10^{20}$ eV[4] was detected in 1991 [30]. This particle, carrying the nickname "*Oh-my-god-particle*", remains the highest-energy particle known to mankind until today. Besides not being able to determine the nature (i.e. the mass and charge) of the CR and besides the large uncertainty on the reconstructed energy, it was estimated that the particle should have originated from a source within a distance of approximately 30 Mpc. Roughly a decade after the detection of the "Oh-my-god-particle" the AGASA experiment reported the absence of a cutoff in the CR spectrum at the highest energies (not shown in Fig. 2.1), using a surface-detector array that covered an area of approximately 100 km^2 [31]. These measurements fuelled the idea of *top-down* scenarios, in which very heavy theoretical particles decay into leptons and quarks [32]. These quarks then hadronize and thus directly produce UHECRs without any preceding acceleration mechanism. The CR flux as reported by AGASA, however, is at the highest energies invalidated by measurements of modern experiments and top-down scenarios are now disfavoured [33].

Another noteworthy ground-based experiment is KASCADE, and its subsequent extension KASCADE-Grande [34]. The stations of the surface detector array of KASCADE used a multi-layer detector system with scintillators and lead to discriminate the muons produced in extensive air showers from other (charged) particles. With this technique the *knee* (see Fig. 2.1) of the CR spectrum and especially its dependence on the primary particle composition was confirmed [35].

The largest experiment to detect UHECRs is the Pierre Auger Observatory (in short *Auger*), located in the southern hemisphere in Argentina. Auger combines a surface-detector array with several fluorescence telescopes to detect both the development of air showers in the atmosphere and their footprints on the ground. The 3000 km^2 surface-detector array is currently undergoing an upgrade under the name of *AugerPrime* [36]. The Pierre Auger Observatory is discussed in detail in Chapter 4. A similar but smaller experiment located in the northern hemisphere is the *Telescope Array* [37]. It also combines a surface-detector array with telescopes to obtain hybrid measurements of extensive air showers. Both the Telescope Array and the Pierre Auger Observatory are able to measure the flux of the highest-energy CRs given their size and exposure during more than 10 years and more than 15 years of data acquisition, respectively.

Future experiments, both ground-based and satellite-borne, are already planned, possibly extending the search for UHECRs in the near and distant future [38, 39].

2.4 POSSIBLE SOURCES OF UHECRS

Whilst the sources for most cosmic rays (up to $\sim 10^{18}$ eV) are identified, for example as Supernovae remnants [2], it was not yet possible to pinpoint single sources of UHECRs [1]. There are, however, some astrophysical objects treated as possible source candidates.

In the search for the sources of UHECRs the Hillas criterion is a handy tool, connecting the size and the magnetic field strength of a possible accelerator [40]. Assuming gradual acceleration, the Larmor radius is a physical limit to the size R of a cosmic accelerator which

[3] 300 nm to 400 nm
[4] 3.2×10^{20} eV $\simeq 10^{20.5}$ eV

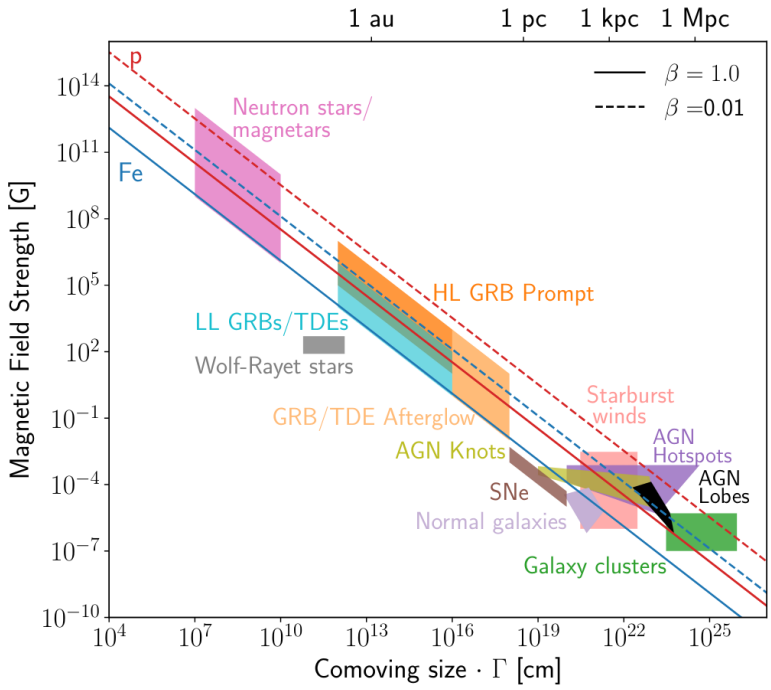


Figure 2.5: Hillas diagram, from Ref. [1]. Possible sources of UHECRs are depicted as function of their characteristic size and corresponding magnetic-field strength. The limits to accelerate protons and ionized iron nuclei by ultra-relativistic (non-relativistic) shock fronts to 10^{20} eV are depicted in solid (dashed) lines.

contains or creates a magnetic field B. A particle that exceeds the energy

$$E_{\max} = q\beta RB\Gamma \tag{2.5}$$

will inevitably escape the accelerator, where q is the charge of the particle, Γ is the Lorentz factor of the motion of the source and β the speed of the shock fronts accelerating the particle, given in units of the speed of light. Fig. 2.5 shows several possible sources of UHECRs as function of their characteristic sizes and magnetic-field strengths, of which we want to highlight three in the following.

Active galactic nuclei (AGNs), that usually host a super-massive black hole emitting relativistic jets, for example, are promising candidates as accelerators for UHECRs. Centaurus A, a radio galaxy with an AGN at a distance of approximately 5 Mpc, is one of the most promising candidates as a single source of UHECRs. A correlation of the arrival directions of UHECRs above an energy of 6×10^{19} eV with AGNs within a distance of 75 Mpc was first reported by the Auger Collaboration, but had to be critically revisited with the acquisition of more data over time [41, 42, 43]. Given the large size and relatively small magnetic field (about 6 orders of magnitude lower than the Earth's magnetic field), AGNs could possibly contain CRs for a long time, accelerating them gradually over millions of years until they escape as UHECRs. An AGN is known as *blazar* if the axis of its jets points towards the Earth. In this case it becomes an even more favourable source candidate since it appears brighter and the CRs would be minimally deflected by the magnetic fields of the AGN [44]. Furthermore, the signature of the UHECRs acceleration would be visible in the gamma-ray spectrum of the blazar. Even though the number of blazars in the galactic neighborhood does not suffice to produce enough UHECRs as seen in data, they are not ruled out as candidates in a scenario of multiple contributing accelerators.

Tidal Disruption Events (TDEs), which are events where a black hole devours a nearby star, usually in the nucleus of a galaxy, have also been recently suggested as possible candidates for UHECR-sources [45]. TDEs are very short-lived and bright events, only lasting a few years or even only a few months. During this time approximately half of the matter of the partaking star is absorbed by the black hole and the other half is emitted, either as relativistic jet or escaping from the accretion disk that is heated in this process. It was shown in Ref. [46] that blazar-like, jetted TDEs fulfill the necessary conditions to accelerate protons up to 10^{20} eV. However, only a few[5] TDEs, and even fewer jetted TDEs, are known at all, the closest candidate being at a distance of $\sim$150 Mpc [47, 48].

With only a few kilometers in diameter, fast rotating jetted neutron stars, known as pulsars or, given an extremely powerful magnetic field, magnetars, are the most compact source candidates for UHECRs. Their magnetic field of $\sim 10^{14}$ G to 10^{15} G [49] is so strong that its magnetic-energy equivalent mass density exceeds the heaviest known elements by several orders of magnitude. Since they are the remnants of massive stars they are rich in heavy elements and can therefore accelerate heavy nuclei to extreme energies [50]. After the preceding Supernova in which a pulsar is formed, the pulsar transfers a substantial amount of its matter and energy to its surroundings within a fraction of a year. In this process particles such as ionized nuclei are accelerated in the pulsar wind. Accelerated nuclei could then escape the system as UHECRs, possibly after only $\mathcal{O}(100)$ days.

In the search for the sources of UHECRs, multi-messenger astronomy is becoming the spearheading scientific discipline, combining the searches for gravitational waves, neutrinos, gamma rays, and CRs [1, 51, 23]. The coincidence of these events or, alternatively, the absence of one or multiple of these messengers will give an indication on the mechanisms to be held accountable for the production UHECRs. It is therefore of utter importance to be able to identify possible light CRs and consequently their origin using data obtained by ground arrays such as Auger.

[5] As of December 2021 only 98 TDEs, including TDE candidates, are known.

CHAPTER III

EXTENSIVE AIR SHOWERS

The particle cascade that develops in the atmosphere upon impact of a CR is known as *air shower*, named by Schmeiser and Bothe (originally in German as "*Luftschauer*") who discovered that the atmosphere acts as a calorimeter for CRs [52]. The discovery of large-scale coincidences of signals in ground detectors led to the name *extensive air shower* (EAS) [11, 53].

Extensive air showers develop along the extended trajectory of the CR, called the *shower axis*, where the *shower core*, which mainly consists of high-energy hadrons, fuels the particle cascade until it eventually hits the ground [2]. Most of the energy of the primary CR is deposited in the atmosphere by bremsstrahlung and ionization; only a small fraction is lost to weakly interacting particles (such as neutrinos) or to particles stopped in the ground. The creation, interaction, and the decay of particles in the air shower are in constant interplay with each other, balancing out at the *shower maximum*, when the maximum number of particles present at the same time is reached.

In this chapter, we explain the basics of extensive air showers and briefly introduce the essential mathematical formulae that are needed to describe the behaviour of the emerging particles. If not explicitly stated otherwise, the notation and the steps outlined in Refs. [2] and [54] are followed. Furthermore, we want to give a brief overview of the methods used to simulate extensive air showers and the state-of-the-art computing techniques used for that purpose.

3.1 BASICS OF CASCADE THEORY

Particle cascades are the result of high-energy particles interacting with matter. The matter acts as calorimeter in which the energy of the initiating particle is converted. In the case of air-shower physics, the calorimeter is given by the atmosphere. Despite the fact that the density of the atmosphere ρ is not homogeneous, the same principles apply for the atmosphere as for regular calorimeters. All distances and distance-dependent quantities can be converted into units of traversed matter, integrated along a (straight) path $\vec{r}$ of length r,

$$X(r) = \int_0^r \rho(\vec{r}(r'))\,\mathrm{d}r', \tag{3.1}$$

that would be proportional to distance given a homogeneous calorimeter. The depth X is usually expressed in units of $\mathrm{g\,cm^{-2}}$ and is measured from the entry point of the primary particle into the calorimeter, i.e. from the top of the atmosphere downwards along the shower axis. The development of the cascade of particles in the atmosphere can be

described by the differential equation

$$\frac{\mathrm{d}N_i(E_i,X)}{\mathrm{d}X} = -\frac{N_i(E_i,X)}{\lambda_i} - \frac{N_i(E_i,X)}{d_i} + \sum_{j=i}^{J} \int_{E_i}^{\infty} \frac{F_{ji}(E_i,E_j)}{E_i} \frac{N_j(E_j,X)}{\lambda_j} \,\mathrm{d}E_j, \tag{3.2}$$

where N_i describes the number of particles of type i at the energy E_i, which are present at the depth X. The characteristic distances, after which a particle of type i either interacts or decays, is given by the interaction length λ_i and the decay length d_i, respectively. The function F_{ji} describes the yield of a particle i with the energy E_i from an interaction of a particle j with energy E_j with an air nucleus.

The development of scenarios and methods to find solutions of Eq. (3.2) that accurately describe data has been challenging for years up to today, with various solutions and comparisons outlined for example in Refs. [54, 55, 56, 57, 58, 59, 60, 61] and [62]. The availability of the increasing computational power in the last decades allows for numerical solutions which are briefly discussed in Section 3.5. In simplified scenarios, ostensive and academic solutions of Eq. (3.2) can be obtained. Such could be, for example, a CR producing n_0 equal particles of energy E_0 instantly on impact, that decay after a mean free path d without contributing to the cascade, resulting in a simple exponentially-decreasing number density of particles, $N(X) = n_0 \exp(-X/d)$. A more realistic scenario describes the electromagnetic cascade of an air shower in such a way that Eq. (3.2) transforms into two coupled differential equations[1], which we discuss in the following.

3.1.1 THE CASCADE EQUATIONS

Assuming $\gamma(W,X)\,\mathrm{d}W$ is the number of photons, γ, within an energy range of $[W, W+\mathrm{d}W)$ at a depth X and $\pi(E,X)\,\mathrm{d}E$ is the number of electrons and positrons[2], $\mathrm{e}^{\pm}$, respectively. Electrons and positrons are treated as equivalent particles in this section, therefore from now on both electrons and positrons are referred to simply as electrons, e. The cascade equations read [54]

$$\frac{\mathrm{d}\gamma}{\mathrm{d}X} = -\frac{\gamma(W,X)}{\lambda_{\mathrm{pair}}} + \int_{W}^{\infty} \pi(E',X) \frac{\mathrm{d}^2 n_{\mathrm{e}\to\gamma}}{\mathrm{d}W\,\mathrm{d}X} \,\mathrm{d}E', \quad \text{and} \tag{3.3}$$

$$\frac{\mathrm{d}\pi}{\mathrm{d}X} = -\frac{\pi(E,X)}{\lambda_{\mathrm{brems}}} + \int_{E}^{\infty} \pi(E',X) \frac{\mathrm{d}^2 n_{\mathrm{e}\to\mathrm{e}}}{\mathrm{d}E\,\mathrm{d}X} \,\mathrm{d}E' + 2\int_{E}^{\infty} \gamma(W',X) \frac{\mathrm{d}^2 n_{\gamma\to\mathrm{e}}}{\mathrm{d}E\,\mathrm{d}X} \,\mathrm{d}W'. \tag{3.4}$$

This scenario is called *Approximation A*; all processes except for bremsstrahlung and pair production are neglected and Eqs. (3.3) and (3.4) are valid under the assumption of asymptotic screening of the atomic nuclei in the atmosphere. A schematic representation of the cascade is depicted in Fig. 3.1. The corresponding *Approximation B*, which is not considered here, takes into account the energy loss of electrons due to collisions with the calorimeter, thus adding another term to Eq. (3.4). The characteristic length for a particle to interact via bremsstrahlung or pair production is given as λ_{brems} and λ_{pair}, respectively. In this picture, each electron steadily radiates a photon whilst losing some of its energy after propagating one interaction length, and each photon interacts after one interaction length, producing an electron-positron pair. Thus, as long as the energy of the individual particles is sufficient,

[1] Also referred to as the *diffusion equations*.

[2] Although by this notation the $\mathrm{e}^{\pm}$ spectrum can be easily mistaken for the number of pions, we adhere to it because it is commonly used in the literature, thus following Ref. [54]. It is noteworthy that this notation precedes the discovery of the pion in 1947.

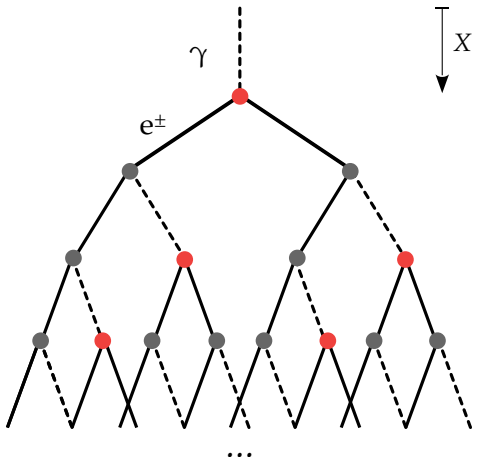


Figure 3.1: Schematic representation of an electromagnetic air shower cascade induced by a photon under Approximation A. Interactions of particles due to bremsstrahlung are indicated as gray vertices, interactions resulting in pair production are indicated as red vertices. Photons are depicted as dashed lines, electrons are depicted as full lines.

the number of particles doubles after each interaction length. The yield of particles of type i by the decay or interaction of particles of type j at the energy E and depth X is written as $n_{j\to i}(E, X)$, for which we introduce the notation in terms of the respective energy dependent probabilities for bremsstrahlung, ϕ, and for pair production, ψ,

$$\phi\left(\tfrac{W}{E'}\right) := E' \,\frac{\mathrm{d}^2 n_{\mathrm{e}\to\gamma}}{\mathrm{d}W\,\mathrm{d}X}, \tag{3.5}$$

$$\psi\left(\tfrac{E}{W'}\right) := W' \,\frac{\mathrm{d}^2 n_{\gamma\to\mathrm{e}}}{\mathrm{d}W\,\mathrm{d}X}, \quad \text{and} \tag{3.6}$$

$$\phi\left(1-\tfrac{E}{E'}\right) := E' \,\frac{\mathrm{d}^2 n_{\mathrm{e}\to\mathrm{e}}}{\mathrm{d}W\,\mathrm{d}X}. \tag{3.7}$$

Using the relative energy of photons and electrons with respect to each other, $v = W/E$ and $u = E/W$, Eqs. (3.3) and (3.4) can be written as

$$\frac{\mathrm{d}\gamma}{\mathrm{d}X} = -\frac{\gamma}{\lambda_{\mathrm{pair}}} + \int_0^1 \pi(W/v, X)\,\phi(v)\,\frac{\mathrm{d}v}{v}, \quad \text{and} \tag{3.8}$$

$$\frac{\mathrm{d}\pi}{\mathrm{d}X} = -\frac{\pi}{\lambda_{\mathrm{brems}}} + \int_0^1 \pi(E/(1-v), X)\,\phi(v)\,\frac{\mathrm{d}v}{1-v} + 2\int_0^1 \gamma(E/u, X)\,\psi(u)\,\frac{\mathrm{d}u}{u}. \tag{3.9}$$

As it can be seen by the denominator, the second term on the right-hand side of Eq. (3.9) is logarithimically divergent for $v \to 1$. The cross section for bremsstrahlung,

$$\frac{1}{\lambda_{\mathrm{brems}}} = \int_0^1 \phi(v)\,\mathrm{d}v, \tag{3.10}$$

however, is infrared divergent in such a way that the expression can be regularized. Consequently, the difference of the two terms and thus Eq. (3.9) remains finite. Since Approximation A is only valid above a certain energy, the divergences occurring at the lower integration limits can be treated with a cutoff.

3.1.2 ELEMENTARY SOLUTIONS UNDER APPROXIMATION A

It can be shown by substitution that Eqs. (3.8) and (3.9) have solutions of the form

$$\pi(E, X) = F_\pi(E)\, f(X) \quad \text{and} \quad \gamma(W, X) = F_\gamma(W)\, f(X) \tag{3.11}$$

such that the dependence on energy and depth factorizes. Upon substitution into the cascade equations we obtain

$$F_\pi(E)\,\frac{\mathrm{d}f(X)}{\mathrm{d}X} = f(X)\,\left[-\frac{F_\pi(E)}{\lambda_{\text{brems}}} + \int_0^1 F_\pi(E/(1-v))\,\phi(v)\,\frac{\mathrm{d}v}{1-v} + 2\int_0^1 F_\gamma(E/u)\,\psi(u)\,\frac{\mathrm{d}u}{u}\right], \tag{3.12}$$

$$F_\gamma(W)\,\frac{\mathrm{d}f(X)}{\mathrm{d}X} = f(X)\,\left[-\frac{F_\gamma(W)}{\lambda_{\text{pair}}} + \int_0^1 F_\pi(W/v)\,\phi(v)\,\frac{\mathrm{d}v}{v}\right], \tag{3.13}$$

where the expressions in square brackets are only dependent on energy. For the function $f(X)$ this implies that the ratio

$$\lambda := \frac{1}{f}\frac{\partial f}{\partial X} \tag{3.14}$$

is constant in X at a fixed energy and therefore

$$f(X) \propto \mathrm{e}^{\lambda X}. \tag{3.15}$$

At this point it is clear that solutions to the cascade equations exist, in which the number of particles change exponentially with depth, while the energy spectra and relative abundance of photons and electrons remain unchanged, as already mentioned earlier in Section 3.1. In a realistic scenario, however, the energy spectra of particles are not constant and will change with the development of the shower[3] so that λ becomes a function of the *shower age* (see Section 3.1.3).

It can be shown again by direct substitution that using the factorization of Eq. (3.11) the Eqs. (3.8) and (3.9) are satisfied by power laws of the form

$$F_\pi(E) = a\,E^{-(s+1)} \quad \text{and} \quad F_\gamma(W) = b\,W^{-(s+1)}, \tag{3.16}$$

with a positive number s and constants a and b. Substituting Eq. (3.16) in Eqs. (3.12) and (3.13) yields

$$\lambda(s)\,a = -A(s)\,a + B(s)\,b, \quad \text{and} \quad \lambda(s)\,b = -C(s)\,a - \frac{b}{\lambda_{\text{pair}}}, \tag{3.17}$$

[3] The relative rate of change $\lambda = (\partial f/\partial X)/f$ must not be mistaken for a length. This unfortunate choice of notation is again historically motivated, as for example used in Ref. [55].

where

$$A(s) = \int_0^1 (1 - (1-v)^s)\, \phi(v)\, \mathrm{d}v, \tag{3.18}$$

$$B(s) = 2\int_0^1 u^s\, \psi(u)\, \mathrm{d}u, \quad \text{and} \tag{3.19}$$

$$C(s) = \int_0^1 v^s\, \phi(v)\, \mathrm{d}v. \tag{3.20}$$

The numerical values of these expressions can be calculated exactly, given the respective energy-loss length of the electrons and the characteristic length for photons to produce an electron pair are known. The energy-loss length of charged particles, or *radiation length*, X_0, is the mean length after which only a fraction of $1/\mathrm{e}$ of the initial energy remains. For particles with an initial energy E_0, the remaining energy at the depth X is given by

$$E(X) = E_0\, \mathrm{e}^{-\frac{X}{X_0}}, \tag{3.21}$$

such that E/X_0 is the energy loss per step length, if energy loss due to ionization is neglected. For electrons in air the radiation length is $X_0 \simeq 37\,\mathrm{g\,cm^{-2}}$. It is convenient to express all distances and lengths in terms of X_0, so that the depth is written as $t := X/X_0$. The unit-less cross section for pair production in air then reads $\sigma_0 := X_0/\lambda_{\mathrm{pair}} \approx 7/9$. Using approximate expressions[4] for the radiation probabilities ϕ and ψ, the expressions for $A(s)$, $B(s)$, and $C(s)$ can be approximated as

$$A(s) = \mu_{\mathrm{E}} + \frac{\mathrm{d}}{\mathrm{d}s} \ln\left(\Gamma(s+1)\right), \tag{3.22}$$

$$B(s) = 2\frac{\sigma_0}{s+1}, \quad \text{and} \tag{3.23}$$

$$C(s) = \frac{1}{s}, \tag{3.24}$$

where $\mu_{\mathrm{E}} \simeq 0.5772$ is the Euler-Mascheroni constant and $\Gamma(x)$ is the gamma function. Given the expressions of Eqs. (3.22) to (3.24), explicit solutions for $\lambda(s)$ can be found. $\lambda(s)$ must satisfy the quadratic equation that is obtained from the condition given in Eq. (3.17),

$$\lambda^2 + (A(s) + \sigma_0)\,\lambda + (A(s)\sigma_0 - B(s)C(s)) = 0. \tag{3.25}$$

The roots of Eq. (3.25) are given by

$$\lambda_1(s) = -(A(s) + \sigma_0) + \sqrt{(A(s) - \sigma)^2 + 4B(s)C(s)}, \quad \text{and} \tag{3.26}$$

$$\lambda_2(s) = -(A(s) + \sigma_0) - \sqrt{(A(s) - \sigma)^2 + 4B(s)C(s)}. \tag{3.27}$$

Thus, the elementary solutions of the cascade equations assuming a power law energy spectrum are given by

$$\pi(E,t) = a_1\, E^{-(s+1)}\, \mathrm{e}^{\lambda_1(s)\,t}, \quad \gamma(W,t) = \frac{a_1 C(s)}{\sigma_0 + \lambda_1(s)}\, W^{-(s+1)}\, \mathrm{e}^{\lambda_1(s)\,t} \tag{3.28}$$

$$\pi(E,t) = a_2\, E^{-(s+1)}\, \mathrm{e}^{\lambda_2(s)\,t}, \quad \gamma(W,t) = \frac{a_2 C(s)}{\sigma_0 + \lambda_2(s)}\, W^{-(s+1)}\, \mathrm{e}^{\lambda_2(s)\,t}, \tag{3.29}$$

where a_1 and a_2 are constants to be defined by boundary conditions and λ_1 and λ_2 are given by Eqs. (3.26) and (3.27).

[4] Expressions can be found in Ref. [54], §13, Eqs. (1.32) to (1.34).

3.1.3 EXAMPLE SOLUTIONS

Ensemble of Particles

Solutions to the cascade equations of course depend on the respective boundary conditions. A trivial example shall be given in the following. If n photons are injected, for example, at the depth $X = 0$, following an energy spectrum with the spectral index $s = 1.7$, the solution is given by $A = 1.4$, $B = 0.6$, $C = 0.6$, $\lambda_1 = -0.4$, and $\lambda_2 = -1.8$, according to Eqs. (3.22) to (3.24). The spectrum of particles at a depth X is thus given by

$$\frac{\mathrm{d}n_{\gamma+\mathrm{e}^\pm}}{\mathrm{d}E}(X) \approx 1.18\, n\, E^{-2.7} \mathrm{e}^{-\frac{X}{85\,\mathrm{g\,cm^{-2}}}}. \tag{3.30}$$

Single Primary Particle

Difficulties arise if one tries to solve Eqs. (3.8) and (3.9) under arbitrary boundary conditions or for different injection spectra, such as for example a single photon with an energy of W_0 injected at $X = 0$,

$$\gamma(W, 0) = \delta(W - W_0). \tag{3.31}$$

For these scenarios the Mellin transformation $\mathfrak{M}$ of the function f,

$$\mathfrak{M}_f(s) = \int_0^\infty \mathrm{d}E\, E^s f(E), \tag{3.32}$$

is a handy tool to derive the solutions for $\pi(E, t)$ and $\gamma(W, t)$. The reason for this is that $\mathfrak{M}_f$ can be more easily determined than the original functions π and γ, which are then given by the inverse transformation, $\mathfrak{M}_f^{-1}$. One can already identify terms of the form of Eq. (3.32) in the right-hand side of Eqs. (3.3) and (3.4), using Eqs. (3.28) and (3.29). Solutions to δ-distribution boundary conditions, like air showers induced by single particles, are thus convolutions of the elementary solutions with the energy spectra from the respective boundary conditions. To demonstrate this the spectrum of electrons resulting from a photon following Eq. (3.31) is calculated in the following. The Mellin transformation of $\pi(E, t)$ given a single incident photon as described by Eq. (3.31) inserted in Eq. (3.4) reads as

$$\mathfrak{M}_\pi(s, t) = \frac{B(s)}{\lambda_1(s) - \lambda_2(s)} W_0^s\, \mathrm{e}^{\lambda_1(s)\, t}. \tag{3.33}$$

The inverse transformation of Eq. (3.33) is given by

$$\pi(E, t)\, \mathrm{d}E = \frac{\mathrm{d}E}{E} \frac{1}{2\pi\mathrm{i}} \int_{-\mathrm{i}\infty+s_0}^{\mathrm{i}\infty+s_0} \mathrm{d}s \left(\frac{B(s)}{\sqrt{s}(\lambda_1(s) - \lambda_2(s))} \right) \exp\left[\lambda_1(s)\, t + s \ln \frac{W_0}{E} + \frac{1}{2} \ln s \right]. \tag{3.34}$$

This integral cannot be solved exactly but only approximately. For the sake of briefness only the essential steps of a possible procedure are outlined here; the details of the explicit calculation can be found in Chapter 15 of Ref. [2] or alternatively in §28 and Appendix A of Ref. [54]. The argument of the exponent is expanded to the second order in a Taylor series around the point $s = \bar{s}$, which solves the equation

$$\frac{\mathrm{d}}{\mathrm{d}s} \left(\lambda_1(s)\, t + s \ln \frac{W_0}{E} + \frac{1}{2} \ln s \right) = 0, \tag{3.35}$$

so that the integral can be treated with the saddle-point method. Furthermore, to simplify the resulting expressions, λ_1 can be approximated [62] as

$$\lambda_1(s) \approx \frac{1}{2}\left(s - 1 - 3\ln s\right), \tag{3.36}$$

and for $t \gg 1$ the terms containing λ_2 can be safely neglected. The term in parenthesis of Eq. (3.34) is assumed to be constant under variation of s and is approximated by its value at $s = \bar{s}$. The resulting integral can be understood as a Gaussian integral that can be evaluated integrating through the saddle point $\bar{s}$,

$$\begin{aligned}\pi(E,t)\,\mathrm{d}E &\approx \frac{\mathrm{d}E}{E}\frac{1}{2\pi\mathrm{i}}\frac{B(\bar{s})}{\sqrt{\bar{s}}\,\lambda_1(\bar{s})}\int_{-\mathrm{i}\infty+\bar{s}}^{\mathrm{i}\infty+\bar{s}}\mathrm{d}s\,\exp\left[\lambda_1(s)t + sy + \frac{1}{2}\ln s\right] \\ &\approx \frac{\mathrm{d}E}{E}\frac{B(\bar{s})}{\sqrt{2\pi}\left(\bar{s}\left(\frac{3}{\bar{s}^2}t - \frac{1}{2\bar{s}^2}\right)^{\frac{1}{2}}\lambda_1(\bar{s})\right)}\exp\left[\lambda_1(\bar{s})t + \bar{s}y + \frac{1}{2}\ln\bar{s}\right],\end{aligned} \tag{3.37}$$

using $y := \ln(W_0/E)$.

The expression obtained for the number of electrons of energy E present at a depth t from a shower induced by a single photon yields two important implications. First, for $E \ll W_0$, π has a maximum with respect to t, approximately when the argument of the exponent of Eq. (3.37) is at its maximum itself, so at the ordinate

$$t_{\max} = y - 1 \approx y. \tag{3.38}$$

This implies the existence of a maximum of the number of particles of a certain energy present at a certain depth. In the longitudinal profile of the shower,

$$\Pi(t) = \int_{E_{\mathrm{cut}}}^{W_0}\pi(E,t)\,\mathrm{d}E, \tag{3.39}$$

which includes all particles with an energy larger than E_{cut}, this maximum is referred to as the *shower maximum*. Second, the parameters s and t are connected by the relation

$$s = \frac{3t-1}{t+2y} \approx \frac{3t}{t+2y}, \tag{3.40}$$

as implied[5] by Eqs. (3.35) and (3.36). This means, approximately, the energy dependence and the dependence on depth of the particle spectrum π can both be expressed in terms of s, which is identified as the *shower age*, or alternatively the shower-age parameter. With increasing shower age s the spectrum quickly steepens and high-energy particles become less abundant in the particle content of the cascade. Generally, the shower age s depends heavily on the boundary conditions of the cascade, and is related to the fractional rate of change of the particles in the cascade. It can be expressed as

$$s(t) = \lambda_1^{(-1)}\left(\frac{1}{\Pi(t)}\frac{\mathrm{d}\Pi(t)}{\mathrm{d}t}\right), \tag{3.41}$$

[5] If the same calculation is done for the differential photon spectrum, or for the electron spectrum as induced by a single electron, the constant "-1" in Eqs. (3.38) and (3.40) is missing. Calculating the photon spectra from a single incident electron, this addend becomes "$+1$"; in general it is safe to follow the approximations in Eqs. (3.38) and (3.40) [2].

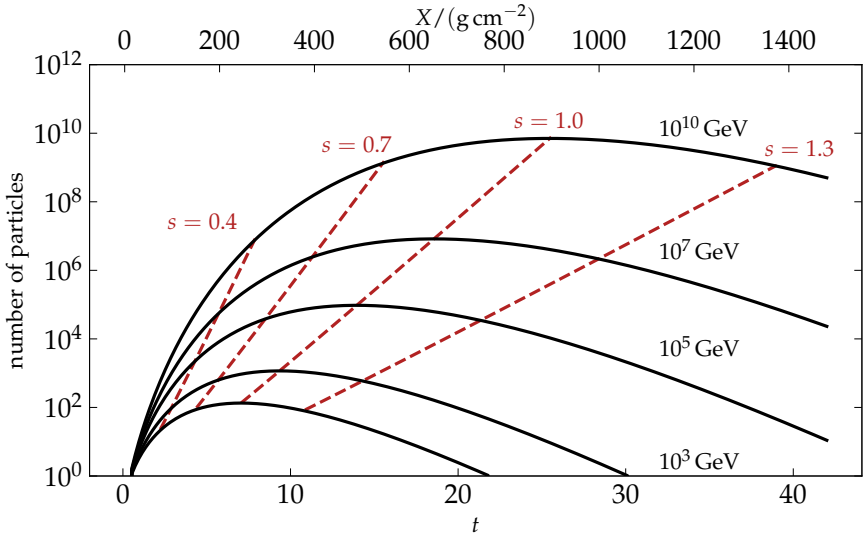


Figure 3.2: The Greisen profile as given in Eq. (3.43) for several values of E_0, ranging from 100 GeV to 10^{10} GeV. The regions of equal shower age for the different profiles are depicted in red for several values of s, where $s = 1$ marks the shower maxima. The upper x-axis shows the absolute slant depth for a mean radiation length of $X_0 \simeq 37\,\mathrm{g\,cm^{-2}}$. The profiles are shown for a total of 42 radiation lengths.

where $\lambda_1^{(-1)}$ is the inverse function of Eq. (3.36) [63]. In general, $s > 0$, with $s = 1$ at the maximum of the cascade and $s > 1$ after the maximum. The fact that the overall development of an air shower can be described solely by the shower age is discussed in detail in Chapter 5. Using the approximation Eq. (3.36) the exponent of Eq. (3.37) can be simplified and the corresponding longitudinal profile of electrons with an energy larger[6] than $E_{\mathrm{cut}} = 100\,\mathrm{MeV}$ can be written as

$$\Pi(t) \approx (2\pi)^{-\frac{1}{4}} \frac{\sigma_0 + \lambda_1(s)}{\lambda_1(s)} \frac{\mathrm{e}}{\sqrt{1.5t+1}} \exp\left[t\,\left(1 - \tfrac{3}{2}\,\ln s\right)\right], \tag{3.42}$$

where s depends on $y = \ln(W_0/E_{\mathrm{cut}})$. A considerable amount of approximations is necessary to arrive at Eq. (3.42), which technically only describes the number of electrons present at the depth t and only for energies $E \gtrsim 100\,\mathrm{MeV}$. Still, the qualitative behaviour of Π with respect to t accurately describes the electromagnetic cascade of air showers.

3.1.4 THE GREISEN PROFILE

To reconstruct the energy of the primary particle starting the cascade of the air shower, the expected number of particles N that can be detected at the ground must be known as a function of the primary energy. Given, for example, a detector array of slightly shielded scintillators such as employed by the Volcano Ranch (see Section 2.3), electrons and positrons with an energy of $E \simeq 20\,\mathrm{MeV}$ and higher would deposit a signal in the surface detectors. It is therefore safe to treat $N(E > 20\,\mathrm{MeV})$ as the total number of electrons. Given such low energies, however, it is expected that particles can no longer be produced in the cascade, i.e. the probabilities for bremsstrahlung and pair production become $\phi = \psi = 0$. This point is reached once the constant energy loss of the individual particles due to ionization is considerably larger than the energy loss from radiation. For the electromagnetic particle cascade

[6] This lower limit in energy is chosen such that Approximation A – neglecting energy loss due to ionization – is well justified.

this energy is $\epsilon_c^{e\gamma} \simeq 87\,\text{MeV}$ and it is assumed that particles of $E \ll \epsilon_c^{e\gamma}$ do not contribute to the cascade anymore. This detail is crucial, especially for the late development of the cascade, where on average the particles carry less energy than $\epsilon_c^{e\gamma}$. Thus, to accurately describe the late development of the cascade also Coulomb scattering and ionization must be taken into account, and Approximation A is no longer valid. The Greisen profile is an approximation for the longitudinal development of electromagnetic cascades, similarly to Eq. (3.42), but with the absolute number of particles calculated under Approximation B of the cascade equations, thus also taking into account the critical energy limit for which the particles do not contribute to the cascade anymore [62]. In the Greisen profile function the depth of the shower maximum is related to the energy of the primary particle and the critical energy by $\beta_0 = \ln(E_0/\epsilon_c^{e\gamma})$ and $y = \beta_0$. Thus, according to the shower age of the Greisen profile as given in Eq. (3.40), the shower age s is mainly driven by low-energy particles. The Greisen profile takes the form

$$N(t) = 0.31\ \beta_0^{-\frac{1}{2}}\ \exp\left[t\ \left(1 - \tfrac{3}{2}\ \ln s\right)\right]. \tag{3.43}$$

Eq. (3.43) was not strictly derived in Ref. [62] but rather numerically motivated, an a-posteriori derivation is given in Ref. [63]. A possible way to derive the Greisen profile is given in Appendix C. In Fig. 3.2 the expected number of electrons and positrons as a function of radiation lengths according to Eq. (3.43) is depicted for several values of E_0. The shower maximum occurs at the shower age $s = 1$, again implying $t_{\text{max}} = y$, and thus implying the depth of the shower maximum

$$X_{\text{max}} = X_0 \ln \tfrac{E_0}{\epsilon_c^{e\gamma}}. \tag{3.44}$$

The total expected number of particles at the shower maximum is given by

$$N_{\text{max}} = 0.31\ \left(\ln \tfrac{E_0}{\epsilon_c^{e\gamma}}\right)^{-\frac{1}{2}}\ \frac{E_0}{\epsilon_c^{e\gamma}} \approx \frac{5.4\,E_0/\text{GeV}}{\sqrt{\ln\left(11.5\,E_0/\text{GeV}\right)}} \approx E_0/\text{GeV}, \tag{3.45}$$

as it can be verified[7] in Fig. 3.2.

Although the right-hand side of Eq. (3.45) yields a good approximation for the order of magnitude of the number of particles produced in an extensive air shower, especially for energies $E_0 \gtrsim 10^{16}\,\text{eV}$, an important observation must be pointed out here. The number of particles produced in an air shower scales not linearly, but slightly less than linearly with the energy of the primary particle, which is an important aspect of Section 3.4.

3.2 LATERAL SPREAD

So far, only the longitudinal development of particles along the shower axis has been considered. In reality, however, particles are emitted under a finite angle α with respect to the shower axis and are scattered during their propagation through the atmosphere. Thus, the air shower will inevitably also develop laterally. In fact, as it has been already mentioned in Section 2.3, although most particles stay concentrated in the shower core, the overall footprint of an air shower at the ground level can reach several kilometers in diameter. Kamata and Nishimura were the first to accurately and analytically describe the angular and lateral distributions of electromagnetic particles in air showers [64, 65, 66]. For that purpose the cascade equations are augmented by additional terms regarding the distance to the shower

[7] The approximation on the right-hand side of Eq. (3.45) is only valid for $E_0 \gtrsim 10^7\,\text{GeV}$.

axis r and the angle α, and are then solved under Approximation B. The mathematical treatment of the three-dimensional cascade equations comprises numerous steps and methods that were specifically derived and performed in a series of three articles; for the sake of briefness they will not be repeated here. Moreover, the functional forms of the lateral and angular distributions derived by Kamata and Nishimura are so unwieldy that they have not been used by the scientific community for experimental purposes.

An approximative form of the lateral distribution of particles derived in Ref. [66] was introduced by Greisen [67], it is thus known as the *Nishimura-Kamata-Greisen*, or NKG function. The NKG function accurately describes the lateral density of the electromagnetic particles of an air shower and is widely used until today, albeit sometimes in modified versions from its original form [68]. The NKG function reads as

$$\varrho(r) = \frac{N}{2\pi r_{\mathrm{M}}^2} \frac{\Gamma(\frac{9}{2} - s)}{\Gamma(s)\Gamma(\frac{9}{2} - 2s)} \left(\frac{r}{r_{\mathrm{M}}}\right)^{s-2} \left(1 + \frac{r}{r_{\mathrm{M}}}\right)^{s-\frac{9}{2}}, \tag{3.46}$$

with the shower age s as given in Eq. (3.40) and the total number of electrons N. The distance r_{M} is the Molière unit that characterizes the lateral spread of particles due to Coulomb scattering. Approximately 90% of the electromagnetic particles in an air-shower are contained within a distance of one r_{M} from the shower axis. For low energy particles, the equivalent column density of r_{M} is $\sim 10\,\mathrm{g\,cm^{-2}}$, which is about 100 m in air at the typical altitudes of modern detector array experiments. Generally r_{M} is given by

$$r_{\mathrm{M}} = \frac{E_{\mathrm{s}}}{\epsilon_{\mathrm{c}}} \frac{X_0}{\rho}, \tag{3.47}$$

where E_{s} is the typical energy loss of the particles due to scattering above the critical energy ϵ_{c}, and ρ is the density of the calorimeter [69]. For electrons the scattering loss is approximately $E_{\mathrm{s}} \simeq 27\,\mathrm{MeV}$. For the NKG function, Greisen proposes a distance-dependent shower age of the form [62]

$$s = \frac{3t}{t + 2\ln\left(E_0/\epsilon_{\mathrm{c}}\right) + 2r/r_{\mathrm{M}}}. \tag{3.48}$$

Data collected by surface detector arrays have shown already in the early days of CR physics that the lateral profile as seen by detectors does not fully agree with Eq. (3.46). An empirical functional form for the lateral profile of all shower particles combined was suggested in Ref. [67],

$$\varrho(r) = \frac{0.4\,N}{r_{\mathrm{M}}^2} \left(\frac{r}{r_{\mathrm{M}}}\right)^{-0.75} \left(1 + \frac{r}{r_{\mathrm{M}}}\right)^{-3.25} \left(1 + \frac{r}{11.4\,r_{\mathrm{M}}}\right), \tag{3.49}$$

where the deviation from Eq. (3.46) is mainly due to muons and anti-muons[8] produced in hadronic air showers. As implied by Eq. (3.47), muons in the air shower behave very differently from electrons, they lose less energy from scattering ionization and thus have a significantly larger radiation length. The lateral and longitudinal distribution of the particles in an air shower that also contains muons is therefore expected to behave very differently from purely electromagnetic[9] showers. Therefore, in the following the general particle content of extensive air showers is discussed.

[8] Anti-muons are treated analogously to muons and are hence also represented by the symbol μ.

[9] "Purely electromagnetic particles" refers to electrons, positrons, and photons.

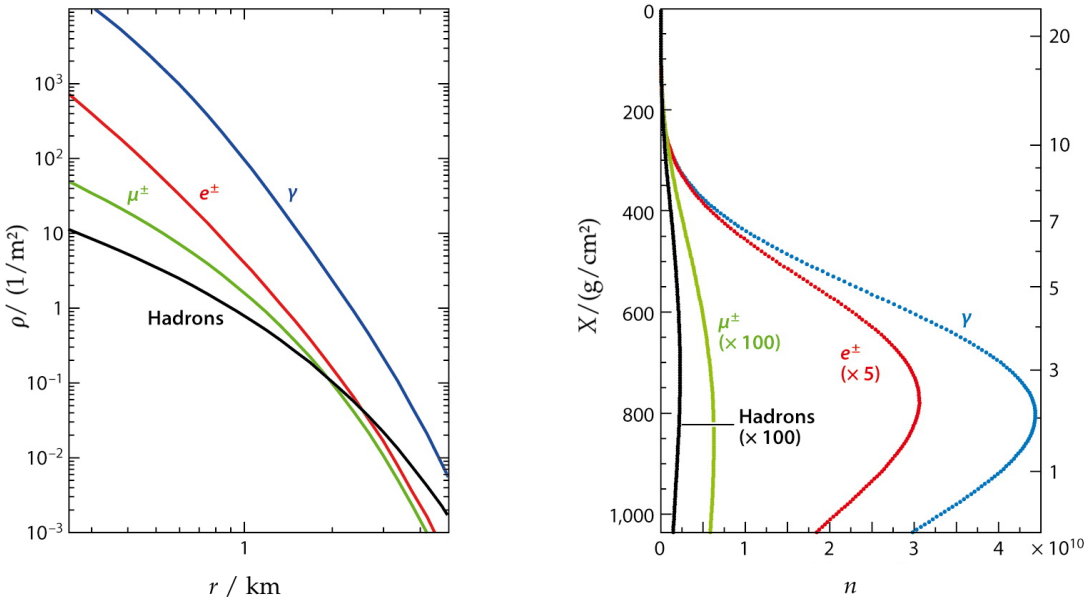


Figure 3.3: Qualitative lateral and longitudinal distributions of particles in an air shower induced by a proton with an energy of 10^{19} eV. Panel (a) illustrates the lateral distribution of the areal density of different types of particles per m^{-2} at the depth of $X = 870\,\mathrm{g\,cm}^{-2}$ while panel (b) illustrates the total number of particles present at the depth X. All electromagnetic particles above an energy of 0.25 MeV and muons and hadrons above 0.1 GeV are taken into account. Adapted from Ref. [12].

3.3 PARTICLE CONTENT

In the previous sections, mostly electromagnetic showers, i.e. showers consisting only of electrons, positrons, and photons, have been considered. Air showers induced by hadronic UHECRs, however, contain all kinds of particles. Given the extreme center-of-mass energy of the first interaction all possible Standard Model particles and interactions[10] must in general be considered in Eq. (3.2). In this case, however, the equations are not analytically solvable anymore and studies on the particle content are mostly driven by simulations (see Section 3.5).

The predominant amount of particles in an air shower is mostly categorized by three particle components: the electromagnetic, the muonic, and the hadronic component, of which the electromagnetic component is by far the largest in terms of numbers of particles and energy content [12]. Fig. 3.3 illustrates the longitudinal and lateral distributions of the different particle components for an air shower induced by proton UHECR. Given a hadron as primary particle, the first interactions are governed by the laws of quantum chromodynamics (QCD). Even though significant progress in the understanding of QCD was made in the last decades, our knowledge of QCD is mostly limited to processes with large momentum transfer, i.e. hard interactions, while in "soft" QCD interactions (i.e. interactions with relatively low momentum transfer) many open questions are still unanswered.

In the first interactions of a hadronic air shower, the most frequently produced particles are charged and neutral pions, $\pi^\pm$ and π^0. Whilst neutral pions almost immediately

[10] Of course also theories beyond the Standard Model could be considered in air-shower physics, however, given that there is currently no method for detecting rare interactions in air showers, these could be tested only qualitatively.

decay into two photons, the charged pions are rather long-lived, propagating several tens of meters in the atmosphere. Heavier and short-lived mesons and baryons are only found in the shower core, where they fuel the hadronic cascade. Secondary hadrons produced in the shower core typically carry a large transverse momentum with respect to the shower axis, therefore producing a relatively wide distribution of particles, as shown in Fig. 3.3 (a). Muons are mostly produced by the decay of charged pions (with a small contribution from the decay of heavier mesons, such as kaons). Their relative abundance therefore peaks in the early cascade and their overall abundance is a good estimator for the amount of hadrons produced in the first interactions. The muonic particle component is rather long-lived and therefore its longitudinal profile is stretched out with respect to the other particle components, as can be seen in Fig. 3.3 (b). Muons propagate rectilinearly without significant energy loss and reach the surface almost unattenuated. For this reason they play a prominent role in surface detector experiments, since they can be detected even when produced in showers induced by CRs at large zenith angles, where most other particles do not reach the ground. As mentioned earlier, the electromagnetic component is by far the dominant particle component of extensive air showers in terms of particle multiplicity and energy deposit. Photons are produced in the cascade continuously by bremsstrahlung, whilst electrons are mostly produced by pair-production and the decay of muons. Since the longitudinal and lateral profiles of the photon and electron component are almost identical, up to a scaling factor, they are often treated as equal particles in air-shower physics.

Neutral and weakly interacting particles such as neutrinos, ν, produced in the decay of pions and muons, do not contribute significantly or not at all to the development of the particle cascade. Their calorimetric energy deposit is neglected and thus they are referred to as *invisible* particles. Above $E_0 \simeq 10^{19}\,\mathrm{eV}$ approximately 10% to 15% of the total energy of the primary CR is lost to invisible particles, so that in good approximation the primary energy is converted by the ionization of the atmosphere or the ground [70, 71].

3.4 THE HEITLER-MATTHEWS MODEL

A simple version of the electromagnetic cascade discussed in Section 3.1 was introduced in Ref. [72] and later extended for hadrons, muons, and neutral particles in Ref. [73]. This model, which is known as the Heitler-Matthews model of hadronic air showers, is a handy tool to describe the interplay of the different particle components in hadronic showers. Stochastic processes, such as the production and decay of particles, are dealt with using average values. The (re-)production of particles is modelled by simple splitting rules, so that an explicit solution of the cascade equation is circumvented. Especially the implications for the expected number of muons are of importance. In the following, the Heitler-Matthews model as introduced in Ref. [73] is briefly discussed.

3.4.1 A MODEL OF HADRONIC CASCADES

A cascade is initiated by a proton with the energy E_0, producing pions in a hadronic interaction. It is assumed that all types of pions, π^+, π^-, and π^0, are produced with equal probability. Thus, in each interaction N_{ch} charged pions $\pi^\pm$ and $N_{\mathrm{ch}}/2$ neutral pions π^0 are produced. After a constant interaction length, $\lambda_{\mathrm{int}} \simeq 120\,\mathrm{g\,cm^{-2}}$, charged pions interact. Neutral pions promptly decay into two photons and therefore do not contribute further to the hadronic cascade; charged pions continue to interact until they reach the critical energy of $\epsilon_{\mathrm{c}}^{\pi} \simeq 20\,\mathrm{GeV}$ where they decay into muons. All particles of the electromagnetic component are treated as equal, splitting into two after a constant interaction length of $\lambda_{\mathrm{int}} = X_0 \ln 2$, until they reach the critical energy of $\epsilon_{\mathrm{c}}^{\mathrm{e}\gamma}$. Below an energy of $\epsilon_{\mathrm{c}}^{\mathrm{e}\gamma}$, collisional

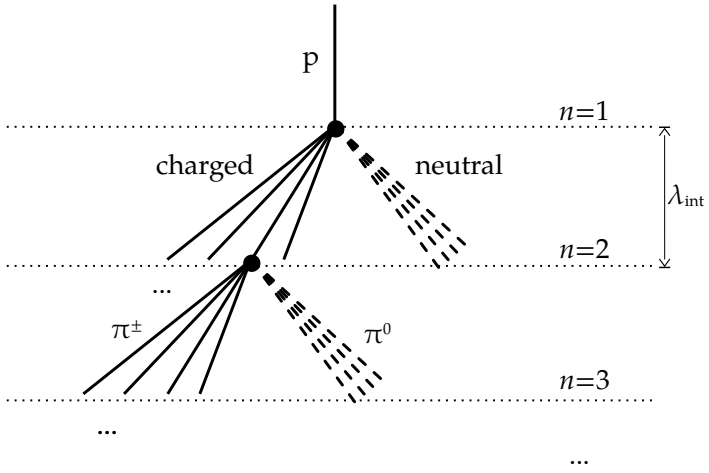


Figure 3.4: Schematic representation of the pionic cascade induced by a proton according to the Heitler-Matthews model. Charged hadrons are depicted in full lines, neutral hadrons are depicted in dashed lines. Adapted from Ref. [2].

energy loss is dominant and suppresses the production of new particles. Energy is always divided evenly among all created particles, energy loss due to radiation or scattering is neglected[11]. A sketch of the hadronic cascade according to the aforementioned splitting rules is given in Fig. 3.4. The energy budget of the shower components after n generations is therefore given by

$$E_{\text{had}} = \left(\frac{2}{3}\right)^n E_0, \qquad E_{\text{e}\gamma} = \left(1 - \left(\frac{2}{3}\right)^n\right) E_0. \tag{3.50}$$

Thus, after 10 interactions only about 1% of the primary energy is left in the hadronic component. The energy of the individual charged pions after n generations is given by

$$E_\pi = \frac{E_0}{\left(\frac{3}{2} N_{\text{ch}}\right)^n}. \tag{3.51}$$

Consequently, the number of interactions necessary for all pions to reach $E_\pi \lesssim \epsilon_{\text{c}}^\pi$ is given by

$$n_{\text{c}} = \left\lceil \frac{\ln\left(E_0/\epsilon_{\text{c}}^\pi\right)}{\ln\left(\frac{3}{2} N_{\text{ch}}\right)} \right\rceil, \tag{3.52}$$

where $\lceil x \rceil$ is the operation of rounding up to an integer value. Since in this model muons are exclusively produced by the decay of charged pions in the last generation, we are left with

$$N_\mu = (N_{\text{ch}})^{n_{\text{c}}} = \left(\frac{E_0}{\epsilon_{\text{c}}^\pi}\right)^{1-\beta}. \tag{3.53}$$

The exponent[12] is given by

$$1 - \beta = \frac{\ln N_{\text{ch}}}{\ln\left(\frac{3}{2} N_{\text{ch}}\right)}. \tag{3.54}$$

[11] Using the splitting rules of the model so far, the result from Section 3.1.3 Eq. (3.38) concerning the maximum of the shower, that is $X_{\text{max}} = X_0 \ln(E_0/\epsilon_{\text{c}}^{\text{e}\gamma})$ can already be reproduced for a purely electromagnetic shower, using that in total $N_{\text{max}} = \exp\left[X_{\text{max}}/\lambda_{\text{int}}\right] = 2^{n_{\text{c}}}$ particles are produced in $n_{\text{c}} = \lceil \ln(E_0/\epsilon_{\text{c}}^{\text{e}\gamma})/\ln 2 \rceil$ steps until all particles reach the critical energy.

[12] In Ref. [73] the notation $(1-\beta) \to \beta$ is used.

Even though no interactions or decay of muons is considered so far, it is clear from Eq. (3.53) that for $\beta > 0$ the number of muons in an air shower is expected to scale slower than linearly with the primary energy of the CR.

The model provides not only a qualitative explanation for the amount of muons produced in an air shower, but also the depth of the shower maximum[13] can be in the Heitler-Matthews model given in terms of N_{ch}. Let X_1 be the depth at which the primary proton first interacts with the atmosphere, then $N_{\mathrm{ch}}/2$ neutral pions are produced, which decay into a total of N_{ch} photons. Each photon then carries an energy of $E_0/(3\,N_{\mathrm{ch}})$ according to Eq. (3.50). Thus, N_{ch} electromagnetic cascades of the same energy are initiated, which produce an overall shower maximum at the depth

$$X_{\mathrm{max}} = X_1 + X_0 \ln\left(\frac{E_0}{3\,N_{\mathrm{ch}}\,\epsilon_{\mathrm{c}}^{\mathrm{e}\gamma}}\right), \tag{3.55}$$

following the implications of Section 3.1.4.

Obviously, the number of charged pions produced in each interaction, N_{ch}, and thus also β, heavily affects all solutions of the discussed model. N_{ch} is a parameter that is governed by the laws of soft QCD and is thus best obtained from simulations. N_{ch} and β are in general dependent on the invariant mass of the (first) interaction and thus on the energy of the primary particle. Numerical values of β obtained from simulations reach from 0.2 down to $\sim$0.07 [2, 73, 74], implying values of N_{ch} between 5 and 220, respectively. These results are in accordance with the corresponding values for β obtained from the method discussed Section 6.4.

3.4.2 IMPLICATIONS FOR DIFFERENT PRIMARY PARTICLES

Considering heavier primary particles to start a Heitler-Matthews-like cascade, according to the superposition model a nucleus of mass A can be treated as A independent protons each with an energy of E_0/A. In this simplistic picture, the resulting A cascades develop simultaneously and individually with respect to each other, however, converting the same amount of total energy as a corresponding proton-induced cascade.

The expected amount of particles N and the expected depth of the shower maximum for showers induced by nuclei with mass number A can be obtained using the substitutions

$$E_0 \to E_0/A, \quad \text{and} \quad N_i^{(A)} \to A\,N_i^{\mathrm{p}}, \tag{3.56}$$

for any particle component i. The respective expressions for the shower maximum, according to Eq. (3.55), and for the expected amount of muons, according to Eq. (3.53), are thus

$$\begin{aligned} X_{\mathrm{max}}^{(A)} &= X_1 + X_0 \ln\left(\frac{E_0}{3\,A\,N_{\mathrm{ch}}\,\epsilon_{\mathrm{c}}^{\mathrm{e}\gamma}}\right) \\ &= X_{\mathrm{max}}^{\mathrm{p}} - X_0 \ln A, \end{aligned} \tag{3.57}$$

and

$$\begin{aligned} N_{\mu}^{(A)} &= A\;\; N_{\mu}^{\mathrm{p}}\big|_{E_0 \to \frac{E_0}{A}} = A\,\left(\frac{E_0}{A\,\epsilon_{\mathrm{c}}^{\pi}}\right)^{1-\beta} \\ &= A^{\beta}\,N_{\mu}^{\mathrm{p}}. \end{aligned} \tag{3.58}$$

On average, considering CRs of the same energy E_0, a shallower shower with more muons is therefore expected from heavier CRs.

[13] Since all air showers are heavily dominated by the electromagnetic component, the shower maximum refers to the maximum of the electromagnetic shower, if not explicitly stated otherwise.

The change in the mean depth of the shower maximum with respect to the order of magnitude of the primary energy, $\mathrm{d}\langle X_{\max}\rangle / \mathrm{d}\lg E_0$, is known as the *decadal elongation rate*. From Eq. (3.57) it is clear that if $\langle A\rangle$ changes over the energy, the mean of the distribution of $X_{\max}$ will also change its behaviour. The evolution of $\langle X_{\max}\rangle$ over several orders of magnitude in energy and possible changes in the elongation rate can thus give information about the average primary mass of CRs. This holds even if fragmentation of heavier nuclei is considered in the first interactions [75].

The relative muon content in an air shower is defined as

$$R_\mu = \frac{N_\mu}{\langle N_\mu^{\mathrm{p}}\rangle}, \tag{3.59}$$

where $\langle N_\mu^{\mathrm{p}}\rangle$ is to be understood as the energy-dependent expected total number of muons at the surface level from a proton-induced air shower[14]. Depending on β, iron-induced air showers, which contain about 30% to 40% more muons than proton-induced air showers, are for R_μ expected to have values around 1.3 to 1.4 for R_μ, while by definition showers induced by protons have $\langle R_\mu\rangle = 1$.

While the Heitler-Matthews model makes a statement only on the average number of particles expected in air showers, the involved processes are highly stochastic. Shower-to-shower fluctuations are thus expected for $X_{\max}$ as well as for R_μ around their corresponding mean values for individual primary particles. In the superposition model, due to Poissonian statistics, the spread of values for $X_{\max}$ is expected to be smaller for heavier primary particles, where multiple individual electromagnetic cascades are involved, than for protons [76]. In both cases the fluctuations depend on the processes involved in the hadronic interactions. Based on simulations, the spread of R_μ from showers of the same primary particle is expected to be $\sim$20%, e.g. values between 0.8 and 1.2 for showers induced by protons [77]. To study and accurately describe shower-to-shower fluctuations simulations of air showers are necessary, as is discussed in the following section.

3.5 SIMULATIONS

Extensive air showers involve highly stochastic processes in the creation and annihilation of the individual particles. Even a complete model, solving a full set of cascade equations according to Eq. (3.2) that includes all Standard Model processes and particles, could not accurately describe the shower-to-shower fluctuations expected in extensive air shower events. The lateral and longitudinal profiles of particles obtained from simulations were studied as a counterpart to the analytical solutions of Sections 3.1.4 and 3.2 as early as the available computing power allowed for it [78]. Simulations that employ the current knowledge of soft QCD by the usage of state-of-the-art models of hadronic interactions yield an irreplaceable added value to air-shower physics [77, 79]. Simulated air shower data are produced either by a pure Monte Carlo (MC) simulation of particles interacting and decaying over time, or by a hybrid simulation that in the process numerically solves the cascade equations. Examples for these kinds of simulation frameworks are CORSIKA [80] and CONEX [81], respectively. Both methods have their own advantages and shortcomings. While pure MC simulations like CORSIKA will accurately reproduce the fluctuations expected from real air-shower events, a vast computing effort is required. For complete MC simulations of extensive air showers, the CPU-time scales approximately linearly with the energy of the primary particle. To overcome this inconvenience, only a representative subset of particles is simulated, weighted, and finally multiplied in a process called thinning

[14] Another definition of the relative muon content R_μ, which can often be found in literature is normalized to $\langle N_\mu^{\mathrm{p}}(E_0 = 10^{19}\mathrm{eV})\rangle$ so that R_μ scales differently with energy, namely linearly with $(E_0/10^{19}\,\mathrm{eV})^{1-\beta}$.

and de-thinning. Hybrid simulation approaches, as employed by CONEX, for example, lack the need for extensive CPU time, but will not be able to account for all possible fluctuations or sub-showers, nor produce instances of single particles instead of expectation values and densities, like full MC simulations do.

3.5.1 EVENT GENERATORS

To produce simulated data, MC frameworks have to employ a model of hadronic interactions as an event generator. In the scattering processes of hadrons, the string model, describing a flux of constant color charge, explains the formation of hadrons by the splitting of strings and the formation of quark-antiquark pairs [82]. This process is in general referred to as hadronization. The cross sections for elastic and inelastic scattering of particles, the treatment of the string formation and breaking, and the employed particle properties drastically effect the final state of particles in simulations. At the highest energies, many details in the treatment of this parameters, however, are still unclear, given the lack of experimental data at above-LHC energies. Three prominent event generators used in air-shower physics are EPOS-LHC [83], QGSJET-II [84], and SIBYLL [85]. Combined with a MC simulation framework, each of them produces slightly different outcomes for the fluctuations and the mean values of observables, such as the shower maximum, the number of muons produced in air showers, or the distribution of the pseudorapidity of particles, as studied in Refs. [79] and [86]. Especially the muons produced in air-showers are of importance in this context, since the different event generators produce notably different amounts of muons for comparable showers. Additionally, the expected lateral distribution of muons is different for each model [79]. Comparing EPOS-LHC and QGSJET-II, for example, EPOS-LHC produces more muons on average than QGSJET-II, which is the model producing the smallest amount of muons in air shower simulations. This is discussed in detail in Chapter 6.

Even though qualitatively the electromagnetic particle component behaves similarly for the three hadronic interaction models, expected distributions of $X_{\max}$ can vary significantly not only between different models but also between different versions of the same model[15]. Therefore, the three aforementioned hadronic interaction models have to be treated individually and with care, since different interpretations of data are possible according to implications from the different hadronic interaction models.

3.5.2 ELECTRON SPECTRA AS OBTAINED FROM SIMULATIONS

Besides accurately describing fluctuations in air showers, simulations can be used to obtain a realistic solution for the expected particle spectra without any strong approximations, such as which were necessary to obtain Eq. (3.37) in Section 3.1.3. The studies performed in Refs. [87, 88] and [89], for example, investigated the particle distributions with respect to space and energy based on full MC simulations. The electron spectrum π of extensive air showers was parametrized using a functional form based on the elementary solutions found in Ref. [54] and given in Eqs. (3.28) and (3.29). The simulated spectra as well as the resulting parametrizations are depicted in Fig. 3.5. It was shown that for $t \gg 0$ the electron spectra can be expressed only in terms of the shower age and the primary energy, independently of the type of the primary particle. This result is consistent with the implications from Sections 3.1 and 3.4. The universality of the electron spectra of air showers is of greater importance when considering the longitudinal profile in terms of energy deposit per step length in units of $\mathrm{g\,cm^{-2}}$. This is discussed in detail in Chapter 5.

[15] An outdated version QGSJET-II is known to produce a smaller elongation rate than current models and far fewer muons than expected in air-shower data [84]. If not explicitly mentioned, the version of QGSJET-II we refer to is the latest, QGSJET-II-04.

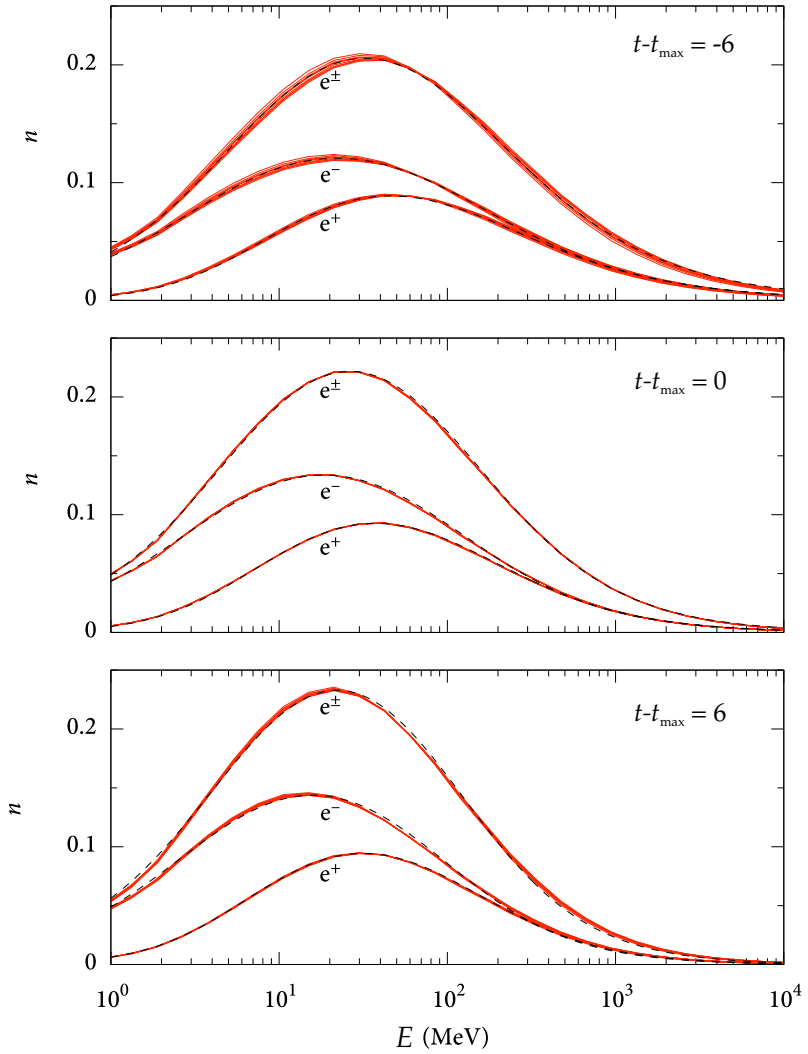


Figure 3.5: Normalized average electron and positron spectra as obtained from simulations for three different shower-evolution stages. Red curves represent simulated distributions for different primaries (p, Fe, and γ) and energies (10^{17} eV, 10^{18} eV, and 10^{19} eV). The corresponding parametrizations are plotted on top as dashed lines. Taken from Ref. [87].

CHAPTER IV

THE PIERRE AUGER OBSERVATORY

The Pierre Auger Observatory (*Auger* for short) is the largest observatory for CRs in the world [90, 91]. Located in the southern hemisphere, near the city of Malargüe in Argentina, the observatory is exposed to the galactic center. The observatory lies at an average elevation of 1400 m above sea level. With its fluorescence detectors (FD) and its surface detector array (SD), Auger employs two independent techniques to detect extensive air showers from CRs [92, 93]. The SD is made up of more than 1600 individual detector stations, covering an area of $\sim$3000 km^2. In clear moonless nights, the atmosphere over this area is overseen by the FD telescopes that are located in four buildings at the opposing edges of the SD. The interplay of the two detector systems allows for hybrid measurements of extensive air showers and for the cross calibration of the two detectors [94]. An SD station and FD building as well as the event display for an example hybrid event are depicted in Fig. 4.1

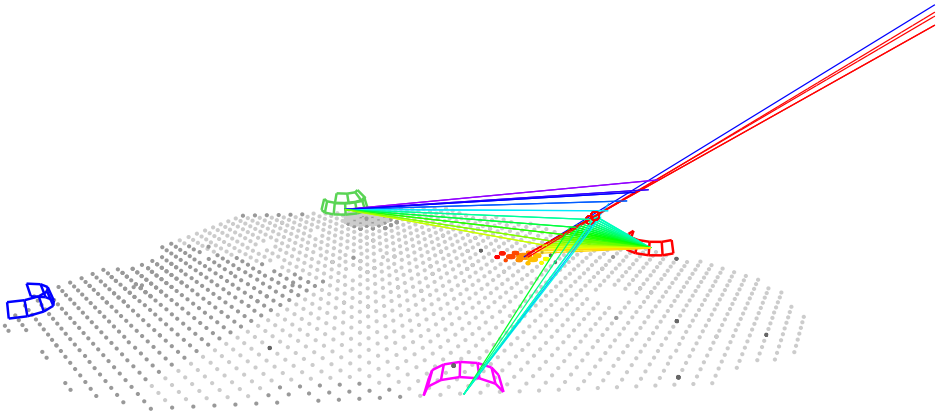
In the search for the highest-energy particles known to mankind the Auger observatory has already produced extraordinary results and is aspiring to produce more high-quality data thanks to an upgrade of its surface-detector array under the name of *AugerPrime*. In this chapter, we will briefly overview the detectors of Auger, as well as the reconstruction of air-shower events. Furthermore, we summarize selected results derived from the data collected by Auger to date and outline the importance of the detector upgrade.

4.1 THE SURFACE DETECTOR ARRAY

The SD of Auger is the largest surface-detector array for CRs today. The main array is arranged in a triangular grid with a spacing of 1500 m between neighboring stations; a sub-grid called the *infill* is employed within the main grid with a spacing of 750 m [93]. The spacing of the main array is chosen in such a way that a reasonably large area can be covered with a limited number of detectors while air showers from UHECRs still trigger enough detector stations for an accurate reconstruction. It is designed to detect CRs of an energy of $\sim 10^{19}$ eV and higher with an expected minimum of five stations per event at these energies.

Each detector station consists of a water-Cherenkov detector (WCD), solar panel, necessary electronics, and a battery. The WCDs employ a cylindrical body of 12 000 ℓ of ultra-high purity water, reflective coating, and three photo multiplier tubes (PMTs), in such a way that charged relativistic particles from air showers illuminate the detector, almost independently of their initial direction of arrival. The hereby produced photo-electrons are digitized and converted into signal S with a sampling time of 25 ns. Atmospheric muons, produced from low-energy CRs, passing through the WCDs are used to calibrate the signal of the detector in real time [95]. A peak can be identified in the histograms of charge produced in the PMTs, which after geometric corrections can be related to the most probable energy deposited by

(a)



(b)

Figure 4.1: Impressions of the Pierre Auger Observatory. A snapshot taken from the plateau *Pampa Amarilla*, where the Pierre Auger Observatory is located, is given in panel (a) [90]. A station of the SD can be seen with the FD building *Los Leones* in the background. (b) shows the Auger event display, with an example hybrid event (Id: 180415052300), detected by the FD and SD in February 2018. The SD is shown as gray dots, FD buildings are indicated by round shapes matching the color of the respective building. Selected stations of the SD are depicted in a time-ordered yellow-red color spectrum, while the path of fluorescence light detected by FD is depicted in a time-ordered rainbow color spectrum. The reconstructed shower axes using data from the FD (three independent reconstructions) and SD are depicted in red and blue, respectively.

one vertically through-going muon. The WCDs are thus calibrated to the signal unit called *vertical equivalent muon* (VEM) in real time[1]. The signal deposited over time in a detector is referred to as *trace* and carries important information about the event. The start time, when a station is beginning to respond to particles from an event, is directly related to the curvature of the shower front that consists of the earliest particles to reach the ground. The overall length and shape of the time-dependent signal is also important for the later reconstruction. A typical trace from a WCD is depicted in Fig. 4.2, in orange. The blue trace is the signal deposited in a scintillator surface detector in units of *minimum ionizing particle* (MIP). These detectors are addressed in the following section.

[1] 1 VEM corresponds to 240 MeV for an Auger-like WCD with 1.2 m height, given an ionization loss of 2 MeV/g cm^{-2}, which is 2 MeV/cm for a body of water.

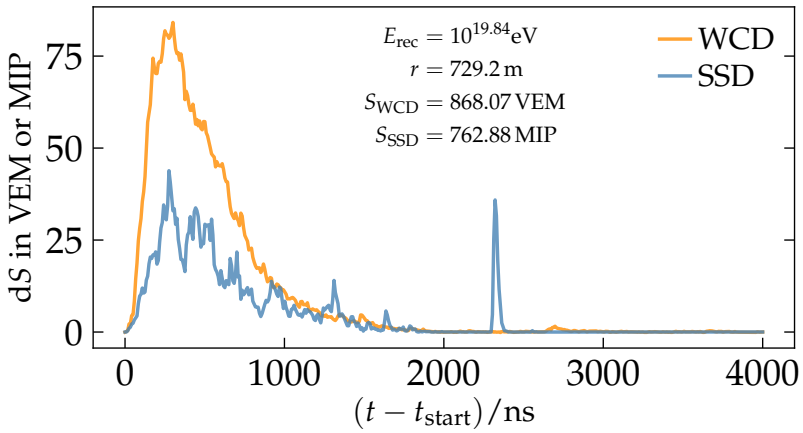


Figure 4.2: Exemplary detector responses to particles over time from a simulated proton shower from detectors of the same station.

An important feature of the WCDs is that because of their longer track length muons deposit more signal on average than other particles do. Extensive air showers with a large relative muon count R_μ will thus produce significantly more signal in WCDs than showers with smaller R_μ, especially beyond a distance of $\sim$1 km to the shower core, where the muon component becomes increasingly relevant (see Fig. 3.3 (a)). The change in the lateral distribution of signal with varying R_μ, however, is subtle and in the reconstruction procedure more signal can generally be understood as a higher particle density and consequently as a larger primary energy of the CR.

4.1.1 AUGERPRIME

To overcome the ambiguity of larger signals in distant WCDs, the stations of the SD are undergoing a detector upgrade. For the updated SD of *AugerPrime*, additional detectors with different detection mechanisms are mounted on the existing detector stations. Scintillator surface detectors (SSDs), each covering an area of $\sim 4\,\mathrm{m}^2$, are placed on top of the WCDs[2] [36]. In each SSD, several plastic scintillator bars of 1 cm thickness are placed. Optical wavelength-shifting fibers are laid out inside the scintillator bars to convert and guide the scintillation light to a photo multiplier. The components of the detector are placed in an opaque aluminum box. A model of an AugerPrime station is depicted in Fig. 4.3. In a manner similar to the WCDs, the SSDs are calibrated in terms of MIP, i.e. by the most probable energy deposited by one vertically through-going minimum-ionizing particle.

The added value of the SSDs for the detector array stems from its difference to the WCDs. In contrast to the WCDs, the response of the SSDs is approximately identical for all charged particles from extensive air showers. Given the large abundance of electromagnetic particles in air showers, the total signal deposited in the SSDs is thus mainly driven by this type of particles. The coincident measurement of particles in different detectors of a single station can thus be used to disentangle the signal of the individual particle components. If the expected lateral distribution of the individual particles is thoroughly understood, the signal from the different components can be disentangled and therefore the relative muon content can be accurately reconstructed using the interplay of the WCD and SSD signals of AugerPrime.

[2] In addition to the SSDs, many of the detector stations are augmented by underground muon detectors and radio antennas on top of the WCDs, directly above the SSDs. Both of these detector types are not directly relevant to the context of this work and are thus not discussed in detail.

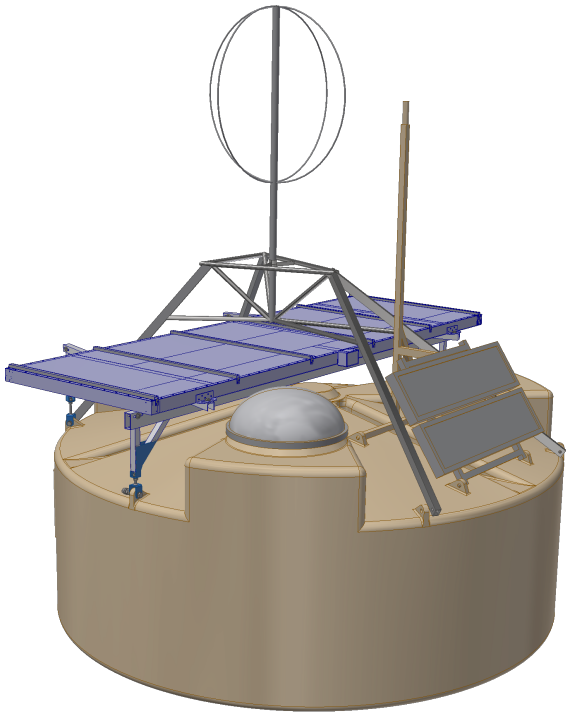


Figure 4.3: Model of an AugerPrime surface detector station with a WCD (ocher), SSD (blue) and radio detector.

4.1.2 SATURATION OF PMTS

A weakness of the SD is that for events from UHECRs the WCDs close to the shower axis often reach their maximum capacity to detect particles. In this case, too many particles emit Cherenkov light inside the tank at the same time and the electronics do not respond linearly to the number produced photoelectrons anymore. The signal then saturates at a constant level at ∼600 VEM for each 25 ns time bin. Even though the total signal that would have been deposited can be estimated to some degree, the time dependent information for these stations is partially lost [96]. This is particularly unfavourable since for ultra-high-energy events the quality of the station traces is badly affected. Furthermore, the precision and accuracy of the reconstruction of the primary energy is aggravated.

In AugerPrime the built-in electronics of the FD and SD are being upgraded. For the SD stations an additional small PMT, with a much higher saturaion level, is added to the stations. In this way, the problem addressed in Section 4.1.2 will be resolved in the future. Furthermore, the signal sampling time will be improved by a factor of 3, from 25 ns to $8.\bar{3}$ ns.

4.2 SD RECONSTRUCTION OF EXTENSIVE AIR SHOWERS

In this section, the reconstruction of events using data collected by the Auger SD [97, 98] is briefly discussed. To reconstruct the primary energy and the arrival direction of CRs several steps are performed.

In each station the timing and the size of the deposited signal is measured. Together with the absolute position of the station in space a four dimensional event is recorded. First a plane, and then a spherical shower front, propagating with the speed of light is fitted to the absolute station positions and individual start times of the signal. In this way, the geometry of the event is reconstructed, yielding the shower axis and the impact point of

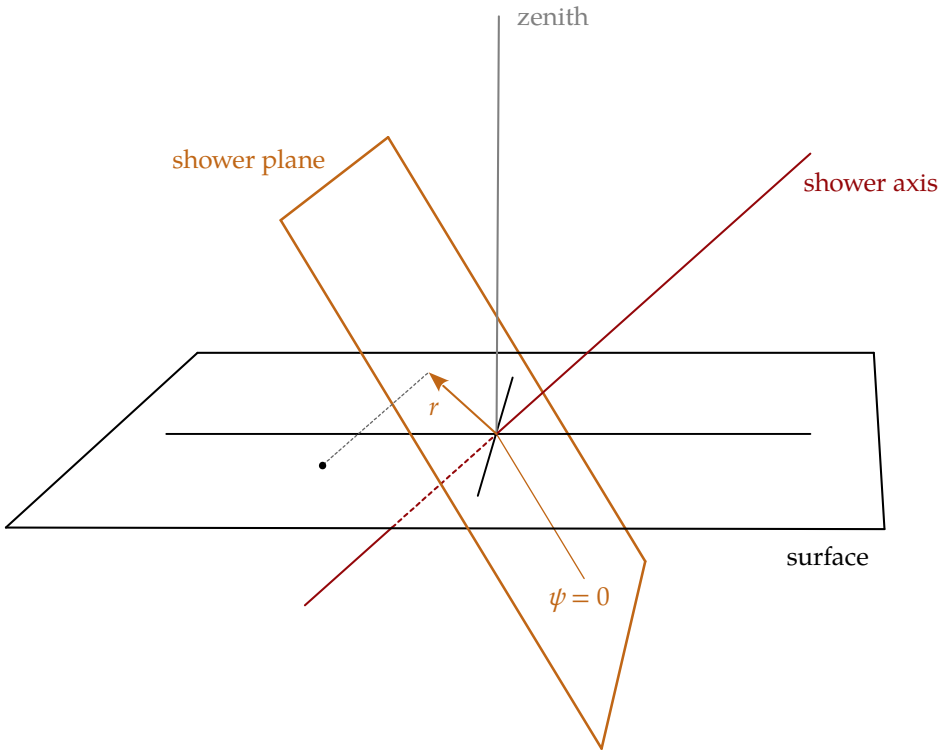
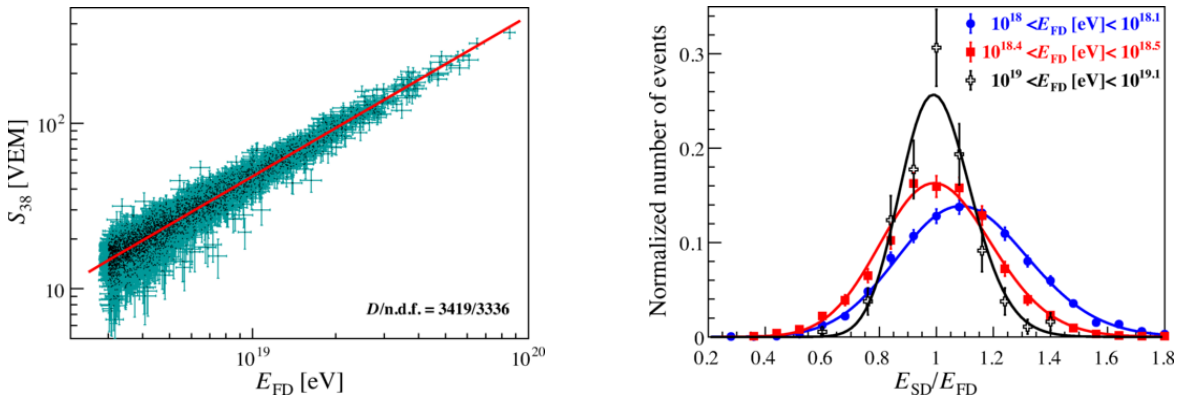


Figure 4.4: Geometry of a reconstructed air-shower event. The ground plane that is perpendicular to the zenith is depicted in black, the shower plane is depicted in orange. The shower axis pierces both planes in a coincident point that is the impact point of the shower core at the ground. An example point on the surface at a shower-plane distance r with $\psi \simeq 200^\circ \equiv -160^\circ$ is depicted as black dot. A set of points for which $\psi = 0$ is depicted in orange in the shower plane.

the shower core on the ground. The distance r of each station to the shower axis, referred to as *shower-plane distance* is thereby fixed. The *shower plane* is defined as the plane that contains the impact point of the shower core at the ground whilst being perpendicular to the shower axis. In the event-based, tilted cylindrical coordinate system, the impact point of the shower core is by definition located at the height $h = 0$ and $r = 0$. The lateral distribution of particles is examined in the shower plane, i.e. in the "perspective" of the CR. The angular distance between the zenith and the shower axis is given by the zenith angle θ. The azimuth ψ is measured in the shower plane between any point and the vertically-downwards projected shower axis. The geometry is illustrated in Fig. 4.4. For reconstructed events, points on the ground are described in terms of their polar shower-plane coordinates, $\vec{r} = r\,\hat{e}_r + \psi\,\hat{e}_\psi$, with unit vectors $\hat{e}_r$ and $\hat{e}_\psi$. Points on the ground for which $\cos\psi > 0$, i.e. on the right-hand side of Fig. 4.4, are referred to as *upstream*, whereas points for which $\cos\psi < 0$ are referred to as *downstream*, in reference to the average direction of particles in a shower.

The energy of an event is reconstructed from the signal size of the footprint of the shower. A lateral distribution function (LDF) $S(r)$ is fitted to the signals from individual detectors as a function of the shower plane radius. An LDF is usually expressed in terms of signal rather than particle density, thus historically various LDFs have been examined to fit shower footprints for different detector types of different experiments [99]. One of the two LDFs used for Auger is motivated by the NKG formula from Eq. (3.46). Using the SD of Auger, the shower size is estimated by the signal predicted by the best-fitting LDF at



(a) (b)

Figure 4.5: Reconstruction of the primary energy with Auger, (a) showing the calibration of S_{38} to the primary energy as reconstructed by hybrid FD, and (b) showing the resolution of the energy as reconstructed by SD with respect to FD. Taken from Ref. [100].

1000 m, S_{1000}. This particular distance is chosen due to geometric considerations regarding the spacing of the array in such a way that the signal fluctuations are expected to be minimal [97]. Depending on the arrival direction of CRs, however, S_{1000} is not expected to be universal due to attenuation effects. Highly inclined showers pass through more matter in the atmosphere until reaching the SD than vertical showers do. In general, fewer particles and less signal is therefore expected from inclined showers. To account for the attenuation of particles by the atmosphere S_{1000} is converted to S_{38}, which is the shower-size equivalent at 1000 m for an air shower induced by a CR with zenith angle $\theta = 38°$ [100]. The constant-intensity cut method is applied for this conversion, which introduces an attenuation factor $f(\theta)$, such that

$$f(\theta)\, S_{38} = S_{1000}. \tag{4.1}$$

$f(\theta)$ is derived under the assumption that the arrival directions of CRs above the atmosphere are approximately isotropic. Histograms of S_{1000} in ranges of equal exposure[3] are compared against each other, assuming that the n-th largest bins correspond to the same amount of total CRs, for various n, thus yielding the empirical attenuation curve $f(\theta)$. Finally, the energy of the CR is assigned by the power law

$$E_0^{\mathrm{rec}} = A\, S_{38}^B, \tag{4.2}$$

where A and B are constants that are calibrated using the energies reconstructed from hybrid FD measurements. The calibration of E_0^{rec} to S_{38} as well as the resulting resolution is depicted in Fig. 4.5. The amount of calorimetric energy deposited in the atmosphere is obtained by the FD, which is directly measuring the fluorescence and Cherenkov light produced in air showers [92, 101]. In the process of calibration, the effect of invisible energy is taken into account (see Section 3.3). The reconstructed energy as well as the event geometry is required for the reconstruction of the primary mass, as it is discussed in Chapter 6.

4.3 HYBRID EVENTS

Hybrid events, i.e. when both the FD and SD take data from the same air shower, are of major importance, especially for the calibration of the detector. For example, using hybrid

[3] Assuming there is no atmospheric attenuation at all, for each energy the same amount of CRs would be detected in every equally-sized range of $\sin^2\theta$.

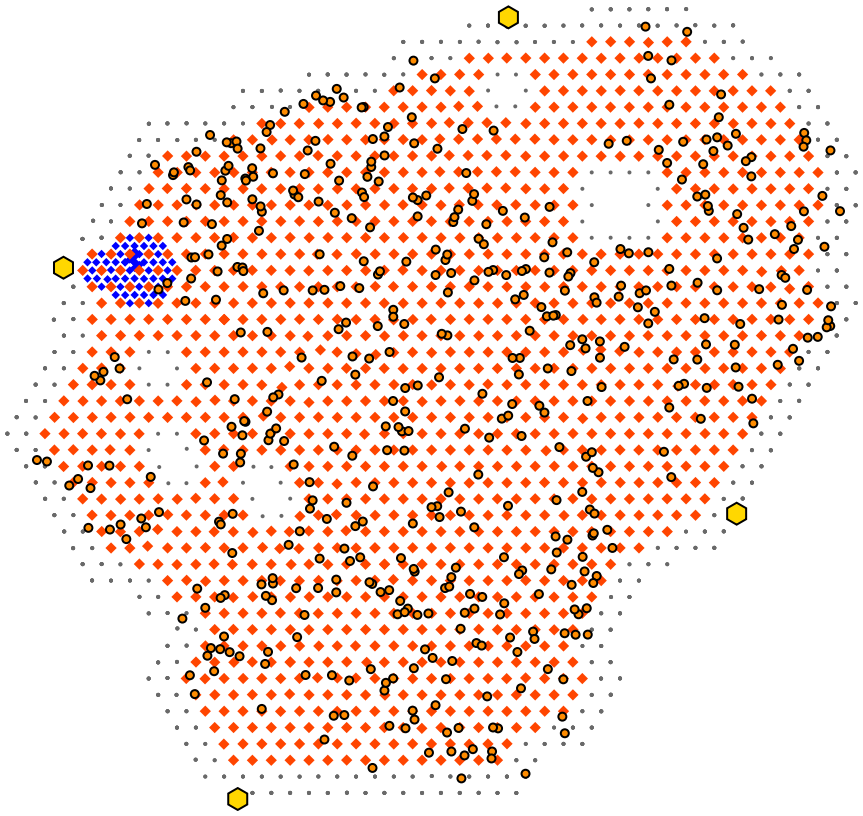
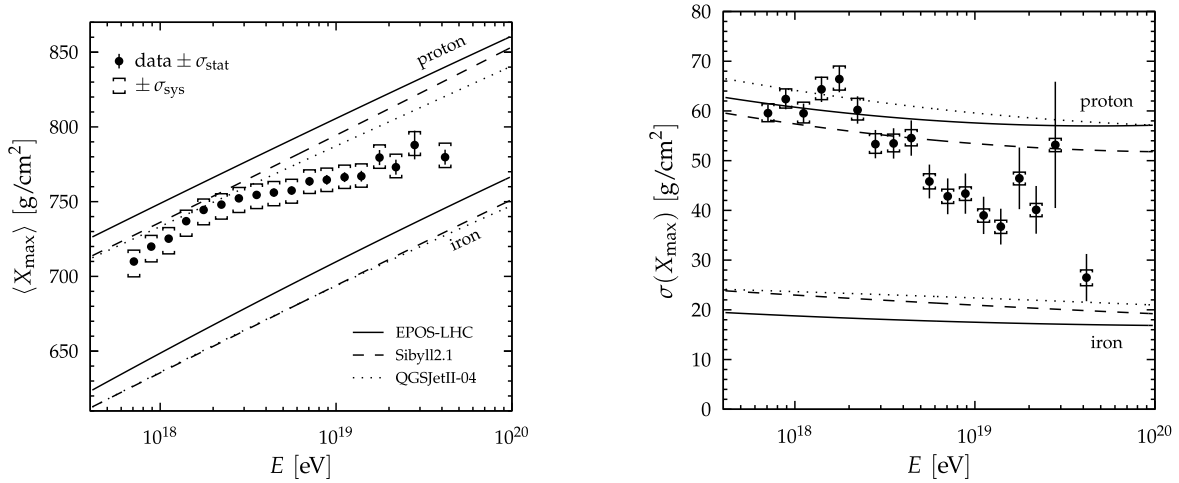


Figure 4.6: Impact point positions of all Golden Hybrid events with $E_{\mathrm{rec}} \geq 10\,\mathrm{EeV}$ in the Auger surface detector array as of 2019. In total, the core positions of 478 events are shown as orange dots with black edge, the SD array as gray markers, the stations to be upgraded as red diamonds, the four FD buildings as large yellow markers, and the infill array in blue.

events, studies on the "invisible" energy, carried away by neutral, non-interacting particles [70], as well as cross-checks with MC simulations are possible .

An event that independently triggers both the SD and FD, and can thus be reconstructed by data from both detectors individually, is labelled *Golden* Hybrid event [102]. A map of the positions of the impact points of shower cores at the ground for Golden Hybrid events with a reconstructed energy above 10 EeV is depicted in Fig. 4.6. As it can be seen in this figure, Golden Hybrid events are rather rare, since the FD can only operate in clear moonless nights, which is only about 15% of the time [91]. At the highest energies, $E \gtrsim 16\,\mathrm{EeV}$, only $\sim$200 Golden Hybrid events have been collected so far. Additional maps of the impact point positions of Golden Hybrid events for various energy thresholds are given in Figs. B.1 to B.3.



(a) (b)

Figure 4.7: First and second moments of the X_{max} distribution as a function of the primary energy as measured by the Auger FD telescopes. Taken from Ref. [104].

4.4 SELECTED RESULTS

Auger has contributed significantly to solve the open questions of CR and astroparticle physics. In this section we highlight those results from the Pierre Auger Collaboration that are of great importance for this work.

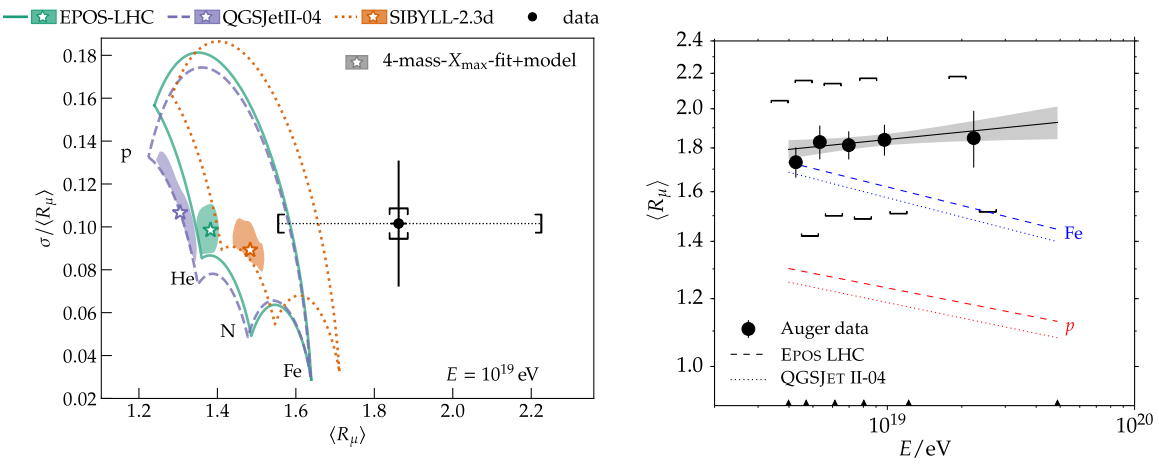
4.4.1 SPECTRUM AND COMPOSITION

In the Auger data, a suppression of the CR flux is visible at approximately $E_0 \simeq 4{\times}10^{19}$eV [103], which is a strong indicator for a GZK-like cutoff of the CR spectrum, as discussed in Section 2.2.2. With increasing energy, the composition of CRs evolves from being proton-dominated to a heavier but mixed composition, as the results from both the FD data and a combined fit of the FD data and the all-particle spectrum imply [16, 104]. The results from Ref. [104] are depicted in Fig. 4.7. Furthermore, the elongation rate found in the data is not constant, but can be described by a line with a breaking point at $\lg(E/\text{eV}) \simeq 18.3$, indicating a change of composition around that energy. At the highest energies, a simple mix of proton and iron nuclei as primary CRs is disfavoured by data when compared to implications from hadronic interaction models [105]. These studies, however, are done statistically, using the first and second moments of the X_{max} distribution. The abundance of protons at the highest energies of the CR spectrum is still unclear. The situation underlines the need for an event-by-event estimator of the primary masses of the CRs.

4.4.2 MUON DEFICIT

When comparing the footprints of extensive air showers with profiles obtained from MC simulations discrepancies arise. The number of muons detected at the ground significantly exceeds the amount of muons expected from MC simulations that employ modern hadronic interaction models [106, 107]. Studies on the amount of muons reaching the ground have been done predominantly for highly inclined ($\theta \gtrsim 60°$) showers, for which the electromagnetic particles are strongly attenuated.

In fact, when only considering R_μ as reconstructed from highly inclined showers, the composition of CRs would be dominated by very rare or unstable nuclei that are signifi-



(a) (b)

Figure 4.8: The muon deficit as measured by Auger. On the left, panel (a) is showing the first and second moment of the distribution of R_μ for CR showers with $E_0 = 10^{19}$ eV; taken from Ref. [107]. On the right, panel (b) is showing the development of $\langle R_\mu \rangle$ with respect to the primary energy; taken from Ref. [106].

cantly heavier than iron. Since this situation is unphysical, the discrepancy is identified as a shortcoming of the hadronic interaction models, which fail to reproduce the amount of muons in air showers as implied by data. This muon deficit is a severe problem, especially for the event-by-event estimation of the mass of the primary CR, since R_μ is a powerful primary mass-dependent observable. Measurements of the relative muon number as measured by Auger are depicted in Fig. 4.8. R_μ as depicted in Fig. 4.8 is based on simulations of proton showers with the QGSJET-II-03 hadronic interactions model. They are are thus derived using a smaller reference value for the expected amount of muons than it is used in this work. The reference predominantly used in this work is obtained from proton shower simulated with the EPOS-LHC model, i.e. given by the dashed red line in Fig. 4.8 (b). A method to circumvent the problems arising from the muon deficit is outlined in Section 6.4.

CHAPTER V

AIR-SHOWER UNIVERSALITY

The symmetry of extensive air showers, which is described as *air-shower universality* or simply *Universality*, implies that around the shower maximum the average development of an air shower is universal. This holds in terms of the particle spectra, the longitudinal profile, and the lateral profile of the shower. The absolute number of particles is determined by the energy of the primary CR, whereas the relative amounts of particle types can be expressed as a function of the relative amount of muons. Thus, an air shower can be uniquely described by the primary energy E_0 and the observables R_μ and $X_{\max}$. For purely electromagnetic cascades, Universality occurs for showers above a primary energy of approximately 10^{14} eV, and for hadronic showers from UHECRs above an energy of approximately 10^{19} eV [78, 88, 89].

In this chapter the principle and the implications of Universality is discussed. Subsidiary particle components are introduced to augment the particle components already mentioned in Section 3.3. The number of particles in each component is described as a function of the relative number of muons in the shower, R_μ. In this way, a unique description of air showers is possible independently of the absolute abundance of the different particle components. We propose a model of the expected signal from air showers based on Universality. Furthermore, we derive a new model of the expected arrival times of particles and the resulting time-dependent expected signal. The model is parametrized and validated using simulated data and put into context with respect to previous work. In contrast to earlier work, all functional forms for the longitudinal and lateral profiles of particles are revisited and made more physically motivated. The model is parametrized using full simulations of detector responses of the AugerPrime surface detector, produced in the Offline software framework.

5.1 THE CONCEPT OF UNIVERSALITY

The concept of universality, or sometimes *similarity* [63, 89], summarizes many regularities concerning the particle spectra, the angular distributions of particles, and lateral and longitudinal profiles of air-showers which were discovered both analytically and in simulations [54, 64, 78, 88]. These regularities, however, are not to be understood as strict physical laws. In the context of air-shower Universality, statements are made about *most* of the particles and for *most* of the showers. Since an air shower is a single realization of the phenomenon that is described by the cascade equations, it will never behave exactly as the solutions for the average shower imply, but given a sufficient number of particles it will do so approximately. Rare air-shower events such as double-bump showers with an anomalous longitudinal profile cannot be described by Universality [108].

Most air showers from UHECRs, however, develop universally. As shown in Ref. [88], for example, the energy spectra of electromagnetic particles as a function of the shower

age are universal in very good approximation (see Section 3.5.2). Furthermore, the lateral distribution of particles in air showers is always well described by an NKG function (see Section 3.2) [68]. Lastly, and most prominently, the relative rate of change of the longitudinal shower profile,

$$\lambda_1(t) = \frac{1}{N(t)} \frac{\mathrm{d}N(t)}{\mathrm{d}t}, \tag{5.1}$$

with $t = X/X_0$, is approximately the same for all showers, if expressed relative to the shower maximum. It is thus invariant under a change of both the energy and the depth of the shower maximum. Especially at very high energies, where fluctuations are less severe due to large particle multiplicity, descriptions of air-showers based on Universality, such as for example given in Ref. [109], are very successful.

Without hesitation, however, these statements are only true for air showers induced by gamma rays, electrons, or positrons. For hadronic showers, which produce substantial amounts of muons and hadrons in the cascade, the lateral distributions of particles are not universal anymore. Depending on the elasticity of the first interaction and on the ratio of neutral to charged pions produced therein, more or less muons are produced in a shower [110]. The fluctuation in the number of muons is enough to affect the longitudinal and lateral profiles of showers significantly [111], even for a single type of primary particles. Moreover, the longitudinal and lateral distribution of particles is systematically different for air showers from heavier primary particles, when compared to showers from lighter primary particles. Thus, it appears as if Universality breaks down. However, identifying and disentangling all the different particle components of a shower, which systematically behave differently from each other, Universality can be effectively restored. This is discussed in the following section.

5.2 THE FOUR-COMPONENT SHOWER MODEL

Historically, the Universality was discovered for purely electromagnetic showers [63, 78, 89]. As discussed in Chapter 3 and Section 5.1, a universal description of hadronic showers, which contain different particle components, is more difficult. Qualitatively, the Heitler-Matthews model, for example, is fit to describe the number of particles expected in hadronic showers, but it does not make detailed statements about the lateral and longitudinal development of the different particle components. Furthermore, the elegance and simplicity, with which the creation and decay of particles is handled in the Heitler-Matthews model, is also the greatest shortcoming of this model, as discussed in Section 3.4. Neither the decay of hadrons nor of muons into electromagnetic particles is included in the model. Decay products from muons and hadrons, however, contribute significantly to the signal in terms of shape and size as detected by surface detector arrays and thus have to be considered in a model based on universality [112, 113]. Studying the difference of the projected impact point of particles onto the ground at a radius r_{proj} and the actual impact point of secondary particles at a radius r, which were produced in in-flight decay, a clear peak was identified in the deposited signal at $r_{\mathrm{proj}} \simeq r$ [114]. In this way, sub-showers from low-energy particles were identified initiating small cascades that contribute to the lateral displacement of shower particles near the ground level. These sub-showers are driven by the momentum and angle of the initiating mother particle. Even though the resulting particles are mostly part of the electromagnetic component, they are accounted for as the component of their mother-particle, from which they decayed. The hereby introduced *jet* component contains all hadrons and visible particles produced in their decay near the ground, which are mainly

pions, electrons, positrons, and photons. Furthermore, electrons produced by in-flight decay of muons follow a lateral distribution which is more similar to the lateral distribution of the muonic component than to the *pure* electromagnetic component. It thus needs to be disentangled as well.

The four particle components that make up the particle content of air showers are henceforth:

- the electromagnetic component, $\mathrm{e}\gamma$, containing all electrons, positrons, and photons,
- the muonic component, μ, containing all muons and anti-muons,
- the hadronic or jet component, $\mathrm{e}\gamma(\pi)$, containing all hadrons and their decay products up to the second generation, and
- the electromagnetic component from muon decays, $\mathrm{e}\gamma(\mu)$, which contains all electromagnetic particles produced by decays of the μ component up to the second generation.

To identify the components of particles in simulated data, the CORSIKA framework allows to save the history of particles with respect to their hadronic interactions [112, 115]. The four particle components are treated individually, except for their dependence on R_μ. To account for different primary particles and for shower-to-shower fluctuations of showers with the same energy, the amount of particles produced in each component is given as a function of R_μ. Initially the relative amounts of particles in each component i is written as

$$R_i = \frac{N_i}{\langle N_i^\mathrm{p} \rangle}, \tag{5.2}$$

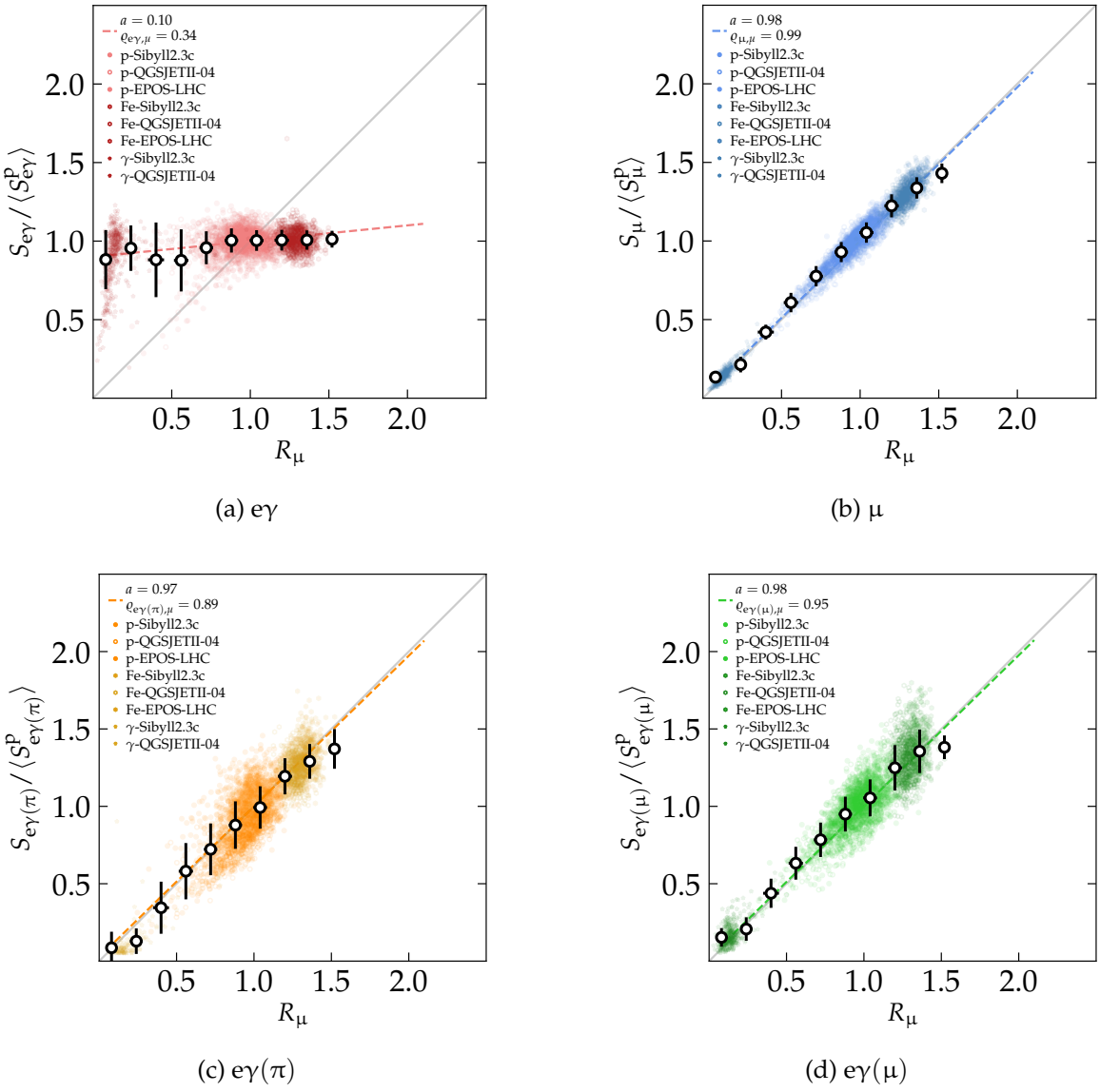
for $i \in \{\mathrm{e}\gamma, \mu, \mathrm{e}\gamma(\pi), \mathrm{e}\gamma(\mu)\}$, using the average amount of particles in a respective proton shower $\langle N_i^\mathrm{p} \rangle$ as a reference. The reference value for each component i is expected to scale with the primary energy according to a power-law

$$\langle N_i^\mathrm{p} \rangle(E) = \langle N_i^\mathrm{p} \rangle|_{E_0 = 10^{19}\,\mathrm{eV}} \left(\frac{E_0}{10^{19}\,\mathrm{eV}} \right)^{\gamma_i}, \tag{5.3}$$

with a corresponding power $\gamma_i \lesssim 1$. As in the Heitler-Matthews model, where the amount of muons is (trivially) linearly dependent on the amount of pions in an extensive air shower, we expect the number of particles of all components to be linearly dependent on R_μ. For any component i we thus write

$$R_i - 1 = a_i (R_\mu - 1), \tag{5.4}$$

such that for $a_i = 1$ the component scales exactly like the muonic component, and for $a = 0$ a component is independent of the amount of muons produced. Eq. (5.4) is written such that the dependence on a_i always vanishes for $R_\mu = 1$. In this case the number of particles in each component is the same as for the average proton shower. As a measure of the relative amount of particles in each component, the relative amount of the signal deposited in Auger WCDs as a function of R_μ is for the four particle components depicted in Fig. 5.1. The signal is normalized to the reference value which is obtained from the average of the signals of proton showers and thus, by definition, from the average signal of a shower with $R_\mu = 1$ (see Section 5.7). The reference signal used for Fig. 5.1 is derived using CORSIKA simulations employing the EPOS-LHC model of hadronic interactions. The behaviour of the MC data depicted in Fig. 5.1 confirms the expectation from Eq. (5.4). All



(a) eγ (b) μ

(c) eγ(π) (d) eγ(μ)

Figure 5.1: Signal deposited in WCDs from the four particle components normalized to the expected average signal from a shower with the same energy and with the same $X_{\max}$, as a function of R_μ. The correlation coefficient ϱ of the normalized signal and R_μ as well as the slope a of the dashed line is given in the respective upper left corner. The data was fitted according to Eq. (5.4). All air showers in the depicted data set are induced by CRs with $E_0 = 10^{19}$ eV.

particle components are linearly dependent on R_μ. A line with slope a_i is fitted to each of the four data sets.

Three important implications can be drawn from the depicted behaviour. Firstly, the number of particles produced in the eγ component of an air shower is approximately independent[1] of R_μ, see Fig. 5.1 (a). In fact, a slight anti-correlation can be examined if only particles at large shower-plane radii are considered, so that $a_{e\gamma} < 0$. Secondly, the components eγ(π) and eγ(μ) in Fig. 5.1 (c) and (d), respectively, scale almost exactly like the μ component, thus $a_{e\gamma(\mu)} \simeq a_{e\gamma(\pi)} \simeq 1$. Lastly, the scaling of the different particle components with R_μ naturally describes the shower-to-shower fluctuations, the amount of particles created

[1] An anti-correlation is expected due to energy conservation. In general, however, it is safe to assume $a_{e\gamma} \simeq 0$.

by different primary particles on average, and the systematic discrepancies of the hadronic interaction models – all of these effects manifest in different values for R_μ. In all four panels of Fig. 5.1 similar groups of data points can be identified from air showers induced by photons on the left, with $R_\mu \simeq 0.1$, iron-induced air showers around $R_\mu \simeq 1.4$ and proton showers at $R_\mu \simeq 1$. Therefore, for any type of primary particle the overall particle content of an air shower can be elegantly described by the sum of its components as a function of R_μ.

If not explicitly stated differently, the statements of the following sections apply to all four particle components individually and independently. For better legibility parameters that are individually fixed for different components will not be indexed with i anymore.

5.3 THE LONGITUDINAL PROFILE

The longitudinal development of air showers describes the number of particles present at a slant depth X. It is almost exclusively driven by the eγ component, which dominates the shower in terms of particle count and energy deposit. The shower maximum, $X_{\max}$, is thus very close to the maximum of the eγ component, at which all electromagnetic particles carry on average an energy of $\epsilon_c^{e\gamma} \simeq 87\,\mathrm{MeV}$. In this section, the profile function used to describe the particle density at a given depth $X \gtrsim X_{\max}$ and distance from the shower axis r is introduced using the shower maximum $X_{\max}$ as a point of reference. Using the shower maximum as a point of reference is crucial for a simple reason. Independently of the primary particle, the absolute value of $X_{\max}$, and even independently of the energy of the CR (assuming $E_0 \gtrsim 10^{19}\,\mathrm{eV}$), the development of all showers is approximately identical around and beyond $X_{\max}$ [116], as discussed in Section 5.1. Thus, the relative shower-depth parameter $\Delta X := X - X_{\max}$ is used to describe the slant depth. In Fig. 5.2 the longitudinal profiles of several air showers generated by CORSIKA using the EPOS-LHC model of hadronic interactions are depicted in terms of absolute number of particles (first row) and calorimetric energy deposit (second row), as well as the ratio of the two (third row), indicating the average energy loss of each particle per step length in units of $\mathrm{g\,cm^{-2}}$. When changing coordinates from X to ΔX (first column to second column in Fig. 5.2), the profiles nicely align on top of each other, not just at the shower maximum but also in vicinity for $|\Delta X| \lesssim 200\,\mathrm{g\,cm^{-2}}$. Even the longitudinal profiles of the depicted proton showers (red), which are strongly fluctuating in terms of $X_{\max}$, show universal behaviour and align with the longitudinal profiles of iron showers (blue) when shown as a function of ΔX. Part of the reason for the strong fluctuations of the longitudinal profiles from proton showers is the large variance of the point of first interaction at the depth X_1, which is removed when $X_{\max}$ is used as a common point of reference. The constant average energy deposit per particle and per depth is an important indicator that in general the assumption of the proportionality of the deposited signal to the number of particles is a valid.

A profile function which accurately fits the longitudinal development of extensive air showers was found in Ref. [117] to test the constant-intensity-cut method, is the Gaisser-Hillas function

$$N(X) = N_{\max} \left(\frac{X - X_1}{X_{\max} - X_1} \right)^{\frac{X_{\max} - X_1}{\lambda}} \mathrm{e}^{-\frac{X - X_{\max}}{\lambda}}. \tag{5.5}$$

Alternatively, if expressed in terms of ΔX, the Gaisser-Hillas function reads as

$$N(\Delta X) = N_{\max} \left(\frac{\Delta X}{X_{\max} - X_1} + 1 \right)^{\frac{X_{\max} - X_1}{\lambda}} \mathrm{e}^{-\frac{\Delta X}{\lambda}}. \tag{5.6}$$

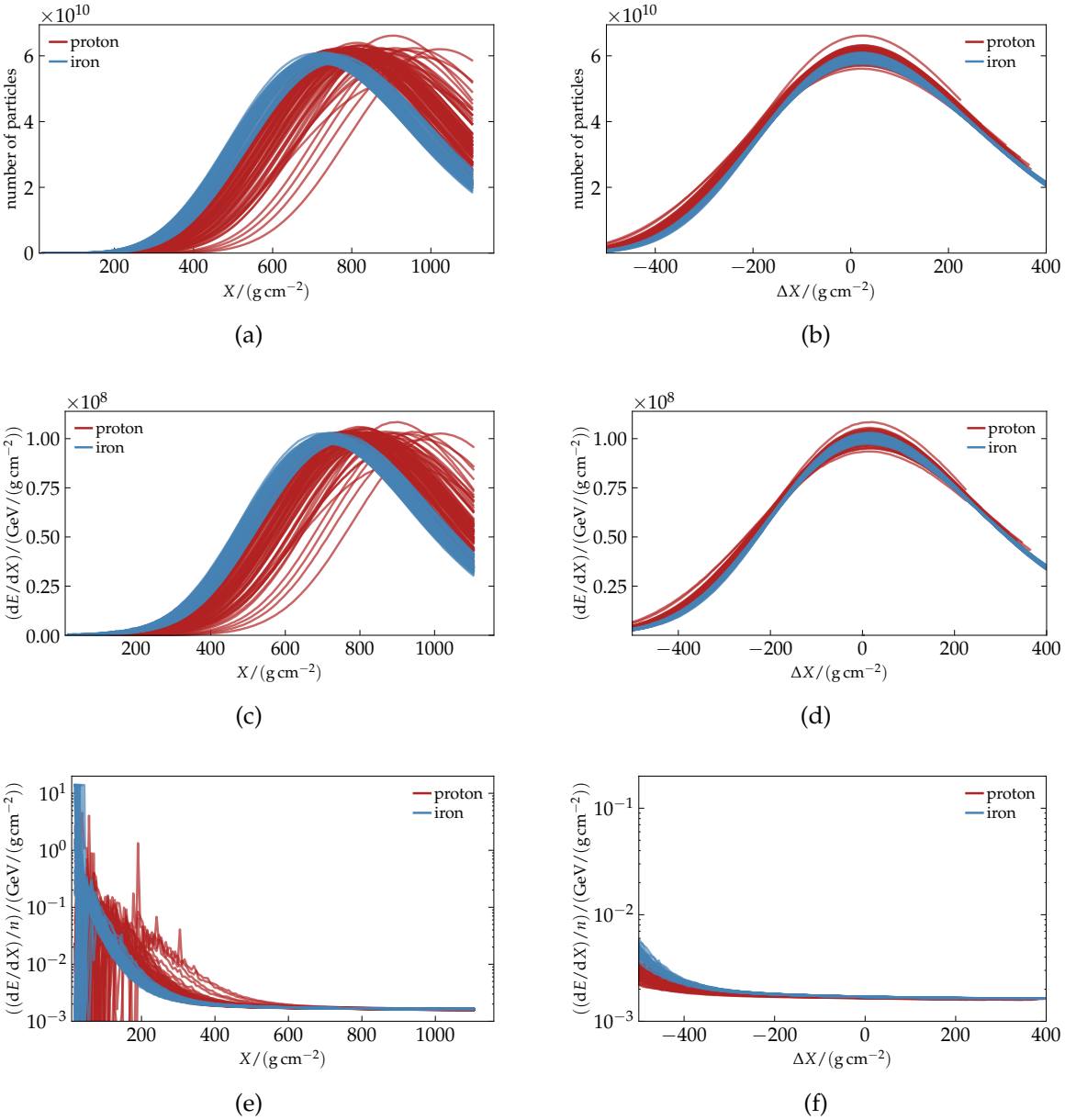


Figure 5.2: Longitudinal profiles of 100 proton- and iron-induced air-showers in red and blue, respectively, simulated with CORSIKA and EPOS-LHC. The primary energy is 10^{19}eV and the showers are inclined with $\theta = 38°$ so that the shower core hits the ground at a depth of $X \simeq 1150\,\mathrm{g\,cm^{-2}}$. Plots (a) and (b) show the number of particles present at a certain depth, whereas (c) and (d) shows the differential energy deposit at the respective depth. Panels (e) and (f) show the average energy deposit per particle at the respective depth, i.e. the ratio of the first and second row. In the first column ((a), (c), and (e)) the ordinate is given by the absolute slant depth X, whereas in the second column ((b), (d), and (f)) the ordinate is the depth expressed in terms of $\Delta X = X - X_{\max}$.

$N_{\max}$ is the maximum number of particles, present at the depth $X_{\max}$, X_1 marks the effective depth of the first interaction, and λ is a parameter, which is related to the interaction length of the considered particles. The Gaisser-Hillas function is closely related to the Greisen profile of electromagnetic particle cascades given in Eq. (3.43), as it was shown already in Ref. [63]. Expressed in terms of radiation lengths, $t = X/X_0$, $t_{\max} = X_{\max}/X_0$, and

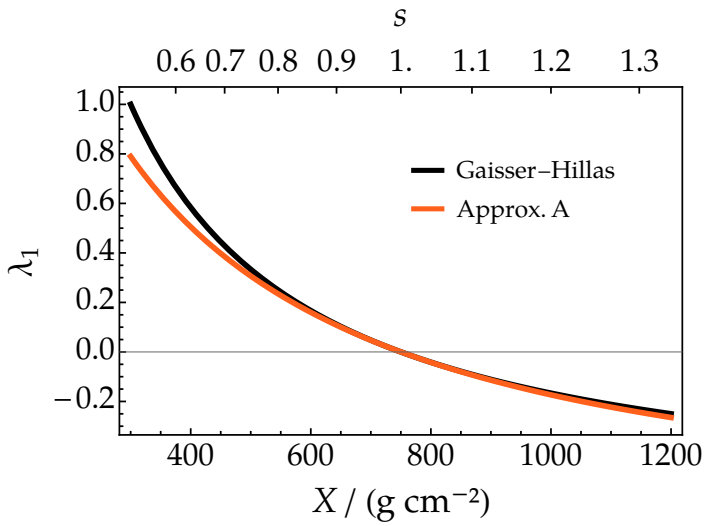


Figure 5.3: Relative rate of change of the longitudinal shower profile λ_1 obtained from the Gaisser-Hillas profile function, compared to λ_1 as implied by the solution of the cascade equations under Approximation A, given in Eq. (3.36). For the Gaisser-Hillas profile $X_{\max}$ is set to $750\,\mathrm{g\,cm^{-2}}$ and the normalized Gaisser-Hillas interaction length Λ is set to $3/2$. $\lambda_1 = 0$ marks the shower maximum, at which $s = 1$.

$\Lambda = \lambda/X_0$, the shower age s at the depth t, as implied by the Gaisser-Hillas function, can be written as

$$s = 1 + \frac{1}{\Lambda}\delta - \frac{4\Lambda - 3}{4\Lambda^2}\delta^2 + \mathcal{O}(\delta^3), \tag{5.7}$$

where $\delta = (t - t_{\max})/t_{\max} = \Delta X/X_{\max}$ and X_1 is set to 0 for the sake of simplicity. The shower age of the Greisen profile, as given in Eq. (3.40), when expanded in terms of δ takes the form

$$s = 1 + \frac{2}{3}\delta - \frac{2}{9}\delta^2 + \mathcal{O}(\delta^3). \tag{5.8}$$

Thus, for $\Lambda \simeq 3/2$ and consequently $\lambda \simeq 3X_0/2$, a reasonable agreement of the two profiles can be found. Values around $\lambda \simeq 58\,\mathrm{g\,cm^{-2}}$ are thus expected for the electromagnetic particle component. Comparing the relative rate of change of the Gaisser-Hillas profile according to Eq. (5.1),

$$\lambda_1^{\mathrm{GH}} = -\frac{1}{\Lambda}\frac{X - X_{\max}}{X - X_1}, \tag{5.9}$$

to λ_1 as given by its classical approximation introduced by Greisen [62], which is given in Eq. (3.36), an overall agreement can be found for $\Lambda = 3/2$. For both cases λ_1 is depicted in Fig. 5.3. For the comparison in Fig. 5.3, X_1 was set to 0. Surprisingly, the agreement of the two functions improves around $X \ll X_{\max}$, if X_1 is set to negative values.

Empirically, the Gaisser-Hillas function successfully describes the longitudinal development of extensive air showers as measured by the Auger FD [101, 104]. It is thus widely used for their reconstruction. In these analyses, however, the focus lies mainly on the reconstruction of $X_{\max}$ and in the process of obtaining a best fit for the shower maximum while λ and X_1 are often allowed to take effective or even unphysical values such as $X_1 < 0$. Direct measurements of the average shower profile of air showers from UHECRs imply a Gaisser-Hillas like profile with $\lambda = (61 \pm 13)\,\mathrm{g\,cm^{-2}}$ and $X_1 = (-121 \pm 172)\,\mathrm{g\,cm^{-2}}$ [118]. For these

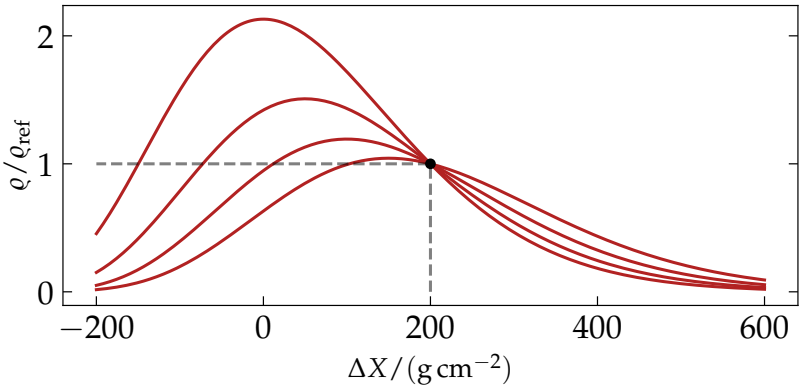


Figure 5.4: The modified Gaisser Hillas function illustrated for the values $\Delta X_{\mathrm{max}} \in \{0, 50, 100, 150\}\,\mathrm{g\,cm^{-2}}$, $\Delta X_1 = -400\,\mathrm{g\,cm^{-2}}$, $\lambda = 50\,\mathrm{g\,cm^{-2}}$ and $\Delta X_{\mathrm{ref}} = 200\,\mathrm{g\,cm^{-2}}$. The point of intersection at ΔX_{ref} is indicated with a black dot.

values, λ_1 from the Gaisser-Hillas function is qualitatively identical to Approximation A, up to a scaling factor that can be absorbed in the parameter Λ. A comparison is given in Fig. B.5. Furthermore, the expected number of particles from a shower reaching the surface is well described by the Gaisser-Hillas function, if the slant depth of the ground level is known relative to X_{max}, as for example studied in Refs. [76, 119].

The Gaisser-Hillas function, however, in its pure form describes only the total number of particles present at a certain depth, i.e. integrated over the plane that is perpendicular to the shower axis at this depth. To accurately describe the areal particle density at a given distance from the shower axis a modified version of Eq. (5.5) is introduced in the following.

5.3.1 THE MODIFIED GAISSER-HILLAS PROFILE

The areal density of particles in an extensive air shower at the depth ΔX and the distance r from the shower axis can be described by the modified Gaisser-Hillas profile[109, 114]

$$\varrho(\Delta X, r) = \varrho(r)_{\mathrm{ref}} \left(\frac{\Delta X - \Delta X_1}{\Delta X_{\mathrm{ref}} - \Delta X_1} \right)^{\frac{\Delta X_{\mathrm{max}} - \Delta X_1}{\lambda}} \mathrm{e}^{-\frac{\Delta X - \Delta X_{\mathrm{ref}}}{\lambda}}. \tag{5.10}$$

The hereby newly introduced quantities $\Delta X_1, \Delta X_{\mathrm{max}}, \Delta X_{\mathrm{ref}}$ and $\varrho(r)_{\mathrm{ref}}$ require individual explanation. The factor $\varrho(r)_{\mathrm{ref}}$ is the reference lateral particle distribution that is addressed in Section 5.4. ΔX_{ref} is a reference depth, which is placed so that $\varrho(\Delta X_{\mathrm{ref}}, r) \equiv \varrho(r)_{\mathrm{ref}}$. ΔX_1 describes the starting point of the profile in terms of ΔX and thus takes negative values. ΔX_{max} is the point of the maximum of the modified Gaisser-Hillas function at the shower-plane distance r, which is in general not coincident with X_{max}. The introduction of ΔX_{max} allows for the description of a retarded shower maximum at finite distances from the shower axis. As derived in Ref. [64] and discussed in Ref. [62], the shower maximum is expected to be increasingly retarded with increasing distance from the shower axis because the vast majority of particles in the cascade are radiated under a small angle with respect to the shower axis. This can be realized by a distance-dependent shower age, where Eq. (3.40) is augmented by a term in the denominator that depends on radius (see Eq. (3.48)), or by the assignment of a local age parameter such as given in Ref. [120]. In the model presented here, this retardation is best described by a radial dependence of ΔX_{max} and λ of Eq. (5.10).

The introduction of the reference depth ΔX_{ref} results in a joint point that is independent of the choice of the other parameters, as it is demonstrated in Fig. 5.4. The overall radial

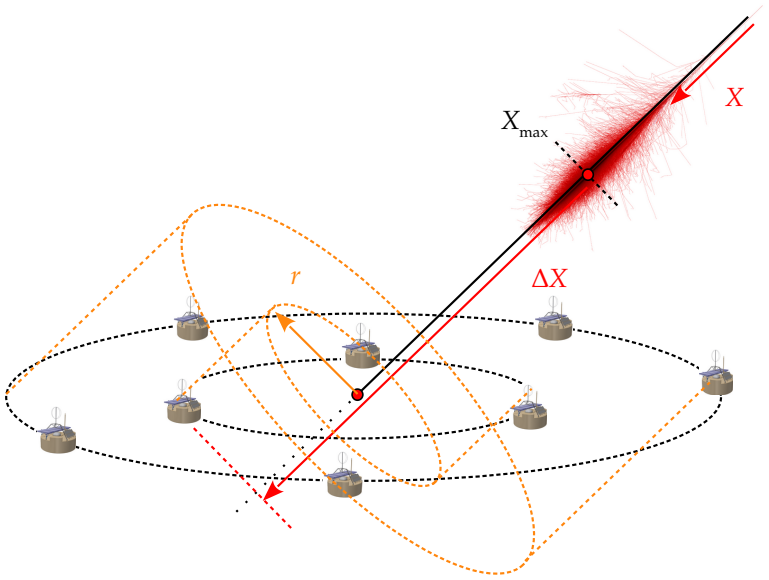


Figure 5.5: Sketch of the geometry that leads to the derivation of the shower depth parameter ΔX. Example points on the surface plane on concentric rings around the shower core at ground level are shown as AugerPrime stations. For the station on the left-hand side on the inner ring, ΔX is measured along the shower axis from X_{max} downwards to the point of perpendicular projection. The shower visualization is taken from Ref. [121].

dependence of the resulting profile, however, is still mainly driven by $\varrho(r)_{\text{ref}}$, such that the dependence on ΔX and r of Eq. (5.10) still approximately factorizes.

5.3.2 THE SHOWER DEPTH PARAMETER

Considering the shower development in three dimensions, as with the modified Gaisser-Hillas profile (Eq. (5.10)), it is not trivial to assign a certain depth ΔX to points that are displaced from the shower axis by a distance r. The value of the depth parameter ΔX for any point in the ground plane is obtained by projecting the point onto the shower axis and measuring the distance of this projection to the shower maximum. The total traversed matter for a particle reaching a point away from the shower core is therefore always larger than the value of ΔX of this point. A sketch is given in Fig. 5.5.

For upstream points in the ground plane and for $\theta \neq 0$ the depth ΔX is measured between X_{max} and a projected point in the atmosphere above the ground on the shower axis. If the shower maximum is sufficiently close to the ground, upstream points on the surface can even be assigned negative values of ΔX. This must not be interpreted as particles from an air shower propagating upwards the shower axis to such points, but that in this case the origin of particles is simply before the shower maximum. For the same geometric reasons, for downstream points ΔX is measured down to a point below the ground. In this case, instead of using the actual density of the ground the density profile of the atmosphere is extrapolated (see Appendix D) to obtain a comparable behaviour with respect to upstream points. In the case of $\theta = 0$ all points on the surface plane are assigned the same value of ΔX.

In general, atmospheric conditions are an important factor to consider when calculating ΔX. Near the surface, where 10 m roughly equals $1\,\mathrm{g\,cm^{-2}}$ in terms of traversed matter, the density of the atmosphere needs to be properly understood to accurately assign the

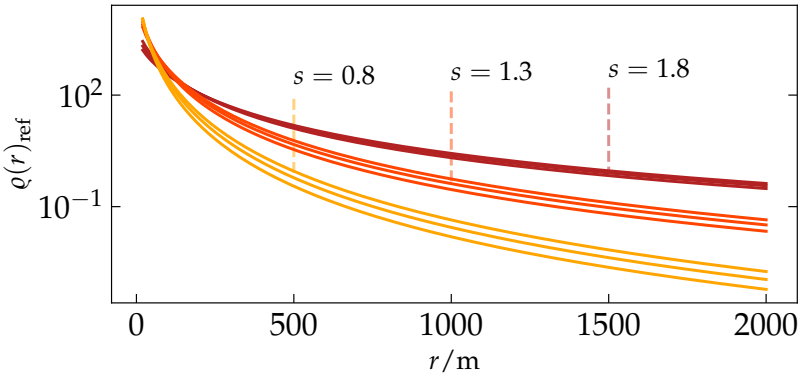


Figure 5.6: NKG profile for fixed N and various values of s with r_{M} set to 80 m, 100 m and 120 m in each case.

correct values for ΔX [122]. This matter is discussed in detail in Appendix D. Assuming an isothermal, exponentially-decreasing density of the atmosphere, ΔX can be analytically calculated with

$$\Delta X = \frac{X_{\mathrm{vg}}}{\cos\theta} \exp\left[-\frac{h + r\cos\psi\,\sin\theta}{h_{\mathrm{s}}}\right] - X_{\mathrm{max}}, \tag{5.11}$$

for any point with a shower-plane distance r, a shower-plane azimuth ψ, and a height h above the ground. The total vertically-integrated mass column X_{vg} of the atmosphere and the scaling height h_{s}, with which the atmospheric density effectively decreases, are in general location- and season-dependent quantities.

5.4 THE LATERAL PROFILE

The lateral distribution of electromagnetic particles in air showers is well described by the NKG function, given in Eq. (3.46). In this work, the reference density $\varrho(r)_{\mathrm{ref}}$ will therefore be described by the NKG function. Using the expected number of particles of a given particle component as N, the reference density reads as

$$\varrho(r)_{\mathrm{ref}} = \frac{N}{2\pi r_{\mathrm{M}}^2} \frac{\Gamma(\frac{9}{2} - s)}{\Gamma(s)\Gamma(\frac{9}{2} - 2s)} \left(\frac{r}{r_{\mathrm{M}}}\right)^{s-2} \left(1 + \frac{r}{r_{\mathrm{M}}}\right)^{s-\frac{9}{2}}. \tag{5.12}$$

As introduced in Section 3.2, r_{M} and s are expected to differ for each component, thus the lateral distribution of particles is significantly different for the four types of particles. The overall lateral distribution of all the particles in an air shower is given by the sum of the distributions of the four contributing particle components.

The NKG function is depicted in Fig. 5.6 for a fixed N. The slope of $\varrho(r)_{\mathrm{ref}}$ hardens significantly with increasing shower age, resulting in an overall broader footprint of the shower. s is individually fixed for each particle component to the value which is obtained from the lateral distribution at $\Delta X = \Delta X_{\mathrm{ref}}$. The development of the lateral profile is then governed by Eqs. (5.10) and (5.12) in terms of absolute scale and slightly in terms of the shape by the implicit dependence of ΔX_{max} and λ on radius. This is an approximation, which is only valid for $0.9 \lesssim s \lesssim 1.3$, that is however very well within the realms of expectations for the Auger SD and showers of $E_0 \gtrsim 10^{19}$ eV since according to Eq. (3.40) this range of s corresponds to $-200\,\mathrm{g\,cm^{-2}} \lesssim \Delta X \lesssim 1000\,\mathrm{g\,cm^{-2}}$. Detectors of the Auger SD are expected to lie very well within this range of ΔX.

The lateral distribution of muons in air showers is expected to be significantly broader than the distribution of electromagnetic particles. The same holds for particles of the jet component, as can be seen in Fig. 3.3. The lateral profile of showers with larger R_μ is thus flatter than those with small R_μ. The change in the slope of the lateral distribution of particles is the main source of the sensitivity on R_μ in the reconstruction algorithm discussed later in Section 6.1.1.

5.5 THE AREAL PARTICLE DENSITY

Combining Eqs. (5.2) to (5.4), (5.10) and (5.12), the expected areal particle density in the shower plane can be given as a sum of the contributions of all particle components. To account for the energy dependence, we use the expected number of particles at the ground for a shower with $R_\mu = 1$ and $E_0 = 10^{19}\,\mathrm{eV}$, written as N_{ref}^{19}, for each component and apply the power law as given in Eq. (5.3). The resulting particle density of each component then reads as

$$\begin{aligned}\varrho(\Delta X, R_\mu) = (a\,(R_\mu - 1) + 1)\left(\frac{E}{10^{19}\,\mathrm{eV}}\right)^{\gamma}\left(\frac{\Delta X - \Delta X_1}{\Delta X_{\mathrm{ref}} - \Delta X_1}\right)^{\frac{\Delta X_{\max} - \Delta X_1}{\lambda}} \mathrm{e}^{-\frac{\Delta X - \Delta X_{\mathrm{ref}}}{\lambda}} \\ \times \frac{N_{\mathrm{ref}}^{19}}{2\pi r_{\mathrm{M}}^2}\frac{\Gamma(\frac{9}{2} - s)}{\Gamma(s)\Gamma(\frac{9}{2} - 2s)}\left(\frac{r}{r_{\mathrm{M}}}\right)^{s-2}\left(1 + \frac{r}{r_{\mathrm{M}}}\right)^{s-\frac{9}{2}},\end{aligned} \tag{5.13}$$

where indices i are omitted for better legibility. The parameters a, γ, ΔX_1, $\Delta X_{\max}$, λ, N_{ref}^{19}, s, and r_{M} define the behaviour of each particle component and need to be determined individually. As outlined in Chapter 3, it is expected that $\gamma_i \lesssim 1$ and that $\gamma_\mu \simeq 1 - \beta$ according to Eq. (3.53).

The total number of particles arriving at an SD station is thus

$$n_{\mathrm{tot}} \propto \sum_i \varrho_i(\Delta X, r), \quad \text{for} \quad i \in \{\mathrm{e}\gamma, \mu, \mathrm{e}\gamma(\pi), \mathrm{e}\gamma(\mu)\}. \tag{5.14}$$

n_{tot} is written in lower case to avoid confusion with the absolute number of particles in the shower. The factor converting the areal density of particles in the shower plane into the number of particles arriving at the detector and is defined by the geometry of the event and the detector.

5.6 THE ARRIVAL TIMES OF PARTICLES AT THE GROUND

The arrival time of particles at the ground is closely related to the longitudinal profile of the shower. Therefore, it can also be universally described for showers from UHECRs, given the depth of the shower maximum and the relative abundance of the particle components are known. Especially for muons, which propagate mostly rectilinearly from the point of their creation to the surface, the dependence of the longitudinal profile on the expected distribution of arrival times is well understood [123]. Connecting each point in time to a slant depth from which particles originate, the flux of particles arriving at ground $\mathrm{d}n/\mathrm{d}t$ is directly governed by the longitudinal profile of the shower,

$$\frac{\mathrm{d}}{\mathrm{d}t} n \propto N(X(t)). \tag{5.15}$$

In this section, the arrival time of muons is derived from the longitudinal development of muons in the shower as a function of time and position of a possible sampling area,

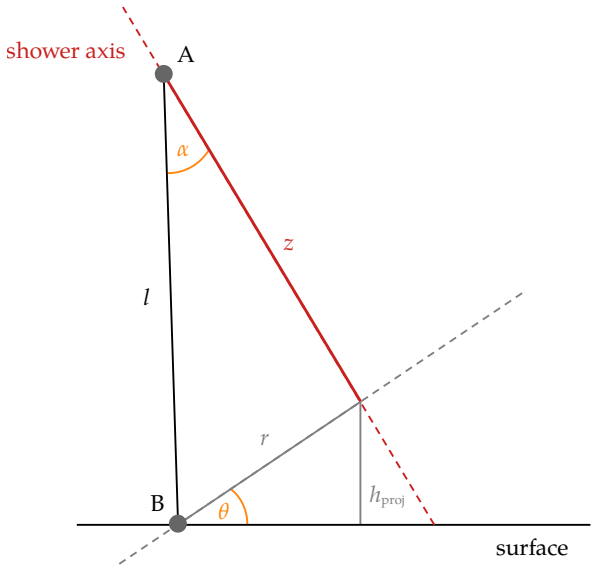


Figure 5.7: Sketch of the geometry for a single particle propagating rectilinearly from its point of creation A in the shower core to the point B on the ground. The plane front containing the point B is depicted as a gray dashed line. The shower axis is depicted as red dashed line. The distance z from point A to the plane front is depicted in red. The angle α between the trajectory of the particle and the shower axis, and the zenith angle θ are depicted in orange.

mostly following the work given in Refs. [124] and [125]. A simplified ansatz will then be introduced to also describe particles, for which the assumption of rectilinear propagation breaks down [126]. This model augments and completes previous universality-based models of the arrival times of particles [127, 128].

5.6.1 THE ARRIVAL TIMES OF MUONS AT THE GROUND

The following steps are performed strictly under the assumption of rectilinear propagation and thus in general hold only for muons, even though in a later section the other particle components are considered as well.

Assume all particles propagate along straight trajectories with the speed of light c. Each particle travels a distance l from its point of origin A to the point B on the surface that is at a distance r from the shower axis. A sketch is given in Fig. 5.7. Using the position of the point B, l can be expressed in terms of the distance of A to the *plane front* of the shower, i.e. the plane which is perpendicular to the shower axis while containing the shower core. The plane front of the shower propagates with the speed of light along the extrapolated trajectory of the CR and coincides with the shower plane when the shower core reaches the ground. Let the distance between the point A and the plane front be z, then

$$l = \sqrt{r^2 + z^2}. \tag{5.16}$$

The time t for the particle to arrive at the surface is defined relatively to the time of the plane front arriving at the same point, $t_{\mathrm{pf}} \equiv 0$, so that $ct = l - z$, thus

$$z = \frac{1}{2}\left(\frac{r^2}{ct} - ct\right). \tag{5.17}$$

Using Eq. (5.17), the height at which a particle was created can be derived from its arrival time relative to the time of the plane front at the ground. For a zenith angle θ, this height is

given by

$$h(t) = z\,\cos\theta + h_{\text{proj}} = \frac{1}{2}\left(\frac{r^2}{ct} - ct\right)\cos\theta + h_{\text{proj}}, \tag{5.18}$$

with

$$h_{\text{proj}} = r\cos\psi\sin\theta. \tag{5.19}$$

Thus, using an isothermal model of the atmosphere, similarly to Eq. (5.11), the origin depth X of particles arriving at time t can be expressed explicitly as

$$X = \frac{X_{\text{vg}}}{\cos\theta}\mathrm{e}^{-\frac{h(t)}{h_{\text{s}}}} = \frac{X_{\text{vg}}}{\cos\theta}\exp\left[-\frac{1}{h_{\text{s}}}\left(\frac{1}{2}\left(\frac{r^2}{ct} - ct\right)\cos\theta + h_{\text{proj}}\right)\right]. \tag{5.20}$$

Aside from dependence of the arrival time on the origin depth, for geometric reasons t is expected to depend on the angle α and consequently on r. This is because the larger α is, the longer the length l of the trajectory becomes. Given the point for which $r = \sin\alpha = 0$, i.e. the shower core at the ground, all particles arrive at the same instance at $t = 0$ in this picture. In terms of time and radius $\sin\alpha = r/l$ can be written as

$$\sin\alpha = \frac{2\,ct\,r}{r^2 + c^2t^2}. \tag{5.21}$$

Furthermore, particles created in the shower core are not expected to be distributed evenly in $\sin\alpha$. In a simplistic approach, $\sin\alpha$ is estimated from the transverse momentum of particles, p_{T}, and the energy spectra [124]. Assuming that the spectrum of N particles can be described by

$$\frac{\mathrm{d}^2N}{\mathrm{d}p_{\text{T}}\,\mathrm{d}E} = N\,\frac{p_{\text{T}}}{Q}\mathrm{e}^{-\frac{p_{\text{T}}}{Q}}\frac{s}{m}\left(\frac{E}{m}\right)^{-(s+1)}\left(1 - \frac{\rho\,\iota\,l}{E}\right)^{\kappa}, \tag{5.22}$$

where $Q \simeq 170\,\text{MeV}$ is the characteristic momentum of hadronic interactions, m is the mass of the particles, s is the shower age, and $\iota \simeq 2\,\text{MeV}/(\text{g}\,\text{cm}^{-2})$ is the energy loss due to ionization. In combination with the path length l, as given in Eq. (5.16), and the atmospheric density ρ, the last term of Eq. (5.22) corresponds to the loss of particles after a distance l, where the spectral slope for muons is given [129] by $\kappa \simeq 0.8$. To find the expression for the number of particles created with an angle α relative to the shower axis, one can use the relation $p_{\text{T}} \simeq E\sin\alpha$. Thus, we obtain the expression

$$\frac{\mathrm{d}^2N}{\mathrm{d}\sin\alpha\,\mathrm{d}E} = E\,\frac{\mathrm{d}^2N}{\mathrm{d}E\,\mathrm{d}p_{\text{T}}}, \tag{5.23}$$

and consequently

$$\frac{\mathrm{d}N}{\mathrm{d}\sin\alpha} = \int_{\epsilon_0}^{\infty} E\,\frac{\mathrm{d}^2N}{\mathrm{d}E\,\mathrm{d}p_{\text{T}}}\,\mathrm{d}E \tag{5.24}$$

$$= N\,s\left(\frac{m}{Q}\right)^{s}(\sin\alpha)^{s-1}\,I(l, \sin\alpha), \tag{5.25}$$

for all particles above a threshold energy ϵ_0. As discussed in Ref. [125], a good approximation for the integral is

$$I(l, \sin\alpha) \approx \Gamma(2 - s, x_0 + y_0) - x_0\kappa\,\Gamma(1 - s, x_0 + y_0), \tag{5.26}$$

with $x_0 = \rho\,\iota\, r/Q$ and $y_0 = m \sin\alpha/Q$. Using Eq. (5.21), I can be expressed in terms of r and t. Performing a change in coordinates, the differential number of particles arriving at the surface at a distance r and time t can thus be written as

$$\frac{\mathrm{d}^2 n}{\mathrm{d}r\,\mathrm{d}t} = \frac{\mathrm{d}N(X(t))}{\mathrm{d}\sin\alpha}\,\mathfrak{J}, \tag{5.27}$$

where the Jacobian $\mathfrak{J}$ is given by

$$\mathfrak{J} = \left|\frac{\mathrm{d}X}{\mathrm{d}t}\frac{\mathrm{d}\sin\alpha}{\mathrm{d}r} - \frac{\mathrm{d}X}{\mathrm{d}r}\frac{\mathrm{d}\sin\alpha}{\mathrm{d}t}\right| = \frac{X(t)}{h_s t}\left(\frac{r^2 - c^2t^2}{r^2 + c^2t^2}\right)^2, \tag{5.28}$$

using the expressions from Eqs. (5.20) and (5.21). Using the Gaisser-Hillas longitudinal profile from Eq. (5.5) as $N(X)$, the expected differential number of particles at the ground is given by

$$\begin{aligned}\frac{\mathrm{d}^2 n}{\mathrm{d}r\,\mathrm{d}t} = N_{\max}\, s \left(\frac{m}{Q}\right)^s I(r,t)\, X_{\mathrm{vg}} \frac{(r^2 - c^2t^2)^2}{h_s\, t\, (r^2 + c^2t^2)^2} \left(\frac{2\,ct\,r}{r^2 + c^2t^2}\right)^{s-1} \\ \times \exp\left[-\frac{h(t)}{h_s} + \frac{X_{\max}}{\lambda} - \frac{X_{\mathrm{vg}}}{\lambda\,\cos\theta}\,\mathrm{e}^{-\frac{h(t)}{h_s}}\right] \left(\frac{X_{\mathrm{vg}}}{X_{\max}\,\cos\theta}\,\mathrm{e}^{-\frac{h(t)}{h_s}}\right)^{\frac{X_{\max}}{\lambda}}.\end{aligned} \tag{5.29}$$

It was verified [125] using MC simulations that Eq. (5.29) accurately describes the arrival times of muons in extensive air showers. In Fig. 5.8 the muon arrival times obtained from a non-thinned CORSIKA shower are depicted together with the predicted arrival times according to Eq. (5.29). Solid lines show the arrival times obtained from a exact numerical calculation of Eq. (5.25), whereas dashed lines are obtained using Eq. (5.26). A similar but simpler functional form of the differential arrival times of particles has been empirically discovered in Ref. [130].

For small sampling areas, comparable to the size of an SD station, the fluctuation in the total number of particles is severe. Also, no sophisticated considerations in terms of the lateral decrease in particle density have been made for the derivation of Eq. (5.29). Therefore, even though the overall shape of the histogram of arrival times of particles is well represented by Eq. (5.29), the total number of particles has to be fixed from a best fit or from a model of the expected particle density such as given by Eq. (5.13).

5.6.2 THE ARRIVAL TIMES OF ELECTROMAGNETIC PARTICLES

For electromagnetic particles the assumption of perfectly rectilinear propagation does not hold, except for the particles in the upper end of the energy spectrum. Consequently, the agreement of Eq. (5.29) with the number of eγ particles arriving at a certain time is not accurate. Even though the qualitative behaviour of $\mathrm{d}n/\mathrm{d}t$ is well reproduced, the values for $X_{\max}$ needed for Eq. (5.29) to match the data disagree with the respective values from the longitudinal profiles [125]. Furthermore, if one tries to fix all occurring parameters from a best fit to the arrival times, the resulting values for $X_{\max}$ and λ are highly correlated and will thus lack any predictive power or interpretability. A simpler and more robust function that describes $\mathrm{d}n/\mathrm{d}t$ as a function of $X_{\max}$ thus needs to be found.

As shown in Ref. [109], the number of particles arriving at time t, can be well represented by a log-normal distribution function, $\mathfrak{ln}$, such that

$$\frac{\mathrm{d}}{\mathrm{d}t}n = n_{\mathrm{tot}}\,\mathfrak{ln}(\mu,\sigma) = \frac{n_{\mathrm{tot}}}{\sqrt{2\pi}\,\sigma\,(t - t_0)}\,\exp\left[-\frac{1}{2\sigma^2}\ln^2\left(\frac{t - t_0}{\mathrm{e}^{\mu}}\right)\right], \tag{5.30}$$

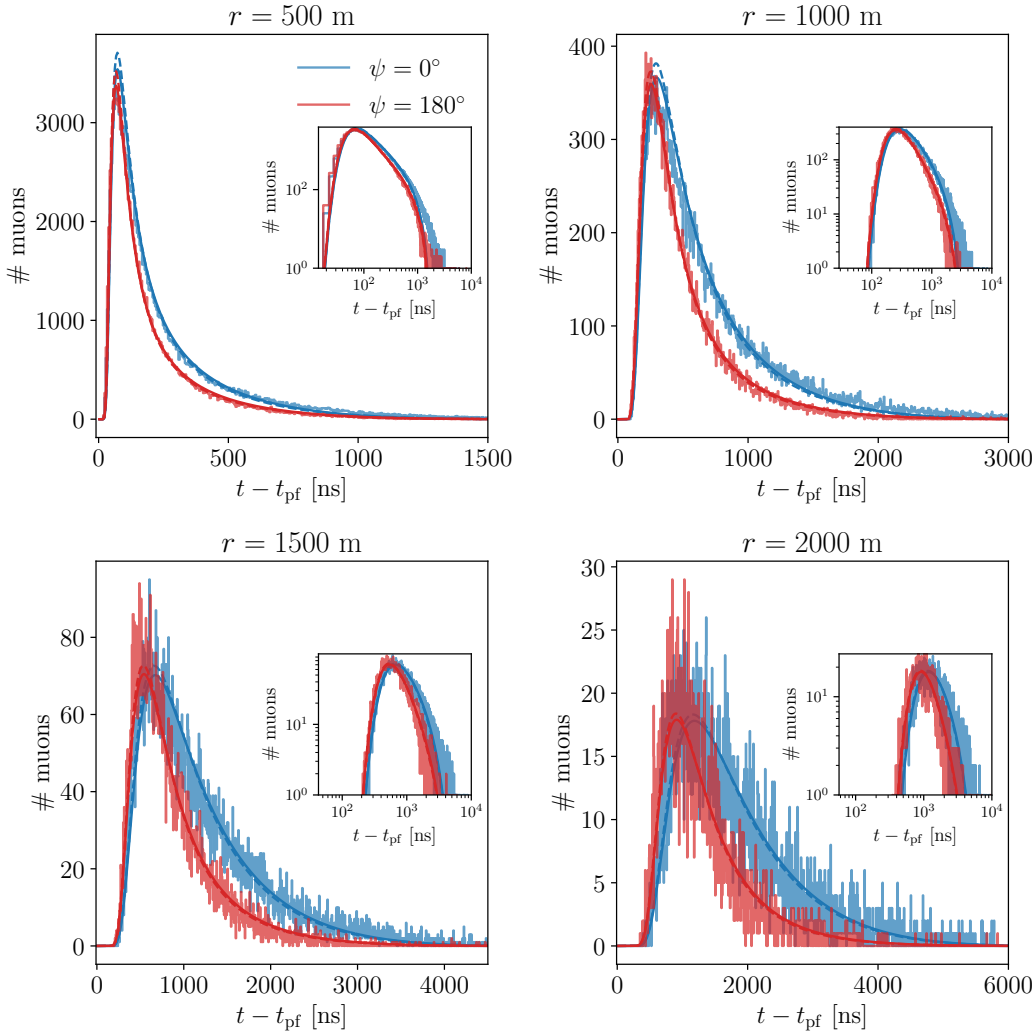
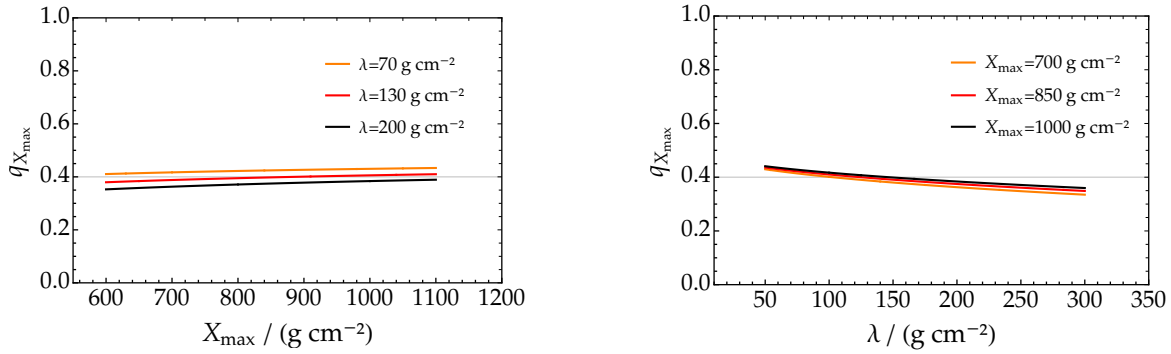


Figure 5.8: Distribution of the arrival times of muons relative to the arrival time t_{pf} of the plane front for different shower plane radii and azimuth. The simulated distributions of arrival times are depicted as histograms and are compared to the predicted distributions of arrival times according to Eq. (5.29), depicted as red and blue lines for upstream and downstream regions, respectively. The data were obtained from a non-thinned CORSIKA shower with $E_0 = 10^{18.5}\,\mathrm{eV}$ and $\theta = 30°$. The total number of particles was obtained from an individual best fit for each sampling area, while X^{μ}_{max} and λ were obtained from a global fit, matching the longitudinal profile of the shower. For each panel the inset depicts the same data in double-logarithmic scale, see Ref. [125].

where t_0 is related to the earliest particles arriving at a given sampling area. The parameters σ and e^{μ} are related to the first and second moment of the log-normal distribution[2]. It was shown in Ref. [125] that for a reasonable set of parameters Eq. (5.29) and Eq. (5.30) are qualitatively identical up to $t \simeq 4000\,\mathrm{ns}$. However, the dependence of σ, μ, and t_0 on X_{max} and all further occurring parameters of Eq. (5.29) is not trivial. All quantities except for X_{max} are assumed constant for all showers and thus σ, μ, and t_0 are going to be described as functions of ΔX.

The log-normal distribution was already successfully used to describe the signal de-

[2] e^{μ} marks the median of the distribution.



(a) (b)

Figure 5.9: Dependence of $q_{X_{max}}$ from the Gaisser-Hillas profile on the absolute values of X_{max} and λ.

posited in SD stations in preceding analyses [127, 128, 131]. Furthermore, it was empirically proven in Ref. [127] that the start times of the signal in SD stations, i.e. the arrival times of the first particles in a respective sampling area, as well as expected values for μ, depend on X_{max}. This is not surprising, since even though the rectilinear propagation approach does not hold for the particles of the eγ component, the arrival times of particles are still expected to be time ordered with respect to the depth of their origin. In this case, even though the behaviour of the arrival times with respect to $\sin\alpha$ is rather chaotic, Eq. (5.15) still holds. This implies that for any sampling area at $r > 0$, the fraction q of the total number n_{tot} of particles that arrived after the time t_q is related to the depth X_q by

$$q = \frac{1}{n_{tot}} \int_0^{t_q} \frac{dn}{dt}\, dt \simeq \frac{1}{\mathfrak{N}} \int_0^{X_q} N(X)\, dX, \tag{5.31}$$

where $\mathfrak{N} = \int_0^\infty N(X) dX$ is the total integrated longitudinal profile and $t = 0$ refers to the time of the plane front. In this picture, the start time t_0 is directly related to the depth of first interaction X_1 that for the sake of simplicity we set to $X_1 = 0$ in the next step. Given the fact that we want to express the arrival times of particles in terms of X_{max}, it is tempting to use $t_{q_{max}}$, i.e. to calculate q_{max} at $X_q = X_{max}$. For the Gaisser-Hillas longitudinal profile this evaluates to

$$q_{X_{max}} = 1 - \frac{\frac{\lambda}{X_{max}}\, \Gamma\left(\frac{X_{max}+\lambda}{\lambda}, \frac{X_{max}}{\lambda}\right)}{\Gamma\left(\frac{X_{max}}{\lambda}\right)}, \tag{5.32}$$

where $\Gamma(y, x)$ is the upper incomplete gamma function. Numerically the $q_{X_{max}}$ is approximately constant with the value $q_{X_{max}} \simeq 40\%$ for any physically reasonable set of values for X_{max} and λ. The explicit behaviour of $q_{X_{max}}$ with respect to X_{max} and λ is depicted in Fig. 5.9.

The time t_{40}, after which 40% of the total number of particles have arrived at a given sampling area, is thus assumed to be directly related to X_{max}. Moreover, t_{40} is assumed to be dependent *only* on the shower geometry and X_{max}. In this way, the vast number of parameters needed to describe the arrival times of particles in Eq. (5.29) is drastically reduced. Eq. (5.11) and Eq. (5.20) can be combined for a point on the ground at $h = 0$, so

that $X_{\max}$ expressed as a function of t_{40} reads as

$$X_{\max} = \frac{X_{\mathrm{vg}}}{\cos\theta} \exp\left[-\frac{1}{2}\left(\frac{r^2}{ct_{40}} - ct_{40}\right)\frac{\cos\theta}{h_{\mathrm{s}}}\right] \mathrm{e}^{-\frac{h_{\mathrm{proj}}}{h_{\mathrm{s}}}} \tag{5.33}$$

$$= (\Delta X + X_{\max}) \exp\left[-\frac{1}{2}\left(\frac{r^2}{ct_{40}} - ct_{40}\right)\frac{\cos\theta}{h_{\mathrm{s}}}\right]. \tag{5.34}$$

Note that in the second line the shower-depth parameter ΔX includes the dependence on the shower-plane azimuth ψ. Inverting this expression, t_{40} can be written as a function of $X_{\max}$,

$$ct_{40} = \sqrt{R_{X_{\max}}^2 + r^2} - R_{X_{\max}}, \tag{5.35}$$

where

$$R_{X_{\max}} = \frac{h_{\mathrm{s}}}{\cos\theta} \ln\left(\frac{\Delta X}{X_{\max}} + 1\right) \tag{5.36}$$

can be identified as the radius of a spherical front of particles originating from the shower maximum.

Because the model described by Eq. (5.35) is based on the assumption that particles reach a particular point on the ground approximately time ordered with respect to their point of origin, it is not valid for regions in which $\Delta X < 0$. For $\Delta X < 0$, particles arriving later than t_{40} but time ordered with respect to their origin in the shower core would propagate upwards the shower axis. This becomes apparent, for example, if the shower maximum is below the ground and $\theta \simeq 0°$. In this case, all points on the surface are located at $\Delta X < 0$ and no particle originating from the shower maximum would reach any station. Consequently, t_{40} is then not directly dependent on $X_{\max}$. In most cases, however, $\Delta X > 0$ and the model is applicable.

The shape parameter μ of the log-normal distribution can be written explicitly as a function of t_{40} and thus, using Eq. (5.35), also as a function of $X_{\max}$. For log-normally distributed arrival times of particles, e^{μ} as a function of t_{40} reads as

$$\mathrm{e}^{\mu} = k(t_{40} - t_0) \tag{5.37}$$

with

$$k = \exp\left[-\sqrt{2}\,\sigma\,\mathrm{erf}^{-1}(2\times 0.4 - 1)\right] \simeq 1.288\,\sigma. \tag{5.38}$$

In fact, e^{μ} can be expressed in terms of any quantile of the distribution, if the argument of the inverse error function in Eq. (5.38) is adjusted accordingly[3]. In this way, the log-normal distribution function of arrival times is given as

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}n &= n_{\mathrm{tot}}\, \ln(t_{40}, \sigma) \\ &= \frac{n_{\mathrm{tot}}}{\sqrt{2\pi}\,\sigma\,(t - t_0)} \exp\left[-\frac{1}{2\sigma^2}\left(\ln\left(\frac{t-t_0}{t_{40}-t_0}\right) + \sqrt{2}\sigma\,\mathrm{erf}^{-1}(2\times 0.4 - 1)\right)^2\right].\end{aligned} \tag{5.39}$$

Eq. (5.35) in combination with Eq. (5.39) has two important implications. First, for very small radii $ct_{40} \simeq 0$, meaning that for $r \to 0$ all particles arrive asymptotically at the same time together with the plane front of the shower. Second, for very inclined showers, i.e.

[3] For t_{50} the expressions reduces trivially to $\mathrm{e}^{\mu} = t_{50} - t_0$.

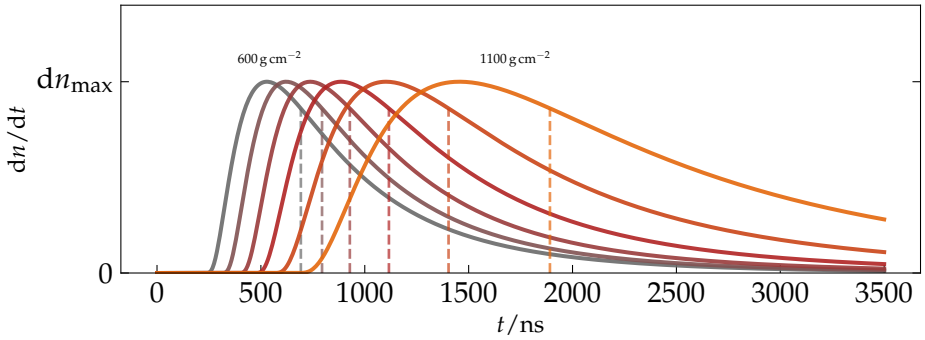


Figure 5.10: Normalized differential arrival times of particles for an example sampling area at $r = 1500\,\mathrm{m}$ and $\psi = \pi/2$ as predicted by Eqs. (5.30) and (5.37). Showers at $\theta = 32^\circ$ are considered for 6 equidistant values of $X_{\max}$ between $600\,\mathrm{g\,cm^{-2}}$ and $1100\,\mathrm{g\,cm^{-2}}$. Corresponding values of t_{40} are indicated by dashed lines. σ is set as a constant to the value of 0.8 for the sake of simplicity. The respective values of t_0 are obtained from the model that is parametrized in Section 5.8.

$\theta \gtrsim 50^\circ$, where $\sec\theta$ and ΔX are rather large, the gradient of t_{40} with respect to ΔX is expected to be small and sensitivity to $X_{\max}$ is lost.

Using Eqs. (5.37) and (5.39), the differential arrival times of particles can be expressed explicitly as a function of $X_{\max}$ and the shower-plane coordinates of the considered sampling area. A depiction of $\mathrm{d}n/\mathrm{d}t$ is given in Fig. 5.10 for different values of $X_{\max}$ for CRs with $\theta = 32^\circ$. The orange curve in Fig. 5.10 corresponds to a shower for which the shower maximum at the slant depth $X_{\max} = 1100\,\mathrm{g\,cm^{-2}}$ is already below the ground.

From Fig. 5.10 it is immediately apparent that the shape of the distribution of arrival times of particles alone does not carry the first-order information on the depth of the shower maximum. For $X_{\max} \lesssim 800\,\mathrm{g\,cm^{-2}}$ the arrival times simply appear shifted from left to right without any significant change in the overall shape. The arrival times are only increasingly stretched out for large values of $X_{\max}$. This effect, however, is reduced if σ is not assumed constant, as it is for Fig. 5.10. Only if time is measured relative to the plane front, or if somehow equivalently the geometry of the shower is taken into account, the first-order information on $X_{\max}$ can be extracted from the arrival times of particles.

5.7 PARAMETRIZATION OF THE AREAL DENSITY

In this section, we address the parametrization of the model given above using simulated detector responses. In contrast to the previous works [109, 128, 132, 133], where an average detector response to particles from CORSIKA showers was simulated, full detector simulations are employed for the parametrization in this work. We describe the preparation of the simulation data, as well as the procedure to fit the model to the data. We performed all the described steps using the Offline [134] software framework and the *universality-v2* toolkit [135].

The parametrization of the model is performed directly in terms of signal instead of particle densities. In this way effects from the detector geometry and the detector response to single particles are directly absorbed in the resulting parameters of the model. This is, however, only possible because for each detector and particle component i we expect the average signal component in each detector to be directly proportional to the density of the

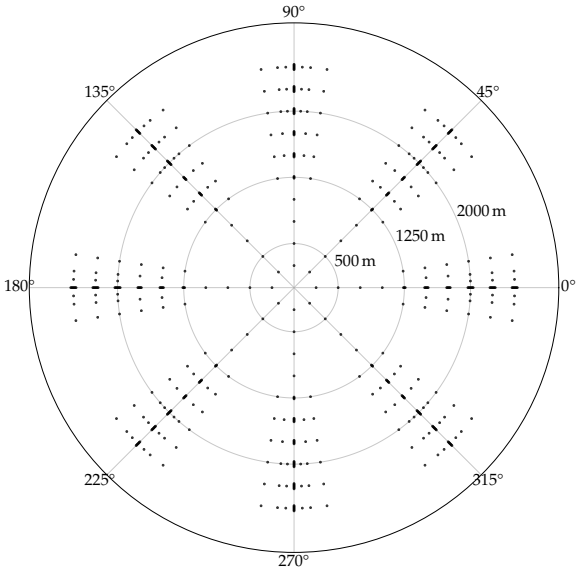


Figure 5.11: Array of simulated virtual detector stations in the shower plane used for the parametrization of the Universality model.

particles of the given component,

$$S_i \propto \varrho_i. \tag{5.40}$$

This very strong assumption, however, is valid because of the universality of the particle spectra, as discussed in Section 3.5.2, and thus because of the average signal deposit per particle being constant with respect to the shower age, as discussed in Section 5.3, especially for the $e\gamma$ component. Furthermore, due to the second-order effects such as differences in the average track lengths of particles in the detectors, we expect slight deviations of all parameters that depend on geometry for different detectors and components. The parametrization of the model is thus performed separately for the SSDs and WCDs as well as for each particle component. For this work, independent parametrizations based on proton shower libraries simulated using the EPOS-LHC, QGSJET-II, and SIBYLL model were produced.

Since the Universality model is parametrized solely on proton showers – and consequently on the average of proton showers – the resulting parametrized signals correspond, by definition, to a shower with $R_\mu = 1$. On a shower-to-shower basis, the signal of the individual components must be scaled according to the respective value of R_μ.

The resulting Universality model based on showers simulated with EPOS-LHC was already discussed in Ref. [136]. All numerical values resulting from the parametrization for all hadronic interaction models can be found in tabular form in Appendix A.

5.7.1 SIMULATION DATA

The simulated detector responses for the AugerPrime SD were obtained with the $\overline{\text{Off}}\underline{\text{line}}$ software framework [137] using proton showers simulated with CORSIKA 7.56 (and upwards) employing EPOS-LHC 1.99, QGSJET-II 04 and SIBYLL 2.3c models. Illustrations in this section show the parametrization of showers generated with EPOS-LHC. The shower libraries for EPOS-LHC and QGSJET-II consist of 2880 simulated showers, each. Both libraries are divided into three discrete primary energies, $\lg(E_0/\text{eV}) \in \{19, 19.5, 20\}$, and eight discrete zenith angles, $\theta \in \{0°, 12°, 22°, 32°, 38°, 45°, 56°, 65°\}$. For each zenith angle and primary energy 120 showers are simulated, which are subdivided into 12×10 for 12

different tabulated monthly atmospheric profiles. The shower library for SIBYLL follows the same scheme, however only consists of CR showers with $E_0 = 10^{19}$ eV.

Each shower is simulated in the AugerPrime detector using the standard SD simulation procedure of the Offline software. The simulated array of stations, however, is significantly modified. To provide the required amount of information on the particle densities of the showers, the SD standard array is augmented by several virtual dense stations. In this way, much more information is obtained from the simulated responses of the SD station detectors to each shower. When using virtual dense stations, usually 10 stations are placed on a ring with shower-plane radius of $r = 1000$ m, equally distributed in ψ. For the parametrization of a model that is valid for all shower plane radii at which stations are expected to detect signal, many more of these dense station rings are required. Dense stations are placed on concentric rings in the shower plane every 250 m from $r = 250$ m to $r = 2500$ m at eight equally distributed azimuth angles[4]. The virtual super-dense station grid is depicted in the shower plane in Fig. 5.11. The stations are placed exactly so that sampling areas in which de-thinned particles from the same origin do not overlap[5]. From $r = 1250$ m outwards, extra stations are placed around a central main station for each sampling area. With these the expected fluctuations of signal due to low particle multiplicities can be examined. In total 528 stations equipped with an SSD and WCD each are simulated for every shower. The simulated data is converted and stored in an HDF5 [139] database, where the library of signal responses is sorted in terms of the properties of the shower and the positions of the stations. For every sampling area a list of the detector responses of the SSDs and WCDs is created for multiple different values of ΔX and R_μ. In addition to the total signal, the signal traces of every detector are stored together with the average trace of every local detector group. The total signal is stored in addition to the signal components, S_i, which are the hypothetical signals from only those particles contained in the respective component.

5.7.2 PARAMETRIZATION OF THE LONGITUDINAL PROFILE

For the sampling areas[6] at $\psi \neq 180°$ and $\psi \neq 0°$, the database of signal components is fitted to the modified Gaisser-Hillas function, written in terms of signal as

$$S(\Delta X, r) = S_{\mathrm{ref}}(r) \left(\frac{\Delta X - \Delta X_1}{\Delta X_{\mathrm{ref}} - \Delta X_1} \right)^{\frac{\Delta X_{\max} - \Delta X_1}{\lambda}} \mathrm{e}^{-\frac{\Delta X - \Delta X_{\mathrm{ref}}}{\lambda}}, \tag{5.41}$$

where again indices i are dropped for better legibility. The value for ΔX_{ref} is fixed to $200\,\mathrm{g\,cm^{-2}}$, which is just a bit larger than the expected values for $\Delta X_{\max}$ of the $\mathrm{e}\gamma$ component, so that $S_{\mathrm{ref}}(r)$ can be parametrized in the next step with sufficient signal even for distant sampling areas. ΔX_1 is fixed to $-600\,\mathrm{g\,cm^{-2}}$ for the $\mathrm{e}\gamma$ and μ components and to $-500\,\mathrm{g\,cm^{-2}}$ for the $\mathrm{e}\gamma(\pi)$ and $\mathrm{e}\gamma(\mu)$ components. ΔX_1 describes the depth at which the signal of a respective component is expected to start for a hypothetical station located at a small distance and at a largely negative ΔX. The fixed values for ΔX_1 are motivated by studies where ΔX_1 was treated as a free parameter of the fit. In this way, for all radii, a set of best fit values for $S_{\mathrm{ref}}(r)$, $\Delta X_{\max}$ and λ, is obtained. The radial dependence of these quantities is fixed according to the steps described in the next section.

[4] To save memory and CPU time the innermost two rings of dense stations, where a vast amount of particles is expected, are not simulated for showers with primary energy of 10^{20} eV.

[5] The process of thinning is briefly mentioned in Section 3.5. To save CPU time and memory, in the late stages of a CORSIKA shower only a subset of particles is actually simulated, but weighted according to the number of particles that are not simulated. In the later simulation of the SD response, particles are multiplied according to their weight and thus the shower is de-thinned. The de-thinned set of particles is distributed in a respective area in the shower plane [138].

[6] Asymmetries of the signal for upstream and downstream stations are discussed in Section 5.7.5.

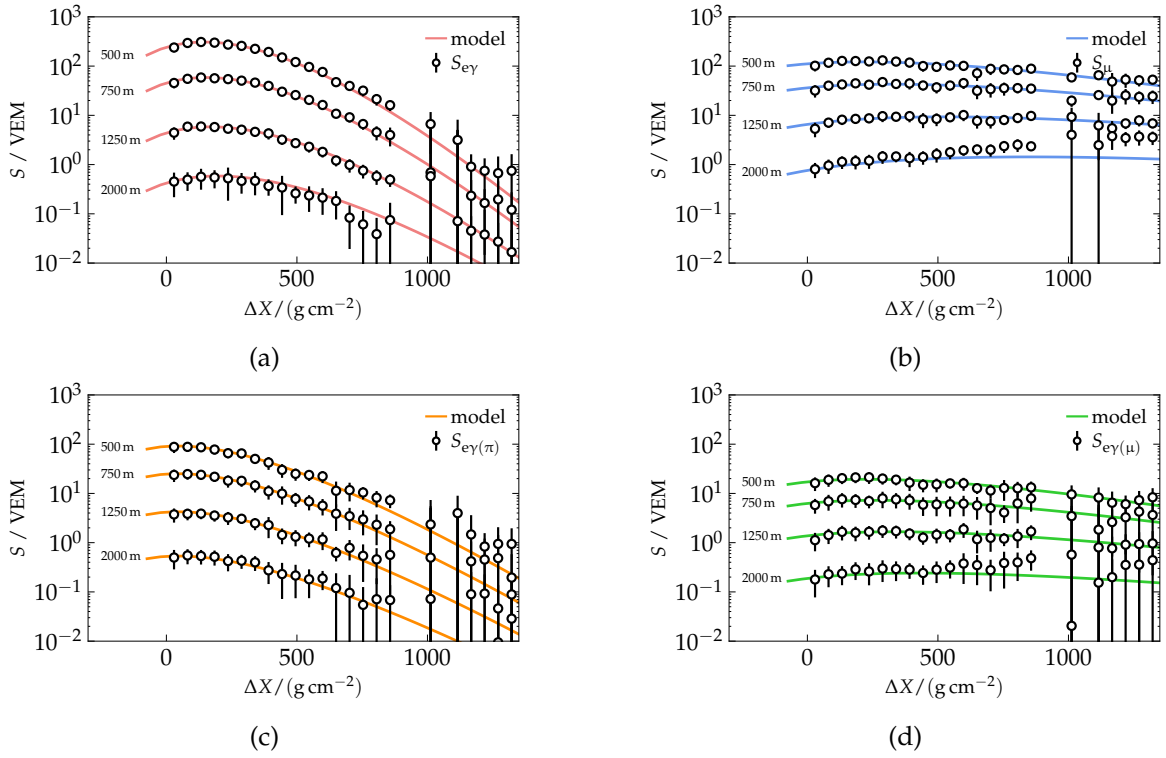


Figure 5.12: Signal response of simulated WCDs as a function of ΔX along with the resulting model according to Eq. (5.10) for the individual components. Error bars indicate one standard deviation of the underlying data.

The data used to perform the fit of the longitudinal development of showers with $E_0 = 10^{19}$ eV is depicted in Fig. 5.12 for the signals deposited in WCDs, and in Fig. 5.13 for SSDs, respectively, along with the signal predicted from the overall resulting model. The same data from showers with $E_0 = 10^{20}$ eV is given in Figs. B.6 and B.7, where the corresponding signal is scaled down by a factor of $10^{-\gamma}$ for better comparison (the values for γ are discussed in Section 5.7.4). Only a subset of the data at representative shower-plane radii is shown in each figure to avoid extensive overlap of the data. Already here it is apparent that the model succeeds to match the data. Note that the agreement of the model with the simulated data of the muonic component, shown in Fig. 5.12 (b), is only poor at large distances and large ΔX where the number of muons reaches approximately only 1 per detector. The discrepancy is not present for higher particle densities at higher primary energies, as it is shown in Fig. B.6 (b).

5.7.3 PARAMETRIZATION OF THE LATERAL PROFILE

The lateral distribution of particles in the Universality model is mainly governed by the shape of $S_{\mathrm{ref}}(r)$ that is given by the NKG function written in terms of signal as

$$S_{\mathrm{ref}}(r) = S_{\mathrm{ref}}^{19} \left(\frac{E_0}{10^{19}\mathrm{eV}}\right)^{\gamma} \frac{1}{2\pi r_{\mathrm{M}}^2} \frac{\Gamma(\frac{9}{2}-s)}{\Gamma(s)\Gamma(\frac{9}{2}-2s)} \left(\frac{r}{r_{\mathrm{M}}}\right)^{s-2} \left(1+\frac{r}{r_{\mathrm{M}}}\right)^{s-\frac{9}{2}}, \tag{5.42}$$

where a constant reference signal, S_{ref}^{19}, is introduced, which is representative for showers with $E_0 = 10^{19}$ eV. S_{ref}^{19}, r_{M}, and s are fixed using the results of $S_{\mathrm{ref}}(r)$ from the previous section. Since Eq. (5.42) is formulated in terms of signal instead of particle density and then fit to an ensemble of best fit results for the signal at a fixed depth ΔX_{ref}, the interpretability of

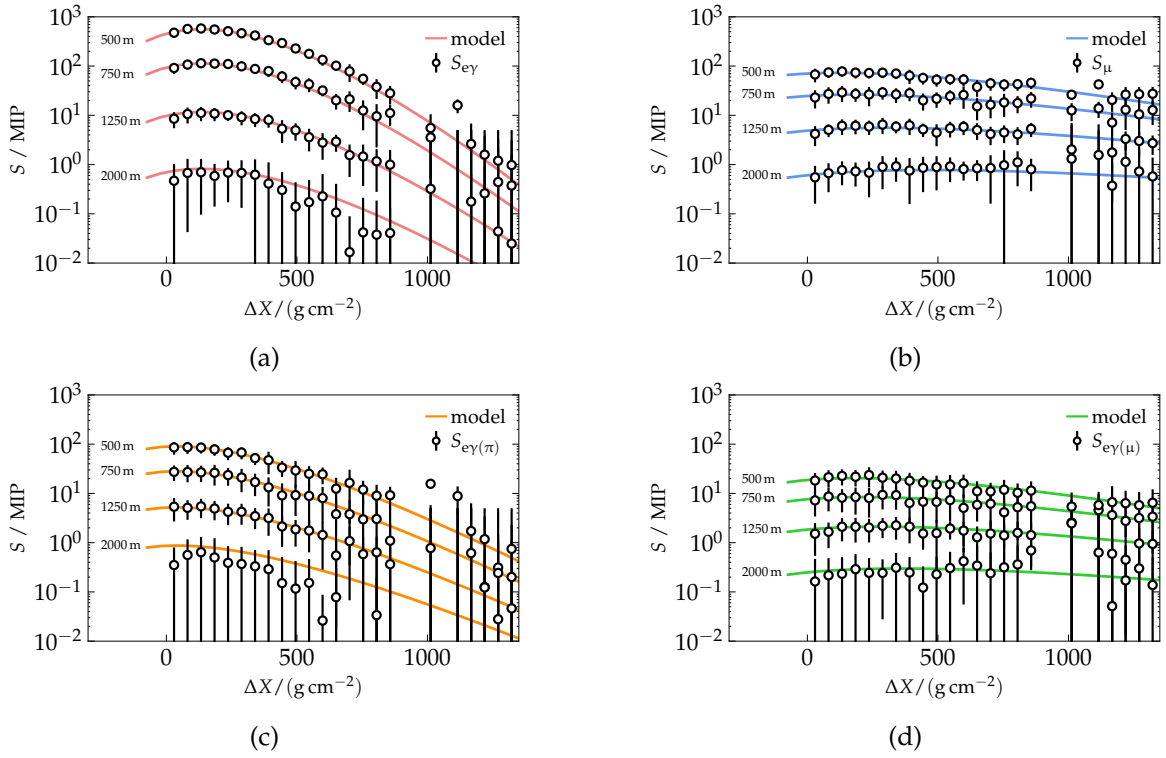


Figure 5.13: Signal response of simulated SSDs as a function of ΔX along with the resulting model according to Eq. (5.10) for the individual components. Error bars indicate one standard deviation of the underlying data.

the obtained best fit values of s and r_M is limited. The lateral distribution of muons and the corresponding $e\gamma(\mu)$ component is expected to differ from the shape of the classical NKG function, as discussed in Section 3.2. To handle the flattening of the lateral distribution by both, kinematic effects and effects from the detector thresholds that do not allow for the detection of arbitrarily small signals, a minor modification to Eq. (5.42) is required. This modification, however, is only relevant for sampling areas where the density of particles is of the order of one per detector. Small detector signals are thus weighted by a sigmoid-like function,

$$\mathfrak{w}(S) = \frac{1}{\pi}\arctan\left(\frac{\ln\left(S_{\mathrm{th}}/S\right)}{\ln c_{\mathrm{th}}}\right) + \frac{1}{2}, \tag{5.43}$$

and continuously attenuated in the region of $S \simeq S_{\mathrm{th}}$ and assymptotically replaced by,

$$S \to S(1-\mathfrak{w}) + \mathfrak{w}\, S_{\mathrm{bg}}, \tag{5.44}$$

where S_{th}, S_{bg}, and c_{th} are fitted individually for each detector for the μ and $e\gamma(\mu)$ components. This effect is only relevant for regions where $S \simeq 1\,\mathrm{VEM}$ or $1\,\mathrm{MIP}$.

As discussed in Section 5.3.1, values of ΔX_{max} and λ are dependent on r and their expected behaviour with respect to the radius needs to be considered. In previous works, ΔX_{max} and λ were fitted simply as polynomials in r and ΔX [109, 133] This, however, can introduce a divergent behaviour of the resulting signal model, if the parameters are not strictly confined. Furthermore, there is no direct physical motivation for these parameters to be quadratically or cubically dependent on radius. If λ is obtained from a best fit to the *total* longitudinal profile of the shower instead of individual components, there is an implicit dependence on θ and ΔX from different particle components dominating the shower

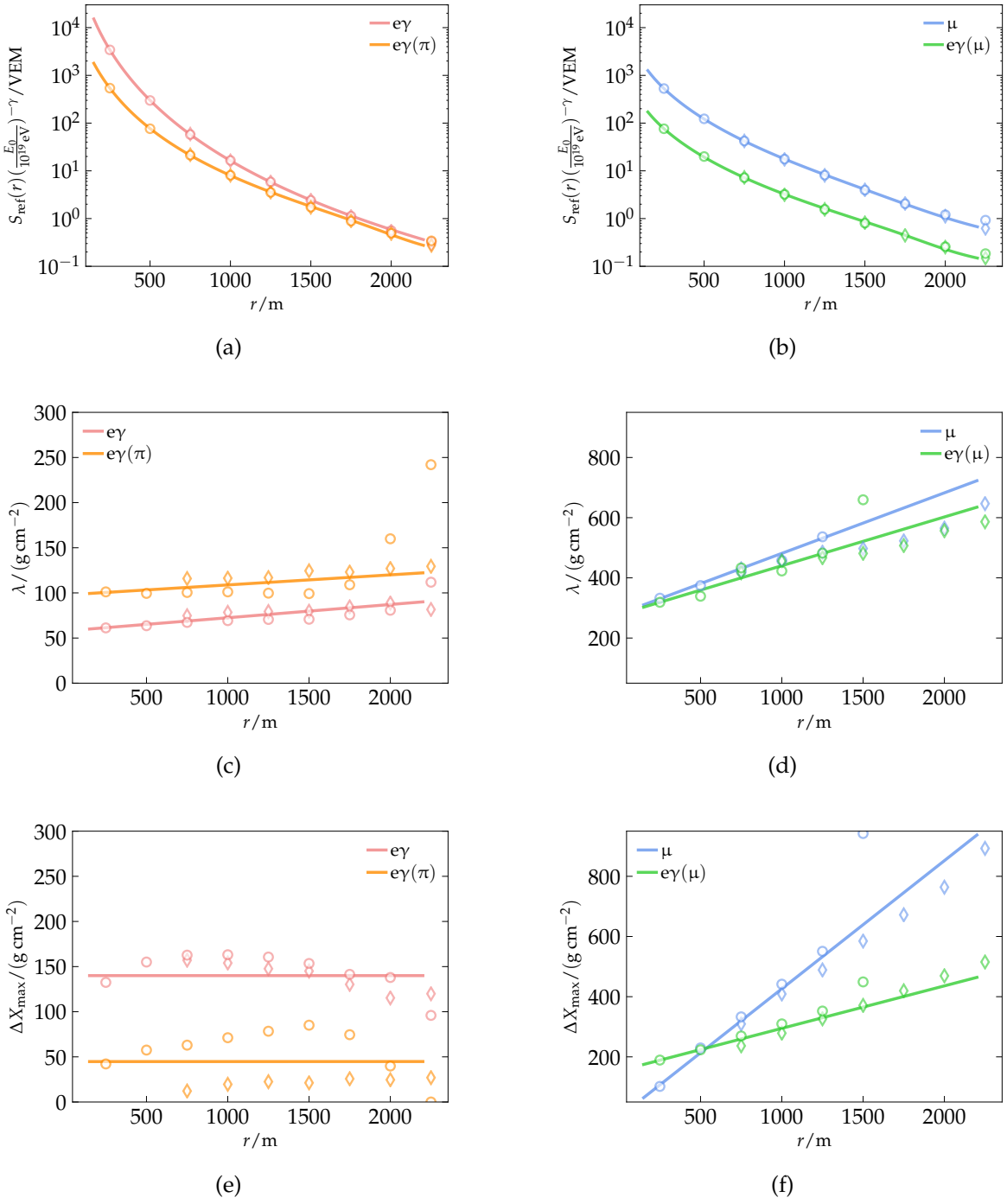


Figure 5.14: Progression of the best fit values of Eq. (5.10) fitted to WCD responses as a function of the radius for the individual components. Results obtained using proton showers with a primary energy of $E_0 = 10^{19}$ eV are depicted as circles and for $E_0 = 10^{20}$ eV as diamonds.

at different depths, i.e. muons taking over at the ground for very inclined showers. This is, however, not the case if λ is fixed individually for each particle component. Additionally, even though λ in general is considered to be energy-dependent, no significant discrepancies for the best fit values of λ that could not be explained by a non-optimal fit result were obtained with respect to showers of 10^{19} eV and 10^{20} eV. Consequently, λ is considered a

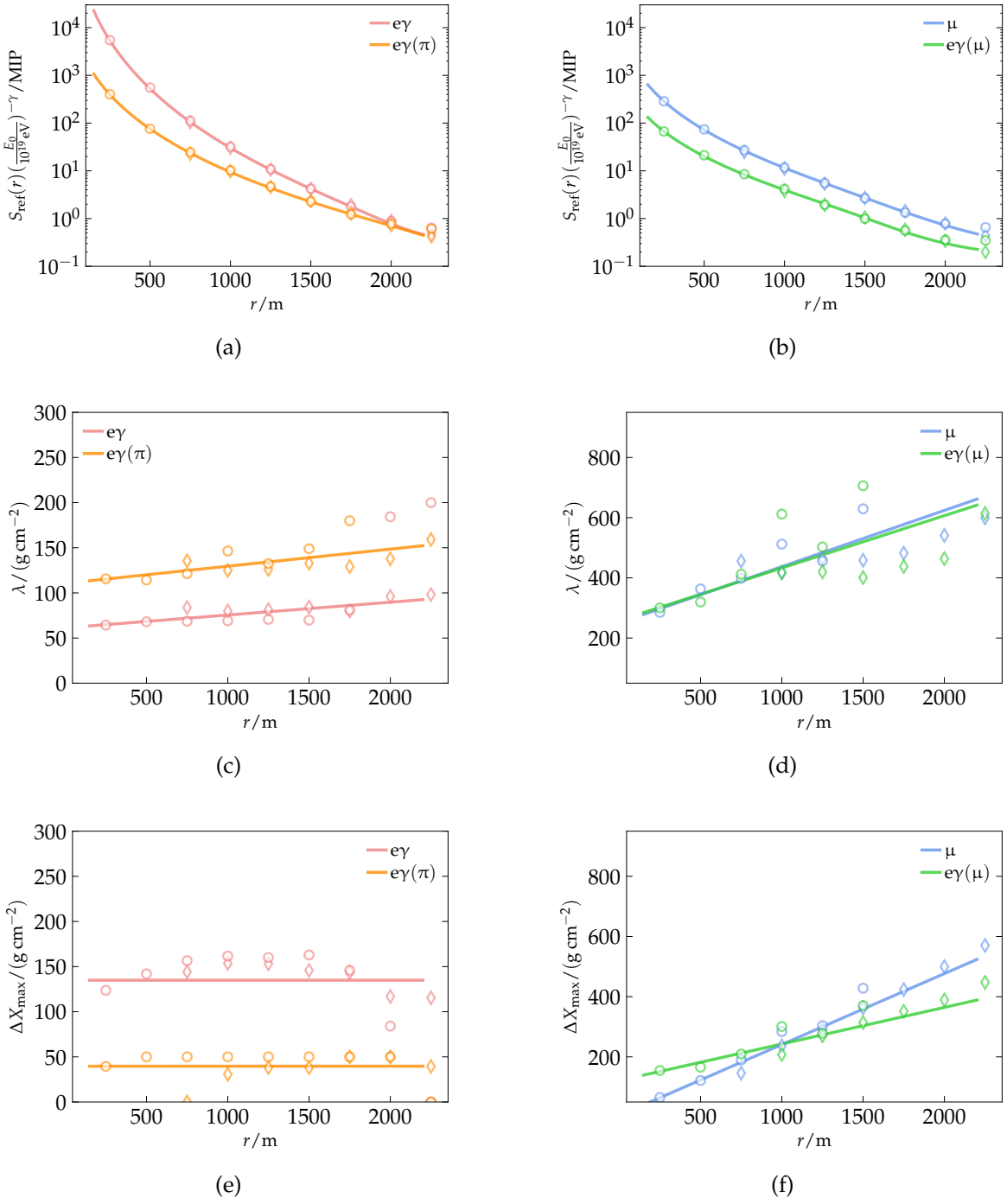


Figure 5.15: Progression of the best fitted values of Eq. (5.10) fit to SSD responses as a function of the radius for the individual components. Results obtained using proton showers with a primary energy of $E_0 = 10^{19}\,\mathrm{eV}$ are depicted as circles and for $E_0 = 10^{20}\,\mathrm{eV}$ as diamonds.

constant with respect to the energy.

Assuming the retardation of the shower maximum is due to particles being radiated from the shower core under an angle α that is constant around $\Delta X \simeq 0$, then ΔX_{max} is linear in radius. Assuming further that at large radii the longitudinal structure of the shower is smeared out, λ is also expected to be monotonically growing with radius. This smearing of

the longitudinal profile is also well described by a linear dependence on r.

The lateral distribution of particles, as resulting from the best fit values for $S_{\mathrm{ref}}(r)$ from Section 5.7.2, is fitted to Eq. (5.42), together with the threshold effect defined in Eq. (5.44). Results from various energies are scaled by a factor $(E_0/(10^{19}\,\mathrm{eV}))^{-\gamma}$ to attain a comparable lateral distribution for all energies. The best fit values of ΔX_{max} and λ are fitted to a monotonically increasing linear function in r. In this way, a model of the areal signal density is obtained for each particle component. The results from the best fits for WCDs are depicted in Fig. 5.14, and for SSDs in Fig. 5.15.

Results from showers of different energies, depicted as circles for 10^{19} eV and diamonds for 10^{20} eV, yield the same behavior, as expected if the principle of universality holds. The detector threshold effect of Eq. (5.44) can be seen, for example, in the slope of the lateral distribution of the μ and $e\gamma(\mu)$ components at $r \simeq 2000\,\mathrm{m}$ in Fig. 5.15 (b). Note that the lateral distributions of the $e\gamma(\pi)$ and $e\gamma(\mu)$ components are significantly flatter than for the $e\gamma$ component in both detectors, thus confirming the necessity to disentangle the signal of the four components for showers with different values of R_μ. The best-fit-results for λ from the $e\gamma$ component agree well with the expectation from Eqs. (5.7) and (5.8), and results for the $e\gamma(\pi)$ component match the attenuation length of light mesons in the atmosphere of approximately $110\,\mathrm{g\,cm^{-2}}$ [2]. Best-fit results for λ for the μ and $e\gamma(\mu)$ component are significantly larger than for the other components, as it is expected from the stretched-out longitudinal profiles seen in Fig. 5.12 (b), (d) and Fig. 5.13 (b), (d), and from the large attenuation length of muons. The retardation of the shower maximum by ΔX_{max} is approximately constant for the $e\gamma$ and $e\gamma(\pi)$ components, but is increasing linearly with radius for the μ and $e\gamma(\mu)$ components. This behaviour is a direct result of the particles in the μ (and consequently also the $e\gamma(\mu)$) component propagating approximately rectilinearly under a certain angle with respect to the shower axis.

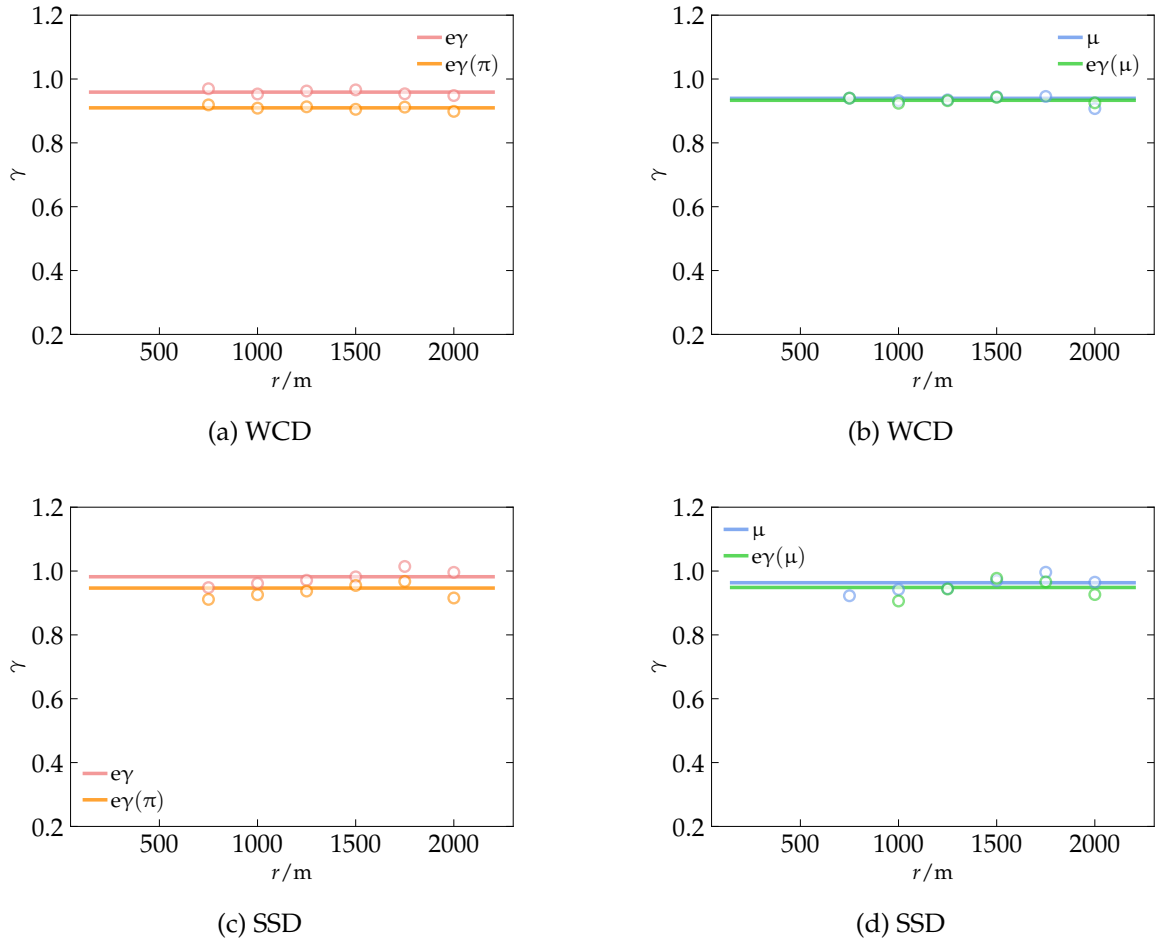
For distances to the shower axis larger than approximately 2250 m significant differences occur for S_{ref}, λ, and ΔX_{max} when obtained from showers with different primary energies. In this region a universal parametrization cannot be performed and the model experiences its limitations.

5.7.4 ENERGY DEPENDENCE OF THE SIGNAL

The dependence of the air-shower signal on the primary energy of the CR is condensed in the parameter γ of Eq. (5.42). In contrast to earlier work, γ is assumed not to be dependent on radius. The reason for this is assumption is twofold. First, if results of $S_{\mathrm{ref}}(r)$ are scaled radially dependent before a lateral distribution function is fitted to all data, then the lateral distribution is not universal with respect to the primary energy. In this case effects from low particle multiplicity might be misinterpreted in the data for lower and moderate primary energies. Second, there is no physical motivation for γ to depend on radius. γ is thus obtained from an average over all radii of the ratio of $S_{\mathrm{ref}}(r)$, according to

$$\gamma = \lg\left(\frac{S_{\mathrm{ref}}(r)}{S_{\mathrm{ref}}(r)|_{E_0=10^{19}\,\mathrm{eV}}}\right) / \lg\left(\frac{E_0}{10^{19}\,\mathrm{eV}}\right), \tag{5.45}$$

based on Eq. (5.3). Thus, values just below 1 are expected for γ for all components. Figuratively speaking, for $\gamma = 0$ the amount of signal of a component stays constant with respect to the primary energy, whereas for $\gamma = 1$ it would scale exactly as the primary energy. The resulting values for γ for the four components, as obtained from signals of simulated WCDs and SSDs, are depicted in Fig. 5.16, together with data obtained from showers simulated with $E_0 = 10^{20}$ eV and $E_0 = 10^{19}$ eV.



(a) WCD

(b) WCD

(c) SSD

(d) SSD

Figure 5.16: Results obtained from the ratios of $S_{\text{ref}}(r)$ for different primary energies as a function of radius according to Eq. (5.45) for the four signal components as deposited in the WCDs in (a) and (b) and in the SSDs in (c) and (d).

5.7.5 CORRECTIONS OF AZIMUTHAL ASYMMETRY

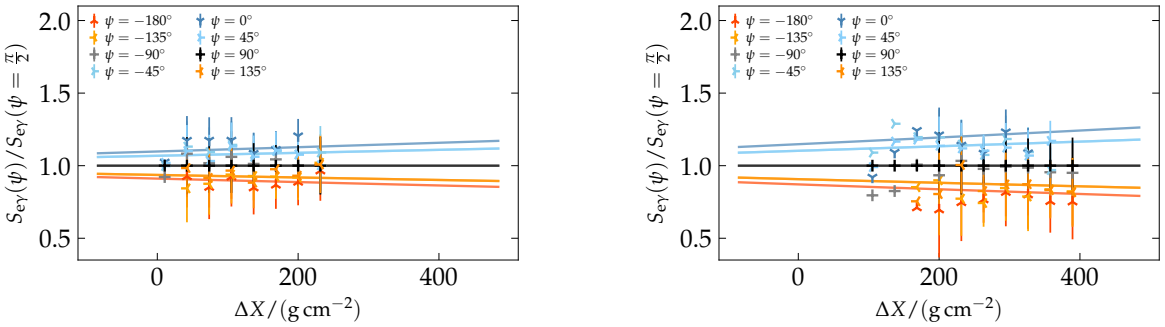
In the shower plane, asymmetries are expected in the signal of the different components for stations in sampling areas with different azimuth ψ. In the first order, this asymmetry is well described by the differences in ΔX and the resulting differences in signal according to Eq. (5.41). Effects arising from the differences in solid angle and length of attenuation of the particles, however, are not considered [140]. These can be covered by a small correction factor c_ψ given by

$$c_\psi = \exp\left[\zeta \, \frac{\Delta X - \Delta X_1}{\lambda} \frac{r}{1000\,\text{m}} \cos\psi \, \sin\theta\right], \tag{5.46}$$

where ζ is $\mathcal{O}(\frac{1}{100})$ for all components. The choice of using 1000 m as reference distance is of course arbitrary. An example of the resulting correction of the ratio of signals of different azimuth angles is given in Fig. 5.17. The asymmetry is strongest for the eγ component, whereas for the μ component no significant azimuthal asymmetry is observed.

5.7.6 THE MODEL OF THE TOTAL EXPECTED SIGNAL

Combining the expressions for the longitudinal and lateral development of the expected shower signal Eqs. (5.41) and (5.42), together with the corrections defined in Eq. (5.44) and



(a) $\theta = 22°$ (b) $\theta = 38°$

Figure 5.17: Asymmetry in azimuth for the signals of the eγ component from showers of $E_0 = 10^{19}\,\mathrm{eV}$ deposited in WCDs at $r = 750\,\mathrm{m}$ for (a) $\theta = 22°$ and (b) $\theta = 38°$. Upstream regions are colored in blue, downstream regions are colored in orange.

Eq. (5.46), the model of the expected signal in a surface detector is complete. The contributions of the individual particle components are scaled in the same manner as in Eq. (5.13). The signal of each component as from the Universality model thus reads as

$$\begin{aligned} S(\Delta X, R_\mu) = (a\,(R_\mu - 1) + 1) \left(\frac{E}{10^{19}\,\mathrm{eV}}\right)^{\gamma} \left(\frac{\Delta X - \Delta X_1}{\Delta X_{\mathrm{ref}} - \Delta X_1}\right)^{\frac{\Delta X_{\max} - \Delta X_1}{\lambda}} \mathrm{e}^{-\frac{\Delta X - \Delta X_{\mathrm{ref}}}{\lambda}} \\ \times c_\psi \frac{S_{\mathrm{ref}}^{19}}{2\pi r_{\mathrm{M}}^2} \frac{\Gamma(\frac{9}{2} - s)}{\Gamma(s)\Gamma(\frac{9}{2} - 2s)} \left(\frac{r}{r_{\mathrm{M}}}\right)^{s-2} \left(1 + \frac{r}{r_{\mathrm{M}}}\right)^{s-\frac{9}{2}} (1 - \mathfrak{w}) + \mathfrak{w} S_{\mathrm{bg}}, \end{aligned} \tag{5.47}$$

and the total signal is given by the sum over all signal components,

$$S_{\mathrm{tot}} = \sum_i S_i(\Delta X, r), \quad \text{for} \quad i \in \{\mathrm{e}\gamma, \mu, \mathrm{e}\gamma(\pi), \mathrm{e}\gamma(\mu)\}. \tag{5.48}$$

5.8 PARAMETRIZATION OF THE EXPECTED TRACE

To create an accurate model of the trace as expected in SD detectors, the model of arrival times described in Section 5.6 needs to be fitted and parametrized to simulated data. The distributions of arrival times of particles, however, are not exactly the same as the traces in a detector, since each particle deposits its signal not in an instance, but with a certain distribution of signal over time. The resulting trace is thus not only scaled, as we assume for the total signal, but is a convolution of the differential arrival times, $\mathrm{d}n/\mathrm{d}t$, and the individual particle response functions. The implications of this circumstance on t_{40} and the shape of the log-normal distributions to be fitted to the traces are not trivial.

Fig. 5.18 shows a model of the detector response to 150 particles, for which each particle j deposits signal according to

$$\frac{\mathrm{d}}{\mathrm{d}t} S_j(t) = \Theta(t - t_j)\,\mathrm{e}^{-\frac{t}{\tau}}\,\mathrm{VEM}, \tag{5.49}$$

after its arrival time t_j, where $\tau = 61\,\mathrm{ns}$, which is the decay time of light in a WCD [141], and Θ is the Heaviside-function. The arrival times of the particles were sampled from a log-normal distribution, depicted in Fig. 5.18 in gray. The histogram of deposited signal, however, is shifted and is better described by a log-normal distribution where the median of the distribution is also shifted by τ, as depicted in red. The shifted, red curve yields a

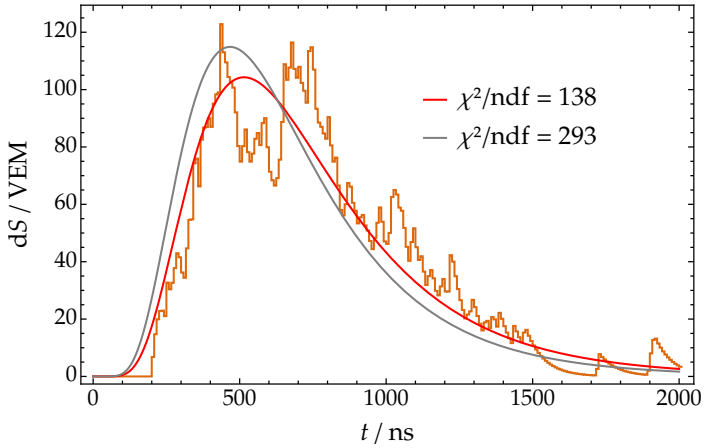


Figure 5.18: Toy model of the detector response to a set of 150 particles, whose arrival times were sampled from a log-normal distribution, depicted in gray. A better-fitting log-normal distribution with shifted median is depicted in red.

significantly better match to the model data in terms of the least-squares deviation. As a result, t_{40} in data is going to be significantly larger than predicted by Eq. (5.35). Furthermore, the perfectly spherical development of the shower front that was assumed in Section 5.6.1 is in general not true. The real shower front is slightly hyperbolically curved and thus a radially-dependent correction is required. Lastly, poor sampling statistics, i.e. when fewer than $\mathcal{O}(100)$ particles arrive at a single detector, will inevitably shift and stretch the best fitting log-normal-like function; this is especially important when considering showers from CRs with different primary energies. As a minimum, 5 VEM or 5 MIP of signal are thus imposed for detectors to be considered in the parametrization of the traces.

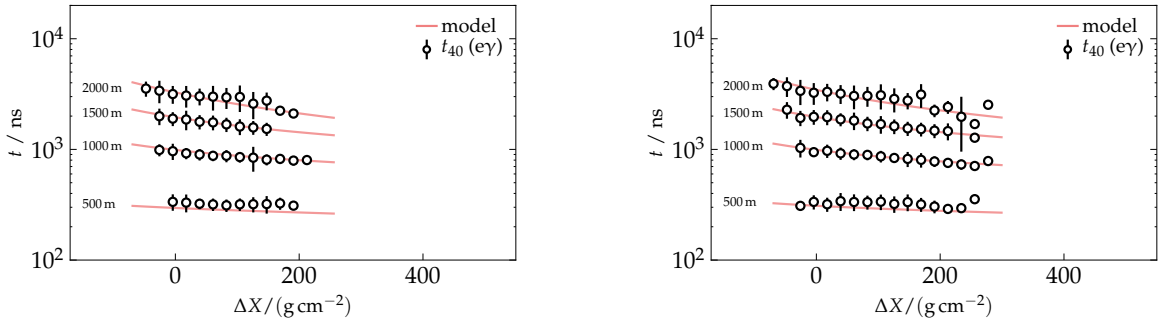
Furthermore, as discussed in Section 5.6, without adjustment the model is not valid in regions of $\Delta X < 0$.

5.8.1 PARAMETRIZATION OF t_{40}

To counter all of the aforementioned effects when parametrizing t_{40} an additive correction t_{hyp} to match the slightly hyperbolic shower front as well as a shift in the ordinate $\Delta X \to \Delta X + \delta X$ is used when fitting to simulated data. Therefore, t_{40} as obtained from data is fit to the function

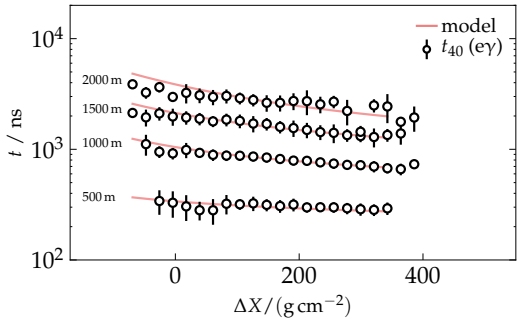
$$ct_{40} = ct_{\mathrm{hyp}} + \sqrt{\left(\frac{h_{\mathrm{s}}}{\cos\theta}\ln\left(\frac{\Delta X + \delta X}{X_{\mathrm{max}}} + 1\right)\right)^2 + r^2} - \frac{h_{\mathrm{s}}}{\cos\theta}\ln\left(\frac{\Delta X + \delta X}{X_{\mathrm{max}}} + 1\right), \quad (5.50)$$

where time is again measured relatively to the time of the plane front of the shower arriving at the respective station at distance r and depth ΔX. As t_{40} (and any other time quantile) according to Eq. (5.50) is expected to explicitly depend on the radius and zenith angle, the data is fitted individually for each particle component, radius and zenith angle. The individual best-fit results for δX and t_{hyp} are then for each component fitted to polynomials in r and $\sin\theta$. Examples for the resulting models for δX and t_{hyp} are depicted for the eγ and μ components in Fig. 5.20. As expected from the different geometric evolution of the components, δX is rather flat and t_{hyp} is strongly increasing with radius for the eγ component, while the exact opposite is the case for the μ component. In this way, a subtle but present change in t_{40} with respect to zenith angle, that was also empirically found and addressed in Ref. [128], is taken care of. The effective speed of the propagation of the particles is fixed to $c = 0.28\,\mathrm{m/ns}$, that is 93% of the speed of light in vacuum. This value was obtained by

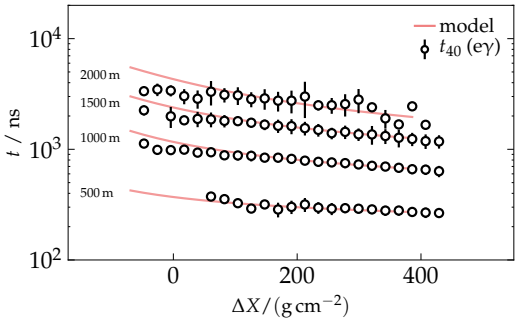


(a) $\theta = 12°$

(b) $\theta = 22°$



(c) $\theta = 32°$



(d) $\theta = 38°$

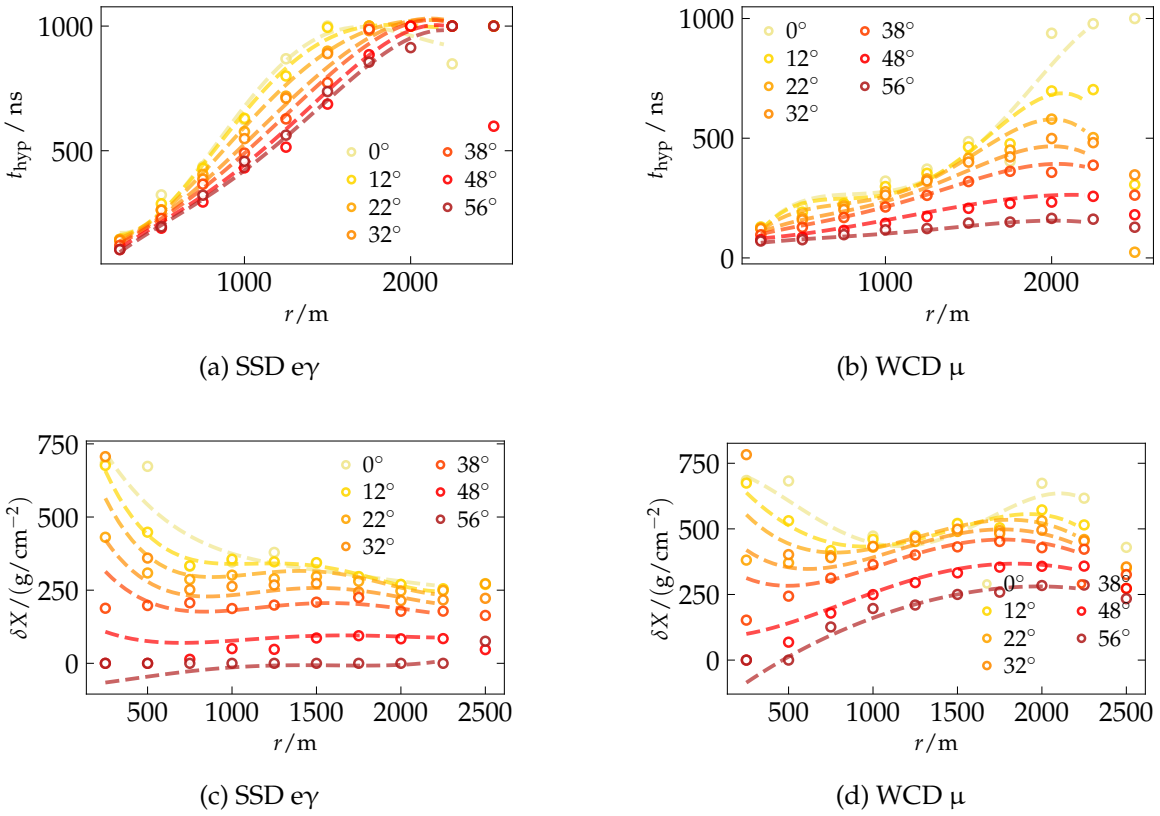
Figure 5.19: Values of t_{40} for the eγ component as a function of ΔX obtained from WCD responses to showers of $E_0 = 10^{19}$ eV and 10^{20} eV and various zenith angles, alongside with the model resulting from the parametrization.

a global best fit but is in general correlated to δX. In this process also other time quantiles have been parametrized the same way, namely t_1, t_5, t_{10}, t_{30}, t_{50}, and t_{70}, each referring to the time after which the indexed number in percent of the total signal has been deposited in a station (t_1 is essentially the start time of the station). In this way, the test of the hypothesis that t_{40} is significantly dependent on X_{max} was validated. These additional parametrizations might be used in a future work.

The values for t_{40} from the traces of the eγ component with respect to the time of the plane front as a function of ΔX are depicted in Fig. 5.19, alongside with the resulting model. The corresponding figures for the other particle components have been moved to the Appendix as Figs. B.8 to B.10, together with the result of the same analysis performed on the total signal, i.e. the sum of the four signal components, given in Fig. B.11. The model only matches the data poorly for inclined showers and beyond a distance of 2000 m from the shower axis. The behavior of t_{40} is qualitatively the same for both the WCDs and SSDs, however, the data show a slightly larger uncertainty when obtained from the SSDs.

5.8.2 PARAMETRIZATION OF σ

The shape parameter σ of the log-normal distribution from Eq. (5.30) can be parametrized either directly, or indirectly in terms of the mean and the standard deviation of the arrival times of signal in a detector, $\mathrm{E}[t]$ and $\mathrm{std}[t]$, respectively. In earlier work, parametrizing σ directly using the mean values obtained for several radii or zenith angles has caused problems because of the apparently chaotic behaviour of σ with respect to ΔX. When expressed in terms of the moments of the log-normal distribution, however, the situation can be im-



(a) SSD eγ

(b) WCD μ

(c) SSD eγ

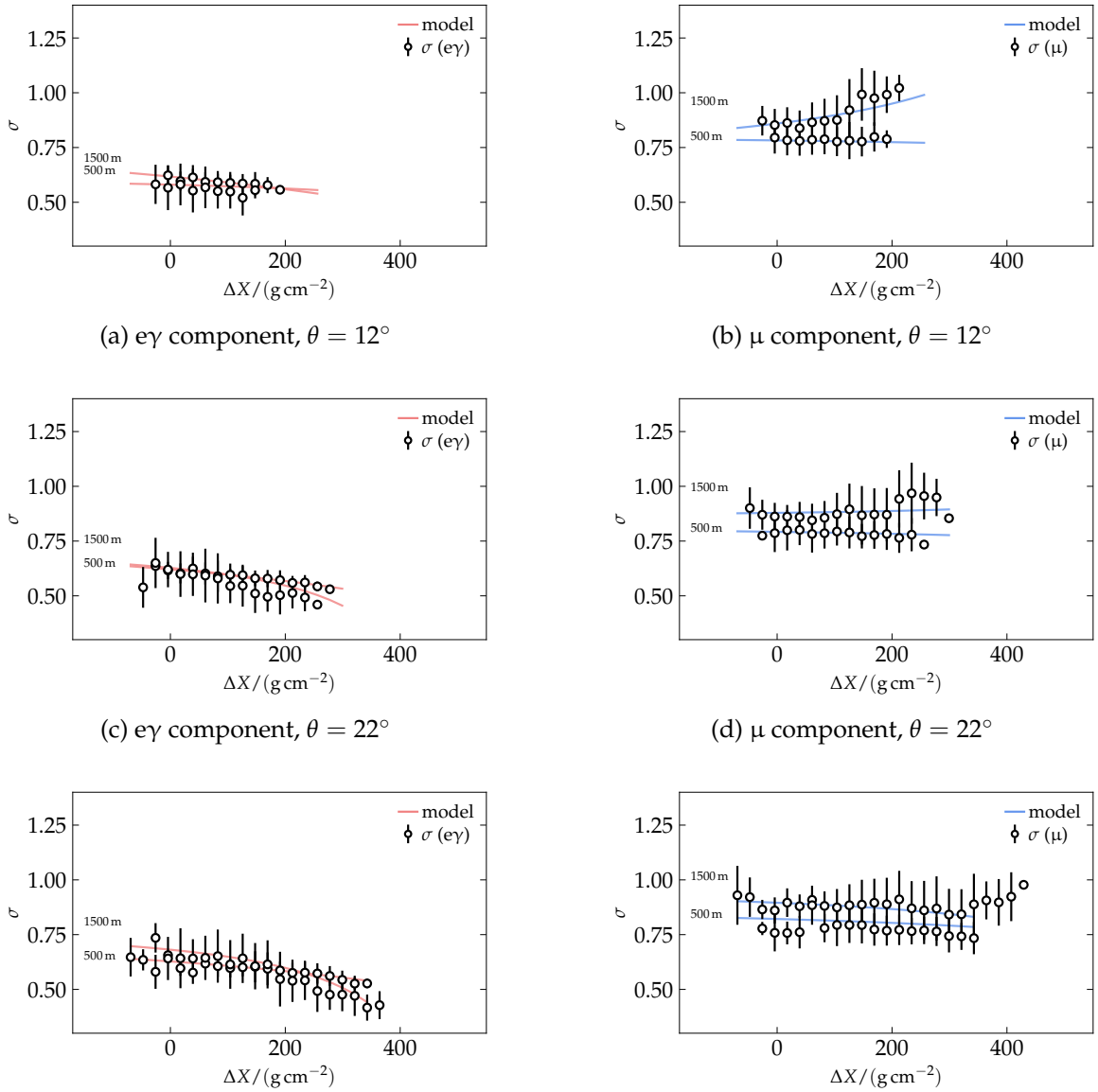
(d) WCD μ

Figure 5.20: Examples for the modelled correction paramters t_{hyp} and δX employed in Eq. (5.50) to match data, showing the results obtained for the eγ component in the SSDs (*left*), and the μ component in the WCDs (*right*). Best-fit values for different radii and zenith angles are depicted as dots, whilst their parametrization is depicted as dashed lines.

proved. The shape parameter is related to the moments of the log-normal distribution by the expression

$$\sigma^2 = \ln\left(1 + \frac{\mathrm{std}[t]}{\mathrm{E}[t]}\right). \tag{5.51}$$

The average of the moments of the log-normal distributions fitted to data, $\mathrm{E}[t]$ and $\mathrm{std}[t]$, can be accurately modelled as linear functions of ΔX. In this case $\mathrm{std}[t]$ is given in absolute time and $\mathrm{E}[t]$ is measured relative to the start time of the signal in a respective detector, t_0, as indicated in Eq. (5.30). σ as obtained from the responses of simulated WCDs to showers of $E_0 = 10^{19}$ eV and 10^{20} eV as well as the resulting model of σ for the eγ and μ components is given in Fig. 5.21. When directly comparing the behaviour of σ from the two components for showers with $\theta = 12°$, see Fig. 5.21 (a), (b), for example, clearly a counter-intuitive behaviour can be observed, especially when considering different radii. Moreover, the behaviour of σ is systematically different for showers from different zenith angles. Still the model is able to accurately reproduce the expected values of σ within the uncertainty of the data. The underlying moments of the log-normal distribution fitted to the responses of WCDs to showers of two representative zenith angles as a function of ΔX are given in Fig. 5.22, again for the eγ and μ components. The respective figure for the SSDs is given in Fig. B.12. Individual data points are depicted explicitly as colored markers to illustrate the spread of the data. Detectors at $r = 1000$ m are picked solely for illustrative purposes, the behaviour of the data is qualitatively the same for all radii. Note that the behaviour of the



(a) eγ component, $\theta = 12°$

(b) μ component, $\theta = 12°$

(c) eγ component, $\theta = 22°$

(d) μ component, $\theta = 22°$

(e) eγ component, $\theta = 32°$

(f) μ component, $\theta = 32°$

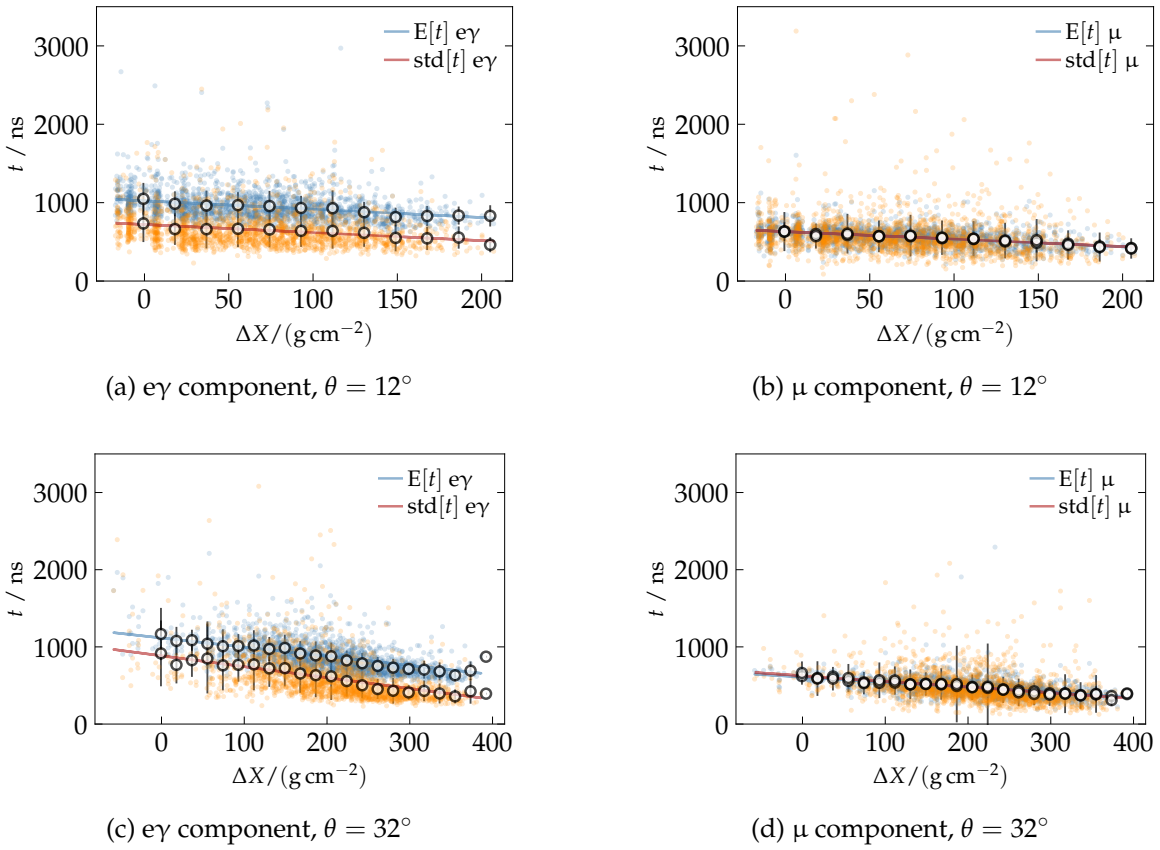
Figure 5.21: Behaviour of the log-normal shape parameter σ for WCD traces with respect to ΔX for the eγ and μ. Three different zenith angles are shown. In each panel the behaviour of σ is shown at two different example radii.

moments of the fitted log-normal distributions depicted in Fig. 5.22 (a), (b) directly relates to the behaviour of the model for σ depicted in Fig. 5.21 (a), and (b).

If not expressed as a function of t_{40}, μ in Eq. (5.37) could also be calculated from the average first and second moment of the traces. In this case, however, the information carried by the absolute start time of the traces is lost.

5.8.3 THE MODEL OF THE EXPECTED TRACE

With the values of both t_{40} and σ parametrized as a function of ΔX and the geometry, the expected trace in a station with the signal start time t_0 is given for each particle component



(a) eγ component, $\theta = 12°$

(b) μ component, $\theta = 12°$

(c) eγ component, $\theta = 32°$

(d) μ component, $\theta = 32°$

Figure 5.22: Behaviour of the moments of the log-normal distribution fitted to the responses of WCDs located at $r = 1000\,\mathrm{m}$ as a function of ΔX. Colored dots depict individual data points, the binned mean is depicted by white markers. The corresponding model is given as line.

as

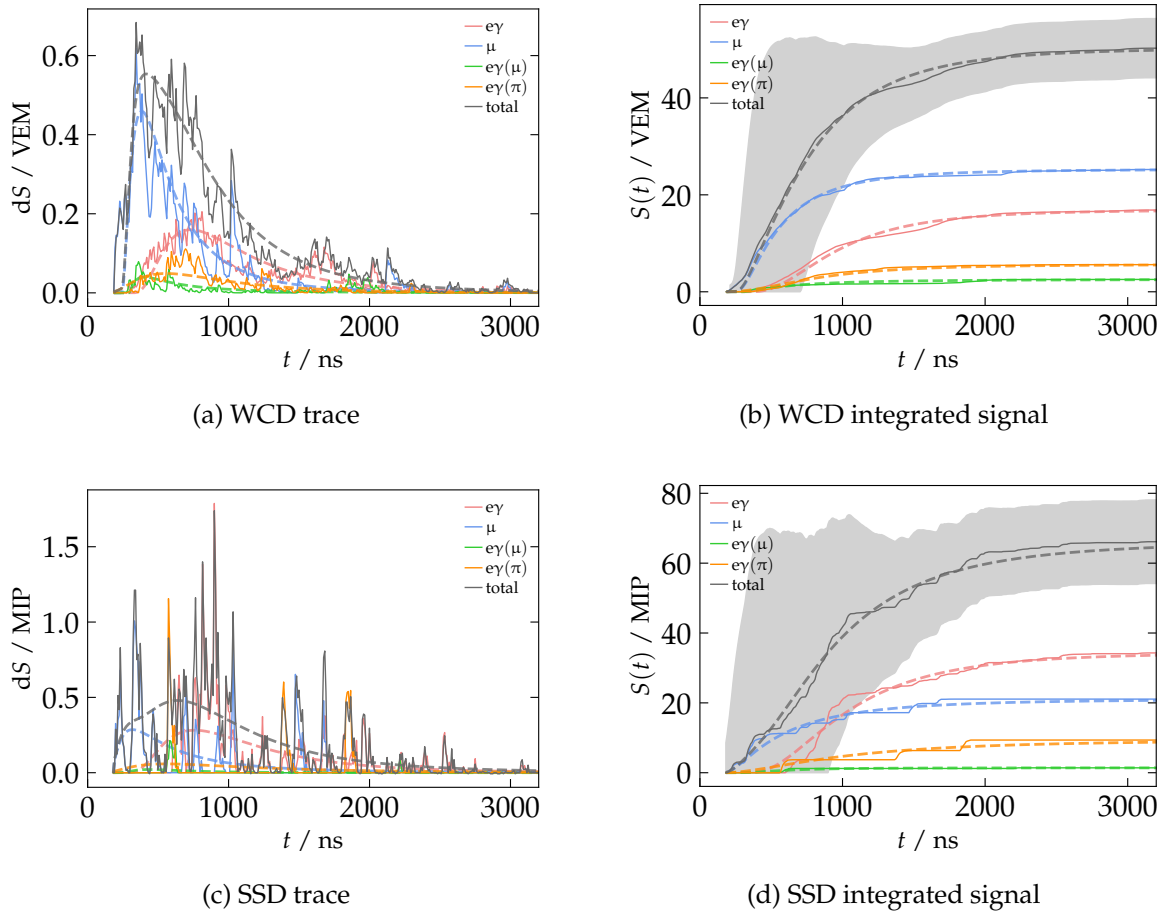
$$\frac{\mathrm{d}}{\mathrm{d}t}S_i(t) = \frac{S_i}{\sqrt{2\pi}\,\sigma\,(t-t_0)}\,\exp\left[-\frac{1}{2\sigma^2}\left(\ln\left(\frac{t-t_0}{t_{40}-t_0}\right)+\sqrt{2}\sigma\,\mathrm{erf}^{-1}(2\times 0.4-1)\right)^2\right], \tag{5.52}$$

with the signal S_i of each component parametrized as in Eq. (5.47). The total signal is given by the sum of the individual time-dependent signal components.

An example of the parametrized time-dependent signal for a WCD and an SSD is given in Fig. 5.23. The time-dependent signal is given in differential form in Fig. 5.23 (a) and (c), and in integrated form in (b) and (d). The total signal of each component is normalized to its simulated value. $S_i(t)$ is depicted as obtained from Eq. (5.52) for each component, with X_{max} set to the MC value of $782\,\mathrm{g\,cm^{-2}}$. The uncertainty band depicted in the panels (b) and (d) is addressed in Section 6.1.2. The simulated shower for which the example signal is depicted was induced by a proton primary with $\lg(E_0/\mathrm{eV}) = 19$ and $\theta = 32°$ using the EPOS-LHC model. The parametrized and simulated trace shape parameters and signal sizes for this particular example are given in Table 5.1.

5.9 VALIDATION

To validate that the parametrized model accurately reproduces the signal one expects from data, the available library of air-shower signals discussed Section 5.7.1 was reproduced by



(a) WCD trace

(b) WCD integrated signal

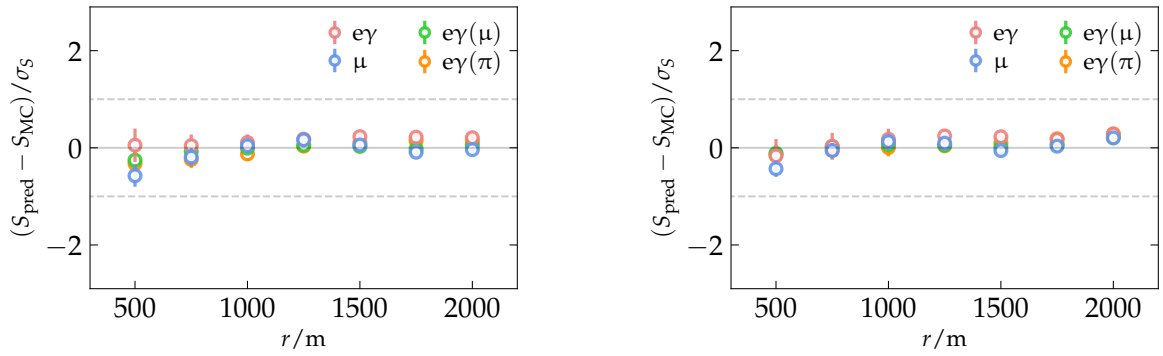
(c) SSD trace

(d) SSD integrated signal

Figure 5.23: Example of the time dependent signal deposited in a WCD and an SSD at $r = 1000\,\mathrm{m}$, from a simulated proton event with $\lg(E_0/\mathrm{eV}) = 19$ and $X_{\max} = 783\,\mathrm{g\,cm^{-2}}$. The start time with respect to the plane-front is at $t_0 = 188\,\mathrm{ns}$.

inserting the respective MC values for the geometry, the energy, $X_{\max}$, and R_μ in the final parametrization. The set of signals were then compared to the respective simulated signals in the shower library. Fig. 5.24 depicts the average residuals of the predicted signal S_{pred} with respect to the MC signal S_{MC}, normalized by the signal uncertainty σ_S of the respective detector and signal size (see Section 6.1.1), given for all showers with a primary energy of $E_0 = 10^{19}\,\mathrm{eV}$. The models used for the signal uncertainty are introduced in Refs. [142] and [143]. The parametrized model of air-shower signals yields a satisfying agreement with the data over all radii. Significant systematic deviations from data are only present in the model of the signal of the μ component for the WCD responses for very small radii, and in the overall model of the signal for all components at very distant stations for the SSD data. While the first can be explained by the divergent nature of the lateral distribution of particles for small radii, the latter are most probably due to the detector trigger effects that were discussed earlier in Section 5.7.2. In the same manner, the models parametrized by showers generated with the QGSJET-II and SIBYLL models have been examined. The respective figures are given in the Appendix in Fig. B.13.

To validate the model of the expected trace shapes, in a similar manner the predicted values of the observables were compared to the simulated data. The residuals of the predicted values for t_{40} and σ with respect to data are depicted in Figs. 5.25 and 5.26. Error bars in both figures show the spread of the data as obtained from the simulated detector responses. Showers of all energies available in the simulation library were considered. Sets



(a) WCD

(b) SSD

Figure 5.24: Validation of the model of the predicted signal S_{pred} in (a) WCDs and (b) SSDs, parametrized with EPOS-LHC showers. The z-score of the predicted signal is given using the signal uncertainty σ_S from the uncertainty model of the respective detector responses. A deviation of the expected signal by ± 1 standard deviation is indicated as gray dashed line.

of residuals for different zenith angles are explicitly plotted on top of each other. Overall the model is fit to reproduce the values of t_{40} and σ on average, considering the spread of the data. Only for the $e\gamma(\mu)$ component, the chaotic behaviour of σ in data aggravates the prediction of the respective values. The overall trace shape, however, only depends little on small deviations of σ and is mostly driven by t_{40}. The same result is obtained for the SSDs.

Concluding, it is thus ascertained that the model of the spatial and temporal distribution of particle densities and signal accurately describes air showers.

5.10 EXPECTED SENSITIVITY OF THE MODEL

To estimate the predictive power of the model with respect to the observables X_{max} and R_μ, the sensitivity ξ is examined from the model. ξ is defined as the absolute of the derivative of the modelled total signal, S, or modelled values of t_{40}, with respect to X_{max} and R_μ, normalized to the standard deviations of S and t_{40} obtained from data, σ_S and $\sigma_{t_{40}}$. σ_S and $\sigma_{t_{40}}$ comprise the shower-to-shower fluctuations as well as the uncertainty of the detector responses but not the uncertainties arising from a non-ideal reconstruction of the energy and event geometry. Considering only absolute values, the derivative with respect to X_{max} is identical as with respect to ΔX, such that

$$\xi_{X_{max}}(S) = \frac{1}{\sigma_S} \left| \frac{\partial S}{\partial \Delta X} \right| . \tag{5.53}$$

Given that all signal components except for the $e\gamma$ component scale approximately identically to the μ component, the total signal can be written as

$$S \simeq R_\mu (S_\mu + S_{e\gamma(\mu)} + S_{e\gamma(\pi)}) + S_{e\gamma} \tag{5.54}$$

and thus the sensitivity of the total signal with respect to the R_μ is given by

$$\xi_{R_\mu}(S) = \frac{1}{\sigma_S} \frac{\partial S}{\partial R_\mu} \simeq \frac{1}{\sigma_S} (S_\mu + S_{e\gamma(\mu)} + S_{e\gamma(\pi)}). \tag{5.55}$$

The estimated sensitivity of the signal with respect to X_{max} and R_μ is numerically evaluated; the results are depicted in Fig. 5.27. It must be emphasized that $\xi_{X_{max}}$ is not constant over all

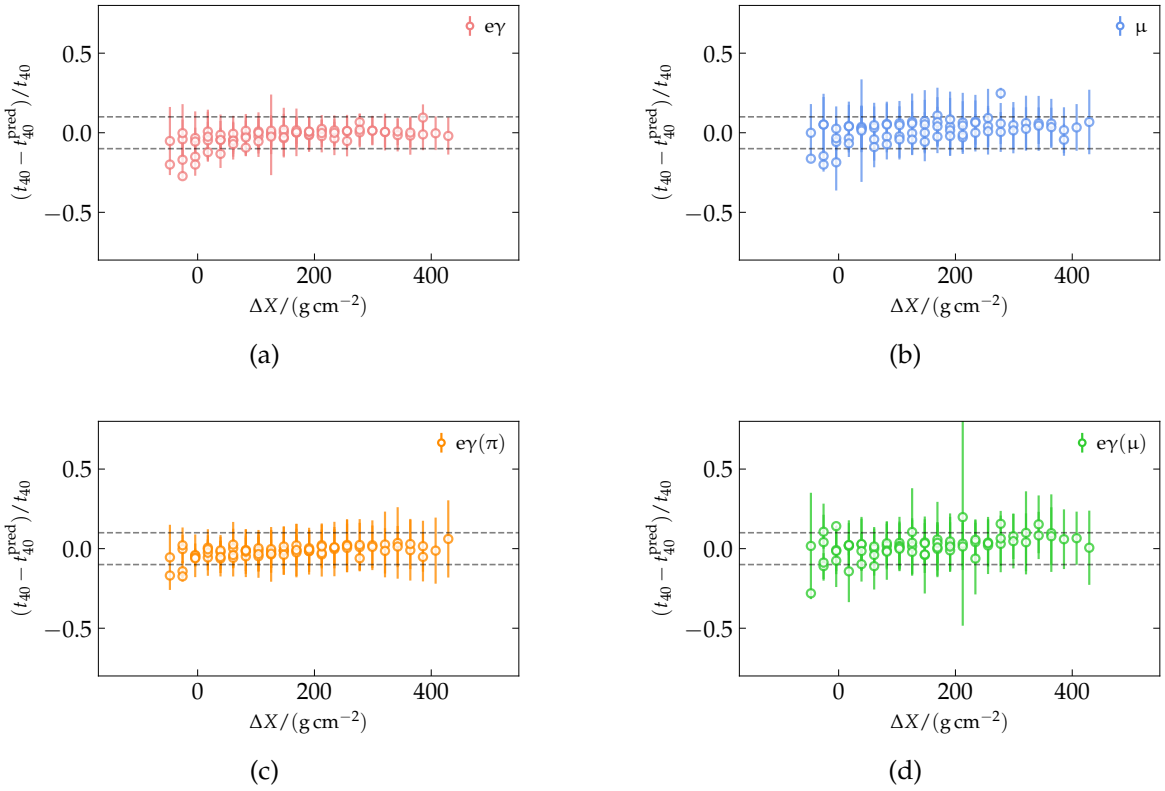


Figure 5.25: Validation of the model of the expected time quantile t_{40} for responses of the WCDs for the four components. Residuals are averaged over all radii and plotted individually for $\theta = 12°$ to $\theta = 38°$. $\pm 10\%$ deviation is marked as gray dashed line.

values of X_{max} and that this analysis did not take into account the corresponding weighting of the values expected for ΔX. Furthermore, since σ_S does not include the uncertainty that arises from a non-ideally reconstructed geometry, the sensitivity for small radii on X_{max} and R_μ is likely overestimated [144]. Still two important conclusions can be drawn from ξ as depicted in Fig. 5.27. Firstly, as clearly visible in Fig. 5.27 (a) and (b), the responses of the WCDs are more sensitive to a change in R_μ and the corresponding signal components than the responses of SSDs. This is not surprising, since muons do on average deposit more signal in the WCDs than in the SSDs, and this is thus an indication that the design of the detector works as intended. Furthermore, $\xi_{R_\mu}(S)$ is rather flat as a function of the distance to the shower axis and no significant dependence of the sensitivity on zenith can be obtained from the model. Even though the μ and corresponding components dominate the signal at a large distances because of their relatively flat lateral distribution, the signal uncertainty becomes significant and counters this effect. Taking into account the systematic uncertainties of the reconstructed geometry that are significant especially for the SSDs, an optimal distance of sensitivity might emerge in the region of $500\,\mathrm{m} < r < 1500\,\mathrm{m}$. Secondly, the sensitivity of the total signal with respect to X_{max}, $\xi_{X_{\mathrm{max}}}(S)$, is significantly smaller than $\xi_{R_\mu}(S)$ for all radii. The reasons for this is that the variation of the signal as a function of X_{max} is mainly governed by the modified Gaisser-Hillas function, see Eq. (5.41), that in the region of ΔX where SD stations are expected does not vary over several orders of magnitude, as the lateral distribution of signal does, see Eq. (5.42). Additionally, given the small absolute variation of $\xi_{X_{\mathrm{max}}}(S)$ over the distance to the shower axis, possible best fitting values of X_{max} and R_μ are strongly correlated. The depth of the shower maximum X_{max} thus cannot be reasonably estimated from the lateral distribution of the signal if R_μ is

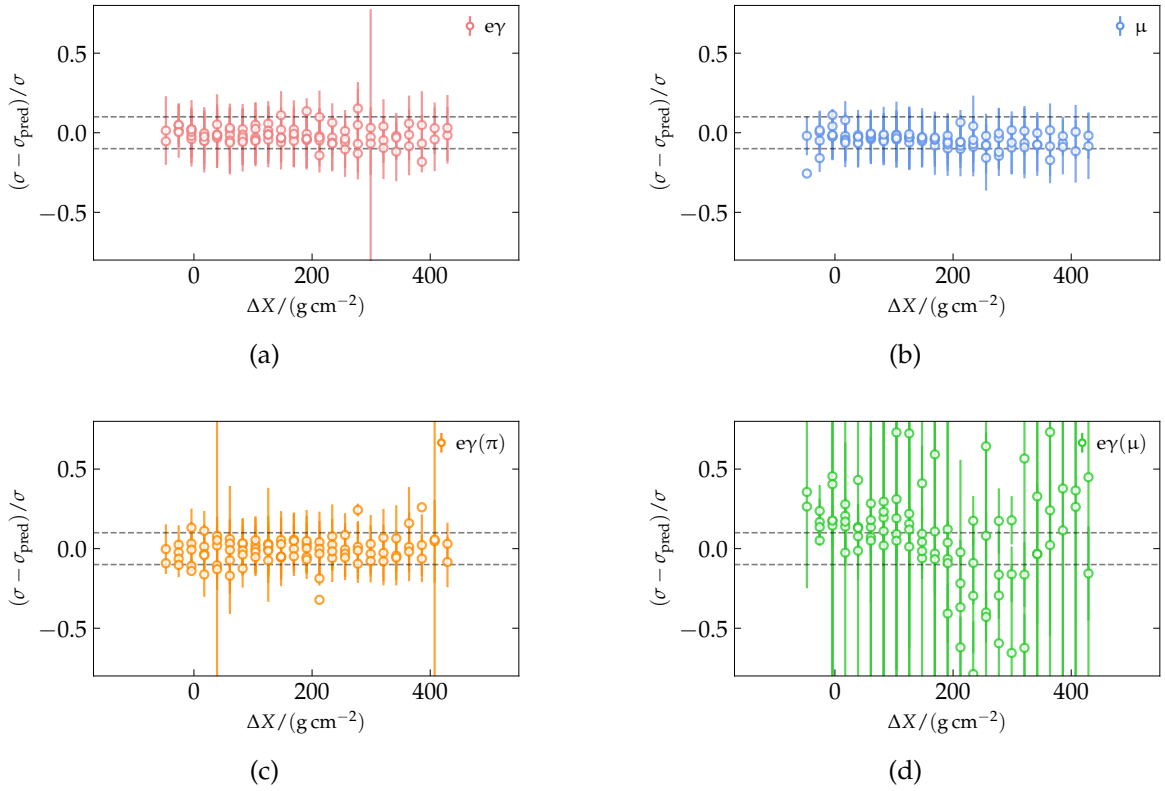


Figure 5.26: Validation of the model of the predicted trace shape parameter σ_{pred} for responses of the WCDs for the four components. Residuals with respect to simulated values are averaged over all radii and plotted individually for $\theta = 12°$ to $\theta = 38°$. $\pm 10\%$ deviation is marked as gray dashed line.

not known. If R_μ is known, values for X_{max} reconstructed from the lateral distribution of the total signal alone would still suffer uncomfortable uncertainties.

Given an accurate reconstruction of the energy and geometry of an air-shower event, the precision with which X_{max} and R_μ can be reconstructed from the total signal of an SD station can be estimated as $\sigma^{\mathrm{approx}} \simeq \xi^{-1}$. A conservative estimate for a station ar $r \simeq 1000\,\mathrm{m}$ is thus

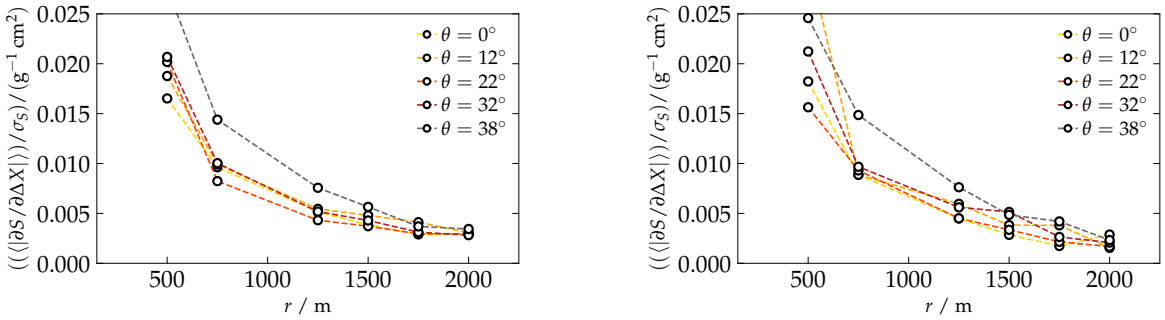
$$\sigma^{\mathrm{approx}}_{R_\mu}(S) \simeq 0.15, \quad \text{and} \quad \sigma^{\mathrm{approx}}_{X_{\mathrm{max}}}(S) \simeq 120\,\mathrm{g\,cm^{-2}}, \tag{5.56}$$

where the latter requires an accurate estimation of R_μ and the first is almost irrespective of the estimation of X_{max}. To put this differently, a variation of X_{max} of $100\,\mathrm{g\,cm^{-2}}$, which is approximately the average separation for iron and proton showers, is expected to have less impact on the total signal than an increase of R_μ of 15%, which is about half of the average separation of iron and protons showers in this observable.

In the same way the sensitivity of t_{40} was numerically evaluated for its dependence on X_{max},

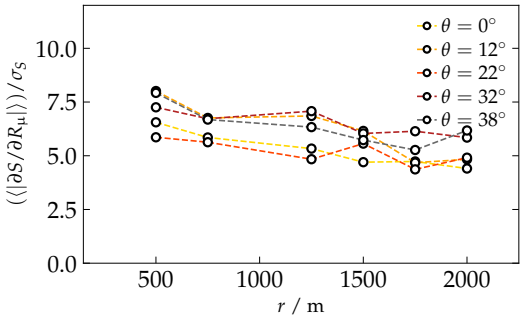
$$\xi_{X_{\mathrm{max}}}(t_{40}) = -\frac{1}{\sigma_{t_{40}}} \frac{\partial t_{40}}{\partial \Delta X}. \tag{5.57}$$

The results are depicted in Fig. 5.28. As already expected from Fig. 5.19, for stations close to the shower axis, t_{40} is almost independent of ΔX and thus X_{max}, the sensitivity is therefore rather low. An optimum emerges for stations around a distance between 1250 m to 1500 m from the shower axis. In this region, given an accurate reconstruction of the energy and

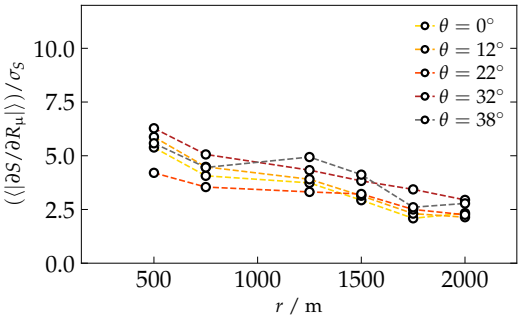


(a) WCD $\xi_{X_{\mathrm{max}}}(S)$

(b) SSD $\xi_{X_{\mathrm{max}}}(S)$



(c) WCD $\xi_{R_\mu}(S)$



(d) SSD $\xi_{R_\mu}(S)$

Figure 5.27: Estimated sensitivity of the total parametrized signal of single WCDs (*left*) and SSDs (*right*) with respect to X_{max} (*top row*) and R_μ (*bottom row*), given a perfect reconstruction of the energy and geometry of an event.

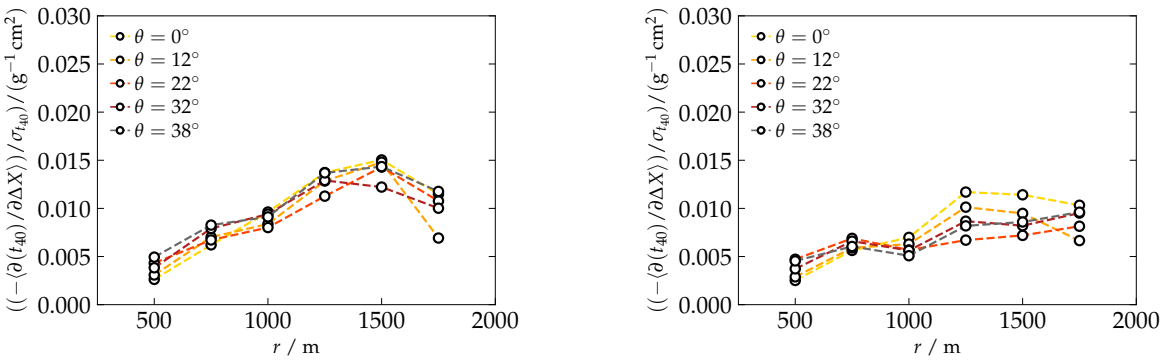
geometry of the event, the possible precision from data collected by a single detector can be conservatively estimated to

$$\sigma^{\mathrm{approx}}_{X_{\mathrm{max}}}(t_{40}) \simeq 70\,\mathrm{g\,cm}^{-2}. \tag{5.58}$$

The same result was obtained in Ref. [126] using the model of the time-dependent signal parametrized in this work in a χ^2 fit to simulated data of single WCDs. In this analysis the sensitivity was averaged over a range of ΔX as well. It is neglected that on average iron showers, for example, will result in larger values of ΔX which yield less sensitivity on X_{max}.

Surprisingly, within a set of simulated data from a single type of primary, e.g. proton showers, t_{40} is not significantly dependent on R_μ. Only for a data set that contains multiple types of primary particles, where X_{max} also systematically changes, t_{40} as parametrized in this work is implicitly dependent on R_μ.

Given its large uncertainty in the data, the trace shape parameter σ alone as modelled in Section 5.8.2 and obtained from data holds no predictive power with respect to either X_{max} or R_μ.



(a) WCD $\xi_{X_{max}}(t_{40})$

(b) SSD $\xi_{X_{max}}(t_{40})$

Figure 5.28: Estimated sensitivity of t_{40} measured from single WCD trace (*left*) and SSD trace (*right*) with respect to X_{max}, given a perfect reconstruction of the energy and geometry of an event.

Table 5.1: Parametrized and simulated values concerning the time-dependent signal components for the example given in Fig. 5.23, along with the simulated signal sizes. The parametrized values are given in angle brackets.

WCD at $r = 1000\,\mathrm{m}$

	eγ	μ	eγ(μ)	eγ(π)	total
S/VEM	16.9	25.2	2.6	5.6	50.3
$(\langle t_{40}\rangle \pm \sigma_{t_{40}})$/ ns	771 ± 68	446 ± 54	464 ± 64	804 ± 52	607 ± 62
t_{40}/ns	795	463	570	696	596
$\langle\mu\rangle$	4.41	4.65	3.17	4.48	4.12
μ	4.50	3.70	4.21	4.26	4.08
$\langle\sigma\rangle$	0.65	0.85	0.95	0.67	0.80
σ	0.66	0.87	1.06	0.56	0.83

SSD at $r = 1000\,\mathrm{m}$

	eγ	μ	eγ(μ)	eγ(π)	total
S/MIP	34.3	21.1	1.4	9.4	66.1
$(\langle t_{40}\rangle \pm \sigma_{t_{40}})$/ns	858 ± 108	519 ± 109	521 ± 173	936 ± 235	758 ± 101
t_{40}/ns	887	470	587	1370	804
$\langle\mu\rangle$	4.58	3.90	4.10	4.79	4.43
μ	4.62	3.73	3.91	4.71	4.41
$\langle\sigma\rangle$	0.65	0.95	0.74	0.65	0.78
σ	0.55	1.07	0.86	0.07	0.72

CHAPTER VI

RECONSTRUCTION OF AIR-SHOWER EVENTS

The reconstruction of depth of the shower maximum X_{max} and the relative muon content R_{μ} solely from the signals of the surface detector stations is the main application of the Universality model of air showers. In this chapter, the reconstruction procedure as well as the accuracy estimated using simulations is discussed. Furthermore, we validate that the reconstructed values obtained from SD data are in accordance with data from FD measurements. For this purpose, the same set of Golden Hybrid events is reconstructed by both the standard Auger FD reconstruction procedure and the Universality reconstruction. The reconstruction is performed using the Offline framework. The reconstructed values of R_{μ} cannot yet be validated with data in the same way, since the data for the analyses given in Refs. [107] and [106] are obtained only from almost horizontal air showers with $\theta \geq 62^{\circ}$. This range in zenith angle is outside of the parameter space for which the Universality model developed in this work is valid.

Additionally, we propose a method to calculate the logarithmic atomic mass number $\ln A$ from the reconstructed values of X_{max} and R_{μ}. This method minimizes the effect of shower-to-shower fluctuations when attempting to discriminate light from heavy CRs, since the physical correlation of X_{max} and R_{μ} is taken into account, as it is discussed in Section 3.4.2.

6.1 THE UNIVERSALITY-BASED RECONSTRUCTION

The Universality reconstruction can be separated in the fit of the lateral distribution of the total signal deposited in the detectors and the time fit of the individual traces. These correspond to the models described in Section 5.7 and Section 5.8, respectively. The total likelihood function $\mathcal{L}$ of the reconstruction method therefore factorizes and the total log-likelihood can be written as

$$\ell = \ln \mathcal{L} = \ell_{\mathrm{lat}} + \ell_{\mathrm{time}}. \tag{6.1}$$

According to the considerations discussed in Section 5.10, ℓ_{lat} is expected to carry the predictive power with respect to R_{μ}, whereas the determination of X_{max} is mostly driven by ℓ_{time}. However, since ℓ_{lat} is also dependent on X_{max}, the reconstruction of X_{max} and R_{μ} cannot be performed at the same time maximizing ℓ, instead the individual parts of the log-likelihood function are maximized after another. This procedure is similar to the methods developed in previous works [128, 132, 133]. In this way, the correlation of the free parameters of the fit is limited. Furthermore, the physical correlation of X_{max} and R_{μ} is considered in ℓ_{lat} to minimize outliers in the reconstruction, in a method similar to the one performed in Ref. [109].

6.1.1 FIT OF THE LATERAL DISTRIBUTION

The Universality reconstruction of an event is performed after the standard reconstruction of the Offline framework, which is described in Ref. [97] Therefore, the reconstructed primary energy and the geometry for all candidate stations is already given to the Universality reconstruction as an input. The treatment of the signal in a detector is very similar to the standard procedure, however, instead of a lateral distribution function the Universality model of the total signal is used as prediction for the total deposited signal. The uncertainty of the signal in a WCD is well described by a Poissonian parametrization of the variance seen in data, where

$$\sigma_S = f(\theta)\,\sqrt{S}, \tag{6.2}$$

with a zenith dependent scaling factor f. The expected distribution of the signal is normally distributed for $S \gtrsim 20\,\mathrm{VEM}$ and follows a Poissonian distribution below that limit. The signal uncertainty of the SSDs can be parametrized in the same manner as Eq. (6.2), the expected distribution of signal, however, follows a truncated normal distribution [143, 145]. The log-likelihood of the lateral distribution of the signal is then given by the sum

$$\ell_{\mathrm{lat}} = \sum_d \ln P_d(S, S_{\mathrm{pred}}), \tag{6.3}$$

with the predicted signal S_{pred} that is a function of X_{max} and R_μ, and P following the outlined probability of the corresponding measured signal. Saturated detectors are in general discarded. The sum over d detectors thus does not contain signals from saturated WCDs, however it does contain the signal from a respective SSD mounted on the same station, if available. Furthermore, stations, for which the Universality model at the average value of X_{max} with respect to energy and $R_\mu = 1$ predicts less than 5 VEM or 5 MIP of signal, are rejected. This is the rough equivalent of a cut at approximately 2000 m for an event from a CR of $10^{19.5}$ eV. This is a conservative limit for the regions in which the parametrized Universality model is not stable anymore and the signal cannot be accurately predicted, as discussed in Section 5.7.3. The log-likelihood is maximized using the MINOU frontend for the MINUIT minimization library [146]. The start value of X_{max} is obtained from the average X_{max} at that primary energy, the start value of R_μ is set to the expected value for protons, which is 1.0 for simulations and approximately 1.3 for data. In this way, a best fit value for both X_{max} and R_μ is obtained.

For showers from high-energy particles at low zenith angles, where the shower maximum is very close to the ground and all candidate stations experience similar values of ΔX, a problem arises from the degenerate nature of the Gaisser-Hillas function. As it can be seen in Fig. 5.4, if ΔX is close to the maximum of the shower, the equivalent points of the ill-defined inverse of Eq. (5.41), $\Delta X(S)$, are close to each other. Different values for X_{max} can thus be reconstructed from the same footprint at the ground even with an ideal detector. Thus, since the lateral variation of the Universality signal model is mainly driven by the NKG function, which is in a good approximation independent of ΔX, at a given distance to the shower axis almost identical values for S are obtained for different ΔX. In these cases, if the shower maximum is very close or even below the ground, only few particles reach distant stations and thus the time information which is addressed in Section 6.1.2 usually contains little about the depth of the shower maximum. In the context of this work, this issue is addressed by a penalty term which is added to ℓ_{lat} and which incorporates the inter-primary correlation of R_μ and X_{max}, which will also be discussed in Section 6.4. Similarly to the parametrization of X_{max} as a function of R_μ that was used in Ref. [109], nonphysical parameter regions of shallow showers with very low muon content (and vice versa) are

thus artificially suppressed. This penalty term is realized by a normal distribution[1] around the average values of $X_{\max}$ as a function of R_μ. The result is a moderately constrained fit of $X_{\max}$. It was thoroughly tested that this term does not override the information given by the temporal distribution of signal in the traces, see Section 6.2.4. The scale of the penalty term remains a fudge factor and will be addressed in future work. In the work presented it is set close to its initial guess of $\xi_{X_{\max}}(t_{40})/\xi_{X_{\max}}(S)$ at 1500 m to optimize the correlation and resolution of $X_{\max}$ with respect to both simulations and Golden Hybrid events, see Section 5.10.

6.1.2 FIT OF THE TIME TRACES

To fit the time dependent signals deposited in the SD detectors, the model as parametrized in Section 5.8 is fitted to the responses of all considered detectors at the same time. Detector candidates for the fit of the time traces, however, are more conservatively selected than for the fit of the lateral distribution. Given the large uncertainties of t_{40} for large radii, detectors beyond a distance of 1800 m to the shower axis are discarded irrespective of their signal. The start time of the signal relative to the arrival time of the plane front in each detector is evaluated from the reconstructed geometry. The time-dependent signal is then fitted to the parametrized model, where $X_{\max}$ is treated as a free parameter. In this stage R_μ and the start value of $X_{\max}$ is fixed by the best-fit result from the procedure described in the previous section.

The time-dependent signal, however, is not fitted in its differential or *trace* form, but as integrated signal deposited over time, $S(t)$. This yields two advantages. Firstly, the data is artificially smoothed, as can be seen in Fig. 5.23, and spikes from single particles appear only as small steps in the integrated signal. Secondly, in this way it is easier to formulate a model of the time-dependent uncertainty of the signal.

Since the differential signal over time can be well described by a log-normal function, as discussed in Section 5.6.2, the time-dependent integrated signal $S(t)$ is expected to follow the expression

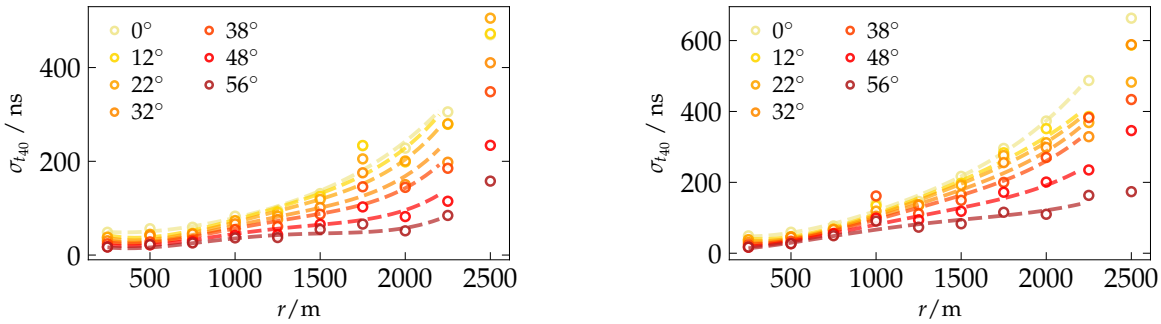
$$S(t) = S_{\text{tot}}\,\mathfrak{Ln}(\mu,\sigma) = S_{\text{tot}}\left(\frac{1}{2}+\frac{1}{2}\,\text{erf}\left(\frac{\ln\left((t-t_0)/\mu\right)}{2\sqrt{\sigma}}\right)\right), \tag{6.4}$$

where $\mathfrak{Ln}$ represents the integrated log-normal distribution function. The start time of the detector relative to the arrival time t_0 of the plane front is determined by the event geometry and the shape parameter μ is a function of t_{40} as defined in Eq. (5.37) [126]. The expression can thus be written as

$$S(t) = S_{\text{tot}}\,\mathfrak{Ln}(t_{40},\sigma) = S_{\text{tot}}\left(\frac{1}{2}+\frac{1}{2}\,\text{erf}\left(\frac{1}{2\sqrt{\sigma}}\ln\left(\frac{t-t_0}{t_{40}-t_0}\right)+\text{erf}^{-1}(-0.2)\right)\right). \tag{6.5}$$

The uncertainty of $S(t)$ is comprised of three effects which we consider to be independent of each other. Firstly, the uncertainty of the total signal σ_S is treated as discussed in Section 6.1.1. Secondly, fluctuations of the baseline of the detectors, which occur independently of events and mark an overall lower limit for the signal uncertainty. These are treated by a constant uncertainty σ_b. Lastly, there is the uncertainty of the arrival time of particles due to shower-to-shower fluctuations and due to low particle multiplicity at distant stations. Both are already contained in the total variance that is found for t_{40} in data. Using all showers available in the simulation library discussed in Section 5.7.1, the standard deviation of t_{40} is simply parametrized in terms of radius and zenith angle for all individual

[1] Additionally to a normal distribution, also a Gumbel distribution and a triangular distribution were tested for this purpose. All three yield qualitatively the same results.



(a) WCD (b) SSD

Figure 6.1: Parametrization of $\sigma_{t_{40}}$ obtained from the variance of the total traces as a function of radius and zenith angle for (a) the WCDs and (b) the SSDs.

particle components but also for the average total traces. The latter is depicted for both the WCDs and the SSDs in Fig. 6.1. As traces become shorter and particle multiplicity decreases with increasing distance of the detectors to the shower maximum, i.e. for increasing zenith angles, also $\sigma_{t_{40}}$ decreases. This effect is caused by both the lower particle multiplicity and the smaller curvature of the shower front at the ground for moderately to strongly inclined showers.

Assuming that the three sources of uncertainty are uncorrelated, we employ a linear propagation of the variance. The total uncertainty of the time dependent signal $\sigma_{S(t)}$ is expressed as

$$\sigma_{S(t)} = \sqrt{\left(\sigma_S \frac{\partial S(t)}{\partial S_{\text{tot}}}\right)^2 + \left(\sigma_{t_{40}} \frac{\partial S(t)}{\partial t_{40}}\right)^2 + \sigma_{\text{b}}^2}. \tag{6.6}$$

The first term evaluates to

$$\sigma_S \frac{\partial S(t)}{\partial S_{\text{tot}}} = \sigma_S \, \mathfrak{Ln}(t_{40}, \sigma), \tag{6.7}$$

with the integrated log-normal distribution function $\mathfrak{Ln}$ in Eq. (6.5); the second term of Eq. (6.6) requires certain attention. It can be shown either by explicit calculation or by the substitution $t \to -t$ for all times, that $\partial_{t_{40}} S(t)$ is equal to the log-normal distribution function multiplied by $(-S_{\text{tot}})$. Thus one finds

$$\sigma_{t_{40}} \frac{\partial S(t)}{\partial t_{40}} = -\sigma_{t_{40}} \, S_{\text{tot}} \, \mathfrak{ln}(t_{40}, \sigma), \tag{6.8}$$

where the log-normal distribution function is expressed in terms of the parametrized values of t_{40} and σ as given in Eq. (5.39). The negative sign is of course removed by the square in Eq. (6.6).

By formulating the uncertainty model of the signal in this way, the detectors with different electronics and different resulting sampling speed are also treated correctly. In the currently employed bin-by-bin Possonian model, where in a time bin of $8.\overline{3}$ ns width 1/3 of the signal is expected compared to a time bin of 25 ns, the uncertainty is over-estimated.

For the individual contributions from different detectors to the log-likelihood, the integrated time traces are normalized to the same length m, so that shorter traces, for example from downstream stations, are not under-represented in the fit. The log-likelihood of the

time-dependent signals in d stations is then given by a sum over all discrete time bins t_j,

$$\ell_{\text{time}} = \sum_d \left(\sum_{j=1}^{m} \mathcal{N}\left(S(t_j) - S_{\text{pred}}(t_j), \sigma_{S(t_j)} \right) \right)_d \tag{6.9}$$

where for convenience a normal distribution $\mathcal{N}$ is used for the probability of the individual signal bins. The first 100 ns of a time trace, usually just before or when the first bit of signal is deposited, are neglected because otherwise the fit is often dominated by the content of these first bins. This is the case when the start time of the signal trace is very different form the expected start time given by the Universality model, which could be due to low particle multiplicity or due to the finite precision of the reconstruction of the event geometry. The uncertainty, however, of the start time that is due to a non-optimal reconstruction of the shower geometry, is not yet considered in the fit of the time traces. For this reason, a tolerance of ± 100 ns is allowed for the predicted time-dependent signal for each detector.

6.2 VALIDATION USING SIMULATIONS

To test the accuracy of the Universality reconstruction of the to X_{max} and R_μ parameters, 64 000 events using showers from four different primaries, namely proton, helium, oxygen, and iron, each generated with CORSIKA and EPOS-LHC are simulated for the AugerPrime detector and reconstructed. In this way, approximately 16 000 showers for each simulated type of primary particle. Unlike the shower libraries used to parametrize the model, the primary energy and zenith angle of each event are not fixed to specific values, but are sampled randomly between $10^{18.6}$ eV and $10^{20.2}$ eV, and 0° and 56°, respectively. The distributions are uniform in $\lg(E_0/\text{eV})$ and $\sin^2\theta$. For $\lg(E_0/\text{eV}) > 20.0$ the simulation library, unfortunately does not contain showers with oxygen as a primary CR.

The MC values for X_{max} are obtained from the longitudinal profiles of the showers as stored by CORSIKA, analogously to the example profiles depicted in Fig. 5.2. X_{max} is given by the best-fit value of the respective parameter from the Gaisser-Hillas function to the longitudinal profiles from CORSIKA[2]. The longitudinal profiles, as stored by CORSIKA, are given in steps of 5 g cm^{-2} of vertical depth, so that the resulting vales for X_{max} carry a systematic uncertainty that is at least 2.5 g cm$^{-2} \times \sec\theta$.

Since the model is parametrized entirely in terms of signal, R_μ, as defined in Eq. (3.59), is assigned to each shower by the magnitude of the μ signal component relative to its expectation value from the average proton shower with $R_\mu = 1$. For this purpose, virtual dense stations are placed in the simulated SD array. Aside from the assignment of the MC value for R_μ, the virtual stations are ignored by the reconstruction procedure. Using the dense station information, R_μ is given by

$$R_\mu = \frac{1}{n} \sum_j^n \frac{S_\mu(\vec{r}_j)}{S_\mu^{\text{ref}}(\vec{r}_j)}, \tag{6.10}$$

where $\vec{r}_j$ are the positions of the n employed dense detectors. Multiple $\overline{\text{Off}}\underline{\text{line}}$ SD simulations of the same CORSIKA showers were used to determine the systematic uncertainty of the MC values for R_μ. With only one ring of dense WCDs at a distance of $r = 1000$ m the systematic uncertainty was evaluated to be $\sim$0.12, which is larger than the aspired precision for the reconstructed R_μ. Using multiple dense rings at three different distances with 4

[2] An additional value for X_{max} is stored in the database that is obtained by numerically interpolating the longitudinal profiles. The maximum of the interpolated profile, however, sometimes significantly differs from the value obtained by a best fit of the Gaisser-Hillas function. Especially when the shower maximum is lying under the ground, the maximum of the interpolated profile does not correspond to X_{max} anymore.

virtual dense stations each, all equipped with SSDs and WCDs and thus 24 virtual detectors in total, the systematic uncertainty on the MC values of R_μ was reduced to $\sim$0.04.

The residuals, to which we refer to in this section are the reconstructed observables relative to the MC values,

$$\Delta X_{\max} = X_{\max}^{\text{rec}} - X_{\max}^{\text{MC}}, \tag{6.11}$$

$$\Delta R_\mu = R_\mu^{\text{rec}} - R_\mu^{\text{MC}}. \tag{6.12}$$

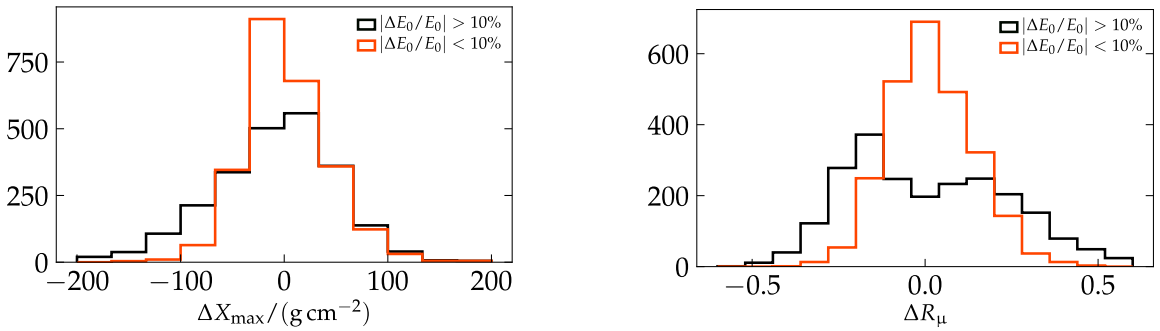
6.2.1 IMPACT OF THE RECONSTRUCTION OF THE PRIMARY ENERGY

The reconstruction of the primary energy is crucial for the Universality reconstruction, since possible best-fit values for E_0 and R_μ are highly correlated, as already discussed in Section 4.1.1 and as it can be seen in Eq. (5.13). Thus, using a single best-fit method it is not possible to precisely reconstruct both the primary energy and R_μ at the same time with reasonable accuracy. In Refs. [128] and [132], for example, this problem was treated with an iterative fit procedure and a strict constraint on the reconstructed primary energy around the MC energy. Given the muon deficit in simulations, which is discussed in Section 4.4.2, the reconstructed energy of the standard reconstruction procedure in Offline is not fit to describe simulated data, as it is calibrated to the energy as reconstructed by FD, see Fig. 4.5 (a). To solve this problem and to examine the benchmarks of the Universality reconstruction, for simulations the energy is set to the MC energy $\pm$10%, to mimic the real SD energy reconstruction and to compare to earlier work [128, 132]. When examining the data, the energy obtained from the standard SD reconstruction is treated as the primary energy.

If one tries to reconstruct the primary energy and R_μ simultaneously without any constraints, the results are rather unsatisfying. This was examined for 5000 simulated proton and iron CRs using the fit procedure outlined in Section 6.1.1, where E_0 was treated as an additional free parameter. Histograms of the residuals of the reconstructed observables $\Delta X_{\max}$ and ΔR_μ are depicted in Fig. 6.2 for CRs with $\lg(E_0/\text{eV}) > 19.5$ and $\theta < 50°$. The data are subdivided in events for which $\Delta E_0 = E_0^{\text{rec}} - E_0^{\text{MC}}$ is smaller or larger than 10% of the MC primary energy. The reconstruction of $X_{\max}$ is less severely affected as the reconstruction of R_μ, since it is mostly driven by the normalized, time-dependent signal. The reconstruction of R_μ, however, with a standard deviation of approximately 0.34 for the exclusive events with $|\Delta E_0/E_0| > 10\%$, would not be sufficient for the discrimination between heavy and light CR. The primary energy therefore needs to be known precisely for any Universality-based reconstruction being able to accurately identify R_μ. This is only the case for SD data above 10^{19} eV, as can be seen in Fig. 4.5.

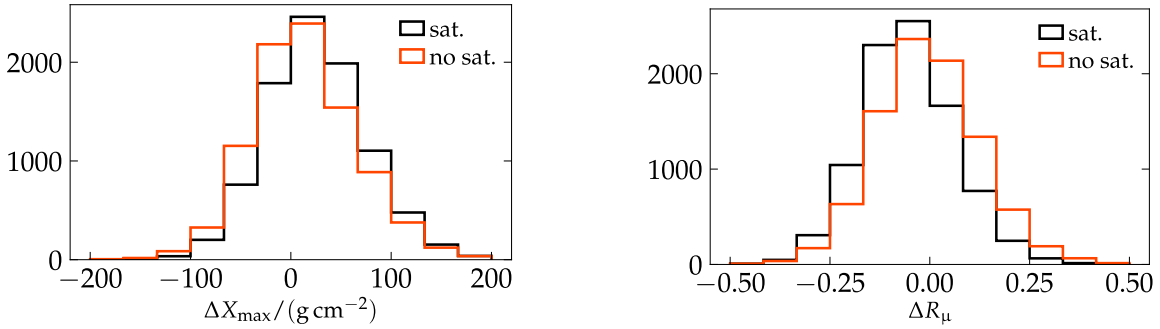
6.2.2 SATURATED EVENTS

Saturated events are events where at least one low-gain channel of a PMT is saturated, i.e. the electronics are overloaded with photoelectrons because too many particles from the air-shower reach the WCD (see Section 4.1). For high-energy events this is usually the case for the WCD closest to the shower axis, or for events with moderate and low energies when the shower core directly hits a detector station. At high energies around $\lg(E_0/\text{eV}) \gtrsim 19.5$, however, the number of particles close to the shower core is so large that the station closest to the shower core saturates approximately 50% of the time. In this case the station multiplicity is of course anyways large, but valuable information about the signal close to the shower axis is lost. In Ref. [128] and [132] a significant difference was reported for the quality of the reconstruction of $X_{\max}$ and R_μ with respect to saturated and non-saturated events, whereas for the reconstruction of $X_{\max}$ the opposite was reported in Ref. [127]. In this work only a minor systematic difference was found in saturated and non-saturated



(a) (b)

Figure 6.2: Residuals of the reconstructed values of $X_{\max}$ and R_μ for events with successful ($|\Delta E_0|/E_0 < 10\%$) and unsuccessful ($|\Delta E_0|/E_0 > 10\%$) energy reconstruction.



(a) (b)

Figure 6.3: Residuals of the reconstructed values of $X_{\max}$ and R_μ for saturated and non-saturated events with $\lg(E_0/\mathrm{eV}) > 19.5$ and $\theta < 50°$.

events for the reconstructed values of $X_{\max}$ and R_μ. Histograms of the residuals $\Delta X_{\max}$ and ΔR_μ are depicted in Fig. 6.3 for CRs with $\lg(E_0/\mathrm{eV}) > 19.5$ and $\theta < 50°$. In this region of energy, approximately half of the events contained at least one low-gain saturated station. The depicted data comprises showers from all four types of primary CRs. Saturated events are thus treated equally to non-saturated events, however with all saturated detectors being ignored by the Universality reconstruction.

6.2.3 EXAMPLE EVENTS

The reconstruction of simulated example events is depicted in Fig. 6.4. Each row corresponds to one simulated air-shower event. For each event the lateral distribution of signal, as well as the time-dependent signal deposited in the SSD and WCD of one example station is depicted.

The lateral distribution of signal in the shower plane (Fig. 6.4 *first column*) shows both the simulated signal in each detector as well as the predicted signal from the Universality parametrization as a function of the shower-plane distance for the two detector types. Detectors which are rejected because of too low signal are depicted in gray. The upper and lower edge of the bands of predicted signal correspond to the shower-plane azimuth values of $\psi = 0$ and $\psi = \pi$, respectively. The difference is caused by the variation in ΔX for stations at the same distance but at different shower-plane azimuth angles, see Eq. (5.11), which results in a different signal prediction by the modified Gaisser-Hillas profiles, see

Eq. (5.41). An additional but subdominant contribution is given by the correction factor c_ψ parametrized by Eq. (5.46).

The panels in the second and third column of Fig. 6.4 depict the simulated time-dependent signal as well as the respective prediction using the best-fit value of X_{max} as obtained from the Universality reconstruction for selected stations. The gray bands indicate the signal uncertainty as given by Eq. (6.6). Details of the reconstruction for the selected events (a), (b), (c), and (d) are given in Table 6.1.

The selected example events show different scenarios for the Universality reconstruction. Event (a), for example, appears to be successfully reconstructed, with the lateral distribution of signal in the respective detectors being very well aligned with the signal predictions. The time-dependent signal of the WCD at the distance of $r = 791\,\mathrm{m}$ perfectly matches the signal prediction for the best fit. The reconstructed value of X_{max}, however, deviates from the MC truth by $33\,\mathrm{g\,cm^{-2}}$. The opposite is the case for the example events (b) and (c). For both the depth of the shower maximum is very well reconstructed, but the prediction of the time-dependent signals appears to match the simulated data rather poorly, especially for the depicted WCD of event (b).

The selection does not showcase particularly well or badly reconstructed events, but rather tries to show typical examples in the data set. Events (a) and (b) are induced by protons, whereas (c) and (d) are induced by iron nucleus as the primary particle.

6.2.4 VALIDATION OF INTRA-PRIMARY CORRELATION

The correlation of X_{max} and R_μ from different types of primary particles has been parametrized using an empiric arctan-like function in earlier work (see Refs. [109] and [127]) to boost the performance of the reconstruction, as mentioned in Section 6.1.1. Given that $\xi_{R_\mu} \gg \xi_{X_{\mathrm{max}}}$ as discussed in Section 5.10, however, it is very well possible that in this case the predictive power for X_{max} is only driven by the reconstruction of R_μ. In this case, X_{max} is well predicted on average for different primaries, but intra-primary shower-to-shower fluctuations (i.e. within a single type of primary particles) cannot be recovered and sensitivity to shower-to-shower fluctuation is lost in terms of X_{max}. For a mix of primary particles, X_{max} can then still be estimated on average by its dependence on the primary energy, that is em-

Table 6.1: Properties of the simulated air-shower events depicted in Fig. 6.4 along with the reconstructed observables and details of the reconstruction. Shower-plane coordinates as well as the start time and χ^2 values of the best fit are given for the selected station of each event.

	$\lg(E_0/\mathrm{eV})$	θ	$X_{\mathrm{max}}^{\mathrm{MC}}$	$X_{\mathrm{max}}^{\mathrm{rec}}$	R_μ^{MC}	R_μ^{rec}
(a)	19.35	23.5°	$862\,\mathrm{g\,cm^{-2}}$	$829\,\mathrm{g\,cm^{-2}}$	1.09	1.04
(b)	19.13	28.1°	$795\,\mathrm{g\,cm^{-2}}$	$812\,\mathrm{g\,cm^{-2}}$	1.14	1.21
(c)	18.85	22.0°	$684\,\mathrm{g\,cm^{-2}}$	$701\,\mathrm{g\,cm^{-2}}$	1.33	1.39
(d)	19.04	42.3°	$734\,\mathrm{g\,cm^{-2}}$	$732\,\mathrm{g\,cm^{-2}}$	1.42	1.30

	r	ψ	t_0	χ^2_{WCD}	χ^2_{SSD}
(a)	791 m	−12.0°	107 ns	303	4451
(b)	1682 m	−15.5°	515 ns	11551	8034
(c)	1050 m	129.0°	162 ns	344	727
(d)	1162 m	10.3°	147 ns	141	797

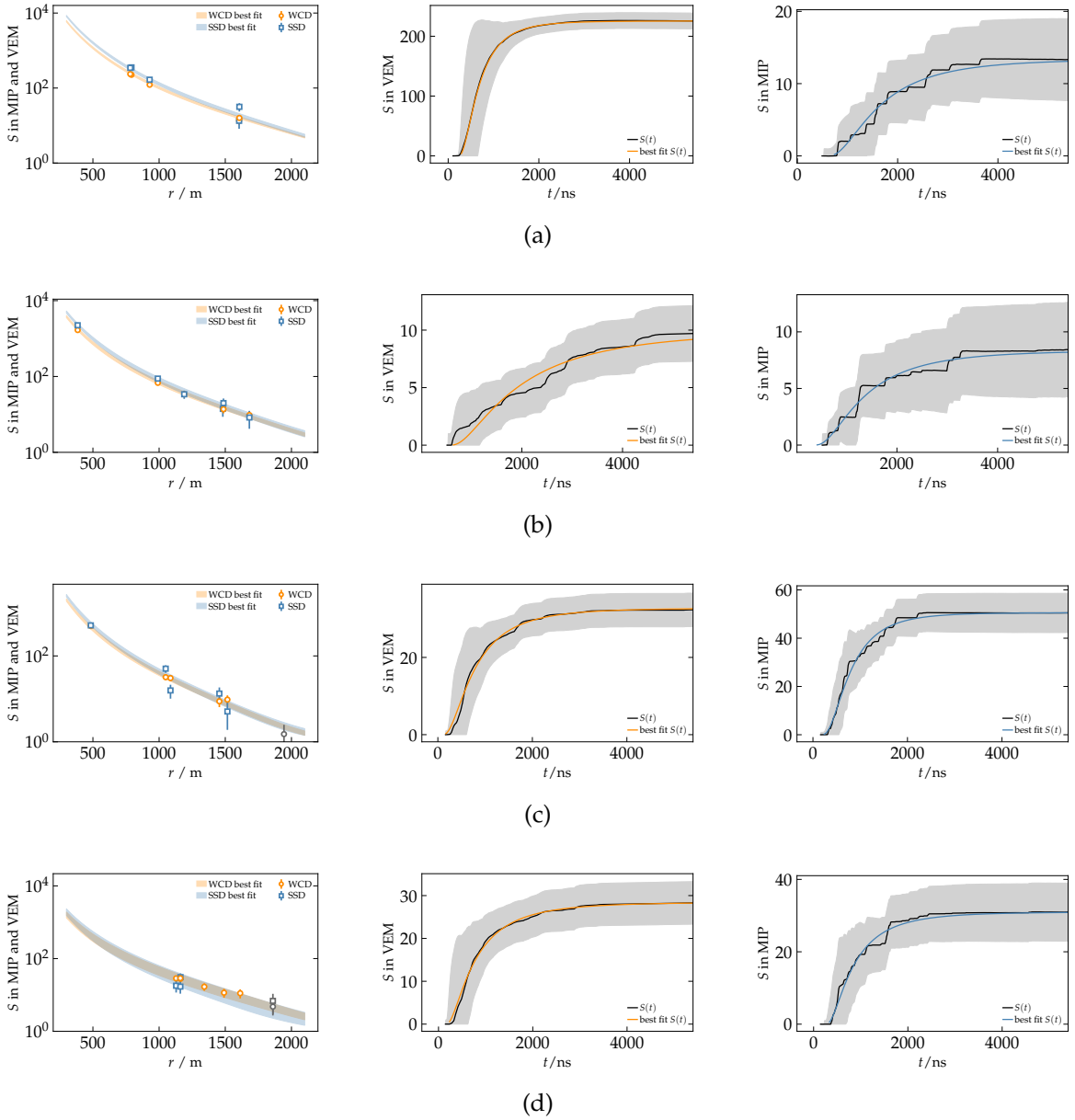


Figure 6.4: Universality reconstruction of four different simulated example events. The first column shows the lateral distribution of signal of all detectors for each event. The second and third columns depict the time-dependent signal deposited in the WCD and SSD of one selected station, respectively. Details are given in Table 6.1.

ployed for the start value of the fit. The reconstructed values of X_{max}, however, would then yield no added value for the recovery of the primary mass of the CR. To validate that both is not the case for the method developed in this work, the intra-primary correlation of the MC and the reconstructed values for X_{max} is examined for proton showers in very narrow windows in zenith angle and energy. The result is depicted in Fig. 6.5. The data comprises on average about 100 showers in each subset. The correlation gradually increases with energy and the Pearson correlation coefficient ϱ exceeds $\sim$0.7 for the highest energies. Thus, the reconstructed values of X_{max} from the Universality reconstruction are indeed correlated with the real depths of the shower maxima and thus yield added value to identify the primary mass of the CRs, independently of R_{μ}.

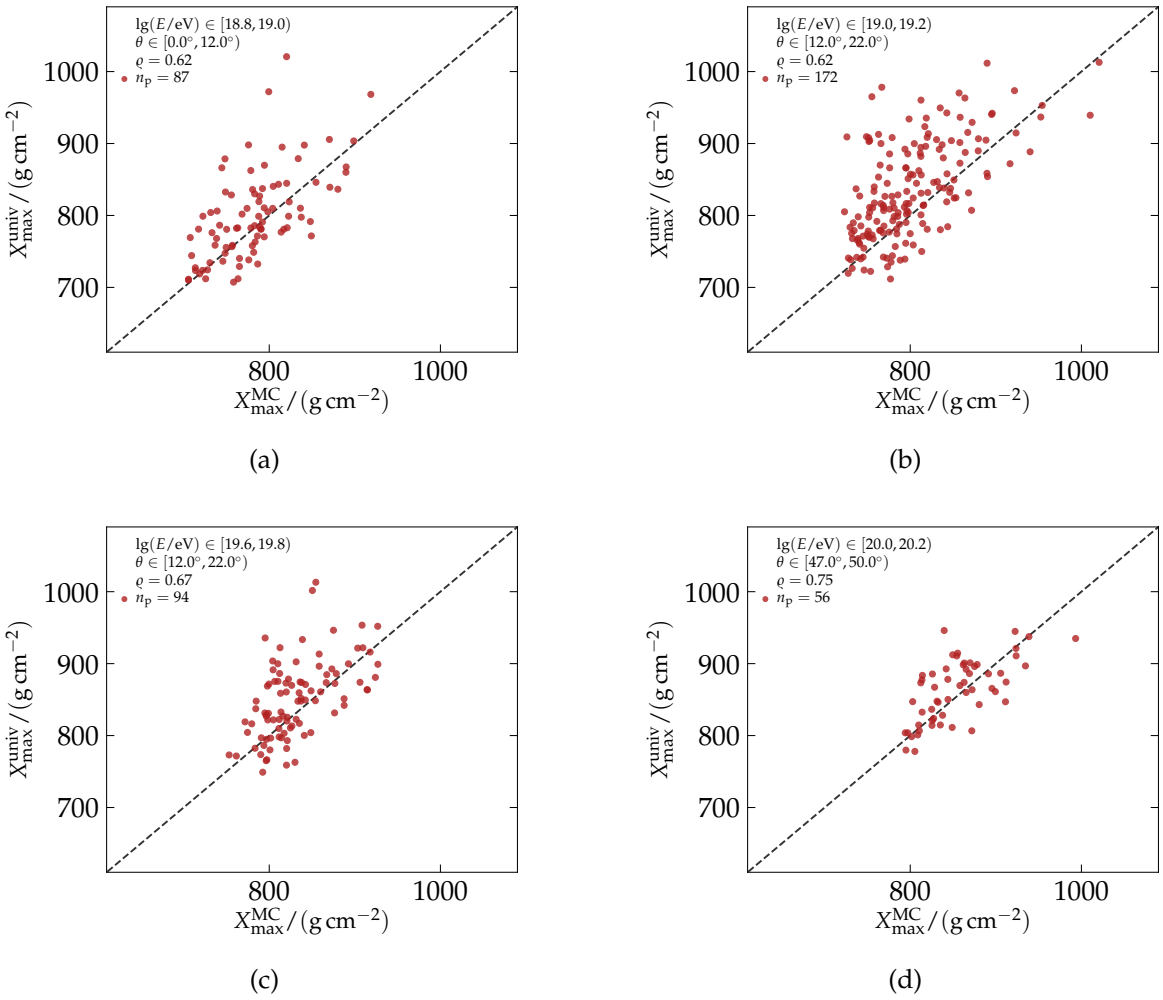


Figure 6.5: Correlation of the real and reconstructed values of $X_{\max}$ from n_p simulated proton showers in increasing ranges of energy and zenith angle. The Pearson correlation coefficient is given in the upper left of each frame, together with the corresponding window in energy and zenith angle from which the showers were sampled. The identity function is depicted as a gray dashed line.

6.2.5 ACCURACY AND SIMULATION DRIVEN CORRECTIONS

The average values obtained for $X_{\max}$ and R_μ from the Universality reconstruction are not entirely unbiased. Even though on average the reconstruction performs well, both $X_{\max}$ and R_μ systematically differ from the MC truth, especially as a function of the zenith angle of the CR. A depiction of the residuals $\Delta X_{\max}$ and ΔR_μ is given in Fig. 6.6. The residuals as a function of $\sin^2\theta$, averaged over the energy, are given in panels (a) and (b). The residuals as a function of $\lg(E_0/\mathrm{eV})$, averaged over the zenith angle, in panels (c) and (d). To maintain accuracy of the Universality reconstruction over the whole parameter space in energy and zenith angle, a bias correction is applied according to

$$\delta x = \sum_i a_i \left(\sin^2\theta\right)^i \qquad = a_0 + a_1 \sin^2\theta + \dots, \tag{6.13}$$

$$\delta x = \sum_i a_i \left(\lg(E_0/\mathrm{eV})\right)^i = a_0 + a_1 \lg(E_0/\mathrm{eV}) + \dots. \tag{6.14}$$

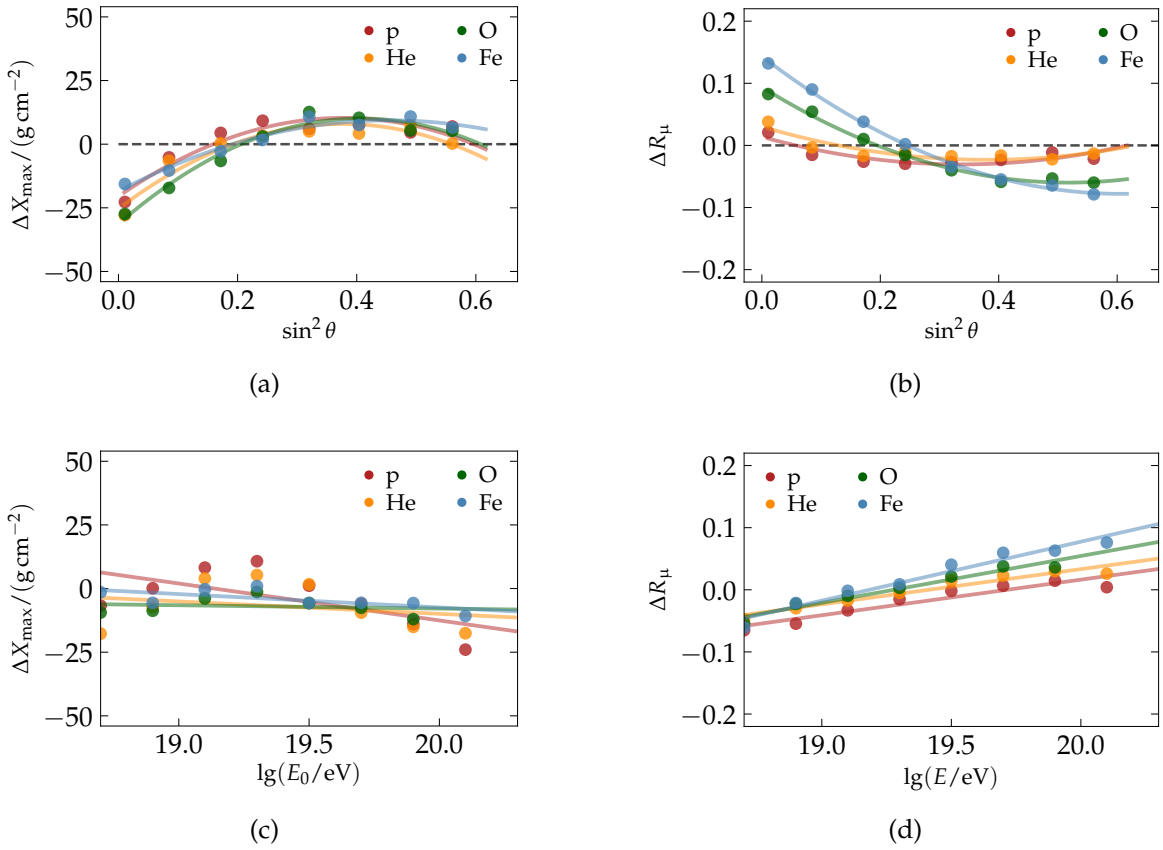


Figure 6.6: Average residuals, $\Delta X_{max} = X_{max}^{rec} - X_{max}^{MC}$ and $\Delta R_\mu = R_\mu^{rec} - R_\mu^{MC}$, as a function of the primary energy and zenith angle. The bias over $\sin^2\theta$ is depicted with a quadratic interpolation, whereas the energy-dependent bias is depicted with a linear interpolation.

For both the reconstructed values of X_{max} and R_μ the corresponding values are added for each event. The numerical values for the bias corrections are given in Table 6.2.

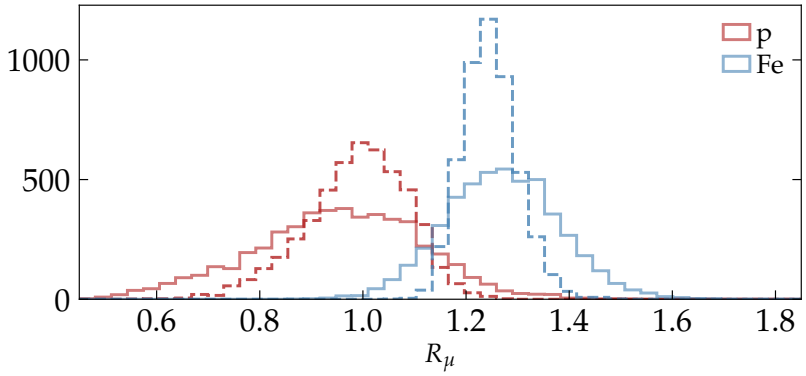
When examining the precision of the Universality reconstruction relative to the simulated truth, no energy-dependent bias correction is applied to the reconstructed values of X_{max}, because X_{max} will be ultimately calibrated using the Golden Hybrid events (see Section 6.3).

Histograms of the reconstructed values of R_μ and X_{max} at the highest energies after the bias correction are given in Fig. 6.7, along with the corresponding distributions of MC values. Both the distributions of the reconstructed R_μ and X_{max} qualitatively match the MC data. The distributions of reconstructed R_μ for iron and proton are given in Fig. 6.7 (a), however, they are visibly smeared with respect to the MC data.

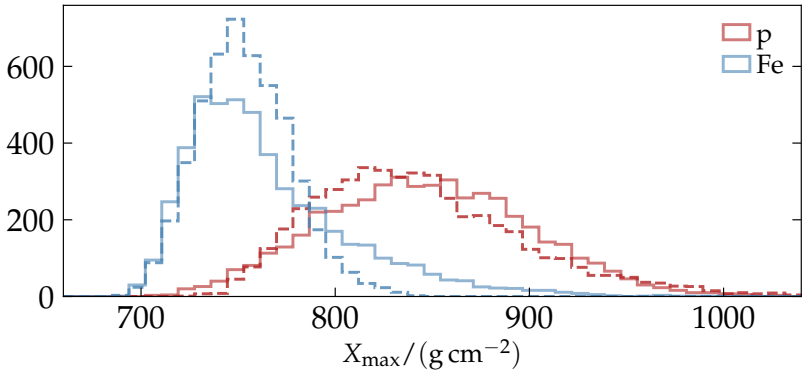
Two-dimensional representations of the data depicted in Fig. 6.6 after the outlined bias correction are given in Figs. B.15 and B.17.

6.2.6 ESTIMATED PRECISION

In this section, the precision of the Universality reconstruction is evaluated using the available simulation library. The bias correction is applied to both the reconstructed values of X_{max} and R_μ, see Section 6.2.5. In a previous work, the precision of a reconstruction algorithm based on a Universality parametrization of showers simulated with QGSJET-II-03 reached $\sigma(X_{max}) \simeq 40\,g\,cm^{-2}$ [127] and $\sigma(R_\mu) \simeq 0.2$ [109] for the 1500 m array of Auger. These results were later improved using strict quality cuts on the data and refined fitting



(a)



(b)

Figure 6.7: Distributions of the reconstructed (full lines) and MC values (dashed lines) of (a) R_μ, and (b) X_{max}, for simulated proton and iron events with $\lg(E_0/eV) > 19.5$ and $\theta < 50°$.

algorithms [128, 132]. At the highest energies, however, the event rate is extremely low and any quality cut might remove valuable data. Thus, in this section, the precision of the Universality reconstruction is examined without any additional quality cuts. The standard deviation of the residuals of the reconstructed values for X_{max} and R_μ are given in Fig. 6.8 as functions of $\lg(E_0/eV)$ and $\sin^2\theta$.

As a function of the zenith angle, the reconstruction of X_{max} appears to perform equally well for all zenith angles up to $\theta \simeq 50°$, with even a slight improvement for very vertical and moderately inclined events for heavy CRs, see Fig. 6.8 (a). The reconstruction of R_μ, on the other hand, shows a clear optimum in the range of zenith angles between $25° < \theta < 35°$ for all types of primary particles, see Fig. 6.8 (b). A possible explanation for the performance for X_{max} is the proximity to the shower maximum for vertical showers, while X_{max} is mostly above ground for heavy primary CR even at very low zenith angles. For rather inclined showers, more stations are triggered by the event, which can also result in a boost of performance up to the point where the information of the curvature of the shower front is lost. Over nearly the whole parameter space in $\sin^2\theta$, the precision of the Universality reconstruction undercuts the extent of the shower-to-shower fluctuations of proton showers. For the average precision in each bin in $\sin^2\theta$, showers for which $\lg(E_0/eV) < 19.0$ were disregarded.

With increasing primary energy, the reconstruction of both X_{max} and R_μ becomes more precise. This is because at higher energies larger signals with smaller relative uncertainty and in general more triggered stations are available for each event. As a function of the

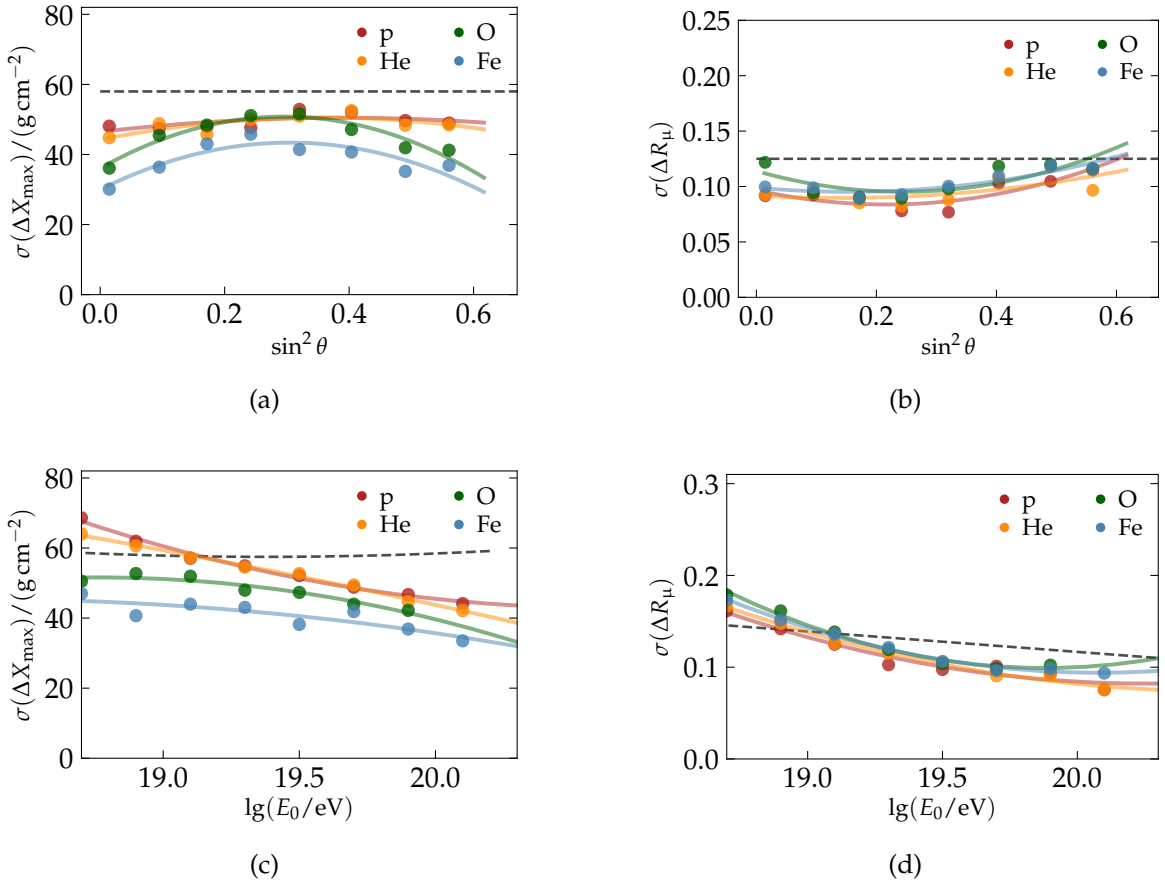


Figure 6.8: Precision of the reconstructed values of X_{max} and R_μ given as standard deviation of the residuals, ΔX_{max} and ΔR_μ, as a function of the primary energy and zenith angle. The extent of the shower-to-shower fluctuations of proton showers is given as a reference as a gray, dashed line. The lowest bin in energy is not included for the average values depicted in panels (a) and (b).

primary energy, the precision is rather poor until the primary energy reaches $\lg(E_0/eV) \simeq 19.5$. Beyond this energy, however, the results are satisfying and allow for an approximate estimation of the primary mass based on X_{max} and R_μ, independently.

The overall precision of the Universality reconstruction could be further improved by a conservative event selection, as it was shown in Refs. [128] and [132]. Additionally, Ref. [132] introduced an approach to reconstruct the event geometry and primary energy, all simultaneously from the Universality parametrization of showers simulated with QGSJET-

Table 6.2: Simulation driven progressive corrections to the reconstructed values of X_{max} and R_μ as functions of $\sin^2\theta$ and $\lg(E_0/eV)$. Corrections are applied according to Eq. (6.14) up to a quadratic term.

		a_0	a_1	a_2
$\delta X_{max}/(g\,cm^{-2})$	$\sin^2\theta$	24.42	−215.3	281
	$\lg(E_0/eV)$	0	0	0
δR_μ	$\sin^2\theta$	−0.064	0.492	−0.422
	$\lg(E_0/eV)$	1.453	0.075	0

II-03. The optimization of the reconstruction algorithm for the Universality parametrizations introduced in this work or parametrizations which are produced analogously, however, is outside the scope of this work.

Two-dimensional representations of the data given in Fig. 6.8 are given in Figs. B.16 and B.18.

6.3 VALIDATION OF THE X_{MAX} RECONSTRUCTION USING FD

Lastly, to test the Universality reconstruction method is able to accurately reconstruct X_{max} with the precision estimated in Section 6.2.6, the Golden Hybrid events are examined. In this case, X_{max} is reconstructed with Universality independently and simultaneously with the FD reconstruction, using only the information provided by the SD data and the SD standard reconstruction. The residuals of the reconstruction considered in this section are thus

$$\Delta X_{\mathrm{max}} = X_{\mathrm{max}}^{\mathrm{univ}} - X_{\mathrm{max}}^{\mathrm{FD}}. \tag{6.15}$$

The efficiency of the Universality reconstruction is rather poor at low and moderate energies, i.e. for $\lg(E_0/\mathrm{eV}) \lesssim 19.2$. Only events at the highest energies are thus considered in this section. Furthermore, strong restrictions are required for the selection of the FD events, such as the *Field-of-View*[3] cut, which ensures that the shower maximum is directly visible in one of the eyes of the FD telescopes. Additionally, a limit on the event-level uncertainty of the reconstructed value for X_{max} is required, which is by default $\sigma_{\mathrm{FD}} < 40\,\mathrm{g\,cm^{-2}}$ for the reconstructed depth of the shower maximum on an event level. For the Universality reconstruction, all events up to a zenith angle of $\theta = 50^\circ$ are taken into account. Misreconstructed events, for which the best fit for X_{max} and R_μ did not converge are disregarded. These make up about 0.5% of all events. Such cases include events containing broken stations with inordinately low or no signal. No further cuts on the data besides the restrictions on the ranges of energy and zenith angle are imposed from the SD reconstruction.

The correlation of the reconstructed shower depths as reported by FD and the Universality reconstruction is depicted in Fig. 6.9. Because of the increasingly low event multiplicity at high energies, the events are not binned in energy but constrained only by a lower limit. Unfortunately, only above energies of $10^{19.3}$ eV a strong correlation is obtained with the data. Surprisingly, the correlation improves the more conservative the event selection is on the side of the FD reconstruction. For panels (b), (c), and (d) of Fig. 6.9, FD events are only selected if they were detected by more than one telescope. At the highest energies, the Pearson correlation of the reconstructed values of X_{max} exceeds 0.5. Without the Field-of-View cut or any restriction of σ_{FD}, only poor agreement is found in the reconstructed values for X_{max} from universality and the FD. Examining the residuals of X_{max} relative to the FD reconstructed data, a negative bias of the order of $\sim 20\,\mathrm{g\,cm^{-2}}$ is apparent. Distributions of the residuals are given in Fig. 6.10. To increase the event multiplicity, the FD event-selection cuts were slightly softened with respect to the data depicted in Fig. 6.9, requiring only one FD telescope for each event and the standard cut of $\sigma_{\mathrm{FD}} < 40\,\mathrm{g\,cm^{-2}}$. This bias is in agreement with the results from the previous work on Universality as well as with recent results from a neural network which is trained to reconstruct X_{max} from SD data in a similar manner [132, 147]. This bias can be corrected for by a constant and an additional zenith-dependent bias correction that is discussed in Section 7.1.1.

[3]The Field-of-View cut is one of the most conservative selection criteria on the quality of the FD data. Only events where the maximum of the longitudinal profile of the shower is directly observed within the field of view of the telescopes are selected. About 20% of all high-energy events fulfill this condition.

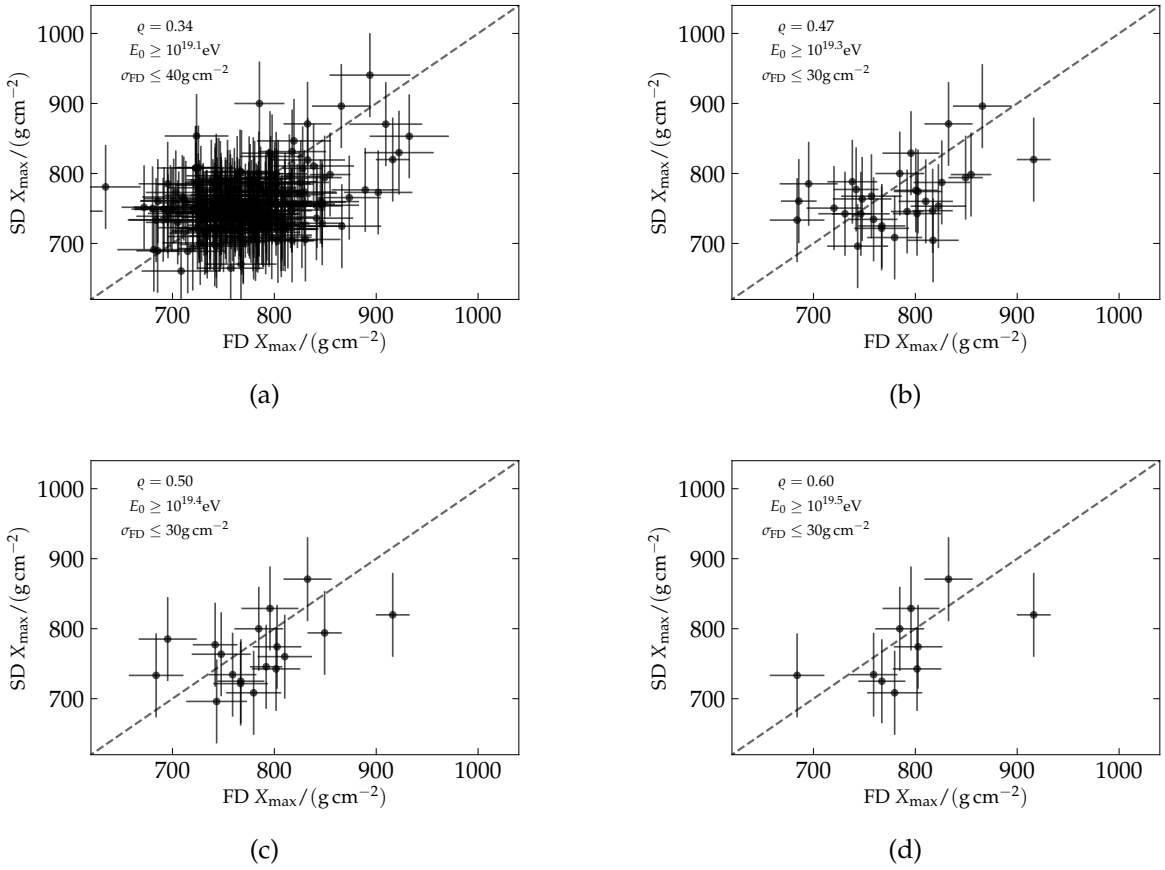


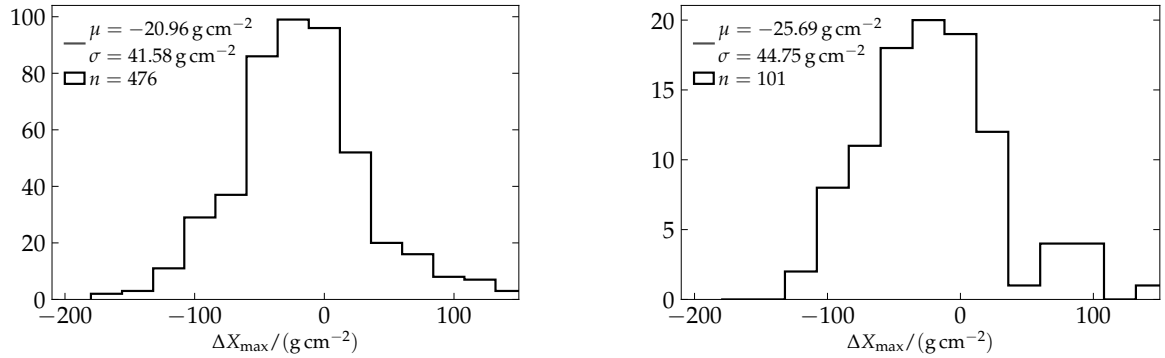
Figure 6.9: Correlation of X_{max} as reconstructed by FD and by the Universality reconstruction. Correlation coefficients ϱ as well as the lower limit on reconstructed primary energy and upper limit on event-based uncertainty σ_{FD} of the FD reconstruction are given in the upper left corner of each panel. Note that for the data shown in panel (a) $\sigma_{\mathrm{FD}} \leq 40\,\mathrm{g\,cm^{-2}}$ whilst for (b), (c), and (d) the limit is set more conservatively to $30\,\mathrm{g\,cm^{-2}}$. Additionally, the events shown in panels (b), (c), and (d) are required to be detected by at least two of the four FD telescopes.

The standard deviation of the residuals of X_{max} relative to FD is $\sim$42 g cm^{-2}, as can be seen in Fig. 6.10. Considering the finite resolution of the FD, which is $\sigma_{\mathrm{FD}}^{\mathrm{sys}} \simeq 16\,\mathrm{g\,cm^{-2}}$ [104], the standard deviation of the residuals of the Universality reconstruction of X_{max} is given by

$$\sigma_{X_{\mathrm{max}}}^{\mathrm{univ}} = \sqrt{(42\,\mathrm{g\,cm^{-2}})^2 - (16\,\mathrm{g\,cm^{-2}})^2} \simeq 39\,\mathrm{g\,cm^{-2}}. \tag{6.16}$$

Thus, the precision of the reconstructed values of X_{max} is estimated to be $39\,\mathrm{g\,cm^{-2}}$. This is in agreement with the results depicted in Fig. 6.8, in which for the whole data set the precision of the reconstruction of X_{max} is $\sim$40 g cm^{-2}.

The average reconstructed values of X_{max} as a function of the SD energy is depicted in Fig. 6.11. Each bin of data contains exactly the same events, reconstructed individually by the Universality reconstruction and the FD telescopes. The results from the Universality reconstruction are treated with an energy-independent bias correction of approximately $20\,\mathrm{g\,cm^{-2}}$. The average simulated values for X_{max} for proton and iron primaries using QGSJET-II and EPOS-LHC are given as a reference. The Universality reconstruction very successfully follows the trend of the average X_{max} reconstructed by the FD as a function of the energy.



(a) $E_0 \geq 10^{19}$ eV (b) $E_0 \geq 10^{19.4}$ eV

Figure 6.10: Distributions of the residuals of X_{max} reconstructed by FD and Universality for n selected events. No bias correction is applied. Sub-captions indicate the lower limits on the SD reconstructed primary energies. Average residuals μ and standard deviations σ of the distributions are given in the upper left corner of each panel.

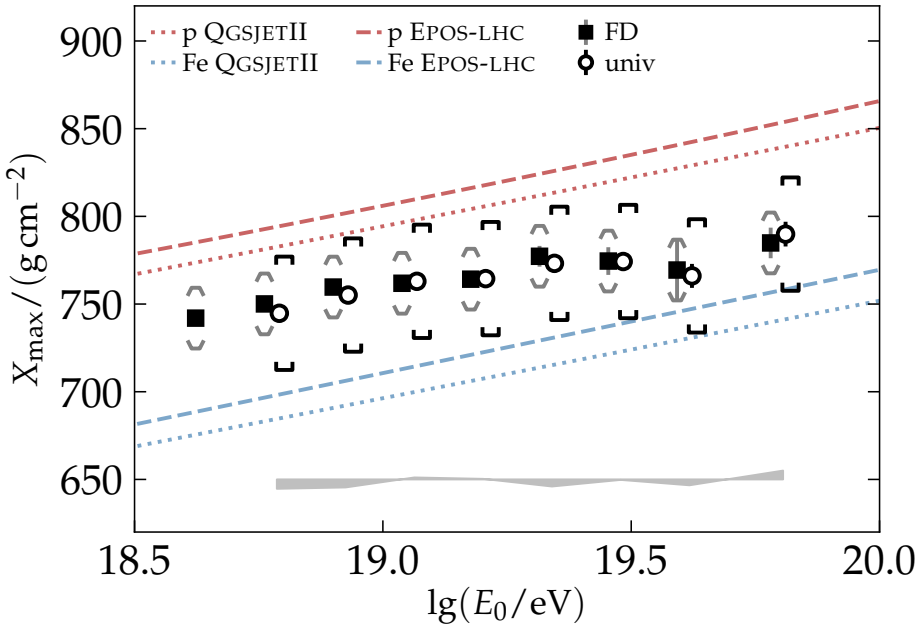


Figure 6.11: Average X_{max} as a function of the reconstructed primary energy E_0. Data comprises the same Golden Hybrid events reconstructed by both the Auger FD and Universality after a bias correction. The markers in each bin are shifted slightly apart to make the comparison easier. The difference in the average X_{max} for each energy is indicated by the gray shape placed at 650 g cm^{-2}. Systematic uncertainties are given as brackets.

6.4 DETERMINATION OF THE ATOMIC MASS NUMBER

As mentioned already in Chapter 2, one of the frontiers of UHECR physics is the determination of the masses and consequently the charges of the CRs. For this task, the observable X_{max} is most prominently used in analyses such as given in Ref. [105], followed by studies on R_μ, such as given in Ref. [107]. Even in an extreme-case scenario with only iron and proton CRs, unfortunately both the distributions for R_μ and especially for X_{max} significantly overlap. Therefore, even with an ideal detector, mass discrimination is difficult using ei-

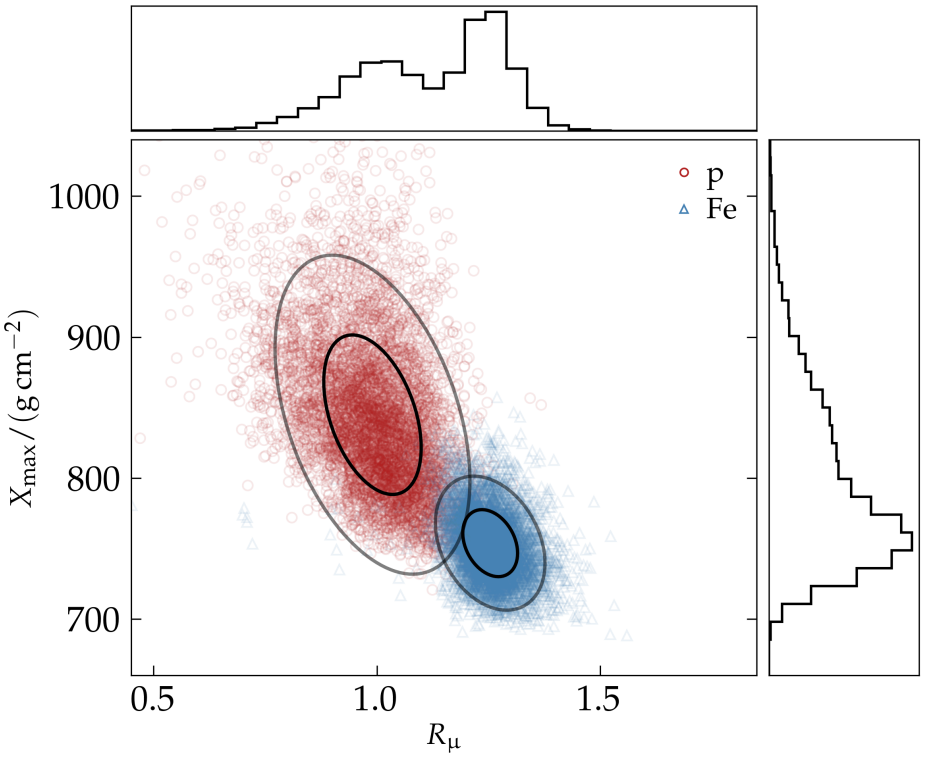


Figure 6.12: Inter-primary correlation of the values of $X_{\max}$ and R_μ. Depicted data comprises MC values of 6000 simulated air showers with primary energies above $10^{19.3}$ eV, for proton and iron, each. Black and gray ellipses show the $1\,\sigma$ and $2\,\sigma$ contours of the data for the respective primary particle.

ther of the two observables. When different primaries are considered, the average values of $X_{\max}$ and R_μ, however, are not independent, as discussed in Section 3.4.2. It was suggested already in Ref. [148] that the two-dimensional information given by both observables when considered at the same time will boost the sensitivity on the primary mass. In fact, when depicted in a two-dimensional way, the distributions of MC values for $X_{\max}$ and R_μ of proton and iron showers can be clearly separated almost at the $2\,\sigma$ level, as can be seen in Fig. 6.12.

In this section, we introduce a method to determine $\ln A$ from the observables $X_{\max}$ and R_μ, based on the combined information of the two observables. Additionally, results on the obtained precision for $\ln A$ based on simulations are presented, alongside with considerations on the estimated precision of the method when performing on data.

6.4.1 THE BASE TRANSFORMATION METHOD

From the Heitler-Matthews model of air showers, which is discussed in Section 3.4, average values for $X_{\max}$ and R_μ can be obtained for any value of $\ln A$, as described by Eqs. (3.53) and (3.55). Inverting these relations, $\ln A$ can be independently estimated by

$$-\lambda \ln A = X_{\max} - X^{\mathrm{p}}_{\max}, \quad \text{and} \quad \beta \ln A = \ln R_\mu - \ln R^{\mathrm{p}}_\mu, \tag{6.17}$$

given $X_{\max}$ or R_μ. Here we identify λ and β as the change of the average of $X_{\max}$ and $\ln R_\mu$, respectively, when $\ln A$ increases by 1. $X^{\mathrm{p}}_{\max}$ and R^{p}_μ denote the average values obtained from proton showers. To reduce the effect of outliers, the analysis in this section is carried out using the respective median values instead of the averages of $X_{\max}$ and $\ln R_\mu$. λ and β

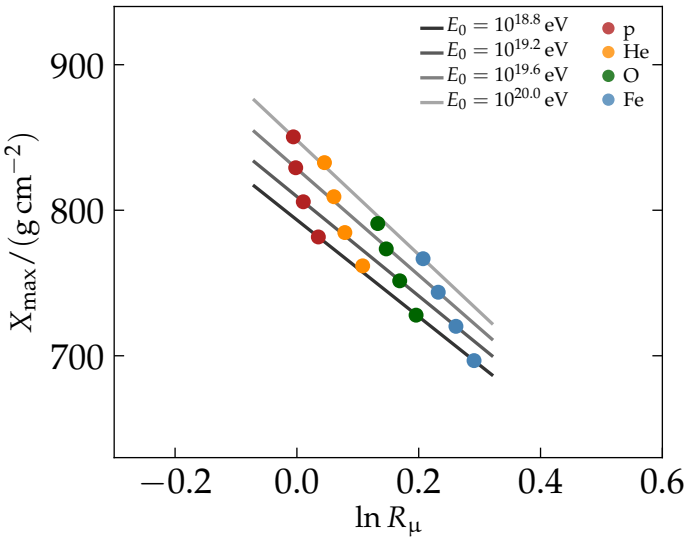


Figure 6.13: Median values for X_{max} and $\ln R_\mu$ as a function of the primary energy and the primary particle. Each data point is calculated from the distribution of 3000 showers generated with EPOS-LHC within a range of ± 0.2 in $\lg(E_0/\mathrm{eV})$ of the indicated energy.

can be calculated as

$$\lambda = -\frac{X^{\mathrm{p}}_{max} - X^{\mathrm{Fe}}_{max}}{\ln(56)} \quad \text{and} \quad \beta = \frac{\ln R^{\mathrm{p}}_\mu - \ln R^{\mathrm{Fe}}_\mu}{\ln(56)}. \tag{6.18}$$

From the data depicted in Fig. 6.13, λ and β can be approximated by a function linear in the logarithmic primary energy, given by

$$\lambda = 21.44\,\mathrm{g\,cm^{-2}} - (\lg(E_0/\mathrm{eV}) - 19)\times 0.234\,\mathrm{g\,cm^{-2}}, \tag{6.19}$$

$$\beta = 0.064 - (\lg(E_0/\mathrm{eV}) - 19)\times 0.009. \tag{6.20}$$

The median values of X_{max} and $\ln R_\mu$ for the four different types of primary particles with different energies are depicted in Fig. 6.13[4]. As expected from the discussion in Section 3.4, X^{p}_{max} is well described by a function linear in $\lg(E_0/\mathrm{eV})$. The numerical values of $\ln R^{\mathrm{p}}_\mu$ are approximately constant. It was verified in Ref. [149] that this holds even for showers from different hadronic interaction models. In fact, except for an overall shift in $\ln R^{\mathrm{p}}_\mu$ and X^{p}_{max}, all occurring parameters are approximately identical when comparing showers from EPOS-LHC and QGSJET-II.

Instead of considering X_{max} and R_μ independently, the joint distribution of values of X_{max} and R_μ, as depicted in Fig. 6.12 tempts us to make use of the apparent anti-correlation, but most importantly, to make use of the lack of the anti-correlation given only a single primary. Whilst in Fig. 6.12 iron showers are frequently found at shallow shower maxima around $700\,\mathrm{g\,cm^{-2}} < X_{max} < 800\,\mathrm{g\,cm^{-2}}$, proton showers extent from $X_{max} \simeq 1000\,\mathrm{g\,cm^{-2}}$ down to $X_{max} \simeq 750\,\mathrm{g\,cm^{-2}}$. In the region of the overlap of the two distributions in X_{max}, the primary particles can still clearly be separated by their individual R_μ. The equivalent holds for the separation in R_μ – in the region of $R_\mu \simeq 1.2$, a shower with a deep shower maximum is likely a proton, while the opposite holds for iron. The combined information of both X_{max} and R_μ is thus considered to estimate $\ln A$. In this way, the effect of shower-to-shower fluctuations is minimized.

[4] As the shower library unfortunately does not contain oxygen showers with primary energies between 10^{20} eV and $10^{20.2}$ eV, this data point does not perfectly match the linear regression for the highest energies.

In the $X_{\max}$-$\ln R_\mu$ plane depicted in Fig. 6.13, for each primary energy a one-dimensional subspace can be identified that contains the medians of the $X_{\max}$-$\ln R_\mu$ distributions for all types of primary particles from proton to iron. We define the base $\{\hat{r}, \hat{x}\}$ parallel to the axes of $\ln R_\mu$ and $X_{\max}$,

$$(\ln R_\mu - \ln R_\mu^{\mathrm{p}})\,\hat{r} + (X_{\max} - X_{\max}^{\mathrm{p}})\,\hat{x} = \begin{pmatrix} \ln R_\mu - \ln R_\mu^{\mathrm{p}} \\ X_{\max} - X_{\max}^{\mathrm{p}} \end{pmatrix}, \tag{6.21}$$

so that the medians of the distribution of proton showers is given by the point $(0,0)^{\mathrm{T}}$ for each energy. For a fixed energy (such as given by gray lines in Fig. 6.13) we define the base vector $\hat{e}_{\ln A}$, such that the ordinate evaluates to the logarithmic mass number of the respective primary at each center of distributions of $X_{\max}$ and $\ln R_\mu$. A second base vector $\hat{e}_\varphi$ is defined to be linearly independent of $\hat{e}_{\ln A}$. Both systems $\{\hat{r}, \hat{x}\}$ and $\{\hat{e}_{\ln A}, \hat{e}_\varphi\}$ then span the whole two-dimensional space in $X_{\max}$ and $\ln R_\mu$, and the points of origin of the two coordinate systems coincide. The matrix $\mathfrak{T}$ can now be defined such that

$$\mathfrak{T}\left((\ln R_\mu - \ln R_\mu^{\mathrm{p}})\,\hat{r} + (X_{\max} - X_{\max}^{\mathrm{p}})\,\hat{x}\right) = \ln A\,\hat{e}_{\ln A} + \varphi\,\hat{e}_\varphi. \tag{6.22}$$

The interpretation of φ is addressed in the following section. Given the conditions that

$$\mathfrak{T}(\beta\,\hat{r} - \lambda\,\hat{x}) = \mathfrak{T}\begin{pmatrix} \beta \\ -\lambda \end{pmatrix} \stackrel{!}{=} 1\,\hat{e}_{\ln A} + 0\,\hat{e}_\varphi, \tag{6.23}$$

$$\mathfrak{T}(\beta\,\hat{r} + \varphi_0\,\hat{x}) = \mathfrak{T}\begin{pmatrix} \beta \\ \varphi_0 \end{pmatrix} \stackrel{!}{=} 0\,\hat{e}_{\ln A} + 1\,\hat{e}_\varphi, \tag{6.24}$$

with a constant φ_0, $\mathfrak{T}$ evaluates to

$$\mathfrak{T} = \frac{1}{\beta(\lambda + \varphi_0)} \begin{pmatrix} \varphi_0 & -\beta \\ \lambda & \beta \end{pmatrix}. \tag{6.25}$$

Using $\mathfrak{T}$, every point on the $X_{\max}$-$\ln R_\mu$ plane can be projected along a line of slope φ_0/β onto the one-dimensional subspace in which $\ln A$ is simply given as the ordinate. The numerical value of $\ln A$ is thus given by

$$\ln A = \frac{1}{\beta(\lambda + \varphi_0)} \left(\varphi_0 \left(\ln R_\mu - \ln R_\mu^{\mathrm{p}}\right) - \beta \left(X_{\max} - X_{\max}^{\mathrm{p}}\right)\right). \tag{6.26}$$

6.4.2 DETERMINATION OF φ_0

For any primary, $X_{\max}$ and $\ln R_\mu$ are distributed with a finite width, thus $\ln A$ cannot be sharply predicted using Eq. (6.26) and the width of the distribution of the calculated values for $\ln A$ is dominated by the shower-to-shower fluctuations in $X_{\max}$ and $\ln R_\mu$. The variance of the distributions is linearly propagated such that

$$\begin{aligned} \sigma_{\ln A}^2 &= \left(\sigma_{X_{\max}} \frac{\partial \ln A}{\partial X_{\max}}\right)^2 + \left(\sigma_{\ln R_\mu} \frac{\partial \ln A}{\partial \ln R_\mu}\right)^2 \\ &= \frac{\beta^2\,\sigma_{X_{\max}}^2 + \varphi_0^2\,\sigma_{\ln R_\mu}^2}{\beta^2\,(\lambda + \varphi_0)^2}, \end{aligned} \tag{6.27}$$

where σ_k is the standard deviation of the distribution of the observable k. σ_k therefore comprises both the natural spread of $X_{\max}$ and R_μ and the resolution of a non-ideal detector. $\sigma_{\ln A}$ can be minimized as a function of φ_0. The optimal value for φ_0 is thus given by

$$\varphi_0 = \frac{\beta^2\,\sigma_{X_{\max}}^2}{\lambda\,\sigma_{\ln R_\mu}^2} \tag{6.28}$$

and $\sigma_{\ln A}$ reduces to

$$\sigma_{\ln A} = \sqrt{\frac{\sigma^2_{\ln R_\mu}\sigma^2_{X_{\max}}}{\lambda^2\sigma^2_{\ln R_\mu} + \beta^2\sigma^2_{X_{\max}}}}. \tag{6.29}$$

For proton and iron showers above an energy of $10^{19.3}$ eV, simulated values of $X_{\max}$ and R_μ are distributed with a width of approximately

$$\sigma^{\mathrm{p}}_{X_{\max}} \simeq 60\,\mathrm{g\,cm^{-2}}, \qquad \sigma^{\mathrm{p}}_{\ln R_\mu} \simeq 0.13, \tag{6.30}$$

$$\sigma^{\mathrm{Fe}}_{X_{\max}} \simeq 20\,\mathrm{g\,cm^{-2}}, \quad \text{and} \quad \sigma^{\mathrm{Fe}}_{\ln R_\mu} \simeq 0.08. \tag{6.31}$$

This corresponds to the width of the distribution of $X_{\max}$ and R_μ as reconstructed by an ideal detector. Given that $\lambda \simeq 21\,\mathrm{g\,cm^{-2}}$ and $\beta \simeq 0.065$, the lower limits on $\sigma_{\ln A}$ are given by

$$\sigma^{\mathrm{p}}_{\ln A} \simeq 1.6 \quad \text{and} \quad \sigma^{\mathrm{Fe}}_{\ln A} \simeq 0.6. \tag{6.32}$$

To obtain such a precision on data, however, φ_0 must be adjusted depending on whether for example proton or iron primaries are observed[5], as it is clearly dependent on $\sigma_{X_{\max}}$ and $\sigma_{\ln R_\mu}$. Since the information about the primary is in general not known a priori, this is impossible. Still, depending on the respective values of $X_{\max}$ and R_μ, an estimate on $\sigma_{X_{\max}}$ and $\sigma_{\ln R_\mu}$ can be given. Assuming a linear behaviour of $\sigma_{X_{\max}}$ with respect to $X_{\max}$ and of $\sigma_{\ln R_\mu}$ with respect to $\ln R_\mu$, both values can be estimated as

$$\sigma_{X_{\max}}(X_{\max}) = \sigma^{\mathrm{p}}_{X_{\max}} + \frac{\Delta\sigma^{\mathrm{p,Fe}}_{X_{\max}}}{\lambda\ln(56)}(X_{\max} - X^{\mathrm{p}}_{\max}), \tag{6.33}$$

$$\sigma_{\ln R_\mu}(\ln R_\mu) = \sigma^{\mathrm{p}}_{\ln R_\mu} - \frac{\Delta\sigma^{\mathrm{p,Fe}}_{\ln R_\mu}}{\beta\ln(56)}(\ln R_\mu - \ln R^{\mathrm{p}}_\mu). \tag{6.34}$$

The absolute difference of the respective widths of the distributions from iron and proton showers are indicated as $\Delta\sigma^{\mathrm{p,Fe}}$.

Using Eqs. (6.33) and (6.34), φ_0 can then be determined according to Eq. (6.28) for all values of $X_{\max}$ and R_μ. $\ln A$ can then be calculated according to Eq. (6.26) for any pair of values of $X_{\max}$ and R_μ. Using this method, the estimated precision given in Eq. (6.32) was verified on MC data, using the data as depicted in Fig. 6.12. The result is given in Fig. 6.14; the values obtained from data perfectly match the estimated optimal precision given in Eq. (6.32).

6.4.3 COMPARISON TO SIMILAR METHODS

The base transformation method is similar to both a principle component analysis, as well as to a linear Fisher discriminant analysis. The advantage of this method with respect to the principle component analysis lies in the data-driven adjustment of φ_0 with respect to $X_{\max}$ and $\ln R_\mu$. It can be easily proven that for a conventional principle component analysis, where the principle components are considered constant with respect to data, i.e. $\varphi_0 \neq$

[5] Simply put, φ_0 quantifies how much $X_{\max}$ and $\ln R_\mu$ are weighted with respect to each other for the determination of $\ln A$, with both contributions evening out approximately at $\varphi_0 \simeq 40\,\mathrm{g\,cm^{-2}}$. For extremely large values of φ_0, only $\ln R_\mu$ contributes to the result of $\ln A$, whereas for $\varphi_0 = 0\,\mathrm{g\,cm^{-2}}$ only the value of $X_{\max}$ is taken into account. For this reason φ_0 is dependent on the spread and consequently on the uncertainty of the respective observables.

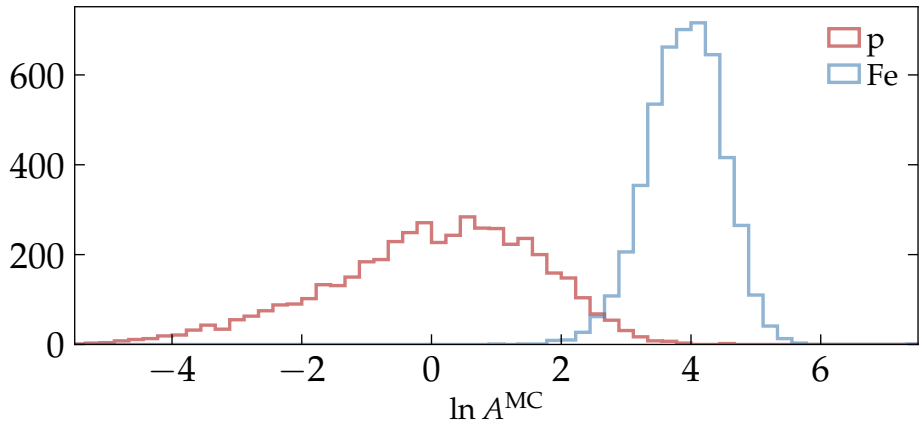


Figure 6.14: Distribution of the obtained values for $\ln A$ using Eq. (6.26) and MC values for X_{max} and R_μ. The standard deviation of the data for protons is given by $\sigma^{\mathrm{p}}_{\ln A} = 1.59$, while for iron $\sigma^{\mathrm{Fe}}_{\ln A} = 0.58$. The mean values of the distributions are located at $\langle \ln A \rangle^{\mathrm{p}} = 0.2$ and $\langle \ln A \rangle^{\mathrm{Fe}} = 4.0$. The same data as depicted in Fig. 6.12 is used.

$\varphi_0(X_{\mathrm{max}}, \ln R_\mu)$, the obtained accuracy on $\ln A$ is worse, since Eq. (6.27) is not minimized in this case. Furthermore, the method is superior to a linear Fisher discriminant, since a continuous value for $\ln A$ is obtained instead of a binary output.

The main advantage of the base transformation method over both methods, however, is its capability of being applied to data. Conventional discrimination methods, including machine-learning approaches, are rigidly optimized on the results of simulated air showers, for which we know that the absolute number of muons is incorrect, see Section 4.4.2. Without any possibility to calibrate on data, a considerable systematic uncertainty thus remains. Furthermore, if the general treatment of the soft-QCD in simulations is not optimal, also X_{max} is expected to behave differently in simulations than in reality. As mentioned earlier, an analysis was carried out in Ref. [149] that indicated that at first order the change in the physics of hadronic interactions could be treated by a simple shift of the parameters $\ln R^{\mathrm{p}}_\mu$ and $X^{\mathrm{p}}_{\mathrm{max}}$, as it was already earlier suggested in Ref. [150]. Results that indicate $\langle \ln A \rangle \gg 4$, such as depicted in Fig. 4.8, can be corrected by such a shift, as $\ln A$ as obtained from Eq. (6.26) is linearly dependent on both $X^{\mathrm{p}}_{\mathrm{max}}$ and $\ln R^{\mathrm{p}}_\mu$.

6.4.4 CALIBRATION USING GOLDEN HYBRID EVENTS

The expected values for $\ln R_\mu$ and X_{max} for proton showers act as the coordinate origin in the base transformation method to reconstruct $\ln A$, so that at this point $\ln A = 0$. From simulations we expect $\ln R^{\mathrm{p}}_\mu = 0$. This is, however, not the case for data because of the apparent muon deficit. Using data from the Golden Hybrid events, $\ln R^{\mathrm{p}}_\mu = 0.26$ is determined for the parametrization of proton showers simulated with EPOS-LHC. This corresponds to $R^{\mathrm{p}}_\mu \simeq 1.3$. Assuming that despite the muon deficit in simulations the average number of muons in an air-shower is given by Eq. (3.58), this implies for iron showers $R^{\mathrm{Fe}}_\mu \simeq 1.6$ on average. β and λ are assumed equal for data and simulations.

For the purpose of the calibration, the values of X_{max} obtained from the FD reconstruction as well as the corresponding values of $\ln R_\mu$ from the Universality reconstruction are compared to the joint distribution of X_{max} and $\ln R_\mu$ obtained from simulations, see Fig. 6.15. At energies below 10^{19} eV a substantial fraction of protons is expected to be present in the particle spectrum of CRs, see Fig. 2.2 and Ref. [16]. Thus, only showers in the region of $18.8 < \lg(E_0/\mathrm{eV}) < 19.0$ are examined to identify the true $\ln R^{\mathrm{p}}_\mu$ in data. At these

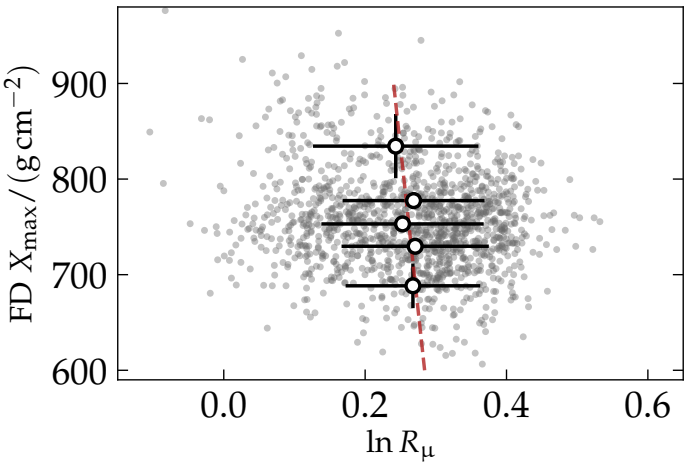


Figure 6.15: Calibration data from the Golden Hybrid events with energies in the primary energy range $18.8 < \lg(E_0/\,\mathrm{eV}) < 19.0$. The depth of the shower maximum is obtained by the FD reconstruction, whereas $\ln R_\mu$ is obtained from Universality. Individual data points are depicted in gray. For 5 subsets reaching from low to high X_{max}, each with equal amount of data, the mean in X_{max} and $\ln R_\mu$ is depicted as black circle, and the standard deviation of the data is given by the error bars. A linear best fit is indicated by a red dashed line.

energies, $\ln R_\mu^p$ can be found as the mean value of $\ln R_\mu$ for which the mean of the distribution of the reconstructed values of X_{max} matches the expectations from simulated protons. The data was subdivided into 5 sets of equal size from low to high X_{max}. From proton showers simulated with EPOS-LHC, the average depth of the shower maximum is expected at $\sim$795 g cm^{-2} for primary energies around $\lg(E_0/\mathrm{eV}) \simeq 18.9$. A linear interpolation of the data depicted in Fig. 6.15 yields the value of $\ln R_\mu^p = 0.26$ at $X_{max} = 795\,\mathrm{g\,cm^{-2}}$.

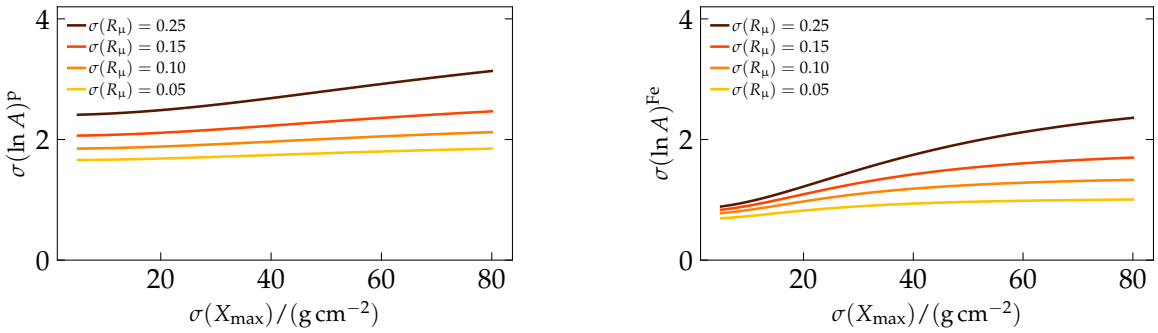
The calibration with the Golden Hybrid events removes most of the systematic uncertainties due to the muon deficit for the reconstruction of $\ln A$. A certain ambiguity, however, remains, because the average reconstructed values from FD for X_{max} heavily depend on the applied quality cuts and thus directly affect $\ln R_\mu^p$ and thus $\ln A$ for each event. Furthermore, depending on which range of X_{max} to consider the average for protons and to calibrate against, $\ln R_\mu^p$ will change. This is a direct impact of the systematic discrepancies in the present hadronic interaction models.

6.5 BENCHMARKS FOR THE RECONSTRUCTION OF $\ln A$

The estimation of the logarithmic atomic mass number of CRs from X_{max} and R_μ is the final goal of the reconstruction procedure. In this section, we examine the accuracy of the reconstructed values of $\ln A$ obtained from simulations using the method outlined in Section 6.4.

As it is discussed in Section 6.4.2, the precision with which $\ln A$ can be reconstructed depends heavily on the size of the shower-to-shower fluctuations for individual primaries. For iron-induced, less fluctuating air showers, the reconstructed value of $\ln A$ will on average be closer to the desired value of $\ln(56) \simeq 4$, given an unbiased estimate of X_{max} and R_μ (see Fig. 6.14). The total widths σ_{tot} of the reconstructed observables X_{max} and R_μ, which depends on the precision of the reconstruction σ_{prec} as well as the contribution of the shower-to-shower fluctuations $\sigma_{sh\text{-}sh}$,

$$\sigma_{tot} = \sqrt{\sigma_{sh\text{-}sh}^2 + \sigma_{prec}^2}, \tag{6.35}$$



(a) (b)

Figure 6.16: Estimated precision of the reconstructed values of $\ln A$ according to Eq. (6.29) as a function of the precision of the reconstructed observables R_μ and X_{max} for (a) proton and (b) iron primaries.

limits the possible precision for the reconstructed values of $\ln A$. The expected precision for $\ln A$ can be estimated as a function of the precision of R_μ and X_{max} according to Eq. (6.29). Using the estimate of the shower-to-shower fluctuations given in Eqs. (6.30) and (6.31), $\sigma(\ln A)$ is given as a function of $\sigma(X_{\mathrm{max}})$ and $\sigma(R_\mu)$ in Fig. 6.16. The result of the estimation shows both a clear increase in the precision of the reconstructed $\ln A$ with increasing precision in the reconstructed X_{max} and R_μ, as well as the limitations given by the shower-to-shower fluctuations (see Eq. (6.32)).

To test the accuracy of the event-by-event reconstruction of $\ln A$, the air-shower events of the simulation library are reconstructed and $\ln A$ is calculated from the reconstructed values of X_{max} and R_μ. The mean bias and resolution as a function of the energy and zenith angle is depicted in Fig. 6.17 for $\Delta \ln A := \ln A^{\mathrm{rec}} - \ln A^{\mathrm{MC}}$. Besides restricting the primary energy and zenith angle to each bin, no further quality cuts are applied on the simulation data.

In contrast to the precision of the reconstruction of X_{max} and R_μ, the resolution of $\Delta \ln A$ is rather flat as a function of energy and zenith angle. Additionally, significant differences in resolution occur for the different primary particles. As already mentioned before, this is mainly due to the natural spread of the observables X_{max} and R_μ, resulting in a rather wide distribution of the "ideal" values (obtained from the MC X_{max} and R_μ) of $\ln A$ for light elements (see Eq. (6.32)). The results depicted in Fig. 6.17 qualitatively match the expectations for the estimated precision, given $\sigma(X_{\mathrm{max}}) \simeq 40\,\mathrm{g\,cm^{-2}}$ and $\sigma(R_\mu) \simeq 0.1$ (see Fig. 6.8), given in Fig. 6.16.

A zenith-dependent bias remains for $\Delta \ln A$ that, however, is addressed for by a data-driven correction, see Section 7.1. A small $\ln A$-dependent bias is apparent at the highest energies, which on average leads to heavy elements being reconstructed even heavier as they are, and light events being reconstructed sightly lighter than they are. Elements of moderate atomic mass, such as oxygen, are reconstructed without any significant energy-dependent bias. The resolution of the reconstructed values of $\ln A$ ranges between approximately $\sigma_{\ln A} \simeq 1$ for iron and $\sigma_{\ln A} \simeq 2.3$ for proton. The precision, with which $\ln A$ can be reconstructed, is thus predominantly driven by the stochastic behaviour of X_{max} and R_μ, i.e. shower-to-shower fluctuations, which is less severe for heavier elements.

A distribution of the reconstructed values for all events above an energy of $10^{19.5}\,\mathrm{eV}$ is given in Fig. 6.18 (a). The distributions of $\ln A$ for the four different types of primaries are expected around the respective real values of

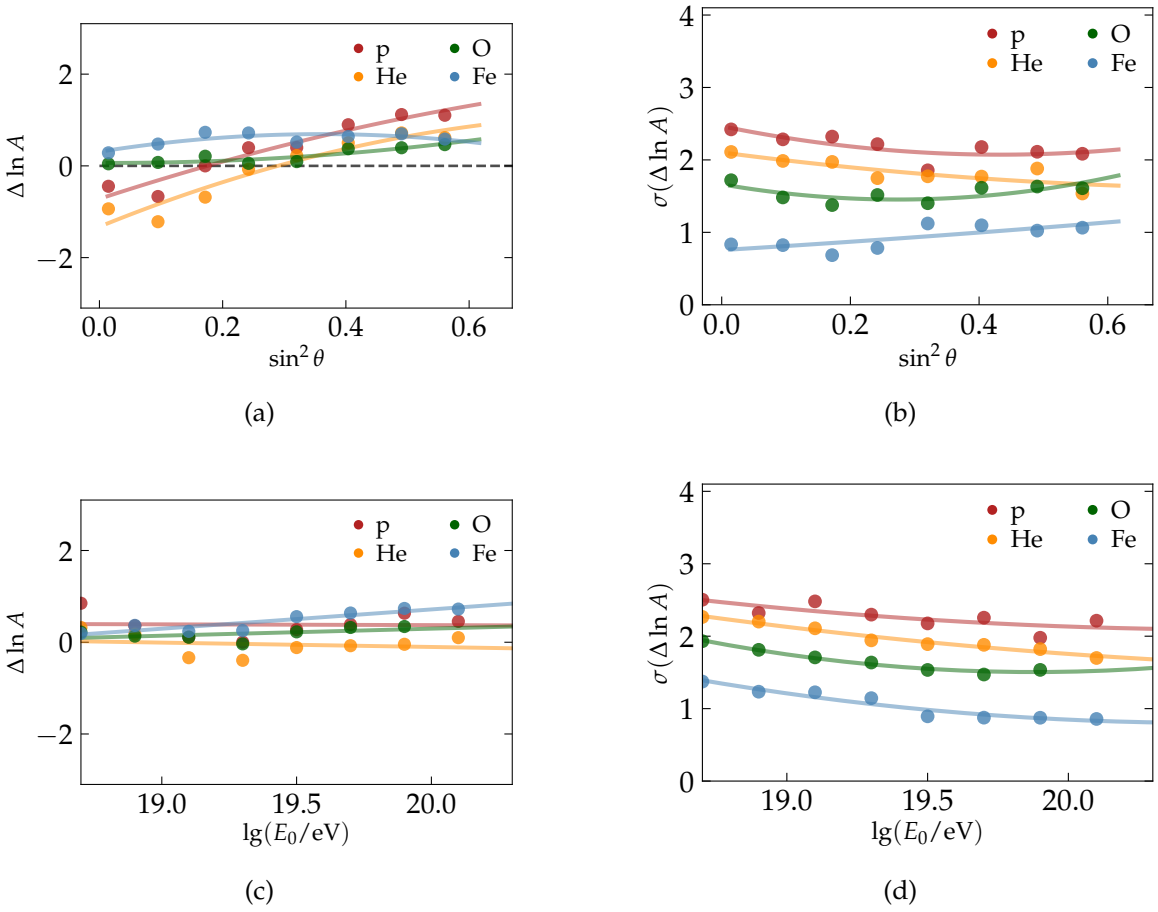


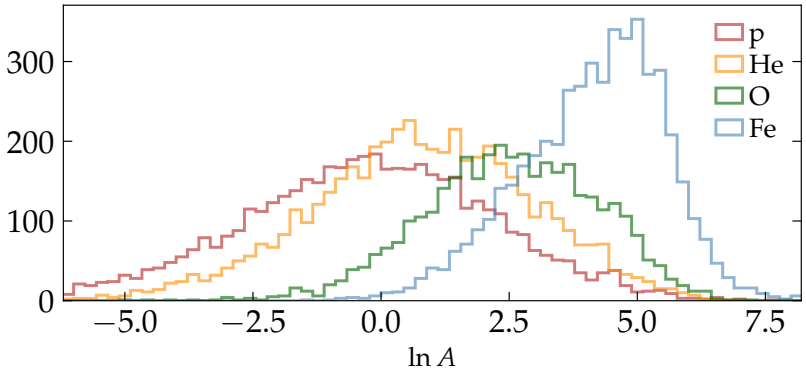
Figure 6.17: Precision of the reconstructed values of $\ln A$ given by the mean residual and its standard deviation as a function of the primary energy and zenith angle. For the dependence on $\sin^2\theta$ the lowest bin in primary energy is not included.

- $\ln A_\mathrm{p} = \ln(1) = 0$,
- $\ln A_\mathrm{He} = \ln(4) \simeq 1.4$,
- $\ln A_\mathrm{O} = \ln(16) \simeq 2.7$, and
- $\ln A_\mathrm{Fe} = \ln(56) \simeq 4.0$.

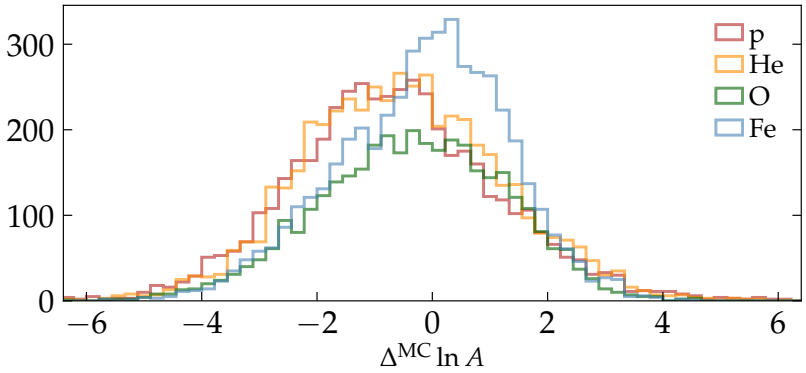
The residuals of the reconstructed values of $\ln A$ with respect to $\ln A$ as calculated from the MC values of X_max and R_μ, $\Delta^\mathrm{MC} \ln A := \ln A(X^\mathrm{rec}_\mathrm{max}, \ln R^\mathrm{rec}_\mu) - \ln A(X^\mathrm{MC}_\mathrm{max}, \ln R^\mathrm{MC}_\mu)$, are depicted in Fig. 6.18 (b). All distributions have approximately the same width, implying the reconstruction of $\ln A$ using the combined information of X_max and R_μ works approximately equally well for all primary particles. As already mentioned, the precision of $\ln A$ for heavier primaries mainly stems from the smaller effect of shower-to-shower fluctuations for these particles.

6.6 IDENTIFYING THE LIGHTEST EVENTS

In this section, the ability of the presented method to identify and identify light CRs in data is discussed. Using the data depicted in Fig. 6.18, the quality to identify protons (or light elements) in a selected subset of the data using a binary decision is examined. The receiver operating characteristic curve (ROC curve) for this binary classification is given in Fig. 6.19 for a mix of protons and one of the three remaining primary particles. It is evident that the separation is best for proton and iron primary particles, since only a minor overlap exists for



(a)



(b)

Figure 6.18: Distributions of the absolute and MC-relative reconstructed values of $\ln A$ for 18 000 CRs, simulated with $\lg(E_0/\mathrm{eV}) > 19.5$ and $\theta < 50°$. Absolute values are depicted in panel (a), whereas (b) shows the residuals with respect to the corresponding values of $\ln A$ which are obtained from the MC values for $X_{\max}$ and R_μ.

the two extremes, iron and proton. A binary decision at a fixed value of $\ln A$ would result in more than 85% of each type of CRs being correctly identified as proton or iron. The result of a binary decision test is given in Table 6.3. The two distributions, however, are very different in shape, with proton showers sometimes being reconstructed at largely negative values of $\ln A$, down to -7. The distribution for iron showers is significantly narrower. The lightest 20% of events in an 1:1 proton-iron scenario therefore almost exclusively consist of proton showers. The abundance of protons in the lightest quantile of data as determined by the Universality reconstruction and the base transformation method is depicted in Fig. 6.20 for all four types of primary particles and different composition scenarios. While the separation of proton and helium is rather poor for any ratio of the two primary particles in the data, an ∼80% pure proton sample can be identified as the 10% of data with smallest reconstructed $\ln A$ even in an 1:10 proton to iron scenario.

The figure of merit,

$$f_{\mathrm{M}} = \frac{|\mu_1 - \mu_2|}{\sqrt{\sigma_1^2 + \sigma_2^2}}, \tag{6.36}$$

is a useful tool to estimate the separation of two normal distributions with means μ_i and widths σ_i. The data depicted in Fig. 6.18 (a) are not normally distributed, but from the

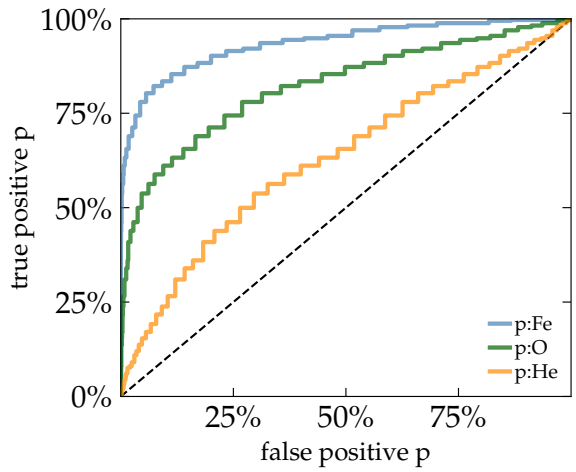
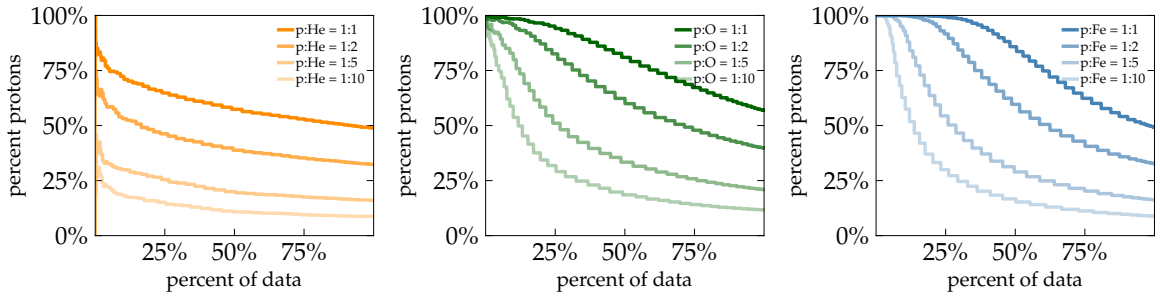


Figure 6.19: ROC curve for the discrimination between proton and helium, oxygen, and iron. The binary decision is based on a cut on reconstructed ln A. The ROC is independent on the mixture of the primary particles.



(a) (b) (c)

Figure 6.20: Proton purity as a function of the size of the lower quantile of the data. Different relative abundances of primary particle mixture ratios are considered for each panel.

respective overlap an equivalent figure of merit for two corresponding normal distributions can be calculated[6]. Two normal distributions with the same overlap as the distributions of ln A depicted in Fig. 6.18 (a) for proton and iron yield a figure of merit of 1.69. For proton and oxygen the equivalent figure of merit evaluates to 1.27, and for proton and helium to 0.27. The separation of proton and helium is thus relatively poor. The separation of proton and iron, on the other hand, is rather satisfying.

[6]The integral over the product of the distributions is evaluated. Then two normal distributions at the same distance with a variable width are considered for the same integral. The width is then set accordingly, so that this integral evaluates to the same value as for the distributions of reconstructed values for ln A.

Table 6.3: Results of a binary classification of proton- and iron-induced showers. The binary classification is applied on the data depicted in Fig. 6.18 using the cut value ln $A = 2.29$.

	p_{MC}	Fe_{MC}
p_{rec}	85.1%	14.9%
Fe_{rec}	13.3%	86.7%

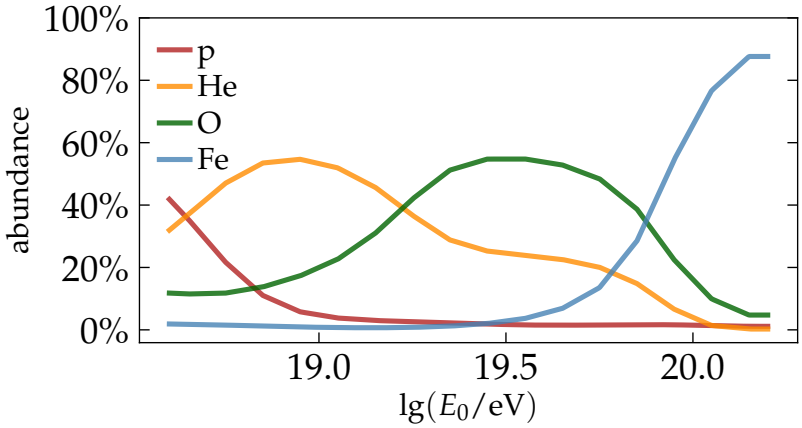


Figure 6.21: Relative abundance of proton, helium, oxygen, and iron nuclei in the CR particle spectrum to match the Auger-mix, which is obtained from the results given in Ref. [151].

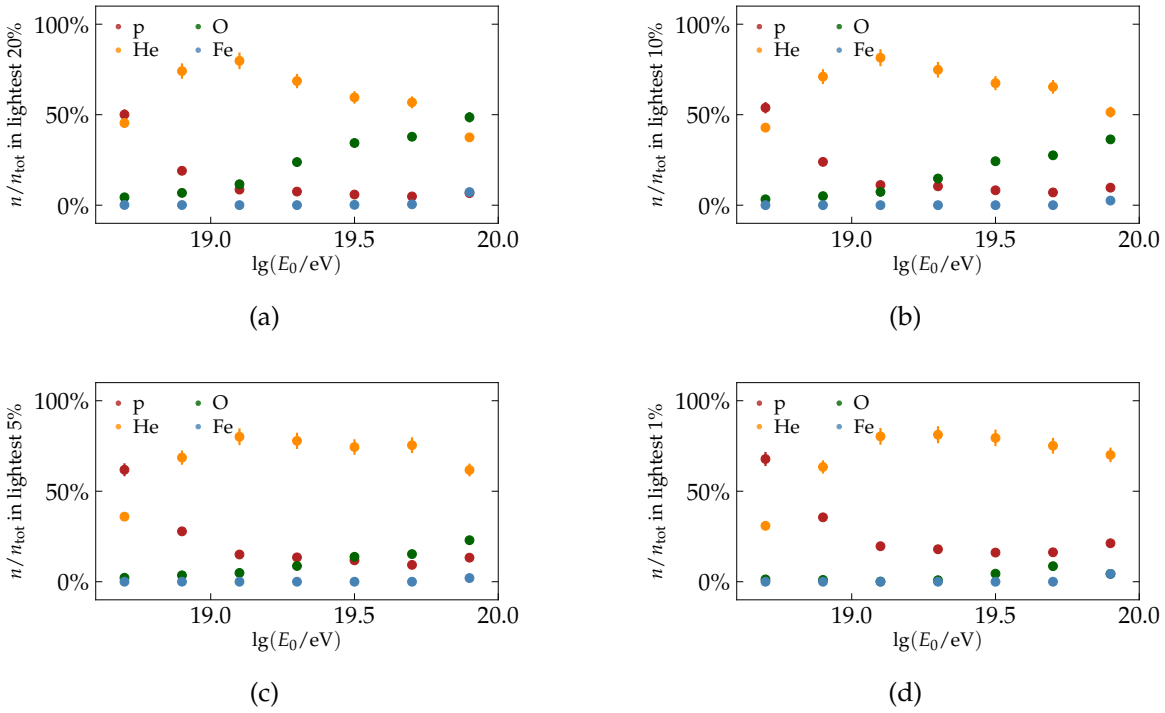


Figure 6.22: Fractional abundance of elements in the Auger-mix data, for different quantile cuts of the reconstructed ln A.

To quantify the ability to identify light CRs in a realistic scenario, a toy MC experiment can be performed using the library of simulated CRs. A data driven result from a combined fit of X_{max} data and the all-particle spectrum was published in Ref. [151]. Using tabulated data from this analysis, the abundance of CRs of each atomic mass can be interpolated to find a realistic composition scenario using only the four types of primary particles considered in the simulation library which is used in this chapter. The result, which is the *Auger-mix*, is given in Fig. 6.21. The utilized tabulated data is given in steps of 0.1 in $\lg(E_0/\mathrm{eV})$ and steps of approximately 1.1 in $\ln A$. Since the simulation library unfortunately does not contain oxygen nuclei at the highest energies, only primary energies below 10^{20} eV are considered.

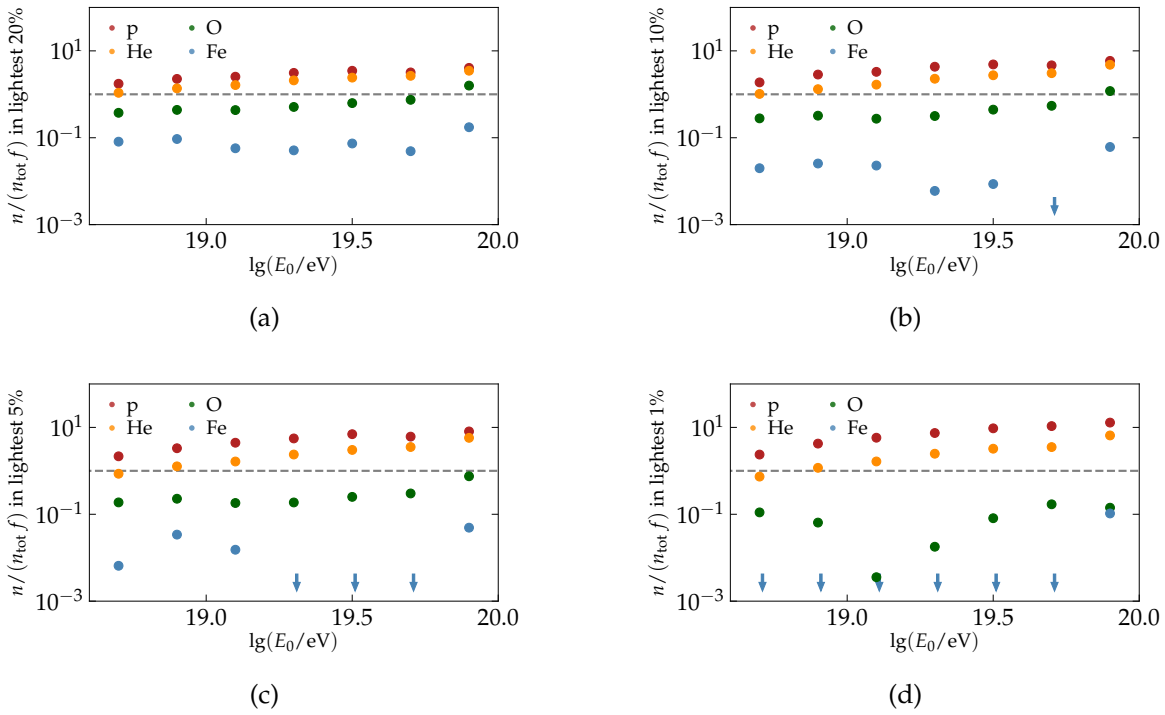
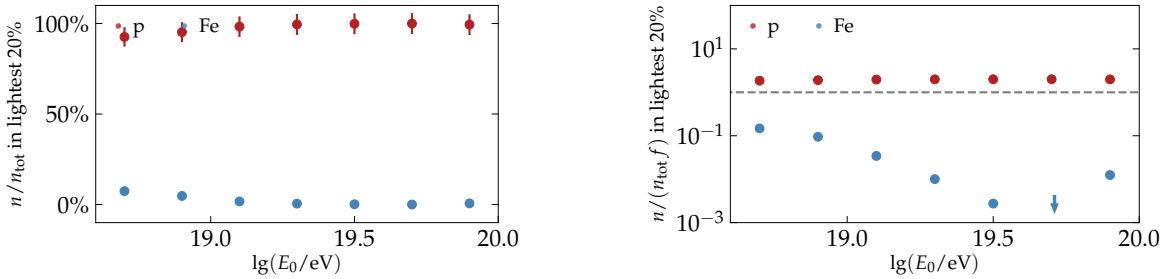


Figure 6.23: Energy-dependent amplification or suppression of elements in the Auger-mix. Lightest events are selected by the Universality reconstruction method.

In a range between $18.6 < \lg(E_0/\mathrm{eV}) < 20.0$ a number of CRs is randomly selected from the simulation library according to the Auger-mix depicted in Fig. 6.21. Approximately 800 events are available for each type of primary particle in each energy bin. From the reconstructed values of X_{max} and R_μ, the mass number $\ln A$ is calculated for each event according to the method outlined in Section 6.4. A fraction of CRs which are identified as the "lightest" is selected from the data set. In this set of n_{tot} lightest events we examine the abundance of the four initial elements, of which n events are identified as part of the lightest quantile, each. This process is repeated 1000 times on the available simulation library, from which the average abundance of each element in the lightest quantile of events can be calculated. The results are depicted in Fig. 6.22. At the highest energies, where the composition is considered extremely iron dominated and only ∼1% of CRs are expected to be protons [151], ∼10% of events in the lightest 10% of data are protons, and ∼20% of the the lightest 1% of data are protons. At lower energies, where more protons are present in the total data, almost 50% of the events identified as the lightest are proton CRs. At the same time, the abundance of iron is extremely suppressed. Even though the composition is substantially iron dominated, iron is almost excluded in the lightest fraction of events at every energy.

The relative abundance of each element of course highly depends on the particle spectrum of CRs arriving at Earth[7]. However, the theoretical amplification or suppression of the abundance of each element in the light quantile of events can be examined. It is given by the relative abundance of events for each primary, n/n_{tot}, divided by the fraction f of CRs that is expected for the primary at the respective energy according to the Auger-mix. The results are given in Fig. 6.23. On average proton is approximately 5 times more abundant in those events that are identified as the lightest, as in the total set of events. At the highest energies, in the 1% of the lightest events a CR is more than 20 times more likely to be a proton than in the whole data set. At the same time, iron is suppressed by approximately a

[7] If there are no proton CRs at the highest energies, none can be identified by the Universality reconstruction.



(a) (b)

Figure 6.24: Energy-dependent relative abundance and amplification or suppression of elements in a 1:1 proton-iron scenario. Lightest events are selected by the Universality reconstruction method.

factor of 1000 at almost all energies in the depicted regions.

The same experiment with a 1:1 mixture of proton and iron and no intermediate elements yields even more promising results. Those are given in Fig. 6.24. In this case, already the 20% of lightest events consist almost exclusively of proton CRs, as it can also be seen in Fig. 6.20.

CHAPTER VII

RESULTS ON DATA

In this chapter, we present results of the Universality reconstruction method applied to the data from the Auger SD and AugerPrime detector.

At first the reconstructed observables are examined if data driven-corrections are necessary. Afterwards we present the results of the reconstruction algorithm discussed in Chapter 6. The presented results include the reconstruction of X_{max}, R_μ, and based on them the estimate of $\ln A$ for CRs detected by the Pierre Auger Observatory. One particular CR identified as light particle is discussed in detail as an example, and the arrival directions of a selection of high-rigidity events are presented.

7.1 DATA DRIVEN CORRECTIONS

Similarly to the analysis in Section 6.2.5 we examine the systematic variation of the reconstructed observables with respect to physical parameters. Especially the dependence of the observables on the zenith angle θ is crucial and we must ensure that there is no zenith-dependent bias in the data. Furthermore, possible effects of the aging of the detectors, the azimuthal arrival direction of the CR, and the season are examined.

7.1.1 ZENITH-DEPENDENT BIAS

Very similarly to the bias correction introduced in Section 6.2.5, the reconstructed values of X_{max} and R_μ are not perfectly constant with respect to the zenith angle. The average reconstructed values for X_{max} and R_μ and the resulting $\ln A$ as a function of the zenith angle are depicted in Fig. 7.1. Albeit different in terms of absolute values, the bias with respect to the average reconstructed values as seen in data shows qualitatively the same behaviour as the bias experienced in simulations, which is depicted in Fig. 6.6. The values of the zenith-dependent bias corrections are given in Table 7.1.

Table 7.1: Data-driven progressive corrections to the reconstructed values of X_{max} and R_μ as functions of $\sin\theta$ and $\lg(E_0/\mathrm{eV})$. Corrections are applied according to Eq. (6.14) with an additional cubic term.

		a_0	a_1	a_2	a_3
$\delta X_{\mathrm{max}}/(\mathrm{g\,cm^{-2}})$	$\sin^2\theta$	17.4	22.9	-589	865
	$\lg(E_0/\mathrm{eV})$	12.0	-	-	-
δR_μ	$\sin^2\theta$	-0.029	-1.17	6.98	-8.537
	$\lg(E_0/\mathrm{eV})$	1.456	0.075	-	-

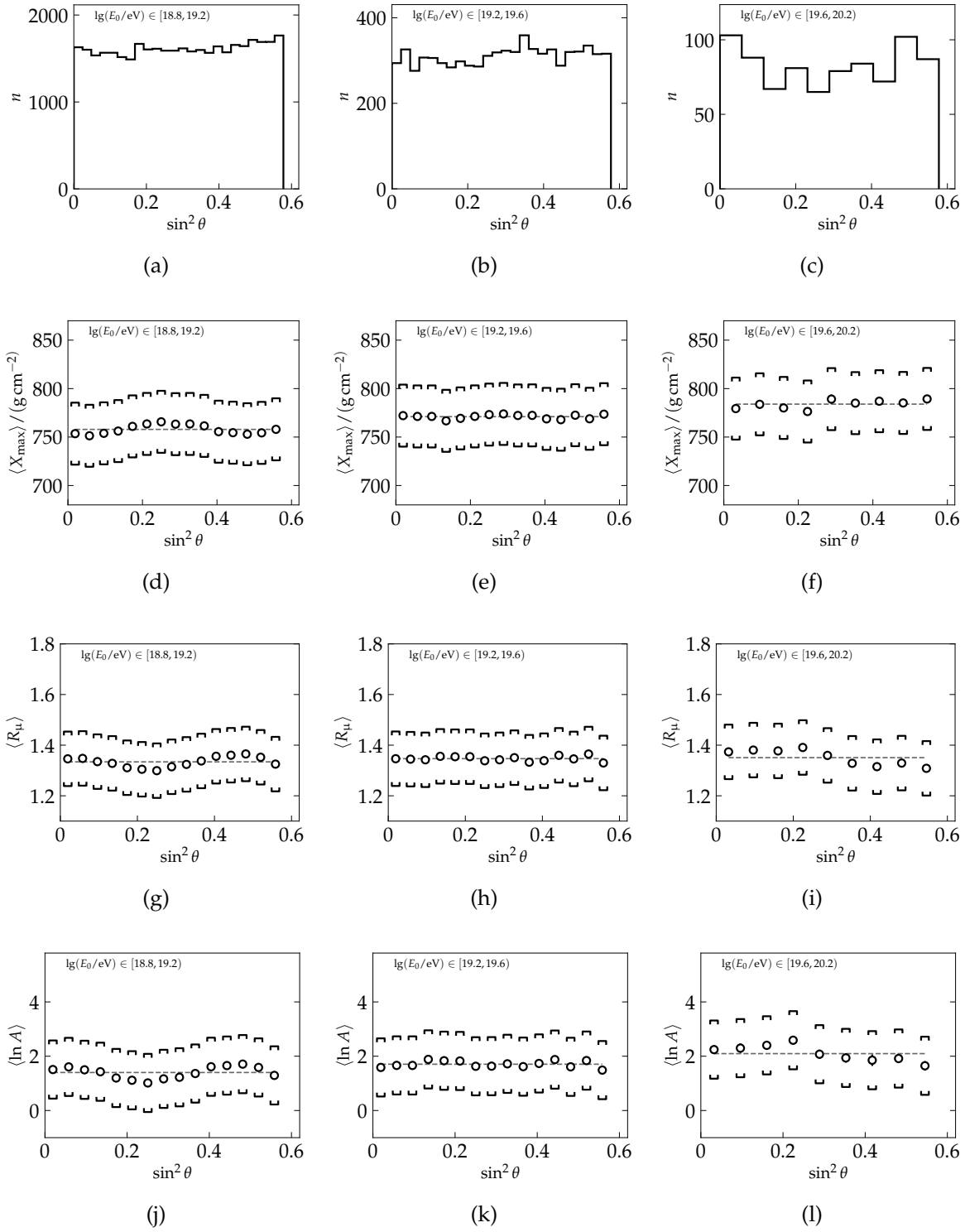


Figure 7.1: Zenith dependence of the reconstructed observables after correction. The depicted data comprise all reconstructed events of the SD data. The respective energy range is given in the upper left corner of each panel. The respective average values are depicted as a reference with a dashed line.

A possible explanation for the zenith-dependent bias in the uncorrected data is the reconstruction of the geometry of the event. The assignment of R_μ heavily depends on the signal as a function of the shower-plane radius of each detector, thus, if the position of the

shower core at the ground is incorrectly reconstructed, also R_μ is shifted. Another possible source for the zenith dependence of the uncorrected values of R_μ is the reconstructed energy that depends on the CIC method (see Section 4.2). The CIC method does not take into account the fact that heavier CRs, which produce significantly more muons than lighter CRs, are more likely to be detected by the Auger SD at large zenith angles. Details on the reconstruction of the geometry and the primary energy are given in Ref. [97]. The reconstruction of the event geometry does also effect the reconstruction of $X_{\max}$, since the timing of the signals, which govern the reconstruction procedure for the depth of the shower maximum, is measured relatively to the time of the plane front (see Section 6.1.2). The time of the plane front, however, is calculated for each station based on the reconstructed geometry of the event. A difference in 50 m in the reconstructed core position, for example, thus results in a change of the plane-front time of approximately 150 ns, which could produce a significant shift of the best-fit value of $X_{\max}$ for the individual detectors (see Fig. 5.10).

Below energies of $\lg(E_0/\mathrm{eV}) \lesssim 19.0$ a zenith-dependent bias remains for both the reconstructed values of $X_{\max}$ and R_μ. On average for all events, however, this bias cancels out. Results on the average values of $X_{\max}$ and R_μ are thus not affected. An event-by-event estimation of the mass of the primary CR, however, could be affected by the residual biases below energies of 10^{19} eV.

7.1.2 AREA OVER PEAK

The aging of the PMTs of the WCDs results in a change of the ratio of the signal peak with respect to the total charge deposited by a vertically through-going muon. This effect and its widespread implications are addressed in a series of internal publications discussed in Ref. [152]. Especially small signals cannot be recovered in an SD station, if the aging of the electronics has progressed to far. To verify that this effect has only a minor influence on the reconstructed values of $X_{\max}$, R_μ, and consequently $\ln A$, the reconstructed observables are examined as functions of the average values of the *Area over Peak* (AoP) of all PMTs that were used in the reconstruction of the event. The results are depicted in Fig. B.21. The average predictions are very constant within their own uncertainty if the number of events per bin is sufficiently. No large systematic drift of the reconstructed observables was found in the data as a function of AoP. This is most probably due to the fact that stations with small signals, which are the most affected by the aging of the electronics, are not considered for the reconstruction. Thus, no AoP-dependent correction is applied.

7.1.3 AZIMUTH

Systematic variations of the reconstructed observables might arise for different azimuth arrival directions, ϕ. This could be due to the effect of Earth's magnetic field or because of the topology of the SD array, that is slightly tilted[1]. CRs with $\phi = 0$ are arriving at the detector from the East, $\phi = \pi/2$ corresponds to CRs arriving from the South. Only at energies below $\lg(E_0/\mathrm{eV}) \lesssim 19.2$ a sinusoidal modulation with a frequency of 1 and an amplitude of $2.6\,\mathrm{g\,cm^{-2}}$ was found for the reconstructed values of $X_{\max}$. Besides that no sinusoidal modulation with respect to ϕ was found in the data for higher energies. Thus, no azimuth-dependent correction is applied. The considered data is depicted in Fig. B.22.

7.1.4 SEASONAL VARIATION

The change in the average air pressure and the resulting variation in the atmospheric density ρ affect the calibration of the SD station detectors and the average reconstructed energy,

[1] The altitude of the array varies between ~1600 m to 1300 m above sea level.

if not properly taken into account [153]. The Universality parametrization and reconstruction presented in this work makes use of a flexible model of the atmospheric density that takes into account the measurements of the atmospheric conditions at the Observatory site. The reconstructed values of X_{max} and R_μ, and consequently also of $\ln A$, show no seasonal variation, as can be examined in Fig. B.23. Thus no further correction is applied.

7.1.5 SYSTEMATIC UNCERTAINTIES

In this section, the underlying systematic uncertainties regarding the reconstruction of X_{max} and R_μ are briefly estimated. The systematic uncertainty on the reconstructed values of $\ln A$ is estimated from the systematic uncertainties affecting X_{max} and R_μ using Eqs. (6.27) and (6.29).

The most basic underlying systematic uncertainty arises from the MC data, which was discussed in Section 6.2. It can be conservatively estimated to 0.04 for R_μ and $5\,\mathrm{g\,cm^{-2}}$ for X_{max}. The choice of the hadronic interaction model, with which to parametrize and run the reconstruction procedure, significantly affects the outcome of the reconstructed values of X_{max} for each shower. As depicted in Fig. 4.7, the different hadronic interaction models produce an absolute offset of up to $15\,\mathrm{g\,cm^{-2}}$ (the difference of the average X_{max} is about $12\,\mathrm{g\,cm^{-2}}$ for proton showers and $15\,\mathrm{g\,cm^{-2}}$ for iron showers, between primary energies of $10^{18}\,\mathrm{eV}$ and $10^{20}\,\mathrm{eV}$.). The normalization of the different contributions to the likelihood function discussed in Section 6.1.1 can shift the reconstructed values of X_{max} for single events. Its contribution to the systematic uncertainty is very conservatively estimated to $12\,\mathrm{g\,cm^{-2}}$ from a small set of showers which was reconstructed using different, extreme values for the described fudge factor. Furthermore, the applied constant and zenith-dependent bias correction of X_{max} reaches $\sim 23\,\mathrm{g\,cm^{-2}}$, especially for very vertical showers.

The systematic uncertainty of the reconstructed values of R_μ due to the muon deficit and the differences in hadronic interaction models is treated with the calibration described in Section 6.4.4 and thus does not further affect the reconstruction of $\ln A$. The applied bias correction to R_μ is estimated to 0.05 on average for the whole parameter space. The by far dominant contribution to the systematic uncertainty of R_μ arises from the uncertainty of the reconstructed energy by the SD detector. At high energies, σ_{E_0} is estimated to about 7% [100]. This directly translates into a systematic uncertainty of R_μ of 7%. The reconstruction of X_{max} is not affected by the systematic uncertainty of the reconstructed energy.

The individual contributions are listed in Table 7.2. The estimated propagated uncertainty of $\ln A$ from the individual contributions is given together with the overall systematic uncertainty of $\ln A$ that arises from the total uncertainties of X_{max} and R_μ. The resulting systematic uncertainty of $\ln A$ is conservatively estimated to $\sigma_{\ln A} \simeq 1.0$.

7.2 EXAMPLE EVENTS INCLUDING SSDS

The current detector upgrade for AugerPrime is not yet completed, some events, however, were already recorded triggering both WCDs and SSDs. In the following three examples are briefly presented. Two are selected randomly from the data, whereas the third is an exemplary high-energy event that triggered the infill of the array and is thus recorded by an extraordinary large amount of detectors.

The selected events are depicted in Fig. 7.2. For each event, the lateral distribution of signal in the shower plane as well as the time-dependent signal of one SSD is shown. In Fig. 7.2 (c) data is depicted from the event 191341958500 detected on 14 March 2019. The event triggered numerous recently installed SSD detectors, providing a successful test of the upgraded surface-detector array. A depiction of the signal recorded from event

191341958500 along with the SSD signal components from the Universality model is given in Fig. B.24.

When examining the whole data set, the events including stations with SSD signal are treated the same as events only recorded by WCDs, even though the systematic uncertainty on the reconstructed primary energy should be significantly smaller for these events. The available SSDs were included in the standard reconstruction procedure according to the work given in Ref. [145].

7.3 RESULTS FROM THE SD DATA SET

In the following we present the results from the Universality reconstruction on the SD data set of Auger from the beginning of data acquisition in 2008 until December 2021. Values for X_{max}, R_{μ}, and $\ln A$ are obtained as described in Chapter 6. Results obtained from the FD measurements as given in Ref. [154] are used for comparison.

Only SD events that fulfill the 5T5Pos2 trigger condition[2] are selected for the data discussed in this chapter. Events up to a reconstructed zenith angle of $\theta = 50^{\circ}$ are considered. The results on the reconstructed average values of X_{max} and R_{μ} are given in Fig. 7.3. Anal-

[2]The T5 status of a station affirms that the station works and is taking data. A 5T5 condition triggers on events with a "hottest" station that is closest to the position of the reconstructed shower core at the ground and is surrounded bz at least 5 working stations. The 5T5Pos2 condition imposes that the position of the reconstructed shower core at the ground has to be inside of a triangle of three neighbouring, functioning stations. A detailed analysis of the 5T5 trigger condition and implications on the reconstructed energy and geometry is given in Ref. [155].

Table 7.2: Estimated systematic uncertainties of the reconstructed observables. Numbers indicated with asterisks are not taken into account for the estimation of the total uncertainty.
* Uncertainty is calculated using Eq. (6.27) for $\varphi_0 \to 0$ or $\varphi_0 \to \infty$.
** Uncertainty is treated with the calibration outlined in Section 6.4.4.

	X_{max}	R_{μ}	$\ln A$
MC	$5\,\mathrm{g\,cm^{-2}}$	0.04	0.2
hadronic interactions	$15\,\mathrm{g\,cm^{-2}}$	0.26**	0.7*
likelihood normalization	$12\,\mathrm{g\,cm^{-2}}$	-	0.5*
bias correction	$23\,\mathrm{g\,cm^{-2}}$	0.05	0.5
SD energy resolution	-	0.09	1.2*
total	30	0.11	1.0

Table 7.3: Details of the reconstruction of the selected example events including SSDs detectors.

	(a)	(b)	(c)
date	2021-03-10	2021-03-13	2019-05-14
Auger Id	210693417400	210726008300	191341958500
SD Id	62456011	62506763	53431864
$\lg(E_0/\mathrm{eV})$	18.67	18.68	19.86
θ	13.4°	31.0°	45.0°
X_{max}	$708\,\mathrm{g\,cm^{-2}}$	$739\,\mathrm{g\,cm^{-2}}$	$791\,\mathrm{g\,cm^{-2}}$
R_{μ}	1.23	1.20	1.33

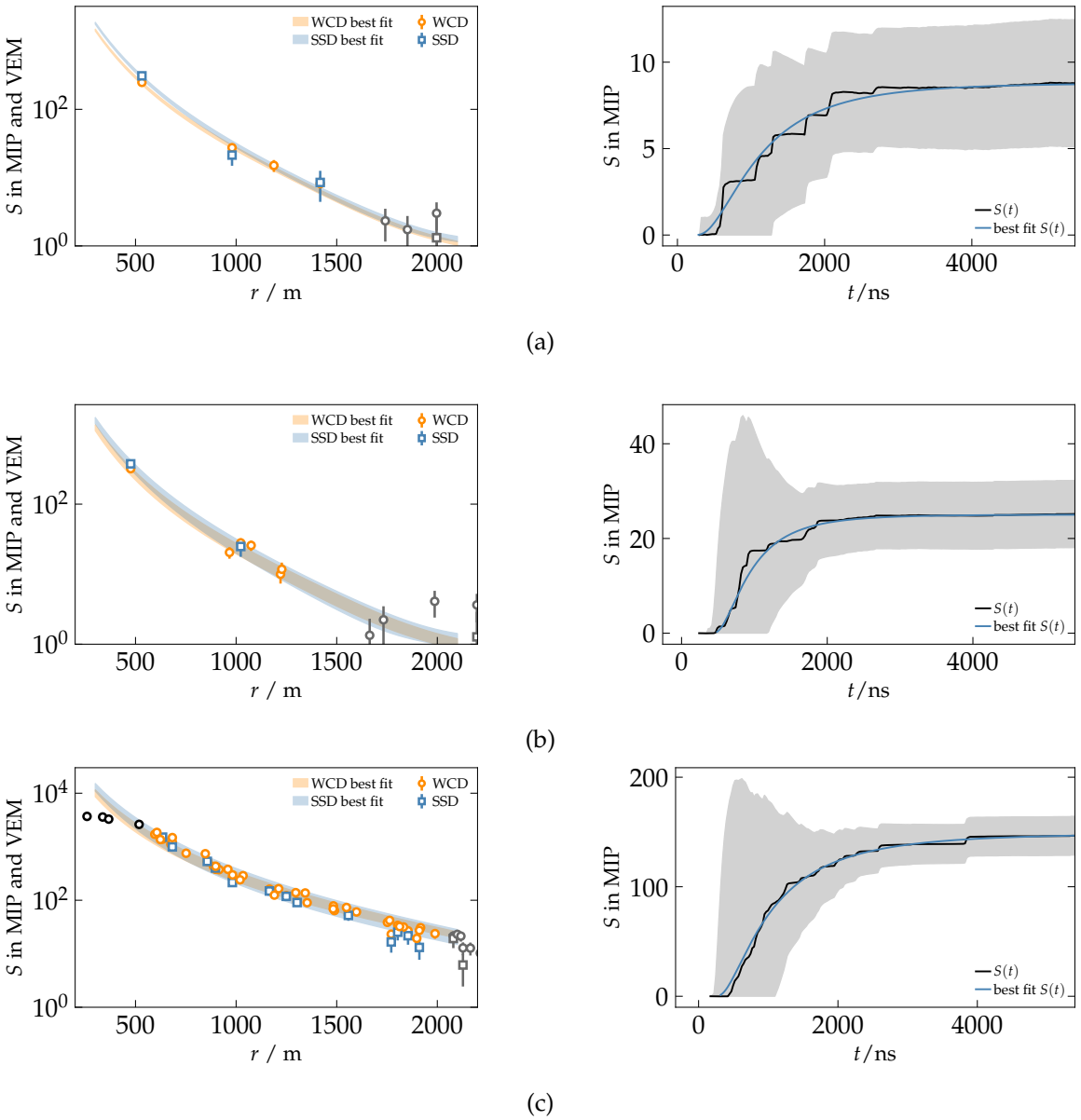


Figure 7.2: Example events including SSD detectors. Each row shows one event with the lateral distribution of signal (*left*) and the time-dependent signal of one SSD (*right*). Black and gray markers represent saturated and low-signal detectors, respectively.

ogously to Figs. 7.3 and 7.4, the same data obtained only from the events that triggered at least two SSDs are given in Figs. B.25 to B.27. The average reconstructed values of $\ln A$, which are calculated from $X_{\max}$ and R_μ for each event, are given in Fig. 7.4. The corresponding standard deviation of the data is given in Fig. 7.5. Additionally, for all depicted energy bins the distributions of the reconstructed values of $\ln A$ are given in Figs. 7.6 and 7.7.

The average $X_{\max}$ as a function of the primary energy as well as the decadal elongation rate as obtained in this work agrees well with the result from Ref. [154]. Between primary energies of $19.1 < \lg(E_0/\mathrm{eV}) < 19.7$ the elongation rate evaluates to $28\,\mathrm{g\,cm^{-2}/decade}$. Above these energies the elongation rate increases to $41\,\mathrm{g\,cm^{-2}/decade}$. The average $X_{\max}$ in each energy bin is slightly different from the FD results. This could be due to the fact that different events contribute to the resulting average values obtained from FD and the Universality reconstruction. At lower energies a systematic discrepancy between FD and

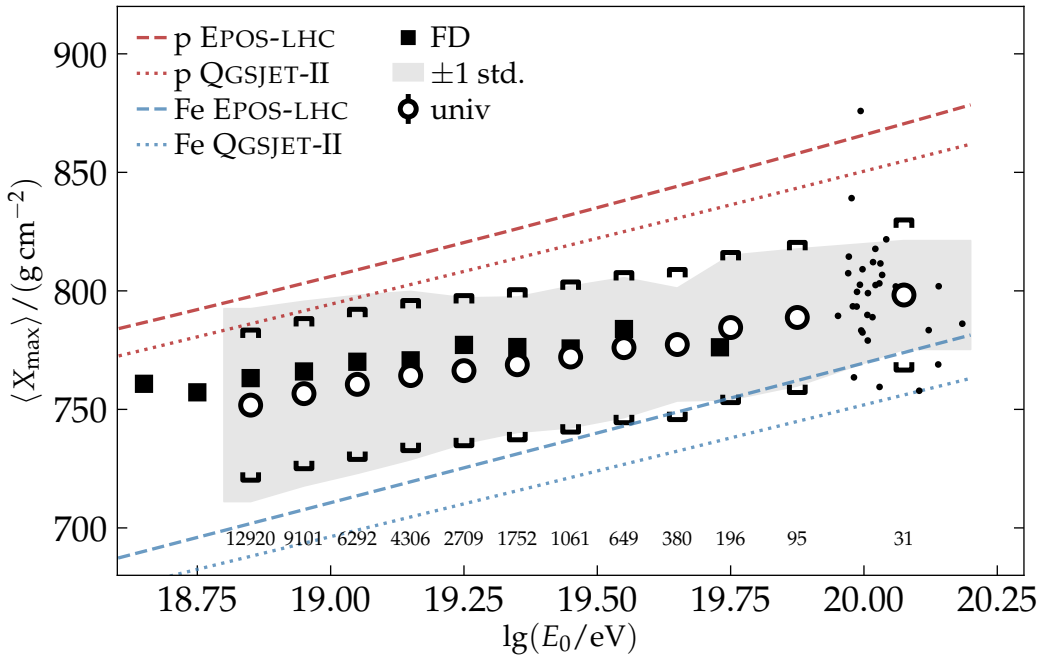
the Universality reconstruction is apparent. Even though the reconstructed values of X_{max} are on average increasing with energy, the relatively low elongation rate disfavours constant compositions and instead favours a change from light towards a heavier CR composition up to $E_0 \simeq 10^{19.7}\,\mathrm{eV}$.

The results on the reconstructed values of R_μ yield similar implications. Even though the average values for R_μ do not match the expectations from simulated air showers (see Section 4.4.2), a clear trend towards a heavier composition is visible between primary energies of $10^{18.8}\,\mathrm{eV}$ and $10^{19.7}\,\mathrm{eV}$. If the reference values of R_μ are rescaled by approximately 30% (see Section 6.4), the reconstructed mean values of R_μ are very well within expectations. Above $E_0 \simeq 10^{19.8}\,\mathrm{eV}$ the average value of R_μ is becoming lower, favouring a lighter composition. This trend towards a lighter composition is less significant but also present in the average reconstructed values of X_{max}.

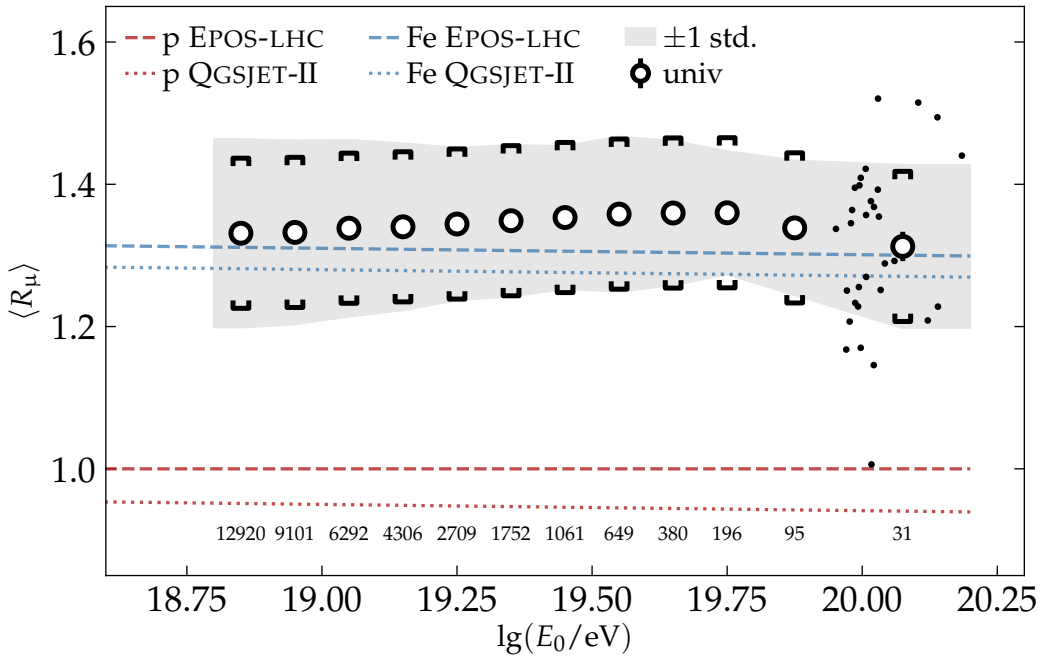
The average reconstructed values of $\ln A$ from the combined information of both X_{max} and R_μ of each event match the result obtained in Ref. [154]. Even though $\ln A$ as reconstructed in this work is slightly lower, the trend of $\langle \ln A \rangle$ with respect to energy was confirmed. The estimated average $\ln A$ resembles the implications of Fig. 7.3 of an increasingly heavier composition up to $E_0 \simeq 10^{19.7}\,\mathrm{eV}$.

The second moment of the distribution of the reconstructed values of $\ln A$ allows for the same interpretation. The standard deviation of the distribution decreases with the primary energy up to $\sim 10^{19.7}\,\mathrm{eV}$. From energies around $10^{19.7}\,\mathrm{eV}$ to the highest the highest energies, the spread of the distribution of $\ln A$ increases slightly from 1.3 to 1.4, indicating a broader mix of the CR composition. The increase, however, is not significant. The standard deviations of the reconstructed values of X_{max} and R_μ are given in Fig. B.28.

A pure iron composition at the highest energies seems unlikely according to both the first and second moment of the distribution in reconstructed $\ln A$. The average logarithmic mass of CRs above an energy of $10^{19.5}\,\mathrm{eV}$ is approximately $\langle \ln A \rangle = 2$ with a width of the distribution of ~ 1.4.



(a)



(b)

Figure 7.3: Average reconstructed values of X_{max} in panel (a) and R_μ in panel (b) as a function of the SD-reconstructed primary energy are shown as circular markers. The mean values for iron and proton from two different hadronic interaction models are shown for reference as dashed lines. The numbers above the x-axis indicate the amount of events in the respective energy bin. The events in the highest-energy bin are depicted also as individual dots. The standard deviation of the underlying data distribution is depicted as a gray band. Systematic uncertainties are shown as brackets. Results from the FD reconstruction of Ref. [154] are given as black squares in the panel (a).

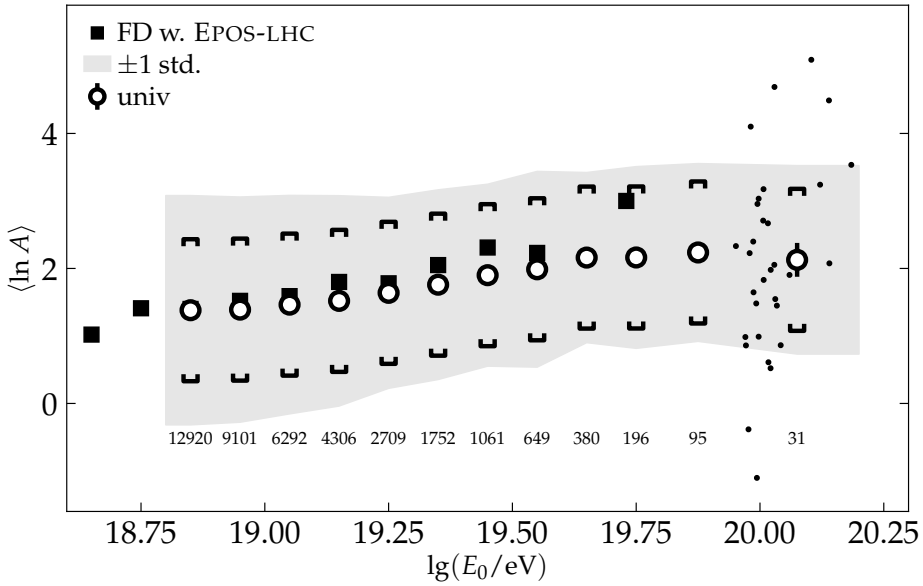


Figure 7.4: Average reconstructed values of ln A as a function of the SD-reconstructed primary energy are shown as circular markers. The numbers above the x-axis indicate the amount of events in the respective energy bin. The events in the highest-energy bin are depicted also as individual dots. The standard deviation of the underlying data distribution is depicted as a gray band. Systematic uncertainties are shown as brackets. ln A obtained from X_{max} reconstructed by FD and converted using the EPOS-LHC model as a reference [154] is indicated as black squares.

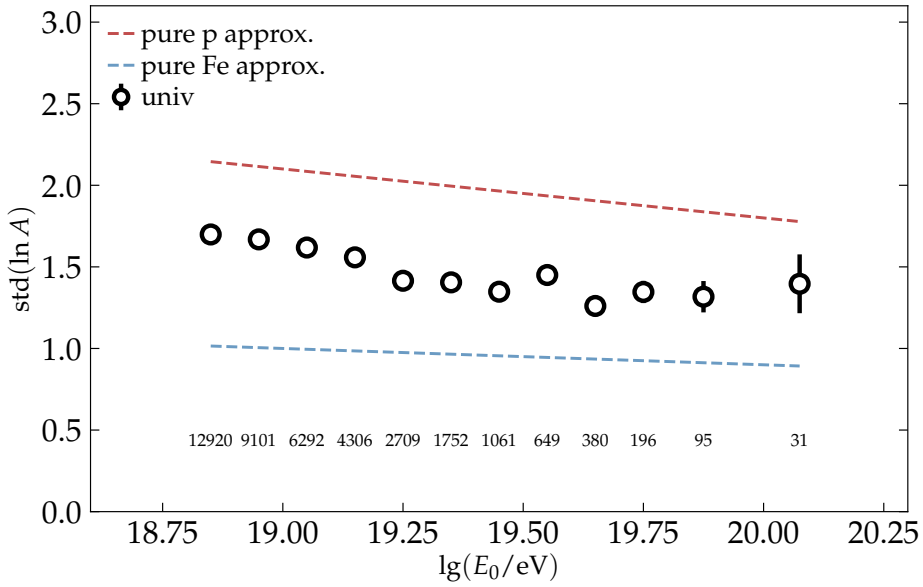


Figure 7.5: Standard deviation of the distribution of the reconstructed values of ln A as a function of the SD-reconstructed primary energy, shown as circular markers. The numbers above the x-axis indicate the amount of events in the respective energy bin. The standard deviation that is expected for a pure proton and iron composition is approximated by the root mean square of the precision given in Fig. 6.17 and the shower-to-shower fluctuations, given in Eq. (6.32).

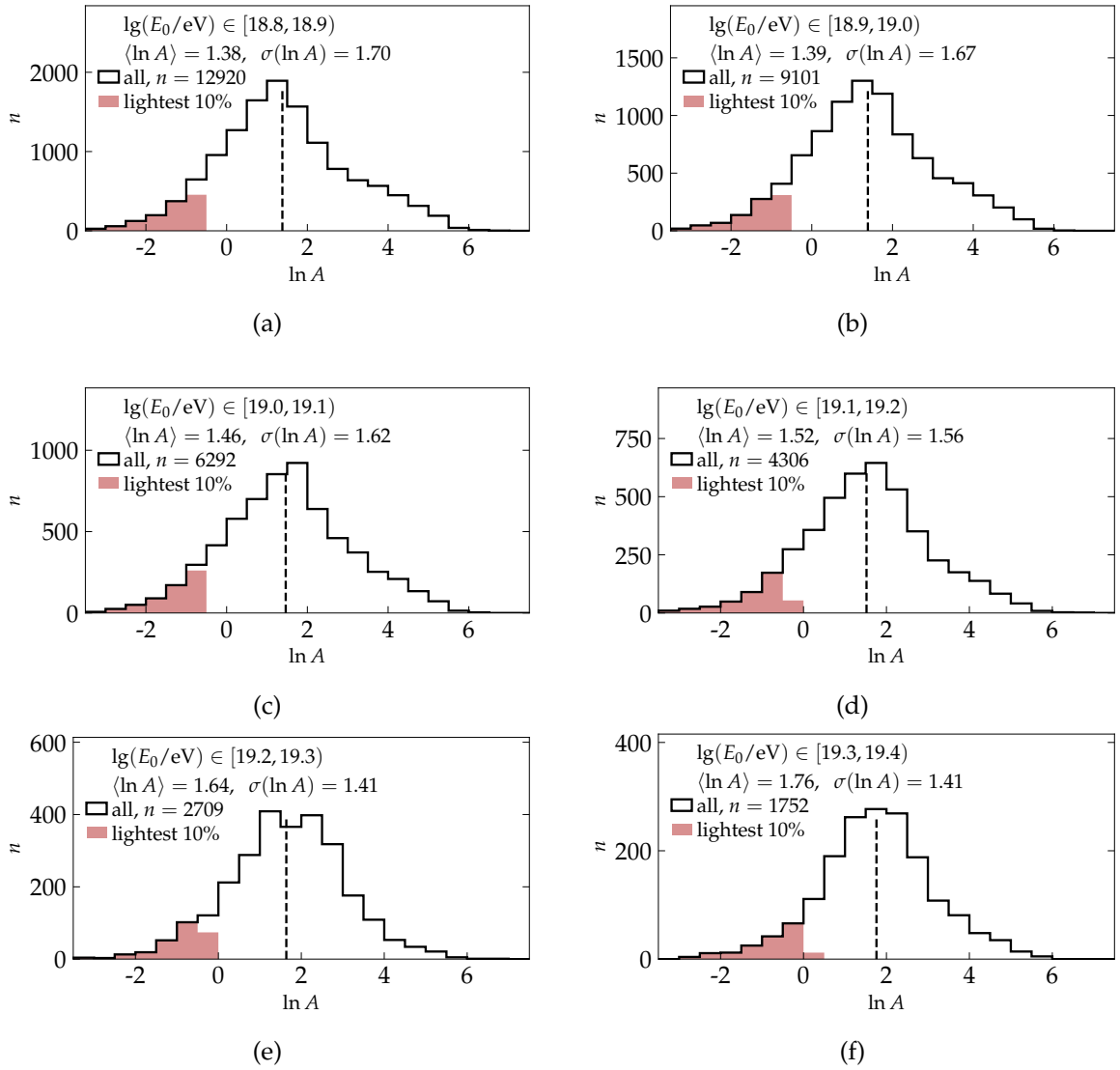


Figure 7.6: Distributions of the event-level estimate of $\ln A$ provided by the Universality reconstruction for the events with a reconstructed primary energy between $18.8 \leq \lg(E_0/\mathrm{eV}) < 19.4$. The considered energy interval is given in the upper-left corner of each panel, as well as the respective mean value and standard deviation of the distribution. The mean value for each panel is indicated with a dashed line. The 10% of events which are identified as lightest by $\ln A$ are shaded in red. The number of events in each panel is given by n.

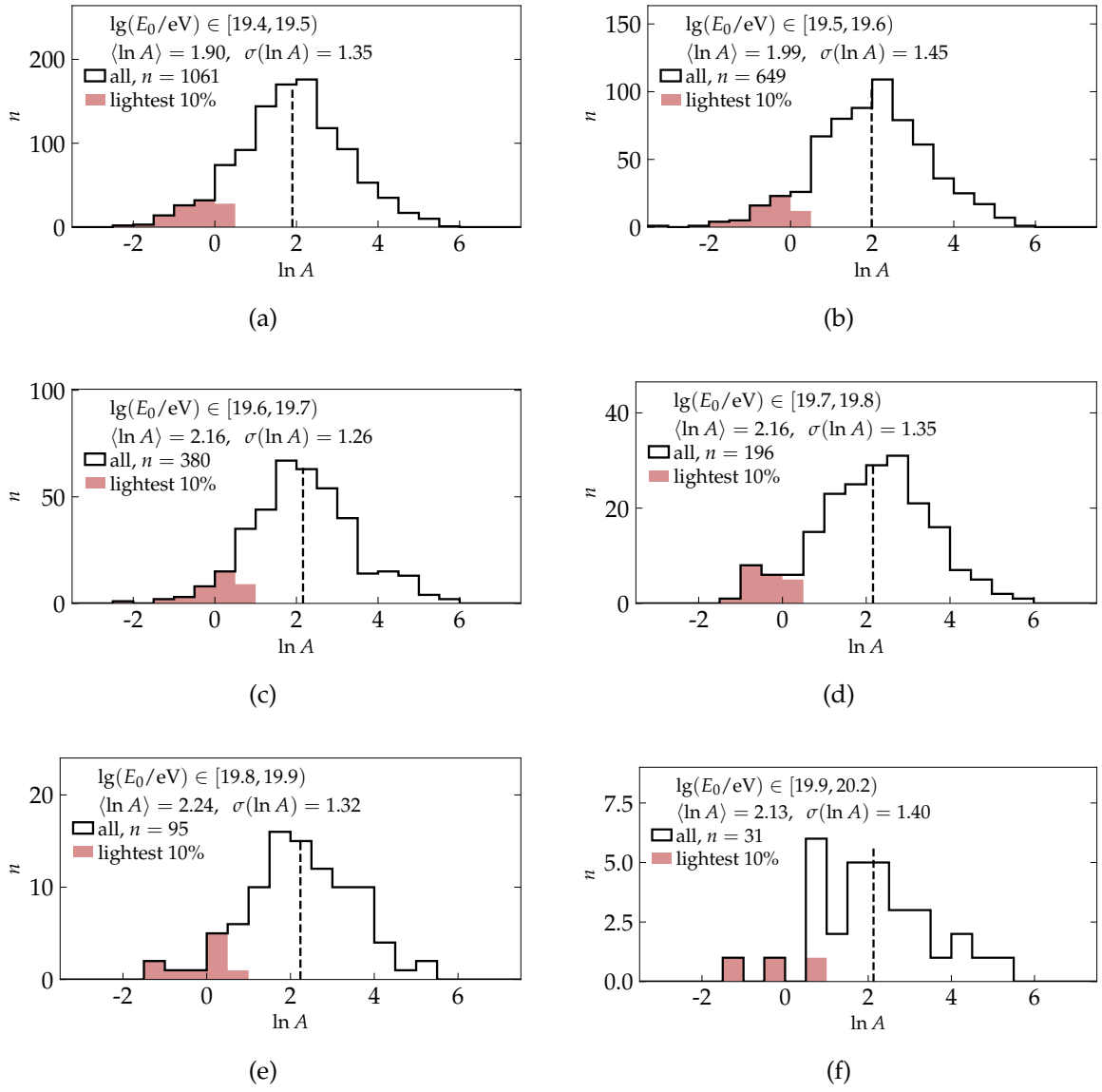


Figure 7.7: Distributions of the event-level estimate of $\ln A$ provided by the Universality reconstruction, for the events with a reconstructed primary energy between $19.4 \leq \lg(E_0/\mathrm{eV}) < 20.2$. The considered energy interval is given in the upper-left corner of each panel, as well as the respective mean value and standard deviation of the distribution. The mean value for each panel is indicated with a dashed line. The 10% of events which are identified as lightest by $\ln A$ are shaded in red. The number of events in each panel is given by n.

7.4 AN EXAMPLE OF A LIGHT UHECR

In this section, we present an example event which is identified as light high-energy CR by the methods presented before.

A CR was recorded on the 23 May 2015 with a reconstructed energy of almost 10^{20} eV. The Universality reconstruction reports a relatively low muon number of $R_\mu \simeq 1.3$, whilst simultaneously a deep shower maximum at $876\,\mathrm{g\,cm^{-2}}$. Both individually indicate that the CR is light. The reconstruction of the event is given in detail in Table 7.4. The combined

information of X_{max} and R_μ results in a reconstructed value for $\ln A$ of -1.3, which within its systematic uncertainty implies that the primary particle was a light particle such as a proton or a helium nucleus. Because of its arrival direction which is close to the position of the large Magellanic cloud, the event is nicknamed *Magellan*[3]. The deposited signal by Magellan is depicted in Fig. 7.8 as a function of the distance of the detectors to the reconstructed shower axis. Magellan triggered a total of seven stations that were considered for the Universality reconstruction and several additional stations that were disregarded because of their large distance to the reconstructed shower axis or because of saturation of the electronics. Because the event was recorded in 2015, only WCDs are present in the set of triggered stations. The time-dependent signal of four stations is depicted in Fig. 7.9, together with the respective prediction from the Universality model.

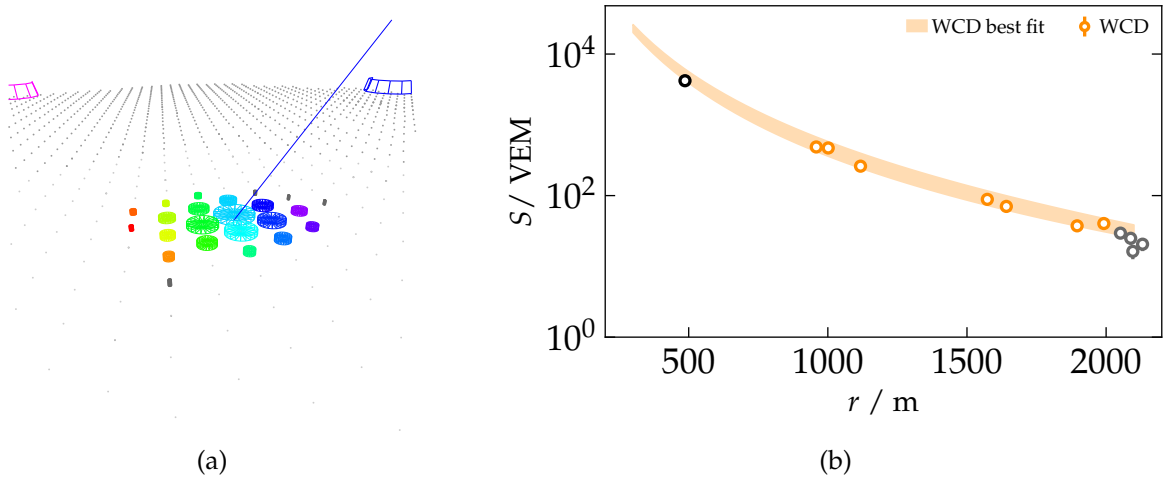
An alternative method to reconstruct the depth of the shower maximum using a deep neural network [147] reports a shower maximum at the depth of $855\,\mathrm{g\,cm^{-2}}$ for this event, which agrees with the result of Universality reconstruction withing its uncertainty. Estimating $\ln A$ only from the information provided by X_{max} as reported by the deep neural network with EPOS-LHC as a reference results in $\ln A \simeq -0.2$.

Magellan is part of the October 2021 Auger Open Data release [156].

[3]The large Magellanic cloud is located at the galactic coordinates $b = -81^\circ$ and $l = -33^\circ$; it is not implied to be the source of this particle.

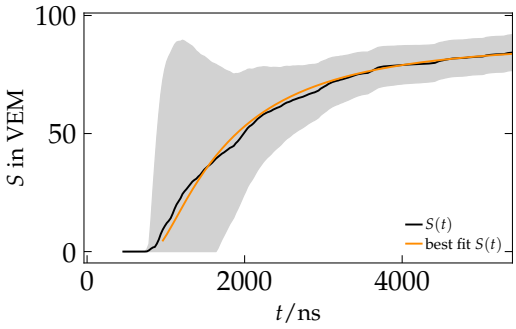
Table 7.4: Details of the reconstruction of event 151433402300. Reconstructed observables are obtained from the SD standard reconstruction and from the Universality reconstruction.

date	2015-05-23
Auger Id	151433402300
SD Id	33002920
GPS second	1116451638
$\lg(E_0/\mathrm{eV})$	19.994 ± 0.014
Auger polar (θ, ϕ)	$(39.7^\circ, 220.1^\circ)$
galactic (l, b)	$(-102^\circ, -33^\circ)$
X_{max}	$(875.9 \pm 30)\,\mathrm{g\,cm^{-2}}$
R_μ	1.26 ± 0.1
$\ln A$	-1.3 ± 1.0

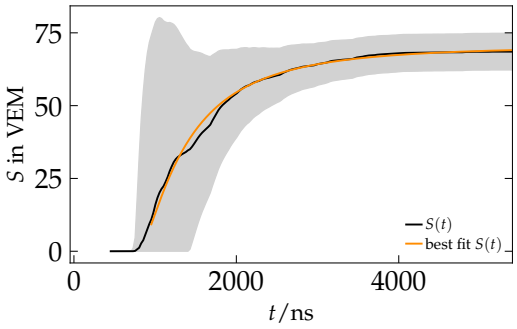


(a) (b)

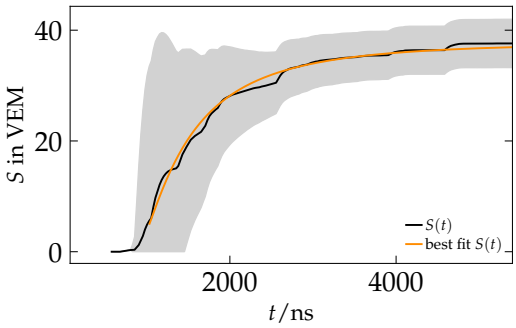
Figure 7.8: Event-display view and lateral distribution of signals from Magellan. The color scheme in panel (a) indicates the start time of the individual stations triggered by Magellan, the signal size is indicated by the amplification of the station sizes; the reconstructed shower axis is depicted in blue. The signals of the WCDs are depicted as a function of the radius in panel (b). A disregarded saturated station is indicated as a black dot. Stations outside of the range for which the Universality model is valid are depicted in gray; these are not taken into account for the Universality reconstruction. The best fit as obtained using the Universality model is depicted as an orange band, with the upper (lower) edge representing the signal prediction at the shower-plane azimuth $\psi = 0$ ($\psi = \pi$).



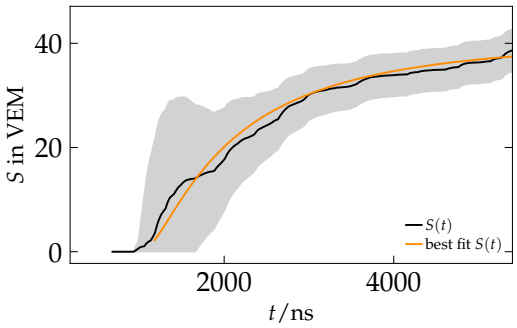
(a) station 873, $r = 1575\,\mathrm{m}$, $\psi = 39°$



(b) station 907, $r = 1641\,\mathrm{m}$, $\psi = 117°$



(c) station 934, $r = 1897\,\mathrm{m}$, $\psi = -129°$



(d) station 871, $r = 1992\,\mathrm{m}$, $\psi = 80°$

Figure 7.9: Signal integrated over time in two stations detecting Magellan. The black lines show the signal as detected by the stations over time; the result of the global best fit for the integrated signal as obtained from the Universality model is given in orange. The estimated time-dependent signal uncertainty is indicated as gray bands.

7.5 HIGH-RIGIDITY EVENTS

To demonstrate the scope of the Universality reconstruction for cosmic-ray astronomy, the 200 highest-rigidity events identified in the data discussed in this chapter are presented. The rigidity R of particles with a charge $Z\,e$,

$$R = \frac{E_0}{Z\,e}, \tag{7.1}$$

governs the trajectory in the intergalactic magnetic fields. The nuclear charge of the CR is estimated by

$$Z\,e \approx \frac{A}{2}\,e. \tag{7.2}$$

For this analysis, if Z evaluates smaller than 1, it is set to its physical lower limit for hadronic CRs, given by 1. In this way, not only is the burning of data for ultra-high energy photon searches avoided, but also is the rigidity estimate less distorted by shower-to-shower fluctuations. The selected events are expected to be minimally deflected and thus their arrival directions potentially point back to their respective sources. The result is depicted in Fig. 7.10. The $\ln A$ estimate of each event is given by the color coding. The same data depicted with a full-range color scale is given Fig. B.29. The data is presented in galactic coordinates using the Aitoff projection. The highest-rigidity events are depicted with small and bright markers. The marker size and transparency increases with decreasing rigidity. In this way, the locatibility of the events is visualized, since the markers for events of lower rigidity cover a slightly larger region of the sky. The shape of the individual markers is not adjusted for the Aitoff projection

Magellan, which is part of the selected data, can be seen as small bright spot near $l \simeq -100^\circ$ and $b \simeq -30^\circ$. A small population of 5 high-rigidity events is apparent slightly south of the location of the Virgo cluster around $l \simeq -60$ and $b \simeq 50^\circ$. Furthermore, from visual inspection a band of CRs can be identified along the Supergalactic Plane and across the region of the Great Attractor.

The detailed analysis of the coincidence of the arrival directions of high-rigidity CRs with astrophysical sources will be addressed in a future work using the results and methods of this thesis.

Figure 7.10: Arrival directions of 200 selected high-rigidity CRs. Details of the illustration are explained in the text. Regions outside of the exposure (including the South Pole) are dashed and shaded in gray. The galactic center is given by a black star.

CHAPTER VIII

SUMMARY AND CONCLUSION

In this work I revisited and extended the work on air-shower universality performed in the last decades (see Chapters 3 and 5). The main application of the Universality model and reconstruction is to determine the mass composition of the ultra-high-energy cosmic rays and to identify weakly charged and minimally deflected cosmic rays on an event-by-event basis.

I highlighted classic solutions of the cascade equations with which the density of electromagnetic particles in extensive air showers can be accurately described. I presented evidence that the widely used Gaisser-Hillas profile matches the analytical solutions of the cascade equations after the earliest stages of the shower development. For a single type of particles, using a combination of the classical solutions to the cascade equations and a modified Gaisser-Hillas profile, one can express the expected areal particle density in air showers relative to the depth of the shower maximum. In this way, the development of individual particle components can be modelled. The particle content of an air shower depends on the type of the primary cosmic ray, with heavy nuclei producing on average more muons than protons. By disentangling the different types of particles in an air shower, a universal description of air showers can be obtained. The abundance of the different types of particles is given as a function of R_μ, which is the number of muons relative to the average number of muons in a proton shower. Using the disentangled particle components and the solutions of the cascade equations I proposed a physics-driven model of the dependence of the expected areal particle densities on the depth of the shower maximum $X_{\max}$ and the relative muon number R_μ. I developed a new model of the temporal distribution of particles arriving at the ground that relates the longitudinal development of an air shower to the distribution of the arrival time of particles. Furthermore, I highlighted the importance of the event geometry when considering the timing of the particles. In this work this is realized by employing the arrival time of a hypothetical plane shower front as a reference.

I proposed a way to parametrize models of the ground particle densities from extensive air showers using Monte Carlo simulations of cosmic ray cascades (see Chapter 5). For this purpose I produced a toolkit to generate models of the spatial and temporal distributions of signal expected in the surface detectors of the Pierre Auger Observatory and its successor, AugerPrime (discussed in Chapter 4), using simulations produced in the $\overline{\text{Off}}\underline{\text{line}}$ framework. The model that is mainly discussed in this work is based on simulations employing the EPOS-LHC model of hadronic interactions. It is accurate within distances of $\sim$2000 m to the shower axis and for showers with zenith angles smaller than $\sim$50°. I developed a reconstruction algorithm employing this model with which $X_{\max}$ and R_μ can be reconstructed with a precision of $\sim$40 g cm^{-2} and $\sim$0.1, respectively, using the surface-detector data only (see Chapter 6). In more detail, the reconstruction is performed in two steps in which first the lateral distribution of the signal is examined to determine R_μ and then the temporal distribution of signal is used to estimate $X_{\max}$; in this process, a steeply falling lateral distri-

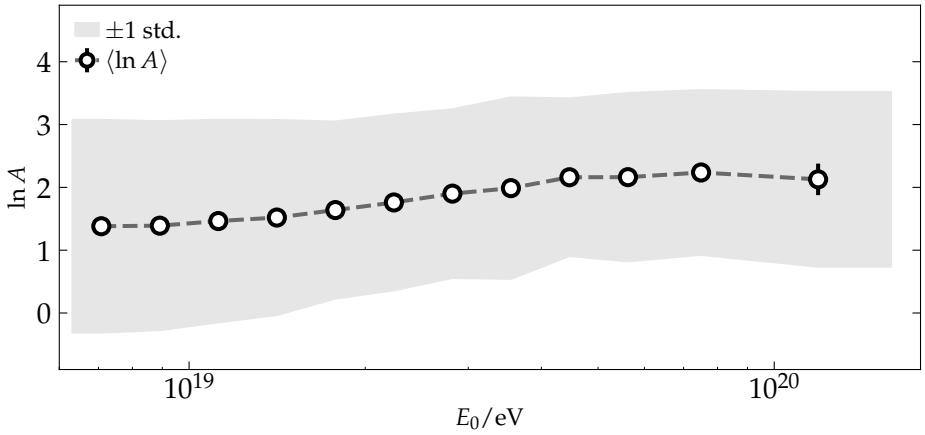


Figure 8.1: Development of the average logarithmic atomic mass $\langle \ln A \rangle$ of cosmic rays as a function of the primary energy E_0. One standard deviation width of the underlying data distribution is depicted as a gray band.

bution of signals corresponds to a small R_μ and particles arriving rather late at the detectors correspond to large $X_{\max}$. The model and reconstruction algorithm presented in this work do not outperform earlier works in terms of precision, but require fewer quality cuts on the data and comprise a novel uncertainty model of the time-dependent signals. An issue of the reconstruction that must be addressed in future work is the surprisingly low standard deviation of the reconstructed values of $X_{\max}$ (see Fig. B.28), which might be an indicator for regression towards the mean. Furthermore, I proposed a method to calculate the logarithmic atomic mass number $\ln A$ of cosmic rays directly from the values of the observables $X_{\max}$ and R_μ (see Section 6.4). Given the precision of the reconstructed values of $X_{\max}$ and R_μ, with this method $\ln A$ can be estimated with a precision of $\sim$1.7, which is sufficient to separate heavy from light cosmic rays on an event-by-event basis. In this method the anti-correlation of $X_{\max}$ and R_μ is employed through a geometric approach to minimize the effect of shower-to-shower fluctuations on the reconstructed primary mass. Additionally, systematic uncertainties arising from the muon deficit, which is apparent in simulated data, are condensed into a single free parameter, which is the average R_μ for protons. Using Golden Hybrid events, where cosmic rays are simultaneously detected by fluorescence telescopes and surface detectors the free parameter is calibrated to the data. The result is a physics-driven estimate of mass that allows for the selection of light cosmic rays at the highest energies. Even though the obtained precision is adequate, I showed that there is a strict physical limit on the possible precision in $\ln A$ arising from the shower-to-shower fluctuations that amounts to 0.6 for iron and 1.6 for protons that cannot be exceeded even with an ideal detector.

Using the reconstructed values of $X_{\max}$ and R_μ (see Chapter 7) the mass composition of cosmic rays was obtained with the proposed method to estimate $\ln A$. The evolution of the average mass of cosmic rays as a function of energy is given in Fig. 8.1. The results show that above an energy of 10^{19} eV an increasingly heavy composition is implied by the data. Between primary energies of $10^{18.8}$ eV and $10^{19.6}$ eV the trend towards a heavier composition, as previously reported by the $X_{\max}$ measurements of the fluorescence detectors, is reaffirmed. The extreme scenario of a pure iron composition of cosmic rays, however, is disfavoured. At the highest energies above $\sim 10^{19.7}$ eV, this trend stops and the composition of cosmic rays is mixed with a constant average mass of $\langle \ln A \rangle \simeq 2$. The standard deviation of the $\ln A$ distribution slightly increases from $\sim$1.3 to $\sim$1.4 above an energy of $\sim 10^{19.7}$ eV,

which could indicate a broader mix in the composition. One particular light cosmic ray candidate found in the data was highlighted in Section 7.4.

The event-level determination of $\ln A$, allows the search for extragalactic accelerators of ultra-high-energy particles. Possible sources of cosmic rays could now be investigated in regions populated most by the arrival directions of light particles. To demonstrate this ability, the arrival directions of 200 ultra-high-energy cosmic rays detected by Auger are visualized in Fig. 8.2. For each cosmic ray the estimated mass is given by the color code of the markers. An example of a sky map of cosmic-ray arrivals with a high-rigidity cut is given in Section 7.5. Furthermore, using this method to separate light from heavy cosmic ray events, the physics of hadronic interactions, such as the proton-air cross section, can be tested using surface-detector data at beyond LHC energies.

Figure 8.2: The arrival directions of 200 high-energy events recorded by the Pierre Auger Observatory, depicted in galactic coordinates using the Aitoff projection and the astronomical coordinate convention. The galactic center is marked as black star. The size of the individual markers increases as a function of the reconstructed logarithmic primary energy. The estimated masses of the cosmic rays are indicated by the color coding, from light (bright red) to heavy (dim blue). Regions of the Universe outside of the exposure are shaded in gray.

APPENDIX A

TABULATED DATA

In this chapter, the numerical results of the parametrizations are presented. The parametrizations of the model are performed on proton showers simulated by CORSIKA employing QGSJET-II-04, EPOS-LHC, and SIBYLL 2.3c as high-energy hadronic interaction model. The respective results are given in Tables A.1 to A.6. The parametrization was performed as described in Section 5.7. The scaling factor a of the individual components was obtained from a global fit using results of all hadronic interaction models and is thus the same for each parametrization. The energy dependence of the individual components that is incorporated in the paramater γ could not be fitted for the showers obtained from SIBYLL 2.3c, since the available library of simulated showers does not contain showers from CRs of different energies. If the model based on SIBYLL 2.3c would be used, it is advised to use γ as obtained from EPOS-LHC.

A Tabulated Data

Table A.1: WCD QGSJET-II

	eγ	μ	eγ(μ)	eγ(π)
a	0.0	1.0	1.0	1.1
γ	0.982	0.963	0.949	0.947
$S^{19}_{\mathrm{ref}}/\mathrm{MIP}$	1.759×10^{10}	7.406×10^{8}	1.102×10^{8}	8.179×10^{8}
$r_{\mathrm{M}}/\mathrm{m}$	355	704	695	164
s	0.49	1.04	1.17	1.32
$S_{\mathrm{th}}/\mathrm{VEM}$	0.9	5.4	1.0	1.11
c_{th}	2.1	1.9	1.9	1.9
$S_{\mathrm{bg}}/\mathrm{VEM}$	0.4	1.0	0.2	0.2
$\Delta X_{\mathrm{max}}/(\mathrm{g\,cm^{-2}})$	$135 + \frac{r}{\mathrm{m}}\,0$	$0 + \frac{r}{\mathrm{m}}\,0.766$	$0 + \frac{r}{\mathrm{m}}\,0.721$	$0.02 + \frac{r}{\mathrm{m}}\,0.019$
$\lambda/(\mathrm{g\,cm^{-2}})$	$61 + \frac{r}{\mathrm{m}}\,0$	$259 + \frac{r}{\mathrm{m}}\,0.258$	$188 + \frac{r}{\mathrm{m}}\,0.539$	$120 + \frac{r}{\mathrm{m}}\,0.025$
ζ	0.03	0.01	0.03	0.1

Table A.2: SSD QGSJET-II

	eγ	μ	eγ(μ)	eγ(π)
a	0.0	1.0	1.0	1.1
γ	0.982	0.963	0.949	0.947
$S^{19}_{\mathrm{ref}}/\mathrm{MIP}$	4.258×10^{10}	4.037×10^{8}	1.122×10^{8}	4.705×10^{8}
$r_{\mathrm{M}}/\mathrm{m}$	630	664	583	215
s	0.23	1.2	1.46	1.47
$S_{\mathrm{th}}/\mathrm{MIP}$	5.0	4.7	2.3	2.6
c_{th}	2.1	1.9	1.9	1.9
$S_{\mathrm{bg}}/\mathrm{MIP}$	0.8	0.7	0.3	0.6
$\Delta X_{\mathrm{max}}/(\mathrm{g\,cm^{-2}})$	$127 + \frac{r}{\mathrm{m}}\,0$	$0 + \frac{r}{\mathrm{m}}\,0.656$	$0 + \frac{r}{\mathrm{m}}\,0.628$	$0 + \frac{r}{\mathrm{m}}\,0.02$
$\lambda/(\mathrm{g\,cm^{-2}})$	$65 + \frac{r}{\mathrm{m}}\,0$	$296 + \frac{r}{\mathrm{m}}\,0.2$	$312 + \frac{r}{\mathrm{m}}\,0$	$120 + \frac{r}{\mathrm{m}}\,0.026$
ζ	0.03	0.01	0.03	0.1

Table A.3: WCD EPOS-LHC

	eγ	μ	eγ(μ)	eγ(π)
a	0.0	1.0	1.0	1.1
γ	0.982	0.963	0.949	0.947
$S^{19}_{\mathrm{ref}}/\mathrm{MIP}$	1.602×10^{10}	7.211×10^{8}	1.042×10^{8}	8.147×10^{8}
$r_{\mathrm{M}}/\mathrm{m}$	410	660	725	174
s	0.4	1.1	1.14	1.33
$S_{\mathrm{th}}/\mathrm{VEM}$	0	3.3	0.5	0.7
c_{th}	1.5	3.0	1.9	1.9
$S_{\mathrm{bg}}/\mathrm{VEM}$	0.00	0.20	0.10	0.10
$\Delta X_{\mathrm{max}}/(\mathrm{g\,cm^{-2}})$	$140 + \frac{r}{\mathrm{m}}\,0$	$0 + \frac{r}{\mathrm{m}}\,0.426$	$154 + \frac{r}{\mathrm{m}}\,0.141$	$45 + \frac{r}{\mathrm{m}}\,0$
$\lambda/(\mathrm{g\,cm^{-2}})$	$58 + \frac{r}{\mathrm{m}}\,0.014$	$280 + \frac{r}{\mathrm{m}}\,0.201$	$278 + \frac{r}{\mathrm{m}}\,0.162$	$98 + \frac{r}{\mathrm{m}}\,0.011$
ζ	0.03	0.01	0.03	0.1

Table A.4: SSD EPOS-LHC

	eγ	μ	eγ(μ)	eγ(π)
a	0.0	1.0	1.0	1.1
γ	0.982	0.963	0.949	0.947
$S^{19}_{\mathrm{ref}}/\mathrm{MIP}$	4.077×10^{10}	4.028×10^{8}	1.078×10^{8}	5.392×10^{8}
$r_{\mathrm{M}}/\mathrm{m}$	649	579	613	658
s	0.2	1.28	1.43	0.89
$S_{\mathrm{th}}/\mathrm{MIP}$	1.2	3.0	1.3	0.5
c_{th}	2.5	2.9	2.1	1.5
$S_{\mathrm{bg}}/\mathrm{MIP}$	0.30	0.13	0.13	0.24
$\Delta X_{\mathrm{max}}/(\mathrm{g\,cm^{-2}})$	$135 + \frac{r}{\mathrm{m}}\,0$	$5 + \frac{r}{\mathrm{m}}\,0.236$	$121 + \frac{r}{\mathrm{m}}\,0.122$	$40 + \frac{r}{\mathrm{m}}\,0$
$\lambda/(\mathrm{g\,cm^{-2}})$	$61 + \frac{r}{\mathrm{m}}\,0.015$	$250 + \frac{r}{\mathrm{m}}\,0.187$	$259 + \frac{r}{\mathrm{m}}\,0.174$	$111 + \frac{r}{\mathrm{m}}\,0.019$
ζ	0.03	0.01	0.03	0.1

Table A.5: WCD SIBYLL 2.3c

	eγ	μ	eγ(μ)	eγ(π)
a	0.0	1.0	1.0	1.1
γ	-	-	-	-
$S^{19}_{\mathrm{ref}}/\mathrm{MIP}$	1.674×10^{10}	7.955×10^{8}	1.180×10^{8}	9.237×10^{8}
$r_{\mathrm{M}}/\mathrm{m}$	350	699	676	163
s	0.50	1.03	1.16	1.30
$S_{\mathrm{th}}/\mathrm{VEM}$	0.8	5.7	1.0	1.2
c_{th}	2.0	1.9	1.9	1.9
$S_{\mathrm{bg}}/\mathrm{VEM}$	0.31	1.00	0.20	0.32
$\Delta X_{\mathrm{max}}/(\mathrm{g\,cm^{-2}})$	$134 + \frac{r}{\mathrm{m}}\,0$	$0 + \frac{r}{\mathrm{m}}\,0.835$	$0 + \frac{r}{\mathrm{m}}\,0.627$	$0 + \frac{r}{\mathrm{m}}\,0.300$
$\lambda/(\mathrm{g\,cm^{-2}})$	$62 + \frac{r}{\mathrm{m}}\,0.010$	$288 + \frac{r}{\mathrm{m}}\,0.301$	$292 + \frac{r}{\mathrm{m}}\,0.206$	$116 + \frac{r}{\mathrm{m}}\,0$
ζ	0.03	0.01	0.03	0.1

Table A.6: SSD SIBYLL 2.3c

	eγ	μ	eγ(μ)	eγ(π)
a	0.0	1.0	1.0	1.1
γ	-	-	-	-
$S^{19}_{\mathrm{ref}}/\mathrm{MIP}$	3.684×10^{10}	4.324×10^{8}	1.184×10^{8}	5.477×10^{8}
$r_{\mathrm{M}}/\mathrm{m}$	602	684	617	324
s	0.26	1.16	1.41	1.25
$S_{\mathrm{th}}/\mathrm{MIP}$	5.1	4.8	2.4	1.8
c_{th}	2.1	1.9	1.9	1.9
$S_{\mathrm{bg}}/\mathrm{MIP}$	0.80	0.68	0.35	0.60
$\Delta X_{\mathrm{max}}/(\mathrm{g\,cm^{-2}})$	$128 + \frac{r}{\mathrm{m}}\,0$	$0 + \frac{r}{\mathrm{m}}\,0.685$	$0 + \frac{r}{\mathrm{m}}\,0.707$	$1 + \frac{r}{\mathrm{m}}\,0.029$
$\lambda/(\mathrm{g\,cm^{-2}})$	$65 + \frac{r}{\mathrm{m}}\,0.008$	$273 + \frac{r}{\mathrm{m}}\,0.254$	$201 + \frac{r}{\mathrm{m}}\,0.382$	$123 + \frac{r}{\mathrm{m}}\,0.024$
ζ	0.03	0.01	0.03	0.1

APPENDIX B

ADDITIONAL FIGURES

In this chapter, we present additional figures to which it is referred to in the text.

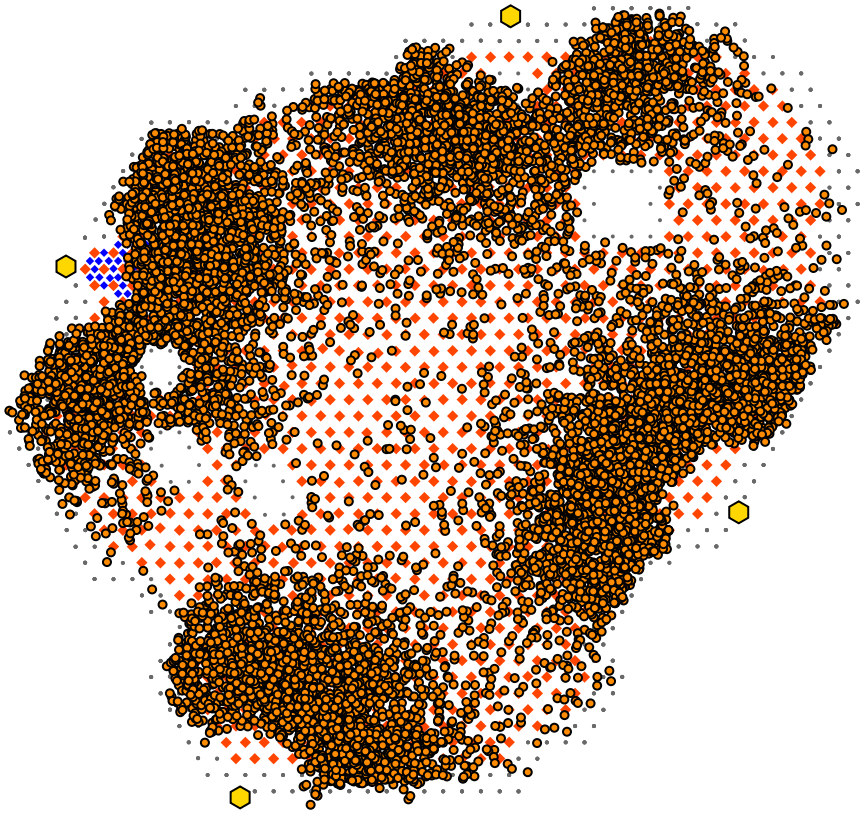


Figure B.1: Same as Fig. 4.6, showing core positions at the ground for 11228 Golden Hybrid events with $E_{\mathrm{rec}} \geq 1\,\mathrm{EeV}$.

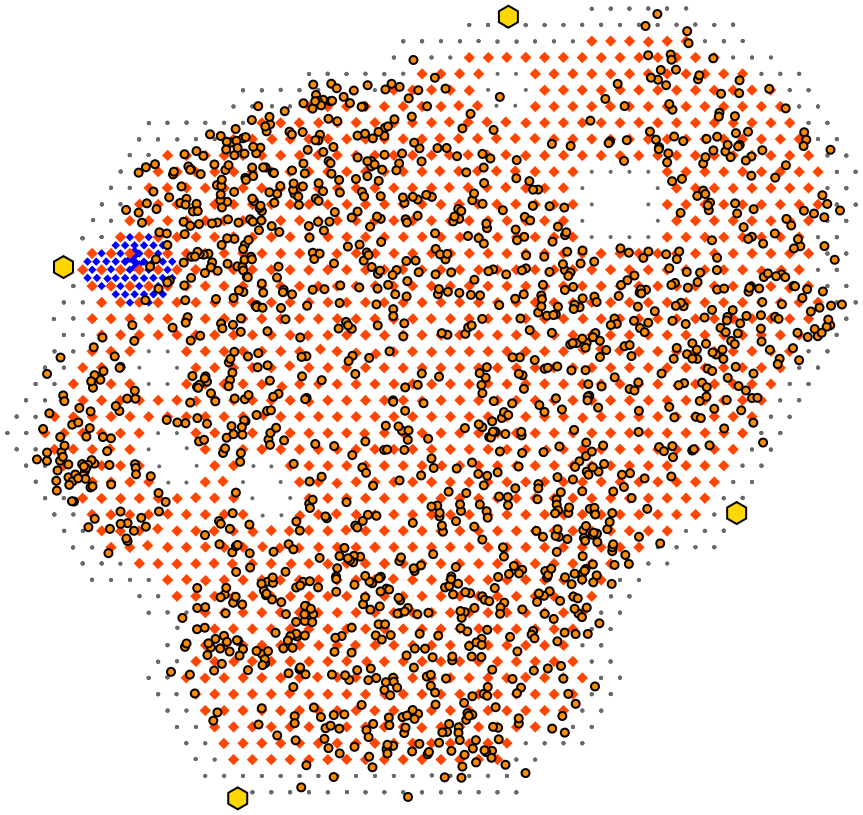


Figure B.2: Same as Fig. 4.6, showing core positions at the ground for 1323 Golden Hybrid events with $E_{\mathrm{rec}} \geq 5\,\mathrm{EeV}$.

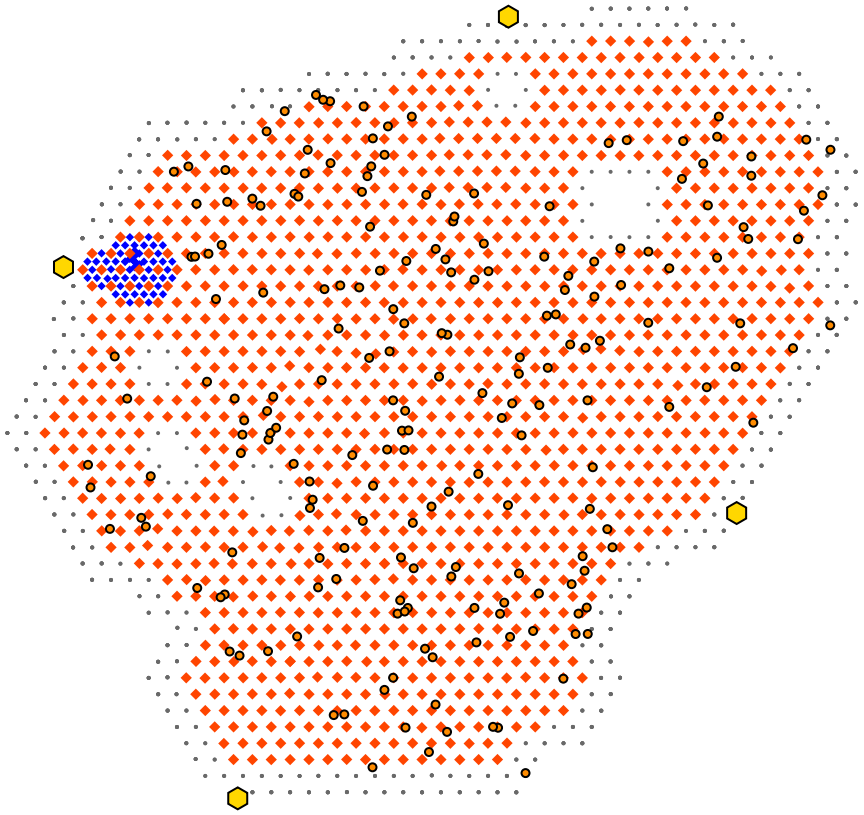


Figure B.3: Same as Fig. 4.6, showing core positions at the ground for 196 Golden Hybrid events with $E_{\mathrm{rec}} \geq 16\,\mathrm{EeV}$.

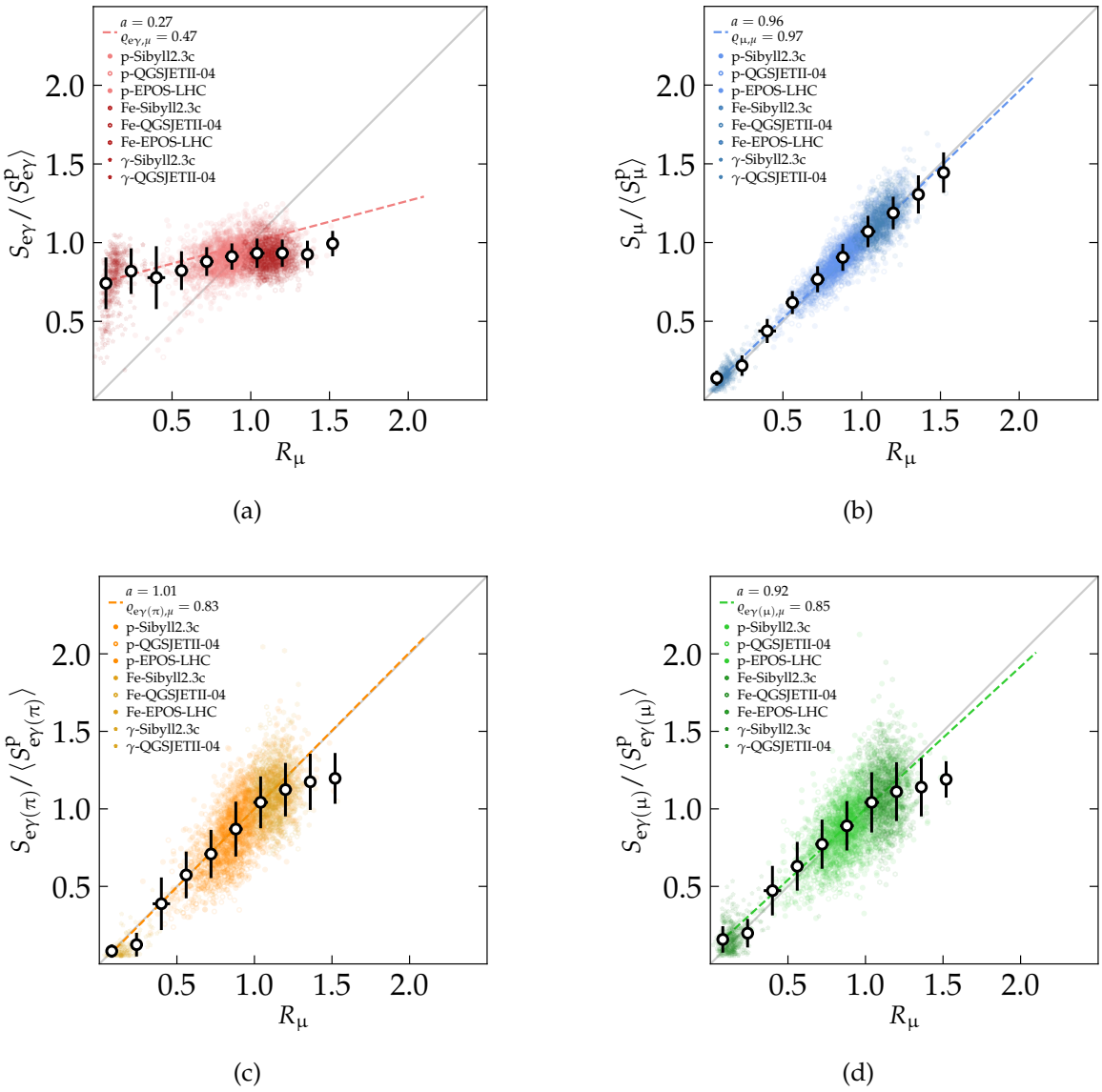


Figure B.4: Same as Fig. 5.1 but using simulated SSD responses.

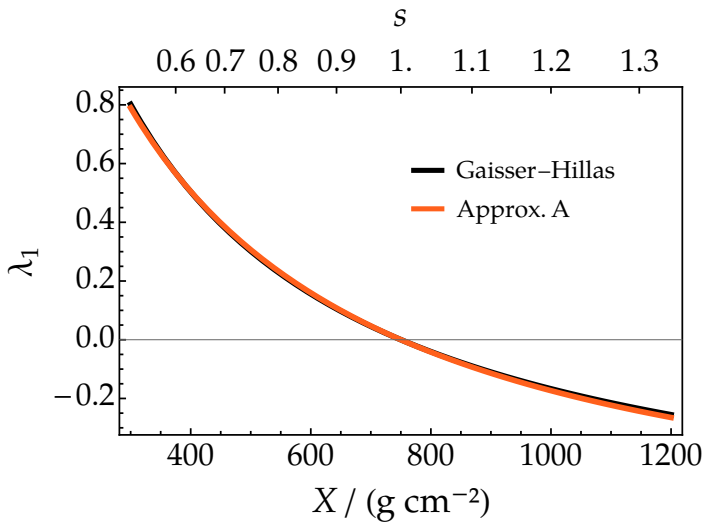


Figure B.5: λ_1 as the relative rate of change of the Gaisser-Hillas profile compared to λ_1 as implied by the solution of the cascade equations under Approximation A, given in Eq. (3.36). For the Gaisser-Hillas profile $X_{\max}$ is set to $750\,\mathrm{g\,cm^{-2}}$, the parameters X_1 and λ are set according to Ref. [118] to $X_1 = -121\,\mathrm{g\,cm^{-2}}$ and $\lambda = 61\,\mathrm{g\,cm^{-2}}$. λ_1 is scaled by a factor of 6/5 for the Gaisser-Hillas function.

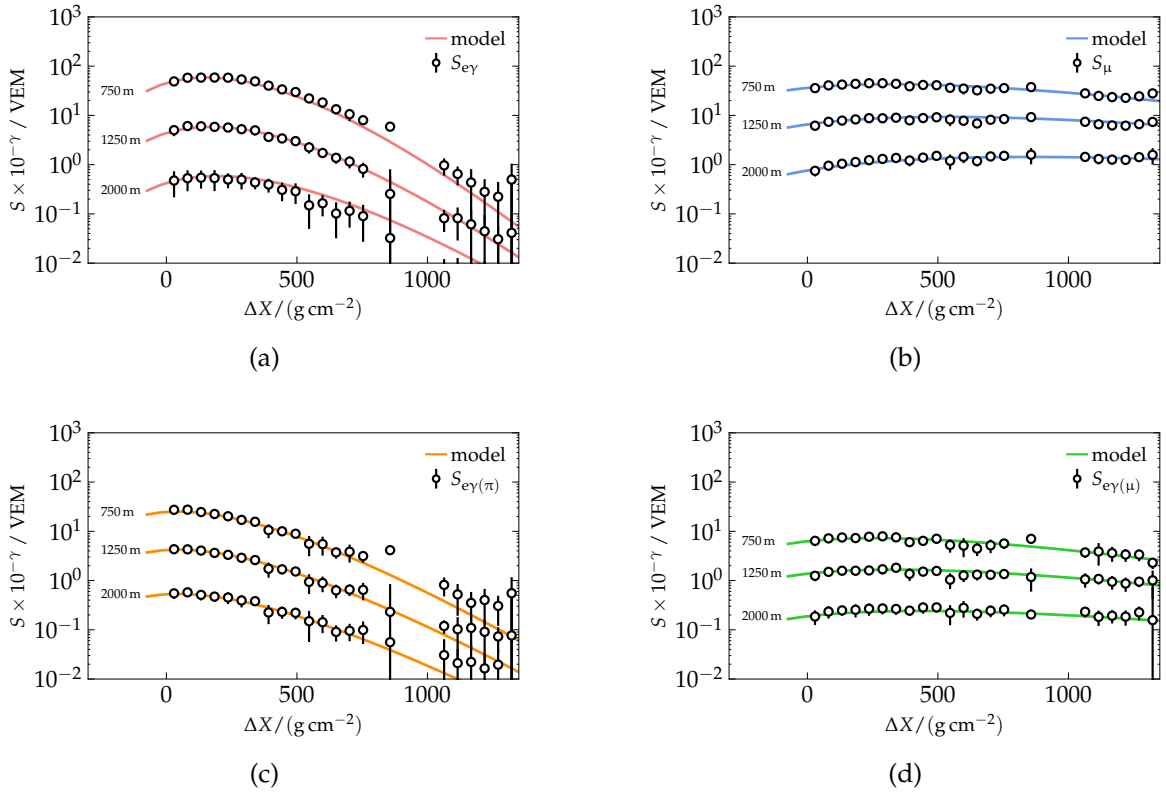


Figure B.6: Signal responses of simulated WCDs as a function of ΔX along with the resulting model according to Eq. (5.10) for the individual components. Same as Fig. 5.12 but for higher primary energies. Error bars indicate one standard deviation of the underlying data.

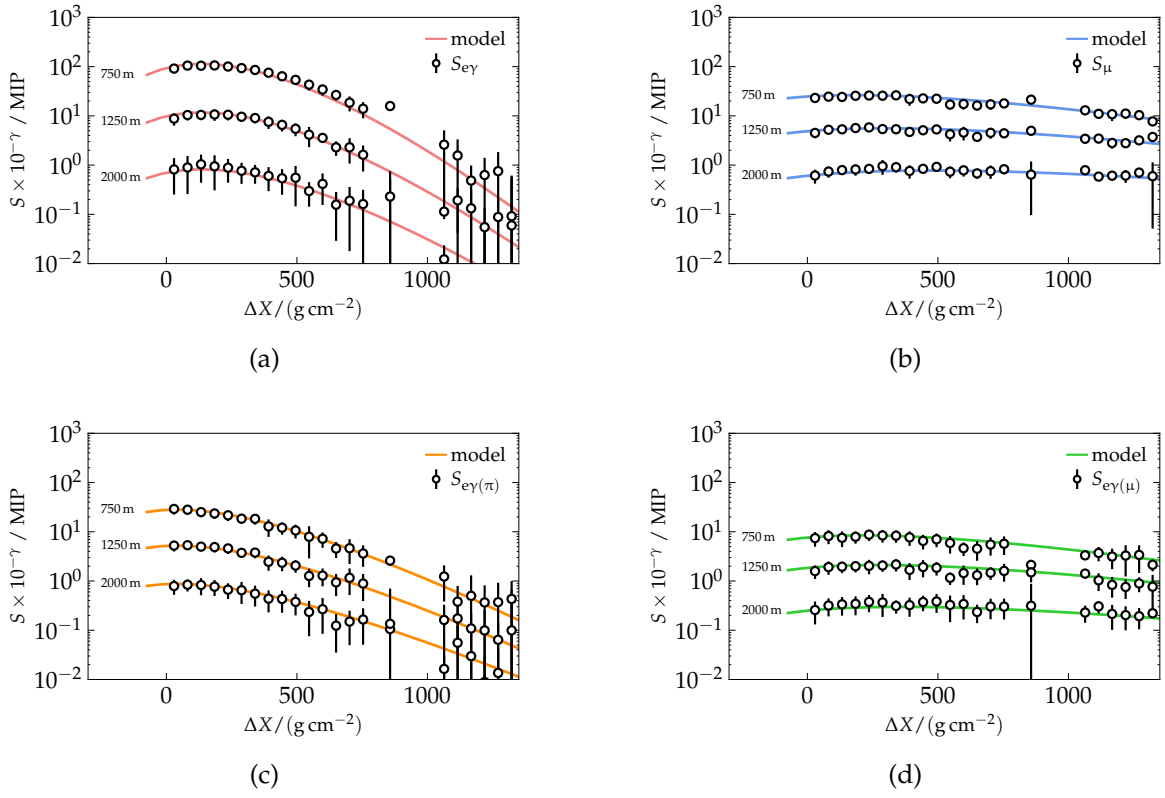
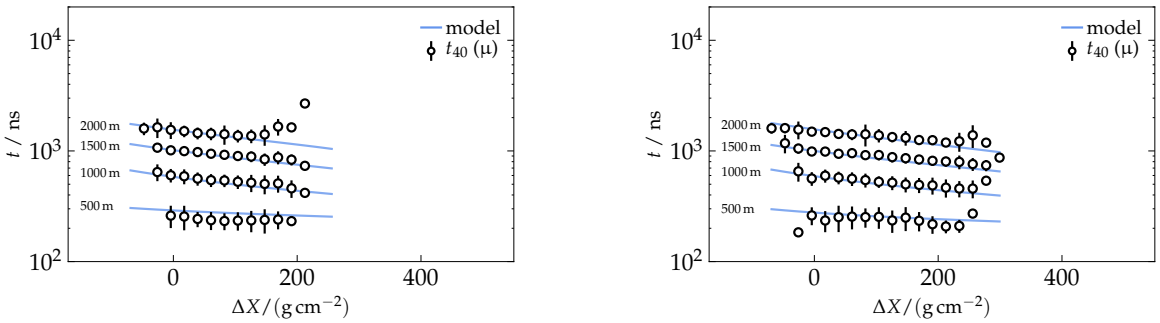


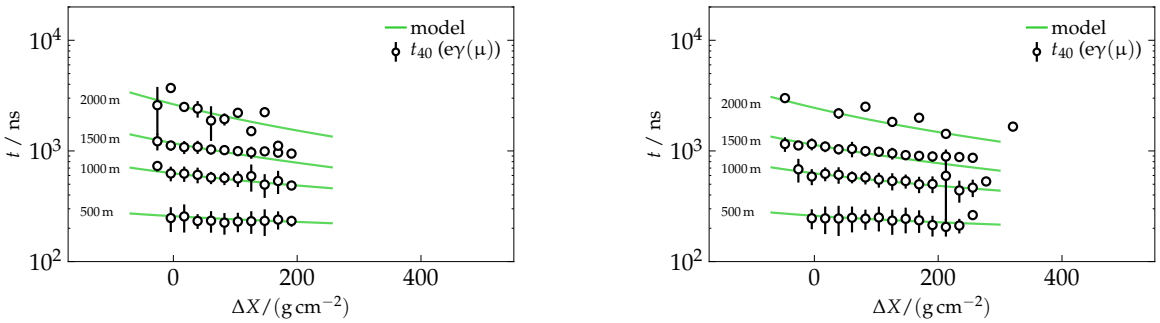
Figure B.7: Signal responses of simulated SSDs as a function of ΔX along with the resulting model according to Eq. (5.10) for the individual components. Same as Fig. 5.13 but for higher primary energies. Error bars indicate one standard deviation of the underlying data.



(a) $\theta = 12°$ (b) $\theta = 22°$

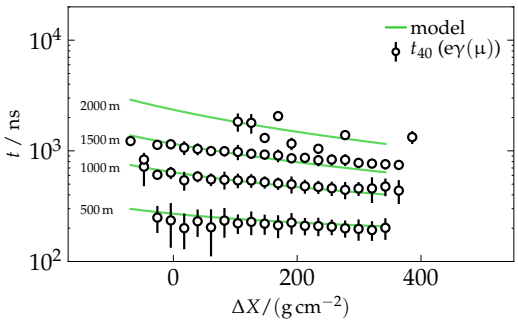
(c) $\theta = 32°$ (d) $\theta = 38°$

Figure B.8: Values of t_{40} for the μ component as a function of ΔX obtained from WCD responses to showers of $E_0 = 10^{19}$ eV and 10^{20} eV and various zenith angles, alongside with the model resulting from the parametrization.

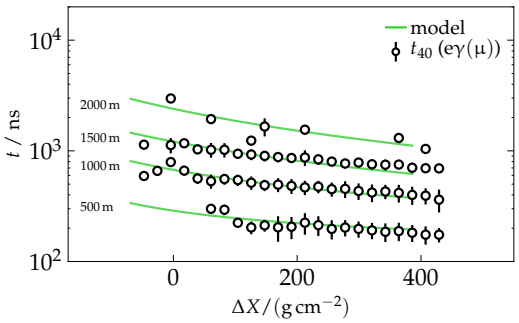


(a) $\theta = 12°$

(b) $\theta = 22°$

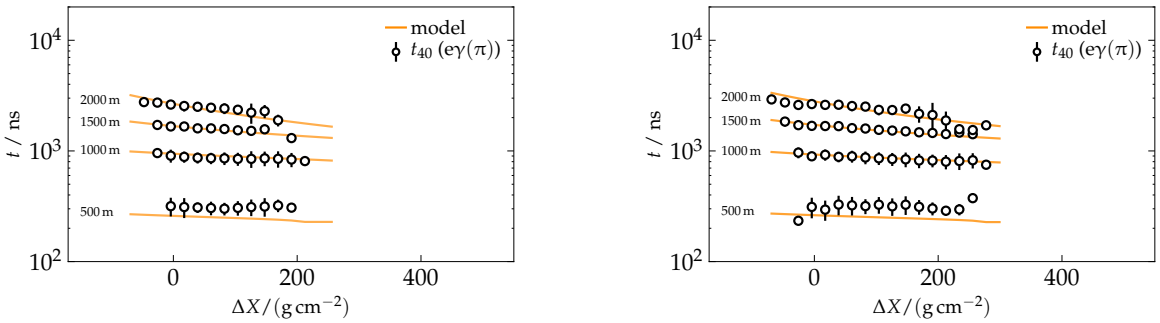


(c) $\theta = 32°$



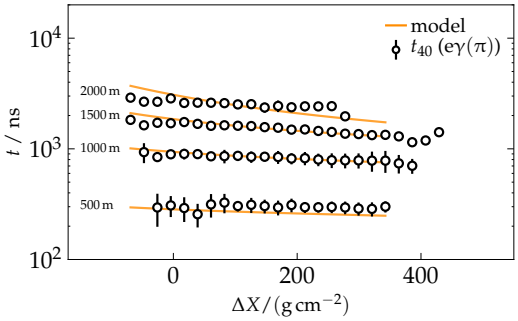
(d) $\theta = 38°$

Figure B.9: Values of t_{40} for the eγ(μ) component as a function of ΔX obtained from WCD responses to showers of $E_0 = 10^{19}$ eV and 10^{20} eV and various zenith angles, alongside with the model resulting from the parametrization.

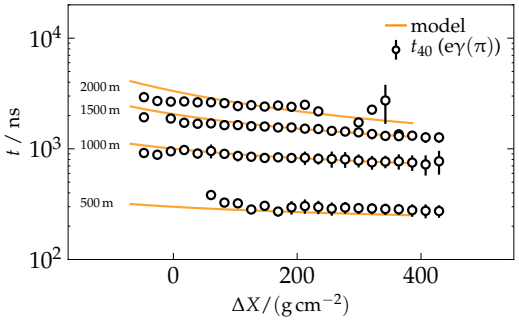


(a) $\theta = 12°$

(b) $\theta = 22°$

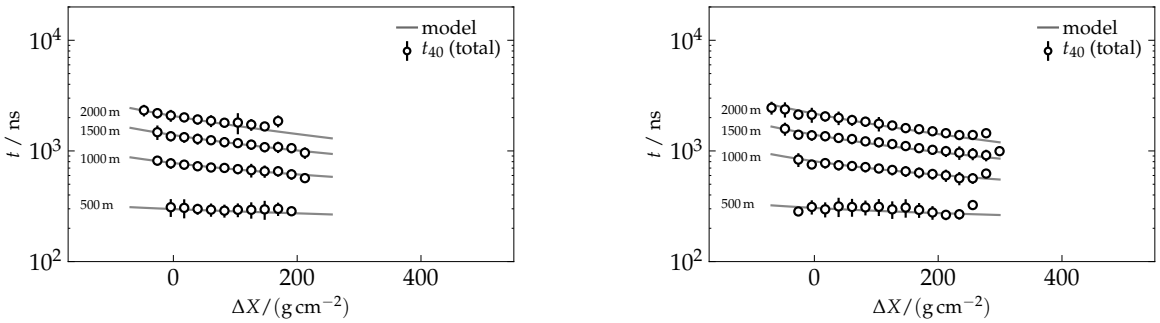


(c) $\theta = 32°$



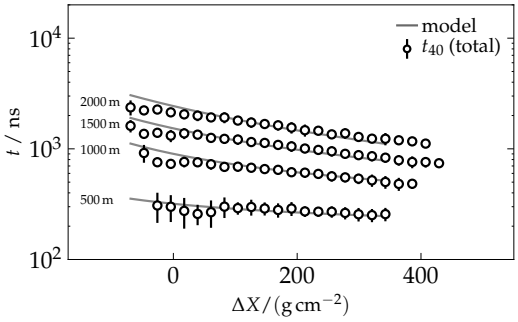
(d) $\theta = 38°$

Figure B.10: Values of t_{40} for the eγ(π) component as a function of ΔX obtained from WCD responses to showers of $E_0 = 10^{19}$ eV and 10^{20} eV and various zenith angles, alongside with the model resulting from the parametrization.

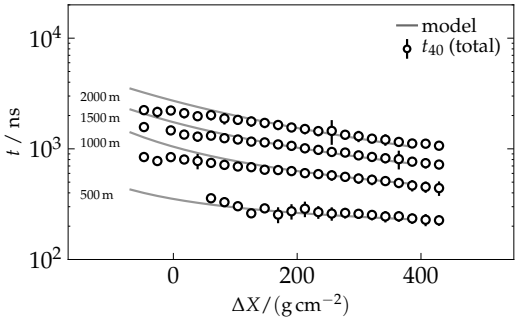


(a) $\theta = 12°$

(b) $\theta = 22°$

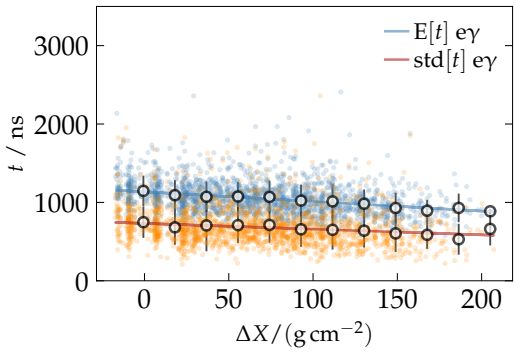


(c) $\theta = 32°$

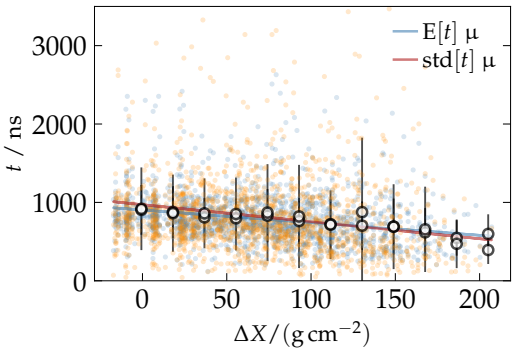


(d) $\theta = 38°$

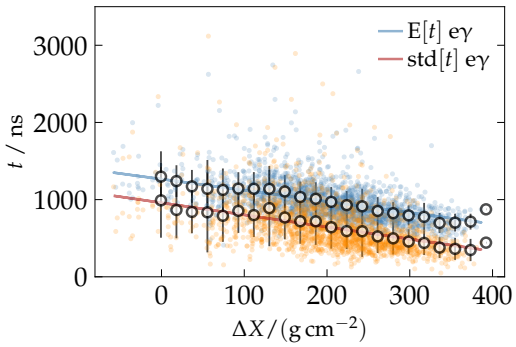
Figure B.11: Values of t_{40} for the sum of all particles components as a function of ΔX obtained from WCD responses to showers of $E_0 = 10^{19}\,\mathrm{eV}$ and $10^{20}\,\mathrm{eV}$ and various zenith angles, alongside with the model resulting from the parametrization.



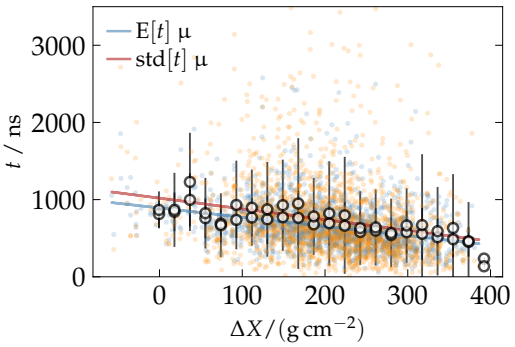
(a) eγ component, $\theta = 12°$



(b) μ component, $\theta = 12°$

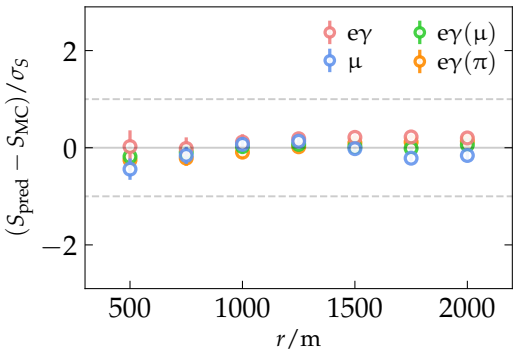


(c) eγ component, $\theta = 32°$

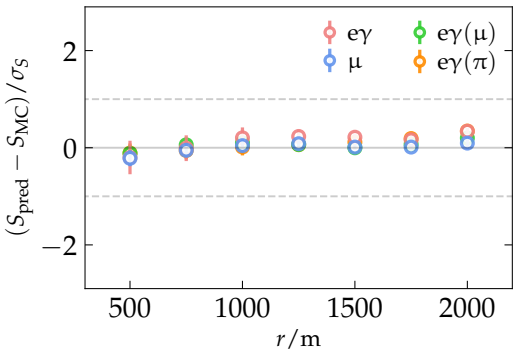


(d) μ component, $\theta = 32°$

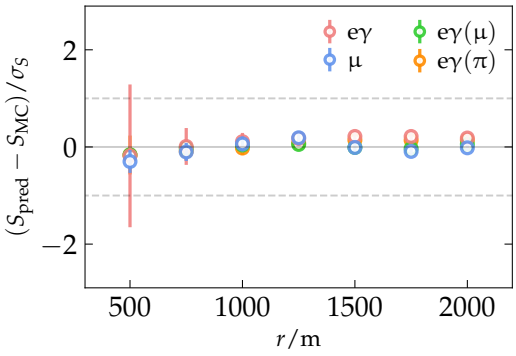
Figure B.12: Behaviour of the moments of the log-normal distribution fitted to the responses of SSDs located at $r = 1000\,\mathrm{m}$ as a function of ΔX. Colored dots depict individual data points, the binned mean is depicted as white markers. The corresponding model is given as line.



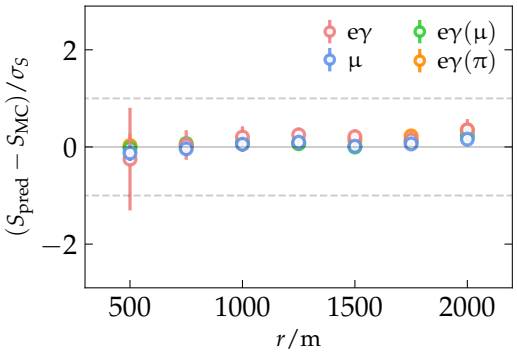
(a) WCD QGSJET-II



(b) SSD QGSJET-II



(c) WCD SIBYLL



(d) SSD SIBYLL

Figure B.13: Validation of the model of the expected signal in WCDs (*left*) and SSDs (*right*), parametrized upon QGSJET-II (*top row*) and SIBYLL (*bottom row*) showers. The average z-score of the predicted signal is given using the signal uncertainty model of the respective detector responses. A deviation of the expected signal by ± 1 standard deviation is indicated as gray dashed line. Validating the model parametrized on showers generated with SIBYLL 2.3c a larger than expected spread of the eγ signals at small radii was seen in the simulated data.

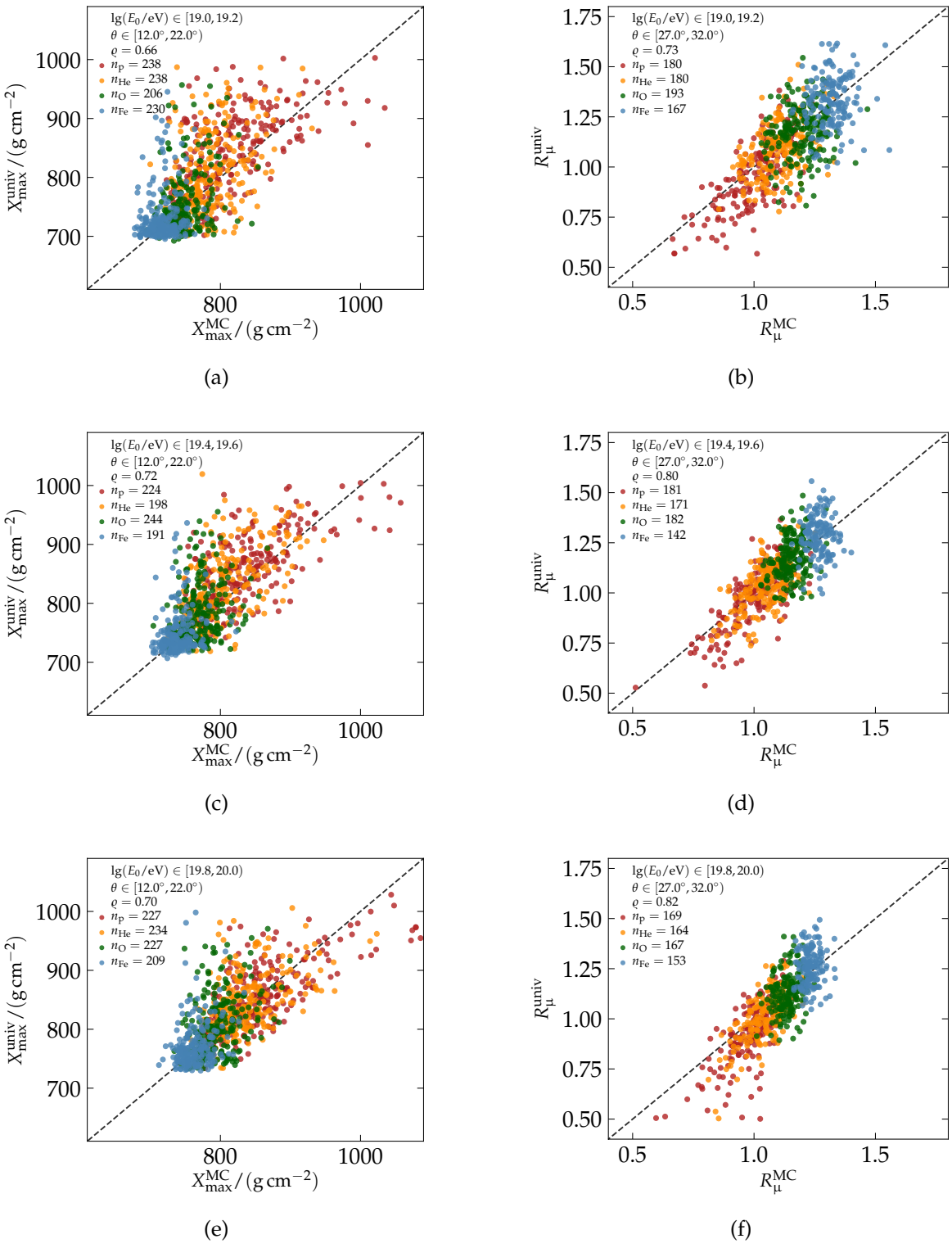
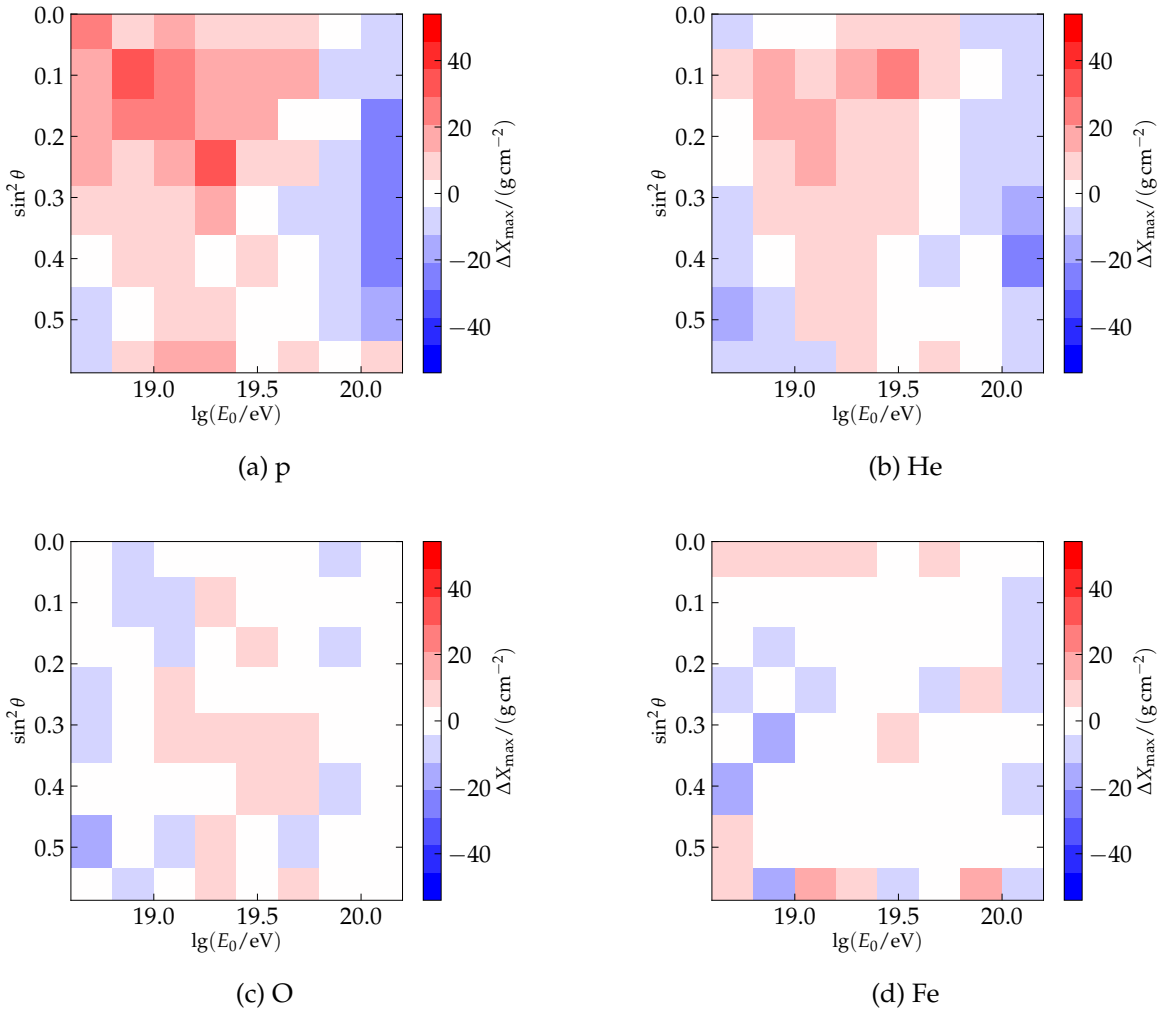


Figure B.14: Correlation of the reconstructed and MC values for X_{max} and R_μ in arbitrarily selected zenith- and energy ranges. The corresponding parameter spaces for E_0 and θ as well as the respective correlation coefficients are given in the upper-left corner of each frame.



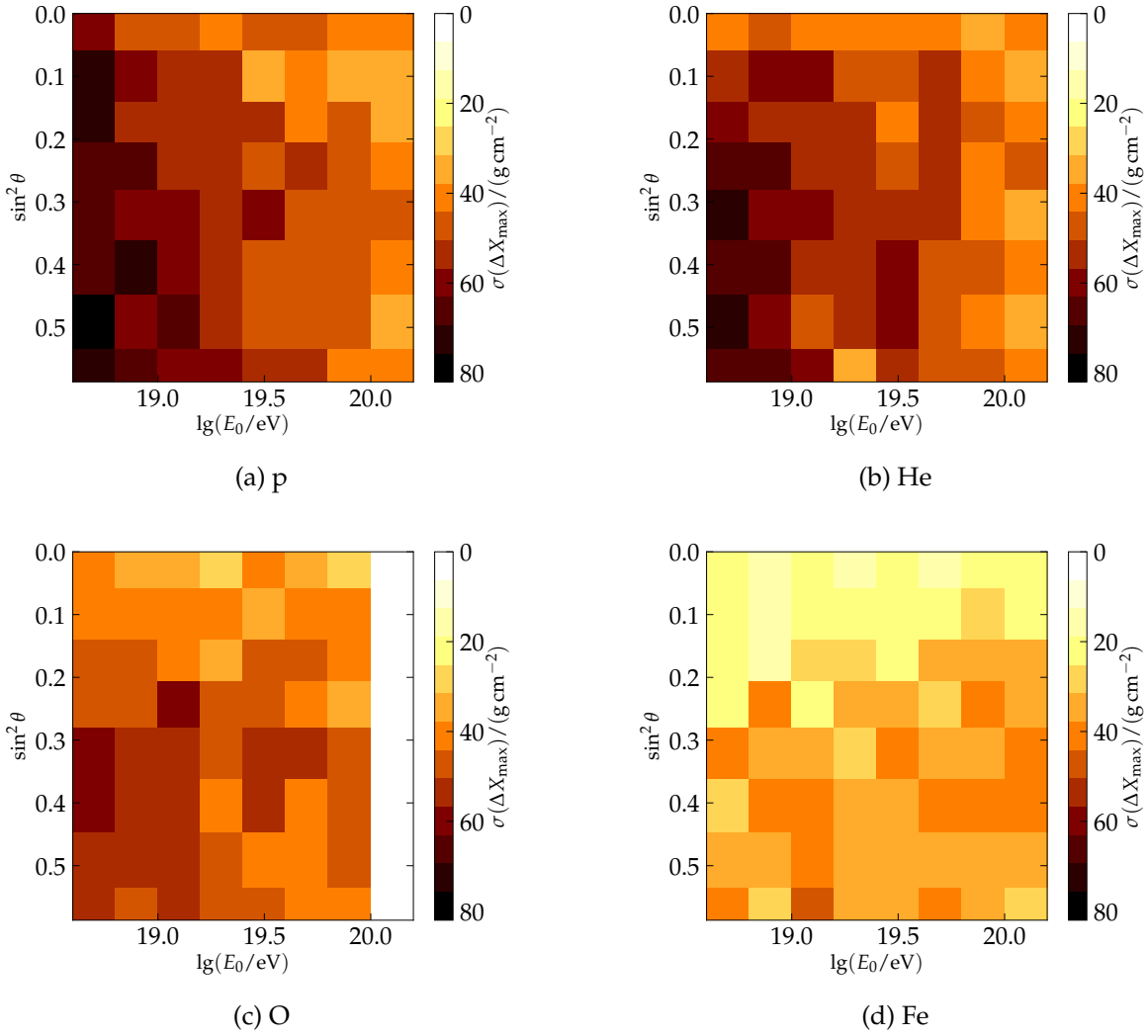
(a) p

(b) He

(c) O

(d) Fe

Figure B.15: Two-dimensional representation of the bias of the reconstructed values for $X_{\max}$ for four different primary nuclei.



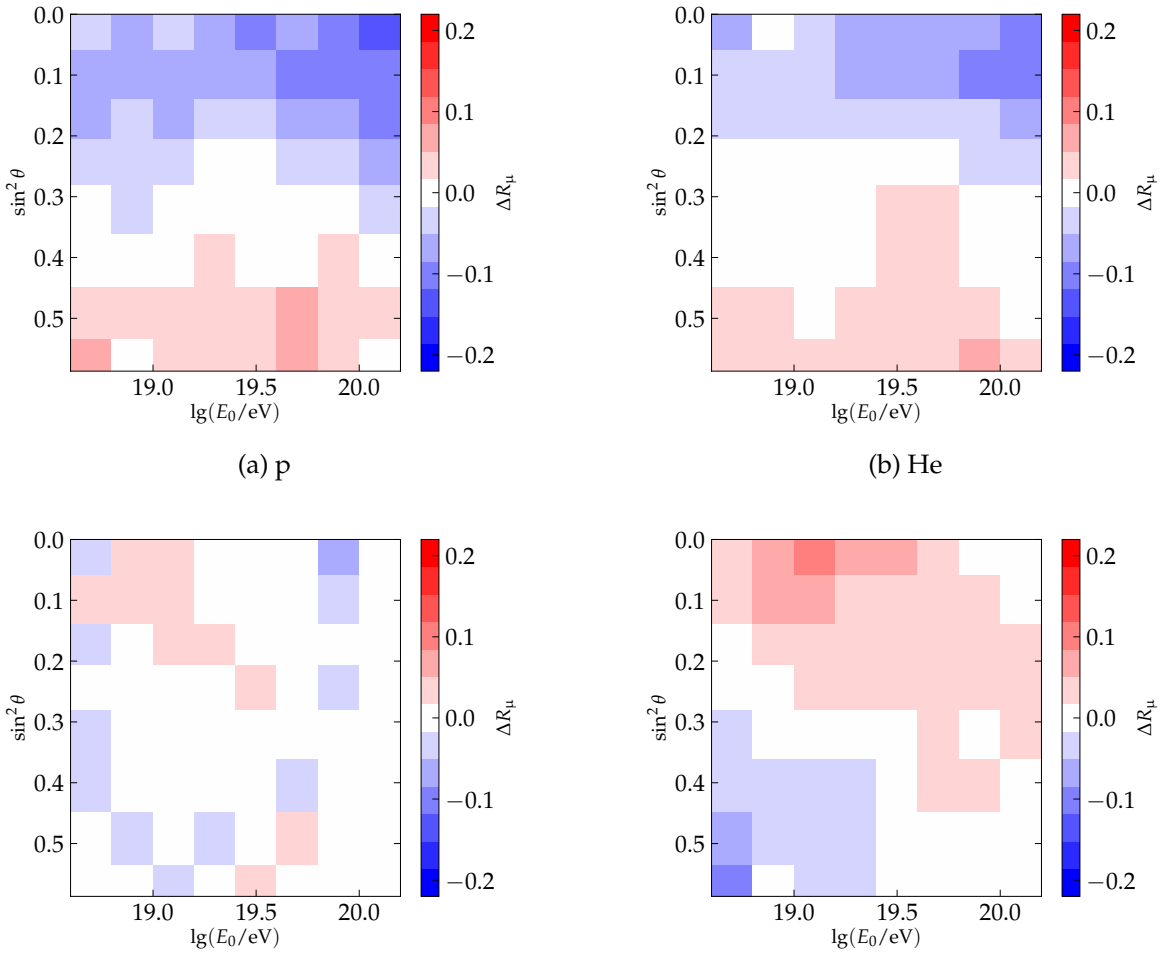
(a) p

(b) He

(c) O

(d) Fe

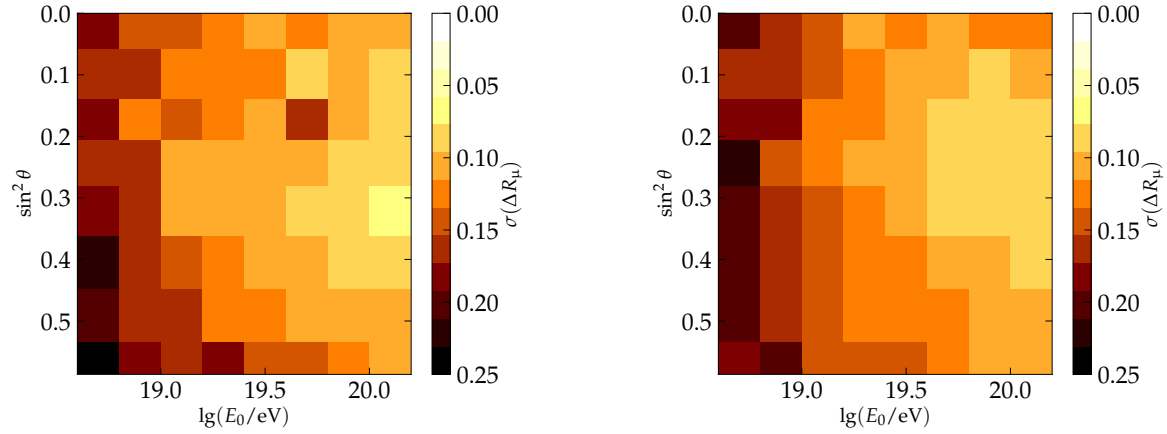
Figure B.16: Two-dimensional representation of the precision of the reconstruction of X_{max} for four different primary nuclei.



(a) p (b) He

(c) O (d) Fe

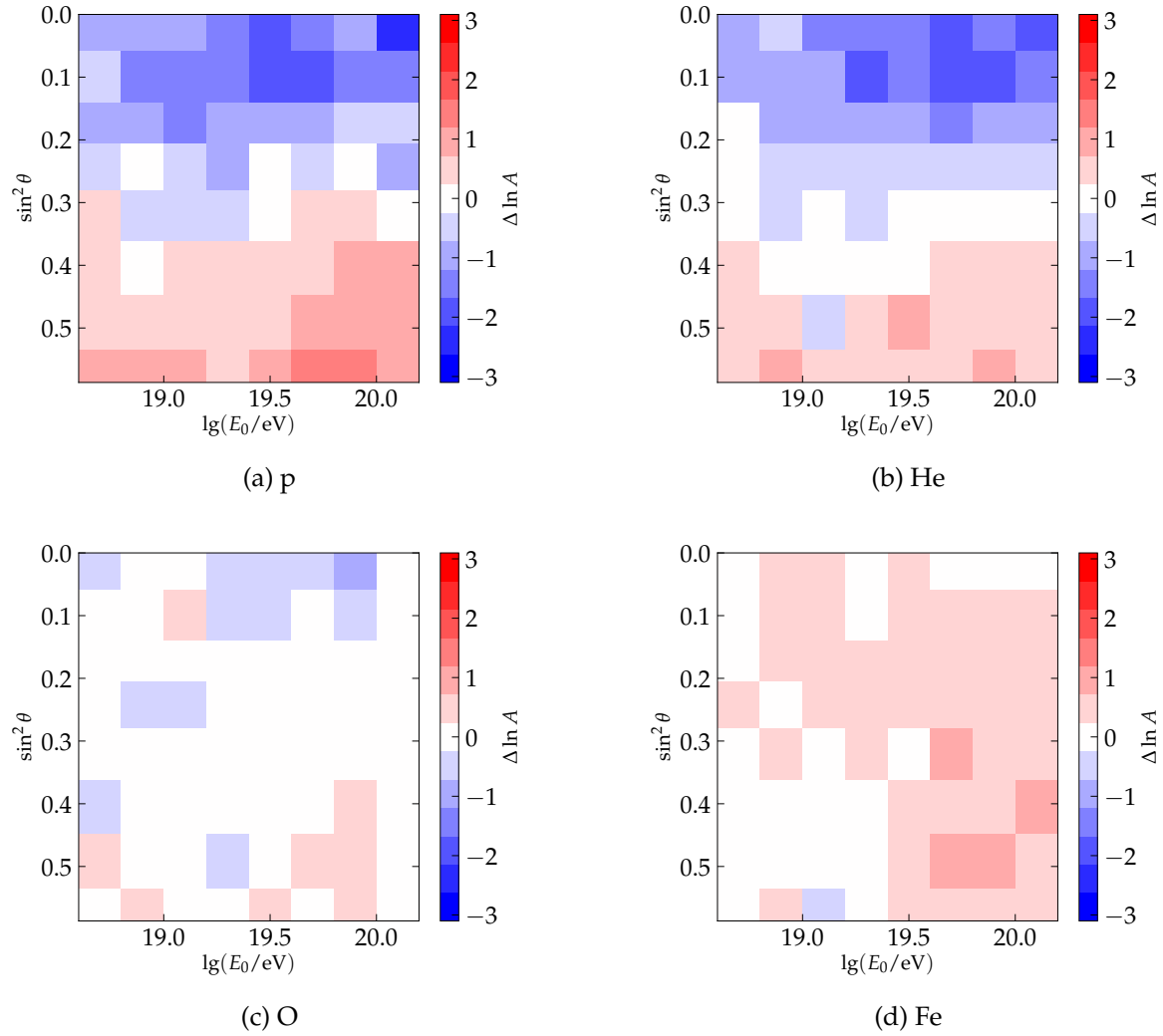
Figure B.17: Two-dimensional representation of the bias of the reconstructed values for R_μ for four different primary nuclei.



(a) p (b) He

(c) O (d) Fe

Figure B.18: Two-dimensional representation of the precision of the reconstruction of R_μ for four different primary nuclei.



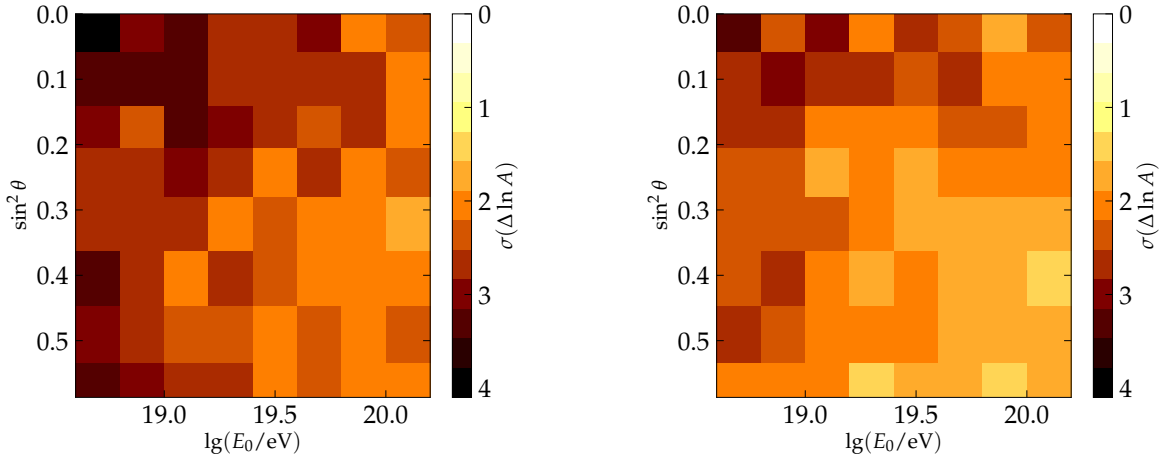
(a) p

(b) He

(c) O

(d) Fe

Figure B.19: Two-dimensional representation of the bias of the reconstructed values for ln A for four different primary nuclei.



(a) p

(b) He

(c) O

(d) Fe

Figure B.20: Two-dimensional representation of the precision of the reconstruction of $\ln A$ for four different primary nuclei.

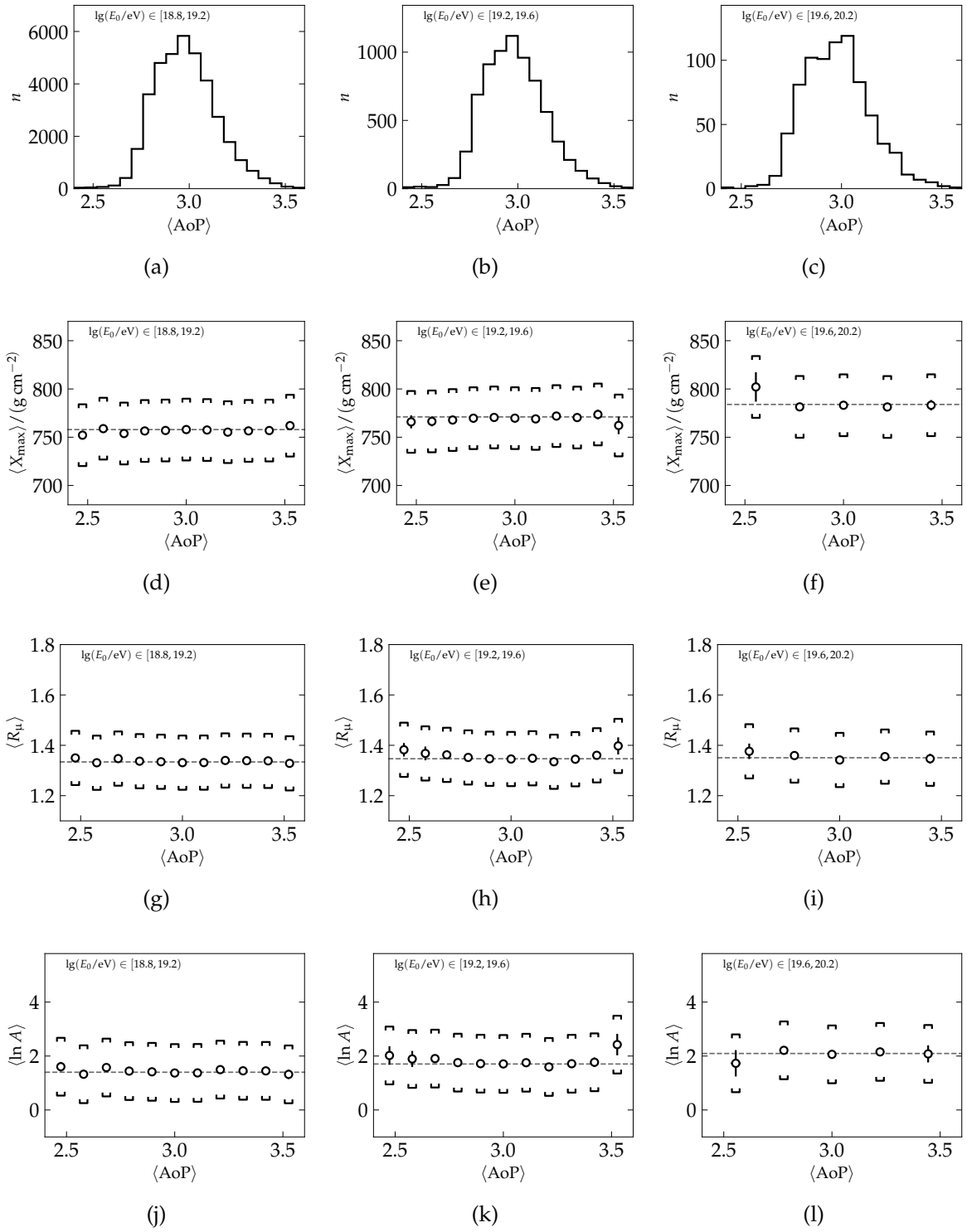


Figure B.21: Effect of the average AoP of the PMTs on the reconstructed observables. The depicted data comprise all reconstructed events of the SD data. The respective energy range is given in the upper-left corner of each panel. The respective average values are depicted as a reference by a dashed line.

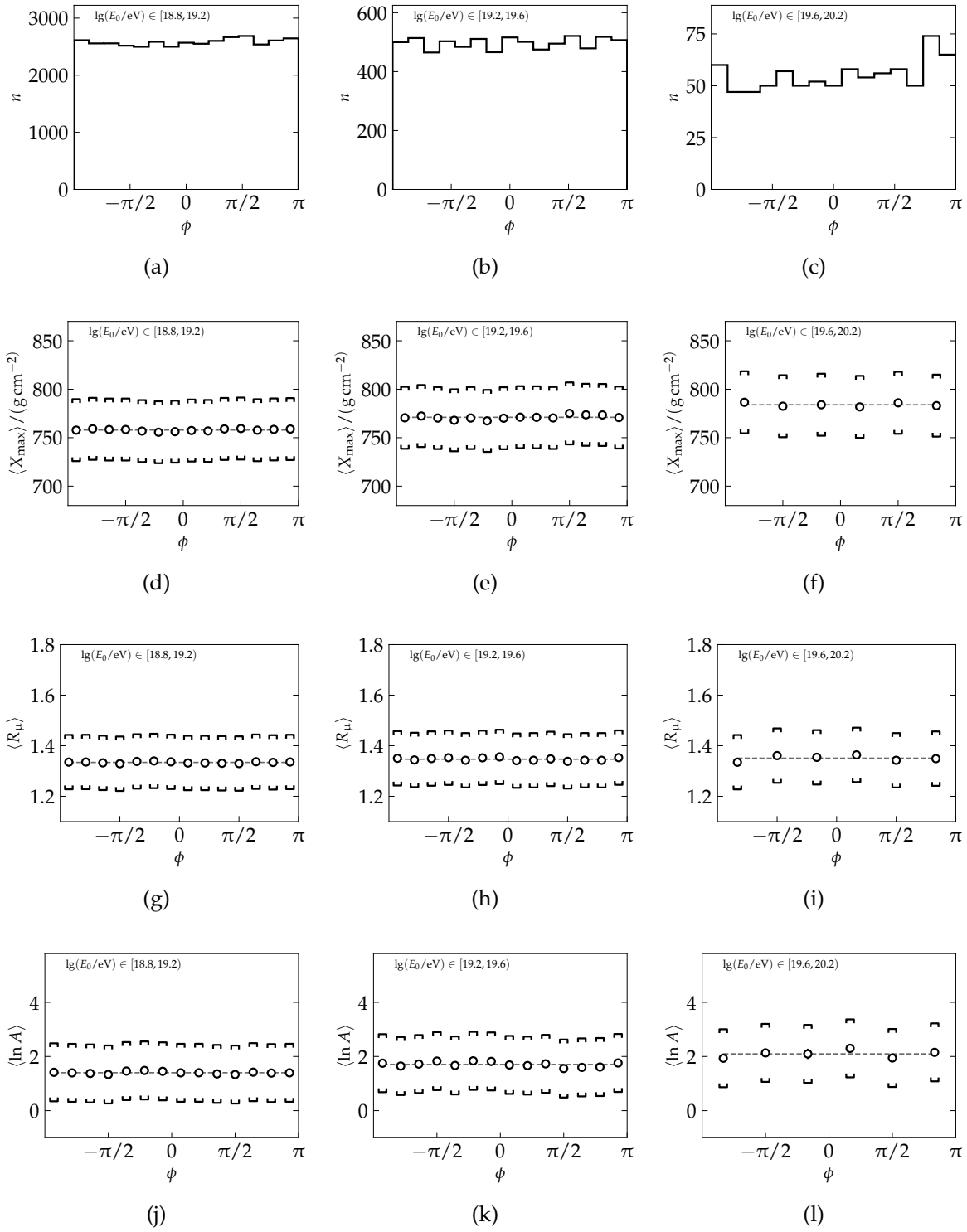


Figure B.22: Effect of the azimuthal direction of CRs on the reconstructed observables. The depicted data comprise all reconstructed events of the SD data. The respective energy range is given in the upper-left corner of each panel. The respective average values are depicted as a reference by a dashed line.

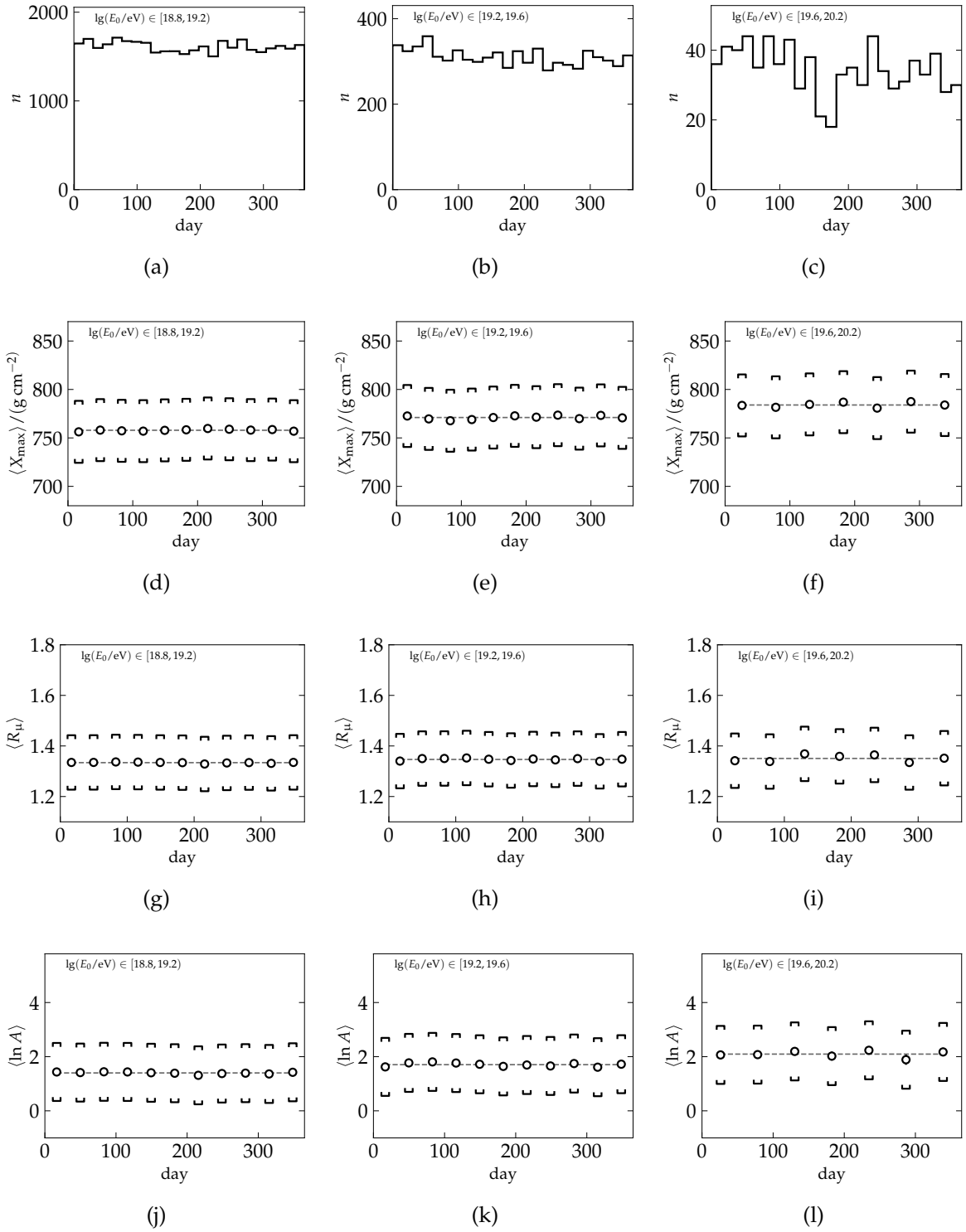


Figure B.23: Seasonal effects on the reconstructed observables as a function of the day of the year. The depicted data comprise all reconstructed events of the SD data. The respective energy range is given in the upper-left corner of each panel. The respective average values are depicted as a reference by a dashed line.

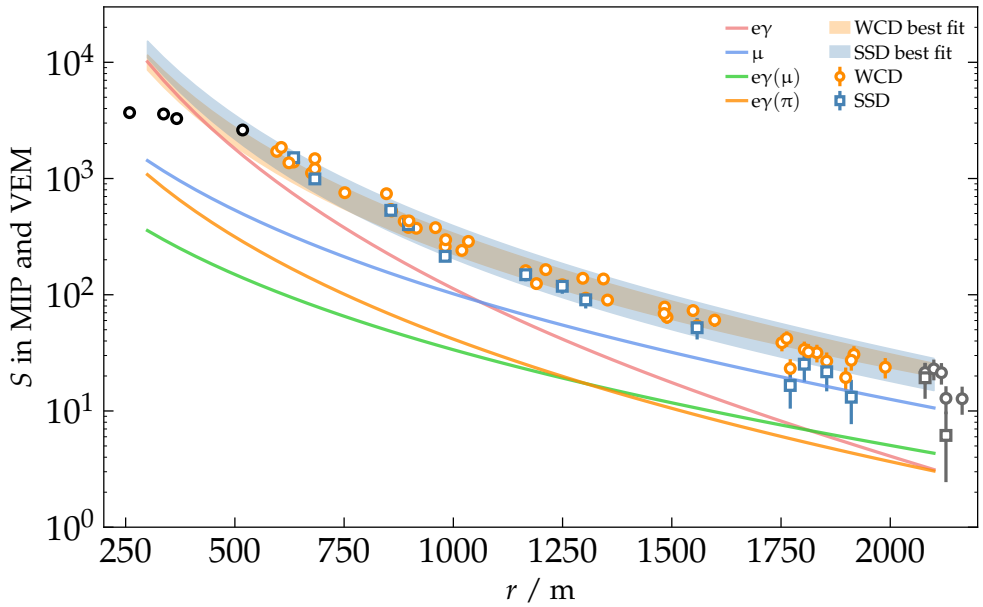


Figure B.24: Reconstruction of event (c) described in Table 7.3, showing the signal of the respective detectors, as well as the best fit using the Universality reconstruction as bands. Additionally, the prediction of the individual SSD signal components from the Universality model are given as a function of the shower-plane distance as colored lines for $\psi = \pm\pi/2$.

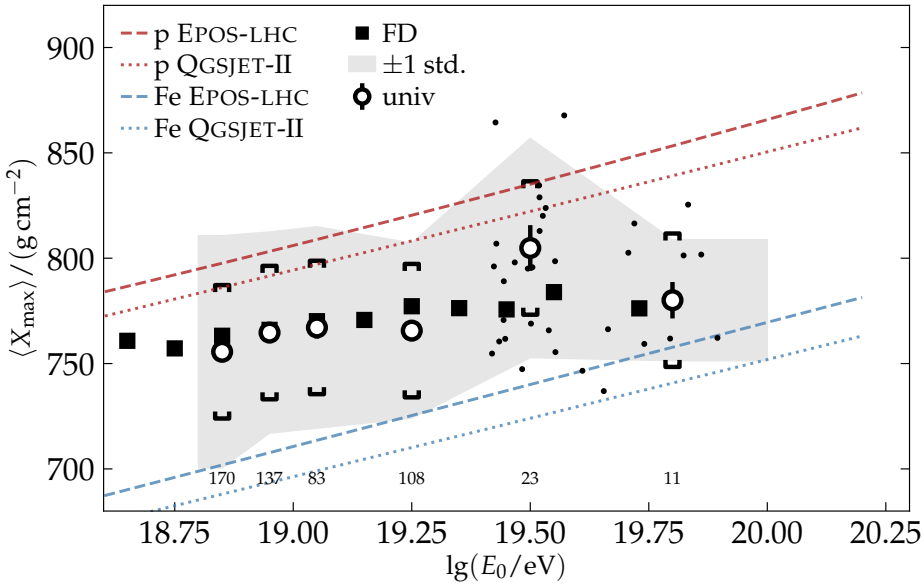


Figure B.25: Same as Fig. 7.3 (a) using only events with data from at least 2 SSDs. The content of the two highest-energy bins is also depicted as individual markers. One event at $(\lg(E_0/\mathrm{eV}),\ X_{\mathrm{max}}/(\mathrm{g\,cm^{-2}})) = (19.4, 995)$ is outside the box of this plot.

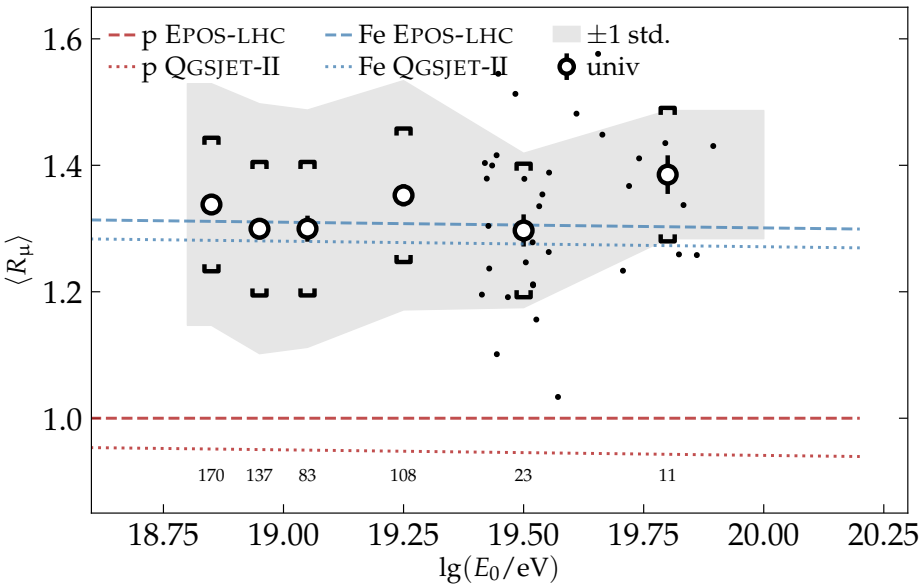


Figure B.26: Same as Fig. 7.3 (b) using only events with data from at least 2 SSDs. The content of the two highest-energy bins is also depicted as individual markers.

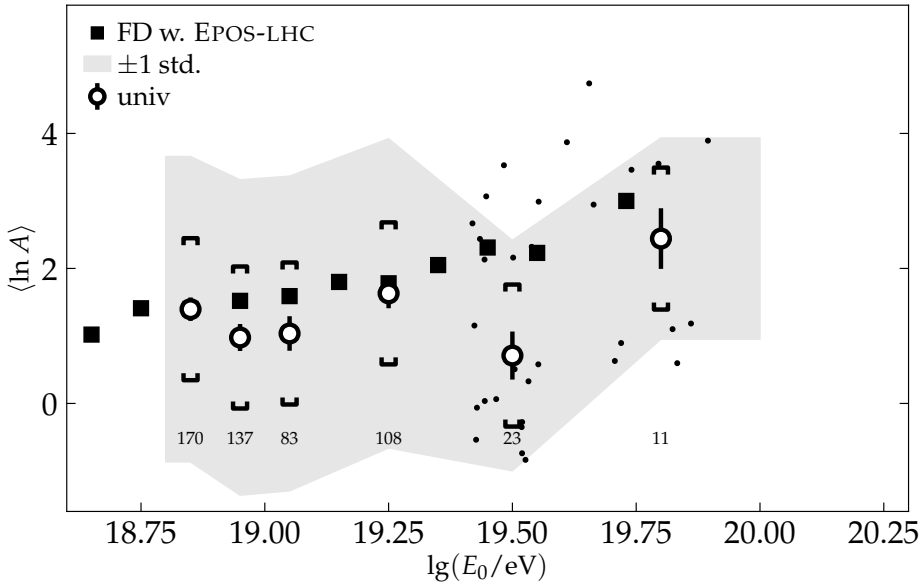
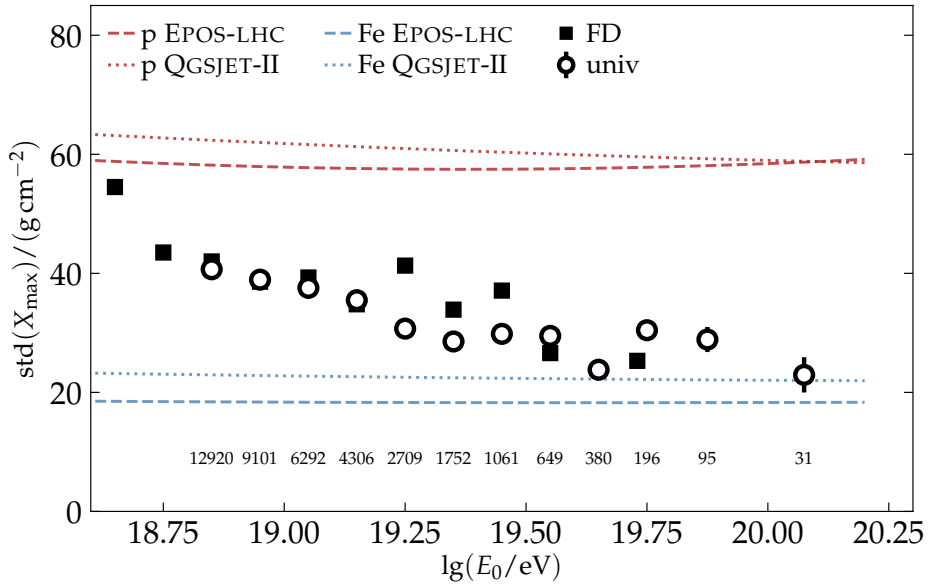
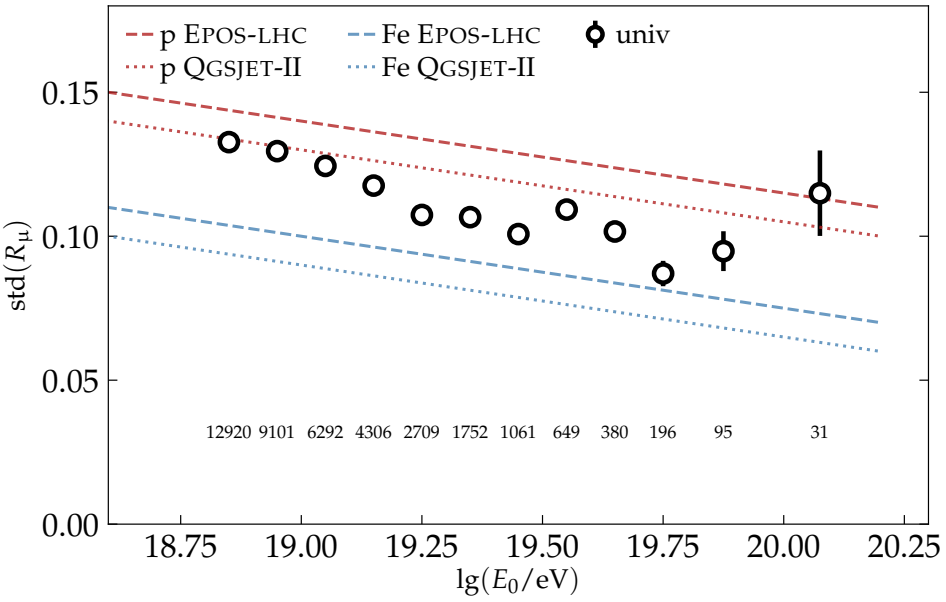


Figure B.27: Same as Fig. 7.4 using only events with data from at least 2 SSDs. The content of the two highest-energy bins is also depicted as individual markers. Three events at $(\lg(E_0/\mathrm{eV}), \ln A) = (19.4, -1.6)$, $(19.4, -2.4)$, and $(19.6, -3.1)$ are outside the box of this plot.



(a)



(b)

Figure B.28: Standard deviation of the distribution of the reconstructed values of X_{max} in panel (a) and R_μ in panel (b) as a function of the SD-reconstructed primary energy, shown as circular markers. Approximated values for iron and proton from two different hadronic interaction models are shown as a reference as dashed lines. The numbers above the x-axis indicate the amount of events in the respective energy bin. Results from the FD reconstruction of Ref. [154] are given as black squares. Note that the depicted data is the raw standard deviation of the data and not the true width, for which the estimated precision is subtracted in quadrature. Even though the standard deviation of the distribution of the reconstructed X_{max} as a function of energy seems to agree well with the results from the FD, a word of caution is needed. The standard deviation of the distribution of the reconstructed X_{max} is smaller than the estimated precision of the method. This could be a remnant of the outlier suppression described in Section 6.1.1 and must be addressed in a future work.

Figure B.29: Arrival directions of 200 selected high-rigidity CRs. Details of the illustration are explained in the text. Regions outside of the exposure (including the South Pole) are dashed and shaded in gray. The galactic center is given by a black star.

B ADDITIONAL FIGURES

APPENDIX C

HISTORIC DERIVATION OF THE GREISEN PROFILE

The renown Greisen profile is a successful approximation for the number of particles in the longitudinal development of electromagnetic cascades. When it was introduced more than 70 years ago, unfortunately, there was no strict derivation provided by its name giver himself. A way to derive the Greisen profile is outlined in Ref. [63], where the cascade equations are solved under Approximation B and the absolute scale of the Greisen profile is numerically verified a-posteriori. In this section, we propose a derivation that might have been the original way to obtain the profile function.

In Ref. [62] Kenneth Greisen introduces the approximation

$$\lambda_1 = \frac{1}{2}\left(s - 1 - 3\ln(s)\right) \tag{C.1}$$

for the relative change of the longitudinal shower profile for electromagnetic cascades. As discussed in Section 3.1.3, Eq. (C.1) is a direct consequence of the conditions for λ_1 that arise if one tries to solve the cascade equations under Approximation A for a single particle initiating the cascade. In Ref. [62] as well as in Ref. [67] λ_1 is expressed as *(sic)*

$$\frac{\mathrm{d}\ln N}{\mathrm{d}t} = \lambda_1(s). \tag{C.2}$$

With the dependence of s on the radiation length of the cascade t, which is well represented by Greisen's approximation given in Eq. (3.40), the differential equation emerging from Eq. (C.2) reads as

$$\frac{\mathrm{d}\ln N}{\mathrm{d}t} = \frac{1}{2}\left(\frac{3t}{t+2y} - 1 - 3\ln\left(\frac{3t}{t+2y}\right)\right). \tag{C.3}$$

Simply integrating the right-hand side of the equation with respect to t, one immediately finds that

$$N(t) = N_0 \exp\left[t\left(1 - \frac{3}{2}\ln\left(\frac{3t}{t+2y}\right)\right)\right], \tag{C.4}$$

with a constant N_0 that needs to be determined. Using the boundary condition that at the shower maximum,

$$N(y) = N_0\, \mathrm{e}^{y} \stackrel{!}{=} N_{\max}, \tag{C.5}$$

the remaining free parameter can be fixed. It is assumed that at this point all particles carry an average energy of $\langle E\rangle = \epsilon_c^{e\gamma}$, thus, as already mentioned in Section 3.1.4, $y = \ln(E_0/\epsilon_c^{e\gamma})$.

C Historic Derivation of the Greisen Profile

The maximum number of particles produced in electromagnetic cascades was calculated under Approximation B by Tamm and Belenky in Ref. [157] shortly before Greisen's series of publications on air showers and was found to be

$$N_{\max}(E) = \frac{0.31}{\sqrt{\ln(E/\epsilon_c^{e\gamma})}} \frac{E_0}{\epsilon_c^{e\gamma}}. \tag{C.6}$$

Given the coincident prefactor of Eq. (C.6) and the Greisen profile, it is likely that the boundary condition was solved to match the solutions of Tamm and Belenky, so that

$$N_0 = \frac{0.31}{\sqrt{\ln(E/\epsilon_c^{e\gamma})}}. \tag{C.7}$$

Thus, using the historic notation $\beta_0 = \ln(E_0/\epsilon_c^{e\gamma})$ and $s = 3t/(t + 2\beta_0)$, the Greisen profile reads as

$$N(t) = \frac{0.31}{\beta_0^{1/2}} \exp\left[t\left(1 - \tfrac{3}{2}\ln s\right)\right]. \tag{C.8}$$

APPENDIX D

A MODEL OF THE ATMOSPHERE FOR AUGER

The Offline framework uses a number of various tabulated atmospheres to approximate the natural conditions for simulations. When handling these tabulated atmospheres, Offline extrapolates between data points and thereby creates an almost continuous scalar field of the atmospheric density. Tabulated atmospheres can be found in the Offline source directory under the following paths:
`FdIdealStandard` in

`Framework/Atmosphere/ParametricXMLProfileModel.xml.in`

and "Bariloche Universality" in

`Tools/ShowerUniversality/UnivParamNS/Atmosphere.cc`

The numerical integration along a path in this scalar field can be computationally challenging and thus slowing down the simulation or reconstruction process. In this chapter, we propose an effective and computationally cheap analytic approximation for the tabulated atmospheric data. The work of this chapter is published in Ref. [122].

Using an isothermal atmospheric density profile, the vertical depth X_v from the top of the atmosphere down to the height h is given by

$$X_\mathrm{v}(h) = \int_h^\infty \rho(h)\,\mathrm{d}h = X_\mathrm{vg}\exp(-(h-h_\mathrm{g})/h_\mathrm{s}). \tag{D.1}$$

using h_g as the ground height above sea level. The slant depth X at the slant distance l from the ground is similarly given by

$$X(l,\theta) = \int_l^\infty \rho(h(l))\,\mathrm{d}l = X_\mathrm{vg}\sec\theta\exp(-l\cos\theta/h_\mathrm{s}). \tag{D.2}$$

Using Eqs. (D.1) and (D.2), the absolute height corresponding to a slant depth X can therefore be expressed as

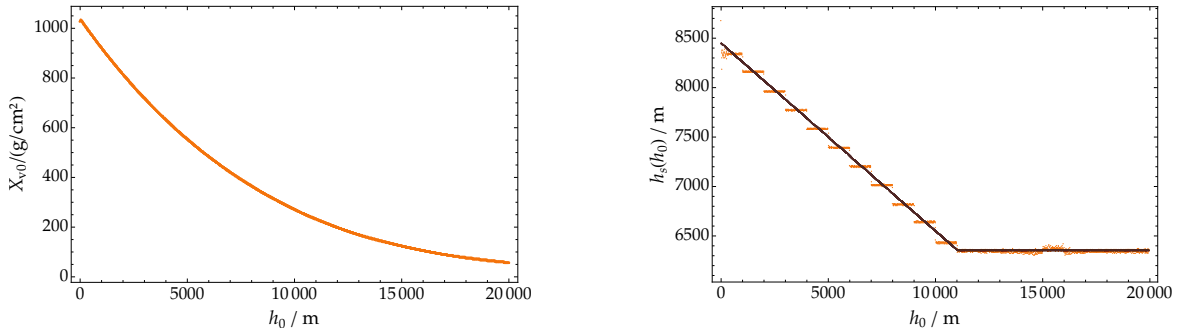
$$h = h_\mathrm{s}\,\ln\left(\frac{X_\mathrm{vg}\,\sec\theta}{X}\right) + h_\mathrm{g}\,. \tag{D.3}$$

We note that the relative derivative of the atmospheric density is

$$\frac{1}{\rho(h)}\frac{\mathrm{d}\rho(h)}{\mathrm{d}h} = -\frac{1}{h_\mathrm{s}}, \tag{D.4}$$

and thus

$$\frac{1}{X_\mathrm{v}}\frac{\mathrm{d}X_\mathrm{v}}{\mathrm{d}h} = -\frac{1}{h_\mathrm{s}}. \tag{D.5}$$



(a) (b)

Figure D.1: (a): Vertically integrated depth X_{v0} from height h_0 above sea-level up to the top of the atmosphere. (b): Local scale height $h_s(h_0)$ at height h_0 above sea-level. A piecewise linear interpolation is depicted in black. Data is extracted from the yearly average `ProfileXML`-model included in Offline.

Even though an isothermal atmosphere is not an accurate approximation, we can assume that the atmosphere can be at each particular height to a good degree approximated with a local scale height which is, following Eq. (D.5), obtained as

$$h_s(h_0) = -\left(\frac{1}{X_v}\frac{dX_v}{dh}\right)^{-1}\Bigg|_{h=h_0}. \tag{D.6}$$

The behavior of the vertically integrated depth for the atmosphere of the `FdIdealStandard` configuration, as well as its local scale height, according to Eq. (D.5), is shown in Fig. D.1.

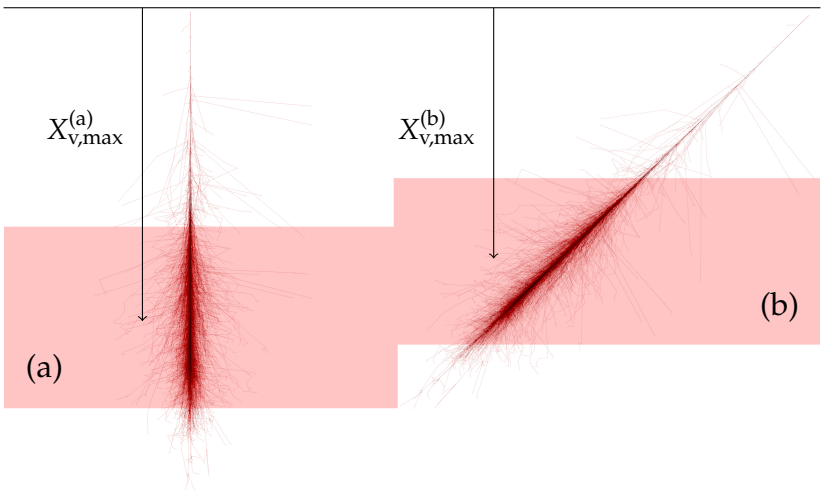


Figure D.2: Shower (a) reaches deeper layers of the atmosphere than shower (b), $X^{(a)}_{v,max} > X^{(b)}_{v,max}$, while both have roughly the same depth in terms of traversed matter and the same (slant) depth of the shower maximum, $X^{(a)}_{max} = X^{(b)}_{max}$. Shower images obtained from Ref. [121].

The atmospheric density, for example as used in Offline, can be parametrized as

$$\rho(h) = \frac{X_{vg}}{h_s}\, e^{-(h-h_g)/h_s}. \tag{D.7}$$

For vertical showers ($\theta \simeq 0°$), this parametrization is an accurate description of the relevant atmospheric profile even when using a constant scale height, since the relevant region of the shower development is thin and near the ground, where the isothermal approximation holds. More inclined showers ($\theta \gtrsim 10°$), however, deposit more of their energy in higher layers of the atmosphere since the vertical depth of the shower maximum scales with $\cos\theta$, while the total depth of the longitudinal development is independent of the zenith angle, as illustrated in Fig. D.2.

As it can be seen in Fig. D.1 (b), in the lower levels of the atmosphere, where the shower maximum $X_{\max}$ is located[1], the scale height h_s behaves linearly as a function of the height. The discontinuous behavior of h_s in Fig. D.1 results from the piece-wise log-linear nature of the tabulated atmospheric data.

To take into account the effects of the non-log-linear behavior of the density in higher layers of the atmosphere, it is possible to correct for the effect described above in the first order by substituting the scale height h_s, i.e.

$$h_s \to h_s + \frac{\mathrm{d}h_s}{\mathrm{d}h} h = h_s + \Delta h_s \, h. \tag{D.8}$$

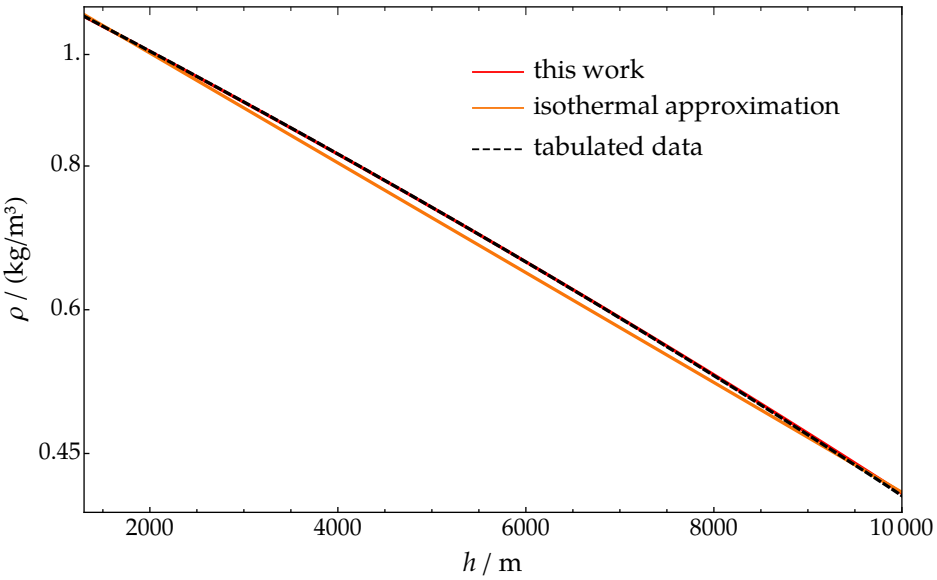
The resulting density profile is depicted in Fig. D.3. Using a progressive scale height, as it is illustrated in Fig. D.3 (b), the parametrized values of the atmospheric density are within 0.2% with respect to the tabulated atmospheric data. Δh_s is unit-less.

Using the substitution of Eq. (D.8), the integral of Eq. (D.2) becomes less elegant to solve. However, fixing the scale height according to Eq. (D.8) and evaluating this integral with a constant but corrected value, $h_s(h_{\max})$, yields reasonable accuracy to calculate the differences in slant atmospheric depths X in regions near the ground or around the height of the maximum, $h_{\max} \approx h(X_{\max})$. Since this method assumes a constant scale height along the whole shower core in the atmosphere, it might be insufficient to accurately describe the complete shower profile as seen by the FD. Nevertheless, the projected shower age parameter ΔX, depends only on the observables between the ground and $X_{\max}$ and can therefore be well approximated with a height-dependent scale height, $h_s(h_{\max})$.

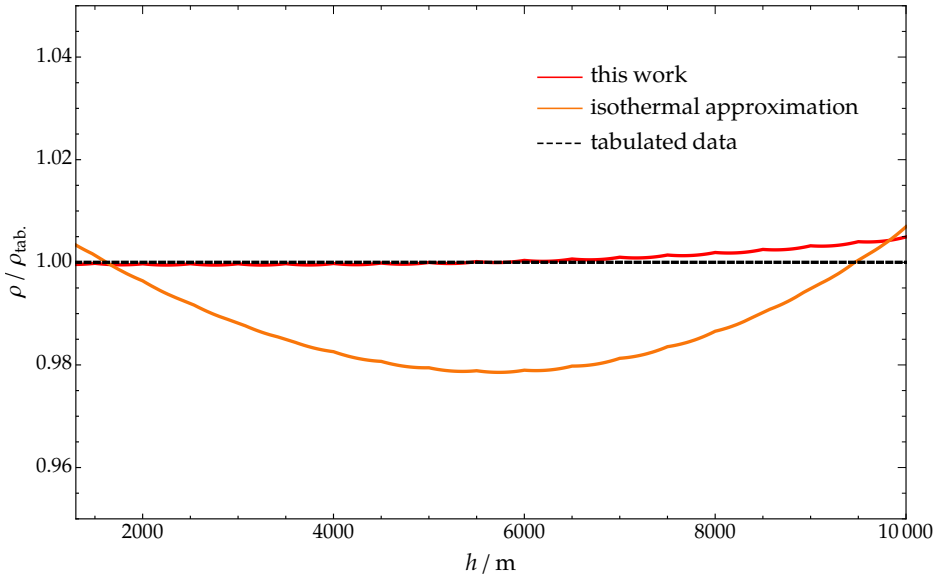
The vertical depth of the atmosphere, X_{vg} was inserted for the respective month, then a fit was performed to find the most suitable values for h_s and Δh_s. The comparison of ΔX calculated in the isothermal approximation with respect to the tabulated piece-wise exponential atmosphere is depicted in Fig. D.4. The relative deviation of the result from the analytical calculation with respect to the calculation using tabulated atmospheric data is on the order of 1% using only the correction stemming from the non-isothermal behavior of the atmosphere. Furthermore, the corrections due to a non-flat atmosphere, as described in Ref. [122], are shown to be of the second order and can thus be neglected. The results are saved in XML-format. A file containing the results for the monthly mean atmospheric profiles obtained by using

```
<configLink id="Atmosphere">
  <AtmosphereInterfaceConfig>
    ...
    <ProfileModel>  MonthlyAvgDB  </ProfileModel>
    ...
  </AtmosphereInterfaceConfig>
</configLink>
```

[1] For $\theta \leq 60°$ and $E \simeq 10^{19}$ eV the shower maximum is within ~1000 m above the ground of the Auger Observatory.



(a)



(b)

Figure D.3: (a): Density of the atmosphere as a function of the height above sea level, using tabulated data (black), an isothermal approximation (orange) and a parametrization using a progressive scale height, as presented in this work (red). (b): The same information normalized to the tabulated data.

in the Offline `bootstrap.xml` (labelled `un2`) and for the CORSIKA monthly atmospheric profiles (labelled `bariloche`) configurations is given the Universality software toolkit [135]. It is important to note that this approximate form of ΔX works well in the region where

$$h_{\text{proj}} \sec\theta = r \cos\psi \tan\theta \ll h_{\text{s}}, \tag{D.9}$$

which is the case for about two crowns in the 1500 m array and for $\theta \lesssim 50^\circ$.

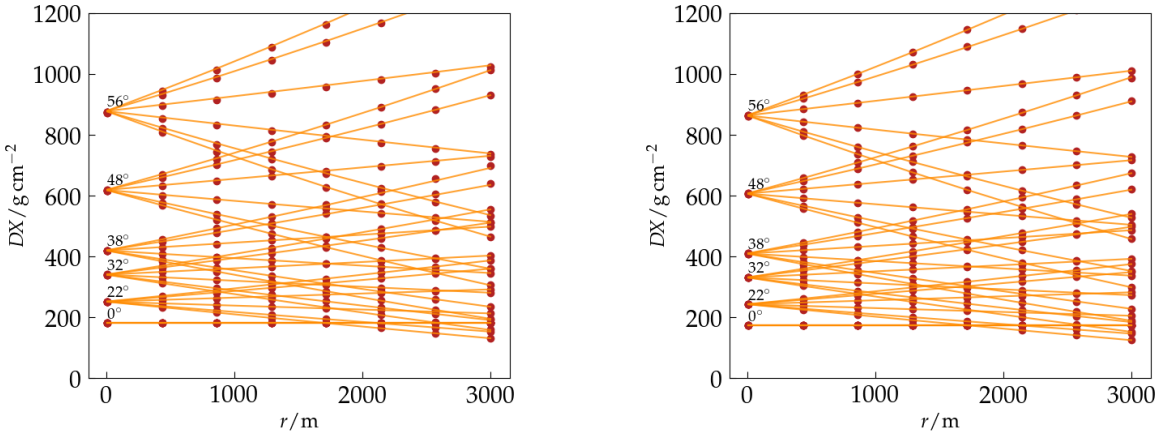


Figure D.4: The shower depth parameter ΔX (here denoted as DX) for different radii, calculated according to Eq. (5.11) (orange lines) and using a tabulated atmosphere and numeric integration in Offline (red dots) for different zenith angles θ (as indicated inside the plots) and $\psi \in \{0°, 30°, 60°, 120°, 150°, 180°\}$. Both illustrations use a mean atmospheric profile for the month of March from the `FdIdeal` configuration, depicted in panel (a), and from the atmosphere used in CORSIKA, in panel (b) respectively. The shower maximum was set to $X_{\max} = 700\,\mathrm{g/cm^2}$. Differences for equivalent data points are on the order of $\simeq 20\,\mathrm{g/cm^2}$ and less.

D A Model of the Atmosphere for Auger

Appendix E

Danksagungen

At this point I would like to take the opportunity to express my gratitude. First and foremost I would like to thank my referees Ralph Engel and Federico Sanchez for allowing me to write this work, and for greatly supporting me with their help and guidance, especially during my first days in Argentina, at the end of my thesis, and beyond the scope of this work. Markus Roth, who invited me to the Auger collaboration in the first place, always guided me through the different stages of this work with his experience and intuition and thus very much deserves my thankfulness. Additionally, I deeply enjoyed supporting him in the *Astroparticle Physics II* lecture, which was an opportunity for which I am also thankful. Furthermore, I would like to express my deep gratitude towards Darko Veberič, who became a dear friend and mentor, who spend countless hours proofreading my text and code, and with whom I eagerly discussed physics dozens of times. Another special thank-you goes to David Schmidt, who guided me especially during my first days at the institute and who was always there for help and support.

I want to say *thank you* to my great office colleagues, especially Christoph Schäfer, Steffen Hahn, and Emily Martins, for bringing joy in the work day and for bearing me through the last years. I want to thank all my colleagues in the DDAp program for the great times we had before and besides the pandemic. A special thank-you of course goes out to my dear friends in Argentina, who made me and my wife feel so welcome and at home there.

Furthermore, I would like to dearly express my gratitude towards Paolo Lipari, Alan Watson, and Thomas Gaisser, who were kind enough to spend some of their valuable time discussing physics with a foreign PhD student during the ICRC 2021.

Des Weiteren möchte ich noch meinen Freunden danken – allen voran Fritz Waitz, der sich mühsam durch die Zeilen einer unfertigen Version dieser Arbeit gekämpft hat. Vielen Dank an alle, die mir in den letzten Monaten eine Stütze waren und mich dabei ab und an abwesend und gedankenversunken aushalten mussten.

Ein besonderes Dankeschön geht natürlich an meine großartige Frau Kim, die mich wie sonst niemand durch die Höhen und Tiefen dieser Arbeit begleitet hat und deren Unterstützung mir immer Halt gegeben hat. Zum Schluss möchte ich von Herzen meiner Familie, insbesondere meinen Eltern Arnd und Helga, danken, ohne deren immerwährende Unterstützung so vieles nicht möglich gewesen wäre.

Typeset with LaTeX[158].
All analysis performed with C++, Mathematica, and python [159, 160, 161].
Illustrations created using Matplotlib [162].

BIBLIOGRAPHY

[1] R. Alves Batista et al. „Open Questions in Cosmic-Ray Research at Ultrahigh Energies“ *Front. Astron. Space Sci.* **6** (2019) p. 23 DOI: 10.3389/fspas.2019.00023 arXiv: 1903.06714 [astro-ph.HE]

[2] T. K. Gaisser, R. Engel, and E. Resconi *Cosmic Rays and Particle Physics* 2nd ed. Cambridge University Press, 2016 DOI: 10.1017/CBO9781139192194

[3] W. F. G. Swann „The History of Cosmic Rays“ *Am. J. Phys.* **29** (1961) p. 811 URL: https://doi.org/10.1119/1.1937624

[4] J. R. Hörandel „Early cosmic-ray work published in German“ (2013) DOI: 10.1063/1.4792540 URL: http://dx.doi.org/10.1063/1.4792540

[5] T. Wulf „Observations on the radiation of high penetration power on the Eiffel tower“ *Phys. Z.* **11** (1910) p. 811

[6] V. F. Hess „Über Beobachtungen der durchdringenden Strahlung bei sieben Freiballonfahrten“ *Phys. Z.* **13** (1912) p. 1084

[7] B. Rossi „On the Magnetic Deflection of Cosmic Rays“ *Phys. Rev.* **36** (1930) p. 606 DOI: 10.1103/PhysRev.36.606 URL: https://link.aps.org/doi/10.1103/PhysRev.36.606

[8] P. Freier et al. „Evidence for Heavy Nuclei in the Primary Cosmic Radiation“ *Phys. Rev.* **74** (1948) p. 213 DOI: 10.1103/PhysRev.74.213 URL: https://link.aps.org/doi/10.1103/PhysRev.74.213

[9] P. Zyla et al. „Review of Particle Physics“ *Prog. Theor. Exp. Phys.* **2020** (2020) p. 083C01 DOI: 10.1093/ptep/ptaa104

[10] B. Rossi „Über die Eigenschaften der durchdringenden Korpuskularstrahlung im Meeresniveau“ *Z. Phys.* **82** (1933) p. 151

[11] P. Auger et al. „Extensive Cosmic-Ray Showers“ *Rev. Mod. Phys.* **11** (1939) p. 288 DOI: 10.1103/RevModPhys.11.288 URL: https://link.aps.org/doi/10.1103/RevModPhys.11.288

[12] R. Engel, D. Heck, and T. Pierog „Extensive air showers and hadronic interactions at high energy“ *Ann. Rev. Nucl. Part. Sci.* **61** (2011) p. 467 DOI: 10.1146/annurev.nucl.012809.104544

[13] A. Coleman „Measurement of the Cosmic Ray Flux near the Second Knee with the Pierre Auger Observatory“ *Proc. Int. Cosmic Ray Conf.* **358** (2019) DOI: 10.22323/1.358.0225 URL: https://doi.org/10.22323/1.358.0225

[14] A. Aab et al. „Observation of a large-scale anisotropy in the arrival directions of cosmic rays above 8×10^{18}eV“ *Science* **357** (2017) p. 1266 DOI: 10.1126/science.aan4338 URL: http://dx.doi.org/10.1126/science.aan4338

[15] J. Abraham et al. „Measurement of the Energy Spectrum of Cosmic Rays above 10^{18} eV Using the Pierre Auger Observatory“ *Phys. Lett. B* **685** (2010) p. 239 DOI: 10.1016/j.physletb.2010.02.013 arXiv: 1002.1975 [astro-ph.HE]

[16] A. Aab et al. „Features of the Energy Spectrum of Cosmic Rays above 2.5×10^{18} eV Using the Pierre Auger Observatory“ *Phys. Rev. Lett.* **125** (2020) p. 121106 DOI: 10.1103/PhysRevLett.125.121106 arXiv: 2008.06488 [astro-ph.HE]

[17] A. Penzias and R. Wilson „A Measurement of Excess Antenna Temperature at 4080 Mc/s“ *Astrophys. J.* **142** (1965) p. 419 DOI: 10.1086/148307

[18] K. Greisen „End to the Cosmic-Ray Spectrum?“ *Phys. Rev. Lett.* **16** (1966) p. 748 DOI: 10.1103/PhysRevLett.16.748 URL: https://link.aps.org/doi/10.1103/PhysRevLett.16.748

[19] G. T. Zatsepin and V. A. Kuz'min „Upper Limit of the Spectrum of Cosmic Rays“ *Sov. J. Exp. Th. Phys. Lett.* **4** (1966) p. 78

[20] P. J. E. Peebles and P. J. Peebles *Principles of physical cosmology* Princeton university press, 1993

[21] R. Aloisio, V. Berezinsky, and S. Grigorieva „Analytic calculations of the spectra of ultra high energy cosmic ray nuclei. II. The general case of background radiation“ *Astropart. Phys.* **41** (2013) p. 94 DOI: 10.1016/j.astropartphys.2012.06.003 URL: http://dx.doi.org/10.1016/j.astropartphys.2012.06.003

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

[23] A. Aab et al. „Multi-Messenger Physics with the Pierre Auger Observatory“ *Front. Astron. Space Sci.* **6** (2019) p. 24 DOI: 10.3389/fspas.2019.00024 arXiv: 1904.11918 [astro-ph.HE]

[24] K.-H. Kampert and A. A. Watson „Extensive air showers and ultra high-energy cosmic rays: a historical review“ *Eur. Phys. J. H* **37** (2012) p. 359 DOI: 10.1140/epjh/e2012-30013-x URL: http://dx.doi.org/10.1140/epjh/e2012-30013-x

[25] J. Chang et al. „An excess of cosmic ray electrons at energies of 300 - 800 GeV“ *Nature* **456** (2008) p. 362 DOI: 10.1038/nature07477 URL: https://doi.org/10.1038/nature07477

[26] S. Caroff „High Statistics Measurement of the Positron Fraction in Primary Cosmic Rays with the Alpha Magnetic Spectrometer on the International Space Station“ *25th European Cosmic Ray Symposium* 2016 arXiv: 1612.09579 [astro-ph.HE]

[27] J. Sokol „Giant space magnet may have trapped antihelium, raising idea of lingering pools of antimatter in the cosmos“ *Science* (2017) DOI: 10.1126/science.aal1067 URL: https://doi.org/10.1126%2Fscience.aal1067

[28] J. Linsley „Evidence for a primary cosmic-ray particle with energy 10^{20} eV“ *Phys. Rev. Lett.* **10** (1963) p. 146 DOI: 10.1103/PhysRevLett.10.146

[29] R. M. Baltrusaitis et al. „THE UTAH FLY'S EYE DETECTOR“ *Nucl. Instrum. Meth. A* **240** (1985) p. 410 DOI: 10.1016/0168-9002(85)90658-8

[30] D. J. Bird et al. „Detection of a cosmic ray with measured energy well beyond the expected spectral cutoff due to cosmic microwave radiation“ *Astrophys. J.* **441** (1995) p. 144 DOI: 10.1086/175344 URL: http://dx.doi.org/10.1086/175344

[31] M. Takeda et al. „Extension of the cosmic ray energy spectrum beyond the predicted Greisen-Zatsepin-Kuz'min cutoff“ *Phys. Rev. Lett.* **81** (1998) p. 1163 DOI: 10.1103/PhysRevLett.81.1163 arXiv: astro-ph/9807193

[32] P. Bhattacharjee and G. Sigl „Origin and propagation of extremely high-energy cosmic rays“ *Phys. Rept.* **327** (2000) p. 109 DOI: 10.1016/S0370-1573(99)00101-5 arXiv: astro-ph/9811011

[33] J. Abraham et al. „Upper limit on the cosmic-ray photon flux above 10^{19} eV using the surface detector of the Pierre Auger Observatory“ *Astropart. Phys.* **29** (2008) p. 243 DOI: 10.1016/j.astropartphys.2008.01.003 arXiv: 0712.1147 [astro-ph]

[34] A. Haungs et al. „The air-shower experiment KASCADE-Grande“ *Nucl. Phys. B Proc. Suppl.* **196** (2009) p. 80 DOI: 10.1016/j.nuclphysbps.2009.09.012

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

[36] A. Aab et al. „The Pierre Auger Observatory Upgrade - Preliminary Design Report“ (2016) arXiv: 1604.03637 [astro-ph.IM]

[37] H. Kawai et al. „Telescope array experiment“ *Nucl. Phys. B Proc. Suppl.* **175** (2008) p. 221 DOI: 10.1016/j.nuclphysbps.2007.11.002

[38] A. V. Olinto et al. „The POEMMA (Probe of Extreme Multi-Messenger Astrophysics) observatory“ *J. Cosmol. Astropart. P.* **06** (2021) p. 7 DOI: 10.1088/1475-7516/2021/06/007 arXiv: 2012.07945 [astro-ph.IM]

[39] J. R. Hörandel „The Global Cosmic Ray Observatory - GCOS“ *Proc. Int. Cosmic Ray Conf.* **395** (2021) p. 027 DOI: 10.22323/1.395.0027

[40] A. M. Hillas „The Origin of Ultrahigh-Energy Cosmic Rays“ *Ann. Rev. Astron. Astrophys.* **22** (1984) p. 425 DOI: 10.1146/annurev.aa.22.090184.002233

[41] G. Helou et al. *NASA/IPAC Extragalactic Database* URL: https://ned.ipac.caltech.edu/

[42] J. Abraham et al. „Correlation of the highest energy cosmic rays with nearby extragalactic objects“ *Science* **318** (2007) p. 938 DOI: 10.1126/science.1151124 arXiv: 0711.2256 [astro-ph]

[43] P. Abreu et al. „Update on the Correlation of the Highest Energy Cosmic Rays with Nearby Extragalactic matter“ *Astropart. Phys.* **34** (2010) p. 314 DOI: 10.1016/j.astropartphys.2010.08.010 arXiv: 1009.1855 [astro-ph.HE]

[44] C. D. Dermer and S. Razzaque „Acceleration of Ultra-High Energy Cosmic Rays in the Colliding Shells of Blazars and GRBs: Constraints from the Fermi Gamma ray Space Telescope“ *Astrophys. J.* **724** (2010) p. 1366 DOI: 10.1088/0004-637X/724/2/1366 arXiv: 1004.4249 [astro-ph.HE]

[45] B. T. Zhang et al. „High-energy cosmic ray nuclei from tidal disruption events: Origin, survival, and implications“ *Phys. Rev. D* **96** (2017) p. 63007 DOI: 10.1103/PhysRevD.96.063007 arXiv: 1706.00391 [astro-ph.HE]

[46] G. R. Farrar and T. Piran „Tidal disruption jets as the source of Ultra-High Energy Cosmic Rays“ (2014) arXiv: 1411.0704 [astro-ph.HE]

[47] M. Del Santo et al. „The puzzling source IGR J17361–4441 in NGC 6388: a possible planetary tidal disruption event“ *Mon. Not. Roy. Astron. Soc.* **444** (2014) p. 93 DOI: 10.1093/mnras/stu1436 arXiv: 1407.5081 [astro-ph.HE]

[48] J. Guillochon and K. Auchettl *The open TDE Catalog* URL: https://TDE.space/

[49] R. C. Duncan and C. Thompson „Formation of very strongly magnetized neutron stars - implications for gamma-ray bursts“ *Astrophys. J. Lett.* **392** (1992) p. L9 DOI: 10.1086/186413

[50] K. Fang, K. Kotera, and A. V. Olinto „Newly-born pulsars as sources of ultrahigh energy cosmic rays“ *Astrophys. J.* **750** (2012) p. 118 DOI: 10.1088/0004-637X/750/2/118 arXiv: 1201.5197 [astro-ph.HE]

[51] B. P. Abbott et al. „Multi-messenger Observations of a Binary Neutron Star Merger“ *Astrophys. J. Lett.* **848** (2017) p. L12 DOI: 10.3847/2041-8213/aa91c9 arXiv: 1710.05833 [astro-ph.HE]

[52] K. Schmeiser and W. Bothe „Die harten Ultrastrahlschauer“ *Annalen Phys.* **424** (1938) p. 161

[53] W. Kolhörster, I. Matthes, and E. Weber „Gekoppelte Höhenstrahlen“ *Naturwiss.* **26** (1938) p. 576

[54] B. Rossi and K. Greisen „Cosmic-ray theory“ *Rev. Mod. Phys.* **13** (1941) p. 240 DOI: 10.1103/RevModPhys.13.240

[55] L. D. Landau and G. Rumer „The cascade theory of electronic showers“ *P. Roy. Soc. Lond. A. Mat.* **166** (1938) p. 213

[56] L. Landau „The angular distribution of the shower particles“ *J. Phys. USSR* **3** (1940) p. 237

[57] H. S. Snyder „Comparison of Calculations on Cascade Theory“ *Phys. Rev.* **76** (1949) p. 1563 DOI: 10.1103/PhysRev.76.1563 URL: https://link.aps.org/doi/10.1103/PhysRev.76.1563

[58] A. Ueda and N. Ogita „On the Structure of Extensive Air Showers“ *Prog. Theor. Phys.* **18** (1957) p. 269

[59] N. Ogita „A Study on the Extensive Air Showers“ *Prog. Theor. Phys.* **27** (1962) p. 105

[60] L. Dedenko and G. Zatsepin „Extensive air showers and cascades process“ *Proceedings of the 6-th ICRC, Moscow* (1960) p. 201

[61] T. Cranshaw and A. Hillas „A model for the interpretation of EAS data“ *Proceedings of the 6-th ICRC, Moscow* (1960) p. 210

[62] K. Greisen „The Extensive Air Showers“ *Progress in Cosmic Ray Physics* **3** (1956) p. 3

[63] P. Lipari „The Concepts of 'Age' and 'Universality' in Cosmic Ray Showers“ *Phys. Rev. D* **79** (2009) p. 063001 DOI: 10.1103/PhysRevD.79.063001 arXiv: 0809.0190 [astro-ph]

[64] J. Nishimura and K. Kamata „The lateral and angular distribution of cascade showers“ *Prog. Theor. Phys.* **5** (1950) p. 899

[65] J. Nishimura and K. Kamata „On the Theory of Cascade Showers, I“ *Prog. Theor. Phys.* **7** (1952) p. 185 DOI: 10.1143/ptp/7.2.185 URL: https://doi.org/10.1143/ptp/7.2.185

[66] K. Kamata and J. Nishimura „The Lateral and the Angular Structure Functions of Electron Showers“ *Prog. Theor. Phys. Supplement* **6** (1958) p. 93 DOI: 10.1143/PTPS.6.93 eprint: https://academic.oup.com/ptps/article-pdf/doi/10.1143/PTPS.6.93/5270594/6-93.pdf URL: https://doi.org/10.1143/PTPS.6.93

[67] K. Greisen „Cosmic Ray Showers“ *Ann. Rev. Nucl. Sci.* **10** (1960) p. 63 DOI: 10.1146/annurev.ns.10.120160.000431 URL: https://doi.org/10.1146/annurev.ns.10.120160.000431

[68] T. Antoni et al. „Electron, muon, and hadron lateral distributions measured in air showers by the KASCADE experiment“ *Astropart. Phys.* **14** (2001) p. 245 DOI: https://doi.org/10.1016/S0927-6505(00)00125-0 URL: https://www.sciencedirect.com/science/article/pii/S0927650500001250

[69] J. Abraham et al. „Atmospheric effects on extensive air showers observed with the Surface Detector of the Pierre Auger Observatory“ *Astropart. Phys.* **32** (2009) p. 89

[70] A. Aab et al. „Data-driven estimation of the invisible energy of cosmic ray showers with the Pierre Auger Observatory“ *Phys. Rev. D* **100** (2019) p. 082003 DOI: 10.1103/PhysRevD.100.082003 arXiv: 1901.08040 [astro-ph.IM]

[71] A. G. Mariazzi „Determination of the invisible energy of extensive air showers from the data collected at Pierre Auger Observatory“ *EPJ Web Conf.* **210** (2019) p. 02010 DOI: 10.1051/epjconf/201921002010 arXiv: 1905.04178 [astro-ph.HE]

[72] W. Heitler *The quantum theory of radiation* vol. 5 International Series of Monographs on Physics Oxford University Press, 1936

[73] J. Matthews „A Heitler model of extensive air showers“ *Astropart. Phys.* **22** (2005) p. 387 DOI: 10.1016/j.astropartphys.2004.09.003

[74] L. Cazon, R. Conceição, and F. Riehn „Probing the energy spectrum of hadrons in proton air interactions at ultrahigh energies through the fluctuations of the muon content of extensive air showers“ *Phys. Lett. B* **784** (2018) p. 68 DOI: 10.1016/j.physletb.2018.07.026 arXiv: 1803.05699 [hep-ph]

[75] J. Engel et al. „Nucleus-nucleus collisions and interpretation of cosmic ray cascades“ *Phys. Rev. D* **46** (1992) p. 5013 DOI: 10.1103/PhysRevD.46.5013

[76] K.-H. Kampert and M. Unger „Measurements of the Cosmic Ray Composition with Air Shower Experiments“ *Astropart. Phys.* **35** (2012) p. 660 DOI: 10.1016/j.astropartphys.2012.02.004 arXiv: 1201.0018 [astro-ph.HE]

[77] J. Knapp et al. „Extensive air shower simulations at the highest energies“ *Astropart. Phys.* **19** (2003) p. 77 DOI: 10.1016/S0927-6505(02)00187-1 arXiv: astro-ph/0206414

[78] A. M. Hillas „Angular and energy distributions of charged particles in electron-photon cascades in air“ *J. Phys. G* **8** (1982) p. 1461 DOI: 10.1088/0305-4616/8/10/016

[79] J. Allen et al. „Air Shower Simulation and Hadronic Interactions“ *EPJ Web Conf.* **53** (2013) p. 01007 DOI: 10.1051/epjconf/20135301007 arXiv: 1306.6090 [astro-ph.HE]

[80] D. Heck et al. „CORSIKA: A Monte Carlo code to simulate extensive air showers“ *Report FZKA* **6019** (1998)

[81] T. Bergmann et al. „One-dimensional Hybrid Approach to Extensive Air Shower Simulation“ *Astropart. Phys.* **26** (2007) p. 420 DOI: 10.1016/j.astropartphys.2006.08.005 arXiv: astro-ph/0606564

[82] H. J. Drescher et al. „Parton based Gribov-Regge theory“ *Phys. Rept.* **350** (2001) p. 93 DOI: 10.1016/S0370-1573(00)00122-8 arXiv: hep-ph/0007198

[83] T. Pierog et al. „EPOS LHC: Test of collective hadronization with data measured at the CERN Large Hadron Collider“ *Phys. Rev. C* **92** (2015) p. 034906 DOI: 10.1103/PhysRevC.92.034906 arXiv: 1306.0121 [hep-ph]

[84] S. Ostapchenko „Monte Carlo treatment of hadronic interactions in enhanced Pomeron scheme: I. QGSJET-II model“ *Phys. Rev. D* **83** (2011) p. 014018 DOI: 10.1103/PhysRevD.83.014018 arXiv: 1010.1869 [hep-ph]

[85] R. S. Fletcher et al. „SIBYLL: An Event generator for simulation of high-energy cosmic ray cascades“ *Phys. Rev. D* **50** (1994) p. 5710 DOI: 10.1103/PhysRevD.50.5710

[86] C. J. Todero Peixoto, V. de Souza, and J. A. Bellido „Comparison of the moments of the Xmax distribution predicted by different cosmic ray shower simulation models“ *Astropart. Phys.* **47** (2013) p. 18 DOI: 10.1016/j.astropartphys.2013.05.011 arXiv: 1301.5555 [astro-ph.HE]

[87] S. Lafebre et al. „Universality of electron-positron distributions in extensive air showers“ *Astropart. Phys.* **31** (2009) p. 243 DOI: 10.1016/j.astropartphys.2009.02.002 arXiv: 0902.0548 [astro-ph.HE]

[88] F. Nerling et al. „Universality of electron distributions in high-energy air showers – Description of Cherenkov light production“ *Astropart. Phys.* **24** (2006) p. 421 DOI: https://doi.org/10.1016/j.astropartphys.2005.09.002 URL: https://www.sciencedirect.com/science/article/pii/S0927650505001325

[89] M. Giller et al. „Similarity of extensive air showers with respect to the shower age“ *J. Phys. G* **31** (2005) p. 947 DOI: 10.1088/0954-3899/31/8/023

[90] The Pierre Auger Collaboration *Homepage* URL: https://auger.org

[91] A. Aab et al. „The Pierre Auger Cosmic Ray Observatory“ *Nucl. Instrum. Meth. A* **798** (2015) p. 172 DOI: 10.1016/j.nima.2015.06.058 arXiv: 1502.01323 [astro-ph.IM]

[92] J. Abraham et al. „The Fluorescence Detector of the Pierre Auger Observatory“ *Nucl. Instrum. Meth. A* **620** (2010) p. 227 DOI: 10.1016/j.nima.2010.04.023 arXiv: 0907.4282 [astro-ph.IM]

[93] I. Allekotte et al. „The Surface Detector System of the Pierre Auger Observatory“ *Nucl. Instrum. Meth. A* **586** (2008) p. 409 DOI: 10.1016/j.nima.2007.12.016 arXiv: 0712.2832 [astro-ph]

[94] R. Pesce „Energy calibration of data recorded with the surface detectors of the Pierre Auger Observatory: An update“ *Proc. Int. Cosmic Ray Conf.* **2** (2011) p. 214 DOI: 10.7529/ICRC2011/V02/1160

[95] X. Bertou et al. „Calibration of the surface array of the Pierre Auger Observatory“ *Nucl. Instrum. Meth. A* **568** (2006) p. 839 DOI: 10.1016/j.nima.2006.07.066 arXiv: 2102.01656 [astro-ph.HE]

[96] D. Veberic „Estimation of the Total Signal in Saturated Stations of Pierre Auger Surface Detector“ *Proc. Int. Cosmic Ray Conf.* **1** (2013) p. 633 URL: https://ui.adsabs.harvard.edu/abs/2013ICRC...33.2193V/abstract

[97] A. Aab et al. „Reconstruction of events recorded with the surface detector of the Pierre Auger Observatory“ *J. Instrum.* **15** (2020) p. 10021 DOI: 10.1088/1748-0221/15/10/P10021 arXiv: 2007.09035 [astro-ph.IM]

[98] P. Billoir *Reconstruction of showers with the Ground Array: status of the "prototype" program* Auger internal note GAP-2000-025 2000 URL: https://www.auger.org/document-centre2/download/16-gap-notes-2000/1680-gap2000-025

[99] M. Nagano and A. A. Watson „Observations and implications of the ultrahigh-energy cosmic rays“ *Rev. Mod. Phys.* **72** (2000) p. 689 DOI: 10.1103/RevModPhys.72.689

[100] A. Aab et al. „Measurement of the cosmic-ray energy spectrum above 2.5×10^{18} eV using the Pierre Auger Observatory“ *Phys. Rev. D* **102** (2020) p. 62005 DOI: 10.1103/PhysRevD.102.062005 arXiv: 2008.06486 [astro-ph.HE]

[101] M. Unger et al. „Reconstruction of Longitudinal Profiles of Ultra-High Energy Cosmic Ray Showers from Fluorescence and Cherenkov Light Measurements“ *Nucl. Instrum. Meth. A* **588** (2008) p. 433 DOI: 10.1016/j.nima.2008.01.100 arXiv: 0801.4309 [astro-ph]

[102] P. Abreu et al. „The exposure of the hybrid detector of the Pierre Auger Observatory“ *Astropart. Phys.* **34** (2011) p. 368

[103] J. Abraham et al. „Observation of the suppression of the flux of cosmic rays above 4×10^{19} eV“ *Phys. Rev. Lett.* **101** (2008) p. 061101 DOI: 10.1103/PhysRevLett.101.061101 arXiv: 0806.4302 [astro-ph]

[104] A. Aab et al. „Depth of Maximum of Air-Shower Profiles at the Pierre Auger Observatory: Measurements at Energies above $10^{17.8}$ eV“ *Phys. Rev. D* **90** (2014) p. 122005 DOI: 10.1103/PhysRevD.90.122005 arXiv: 1409.4809 [astro-ph.HE]

[105] A. Aab et al. „Depth of maximum of air-shower profiles at the Pierre Auger Observatory. II. Composition implications“ *Phys. Rev. D* **90** (2014) p. 122006 DOI: 10.1103/PhysRevD.90.122006 arXiv: 1409.5083 [astro-ph.HE]

[106] A. Aab et al. „Muons in Air Showers at the Pierre Auger Observatory: Mean Number in Highly Inclined Events“ *Phys. Rev. D* **91** (2015) p. 032003 DOI: 10.1103/PhysRevD.91.032003 arXiv: 1408.1421 [astro-ph.HE]

[107] A. Aab et al. „Measurement of the Fluctuations in the Number of Muons in Extensive Air Showers with the Pierre Auger Observatory“ *Phys. Rev. Lett.* **126** (2021) p. 152002 DOI: 10.1103/PhysRevLett.126.152002 arXiv: 2102.07797 [hep-ex]

[108] C. Baus et al. „Anomalous Longitudinal Shower Profiles and Hadronic Interactions“ **2** (2011) p. 206 DOI: 10.7529/ICRC2011/V02/1130 arXiv: 1111.0504 [astro-ph.HE]

[109] M. Ave et al. „A generalized description of the signal size in extensive air shower detectors and its applications“ *Astropart. Phys.* **87** (2017) p. 23 DOI: https://doi.org/10.1016/j.astropartphys.2016.11.008 URL: https://www.sciencedirect.com/science/article/pii/S092765051630192X

[110] R. Ulrich, R. Engel, and M. Unger „Hadronic Multiparticle Production at Ultra-High Energies and Extensive Air Showers“ *Phys. Rev. D* **83** (2011) p. 054026 DOI: 10.1103/PhysRevD.83.054026 arXiv: 1010.4310 [hep-ph]

[111] R. Conceicao et al. „The interplay between the electromagnetic and the muonic longitudinal profile at production“ *EPJ Web Conf.* **52** (2013) p. 03004 DOI: 10.1051/epjconf/20125203004 arXiv: 1301.0507 [hep-ph]

[112] M. Ave et al. „Extensive Air Shower Universality of Ground Particle Distributions“ *Proc. Int. Cosmic Ray Conf.* International Cosmic Ray Conference **2** (2011) p. 178 DOI: 10.7529/ICRC2011/V02/1025

[113] F. Schmidt et al. „A Model-Independent Method of Determining Energy Scale and Muon Number in Cosmic Ray Surface Detectors“ *Astropart. Phys.* **29** (2008) p. 355 DOI: 10.1016/j.astropartphys.2008.03.004 arXiv: 0712.3750 [astro-ph]

[114] M. Ave et al. *Prediction of the tank response $S(r, DX, E|\theta, \psi)$ from shower universality* Auger internal note GAP-2011-087 2011 URL: https://www.auger.org/document-centre2/download/3-gap-notes-2011/224-gap2011-087

[115] D. Heck and R. Engel „The EHISTORY Option of the Air-Shower Simulation Program CORSIKA“ *Report FZKA* **7495** (2009)

[116] P. Lipari „Universality in the longitudinal development of Cosmic Ray showers“ *Nucl. Part. Phys. Proc.* **279** (2016) ed. by G. Cataldi, I. De Mitri, and D. Martello p. 111 DOI: 10.1016/j.nuclphysbps.2016.10.016

[117] T. K. Gaisser and A. M. Hillas „Reliability of the Method of Constant Intensity Cuts for Reconstructing the Average Development of Vertical Showers“ *Proc. Int. Cosmic Ray Conf.* **8** (1977) p. 353 URL: https://ui.adsabs.harvard.edu/abs/1977ICRC....8..353G

[118] A. Aab et al. „Measurement of the average shape of longitudinal profiles of cosmic-ray air showers at the Pierre Auger Observatory“ *J. Cosmol. Astropart. P.* **03** (2019) p. 018 DOI: 10.1088/1475-7516/2019/03/018 arXiv: 1811.04660 [astro-ph.HE]

[119] D. Veberič „Lambert W function for applications in physics“ *Computer Physics Communications* **183** (2012) p. 2622

[120] M. F. Bourdeau, J. N. Capdevielle, and J. Procureur „The lateral age parameter in extensive air showers“ *J. Phys. G* **6** (1980) p. 901 DOI: 10.1088/0305-4616/6/7/013

[121] F. Schmidt and J. Knapp *CORSIKA Shower Images* 2005 URL: https://www-zeuthen.desy.de/~jknapp/fs/showerimages.html/

[122] M. Stadelmaier et al. *A Useful Model of the Atmosphere* Auger internal note GAP-2021-010 2021 URL: https://www.auger.org/gap-notes/download/145-gap-notes-2021/5278-gap2021-010

[123] A. Aab et al. „Muons in Air Showers at the Pierre Auger Observatory: Measurement of Atmospheric Production Depth“ *Phys. Rev. D* **90** (2014) p. 012012 DOI: 10.1103/PhysRevD.90.012012 arXiv: 1407.5919 [hep-ex]

[124] L. Cazon, R. A. Vazquez, and E. Zas „Depth development of extensive air showers from muon time distributions“ *Astropart. Phys.* **23** (2005) p. 393 DOI: 10.1016/j.astropartphys.2005.01.009 arXiv: astro-ph/0412338

[125] M. Pothast and M. Stadelmaier *An analytical approach to the time-dependent muon signal in the surface detectors* Auger internal note GAP-2021-057 2021 URL: https://www.auger.org/document-centre2/download/145-gap-notes-2021/5389-gap2021-057

[126] M. Stadelmaier et al. *A model of the time dependent signal of the surface detector* Auger internal note GAP-2021-011 2021 URL: https://www.auger.org/document-centre2/download/145-gap-notes-2021/5279-gap2021-011

[127] M. Ave, M. Roth, and A. Schulz „A generalized description of the time dependent signals in extensive air shower detectors and its applications“ *Astropart. Phys.* **88** (2017) p. 46 DOI: https://doi.org/10.1016/j.astropartphys.2017.01.003

[128] A. Schulz *Measurement of the Energy Spectrum and Mass Composition of Ultra-high Energy Cosmic Rays* Dissertation Karlsruhe Institute of Technology, 2016

[129] L. Cazon et al. „Time structure of muonic showers“ *Astropart. Phys.* **21** (2004) p. 71 DOI: 10.1016/j.astropartphys.2003.12.009 arXiv: astro-ph/0311223

[130] A. Aab *Air Shower Signal Shape Analysis: A New Method to Separate Ultra-High-Energy Photons from Cosmic Rays* Dissertation Radboud University Nijmegen, 2020

[131] D. Maurel *Mass composition of ultra-high energy cosmic rays based on air shower universality* Dissertation Karlsruhe Institute of Technology, 2013

[132] A. Bridgeman *Determining the Mass Composition of Ultra-high Energy Cosmic Rays Using Air Shower Universality* Dissertation Karlsruhe Institute of Technology, 2018

[133] J. Hulsman *Hybrid Universality Model Development and Air Shower Reconstruction for the Pierre Auger Observatory* Dissertation Karlsruhe Institute of Technology, 2020

[134] S. Argiro et al. „The Offline Software Framework of the Pierre Auger Observatory“ *Nucl. Instrum. Meth. A* **580** (2007) p. 1485 DOI: 10.1016/j.nima.2007.07.010 arXiv: 0707.1652 [astro-ph]

[135] M. Stadelmaier *universality-v2 repository* 2021 URL: https://https://gitlab.ikp.kit.edu/shower-universality/universality-v2

[136] M. Stadelmaier et al. „A complete model of the signal in surface detector arrays and its application for the reconstruction of mass-sensitive observables“ *Proc. Int. Cosmic Ray Conf.* **395** (2021) p. 432 DOI: 10.22323/1.395.0432

[137] D. Schmidt „AugerPrime implementation in the Offline simulation and reconstruction framework“ *Proc. Int. Cosmic Ray Conf.* **301** (2017) p. 185 DOI: 10.22323/1.301.0353 URL: https://pos.sissa.it/301/353

[138] B. T. Stokes et al. „Dethinning Extensive Air Shower Simulations“ *Astropart. Phys.* **35** (2012) p. 759 DOI: 10.1016/j.astropartphys.2012.03.004 arXiv: 1104.3182 [astro-ph.IM]

[139] The HDF Group *Hierarchical data format version 5* URL: http://www.hdfgroup.org/HDF5

[140] Q. Luce et al. „On the need for unbiasing azimuthal asymmetries in signals measured by surface detector arrays“ *Proc. Int. Cosmic Ray Conf.* **395** (2021) p. 435 DOI: 10.22323/1.395.0435

[141] R. Bardenet, B. Kégl, and D. Veberič *Single muon response: The signal model* Auger internal note GAP-2010-110 2010 URL: https://www.auger.org/gap-notes/download/6-gap-notes-2010/804-gap2010-110

[142] L. M. Bueno, P. Billoir, and I. Maris *Signal variance for the TOTd and MoPS triggers* Auger internal note GAP-2014-035 2014 URL: https://www.auger.org/document-centre2/download/23-gap-notes-2014/1931-gap2014-035

[143] T. Schulz *Enhanced UHECR Event Reconstruction by means of Sampling Lateral Distributions with Multiple Surface Sub-Detectors* Auger internal note GAP-2018-025 2018 URL: https://www.auger.org/gap-notes/download/107-gap-notes-2018/4657-gap2018-025

[144] T. Schulz et al. *Propagating SD Core Uncertainties in Subordinate Reconstructions* Auger internal note GAP-2020-068 2020 URL: https://www.auger.org/document-centre2/download/144-gap-notes-2020/5224-gap2020-068

[145] Á. Taboada Núñez *Analysis of the First Data of the AugerPrime Detector Upgrade* Dissertation Karlsruhe Institute of Technology, 2020 DOI: 10.5445/IR/1000104548

[146] F. James and M. Roos „Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations“ *Comput. Phys. Commun.* **10** (1975) p. 343 DOI: 10.1016/0010-4655(75)90039-9

[147] J. Glombitza and M. Erdmann *Deep Neural Network for the Reconstruction of the Shower Maximum Xmax using the Water-Cherenkov Detectors of the Pierre Auger Observatory* Auger internal note GAP-2020-005 2020 URL: https://www.auger.org/gap-notes/download/144-gap-notes-2020/5068-gap2020-005

[148] P. Younk and M. Risse „Sensitivity of the correlation between the depth of shower maximum and the muon shower size to the cosmic ray composition“ *Astropart. Phys.* **35** (2012) p. 807 DOI: 10.1016/j.astropartphys.2012.03.001 arXiv: 1203.3732 [astro-ph.HE]

[149] M. Novotný *On the muon-dependent elongation rate of extensive air showers* Auger internal note GAP-2022-006 2022 URL: https://www.auger.org/gap-notes

[150] A. A. Lagutin, R. I. Raikin, and T. L. Serebryakova „Air shower universality in the energy range of 10^{14} to 10^{22} eV“ *Bull. Russ. Acad. Sci. Phys.* (2013) p. 623 DOI: 10.3103/S1062873813050353

[151] P. Abreu et al. „Combined fit of the energy spectrum and mass composition across the ankle with the data measured at the Pierre Auger Observatory“ *Proc. Int. Cosmic Ray Conf.* **395** (2021) p. 311 DOI: 10.22323/1.395.0311

[152] P. Billoir *What is ageing in the tanks of the Surface Detector?* Auger internal note GAP-2014-038 2014 URL: https://www.auger.org/gap-notes/download/23-gap-notes-2014/1934-gap2014-038

[153] B. Dawson and V. Verzi *Update on the Auger Energy Scale* Auger internal note GAP-2010-120 2010 URL: https://www.auger.org/gap-notes/download/6-gap-notes-2010/814-gap2010-120

[154] A. Yushkov, P. A. Collaboration, et al. „Mass Composition of Cosmic Rays with Energies above $10^{17.2}$ eV from the Hybrid Data of the Pierre Auger Observatory“ *Proc. Int. Cosmic Ray Conf.* **358** (2019) p. 482

[155] M. Stadelmaier, D. Veberič, and M. Roth *Systematic Effects of the 5T5 Posterior Trigger* Auger internal note GAP-2019-024 2019 URL: https://www.auger.org/gap-notes/download/142-gap-notes-2019/4900-gap2019-024

[156] *Magellan in the Auger Open Data Release* 2021 URL: https://opendata.auger.org/display.php?evid=151433402300

[157] I. Tamm and S. Belenky „The Energy Spectrum of Cascade Electrons“ *Phys. Rev.* **70** (1946) p. 660 DOI: 10.1103/PhysRev.70.660 URL: https://link.aps.org/doi/10.1103/PhysRev.70.660

[158] L. Lamport „A Document Preparation System LATEX“ *User's guide & Reference manual* (1986)

[159] B. Stroustrup *The C++ programming language* Pearson Education India, 2000

[160] W. R. Inc. *Mathematica, Version 12.0*

[161] G. Van Rossum and F. L. Drake *Python 3 Reference Manual* Scotts Valley, CA: CreateSpace, 2009

[162] J. D. Hunter „Matplotlib: A 2D graphics environment“ *Computing in Science & Engineering* **9** (2007) p. 90 DOI: 10.1109/MCSE.2007.55