Measurement of the Energy Spectrum and Mass Composition of Ultra-high Energy Cosmic Rays

Zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

an der Fakultät für Physik des
Karlsruher Instituts für Technologie (KIT)

genehmigte

Dissertation

von

Dipl.-Phys. Alexander Schulz
aus Kandel

Tag der mündlichen Prüfung: 29. Januar 2016

Druckdatum dieser Version: 08. März 2016

Referent: Prof. Dr. Dr. h.c. Johannes Blümer

Korreferent: Prof. Dr. Marc Weber

Betreuer: Dr. Markus Roth

Abstract

The physics of cosmic rays has been a subject of study for almost one hundred years, and, throughout this time, our understanding of the measured spectrum and theoretical ideas of the physics behind it have vastly improved. We have ideas for potential sources (for example supernova shock fronts, gamma-ray bursts and active galactic nuclei) as well as acceleration mechanisms. As with human-made accelerators, we are progressing towards measuring even higher energies. This progress is mainly lead by an increase in understanding and new technological possibilities. Nowadays, the highest energetic cosmic rays with energies of 10^{20} eV are still far outside the grasp of the Large Hadron Colliders (roughly a factor of one hundred between the center-of-mass energies) and, in addition to astrophysical questions, we want to understand more about particle interactions and cross-sections at the highest energies. To understand the origin of the highest energetic particles, we need to understand more about their mass composition and energy spectrum. These are the main topics that are discussed in this work. Two things are of major importance for the study of UHECRs[1] and air showers. Firstly, we need detectors with cutting edge technology, and, secondly, we need a vast number of detectors spread over an enormous area to measure particles at fluxes as low as one particle per century and square kilometer. The *Pierre Auger Observatory* (Auger) is a powerful detector that covers an area of more than 3000 km^2 in Argentina. Both surface and fluorescence techniques are employed as complementary methods to measure extensive air showers. In this work, I present studies that make use of the full set of data measured with Auger[2] from 2004 until 2014.

The Auger fluorescence detector directly observes the longitudinal development of air showers. In particular, the depth of shower maximum is measured, from which information about the primary mass is inferred. However, it has a limited duty cycle of about 15 % and the event statistics at highest energies are vastly limited. Within this work, I extend and make use of a new analysis technique to infer information about the mass composition from data measured with the surface detector at a duty cycle of nearly 100 %. The new analysis technique is based on the paradigm of shower universality. Within this method, the unique properties of different particle components are exploited to describe air showers as a function of their primary energy, mass and geometry. A major part of this work was to advance the study of a model to describe the temporal structure of secondary particles on the ground as measured by Auger surface detectors. The method is easily extendable to other detector types and is of essential importance for the upgrade of Auger and future analyses. The derived results on the mass composition of UHECRs are compatible with the

[1] ultra-high energy cosmic rays
[2] Pierre Auger Observatory

results obtained from fluorescence measurements in the common energy range. Additionally, the results indicate a trend towards a lighter composition at the very end of the energy spectrum.

The standard reconstruction of air showers is based on an empirical description of the overall lateral distribution of secondary particles on the ground. For each event, the expected lateral size at a certain distance to the shower core is a robust estimator of the primary energy. It is calibrated using a subset of events that are measured with both surface and fluorescence detectors. The fluorescence detector provides a calorimetric measurement of the primary energy. Before the calibration, the lateral sizes depend strongly on the shower inclination. An average correction based on the Constant Intensity Cut method is applied to correct for this dependence. In the course of this work, these analysis steps were studied for both the standard surface detector and an extension of nested detectors to detect primaries at lower energies. A consistent picture of the energy spectrum over three decades of energy from these two measurements is derived. The results extend previous analyses and are compatible with previous results. A novel search for a dependence of the flux of UHECRs on the incoming direction was undertaken. The results are consistent with a large-scale anisotropy in the arrival direction of particles as obtained with an independent analysis. To achieve an unbiased outcome, the measured energies were corrected for distortions due to atmospheric effects and the influence of the geomagnetic field. As a novel and complementary approach, the flux of UHECRs was derived with the shower universality approach, yielding a similar result to the standard method. This sets an important basis for future analyses of spectra for different mass groups, possibly making use of data from the upgraded detectors of AugerPrime.

Zusammenfassung

Seit ihrer Entdeckung vor etwa 100 Jahren wird die kosmische Strahlung als essentielle Schnittstelle zwischen Astro- und Teilchenphysik immer weiter und tiefer untersucht. Unser Verständnis des Teilchenflusses und theoretischer Ideen zu dessen Erklärung hat sich stetig weiterentwickelt. Heutzutage steht fest, dass die höchstenergetischen Teilchen mit Primärenergien von mehr als 10^{18} eV ihren Ursprung in den gewaltigsten und energiereichsten Objekten des Universums finden müssen, zum Beispiel in supermassiven schwarzen Löchern oder sogenannten Gamma-Ray-Bursts. Durch unser sich stetig verbesserndes Verständnis sowie neue technologische Möglichkeiten wurden in den letzten Jahren immer mehr Messungen bei den höchsten Teilchenenergien ermöglicht. Dabei übersteigen die Schwerpunktenergien, die bei Wechselwirkungen der höchstenergetischen Teilchen mit Energien von 10^{20} eV in unserer Atmosphäre auftreten, die im modernsten Teilchenbeschleuniger auf der Erde, dem LHC, realisierbaren Energien, noch um einen Faktor hundert. Gerade deshalb ist die Untersuchung kosmischer Strahlung für die Teilchenphysik von höchster Bedeutung. Eine entscheidende Frage der Astrophysik dreht sich um den Ursprung und die Beschleunigungsmechanismen der höchstenergetischen Teilchen. Um diese Frage zu beantworten, ist eine genaue Messung des Energiespektrums und der Massenzusammensetzung der Teilchen von grundlegender Bedeutung. Die Bestimmung dieser Größen mit modernen Analysemethoden stellt das Thema dieser Arbeit dar. Um Messungen bei den höchsten Energien und kleinsten Teilchenflüssen (weniger als ein Teilchen pro Jahrhundert und Quadratkilometer) zu ermöglichen, ist eine gewaltige Detektorfläche nötig. Das Pierre-Auger-Observatorium in Argentinien ist der größte Detektor für ultrahochenergetische kosmische Strahlung mit einer Detektorfläche von mehr als $3000\,\mathrm{km}^2$. Dabei kommt sowohl ein Oberflächendetektor, bestehend aus mehr als 1600 Wasser-Cherenkov-Detektoren, als auch ein Fluoreszenzdetektor zur Messung von Luftschauern zum Einsatz. Diese Arbeit beschäftigt sich mit der Rekonstruktion und Analyse von Daten, die zwischen 2004 und 2014 mit dem Pierre-Auger-Observatorium gemessen wurden.

Der Fluoreszenzdetektor erlaubt die direkte Beobachtung der longitudinalen Entwicklung eines Luftschauers, insbesondere des Maximums der Schauerentwicklung, welches stark mit der Primärmasse korreliert ist. Allerdings sind Messungen mit dem Fluoreszenzdetektor nur an klaren, mondlosen Nächsten möglich, weshalb nur in etwa 15 % der Zeit eine Messung stattfindet. Dadurch ist die Statistik im höchsten und besonders interessanten Energiebereich oberhalb von $10^{19.5}$ eV sehr begrenzt. Basierend auf dem Prinzip der Schauer-Universalität wurde eine neue Analysemethode weiterentwickelt, die es erlaubt Rückschlüsse über die Primärmasse aus Messungen des Oberflächendetektors abzuleiten. Dieser misst Ereignisse fast 100 % der Zeit. Neben der Weiterentwicklung der Ereignisrekonstruktion,

war die Modellierung der zeitlichen Struktur verschiedener Teilchenkomponenten im Oberflächendetektor ein wichtiger Bestandteil dieser Arbeit. Die entwickelten Methoden sind von fundamentaler Bedeutung für zukünftige Analysen mit neuen Detektortypen, zum Beispiel den neuen Detektoren von AugerPrime. Die im Rahmen dieser Arbeit aus Daten des Oberflächendetektors abgeleitete Massenzusammensetzung der ultrachochenergetischen Strahlung stimmt sehr gut mit Ergebnissen der Fluoreszenzmessung überein. Zusätzlich gibt es Hinweise dafür, dass die Massenzusammensetzung bei höchsten Energien leichter ist als erwartet. Obwohl diese Beobachtung erst eine geringe statistische Signifikanz von $2\,\sigma$ aufweist, ist sie von großer Bedeutung für die Astrophysik und Teilchen-Astronomie bei höchsten Energien.

Die Standardrekonstruktion der mit dem Oberflächendetektor gemessenen Luftschauer basiert auf der empirischen Beschreibung der Lateralverteilung aller Sekundärteilchen eines Luftschauers, die den Boden erreichen. Die Größe der Lateralverteilung bei einem bestimmten Abstand zum Zentrum des Luftschauers stellt für jedes gemessene Ereignis eine robuste Schätzung der Primärenergie dar. Die Zenitwinkelabhängigkeit dieses Energieschätzers wird mithilfe der Constant-Intensity-Cut-Methode entfernt. Die daraus abgeleiteten Energieschätzer werden mit Energien kalibriert, die mit dem Fluoreszenzdetektor bestimmt wurden. Im Kontext dieser Arbeit wurden diese Analysemethoden weiterentwickelt und auf die Messungen mit Oberflächendetektoren angewandt. Während mit dem normalen Oberflächenfeld des Pierre-Auger-Observatoriums aufgrund dessen Energieschwelle nur Ereignisse mit Energien oberhalb von $3\times10^{18}\,\mathrm{eV}$ gemessen werden können, können mit der Infill-Erweiterung Energien bis hinunter zu $10^{17.2}\,\mathrm{eV}$ erreicht werden. In dieser Arbeit wurde eine konsistente Analyse der Datensätze beider Detektoren durchgeführt. Dies ermöglichte die Bestimmung eines kombinierten Flusses der höchstenergetischen Strahlung über drei Dekaden in der Energie. Die erhaltenen Resultate sind in Übereinstimmung mit vorherigen Publikationen der Pierre-Auger-Kollaboration. Zusätzlich dazu zeigen sich interessante und weitestgehend unerforschte Strukturen im Fluss bei Energien um $10^{17.5}\,\mathrm{eV}$. Eine neue Suche nach einer Abhängigkeit des gemessenen Spektrums von der Ankunftsrichtung der Teilchen wurde durchgeführt. Es wurde dabei keine Änderung der spektralen Form festgestellt. Allerdings zeigt sich ein Überschuss an Ereignissen aus der südlichen Hemisphäre. Dieser Überschuss ist in Übereinstimmung mit einer gemessenen Dipol-Anisotropie, welche in einer separaten Analyse von Auger-Daten festgestellt wurde. Um eine korrekte Analyse dieser Abhängigkeiten zu ermöglichen, wurde der Einfluss von Wettereffekten und des geomagnetischen Feldes auf die Verteilung von Sekundärteilchen am Boden berücksichtigt. Zusätzlich zur Standardrekonstruktion, wurde der gesamte Teilchenfluss auch mittels Schaueruniversalität bestimmt. Es ergeben sich sehr ähnliche Resultate. Dies legt den Grundstein für zukünftige Analysen des Flusses für unterschiedliche Massengruppen. Dabei werden Daten des aufgerüsteten Detektors von AugerPrime von entscheidender Bedeutung sein.

Acronyms

This is a list of acronyms used within this work sorted alphabetically according to the short version.

AERA Auger Engineering Radio Array . . . 17
AGNs active galactic nuclei . . . 1
AMIGA Auger Muon Detectors for the Infill Ground Array
AoP area-over-peak . . . 129
ASCII Auger Scintillator for Composition - II
Auger Pierre Auger Observatory . . . iii
CIC Constant Intensity Cut . . . 109
c.d.f. cumulative distribution function . . . 41
CMB cosmic microwave background radiation . . . 4
CR cosmic ray . . . 1
EAS extensive air shower . . . 2
FADC flash analog to digital converter . . . 15
FD Fluorescence detector
FoV field of view
GRB gamma-ray burst . . . 1
GPS Global Positioning System . . . 65
HEAT High Elevation Auger Telescopes . . . 15
SD-750 750 m SD vertical
ICRC International Cosmic Ray Conference . . . 146
LDF lateral distribution function . . . 31
LHC Large Hadron Collider . . . 204
MD Muon detector . . . 17
MoPS Multiplicity of positive steps . . . 64
p.d.f. probability density function . . . 71
PE photo-electron . . . 14
PMT photo-multiplier tube . . . 12
SD-1500 1500 m SD vertical . . . 12
SD Surface detector . . . 12
SNR supernova remnant . . . 1
TA Telescope Array . . . 5
ToT time-over-threshold trigger
ToTd time-over-threshold deconvoluted trigger . . . 64
UHECR ultra-high energy cosmic ray . . . iii
VAOD vertical aerosol optical depth . . . 134
VCT vertical centered through-going . . . 12
VEM vertical-equivalent muon
WCD water-Cherenkov detector . . . 2

Contents

CHAPTER 1

Introduction

Cosmic rays are charged subatomic particles. They reach Earth from both galactic and extragalactic sources; the spectrum of particles extends up to energies of 10^{20} eV, which is comparable to the kinetic energy of macroscopic objects like a tennis ball after a serve. At an energy of 10^{16} eV, the rate of particles is one particle per square meter and year. Towards higher energies, the flux decreases according to a steep power law. The highest energetic particles interact with molecules in Earth's atmosphere and create extensive air showers. The study and reconstruction of these showers are the topics of this work. The total energy density of CRs[1] in interstellar space amounts to about $1\,\mathrm{eV/cm^3}$ and is thus comparable to, but a factor of three to four larger, than energy densities of starlight, galactic magnetic fields or the *cosmic microwave background radiation* (CMB). There is evidence that the highest energy UHECRs are produced in the most powerful environments of the universe, among which are shock fronts of SNRs[2], AGNs[3] or even GRBs[4]. The most abundant CRs are proton and helium at lower energies. Also, there are electrons and a minuscule fraction of positrons and other antimatter. In the following discussions, I will focus on hadronic UHECRs with energies above 10^{17} eV.

The goal of studying UHECRs is to gain insight into physical mechanisms, e.g. particle acceleration, of the universe at highest energies on both the macroscopic and microscopic scale. For example, we learn about the structure and evolution of the universe and galactic and intergalactic magnetic fields. On a microscopic scale, the study of UHECRs is the study of particle physics at the highest energies. A measurement of particles with energies of 10^{20} eV corresponds to a center-of-mass energy that is a factor of 100 larger than accessible at the *Large Hadron Collider* (LHC). Fundamental questions about hadronic interactions and cross-sections are answered. More exotic topics involve new particles and even new physical mechanisms that might be uncovered.

[1] cosmic rays
[2] supernova remnants
[3] active galactic nuclei
[4] gamma-ray bursts

1.1 Cosmic-rays

The following sections start with a brief history of CRs, followed by ideas about candidate sources as well as an overview about the current results in the field. The last section will cover the physics of EASs[5]. The focus in these introductory sections is put on UHECRs with primary energies above 10^{17} eV. This is the energy range covered by measurements presented in this work.

1.1.1 A brief history

At the beginning of the 20th century, radioactivity was established, but the popular belief was that the source of ionizing radiation was Earth itself. In 1912, motivated by previous attempts from Pacini and other physicists, Victor Hess used hot air balloons to transport electro-meters up in the air and to measure the degree of ionizing radiation as a function of height [1, 2]. Surprisingly, he found that the degree of ionization increases with height. This led him to conclude that there might be extraterrestrial sources of particles that ionize our atmosphere once they enter. His claim proved true and he received the Nobel Prize in 1936 for the discovery of CRs.

Follow-up experiments were able to establish that the detected primary CRs are almost always positively-charged particles and mainly protons [3]. In 1937, Pierre Auger was the first to measure the lateral distribution of air showers by placing Geiger counters at different positions and by observing coincident triggers [4]. This was motivated by previous experiments of Bothe and Kolhörster [3]. Auger had proven the existence of extensive air showers, for which Heitler later formulated a first theoretical description in 1954 [5, 6] as will be discussed in Section 1.1.4.

At the beginning of the 1960s, MIT physicists at the Volcano Ranch experiment were the first to measure high-energy CRs. They used an array of scintillators to observe the direction and particle content on ground. One of the highest energetic particles ever measured was found in February 1962 [7]. With an estimated energy above 10^{20} eV, this was a milestone for UHECR physics and created the demand for new and larger experiments to solve puzzles at the highest energies [8].

A theoretical milestone was the independent discovery of a flux suppression by both Greisen [8] as well as Zatsepin and Kuz'min [9] in 1966. They predicted an interaction of UHECR protons with photons of the cosmic microwave background. At proton energies above $\approx 5\times10^{19}$ eV, this leads to a significant energy loss of the primary particle. This is called the GZK cut-off.

The first array of WCDs[6] was employed at the Haverah Park experiment in North Yorkshire. Using a 12 km^2 array, UHECRs were measured for about 20 years [10]. A rich number of events and vital information about air showers was collected during that time. First evidence for the flux suppression was found by the High Resolution Fly's eye (HiRes) collaboration using fluorescence telescopes to measure particles above 10^{17} eV in energy [11, 12]. However, the AGASA experiment employing Surface detectors could not confirm these findings [13]. Solving this ambiguity with higher statistics and better detector techniques was one motivation behind the construction of the *Pierre Auger Observatory* (Auger). During the last ten years, the puzzle was solved in favor of the suppression, though it is still unclear if the feature is due to the GZK cut-off or simply due to the vanishing acceleration power of sources at the end of the spectrum. This is further addressed in the following sections.

[5] extensive air showers
[6] water-Cherenkov detector s

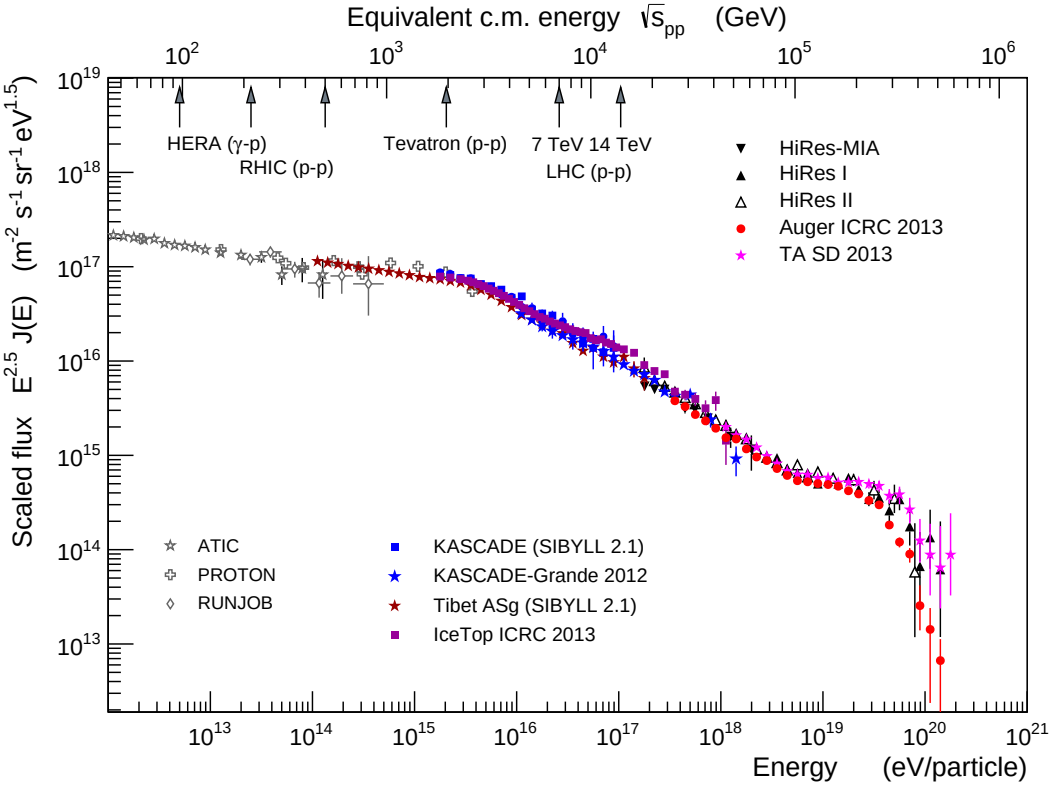


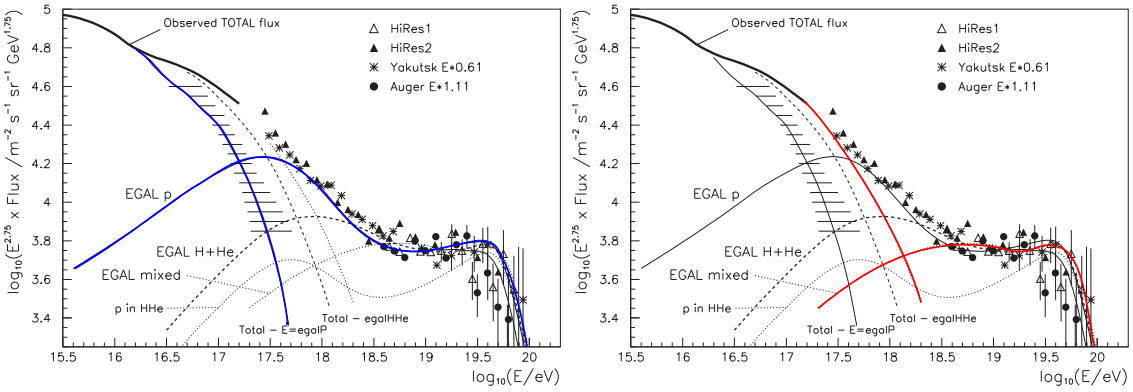
Figure 1.1: Measurements of the flux of CRs over almost ten decades in energy from various experiments (courtesy of Ralph Engel). The data are scaled with $E^{2.5}$ to better visualize spectral features. The most prominent spectral features are further discussed in this section.

1.1.2 The flux and its features

The flux of CRs is well described with a steeply falling power law. For example, the acceleration of particles in diffuse shock fronts leads to such a scaling with energy [14]. The measured flux is depicted in Fig. 1.1. Shown are results from different experiments in the energy range 10^{12} eV to 10^{20} eV. The data are scaled with $E^{2.5}$ to enhance the visibility of the spectral features. Both the axes of scaled flux and energy are logarithmic. Three spectral features are clearly visible: a steepening of the flux at 3×10^{15} eV to 5×10^{15} eV (*the knee*), a flattening of the spectrum around 5×10^{18} eV (*the ankle*) and a strong flux suppression above energies of 4×10^{19} eV. These features are briefly discussed in the following sections. The measurement of the ankle and the flux suppression using current Auger data are topics of this work and are discussed in depth in Chapter 4.

The knees

There are several knee-like structures in the measured spectrum. The first steepening occurs at an energy of 3×10^{15} eV. Measurements with the KASCADE experiment have helped to establish that this observation corresponds to a loss of light elements in the all particle flux [15, 16]. A more recent analysis of KASCADE-Grande data indicates another (less pronounced) steepening at 8×10^{18} eV [17]. This feature can be attributed to an extinguishing heavy component in the flux. Detailed analyses also reveal knees for the elements with intermediate masses. The different fluxes add up to the all particle spectrum in such a way that the individual knees are hardly visible. These observations fit well to the scenario of



(a) Pair production dip **(b)** Mixed composition

Figure 1.2: Visualization of two scenarios to describe the ankle feature in the flux of UHECRs. The scenarios are detailed in the text (from [18, 19]).

diffuse particle acceleration in shock fronts of supernova remnants. In this most popular explanation, the knees are charge-dependent cut-offs of the flux of galactic CRs. They define the drop of acceleration possibilities for galactic CRs. The maximum energy of particles with charge Z is estimated to be $E_{\mathrm{max}} \approx 3\,Z\,10^{15}$ eV. This fits well with the observed features in the flux. Alternative theories to describe the knees with new physics or unexpected changes in hadronic interactions are strongly disfavored by the experimental results.

The ankle

While the knees in the spectrum are well understood as features of the flux of galactic CRs, the extragalactic component is relevant to describe the ankle at 5×10^{18} eV. In particular, the ankle region is understood as the region in which a transition from galactic to extragalactic radiation occurs. Two popular scenarios to describe the ankle are depicted in Fig. 1.2 and described in the following:
In the dip scenario plotted in Fig. 1.2a, the flux of extragalactic protons is already dominant at energies below the ankle, but it is suppressed due to e^+e^- pair-production processes with photons from the CMB[7] [18]. The galactic flux of iron is required to drop quickly in this scenario and the overall fraction of protons in the spectrum needs to be larger than 80 %.
In the mixed-composition scenario shown in Fig. 1.2b, the galactic component is still dominant before the ankle and the transition to extragalactic particles occurs at the ankle [19]. A very new model gives an accurate description of the ankle, as well as the observed suppression in the flux of UHECRs [20]. A critical part of this model are photo-disintegration effects of nuclei in the region surrounding the source. In this scenario, the flux below the ankle is dominated by protons knocked off from higher energy nuclei, while the flux above the ankle is dictated by the injected spectrum and propagation effects. At the heart of the model is the synergy of the interaction and escape times of nuclei with different masses and energies as depicted in the left figure in Fig. 1.3. At lower energies, heavier injected nuclei are disintegrated because the escape time is much larger than the interaction time. However, above the ankle energy, an increasingly larger fraction of heavy nuclei can escape before interacting. The middle figure in Fig. 1.3 represents the flux of injected particles when a pure silicon composition is assumed. From the described assumptions, an ankle-like feature naturally appears as illustrated in the right figure. Different mass groups of elements

[7] cosmic microwave background radiation

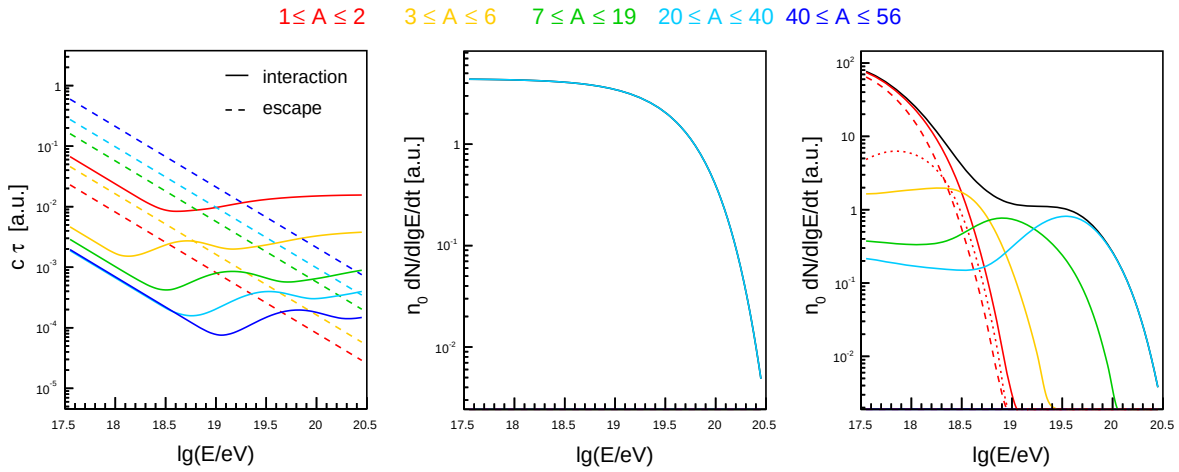


Figure 1.3: (Left) Interaction and escape times for different mass groups as illustrated with different colors. (Middle) The injected flux of Si ($A' = 28$) elements. (Right) The flux of escaping nuclei and nucleons (from [20]).

are depicted as stated above the figures. Using this model together with realistic assumptions on magnetic and photon fields around sources allows the authors to reproduce the spectrum and mass composition of UHECRs as measured with Auger. This comparison is shown in Fig. 1.4. Only a single nuclear species is injected. The best description of data is obtained with a pure silicon spectrum. The maximum energy of the accelerators is found to be $E_{\mathrm{max}} = Z\,10^{18.5}$ eV. Other assumptions and fit parameters are stated in [20].

Flux suppression

On the one hand, the flux suppression can be understood as a propagation effect. The interaction of UHECR protons with energies above $5{\times}10^{19}$ eV and photons from the CMB leads to a Δ-resonance. The two relevant processes are:

$$\gamma_{\mathrm{CMB}} + p \rightarrow \Delta^{+} \rightarrow p + \pi^{0}\,, \tag{1.1}$$

$$\gamma_{\mathrm{CMB}} + p \rightarrow \Delta^{+} \rightarrow n + \pi^{+}\,. \tag{1.2}$$

Each of these interactions leads to an effective energy loss of 20 % for the primary proton. Due to the mean free path of this interaction, extragalactic protons traveling farther than distances on the order of 100 Mpc, and with energies above the threshold, will never be observed on Earth. This defines the GZK horizon. The energy loss as a function of distance is depicted in Fig. 1.5. Photo-dissociation processes, e.g. the giant dipole resonance, lead to similar energy losses for iron nuclei. Thus, proton and iron nuclei actually have similar propagation characteristics in this respect.

On the other hand, the flux suppression might just be experimental evidence of the maximum energy of accelerators. Charge dependent cut-offs of the different extragalactic components are then expected (similar to the case of the galactic components).

Two scenarios to describe current data are depicted in Fig. 1.6. Data measured by TA[8] are fit to a pure proton model, taking into account the GZK suppression. Auger data are described with a mixed model assuming a charge-dependent maximum energy of $E_{\mathrm{max}} = Z \times 10^{18.7}$ eV [22]. A multitude of different theoretical explanations exist. The measurement of the mass composition of UHECR is the vital ingredient to decide between different scenarios. Both, the current Auger results described in Section 1.2.8 and the results

[8] Telescope Array

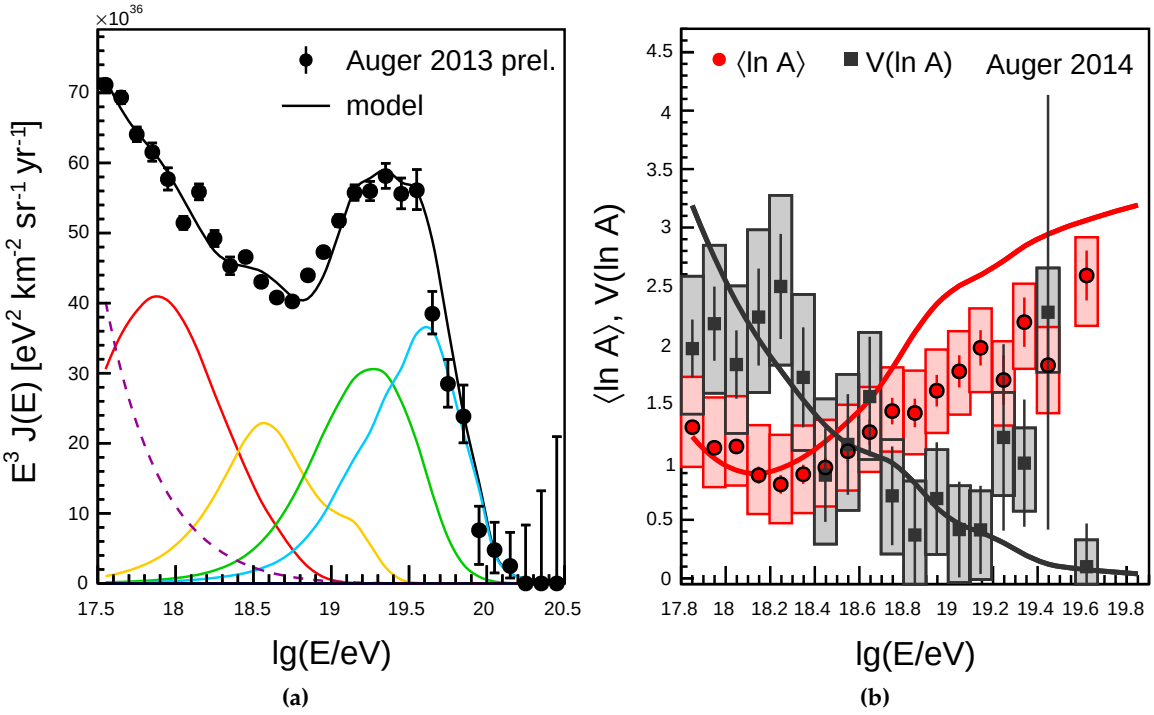


Figure 1.4: (a) Comparison of the predicted flux of UHECRs to a recent combined Auger spectrum. The flux is scaled with E^3. The color scale for the fluxes of the individual element groups is identical to Fig. 1.3. (b) Prediction of the average logarithmic mass and its fluctuations. A comparison to recent Auger results from [21] is given. Statistical uncertainties of the measurement are represented with error bars, while systematic uncertainties are depicted with boxes. (Plots from [20]).

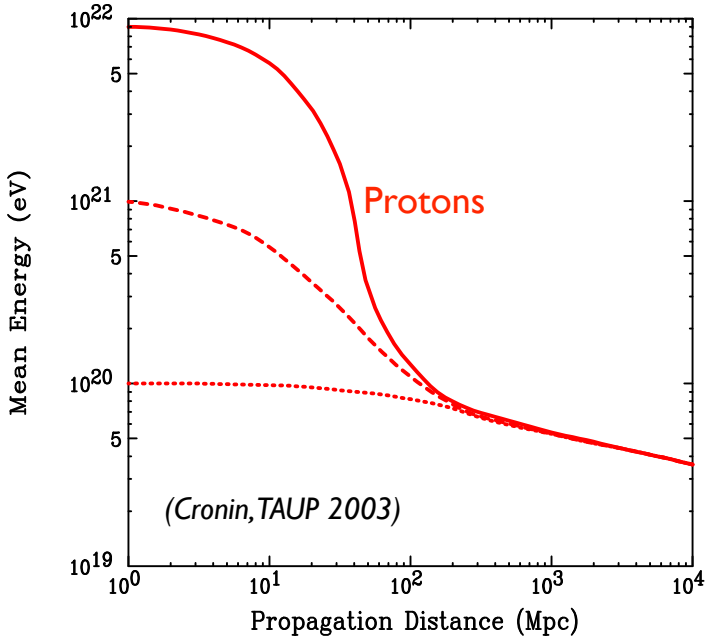


Figure 1.5: Energy losses of protons due to the interaction with photons from the CMB (reference stated in plot).

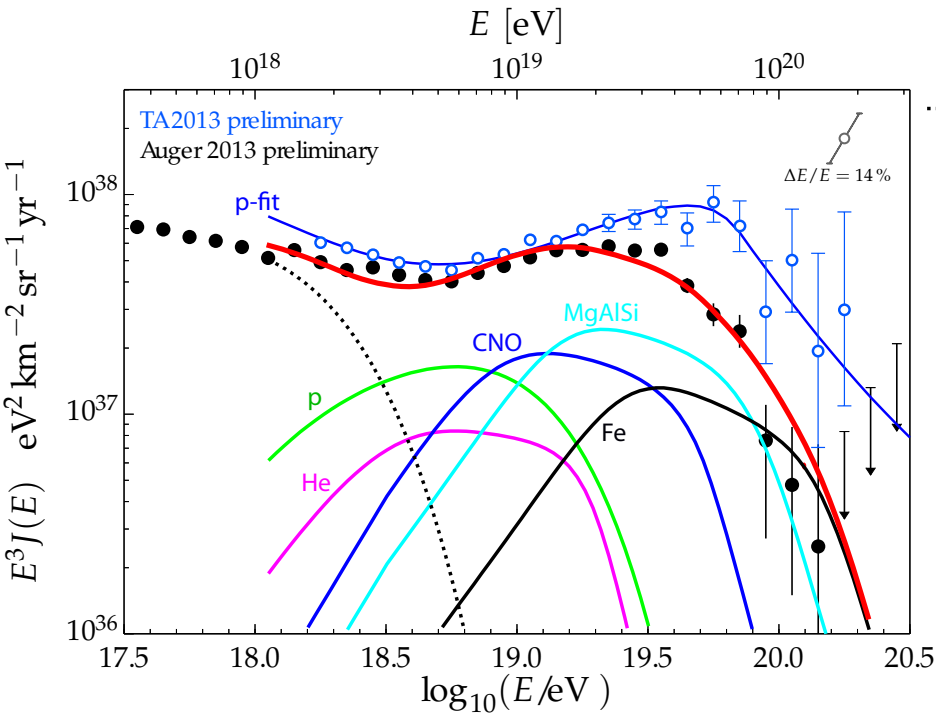


Figure 1.6: Recent measurements of the flux of UHECRs as published by TA and Auger [23, 24]. Data measured by TA are fit to a model of extragalactic proton sources, distributed cosmologically according to $(1+z)^{4.4}$ and injecting a power-law distribution at the sources following $E^{-2.39}$ (blue line). The Auger data are compared to a model assuming a maximum acceleration energy $E_{\max} = Z \times 10^{18.7}$ eV with injection spectra $\gamma = 1$ and an enhanced galactic CR composition from [25]. An additional galactic components is plotted as dotted black line (plot and description from [22]).

derived in this work (see Chapter 5) strongly disfavor a pure proton scenario at the highest energies.

1.1.3 Candidate sources

From the very small flux of UHECRs at the highest energies of one particle per km^2 sr yr, we know that the sources of these particles are very rare. The rather isotropic distribution of particles suggests that the accelerators are mostly of extragalactic nature, as a galactic source would lead to a distinct hotspot of particles from a certain direction in the sky. Also, considering distances of extragalactic objects on the scale of Mpc or more, various propagation effects have to be considered, as discussed in the previous section.

Sources of UHECRs are summarized in the Hillas plot in Fig. 1.7. The maximum particle energy is roughly $E \approx Z\,B\,R$, with the charge Z of the primary particle, the magnetic field strength B and the size of the accelerating system R. This is the basis for the categorization of sources in the Hillas plot. The solid diagonal line indicates the requirement for the acceleration of a 10^{20} eV proton, while the dashed line represents iron primaries. It is apparent that only a few sources are able to accelerate particles up to the highest energies. Among the most promising candidates are active galactic nuclei (massive black holes) and gamma ray bursts. More exotic scenarios with pulsars (neutron stars with strong magnetic fields) exist as well.

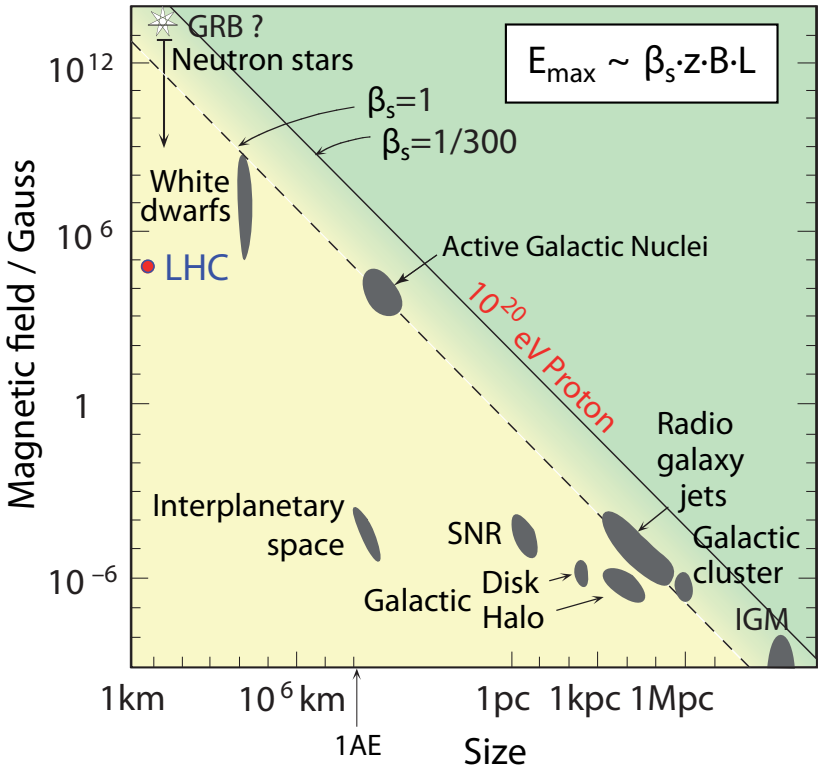


Figure 1.7: Possible candidate sources for UHECRs illustrated in the Hillas plot. A detailed description is given in the text (from [19]).

1.1.4 Extensive air showers

Upon entering Earth's atmosphere, a cosmic ray interacts with nuclei in the air and produces secondary particles. The process repeats and an extensive air shower is formed.

It is reasonable to divide air showers into three particle components: the electromagnetic component that carries 90 % of the energy ($e^{\pm}$, γ), the muonic component, which propagates towards the ground with a low probability of interacting after its formation ($\mu^{\pm}$) and the hadronic component that is dominant in the early part of the air shower development and feeds the other components with energy ($ß^{\pm}, ß^{0}, \mathrm{p}, \mathrm{n}, \mathrm{K}^{\pm}, \mathrm{K}^{0}$). A brief discussion of these components based on the Heitler (or cascade) model follows [5, 26]. A more detailed overview of components used in the air shower universality analyses is given in Section 2.2.

The electromagnetic shower

The development of the electromagnetic component can be well understood by taking two physical processes into account (depicted in Fig. 1.8 on the right hand side). When interacting with a nucleus, a photon creates an $e^{+}e^{-}$ pair. The $e^{\pm}$ themselves again create photons through Bremsstrahlung while being affected by the field of a nucleus (a virtual photon exchanged in the Feynman diagram). It is assumed that an electromagnetic particle interacts after one splitting length $\mathrm{d} = \ln(2)\lambda_r$, where $\lambda_r \approx 37\,\mathrm{g\,cm^{-2}}$ is the electromagnetic radiation length in air. In every step of this simple electromagnetic Heitler model, the particle number is doubled, and the energy is shared equally between the particles. So, after n steps, the shower consists of 2^n particles. This process continues until the individual particle energies drop below a critical energy E_c^e. At this point, energy losses due to inelastic

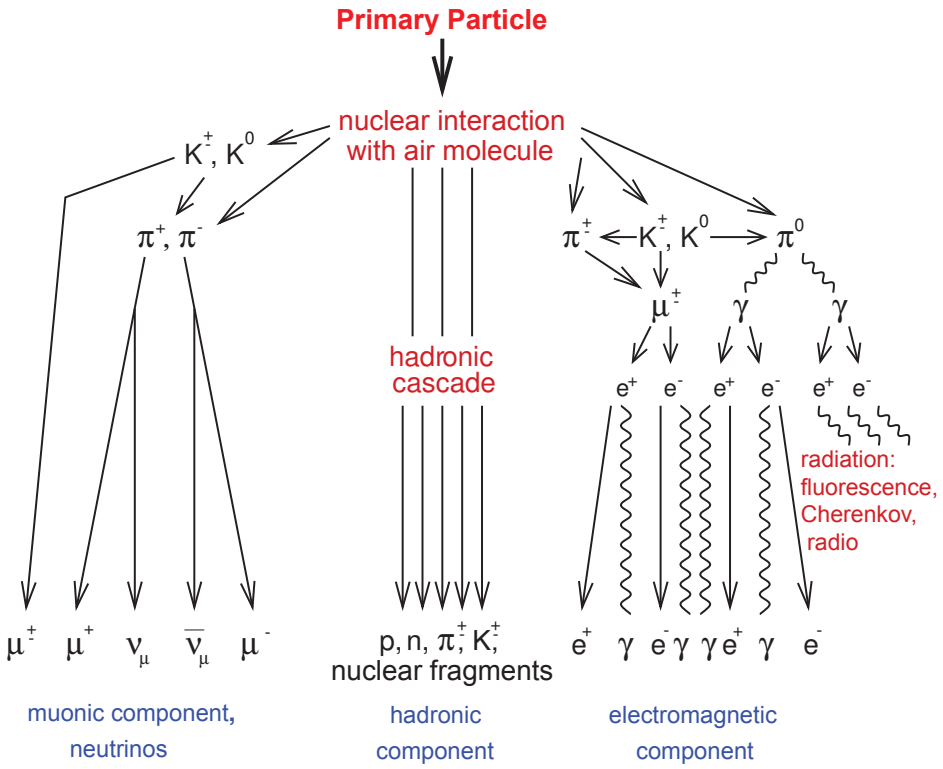


Figure 1.8: Illustration of the particle components during an air shower development.

collisions start to dominate and no further particles are created. In air, the critical energy amounts to about 85 MeV.

This simple model is not completely accurate; for example, an $e^{\pm}$ will in most cases not deposit exactly half of its energy into a single Bremsstrahlung photon. However, two important facts of electromagnetic showers are well accounted for: the total number of electromagnetic particles is proportional to the initial energy $N_{\max} \propto E_0$ and the atmospheric depth of maximum shower development $X_{\max}$ is proportional to $\log(E_0)$.

The hadronic shower

The microscopic interactions of the hadronic component are less well known, since soft processes with small momentum transfer cannot be calculated in the fundamental theory of Quantum Chromodynamics. Nevertheless, it is possible to extend the electromagnetic Heitler model to qualitatively describe the hadronic development. More accurate results are obtained with air shower simulations that are based on the extrapolation of known cross-sections to higher energies.

Hadronic interaction dominate the early part of the shower development. When an incoming proton interacts with an air molecule, $a \cdot n_{\text{mult}}$ charged pions $\pi^{\pm}$ and $(1-a) \cdot n_{\text{mult}}$ neutral pions π^0 are created. The pion charge ratio $1-a$ is typically assumed to be $1/3$. Neutral pions almost instantly decay into two photons. These photons initiate electromagnetic sub showers. The charged pions continue to interact and produce more pions until the individual pion energies cross the critical energy E_c^{π}. This critical energy is the energy at which the decay length of the charged pions becomes smaller than the distance to the next interaction point. E_c^{π} scales with the atmospheric density. Typical values

for air showers are 20 GeV to 30 GeV. Reaching E_c^π, the charged pions will decay into muons and neutrinos:

$$\pi^+ \rightarrow \mu^+ + \nu_\mu \tag{1.3}$$
$$\pi^- \rightarrow \mu^- + \bar{\nu}_\mu. \tag{1.4}$$

The number of generations n_c to reach E_c^π is

$$n_c = \frac{\ln(E_0/E_c^\pi)}{\ln(n_{\text{mult}})}\,. \tag{1.5}$$

n_c is typically between four and seven for air showers [27].
The primary energy can be calculated according to:

$$E_0 = E_c^e N_{\max} + E_c^\pi N_\mu. \tag{1.6}$$

These considerations hold for proton induced air showers. A superposition model is used to describe showers of heavier nuclei. The shower of a nucleus with atomic number A and total energy E is described as A showers induced by protons with energies E/A. It follows that the depth of maximum development of an iron shower with energy E will be higher in the atmosphere than the one of a proton shower with identical energy E. The neutrinos created in these reactions are part of the invisible energy. This component will stay unobserved and is estimated from simulations to be about 5 %.

The depth of shower maximum

Most of the energy of an air shower is deposited by electrons and positrons. Their development is best described as a function of traversed air mass or grammage:

$$X = \int_z^\infty \rho(\vec{r}(z'))\,\mathrm{d}z'\,. \tag{1.7}$$

It is the integration of the air density along the shower axis. The point along the shower axis at which the energy deposit is maximal is called the depth of shower maximum $X_{\max}$. It depends logarithmically on both primary energy and mass and is used as an estimator of the primary mass within this work. An extension of the considerations from the Heitler model leads to a generalized superposition model [28]:

$$\langle X_{\max}\rangle = X_0 + D\lg\left(\frac{E}{E_0 A}\right) + \xi \ln A + \delta \ln A \lg\left(\frac{E}{E_0}\right), \tag{1.8}$$

where X_0 is the mean depth of proton showers at an energy of E_0 and D is the elongation rate, i.e. the change of $\langle X_{\max}\rangle$ per decade in energy. The parameters ξ and δ describe deviations from the ideal superposition model. The dispersion of $X_{\max}$ is expected to be only influenced by shower-to-shower fluctuations:

$$\sigma^2(X_{\max}) = \sigma^2_{\text{sh}}(\ln A)\,. \tag{1.9}$$

The longitudinal profiles and depth of shower maxima for different particle components are depicted in Fig. 1.9. The plot corresponds to a CORSIKA simulation of an air shower induced by a proton with energy of 10^{20} eV.

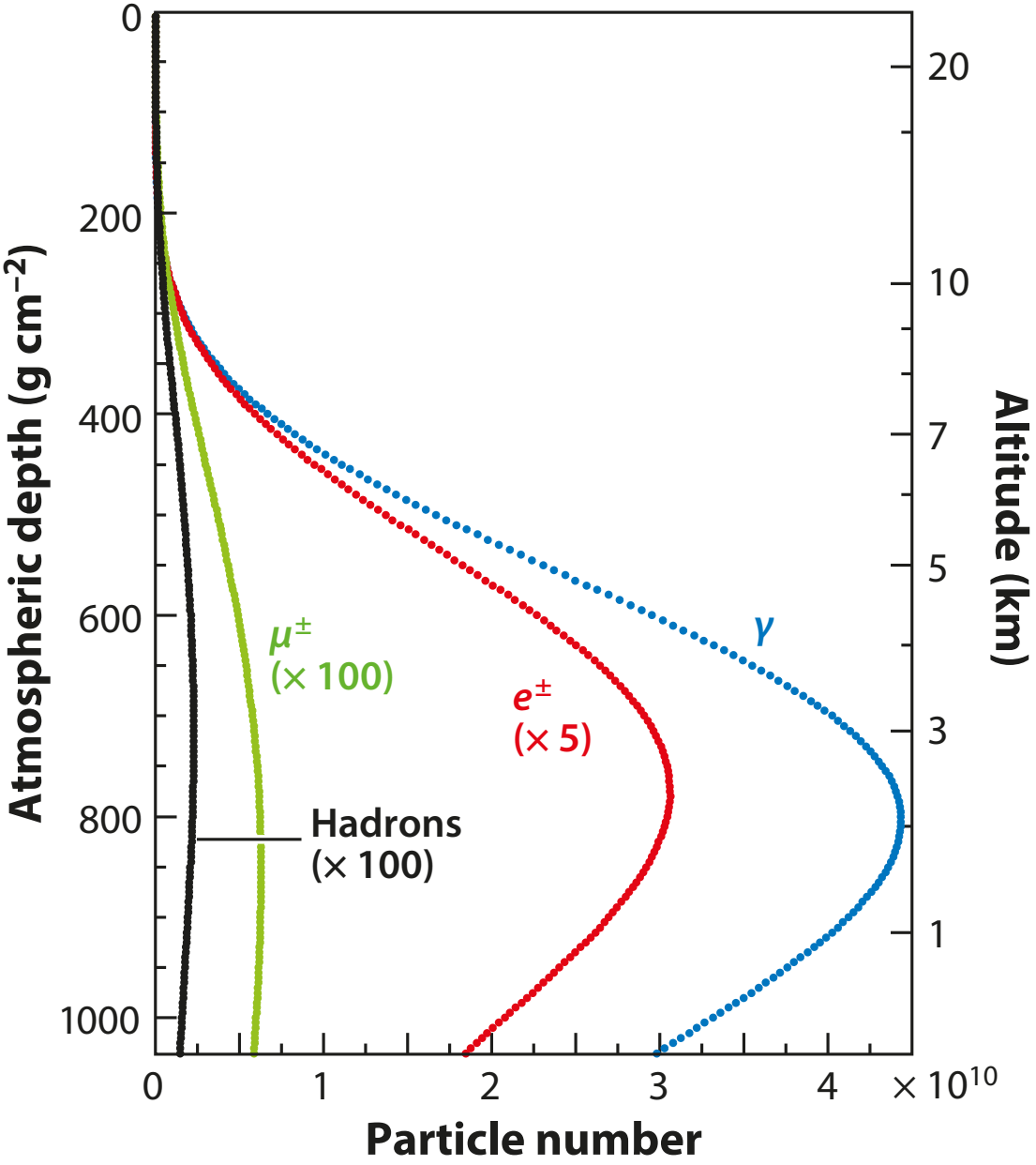


Figure 1.9: Longitudinal profiles of different particle components; from the simulation of a proton induced air shower with an energy of 10^{19} eV (from [29]).

1.2 The Pierre Auger Observatory

The *Pierre Auger Observatory* (Auger) [30] is the world-leading hybrid detector to measure UHECRs. The experiment is located in the Argentinian Pampa, close to the old gold-digger town Malargüe. As a precision instrument, Auger was built to study properties of CRs at the very highest energies (above $E = 10^{19}$ eV). To access these regions of very low particle fluxes, the baseline design of the observatory consists of a *Surface detector* (SD) covering an immense area of more than 3000 km^2. In addition, a *Fluorescence detector* (FD) consisting of 27 telescopes observes light in the atmosphere above the SD[9]. The presence of these two complementary detectors allows us to reconstruct properties of the primary particle with minimal use of simulations. The location of the observatory is optimal for the experiment. It is situated on a plateau close to the Andes at an average height of $\approx$ 1400 m, corresponding to $\approx$ 880 g cm^{-2} in vertical atmospheric overburden, with a maximum deviation from west to east of about 300 m in height. It rains rarely, the atmosphere is clear and the amount of light pollution is comparably small. Several important discoveries and publications were made since the start in 2004, see Section 1.2.8 for selected results relevant to this work. Various extensions to the observatory have been built or are under construction. In the following sections, I will summarize the baseline design, finished and ongoing extensions, as well as plans to upgrade the Observatory.

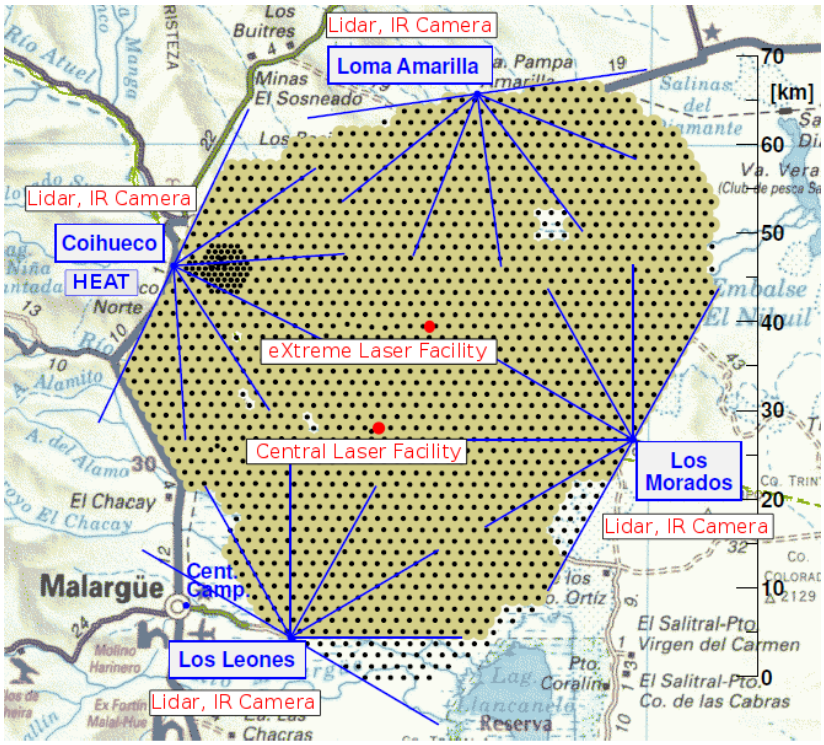
1.2.1 Surface detector

The standard SD is abbreviated with SD-1500[15]. It is composed of a tessellated hexagonal grid of more than 1600 WCDs separated from each other by 1500 m [32]. Deployment of the SD-1500 started in 2004 and was finished in 2008. The threshold of full trigger efficiency of the SD-1500 is reached for primary energies above 3×10^{18} eV [33]. Details about the event selection and reconstruction are given in Section 3.3.

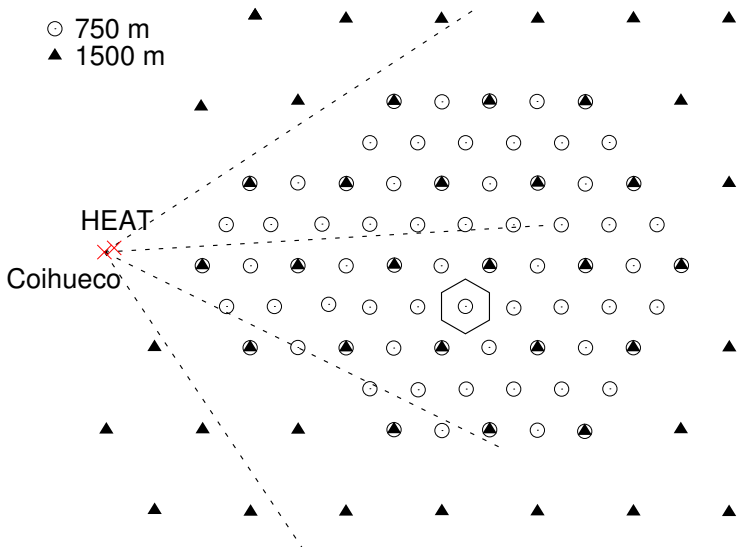
Each WCD has an area of 10 m^2, a height of 1.2 m and is filled with 12 tons of ultra-pure water. The water is contained in a sealed liner with a reflective inner surface. Light created within the water is scattered, reflected and finally amplified and observed by three 9 inch diameter PMTs[16] that are mounted symmetrically at 1.2 m distance from the center of the tank. They look vertically downward into the water through clear windows of polyethylene. Each SD station is powered autonomously with a solar power system for the PMTs and electronics. A schematic overview of an SD station is shown in Fig. 1.12a.

Due to its relevance for this work, I will continue with summarizing critical hardware aspects of SD stations. Two outputs are available for each PMT. Firstly, an AC coupled anode signal also referred to as low gain channel after digitization. Secondly, the amplified signal of the last dynode, also called high gain channel, is provided as output. There is a factor of 32 in charge gain between the amplified dynode output and the anode. This enables a higher dynamic range for the signal measurement and is motivated by the exponential increase of particle densities close to the shower core. A wide variety of particle densities ranging from 1000/µs close to the core to 1/µs far away needs to be covered. The analog signals are filtered and digitized by two 10 bit 40 MHz semi-flash ADCs, resulting in 25 ns time traces for both low and high gain channels.
By definition, signals are calibrated in units of *vertical-equivalent muon* (VEM). 1 VEM corresponds to the signal induced by the Cherenkov light from a muon that passes through a tank in a vertical trajectory (a VCT[18] muon). A conversion of both integrated charge

[9] Surface detector
[15] 1500 m SD vertical
[16] photo-multiplier tubes
[18] vertical centered through-going



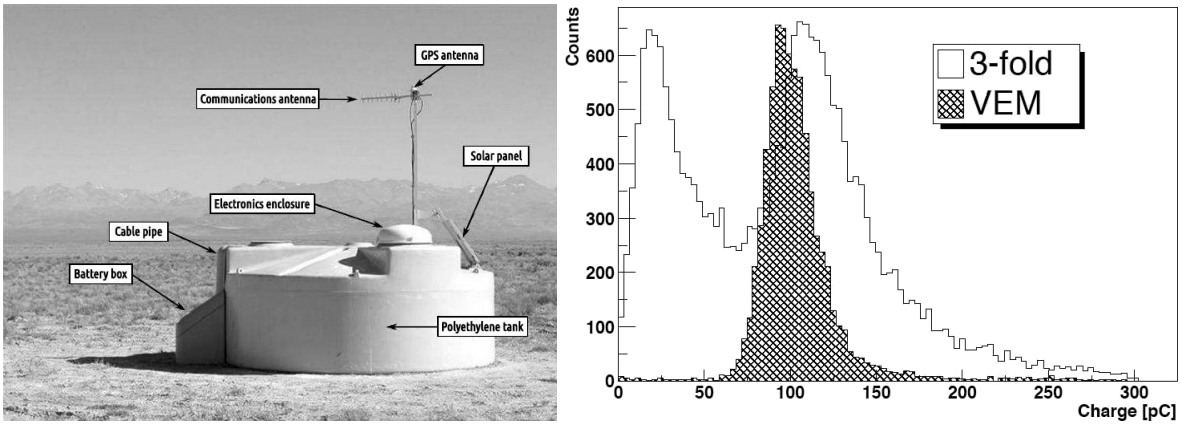
(a)



(b)

Figure 1.10: (a) Overview of Auger, including an SD station and an FD[10] building. Also shown are essential facilities for atmospheric monitoring as described in Section 1.2.3 (adapted from [30]). (b) Schematic overview of the SD-750[11] array as described in Fig. 1.10b. Indicated are the FoV[12]s of the FD telescopes Coihueco and HEAT (from [31]).

Figure 1.11: The SD-750 station Kathy Turner (Id 1764) with an ASCII[13] 2 m^2 prototype mounted on top (see Section 1.2.7). Also visible are additional solar panels and batteries to cope with the increased power consumption. Three AMIGA[14] access tubes are visible in the forefront (personal photograph, November 2014).



(a) (b)

Figure 1.12: (a) Schematic of an SD station (from [30]). (b) Average charge spectrum from SD stations triggering either on coincidences of the 3 PMTs or selected vertical muons. Used to infer the VEM[17] calibration as detailed in the text (from [30]).

and differential FADC traces into VEM is provided. All recorded signals are converted into units of VEM prior to air shower analyses, e.g. the study of lateral distributions. The dependence of VEM on the actual muon energy is on the percent level as long as the muon passes through the tank. On average, a VCT muon will induce 79 PEs[19] in one PMT. Calibration histograms are plotted in Fig. 1.12b. The hatched histogram includes only the charge spectrum in a PMT from vertical muons. To ensure this, triggers were provided from vertically aligned plastic scintillators on the top and on the bottom of a

[19] photo-electrons

test tank. The open histogram represents the spectrum in an SD when the three PMTs are required to trigger in coincidence. The muon peak in the open histogram is at a slightly larger value because inclined muons traverse more water and produce more Cherenkov light. Determining the peak position in these high-gain charge histograms leads to an expression of integrated signal in units of VEM-charge. The peak charge of 100 pC results from the equation $C_{\text{peak}} = \frac{n_{\text{PE}}}{\text{VCT}}\, e\, G\, A \approx 100\,\text{pC}$, with the average number of PEs per VCT as mentioned above, the average PMT gain $G \approx 2\times10^5$, the dynode amplification factor $A \approx 40$ and the elementary charge e. The conversion of time-binned signals is obtained from pulse-height histograms and leads to a unit of VEM-peak. These subtle differences are accounted for in the calibration procedure [34]. Online calibration data for each tank and PMT are provided once every minute.

There are two effects that hinder the measurement of very large signals [35]. Firstly, the FADC[20] has a dynamic range of 10 bit. Approximately 50 channels are reserved for the baseline offset, and thus, there are about 950 channels left for the signal range. Typically 1 VEM corresponds to about 50 channels (in the pulse-height histogram) in the high gain and 1.6 channels in the low gain. Overflow of the dynode readout will therefore occur at ≈ 20 VEM and at ≈ 600 VEM in case of the anode.
Secondly, the response of the PMTs starts to be non-linear above currents of 50 mA corresponding roughly to the onset of the overflow of the anode readout as mentioned before. A mixture of these effects leads to saturated signals.

During the years of SD measurement, methods were developed to successfully recover saturated signals and estimate the true signal [35–38].

750 m SD infill array

To substantially decrease the threshold of full efficiency to less then 3×10^{17} eV for showers arriving with zenith angle below 55°, a nested array of 71 (49 additional) WCDs, distanced 750 m from each other, was built [31]. This SD-750 array covers an area of 27 km^2 and is located close to the FD sites Coihueco and HEAT, which both overlook the SD-750 with three telescopes (Section 1.2.2). Data taking with the first hexagon started in September 2007 and the array was fully deployed by September 2012. A schematic overview is shown in Fig. 1.10b. In July 2013, new station triggers were installed in all SD stations. This update enhances the sensitivity of individual stations to small signals in the 1 VEM regime and results in a threshold of full efficiency for the SD-750 of $\approx 10^{17.2}$ eV, see e.g. [39].

1.2.2 Fluorescence detector

The standard FD consists of 24 telescopes at four sites surrounding the SD: Los Leones, Los Morados, Loma Amarilla and Coihueco [40]. At each site, there is an FD building with six telescopes in clean climate-controlled rooms. Shown in Fig. 1.13a is the FD building at Los Leones during daytime. The shutters are usually closed during day. Each telescope has a FoV of 30° × 30° in azimuth and elevation. Together, the six telescopes in a building cover 180° in azimuth. Three additional telescopes with an elevated FoV were built 180 m in front of the FD site at Coihueco. This extension is called HEAT[21] [41]. Its telescopes are similar to the standard ones except that they can be tilted upwards by a maximum of 29° in elevation. This allows the observation of shallower showers induced by primaries with lower energies. These showers are outside the FoV of the standard FD. While the standard

[20] flash analog to digital converter
[21] High Elevation Auger Telescopes

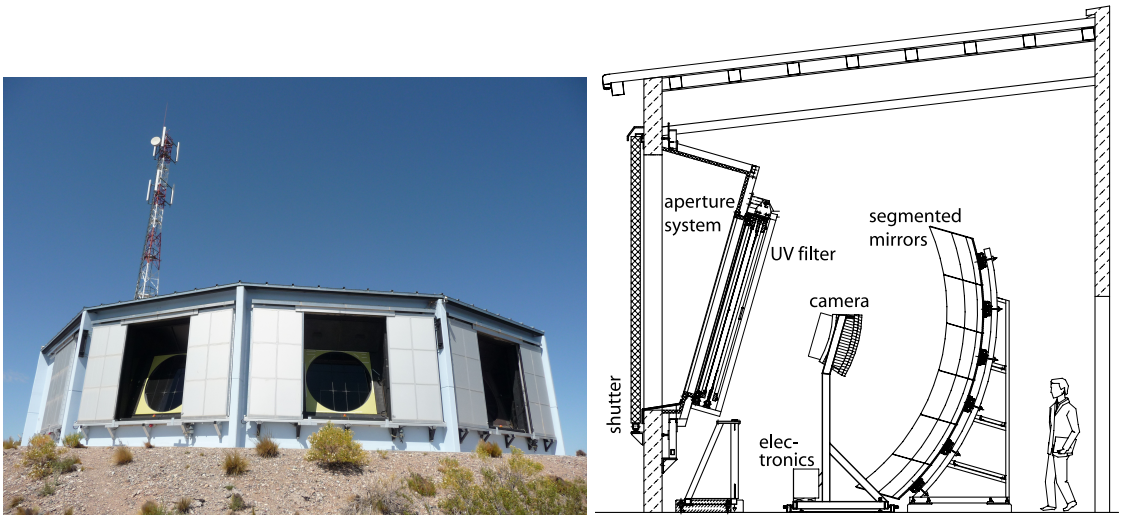


Figure 1.13: (a) FD building Los Leones with oppen shutters during the day (from [30]). (b) Schematic picture of a FD telescope (from [30]).

FD acts as counterpart to the standard SD, HEAT forms a hybrid detector together with the SD-750, which is nicely overlooked by HEAT and Coihueco (see Fig. 1.10b).

A schematic view of an FD telescope is depicted in Fig. 1.13b. Secondary electromagnetic particles in air showers excite nitrogen molecules in the air. As a de-excitation, they isotropically emit fluorescence light with a peak in spectral density at wavelengths of 337 nm and 357 nm. If an FD telescope is close enough to observe the air shower and atmospheric conditions are sound, the potentially scattered fluorescence photons will enter through a circular diaphragm with a diameter of 1.1 m. A glass window in the aperture acts as a UV filter; this increases the signal-to-noise ratio for the measurements. The light is focused on the actual camera by a 13 m^2 segmented mirror. The reflectivity of the mirror is above 90 % in the UV range of interest. The camera consists of a grid of PMTs arranged into 22 rows and 20 columns. Each PMT is of hexagonal shape with a side-to-side distance of 45.6 mm. Winston cones collect the light to the active cathode of the PMTs. Regular cleaning of the UV filter from the outside has been performed throughout the years. Less frequently, filter, mirror and corrector ring have been cleaned from the inside. Ongoing studies show that dust on the various parts of the system has a major impact on the level of 10 % on the final energy scale of Auger [42]. The FD operation is limited to clear, moonless nights, resulting in a duty cycle of around 15 %.

1.2.3 Atmospheric monitoring

Facilities for atmospheric monitoring are very important for the calorimetric measurements using the FD and moderately important for measurements with the SD. Most facilities are plotted in the overview in Fig. 1.10a. Atmospheric parameters like temperature, humidity and pressure of air influence the longitudinal development of air showers and the amount of emitted fluorescence light. Secondary particle densities on ground are affected as well.

On each FD site, there are, among other small facilities, a lidar station, a ground-based weather station and an infrared camera for cloud monitoring [43]. Pulsed UV lasers are operated at the lidar stations. Shots are fired at various rates and into various directions. The amount of aerosols and clouds is estimated from the signature of back-scattered light.

The light is collected with separate mirrors and PMTs at each station, and the laser shots are fired outside the FoV of active telescopes. Thus, this monitoring does usually not interfere with FD measurements.

Two laser facilities CLF and XLF are operated in the center of the SD array. They are used to measure the aerosol contamination in the line of sight of each of the FD telescopes four times per hour.

In addition to measurements with local weather stations, the Global Data Assimilation System (GDAS) provides valuable continuous information about atmospheric parameters [44].

1.2.4 AMIGA muon detector

The *Auger Muon Detectors for the Infill Ground Array* (AMIGA) enhancement was designed for the direct measurement of the muon content of air showers. It consists of the SD-750 (Fig. 1.10b) and scintillators, which are buried a few meters next to the SD stations and measure muons independently [45]. The collection of scintillators is referred to as MD[22]. Entry tubes to these scintillators are shown in Fig. 1.11. At the end of 2014, the deployment of the first unitary cell of scintillators was finished and the detectors are taking data. Each MD station is made of a $30\,\mathrm{m}^2$ scintillator subdivided into four modules, two of which have an area of $10\,\mathrm{m}^2$ and $5\,\mathrm{m}^2$, respectively. Twin MD stations were deployed at two locations in the unitary hexagon to study systematics and accuracies. The scintillators are buried at a depth of 2.3 m corresponding to an overburden of $\approx 540\,\mathrm{g\,cm}^{-2}\,\sec\theta$. This effectively shields the detectors from electromagnetic particles and imposes a cut-off in the energy spectrum of vertical muons at 1 GeV. The scintillators are triggered by SD-750 stations and provide digital counting of muons irrespective of their energies. In light of the Auger upgrade (Section 1.2.7) the plan is to finish the deployment of AMIGA scintillators in the whole SD-750. This will enable a powerful and versatile measurement of air showers and their muon content up to the ankle in the primary energy spectrum and slightly above.

1.2.5 Radio detector

A very interesting subject is the study of radio emission from air showers. It was proven some years ago that it is experimentally feasible and promising to study the emission of air showers in the MHz-regime [46]. Recently, it was shown by the LOFAR collaboration that the measurement of the radio footprint allows the reconstruction of the energy of the primary particle as well as quantities that are related to its mass, like the depth of shower maximum or the shower curvature [47].

The AERA[23] is built within the SD-750 array to enable a study of the radio emission of air showers together with measurements from SD, FD and MD. Different antenna types and distances are studied at the moment. Current results are given in [48–51]. Some first multi-hybrid events exist and are currently studied.

With a duty cycle of nearly 100 %, low costs and definite possibilities to reconstruct primary properties, the radio emission is a very promising candidate for the future measurement of UHECRs at the highest energies. From the point of view of simulations, radio is very well understood because one does not have to deal with muons and the remaining uncertainties in hadronic interactions.

[22] Muon detector
[23] Auger Engineering Radio Array

1.2.6 Further activities

Various other research topics are studied at Auger. To mention a few of them, there are studies of geophysics topics like lightnings and related phenomena. Radio emission in the GHz-regime is studied with a new set of antenna, the EASIER project. The various instruments for atmospheric monitoring allow detailed studies of atmosphere physics.

1.2.7 Upgrade of the Observatory

Efforts to upgrade the observatory are now in an advanced planning stage [30]. With the general goal to enhance the sensitivity of the measurement to the primary mass composition, several upgrade proposals were studied in detail. This includes both the study of actual prototypes in the field as well as the study of detector simulations to quantify the performance of each design with respect to science goals. For example, the resolution in the measured number of muons at specific distances to the core. The collaboration has opted for *Auger Scintillator for Composition - II* (ASCII) as its choice for the upgrade in disfavor of segmented WCDs. The latter would have required opening up each WCD in the course of the upgrade, but had a similar physics performance regarding the defined science goals compared to scintillators.

Better knowledge of the muon content in air showers is essential to improve the sensitivity to the primary mass. The approach with ASCII is to measure mostly the vast electromagnetic component of air showers with scintillators mounted on top of SD stations (see Fig. 1.11). The WCDs are already very sensitive to muons, and thus the scintillator measurements will help to unravel the pure electromagnetic and muonic signals. This can either be done on the single station or the event level. Due to a large amount of electromagnetic particles in air showers, the area of the scintillators can be rather small. The new detectors will either be $2\,\mathrm{m}^2$ or $4\,\mathrm{m}^2$ in area. To collect as many events as possible, scintillators will be installed in the whole SD array. The final technical design is ongoing at the moment, and the deployment of detectors will happen during the next few years.

The plan also includes an upgrade of station electronics. This does not only cover the support of additional PMTs for the new detectors, but also a faster sampling rate of 120 MHz. The installation of an additional small PMT in the WCDs is discussed as an option to measure signals closer to the core without saturation.

1.2.8 Selected results

This section gives an overview about current, selected results obtained from data measured at Auger. The focus will be on the most relevant results in connection to this work. A rather complete list of physics results and technical reports of the last three years is given by [21, 28, 43, 44, 49–85].

The measurement of the flux of UHECRs using ten years of data is shown in Fig. 1.14. The plot is extracted from [86]. The flux measurement combines measurements with the SD-750 and SD-1500, as well as inclined events with zenith angles above 60° measured with the SD-1500. Also included in the combined spectrum is the flux measured with hybrid events. The spectrum shows a clear flattening at an energy of $10^{18.7}$ eV; a feature denoted as the ankle. The flux suppression at highest energies is established with a significance of more than $30\,\sigma$. A detailed comparison with the flux measurement derived in this work is given in Chapter 4.

The first two central moments of the measured distribution of the depth of shower maximum is shown in Fig. 1.15 [21]. There is a clear trend towards heavier elements at high energies, while there are no events at the highest energies due to the limited duty cycle of

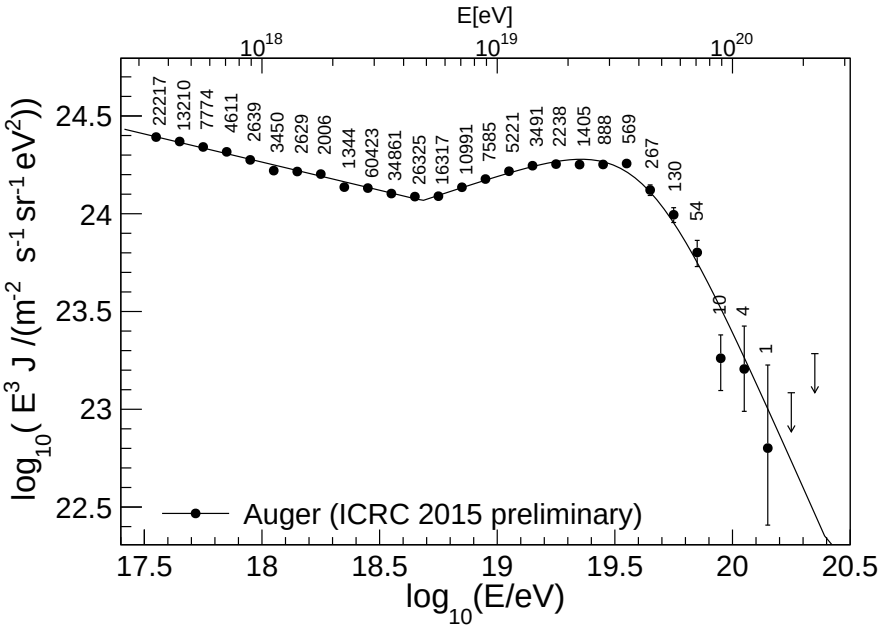


Figure 1.14: Results on the measurement of the flux of UHECRs from Auger data. Shown is a combined spectrum composed of three measurements with SDs and a hybrid measurement. The overall systematic of the energy scale is 14 % (from [86]).

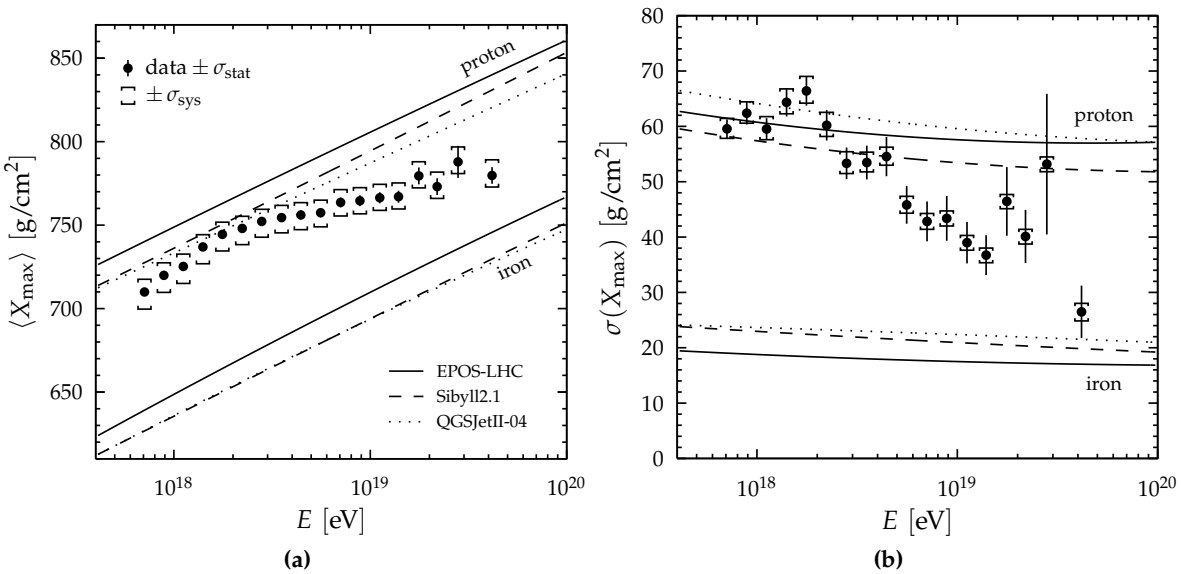


Figure 1.15: (a) The average X_{max}, $\langle X_{max} \rangle$, as a function of the primary energy derived from hybrid measurements. (b) The second central moment of X_{max}, $\sigma(X_{max})$ as a function of the primary energy. Model lines from current air shower simulations for different primary species are included for comparison (from [21]).

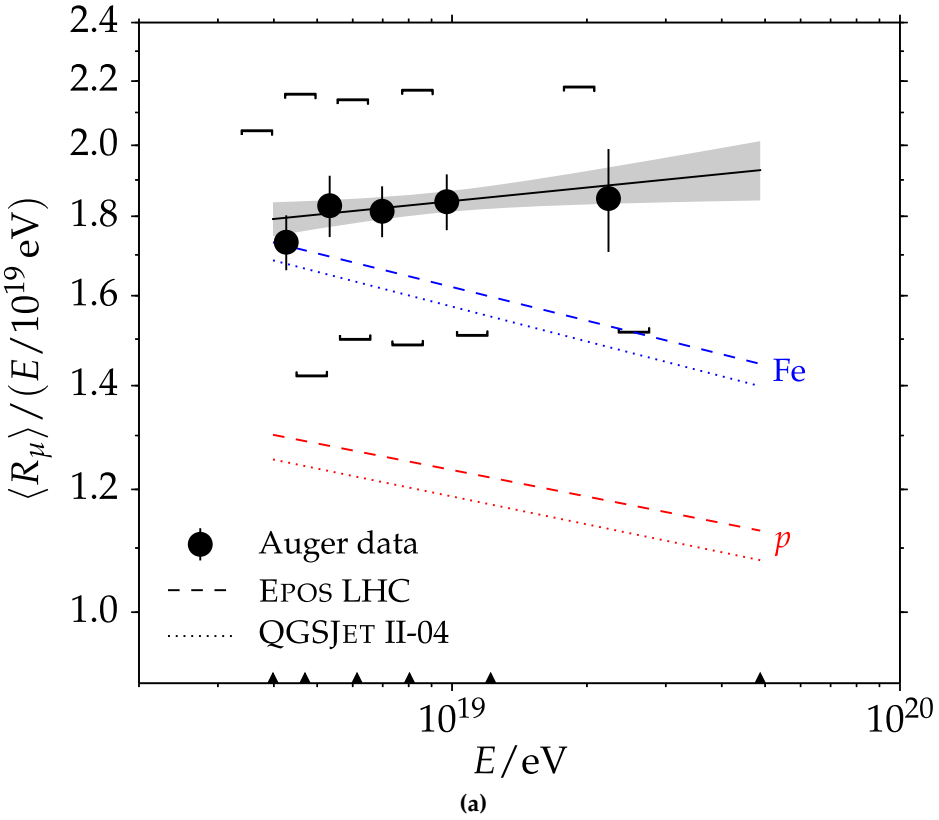
the FD. Using SD data, the range of the highest energies is explored in detail in this work. The results and discussion are given in Chapter 5.

Different ways to determine the amount of muons in air showers with respect to a simulation reference were used on Auger data. One method is the shower universality reconstruction discussed and developed in this work. The results of two other methods are depicted in Fig. 1.16 and briefly discussed here. The result obtained from inclined

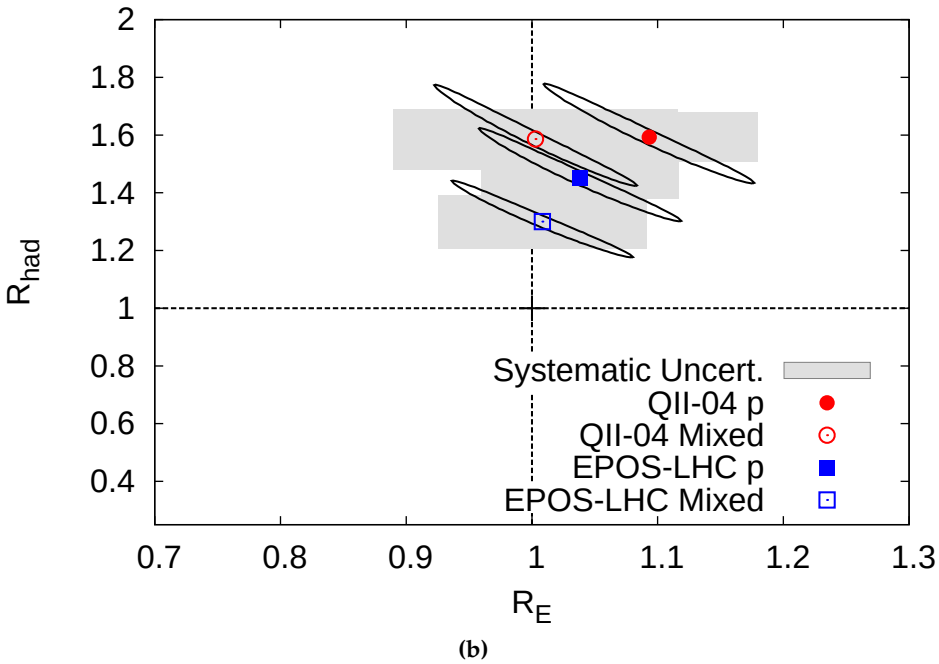
air showers is shown in Fig. 1.16a [83]. Presented is a direct measurement of the muon content with respect to predictions of two contemporary hadronic interaction models. The absolute reference of 1.0 corresponds to proton at 10^{19} eV and simulated with QGSJET-II.03. It is apparent that, while the new interaction models predict a larger amount of muons compared to the old ones, data indicates an even larger relative number close to 2.0. This muon excess is currently not understood. None of the contemporary hadronic interaction model predicts an amount of muons as seen in data. The abundance of muons scales with:

$$N_\mu \propto \left(\frac{E_0}{\zeta_c^\pi}\right)^\beta . \tag{1.10}$$

The slope β is $\approx 0.9 - 0.95$, depending on the hadronic interaction model and details of the simulations [26]. Thus, a mismatch in the muon number between data and simulations could also be attributed to a mismatch in the overall energy scales or a mixture of both effects. This question was addressed with an analysis of vertical showers measured at Auger [87]. To disentangle the energy and muon scales, data and simulations are compared at different zenith angles, thereby exploiting the individual attenuation characteristics of particle components in air showers, i.e. a much weaker attenuation of the muonic component. A scaling factor R_E modifies the ground particle distributions with respect to the overall energy, while the scaling R_{had} only affects the hadronic (muonic) component. The results are depicted in Fig. 1.16b. The statistical uncertainty ellipses reflect the strong anti-correlation between the energy and the number of muons. The gray boxes represent the systematic uncertainties of the analysis. Independent of the hadronic interaction model, the results indicate that, in fact, the number of muons in simulations is underestimated by an amount of 30 % to 60 %, while the overall energy scale is correct. Results on the number of muons derived in this work are presented in Chapter 5. They are compatible with the results discussed in this section.



(a)



(b)

Figure 1.16: (a) Measurement of the relative muon content in very inclined air showers with zenith angles above 62° (from [83]). (b) Measurement of the overall energy rescaling R_E and rescaling of the hadronic energy R_{had} to match air shower measurements with up-to-date simulations (from [87]). The muon numbers obtained with all current hadronic interaction models need to be scaled up to match data.

CHAPTER 2

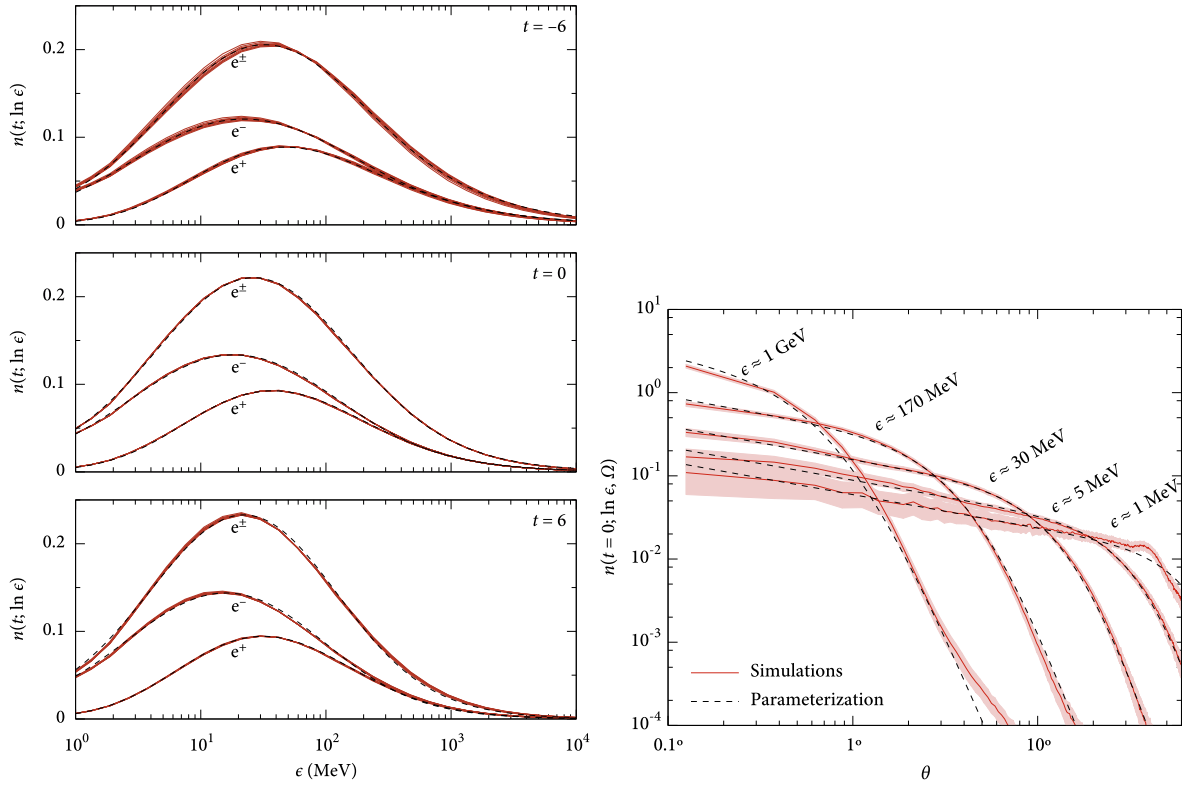
Air shower universality

As an *extensive air shower* (EAS) develops, the cascade of secondary particles can be described as consisting of four components, each of which exhibits a unique yet universal development characteristic. With high energy cosmic ray showers, the shower-to-shower fluctuations of each of these components is minimal when compared to their general development. Thus, it becomes possible for showers to be reconstructed with respect to macroscopic parameters - like energy, the shower maximum or the overall muon content. Denoted air shower universality, this phenomenological method allows for studies of the mass composition or anisotropy with great event statistics when compared to pure FD measurements.

The concept of universality will be briefly explained in the next section. Besides its profound implications for particle physics, shower universality allows one to construct models of the expected average size and time structure of signals or particle energy densities at different development stages of air showers and for various particle detectors. The signal model developed over the last ten years is introduced in Section 2.3. As part of this work, a new model to describe the time structure of signals in WCDs was studied. I will give a detailed description of that in Section 2.4. The prediction of different time quantiles will be shown, and dependencies on the primary mass and hadronic interaction model are investigated. Together, the signal and time model are used to reconstruct air showers and proeprties of the primary particle. The statistical reconstruction algorithm and tests with Monte Carlo events are explained in Section 3.4. Finally, results on the estimation of the shower maximum and the relative muon number of events measured at Auger are discussed in Chapter 5. Energy spectra reconstructed with the universality approach are presented in Section 4.11 and compared to the ones obtained with an updated standard reconstruction developed in this work.

2.1 The concept of air shower universality

The average properties of an EAS depend mostly on the primary energy and the stage of shower development (as explained in the text below). On first order, there is no dependence on the primary mass or the shower geometry. This phenomenon is called shower universality. We point to [88] and references therein. The universality of the electromagnetic component of air showers has been thoroughly analyzed and described in literature [88–93]. For example, distributions of secondary particle energies and angular distributions



(a) Energy distributions **(b)** Angular distributions

Figure 2.1: Universality of the electromagnetic component of air showers. Shown is the universal behavior of the normalized energy distributions and angular distributions of secondary particles (from [93]). Detailed information is given in the text.

are given in Fig. 2.1. These plots are taken from [93]. In that work, the relative evolution stage t is defined as:

$$t := \frac{X - X_{\max}}{X_0}, \tag{2.1}$$

with the radiation length of electrons in air of $X_0 \approx 36.7\,\mathrm{g\,cm^{-2}}$. As such, the shower maximum is at $t = 0$. Positive values of t represent a development stage after the shower maximum while negative values state that the maximum in development was not reached yet. Distributions of secondary particles were found to be particularly universal when expressed as a function of t as opposed to slant depth X or other variables given by combinations of X and $X_{\max}$.

Normalized average energy distributions for electrons, positrons and their sum are shown in Fig. 2.1a. The red curves in the background represent simulated distributions for different primary species (p, Fe and γ) and different energies (10^{17} eV, 10^{18} eV and 10^{19} eV). The dashed lines are parameterizations from [91]. The plot is composed of three plots for young showers (top), showers at the maximum of their development (middle) and old showers (bottom). Within each of the three shower stages, the distributions among the different primaries and energies are very universal. The increased deviations for very young showers are mostly due to variations in the primary energy. It does not reflect a breakdown of the universal behavior among different primary species. In all the distributions, the most probable energy of secondary particles is in the range 10 MeV to 30 MeV. Even the distributions for different shower stages are very similar among each other.

Normalized average angular distributions for electrons are shown in Fig. 2.1b. 20 proton showers with primary energies of 10^{18} eV were used to obtain the simulated distributions.

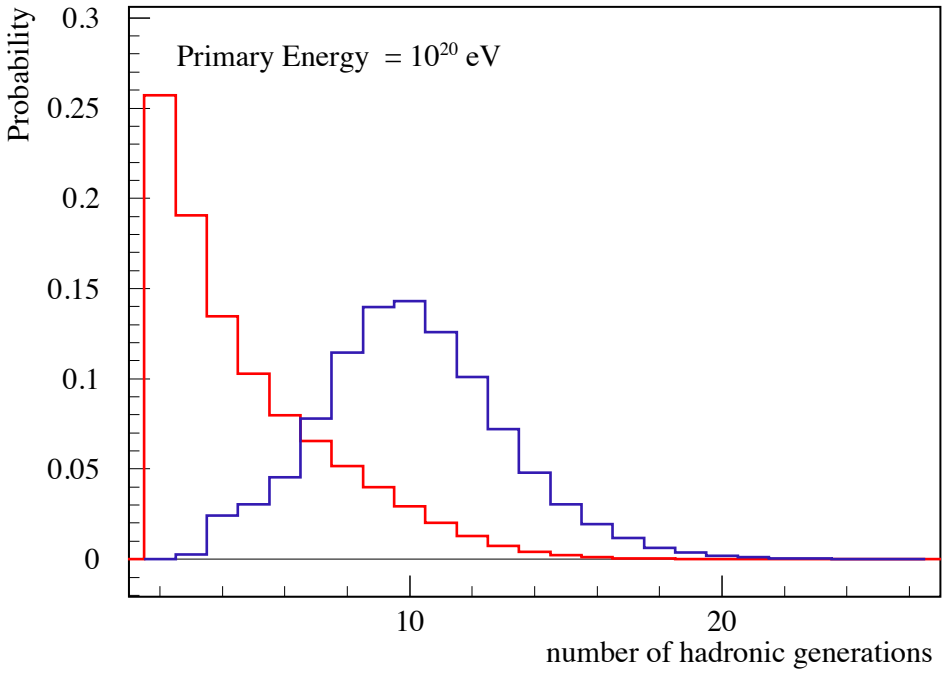


Figure 2.2: The average number of hadronic generations for the electromagnetic component (red) and muons (blue). The plot is the result from a CORSIKA simulation of a proton primary with energy 10^{20} eV (courtesy of Tanguy Pierog).

The red bands reflects the distributions with $3\,\sigma$ statistical uncertainty margins. For each secondary particle energy, the dashed lines represent parameterizations that only depend on secondary energy ϵ and momentum angle θ. The angle θ specifies the angle between the momentum vector of the secondary particle and the shower axis. With increasing secondary particle energy, most of the particles are confined very close to the shower axis. There is a significant drop in the frequency above a cutoff angle that depends on energy $\theta_c(\epsilon)$.

As it is further discussed in [93], the universal behavior of the electromagnetic shower extends to the outward momentum distribution, the lateral distribution, the delay time distribution and the shape of the shower front. Suitable parameterizations were found to describe these distributions. Limitations occur, for example, in the description of distributions for large time delays. In this case, the parameterizations do not reflect the simulations accurately. In most other cases, accurate and universal parameterizations were found. In some cases, for example in the case of the lateral distribution, the shower stage t must be added as an explicit variable to obtain a universal description.

The existence of these universal behaviors is rooted in the superposition model and a large number of interactions in an air shower. Consecutive interactions wash out individual statistical fluctuations and lead to the convergence towards universal distributions. See Fig. 2.2 for an overview about the average number of hadronic generations during the formation of air showers.

For the current universality models, the shower stage is described with ΔX (see Fig. A.1b). It is the atmospheric overburden between shower maximum X_{max} and the projected position of the station in the shower plane. The atmosphere is integrated parallel to the shower axis. As a consequence, ground stations in the early and late part of a shower have different corresponding ΔX because they are at a different development stage of the shower.

The concept of universality can be extended to showers initiated by hadrons with $A > 1$ through the introduction of an additional parameter, the relative muon content N_μ or R_μ.

With this extension, a universal description of showers initiated by protons, hadrons and photons is achieved [94–96]. The nomenclature R_μ is used throughout this work.

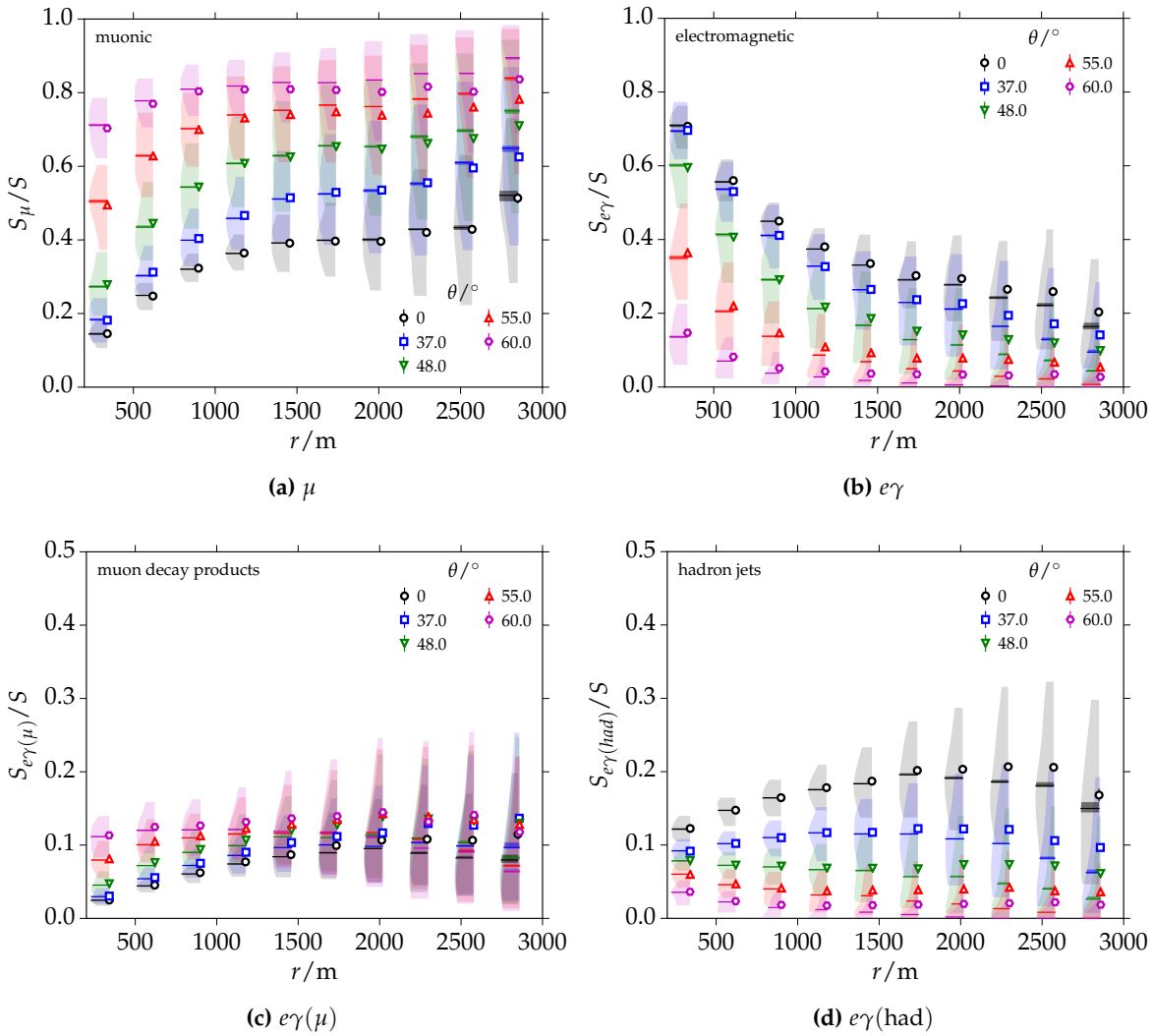
2.2 The different particle components

For the current signal and time models developed in the context of Auger data and simulations [97], particles are divided into these four particle components:

1. Muons μ^-, μ^+ $[\mu]$
2. Electromagnetic particles from high-energy π^0 decays $[e\gamma]$
3. Electromagnetic particles from muon decay or interactions $[e\gamma(\mu)]$
4. Electromagnetic particles from the decay of low-energy hadrons $[e\gamma(\mathrm{had})]$

The abbreviations in square brackets are used within this work to refer to these signal contributions. In each case, the nomenclature electromagnetic particles is used to refer to electrons, positrons and photons (while muons are also electromagnetically interacting particles, they are treated separately due to their distinct interaction characteristics in air showers). The second component refers to particles from the standard electromagnetic cascade (from π^0 decays in the early shower). The fraction of each of the signal components with respect to the total signal deposited in a WCD are shown in Fig. 2.3. Included are simulations for the three primaries proton, carbon and iron, and at the two energies $10^{19.5}$ eV and 10^{20} eV. Looking at the fraction of the muonic signal in Fig. 2.3a, it is apparent that there is a very strong dependence on zenith angle and a less pronounced dependence on radial distance. The increase in zenith angle reflects an increased average distance to the shower maximum, so a larger ΔX. The electromagnetic shower component attenuates much more quickly than the muonic component. As such, the muonic signal is dominant in the late stage of shower development (this does not directly relate to the particle numbers because an average muon deposits significantly more signal in a WCD than an electron, positron or photon). The fraction of the electromagnetic signal in Fig. 2.3b shows the opposite behavior than the muonic signal. The increase at largest radial distances is a consequence of the two simulated primary energies and a transition between the two. The $e\gamma(\mu)$ and $e\gamma(\mathrm{had})$ components in Fig. 2.3c and Fig. 2.3d typically each contribute less than 10 % to the total signal. The behavior of the muon decay products is very similar to the muonic component itself while the scaling of the hadron jet component with zenith angle is analogous to the pure electromagnetic component.

The simulated lateral distributions of the electromagnetic component for different primary species are depicted in Fig. 2.4a. To obtain a universal behavior at different atmospheric heights, the lateral distance is expressed in terms of the Molière radius r_m as $x := r/r_m$. The Molière radius spans the area (volume) around the shower axis that contains roughly 90 % of the shower energy deposit [98]. The lines for secondary energies of 1 GeV are at their actual position while the other sets of lines are shifted up by multiples of 10. The higher the energy of secondary particles, the more are they found close to the core. Moreover, the distributions show a universal behavior for the first peak closer to the core. Particles in this region are created through the main electromagnetic cascades of bremsstrahlung and pair-production, and there is no significant dependence on primary species or energy. However, there is no universal behavior for the second bulk farther from the core. Looking at Fig. 2.4b, this shape is identified to result from the decay of charged pions. By switching off this decay channel in the simulation of showers from proton primaries, the lateral distributions agree very well with the ones from γ's. 20 photon and

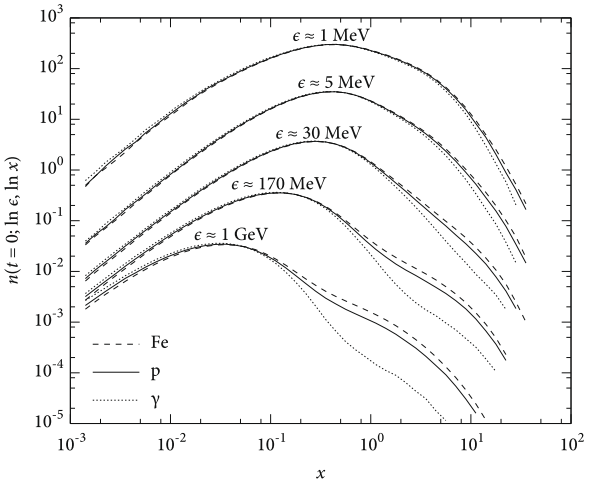


(a) μ **(b)** $e\gamma$ **(c)** $e\gamma(\mu)$ **(d)** $e\gamma(\mathrm{had})$

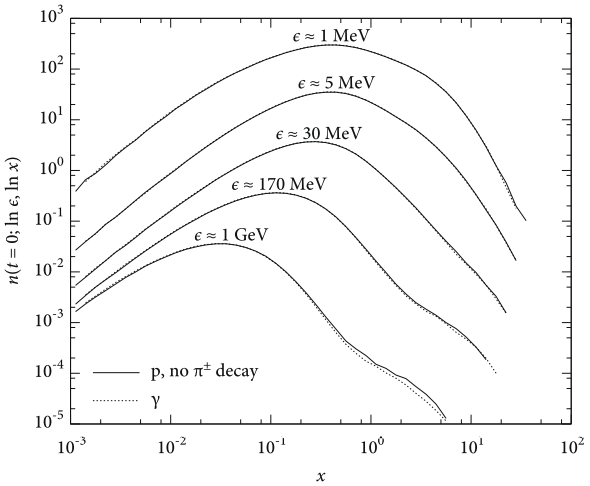
Figure 2.3: The fraction of the signal from a certain particle component with respect to the total signal deposited in a Auger WCD as a function of radial distance. Distributions for different zenith angles are given. Simulations for the primary particles p, C, Fe and energies $10^{19.5}$ eV, 10^{20} eV are included. The open markers indicate mean values for each bin, while median values are shown with the small vertical lines. A kernel density estimate of the full distribution of ratios within each radial bin is visualized with the filled shape around the mean (see Appendix A.3). The $1\,\sigma$ regimes are drawn. Note that the value range in the bottom two plots was halved due to the lower signal fractions of these two particle components.

proton showers at an energy of 10^{17} eV are used in this comparison. This is a vital observation on the universality of air showers and points to the fact that the description of muonic and hadronic components and decays in air showers is vital to reach a universal description.

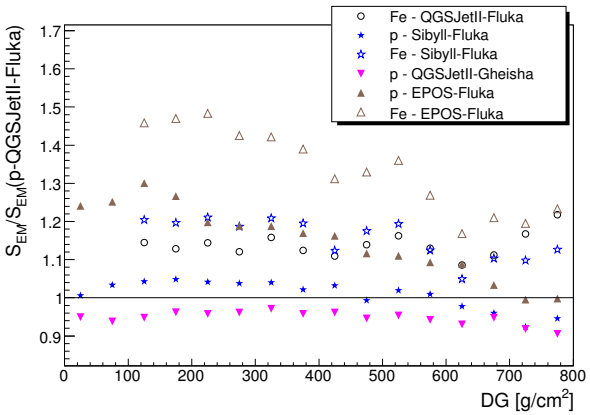
The ratio of the simulated electromagnetic signal in a WCD at radial distance of 1000 m to a reference signal from simulations with QGSJET-II.03 is shown in Fig. 2.4c. Using other interaction models, especially an old version of the EPOS interaction model (1.61), results in a large discrepancy of signals. Deviations up to 50 % are visible. Electromagnetic particles from muon decay are not counted in this comparison, but a further distinction between electromagnetic particles was not done yet. This three-component universality is strongly broken. A significant hint towards a better description is that also the muonic signals are



(a) Lateral distributions with $\pi^{\pm}$-decay



(b) Lateral distributions without $\pi^{\pm}$-decay



(c) Three component universality

Figure 2.4: (a) Simulated lateral distributions for p, Fe and γ primaries. (b) Equivalent lateral distributions for γ primaries and protons when disallowing $\pi^{\pm}$-decay. Plots from [93]. (c) The electromagnetic signal in showers simulated with different interaction models with respect to the average parametrized signal in QGSJET-II.03 simulations. Using only three particle components, a universal prediction of shower signals is not obtained (from [99]).

in a strong disagreement and that the structure and ordering of this bias with respect to interaction model and primary species are similar to the one of the disagreement in the electromagnetic signals. That implies that the classification of a fourth particle component, which is related to the overall muon content of an air shower, is necessary to universally describe air shower signals.

The introduction of the electromagnetic component from the decay of low-energy hadrons close to the ground allows one to construct a more robust and universal description of air shower signals. The particles in this group are identified through the projected position of the momentum of the mother particle onto the ground. The distance of this projected position to the shower core is denoted as r_{proj}. A schematic is given in Fig. 2.5a. A histogram of signals relating to r_{proj} is shown in Fig. 2.5b. Particles are from a simulated proton shower at $E = 10^{19}$ eV and with a zenith angle of 45°. For each radial distance there is a distinct peak close to the border of the sampling area. It gets more pronounced and important for larger radial distances. The cause for this peak are electromagnetic particles from hadrons with large traverse momenta. The identification of particles using this

criterion requires a CORSIKA simulation with enabled particle history. For the shower libraries used and described within this work, another method is used to select particles (the showers were not simulated with the CORSIKA history option). A histogram of the hadronic generation count vs. particle weight from the thinning algorithm is shown in Fig. 2.5c. Included are particles from a simulations with a sampling area around the radial distance of 400 m. The plot looks nearly identical for other radial distances. The expected signal in a WCD is indicated with the color of each bin according to the color scale. By convention, particles with hadronic generation count larger than 50 are electromagnetic particles from muon decay. For a smaller hadronic generation, there is a distinct peak for particle weights smaller than 500 and an extended peak at larger weights. The distinct peak at small generation counts is due to particles from the hadron jet component. Hadrons have a smaller maximum weight and their decay products will inherit this weight at their production. As these particles are created close to the ground, they only propagate through a small integrated density and their weights will stay significantly smaller than the ones of particles from the main electromagnetic cascade. To verify this hypothesis, electromagnetic particles with large r_{proj} are manually removed. The resulting plot in Fig. 2.5d confirms the correct identification of particles, only the pure electromagnetic component is left. The criterion based on these illustrations is currently used to identify the particle components. It is important to note that the classification of the hadron jet component has dependencies on the settings of CORSIKA simulations (for example, the maximal particle weight and the chosen thinning levels). If these settings are chosen differently as discussed in Appendix E, the criteria to select particle components need to be updated. Furthermore, an investigation of the criteria for simulations with the new hadronic interaction models, specifically QGSJET-II.04 and EPOS-LHC, would be worthwhile. Lateral distributions of the four signal components are plotted in Fig. 2.5e. The solid lines correspond to the cuts based on r_{proj}, while the dashed lines refer to the cuts in hadronic generation count and particle weight. There are only minor differences except for the hadron jet component at very small distances to the core. The clear correlation between the signal from particles of the jet component and the relative muon scale of air showers is shown in Fig. 2.5f. This correlation is exploited in the universal signal model described in the next section.

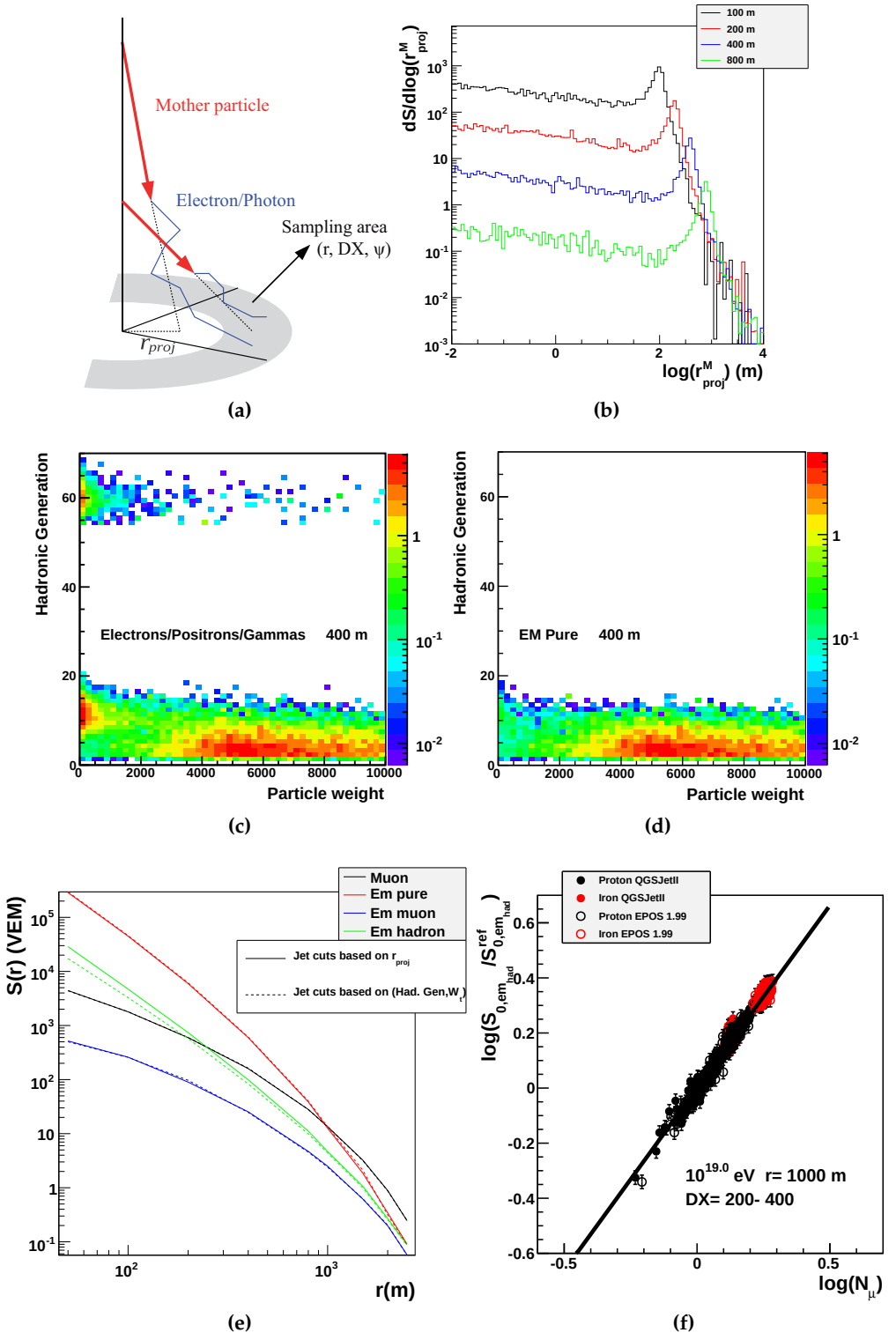


Figure 2.5: Illustrations and figures for the classification of the fourth particle component, the electromagnetic particles from hadron jets. Plots are taken from [97]. They are detailed in the text.

2.3 A model of ground signals in WCDs

Exploiting the paradigm of shower universality and the classification of particle components, a model to describe signals in the Auger WCDs was developed and later extended to new scintillator and muon detectors [97, 100]. This work deals with the analysis of current Auger data measured with WCDs, and, as such, my discussion will focus entirely on this detector type. Nevertheless, The universal description of signals in the other detectors is of high importance for the upgrade of Auger (see Section 1.2.7).

The library used to parametrize the signal model consists of proton and iron induced showers simulated with QGSJET-II.03 and EPOS-1.99. A fractional energy of 10^{-6} with respect to the primary energy was set as thinning level. This corresponds to the standard setting for simulations of showers from UHECRs. For each combination of primary and interaction model, showers with the zenith angles 0°, 12°, 25°, 36°, 45°, 53° and 60° and the energies $10^{18.6}$ eV, 10^{19} eV, $10^{19.5}$ eV and 10^{20} eV were simulated. Furthermore, 12 monthly atmospheric models were used. To account for shower-to-shower fluctuations, 10 showers for each of the combinations were created. More details are described in [97].

To parametrize the signal S in a WCD, the signal S_0 in an ideal spherical detector with an area of $10\,\mathrm{m}^2$, regardless of the incoming direction of the shower particles, is studied first. This is done for each particle component separately, and the detector response of the WCD to vertical incident particles is used. To minimize statistical fluctuations, all particles within a sampling region are used. In the next step, asymmetries due to the detector geometry and the particle production are taken into account. The asymmetries themselves are parametrized, and, together with a parametrization of S_0, the expected signal in a real detector is described.

Details of the derivation of S_0 will be omitted here. The fundamental part of the model is the dependence of S_0 on the distance to the electromagnetic shower maximum ΔX, which defines the longitudinal development of air shower signals. It is described with a Gaisser-Hillas type function

$$S_0(\Delta X, E) = S_{\max} \left(\frac{E}{10^{19}\,\mathrm{eV}} \right)^{\gamma} \left(\frac{\Delta X - \Delta X_0}{\Delta X_{\max} - \Delta X_0} \right)^{\frac{\Delta X_{\max} - \Delta X_0}{\lambda(E)}} \exp\left(\frac{\Delta X_{\max} - \Delta X}{\lambda(E)} \right), \qquad (2.2)$$

with $\lambda(E) = \lambda_0 + f_\lambda \lg\left(E/10^{19}\,\mathrm{eV}\right)$. The description is obtained for each particle component separately. The energy evolution f_λ is 0 for the muonic component and for the muon decay products. To fit the description of $S_0(\Delta X, E)$, a likelihood method was used. The fluctuations in S_0 were obtained from the distributions of all signals within windows of $\pm 20\,\mathrm{g\,cm^{-2}}$; they were found to be approximately independent of energy. Results of the longitudinal parametrization $S_0(\Delta X)$ are shown in Fig. 2.6 for muons and for the pure electromagnetic component. A radial distance of 1000 m is chosen and different zenith angles are color-coded as specified in the label of the figure. Showers for different primary energies are also included. The evolution of signals with energy is corrected for via $\lambda(E)$. The parameterizations are depicted as black lines. It is evident that the attenuation of the muonic component is much weaker than the one of the electromagnetic component. The muon decay products behave very similarly to the muons, while the hadron jet component behaves similarly to the pure electromagnetic component. Deviations in the description of S_0 are smaller than 5 % for almost all regions of the parameter space.

The parameters of the longitudinal description $S_{\max}, \Delta X_{\max}, \gamma, \lambda_0$ and f_λ additionally depend on the distance to the shower core. Parameterizations for these dependencies were found. For example, $S_{\max}(r)$ is described with a power-law LDF[1]. At an energy of 10^{19} eV,

[1] lateral distribution function

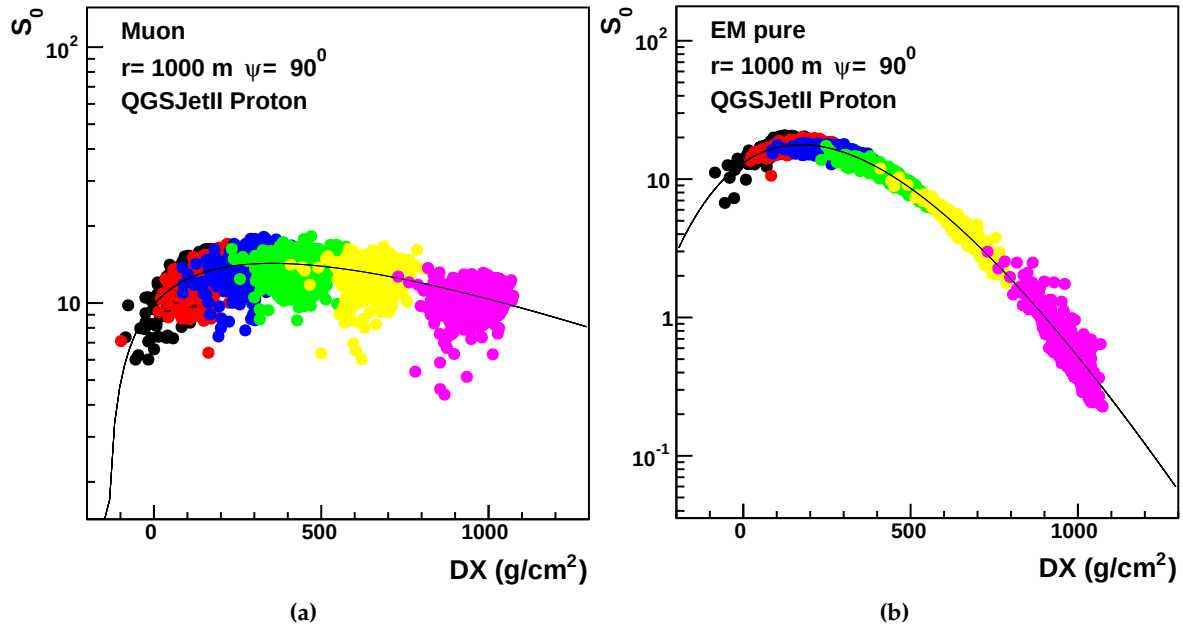


Figure 2.6: The longitudinal dependence of signals in an ideal detector for different sets of zenith angles (black: 12°, red: 25°, blue: 36°, green: 45°, yellow: 53° and magenta: 60°). Simulations are compared to the parametrization derived in [97] and briefly described in this section. (a) Muons (b) Pure electromagnetic component.

this LDF is independent of the primary species and hadronic interaction model. Together, the parameterizations of the longitudinal and the lateral distributions of the ideal signal establish the model $S_0(\Delta X, r, E)$.

The signal in a real detector is approximately given by:

$$S(r, \Delta X, E, \theta, \psi) = S_0(r, \Delta X)\ f_{\text{trunc}}(r, \theta, \psi)\ A_{\text{eff}}(r, \Delta X, \theta, \psi). \tag{2.3}$$

Two modifications to the signal in an ideal detector are applied. f_{trunc} denotes truncation asymmetries. These occur because of the presence of the ground and the suppression of particle propagation and production. For example, considering a detector in the late part of an inclined shower (ψ = 180°) and at a large distance to the core, particles created along the shower axis and at similar values of ΔX will be effectively suppressed because of the presence of the ground. A small dependence of f_{trunc} on ΔX itself is negligible and was ignored. A_{eff} describes geometrical asymmetries due to the non-spherical shape of the detector. It also includes the fact the particles with certain transverse momenta will not reach a detector at a certain position:

$$A_{\text{eff}}(r, \Delta X, \theta, \psi) = T_{\text{mod}}(\theta_p(p_z^{\text{eff}}, \theta, \psi))\ A_{\text{mod}}(\theta_p(p_z^{\text{eff}}, \theta, \psi)) \int_{p_z^{\text{cut}}(\theta_p=90°|\theta,\psi)}^{1} \frac{\mathrm{d}S_0}{\mathrm{d}p_z}\,\mathrm{d}p_z. \tag{2.4}$$

The transverse momentum p_z^{cut} describes the minimal momentum that a particle needs in order to reach a detector at a certain radial distance and geometry. p_z^{eff} is a median value between the cut value and 1. This approximation works well with an accuracy of 5 %. Different atmospheric profiles are taken into account during the calculation of the atmospheric overburden ΔX.

Using these parameterizations, the total signal is given by:

$$S(r, \Delta X, E, \theta, \psi) = S_{e\gamma} + R_\mu \left(S_\mu + S_{e\gamma(\mu)}\right) + R_\mu^{\gamma(r)} S_{e\gamma(\text{had})}. \tag{2.5}$$

The exponent $\gamma(r)$ is modeled as a function of core distance. The correlation can be extracted from Fig. 2.5f. An example of the different LDFs is illustrated in Fig. 2.5e. R_μ is a relative measure of the muon number in an air shower:

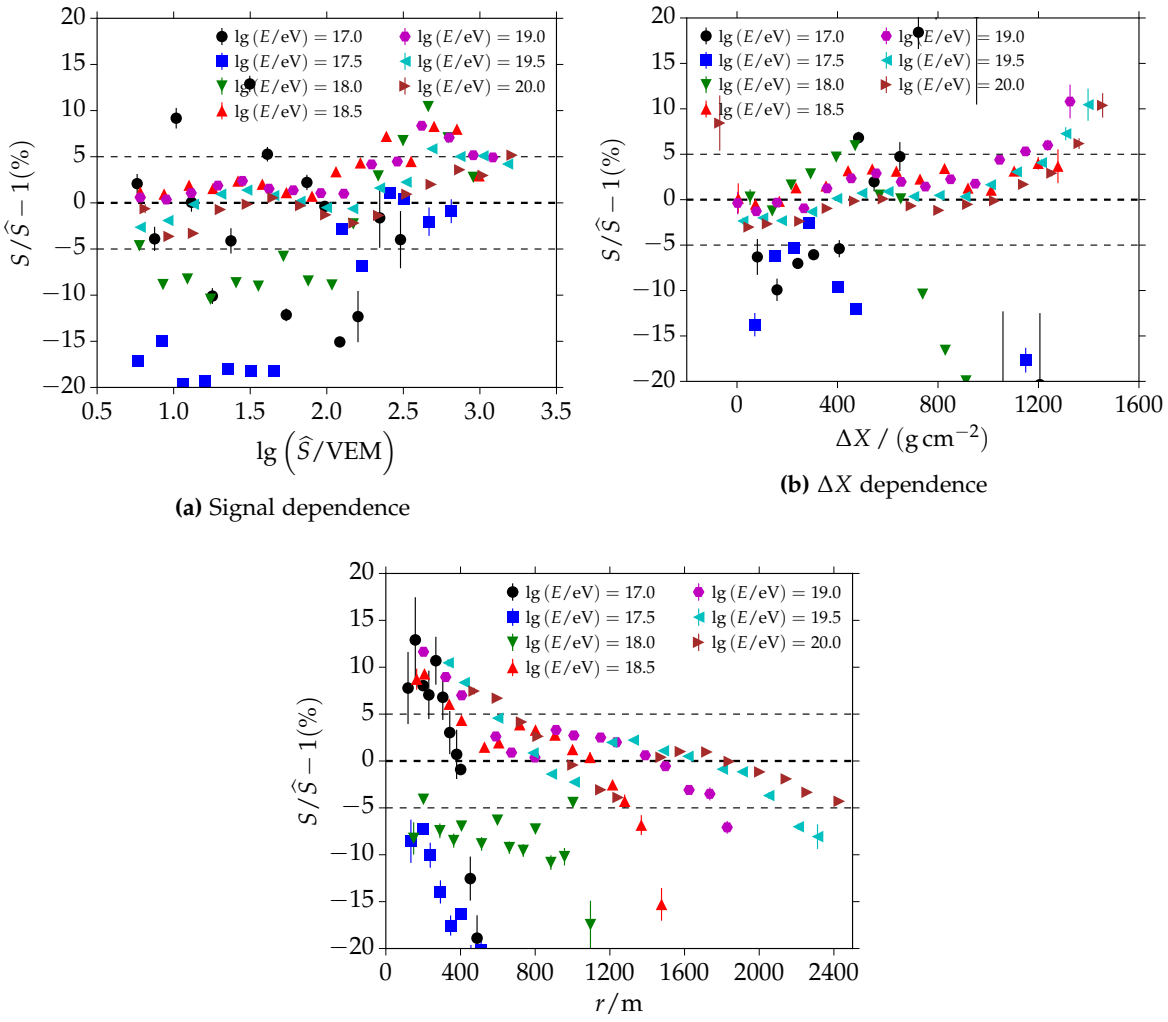
$$R_\mu(E) := \frac{S_{0,\mu}\left(r = 1000\,\mathrm{m}, \Delta X = 400\,\mathrm{g\,cm^{-2}}, E\right)}{S^{\mathrm{ref}}_{0,\mu}\left(r = 1000\,\mathrm{m}, \Delta X = 400\,\mathrm{g\,cm^{-2}}, E = 10^{19}\,\mathrm{eV}\right)}. \tag{2.6}$$

The reference $S^{\mathrm{ref}}_{0,\mu}$ is the signal of a proton shower at 10^{19} eV simulated with QGSJET-II.03 and at local shower azimuth of $\psi = 90°$.

The signal model was parametrized for primary energies above $10^{18.6}$ eV. As part of this work, I reinvestigate the accuracy of the signal model at large energies and also analyze the quality of an extrapolation to lower energies. A set of independent simulations with energies down to 10^{17} eV is used. This set of simulations is detailed in Appendix E.1 and it is used to derive the time model in Section 2.4.

To analyze the accuracy of the signal model, each simulated signal S in the shower library is compared to its model prediction $\widehat{S}$ from Eq. (2.5). The quantity of comparison is the relative residual $S/\widehat{S} - 1$. Due to limitations in the parameterization of the signal model, the model is valid for radial distances between 100 m and 2000 m. Only signals of stations within 2500 m from the core are used for the comparison. In addition, a cut on the expected signal of $\widehat{S} > 5$ VEM is applied in order to avoid a distortion of the signal distribution at low signals due to the trigger threshold of the WCD. Depending on the signal shape, a WCD saturates in the signal range 1000 VEM to 2000 VEM. Stations with expected signals larger than 2000 VEM and low-gain saturated stations are neglected. To account for differences in the relative muon content R_μ, the muon content of each simulated event is estimated from the dense ring of stations at a radial distance of 1000 m. This calculation is done with respect to the reference primary and interaction model as stated in Eq. (2.6). The estimate is used in the calculation of the total expected signal.

Relative signal residuals for different primary energies are shown in Fig. 2.7. The residuals in Fig. 2.7a are plotted as a function of the expected signal $\widehat{S}$. Different markers and colors represent different primary energies (exact Monte Carlo energies, not bins). The uncertainty of the mean values of each bin are depicted with error bars. Above primary energies of $10^{18.5}$ eV, deviations are within ± 5 %. There is a trend with the expected signal itself: small signals are on average slightly overestimated while large signals are underestimated. The accuracy of the model is very good in the most relevant and intermediate signal range. The predictions show a larger bias of 5 % at a primary energy of 10^{18} eV and even 15 % at an energy of $10^{17.5}$ eV. The accuracy of the model increases again towards lower energies (10^{17} eV). This is an artifact due to the extrapolation of the signal model. The ΔX dependence in Fig. 2.7b shows a very similar structure. The largest deviations occur at lowest primary energies and for very large distances to the shower maximum. The signal residuals are plotted as a function of distance to the core in Fig. 2.7c. Again, the signal model gives a robust prediction in the intermediate radial range with deviations below 5 %. Larger biases are apparent for stations very close and very far from the core. There are again larger differences for the primary energies $10^{17.5}$ eV and 10^{18} eV, while signals of showers with an energy of 10^{17} eV are described with deviations below 10 %. Thus, using the independent set of simulations, the accuracy of the signal model of 5 % for primary energies above $10^{18.5}$ eV is confirmed. At lower energies, the accuracy decreases to less than 15 %. These deviations are still within a reasonable range for this work, but a detailed parameterization of the signal model at lower energies should be studied in the future. This will be very important for studies of AMIGA data with shower universality. A universal description of air shower signals down to energies of 10^{17} eV is possible, while universality



(a) Signal dependence

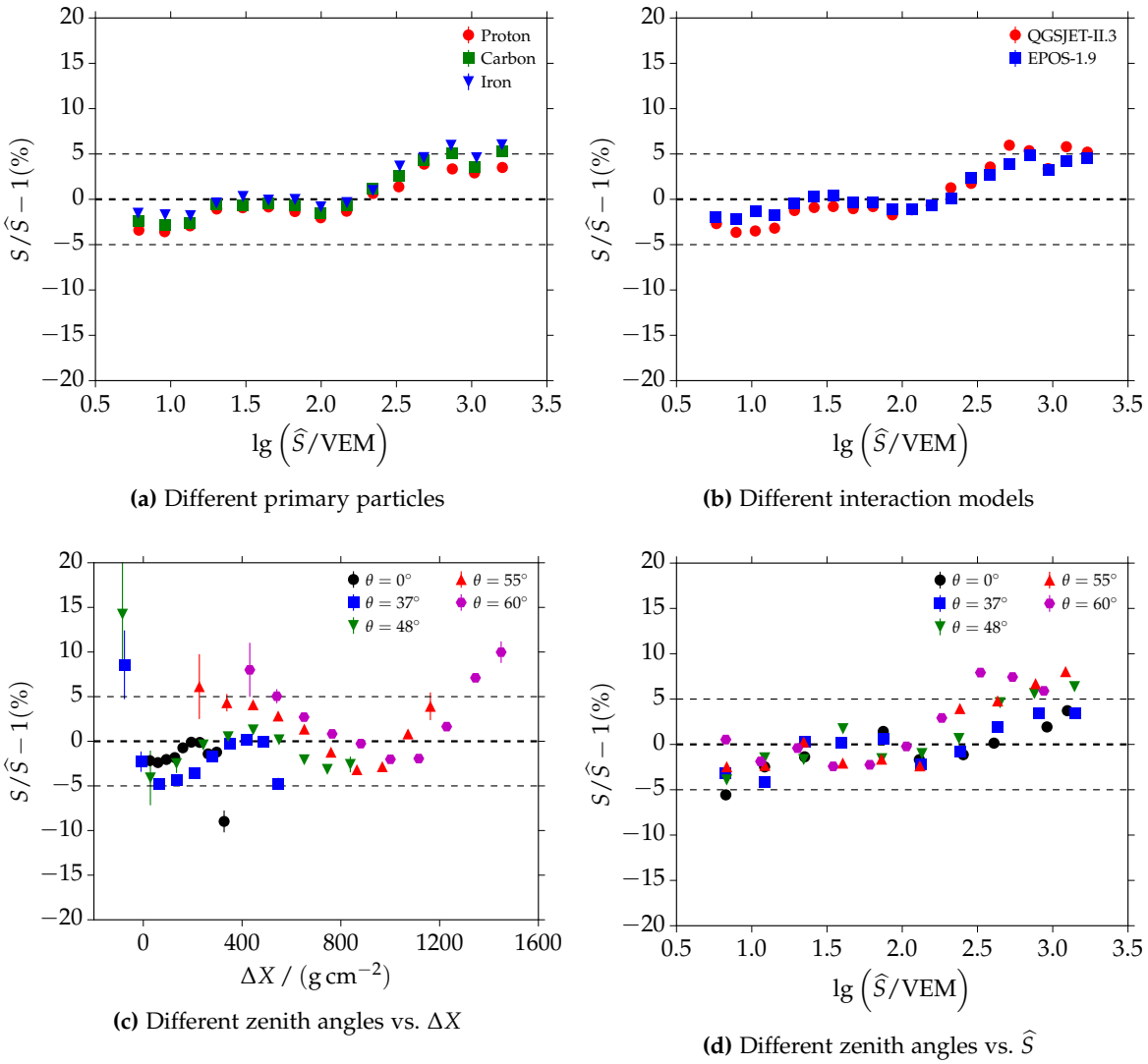
(b) ΔX dependence

(c) Radial dependence

Figure 2.7: Comparison of simulated signals to the model prediction from the universality parametrization presented in this section. The relative difference of signals is shown with respect to different shower parameters and for different primary energies.

might break down at even lower energies due to significantly larger fluctuations in the development of air showers.

Further signal residuals are depicted in Fig. 2.8. The residuals for different primary particles and interaction models in Fig. 2.8a and Fig. 2.8b show a similar trend as a function of expected signal. There is no bias for different primary particles and interaction models; all of the signals are well described by the model. Residuals for different zenith angles are shown in Fig. 2.8c as a function of ΔX and in Fig. 2.8d as a function of expected signal. There are no significant biases for particular zenith angles. However, I want to point out the non-trivial trends of the ΔX-dependence for a certain zenith angle. These are residual dependencies that are not fully taken into account in the parameterization of the signal model. A parameterization as a function of ΔX and zenith angle is very challenging because different ranges in ΔX are covered for different zenith angles. This is a consequence of the shower geometry. Only the overlapping ranges provide information about both dependencies at the same time without loss of correlations. For future analyses,



(a) Different primary particles

(b) Different interaction models

(c) Different zenith angles vs. ΔX

(d) Different zenith angles vs. $\widehat{S}$

Figure 2.8: Comparison of simulated signals to the model prediction from the universality parametrization presented in this section. The relative difference of signals is shown with respect to different air shower parameters.

I want to recommend a shower library with at least eight different zenith angles in order to approach this parameterization.

In the context of simulations with different hadronic interaction models, it is particularly interesting to study the ratio of the component signals with respect to the total signal. This ratio differs for different interaction models as does the fraction between these ratios for different models. Additional plots on this topic are included in Appendix B.2.

2.4 A time model of signals in WCDs

This section describes the analysis of the arrival time distributions of particles from the different signal contributions in a WCD and the derivation of a model to describe these distributions. The model exploits the paradigm of shower universality and is used in the reconstruction of events as described in Section 3.4. It was developed as part of this thesis and as an extension of a previous model that was presented in [96].

The procedure to parametrize the time model is based on many previous efforts, most prominently on [94, 97, 100–104]. The motivation to create a new model is to improve on some aspects of the old description [101], especially to extensively study dependencies on primary energy, hadronic interaction model and zenith angle. Also, the overall reconstruction bias and resolution of $X_{\max}$ should be minimized. There are other technical aspects that will be mentioned in the later course of this section.

CORSIKA and $\overline{\text{Off}}\underline{\text{line}}$ detector simulations as described in Appendix E.1 are used for this analysis. Particles are divided into four separate components as described in Section 2.2. The SD detector simulation includes the hardware properties and calibration as discussed in Section 1.2.1. In particular, time traces of signals are simulated in 25 ns bins. Saturation effects are not simulated for time traces of specific particle components, only for the total trace. In addition, an unsaturated version of the total trace is stored. The traces for different particle components are from now on referred to as *component traces*.

I will follow the convention that the total signal (charge) $S_i := S_i(t_{100})$ of a particle component i is obtained from its instantaneous time trace $\mathrm{d}S_i/\,\mathrm{d}t(t)$ as follows:

$$S_i = \int_{t_0}^{t_{100}} \frac{\mathrm{d}S_i}{\mathrm{d}t}(t)\,\mathrm{d}t, \tag{2.7}$$

with non-zero signal in the time range t_0 to t_{100} (100 % quantile), i.e. $\int_{-\infty}^{t_0} \frac{\mathrm{d}S_i}{\mathrm{d}t}(t)\,\mathrm{d}t = \int_{t_{100}}^{\infty} \frac{\mathrm{d}S_i}{\mathrm{d}t}(t)\,\mathrm{d}t \equiv 0$. As the total signal is calibrated in units of VEM, $\dot{S}_i = \mathrm{d}S_i/\,\mathrm{d}t$ has units of VEM/ns (note that the actual unit VEM already includes time as being proportional to charge, thus VEM/ns is actually a current). The time-dependent signal in a WCD depends on the arrival time distribution and the detector response to these particles. The latter includes both the time and signal response to a particle and depends on particle type, energy, momentum and incoming direction.

On average, one *vertical centered through-going* (VCT) muon at 1 GeV creates 79 *photoelectron* (PE) in a PMT. For an arbitrary muon, the expected number of PEs depends mainly on the muon track-length and, to a much lesser extent, on the muon energy. The actual number of emitted Cherenkov photons for one VCT is on the order of 50 000. The measured number of PEs is a Poisson random variable. This is due to statistical fluctuations in the production of Cherenkov photons along the muon trajectory. After emission, the Cherenkov photons can be reflected several times before entering a PMT. Most of them are absorbed or do not displace a PE from the cathode (PMT efficiency typically below 30 %). The arrival time distribution of photons from one VCT muon in a PMT increases about linearly in the first $\approx$ 15 ns and decreases exponentially afterwards. Furthermore, the spectrum of signal caused by a single PE can be approximated with a Gamma distribution. The actual value is digitized afterwards with a FADCs. Thus, an analytical description of the expected time trace from an observed muon leads to a complicated convolution of several distributions that are by itself often approximations. First, successful attempts to find a fitting parametrization are described in [105, 106]. The semi-analytical model of the detector response to a single VCT muon works on average, but still the measured fluctuations are underestimated (see Fig. 3 in [106]). Reasons for this are not fully understood, but an important factor is the energy-dependence of the muon yield, which is still studied.

When measuring real air showers, we have a multitude of particles of different type and energy entering the detector at different positions and with different momentum vectors. An analytical expression of the detector response is, at the very least, hard to derive under these conditions. To speed up time-expensive simulations with GEANT4, the detector response was tabulated [107].

In consideration of a parametrization of time shapes, it would be most elegant to have a model that describes the flux of secondary particles on ground (with a fixed height above sea level or as a function of height) as a function of particle type i, secondary energy ϵ_i, momentum vector $\vec{p}_i$ and primary characteristics like energy E, mass A and zenith angle θ. The geometry within a plane perpendicular to the shower axis is also relevant, mostly the distance to the shower core r and the shower azimuth angle ψ (see Fig. A.1a). Thus, the desired model would read:

$$j_i = \frac{\mathrm{d}N_i}{\mathrm{d}t\,\mathrm{d}S\,\mathrm{d}\epsilon_i\,\mathrm{d}\vec{p}_i}\,(t, \epsilon_i, \vec{p}_i | E, A, \theta, r, \psi) \tag{2.8}$$

with the surface element $\mathrm{d}\vec{S}$ perpendicular to the particle direction. This holds purely on the level of secondary particles prior to the detector response. For simplicity, various other dependencies, like the state of the atmosphere, are neglected here.
With a known detector response $R_i(t, t', \epsilon_i, \vec{p}_i)$ for particles of type i, the time-dependent signal can be expressed as:

$$\begin{aligned}\frac{\mathrm{d}S}{\mathrm{d}t}(t|E, A, \theta, r, \psi) = \sum_i \int \mathrm{d}t' \int \mathrm{d}S \int \mathrm{d}\epsilon_i \int \mathrm{d}\vec{p}_i\, R_i(t, t', \epsilon_i, \vec{p}_i) \\ \cdot \frac{\mathrm{d}N_i}{\mathrm{d}t\,\mathrm{d}S\,\mathrm{d}\epsilon_i\,\mathrm{d}\vec{p}_i}\,(t, \epsilon_i, \vec{p}_i | r, \psi, E, A, \theta)\,,\end{aligned} \tag{2.9}$$

with the sum running over different particle components. This equation can be simplified when introducing momentum and energy thresholds (specific to the detector in question), as well as coordinate transformations similar to what has been done in [97, 100]. This is not the purpose here, further remarks about possible analyses in this direction are included in Appendix A.4.1.

As we do not have an analytical expression of the detector response and the main goal is to reconstruct properties of the primary particle, the approach followed for this work was to find a model for the temporal signal *after* the simulation of the detector response. The procedure is not limited to a specific type of detector. Thus, I seek a parametrization of:

$$\frac{\mathrm{d}S_i}{\mathrm{d}t}(t) = \frac{\mathrm{d}S_i}{\mathrm{d}t}(t | r, \Delta X, \theta, \psi, E) \tag{2.10}$$

with the distance to the shower maximum ΔX (schematically shown in Fig. A.1b). This model is used in air shower reconstructions together with the signal model (described in Section 2.3). Particles are divided into distinct components as given in Section 2.2. A parametrization of Eq. (2.10) is created independently but similarly for each of the particle components. The mass of the primary particle A is not an explicit parameter of the model. Expressed as a function of ΔX, the dependence of the model on the primary species A is weak as will be shown later in Section 2.4.4. In Chapter 5, reconstructed shower maxima X_{max} and relative muon numbers R_μ will be used to infer information about the mass of primary particles.

The ansatz of the model is to either use a log-normal (Appendix F.1.1) or a generalized gamma distribution (Appendix F.1.2) to parametrize $\frac{\mathrm{d}S_i}{\mathrm{d}t}(t)$. While the log-normal has two parameters m and s (proportional to mean and standard deviation, but not exactly equivalent), the generalized gamma distribution has three, denoted as m, s and ℓ. For $\ell = 0$,

the generalized gamma distribution reduces to a log-normal distribution with equivalent parameters m and s. Depending on the quality of fit to simulated traces, either of the two functions is chosen. The model function is abbreviated as $f_{\text{trace}}(t)$. Other functions to describe the time distribution of particles were tested. Among these are the (generalized) Gumbel distribution, the Moyal distribution, the Gamma distribution, the Inverse-gamma distribution, the Levy distribution and the Frechet distribution [108]. None of those distributions yield a better description of time traces.

The following explanations will explain the parametrization of the time model for WCDs. The start of each simulated trace is determined as the FADC bin in which the calibrated signal exceeds 0.1 VEM. This is fulfilled by a single muon if it is not strongly corner-clipping. The algorithm is run in simulations in the same manner as during the reconstruction of data. It accounts for accidental peaks in traces or fluctuations to give a robust estimation of the start bin (for example, isolated bins due to background particles are rejected). The total simulated trace with a fluctuating baseline is analyzed. There is no baseline simulation for the component traces; the simulated number of PEs are written into them without the simulation of electronic fluctuations. An example trace with the estimated start bin is shown in Fig. 2.9a. A small portion of deposited signal might arrive earlier than the estimated start time. Fig. 2.9b shows are rare case with signal entries in two bins prior to the estimated start time. Due to fluctuations and the electronics simulation, small signals are unlikely to stem from actual particles in the detector. However, some particles arrive earlier than the plane front time, which should not be possible. It is currently unclear if this is due to biases from the method of shower resampling or due to bugs in the Offline framework. More details are given in Section 2.4.6. All times are expressed relative to the time when the shower core hits the ground. For the parametrization of the time model, times are calculated relative to the first possible arrival time of a particle. As shown in [100], the shower front can be parametrized as a parabola with a different curvature depending on the particle component. The quantity of interest is the difference in height between the effective origin of times and the shower maximum, which depends on the particle component and properties of the primary particle. For the muonic component of proton showers at different zenith angles, the height difference between this effective origin of times and the shower maximum is plotted in Fig. 2.9c as a function of X_{max}. There is a linear decrease with increasing X_{max}. This is connected to the energy spectrum of muons reaching the detector. The actual start times from the curvature model are plotted in Fig. 2.9d. Due to the fact that muons propagate almost without interaction, they will arrive prior to the more frequently scattered electrons, positrons or photons. When a muon decays close to the ground, its decay products ($e^{\pm}$ detectable) will arrive almost together with the first muons though (if they appear to arrive even earlier in Fig. 2.9d, it is due to small inaccuracies of the model). The curvature model for muons is almost identical to a curved shower front originating at the first interaction point X_0. The start times of individual component traces are then expressed relative to the expected start time for the respective particle component.

In the next step, traces are divided into different bins with respect to core distance r, distance to shower maximum ΔX, primary energy E, zenith angle θ and azimuth angle ψ. The employed Monte Carlo shower library has fixed primary energies and zenith angles as described in Appendix E.1. As such, they are chosen as bins. The binning in radial and azimuthal dimension is depicted in Fig. 2.10. Stations are positioned either at fixed distances to the core (in rings of dense stations) or at random positions depending on the random location of the Monte Carlo core within the array. Bins are chosen centered at the locations of dense stations and in between. A compromise between a uniform bin width and a reasonable number of entries is chosen. The distribution of ΔX depends strongly on the particular (r, ψ, θ, E) bin (concerning the range of values), but it is rather uniform

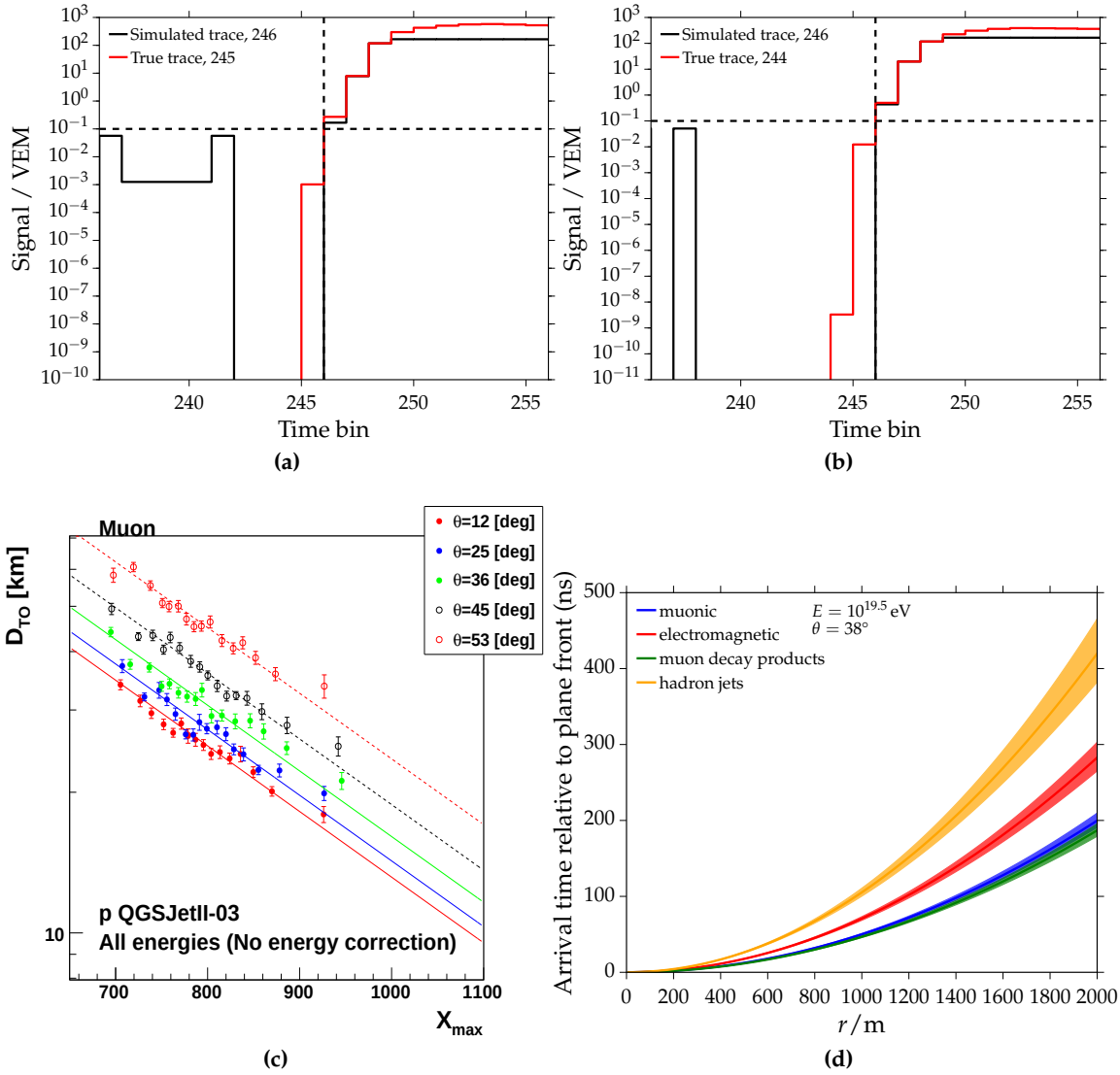


Figure 2.9: (a) Example of a simulated trace in an WCD. The simulated trace with baseline simulation and saturation is plotted as a black histogram, while the true trace is shown in red. Baseline fluctuations are visible as black boxes before the actual start of the trace in bin 245. A small fraction of signal is apparent before the start time bin. To pronounce the start of the trace, the histogram is given in logarithmic units. (b) In this example, there are two bins with small signal entries before the estimated start bin. Baseline fluctuations are visible before the actual start of the particle trace. (c) Distance from the effective origin of times to the shower maximum in km as a function of the shower maximum and for different zenith angles. (d) The start time model as a function of core distance. Times are given relative to the arrival time of a plane front traveling with speed of light c. The model is plotted for a fixed energy and zenith angle as denoted in the plot. The four signal components are visualized with different colors. The shaded bands around each model line represent the variation due to different azimuthal angles.

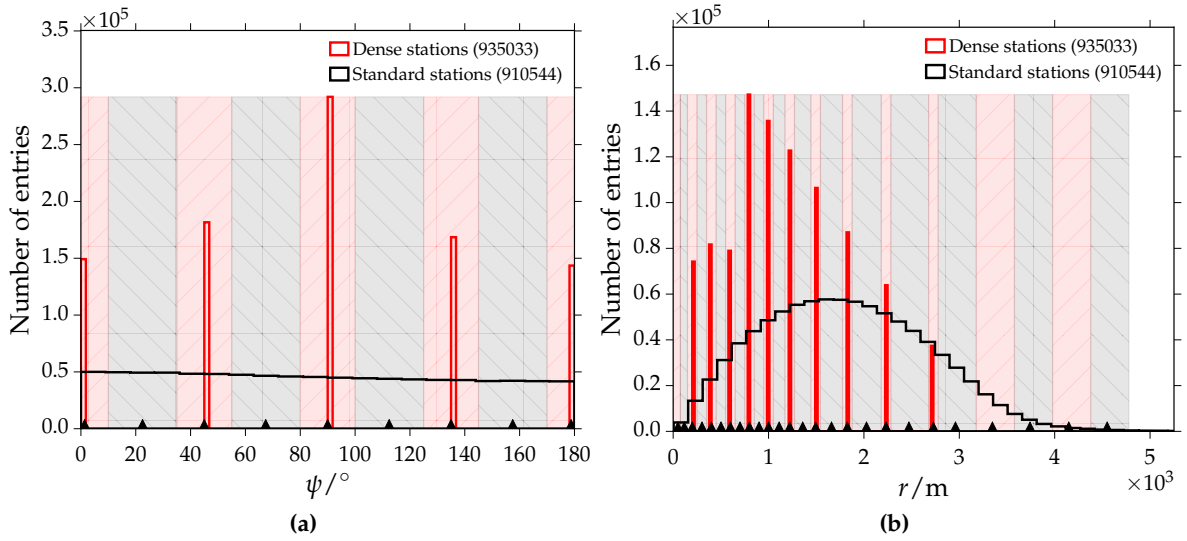


Figure 2.10: Distribution of stations in the Monte Carlo library as a function of local shower azimuth ψ in (a) and radial distance to the shower core r in (b).

in most cases. ΔX bins are chosen such that they include more than 30 and less than 60 traces. To avoid distortions due to trigger thresholds, only traces with an expected total signal above 4 VEM are considered. The expected signal is calculated from the signal model discussed in Section 2.3. This selection removes 28 % of the 1.8×10^6 traces. To reduce statistical and thinning fluctuations, a weighted average of traces within each $(r, \psi, \theta, E, \Delta X)$ bin is calculated. The traces are all binned in intervals of 25 ns, but they are not in phase. First, the traces are resampled to be in phase with a common set of time bins $\{t_j\}_{j=1\ldots n}$ with $t_1 > 0$. Then a weighted average of all n_{trace} normalized traces in time bin x is calculated (for particle component i):

$$\langle \tilde{S}_i \rangle (t_x) = \frac{\sum_k^{n_{\text{trace}}} \tilde{S}_i^k(t_x)\, S_i^k}{\sum_k^{n_{\text{trace}}} S_i^k} \tag{2.11}$$

with the normalized signal bin $\tilde{S}_i^k(t_x) = \frac{\dot{S}_i^k(t_x)}{\int \mathrm{d}t\, \dot{S}_i^k(t_x)} =: \frac{\dot{S}_i^k(t_x)}{S_i^k}$.
Simplifying Eq. (2.11) leads to:

$$\langle \tilde{S}_i \rangle (t_x) = \frac{\sum_k^{n_{\text{trace}}} \dot{S}_i^k(t_x)}{\sum_k^{n_{\text{trace}}} S_i^k} \tag{2.12}$$

$$= \frac{\langle \dot{S}_i \rangle (t_x)}{\int \langle \dot{S}_i \rangle (t)\, \mathrm{d}t} \tag{2.13}$$

meaning that one can either normalize the traces and calculate a weighted average or calculate the mean of traces and normalize it afterwards, the results are equivalent. The uncertainty of a signal bin is estimated from the standard deviation of the mean of all normalized signals within the bin:

$$\sigma_i(t_x) = \frac{1}{n} \sqrt{\mathrm{Var}[\tilde{S}_i(t_x)]}. \tag{2.14}$$

In case of a small number of traces n, this procedure yields an underestimated standard deviation. To cope with that, a correction as described in Appendix F.2.3 is applied. For most sample sizes, this is a correction on the percent level or smaller.

The average traces are fitted with a chi-squared fit. For simplicity, I will denote the normalized signal in bin j with S_j and its uncertainty σ_j. Instead of comparing the value of the time model $f_{\text{trace}}(t_j)$ with S_j, the model is integrated within the bin to properly take into account its nonlinear shape. The model estimate for bin j is thus:

$$\hat{S}_j = \frac{1}{\Delta t} \int_{t_j - \Delta t/2}^{t_j + \Delta t/2} \mathrm{d}t \, f_{\text{trace}}(t) = \frac{1}{\Delta t} \left(F_{\text{trace}}(t_j + \Delta t/2) - F_{\text{trace}}(t_j - \Delta t/2) \right), \tag{2.15}$$

with the width of the time bin $\Delta t = 25\,\text{ns}$ and the c.d.f.[2] $F_{\text{trace}}(t)$.
The χ^2-function is constructed as follows:

$$\chi^2 = \sum_{j=1}^{n} \left(\frac{S_j - \hat{S}_j}{\sigma_j} \right)^2. \tag{2.16}$$

Due to deviations of the time shape at large times, only time bins below the 95 % time quantile are taken into account. Eq. (2.16) is minimized with the software packages NLopt or Minuit-2 [109, 110], which give identical results. As the minimization of the generalized gamma distribution is rather sensitive with respect to start values, the following procedure turned out to be reasonable:

1. Estimation of start parameters m_0, s_0 from the mode and median of the average trace.
2. Fit of a log-normal model with start parameters m_0, s_0 to obtain m_1 and s_1 (and a covariance matrix of the parameters).
3. Estimation of the third parameter ℓ_0 depending on the χ^2 probability of the log-normal fit. ℓ_0 is close to zero for a very good fit and increases towards 1 for a bad fit ($\ell > 1$ leads to numerical problems).
4. Fit of a generalized gamma model with start parameters m_1, s_1 and ℓ_0. New estimates m_2, s_2 and ℓ_2 are obtained.

Examples of fits to average simulated traces in WCDs are depicted in Fig. 2.11. The different plots reflect different particle components as stated in the plots and legend. In this case, with a primary energy of $10^{19}\,\text{eV}$ and an average distance of 700 m from the core, the log-normal model works well, and the fits with a generalized gamma distribution yield no improvement. Note that in case of an equivalent χ^2, the χ^2/n_{dof} using the generalized gamma distribution is larger as there is one less degree of freedom. In each plot, the shaded distributions in the background show the variation of the 29 simulated traces within the time bins. The darker regions represent the $1\,\sigma$ spread, while the transparent regions extend by $\pm 3\,\sigma$. The black triangles at the bottom of the plots represent the t_{10}, t_{50} and t_{90} time quantiles. The integral signal of the average trace up to these points is 10 %, 50 % and 90 % of the total signal. The small inlet plots illustrate the distributions of the component signals. The air showers for those examples were simulated with a zenith angle of 0. The stations are located in the early part of the shower around $\psi = 0$. Due to the small zenith angle and relatively large energy, the average shower maximum is very close to the ground. On average, the generalized gamma distribution describes the simulated traces slightly better than the log-normal distribution as can be seen in the χ^2/n_{dof}-distributions plotted in Fig. 2.12. The difference is very small though and uncertainties in the description of the third shape parameter ℓ are large. Thus, a log-normal description seems favorable at this point of the analysis. In case of an optimal description of the time traces, $\langle \chi^2/n_{\text{dof}} \rangle \approx 1$

[2] cumulative distribution function

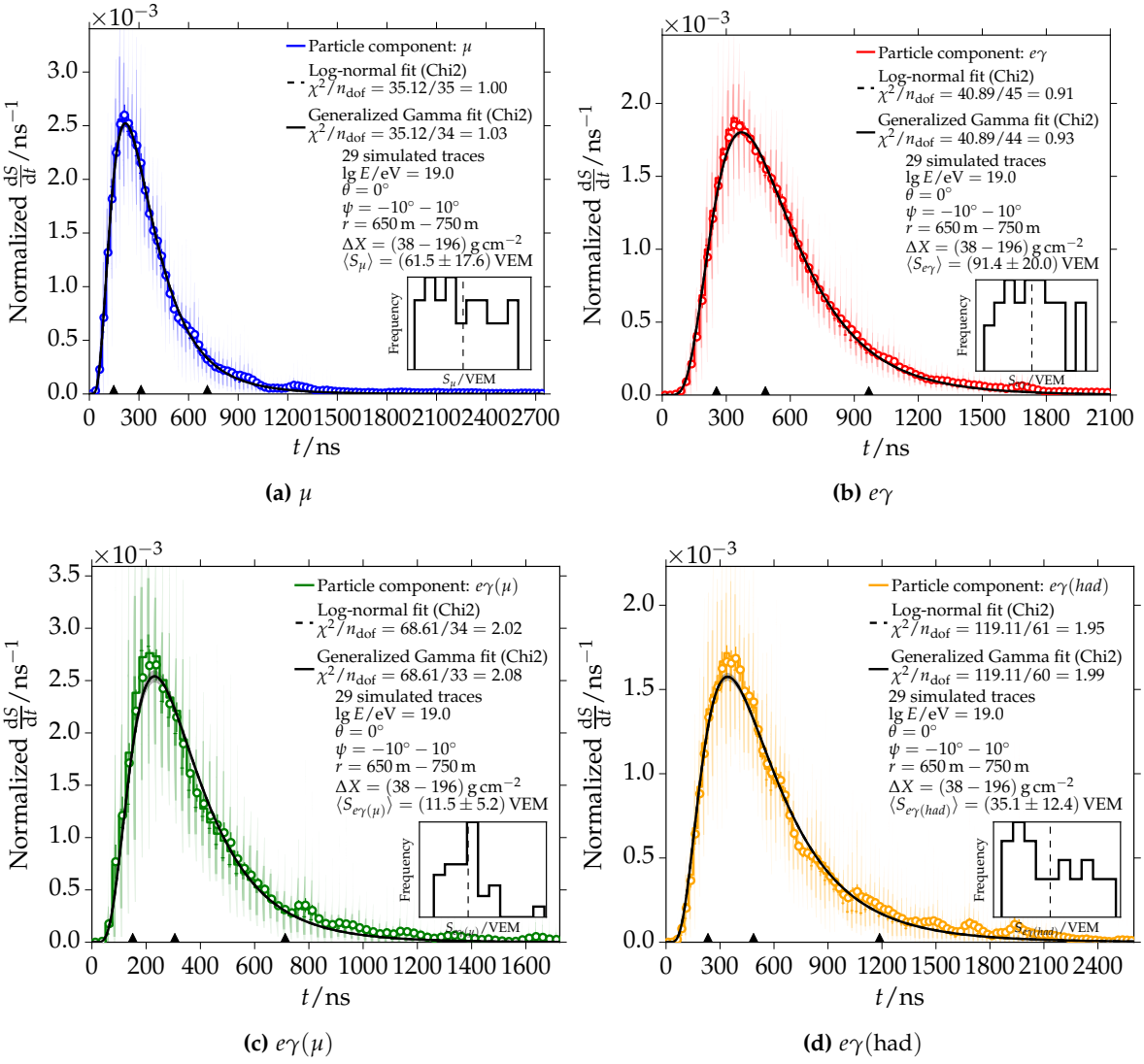
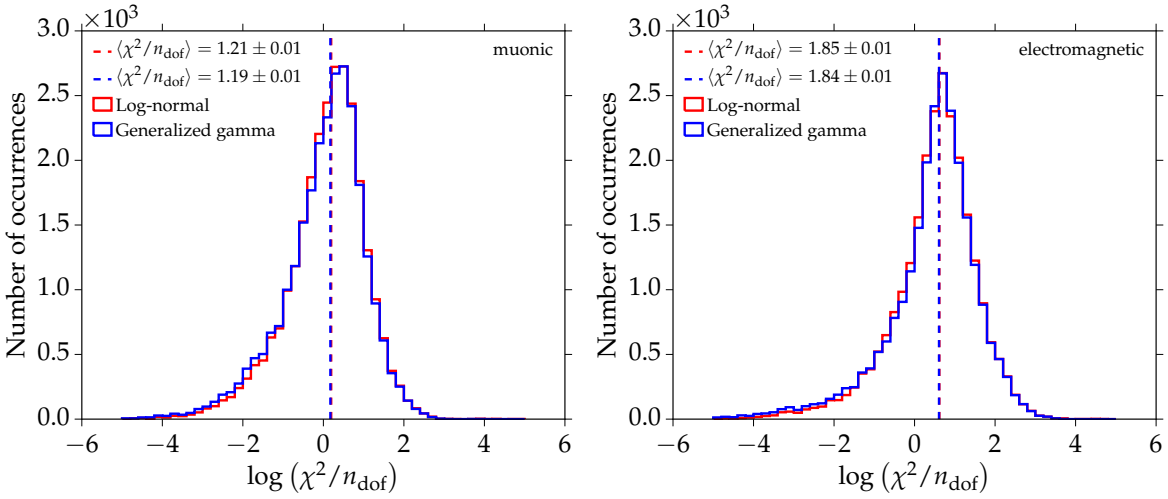


Figure 2.11: Average time distributions of ground signals in air showers initiated by primary particles with an energy of 10^{19} eV and a zenith angle of 0. Included primary species are proton, carbon and iron. Shown are average traces for specific distances to X_{max}. Particle components: (a) muons, (b) pure electromagnetic, (c) electromagnetic from muon decay and (d) electromagnetic from hadron jets.

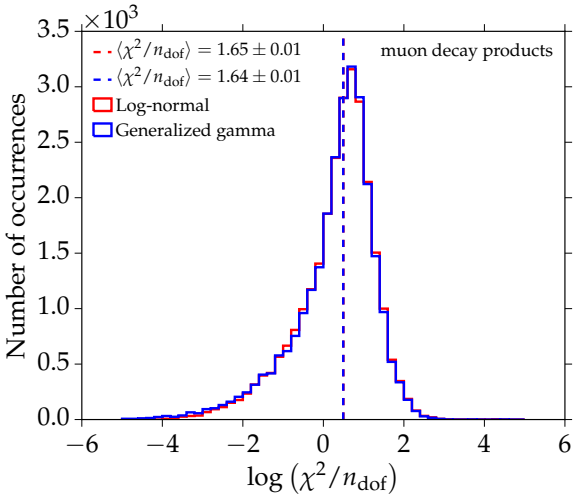
should hold. This is not the case for all but the muonic shower component. Thus, either the model description is not optimal for the other particle components or the fluctuations are underestimated. As shown in Fig. 2.12e for the pure electromagnetic component, the fit quality decreases for large radial distances. Furthermore, the reduced χ^2-values are on average larger for traces with very high signals. In these cases, the fluctuations in the trace are small even at large times, and a deviation of the model will result in very large residuals.

For the following analysis, outliers of more than $6\,\sigma$ with respect to the median of the χ^2/n_{ndof} distribution for a certain particle component are rejected.

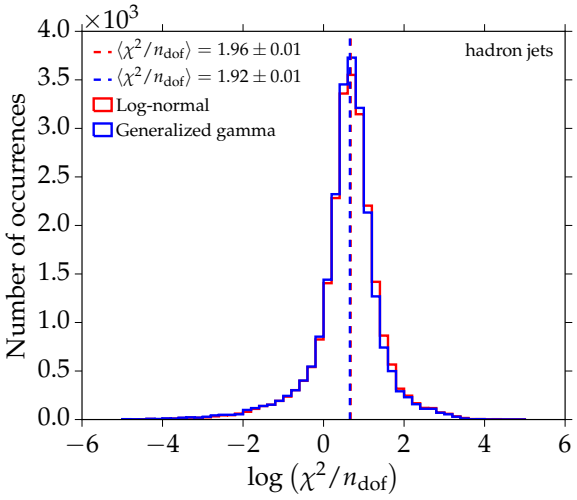


(a) μ

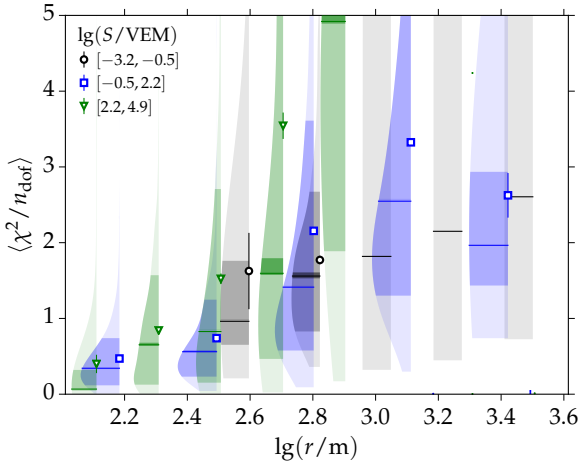
(b) $e\gamma$



(c) $e\gamma(\mu)$



(d) $e\gamma$(had)



(e) Radial dependence

Figure 2.12: (a) - (d): Histograms of the reduced χ^2 distributions for the fits to average traces. The vertical lines represent the median values of the distributions. (e) Radial and signal dependence of the fit quality for the pure electromagnetic component.

2.4.1 Modeling of shape parameters

The next step in the model building is the parametrization of the shape parameters that are obtained from the fits to average traces. For a description with the log-normal distribution, this implies the parametrization of mean m and width s:

$$\begin{aligned} m &= m(r, \Delta X, \psi, \theta, E), \\ s &= s(r, \Delta X, \psi, \theta, E). \end{aligned} \tag{2.17}$$

The general functional form is a mixture of physics motivated and empirical quantities. It reads:

$$\begin{aligned} f_{\mathrm{m,s}}(\Delta X, \psi, \theta, E) &= f_{\Delta X}(\Delta X) + f_{\mathrm{geo}}(\theta, \psi, \Delta X) + f_{\lg E}(\lg E, \Delta X) \quad \text{with} \\ f_{\Delta X}(\Delta X) &= a_{\Delta X} + \Delta X_{\mathrm{ref}}(b_{\Delta X} + \Delta X_{\mathrm{ref}}(c_{\Delta X} + d_{\Delta X}\,\Delta X_{\mathrm{ref}})), \\ f_{\mathrm{geo}}(\theta, \psi, \Delta X) &= \sin\theta(a_{\mathrm{geo}}\,\cos\psi + b_{\mathrm{geo}}\,\Delta X_{\mathrm{ref}}), \\ f_E(\lg E, \Delta X) &= \lg E_{\mathrm{ref}}(a_{\lg E}\, + b_{\lg E}\Delta X_{\mathrm{ref}}), \\ \Delta X_{\mathrm{ref}} &= \Delta X / (750\,\mathrm{g\,cm^{-2}}) \quad \text{and} \\ \lg E_{\mathrm{ref}} &= \lg(E/\mathrm{eV}) - 19. \end{aligned} \tag{2.18}$$

It was interactively found while fitting the model to simulations and observing features. Depending on the particle component and parameter, the number of actually fitted parameters is reduced by fixing higher order contributions to zero. The model holds for a specific core distance range with width Δr. An iterative procedure is used to obtain a full parametrization. Firstly, Eq. (2.18) is fit in each of the available core distance bins, and, secondly, analytical expressions for the fitted parameters as a function of r are found. A global fit was attempted, but did not result in an improved parametrization. Still it could be beneficial to study a global fit in the future to better include minor correlations in the final model.

The dependencies of m and s are described as a function of ΔX, $\lg E$, θ and ψ according to Eq. (2.18). For each radial interval depicted in Fig. 2.10b, a chi-squared fit is constructed, taking into account the values and uncertainties from the fits to the average traces. For the mean parameter m of the log-normal, the following parameters of f_m are fit:

$$\begin{aligned} &\mu: \; a_{\Delta X},\; b_{\Delta X},\; c_{\Delta X},\; a_{\mathrm{geo}}, \\ &e\gamma: \; a_{\Delta X},\; b_{\Delta X},\; c_{\Delta X},\; d_{\Delta X},\; a_{\mathrm{geo}},\; a_{\lg E}, \\ &e\gamma(\mu): \; a_{\Delta X},\; b_{\Delta X},\; c_{\Delta X},\; a_{\mathrm{geo}}, \\ &e\gamma(\mathrm{had}): \; a_{\Delta X},\; b_{\Delta X},\; a_{\mathrm{geo}},\; a_{\lg E}. \end{aligned} \tag{2.19}$$

The parameters that are not listed for a certain particle component are not included in the parameterization. They were excluded from extensive analyses of parameter residuals as shown in Section 2.4.2 (with the help of examining linear correlation coefficients) and from physics arguments. For example, muons propagate almost directly from their production point at the shower axis to their point of detection in the WCD. The overall muonic signal will change with primary energy, but the average delay with respect to the first arrival time of a muon is not expected to vary with primary energy, see e.g. [111]. This is a direct result from the universality of longitudinal shower profiles. Furthermore, in case of muons, the geometrical asymmetry quantified with a_{geo} is expected to be small.

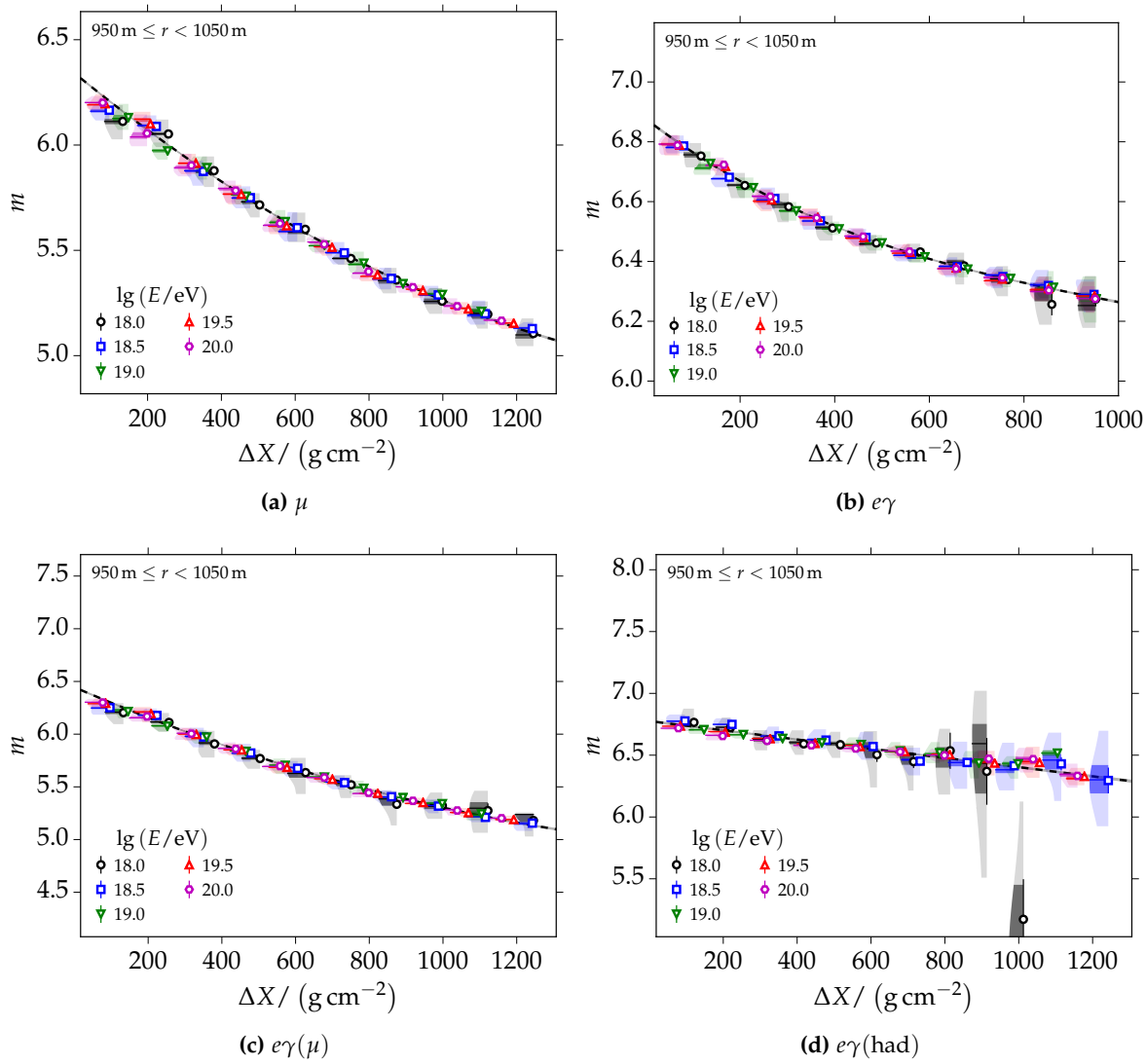
Results on the ΔX dependence of m for the different particle components are shown in Fig. 2.13. Fits in a radial range of $950\,\mathrm{m} \leq r < 1050\,\mathrm{m}$ are included. The profiles are shown for different logarithmic energies. The primary species is averaged over in this

step. The profiles are corrected for geometrical and energy dependencies; only the ΔX dependence is plotted. Independent of the particle component, it is apparent that the bulk of the signal arrives earlier for large distances to the shower maximum (with respect to the first possible arrival time of a particle). The change over the ΔX range is largest for muons and muon decay products. The smallest change is found in the time distributions of the hadron jet component. One explanation for this behavior is that the farther away one detects particles from the shower maximum the more particles come from the same effective origin. Regarding muons, low energy particles will decay earlier, and only high energetic muons will reach a detector far from the shower maximum. These muons are mostly produced close to each other (in time and space) in the early shower development. Thus, the distortion in the arrival times is smallest, which reduces the average delay m (and also the spread s). Hadron jets are always produced close to the detector, so the change is expected to be smallest for this particle component. The points with larger scatter in Fig. 2.13b and Fig. 2.13d are due to small event statistics at lowest energies and far from the shower maximum. A third degree polynomial is used to describe the ΔX dependence of the pure electromagnetic component (see Eq. (2.19)). Due to large deviations above $\Delta X > 1000\,\mathrm{g\,cm^{-2}}$, the ΔX range is limited to smaller values. This restriction leads to a better description of the simulations in the important physics range. The models and $1\,\sigma$ uncertainties are indicated with the dashed lines. The description works well over the whole ΔX range and for all considered energies. Simulations for energies below $10^{18}\,\mathrm{eV}$ are not present in these plots because there are no signals that are large enough to fulfill $\widehat{S} > 4\,\mathrm{VEM}$ at radial distances around 1000 m. These energies are present in the parameterizations at smaller radial distances. Additional plots at different radial distances and for different shower geometries are included in Appendix B.3.1.

For the spread s of the log-normal, the following parameters of f_s are used to describe simulated air showers:

$$\begin{aligned}
&\mu: \; a_{\Delta X},\, b_{\Delta X},\, c_{\Delta X},\, a_{\mathrm{geo}},\, a_{\lg E},\\
&e\gamma: \; a_{\Delta X},\, b_{\Delta X},\, c_{\Delta X},\, a_{\mathrm{geo}},\, b_{\mathrm{geo}},\, a_{\lg E},\\
&e\gamma(\mu): \; a_{\Delta X},\, b_{\Delta X},\, a_{\mathrm{geo}},\, b_{\mathrm{geo}},\, a_{\lg E},\\
&e\gamma(\mathrm{had}): \; a_{\Delta X},\, b_{\Delta X},\, a_{\mathrm{geo}},\, b_{\mathrm{geo}},\, a_{\lg E}.
\end{aligned} \tag{2.20}$$

Results for the radial range $950\,\mathrm{m} \leq r < 1050\,\mathrm{m}$ are shown in Fig. 2.14. Again the fitted parameterizations are indicated by the black dashed lines. For muons and muon decay products, the spread of the time distribution reduces with the distance to X_{max}. It is fairly constant for the pure electromagnetic component and even increases for hadron jets. Unlike in the case of the mean parameter m, this behavior changes for different radial intervals and is thus hard to interpret. For the muons, the spread always reduces with ΔX when $r < 350\,\mathrm{m}$. For $r > 350\,\mathrm{m}$, the opposite behavior is observed. This might be due to the fact that muons with lower energies, which increase the spread in the time distribution, are most likely found close to the shower axis. Though energies dependencies are taken into account according to Eq. (2.18) and Eq. (2.20), the spread for different energies is larger than for the mean parameter. This results from increased fluctuations and difficulties in the parameterization. To illustrate the importance and impact of the zenith angle corrections, a comparison of models with and without corrections are given in Fig. 2.15. For the muonic component, Fig. 2.15a shows the full model including the correction, while the zenith angle correction is switched off in Fig. 2.15b. A much larger spread between fit values for the different zenith angles is visible. The same is true to an even larger extend for the electromagnetic component as shown in the other two plots. The magnitude of the zenith angle correction is given by the parameter a_{geo}. Its size and variation with radial distance is



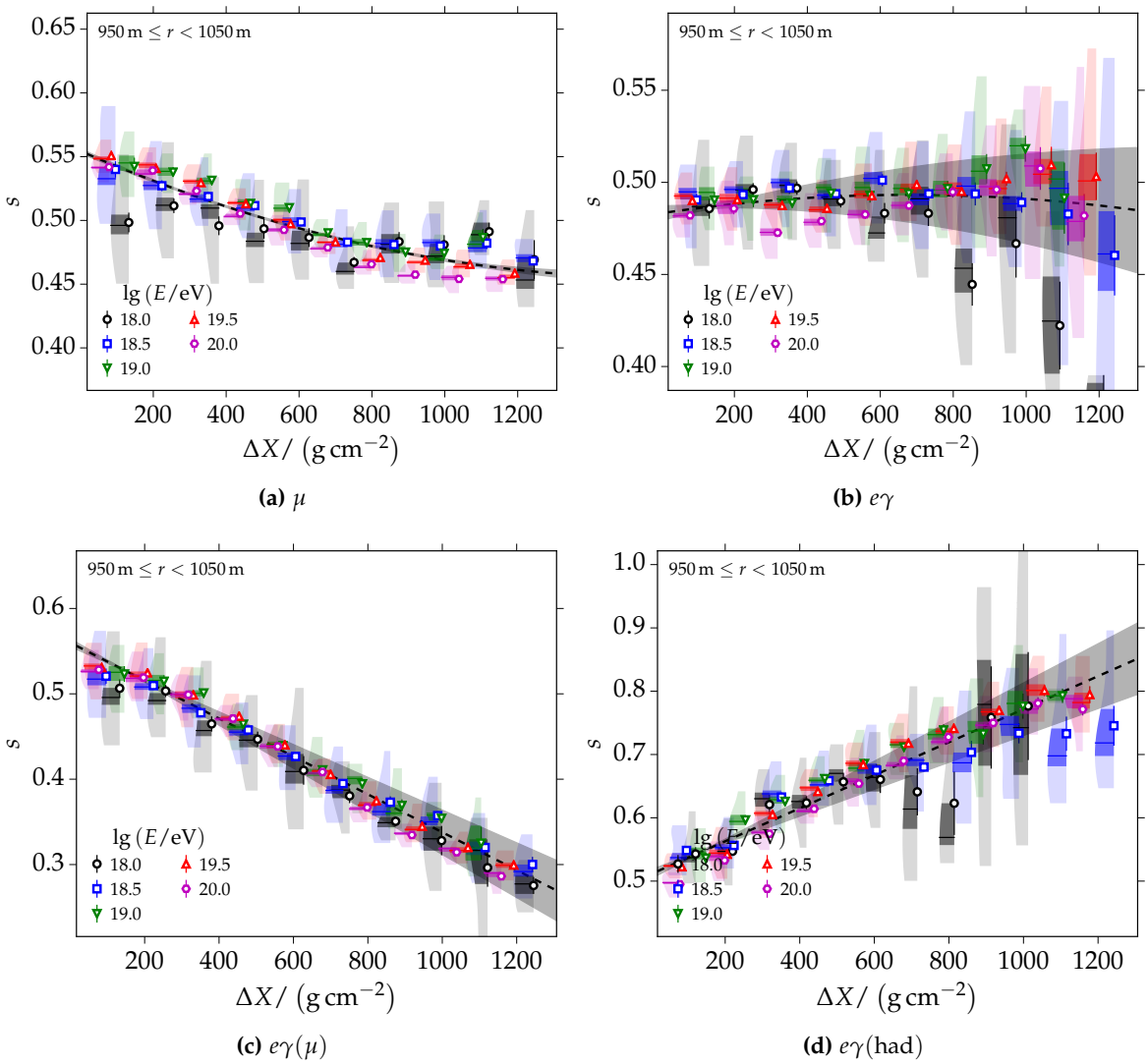
(a) μ **(b)** $e\gamma$ **(c)** $e\gamma(\mu)$ **(d)** $e\gamma$(had)

Figure 2.13: Model description of the ΔX dependence of the mean parameter m for the different signal contributions. Different primary energies are indicated with different colors and markers. Only stations in the radial distance range 950 m $\leq r <$ 1050 m are included.

shown in Fig. 2.16 for the muonic component and in Fig. 2.17 for the pure electromagnetic component. After the modeling of all ΔX dependencies, each of the parameters listed in Eq. (2.19) and Eq. (2.20) is parametrized as a function of core distance r. Due to large fluctuations and potentially remaining trigger biases, only radial distances smaller than 2250 m are considered in this step of the analysis. The following polynomial in r is used for the description:

$$f_r(r) = \sum_{i=0}^{n} a_i \left(\frac{r}{1000\,\mathrm{m}}\right)^i . \tag{2.21}$$

The order of the polynomial n depends on the parameter of question and its observed change with r. Fits for all polynomial orders from 0 to 10 are attempted, and the best choice is selected with a leave-one-out cross-validation method (LOOCV). With this method, the correct model is determined from the statistical fluctuations of the points around the model



(a) μ **(b)** $e\gamma$

(c) $e\gamma(\mu)$ **(d)** $e\gamma(\text{had})$

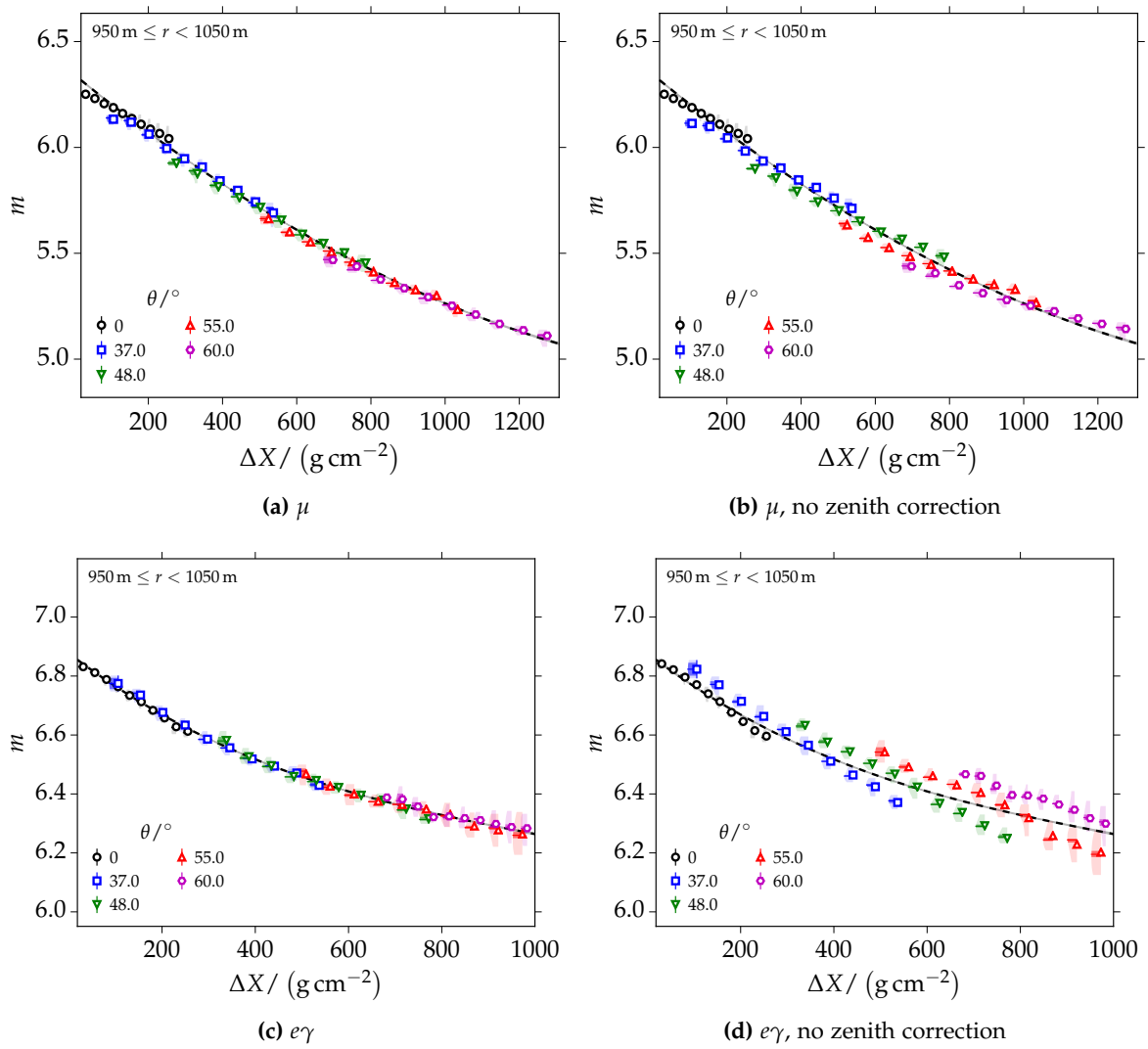
Figure 2.14: Model description of the ΔX dependence of the spread parameter s for the different signal contributions. Different primary energies are indicated with different colors and markers. Only stations in the radial range 950 m $\leq r <$ 1050 m are included.

and the residual bias of the model. Overfitting of the points is thereby avoided. The LOOCV value is calculated for each polynomial with order $k = 0 \ldots n$:

$$\text{LOOCV}(k) = \sum_j \left(y_j - \tilde{f}_j(x_j)\right)^2, \tag{2.22}$$

with the sum running over all j points that are described with the polynomial. The function $\tilde{f}$ results from a chi-squared fit of the polynomial to all points except the point x_j. As such, the value LOOCV represents the sum of quadratic residuals for each point with respect to the polynomial that is fit to all but the respective point. The LOOCV-values are calculated for all $n + 1$ polynomials. The best statistical description of simulated points is then given by the model with minimal LOOCV.

Results on the observed and parametrized radial dependencies of the parameters of the mean values of muon time distributions are shown in Fig. 2.16. The first plot Fig. 2.16a represents the radial dependence of $a_{\Delta X}$, which is the constant parameter in Eq. (2.18). Each of the points in the plot is a result of the fit of Eq. (2.18) in the respective radial range.

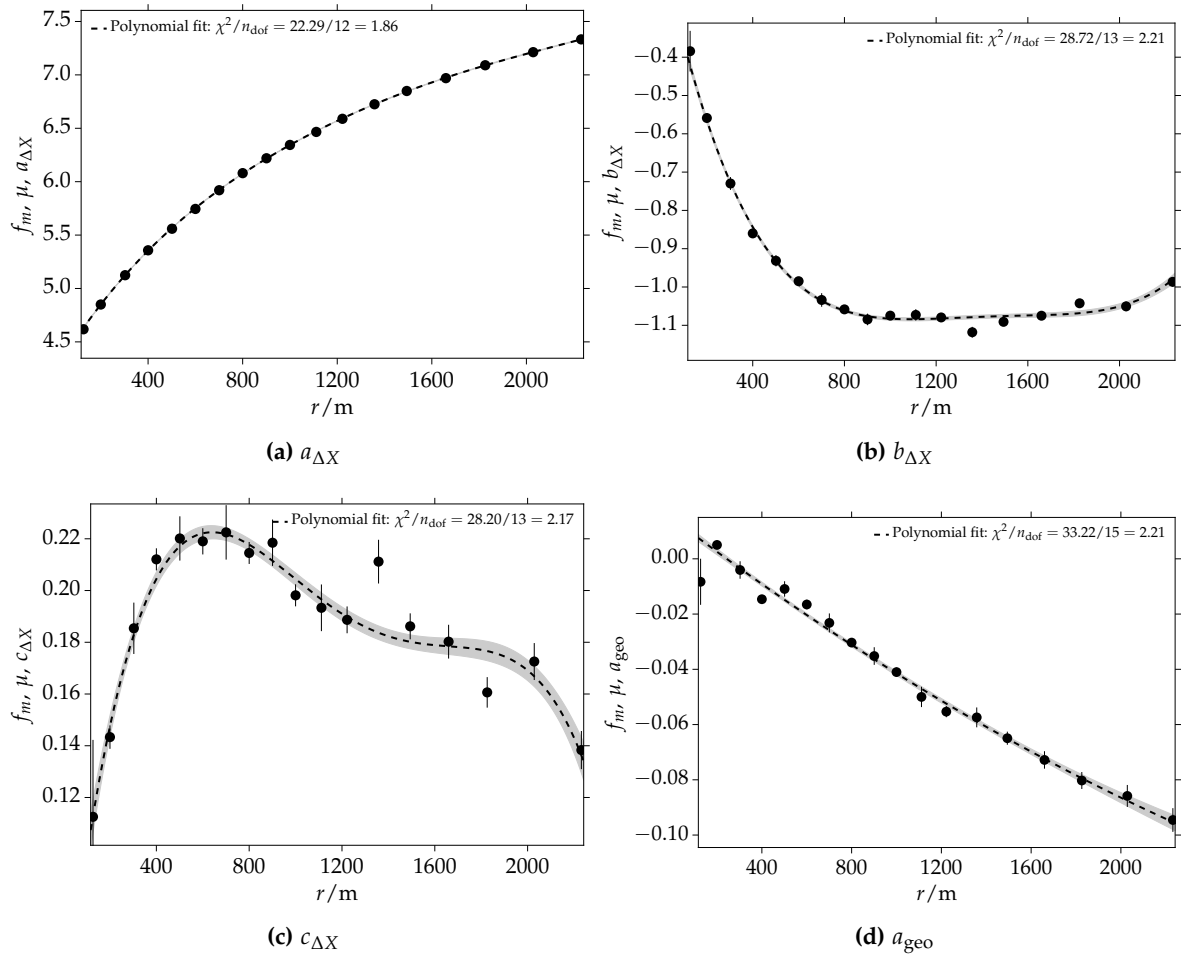


(a) μ **(b)** μ, no zenith correction

(c) $e\gamma$ **(d)** $e\gamma$, no zenith correction

Figure 2.15: Illustration of the zenith angle dependence of the model description of m for the muonic and pure electromagnetic particle components. The plots (a) and (c) visualize the ΔX dependence including zenith angle and energy corrections, while (b) and (d) are respective plots without the corrections.

There is a clear increase of $a_{\Delta X}$ with the distance to the core. This means that the bulk of particles farther from the core is more delayed with respect to the fist particle arrival time. This is understood by taking into account that the average muon energy always decreases with the distance to the core, while the energy of the first muon that dictates the curvature model decreases by a smaller amount [111]. The parametrization describes all points in the radial range 200 m to 2200 m very well. In this case, the best fit is obtained with a sixth order polynomial. The radial dependence of the linear slope parameter $b_{\Delta X}$ is depicted in Fig. 2.16b. As previously shown in Fig. 2.13a, negative values indicate that the mean delay always decreases with increasing ΔX. The rate of this decrease increases up to a radial distance of 800 m and stays relatively constant afterwards. The quadratic slope of the ΔX dependence $c_{\Delta X}$ mostly changes close to the core and is constant elsewhere as shown in Fig. 2.16c. The parameter a_{geo} describes the strength of the geometrical asymmetry $\propto \sin\theta\cos\psi$. The radial dependence of this parameter is visualized in Fig. 2.16d together with the parametrization. Interestingly, the asymmetry vanishes close to the core and shows



(a) $a_{\Delta X}$ **(b)** $b_{\Delta X}$

(c) $c_{\Delta X}$ **(d)** a_{geo}

Figure 2.16: Fits and profiles of the radial dependencies of the mean parameter m for the muonic shower component. Different plots correspond to different parameters of the model describing m. Each of the points is the result of a fit according to Eq. (2.18).

a roughly linear increase with increasing distance to the core. First order asymmetries are already included in ΔX. Therefore, this asymmetry is connected to the average momentum vector of muons entering the detector. Due to the shower geometry, particles entering an early detector ($\psi \approx 0$) are on average slightly more inclined with respect to the ones entering a late detector ($\psi \approx 180°$). The magnitude of this effect increases with zenith angle and r. Thus, particles in the early detector are from a more limited development range and the mean delay should be smaller. This is indeed reflected by the fact that $a_{\mathrm{geo}} < 0$ at all radial distances. I note that the main radial dependence of m is given by the change of $a_{\Delta X}$. Results on the radial dependencies of the model parameters to describe the average time shape of the pure electromagnetic component are depicted in Fig. 2.17. The dependence of $a_{\Delta X}(r)$ is very similar to the one obtained for muons. On average, particles are more delayed farther away from the core. In th case of electromagnetic particles, this is easily explained when taking into account that the particles farther from the core experience more scattering as they traverse a larger grammage before reaching the detector. The dependencies of $b_{\Delta X}$ and $c_{\Delta X}$ are similar to the ones for muons. Interestingly, the geometrical asymmetries in Fig. 2.17d are very different. While they also vanish very close the the core, there is a rather linear increase for radial distances $r < 1200\,\mathrm{m}$ and a sudden drop at even larger distances. The second asymmetry parameter b_{geo} describes the magnitude of a change $\propto \Delta X \sin\theta$. This is correlated to $b_{\Delta X}$ and $c_{\Delta X}$, which might explain the rather bad fit quality of $b_{\Delta X}$ as

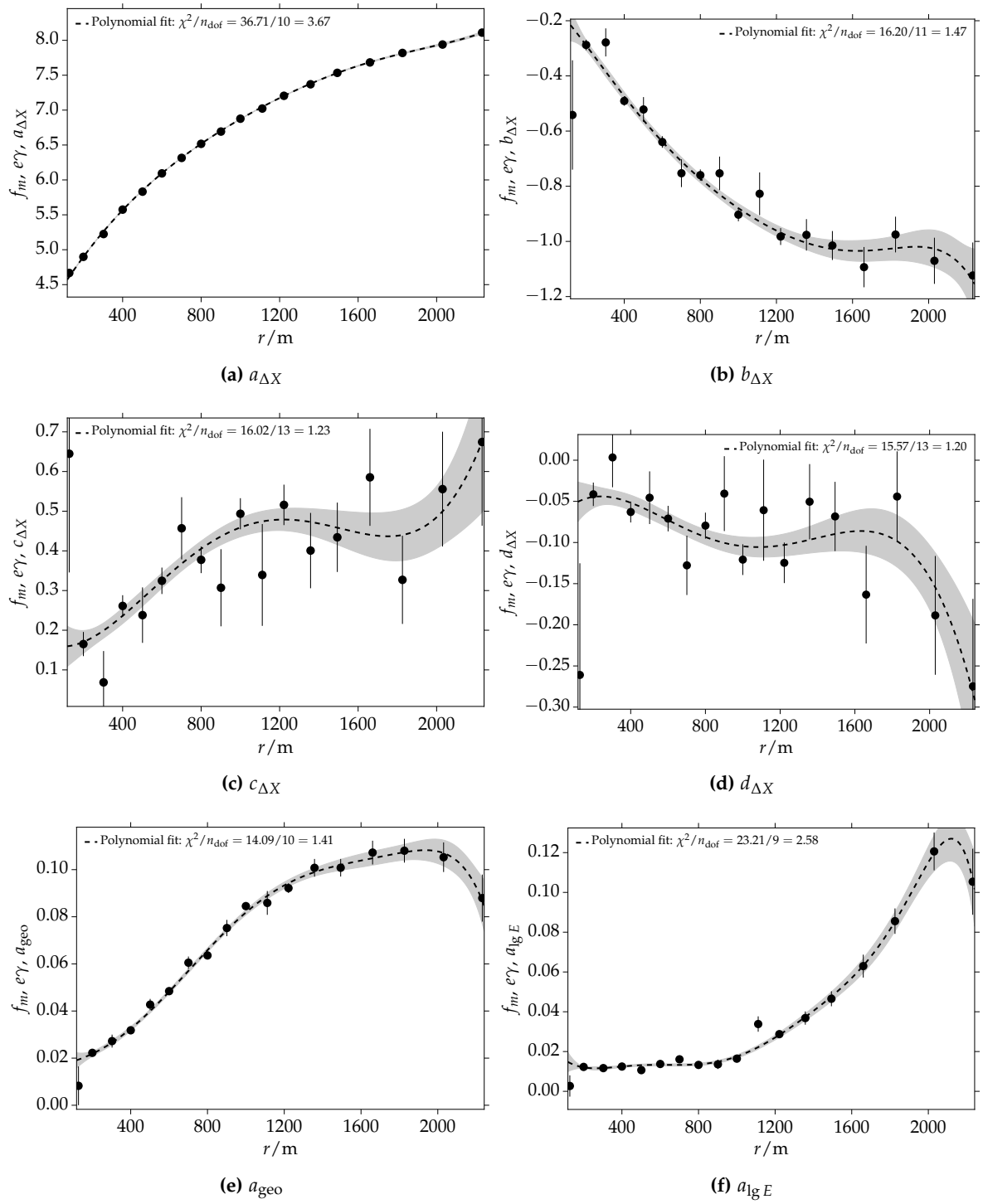
shown in the plot in Fig. 2.17b. Due to correlations, there are compensating fluctuations in these plots. This is a consequence of the staged fit procedure and can only be avoided with a global fit of all dependencies at the same time. The radial dependence of $a_{\lg E}(r)$ is parametrized in Fig. 2.17f. The energy dependence is small and constant for $r < 800\,\mathrm{m}$ and increases at larger distances to the core. The dependencies of the mean parameter of the other two particle components show very similar structures. They are included in Appendix B.3.2 together with models for the spread parameter and tables with numerical values of all the model parameters.

2.4.2 Parameter model residuals

Combining the parameterizations from Eq. (2.18) and Eq. (2.21) leads to a complete descriptions of the mean $m(r, \Delta X, \psi, \theta, \lg E)$ and width parameters $s(r, \Delta X, \psi, \theta, \lg E)$ of the log-normal description of time distributions. To analyze the accuracy of the parameterization, each individual m and s (from the fits to average traces) is compared to the model prediction $\widehat{m}$ and $\widehat{s}$. The relative residuals $\frac{m-\widehat{m}}{\widehat{m}}$ and $\frac{s-\widehat{s}}{\widehat{s}}$ are studied as a function of all relevant quantities and physics parameters. For all following plots throughout this thesis, the comments in Appendix A.2 apply.

Selected residuals for the mean parameter of muonic time traces are shown in Fig. 2.18. $\pm 1\,\%$ lines are depicted in each plot to help interpret the accuracy of the model. The most important information about the mean residual is shown with the different markers, while the violins in the background give additional and important information about the distributions within each bin (see Appendix A.3). More than $3\,\sigma$ outliers within each bin are plotted as small dots above and below the distributions. The residuals of the mean parameter are shown as a function of ΔX for different radial distances in Fig. 2.18a. In this representation, the model gives a perfect description up to highest radial distances and largest ΔX. Biases are below $0.5\,\%$. The uniformity of the residuals as a function of azimuth angle are depicted in Fig. 2.18d. There is no remaining dependence on zenith angle or azimuth angle. Both stations in the early and late part of the shower are described equally well. The residuals are shown as a function of radial distance and for different primary energies in Fig. 2.18c. The model gives a good description of time distributions for all primary energies. Note that an explicit energy dependence was not taken into account in the parametrization. It is interesting to note that the largest bias is apparent for primaries with energies $10^{17.5}\,\mathrm{eV}$. An identical behavior is visible in the signal model in Fig. 2.7. This suggests that there are non-trivial changes and dependencies of signals and time distributions at this primary energy. However, for the time model discussed here, this bias is still below $1\,\%$ and not significant. An important residual is shown in Fig. 2.18b. Relative differences are shown as a function of ΔX and for different zenith angles. The distributions are rather flat and within $1\,\%$ for the individual zenith angles. Some non-trivial structure is apparent at lowest zenith angles and ΔX. This reflects some minor difficulties in the description of time shapes for very small distances to the shower maximum and is left as a future task. It does not pose a problem for the desired accuracy of the presented model.

Relative residuals for the mean parameter of the pure electromagnetic component are included in Fig. 2.19. The dependence on ΔX is shown for different radial distances in Fig. 2.19a. There is a very good description for radial distances smaller than 2000 m. At larger radial distances, the description of the model shows deviations above $1\,\%$, but the $1\,\sigma$ regions are almost as large. The dependence on azimuth and zenith angle is depicted in Fig. 2.19d. A small residual modulation with azimuth angle on the order of $0.5\,\%$ is visible. This modulation might be reduced by taking into account higher order geometrical corrections or by building the model with a library with more zenith angles in order to have a



(a) $a_{\Delta X}$ **(b)** $b_{\Delta X}$

(c) $c_{\Delta X}$ **(d)** $d_{\Delta X}$

(e) a_{geo} **(f)** $a_{\lg E}$

Figure 2.17: Fits and profiles of the radial dependencies of the mean parameter m for the pure electromagnetic shower component. Different plots correspond to different parameters of the model describing m. Each of the points is the result of a fit according to Eq. (2.18).

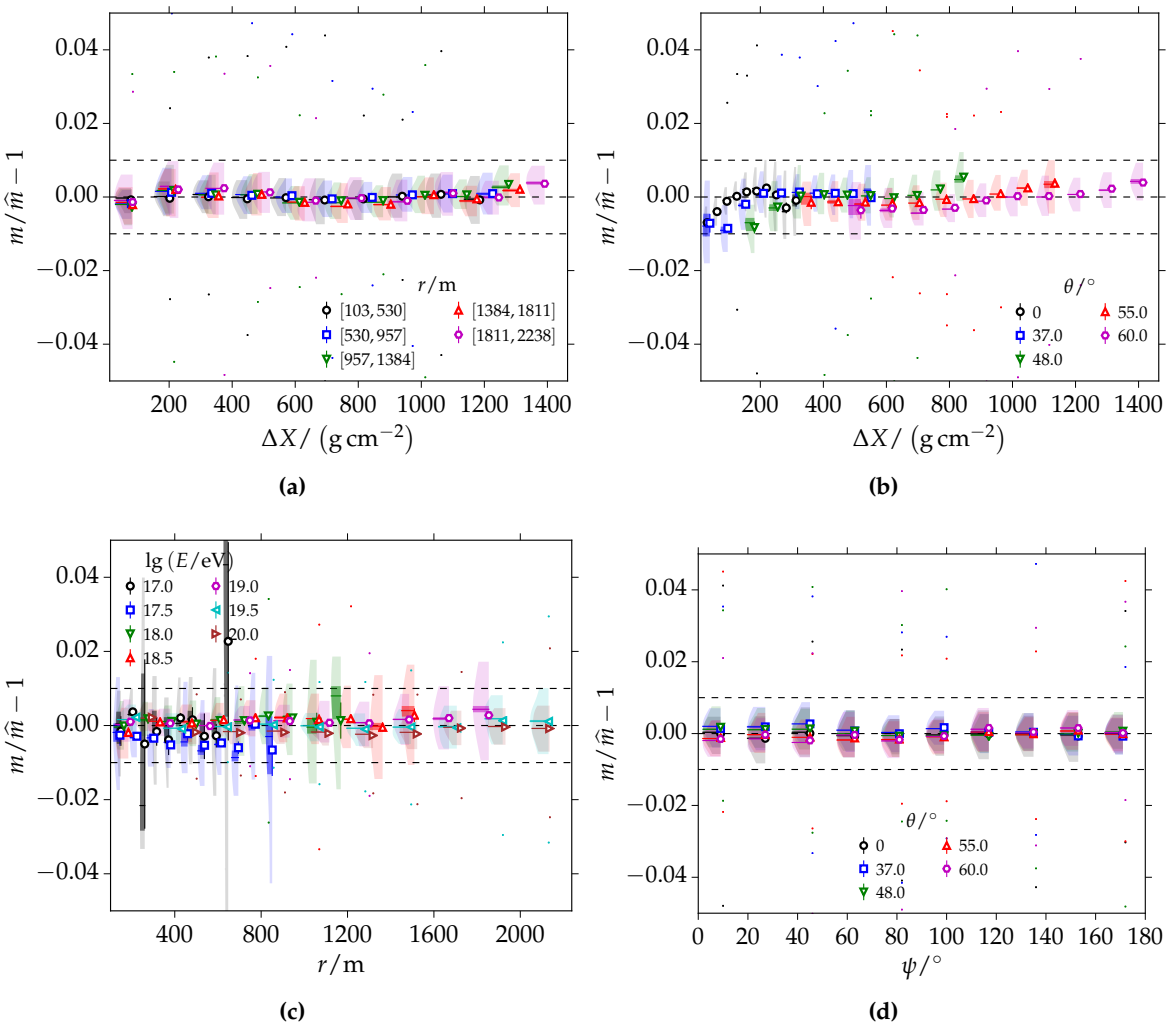


Figure 2.18: Relative residuals of the model description of the mean parameter m of muonic traces. Residuals are given as a function of different quantities of interest.

better handle on these dependencies. The radial dependence for different primary energies is shown in Fig. 2.18c. In contrast to the model for muons, a linear energy dependence is fitted. The result demonstrates that there is no remaining bias for any of the primary energies, even for low energies. The residuals for different zenith angles are unbiased within $\pm 2\,\%$ as a function of ΔX in Fig. 2.19b.

Residuals for the other particle components and residuals for the width parameter s are given in Appendix B.3.4. The size of the relative residuals of m are typically below 1 % for all particle components and dependencies. The model of s is much harder to control. Deviations of up to 5 % occur for muons and the electromagnetic component. For the other two particle components, most of the deviations are within 5 %, but larger deviations of up to 15 % are apparent in a few residuals.

2.4.3 Model predictions and time quantiles

After the construction and parameterization of the time model from fits to average traces, each individual simulated trace can be compared to the prediction from the time model. This is an important check in order to see features and deviations in individual measured

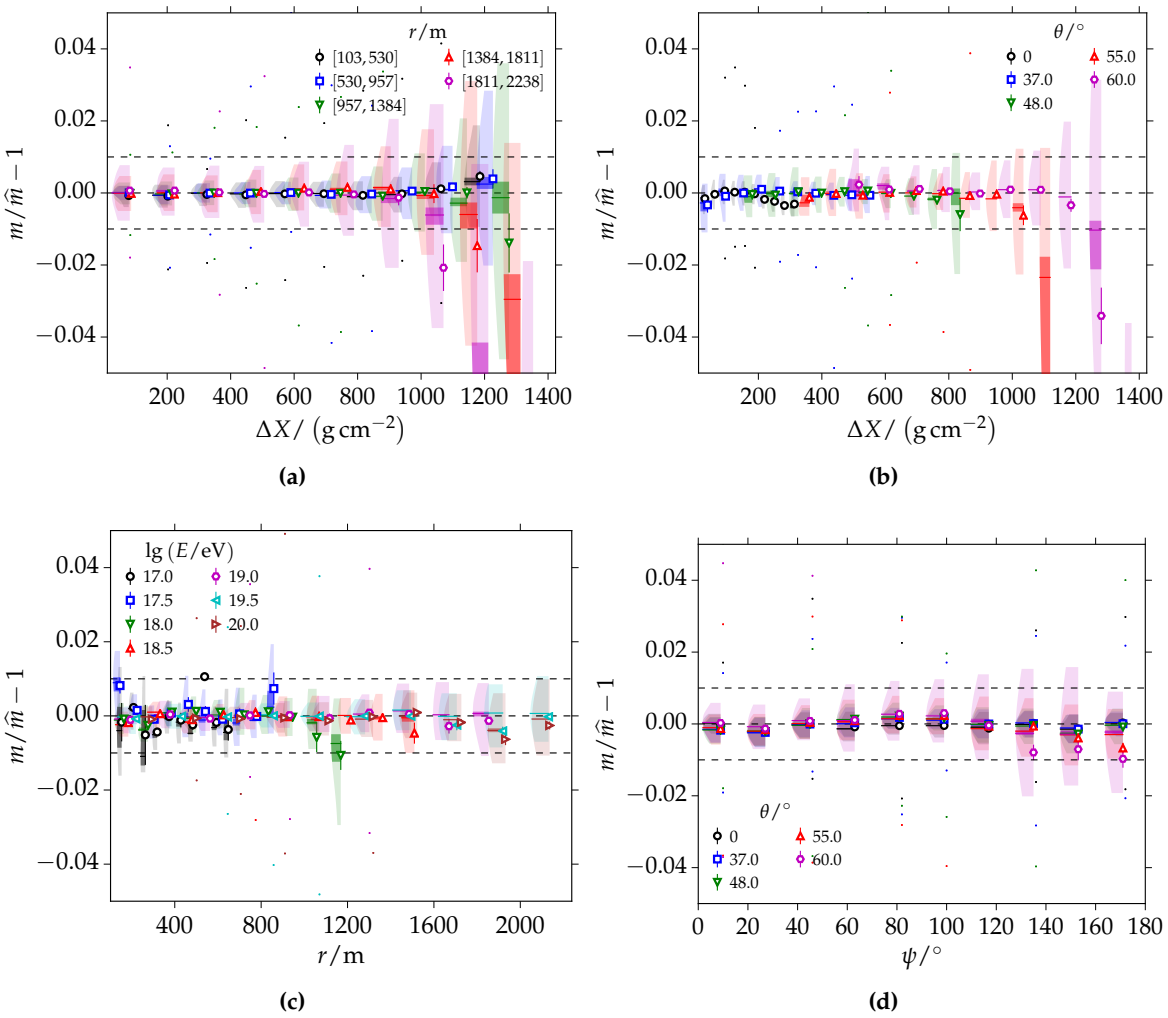
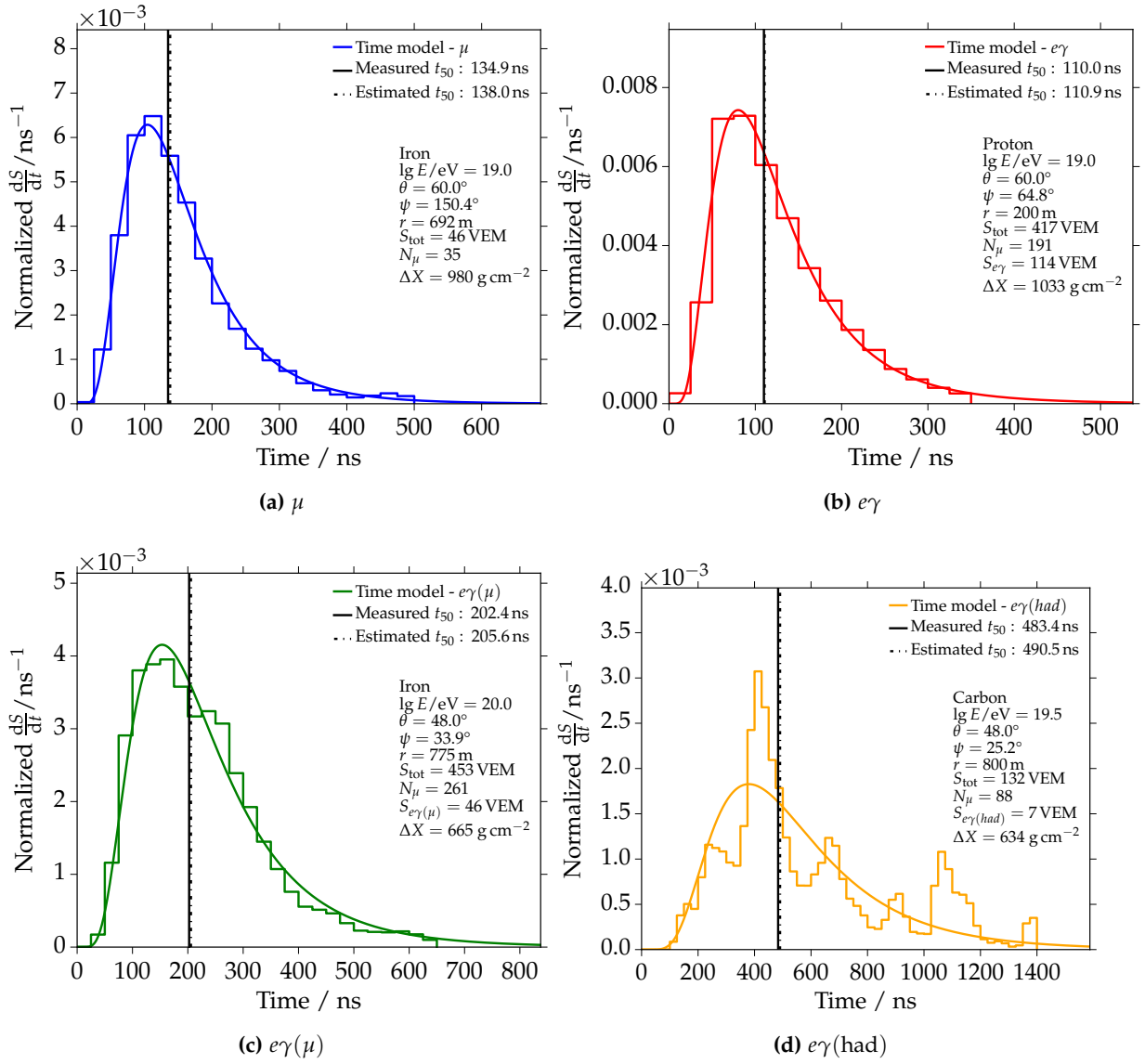


Figure 2.19: Relative residuals of the model description of the mean parameter m of pure electromagnetic particle traces. Residuals are given as a function of different quantities of interest.

and expected distributions. Traces for the four particle components and for different primary energies and species are shown in Fig. 2.20. For muons, the example trace is from an iron shower at an energy of 10^{19} eV and a zenith angle of 60°. The WCD was located 692 m from the core and in the late part of the shower. The true simulated number of muons N_μ is 35. The model gives an accurate description of the simulated trace. The difference between the measured and expected 50 % quantile is very small with $\Delta t_{50} = 3.1$ ns. An example for a proton shower and a detector very close to the core ($r = 200$ m) is shown for the pure electromagnetic component in Fig. 2.20b. Predictions for larger core distances and smaller signals are given for the other two components in Fig. 2.20c and Fig. 2.20d. Large fluctuations are visible in the trace from the hadron jet component. The component signal is very small with only 7 VEM. The traces for these examples have been randomly selected from a subsample of traces with expected signal larger than 10 VEM and primary energies starting at 10^{19} eV.

To be able to better quantify differences in individual traces, it makes sense to look at the mean values and distributions of residual differences between the simulated and expected time quantiles. I will focus especially on the 10 % quantile t_{10} and the 50 % quantile t_{50}. The description of the late time quantile t_{90} was studied as well. However, there are



(a) μ **(b)** $e\gamma$ **(c)** $e\gamma(\mu)$ **(d)** $e\gamma$(had)

Figure 2.20: Individual simulated traces for different particle components are compared to the model predictions from the component-dependent time model. Different primaries, energies and shower parameters are used for each of the plots. The information is given within each plot. The solid vertical lines indicate the measured t_{50}, while the dashed lines represent the model predictions $\hat{t}_{50}$.

large deviations in the description of the late traces that could not be resolved with the current log-normal ansatz. Only the description of the late trace of the pure electromagnetic particles is unbiased on average. In case of the late time quantile for muons, the model only works well for radial distances $r < 500\,\mathrm{m}$. Additional plots are included in Appendix B.3.4.

A comparison of quantiles for the muonic shower component is shown in Fig. 2.21. The t_{50} quantile is shown as a function of ΔX and for different radial distances in Fig. 2.21a. Measured (simulated) values are represented with markers, while the model predictions are shown with dashed lines. To better guide the eye, continuous model lines are drawn. However, this is merely a connection of the predicated points and not a direct parameterization. A nice agreement between simulation and prediction is apparent. The average t_{50} decreases with increasing distance to the shower maximum (compare with Fig. 2.13) and increases with radial distance (compare with Fig. 2.16 and explanations in the text).

The dependence of t_{50} on core distance is shown in Fig. 2.21b. Different markers and colors represent different zenith angles. Open markers include simulations at an energy of 10^{19} eV, while filled markers refer to energies of 10^{20} eV. Again, a clear increase with radial distance is apparent. In addition, t_{50} decreases with increasing zenith angle, mostly as a consequence of the dependence on ΔX as depicted in Fig. 2.21a. There is a universal behavior for different primary energies. The small difference for the different energies results from the fact that the average ΔX at larger energies is smaller and t_{50} is thus slightly larger. Absolute residuals as a function of ΔX and for different energies are shown for t_{10} and t_{50} in Fig. 2.21c and Fig. 2.21d, respectively. The transparent regions indicate the $1\,\sigma$ spread of simulations within each bin in ΔX. No significant deviations from zero are visible. Residuals for different component signals and ranges in ΔX are shown in Fig. 2.21e and Fig. 2.21f. The fluctuations increase strongly with decreasing signal. Again, there are no apparent biases. Due to memory limitations, a random selection of 2×10^5 simulated traces was used for each plot in this section.

Equivalent figures for the time quantiles from the pure electromagnetic shower component are shown in Fig. 2.22. There is a distinct energy dependence in Fig. 2.22b, which is correctly taken into account by the time model. On average, the residuals show no large biases. However, it is visible in Fig. 2.22f that there are larger deviations in the description of traces at large distances to the core. It is visible in Fig. 2.22f that the accuracy of the description improves with increasing component signal. A cut on the electromagnetic component signal of 10 VEM was applied for these plots. There are problems in the description of smaller signal traces and larger biases are visible. This is also true for all other particle components. Fluctuations in time quantiles at lower signal sizes are not symmetric. It is more likely that t_{10} or any other time quantile fluctuates to larger times with respect to the average model than to smaller times. In addition to model inaccuracies, this is one relevant factor that explains deviations in some of the residuals.

Quantiles and residuals of the muon decay products are shown in Fig. 2.23, while the respective plots for the hadron jet component are illustrated in Fig. 2.24. On average, both parameterizations work well, but due to the smaller average signal sizes, fluctuations and deviations are larger, especially at large radial distances. Again, the figures Fig. 2.23e, Fig. 2.23f, Fig. 2.24e and Fig. 2.24f illustrate that the description works perfectly at large component signals and fluctuations to later times cause some of the deviations at smaller signals.

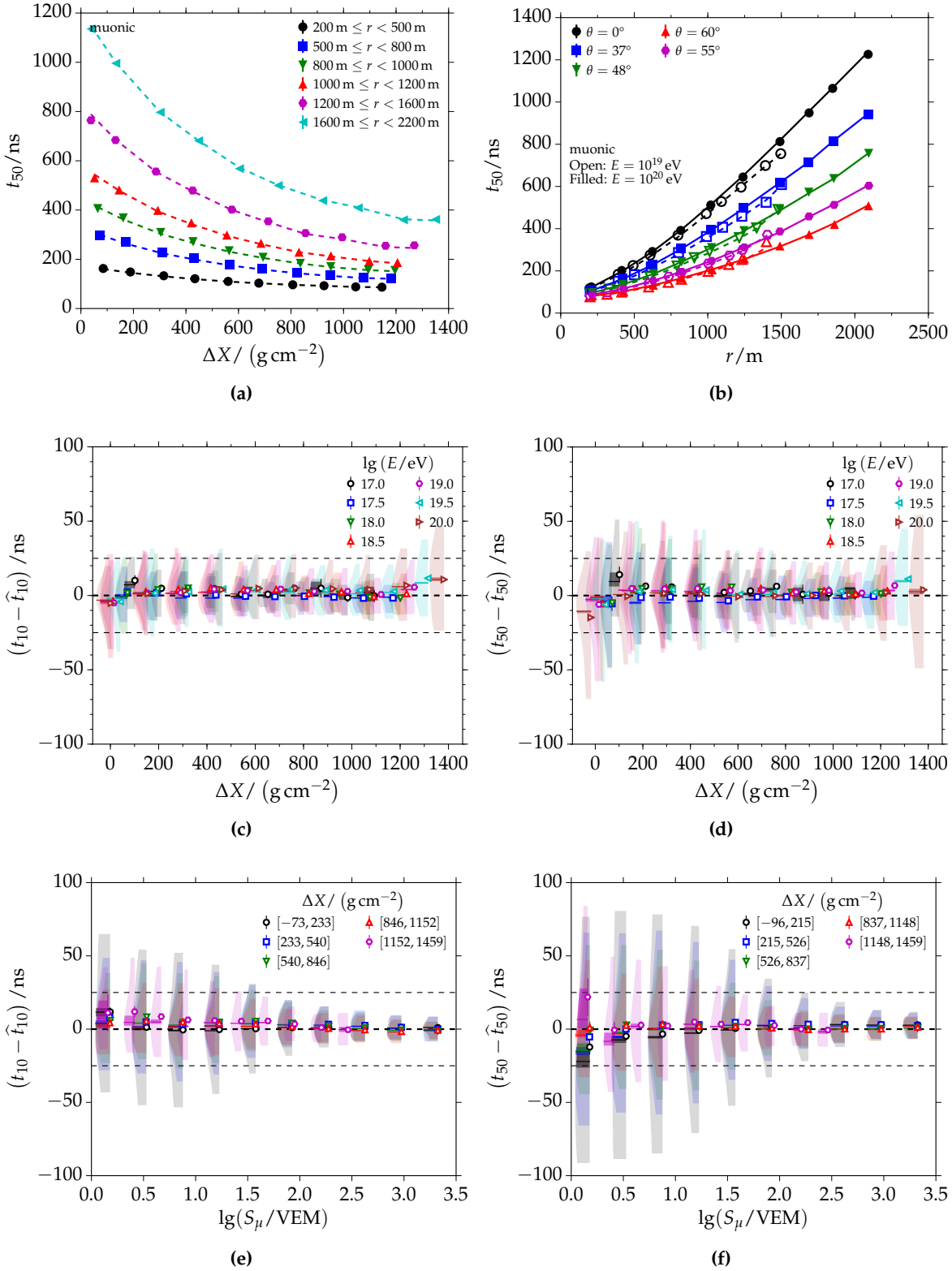


Figure 2.21: Comparison of simulated and estimated time quantiles for the muonic signal component.

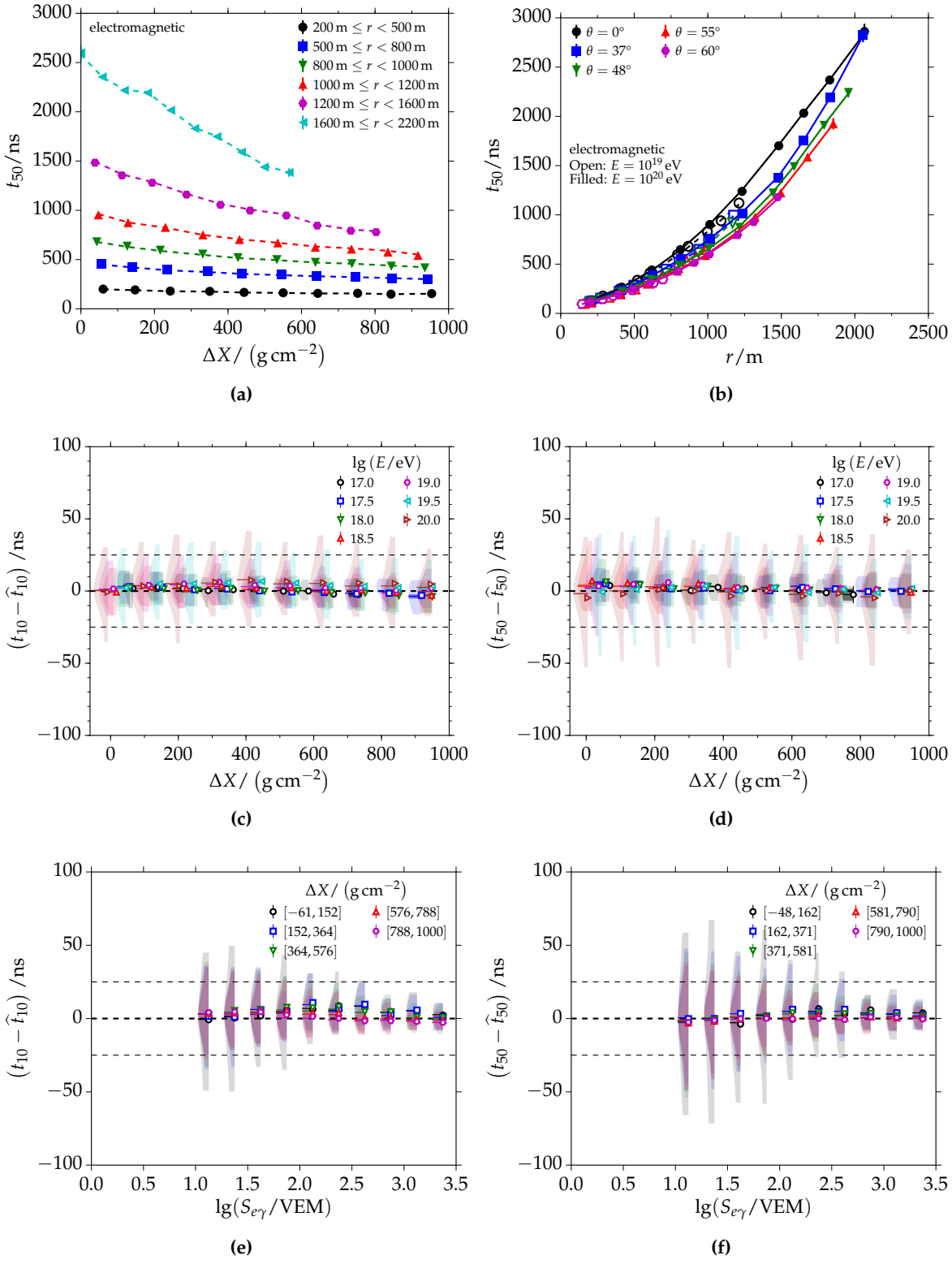


Figure 2.22: Comparison of simulated and estimated time quantiles for the pure electromagnetic signal component.

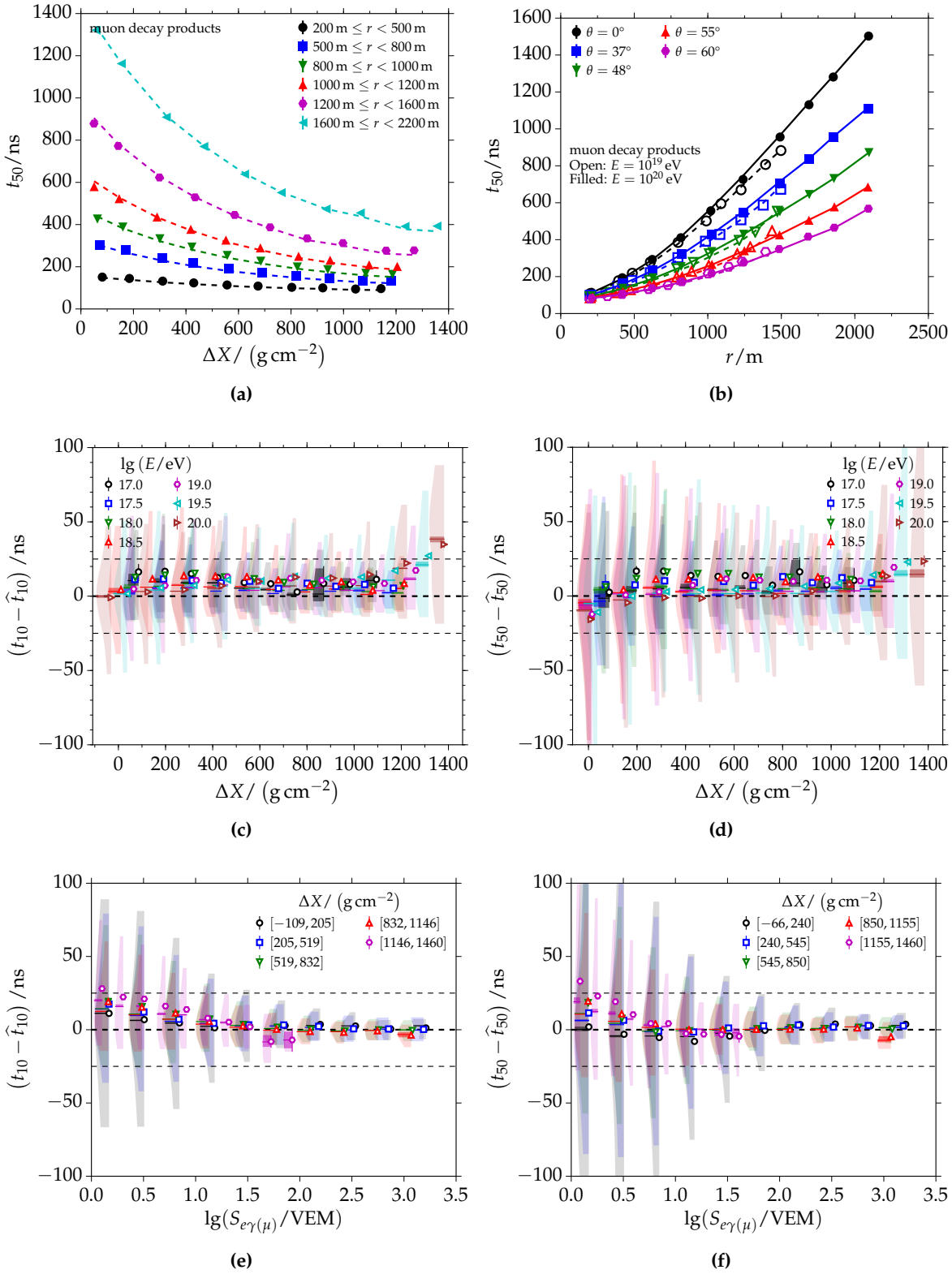


Figure 2.23: Comparison of simulated and estimated time quantiles for the muon decay signal component.

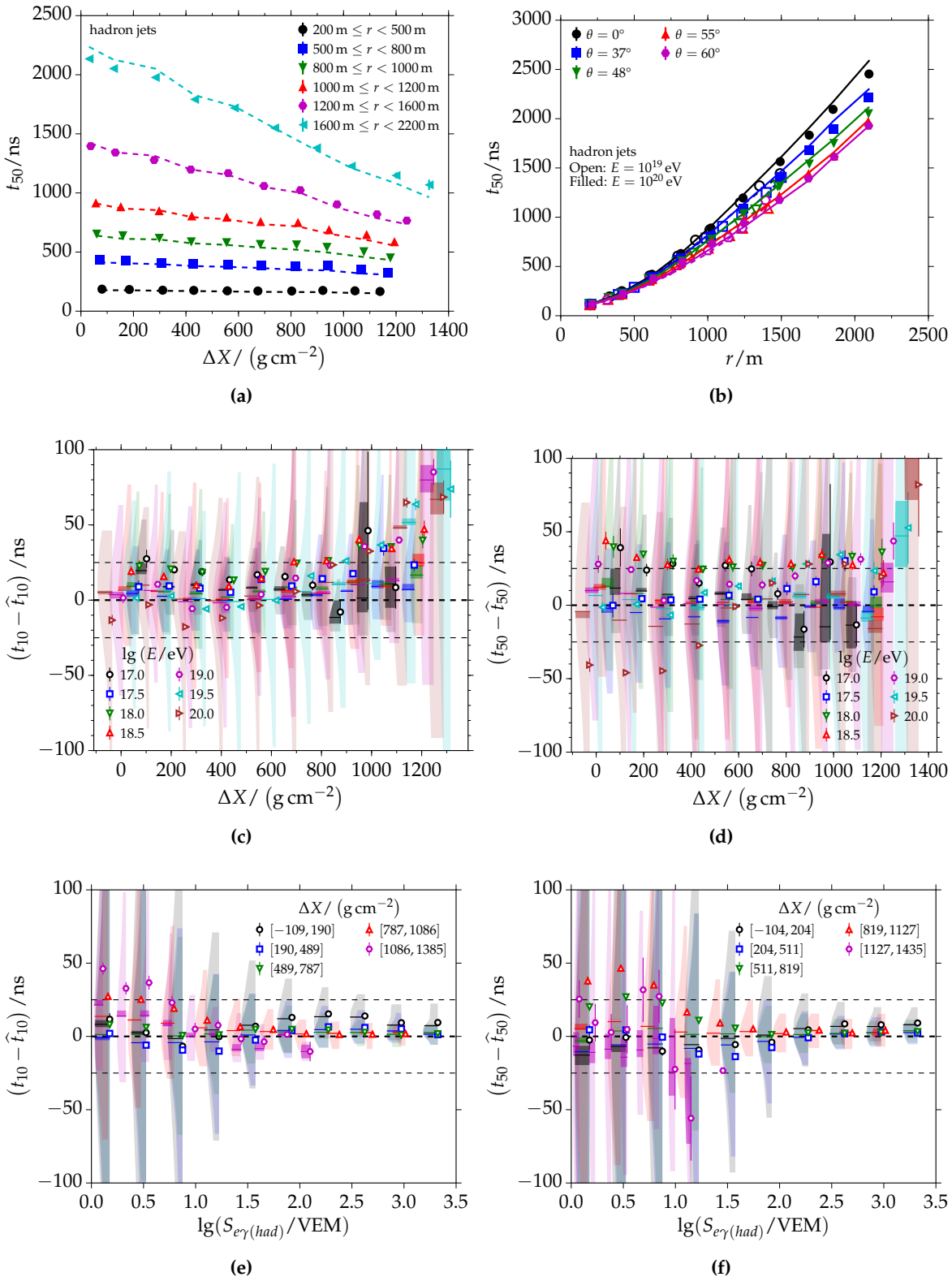


Figure 2.24: Comparison of simulated and estimated time quantiles for the hadron jet signal component.

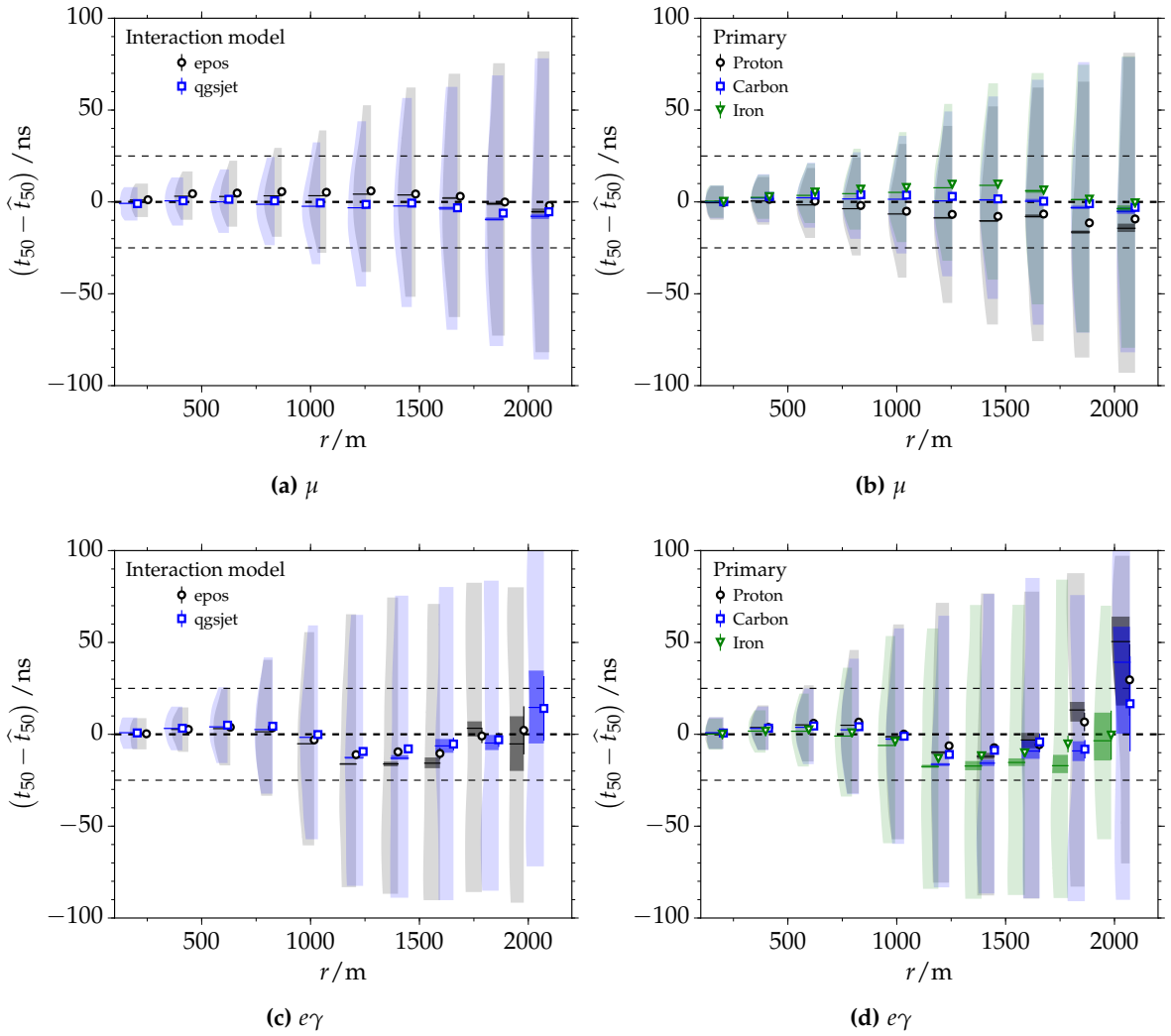
2.4.4 Systematics

The time model was parameterized with average traces. This includes an average over the simulated primary species proton, carbon and iron and the hadronic interaction models QGSJET-II.03 and EPOS-1.99. In addition, different monthly atmospheres were used in the simulation of air showers. In the following, the systematic uncertainties from these factors based on their impact on selected time quantiles is discussed.

The difference in the estimation of t_{50} for different primary particles and interaction models is shown in Fig. 2.25. In all cases, differences in t_{10} are of a similar structure but smaller magnitude. The dependence of t_{50} of the muonic component for different interaction models is shown in Fig. 2.25a. For all core distances, deviations are very small on the order of 5 ns. This is negligible within the accuracy of the model. Differences between primary particles as shown in Fig. 2.25b are however relevant and on the order of 10 ns at a radial distance of 1000 m and even larger at larger distances. These differences are caused by small differences in the time structure of traces for the different primary particles and thus small differences in model parameters m and s. These depend on radial distance. It is possible to compensate for these biases by introducing offset parameters to m that depend on primary type and distance to the core (and maybe other parameters). As the models are used in the reconstruction of measured air showers and the composition is not reliably known, I pursue a different approach for this thesis. The dependence of shape parameters and quantiles on primary species and interaction model will effectively lead to small reconstruction differences in X_{max}. These are quantified and treated as systematic uncertainties in Section 3.4.5. Dependencies of t_{50} of the pure electromagnetic particle traces on primary type and interaction model are shown in Fig. 2.25c and Fig. 2.25d. The systematic impact of both is negligible in this case. The bias at larger radial distances is mostly due to problems in the description of traces with small signals and at large radial distances. The impact of this inaccuracy on the air shower reconstruction is small because the fraction of the electromagnetic signal at large radial distances is negligible with respect to the muonic signal. The dependence of the estimation of t_{50} on the monthly atmosphere is shown in Fig. 2.26. There is a small modulation on the order of 3 ns for both the muonic and pure electromagnetic components, which is negligible compared to both the model accuracy and the other systematics mentioned before. Additional plots with systematics for the other particle components are included in Appendix B.3.4. I want to note that the dependence of t_{50} on the monthly model is a higher order effect. The difference in ΔX due to differences in the densities of atmospheric profiles is taken into account for the calculation of ΔX values. Thus, the parameterized model includes these dependencies.

2.4.5 Correlation to the shower maximum

Instead of only using a correlation to the electromagnetic shower maximum X_{max}, it is interesting to study signal and time parameterizations as a function of the muonic shower maximum X^{μ}_{max}. For example, the muon and muon decay components are expected to correlate more strongly with X^{μ}_{max}, and, taking this into account, the precision of the signal and time parameterizations can potentially be improved. Previous attempts to extend models in that respect are documented in [100, 103]. The main drawback that lead to the decision not to use X^{μ}_{max} in the current model is that an event-by-event estimate of the muonic shower maximum is not available in the current simulation libraries. It is possible to use models for X^{μ}_{max} as a function of primary energy and species, but model dependencies and inaccuracies need to be studied in further detail. These will directly propagate into resulting models and need to be thoroughly quantified.



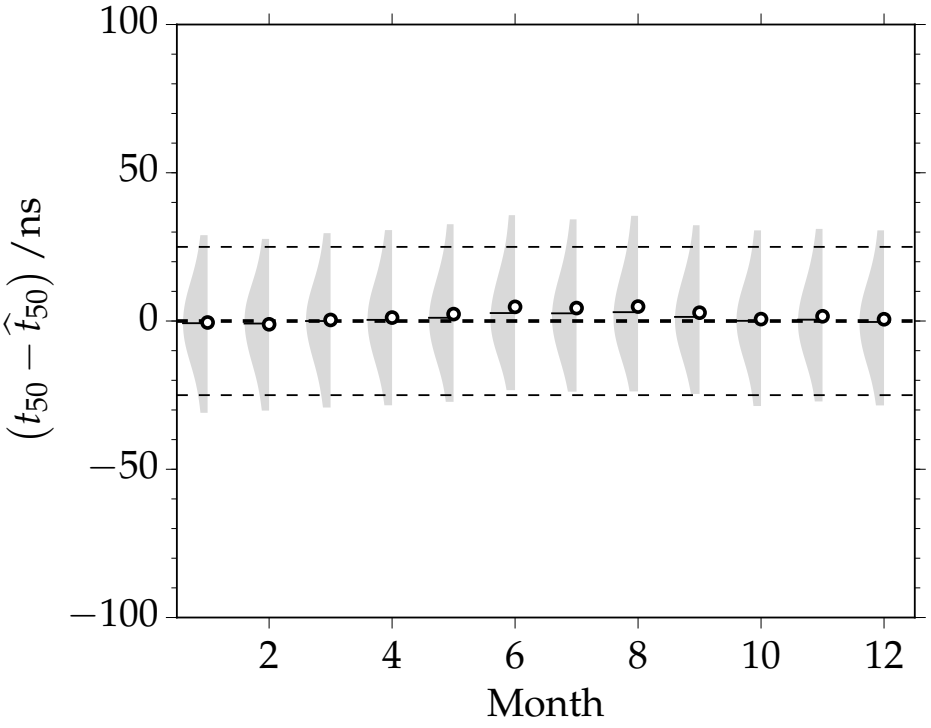
(a) μ **(b)** μ

(c) $e\gamma$ **(d)** $e\gamma$

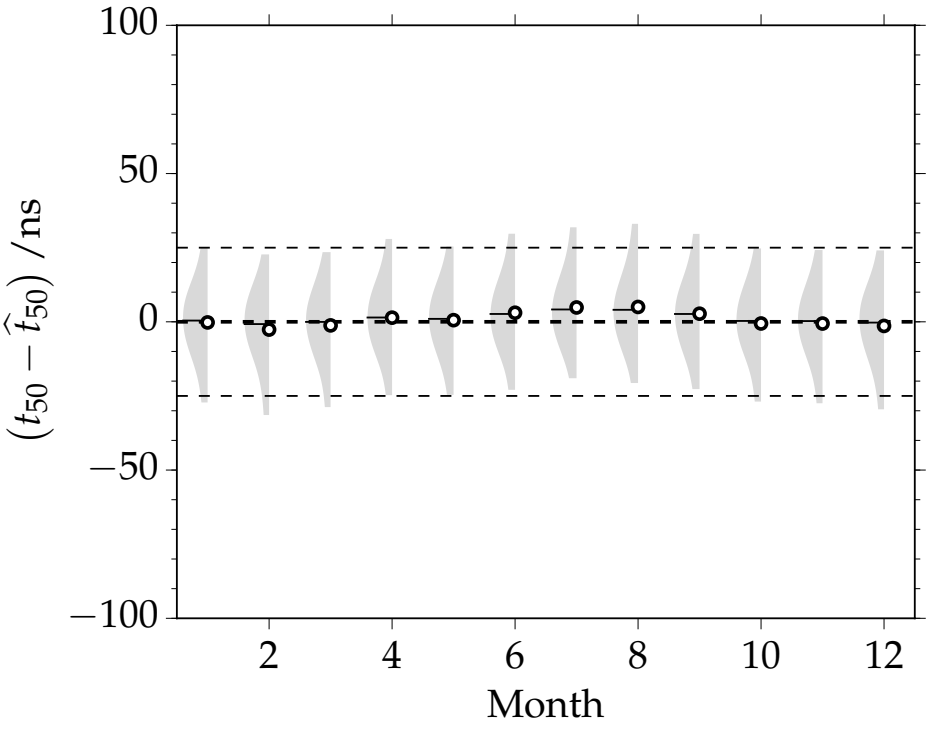
Figure 2.25: Systematic influence of different primary particles and interaction models on the estimation of time quantiles. The model expectations are subtracted from the simulated values; the difference is shown as a function of distance to the core.

2.4.6 Sources of time biases

The impact of the size of the resampling regions (used together with thinned air shower simulations) on biases in the average arrival time of muons was recently studied in [112]. The study demonstrates that there are non-negligible biases when the standard radial extend of 10 % around a station position is used for the resampling. The reduction of the size of the region to only 5 % results in a significant reduction of biases. For the simulations used in this work, the size was reduced to 5 % in radial dimension. Time biases occur in the resampling because of two reasons [113]. At first, the resampling regions are larger than the actual area of the detector. Due to the steep power-law LDF, the density of particles is larger at smaller radial distances. As such, too many particles that arrive too early are injected in a detector. At second, the actual arrival times of particles with a certain distance to the station are corrected for the time delay of a plane front propagating this distance according to the shower geometry. The problem is that the shower front is curved and the curvature even depends on particle type. These effects are currently not taken into account.



(a) Muons



(b) Pure electromagnetic component

Figure 2.26: Variation in the t_{50} time quantiles for the muonic and pure electromagnetic particle components as a function of month. For each month, different atmospheric models were used in the simulation of air showers.

CHAPTER 3

Reconstruction of air showers from UHECRs

The *Pierre Auger Observatory* (Auger) is described in detail in Section 1.2. The observatory utilizes a hybrid technique to detect air showers with both surface and fluorescence detectors (SD, FD). The hybrid approach enables a mostly simulation-independent way of calibrating signals measured with the SD to energies measured with the FD. This energy calibration procedure is further explained in Section 4.7. The calorimetric energy measurement with the FD defines the overall energy scale of the experiment. The actual detection and reconstruction of air showers with the SD is completely independent of the FD. The two SD reconstruction methods used in this work are introduced and described within this chapter. The standard SD will be referred to as SD-1500 and the 750 m infill extension (Section 1.2.1) with SD-750. For example, the terminology SD-1500 data will be used to name data measured with the SD-1500.

When an air shower is recorded with the SD, it leaves a well defined footprint of signals (see for example Fig. 3.1b). Each individual SD station utilizes several hardware and software triggers to distinguish air shower signals from the constant background of, for example, atmospheric muons from lower energy CRs. These local station triggers are described in Section 3.1.1. To be further considered as an air shower event, multiple stations need to trigger in a certain pattern in both space and time. This physics event selection is described in Section 3.1.3. Furthermore, quality cuts ensure the proper functioning of a station at the time of the event as well as the availability of individual PMTs in each station. These quality cuts are discussed in Section 3.1.4.

The standard air shower reconstruction starts with the determination of the shower arrival direction and the impact point on ground (Section 3.3.1)). Afterwards, the lateral distribution of station signals on ground is fit with a *lateral distribution function* (LDF) as described in Section 3.3.2. For each event, the expected signal of the LDF at a certain distance to the core is used as an estimate of the primary energy. This is discussed in detail in the context of the energy spectrum in Chapter 4. The standard reconstruction exploits the trigger times of stations and the total signals deposited by secondary particles of air showers. The goal is the reconstruction of the energy and arrival direction of the primary particle. This is achieved with minimal use of simulations and small systematic uncertainties. It remains the standard method to derive the flux of UHECRs as presented in Chapter 4.

The universality reconstruction described in Section 3.4 was developed as part of [101] and further developed in [114, 115]. It is a novel method that aims at extracting information about the primary mass from the wealth of SD information, in addition to the quantities reconstructed with the standard approach. It exploits unique physics properties of individual particle components, especially their signal and time patterns in individual stations. These signal and time structures are described with the universality signal and time models that were discussed in the previous chapter Chapter 2. This new reconstruction method allows the event-by-event determination of both the depth of shower maximum X_{max} and the relative muon content of an air shower R_μ. Both of these quantities are tightly coupled with the primary mass. Results on the reconstruction of these quantities using the current dataset measured with Auger are presented in Chapter 5.

Both of these reconstruction methods are implemented in the $\overline{\mathrm{Off}}\underline{\mathrm{line}}$ software [116]. It is an extensive framework for the reconstruction and simulation of air showers, and was developed within the Pierre Auger Collaboration. As part of this work, the software was extended and further refined. Most of the work was put into extending and refining the universality reconstruction described in Section 3.4. While the general reconstruction code was already available, the changes I did include the implementation of the time model as presented in Section 2.4, various updates to the reconstruction algorithm and the general framework, for example additional reconstruction types.

3.1 Triggers and event selection

To be recognized as an air shower, each event needs to pass trigger and selection criteria on the level of individual stations and on the event level consisting of multiple triggered stations. These different triggers and selections are described in the following sections.

3.1.1 Local station triggers

Hardware triggers in each local station are the first element of the trigger chain, the T1 triggers. It asks for a coincident signal above a threshold of 1.75 VEM in all three PMTs, or for 12 FADC bins above 0.2 VEM in a time window of 120 time bins in at least two PMTs.

The next trigger level is specified in the local station software and denoted as the T2 trigger. It requires signals to fulfill either of these two criteria:

1. Threshold trigger (Thr2): requires signals above 3.2 VEM in all PMTs. This mainly triggers on short muonic signals, relevant mostly in the context of very inclined showers. The trigger rate is about 20 Hz.

2. *time-over-threshold trigger* (ToT): signals in two PMTs with more than 13 bins above 0.2 VEM and within a time window of 120 bins. This trigger produces a rate of 1 Hz to 5 Hz and works optimally for the discrimination of air shower signals from the background. It is of most relevance for vertical showers with zenith angles below 60°.

If a station fulfills one of the two T2 criteria, the information is sent to the central data station (CDAS) for further checking of an event trigger.

To further enhance the sensitivity to small signals, two new station triggers were introduced starting in June, 2013. With the ToT-deconvoluted trigger (ToTd[1]), the ToT is optimized using a prior deconvolution of the time traces [117, 118]. This provides enhanced sensitivity to small muonic signals. The MoPS[2] trigger was added in order to increase the

[1] time-over-threshold deconvoluted trigger
[2] Multiplicity of positive steps

sensitivity to small electromagnetic signals ([119] and refs. within). More discussions on the new triggers are given in the context of the flux measurement with SD-750 events as discussed in [39]. The data used in this work do not include the new station triggers, as the dataset is still very limited, and all stages of the air shower reconstruction need to be validated. This includes the re-parametrization of models to account for the addition of stations with small signals far from the shower core.

3.1.2 Central data station trigger

With the T3 trigger, time coincidences of T2 triggers that were sent to the central data station, are examined. This aims at the identification of individual triggered stations that belong to a common event, potentially an air shower. These two different station patterns are allowed:

1. A 3-fold condition (T3-3ToT): requires a coincidence of three neighboring stations with ToT triggers. The permitted time window for differences in the trigger times depends on the distance of the stations. The window is optimized to decrease the chance of random coincidences. In particular, stations have to be part of the first two crowns around the hottest station (a crown is defined as the set of all stations with the same distance to a central station). Vertical showers mostly fulfill this criterion.

2. A 4-fold condition: requires four coincident stations with T2 (Thr2 or ToT). The distance requirements are loosened such that the fourth station needs to be within four crowns around the central station instead of two. This condition is mainly relevant for very inclined showers.

The T3 online trigger was optimized to be a good compromise between selection efficiency and purity. All events that fulfill this trigger are recorded for later analysis. Accordingly, events that do not fulfill the T3 criterion are not saved. A refined offline selection as described in the following section is applied to all saved events. A visualization of the trigger criteria is shown in the next section in Fig. 3.1a.

3.1.3 Physics event selection

The T4 physics selection is used to find real air shower events in the wealth of stored data. Fake events from random coincidences of accidentally triggered stations or from lightnings are rejected with an efficiency of more than 99.99 %. The T4 selection requires a compact alignment of stations similar to the T3 online trigger described in the previous section. It can be seen as a stricter version of the T3.

There are two allowed geometries as depicted in Fig. 3.1a: a compact alignment of at least three stations within the first two crowns (T4-3ToT) or a more loose alignment of at least four stations within the first four crowns (T4-4C2). Time compatibility of the triggered stations with respect to a planar shower front moving at the speed of light is required. A first estimate of the shower geometry is derived at this stage. A variance of arrival times is expected due to sampling fluctuations and uncertainties in the absolute GPS[3] times of individual stations. Stations with incompatible timing are rejected as random stations. In most cases, they are triggered by background muons.

[3]Global Positioning System

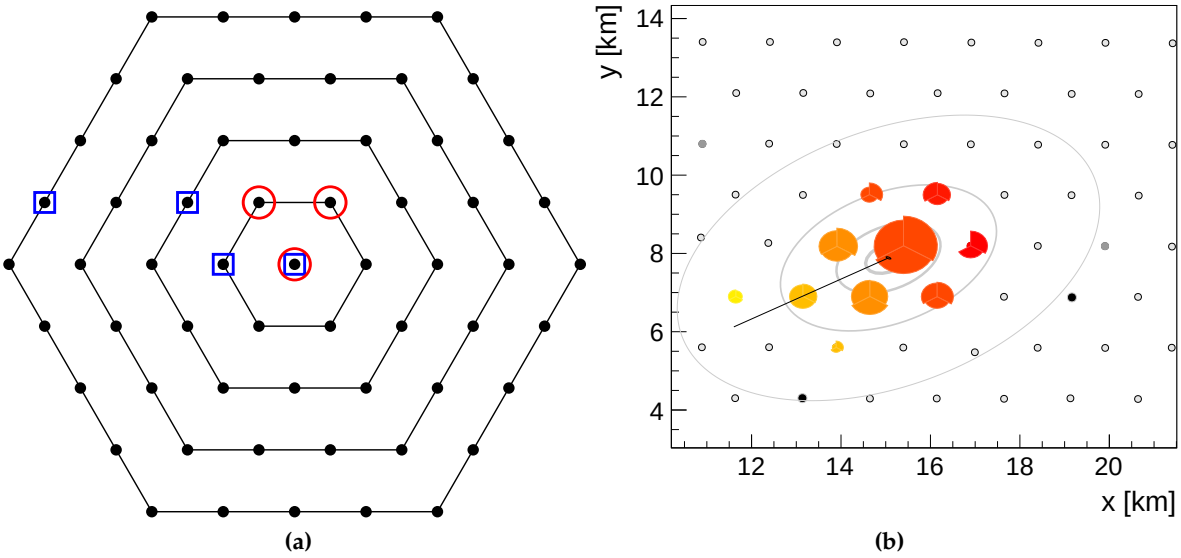


Figure 3.1: (a) Example of station configurations fulfilling the T4-3TOT condition (red) and the 4C1 condition (blue) (from [120]). (b) Example footprint of an air shower in the SD-1500 array. The x- and y-axes represent east and north, respectively (from Offline EventBrowser).

3.1.4 Station and photomultiplier quality selections

Each station transfers its state of availability to the central data station at a rate of 100 Hz. This information about the activity of individual stations is stored in separate files (with respect to the actual event data). They are called T2 files. On the one hand, it is used to calculate the geometrical exposure (described in Section 4.2). On the other hand, the activity information is used during the reconstruction of events with Offline. It ensures that stations that participate in the reconstruction were functional at the time of the event. In this step, the availability information is merged with air shower data. Analyses on the level of single PMTs allow us to reject non-functioning PMTs in the reconstruction as well. A station signal is a well defined function of an arbitrary number of active PMTs, but the fluctuations of the measured signal decrease with the addition of PMTs.

The footprint of an example event in the SD-1500 is shown in Fig. 3.1b. Shown with colored wedges are the signals of individual PMTs. Each stations is represented by three wedges if all the PMTs are functional. Asymmetries depending on the arrival direction of the air shower are visible. The overall size of each station is proportional to its total signal. The trigger times are encoded with different colors; early stations are drawn in yellow, while late stations are drawn in red. The reconstructed arrival direction of the air shower is indicated with the solid black line.

3.2 Fluctuations of the SD measurements

The process of measuring air showers with the SD underlies multiple statistical and systematic effects. An exact study of these uncertainties is required to properly reconstruct air showers and to correctly estimate the uncertainties of the reconstructed quantities.

3.2.1 Shower-to-shower fluctuations

Due to the nature of particle interactions, the evolution of air showers is prone to statistical fluctuations. Identical primary particles that create showers within identical atmospheres will lead to different ground signals. These differences are hard to quantify on the level of individual stations, but they are quantified on the level of event quantities like the shower size in Chapter 4. The typical size of variations of the measured shower size due to shower-to-shower fluctuations is 10 %. Shower-to-shower fluctuations do not need to be taken into account in the event reconstruction because all station signals are subject to the same variation.

3.2.2 Sampling fluctuations

Due to sampling fluctuations within a WCD, the measurement of signals with SD stations underlies statistical variations. The corresponding relative uncertainty decreases with increasing signal and scales with $1/\sqrt{S}$. This reflects the underlying Poissonian fluctuations of the counting of particles. The overall uncertainty is estimated from the observed variations in the signal measurement of detectors at roughly the same location of the shower [119, 121]. These detectors are called multiplets and are placed at various positions in the SD-750. Due to the increase in the fraction of the muonic with respect to the total signal with increasing zenith angle (see Fig. 2.3a), the average signal uncertainty also increases with zenith angle as $\propto \sec\theta$. The uncertainty is depicted in Fig. 3.2 as a function of signal and for different zenith angle intervals. The dashed lines represent the model predictions. A heuristic Poisson factor is used to convert signals into effective particle numbers in Section 3.3.2. The signal uncertainty is currently described with this model [119]:

$$f_S := \frac{\sigma[S]}{\sqrt{S}}(\theta) = 0.865\ (1 + 0.593\ (\sec\theta - 1.22))\,. \tag{3.1}$$

The derived Poisson factor is given by:

$$p(f_S) = t + \frac{1-t}{\sqrt{f_S}}, \tag{3.2}$$

with $t := 1/\left(1 + e^z\right)$ and $z := 40\ (0.98 - f_S)$. Plots and further details to the Poisson factor are given in Appendix B.5.1. Both the signal uncertainty and the Poisson factor are used and implemented for the event reconstruction in Offline. A model of the variance of station start times is estimated in a similar fashion [122].

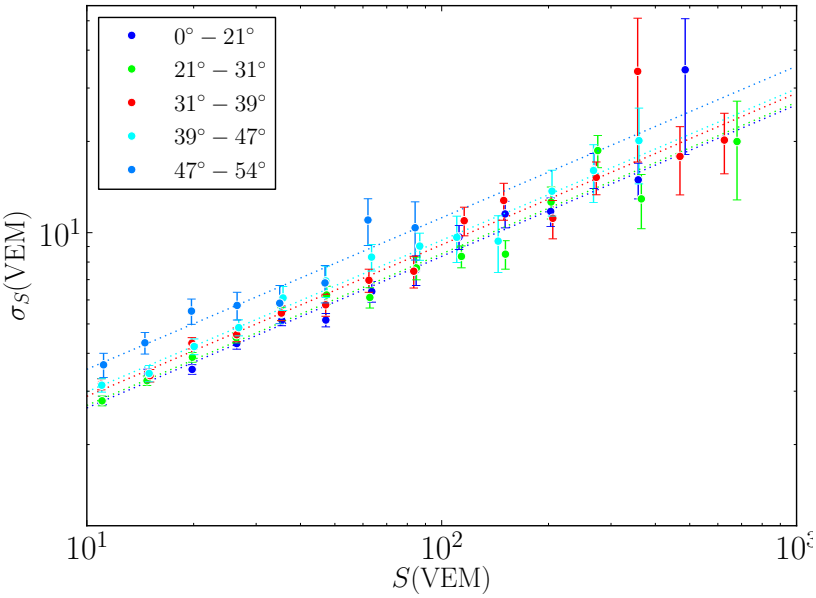


Figure 3.2: The uncertainty model for signals measured with Auger WCDs (taken from [121]). The model was derived from the measurement with multiplet detectors in the SD-750 array.

3.3 Standard air shower reconstruction

The standard air shower reconstruction was developed for the reconstruction of the energy and arrival direction of the primary UHECR. Its foundations are a fit of the shower geometry, taking into account the signals and trigger times of individual SD stations, and the estimation of a lateral distribution function that describes the radial dependence of the shower signal with respect to its reconstructed impact point. I will refer to the latter as core from heron. The air shower reconstruction as implemented in Offline is described in detail in [120].

The expected signal at an optimal radial distance, which depends on the spacing of stations, gives a robust estimate of the primary energy. It is robust in the sense that it depends very weakly on the assumption of the functional form of the LDF and on the primary mass. This is an advantage for obtaining the energy but a disadvantage for obtaining the mass. A novel reconstruction method based on the paradigm of shower universality was developed to overcome some of these issues (see Section 3.4).

3.3.1 Finding the shower geometry

As a first approximation, the arrival direction of a shower is reconstructed by fitting the start times to a plane front moving at the speed of light c along the shower axis $\hat{a}$:

$$\vec{x}(t) - \vec{b} = -c(t - t_0)\hat{a}. \tag{3.3}$$

The signal weighted barycenter of triggered stations is set as the spatial origin $\vec{b}$, while the weighted time average is used as time origin t_0.

A better description of the shower front is given by a spherical model. This description is used in a next step if there are at least four triggered stations and if the fit of an LDF succeeded. The evolution of a spherical front with speed of light is visualized in Fig. 3.3a. The relevant equation reads:

$$c(t_i - t_0) = |\vec{x}_0 - \vec{x}_i|\,, \tag{3.4}$$

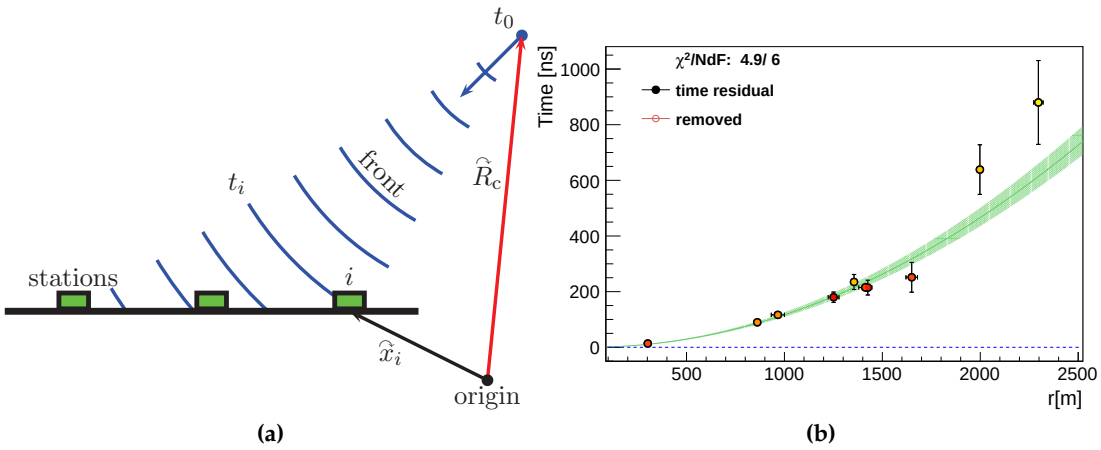


Figure 3.3: (a) Schematic depiction of the arrival of a spherical shower front at the SD and ground. (b) Example of a time fit of triggered signals to a curved shower front. Uncertainties from the time variance model are represented with error bars.

with the station trigger times t_i and positions $\vec{x}_i$. The vector $\vec{x}_0$ and t_0 are space and time coordinates of a virtual origin of the spherical shower front. This fit enables the estimation of the radius of curvature R_c through $\vec{x}_0 = \vec{c} + R_c\hat{a}$. $\vec{c}$ is the core position that needs to be estimated beforehand. The system of equations arising from Eq. (3.4) is solved with a linear approximation followed by a full non-linear optimization.

Only the very first arriving particles are expected to originate close to the point of first interaction; particles arriving later will have an effective origin closer to the shower maximum. This also depends on particle components and their respective longitudinal development as discussed in Section 2.4.

3.3.2 Obtaining the lateral distribution function

After the previous estimation of the shower arrival direction, an LDF fit is attempted. The impact point, arrival direction and the lateral distribution are fit together. The heart of this procedure is a maximum-likelihood fit taking into account probabilities for non-triggered stations and saturated stations very close to the core. Saturation is caused by an overflow of the FADC read-out electronics with a finite range and a transition to non-linear behavior of the PMTs (as mentioned in Section 1.2.1). Most of the saturated signals are recovered as explained in [35, 37, 123]. A signal spectrum of non-recovered and recovered stations is depicted in Fig. 3.5a. It is apparent that the spectrum of recovered signals nicely extends the one obtained with non-saturated stations. The functional form is a convolution of the energy spectrum and the lateral distribution of air shower signals, which is roughly described with a power law with changing spectral index.

Different functional forms were employed and tested to describe air shower signals on the ground. The current best choice which is also implemented in Offline is a modified Nishimura-Kamata-Greisen (NKG) function [124, 125]. In general, it reads:

$$\begin{aligned} S(r) &= S(r_{\text{opt}})\, f_{\text{LDF}}(r) \\ &= S(r_{\text{opt}}) \left(\frac{r}{r_{\text{opt}}}\right)^{\beta} \left(\frac{r + r_1}{r_{\text{opt}} + r_1}\right)^{\beta+\gamma} . \end{aligned} \tag{3.5}$$

The optimal distance r_{opt} depends strongly on the station spacing and was estimated to be 1000 m for the SD-1500 and 450 m for the SD-750 array [38, 126, 127]. As such, the shower size for SD-750 events is S_{450} and the one for SD-1500 data is S_{1000}. The distance r_1 and

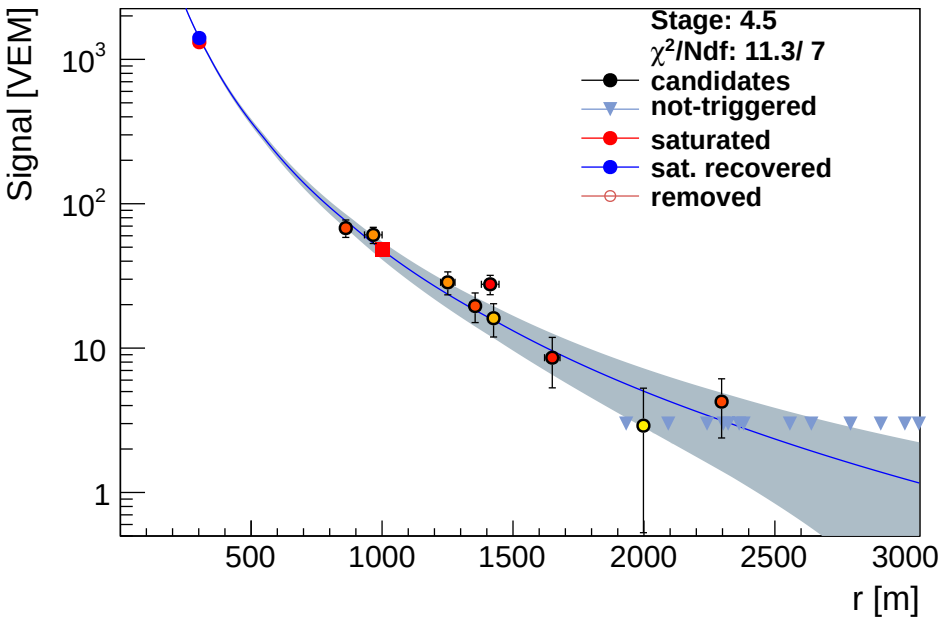


Figure 3.4: Fit of the lateral distribution function to the measured signals of an SD-1500 event with reconstructed energy of 1.43×10^{19} eV and zenith angle of 51° (Event time: 01. January 2014, Sd Id: 24909163, figure from Offline EventBrowser).

spectral index γ allow for more flexibility of the function far from the core. This reflects the transition between the lateral distributions of the electromagnetic and muonic components. The latter dominates far from the core. The current best choice as estimated from fitting data is $r_1 = 700\,\mathrm{m}$.

An example of an LDF fit is depicted in Fig. 3.4. This particular example event includes one saturated station with a successfully recovered signal. The recovered signal is used in the fit of the LDF in addition to the non-saturated signals and the non-triggered stations. There is a major caveat with the construction of a maximum-likelihood fit to obtain the LDF. The WCDs provide information about PEs induced by Cherenkov photons. The actual number and energy spectrum of secondary particles inducing a certain signal is, however, not known and very difficult to estimate. This is not an easily invertible problem due to the sensitivity of the WCD to various particle types and similar responses. In order to construct the likelihood, an effective conversion from signal in units of VEM to a number of particles is used. This effective conversion depends mainly on the zenith angle of the shower and is derived from the signal uncertainty model as shown in Section 3.2.2. With the Poisson factor p, we have the conversion:

$$n_{\mathrm{eff}} = p\,(S/\mathrm{VEM})\ . \tag{3.6}$$

The factor p is described in Eq. (3.2) and depicted in Fig. 3.5b as a function of zenith angle. At very large zenith angles, muons are the dominating particle species, and as the signal of one VCT muon is on average 1 VEM, the factor approaches one. The electromagnetic component is important at smaller zenith angles and around 20 particles account for 1 VEM. Thus, the factor increases for smaller zenith angles. I corrected a previously wrong implementation of the factor within the scope of this work, details are stated in Appendix B.5.1. The factor is estimated from the signal uncertainty model $f_s(\theta)$ as $p = \max\left(1, 1/f_s^2(\theta)\right)$.

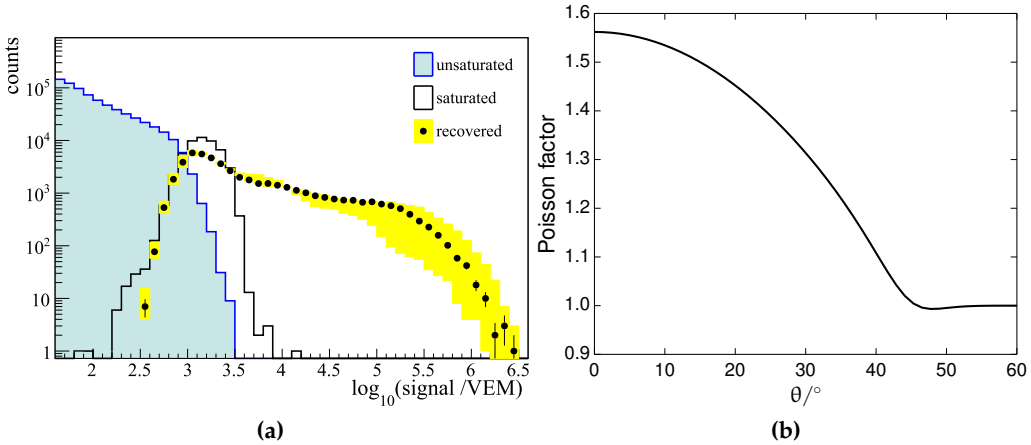


Figure 3.5: (a) Spectrum of signals measured by triggered SD-1500 stations in 5T5 events. Unsaturated stations are shown with the blue histogram, while saturated ones are plotted in black. Recovered signals are drawn with black points and a yellow area to represent systematic uncertainties of the recovery method. The recovered spectrum extends the spectrum of unsaturated stations and follows the power-law shape of the energy spectrum (folded with radial dependencies) (from [35]). (b) The Poisson factor used to convert signals to effective particle numbers as explained in the text. The factor is shown as a function of zenith angle.

With the established conversion to particle numbers, the log-likelihood function $\ln \mathcal{L}$ for the LDF fit is constructed as follows:

$$\ell = \ln \mathcal{L} = \sum_i^{\text{signal}} \ln f_{\text{P}}(n_i, \mu_i) + \sum_i^{\text{sat.}} \ln F_{\text{sat}}(n_i, \mu_i) + \sum_i^{\text{zero}} \ln F_{\text{zero}}(n_i, \mu_i)\,. \tag{3.7}$$

The function has three contributions:

1. *Non-saturated signals*: Poissonian statistics according to $f_p(n_i, \mu_i) = \mu_i^{n_i} \exp(-\mu_i)/n_i!$
2. *Saturated signals*: If the signal recovery succeeds, the recovered signal is used in the fit. Otherwise, the saturated signal is used as a lower limit in the likelihood.
3. *Non-triggered stations*: A signal of roughly 3 VEM is required to trigger single stations ($\approx$ 1 VEM with the new station triggers ToTd and MoPS). Thus, Poissonian probabilities below this threshold are summed up for each of the non-triggered stations that enter the likelihood.

Strictly speaking, the Poissonian probability is not defined for non-integer values of n_i, but it is used for arbitrary real values in this context. It can be shown that this *continued Poissonian* does not affect the properties or normalization of the p.d.f.[4] and is thus legitimate. A distinction between small Poissonian signals and large Gaussian signals is irrelevant because the continued Poissonian gives an exact description of a Gaussian for signals $\gtrsim$ 20 VEM. A distinction in the code would only make sense in terms of numerical or performance reasons.

The LDF (Eq. (3.5)) by itself has three parameters: the shower size S_{450} or S_{1000} and the slopes β, γ. Together with the arrival direction and the impact point, the number of parameters increases to at least eight. The actual fit in Offline is divided into several stages in which some of the parameters are fixed to previous estimates. This reduces the free

[4] probability density function

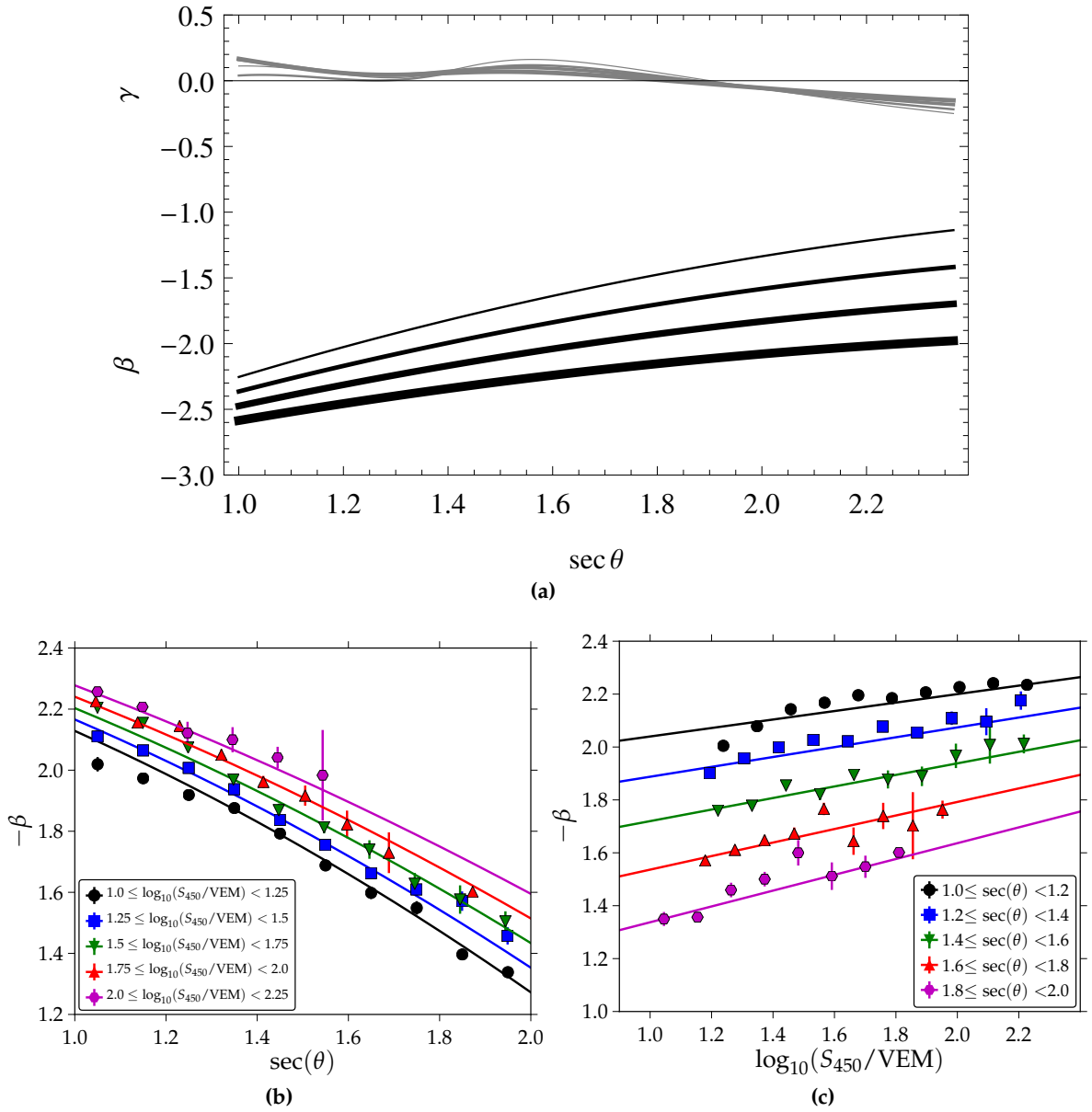


Figure 3.6: (a) Parametrization of the LDF slope parameters β and γ as a function of $\sec\theta$ for events measured with the SD-1500. The exponents are drawn for different shower sizes S_{1000} =1 VEM, 10 VEM, 100 VEM and 1000 VEM with increasing line thickness. (b) Parametrization of the LDF slope β for events measured with the SD-750 array as a function of $\sec\theta$. (c) Same as before but as a function of $\lg S_{450}$ (from [127]).

parameters of the problem and improves convergence. If there are less than four candidate stations in the fit, the slopes β and γ cannot be fitted, and values from parameterizations are used. These parameterizations are obtained from the analysis of large multiplicity events or, alternatively, global fits in which the information of all events are exploited at the same time [127, 128]. The best set of parameters for β and γ in the case of SD-750 events were obtained with a global event fit as described in [127]. This parametrization is used within this work and in [39]. The parametrization for the SD-1500 is older and has been obtained with a previous analysis method in [129]. It would be worthwhile to apply the analysis of a global event fit to events measured with the SD-1500 in order to improve and update existing parameterizations of the LDF parameters. In particular, the accuracy of

the LDF parametrization for different core distances could be improved significantly. At the same time, the global event fit would enable a parametrization of azimuthal signal asymmetries (building on top of analyses of simulations). These occur due to differences in the distributions of ground particles in the early and late part of the shower and are not taken care of in the current fit, leading to biases in the reconstructed core position. The core bias increases with increasing zenith angle. This is highly relevant for the analysis of radio events, in which reconstructed SD parameters are used. Biased parameters will thus lead to biases in derived quantities.

To summarize, the standard reconstruction enables the estimation of the shower geometry and the impact point of the shower, as well as shower sizes at an optimal distance from the shower core. The shower size is S_{450} for SD-750 data and S_{1000} for SD-1500 data. In this work, the shower sizes are corrected for their zenith angle dependency in Section 4.5 and calibrated to the energy scale given by the FD measurement in Section 4.7. The reconstructed energies are used to derive the flux of UHECRs as discussed in Chapter 4.

3.4 Universality reconstruction

The universality reconstruction is an analysis technique based on the signal and time models that are discussed in Section 2.3 and Section 2.4. Compared to the standard reconstruction discussed in the previous section, SD information is exploited more deeply in order to be able to reconstruct information about the primary mass. This includes information about the time structure of particle traces. The reconstruction is implemented as a separate module in the $\overline{\text{Off}}\underline{\text{line}}$ framework and optionally uses parameters from the standard reconstruction, for example the reconstructed primary energy.

The reconstructions is composed of several aspects: a fit of the distribution of time traces , the start times of triggered stations and the lateral distribution of the total measured signal. Depending on the signal size of individual stations, a station contributes to one or two of these domains.

The universality fit is a combined reconstruction of the core position $\vec{x}$, the relative core time t_c, the shower arrival direction (θ, ϕ), energy E, shower maximum X_{max} and relative muon number R_μ. This amounts to 9 parameters if each one is set as a free parameter in the fit. Shower maximum, energy and muon number are highly correlated and are very difficult to fit at the same time, except for events with very large station multiplicity. By default, the energy is fixed to the result of the standard reconstruction, while X_{max} and R_μ are estimated.

The total log-likelihood reads:

$$\ell = \ln \mathcal{L} = \ell_{\text{shape}} + \ell_{\text{start}} + \ell_{\text{LDF}}\,, \tag{3.8}$$

with the individual likelihood contributions explained in the following sections in the order in which they are listed.

3.4.1 Fit of arrival time distributions

Traces with at least five signal bins above 0.5 VEM/bin contribute with their entire time trace to the reconstruction. Each time bin is compared to the universality prediction, which is obtained by combining the time and signal responses for the four particle components. I will denote the measured instantaneous signal in bin i as s_i and the prediction as $\hat{s}_i$. t_i will refer to the time at the bin center i with respect to the arrival time of a plane front at the position of the station. The start time t_s describes the center of the start bin with respect to the plane front time. The offset of the arrival times of first particles with respect to the

plane front is given by the curvature model T_c, which depends on the particle component p, shower geometry and properties of the primary particle:

$$T_c^p = T_c^p(r, X_{\max}, \lg E, \theta, \phi) \tag{3.9}$$

The curvature model is discussed in more detail in Section 2.4. The relevant times of a time trace are depicted in Fig. 3.7a. Shown is a typical log-normal time distribution. The plane front time as well as the first particle time are indicated. To take a non-linear shape of the time model into account, the model is integrated within the bin of question prior to comparing it to the measured signal. This is done by evaluating the c.d.f. of the time model at the edges of the bin, subtracting the resulting values and dividing by the bin width. The measured quantity and the estimate for bin i are thus denoted as S_i and $\hat{S}_i$. The model estimate is obtained from the normalized time model f^p_{trace} and the signal model f^p_S:

$$\hat{S}_i = \hat{S}(t_i) = \sum_{p=1}^{4} f^p_{\text{trace}} f^p_S \tag{3.10}$$

with the sum running over the four particle components. The model dependencies of f^p_{trace} and f^p_S are omitted for readability: $f(r, \theta, \psi, \lg E, \Delta X, R_\mu, p, \text{month})$.

The actual log-likelihood is then given by a sum over the logarithm of normal distributions for each time bin with mean values given by the model predictions and uncertainties according to the signal uncertainty model detailed in Eq. (3.1):

$$\ell_{\text{shape}} = \sum_{s}^{\text{stations}} \sum_{i} \ln \mathcal{N}(S_{s,i}, \mu = \hat{S}_{s,i}, \sigma = \sigma_{\text{SD}}(\hat{S}_{s,i})) \tag{3.11}$$

Saturated events

A saturated event is defined as an event with at least one low-gain saturated station. Time traces from low-gain saturated stations are treated differently. They still contribute to the shape fit, but only with the non-saturated part of their trace, i.e. all time bins that do not belong to the plateau. At the moment, this partial trace is comprised of all bins with signal entries $S_i < 0.95 \max_j S_j$. The plateau of saturated bins is thereby excluded. An example is depicted in Fig. 3.7c.

3.4.2 Estimation of the start time

Traces with small signals or saturation cannot be used in shape fits without introducing potential biases. In order to still exploit information from the time structure of these measurements, the station trigger times are compared to model predictions of the arrival time of first particles. As discussed earlier, the particles arriving earliest are typically muons. As such, the model prediction is obtained from the time model f_{time} for the muonic component through an *extreme value transformation* to calculate the p.d.f. of the first arriving particle. Mathematically, it reads:

$$f_{\text{first}}(t) = n_\mu \left(1 - \int_0^t \mathrm{d}t' f^\mu_{\text{time}}(t')\right)^{n_\mu - 1} f^\mu_{\text{time}}(t) \tag{3.12}$$

The time t is expressed relative to the time given by the curvature model for muons. n_μ is estimated from the signal model of the muonic component, taking the track-length in the detector into account. In taking only the model for muons into account, the p.d.f. is an

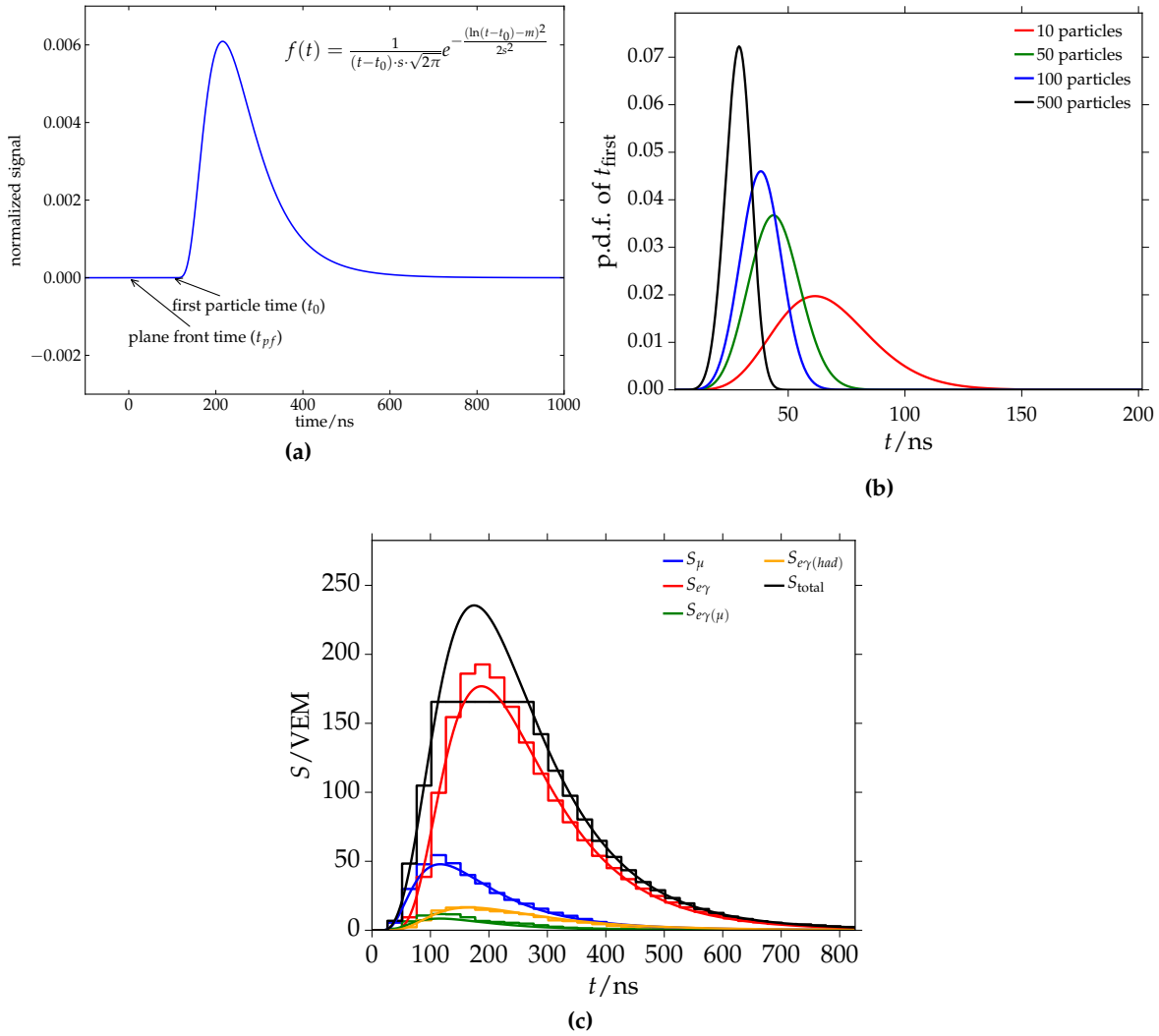


Figure 3.7: (a) Explanation of the relevant times for a particle time trace. Indicated are the plane front time and the time of first particle arrival. (b) Example of distributions of the arrival time of first particles for different number of particles, and thus, different magnitudes of sampling delay. This plot represents a toy example assuming a log-normal trace with semi-realistic parameters of $m = 5$ and $s = 0.55$ (compare to Section 2.4). (c) Example shape fit of a low-gain saturated particle trace. Only the total trace in black is actually fitted. The histograms of the other particle components are compared to the predictions from the time model.

approximation in itself. An example of distributions of arrival times for different number of particles or signal sizes is illustrated in Fig. 3.7b. The average delay in the arrival time of the first particle increases with decreasing signal size or number of particles. The log-likelihood for the fit of start times is then constructed as follows:

$$\ell_{\text{start}} = \sum_{s}^{\text{stations}} \ln f^{s}_{\text{first}}(t^{s}_{\text{start}}) \tag{3.13}$$

with the sum including all stations participating in the start time fit and t^{s}_{start} the respective start time determined by an algorithm in Offline. The distribution of the expected arrival time of the first particle varies from station to station, as does the actual measured quantity. The extreme value transformation in Eq. (3.12) allows one to construct a p.d.f. of the earliest particle arrival time. By construction, this p.d.f. only takes statistical sampling

delays into account. In reality, the arrival time of the earliest muon is the time that the particle needs to propagate from the point of its creation, which is very close to the first interaction point X_0. As described at the beginning of Section 2.4, the start time model is a curved shower front with an effective origin of particles along the shower axis. This origin depends on the distance to the shower maximum, the shower zenith angle and the particle component. However, this description does not take into account fluctuations in the depth of the first interaction point. These fluctuations are only relevant for very early particle times, but they drive the fluctuations of the arrival time of the first muon. The width of the distribution given by Eq. (3.12) is thus underestimated. Reconstructions of simulated showers as presented in Section 3.4.5 have shown that the use of the start time fit leads to a bias of the reconstructed X_{max}. This bias is on the level of $30\,\mathrm{g\,cm^{-2}}$ at 10^{19} eV and decreases with increasing energy. This problem is likely to be resolved when the fluctuations in X_0 are properly taken into account in the p.d.f.. For this work, I am not using the start time fit in the universality reconstruction. The criterion for stations to participate in the shape fit was relaxed such that the amount of information added with the start time fit is negligible.

3.4.3 LDF fit

Stations that do not participate in the shape fit are included in an LDF fit that compares the total measured signal to the expectation from the signal model. Also non-triggered stations are taken into account. The log-likelihood is constructed as in the case of the standard reconstruction:

$$\ell_{\mathrm{LDF}} = \sum_{s}^{\mathrm{signal}} \ln f_{\mathrm{P}}(n_s, \mu_s) + \sum_{s}^{\mathrm{sat.}} \ln F_{\mathrm{sat}}(n_s, \mu_s) + \sum_{s}^{\mathrm{zero}} \ln F_{\mathrm{zero}}(n_s, \mu_s)\,. \tag{3.14}$$

The function itself has three contributions:

1. *Non-saturated signals*: Poissonian statistics according to $f_p(n_s, \mu_s) = \mu_s^{n_s} \exp(-\mu_s)/n_s!$
2. *Saturated signals*: If the signal recovery succeeds, the recovered signal is used in the fit. Else, the saturated signal is used as a lower limit in the likelihood.
3. *Non-triggered stations*: A signal of roughly 3 VEM is required to trigger single stations (≈ 1 VEM with the new station triggers ToTd and MoPS). Thus, Poissonian probabilities below this threshold are summed up for each of the non-triggered stations that enter the likelihood.

As in the standard reconstruction, the signal uncertainty model described in Section 3.2.2 is used to convert signals into effective particle numbers.

3.4.4 Example reconstructions

A simulated event

To validate and visualize the universality reconstruction, this section describes the results of reconstructing a simulated proton shower with a true energy of $10^{19.5}$ eV and a zenith angle of 36°. The event with the Sd Id 204881 is part of the continuous shower library as described in Appendix E.3. The reconstruction type used for this example is discussed in Section 3.4.5 (Fitting everything but the energy).

The relative muon content R_μ is reconstructed to 0.990 ± 0.084 compared to a true value of 1.0 and the shower maximum is estimated to be $X_{\mathrm{max}} = (820 \pm 30)\,\mathrm{g\,cm^{-2}}$ with a true

value of 842.27 g cm^{-2}. Both quantities are well reconstructed within the statistical uncertainties. The energy is fixed to the Monte Carlo value. The shower arrival direction is estimated with a small angular deviation of 0.5°, while the reconstructed impact point is reconstructed 23 m from the true position. These are typical deviations within the expected range of fluctuations. The footprint of stations is visualized in Fig. 3.8a. The size of the stations is proportional to their logarithmic signal. The blue right-pointing triangle gives the reconstructed core position, which needs to be compared to the true position depicted in with the red, upward-pointing triangle. The standard reconstruction yields the core shown as black triangle, which shows a larger deviation from the true value. There are three stations at roughly the same distance to the core and with similar signals of 364 VEM, 244 VEM and 168 VEM. In total, there are 11 triggered stations and 10 of those contribute to the shape fit. The station distribution is not optimal for the universality reconstruction, because the three hottest stations are at roughly the same distance to the core. This situation is similar to the one of saturated events as further discussed in Section 3.4.5. The LDFs are shown in Fig. 3.8b. In the universality reconstruction, the total LDF is compared to the sum of the expected distributions of the four signal components (derived from the signal model discussed in Section 2.3). As such, only the black points are fit. Distributions for the individual components are shown with respect to the model predictions. The shaded regions around the model lines represent their azimuthal asymmetries. After the muonic signals, subsequent points are slightly shifted to larger radii to better visualize the error bars. In all cases, the universality models give an accurate description of the component signals. The signal uncertainty model described in Section 3.2.2 is used to calculate the uncertainties on the total and the component signals. For the component signals, this is done to give an estimate of the uncertainty and is not strictly valid because the uncertainty model gives an empirical description of the fluctuations of the total signal only. The remaining four plots visualize the results of the shape fit for the four hottest stations. In each case, only the total trace is actually fit as described in Section 3.4. The individual particle traces are compared to the expected time distributions from Section 2.4.

A real event

This section covers the description of the result of the reconstruction of the event with the SD Id 14837428. The event was recorded on May 24th, 2012. Its longitudinal development was simultaneously and independently observed with the FD telescopes Loma Amarilla (LA) and Coihueco (CO). Additionally, the lateral extent was recorded with 14 stations of the SD-1500. The event is part of the sample of golden hybrid events that pass the quality cuts for the calibration procedures for universality (Section 5.2) and the energy spectrum (Section 4.7). The recorded footprint of the UHECR event in the SD-1500 and both FD telescopes are visualized in Fig. 3.9a. The profiles of the longitudinal energy deposit as measured with LA and CO are shown in Fig. 3.9b and Fig. 3.9c. The corresponding reconstructed depths of shower maximum are $X^{\mathrm{FD}}_{\mathrm{max}} = (728\pm5)\,\mathrm{g\,cm^{-2}}$ (LA) and $X^{\mathrm{FD}}_{\mathrm{max}} = (764\pm19)\,\mathrm{g\,cm^{-2}}$ (CO). Due to the distance to the telescope, the profile recorded with CO is shallow, resulting in larger reconstruction uncertainties. The reconstructed energies are $E_{\mathrm{FD}} = (12.5 \pm 0.3 \pm 0.6)$ EeV (LA) and $E_{\mathrm{FD}} = (13.9 \pm 1.0 \pm 1.0)$ EeV (CO). The primary energy reconstructed with the SD-1500 and the standard reconstruction is $E_{\mathrm{SD}} = (12.9 \pm 0.5 \pm 0.9)$ EeV. The result obtained with the universality reconstruction is $E_{\mathrm{Univ}} = (12.7\pm0.6)$ EeV, which is compatible with the other results. The universality reconstructed depth of shower maximum is $X^{\mathrm{Univ}}_{\mathrm{max}} = (749\pm18)\,\mathrm{g\,cm^{-2}}$. This is compatible within $1\,\sigma$ to the weighted average of the FD measurements of $\langle X^{\mathrm{FD}}_{\mathrm{max}}\rangle = (730\pm9)\,\mathrm{g\,cm^{-2}}$. The results of the individual geometry reconstructions agree well with each other, yielding an average zenith angle of (37.1±0.3)°

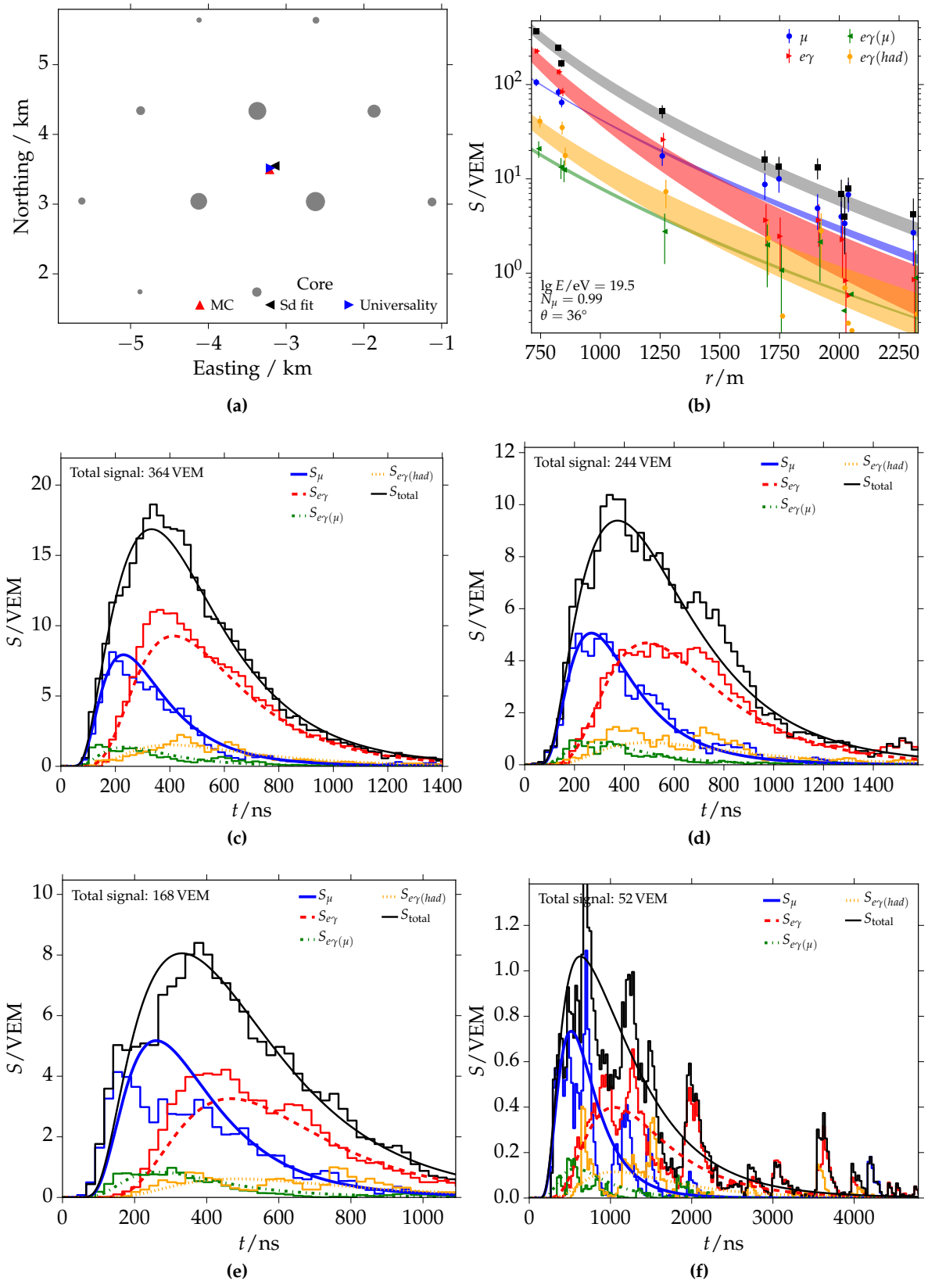


Figure 3.8: Universality reconstruction of a simulated proton shower with an energy of $10^{19.5}$ eV. (a) Footprint of the triggered stations in the ideal array together with the Monte Carlo and the reconstructed core positions. (b) Fit of the LDF. Only the total LDF shown in black is fit. A comparison of the individual component LDFs and the model predictions is given. More details are given in the text. (c) - (f): Results of the shape fits for the four hottest stations. Only the total traces are fit, the other ones are plotted for comparison. (f) is a lower signal trace, which explains the increased fluctuations.

and an average azimuth angle of $(24.1 \pm 0.4)^\circ$. The estimates obtained with the universality reconstruction are $\theta = (37.4 \pm 0.3)^\circ$ and $\phi = (23.9 \pm 0.4)^\circ$. The footprint of SD-1500 stations and reconstructed core positions are depicted in Fig. 3.10a. The fit of the LDF is shown in Fig. 3.10b. The measured signals are fit to the sum of the expectations for the four signal components, which is depicted as a black line. The shaded region represents the azimuthal asymmetry of the prediction. The predictions for the LDFs of the individual particle components are indicated with different colors. Fits to the measured traces of the four hottest stations of the event are shown in Fig. 3.10c to Fig. 3.10f. The hottest station with a signal of 427 VEM has the largest weight in the reconstruction, resulting in a very accurate prediction of the total trace as plotted in Fig. 3.10c. While the lower signal traces show larger fluctuations, their general trends are well accounted for by the model.

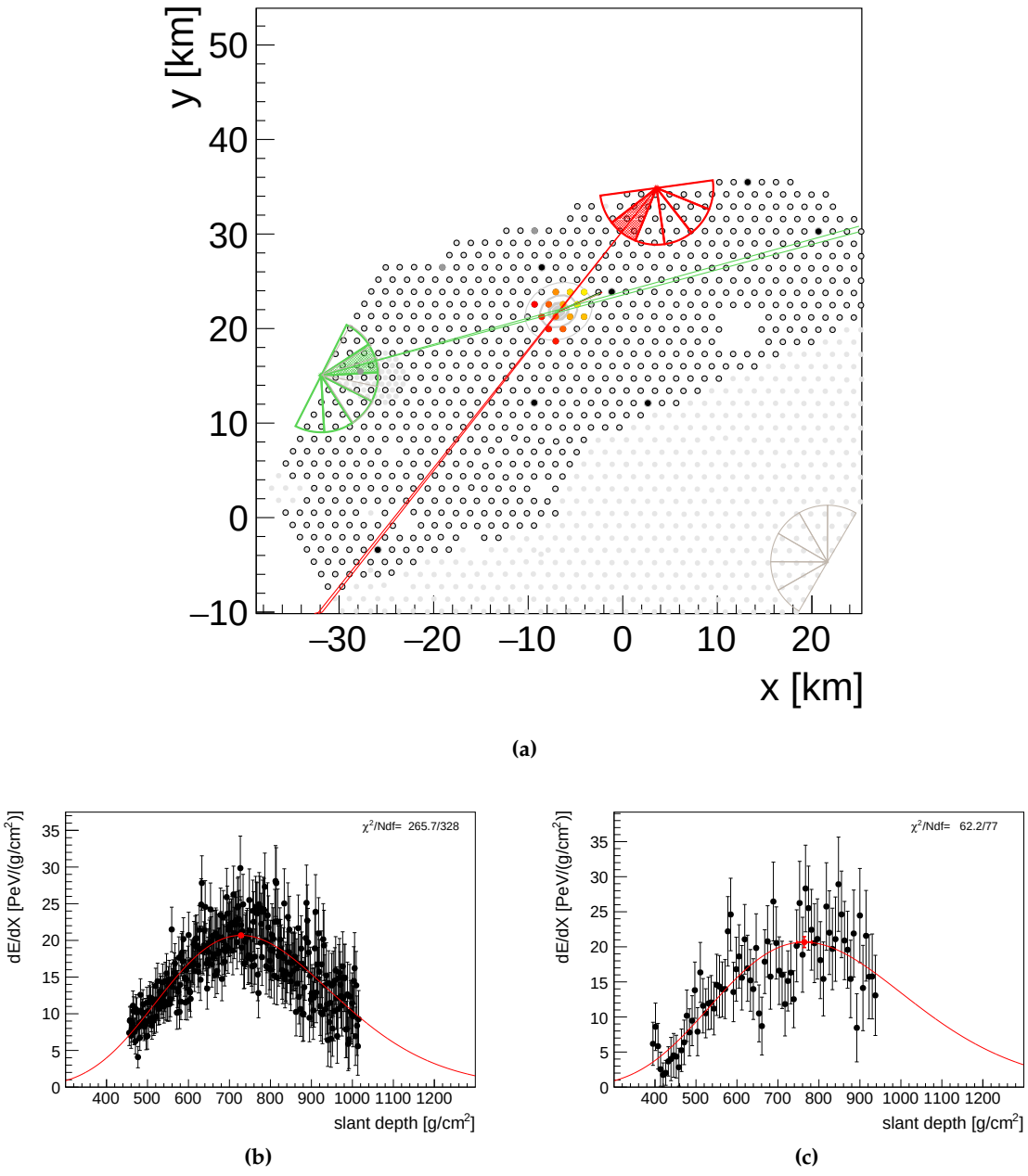


Figure 3.9: Visualizations from the Offline EventBrowser for the reconstructed event with SD Id 14837428. The event was recorded by two FD eyes as well as the SD-1500. It was reconstructed with both the standard and the universality approach. (a) Schematic view of the footprint of the shower in the SD-1500 array. The FD telescopes Coihueco and Loma Amarilla are indicated with differently colored wedges. (b) Profile of the differential energy deposit as a function of slant depth, as measured and reconstructed with the Loma Amarilla FD telescope. (c) Energy deposit profile as measured with Coihueco.

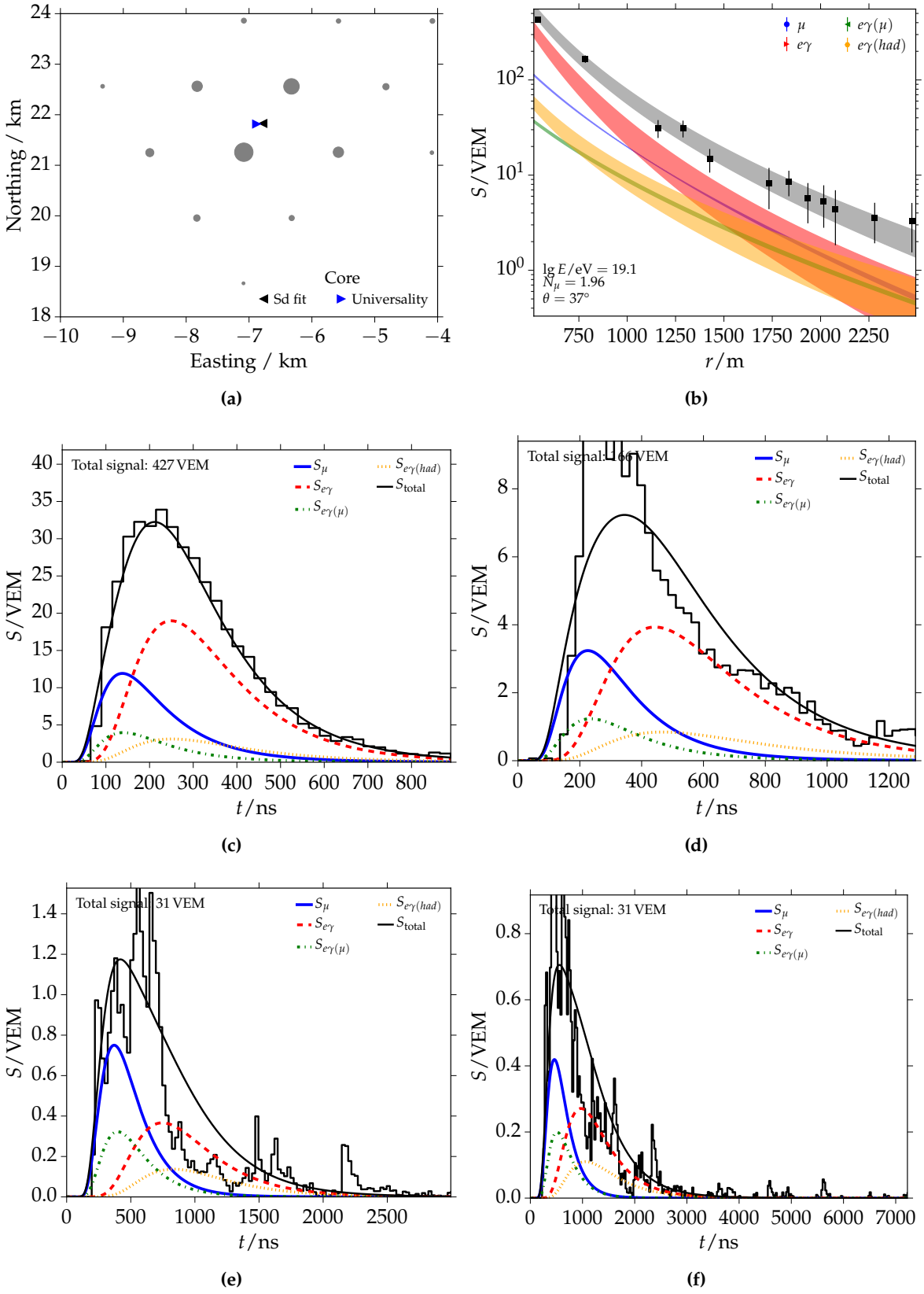


Figure 3.10: Universality reconstruction of a real event with a reconstructed energy of 1.3×10^{19} eV (SD Id: 14837428). (a) Footprint of the triggered stations in the SD-1500 array together with the reconstructed core positions. (b) Fit of the LDF. Only the total LDF shown in black is fit. Shown are the predicted distributions for the different particle components. (c) - (f): Results of the shape fits for the four hottest stations. Only the total traces are fit, the other ones are plotted for comparison.

3.4.5 Validation of the universality reconstruction using simulations

The validation of the universality reconstruction consists of the study of biases and resolutions of the reconstructed quantities. For this purpose, the events in the shower library described in Appendix E.3 were reconstructed along the description in Section 3.4. The reconstruction code is part of the Offline framework. The library of simulated events consist of proton and iron induced showers simulated with QGSJET-II.03. This library is independent of the ones that were used to build the signal and time models.

The quantities estimated in the current universality reconstruction are the core position $\vec{x} = (x, y, z)^T$, the relative core time t_c, the shower arrival direction (θ, ϕ), the primary energy E, the shower maximum $X_{\max}$ and the relative muon content R_μ. Due to the strong correlation between the relative muon number and the primary energy (see e.g. [104]), these quantities are not estimated at the same time. In the standard universality reconstruction, the energy is fixed to the estimate from the standard LDF reconstruction. For all following plots throughout this thesis, the comments in Appendix A.2 apply.

Fitting only the shower maximum

To check the signal and time models discussed in Chapter 2, only $X_{\max}$ is reconstructed; the other quantities are fixed to true values (Monte Carlo values, short: MC) and are not allowed to vary during the fit. The result of this reconstruction is depicted in Fig. 3.11a. Shown are mean profiles of the event-by-event difference between the reconstructed and true $X_{\max}$ (denoted as $X_{\max}^{\mathrm{MC}}$) of non-saturated events as a function of $\lg E$ and for different ranges in zenith angle θ. On average, the reconstruction yields an unbiased estimate of $X_{\max}$. However, $X_{\max}$ is slightly underestimated for zenith angles above 50°, especially for energies below 10^{19} eV. The probable reasons are inaccuracies in the signal and time models at largest zenith angles, where inaccuracies in the parametrization are largest. Another possible but less likely reason is a reconstruction issue that only biases results at high zenith angles. An unbiased estimation of $X_{\max}$ for $\theta < 50°$is also true for saturated events as shown in Fig. 3.11b. In either case, there is no dependence on primary energy. For an unbiased reconstruction in the complete range of energy and zenith angle, I am only considering events with zenith angles from 0° to 50° for all following analyses. Considering all events with $\theta < 50°$, there is no reconstruction bias and no dependence on primary energy or zenith angle (also not on other quantities). In the current reconstruction, all measured times are shifted by −14.5 ns to account for occurring biases when the timing is fixed to the true timing. This is however not relevant for other reconstruction types in which the relative core time t_c is anyway a fit parameter and absolute offsets do not matter. The residual bias for proton and iron primaries is depicted in Fig. 3.13a. The $X_{\max}$ of proton primaries is on average underestimated by $6\,\mathrm{g\,cm^{-2}}$, while the one of iron primaries is overestimated by the same amount. This difference is a consequence of the slight dependence of the time model parameters on the primary type itself. The time model was parameterized for an average of the primary species p, C and Fe. For example, the problem is illustrated in Fig. 2.25b with the impact on the t_{50} quantile. It is mostly relevant for the time model of muons and particles from muon decay. Residuals for the other particle components do not show the same offsets for different primary species (and interaction models). It is easy to account for this by introducing an offset in the time model parameters of the muonic particle components. This offset depends on the primary species and eliminates the $X_{\max}$ biases. It is discussed in the following section.

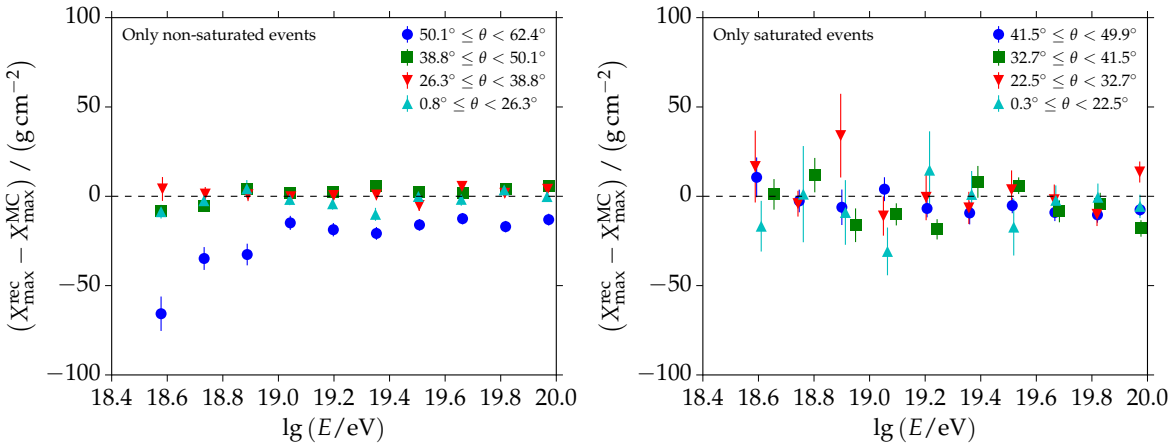


Figure 3.11: (a) The mean bias in the reconstructed shower maximum $X_{\max}$ as a function of $\lg E$ and for different ranges in zenith angle. Only SD information is used to infer $X_{\max}$ as described in this section. The plot results from reconstructions of only $X_{\max}$, all other variables are fixed to true quantities. Only non-saturated events are considered. (b) The same plot for saturated events and zenith angles below 50°. In each case, the error bars represent the uncertainty of the mean.

Composition dependence of model parameters

For a log-normal distribution, the dependence $t_{50} = e^m$ holds independent of s (see Fig. 3.12a), with the mean and width parameters m and s. As such, the difference in the t_{50} estimates due to primary species are easily translated into offsets in the m parameter. The relevant relation is:

$$\Delta m = \frac{\Delta t_{50}}{t_{50}}. \tag{3.15}$$

Based on the differences between proton and iron in Fig. 2.25b, the offset Δm is quantified as a function of core distance. The relative bias in $\frac{\Delta t_{50}}{t_{50}}$ for proton is subtracted from the one of iron showers for a common number of bins in r. The result is shown in Fig. 3.12b for the muonic particle component. An average of simulations calculated with QGSJET-II.03 and EPOS-1.99 is used. The offset increases from 0.01 at 250 m to 0.03 at 1000 m and above. On average, it is slightly larger for smaller zenith angles and higher energies. From the dependence of $m(\Delta X, r = 1000\,\mathrm{m})$ illustrated in Fig. 2.13a, I extract the slope $\frac{\Delta X_{\max}}{\Delta m} = 1000\,\mathrm{g\,cm^{-2}}$. At a smaller radial distance of 200 m, the slope is $500\,\mathrm{g\,cm^{-2}}$. Taking into account that the estimation of $X_{\max}$ is driven by stations in the radial range 200 m to 1000 m, the expected bias between the $X_{\max}$ reconstruction of iron and proton showers is $\Delta X_{\max} \approx 0.02 \cdot 700\,\mathrm{g\,cm^{-2}} = 14\,\mathrm{g\,cm^{-2}}$, with the average values $\langle \Delta m \rangle = 0.02$ and $\left\langle \frac{\Delta X_{\max}}{\Delta m} \right\rangle = 700\,\mathrm{g\,cm^{-2}}$. The observed average bias of $\Delta X_{\max} = 12\,\mathrm{g\,cm^{-2}}$ (twice the amount of Fig. 3.13e) is in very good agreement with that derivation. A slight decrease of the bias with increasing zenith angle is also in agreement with $\Delta m \propto \sec\theta$. This is only a qualitative argument though, the exact radial dependence of $m(\Delta X)$ and the slope of the LDFs need to be considered for a full calculation. The correction of this offset in the analysis of data is discussed in Section 5.2. As the composition in data is not accurately known and changes with energy, it is however relevant to quantify the systematic uncertainty arising from the dependence of the reconstruction bias on the primary species. This systematic uncertainty on the estimated $X_{\max}$ is shown in Fig. 3.13e. The average is $6\,\mathrm{g\,cm^{-2}}$. It is obtained from a chi-squared fit of a constant to half of the difference between the average $X_{\max}$ bias of

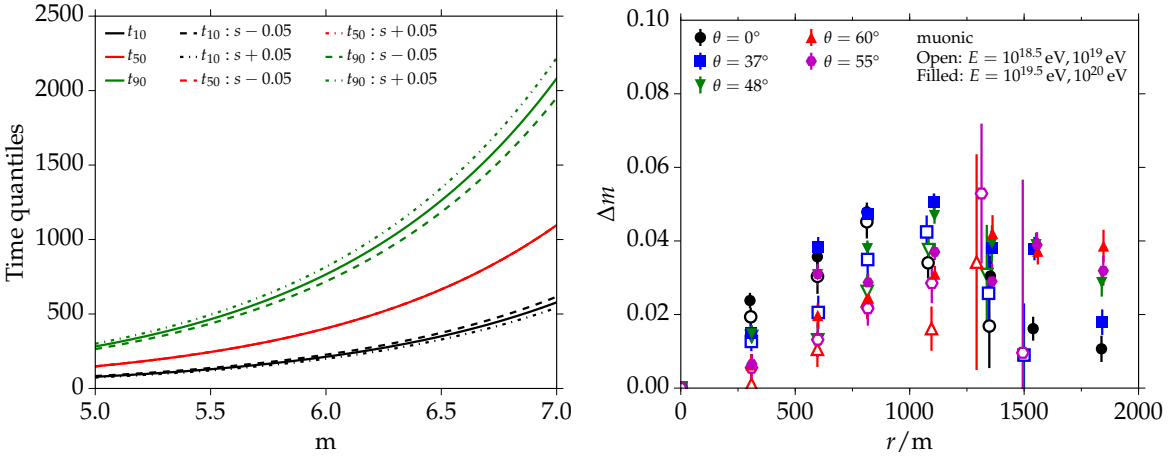


Figure 3.12: (a) Different time quantiles of the log-normal distribution as a function of the mean parameter m and for different realizations of the width s. Realistic parameters from the muon time model are chosen. (b) The offset in the mean parameter m resulting from differences between proton and iron showers as a function of radial distance and for different zenith angles. Two different sets of energies are shown with open and filled markers.

proton and iron. It is accounted for in the data analysis in Chapter 5. The resolution of X_{max} in this optimal scenario is shown in Fig. 3.13d as a function of energy and for different primaries. X_{max} is mainly estimated from the dominant muonic signal component. As the muonic signal is on average larger in iron showers, $\sigma\left[X_{\mathrm{max}}\right]$ is smaller in this case. The resolution approaches $15\,\mathrm{g\,cm^{-2}}$ at highest energies. This describes a lower limit on the achievable resolution using the shower universality reconstruction for signals measured with WCDs.

Estimating the muonic time model offset in simulations

The time model offset Δm_μ is fit to describe data in Section 5.2. This is an important step in the calibration of the universality reconstruction with data because the composition changes with energy and the simulations might not describe data perfectly. The offset describes the difference with respect to the mean parametrization of the time model due to a different primary species. To test the method of deriving the offset with simulations, all quantities are fixed to the true quantities. In a first step, the core position and R_μ are estimated based on the signal model only. In the next step, the overall timing and Δm_μ are fit. The result is shown as a function of energy and zenith angle in Fig. 3.14. Results for proton induced showers are shown in red, while iron is drawn in blue. There is a small positive bias of 0.01 for iron events and a small negative bias of −0.01 for proton events. Energy and zenith angle dependencies are negligible. This matches with the expectations from Fig. 3.12b. The time model used in the reconstruction is based on an average of proton, carbon and iron primaries. To estimate the time model offset on an event-by-event basis, the radial dependence is fixed to the following function:

$$\Delta m_\mu(\lg E, \theta, r) = \Delta m_\mu(\lg E, \theta) \cdot \frac{1 - \exp\left(-2.65\, r/2000\,\mathrm{m}\right)}{1 - \exp\left(-2.65\right)}\,. \tag{3.16}$$

The radial dependence was derived in a separate note [130] and verified with the analysis of the offset in simulations in Section 3.4.5.

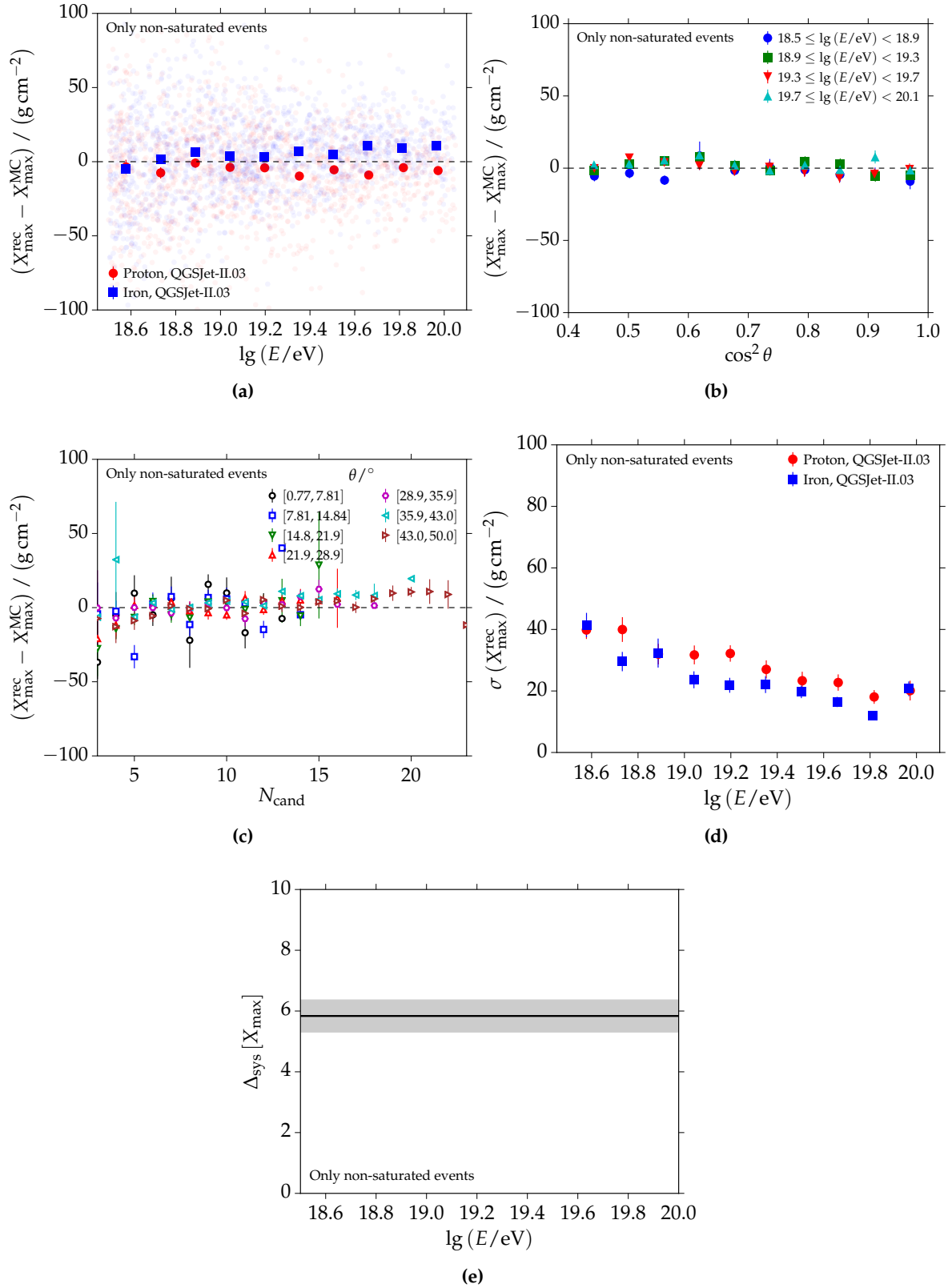


Figure 3.13: Biases and resolutions of the reconstructed X_{max} using a universality reconstruction with fixed variables as described in the text. (e) Systematic uncertainty on X_{max} derived from the difference in the reconstruction between proton and iron induced showers.

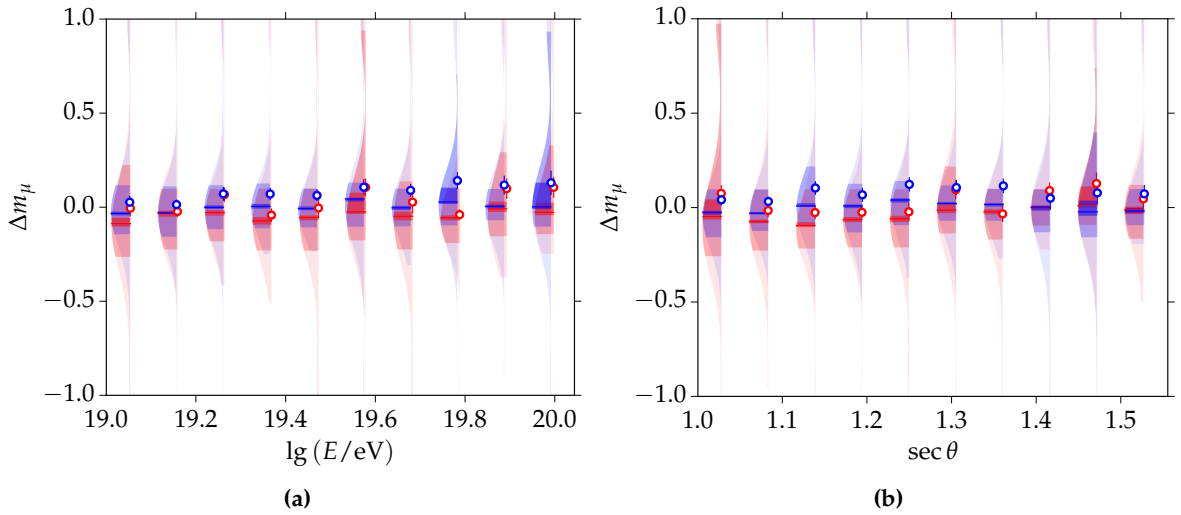
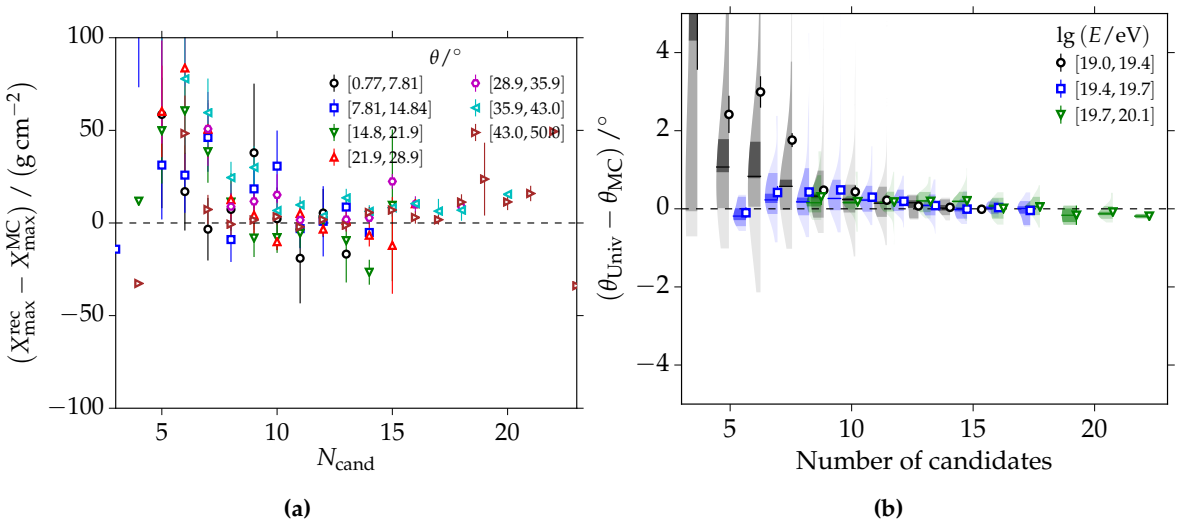


Figure 3.14: Results on the estimation of the time model offset for the muonic particle components in simulations. Results for proton/iron primaries are shown in red/blue. Plot of the time model offset as a function of (a) energy and (b) zenith angle.

Fitting everything but the primary energy

In the established reconstruction routine [101], all variables except the energy are fit at the same time. The primary energy is fixed to the result of the standard reconstruction. In this classic reconstruction, there are eight free parameters in the reconstruction. To study biases and resolutions, all simulated events are reconstructed. The bias for events with $\theta > 50°$ is strongly pronounced and as such these events are omitted in the following comparison. The X_{max} bias is shown in Fig. 3.15a as a function of the number of candidate stations and for different ranges in zenith angle. It is apparent that X_{max} is largely overestimated for $N_{\mathrm{cand}} < 9$. Due to the large degrees of freedom, there is not enough information in the reconstruction to allow an unbiased estimation of X_{max} with less stations contributing to the fit. A bias in the reconstructed zenith angle for low station multiplicities is present as well as shown in Fig. 3.15b. A cut on $N_{\mathrm{cand}} > 8$ seems necessary in order to obtain an unbiased result when all variables except the energy are fit. Applying this cut greatly reduces biases and improves the resolution of reconstructed quantities. The remaining bias in the reconstructed X_{max} is shown in Fig. 3.16a as a function of energy and for different primary species. Again, there is a small positive bias for iron events due to the time model offset. Events with energies below 10^{19} eV are mostly lost due to the cut on the number of candidate stations. The bias for proton events at lower energies might be due to the small event statistics. The bias in the reconstructed R_μ is shown in Fig. 3.16b. On average, there is a small negative bias for proton events, which is statistically not significant. The resolution of X_{max} is shown in Fig. 3.16c. On average, it drops from $50\,\mathrm{g\,cm^{-2}}$ at 10^{19} eV to $25\,\mathrm{g\,cm^{-2}}$ at 10^{20} eV. The resolution is $5\,\mathrm{g\,cm^{-2}}$ to $10\,\mathrm{g\,cm^{-2}}$ larger for proton events. The resolution of R_μ is shown as a function of $\lg E$ in Fig. 3.16d. It drops from 0.25 at 10^{19} eV to 0.1 at 10^{20} eV. It is of similar size for proton and iron events. The evolution of the reconstructed R_μ with energy is depicted in Fig. 3.16e. There is no significant change with energy for both proton and iron showers, while the average values are $R_\mu \approx 1.05$ for proton and $R_\mu \approx 1.45$ for iron.



(a) **(b)**

Figure 3.15: Biases and resolutions of the reconstructed X_{max} using a universality reconstruction of all quantities but the primary energy. The figures are further described in the text.

SD-750 simulations with post-LHC interaction models

The events in the continuous shower library described in Appendix E.4 were simulated and reconstructed with an ideal SD-750 reconstruction [131]. The universality reconstruction developed within this work was employed. The hadronic interaction model QGSJET-II.04 was used to simulate these events. This makes it particularly interesting to study biases in the context of new interaction models. Different plots on the reconstruction bias of X_{max} are included in Fig. 3.17. In this reconstruction, the primary energy was fixed to the true value, all other quantities were fitted. The bias is shown as a function of energy and for different primaries in Fig. 3.17a. There is a negligible bias above primary energies of $10^{18.5}$ eV and a larger bias below. The signal model is not valid below this energy (see Section 2.3), leading to the observed reconstruction biases. On the other hand, the universality reconstruction yields an unbiased estimate of X_{max} at energies above $10^{18.5}$ eV, even for simulations with the new interaction models. There is a distinct dependence with zenith angle in Fig. 3.17b. This is partially due to dependencies on the number of candidate stations shown in Fig. 3.17c and Fig. 3.17d. However, there seems to be an issue with the angular dependence that is not present in reconstructions of QGSJET-II.03 simulations. This indicates that the time model should be re-investigated using new simulations. It should be noted that a large number of candidate stations above 20 is required to get an unbiased estimation of X_{max} (compared to the case of the SD-1500 reconstruction discussed in the previous section). One factor might be the fraction of the muonic with respect to the total signal, which is smaller at lower energies. As the muonic signal drives the estimation of X_{max}, this might cause a larger required number of stations.

Saturated events

The issue with events that include low-gain saturated stations is mostly a geometrical one. The core is located very close to a central station and the radial distances of the other stations are very similar, as are their signals and relative signal contributions. Overall, there is less information content to be used for the reconstruction than in the case of normal events. This results in a larger resolution of all reconstructed quantities. Thus, it is essential to exploit the information of the central saturated station in the shape fit. This is done

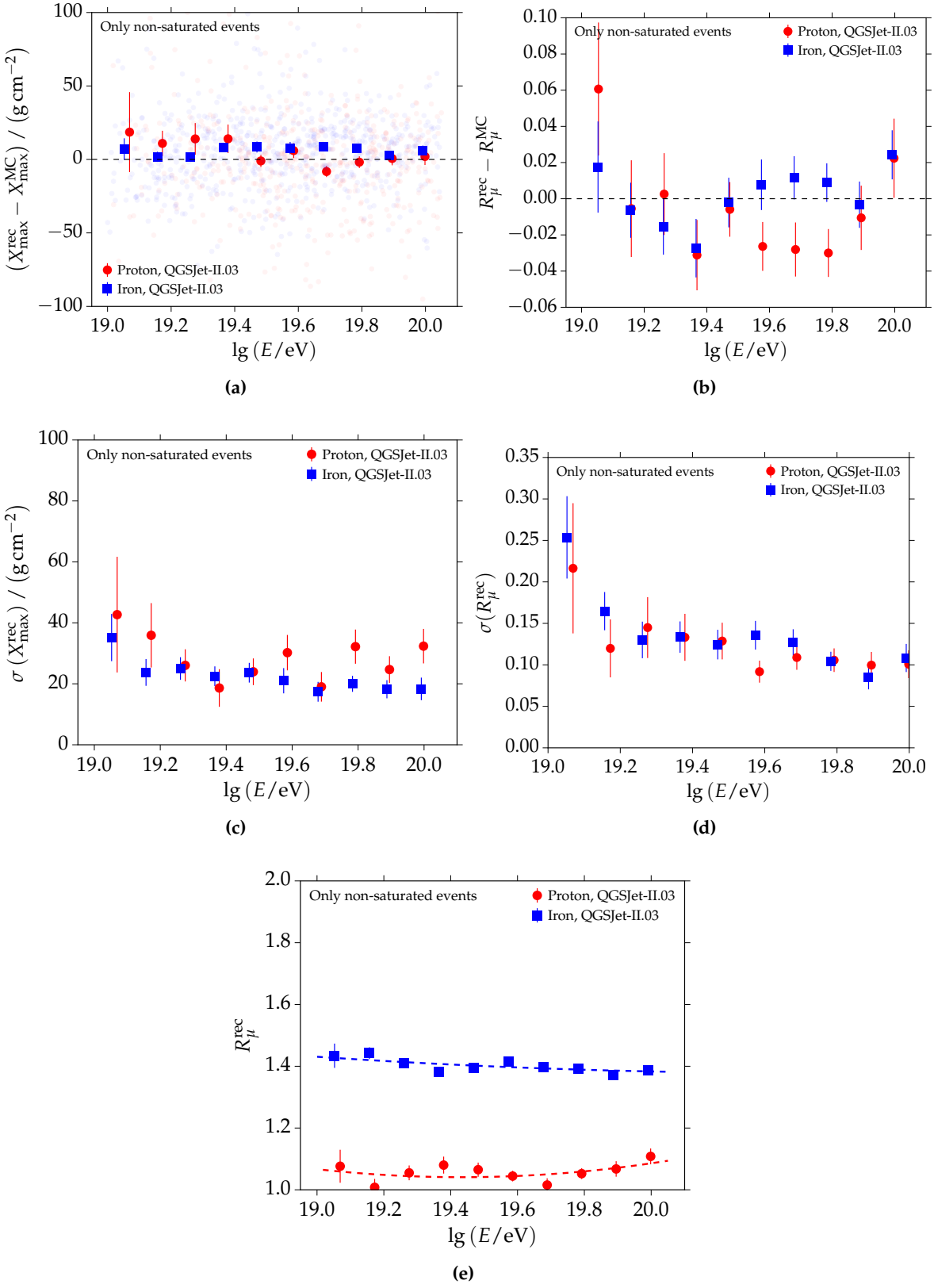


Figure 3.16: (a) - (d) Biases and resolutions of the reconstructed X_{max} and R_μ using a universality reconstruction in which only the energy is fixed to the result from the standard reconstruction. All other parameters are fitted at the same time. (e) Evolution of R_μ as a function of energy and for different primary species proton and iron, simulated with QGSJET-II.03.

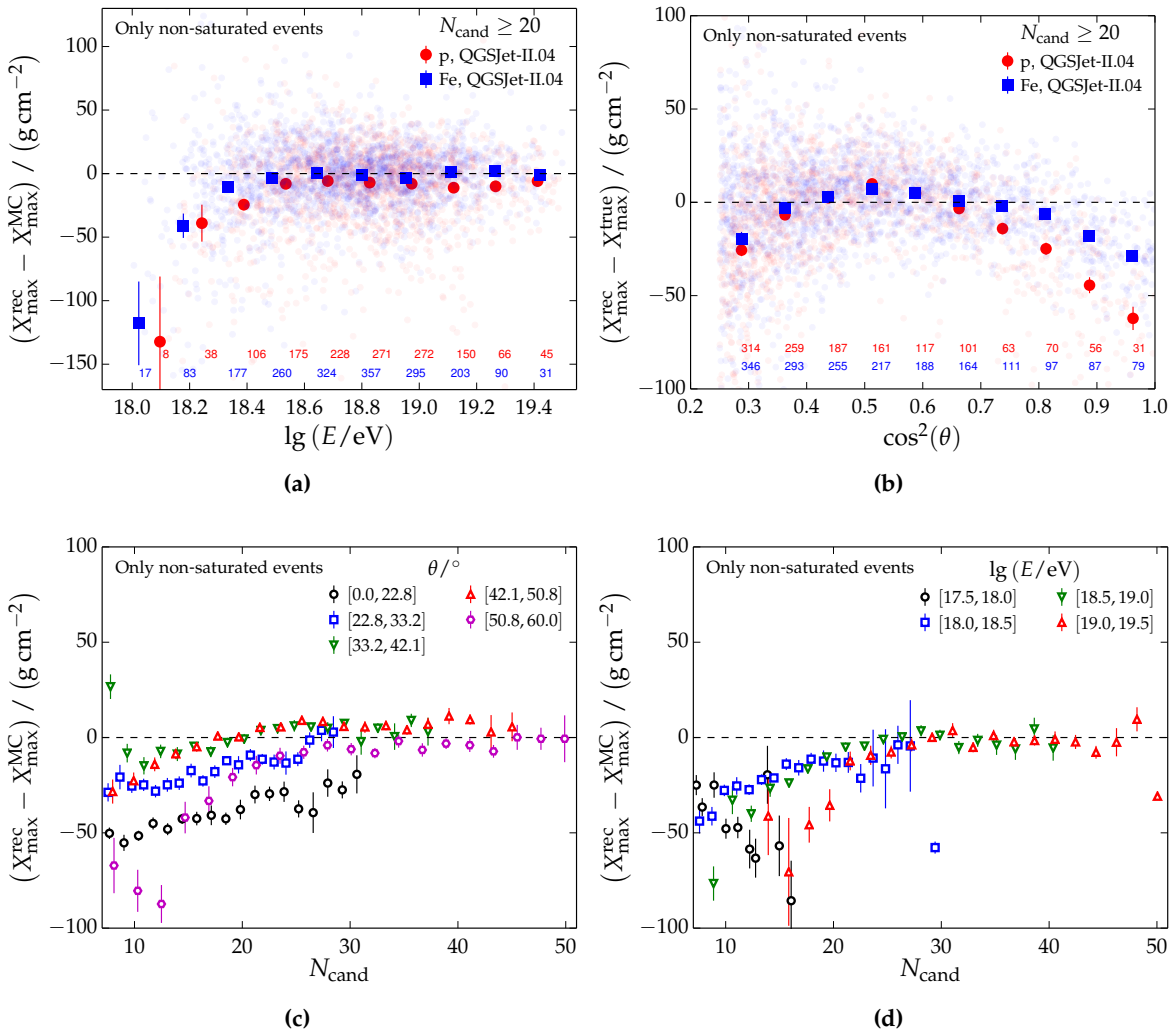


Figure 3.17: Biases and resolutions of the reconstructed X_{max} applying the universality reconstruction to simulated SD-750 events.

with a partial shape fit to the non-saturated part of the total particle traces (as described in Section 3.4.1). This greatly improves the resolution in the reconstruction with respect to not using the saturated station. The impact on the resolution of X_{max} from this change is on the level of $30\,\text{g cm}^{-2}$. Details on biases and resolutions when reconstructing saturated events are given in Section 3.4.6.

3.4.6 An iterative reconstruction method

As part of this work, an iterative reconstruction method was developed in order to improve the resolution of reconstructed quantities and to allow an unbiased reconstruction at a smaller number of candidate stations. The standard reconstruction described in Section 3.3 is run prior to the universality reconstruction. The subsequent universality reconstruction steps of the iterative procedure are as follows:

1. Estimation of the core position and energy while X_{max} and R_μ are kept fixed to true or model values. The geometry is fixed to the true geometry for simulations or the one obtained from the standard reconstruction for data. Only the signal likelihood is used. The energy is constrained to $\pm 15\,\%$ of the true or the SD energy.

2. Repeat 1) with updated model predictions of X_{max} and R_μ given the new estimate of the primary energy E. The energy is constrained to $\pm 15\,\%$ of the previous result.

3. The relative core time t_c is fit together with X_{max}, while all other quantities are fixed (in a first step, only the relative core time is fit).

4. R_μ, energy and core position are fixed, and X_{max}, t_c and the geometry are fit together. The geometry is weakly constrained to previous fit values (with angular standard deviations of $1.5^\circ \approx 3\,\sigma$).

5. R_μ is fit while all other quantities are fixed to previous fit results. Only the signal likelihood is used.

6. X_{max} is fit together with the shower geometry. All other quantities are fixed to previous results. The geometry is moderately constrained to previous values (with angular standard deviations of 1.0°).

7. The energy is re-fitted after fixing all other quantities to previous fit results. Both the signal and time likelihoods are used.

8. R_μ is re-fitted after fixing all other quantities to previous fit results. Only the signal likelihood is used.

Normal distributions with the given standard deviations are used for all constraints. In reconstructions of data, the model for $\langle X_{\mathrm{max}} \rangle(E)$ is taken from [21], the current Auger result from measurements with the FD. It is given by

$$\langle X_{\mathrm{max}} \rangle(E) = \begin{cases} X_{\mathrm{max},0} + (\lg(E/\mathrm{eV}) - \lg(E_0/\mathrm{eV}))\, D_{10}^0 & ; E < E_0 \\ X_{\mathrm{max},0} + (\lg(E/\mathrm{eV}) - \lg(E_0/\mathrm{eV}))\, D_{10}^1 & ; E \geq E_0\,. \end{cases} \tag{3.17}$$

$\lg(E_0/\mathrm{eV}) = 18.27 \pm 0.04$ is the energy at which the elongation rate changes. The other parameters are $X_{\mathrm{max},0} = (746.8 \pm 2.1)\,\mathrm{g\,cm^{-2}}$, $D_{10}^0 = (86.4 \pm 5.0)\,\mathrm{g/cm^2/decade}$ and $D_{10}^1 = (26.4 \pm 2.5)\,\mathrm{g/cm^2/decade}$. Only statistical uncertainties are given here, systematic ones are listed in [21].
The model for $\langle R_\mu \rangle$ is detailed in Section 5.2.

The iterative fit enables a simultaneous reconstruction of the primary energy E, the depth of shower maximum X_{max} and the relative muon content R_μ. This is possible through the use of constraints and parameterizations. Note that X_{max} and R_μ are not constrained during the reconstruction procedure. The performance of the reconstruction is again investigated with Monte Carlo simulations of the SD-1500 (using the continuous library described in Appendix E.3). Asymmetries in the LDFs and time distributions of particles are taken into account in the signal and time models. As such, the universality reconstruction enables an unbiased estimation of the shower core as shown in Fig. 3.18a for events with zenith angles below 38° and for events with larger angles in Fig. 3.18b. The radial extent represents the bias from the true core position. Individually reconstructed core positions are depicted with red dots, and the unbiased average is shown with the red square. In contrast, the LDF fit of the standard reconstruction (Section 3.3.2)) leads to biased core positions as shown with the blue markers. This bias is present because azimuthal signal asymmetries are not taken into account in the fit of the LDF. The resulting average bias increases with zenith angle. All events are included in these plots, also events with saturated stations. The resolution of the reconstructed core position is depicted in Fig. 3.18c for events with zenith angles below 38° and in Fig. 3.18d for the other events. The core resolution decreases strongly with energy and slightly with zenith angle. While the universality and

the standard reconstruction yield a comparable core resolution for events with low zenith angles, the universality reconstruction performs better at large zenith angles. The angular resolution of both reconstructions is compared in Fig. 3.18e. On average, the resolution is on the order of 0.5°, decreasing to about 0.3° at highest zenith angles. For most zenith angles, the universality reconstruction yields a slightly larger resolution. The difference at high energies is mostly caused by saturated events. Excluding these events leads to much more similar resolutions at the highest energies. The angular resolution is calculated as following:

$$\sigma_{\text{ang}} = \frac{3}{2\sqrt{2}} \sqrt{\sigma^2(\theta) + \sin^2(\theta)\,\sigma^2(\phi)}\,, \tag{3.18}$$

with the resolutions of the reconstructed zenith and azimuth angles $\sigma(\theta)$ and $\sigma(\phi)$.

Further results on accuracies and biases are depicted in Fig. 3.19. The bias in the reconstructed energy is shown as a function of $\lg E$ in Fig. 3.19a and as a function of the number of candidate stations in Fig. 3.19b. When averaging over the two primaries, there is a small positive bias of 2 % below $10^{19.2}$ eV and no bias at higher energies. Only non-saturated events are included in these profiles. The bias in the reconstructed X_{max} is shown as a function of energy, and for the primary species proton and iron in Fig. 3.19c. The same quantity is plotted as a function of zenith angle in Fig. 3.19d. Overall, there is a small bias of $-5\,\text{g}\,\text{cm}^{-2}$, independent of energy, zenith angle, and primary species. This value represents the overall median bias (the median values and their $1\,\sigma$ uncertainties are depicted with lines and dark contours; the mean values are represented with markers and error bars). It is apparent that there are more positive outliers at lowest energies. These pull the mean values up while the median values stay unbiased. This is a consequence of the constrained fit of events with a low number of candidate stations. The absolute bias in R_μ is shown in Fig. 3.19e as a function of energy. It is negligible at high energies and 3 % below 10^{19} eV. Overall, the iterative method yields reconstructions with minimal biases. No restriction on the number of candidate stations is required, but upward fluctuations at low energies need to be considered, e.g. by using median profiles instead of mean profiles, or by applying an outlier rejection. A method to reject those outliers for the data analysis is discussed in Chapter 5. The resolution of X_{max} is shown in Fig. 3.20a. On average, it is $50\,\text{g}\,\text{cm}^{-2}$, $30\,\text{g}\,\text{cm}^{-2}$ and $18\,\text{g}\,\text{cm}^{-2}$ at $10^{18.5}$ eV, 10^{19} eV and 10^{20} eV. Depending on energy, the resolution of proton showers is up to $20\,\text{g}\,\text{cm}^{-2}$ larger. The energy resolution is depicted in Fig. 3.20b as a function of energy and for the primaries proton and iron. At 3×10^{18} eV, the resolution ranges from 15 % to 20 % depending on the primary. It drops to under 10 % for energies above 10^{19} eV. The resolution of the reconstructed R_μ is given as a function of energy in Fig. 3.20c. It is 0.08, 0.05 and 0.03 at $10^{18.5}$ eV, 10^{19} eV and 10^{20} eV. The resolutions are obtained from robust estimates of the standard deviation (see Appendix F.2.1). The uncertainties of the standard deviations themselves are obtained from bootstrapping the calculation of the standard deviation [132, 133].

Problems with the reconstruction of saturated events were mentioned in Section 3.4.5. For the iterative method, an issue occurs during the first two fit stages. The estimation of the energy from the signal LDFs is difficult due to the degenerate radial distribution of stations. Due to that, the energy is fixed to the true value or the value from the standard reconstruction. The energy is still fit at stage 7). This change enables a more robust estimation of X_{max}. Results on biases are depicted in Fig. 3.21. The bias in the reconstructed energy is depicted in Fig. 3.21a as a function of the Monte Carlo energy. While the median residual is unbiased over the whole energy range, the mean residual is biased below 10^{19} eV, because there are more outliers with overestimated energies. As depicted in Fig. 3.21c, the reconstructed X_{max} is biased below $10^{19.5}$ eV. Also the bias in the reconstructed R_μ increases strongly below 10^{19} eV, as visualized in Fig. 3.21e. The corresponding robust standard devi-

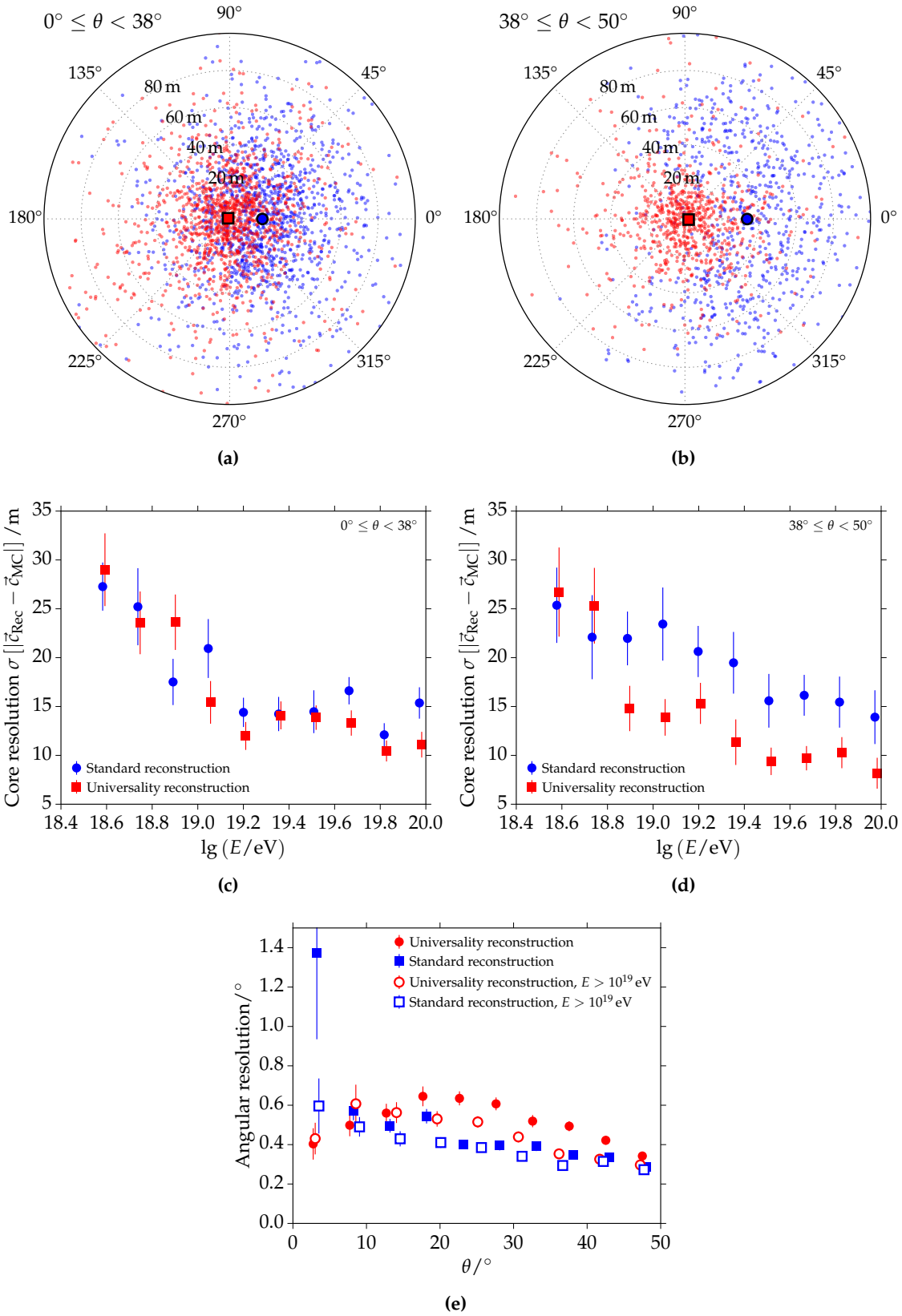


Figure 3.18: (a), (b) Comparison of reconstructed core positions using the universality and the standard reconstruction. The two plots contain events in different zenith angle intervals. On average, the universality reconstruction yields an unbiased core position, while the standard reconstruction gives a biased result (see text). (c), (d) The resolution of the reconstructed core position for the two different reconstruction methods and as a function of energy. (e) The angular resolution as a function of zenith angle and for different energy ranges.

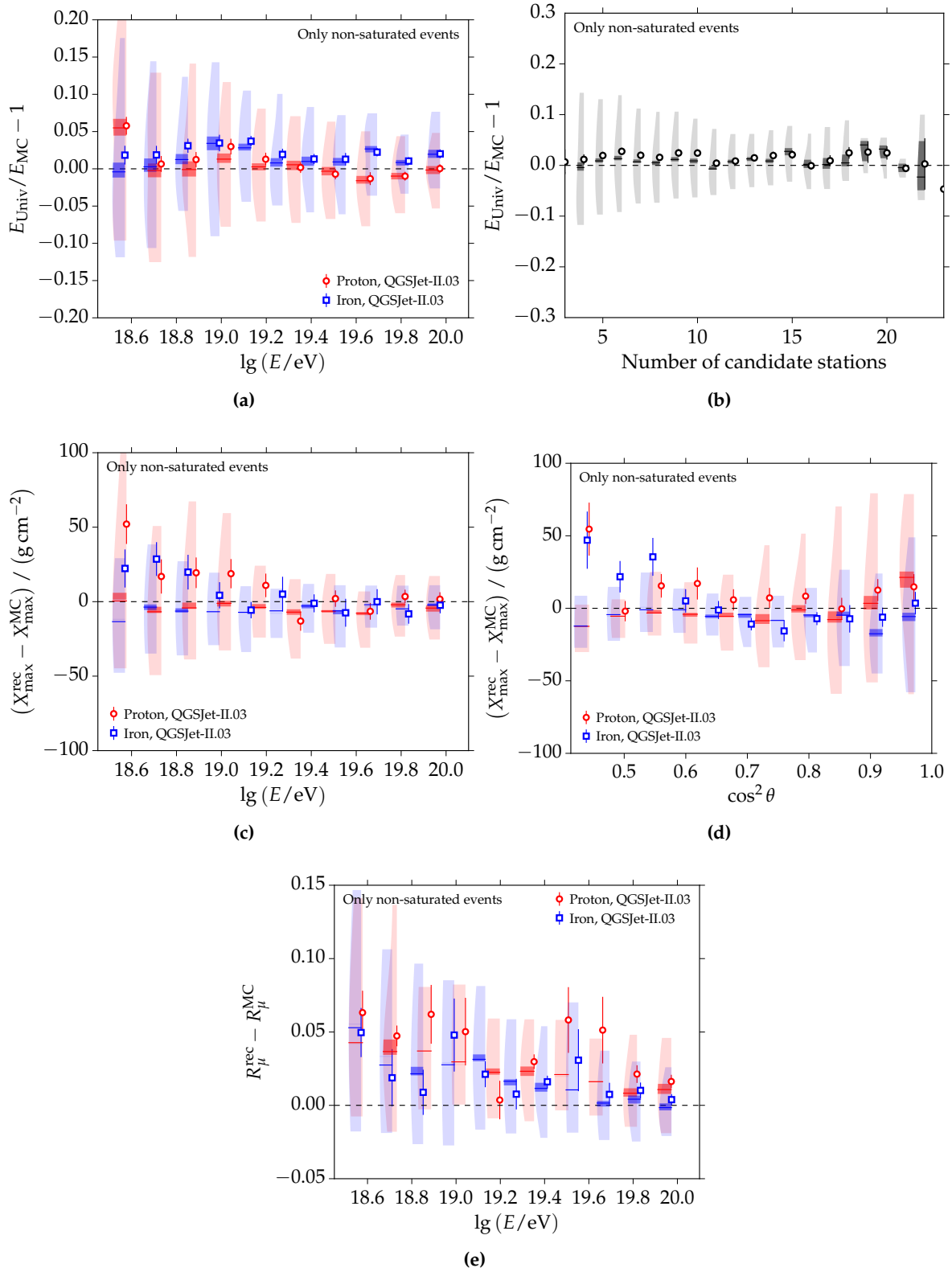


Figure 3.19: Biases of reconstructed quantities using the iterative universality fit. Shown are reconstructions of non-saturated events.

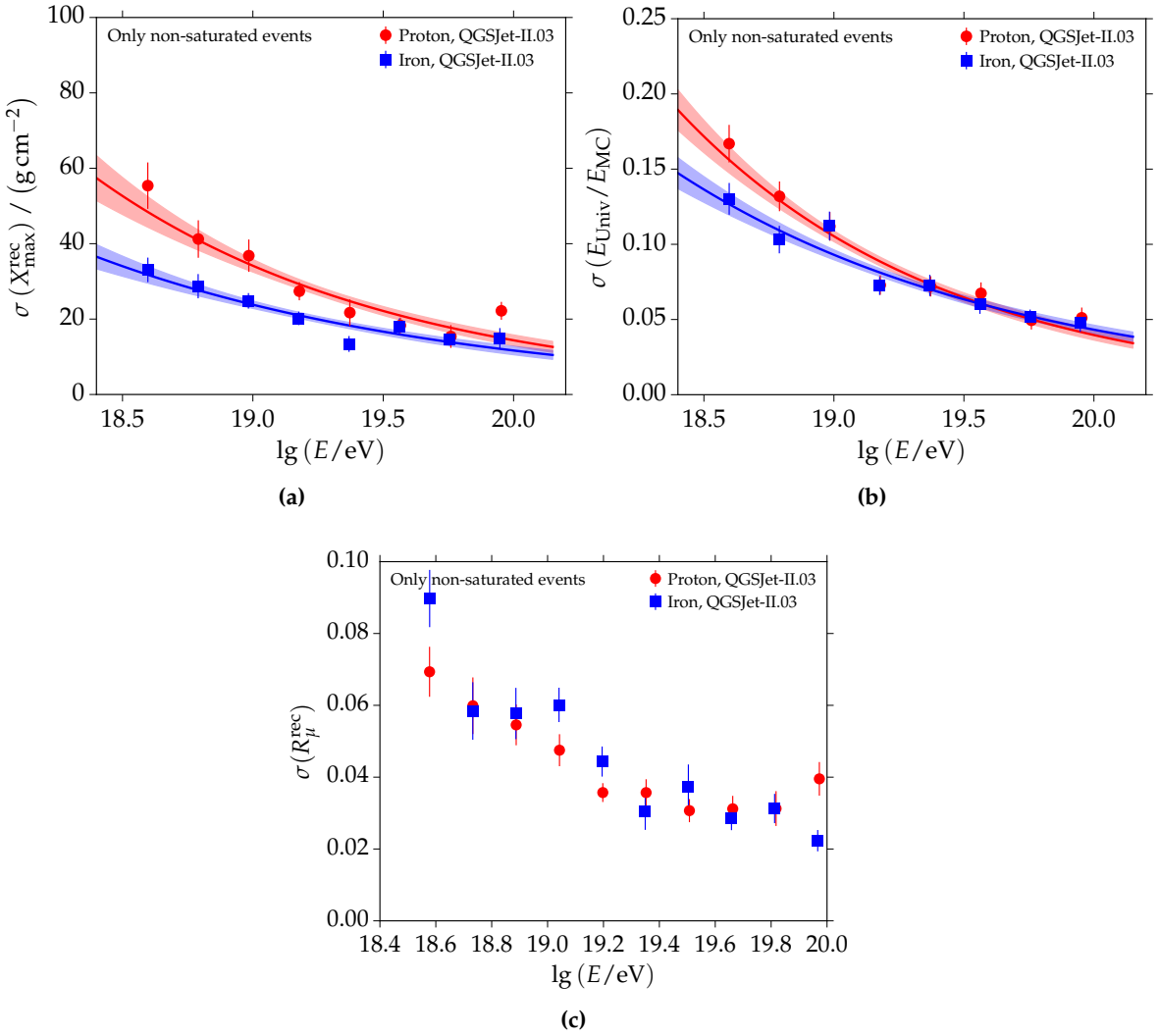


Figure 3.20: Resolution in reconstructed quantities using the iterative universality fit. Shown are reconstructions of non-saturated events.

ations are depicted in Fig. 3.22. They increase strongly below 10^{19} eV and are larger than the resolutions obtained for non-saturated events in Fig. 3.20. These results imply that, for the data analysis, saturated events should only be used above energies of $10^{19.5}$ eV, or handled very carefully below this energy.

3.4.7 Reconstruction efficiency

The reconstruction efficiency is defined as the ratio of the number of events that were successfully reconstructed with respect to all events. As the universality reconstruction depends on the standard reconstruction, which is run before it, the efficiency $p(\mathrm{Univ}|\mathrm{SD})$ gives the conditional probability that the universality reconstruction succeeds if the standard one did. This efficiency is plotted in Fig. 3.23a for the iterative reconstruction and in Fig. 3.23b for the classic reconstruction. The figures include all events that pass the selection criteria as specified in Section 4.1. The reconstruction efficiency is 100 % at all energies. Note that events that are registered as outliers in later analysis steps are not counted as

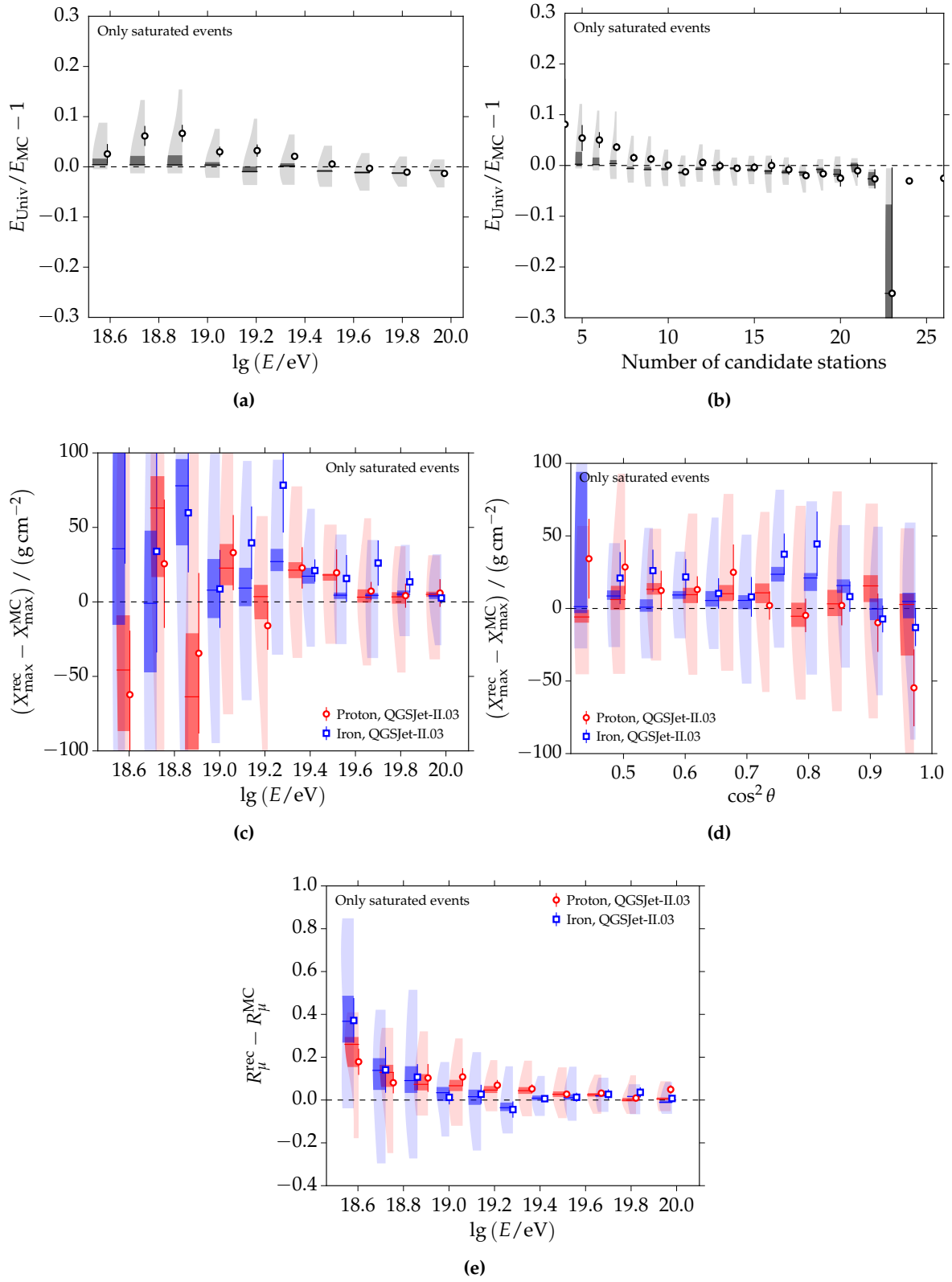


Figure 3.21: Biases of reconstructed quantities using the iterative universality fit. Shown are reconstructions of saturated events.

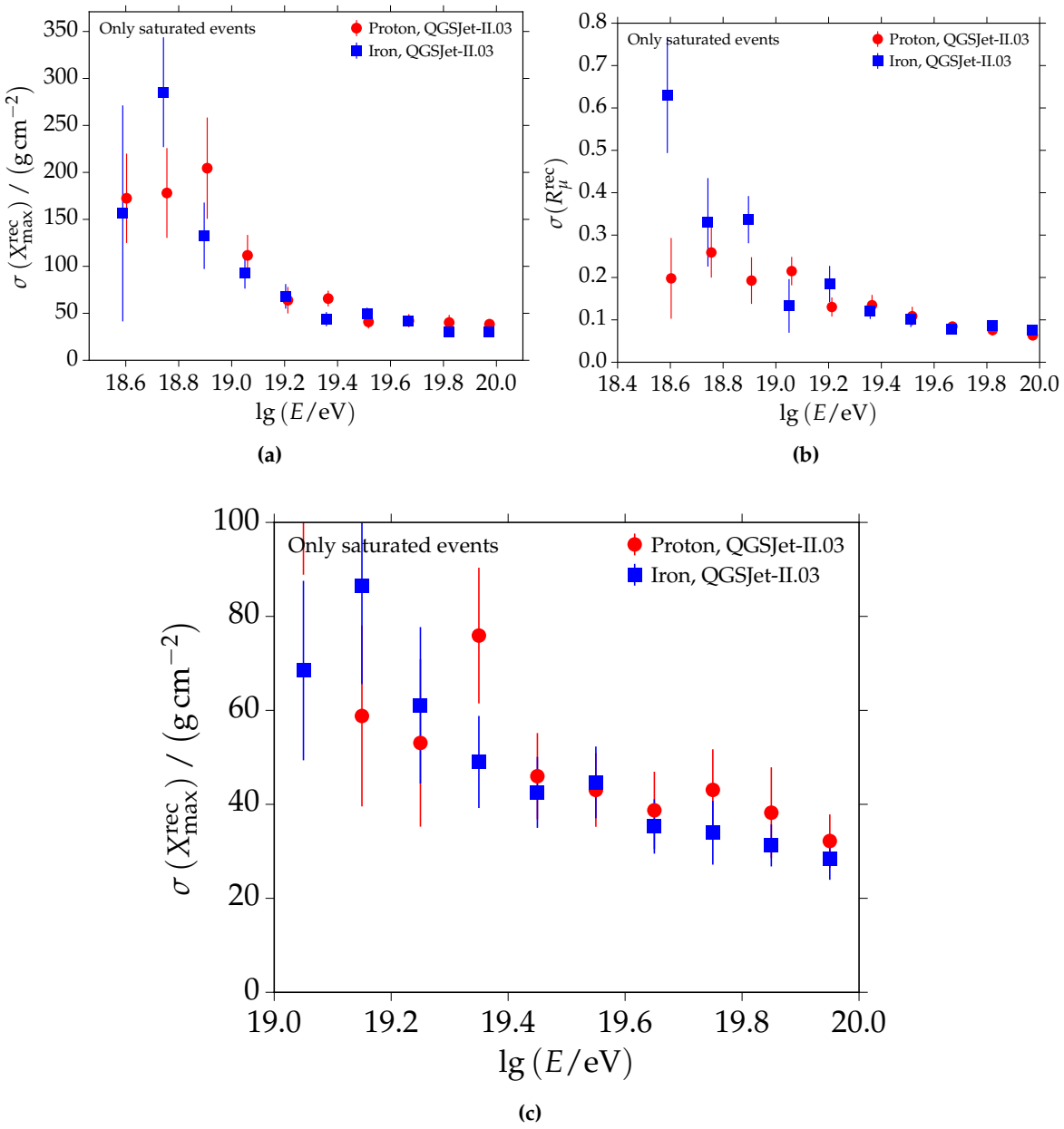
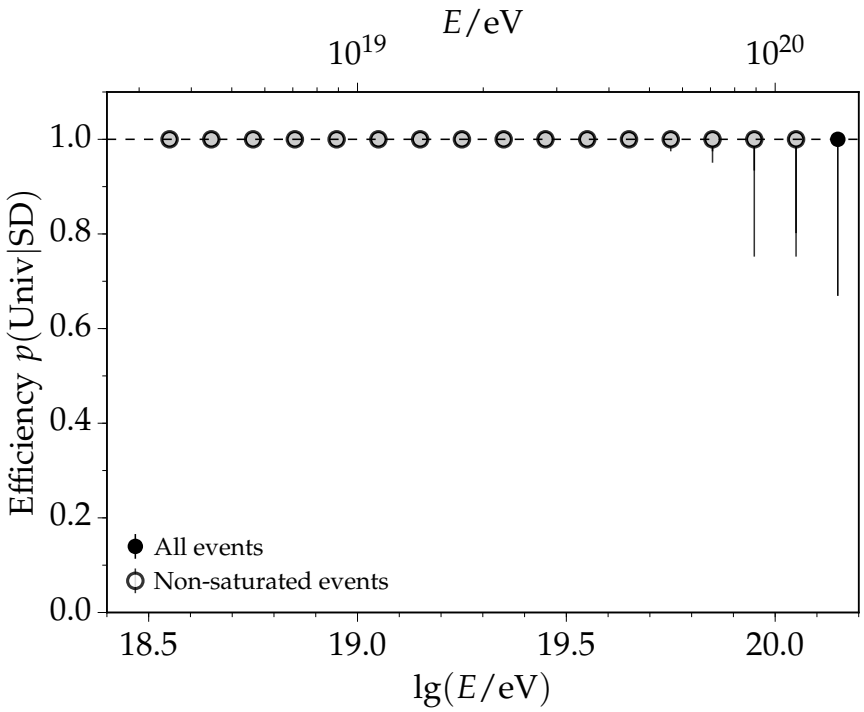


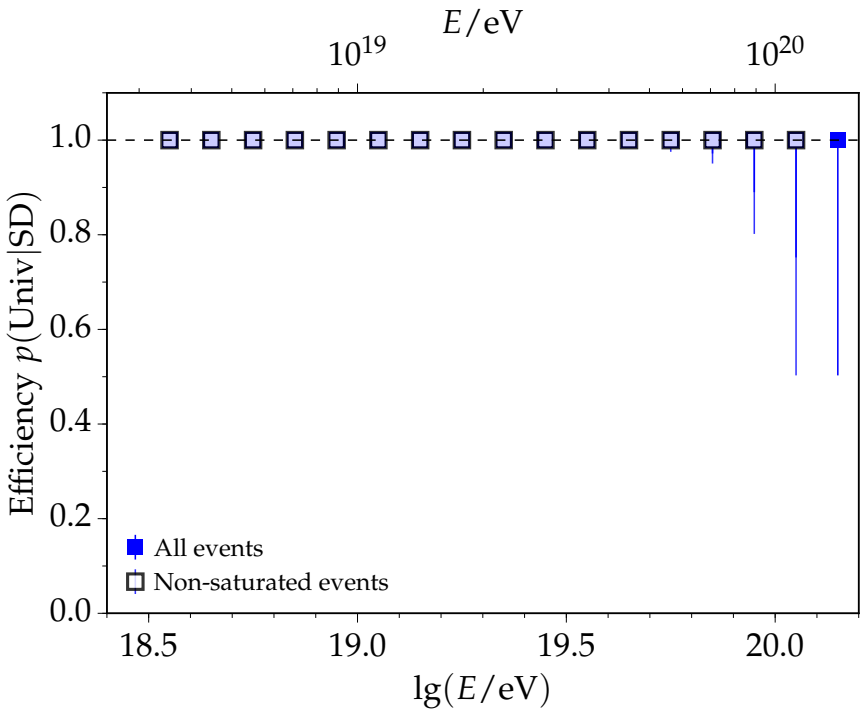
Figure 3.22: Resolution in reconstructed quantities using the iterative universality fit. Shown are reconstructions of saturated events.

failed reconstructions in these figures. The efficiencies for the standard reconstruction of SD-750 and SD-1500 events are detailed in Section 4.4.

3.4.8 Comparison of signals



(a)



(b)

Figure 3.23: Efficiency of the universality reconstruction. Given is the conditional probability that the reconstruction succeeds after a successful standard reconstruction. (a) Iterative universality reconstruction. (b) Classic universality reconstruction (only the energy is fixed).

CHAPTER 4

The flux of UHECRs

The analyses of the flux (or energy spectrum) of events measured with the SD-1500 and SD-750 are presented in this chapter. The former are grouped into vertical events with zenith angles below 60° and inclined events with zenith angles above 60°. Due to event statistics, the current analysis of SD-750 events is limited to vertical events. Air shower physics changes significantly with increasing zenith angle. Due to the increasing atmospheric overburden, nearly only muons reach the ground in very inclined showers, while the electromagnetic component is significant and important to consider in the analysis of vertical showers. Due to this, separate reconstruction algorithms were developed for the analysis of vertical and inclined events. The standard reconstruction of vertical events is discussed within this work in Section 3.3. The reconstruction of inclined events is not discussed in this work, a number of references are [75, 134–136]. However, results based on the latest reconstruction of inclined events are presented and compared to results of this work.

Vertical events recorded with the SD-1500 and SD-750 arrays are analyzed in a similar manner: The first step involves the fit of the shower geometry and lateral distribution as described in Section 3.3. This is only possible for events with sufficiently large energies, because at least three triggered stations are required for a reconstruction of the arrival direction and the impact point on ground. For each event, the expected signal at an optimal distance to the reconstructed shower core is taken as an estimator of the primary energy. The optimal distance is the distance to the core at which the lack of knowledge of the lateral distribution of showers has the smallest impact on the estimation of the shower size. It is the point of the LDF that is best constrained by surrounding stations. Therefore, it strongly depends on the distance between stations (see [126] for a deeper discussion). The grid spacing of 750 m (1500 m) of the SD-750 (SD-1500) leads to an optimal distance of 450 m (1000 m). I then correct for the zenith angle dependence of the estimated shower sizes by employing the Constant Intensity Cut method (see Section 4.5). Using a set of hybrid events from both SD and FD, the corrected SD shower sizes are calibrated to primary energies from a calorimetric measurement with the FD. This energy calibration procedure is detailed in Section 4.7. The resulting energy spectra are presented in Section 4.8. Motivated by large differences in the measured flux of UHECRs above 10^{19} eV, a study of a possible dependence of the flux on the incoming direction of particles is given in Section 4.9. The statistical method of forward-folding is used to correct for event migrations due to the finite detector resolution and shower-to-shower fluctuations. I describe this procedure and derived results in Section 4.10. Comparisons to other spectra in the energy range of

interest are presented as well. Energy spectra derived from reconstructions using shower universality are discussed in Section 4.11. This is a novel comparison with important implications for future analyses. A new method to combine different fluxes is explained in Section 4.12. Statistical and systematic uncertainties of the different measurements are taken into account. Related topics, e.g. a toy validation of the forward-folding method, are discussed in Appendix B.7.1.
The following sections describe details of the flux measurement including events with zenith angles below 60°.

4.1 Event selection

In addition to being successfully reconstructed (Section 3.3), all recorded events are required to pass the 6T5-prior selection. This necessitates that the station with the largest signal in an event is surrounded by six active and functional stations at the time of triggering. Though, not all of these stations need to have triggered. The criterion is visualized in Fig. 4.1a. It should be distinguished from the 6T5-posterior criterion used in the analysis of inclined events. In that case, the station closest to the reconstructed shower core is required to be surrounded by six functional and active stations. The 6T5-prior and 6T5-posterior criteria are different only for events with large zenith angles. For smaller zenith angles, the hottest station is almost always the station closest to the reconstructed shower core.

Other quality cuts on the individual station and PMT level were established in the Offline software as part of this work. To be fully consistent with the exposure calculation (as discussed in Section 4.2), stations that are inactive - according to T2 information - at the time of the event are rejected in the reconstruction of events. This is highly relevant, because bad stations (inactive) and silent (functional but non-triggered) stations are important to distinguish, both for the reconstruction and for the calculation of the exposure. In addition to these quality cuts on the station level, individual PMTs are possibly flagged as bad. By default, the integrated signal of a station is obtained from the average VEM-trace of all three PMTs. In the event of one or two non-functional PMTs, these are just omitted in this process. There is no corresponding incident in the current data, but if all three PMTs were bad (and the station itself was not already flagged as bad), the station would be flagged as a bad station and not taken into account in the reconstruction. During the reconstruction of events, the quality cuts on the PMT level are applied with the Offline module SdPMTQualityChecker, which I developed for this thesis.

The event rate and exposure are constantly monitored and analyzed. If there are large discrepancies between the actual and expected rate of events or if the SD is particularly unstable within a certain time period, e.g. during thunderstorms, the corresponding time period is defined as a bad period and neglected in the analysis. This also includes incidents like the communication crisis in 2009 [137]. Furthermore, a few individual stations are manually rejected because their software or hardware was replaced with elements to test. All of the mentioned cuts are applied equivalently when the exposure is calculated in Section 4.2. Detailed information about data acquisition and quality selection are discussed in [33, 138, 139].

The configuration files that I used for the selection of events in this work are listed in Appendix D.1.2. If not stated otherwise, I am considering data from the time period 1. January 2004 until 31. December 2014. The number of events and other detailed information can be extracted from the flux tables in Appendix C. The events were reconstructed with Offline (distributed as observer production v10r0 [140]). The raw Auger data files were recently reprocessed to fix bugs in the merging of events. These problems resulted in a loss

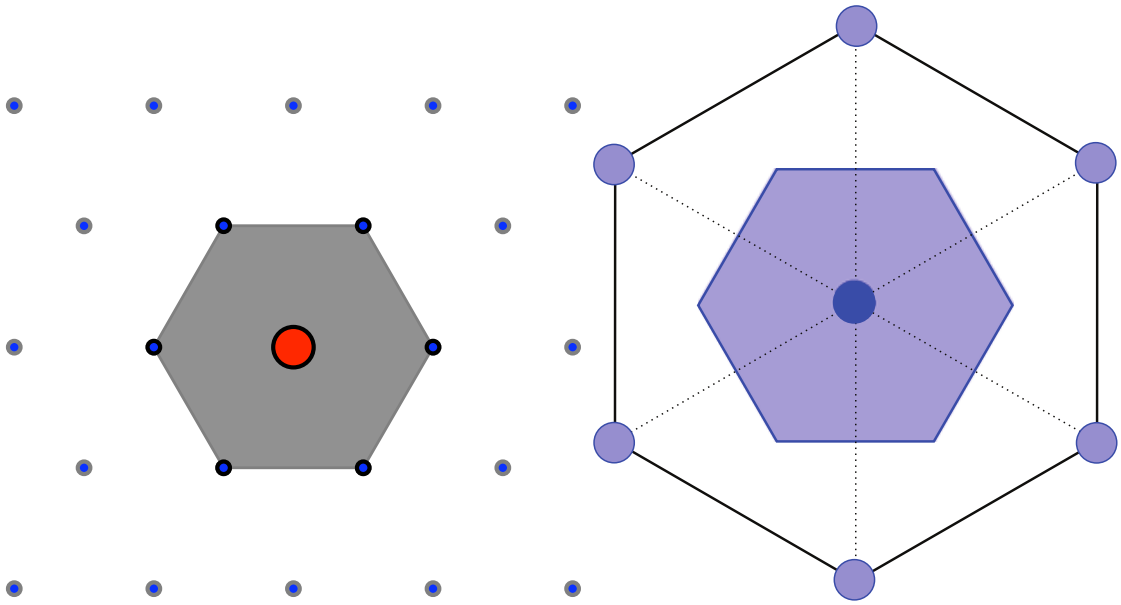


Figure 4.1: (a) Visualization of the 6T5 criterion required for events contributing to the analysis of the flux of UHECRs presented in this work. (b) Depiction of the elementary hexagon area corresponding to active 6T5 cells.

of events since 2013. I am using the updated input files for this work. The internal version number for these files is v1r9.

4.2 Geometrical exposure

A necessary ingredient to determine the flux of UHECRs is a precise knowledge of the experimental exposure (the time-integrated aperture) for both measurements with the SD-750 and SD-1500. The calculation is based on an integration of available elementary hexagons as a function of time. This follows the 6T5 criterion discussed in the previous section. A detailed description is given in [33]. The knowledge of the exact placement of each detector together with the T2 files allows one to calculate the number of active elementary hexagons at a certain second multiplied with the hexagonal cell area A_{cell} (as depicted in Fig. 4.1b). This gives the instantaneous aperture $\mathrm{d}A_{\text{6T5}}$ after taking into account the projection in the shower plane. For the cell areas we have

$$A_{\text{cell}} = \frac{\sqrt{3}}{2} d^2 \tag{4.1}$$

$$A_{\text{cell}}^{750\,\text{m}} = \frac{\sqrt{3}}{2} (0.75\,\text{km})^2 = 0.487\,\text{km}^2 \tag{4.2}$$

$$A_{\text{cell}}^{1500\,\text{m}} = \frac{\sqrt{3}}{2} (1.5\,\text{km})^2 = 1.949\,\text{km}^2.$$

The instantaneous effective cell area reads

$$\mathrm{d}A_{\text{6T5}} = A_{\text{proj}} = A_{\text{cell}} \cos\theta\, \mathrm{d}\Omega = -A_{\text{cell}} \cos\theta\, \mathrm{d}\cos\theta\, \mathrm{d}\phi. \tag{4.3}$$

Integration over solid angle leads to

$$\begin{aligned} A_{\mathrm{6T5}} &= A_{\mathrm{cell}} \int_0^{2\pi} \mathrm{d}\phi \int_{\cos\theta_{\mathrm{max}}}^{\cos\theta_{\mathrm{min}}} \mathrm{d}\cos\theta \cos\theta \\ &= A_{\mathrm{cell}}\, \pi \left(\sin^2\theta_{\mathrm{max}} - \sin^2\theta_{\mathrm{min}}\right) \\ A_{\mathrm{6T5}}^{750\,\mathrm{m}} &= 1.0269\,\mathrm{km}^2\,\mathrm{sr} \\ A_{\mathrm{6T5}}^{1500\,\mathrm{m}} &= 4.5912\,\mathrm{km}^2\,\mathrm{sr}, \end{aligned} \tag{4.4}$$

using the respective zenith angle ranges for SD-750 and SD-1500. The total exposure is then obtained by integrating over time:

$$\mathcal{E} = \int_{t_0}^{t_1} \mathrm{d}t\; A(t), \tag{4.5}$$

with $A(t) = N(t) A_{\mathrm{6T5}}$. $N(t)$ is the number of active hexagons in the time period $[t - \mathrm{d}t, t + \mathrm{d}t]$. In T2 files, the activity of each station is recorded at each second. Thus, the calculation of the exposure simplifies to

$$\mathcal{E} = \sum_i A(t_i) \tag{4.6}$$

with the sum running over each second in the time period of interest and $\Delta t = 1\,\mathrm{s}$ not explicitly stated.

As part of this and a previous work [127], I developed a software that performs the exposure calculations for both SD-750 and SD-1500. The results are compatible with the official acceptance calculation of Auger [141] to less than 0.1 %. The resulting exposure for SD-1500 in the time period 1. January 2004 - 31. December 2014 is $42\,520\,\mathrm{km}^2\,\mathrm{yr\,sr}$. For SD-750 the value is $153.19\,\mathrm{km}^2\,\mathrm{yr\,sr}$, considering the time period 1. August 2008 - 31. December 2014; prior to the starting date, the SD-750 consisted of only one hexagon of stations and thus earlier data are not taken into account. From the variation of the energy spectrum with time and other consistency checks, the relative systematic uncertainties on the exposure calculation were found to be 5 % for the SD-1500 and 3 % for the SD-750 (see also Section 4.8.4). Plots of the resulting integrated exposure and average hexagons are shown in Fig. 4.2. The shaded areas in the background indicate bad periods that are removed during the calculation of the exposure. The SD array has a small tilt with respect to the reference ellipsoid representing Earth. The tilt of the array induces a sub-percent effect on the exposure calculation that can be relevant for studies, e.g. on the large scale distribution of events, that are very sensitive to small differences in the exposure [81]. For the purpose of this work, these effects are negligible and will be neglected.

A number of refined, ongoing analyses attempt to control the event rate as a function of time. The reasons for changes of the rate versus time are not fully understood. The recent progress is detailed in [142].

4.3 Corrections due to atmospheric conditions and the geomagnetic field

The density and distribution of secondary particles on ground are affected by varying atmospheric conditions and the geomagnetic field. The shower sizes S_{450} and S_{1000} used as energy estimators are thus indirectly affected and such are the energies inferred from these. These effects need to be taken into account in order to obtain an unbiased angular dependence of the flux of CRs as presented in Section 4.8. It is especially important for the search of a declination dependence of the energy spectrum in Section 4.9.

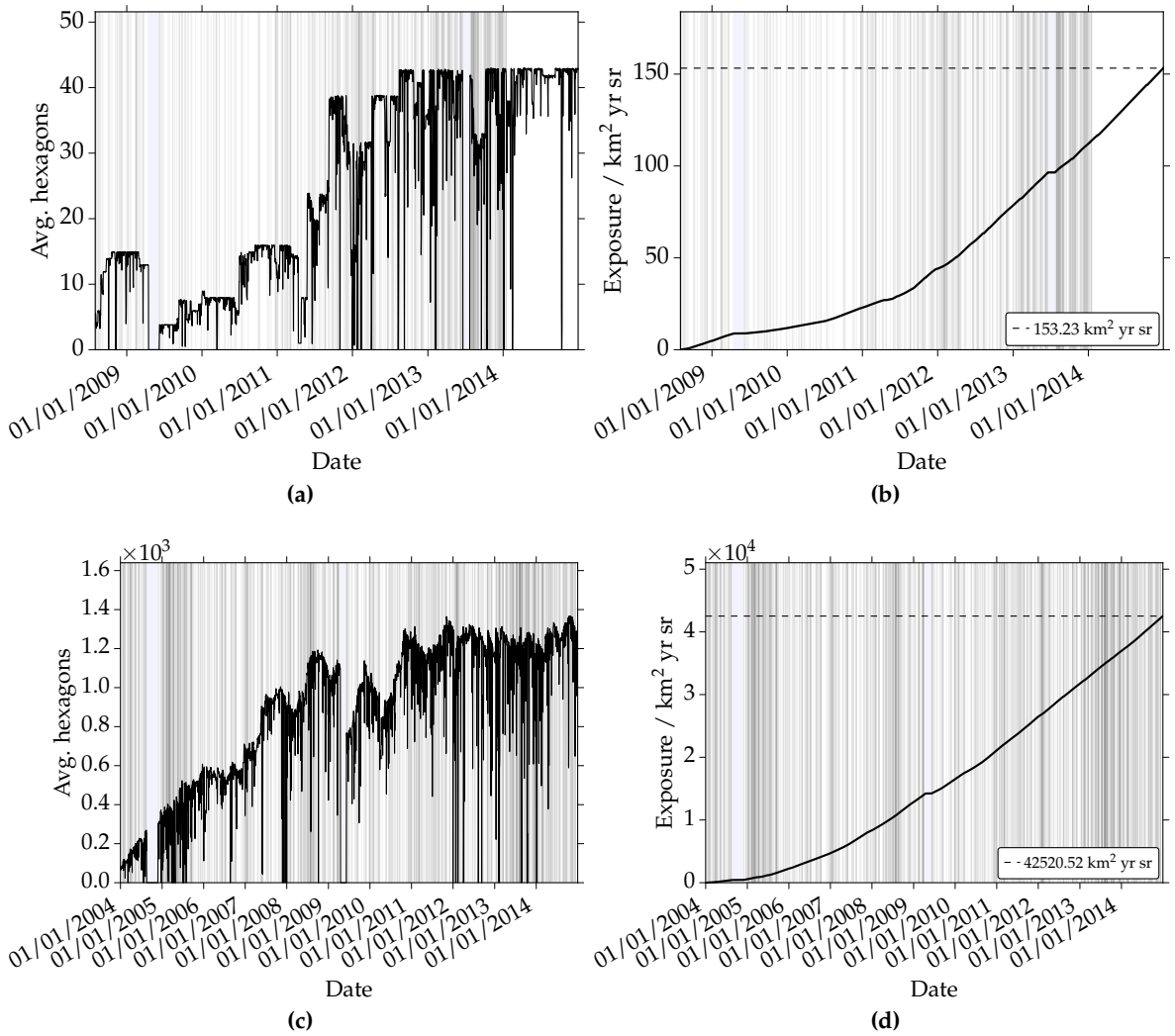


Figure 4.2: The average number of hexagons as a function of time for (a) SD-750 and (c) SD-1500. The integrated experimental exposure as a function of time for (b) SD-750 and (d) SD-1500. The shaded regions in the background of the plots represent bad time periods that are excluded in the analysis.

Atmospheric conditions affect the shower sizes due to two major effects. Firstly, the longitudinal development depends on the amount of traversed atmosphere, and this amount is determined by the air pressure as a function of height (above sea level). If the pressure is consistently larger than on average, the showers will effectively be older or in a later stage of development when they reach the ground. Secondly, the air density affects the Molière radius (the circle around the shower core that encloses 90 % of the energy content of the air shower) and thereby the lateral distribution of showers. These effects have been studied for the shower sizes S_{450} and S_{1000}. They are on the order of ± 1 %, depending mostly on the ground temperature and pressure at the time of the event or shortly before [44, 143–150]. For this work, I adapted the corrections as they are used in [63]. The current correction is zenith dependent and is based on an interpolation of atmospheric values collected at the observatory, in order to provide a valid calculation for each event. Typically, atmospheric conditions are monitored at least every few hours.

Charged secondary particles of air showers are deflected in the Earth's geomagnetic field, which changes the measured shower sizes on the ground. The influence on the

shower size S_{1000} was studied in [59] using air shower simulations. The geomagnetic field introduces a strong azimuthal modulation in S_{1000} with a rather small dependence on the zenith angle θ. The maximal amplitude of the change of S_{1000} is on the order of 2 %. The main modulation is with the azimuthal angle as this determines the angle between the particle direction and the direction of the geomagnetic field. The correction is minimal at an azimuth angle of around 90° as most of the particles propagate parallel to the geomagnetic field. The correction on S_{1000} as used within this work is implemented in Offline. The effect on S_{450} was not studied yet, and no correction is applied.

Figures of the geomagnetic and weather corrections are plotted in Fig. 4.3. Not accounting for the geomagnetic field leads to an effective overestimation of primary energies for particles coming from the southern sky. This is shown in Fig. 4.3c. It is mostly important for the search of a declination dependence in Section 4.9. On average, the amplitude of the weather corrections is largest in the southern winter and smallest in the southern summer, as depicted in Fig. 4.3d. The total correction is on the order of 2 %, with a slight dependence on zenith angle and no dependence on primary energy. This is visualized in Fig. 4.3e. The impact of shower size corrections on the attenuation and energy calibration parameters is discussed in the next two sections. An additional systematic uncertainty of 3.5 % is attributed to the flux measurements in Section 4.10. It stems from uncertainties in the weather and geomagnetic corrections. Details are discussed in e.g. [38].

4.4 Trigger efficiencies

Events have to pass the physics selection (T4 selection) as described in Section 3.1.3 to contribute to the analysis. Simulations are used to evaluate the probability that an event with a certain energy and zenith angle passes the physics selection. This involves the repeated simulation of air showers at different zenith angles and energies in the SD-750 and SD-1500 arrays. A counting of selected events with respect to the number of input events gives the trigger probability. These analyses have been performed for SD-750 data in [127] and SD-1500 data in [38].
The following function was adopted to describe the T4 efficiencies for both sets of data:

$$\epsilon(S,\theta) = \frac{1}{2}\left(1 + \mathrm{erf}\left(\frac{\lg S - a(\theta)}{b}\right)\right), \tag{4.7}$$

with the error function $\mathrm{erf}(y) := \frac{2}{\sqrt{\pi}} \int_0^y dx\, e^{-x^2/2}$ and the parameters a and b. The parameter a is a function of zenith angle and b has no dependencies:

$$a(\theta) = a_0 + a_1 \cos^2\theta + a_2 \cos^4\theta + a_3 \cos^6\theta. \tag{4.8}$$

Contours for different efficiencies are shown in Fig. 4.4. A table with the numerical values for the efficiency models is given in Table 4.1. At a certain zenith angle, the contour represents the lowest energy at which the desired efficiency is reached. While the models were parametrized as a function of the energy estimates, the current energy calibrations for data are used to calculate the energies in the plot (see Section 4.7). As such, the energy scale in data and not the Monte Carlo scale is shown. Contours for three different efficiencies ϵ of 0.9, 0.95 and 0.99 are plotted. The dashed lines represent the energy thresholds of full trigger efficiency of 3×10^{17} eV for the SD-750 and 3×10^{18} eV for the SD-1500. The non-trivial dependence of the efficiency with zenith angle is a result of the mixture between air shower attenuation and geometrical effects. Showers measured with the SD-750 develop higher up in the atmosphere and the average distance to the shower maximum is larger. This explains why the efficiency drops more quickly for larger zenith angles for SD-750 events

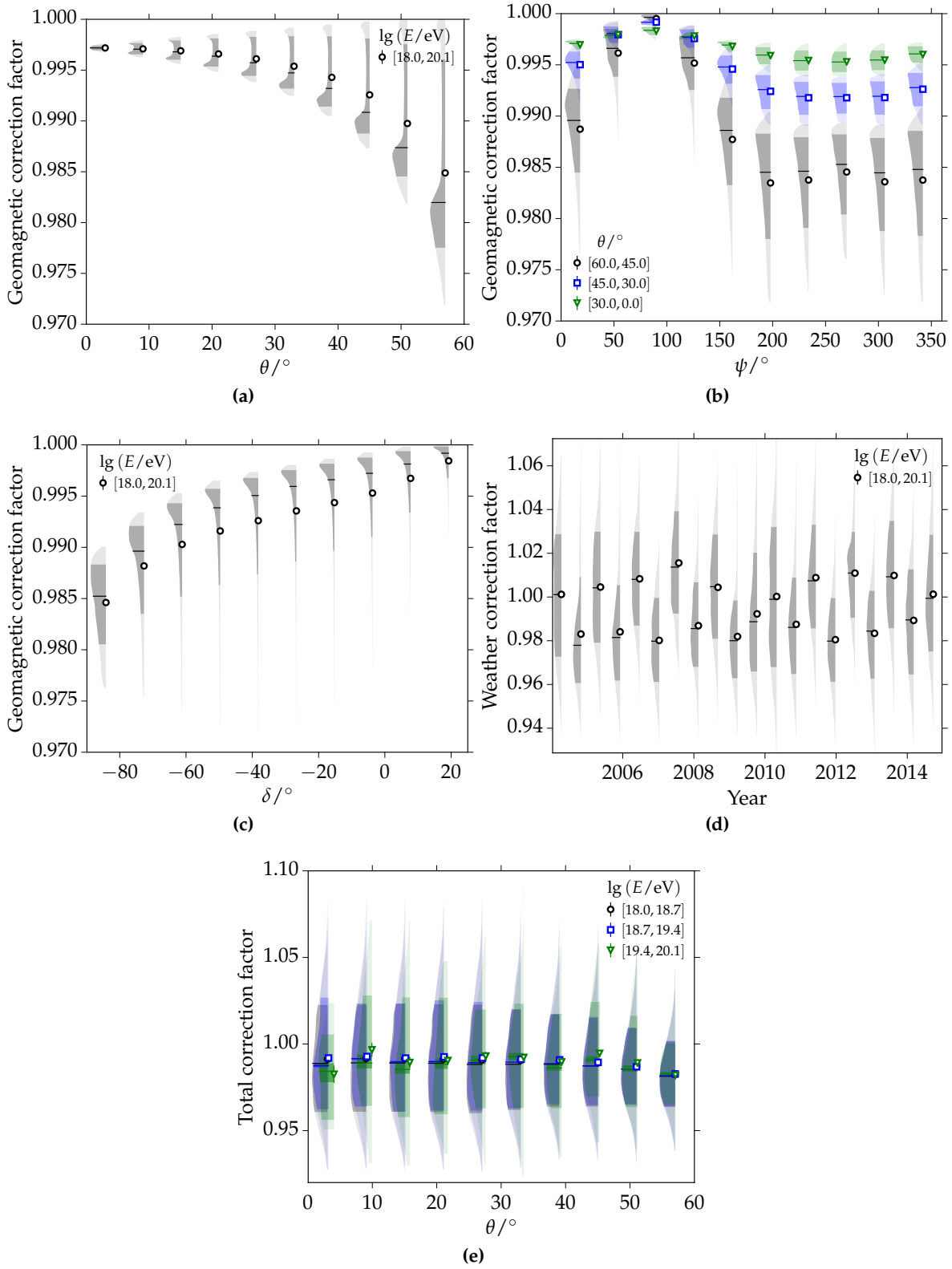
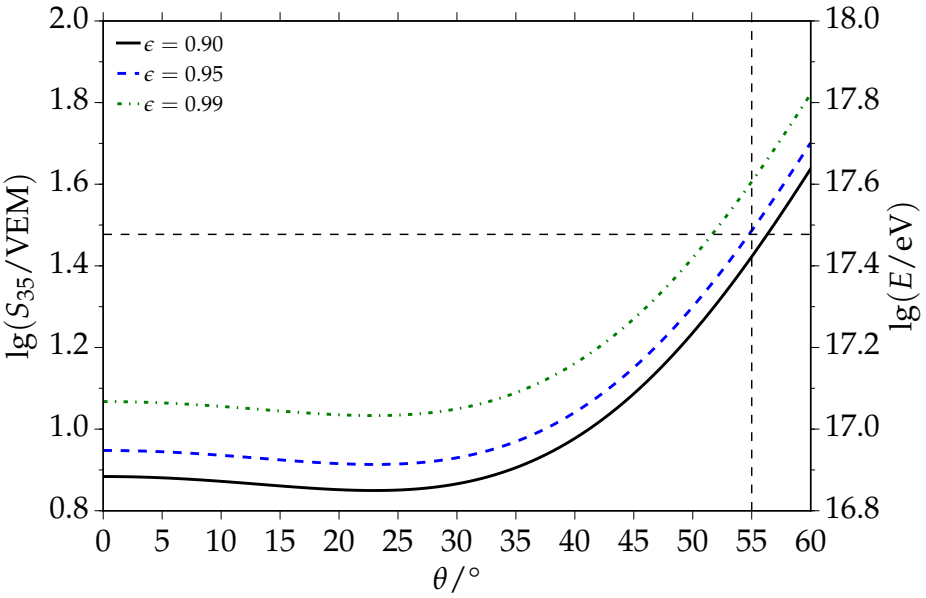


Figure 4.3: Geomagnetic correction factor to S_{1000} as a function of (a) zenith angle, (b) azimuth angle and (c) declination angle. (d) Weather correction factor as a function of time. (e) The total correction as a function of zenith angle and for different ranges in primary energy.

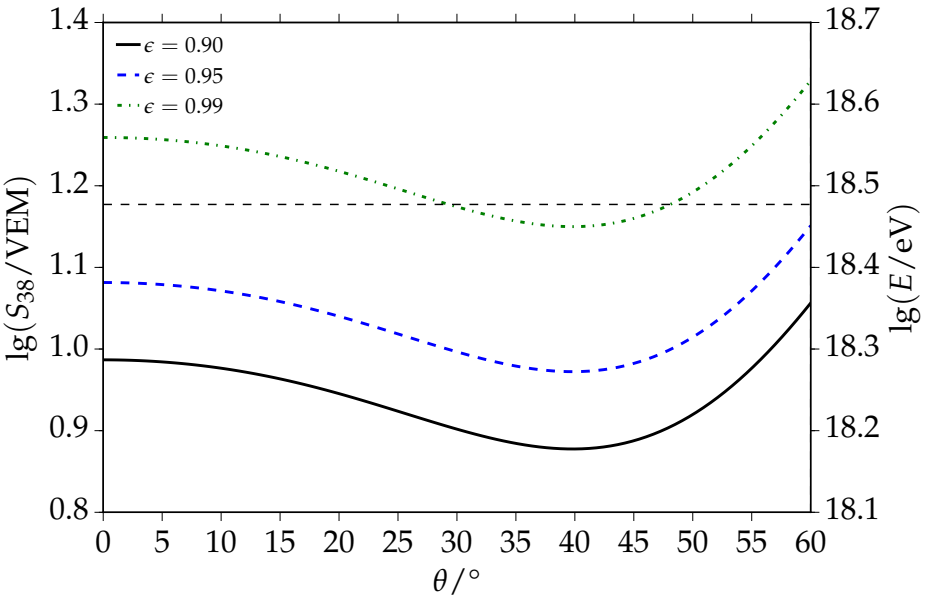
Table 4.1: Parameters of the T4 trigger efficiency models for the measurement of air showers with the SD-750 and the SD-1500.

Parameter	a_0	a_1	a_2	a_3	b
SD-750	2.39 ± 0.06	-4.86 ± 0.32	4.10 ± 0.56	-0.98 ± 0.31	0.249 ± 0.004
SD-1500	1.19 ± 0.06	-2.59 ± 0.35	3.26 ± 0.59	-1.2 ± 0.31	0.369 ± 0.004

as compared to SD-1500 events. As can also be seen in Fig. 4.4a, an efficient measurement of the SD-750 flux down to $10^{17.2}$ eV is possible in the zenith angle range from 0° to 40°. The model lines shown in Fig. 4.4 represent the conservative case of proton simulations. As they are mostly used to construct binary efficiency thresholds, a conservative choice is best to avoid introducing systematic uncertainties on the energy spectrum. Due to the larger muon content of air showers initiated by primaries with $A > 1$, the detection efficiency for these showers is usually larger than for proton primaries. A comparison between efficiency lines for proton and iron is shown in Fig. 4.5. Both plots for SD-750 and SD-1500 simulations are included.

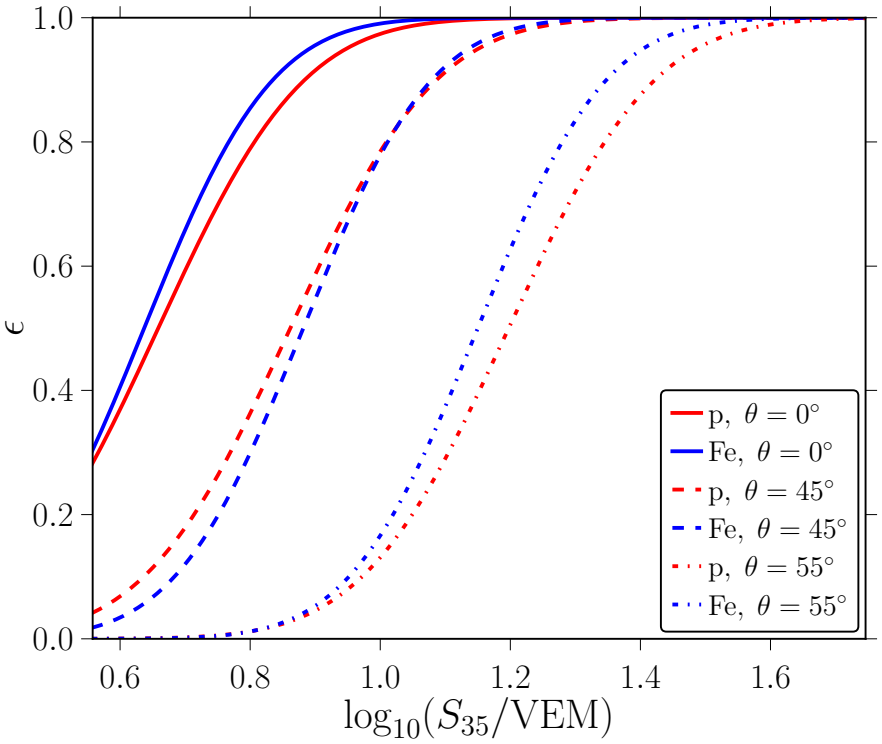


(a) SD-750 simulations

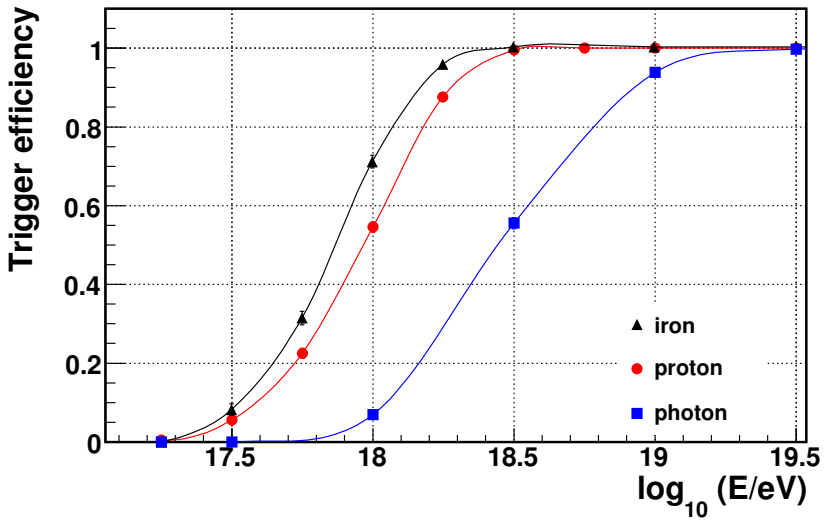


(b) SD-1500 simulations

Figure 4.4: Trigger efficiency lines derived from simulations. The horizontal black dashed lines represent the energies of full trigger efficiency used in the standard analysis.



(a) SD-750 simulations



(b) SD-1500 simulations

Figure 4.5: (a) T4 efficiency lines as a function of energy estimate for SD-750 simulations. Simulation of proton and iron nuclei are compared. Plot taken from [127]. (b) T4 efficiency lines as a function of energy and primary particle for SD-1500 simulations. Plot from [101].

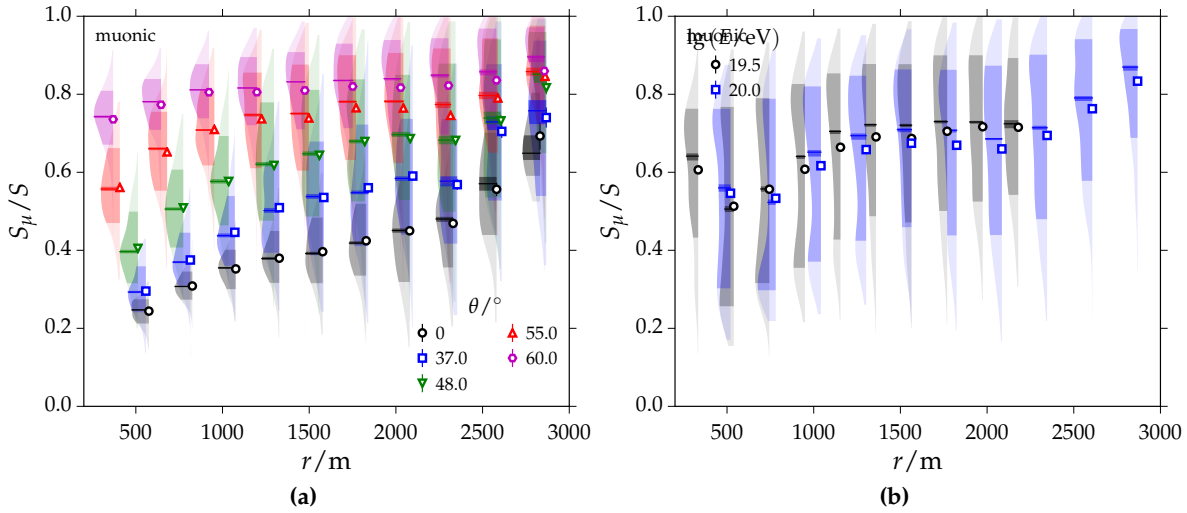


Figure 4.6: Relative contribution of the muonic signal in a WCD as a function of the distance to the shower core. Violin profiles are shown for different zenith angles in (a) and for different primary energies in (b).

4.5 Constant Intensity Cut

The development of EASs is well described as a function of traversed atmospheric overburden ΔX (as used in universality analyses discussed in Chapter 2). The size of air shower signals at ground level depends strongly on the zenith angle of the shower, not only on energy and mass of the primary particle. Until reaching the ground, an air shower with zenith angle of 60° traverses about twice the amount of atmosphere compared to a vertical shower. Strictly speaking, these considerations are crucial for the dominant electromagnetic part of an EAS. Frequent interactions are the foundation of this shower component, and thus the density of traversed atmosphere is of most relevance. Muons interact more rarely, and the chance of decaying, which is proportional to the traversed distance (not overburden), is mostly relevant. These facts introduce a non-trivial dependence of development characteristics on the muon fraction of an EAS. To visualize these facts, the fraction of the muonic signal to the total signal is shown in Fig. 4.6. Overall, ground signals of EASs from identical primary particles as measured with WCDs are decreasing with an increasing zenith angle (at least up to zenith angles above 60°). The CIC[1] method provides an empirical way to correct for the attenuation of air shower signals [151–157]. We assume an isotropic arrival of primary particles above a certain primary energy. The following calculation proves that the same number of events is expected in bins of equal $\mathrm{d}\cos^2\theta$:

$$\begin{aligned} \frac{\mathrm{d}N}{\mathrm{d}E\,\mathrm{d}\Omega\,\mathrm{d}t\,\mathrm{d}A_{\mathrm{eff}}} &= \frac{\mathrm{d}N}{\mathrm{d}E\,\mathrm{d}\cos\theta\,\mathrm{d}\phi\,\mathrm{d}t\,\mathrm{d}A\cos\theta} \\ \Rightarrow \frac{\mathrm{d}N}{\mathrm{d}\cos^2\theta} &= -\frac{\mathrm{d}N}{\mathrm{d}\sin^2\theta} \propto \text{const.}\,. \end{aligned} \tag{4.9}$$

The main point is the introduction of the effective detector area $A_{\mathrm{eff}} = A\cos\theta$ that results from projecting the stations into the shower plane perpendicular to the arrival direction of the shower. The energy dependence was intentionally left away in the last equation in order to emphasize that the principle holds for both differential and integrated fluxes. However, this is a purely mathematical argument. A change in primary composition as a function

[1] Constant Intensity Cut

of energy violates the assumption of an angular correction that is independent of energy. The standard CIC is thus an approximation that works well in the energy range of interest (and maybe also by chance given the specific observation level of Auger and thus a certain average distance to the shower maxima).

Shower sizes are by convention normalized to the expected shower size of an identical shower arriving under the median zenith angle. For the standard SD-1500, the median zenith angle is 38° and shower sizes are expected signals (from the fitted LDF) at a distance of 1000 m from the core:

$$S_{38}(E, A) = S_{1000}(\theta, E, A) / f_{\mathrm{CIC}}(\theta) \,. \tag{4.10}$$

As mentioned, this separation of the angular and the other dependencies does not necessarily hold under a change of composition with energy. The CIC is an effective correction for a specific observation height, energy-dependent primary composition and the energy range of interest. The dependence on the unknown primary mass A is usually dropped:

$$S_{38}(E) = S_{1000}(\theta, E) / f_{\mathrm{CIC}}(\theta) \,. \tag{4.11}$$

Different functional forms for the attenuation function f_{CIC} have been tested, see e.g. [38]. The current best choice is a third order polynomial in x, whereas x is defined as:

$$x := \cos^2\theta - \cos^2 38^\circ \,. \tag{4.12}$$

In the analysis of events recorded with the SD-750 array, the shower size is S_{450}, and the median zenith angle is 35°. Except for that, Eq. (4.10) and Eq. (4.12) are formulated equivalently.

Different methods have been used to derive the attenuation function from measured data. The *countdown-method* employs an equidistant binning in $\sin^2\theta$. Inside each bin, one iterates backwards through the distribution of shower sizes (starting with the largest size) until a certain number of sizes is reached. The intensity or number is chosen a priori. The average shower size at this intensity varies as a function of zenith angle, which allows to obtain the attenuation just from fitting this dependence. An uncertainty on the shower sizes can be obtained with a bootstrap method [155] or an analytical error propagation [38]. The result of this procedure is an attenuation function at a given intensity. The actual CR primary energy corresponding to this intensity is unknown a priori and the method to determine the attenuation function is agnostic towards it. The corresponding energy is estimated after the energy calibration procedure as detailed in Section 4.7.

I am using a slightly different approach to obtain the attenuation function, as discussed in a recent note [157]. The idea is to require uniformity of the number of events above a certain threshold in shower size. In mathematical terms, the following χ^2-function is minimized:

$$\chi^2 = \sum_{i=1}^{k} \frac{(n_i - \langle n \rangle)^2}{\langle n \rangle} \,. \tag{4.13}$$

The sum runs over k different zenith angle bins. Both quantities n_i and $\langle n \rangle$ implicitly include the attenuation function and thus the fit parameters. n_i is defined as:

$$n_i = \int_{S_{38}^0}^{\infty} \mathrm{d}S_{38} \frac{\mathrm{d}N}{\mathrm{d}S_{38}} \,, \tag{4.14}$$

where only events in the zenith interval $[\theta_i, \theta_{i+1}]$ corresponding to bin i are considered. The average number of events is given by:

$$\langle n \rangle = \frac{1}{k} \sum_{i}^{k} n_i \,. \tag{4.15}$$

The shower size S_{38}^0 is chosen a priori. S_{38} is defined via the attenuation function $f_{\mathrm{CIC}}(\theta;\vec{p})$ as specified in Eq. (4.10). The parameter vector given here ultimately enters $\chi^2(\vec{p})$. To check for an energy dependence of the attenuation function, the analysis can be repeated for different values of S_{38}^0, as performed in [157]. There is some non-trivial energy-dependence that should be closely investigated and incorporated into future analyses.

A problem with this method is the discreteness of data. A corrected shower size is either smaller or larger than S_{38}^0. This means that sizes will fluctuate in and out of the calculation of $\chi^2(\vec{p})$ when the CIC parameters are varied, given that each size is either counted or not in Eq. (4.14). To avoid this, the discrete counting of events is replaced by:

$$n_i = \int_0^\infty \mathrm{d}S_{38} K(S_{38}, S_{38}^0, \sigma[S_{38}]) \frac{\mathrm{d}N}{\mathrm{d}S_{38}}\,, \tag{4.16}$$

with a kernel function $K(S_{38}, S_{38}^0, \sigma[S_{38}])$ that effectively smears the distribution. One can write the dependence on parameters more explicitly as $K = K(S_{38}(\theta;\vec{p}), S_{38}^0, \sigma[S_{38}(\theta,\vec{p})])$. To smooth out the distribution, K assigns a real number to each event depending on its distance to S_{38}^0 in terms of corrected shower size. Given the reconstruction uncertainties $\sigma[S_{38}]$ (as discussed in Section 4.7.3 and Section 4.10), K is constructed as the distance of S_{38} to S_{38}^0 in terms of the uncertainty while assuming normally distributed values:

$$K(S_{38}, S_{38}^0, \sigma[S_{38}]) = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{S_{38} - S_{38}^0}{\sqrt{2}\,\sigma[S_{38}]}\right)\right]\,. \tag{4.17}$$

With the c.d.f. of the normal distribution on the right hand side. It has been found empirically that an approximation to K can be used without altering the resulting parameters in a significant way. The support of K as stated in Eq. (4.17) is infinity, and the calculation of the error function is rather expensive. It is thus useful to introduce an approximation of K with finite support (of a few σ) and a functional form that allows for simpler calculation. The chosen approximation K_a is:

$$K_a(x,\Delta) = \int_{-\infty}^{x} \rho(x',\Delta)\;\mathrm{d}x' = \begin{cases} 0 & ;\, x < -\Delta \\ (1 + x/\Delta)^2/2 & ;\, -\Delta \leqslant x < 0 \\ 1 - (x/\Delta - 1)^2/2 & ;\, 0 \leqslant x < \Delta \\ 1 & ;\, \Delta \leqslant x \end{cases}\,, \tag{4.18}$$

with $x := S_{38} - S_{38}^0$ and $\Delta = \sigma[S_{38}]$. Both functions are plotted in Fig. 4.7 for a realistic choice of uncertainties of $0.1\,\sigma[S]/S$. There is no relevant difference in the fitted parameters when the approximation is used instead of the error function. The finite support of the approximate function allows to remove a large amount of data before fitting, as sizes that are significantly smaller than the chosen threshold will never contribute to the calculation. This considerably speeds up the minimization. Furthermore, it imposes no relevant difference on the fit results if the uncertainties $\sigma[S_{38}]$ are replaced by approximative values.

I checked that the analysis method presented here gives equal or better results than previously used methods, e.g. [38, 155].

The calculation of statistical uncertainties based on $\Delta\chi^2 = 1$ and Eq. (4.13) was cross-checked with manually bootstrapping the signal distribution and re-fitting the attenuation function. Both methods lead to compatible uncertainties.

The results of the CIC analysis applied to SD-750 and SD-1500 events are discussed in the next two sections.

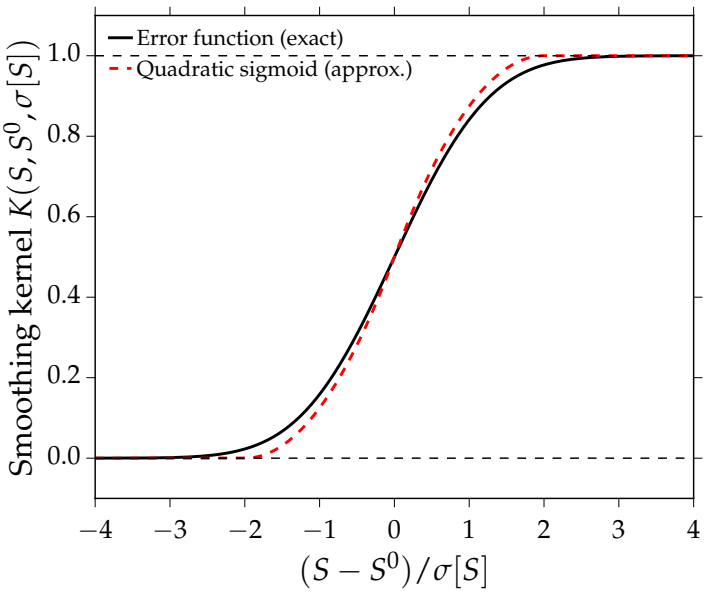


Figure 4.7: A plot of the quadratic sigmoid used to smooth the event distribution in the minimization for the CIC. The exact function (Gaussian c.d.f. or error function) is plotted for comparison.

4.5.1 Attenuation function for SD-750 events

The analysis procedure to obtain the attenuation function is applied to data collected with the SD-750. Data are cut according to the criteria for quality selection described in Appendix D.1.

The shower sizes S_{450} are normalized to sizes of showers with a median zenith angle of 35°, taking into account the range of zenith angles from 0° to 55°. I refer to these normalized sizes as S_{35}. For the fit of the attenuation function, I only consider events with $S_{35} > 45\,\mathrm{VEM}$. This restricts data to a region of full trigger efficiency over the range of zenith angles. This cut leads to the selection of primaries with energies above 6×10^{17} eV.

There are 11 986 events that fulfill the selection criteria for the fit. I use 200 bins in $\cos^2\theta$, thereby limiting the number of events within each bin to approximately 60. This is well above the critical number to be able to approximate the Poisson distribution with a Gaussian.

To correct data, a third order polynomial is minimized in the expression Eq. (4.13). The resulting attenuation function can be written as:

$$f(x) = 1 + (1.605 \pm 0.039)\, x + (-1.51 \pm 0.10)\, x^2 + (-1.92 \pm 0.53)\, x^3 \qquad (4.19)$$

with x as defined in Eq. (4.12). The correlation of the three fitted parameters is above 97 % in all cases. The attenuation curve is shown in Fig. 4.8a as a function of $\sec\theta$. This is a natural way of plotting as the amount of atmosphere traversed by a shower is proportional to $\sec\theta$. For comparison, the attenuation function used in the last major conference contribution [31] is shown as red, dashed line. The points and uncertainties are obtained from a bootstrap re-sampling of the measured event distribution within each zenith angle interval. They are plotted for comparison and do not directly represent the points that are fitted. As they match exactly with the function, the points provide an additional confirmation of the result. A relative residual of both attenuation functions is given in Fig. 4.8b. The shaded line represents the $1\,\sigma$ uncertainty contour of the new function. The functions are well compatible to within 5 % or $2\,\sigma$. The distribution of uncorrected shower sizes S_{450} for four different zenith

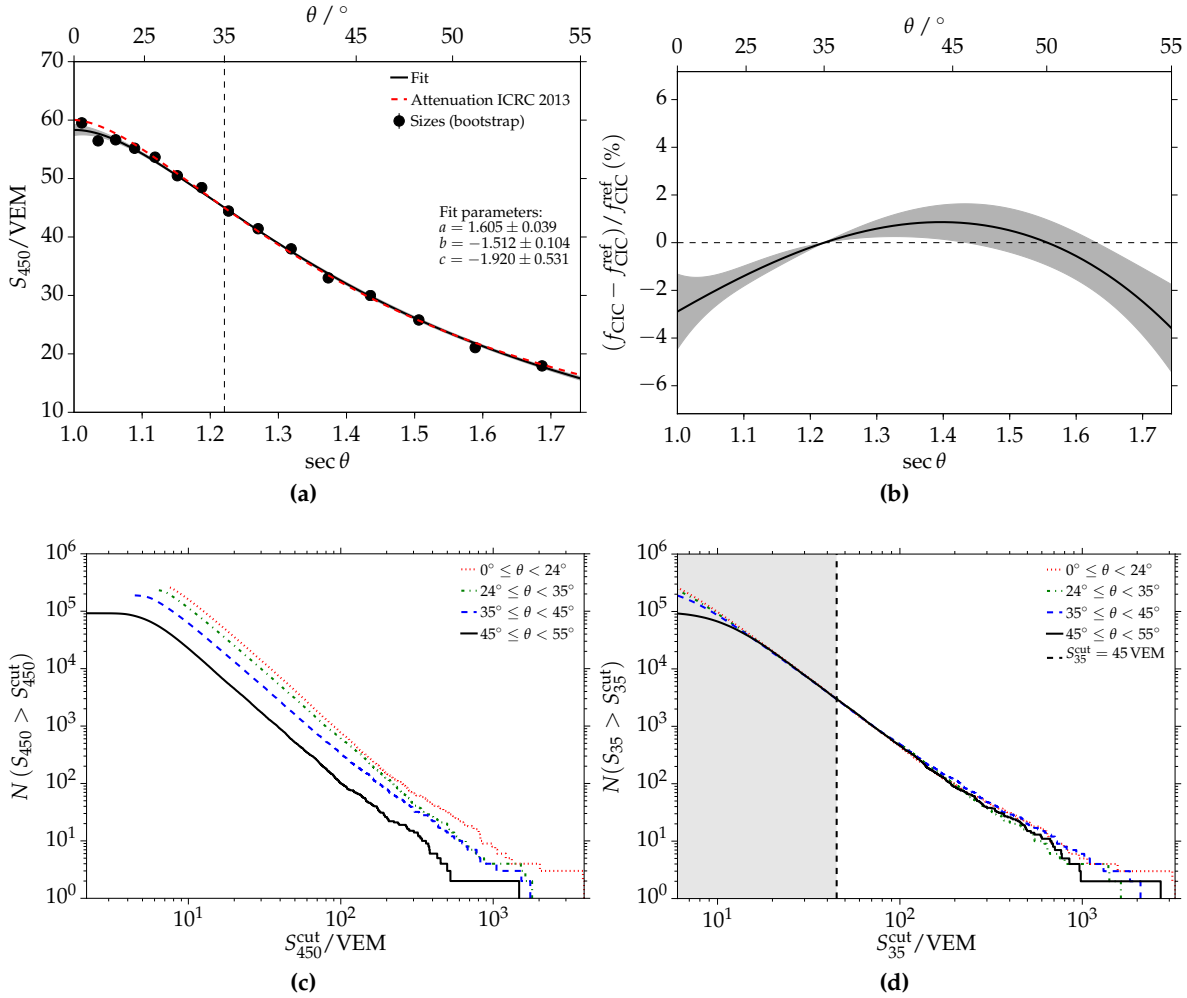
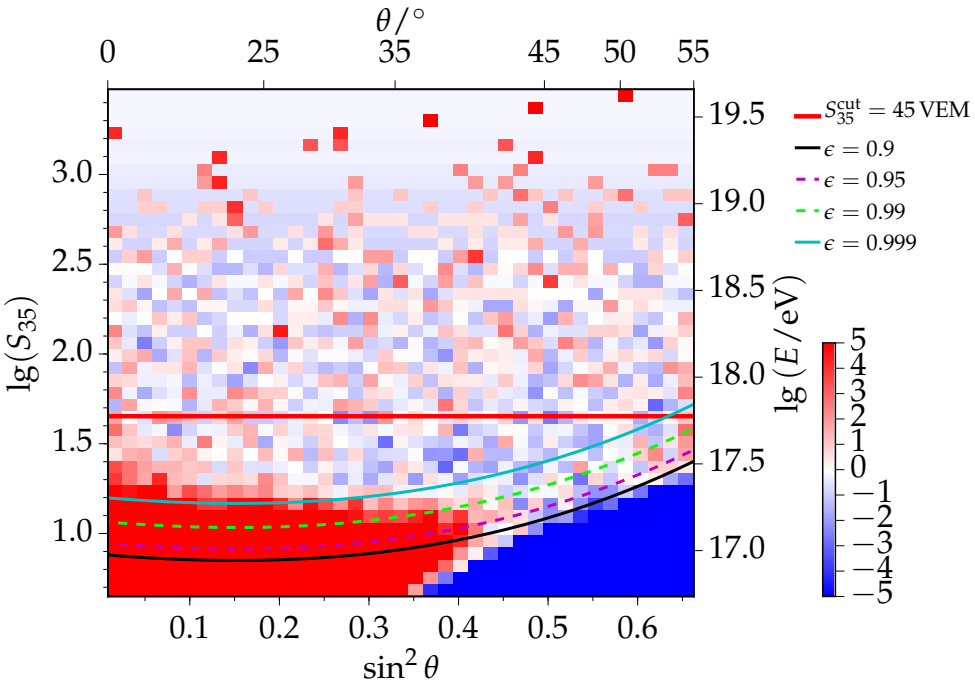
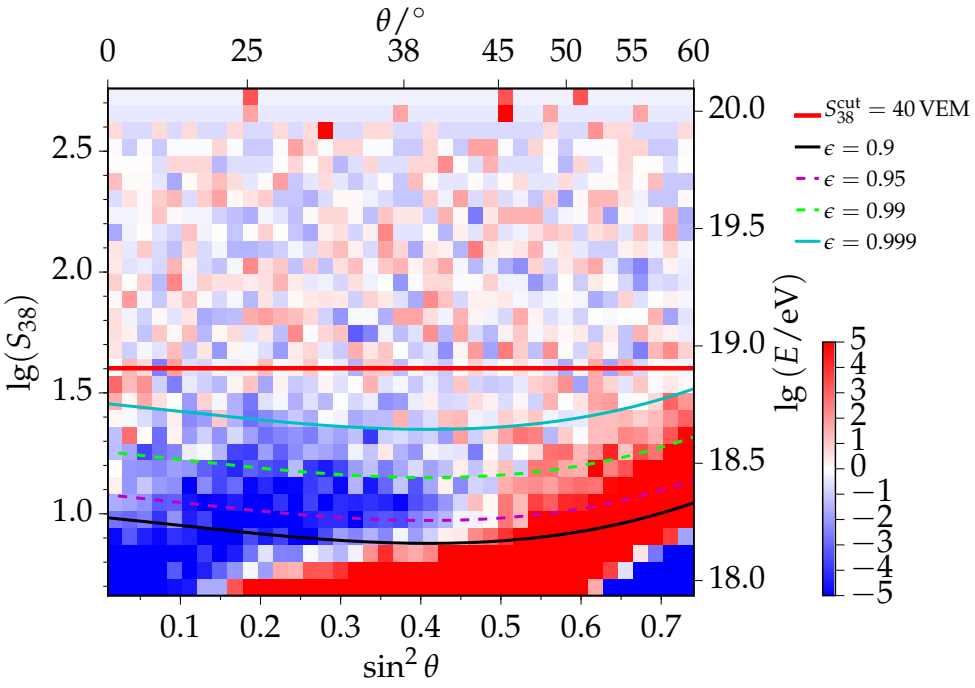


Figure 4.8: (a) The resulting attenuation function for events measured with the SD-750. The attenuation function used at the ICRC 2013 is plotted for comparison. (b) Relative difference of the new attenuation function to the previous one. The shaded area represents the $1\,\sigma$ uncertainty of the new model. (c), (d) Uncorrected and corrected shower size distributions for different zenith angle ranges.

intervals is given in Fig. 4.8c. Each point in the histogram represents the integrated number of events (intensity) above the corresponding S_{450}. The signal distributions vary due to the different shower geometry and attenuation for different zenith angles. After correction with the attenuation function Eq. (4.19), the signal distributions agree well with each other as depicted in Fig. 4.8d; this is expected for a valid correction. A better method to visualize the quality of the CIC correction is a two-dimensional event distribution as shown in Fig. 4.9a. The color of each bin represents the deviation from the average number of events within the chosen bin $\Delta \lg S_{35}$. The difference is given in units of statistical standard deviation σ_{stat}. A perfect correction without residual energy dependencies is expected to result in a perfectly flat distribution in $(\Delta \lg S_{35}, \Delta \cos^2\theta)$. Judging from Fig. 4.9a, this is indeed the case. All the fluctuations for numbers of events above the efficiency threshold are on the order of a few standard deviations. T4 efficiency models as derived from simulations are overlaid in Fig. 4.9a. Efficiencies ϵ closer to 1 lead to a more conservative restriction of data to the regime of full efficiency.



(a) SD-750 data



(b) SD-1500 data

Figure 4.9: Two-dimensional event distribution of energy estimates. T4 trigger efficiency models from simulations are shown for comparison; different values of ϵ reflect different efficiency values. (a) SD-750 data, (b) SD-1500 data. The deviation of event numbers from the mean within each signal bin is shown with colors (see text for details). The red solid lines depict the energy thresholds above which events are used for the fit of the attenuation function.

4.5.2 Attenuation function for SD-1500 events

For the analysis of SD-1500 data, shower sizes S_{1000} are normalized to the shower size of identical showers with median zenith angle of 38°. I will refer to the corrected shower size as S_{38}. The analysis procedure is identical to the analysis of SD-750 events as discussed in the previous section. 16 106 events above a threshold of $S_{38} > 40\,\text{VEM}$ are used in the fit. This results in 80 events per bin in $\cos^2\theta$. The cut in S_{38} is equivalent to a cut in primary energy above $8.1\times10^{18}\,\text{eV}$. While a cut closer to the energy threshold of full efficiency at $3\times10^{18}\,\text{eV}$ would add a significant amount of events to the fit, there are disadvantages that need to be considered. Below primary energies of $6\times10^{18}\,\text{eV}$, the flux of UHECRs shows deviations at largest zenith angles (as shown in Fig. 4.41b). This is possibly caused by event migrations due to the finite (and zenith dependent) resolution of the measurement. By construction, the attenuation function does not depend on energy. The derivation of the correction at lower energies would thus introduce a bias for the correction at higher energies. The importance of an unbiased energy spectrum above $6\times10^{18}\,\text{eV}$ clearly outweighs the correctness at lower energies, which leads to the choice of the cut value in this work. The introduction of an energy dependence of the CIC correction is a reasonable possibility for the future.

The attenuation function resulting from a minimization of Eq. (4.13) is given by:

$$f(x) = 1 + (0.970 \pm 0.043)\,x + (-1.618 \pm 0.084)\,x^2 + (-1.30 \pm 0.45)\,x^3 \tag{4.20}$$

The results are visualized in Fig. 4.10. The attenuation function is plotted in Fig. 4.10a, while a residual comparison to the previous model is presented in Fig. 4.10b. The agreement of both functions is very good. All differences are within 2 % or $1\,\sigma$. The largest deviations occur at the borders of the zenith angle range 0° to 60°. The distributions of the uncorrected and corrected shower sizes are shown in Fig. 4.10c and Fig. 4.10d, respectively. After correcting the shower sizes, the distribution of the number of events is constant with zenith angle as visualized in Fig. 4.9b. As in the case of SD-750 events, this holds only above the threshold of full trigger efficiency as indicated by the efficiency lines.

4.5.3 The impact of shower size corrections

This section gives a quantification of the impact of shower size corrections (see Section 4.3) on the attenuation functions. I am applying geomagnetic and weather corrections to S_{1000} and only weather corrections to S_{450}. A justification for this is given in Section 4.3. The analysis of the attenuation functions is repeated with the corrected shower sizes. The results for S_{450} or shown in Fig. 4.11. The impact of the weather corrections on the SD-750 attenuation function is below 0.1 % or $0.1\,\sigma$ in the zenith angle range. The final attenuation function is:

$$f(x) = 1 + (1.602 \pm 0.039)\,x + (-1.50 \pm 0.10)\,x^2 + (-1.88 \pm 0.53)\,x^3 \tag{4.21}$$

Equivalently, the attenuation function for the SD-1500 is shown in Fig. 4.12. The combined impact of weather and geomagnetic corrections is below 0.5 % or $0.8\,\sigma$. This is a larger difference than in the SD-750 case and is relevant for the detailed analysis of angular dependencies of the flux. The attenuation function is:

$$f(x) = 1 + (0.973 \pm 0.044)\,x + (-1.632 \pm 0.082)\,x^2 + (-1.23 \pm 0.45)\,x^3 \tag{4.22}$$

These functions are used to correct data for the following analyses. The two-dimensional event distributions in Fig. 4.13 are basically unchanged with respect to the previous ones (see Fig. 4.9). The corrections work very well in the energy ranges of interest and in the considered ranges of zenith angle.

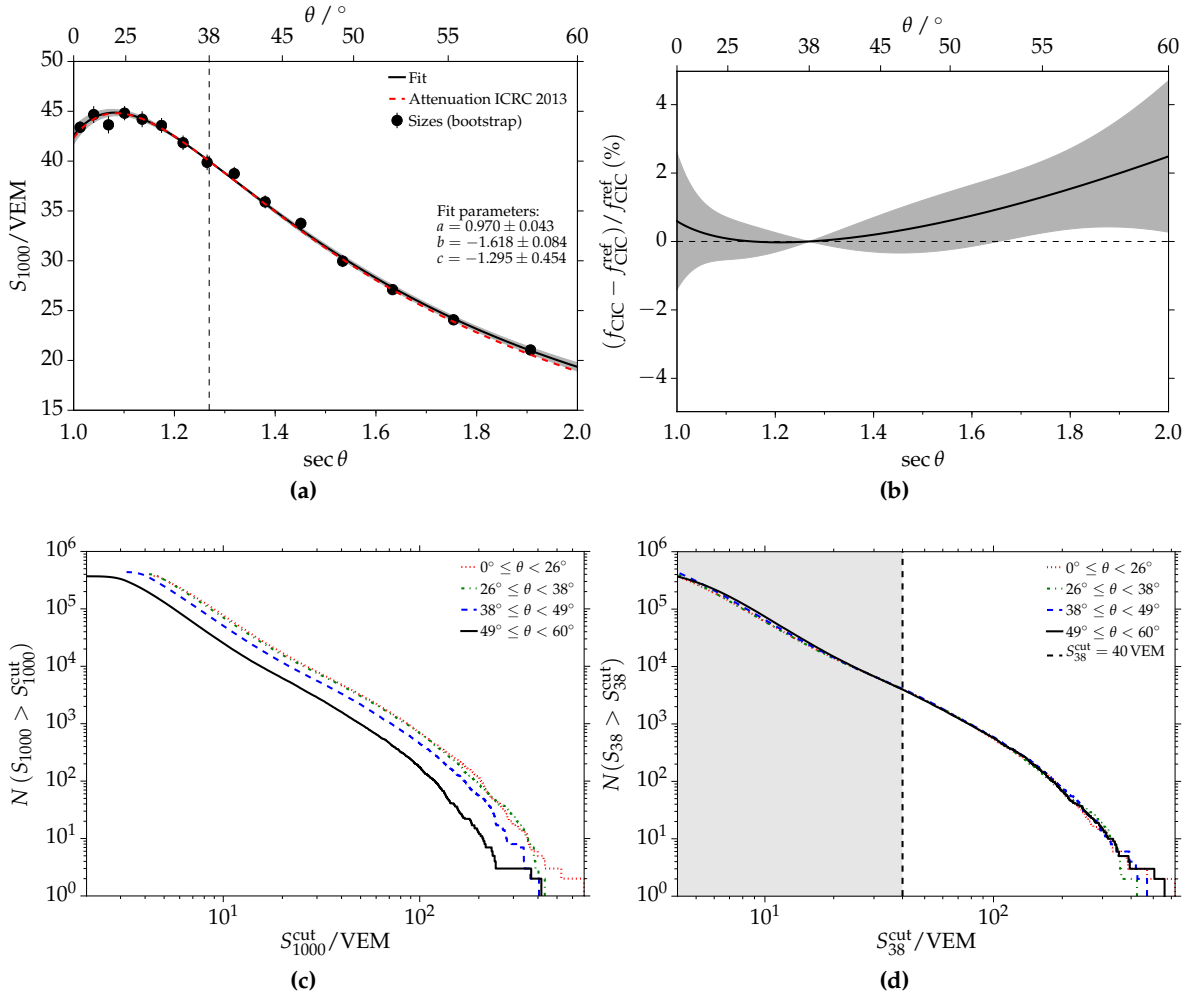


Figure 4.10: (a) The resulting attenuation function for events measured with the SD-1500. The attenuation function used at the ICRC 2013 is plotted for comparison. (b) Relative difference of the new attenuation function to the previous one. The shaded area represents the 1 σ uncertainty of the new model. (c), (d) Uncorrected and corrected shower size distributions for different zenith angle ranges.

4.5.4 Energy dependency of the CIC

The energy dependence of the attenuation function derived from SD-1500 data was recently studied in [157]. The authors looked at the evolution of the parameters of the attenuation function as a function of energy or S_{38}. As there are changes in the parameters that exceed statistical uncertainties, these topics need further investigation. Implications about mass composition can possibly be extracted from an energy dependent CIC, but the current plan is not to introduce an energy dependence in the established analysis of the energy spectrum, consisting of CIC followed by the energy calibration.

Previous analyses on the energy dependence of CIC parameters obtained from SD-750 data exist as well, e.g. [127]. The results are not fully conclusive yet, but the magnitude of an energy dependence seems to be smaller compared to in SD-1500 data.

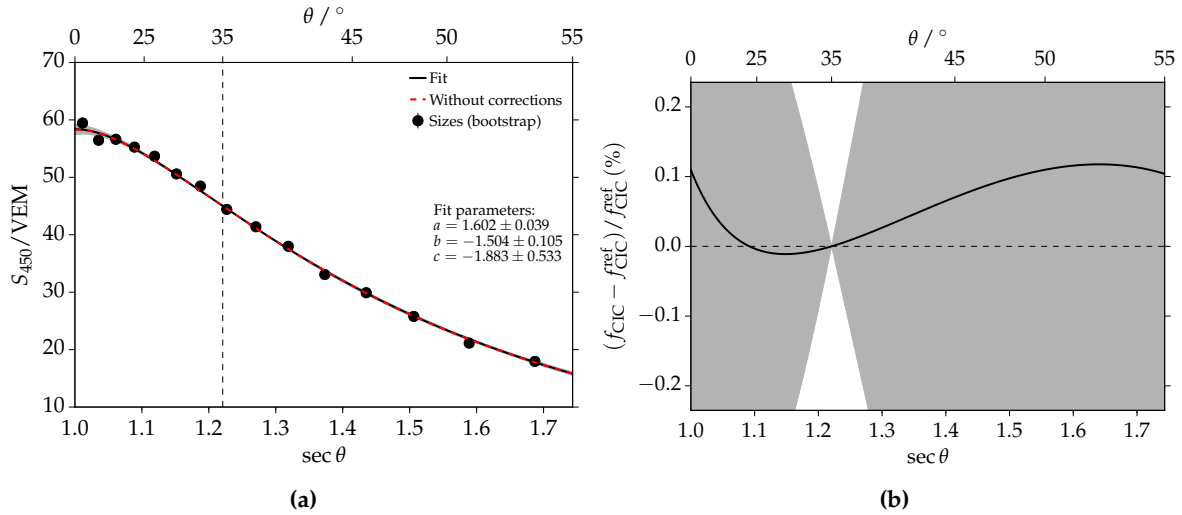


Figure 4.11: (a) The attenuation function fitted to corrected SD-750 data in comparison to the function that describes uncorrected data. (b) A residual comparison of both functions. The $1\,\sigma$ spread is given as shaded region.

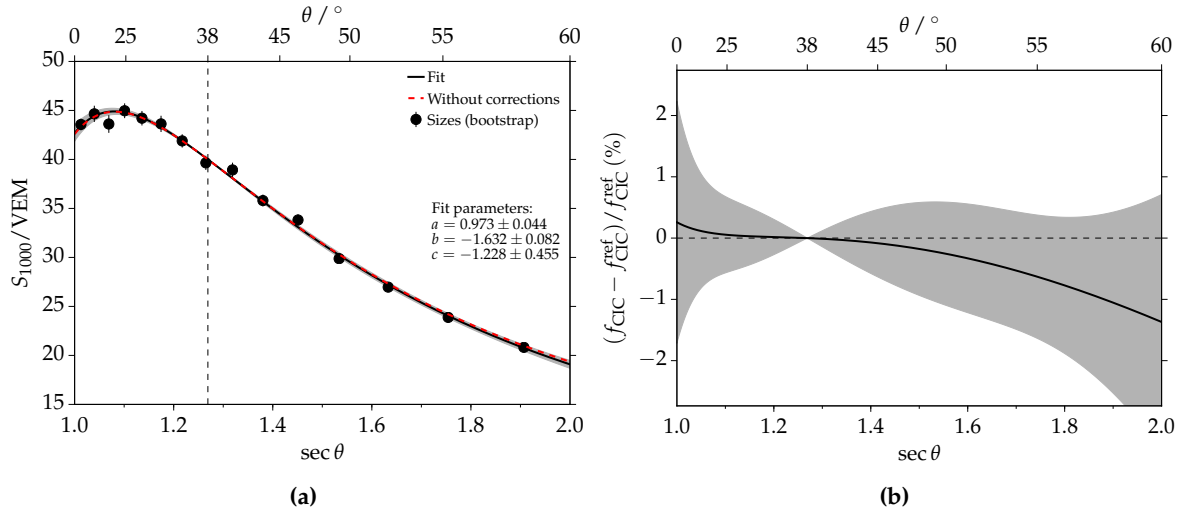
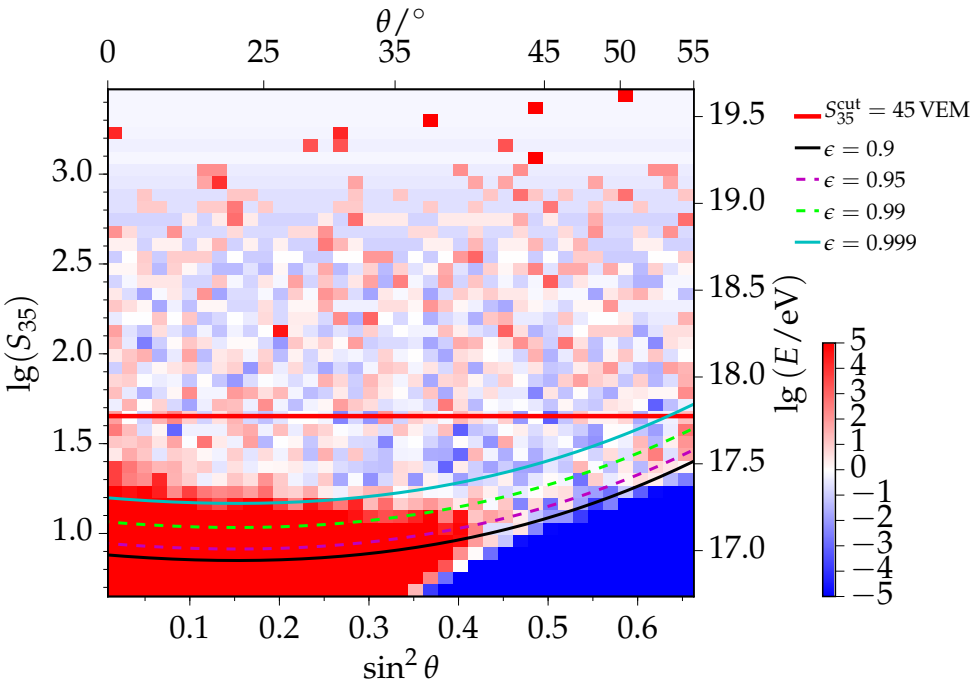


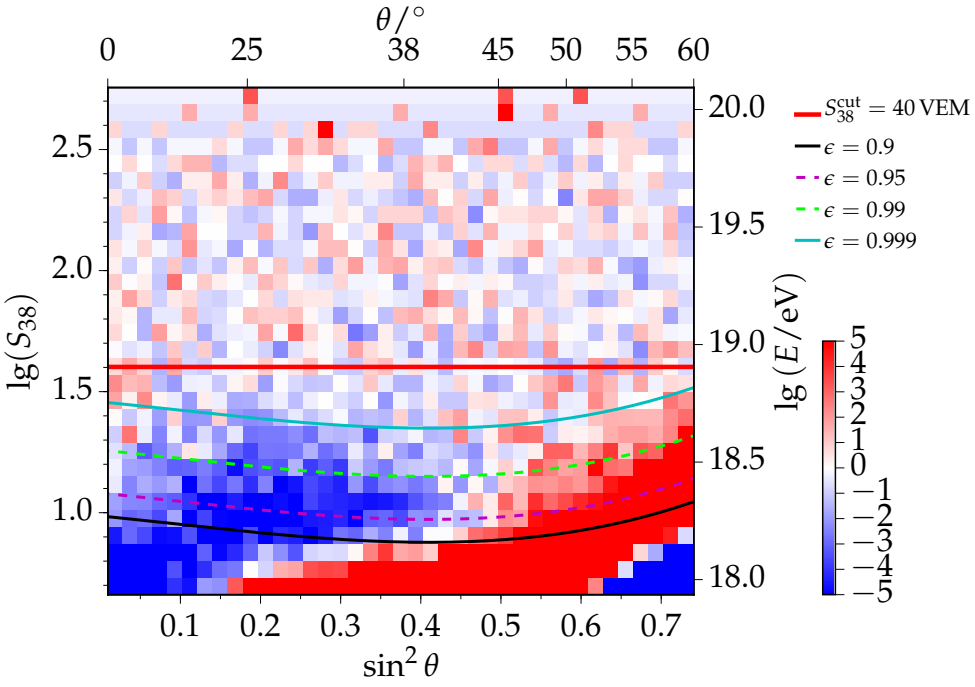
Figure 4.12: (a) The attenuation function fitted to corrected SD-1500 data in comparison to the function that describes uncorrected data. (b) A residual comparison of both functions. The $1\,\sigma$ spread is given as shaded region.

4.5.5 Functional forms and air shower physics

A variety of different empirical functions were used in the attempt to describe the attenuation of air shower signals [38]. The current third order polynomial results in an excellent description of data recorded at Auger. Exploiting the concept of air shower universality as described in Chapter 2, it should be possible to derive the average attenuation function. This would be based on the knowledge of the different particle components in an air shower and their dependencies on properties of the primary particle. Models for the flux of UHECRs and their mass composition need to be used for this. See Chapter 5 for details on the latter. The evolution of the average muon content as discussed in Section 5.2 is critical for the understanding of the attenuation function based on shower universality principles.



(a) SD-750 data



(b) SD-1500 data

Figure 4.13: Two-dimensional event distribution of energy estimates. Weather and geomagnetic effects are accounted for. T4 trigger efficiency models from simulations are shown for comparison; different values of ϵ reflect different efficiency values. (a) SD-750 data, (b) SD-1500 data. The deviation of event numbers from the mean within each signal bin is shown with colors (see text for details). The red solid lines depict the energy thresholds above which events are used for the fit of the attenuation function.

4.5.6 Matching energy scales

The CIC method provides an independent way to match the energy scales of different measurements that are sensitive in an overlapping energy range. For example, the scale of energy estimates S_{38} and N_{19} of the vertical and inclined events have been matched in [156]. This provides a separate way of comparing energies with respect to comparing them after performing calibrations with independent sets of golden hybrid events. Thus, I would like to strongly encourage this kind of analysis in future works.

4.6 Change of the SD energy scale with time

There is a number of reports on the change of the SD detector response with time or the age of the detector. For example, a significant change of the event-by-event energy differences between SD-1500 and FD measurements with time was observed in [158]. The concern is focused on the change of the response to the electromagnetic part of the signal in connection with particularly cold freezing periods. During these periods, some properties of the WCDs might have changed. Candidates are the water itself or the surface of the tyvek material that surrounds the water. Several hardware and simulation studies to find out the physics reason for these changes are currently ongoing. Within the scope of this work it is important to study and quantify these changes as both the total signal as well as the relative electromagnetic and muonic signal contributions could change with time, even if the flux and compositions of primary UHECRs remain unchanged. In case of such a change, energy spectrum and composition analyses are likely to be systematically affected. These systematics need to be quantified and possibly corrected for (see Section 4.8.4). The CIC method is used to study the impact of the aging effects. While the usual CIC method is used to get a zenith angle independent energy estimate for each reconstructed shower, the method is applied to events in different time and tank age ranges in order to analyze changes of the energy scale or angular dependencies. The CIC method itself is explained in Section 4.5. The reference attenuation function for SD-1500 data is presented and derived in Section 4.5.2.

This section is structured as follows: the extension of the CIC method for the purpose of this study is explained in Section 4.6.1 and the main results are presented in Section 4.6.2. Time dependencies of other interesting quantities, e.g. the average number of saturated stations or ratios of event numbers with certain trigger types, will be discussed at the end of the section. Data from the official start of measurement at the 01.01.2004 until 20.02.2012 are included in this analysis. Due to time constraints, the analysis was not updated to include events of the full data period is stated in Section 4.1. There is no indication that the obtained results would change with an updated data set or reconstruction, i.e. the result is independent of the overall energy scale of Auger. The events were reconstructed with $\overline{\text{Off}}\underline{\text{line}}$ (distributed as v7r6 observer on the Auger observer web page [140]). For all analysis steps, if not otherwise stated, the data are selected according to the quality selection described in Section 4.1.
This section is adapted from [159], an Auger internal note that I have written as a part of this thesis.

4.6.1 Extended CIC method

To study a possible time dependence of the SD-1500 detector response, the fitted shower sizes and attenuation parameters, the SD-1500 events are split into independent sets covering different time ranges. The CIC analysis is applied separately to events in each of these

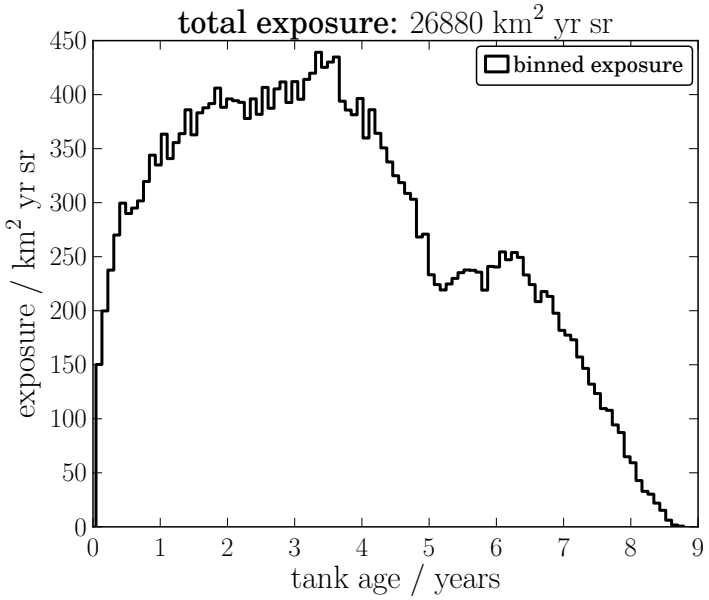


Figure 4.14: Differential exposure of the surface detector for vertical events as a function of tank age. Tank age is defined as the age of the tank with the hottest signal in an event.

time ranges. Time range can refer to an actual time period or the age of tanks. The exposure $\mathcal{E}_i$ for each range t_i must be known. It was calculated in daily steps with a stand-alone software developed for SD-1500 and SD-750 exposure calculations based on T2 information (see Section 4.2). The differential exposure as a function of tank age is plotted in Fig. 4.14. It is obtained from a resorting of active hexagons according to tank age, which is defined is the age of the hottest tank at the time of the event. The following formula based on the CIC assumption is exploited to analyze the change of the average energy (estimate) with time:

$$\int_{S_{38}^{\mathrm{cut}}(t_i)}^{\infty} \frac{\mathrm{d}N}{\mathrm{d}S_{38}}\,\mathrm{d}S_{38} = a\,\mathcal{E}(t_i), \tag{4.23}$$

for each time interval t_i. Only events in this time interval contribute to the integral. a is a relative intensity that will be varied to scan different cut energies. a will be referred to as the cut intensity and it is expressed in relative units of exposure. Eq. (4.23) can be directly used to analyze the change of the energy estimate S_{38}. The equation must be extended to perform a simultaneous shape fit for each time bin. Following from the CIC assumption, the relevant change is a binning in n equidistant $\sin^2\theta$-bins denoted by $\{\theta_j\}_{j=1\ldots n}$. Then, we have for each $j \in \{1\ldots n\}$ and time interval t_i:

$$\int_{S_{1000}^{\mathrm{cut}}(t_i,\theta_j)}^{\infty} \frac{\mathrm{d}N}{\mathrm{d}S_{1000}}\,\mathrm{d}S_{1000} = \frac{a}{n}\,\mathcal{E}(t_i,\theta_j). \tag{4.24}$$

The exposure for the specific time and zenith angle interval is calculated according to:

$$\mathcal{E}(t_i,\theta_j) = \mathcal{E}(t_i)\,\frac{\sin^2\theta^j_{\mathrm{max}} - \sin^2\theta^j_{\mathrm{min}}}{\sin^2\theta_{\mathrm{max}} - \sin^2\theta_{\mathrm{min}}} \tag{4.25}$$

with $\theta_{\mathrm{min}} = 0$ and $\theta_{\mathrm{max}} = 60^\circ$. For a fixed time interval t_i, the cut shower sizes $S_{1000}(\theta)$ are fit. The time dependence of the energy estimate S_{38} is obtained as well as the change of the attenuation function $f_{\mathrm{CIC}}\theta$ with time. The time range itself can be divided into bins in two reasonable ways: a binning in equidistant time ranges and a binning in equidistant exposure. Both methods lead to similar results. A binning in equal exposure is chosen to produce the following results.

4.6.2 Results

In this section, I present the results of the time dependence of SD-1500 events. The method can be applied to events measured with the SD-750 in a very similar manner (see also Section 4.5.1). To quantify possible changes of the SDs due to aging or freezing events, the main focus is to present results as a function of tank age. For each event, its corresponding tank age is defined as the time difference between the occurrence of the event and the deployment date of the hottest tank. The latter is extracted from T2 information and deployment databases. Dates are specified on the web page of the Auger acceptance group [141].

Signal spectra for different tank ages

To get an overview of the impact of tank aging on the measurement, the S_{38} spectra are divided into five tank age bins that cover equal amounts of exposure. The spectra for different energy ranges are depicted in Fig. 4.15. The energy estimates S_{38} are calculated with the reference attenuation presented in Section 4.5.2. A clear structure is visible in the trigger regime, depicted in Fig. 4.15a. The intensity decreases with tank age, indicating that older tanks respond with smaller signals to the same incoming particles, resulting in a loss of events with low signals close to the absolute trigger threshold. At larger signals, on the other hand, the behavior is reversed and the intensity increases with tank age. This could result from a drift of the VEM scale itself, leading to a bias in the signal assignment that increases with tank age. While this is only one explanation, there might be other effects that lead to the observed behavior. A detailed look on how the signals above the trigger threshold drift will be presented in the next section.

Change of the signal with tank age

To analyze a change of the signal with tank age, the angular shape is fixed to the reference attenuation and the signals at a fixed intensity are calculated as a function of tank age. The procedure is repeated for different intensities corresponding to different energy thresholds. The results are depicted in Fig. 4.16. At each intensity, an increase of the signal scale with tank age is observed. The relative change per year of tank age is $\Delta E/E \approx 1\,\%$ as depicted in Fig. 4.17. There is only a slight increase with energy, which is not statistically significant. The current energy calibration (see Section 4.7) was used to convert energy estimates to energies. In this step, there is no additional time dependence of the FD energy scale, because a fixed calibration function is used. This means that the given average energy might be biased, but the relative change is not. The uncertainty in the relative change increases with tank age as the energy threshold increases and thus the event statistics decrease. It should be noted that the increase shown in Fig. 4.16 seems to flatten in the last bin. At the moment it is undecided if this is just a statistical effect or an actual feature. The trend is not as clearly visible when the attenuation functions are fitted at the same time, as discussed in the following section.

Simultaneous fit of the attenuation function

To incorporate a change of the angular shape, the full CIC method is applied for each tank age, meaning that the attenuation is fitted for each tank age instead of using the reference function. The change of the signal itself can be compared to the results when the attenuation is fixed. It is shown in Fig. 4.18. At each cut intensity, to analyze the tank age dependence of S_{38}, a constant and a linear model are fit to data. The difference in χ^2 values is translated

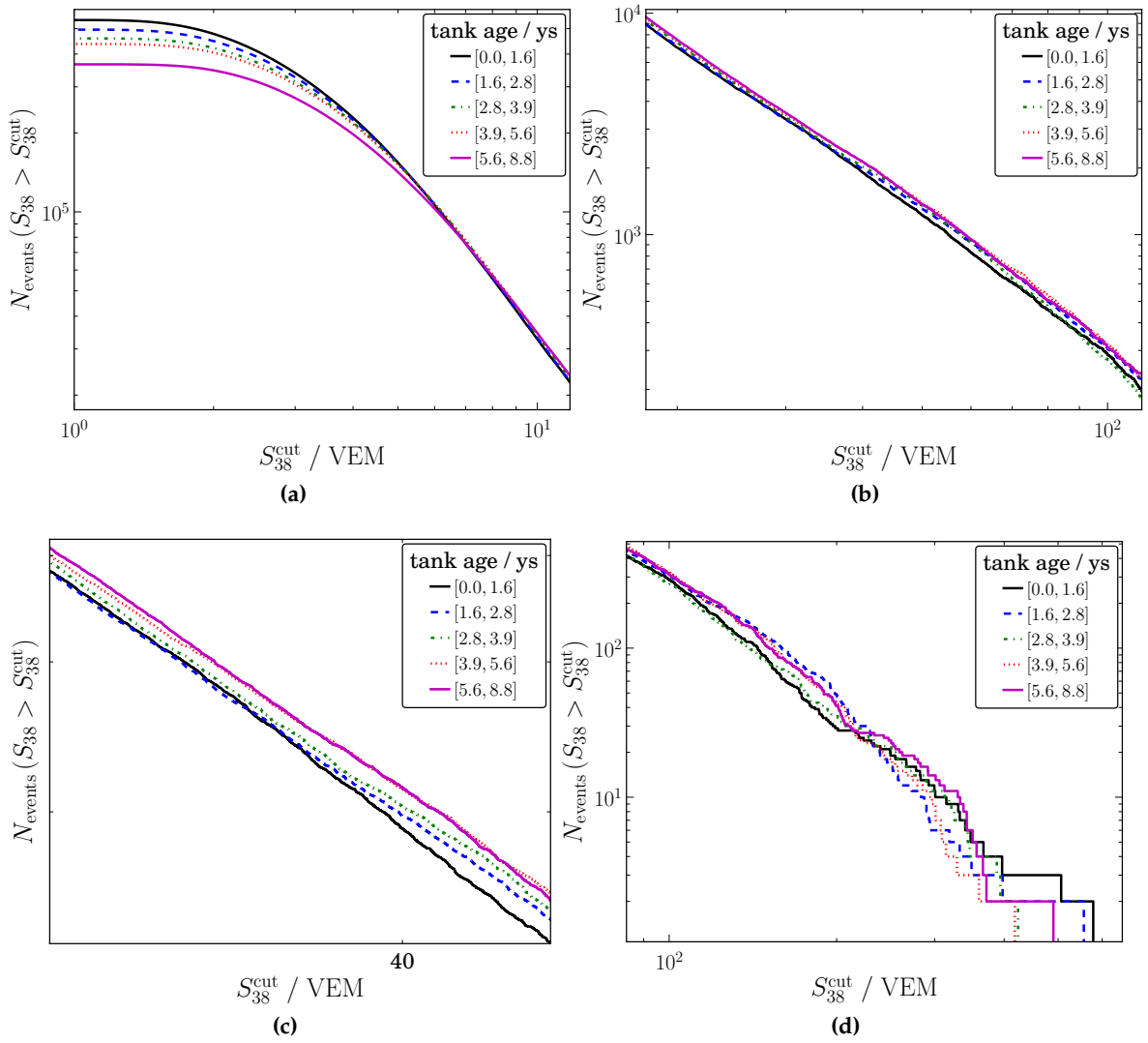


Figure 4.15: The S_{38} spectra for five tank age bins covering equal amounts of exposure. For visibility, the spectrum is divided into different energy ranges and shown in separate sub plots. A different behavior with tank age is observed for different energy ranges.

into a statistical significance n in terms of units of σ in favor of the linear model compared to the constant model. From the difference of the log-likelihood values $\log \mathcal{L}$ it follows:

$$\Delta \log \mathcal{L} = \frac{\Delta\chi^2}{2} = \frac{(\mu - X)^2}{2\sigma^2} = \frac{(n\sigma)^2}{2\sigma^2} = \frac{n^2}{2}, \tag{4.26}$$

$$\Longrightarrow n = \sqrt{\Delta\chi^2}. \tag{4.27}$$

A detailed argumentation is, for example, presented in [160]. The argumentation is not restricted to the case of normally distributed random variables, but it might be a rough approximation otherwise.

In each plot, the reduced χ^2 value of the constant model and the statistical significance of the linear model with respect to the constant is given.

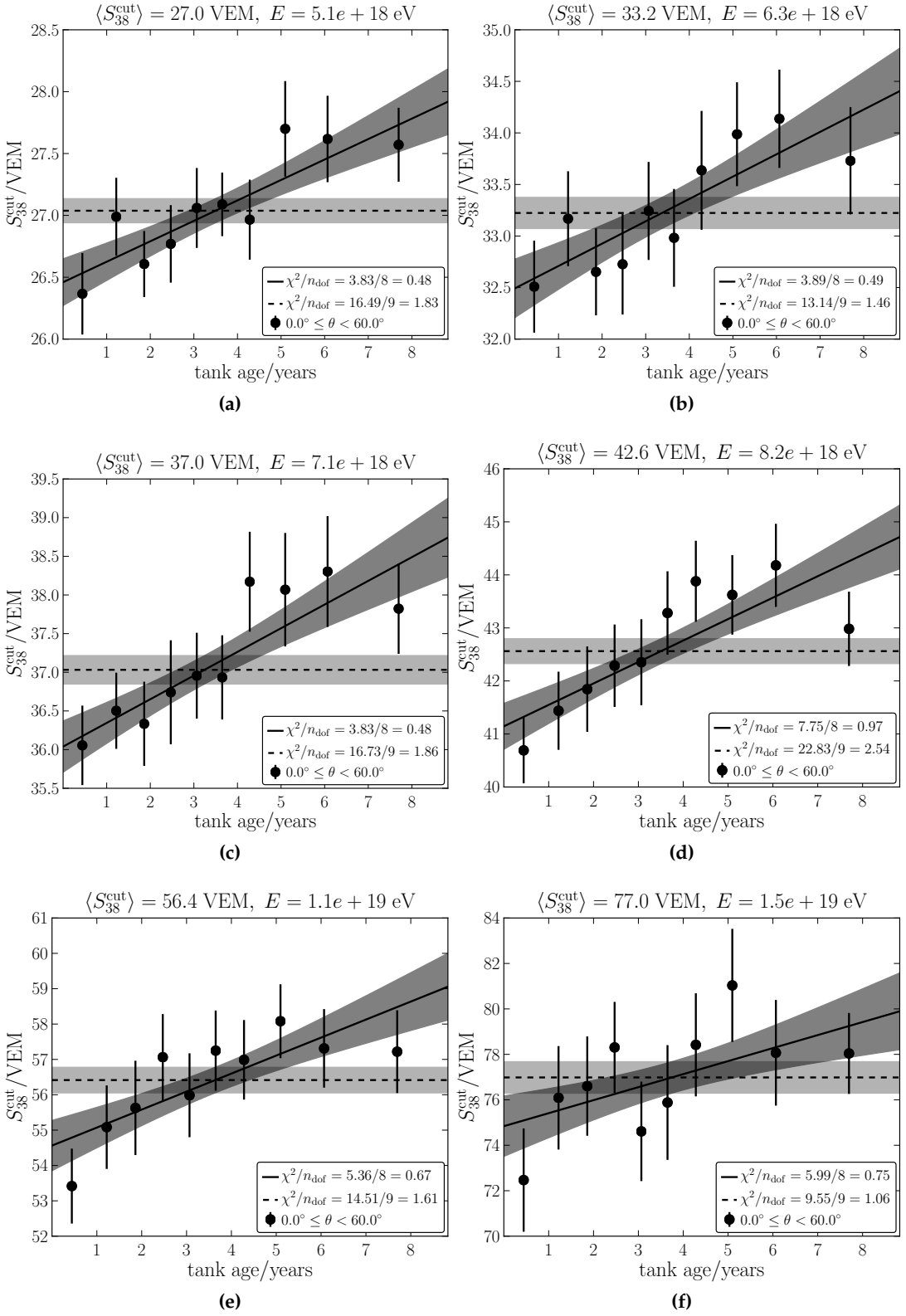


Figure 4.16: S_{38} is shown as a function of tank age for different cut energies. A constant and linear model is fit to the changes of S_{38} with tank age. Both models are depicted with their $1\,\sigma$ uncertainty bands. Judging from the χ^2- and leave-one-out cross-validation-values, the linear model is always favored. Translated into a statistical significance, the difference to a constant amounts to $(3-4)\,\sigma$.

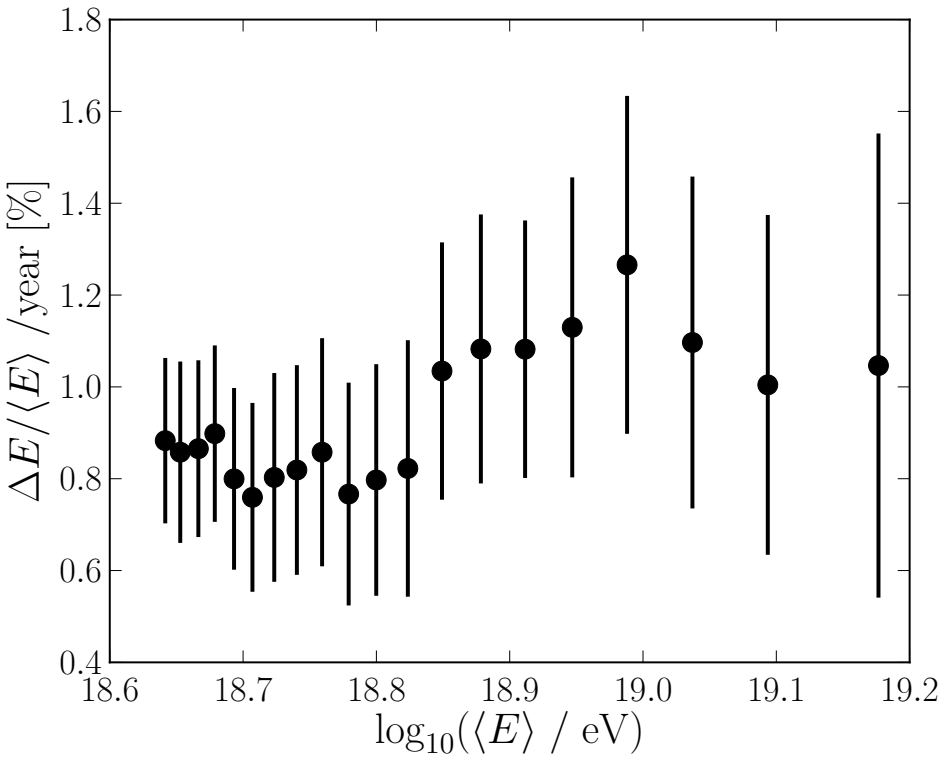


Figure 4.17: The relative change of the SD-1500 energy, assembled with the CIC method with a fixed attenuation function. At every intensity, the analysis is performed separately for each tank age bin.

Change of the angular shape with tank age

In the previous sections, the change of S_{38} with tank age was presented. There are corresponding fits of the attenuation function for each individual fit of S_{38}. In this section, the change of the attenuation function with tank age is investigated. The attenuation functions are shown in Fig. 4.19 for three arbitrary energies of 5×10^{18} eV, 8×10^{18} eV and 1.5×10^{19} eV corresponding to the fits in Fig. 4.18a, Fig. 4.18d and Fig. 4.18f. For each energy, the functions are plotted together with the reference function and as residual of the absolute values with respect to the reference function. In order to check for a distinct time dependence of the attenuation function at a fixed energy, it is more suitable to look at the absolute differences to the reference attenuation function $f^{\text{ref}}_{\text{CIC}}(\theta)$:

$$\Delta_{\text{abs}} = |f_{\text{CIC}}\theta - f^{\text{ref}}_{\text{CIC}}(\theta)|, \tag{4.28}$$

which are expected to follow a half-normal distribution with expectation $\mathrm{E}[X] = \sigma\sqrt{\frac{2}{\pi}}$ and variance $\mathrm{Var}[X] = \sigma^2\left(1-\frac{2}{\pi}\right)$, where the uncertainty of the original (reference) function is denoted by σ.

For different energies, this is depicted in Fig. 4.19b, Fig. 4.19d and Fig. 4.19f. At each energy, the actual coverage is given together with the expected coverage of 0.68. Judging from these plots, there is no significant change of the attenuation function with time (or energy). Except for some outliers at certain time bins, there seems to be no striking trend. This is also the case for other energies that aren't shown here explicitly.

As a complementary view on the problem, I present the results of another method to analyze a possible dependence of the attenuation function with tank age. The first step is to select all events with zenith angles smaller or equal than 38°. This sub-sample of

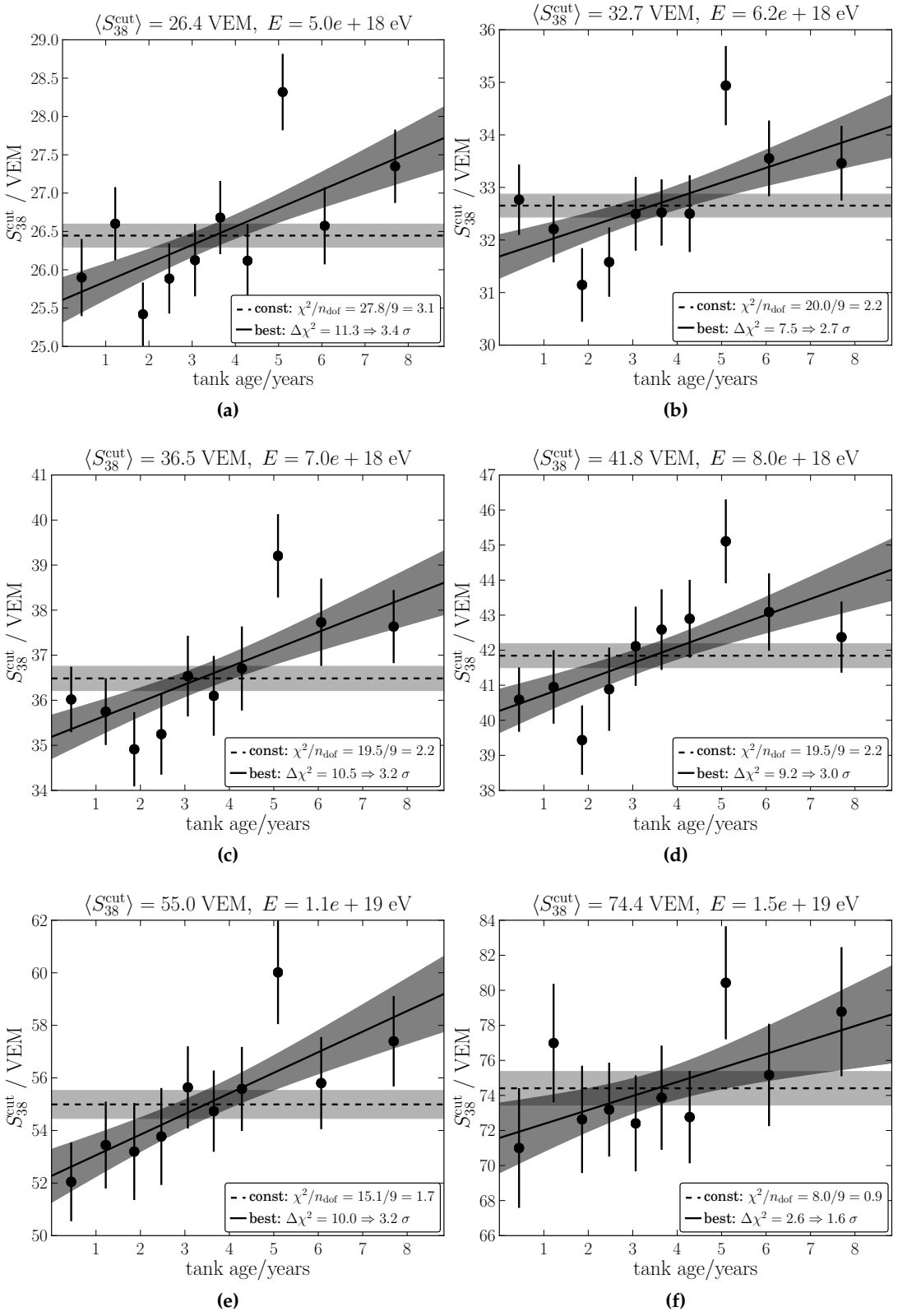


Figure 4.18: S_{38} as a function of tank age for different cut energies. The attenuation is fitted separately for each tank age bin. These plots should be compared to the fits using a fixed angular shape, depicted in Fig. 4.16.

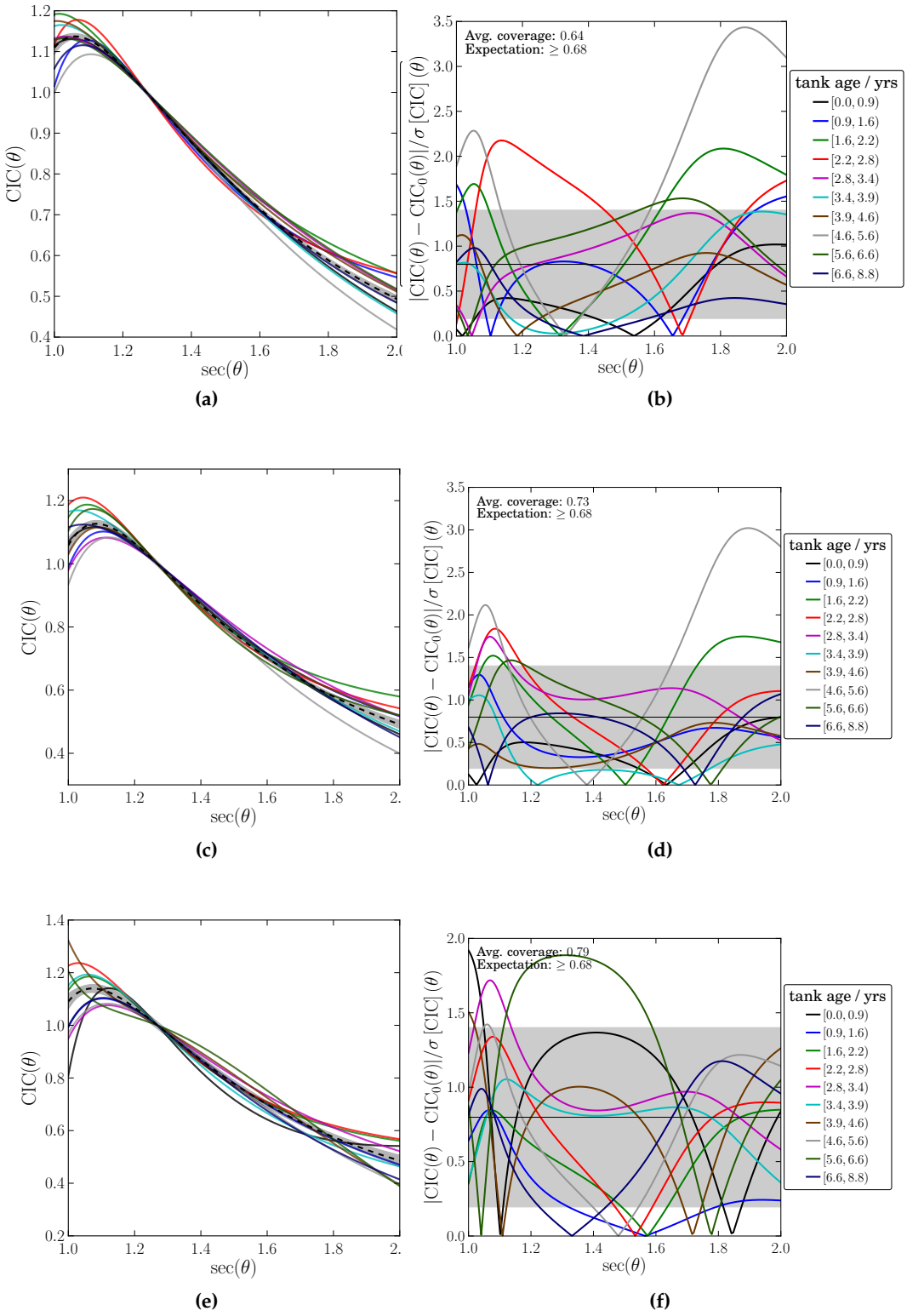


Figure 4.19: The resulting attenuation functions for the different tank age bins. Two adjacent plots include the same functions and belong to the same energy. The energies are stated in the text. In the left plots, the reference attenuation function is depicted with the dashed black line for comparison. In the right plots, all functions are shown relative to the reference attenuation function.

events consists of N_{38} events in total. Afterwards the tank age range is divided into ten bins that include the same amount of events $\frac{N_{38}}{10}$. Then I check if there is a change in the number of events above 38° with respect to these fixed tank age bins. Advantages of this analysis are that it is agnostic to the knowledge of the exposure or the CIC assumption. The results for different energy ranges are shown in Fig. 4.20. Only the first plot of events in the trigger regime below 3×10^{18} eV reflects a statistically significant deviation from a constant behavior. This supports the hypothesis of a weaker response to a signal component affecting the trigger rate (compare to Section 4.6.2). At other energies, no statistically significant behavior is apparent.

Number of saturated stations

An idea to study a possible change of the detector response to the electromagnetic component of air showers is to look at the number of saturated stations as a function of time and tank age. As the PMT saturation is mainly caused by strong electromagnetic signals, a weaker response to these signals is expected to result in a decrease of the number of saturated stations.

The average number of low-gain saturated stations in events with energies above 3×10^{18} eV is depicted as a function of tank age in Fig. 4.21. The events are split into three sub-samples depending on the number of freezes that a tank has experienced. The information about the freezing dates are taken from [158]. According to that information, the first freezing occurred at 13. July 2007 and the second one at 21. July 2010. While there is a quite obvious decrease of the average number of saturated stations for tanks that experienced two freezes, there is no clear trend with tank age except for that. The average decreases with the number of freezes though, which indicates that the freezing events might cause some kind of damage to the tanks.

Trigger rates

In analogy to the previous section, it is of interest to look for a change in the frequency of certain triggers as a function of tank age and freezing events. For vertical events, the ToT is the most common trigger type, fulfilled in more than 95 % of the cases. By definition, the ToT is optimized to trigger long traces as they are produced by electromagnetic particles, while the threshold triggers search for short, strong peaks as produced by muons. In case of a decreasing response to the electromagnetic component of air showers, one would expect a decrease of the ToT frequency compared to the threshold triggers. The trigger frequencies for the different trigger types are plotted in Fig. 4.22. The fitted lines are mainly plotted to guide the eyes. However, the results are compatible with our expectations. The ToT frequency decreases with increasing age of the tanks. Note that the age is again defined as the average age of the three hottest tanks. However, for a fixed tank age, there seems to be no obvious trend with the number of freezes, which is not intuitive when taking into account the previous results.

4.6.3 Conclusions

The application of the Constant Intensity Cut method to different time and tank age intervals allows to conclude that there is a moderate change of the SD detector response with time. That trend is observed in SD-1500 data. On average, the energy scale changes by 1 % per year. The relative change shows a slight increase with energy itself, which is statistically not significant. The attenuation function f_{CIC} does not seem to change significantly as a function of time or tank age. Analyses of the number of low gain saturated stations or

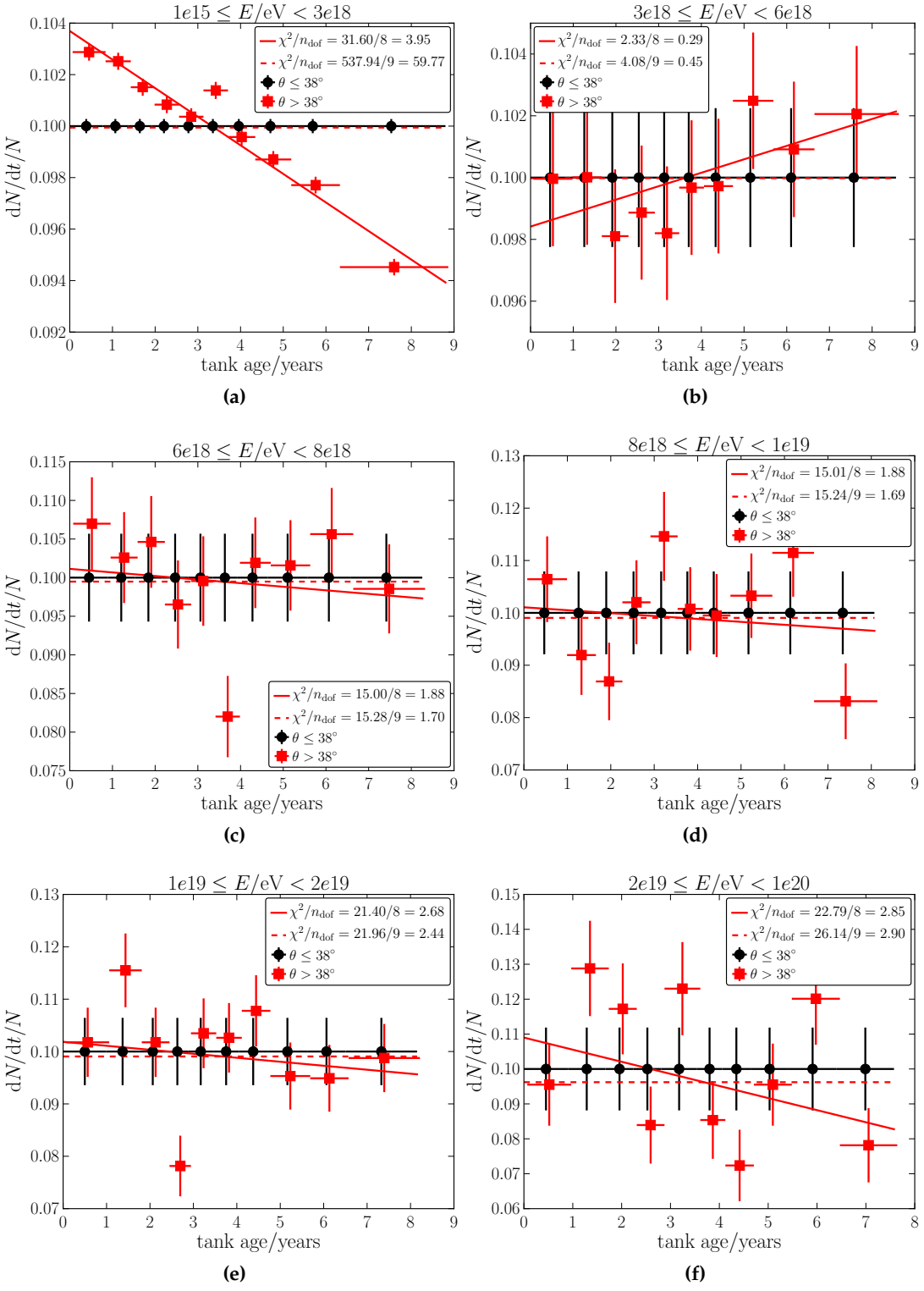


Figure 4.20: Checks of a tank age dependence of the angular shape with two sub-samples of SD-1500 events: events with zenith angles smaller or equal to 38° (black circles) and events with zenith angles above 38° (red squares). Different energy ranges are plotted as stated in the plot titles. The red points are slightly shifted along the x-axis for visibility. For each energy range both a constant and linear model are fitted to the red points to analyse the tank age dependency. Except for the lowest energies, there is no statistically significant deviation from a constant.

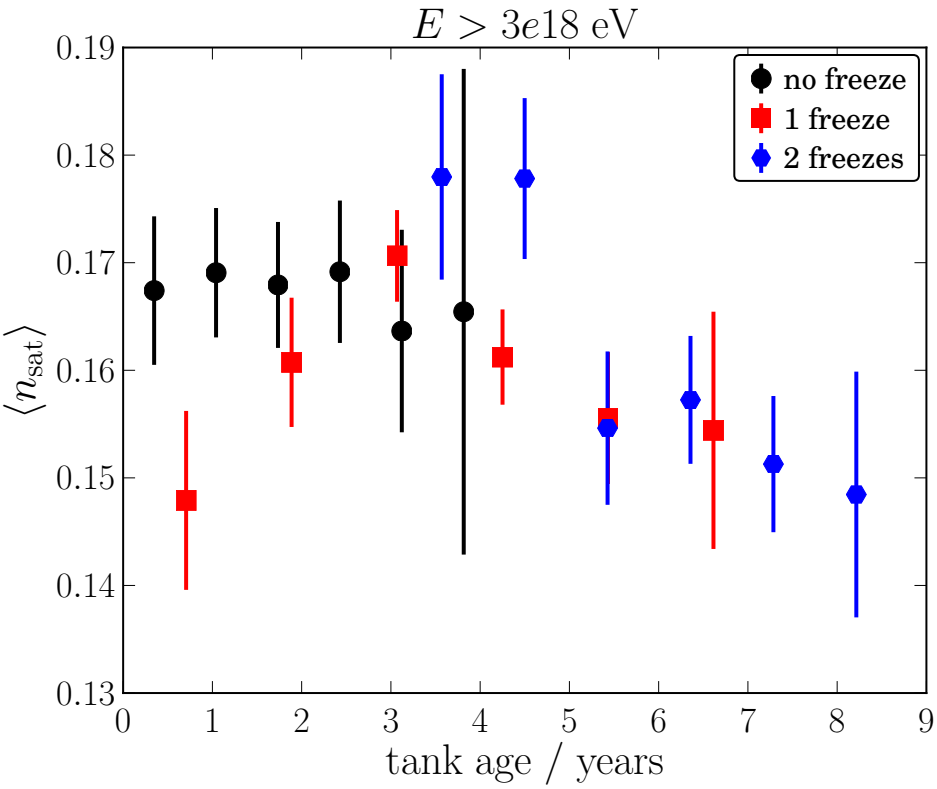
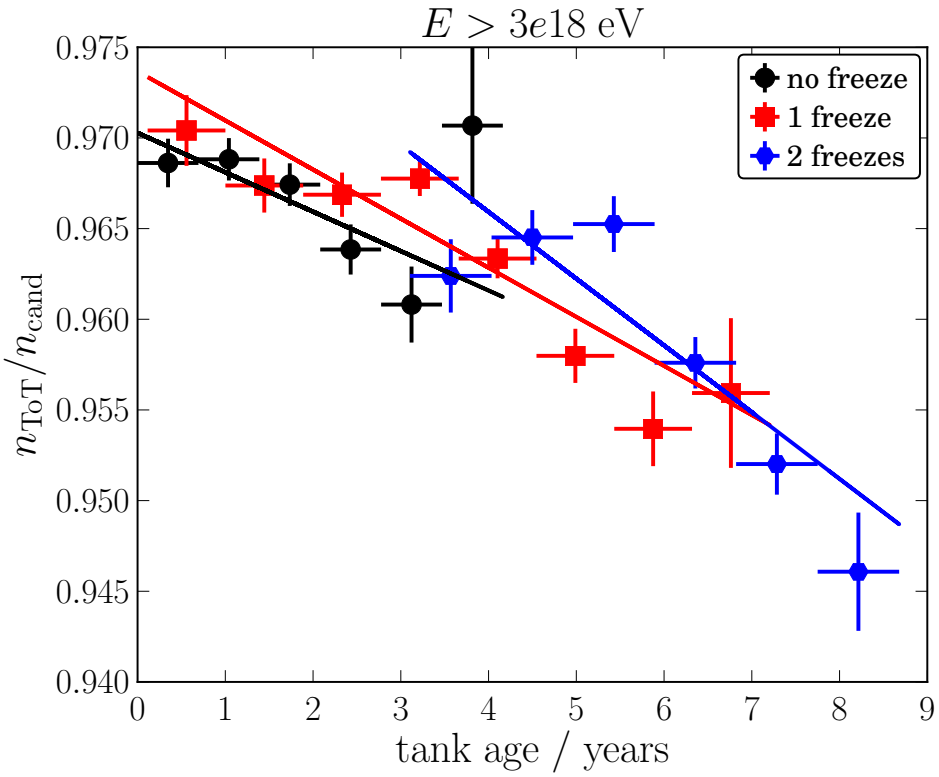


Figure 4.21: The average number of low gain saturated stations in an event as a function of tank age.

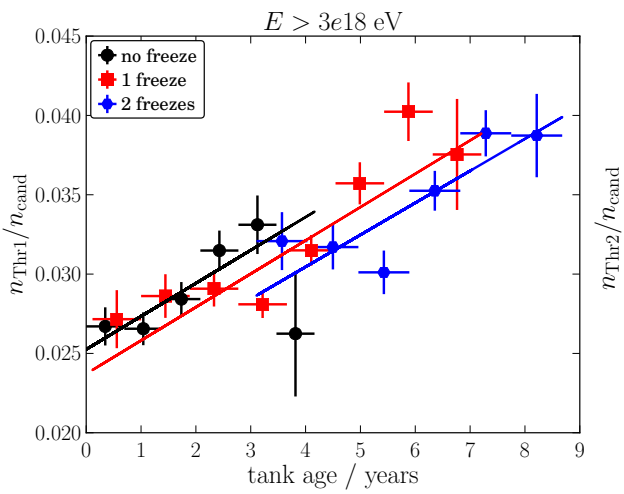
trigger rates as functions of tank age support a change of the detector response with tank age. These results are correlated with the change of AoP[2] with time and, in particular, indicate a shift in the detector response to the electromagnetic component.

A preliminary update of the time dependence was obtained from data of the full time period (see Section 4.1). Results are shown in Fig. 4.23. Shown is the ratio of the number of measured events N to the number of expected events $\hat{N}$ as a function of time. The expected number of events $\hat{N}$ is derived from the total number of events and the fractional exposure for a respective time period. Only events with $E > 10^{18.8}$ eV are shown in Fig. 4.23a, while profiles for two energy thresholds are depicted in Fig. 4.23b. It is interesting to note that the last three bins are very well described by a constant and the first few years of measurement deviate most strongly from an overall constant behavior. There is no significant change in the time dependence for different energy thresholds.

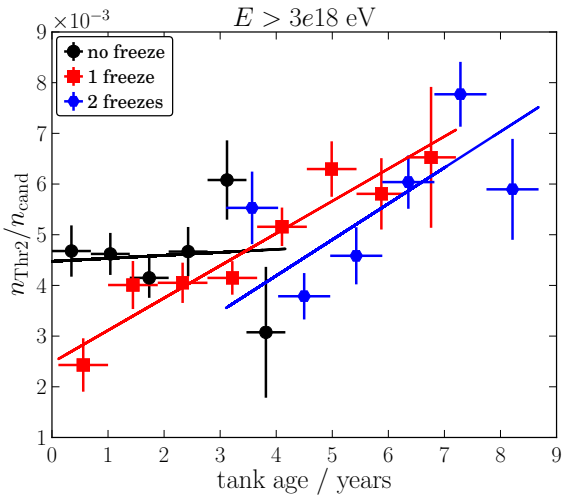
[2] area-over-peak



(a) Time-over-threshold

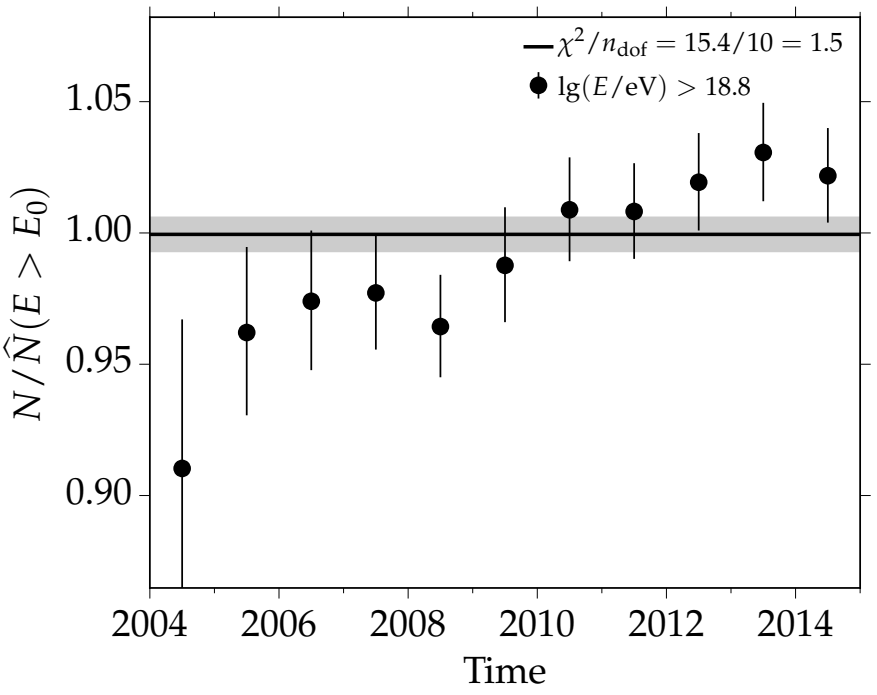


(b) Threshold 1

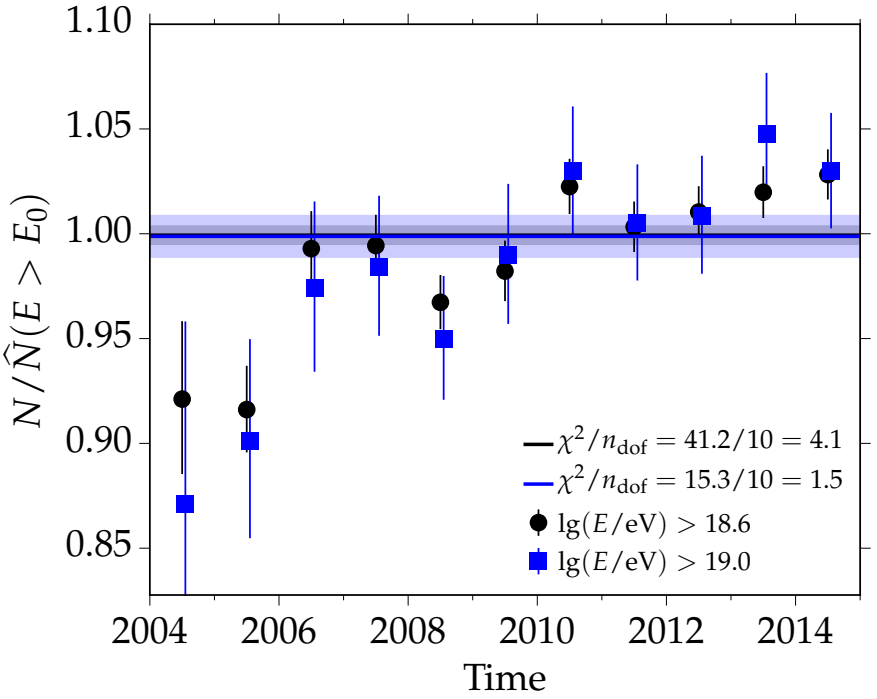


(c) Threshold 2

Figure 4.22: Relative frequency of trigger types as functions of tank age and for different numbers of freezing events. Fitted lines are plotted to guide the eyes.



(a) Events with $E > 10^{18.8}$ eV



(b) Two different energy thresholds

Figure 4.23: Time dependence of the event rate and the energy scale of events measured with the SD-1500.

4.7 Energy calibration

In order to reconstruct primary energies from SD measurements, a calibration with an independent measurement needs to be carried out, if one does not want to rely solely on simulations. The FD provides a reliable measurement of the calorimetric energy of an air shower. Thus, I use a subset of events - called golden hybrid events - that were observed by both the SD and FD in order to calibrate the SD energy estimates $S = S_{\mathrm{SD}} = \{S_{35}, S_{38}\}$ to the energy E_{FD}.

Thus, the goal is to find the energy calibration function $S_{\mathrm{SD}} = S(E)$. In principle, the shower size S is a function of the primary mass A. This dependency cannot be taken into account explicitly because A is not known on an event-by-event basis. To still ensure an unbiased calibration function, the event selection needs to ensure an unbiased mass composition. For the SD measurement, a restriction to primaries with energies above full efficiency fulfills this criterion. For the FD measurement the definition of a fiducial FoV is necessary. This is briefly described in Section 4.7.2.

In the following section, I want to give a short overview of the simplified likelihood method that I am using to obtain the energy calibration function for events recorded with the SD-750 and the SD-1500. The results for the current sets of data are given in Section 4.7.5. Essential quality cuts are described in Section 4.7.2. The agreement of the derived SD-750 and SD-1500 energy scales is analyzed in Section 4.7.6 using a subset of events that were independently reconstructed with both arrays and their respective analyses.

4.7.1 Simplified likelihood method

The foundation of the likelihood fit is the p.d.f. $f(S_{\mathrm{SD}}, E_{\mathrm{FD}})$. It includes all components critical to the physics of the measurements.
The underlying distribution is the distribution of events detected by the FD, the hybrid distribution $h(E, \theta)$. This is the *true* distribution of events prior to influences of detector resolutions or efficiencies. One can write

$$f_1(E, S, \theta) = \delta(S - S(E))\; h(E, \theta), \tag{4.29}$$

with the Dirac delta distribution $\delta(S - S(E))$. The introduction of shower-to-shower fluctuations of SD shower sizes leads to:

$$f_2(E, S_{\mathrm{sh}}, \theta) = \int_0^\infty \mathrm{d}S\; s(S_{\mathrm{sh}}|S, \theta)\; \delta(S - S(E))\; h(E, \theta). \tag{4.30}$$

The shower size S_{sh} deviates from the true value S due to fluctuations in the shower development. These are either due to fluctuations in the height of first interaction or due to fluctuations in the interaction and propagation of secondary particles in the air shower. Up to now, f_2 does not contain any kind of detection efficiencies or detector resolutions. I am introducing the detector kernels $k_{\mathrm{SD}}(S_{\mathrm{SD}}|S_{\mathrm{sh}}, \theta)$ and $k_{\mathrm{FD}}(E_{\mathrm{FD}}|E, \theta)$. Including those resolution models into the previous equation gives us

$$f_3(E_{\mathrm{FD}}, S_{\mathrm{SD}}, \theta) = \int_0^\infty \mathrm{d}E \int_0^\infty \mathrm{d}S_{\mathrm{sh}}\; k_{\mathrm{FD}}(E_{\mathrm{FD}}|E, \theta)\; k_{\mathrm{SD}}(S_{\mathrm{SD}}|S_{\mathrm{sh}}, \theta)\; f_2(E, S_{\mathrm{sh}}, \theta). \tag{4.31}$$

Finally, one has to consider the detection efficiencies of both SD and FD: $\epsilon_{\mathrm{SD}}(S_{\mathrm{SD}}, \theta)$ and $\epsilon_{\mathrm{FD}}(E_{\mathrm{FD}}, \theta)$. These are both functions of the signals after the detector resolution, since these determine if more or less signal is observed. The complete and normalized p.d.f. reads:

$$f(E_{\mathrm{FD}}, S_{\mathrm{SD}}, \theta) = \frac{\epsilon_{\mathrm{SD}}(S_{\mathrm{SD}}, \theta)\; \epsilon_{\mathrm{FD}}(E_{\mathrm{FD}}, \theta)\; f_3(E_{\mathrm{FD}}, S_{\mathrm{SD}}, \theta)}{\int_0^{\pi/2} \mathrm{d}\theta \int_0^\infty \mathrm{d}E_{\mathrm{FD}} \int_0^\infty \mathrm{d}S_{\mathrm{SD}}\; \epsilon_{\mathrm{SD}}(S_{\mathrm{SD}}, \theta)\; \epsilon_{\mathrm{FD}}(E_{\mathrm{FD}}, \theta)\; f_3(E_{\mathrm{FD}}, S_{\mathrm{SD}}, \theta)}. \tag{4.32}$$

This p.d.f. gives a full description of the measurement process and is explained in more detail in [161]. It is computationally very expensive though and minimization processes take a long time. A few reasonable approximations are introduced in order to simplify the p.d.f. and speed-up the minimization. These approximations are described in detail in [162, 163].

The first approximation concerns the treatment of trigger efficiencies of FD and SD. Trigger efficiency curves are replaced by binary energy thresholds above which full detection efficiency is ensured for both SD and FD. These thresholds of full efficiency are 3×10^{17} eV for the SD-750 and 3×10^{18} eV for the SD-1500.

The second approximation simplifies the convolution of shower-to-shower and SD sampling fluctuations. In the approximation, both are modeled as independent normal distributions. For a detailed description, I refer to [162, 163].

The third approximation is of most practical use and denotes the description of the hybrid distribution $h(E,\theta)$ with a bootstrap estimate. This removes the need to parametrize this distribution. The p.d.f. is estimated directly from data as

$$h(E,\theta) \approx \frac{1}{N}\sum_{i}^{N}\delta(E - E_{\mathrm{FD},i})\delta(\theta - \theta_i), \tag{4.33}$$

with the index i running over all N hybrid events that pass the FD quality cuts specified in Section 4.7.2. It is the distribution of potential hybrid events that fall into the array of the respective SD, but an independent SD trigger is not required.

The impact of the approximations on the accuracy and precision of the resulting energy calibration parameters was studied and no significant difference to the exact method was found [162].

The final simplified log-likelihood function reads:

$$\log\mathcal{L} = \sum_{k}\log\left[\sum_{i}\frac{e^{-\frac{1}{2}\frac{\left(E_{\mathrm{FD},k}-E_{\mathrm{FD},i}\right)^2}{\sigma^2_{\mathrm{FD},i}}}}{\sigma_{\mathrm{FD},i}}\frac{e^{-\frac{1}{2}\frac{\left(S_{\mathrm{SD},k}-S(E_{\mathrm{FD},i})\right)^2}{\sigma^2_{\mathrm{SD}}(S(E_{\mathrm{FD,i}}),\theta_i)+\sigma^2_{\mathrm{sh}}(E_{\mathrm{FD},i})}}}{\sqrt{\sigma^2_{\mathrm{SD}}(S(E_{\mathrm{FD,i}}),\theta_i)+\sigma^2_{\mathrm{sh}}(E_{\mathrm{FD},i})}}\right]. \tag{4.34}$$

The index k runs over selected golden hybrid events, while the index i iterates over hybrid events entering the bootstrap estimate of the hybrid p.d.f. in Eq. (4.33). The actual calibration function that enters the minimization is $S(E_{\mathrm{FD}})$. In the next section, I will describe the selection criteria for events that contribute to the energy calibration. Results for data measured with SD-750 and SD-1500 are given in Section 4.7.5.

4.7.2 Quality selection

Events that pass the SD selection criteria, as discussed in Section 4.1, are further considered for the energy calibration method. Among these events, only golden hybrid events that pass a strict FD selection are used to derive the calibration function. The SD and FD quality cuts are listed in Appendix D.1.2. Due to their importance, the FD quality cuts are briefly described in the following. A more detailed description of quality cuts and their derivation is given in [21]. The names of the cuts are taken from the Offline framework.
The cuts are (*Name* (value)):

- *eyeCut* (100000) - only used for analysis of SD-750 data
  Select and use only showers reconstructed with the HeCo merged eye. This is a merging of the individual Coihueco and HEAT eyes.

- *hasMieDatabase*
  Atmospheric conditions are constantly monitored at Auger. The information about atmospheric parameters critical to the reconstruction of the longitudinal profile of air showers are regularly analyzed, processed and stored in dedicated databases. This selection criterion enforces the existence of these database entries for the event.

- *minLgEnergyFD* (10^{-20} eV)
  The reconstructed energy is zero in very rare cases, like when the profile reconstruction failed. These events are rejected.

- *skipSaturated*
  Pixels of the PMT camera saturate above a certain threshold of light intensity. The time-dependent light intensity from the shower profile cannot be reliably reconstructed in these cases. Events with saturated pixels are thus discarded.

- *badFDPeriodRejection*
  Events within periods with known issues in the FD operation are rejected.

- *!badPixels* (1)
  Pixels of the PMT camera occasionally have bad calibration parameters. Events for which this is the case are rejected.

- *minMeanPixelRMS* (17)
  This new criterion was introduced in order to reject events that are recorded during the process of closing the shutters in front of the FD telescope. The variance in background light is significantly smaller when the shutters are closed. This enables a selection based on this quantity.

- *maxVAOD* (0.1)
  The amount of aerosols in the atmosphere is quantified in terms of VAOD[3]. A value below 0.1 is required to ensure low air contamination with aerosols.

- *LidarCloudRemoval* (25)
  Clouds are monitored with the LIDAR systems at Auger [164]. Only events with a cloud fraction below 25 % are considered.

- *MinCloudDepthDistance* (-50, 50)
  This cut ensures that the maximum of shower development was not too close to clouds.

- *MaxCloudThickness* (100)
  Events with too thick clouds are rejected (in units of $\mathrm{g\,cm^{-2}}$)

- *xMaxObsInExpectedFOV* (40, 20)
  The FD telescopes have a limited FoV. The reconstructed shower maximum is required to be within this FoV. Events, for which only the rising or falling edge of the longitudinal profile was recorded, are rejected.

- *xMaxError* (40)
  The reconstruction uncertainty $\sigma[X_{\mathrm{max}}]$ is required to be below $40\,\mathrm{g\,cm^{-2}}$.

[3] vertical aerosol optical depth

- *energyTotError* (0.18)
  The relative uncertainty in the reconstructed total energy is required to be below 18 %. This and the previous selection ensure well-reconstructed events.

- *profileChi2Sigma* (2.5, -1.1)
  The idea behind this rather sophisticated cut is to reject profiles that do not have a distinctive Gaisser-Hillas shape. The χ^2 probability of a Gaisser-Hillas fit to the longitudinal profile is compared to the probability of a linear fit. Profiles for which the Gaisser-Hillas description seems unlikely are rejected.

- *maxDepthHole* (20)
  Events with a hole or holes larger than $20\,\mathrm{g\,cm^{-2}}$ in the longitudinal profile are discarded.

- *maxCoreTankDist* (750)
  The maximum allowed distance between the reconstructed shower core (hybrid) and the next SD station is 750 m. When requiring the 6T5 criterion for an SD event, this is always fulfilled, except if the shower core is largely misreconstructed.

- *FidFOVICRC13* (40, 20)
  The fiducial FoV selection ensures that the (priorly unknown and true) X_{max}-distribution is not distorted by the limited FoV of the telescopes. This is visualized in Fig. 4.24. Depending on the mass of the primary particle, showers develop more shallow or deeply in the atmosphere. The true overall X_{max}-distribution is equal to the sum of the distribution for individual masses. In this selection, only geometries - relating to the incoming direction and distance of showers - are allowed for which the effective FoV of telescopes has a minimal impact on the measurement. Speaking in extremes, only geometries are allowed for which the measurement of profiles initiated by proton or iron primaries would be equally possible. Thereby, the selection ensures that the primary composition is not distorted by the measurement. A very nice explanation is given in Fig. 2 of [21]. Details about the construction of the fiducial geometry are explained there as well. This is a very strict cut that removes about two thirds of the remaining events after the previous cuts. By its nature, the cut depends on the other quality cuts and needs to be re-tuned once larger changes to them are made. Further details are also found in [165].

A large number of these cuts is optimized with data itself. The figure of merit is the flatness of the relative energy difference $\frac{E_{\mathrm{SD}}-E_{\mathrm{FD}}}{E_{\mathrm{FD}}}$. This quantity is studied for different realizations of cut values, and the optimal cut value is chosen at the point where biases start to appear. The event statistics are maximized as much as possible in this procedure. More details are given in Appendix D.1.2.

4.7.3 Resolutions and biases

The FD energy resolution and the resolution of the SD energy estimates are essential for the determination of the calibration function. They both enter the likelihood in Eq. (4.34). A good estimate of the uncertainties $\sigma[E_{\mathrm{FD}}]/E_{\mathrm{FD}}$ and $\sigma[\hat{S}]/\hat{S}$ is given by the reconstruction uncertainties. These include statistical (fit) and systematic components, for example the uncertainty in an energy estimate due to variations in the slope of the LDF. Shower-to-shower fluctuations are not included. A way to quantify the total resolution, including shower-to-shower fluctuations, is to use air shower simulations and to analyze the migration matrices or resolution models (see Section 4.10). The reconstruction uncertainties obtained

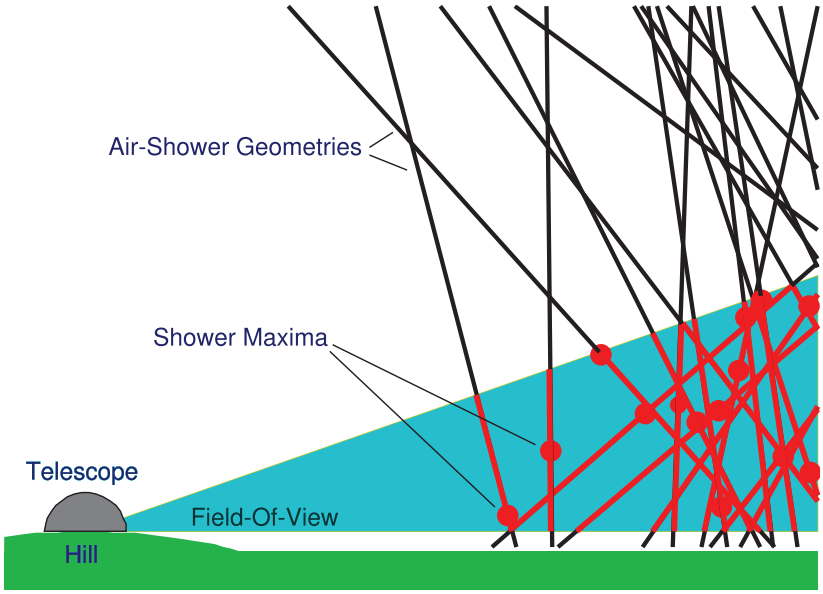


Figure 4.24: A sketch of different showers reaching the field of view of a fluorescence telescope. Only certain geometries can be measured (taken from [166]).

from data itself are shown in Fig. 4.25. The plots Fig. 4.25a and Fig. 4.25c give the relative resolution of E_{FD} for SD-750 and SD-1500 data, respectively. The profiles are given as a function of logarithmic FD energy and for different zenith angles. Overall, the resolution of E_{FD} is weakly dependent on energy and zenith angle. An average value of 7.6 % gives a reasonable description. The resolution is well constrained and below 10 % for almost all events. As visible in the distributions, only a few events show larger uncertainties. The plots Fig. 4.25b and Fig. 4.25d illustrate the relative resolutions of S_{35} and S_{38}. For SD-750 data, the resolution is slightly above 10 % at 3×10^{17} eV and drops below 5 % above 3×10^{18} eV. At largest zenith angles, the resolution is on average 1 % to 2 % larger than at small zenith angles. The behavior is very similar for SD-1500 data. Here, the resolution is around 12 % at 3×10^{18} eV and drops below 5 % above 3×10^{19} eV. The dependence on zenith angle is less pronounced as in the case of SD-750 data but still visible. All the energy and zenith angle dependencies are taken into account during the estimation of the calibration function.

Reconstruction biases on the SD and FD side need to be quantified and taken into account for the energy calibration. No significant biases were found in the analysis of SD-1500 data. However, there is a significant bias of the reconstructed FD energy for SD-750 data. This bias was estimated with simulations and is on the order of −20 % at 10^{17} eV, −10 % at 3×10^{17} eV and less than −1 % above 3×10^{18} eV. At the lowest energies, this is a highly significant bias. The reconstructed calorimetric energy is underestimated with respect to the Monte Carlo energy. At the current stage, it is unclear if the source of the bias is a reconstruction issue or a problem with the FD simulation using the Offline framework. Thus, it is not clear if the full bias is present in data itself. Ongoing studies indicate that half of the bias might be present in data. Therefore, the strategy chosen for this work and [86] is to correct for half of the energy bias and add the impact of the remaining half as a systematic uncertainty on the derived flux. The correction function for half of the energy bias is illustrated in Fig. 4.26a. The systematic uncertainty on the flux measurement from the bias correction is indicated in Fig. 4.26b. It is on the order of 10 % at 3×10^{17} eV and decreases linearly at higher energies. Details and a comparison to other contributions to the overall flux uncertainty are presented in Section 4.10.

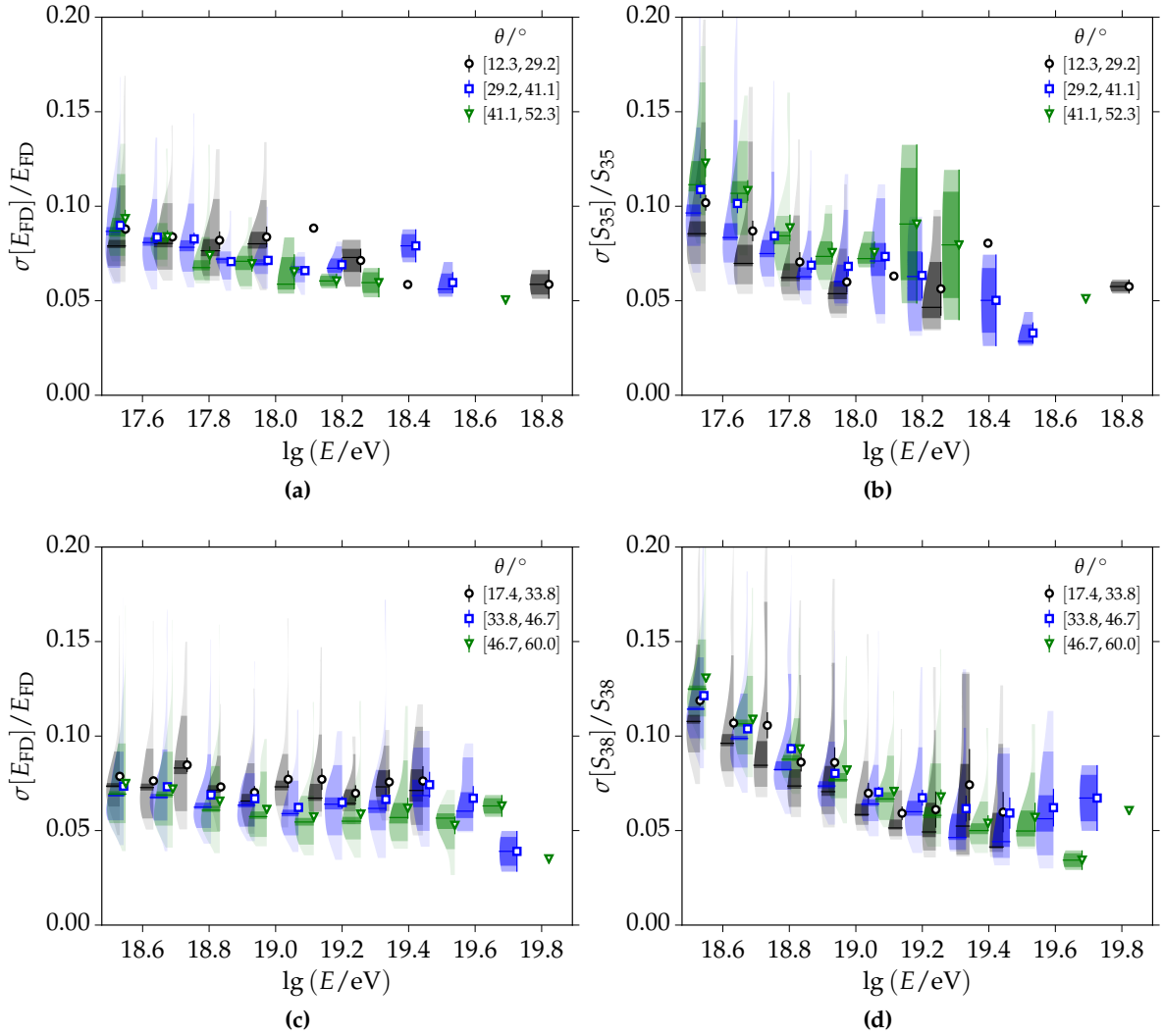


Figure 4.25: (a) Relative FD energy resolution as a function of logarithmic energy. The energy range and events measured with the SD-750 are considered. (b) The relative resolution of S_{35} as a function of $\lg E$. (c) The resolution of E_{FD} as a function of $\lg E$ for SD-1500 energies. (d) The relative resolution of S_{38} as a function of $\lg E$. In each of the plots, the resolutions are given for different ranges of zenith angle.

Another relevant point are differences in the energy scale between different FD telescopes. While, ideally, all FD eyes should have the same energy scale, it was found that energies of showers observed with the Coihueco eye are systematically larger by 5.3 %. To account for this, these energies are scaled down by 5.3 % in both the analysis of SD-750 and SD-1500 data. The energy differences observed in stereo events are plotted in Fig. 4.27. A direct comparison between Coihueco and HEAT in downward mode was analyzed. In this mode, the telescopes have identical FoVs and observe the same air showers. The energy scales should match exactly. Due to its newer sampling and DAQ, HEAT can be assumed to provide the more accurate energy measurement. Thus, energies measured with Coihueco are scaled down. This has the further advantage that energies of showers observed with the other eyes do not need to be scaled up. The remaining differences are then compatible to zero within statistical and systematic uncertainties.

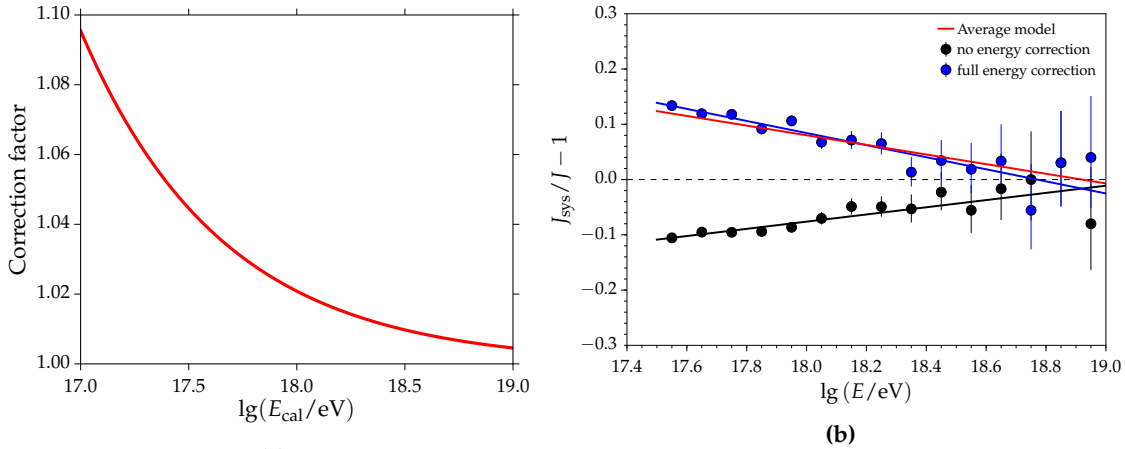


Figure 4.26: (a) Correction function to describe half of the FD energy bias as observed in simulations. SD-750 data are corrected for this bias. (b) Systematic uncertainty on the SD-750 flux measurement due to the uncertainty in the bias correction.

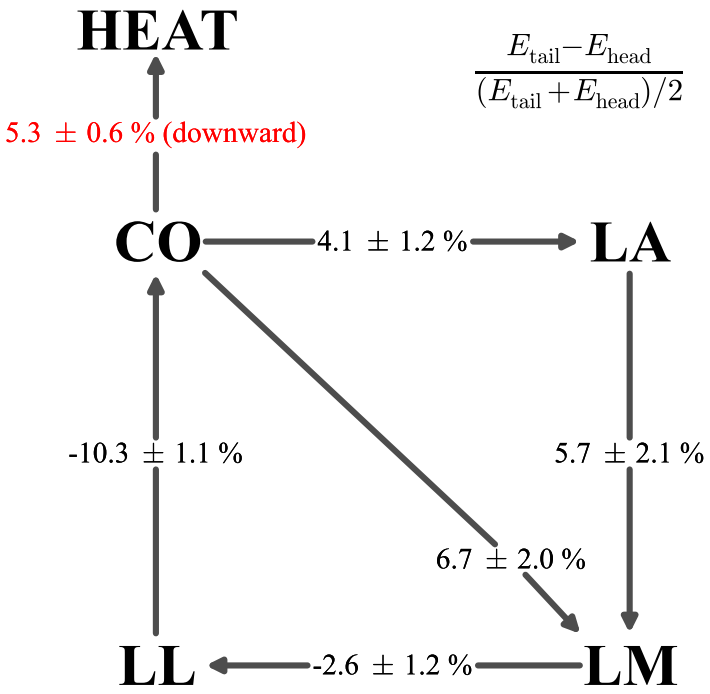


Figure 4.27: The relative energy difference between different FD eyes as estimated from stereo events that were observed with both eyes (taken from [165]).

4.7.4 Calibration function

The relation between FD energy E_{FD} and SD energy estimates $\hat{S} = S_{35}, S_{38}$ is very well-described by the following power-law:

$$E_{FD} = A \left(\frac{\hat{S}}{\text{VEM}} \right)^{B} . \tag{4.35}$$

This relationship can be derived from simple assumptions on the development of air showers (see Section 1.1.4). A change in the primary mass composition with energy should inflict

a change in the calibration parameters with energy. This was studied with inclined showers in [83]. A trend towards heavier composition at the highest energies is apparent, but no significant statement can be made within current event statistics. In vertical events, due to the mixture of the electromagnetic and muonic shower components, a similar statement is much harder to infer. More comments on this are given in Fig. 4.32.

4.7.5 Results

After quality selection 468 SD-750 and 1750 SD-1500 golden hybrid events remain. The energy calibration functions are independently derived by maximizing the likelihood in Eq. (4.34).
The resulting energy calibration function for SD-750 events is:

$$E(S_{35}) = (12.75 \pm 0.62) \times 10^{15}\,\mathrm{eV} \left(\frac{S_{35}}{\mathrm{VEM}}\right)^{(1.006\pm0.013)} . \tag{4.36}$$

The golden hybrid events and the calibration function are plotted in Fig. 4.28a. The resulting function describes data well in the considered energy range. Residuals for different energy and zenith angle intervals are shown in Fig. 4.29c and Fig. 4.29d. No significant deviation from zero is apparent. The event with the Auger ID 112636701200 was manually removed from the calibration procedure. It was a strong outlier in all residuals. Inspection of the event reveals that its core is very close to the border of the SD-750 array, the station geometry is degenerate, and the reconstructed zenith angles between FD and SD-750 differ strongly. Due to these reasons, the event should be discarded. The shower-to-shower fluctuations $\sigma_{\mathrm{sh}}(E)$ are fitted in the calibration procedure. Due to limited event statistics, only a constant is fitted to SD-750 data. This constant is optimized besides A and B in the calibration fit. The result is a 10 % resolution due to shower-to-shower fluctuations as shown in Fig. 4.29a. The statistical uncertainty on the calibration function is visualized in Fig. 4.29b. It is below 4 % at all energies. The energy calibration function derived for SD-1500 data is:

$$E(S_{38}) = (0.180 \pm 0.004) \times 10^{18}\,\mathrm{eV} \left(\frac{S_{38}}{\mathrm{VEM}}\right)^{(1.030\pm0.006)} . \tag{4.37}$$

Equivalently as for SD-750 data, the resulting calibration function is compared to the golden hybrid events in Fig. 4.28b. Residuals with respect to energy and zenith angle are given in Fig. 4.30c and Fig. 4.30d. The model describes all data very well. A linear model was chosen to describe the shower-to-shower fluctuations $\sigma_{\mathrm{sh}}[E]$. The resulting function is plotted in Fig. 4.30a. Fluctuations are 13 % at 3×10^{18} eV and below 10 % above 10^{19} eV. The statistical uncertainty on the calibration function is below 2 % at all energies.

After fitting the calibration functions, the fluctuation of events around the model define the total energy resolution. Profiles of these resolutions for SD-750 data are shown in Fig. 4.31a and Fig. 4.29d. The black, rectangular points represent the total energy resolution, including SD and FD resolutions. Assuming a constant FD energy resolution of $\sigma[E_{\mathrm{FD}}]/E_{\mathrm{FD}} \approx 7.6\,\%$ (see Section 4.7.3) and subtracting it from the points leaves us with the SD-only energy resolution. This contribution is shown with red, circular markers. The average SD resolution is 15 %. This should match with the numbers and plots introduced before. At 3×10^{17} eV, the total resolution is very close to 15 %. Shower-to-shower fluctuations are 10 % (from Fig. 4.29a), while the detector resolution is also 10 % as depicted in Fig. 4.25b. The quadratic sum of these uncertainties gives 14 %, which matches nicely with the total resolution. The zenith angle dependence of the resolution in Fig. 4.31b shows some non-significant deviation from a constant. This behavior hints towards a problem

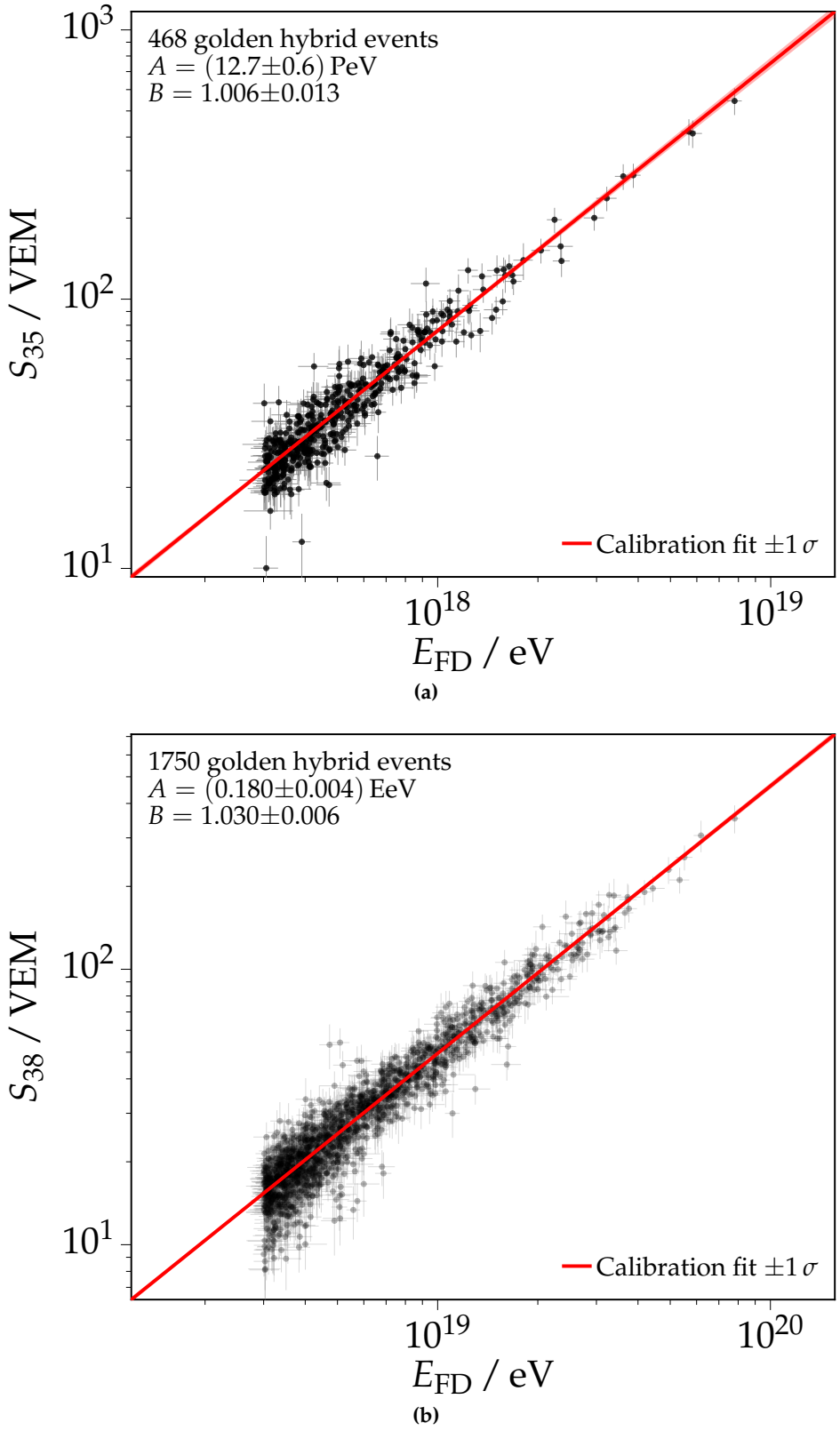


Figure 4.28: (a) Data and fit of the SD-750 energy calibration. (b) Data and fit of the SD-1500 energy calibration. Energy estimates are corrected for weather and geomagnetic effects (see Section 4.3). See Appendix B.6 for a comparison to results without using these corrections.

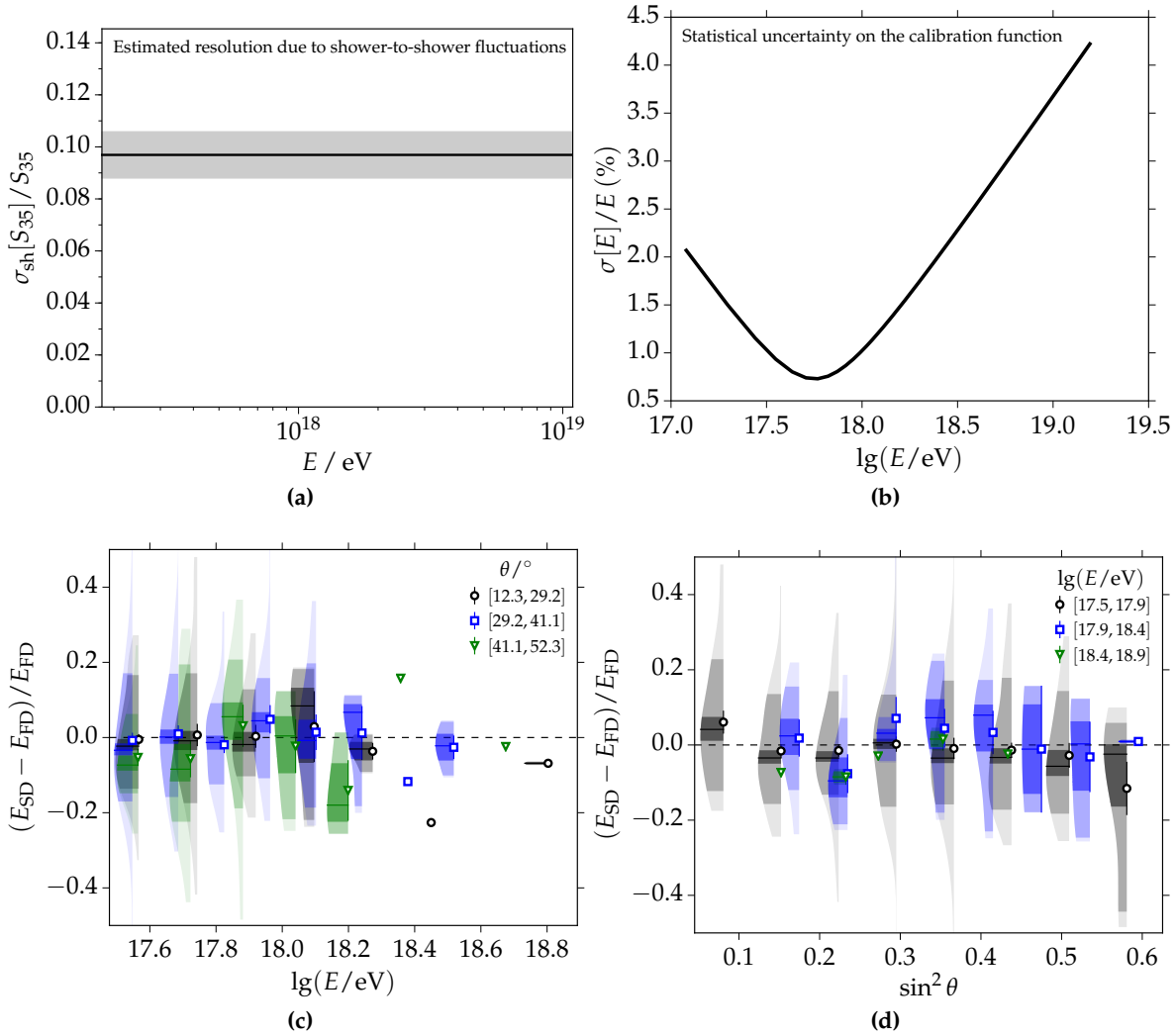


Figure 4.29: (a) Shower-to-shower fluctuations estimated in the energy calibration procedure. A constant model is chosen to describe SD-750 data. The result is a resolution of 10 %. (b) The statistical uncertainty on the calibration function as a function of energy. (c) Residuals between estimated and fitted energies as a function of lg E and for different zenith angles. (d) Residuals between estimated and fitted energies as a function of $\sin^2\theta$ and for different energy ranges.

with either the SD or FD reconstruction at certain zenith angles and should be further observed with more event statistics.

Results on the resolution of SD-1500 events are shown in Fig. 4.31c and Fig. 4.31d. Again, subtracting a constant $\sigma[E_{FD}]/E_{FD} = 7.6\,\%$ leads to the red, circular points. At 3×10^{18} eV, the total resolution is around 17 %. The relative shower-to-shower fluctuations at this energy are 13 % (from Fig. 4.30a) and the detector resolution is 12 %. Summing these contributions quadratically leads to 17.5 %. On average, the SD-1500 energy resolution is 16 %.

4.7.6 Validation with common events

A subset of events is recorded with both SD-750 and SD-1500, as the former represents a nested array within the SD-1500. The stations that are part of the SD-1500 contribute to both reconstructions, while the nested array of stations contributes only to the SD-750

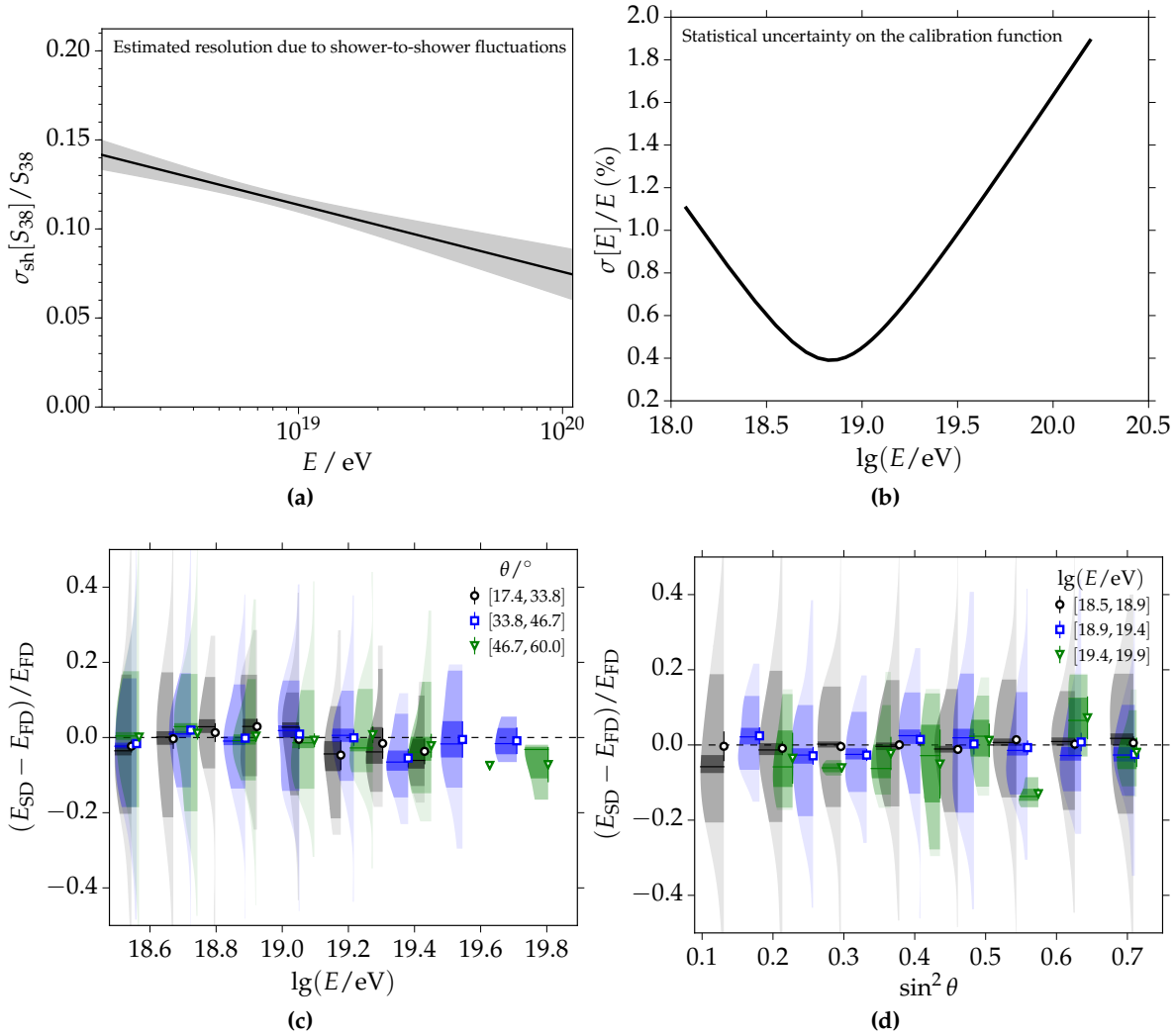


Figure 4.30: (a) Shower-to-shower fluctuations estimated in the energy calibration procedure. A constant model is chosen to describe SD-1500 data. (b) The statistical uncertainty on the calibration function as a function of energy. The uncertainty is less than 2 % at all energies. (c) Residuals between estimated and fitted energies as a function of lg E and for different zenith angles. (d) Residuals between estimated and fitted energies as a function of $\sin^2\theta$ and for different energy ranges.

reconstruction. Otherwise, the reconstruction and analysis of these events are completely separate. The comparison of reconstructed energies of common events provides an important cross-check of the individual energy calibrations. This comparison is shown in Fig. 4.32. Using the energy calibration parameters derived in the previous section, there are 321 common events above the energy threshold for the full efficiency of the SD-1500 of 3×10^{18} eV. A residual comparison of the energies of these events is given in Fig. 4.32a. The energy residuals are divided into five equidistant bins in lg E. The number of events within each energy range is provided below the kernel density profiles in the plot. On average, there is a −8.2 % bias in the overlapping energy range, relatively constant with energy. Energies reconstructed with the SD-750 are systematically smaller. The shaded regions around zero represent different statistical and systematic uncertainties. The largest contribution in gray is the statistical uncertainty on the SD-750 energy calibration, which can only account for half of the deviation at lowest energies. The blue region represents

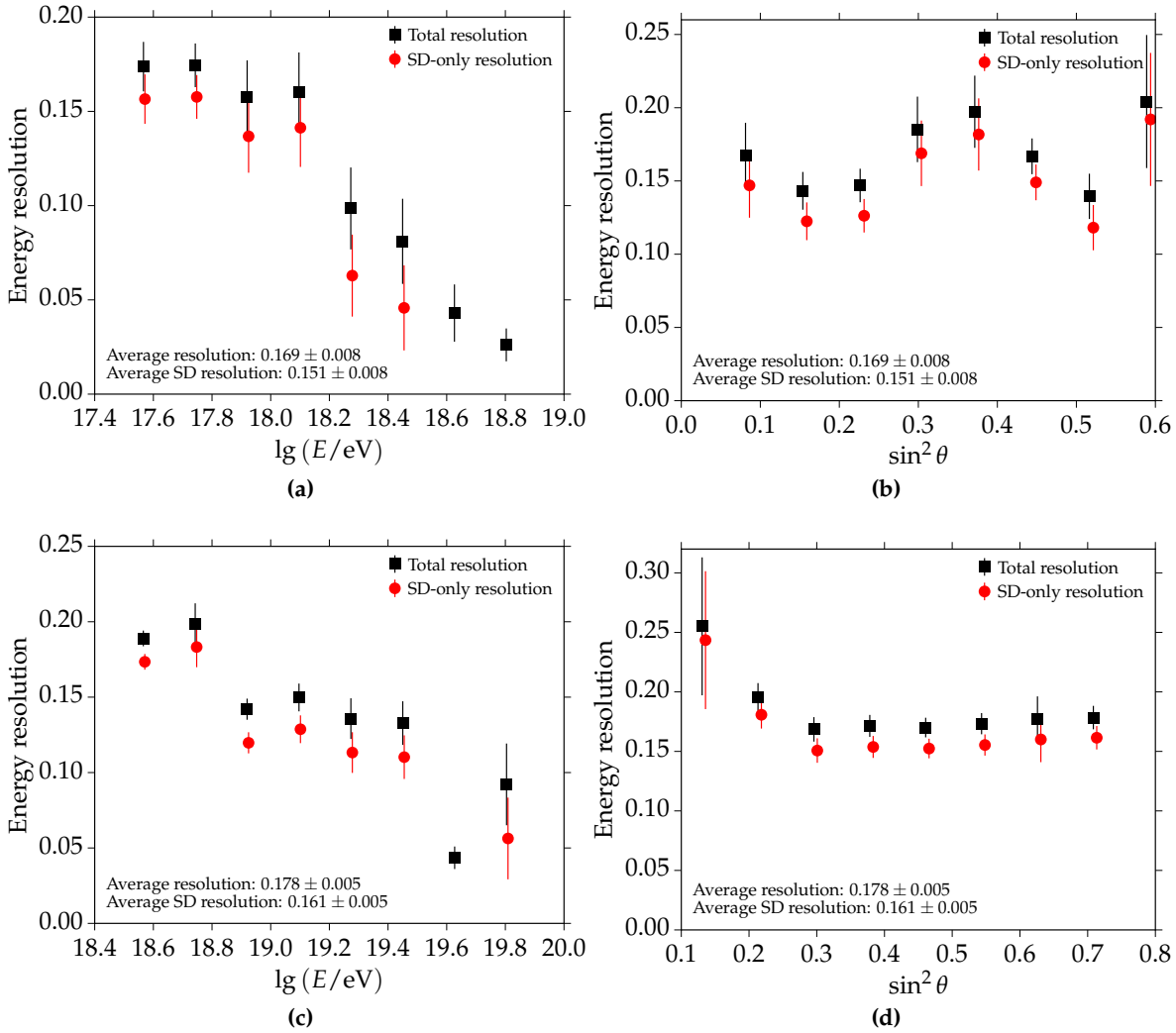
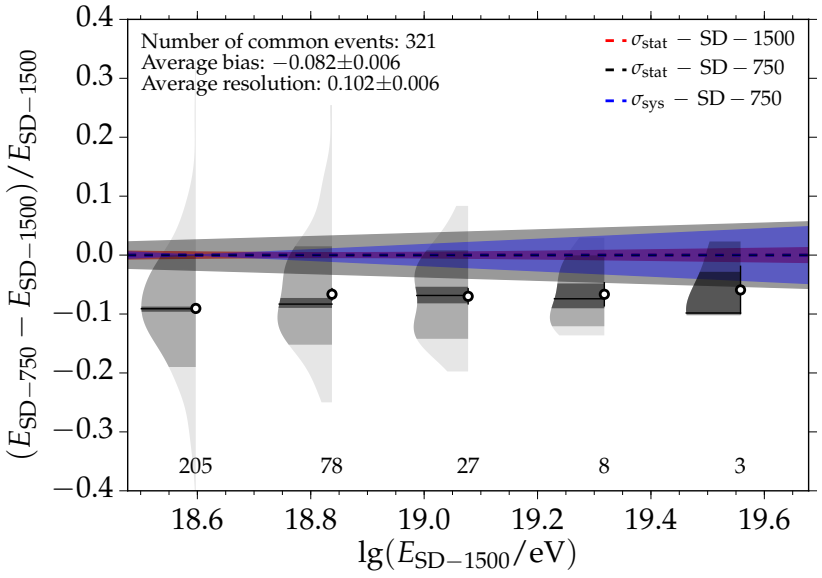


Figure 4.31: Relative energy resolution in SD-750 data as a function of (a) energy and (b) zenith angle. The relative energy resolution in SD-1500 data as a function of (c) energy and (d) zenith angle. The SD-only resolutions are calculated after subtraction of the average FD energy resolution, as further described in the text.

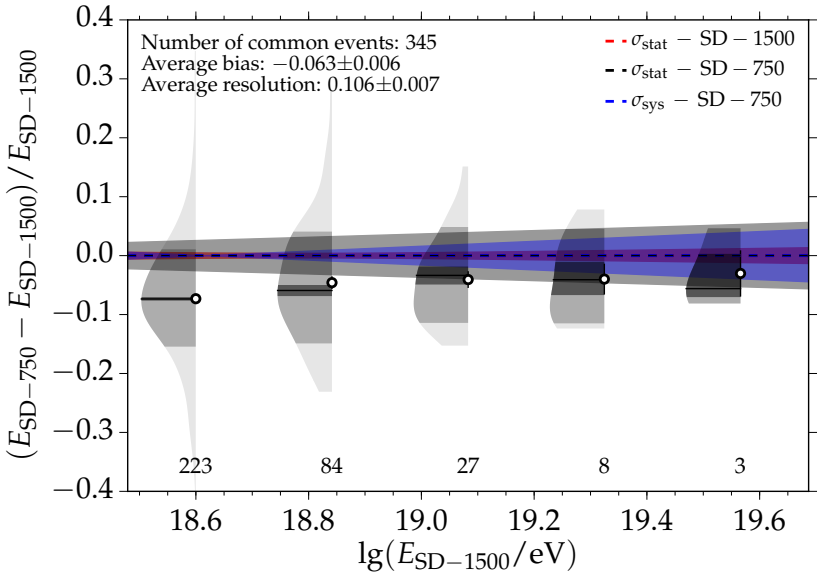
the systematic uncertainty from the energy bias correction mentioned in Section 4.7.3. It is minimal around $10^{18.6}$ eV. The statistical uncertainty in the calibration function of the SD-1500 is negligible (drawn in red). Overall, the different uncertainties cannot account for the observed bias between SD-750 and SD-1500 energies.

For comparison, using the latest published energy calibrations (without shower size corrections) [86] leads to a similar bias of −6.3 % as shown in Fig. 4.32b.

The average relative resolution of the SD-750 and SD-1500 energy differences is 10 %. This matches very well with the SD-1500 detector resolution of slightly above 10 % at 3×10^{18} eV. Shower-to-shower fluctuations are irrelevant in this comparison because the same showers are considered (and only separately reconstructed). Furthermore, the SD-750 detector resolution is negligible at these energies. Various checks have been performed in order to understand the origin of the observed energy differences. No improvement was found when only certain telescopes were selected, for example a calibration with only Coihueco data in both the calibrations of SD-750 and SD-1500 events. According to other analyses, the energy differences found within stereo events can be trusted (see



(a) Standard calibrations of this work



(b) Energy calibrations from ICRC 2015

Figure 4.32: Relative energy difference between SD-750 and SD-1500 reconstructions for a subset of common events. The difference between the plots (a) and (b) is further described in the text.

Section 4.7.3). Dust and dirt on the telescope mirrors and other telescope components, e.g. filters, and especially differences in the contamination for different eyes can contribute to reconstruction differences and consequently differences in the energy scales [42]. This is outside the scope of this work. A different approach to explain the energy differences is described in the next section.

4.7.7 Constrained energy calibration for SD-750 data

The mass composition of UHECRs changes significantly in the energy range considered in the calibration analysis (see e.g. Fig. 1.15). In consequence, the assumption of a simple power-law to describe the relation between S_{35} or S_{38} and E_{FD} might no longer hold.

Tests to extend the energy calibration function to describe data better are presented in Appendix B.6.1.

Due to the very limited amount of SD-750 events above 10^{18} eV, the slope of the calibration function is completely determined at lower energies. At the same time, a change in the slope to match the energies of SD-750 and SD-1500 events in Fig. 4.32a might have negligible impact on the energies where the bulk of events for the SD-750 energy calibration are. To test this, I added the common SD events to the calibration method. The respective resolution of $E(S_{38})$ is used instead of the resolution of E_{FD}. In this extended energy calibration fit, the energy estimates are calibrated with golden hybrid events and, at the same time, a constraint enforces the compatibility to energies measured with the SD-1500. Technically, it is a sum of two separate log-likelihood functions as specified in Eq. (4.34). The situation and the resulting fit are visualized in Fig. 4.33a. The golden hybrid events are given by the blue rectangular points, while the common SD events are shown with black circular markers. The red line represents the result of the analysis, while the standard energy calibration function for the SD-750 is drawn as blue dashed line (Eq. (4.36)). The resulting calibration function is:

$$E(S_{35}) = (11.58 \pm 0.25) \times 10^{15}\,\mathrm{eV} \left(\frac{S_{35}}{\mathrm{VEM}}\right)^{(1.035\pm0.004)} . \tag{4.38}$$

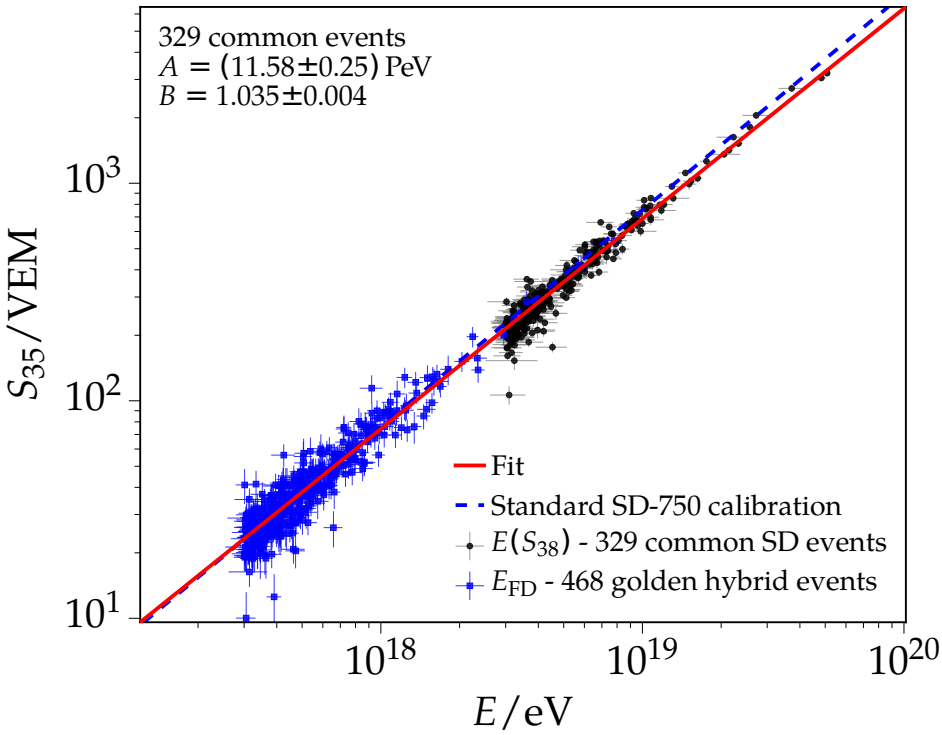
The difference between the two fits is small in the energy range of the golden hybrid events and rather large at the highest energies. The fit parameters are not very compatible to the ones in Eq. (4.36); there is a difference of about $3\,\sigma$. The energy difference of common events is reduced to only 1.3 %, which is a significant improvement. To quantify the effect on the events at lower energies, the energy residuals and resolutions are plotted in Fig. 4.34. These plots have to be compared to the ones in Fig. 4.29 and Fig. 4.31. While there is a slight increase in the relative energy differences and an increase of the relative resolution of 1 %, no significant impact is found. The calibration function Eq. (4.38) is therefore used for the following analyses. The difference to the calibration function in Eq. (4.36) should be considered as a systematic. Propagating half of the difference into the flux leads to the systematics shown in Fig. 4.35c. The systematic is described well with the linear function:

$$\frac{\Delta J}{J}(E) = (1.30\pm0.05) + (-0.073\pm0.003)\,\lg E. \tag{4.39}$$

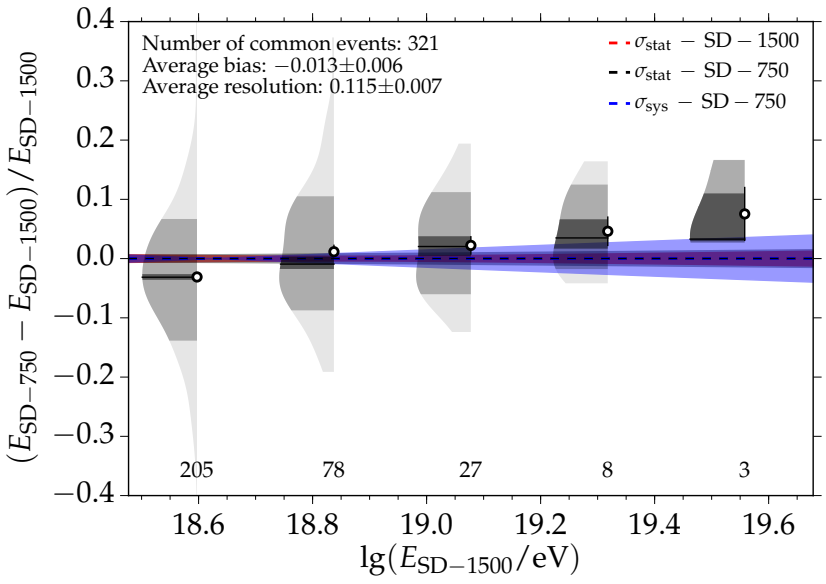
The impact on the energy bias correction from Section 4.7.3 is estimated. Two separate calibration functions using the full bias correction and no bias correction are derived. The results and fit parameters are shown in Fig. 4.35b and Fig. 4.35a, respectively. The difference in the derived flux using these parameters to the standard result are shown in Fig. 4.35d. It is well described with the following function:

$$\frac{\Delta J}{J}(E) = (1.29\pm0.07) + (-0.067\pm0.006)\,\lg E. \tag{4.40}$$

The total systematic uncertainty on the SD-750 flux from the energy calibration is given by the quadratic sum of the two contributions. It is plotted in Fig. 4.35e. The relative systematic is 12 % at 3×10^{17} eV, drops to 7 % at 3×10^{18} eV, and increases again at higher energies.



(a) Combined energy calibration fit



(b) Energy differences

Figure 4.33: (a) Energy calibration fit of the SD-750 events including events shared with SD-1500. (b) Relative energy differences of common events using the calibration function derived with the modified fit.

4.7.8 Comparison to recent results

The energy calibration parameters for the ICRC[4] [86] were produced in collaboration to this work. They are given in Table 4.2. The relation between E_{FD} and the three SD energy estimates is shown in a common representation in Fig. 4.36.

[4]International Cosmic Ray Conference

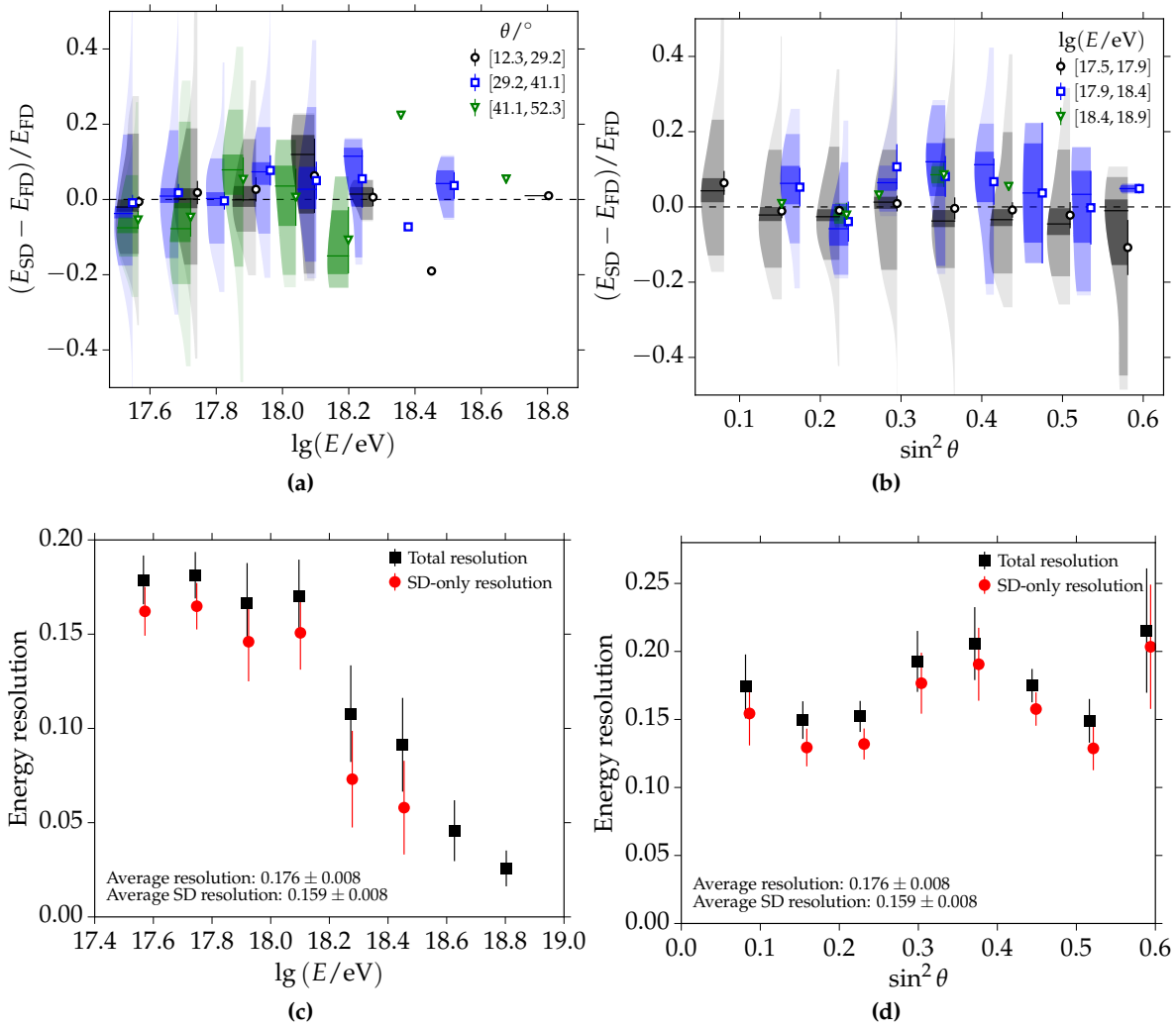
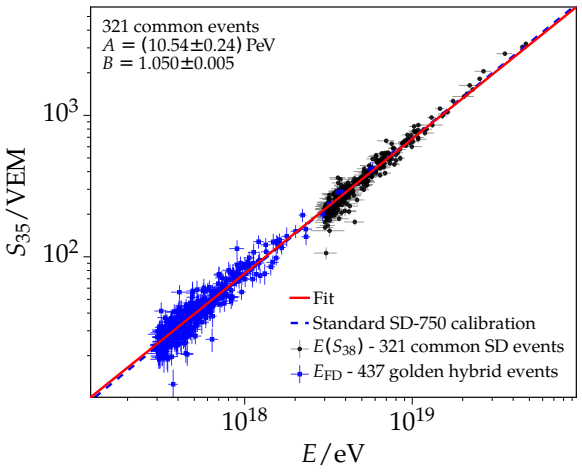


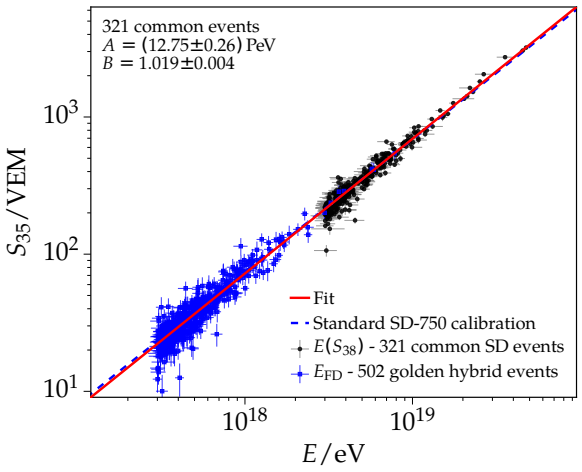
Figure 4.34: Energy residuals after the energy calibration as a function of (a) energy and (b) zenith angle. The energy resolution as a function of (c) energy and (d) zenith angle.

Table 4.2: The energy calibration parameters from the latest publication [86].

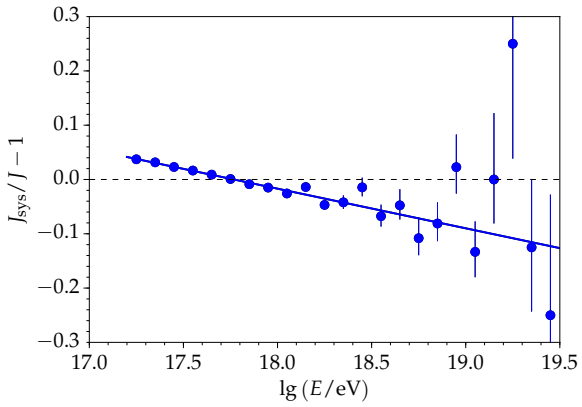
	SD-1500	SD-1500 inclined	SD-750
Energy calibration (A)	(0.187±0.004) EeV	(5.71±0.09) EeV	(12.87±0.63) PeV
Energy calibration (B)	1.023 ± 0.006	1.01 ± 0.02	1.013 ± 0.01



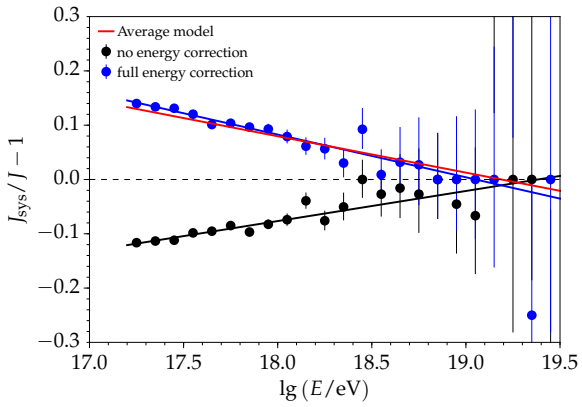
(a) No energy bias correction



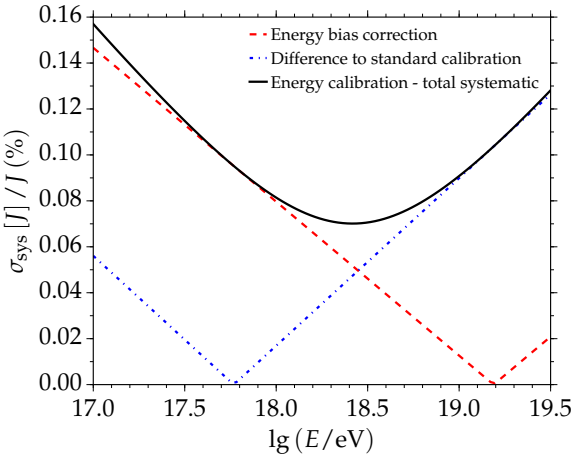
(b) Full energy bias correction



(c) Flux systematic from the difference in the calibration functions



(d) Flux systematic from the energy bias correction



(e) Overview of the flux systematics

Figure 4.35: Additional energy calibration plots to study systematics of the SD-750 calibration. The systematic uncertainties on the flux derived from the energy calibration are shown in (c) - (e).

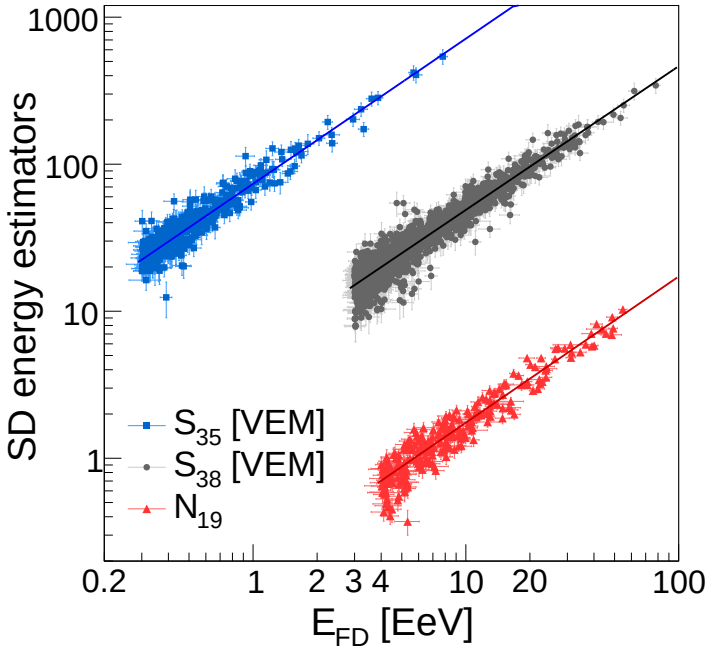


Figure 4.36: Golden hybrid data and the energy calibration curves for the three SD measurements as further described in the text (from [86]).

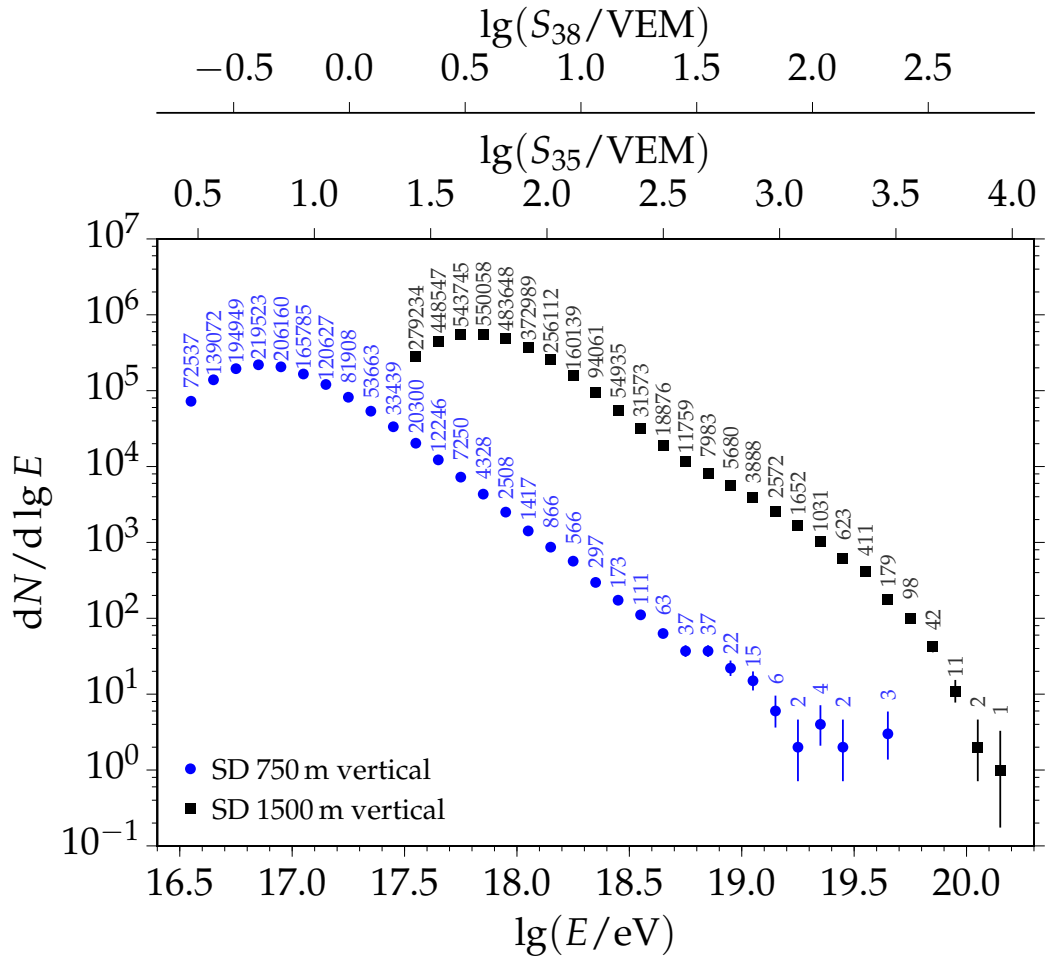


Figure 4.37: Spectrum of raw event numbers for both SD-750 and SD-1500 measurements. The two top axes give the spectrum as a function of the respective energy estimates S_{35} and S_{38}. The bottom axis represents the calibrated energies using the energy calibration parameters derived in Section 4.7.

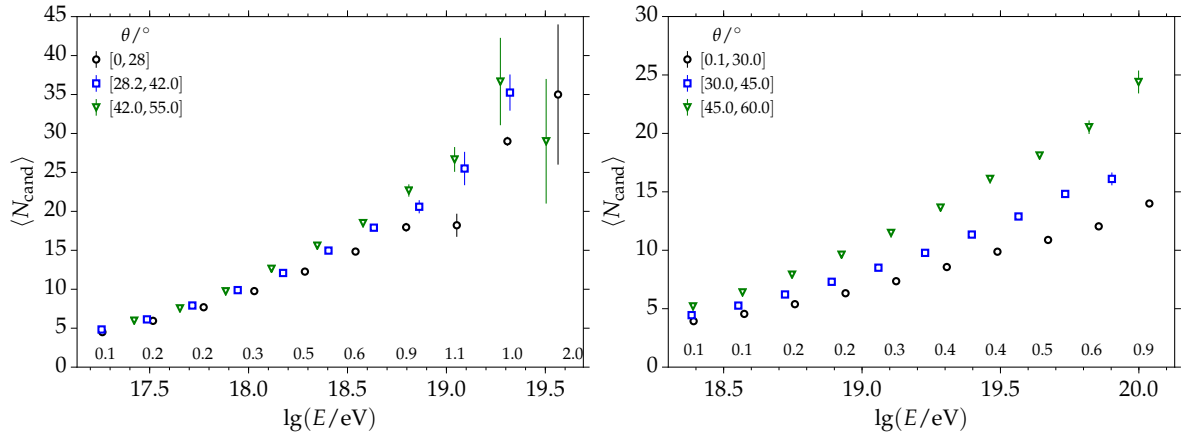
4.8 Energy spectrum

The energy spectrum of UHECRs is the differential event rate with respect to energy and exposure:

$$J(E) = \frac{\mathrm{d}N}{\mathrm{d}E\,\mathrm{d}\mathcal{E}} = E\frac{\mathrm{d}N}{\mathrm{d}\lg E\,\mathrm{d}\mathcal{E}}, \tag{4.41}$$

with $\mathrm{d}\mathcal{E} = \mathrm{d}t\,\mathrm{d}A\,\mathrm{d}\Omega$. The previous sections have introduced all relevant analysis steps prior to the derivation of the flux. A spectrum of the number of events as a function of $\lg E$ and the energy estimates for SD-750 and SD-1500 events is plotted in Fig. 4.37. The number of events within an energy bin is stated above the respective point.

The average number of candidate stations is plotted in Fig. 4.38. The number increases significantly with energy, with up to 40 stations in SD-750 events at the highest energies. It is also visible that, due to the projection of stations into the shower plane, the average number of stations increases with increasing zenith angle. The numbers at the bottom of the plots represent the average number of low-gain saturated stations in an event. On average, there are two saturated stations in SD-750 energies at highest energies and one saturated station in SD-1500 events. This difference reflects the difference in the spacing of detectors. As described in Section 3.3.2, a recovery method is used to be able to use the information



(a) SD-750 data **(b)** SD-1500 data

Figure 4.38: The average number of candidate stations in events as a function of $\lg E$ and for different intervals in reconstructed zenith angle. The numbers at the bottom represent the average number of saturated stations in the same energy intervals.

of saturated stations in the fit of the LDF. The next step from the raw event spectrum to the flux stated in Eq. (4.41) is to take into account the exposure derived in Section 4.2. Additionally, in case of SD-750 events, an energy-dependent aperture is derived. This is detailed in the next section.

4.8.1 Energy-dependent aperture

The T4 trigger efficiency for SD-750 events was thoroughly studied, which allows the construction of conservative regions of full trigger efficiency. The efficiency model for SD-1500 should be re-investigated before employing this analysis method. The regions of full efficiency are constructed by varying the zenith angle range as a function of energy. Instead of dividing the number of events by an energy-independent exposure $\mathcal{E}$ (as derived in Section 4.2), an energy-dependent exposure $\mathcal{E}_i$ is derived:

$$\mathcal{E}_i = \mathcal{E} \cdot \xi_i. \tag{4.42}$$

The efficiency correction ξ_i represents the influence of the efficiency cut on energies and zenith angles within the bin i. ξ_i is the distribution (flux) weighted ratio of selected and total events. The underlying event distribution is given by:

$$f(E,\theta) = \frac{\mathrm{d}^2 N}{\mathrm{d}E\,\mathrm{d}\theta} = \alpha E^{\gamma} \sin\theta\cos\theta. \tag{4.43}$$

The energy dependence is described with a power-law and the zenith angle distribution includes a $\cos\theta$ from the effective area and the $\sin\theta$ from the solid angle element. The total number of events inside a bin is expressed as:

$$\begin{aligned} N_{\mathrm{tot}} &= \int_{E_0}^{E_1} \mathrm{d}E \int_{\theta_0}^{\theta_1} \mathrm{d}\theta\, f(E,\theta) \\ &= \int_{E_0}^{E_1} \mathrm{d}E \int_{\theta_0}^{\theta_1} \mathrm{d}\theta\, \alpha E^{\gamma} \sin\theta\cos\theta \\ &= \frac{\alpha}{\gamma+1}\left(E_1^{\gamma+1} - E_0^{\gamma+1}\right) \cdot \frac{1}{2}\left(\sin^2\theta_1 - \sin^2\theta_0\right), \end{aligned} \tag{4.44}$$

with the zenith angle range considered for the analysis of SD-750 data of $[\theta_0, \theta_1] = [0°, 55°]$. The limited phase space is derived using the binary efficiency cut $\tilde{\epsilon}$, which is used in the same way for the selection of events:

$$\tilde{\epsilon}(E,\theta) = \begin{cases} 1 & , \epsilon(E,\theta) \geq 0.99 \\ 0 & , \epsilon(E,\theta) < 0.99 \end{cases} , \tag{4.45}$$

with the conservative efficiency model $\epsilon(E,\theta)$ presented in Section 4.4. Incorporating the efficiency into the integral above yields:

$$\begin{aligned} N_{\text{sel}} &= \int_{E_0}^{E_1} \mathrm{d}E \int_{\theta_0}^{\theta_1} \mathrm{d}\theta \; f(E,\theta)\; \tilde{\epsilon}(E,\theta) \\ &= \int_{E_0}^{E_1} \mathrm{d}E \int_{\theta_0(E)}^{\theta_1(E)} \mathrm{d}\theta \; \alpha E^{\gamma} \sin\theta \cos\theta \\ &= \frac{1}{2} \int_{E_0}^{E_1} \mathrm{d}E \; \alpha E^{\gamma} \left[\sin^2(\theta_1(E)) - \sin^2(\theta_0(E))\right] . \end{aligned} \tag{4.46}$$

In the second step, the efficiency $\tilde{\epsilon}$ is absorbed into lower and upper zenith angle boundaries $\theta(E)$ exploiting the simple binary form of $\tilde{\epsilon}$ (Eq. (4.45)). The exposure factor ξ_i for the bin i including the energy range $[E_0, E_1]$ and zenith angle range $[\theta_{\text{min}}, \theta_{\text{max}}]$ is finally given by:

$$\begin{aligned} \xi_i &= \frac{N_{\text{sel}}}{N_{\text{tot}}} \\ &= \int_{E_0}^{E_1} \mathrm{d}E \; E^{\gamma} \left[\sin^2(\theta_0(E)) - \sin^2(\theta_1(E))\right] \\ &\quad \cdot \frac{(\gamma+1)}{\left(E_1^{\gamma+1} - E_0^{\gamma+1}\right) \cdot \left(\sin^2\theta_{\text{min}} - \sin^2\theta_{\text{max}}\right)} . \end{aligned} \tag{4.47}$$

Using this method, the resulting energy-dependent exposure for the SD-750 measurement is shown in Fig. 4.39. The error bars represent the systematic uncertainty on the exposure calculation of 3 %. The efficiency correction is significant below $10^{17.6}$ eV.

4.8.2 Raw energy spectra

The exposure calculation from the previous section is used to calculate the raw energy spectra. These spectra are not corrected for migration effects due to the finite detector resolution and fluctuations in the shower development, as it is done in Section 4.10. The flux measured with the SD-750 is plotted in Fig. 4.40a, while the one measured with the SD-1500 is shown in Fig. 4.40b. Both measurements are compared in Fig. 4.40c. For visualization purposes, the SD-1500 points are shifted by $\lg E = 0.01$ on the x-axis (after the energy rescaling). From now on, this is always done in comparison plots without further notice.

Each point is plotted at the center of the respective $\lg E$-bin and is divided by the energy interval $\mathrm{d}E = E_{\text{up}} - E_{\text{low}}$, with up and low indicating the upper and lower bin edges. For historical and comparative purposes, a binning in steps of $\mathrm{d}\lg E = 0.1$ is chosen. The error bars indicate the $1\,\sigma$ statistical uncertainties. If the number of events within a bin exceeds 20, asymmetric Poisson uncertainties are used. For smaller statistics, two-sided Feldman-Cousins limits with a coverage of 68 % are calculated [167]. The upper limits at the end of the SD-1500 flux represent one-sided Feldman-Cousins limits for no measured events, no background expectation, and a coverage of 84 %.

To enhance the spectral features, each point is scaled by a factor of E^3. A flattening of the spectrum around $10^{18.7}$ eV is apparent in both energy spectra. A steep suppression of

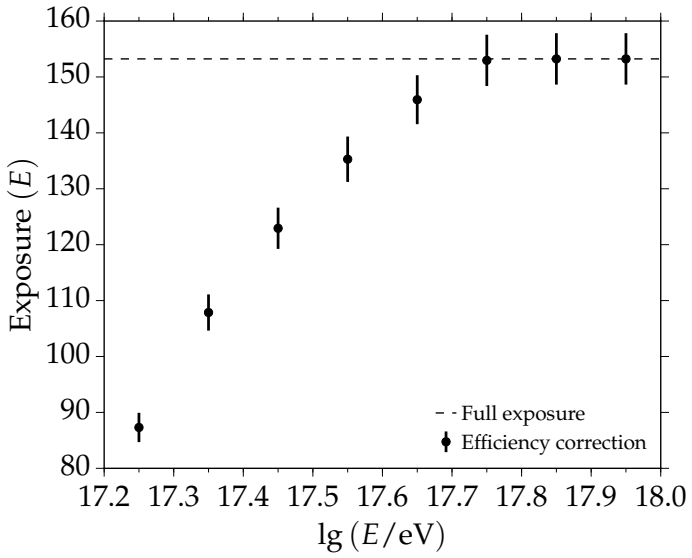


Figure 4.39: Energy-dependent exposure for the SD-750 resulting from the T4 trigger efficiency. The derived efficiency models are used to calculate conservative, efficient regions of zenith angle, which vary as a function of energy.

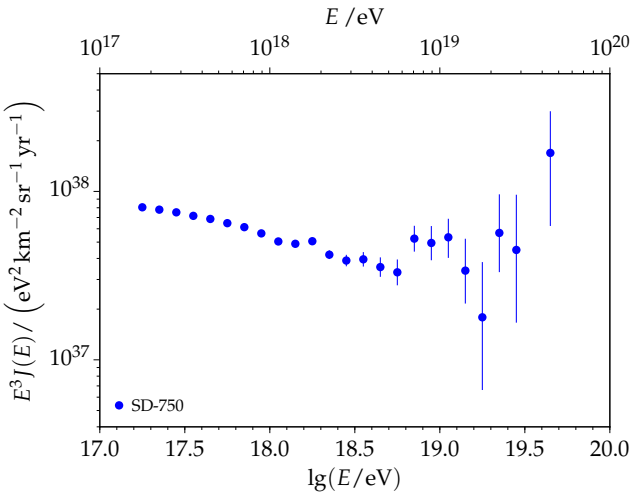
the energy spectrum measured with the SD-1500 occurs above $10^{19.5}$ eV. A more detailed discussion of the spectral slopes and features is presented after the correction of the energy spectra in Section 4.10.

4.8.3 The flux for different zenith angles

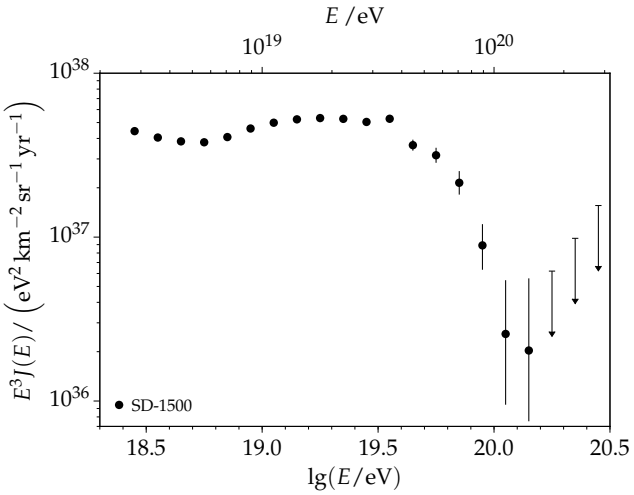
The ground signals of air showers from identical primaries will differ for different zenith angles. Within this work, the CIC method is used to correct for this dependence. The method is described in Section 4.5. Separate attenuation functions are derived for SD-750 and SD-1500 data. The measured flux is divided into different ranges of zenith angle to check for remaining angular dependencies. The results for four bins in zenith angle are shown in Fig. 4.41. The bins are chosen equidistantly in $\mathrm{d}\cos^2\theta$, such that the exposures of the four spectra are identical. Each of the individual fluxes J_i is plotted relative to the average flux J.

Deviations in the SD-750 spectrum in Fig. 4.41a are small and on the order of 5 %. The largest deviation is apparent at smallest zenith angles and below $10^{17.7}$ eV. This deviation does not depend on the details of the efficiency correction and could be caused by a slight energy dependence of the attenuation function itself. The attenuation function is derived for events above an energy of $10^{17.8}$ eV, thus not taking into account possible structures in data below this energy. However, it is more important to describe data at higher energies well which is the case judging from Fig. 4.41a. Moreover, a derivation of the attenuation function at lower energies is problematic because full trigger efficiency is not given and the assumption of a uniform rate of events versus $\cos^2\theta$ breaks down.

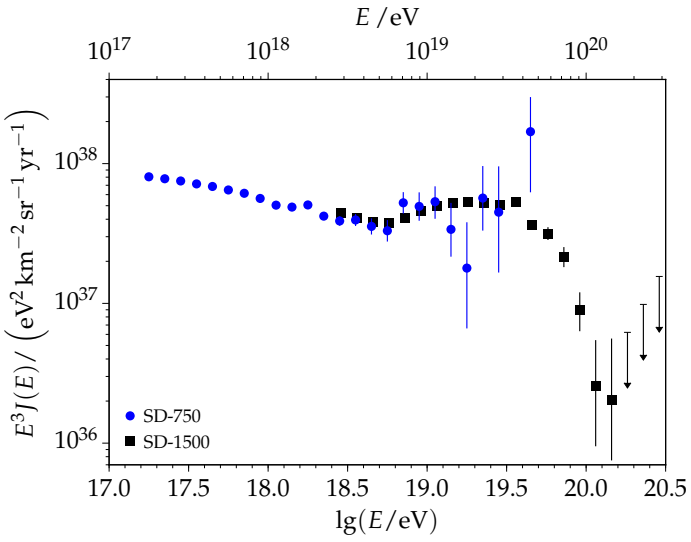
There are larger deviations in the SD-1500 spectrum as shown in Fig. 4.41b. These deviations are on the order of 10 % and are most likely due to an energy-dependence of the attenuation function, which is derived for events with energies above $10^{18.9}$ eV. This dependence itself could be caused by zenith angle dependent migration effects or systematic deviations in the LDF at certain zenith angles. The average flux is not affected by this. Still, in future analyses, the correction of these dependencies should be considered.



(a) Flux measured with the SD-750

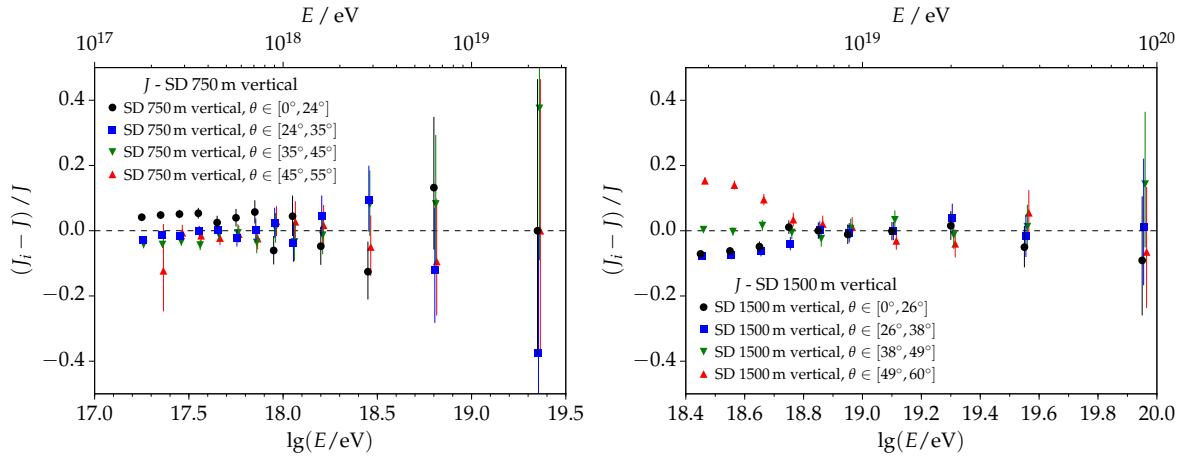


(b) Flux measured with the SD-1500



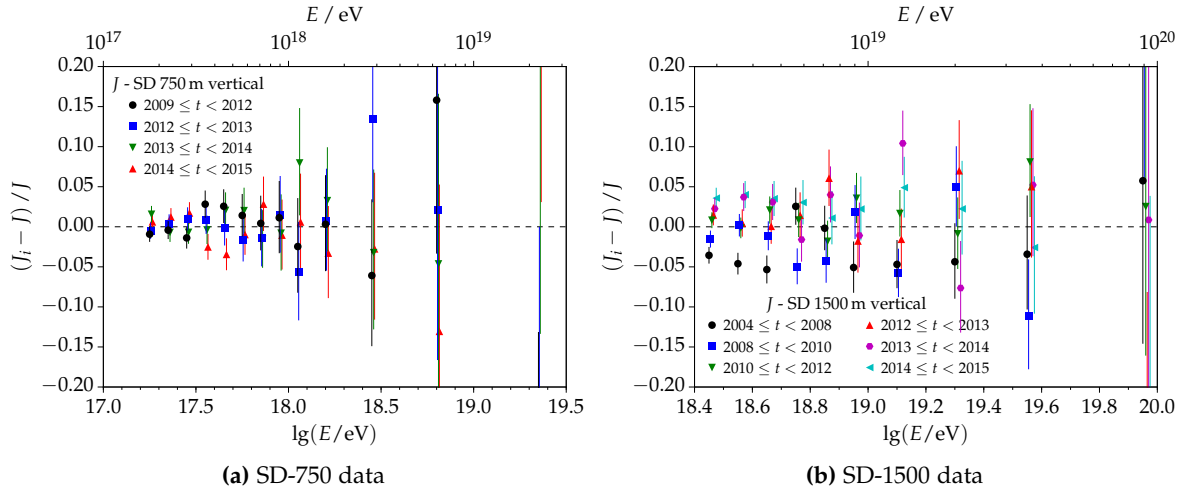
(c) Comparison of the measurements

Figure 4.40: The raw (uncorrected) energy spectra measured with the SD-750 and SD-1500. A comparison between the two spectra is shown in Fig. 4.40c. The points of the SD-1500 spectrum in (c) are shifted to the right with respect to the bin centers by 0.01 in logarithmic energy to increase the visibility. The same binning is used for both spectra. The same applies to similar plots that appear later in this thesis.



(a) SD-750 data **(b)** SD-1500 data

Figure 4.41: The measured flux of UHECRs divided into different intervals of zenith angle.



(a) SD-750 data **(b)** SD-1500 data

Figure 4.42: The measured flux of UHECRs divided into different time periods.

4.8.4 The flux for different time periods

To quantify changes in the flux with time, the measured flux is divided into different time periods. The results for SD-750 data are shown in Fig. 4.42a. All statistically significant fluctuations are within 3 %. This is compatible with the systematic uncertainty attributed to the exposure calculation in Section 4.2.

The results for SD-1500 data in Fig. 4.42b show a similar behavior. Most of the deviations are within 3 %. The flux for the time period 2004 - 2008 is an exception. There is a stronger deviation for this time period, which is related to the observed change of the energy scale with time (see Section 4.6 and especially Fig. 4.23). The origin of this change with time is still unclear. The change of the SD energy scale with time is taken into account as a systematic uncertainty in the overall uncertainty of the energy scale [168]. For this work, I opt to additionally increase the systematic uncertainty on the exposure calculation attributed to SD-1500 data to 5 %.

Table 4.3: Summary of the experimental parameters describing data of the different measurements at Auger. Numbers of events are given above the energies corresponding to full trigger efficiency (adapted from [86]).

	SD-1500	**SD-750**
Data taking period	01/2004 - 12/2014	08/2008 - 12/2014
Exposure in $\mathrm{km}^2\,\mathrm{yr}\,\mathrm{sr}$	42 520±2130	153±5
Zenith angles	0° to 60°	0° to 55°
Threshold energy $E_{\mathrm{eff}}/\mathrm{eV}$	3×10^{18}	3×10^{17}
No. of events ($E > E_{\mathrm{eff}}$)	96 529	52 248
No. of events (golden hybrids)	1750	468 (+329, combined)
Energy calibration (A)	(0.180±0.004) EeV	(11.58±0.25) PeV
Energy calibration (B)	1.030±0.006	1.035±0.004

4.9 Search for a declination dependence

Current results of the measurement of the flux of UHECRs reported by Auger and TA show striking differences, in particular at the highest energies, which are of most interest to contemporary studies [86, 169].

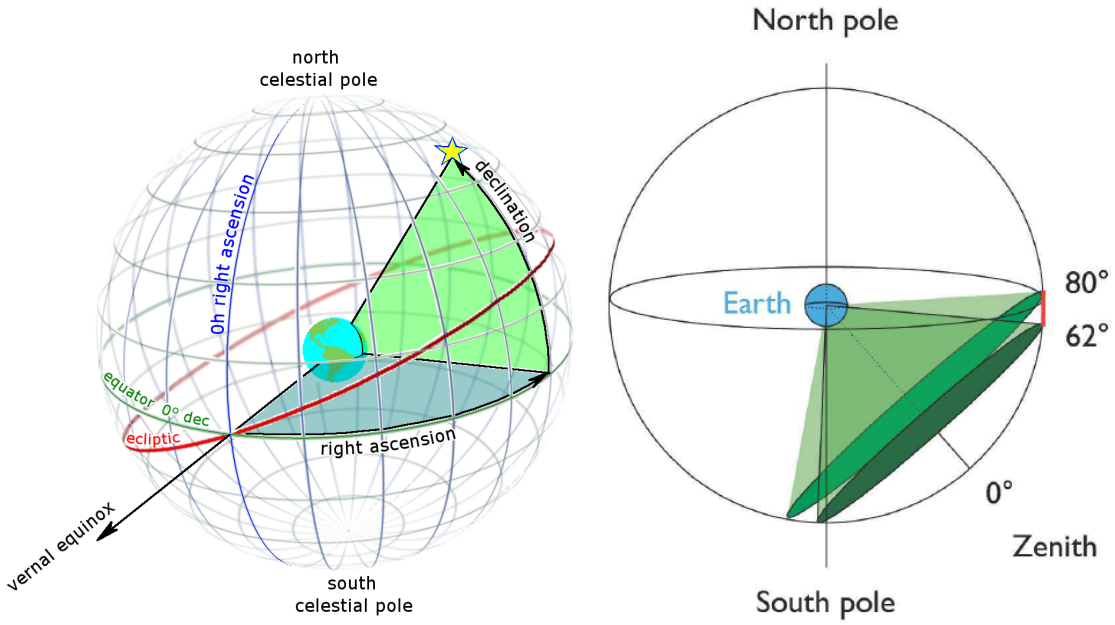
While these differences could still be explained by differences in the overall energy scale of the experiments (Auger: 14 %, TA: 21 %), other experimental or systematic effects, or differences in the analyses, it is particularly interesting to study if there is an actual dependence of the flux of UHECRs on the incoming direction of primary particles. Such an observable difference would set fruitful ground for various astrophysical interpretations related to the distribution of sources of UHECRs. Auger allows us to observe a wide range in declination from −90° to 25° (even up to 46° using inclined events) and thereby offers an excellent set of data to study these effects. I want to present results of the search for a dependence of the measured flux on the incoming direction in recent Auger data. The covered range in declination partly overlaps with the range of TA from −6° to 84°. An SD measurement up to a zenith angle of 45° was used to obtain the TA data presented in this analysis [169].

This note starts with a short description of the used sets of data and selection criteria and proceeds with a description of the experimental exposure in equatorial coordinates. The validity of the CIC method (e.g. [127, 155]) is shown afterwards, followed by detailed results of the integrated and differential flux of cosmic rays in different declination intervals. Results for the different SD arrays are given and a combined flux result is presented and compared to recent energy spectra.

This section follows the description in [170], a previous note that I have written for this thesis. The results were updated to the current set of data.

4.9.1 Data and quality selection

The data and quality selection detailed in Section 4.1 is used. Events up to a zenith angle of 60° are included in the analysis of SD-1500 events. Due to limitations of the CIC method, the angular range for the SD-750 analysis is restricted to an upper angle of 55° (as discussed in Section 4.5.1. Details about the different Auger data sets derived in this work are summarized in Table 4.3. In particular, corrections due to weather and geomagnetic effects as discussed in Section 4.3 are used to correctly derive the angular dependencies.



(a) Schematic of the equatorial coordinate system (from Wikipedia).

(b) A view of local coordinates at the latitude of Auger.

Figure 4.43: Schematics of equatorial and local coordinates.

4.9.2 Exposure and distributions of arrival directions of air showers

The integrated exposure for the SD-1500 energy spectrum in the time range of this study (beginning of 2004 until the end of 2014) sums up to $(42\,520 \pm 2130)\,\mathrm{km^2\,yr\,sr}$. It is obtained from a purely geometrical, time-dependent counting of active elementary hexagons taking into account down-times of single detectors or parts of the array (see e.g. Section 4.2 and [33]). For the SD-750 measurement, the integrated exposure from August 2008 until the end of 2014 is $(153 \pm 5)\,\mathrm{km^2\,yr\,sr}$. The exposure corresponding to the presented energy spectrum measured by TA equals $5400\,\mathrm{km^2\,yr\,sr}$ for a measurement over six years [171].

I am interested in expressing the exposure in terms of declination in equatorial coordinates (see Fig. 4.43a). By exploiting that the SD is continuously operating during each day (e.g. it is not systematically off during fixed hours), one can neglect a dependence on right ascension. Auger is located in the southern hemisphere at an average latitude of $l = -35.21°$, while TA is located in the northern hemisphere at a latitude of 39.3°.

To formulate the exposure as a function of declination, I furthermore exploit that the respective detectors are fully efficient up to a zenith angle of θ_{max}. The relative exposure as a function of declination δ can be formulated as (see also [172]):

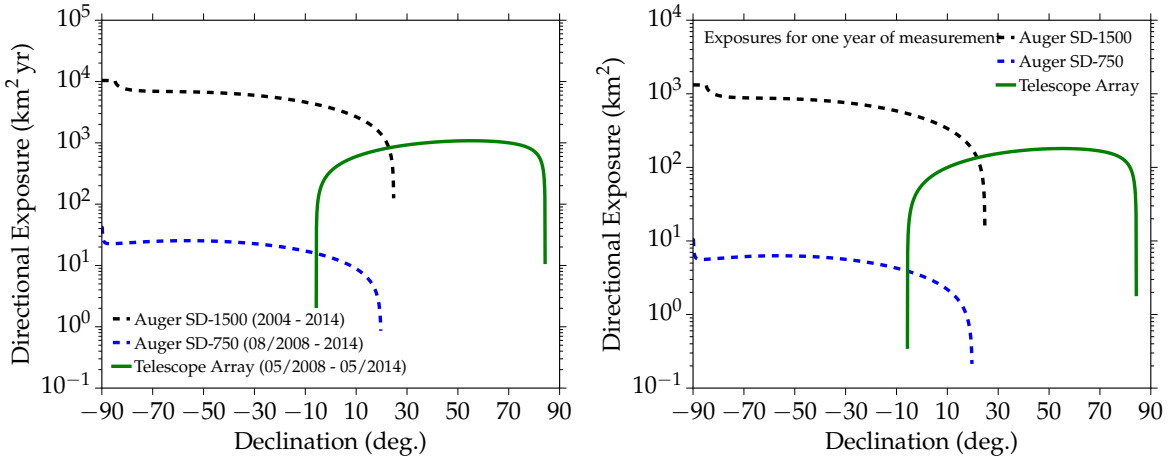
$$\mathcal{E}(\delta) \propto \cos l \cos \delta \sin \beta + \beta \sin l \sin \delta, \tag{4.48}$$

with β given by:

$$\beta = \begin{cases} 0 & , \quad \xi > 1 \\ \pi & , \quad \xi < -1 \\ \cos^{-1} \xi & , \quad \text{else} \end{cases} . \tag{4.49}$$

The variable ξ is expressed in terms of the maximum zenith angle, declination, and latitude as:

$$\xi = \frac{\cos \theta_{\mathrm{max}} - \sin l \sin \delta}{\cos l \cos \delta}. \tag{4.50}$$



(a) Overall exposure

(b) Yearly exposure

Figure 4.44: (a) The experimental exposure for the different data sets in for specified time intervals. (b) The same figure for yearly exposures. A yearly average was calculated for the different data sets. For the Auger exposures an average over the most recent four years is taken, while the TA exposure is divided by six.

The directional exposure as a function of declination has been calculated using these formulas. The results for the different spectra are depicted in Fig. 4.44. The logarithmic scale was chosen to show all exposures in a direct comparison. The exposure is normalized such that by integrating over the sphere

$$\mathcal{E}_{\text{tot}} = \int \mathrm{d}\Omega\, \epsilon(\delta) = 2\pi \int_{-\pi/2}^{\pi/2} \mathrm{d}\delta\, \mathcal{E}(\delta) \cos\delta \tag{4.51}$$

the resulting exposure equals the total exposure for the specific measurement and time range. For the presented measurements, Auger and TA have an overlap in declination from -5.7° to 24.8°. The total exposure of Auger in this declination interval sums up to $8600\,\mathrm{km}^2\,\mathrm{yr}\,\mathrm{sr}$ while the one from TA is $3692\,\mathrm{km}^2\,\mathrm{yr}\,\mathrm{sr}$.

A detailed look at the distributions of incoming directions in different coordinate systems helps to understand the observed data. For example, the measurement of inclined air showers from 62° to 80° does not include the celestial south pole. This is visible in the schematic Fig. 4.43b. Due to the particular zenith angle interval and the latitude of Auger, a ring around and excluding the pole is described.

Different distributions of incoming directions as functions of other directional angles are depicted in Fig. 4.45. Included are events measured with the SD-1500. In all plots, the violins give the distribution of events inside a particular bin via a kernel density estimate of the actual distribution. The mean values are represented by markers, while median and median standard deviations are shown with black bars and dark areas spanning from $-1\,\sigma$ to $1\,\sigma$. For example, in Fig. 4.45a, one can see that only the more inclined showers arrive from very low or high declinations, as expected. The transition from north to south in azimuth of local coordinates is shown in Fig. 4.45d. These distributions are expected from coordinate transformations and the position of Auger. Similar plots for the other Auger spectra are included in Appendix B.8.4.

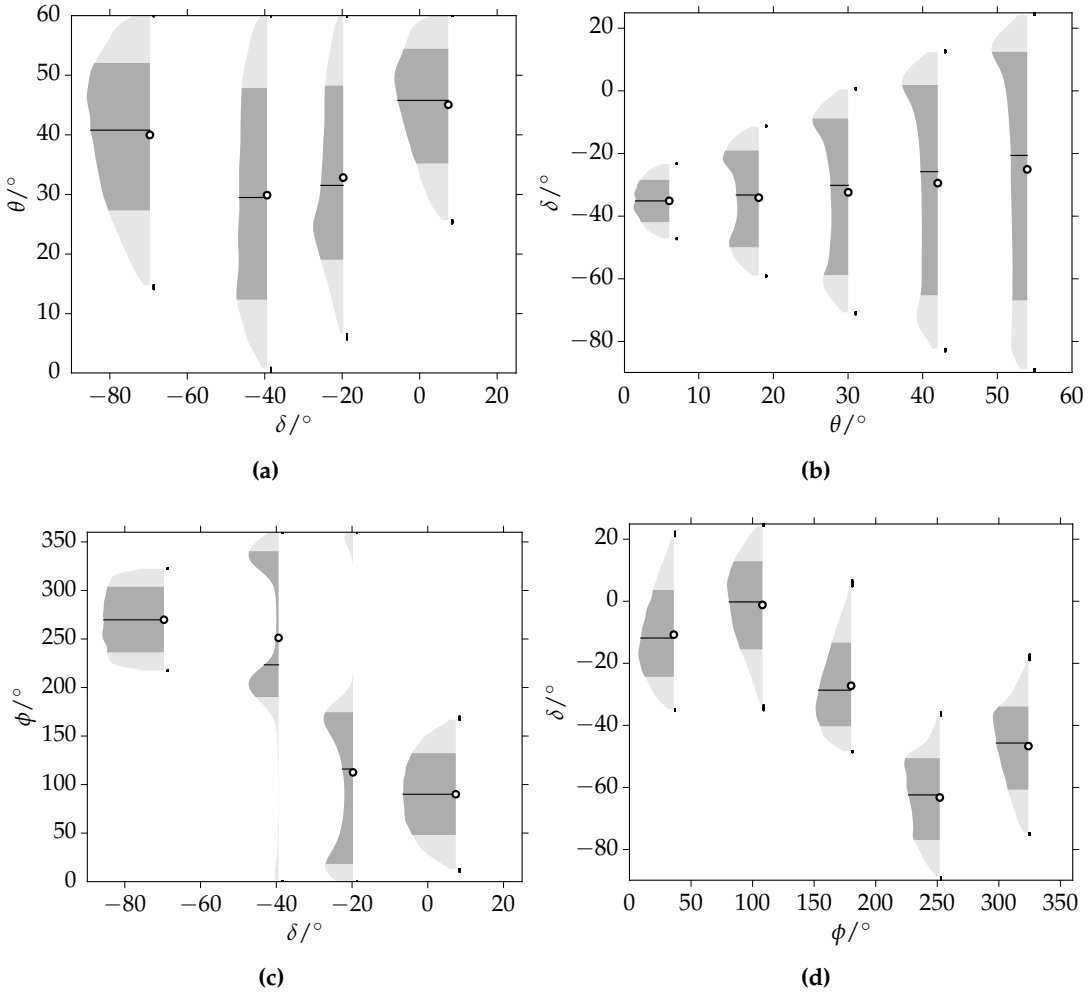


Figure 4.45: Distributions of incoming directions for the measurement of vertical showers with the SD-1500 array.

4.9.3 Validation of the CIC method

For the measurement of vertical events with the SD-750 and SD-1500 arrays, the CIC method is used to empirically correct for the attenuation of air shower signals in the atmosphere and the different station geometries in the shower plane (see Section 4.5). The method is based on the assumption of an isotropic flux of cosmic rays, at least up to a certain energy at which the attenuation functions and average shower sizes are obtained. A difference in the flux of cosmic rays between particles arriving from the north/south would firstly undermine the CIC method and secondly wash out possible differences in the intensities and differential spectra presented in the following sections.

To check the validity of the CIC method, two subsamples of data measured with the SD-1500 are chosen:

- Events from the north: Events arriving in the local azimuthal range from 60° to 120° that were observed during 4 months of austral summer (November-February). Events with high declination are thus enriched (see Fig. 4.43).
- Events from the south: Events arriving in local azimuthal range from 240° to 300° that were observed during 4 months of austral winter (May-August). Events with small declination are thus enriched.

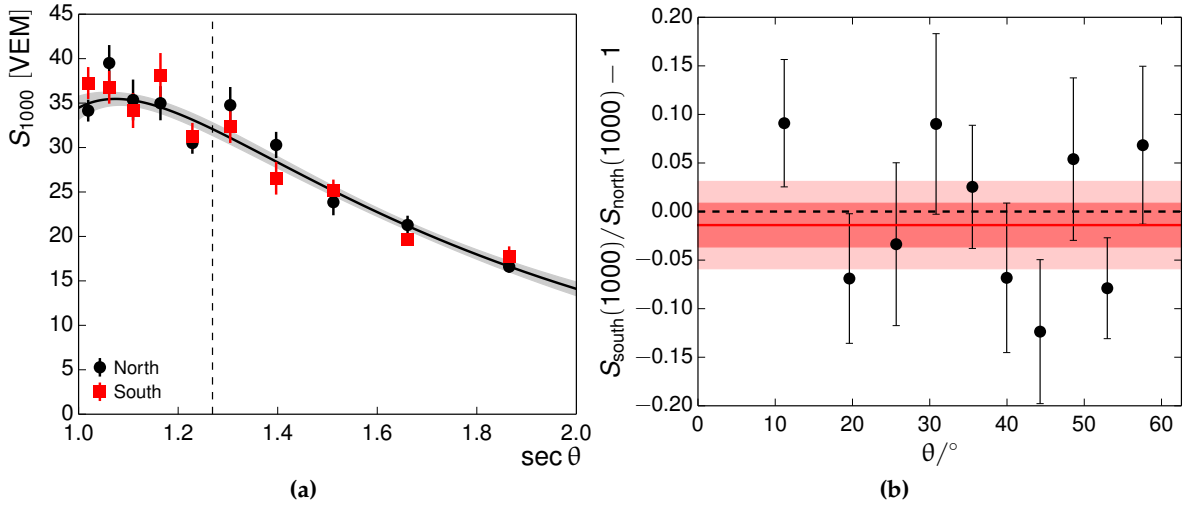


Figure 4.46: (a) The average shower sizes extracted for both samples of events arriving from north or south. The fitted attenuation function for the subsample of events arriving from north is shown for comparison. (b) The residual of the points in (a). The dark red contour reflects the $1\,\sigma$ uncertainty regime that is well compatible with zero.

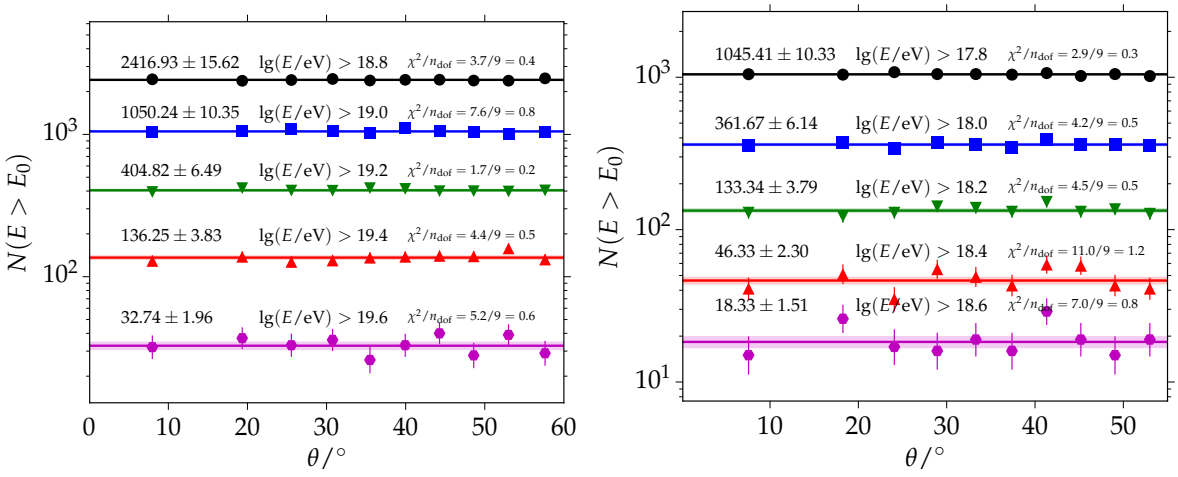
As based only on restrictions of the time and azimuthal ranges, these selections do not cause zenith-dependent biases. Separate attenuation functions are obtained for each of set of events. They are compared in Fig. 4.46. The shower sizes are shown in Fig. 4.46a, while a relative residual is plotted in Fig. 4.46b. Both attenuations agree well within statistical uncertainties. This confirms that the CIC method can be used for the subsequent analyses and already suggests that there is no strong dependence of the measured spectra on the incoming direction.

4.9.4 Checks for directional dependencies in the intensity spectra

Analogously to studying the integrated number of cosmic ray events above a certain energy threshold in equidistant bins of $\cos^2\theta$ to validate the attenuation function (see Fig. 4.47, I divide the data into different intervals of declination that cover the same amount of integrated exposure. This is done in a numerical way by exploiting the dependence of the exposure on declination as depicted in Eq. (4.48), and employing a simple bisection search to a relative tolerance of better than 0.05 %. The resulting intensity spectra are shown in Fig. 4.48 for the different data sets and various energy thresholds above full detector efficiency. For a certain energy threshold, only events above the respective threshold are counted:

$$N(E_0) = \int_{E_0}^{\infty} \mathrm{d}E \, \frac{\mathrm{d}N}{\mathrm{d}E}. \tag{4.52}$$

For the SD-1500 measurement, data are divided into ten exclusive declination intervals that cover an equal amount of exposure of 4252 $\mathrm{km}^2\,\mathrm{yr}\,\mathrm{sr}$. For the SD-750, each of the ten bins covers 15.3 $\mathrm{km}^2\,\mathrm{yr}\,\mathrm{sr}$, taking into account events with zenith angles below 55°. Constant fits to the profiles are used to quantify possible deviations from a invariable behavior. No statistically significant deviation is apparent from the χ^2/n_{dof} values and the plots themselves. Interesting structures are visible though. For example, the intensity profile for SD-1500 data with energies above $10^{18.9}$ eV is depicted in Fig. 4.48c. A transition to a lower rate of events is visible at a southern declination of $\delta \approx -30°$. This is further discussed in Section 4.9.6.



(a) SD-1500 **(b)** SD-750

Figure 4.47: Measured intensity profiles as a function of $\cos^2\theta$ for SD-750 and SD-1500 data. No statistically significant deviation from a constant behavior is apparent.

4.9.5 Declination dependence of the differential flux

The differential flux of cosmic rays

$$J(E) = \frac{\mathrm{d}N}{\mathrm{d}E\,\mathrm{d}A\,\mathrm{d}t\,\mathrm{d}\Omega} \tag{4.53}$$

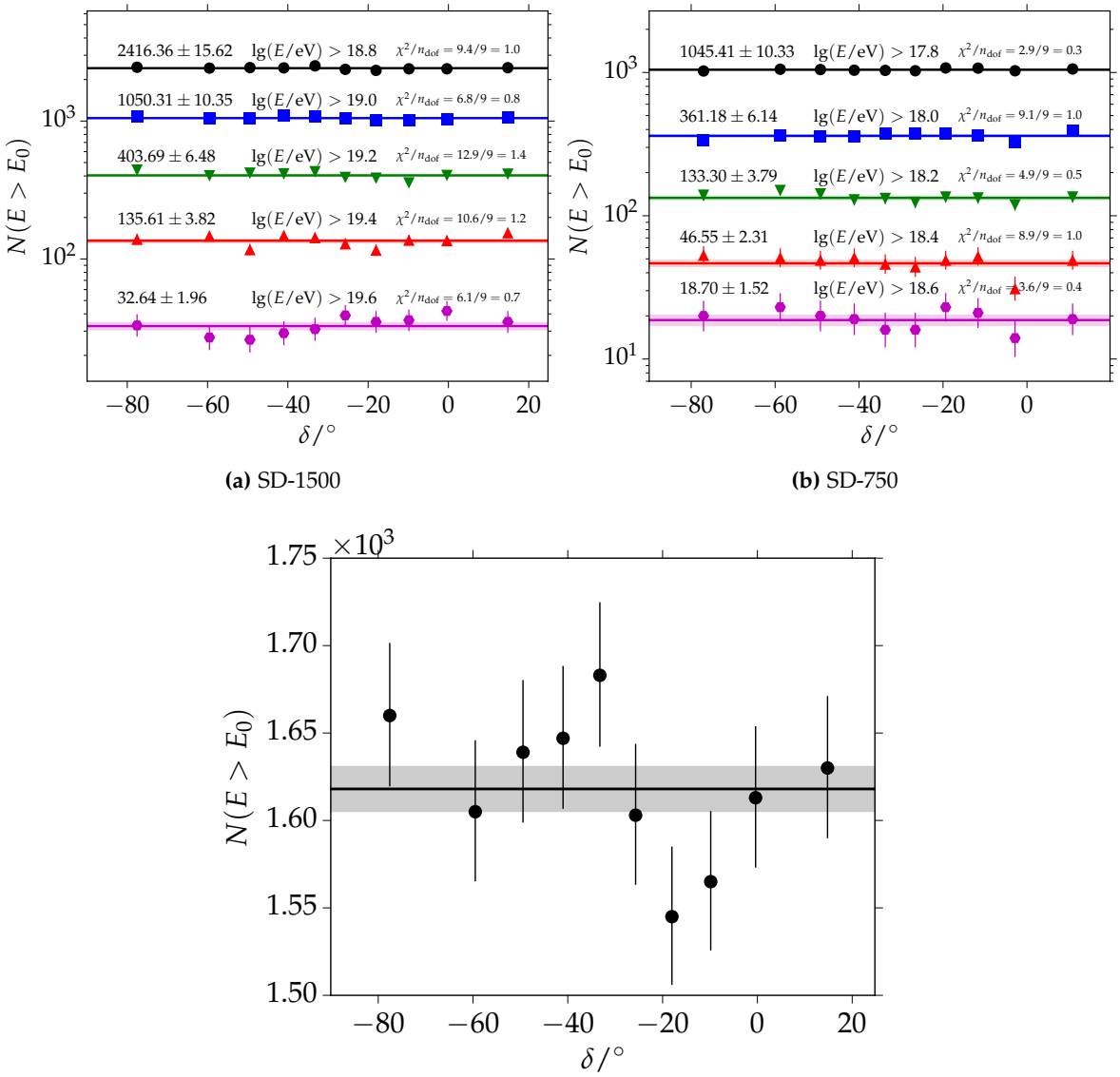
is divided into different intervals of declination to search for a dependence on the incoming direction. Four bins in declination that cover the same amount of exposure are chosen for SD-1500 data, while only three bins are used in case of SD-750 data due to the smaller event statistics. Because of the different zenith angle ranges of the measurements, different declination ranges follow from the equipartition. For the SD-1500, the edges of the declination ranges are −90°, −49.3148°, −29.4746°, −10.0059° and 24.79°. They include a correction for the small tilt of the array as discussed in [74]. The flux is divided into these declination intervals in Fig. 4.49a and Fig. 4.49b. The resulting relative residuals with respect to the overall energy spectrum (including events in the whole angular range) are depicted in Fig. 4.49c for SD-1500 data and in Fig. 4.49d for SD-750 data. Again, there are no significant deviations from the average flux as a function of energy and declination. To quantify this further, the next subsection deals with fits to the sub-spectra measured with the SD-1500 and an analysis of the fit parameters as a function of declination. I note that there is feature in the SD-750 flux: in the energy range $10^{18.0}$ eV to $10^{18.1}$ eV, there is a distinct drop in the flux of the most southern events. This significance of this feature is on the level of $3\,\sigma$.

Fitting the vertical energy spectra

To characterize the spectral features, I describe the data with a power law below the ankle E_a: $J(E) \propto E^{-\gamma_1}$. Above the ankle, the data are fit to a power-law with a smooth suppression at the highest energies:

$$J(E) = a \left(\frac{E}{E_\mathrm{a}}\right)^{-\gamma_2} \left[1 + \left(\frac{E_\mathrm{a}}{E_\mathrm{s}}\right)^{\Delta\gamma}\right] \left[1 + \left(\frac{E}{E_\mathrm{s}}\right)^{\Delta\gamma}\right]^{-1}, \tag{4.54}$$

where γ_1 and γ_2 are the spectral indices below and above the ankle. E_s denotes the energy at which the flux drops to half its value when extrapolating $J \propto E^{-\gamma_2}$ without a flux



(a) SD-1500 **(b)** SD-750

(c) SD-1500, $E > 10^{18.9}$ eV

Figure 4.48: Measured intensity profiles as a function of declination in equatorial coordinates extracted for SD-750 and SD-1500 data. No statistically significant deviation from a constant behavior is apparent.

suppression. $\Delta\gamma$ describes the difference in the spectral index before and after the flux suppression. To fit the spectral models, a Poisson maximum-likelihood fit is used. The minimization package MINUIT-2 is used to perform the minimization [110]. The fitted models are compared to the four sub-spectra of the SD-1500 measurement in Fig. 4.51a. No unfolding of the detector resolution effect is used for these fits. The parameters of the different fits are summarized in Table 4.4. They are also depicted as a function of the different declination intervals in Fig. 4.50. I want to note that for the logarithm of the flux normalization $\lg a$ as shown in Fig. 4.50a, a systematic uncertainty of 0.03 from the uncertainty in the exposure estimation needs to be taken into account. Each of the error bars represents the statistical uncertainty from the fit to data in a single declination interval. The mean values and their $1\,\sigma$ statistical uncertainties are depicted with lines and shaded regions. Except for the overall flux normalizations depicted in Fig. 4.50a, there is no

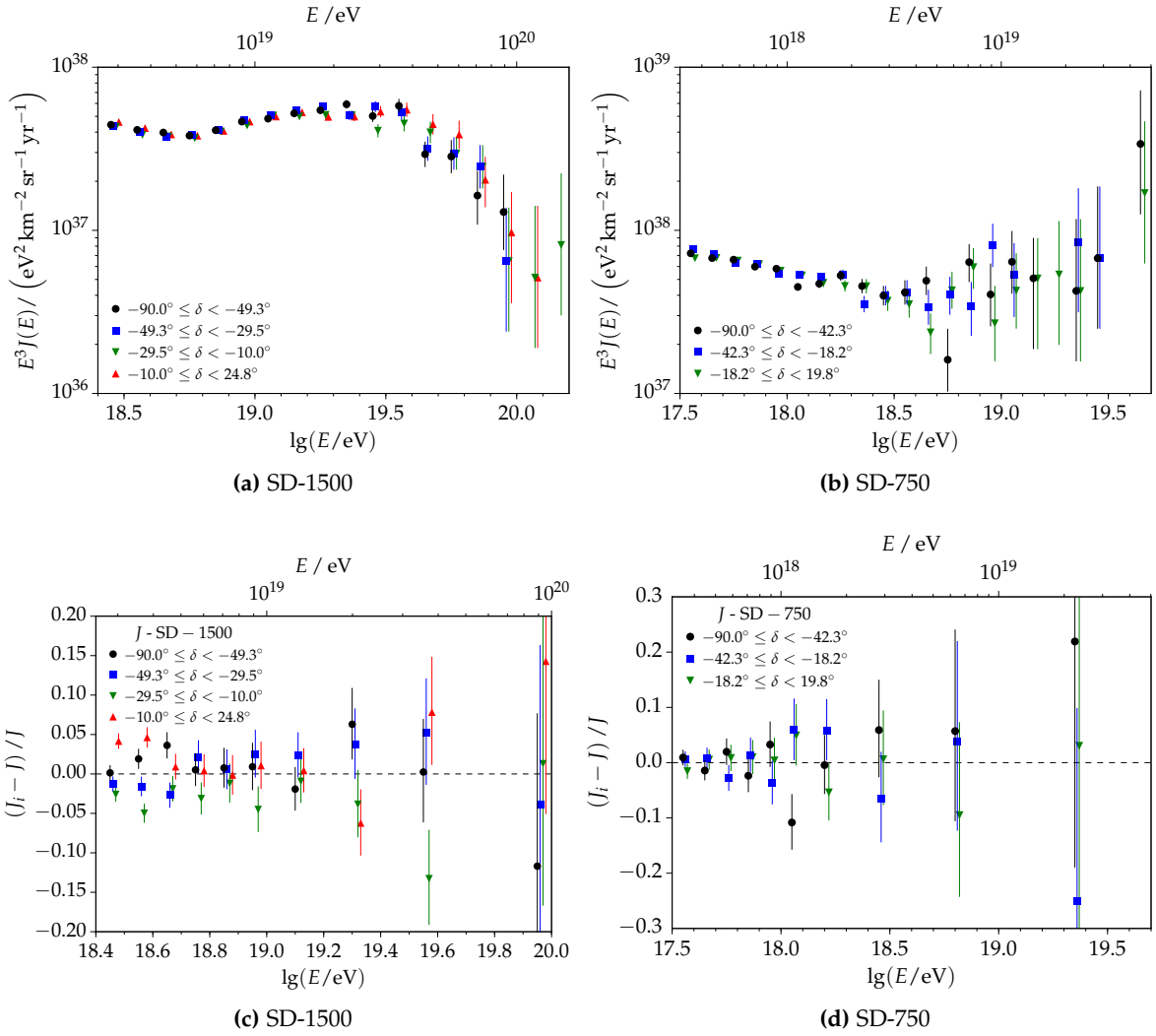


Figure 4.49: The SD-1500 and SD-750 energy spectra divided into different declination intervals. (a) and (b): the rescaled flux in log-log representation; (c) and (d): relative residuals with respect to the average spectrum.

Table 4.4: Parameters and uncertainties from fits to the differential energy spectra for events in different declination intervals. Only data measured with the SD-1500 are included.

Declination range	−90° to −49°	−49° to −29°	−29° to −10°	−10° to 25°
$\lg(a/\mathrm{a.u.})$	−17.866±0.002	−17.880±0.002	−17.890±0.002	−17.860±0.002
$\lg(E_a/\mathrm{eV})$	18.71±0.01	18.68±0.01	18.70±0.01	18.70±0.02
γ_1	−3.32±0.02	−3.34±0.02	−3.32±0.02	−3.38±0.02
γ_2	−2.63±0.05	−2.58±0.05	−2.55±0.07	−2.66±0.06
$\lg(E_s/\mathrm{eV})$	19.64±0.04	19.59±0.05	19.54±0.07	19.67±0.05
$\Delta\gamma$	3.0±0.5	2.6±0.4	2.2±0.3	2.9±0.7
χ^2/n_{dof}	1.10	0.81	1.40	1.04

Table 4.5: Parameters, with statistical and systematic uncertainties, of the model describing the SD-1500 spectrum and the SD-1500 spectrum for the high declination range only. Only statistical uncertainties are given for the latter, the systematic uncertainties are identical to the ones of the overall SD-1500 spectrum.

Parameter	SD-1500	SD-1500 high declination
$\lg(E_a/\mathrm{eV})$	$18.71 \pm 0.01 \pm 0.03$	18.80 ± 0.02
γ_1	$3.22 \pm 0.01 \pm 0.05$	3.10 ± 0.02
γ_2	$2.52 \pm 0.04 \pm 0.05$	2.45 ± 0.20
$\lg(E_s/\mathrm{eV})$	$19.55 \pm 0.03 \pm 0.03$	19.51 ± 0.20
$\Delta\gamma$	$2.5 \pm 0.2 \pm 0.2$	2.3 ± 0.6

statistically significant trend in any of the parameters. Furthermore, there is no evidence for a systematic change in the spectral features at higher declinations.

A comparison between the SD-1500 spectrum in the high declination interval between 0 and 25° with the overall SD-1500 spectrum and the TA spectrum is shown in Fig. 4.51b. As expected from the previous plots, the SD-1500 spectrum for the high declination range matches very well with the overall spectrum derived in this work. Both spectra are corrected for a smearing due to the detector resolution with a forward-folding approach (see Section 4.10). Both Auger SD-1500 spectra show a clear (energy-dependent) difference to the TA spectrum that increases up to 25 % at highest energies (see Fig. B.35). The parameters of the flux models are summarized for comparison in Table 4.5.

4.9.6 Comparison of the event rate for two declination intervals

There is a difference in the overall rate of events from the northern and southern part of the sky. This difference is quantified using two declination intervals that cover equal exposure. The declination intervals are −90° to −29.4746° and −29.4746° to 24.79°. The difference in the event rate is visualized in Fig. 4.52. Above an energy of 8×10^{18} eV, there significantly more events from the south are observed. The shaded regions represent the expectation from the measured dipole. This is a result from a recent but independent analysis of data measured with the SD-1500 [81].

4.9.7 Conclusion

Differences between the recent Auger and TA energy spectra motivate the search for a declination dependence in the flux of cosmic rays. The full set of Auger SD-1500 data until the end of 2014 was analyzed for such a dependence. Both the integral and differential energy spectra of the SD-1500 and SD-750 measurements were studied. The parameters of fits to the different spectra were studied as a function of declination and an SD-1500 spectrum in the declination interval between 0 and 25° was created and compared to the current Auger and TA spectra. While there are some interesting statistical features, there is no significant indication of a dependency of the spectral features at highest energies on declination. As it is, the question about the difference between the Auger and TA spectra cannot be answered with this. However, it would be necessary to also study the TA spectrum in the overlapping declination interval to draw final conclusions. Studies concerning the systematics and the energy scales of the two experiments are ongoing. We observe a difference in the overall event rate from the southern and northern skies. More events from the southern sky are measured. The significance of this excess depends on the

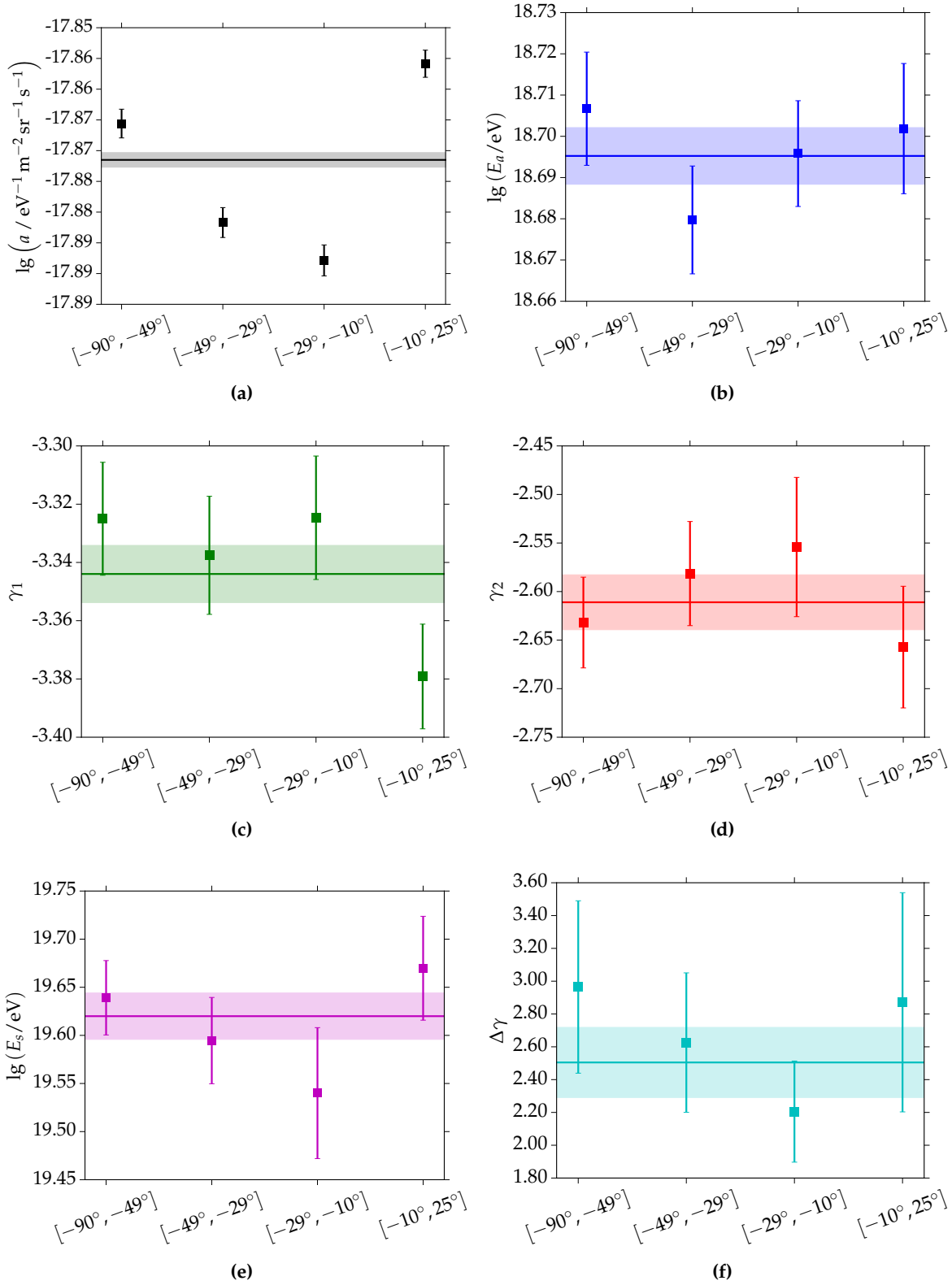


Figure 4.50: Different model parameters describing the energy spectra as a function of the declination interval.

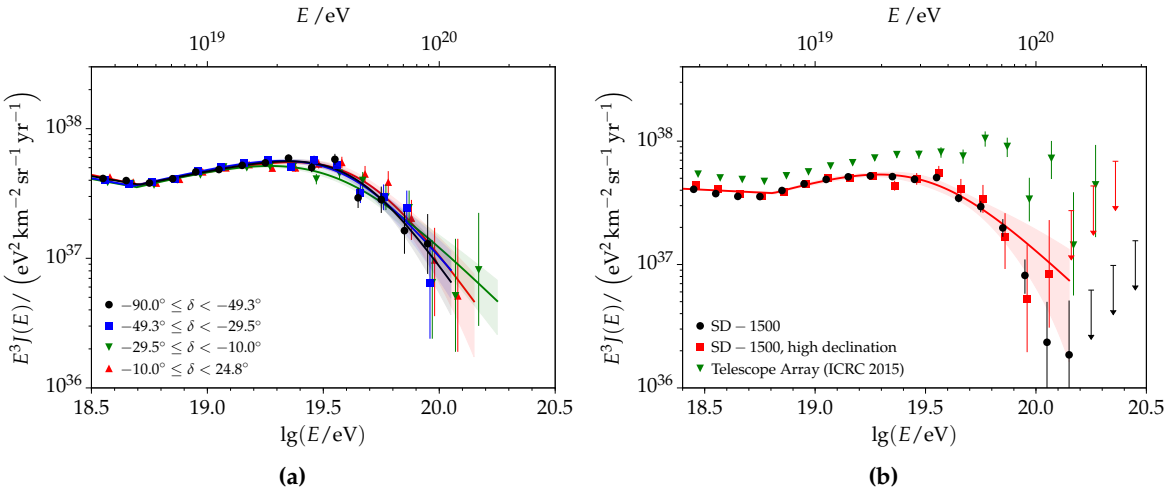


Figure 4.51: (a) The SD-1500 energy spectrum in different declination intervals. Fits lines to the different spectra are drawn for comparison. (b) Comparison between the SD-1500 flux for the high declination interval 0° to 25°, the overall SD-1500 spectrum and the current TA spectrum. A fit to the high declination spectrum is depicted with the red line.

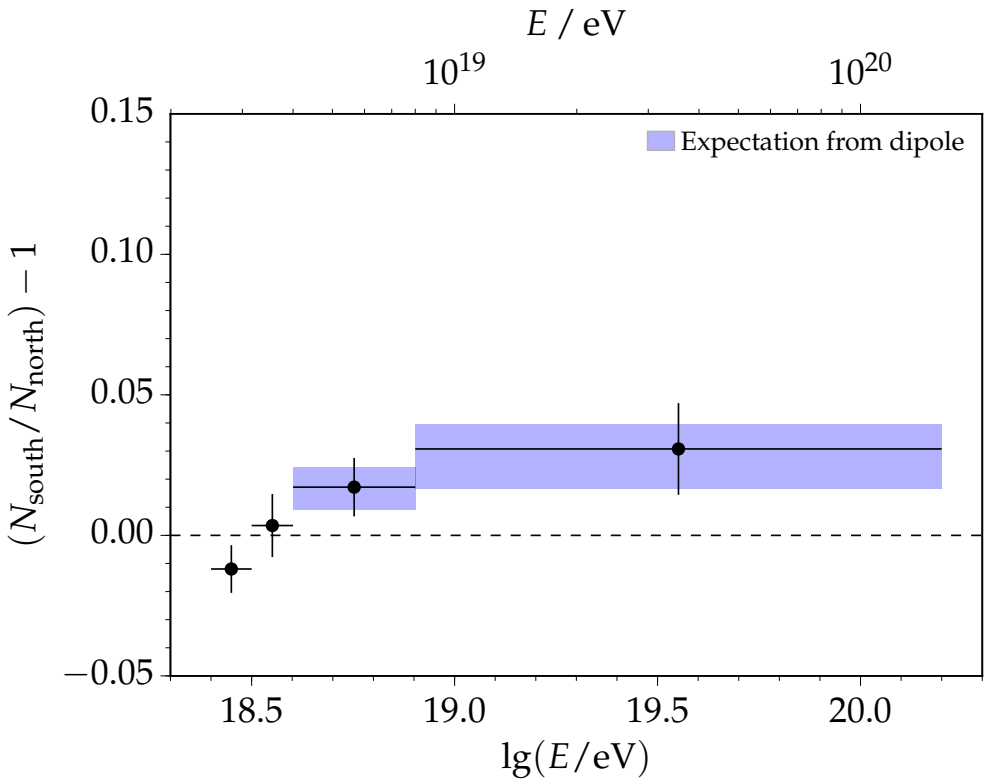


Figure 4.52: The ratio of events from two different ranges in declination: a comparison between events from south and north. The measurement is compared to the expectation from the measured large scale anisotropy [81]. Only data measured with the SD-1500 are included.

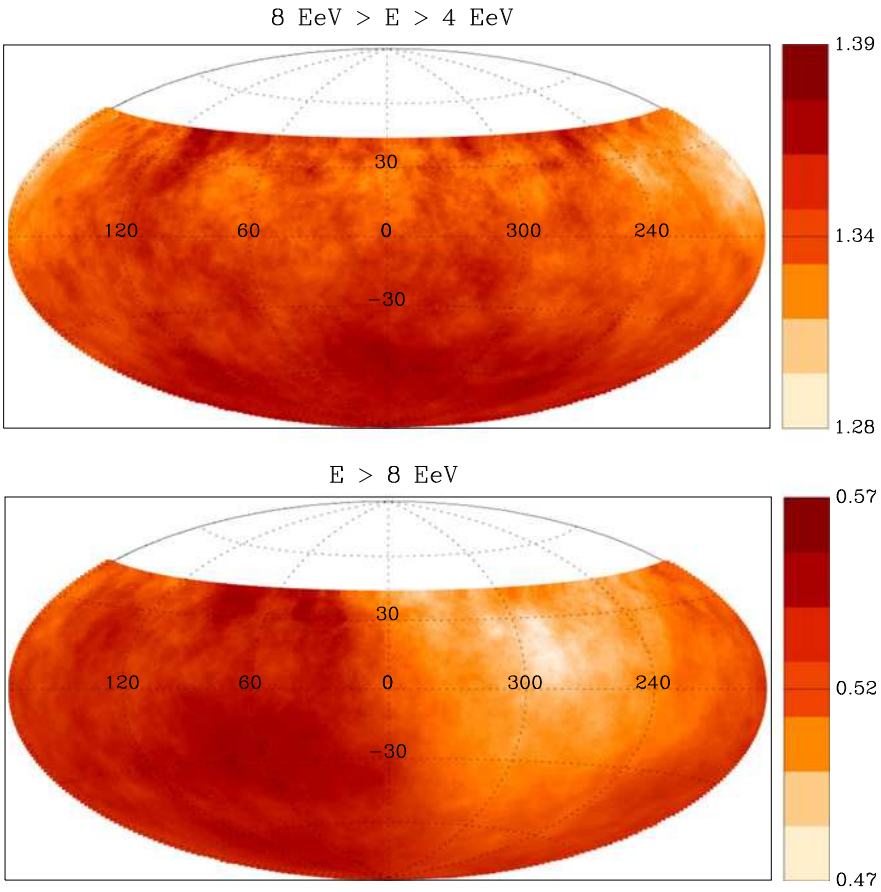
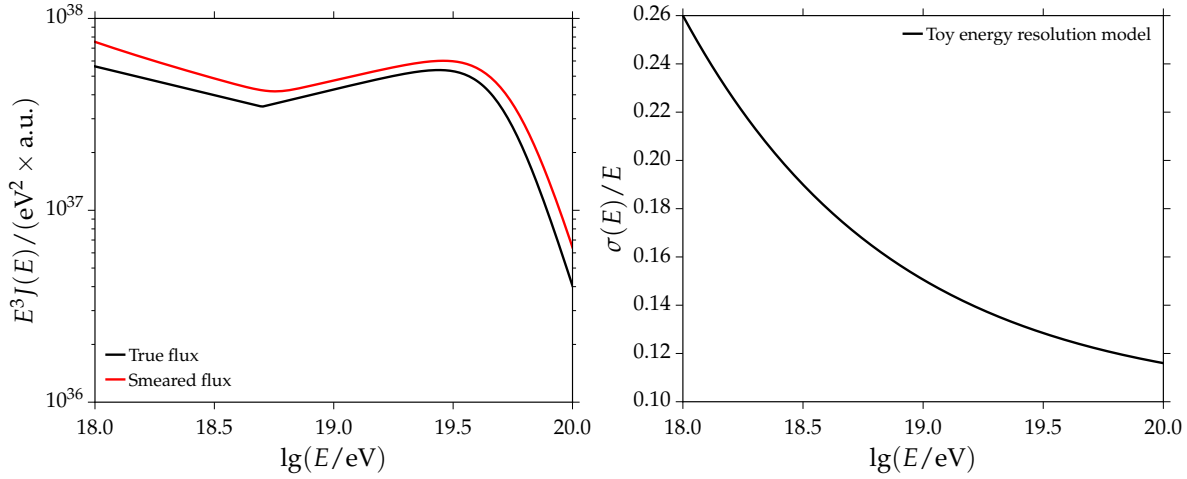


Figure 4.53: Depiction of the measured large scale anisotropy of events observed with Auger. Data are divided into two energy ranges. The dipole anisotropy occurs above primary energies of 8×10^{18} eV (from [81]).

energy interval and reaches $2\,\sigma$. The combined significance for all events above 4×10^{18} eV is $3\,\sigma$. This should be compared to the skymaps visualized in Fig. 4.53. They are taken from the current large scale anisotropy study detailed in [81]. The lower map, for events with energies above 8×10^{18} eV, visualizes a clear excess of events from the south (the darker regime). The rate of events in the sky is compatible to a dipole with a hot spot at $(-39\pm13)^\circ$ in declination and $(95\pm13)^\circ$ in right ascension. The dipole expectation regimes in Fig. 4.52 are calculated according to this measured dipole as explained in [81].



(a) True and smeared toy energy spectrum

(b) Toy resolution model

Figure 4.54: Illustration of migration effects.

4.10 Correction for event migrations

The raw flux of UHECRs presented in Section 4.8.2 is not corrected for event migrations. Due to the steep power-law-like flux and the finite energy resolution $\sigma[E]$, there are always more events that fluctuate from lower to higher energies than vice versa. Because of that, the measured flux is always larger than the true and unknown flux. This is illustrated in Fig. 4.54a. In this example, a very similar spectrum to the measured one is assumed and smeared with a Gaussian kernel. The relative energy resolution is plotted in Fig. 4.54b. A similar effect is present in both the data measured with the SD-750 and the SD-1500. The measured flux needs to be corrected to obtain an estimate of the true flux. A sound knowledge of the scale of shower-to-shower fluctuations and sampling fluctuations is essential.

The different contributions to the total detector resolution were estimated from data and with dedicated sets of simulations. As described in Section 4.7, the magnitude of sampling fluctuations and shower-to-shower fluctuations can be estimated from data themselves. The event-by-event estimate of the sampling fluctuations σ_{det} is given as the quadratic sum of these different contributions:

$$\sigma^2_{\mathrm{det}} = \sigma^2_{\mathrm{LDF,stat}} + \sigma^2_{\mathrm{LDF,sys}} + \sigma^2_{\mathrm{CIC}} + \sigma^2_{\theta}, \tag{4.55}$$

with the statistical uncertainty from the LDF fit $\sigma_{\mathrm{LDF,stat}}$, the propagated contribution from the uncertainty on the reconstructed zenith angle θ and systematic uncertainties from the LDF parametrization $\sigma_{\mathrm{LDF,sys}}$. The term σ_{CIC} stems from the uncertainty on the attenuation function. Given the attenuation function $f_{\mathrm{CIC}}(\theta)$ and the energy calibration relation $E(S) = AS^B$, these uncertainties are easily propagated into uncertainties on the reconstructed primary energy. In case of the shower size of SD-1500 data, S_{1000}, this reads:

$$\sigma^2[E] = \left(\frac{\partial E}{\partial S}\right)^2 \sigma^2[S_{1000}] + \left(\frac{\partial E}{\partial \theta}\right)^2 \sigma^2[\theta]. \tag{4.56}$$

The further derivation of this formula is obtained by plugging in $f_{\mathrm{CIC}}(\theta)$ and $E(S)$. Due to the separation of CIC and the energy calibration, this simple relation between relative uncertainties holds:

$$\frac{\sigma[E]}{E} = B\frac{\sigma[S_{1000}]}{S_{1000}}, \tag{4.57}$$

with the slope B of the energy calibration function. An equivalent relation holds for the shower size S_{450} used in the analysis of SD-750 data.

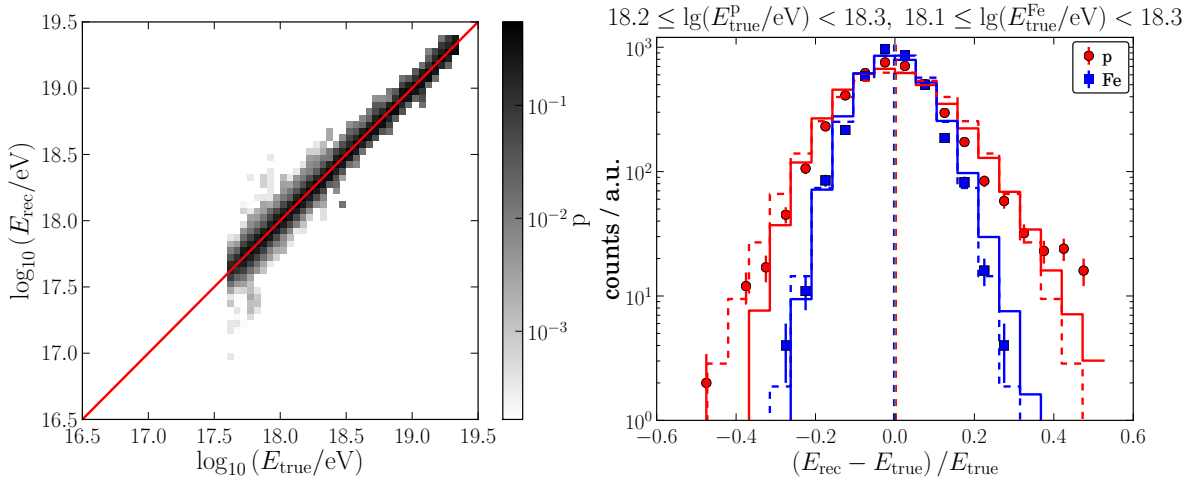
The energy calibration method presented in Section 4.7 allows one to estimate the shower-to-shower fluctuations $\sigma_{\mathrm{sh}}(E)$ together with the calibration parameters A and B. Results on the fluctuations for data measured with the SD-750 and SD-1500 are given in Section 4.7.5. For SD-750 data, σ_{sh} is well described with a constant of 10 %, while SD-1500 data are better described with a linear model that decreases with energy. An estimate for the total detector resolution is given by:

$$\sigma^2_{\mathrm{tot}}(E,\theta) = \sigma^2_{\mathrm{det}}(E,\theta) + \sigma^2_{\mathrm{sh}}(E). \tag{4.58}$$

A derivation of the shower-to-shower fluctuations from data as a function of zenith angle is not possible within the current event statistics.

A caveat concerning the estimation of the resolution from data is that one relies strongly on the correctness of uncertainties and probability models that are used in the reconstruction. One critical example is the Poisson factor used in the fit of the LDF (see Section 3.3.2 and Appendix B.5.1). This factor is used to convert signals in units of VEM into an effective number of particles for the Poisson fit. As shown in Appendix B.5.1, a small change in this factor results in a change of the uncertainties on shower sizes of 10 % or more. Another problem occurs when certain fit parameters are fixed in the reconstruction (or certain stages). In the LDF fit, the slope of the lateral distribution is not fitted for events with station multiplicities below 5. This results in an underestimation of the uncertainty on the shower size. The systematic uncertainty $\sigma_{\mathrm{LDF,sys}}$ represents a parametrization of this contribution and needs to be added [129, 173]. Due to these complications, a derivation of the energy resolution from simulations is of vital importance. Within this work, the Monte Carlo library described in Appendix E.4 is used to derive the resolutions for the SD-750 measurement, while the library detailed in Appendix E.3 is used for the SD-1500 measurement. Both libraries are continuous in energy and zenith angle and are comprised of showers from the primary particles proton and iron. An identical event selection as for data (see Appendix D.1.2) ensures the quality of simulated events.

A method to derive the energy resolution from simulations is to study the fluctuations of the reconstructed energy E_{rec} with respect to the true energy E_{true}. The true energy is either the Monte Carlo energy E_{MC} or the simulated FD energy E_{FD}. The bias of the latter is studied first and corrected for. There is no relevant deviation in SD-1500 simulations. The bias found in SD-750 simulations is significant (larger than 5 %) at energies below 10^{18} eV and is described in Section 4.7.3. To calculate the reconstructed energies, a direct energy calibration method, analogous to the one described in [38], is used. This is done independently for events from proton and iron primaries and leads to energy calibration and attenuation parameters for simulated events. These are different from the parameters obtained from data because the energy scales are different. There are non-trivial dependencies on the hadronic interaction model and other parameters of the simulations. Details are given in Appendix B.8.2. An example of a migration matrix obtained with SD-750 simulations is shown in Fig. 4.55a. A 50/50 mix of proton and iron simulations is used for that example. Separate matrices for the two different primaries are given in Appendix B.8.2. The migration matrix represents the probability to measure a primary energy E_{rec} given the true energy E_{true}. It is the conditional probability $p(E_{\mathrm{rec}}|E_{\mathrm{true}})$. The energy resolution is the width of this matrix. For a particular slice in energy E_{true}, histograms of the relative energy differences are plotted in Fig. 4.55b. The red circular markers represent proton simulations, while the blue and rectangular points correspond to iron simulations. Depending on the energy interval, the resolution for iron primaries is 2 % to 5 % smaller than the resolution for proton. This is related to the larger muon content in air showers initiated by iron primaries



(a) Migration matrix for SD-750 simulations

(b) Histogram of energy differences

Figure 4.55: Migration matrix and energy differences in SD-750 simulations.

(and thus the larger amount of photo-electrons). Both the distributions of proton and iron events in Fig. 4.55b are compared to predictions using a normal (dashed) or a log-normal (line) kernel. On average, data are better described with normal fluctuations in energy. The detector kernel is $\mathcal{N}(E_{\rm rec}|E_{\rm true}, \sigma_{\rm tot}[E_{\rm true}])$.

Assuming a functional form for $\sigma_{\rm tot}(E)$, a model of the resolution is obtained by maximizing the following log-likelihood:

$$\log \mathcal{L} = \sum_i^n \log \mathcal{N}(E_{{\rm Rec},i}|E_{{\rm True},i}, \sigma_{\rm tot}[E_{{\rm True},i}]). \tag{4.59}$$

The sum runs over all n simulated events. $\mathcal{N}$ is the normal distribution. This functional form describes the resolution well:

$$\frac{\sigma[E]}{E} = a + b\sqrt{\frac{E}{E_0}}\,. \tag{4.60}$$

E_0 should be chosen to minimize the correlations of the parameters a and b.

The plots in Fig. 4.56 summarize the energy resolutions for SD-750 and SD-1500 data and provide a detailed comparison between models obtained from simulations and data. Resolutions for the measurement with the SD-750 are visualized in Fig. 4.56a. The sampling fluctuations, as estimated from data, are shown with black circular markers. This includes the different components of the uncertainty as stated in Eq. (4.56). As a function of energy, the sampling fluctuations decrease from around 12 % at 3×10^{17} eV to less than 5 % above 3×10^{18} eV. This is due to the overall decrease in signal sizes and the increase of the corresponding fluctuations. The slight dependence on zenith angle reflects the attenuation of shower signals with increasing zenith angle. The shower-to-shower fluctuations $\sigma_{\rm sh} = (10\pm1)\,\%$ obtained with the energy calibration method are shown with the black dashed line. The shaded area around it represents the $1\,\sigma$ statistical uncertainty. Adding the sampling and shower-to-shower fluctuations from data leads to an estimate of the total resolution as shown with the black rectangular makers. It should be compared to the model of the total resolution derived from simulations. The latter is plotted with the black solid line and the corresponding uncertainty region. This model is taken from [174] and also used for the derivation of the flux in [86]. There is an excellent agreement between the two

separate estimates of the total resolution σ_{tot}. It ranges from (17±2) % at 3×10^{17} eV to less than (5±1) % above 10^{19} eV. Due to the trigger efficiency, the model from simulations is obtained from events with energies above 3×10^{17} eV. Deviations below this energy could reflect inaccuracies in the functional form or selection effects. Within the current event statistics, a reasonable fit of the shower-to-shower fluctuations as a function of energy is hard to obtain. A slight decrease with energy is expected, which would possibly lead to improved agreement of data and simulation in Fig. 4.56a.

The comparison of SD-1500 data and simulations is shown in Fig. 4.56b. In this case, the shower-to-shower fluctuations decrease with energy. Again there is a slight dependence of the sampling fluctuations on zenith angle, which is less distinct than in the case of SD-750 data. The total resolution ranges from (19±2) % at the threshold of full trigger efficiency of 3×10^{18} eV to less than (14±2) % above $10^{19.5}$ eV. At highest energies, there is a small mismatch between the estimate of the total resolution from data and simulations. Though being still well within statistical uncertainties, this difference could reflect the change in mass composition at highest energies (with a smaller resolution when the composition gets heavier) or an underestimation of the resolution in data. The discrepancy could also be due to a broadening effect when two distributions with different mean values (proton and iron) are added. Due to its higher credibility, the model obtained from simulations is used for the further analysis. The connected uncertainties are propagated into uncertainties on the measured and unfolded flux.

The derived resolution models are used to correct the measured flux for migration effects. The exact functional forms and fitted parameters are listed in Appendix B.8.2.

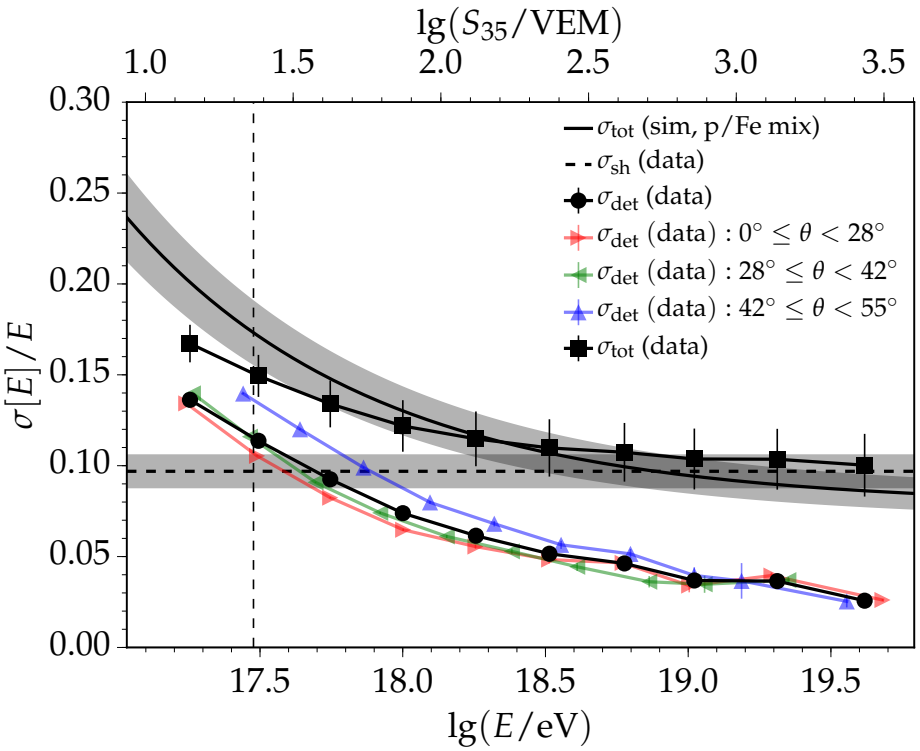
The general problem of unfolding can be stated as follows:

$$J'(E') = \int \mathrm{d}E\; K(E'|E, \sigma[E])\; J(E), \tag{4.61}$$

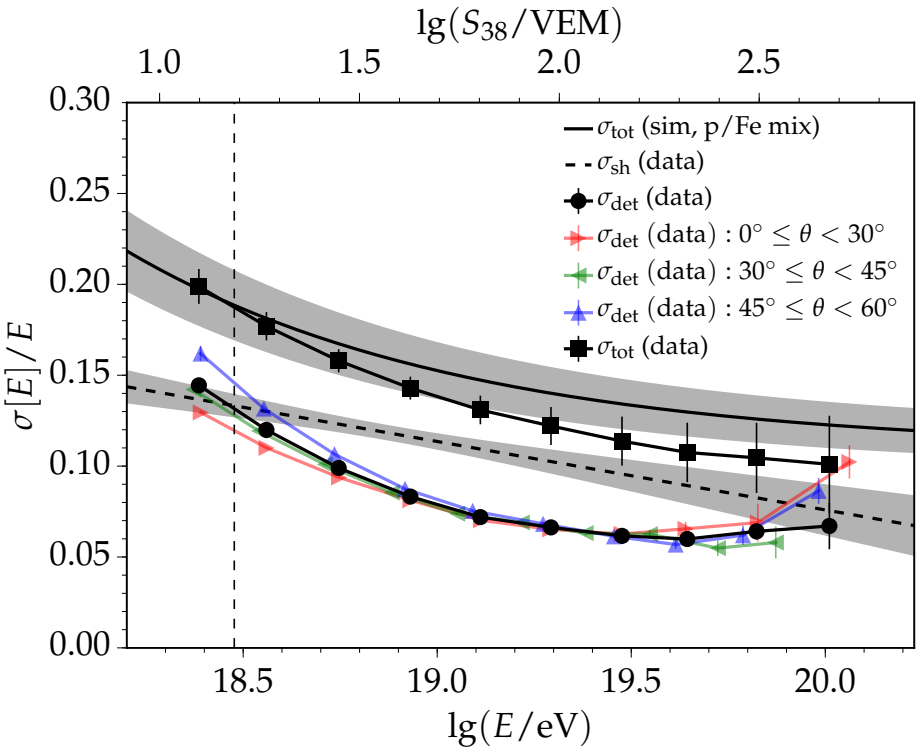
with the observed quantities E' and J'. The distribution K is the detector kernel. It represents the energy resolution with a suitable probability distribution. There are various methods to unfold or solve Eq. (4.61) with respect to the true flux $J(E)$. With most of these unfolding methods, one encounters problems like the introduction of additional fluctuations to the unfolded data [175, 176]. Regularization methods have been developed in order to solve these problems to some degree [177, 178]. Still, the results of these methods are difficult to control and interpret. The use of black-box analyses should not be favored for scientific purposes. In this work, I am using the forward-folding method to fit the spectra. In this method, a suitable model to describe the true flux $J(E)$ is chosen and folded with the detector kernel. This results in the smeared flux which is fitted to data. Thus, the folding is performed in each minimization step in order to smear the true model after a change in the model parameters. This method does not inherit any of the downsides of the other unfolding methods. Its only limitation is the need for a prior knowledge on the model to describe the flux. As the shape and spectral features of the flux only change minimally with the unfolding, this is not a real limitation and the model to describe data well can already be tuned with the measured flux. The statement that the shape of the flux is only minimally influenced by the folding does only hold in this specific case, given the functional form of the resolution and the shape of the flux.

The procedure to fit and correct the measured flux $J'(E')$ starts with an appropriate binning of data. Data are divided into $i = 1 \ldots n$ bins in energy $\lg E'$. For each bin we have:

$$J'_i = \frac{\mathrm{d}N'_i}{\Delta E'_i\; \mathcal{E}_i} = \frac{1}{E'_i \log 10} \frac{\mathrm{d}N'_i}{\Delta \lg(E'_i)\; \mathcal{E}_i}. \tag{4.62}$$



(a) SD-750 data and simulations



(b) SD-1500 data and simulations

Figure 4.56: A comparison of resolutions and resolution models derived from data and simulations. The resolution due to sampling fluctuations derived from data is given for different ranges in zenith angle.

$\mathcal{E}_i$ is the exposure for bin i. The measured quantity to fit is:

$$n'_i = \frac{\mathrm{d}N'_i}{\Delta \lg E'_i} = \log 10 \; E'_i \; J'_i. \tag{4.63}$$

The model prediction $\mu_i = \mathcal{E}_i \; n_i$ for bin i is given by:

$$\mu_i(E'_i) = \int \mathrm{d}E \; \mathcal{N}(E'_i|E, \sigma(E)) \; J(E|\vec{p}). \tag{4.64}$$

$J(E|\vec{p})$ is the model for the true flux with m parameters denoted as $\vec{p}$. For simplicity I will neglect the vector of parameters $\vec{p}$ in the following equations. For numerical reasons, the actual integrations are done in logarithmic space:

$$\mu_i = \log 10 \int \mathrm{d} \lg E \; \mathcal{N}(E'_i|E, \sigma(\lg E)) \; J(\lg E) \; E. \tag{4.65}$$

For simplicity, the $\log 10$ will be absorbed into μ_i. Taking into account the detection efficiency $\epsilon(\lg E', \theta)$ leads to:

$$\mu_i = \int \mathrm{d} \lg E \; \mathcal{N}(E'_i|E, \sigma(\lg E)) \; J(\lg E) \; E \; \epsilon(\lg E', \theta). \tag{4.66}$$

Also, the resolution model might depend on zenith angle; as such, an integration over zenith angle is required:

$$\mu_i = \frac{\int \mathrm{d} \lg E \int \mathrm{d}\theta \; \mathcal{N}(E'_i|E, \sigma(\lg E, \theta)) \; J(\lg E) \; E \; \epsilon(\lg E', \theta) \; \sin\theta \cos\theta}{\frac{1}{2}\left(\sin^2\theta_{\mathrm{max}} - \sin^2\theta_{\mathrm{min}}\right)}. \tag{4.67}$$

Here I explicitly state the $\sin\theta\cos\theta$ dependence of $J(E)$ and divide by the integral over the zenith angle range covered by the measurement, which leads to the denominator in the above equation. For simplicity, this will be absorbed into the left hand side in μ_i. As stated in Section 4.8.1, I am restricting myself to zenith angle ranges for which full efficiency is ensured. Depending on energy E'_i, this leads to the range $\theta_0 - \theta_1$. An appropriate energy interval $\lg E_0 - \lg E_1$ around $\lg E'_i$ is chosen to take into account event migrations from lower and to higher energies. These changes lead to the following simplification:

$$\mu_i = \int_{\lg E_0}^{\lg E_1} \mathrm{d} \lg E \int_{\theta_0(E'_i)}^{\theta_1(E'_i)} \mathrm{d}\theta \; \mathcal{N}(E'_i|E, \sigma(\lg E, \theta)) \; J(\lg E) \; E \; \sin\theta\cos\theta. \tag{4.68}$$

In praxis, the integrations are carried out for the energy estimate S instead of energy. The major advantage is that the resolution and efficiency models can be parametrized as a function of energy estimate and are thereby independent of the overall energy scale (which is prone to regular changes). Using the relation $\mathrm{d} \lg E = B \, \mathrm{d} \lg S$, with the slope parameter B of the energy calibration, leads to:

$$\mu_i = B \int_{\lg S_0}^{\lg S_1} \mathrm{d} \lg S \int_{\theta_0(E'_i)}^{\theta_1(E'_i)} \mathrm{d}\theta \; \mathcal{N}(S'_i|S, \sigma(\lg S, \theta)) \; J(\lg E) \; E \; \sin\theta\cos\theta. \tag{4.69}$$

This finally allows one to construct the Poisson log-likelihood as follows:

$$\begin{aligned} \log \mathcal{L}(\vec{p}) &= \sum_i \log \mathrm{Pois}(n'_i|\mu_i(\vec{p})) \\ &= \sum_i n'_i \log \mu_i + \mu_i, \end{aligned} \tag{4.70}$$

dropping constant terms in the second equation; this should only be done when a singular fit is performed and not in case of combined fits as described in Section 4.12. In the latter case, the constant terms are relevant again due to the different numerical magnitude of log-likelihoods that are summed. The minimization of $-\log\mathcal{L}$ from Eq. (4.70) is done with MINUIT-2 [110] or NLOPT [179]. The two-dimensional integrations are evaluated numerically with multi-dimensional adaptive integration routines from the ROOT software. Different algorithms were tested to produce consistent results. The validity of the forward-folding method is proven with a toy study in Appendix B.7.1.

Applying the forward-folding method to the spectra measured with the SD-750 and SD-1500 leads to the results shown in Fig. 4.57. To better visualize the spectral features, the spectra are scaled with E^3. Shown is the forward-folded flux $J(E)$ obtained from the measured flux $J'(E)$ via the folding correction function $c(E)$. The latter is defined as the ratio between the true fitted flux model and the smeared model:

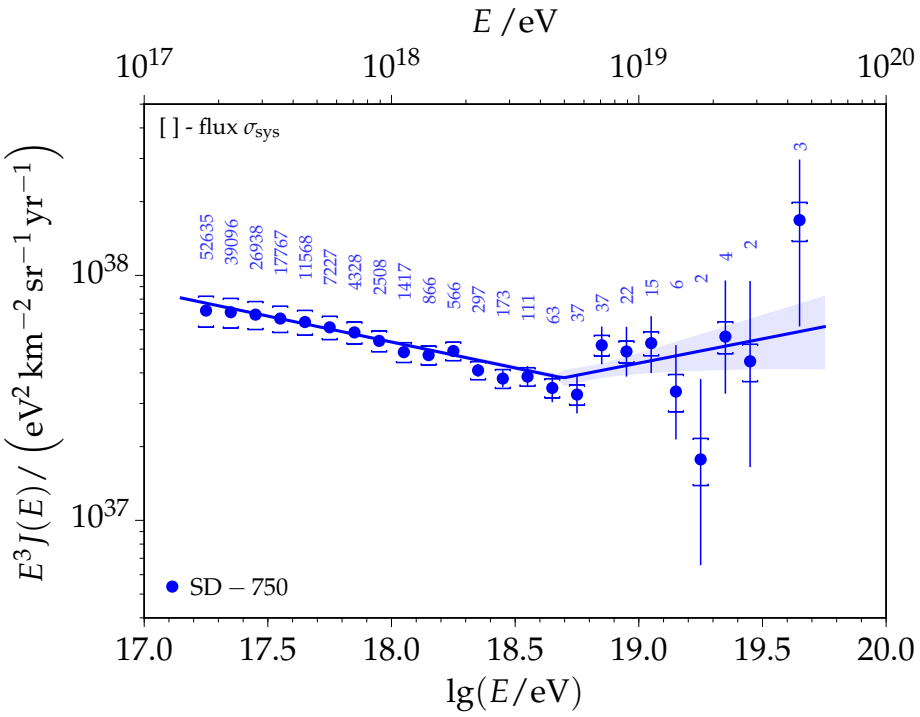
$$c(E) = \frac{J(E)}{J'(E)}\,. \tag{4.71}$$

Given this correction function, the measured flux in each bin is corrected by the value at the center of the $\lg E$ bin. I note that this implicitly contains a correction for binning effects (meaning that the center of the bin does not coincide with the average of the entries within the bin) because the fit includes an integration over the distribution in Eq. (4.69). The SD-750 spectrum in Fig. 4.57a was fit for energies above 3×10^{17} eV. A simple broken power-law with the slope γ_1 below the ankle (E_a) and another slope γ_2 above the ankle was used. There is a distinct change of the spectral index at 18.69 ± 0.07 in $\lg E$. The spectrum gets less step and changes its slope from -3.21 ± 0.01 to -2.80 ± 0.15. The fit parameters and uncertainties are given in Table 4.6. The second uncertainties represent systematic uncertainties from varying the energy-dependent flux systematics by $\pm1\,\sigma$. The blue line and area represent the fit and its $1\,\sigma$ statistical uncertainty. The rectangular caps below and above each point represent the total systematic uncertainty on the flux. This uncertainty is highly correlated among different bins. The number of measured events within a bin is stated above.

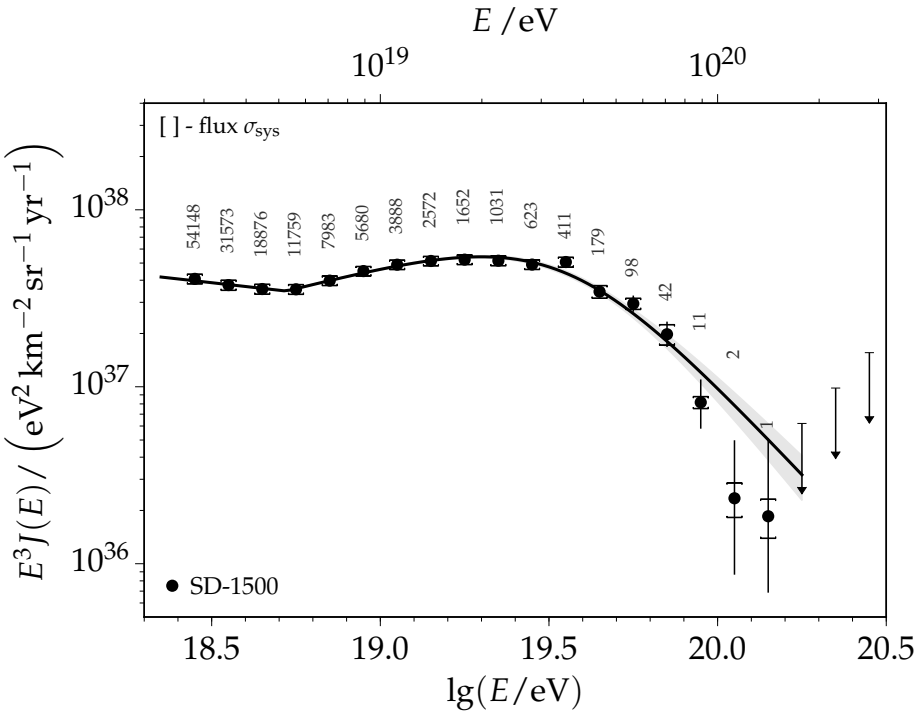
The SD-1500 spectrum is shown in Fig. 4.57b. The fit was done for events with energies above $10^{18.4}$ eV. Events at lower energies were taken into account to estimate the impact of event migrations from the region of partial trigger efficiency. A model with a hard break at the angle and smooth suppression at the highest energies is used (see Appendix B.8.1). The ankle position is fit at 18.721 ± 0.005 in logarithmic energy. The spectral slope changes from -3.21 ± 0.01 to -2.51 ± 0.04. A strong suppression of the spectrum occurs at highest energies. The flux drops to half of its value without suppression at $\lg(E_s/\text{eV}) = 19.54\pm0.03$. The change in the spectral index is 2.53 ± 0.20. Again, the parameters are summarized in Table 4.6. Note that for the SD-1500, the systematic uncertainty on the flux has a negligible energy dependence. Therefore, a systematic uncertainty is derived only for the flux normalization.

The folding correction functions are plotted in Fig. 4.58. The correction for SD-750 data is shown in Fig. 4.58a. The measured flux is corrected down by 10 % at 3×10^{17} eV. At higher energies, the corrections becomes small and negligible. There is a strong deviation from the model for the first two bins. This is due to a change in the spectral shape, which is not taken into account in this fit. The correction function for SD-1500 data is similar. The magnitude of the correction is 10 % at 3×10^{18} eV and smaller at higher energies. Above $10^{19.5}$ eV, the correction increases again due to the significant steepening of the spectral slope. The energy resolution decreases further at highest energies.

A comparison of the energy spectra is shown in Fig. 4.59. In Fig. 4.59a, the spectra are plotted in the usual E^3J representation. Together, the measurements provide a unique



(a) SD-750 flux



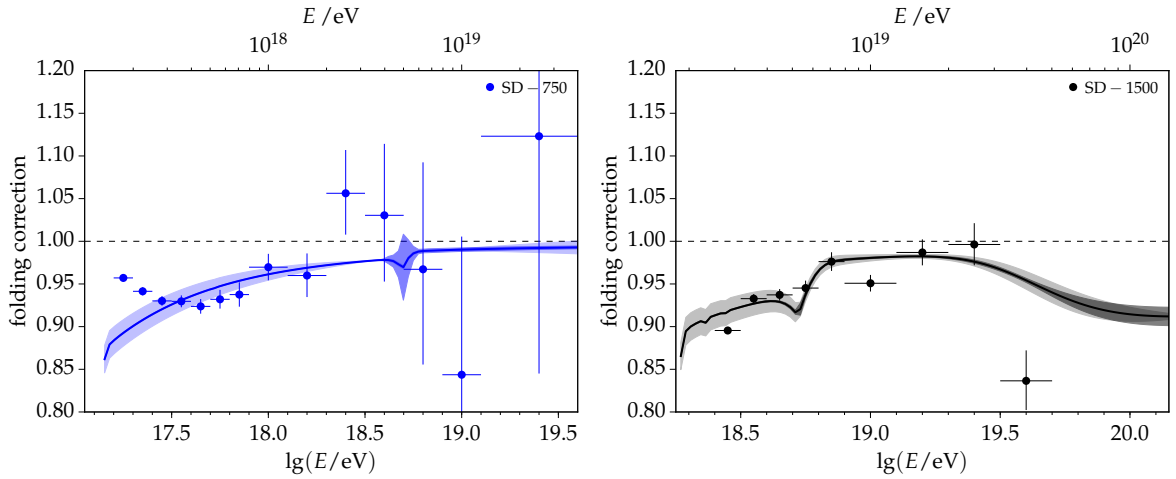
(b) SD-1500 flux

Figure 4.57: Energy spectra corrected for migration effects. The data points are plotted and compared to the fit models. The number of raw events within a bin is stated above the corresponding bin.

Table 4.6: Overview of the spectral parameters from fits to SD-750 and SD-1500 data. The spectra are corrected for migration effects as described in this section. The second uncertainties reflect systematic uncertainties from varying the (partially energy-dependent) flux systematics by $\pm 1\,\sigma$.

Flux	$\lg\left(a/\left(\mathrm{km}^2\,\mathrm{yr}\,\mathrm{sr}/\mathrm{eV}\right)\right)$	$\lg(E_a/\mathrm{eV})$	γ_1
SD-750	$-15.469 \pm 0.002 \pm 0.04$	$18.69 \pm 0.07 \pm 0.03$	$-3.21 \pm 0.01 \pm 0.03$
SD-1500	$-17.912 \pm 0.001 \pm 0.02$	18.721 ± 0.005	-3.21 ± 0.01

Flux	γ_2	$\lg(E_s/\mathrm{eV})$	$\Delta\gamma$
SD-750	$-2.80 \pm 0.15 \pm 0.2$	-	-
SD-1500	-2.51 ± 0.04	19.54 ± 0.03	2.53 ± 0.20

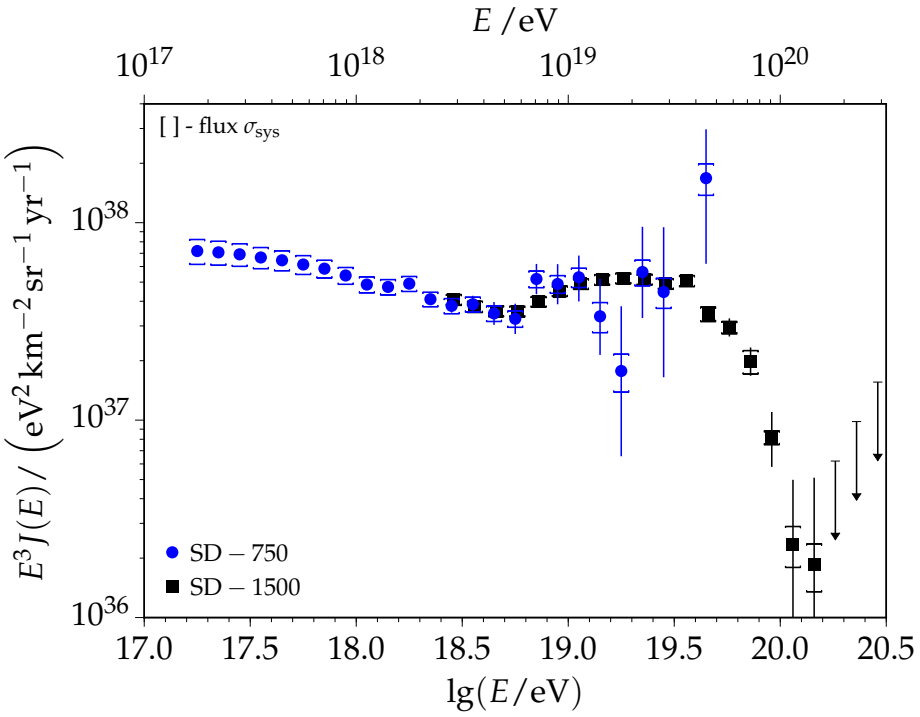


(a) SD-750 corrections

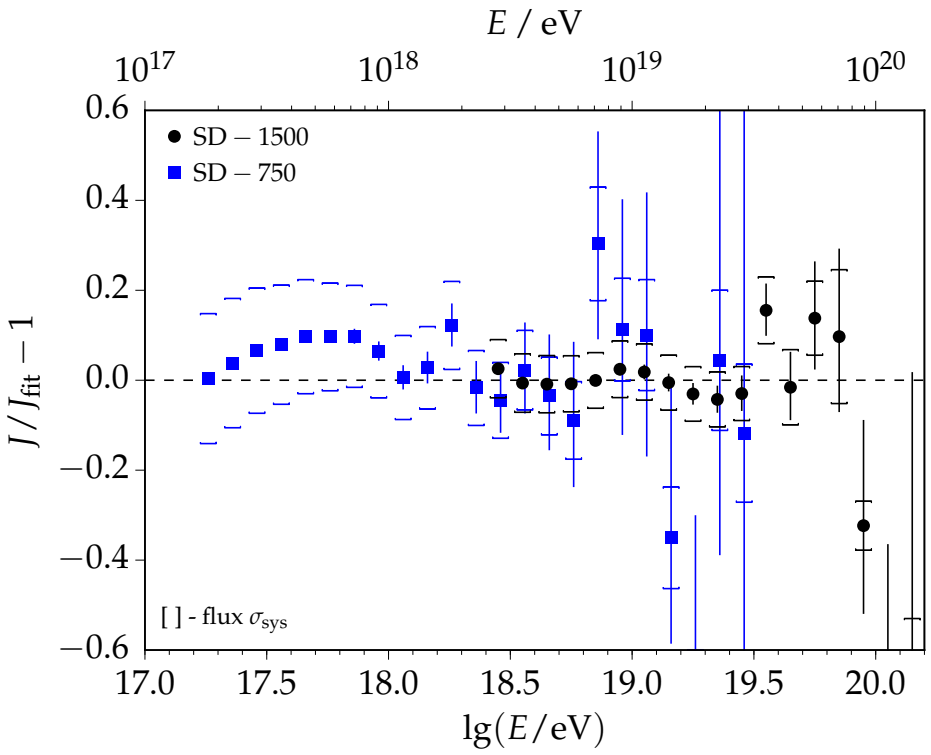
(b) SD-1500 corrections

Figure 4.58: Folding correction factors obtained from the forward-folding fits. With respect to the energy spectra in Fig. 4.57, the points were slightly re-binned for visualization purposes.

measurement of the flux of UHECRs over three decades in energy. There is an excellent agreement in the overlap region. Due to event statistics in SD-750 data, the shape of the flux above the ankle energy varies slightly. Another comparison of the spectra is shown in Fig. 4.59b. Both spectra are divided by the expected flux from the fit to SD-1500 data and subtracted with 1. Thus, both of the spectra are shown relative to the flux model that fits the events measured with the SD-1500. Also from this representation, it is apparent that there is a nice agreement of the spectra in the overlapping energy region, even only within statistical uncertainties. Both measurements agree on the spectral slope below the ankle of -3.21. The SD-750 flux shows a few interesting structures. Starting at lowest energies, the spectrum seems to steepen twice with the first steepening occurring around $10^{17.6}$ eV and the second one around $10^{18.0}$ eV. There is also the structure of a tightly located increase in the flux at $10^{18.3}$ eV. With respect to the surrounding bins, this feature has a statistical significance of $2\,\sigma$. The systematic uncertainties are correlated and do not count in this evaluation. Additional residual plots of the measured energy spectra are shown in Fig. 4.60. In Fig. 4.60a, the measured flux is divided by a power law with a spectral index of -3.21, as obtained from the fits to data in the energy range before the ankle. The normalization is chosen to match the measured flux in the bin centered around $E = 10^{18.55}$ eV. For a better visualization, the data measured with the SD-750 was divided into fewer bins above the



(a) SD-750 and SD-1500 flux



(b) Residual comparison

Figure 4.59: Comparison of the energy spectra measured with the SD-750 and SD-1500 arrays. There is an excellent agreement in the overlapping energy range around the ankle feature in the flux of UHECRs.

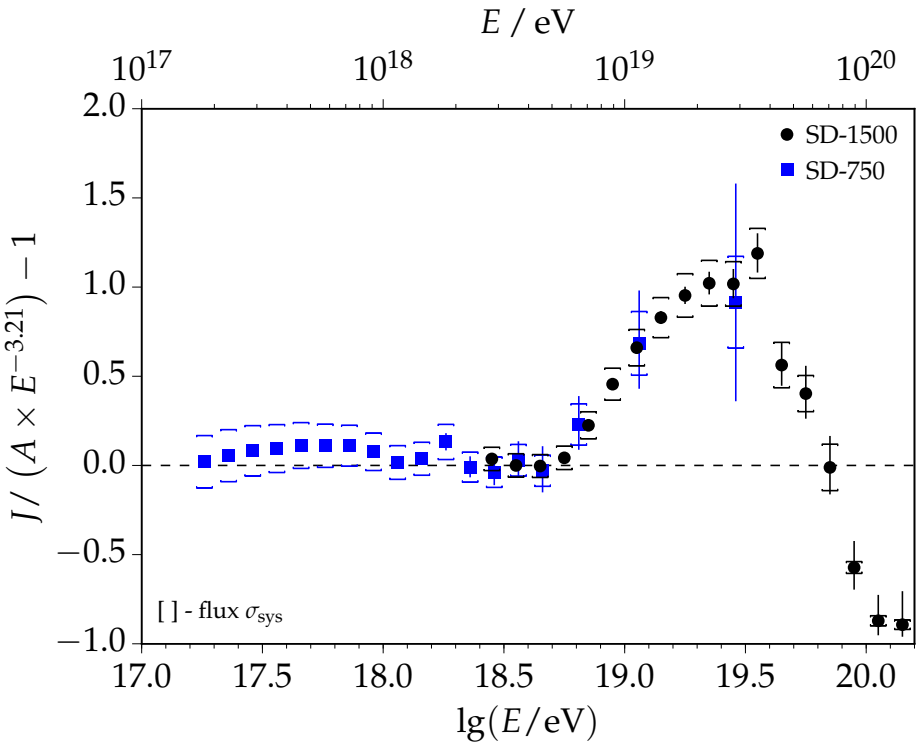
ankle energy. The plot pronounces the transition to a harder spectrum at the ankle; the nice agreement between both energy spectra is well visible. Another representation is included in Fig. 4.60b. In that case, the data are shown with respect to a model with fixed spectral slope of −2.51, the measured value after the ankle. The flux model is chosen to match the measured spectrum at 10^{19} eV.

The statistical and the different systematic uncertainties on the flux measurements are depicted in Fig. 4.61. Looking at the systematic uncertainties on the flux measured with the SD-750 in Fig. 4.61a, it is visible that the largest contribution to the systematic uncertainty is the systematic from the energy calibration as discussed in Section 4.7.7. It results from the difference between different energy calibrations and mainly from the uncertainty in the bias correction of reconstructed FD energies E_{FD}. The statistical uncertainty attributed to the forward-folding procedure stems from the fit, while the systematic uncertainty reflects the uncertainty on the energy resolution model. Both of these contributions are below 2 % for all energies. At highest energies, the statistical uncertainty on the calibration function becomes relevant. This is due to the low event statistics though and each change in the calibration parameters can lead to the migration of a significant number of events. The uncertainties on the SD-1500 flux are much smaller. All contributions to the systematic uncertainty are below 5 % at almost all energies. The total systematic uncertainty on the flux is 8 %.

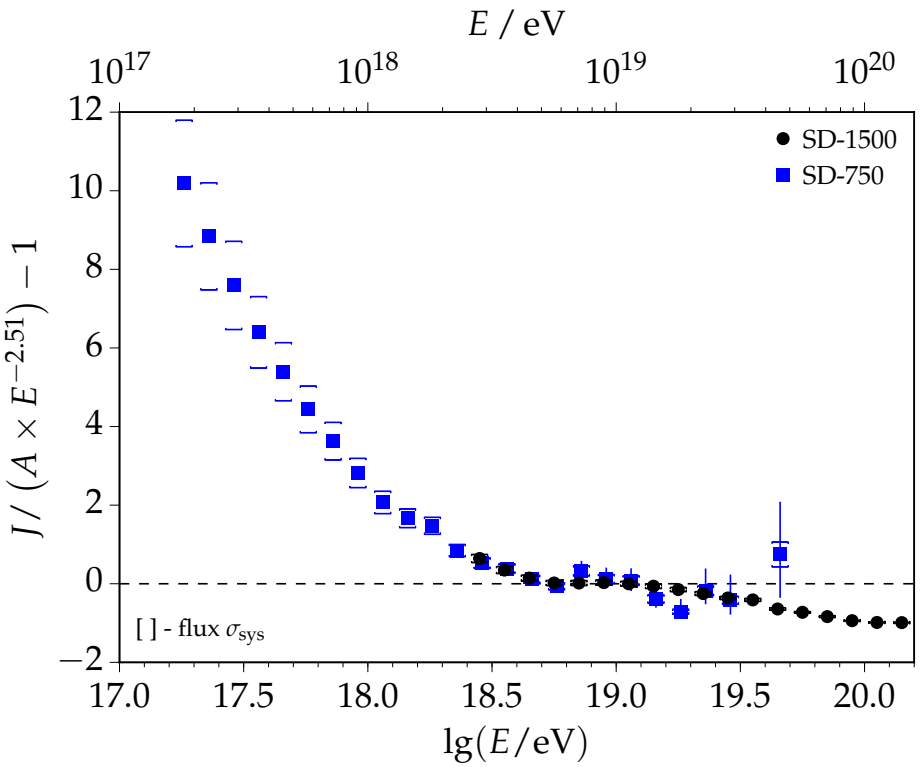
Both of the energy spectra share the overall systematic uncertainty on the revised energy scale of Auger of 14 %, with a minimal dependence on energy [168]. The different contributions to this uncertainty are from the knowledge of the fluorescence yield (3.6 %), atmospheric conditions (3 % to 6 %), absolute detector calibration (9 %) and shower reconstruction (6 %) [168]. The invisible energy is calculated with a new, simulation-driven but model-independent method with an uncertainty of 1.5 % to 3 % [180].

4.10.1 Spectral shape of the SD-750 flux

The rather intricate shape of the flux derived from SD-750 data is analyzed with additional fits and plots. First, the spectrum for energies above 10^{18} eV is forward-folded and fit to a broken power law (as described in Section 4.10). The resulting spectral slopes before and after the ankle at $\lg(E_{\mathrm{ankle}}/\mathrm{eV}) = 18.66 \pm 0.10$ are -3.24 ± 0.04 and -2.74 ± 0.20. A residual of the measured flux of all events relative to that model is depicted in Fig. 4.62a. For visual purposes, the data at high energies is grouped into fewer bins. Above 10^{18} eV, the data and model agree well. The flux in the energy bin centered around $10^{18.35}$ eV shows a $2\,\sigma$ upward fluctuation. No further structure of this feature, e.g. as a function of declination, was found. The first two energy bins below 10^{18} eV indicate a slight steepening of the flux. At lower energies, the flux flattens to roughly the same spectral index than before. However, another change in the spectral index occurs at $10^{17.5}$ eV. A first study indicates a change of the slope to −3.05. This is in agreement with an earlier analysis of the SD-750 spectrum reported in [127]. The difference to the fit model in units of the statistical uncertainty is detailed in Fig. 4.62b. The first two points are at values below −4. The plot reflects the systematic deviation of the flux from the fit model for energies below 10^{18} eV. However, the systematic uncertainties are very large; this does not rule out the presence of an energy-dependent bias. For a detailed study, the flux is divided into ten different zenith angle ranges in Fig. 4.62c. Each of the individual spectra covers a tenth of the total exposure. It is visible that the first two spectra, including events with zenith angles below 21°, deviate the most from the other spectra at energies below $10^{17.7}$ eV. While these deviations are below 10 %, they are statistically significant due to a large number of events. The deviations are visible in 20 % of the measured events. The explanation is an energy dependence of the

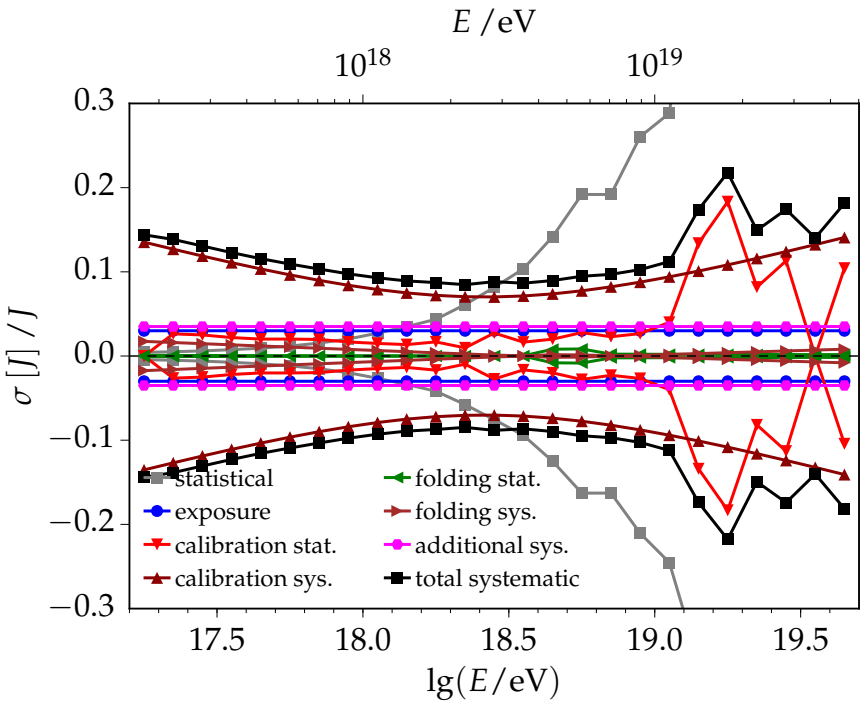


(a) Flux relative to a model with slope before the ankle

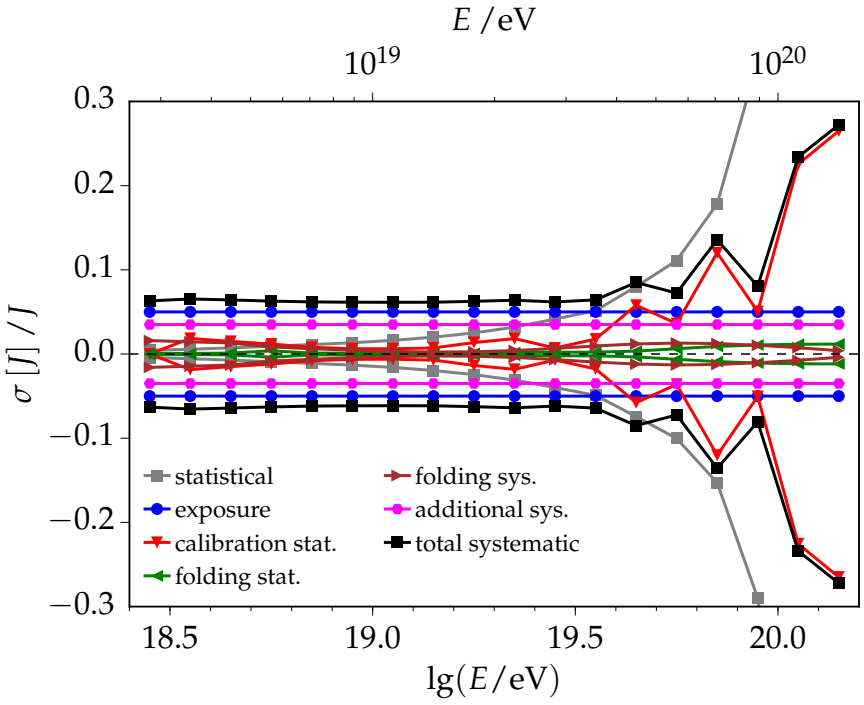


(b) Flux relative to a model with slope after the ankle

Figure 4.60: Comparison of the energy spectra measured with the SD-750 and SD-1500 arrays. Shown are relative residuals with respect to power-law flux models with fixed slopes.



(a) SD-750 data



(b) SD-1500 data

Figure 4.61: The different statistical and systematic uncertainties for the energy spectra derived in this work.

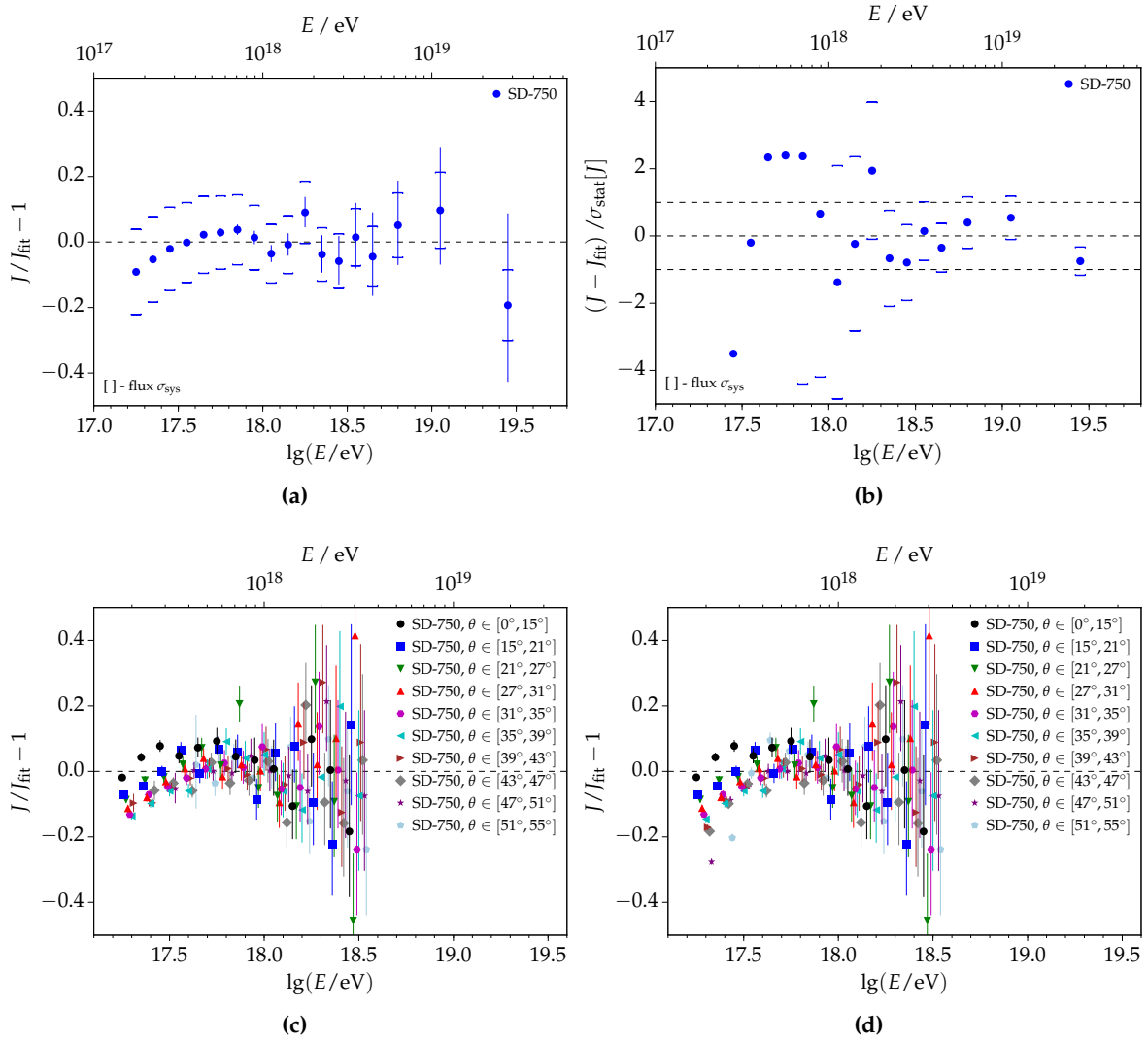


Figure 4.62: Residual representation of the unfolded SD-750 spectrum. (a) Relative residual and (b) Residual with the distance to the fit model in units of statistical uncertainty. (c) Residual representation of the SD-750 spectrum for ten different zenith angle ranges. The flux in each range is plotted with respect to the model that describes the overall flux. (d) The previous plot without taking into account the energy-dependent aperture due to the drop in T4 efficiency at lower energies.

attenuation function derived in Section 4.5.1. An over-density of events at low energies and zenith angles is apparent in Fig. 4.13a. To verify that the measured behavior is not related to a problem with the T4 efficiency (Section 4.4), the residual is repeated without efficiency correction in Fig. 4.62d. Not taking into account the loss of efficiency at low energy and high zenith angle leads to larger deviations in the first few bins. Calculating and applying an energy-dependent aperture works as expected and brings the individual spectra into much better agreement with each other. Note that each of the individual sub-spectra shows a break at an energy of around $10^{17.5}$ eV. Also note that the spectrum for the lowest zenith angle range is not affected by an efficiency correction. Attempts to further verify and fit the shape of the flux at lower energies should be the task of future analyzes.

4.11 The SD-1500 flux derived with shower universality

The universality reconstruction is described in Section 3.4. It is used to derive results on the mass composition in Chapter 5. At the same time, it is employed to obtain the measured flux of UHECRs separately to the standard reconstruction. In particular, the energy scale is driven by the calibration of the average muon content as discussed in Section 5.2. The energy is still fit freely during the reconstruction, but, due to substantial correlations between the two quantities, it is not estimated simultaneously with the relative muon content R_μ. Thus, the model of R_μ in the fit strongly influences the resulting event-by-event estimate of the energy. A CIC correction method is not used. The energy spectrum for events measured with the SD-1500 and reconstructed with the universality approach is plotted in Fig. 4.63a in comparison to the flux derived with the standard approach. A forward-folding method to correct for event migrations is not applied. Only events with zenith angles below 50° are considered in the universality spectrum (see Section 5.1). However, no additional cuts are applied. The resulting exposure is $(33\,270 \pm 1663)\,\mathrm{km^2\,yr\,sr}$. Both spectra agree perfectly with each other up to a reconstructed energy of 10^{19} eV. At higher energies, there is a linear increase in the relative difference of the estimated fluxes. As depicted in Fig. 4.63b, the difference increases to (20 ± 15) % at 10^{20} eV. It might be related to the SD-1500 energy calibration and the fact that a change in the slope of the calibration function at highest energies is not taken into account (compare to the problem of the SD-750 calibration detailed in Section 4.7.7). Further studies on this matter are necessary. It is unlikely that the difference in the energy spectra is due to the energy resolution because the study of simulations and golden hybrid events indicates that the resolutions achieved with the standard and universality reconstructions are very similar. The spectrum derived with shower universality is divided into different zenith angle ranges in Fig. 4.63c. Below energies of $10^{18.8}$ eV, there are more events at large zenith angles. This is caused by event migrations and a slight increase in the energy resolution for large zenith angles (compare to Fig. 4.31c). Above energies of $10^{19.5}$ eV, the spectrum in the highest zenith angle regime is consistently larger than the other sub-spectra. Though not statistically significant, this should be studied in more detail. It is directly related to the energy differences as depicted in Fig. 4.63d. While there is an agreement between the reconstructed energies of the universality and standard reconstruction for small zenith angles, there is a difference of up to 5 % to 10 % at large zenith angles. The integrated number of events above different energy thresholds are shown as a function of zenith and azimuth angle in Fig. 4.64. An identical number of events is expected within each bin. To quantify deviations from a constant behavior, each of the profiles is fit with a chi-squared fit. The corresponding χ^2/n_{dof} values are stated within the plots. Judging from Fig. 4.64a, the profiles indicate that there is an increase in the intensity with increasing zenith angle. This indicates the necessity of additional quality cuts (similar to the X_{max} analysis in Section 5.1). Further studies on this issue are required.

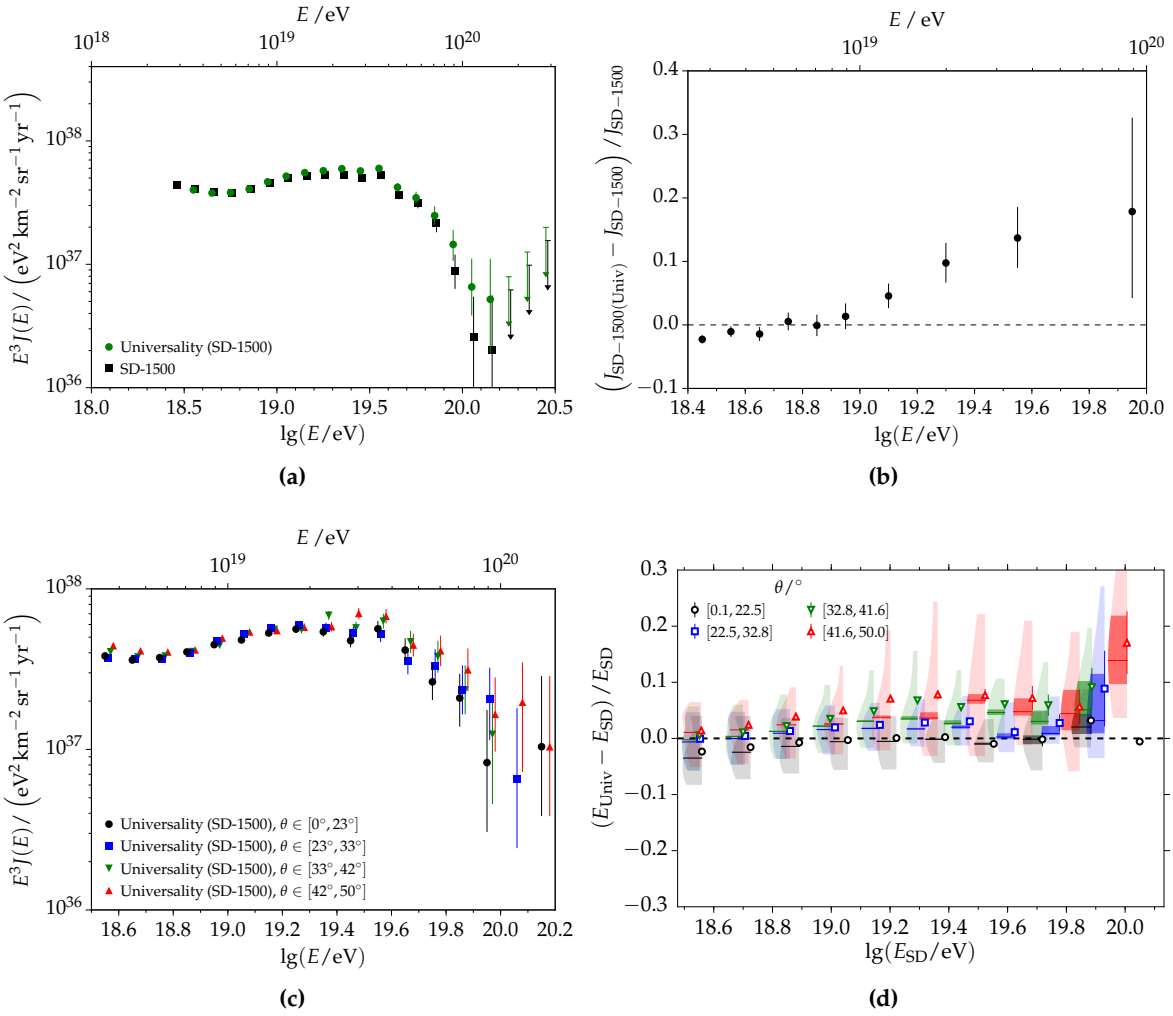
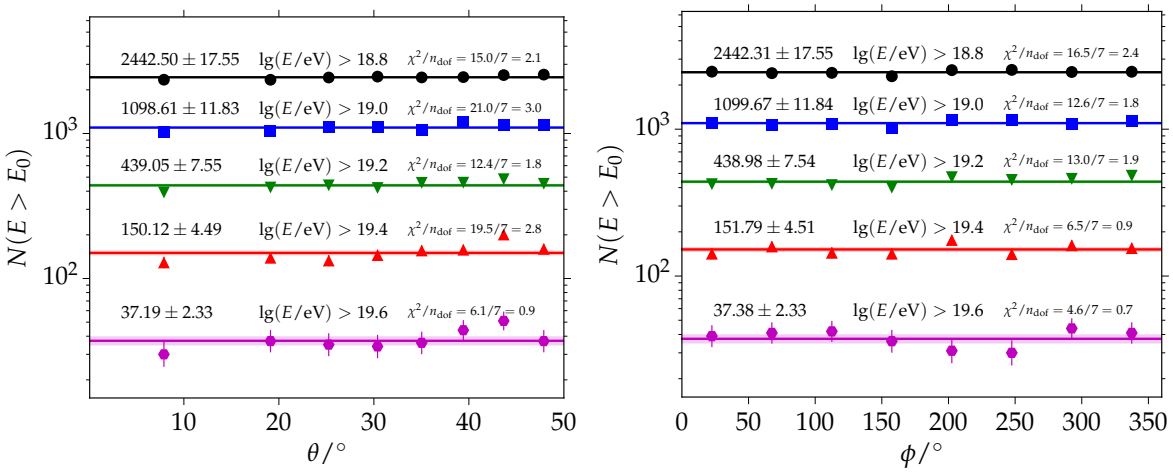


Figure 4.63: (a) The energy spectrum derived with the universality reconstruction compared to the flux obtained with the standard approach. (b) A profile of the relative flux differences with optimized energy bins. (c) The flux of SD-1500 events reconstructed with shower universality for different ranges in zenith angle. (d) Relative difference in the reconstructed energies as a function of the energy determined with the standard reconstruction. The data are separated into different zenith angle ranges.

4.12 Combining flux measurements

The energy spectra presented in Section 4.10 have different systematic uncertainties attributed to their flux points. These are depicted in Fig. 4.61a for the measurement of SD-750 data and in Fig. 4.61b for the SD-1500 measurement. Some of these uncertainties depend on energy itself, for example the contribution on the SD-750 systematic from the energy calibration as discussed in Section 4.7.7. Furthermore, the SD measurements share the same uncertainty of the overall energy scale of 14 % (independent of energy). To combine the energy spectra, a likelihood method was developed that takes the different systematic uncertainties into account. The general log-likelihood for the combination of n spectra reads:

$$\log \mathcal{L}_{\text{comb}} = \sum_{i=1}^{n} \log \mathcal{L}_i^{\text{Pois}} + \log \mathcal{L}_i^{\text{Ecal}} + \log \mathcal{L}_i^{\text{Norm}} \,, \tag{4.72}$$



(a) Zenith **(b)** Azimuth

Figure 4.64: Integrated number of events with energies larger than the respective energy threshold as a function of zenith and azimuthal angle.

with the Poisson log-likelihood $\log \mathrm{L}^{\mathrm{Pois}}$ and the normalization term $\log \mathcal{L}^{\mathrm{Norm}}$. To make use of all correlations, the energy calibration can be included in the fit of the flux model. In praxis, this is done by adding the respective likelihood for the energy calibration $\log \mathcal{L}_i^{\mathrm{Ecal}}$ as detailed in Section 4.7. With the current data, the statistical uncertainty on the energy calibration is 1 % to 3 % (as shown in Fig. 4.29b and Fig. 4.30b). Thus, it can be neglected in the combination procedure because the other systematic uncertainties dominate.

The log-likelihood of the spectrum combination for energy-independent systematic uncertainties is:

$$\log \mathcal{L} = \sum_{i=1}^{n} \sum_{j=1}^{k_i} \log \mathcal{L}_{ij}^{\mathrm{Pois}} + \frac{1}{2} \left(\frac{(a_i - 1)^2}{\sigma_i^2} \right) , \tag{4.73}$$

with k_i indicating the number of bins of spectrum i. σ_i is the total energy-independent systematic uncertainty on the flux of spectrum i. a_i denotes the overall normalization factor of spectrum i. It is determined during the fit such that the best combination of spectra in terms of their systematic uncertainties is achieved. The Poisson log-likelihood for spectrum i and energy bin j is given by:

$$\mathcal{L}_{ij}^{\mathrm{Pois}} = \frac{\mu_j^{n_j} \exp\left(-\mu_j\right)}{n_j!} , \tag{4.74}$$

with the number of measured events n_j and the number of expected events μ_j. If a standard fit of the measured flux is done, μ_j is given directly by the flux model. In the forward-folding method μ_j is determined through the convolution of the flux model with the detector kernel and efficiency models as formulated in Eq. (4.69).

Employing the described method to combine the flux measurements derived from SD-750 and SD-1500 data, results in the following set of flux parameters:

$$\begin{aligned} a_0 &= -17.89 \pm 0.02 \pm 0.02\,, \\ \lg(E_a/\mathrm{eV}) &= 18.72 \pm 0.01 \pm 0.01\,, \\ \gamma_1 &= -3.20 \pm 0.01 \pm 0.04\,, \\ \gamma_2 &= -2.52 \pm 0.03 \pm 0.02\,, \\ \lg(E_s/\mathrm{eV}) &= 19.56 \pm 0.03 \pm 0.02\,, \\ \Delta\gamma &= 2.6 \pm 0.2 \pm 0.2\,, \\ a_{SD-1500} &= 1.047 \pm 0.041\,, \\ a_{SD-750} &= 0.978 \pm 0.038\,. \end{aligned} \tag{4.75}$$

The normalization factors a_{SD-750} and $a_{SD-1500}$ are close to zero in terms of their statistical uncertainties. The spectral parameters are in agreement with the parameters obtained from individual fits to the energy spectra in Table 4.4. The second uncertainties in Eq. (4.75) represent systematic uncertainties due to differences in the parameters of the combined fit model to the results from fitting the individual spectra and from systematic uncertainties of the individual flux measurements. The resulting combined SD spectrum is depicted together with the individual spectra in Fig. 4.65b. The latter are multiplied with the respective normalization factors from the previous equation. The combined spectrum is shown together with the fit and the number of events within each bin in Fig. 4.65a. Due to the structure of the SD-750 flux, only events with energies above 10^{18} eV were taken into account to determine the fit parameters, including the normalization factors. After correction with the forward-folding factors and the normalization factors, the combined flux points are calculated as the weighted average of the flux points of the individual spectra. The weight for each energy bin is given by the number of measured events n_j. In the figures, to total systematic uncertainty on the flux is indicated with lower and upper square brackets. It is also calculated as the weighted average of the total systematic uncertainty of the individual spectra.

Energy-dependent systematic uncertainties can be included in the combination procedure with the inclusion of nuisance parameters (see e.g.[181]). For example, the modified likelihood of flux i with the energy bins $j = 1...k$ reads:

$$\mathcal{L}_i = \prod_{j=1}^{k} \mathrm{Pois}(n_j|\nu_j) \cdot \mathcal{N}(\nu_j|\mu_j, \sigma[\mu_j])\,. \tag{4.76}$$

ν_j is the nuisance parameter corresponding to the energy bin j. It is included to describe the variation around the expected number of events μ_j due to the total systematic uncertainty on the flux $\sigma[\mu_j]$. This is modeled with a normal distribution $\mathcal{N}(\nu_j|\mu_j, \sigma[\mu_j])$. ν_j itself is then used as the expected number of events in the Poisson likelihood which models only the statistical fluctuations. The k nuisance parameters ν_j are not real physics parameters. They are eliminated through requiring the boundary conditions:

$$\frac{\mathrm{d}\mathcal{L}_j}{\mathrm{d}\nu_j} = 0\,. \tag{4.77}$$

An explicit calculation from Eq. (4.76) leads to the following expression for the nuisance parameter ν_j (the one physical solution of a quadratic equation):

$$\nu_j = \frac{1}{2}\left(\mu_j - \sigma_j^2 + \sqrt{(\mu_j - \sigma_j^2)^2 + 4 n_j \sigma_j^2}\right)\,. \tag{4.78}$$

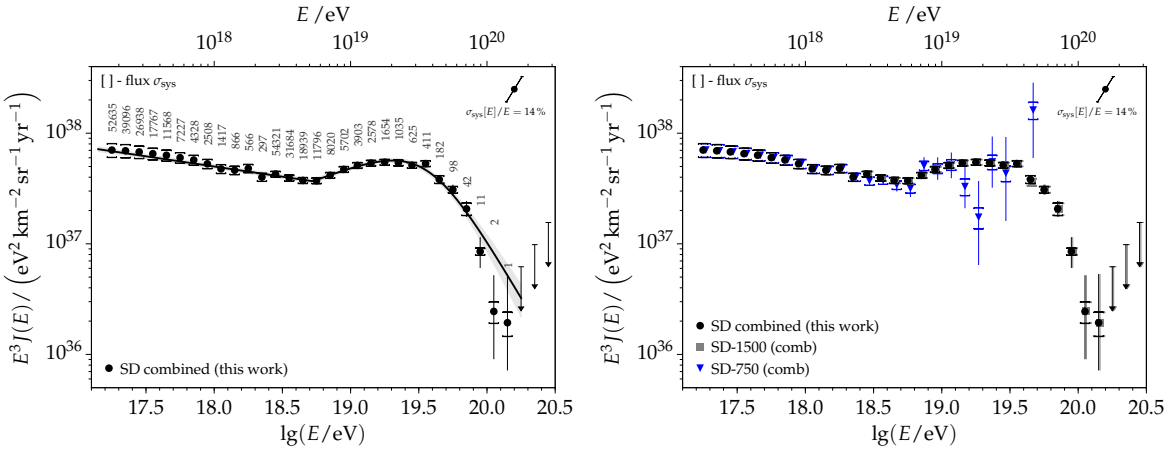


Figure 4.65: The combined SD-750 + SD-1500 energy spectrum derived within this work. The arrow in the upper right corner indicates the overall uncertainty of the energy scale of Auger. Statistical uncertainties are depicted with errorbars; systematic uncertainties are shown with square brackets. In (b), the combined spectrum is compared to the individual spectra derived from SD-750 and SD-1500 data. These spectra are multiplied with the normalization factors resulting from the spectrum combination as stated in Eq. (4.75).

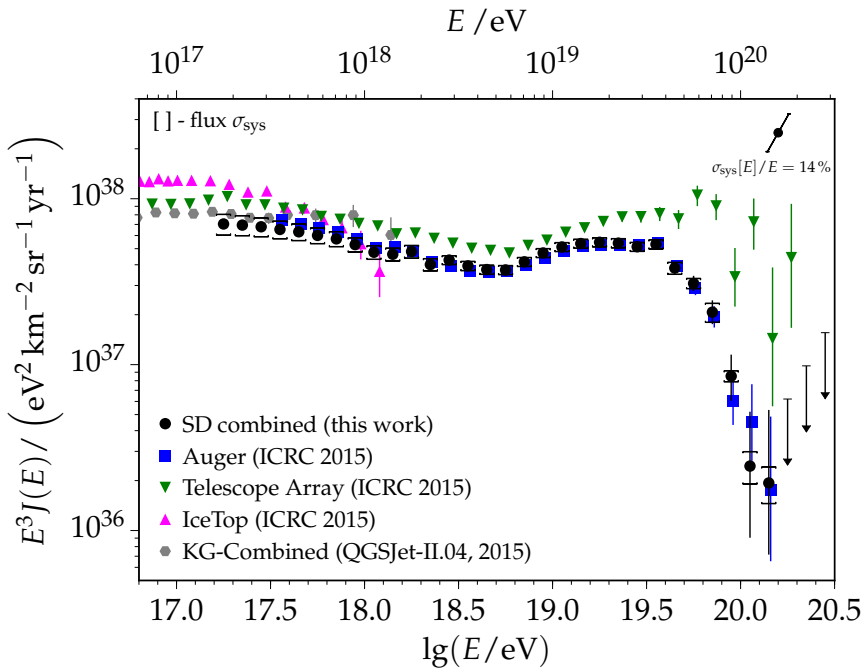
It is thereby eliminated from the fit as an explicit parameter. This method can be used to include energy-dependent uncertainties of the flux as well as the energy itself. Further studies and tests of the method are necessary in order to make it applicable to the study of UHECRs.

4.13 Comparison to other recent measurements

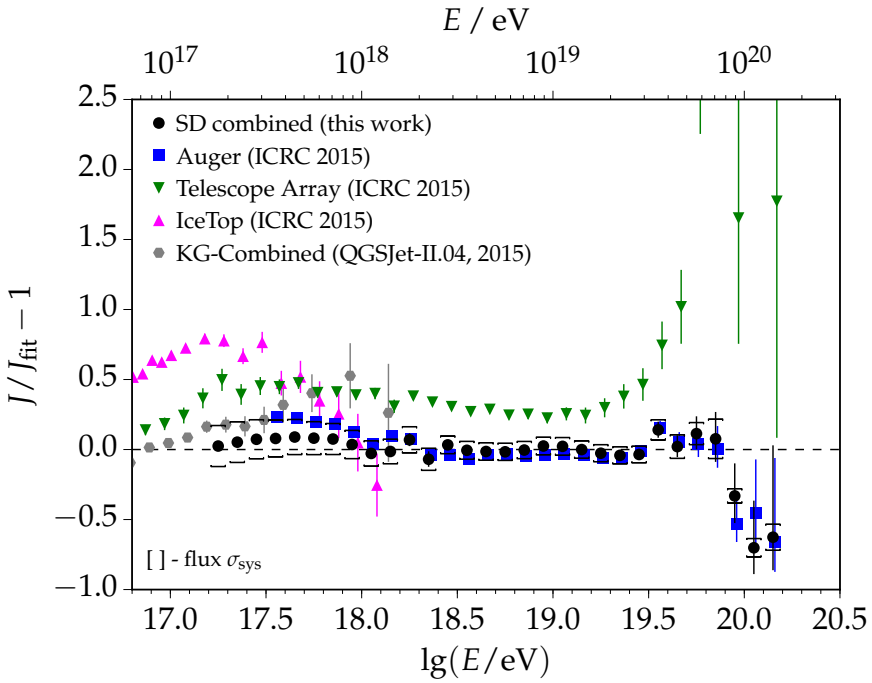
In the following, the results obtained from Auger data within this thesis are compared to the recent published Auger result [86], as well as results from KASCADE-Grande [182], IceTop [183] and TA [169]. The combined SD spectrum is obtained as described in the previous section. It is compared to other measurements in Fig. 4.66. There is an excellent agreement with the combined Auger spectrum [86] and a reasonable match with the measurement of KASCADE-Grande. The flux measured by the TA collaboration is systematically larger. This indicates an offset of the overall energy scale of up to 20 %, which is still in accordance with the systematic uncertainties on the energy scale of 14 % (21 %) of Auger (TA). Except for that, the spectral features are in accordance. The recent measurement using the IceTop detector at the South Pole is not in agreement with the other results. This difference is probably related to the use of old air shower simulations and assumptions on the mass composition. For a more quantitative comparison, all spectra are shown relative to the model that describes the combined SD flux in Fig. 4.66b.

4.14 Comparison to astrophysical scenarios

In Fig. 4.67, the combined SD spectrum is compared to predicted fluxes from different astrophysical models. Fluxes from a pure proton scenario are depicted in red while pure iron scenarios are shown with blue lines. The solid lines represent simulation results assuming a maximum injection energy of sources of $10^{20.5}$ eV. The dashed lines correspond to



(a)



(b)

Figure 4.66: Comparison of the combined Auger SD spectrum derived within this work with other measurements. The arrow in the upper right corner indicates the overall uncertainty of the energy scale of Auger. (a) Comparison of the flux measurements in the usual E^3J representation. (b) All measurements are plotted relative to the model that describes the combined SD spectrum.

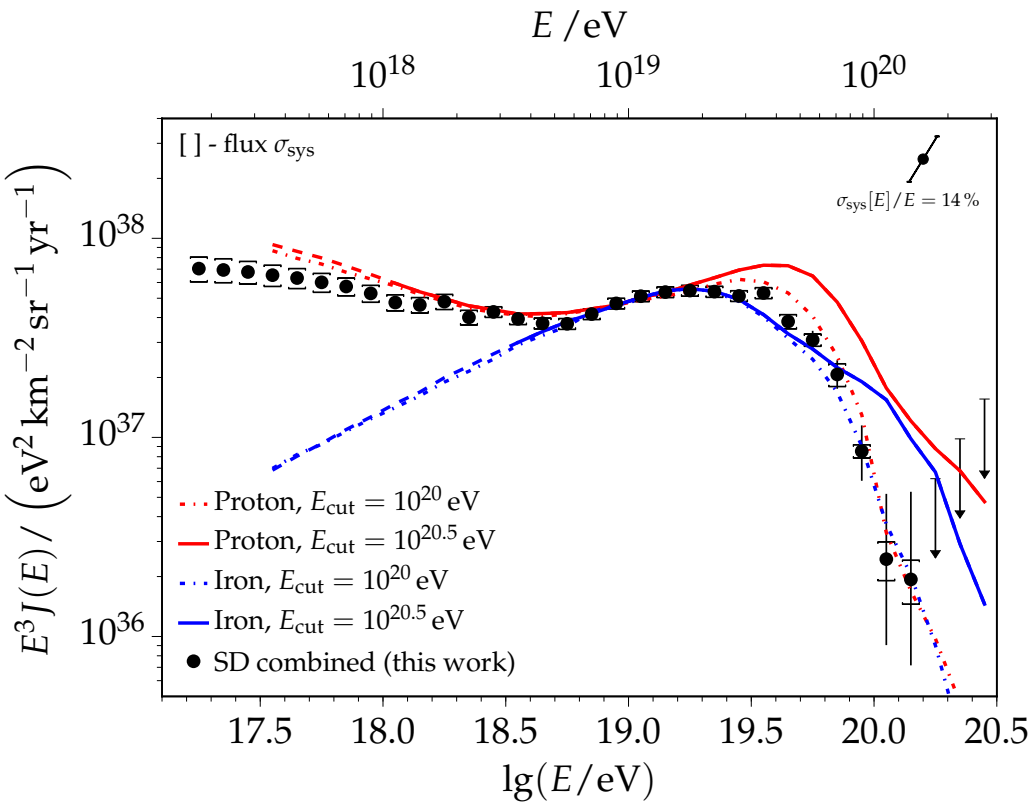


Figure 4.67: Comparison of the combined SD spectrum measured with Auger and derived in this work to fluxes from different astrophysical scenarios. The scenarios are further detailed in the text.

a maximum energy of 10^{20} eV. Above these energies, the injection spectra are exponentially suppressed. The spectral index of the injection spectra, as well as cosmological parameters, were chosen to give a reasonable match to the measurement (see also [23]). The model lines were calculated and verified using two different astrophysical simulation codes [184, 185]. It is apparent that the results of both pure composition scenarios with a maximum energy of 10^{20} eV match the measured flux well above primary energies of $10^{18.5}$ eV. Increasing the maximum injection energy leads to a worse agreement with data. The energy spectrum alone is not enough to decide between different composition scenarios. Therefore, a measurement of the mass composition of UHECRs as discussed in the next chapter is crucial for the selection between various astrophysical scenarios. Further astrophysical scenarios are discussed in Section 1.1.

CHAPTER 5

The mass composition of UHECRs

Determining the mass composition of UHECRs is one of the most important pursuits in modern astroparticle physics. The direct measurement of primary particles at lower energies ($E \lesssim 10^{14}$ eV) enables an estimation of the mass composition on an event-by-event basis. Due to the minuscule particle intensity at the highest energies, a direct measurement of particles is not worthwhile, and the mass composition needs to be inferred from properties of EASs. Due to fluctuations in the first few hadronic interactions of the shower cascade, an event-by-event estimation of the mass composition is no longer possible. Instead, the composition needs to be inferred statistically from an ensemble of measured air showers. The exact knowledge of the composition at the highest energies is crucial for our understanding of acceleration mechanisms in sources and the propagation characteristics of particles. The latter includes knowledge of the magnitude of galactic and extragalactic magnetic fields, and photon densities in interstellar space.

The depth of the electromagnetic shower maximum X_{max} provides one of the best correlations to the primary mass A ($X_{\mathrm{max}} \propto \ln A$, see Section 1.1.4). Using the FD of Auger, the longitudinal development of showers is observed, and X_{max} is directly measured [21, 80, 165]. This measurement of X_{max} is very reliable, mostly model-independent, and provides a good resolution on the order of 25 g cm^{-2} at the highest energies. For comparison, the difference in the average X_{max} of proton and iron showers is 100 g cm^{-2}. However, the duty cycle of the FD measurement is limited to about 15 % of the time. Due to that and the steeply falling flux of UHECRs, there is no FD data above $10^{19.7}$ eV at the current level of event statistics. To study the mass composition at higher energies, the shower universality approach is employed. Using the universality reconstruction, the full set of data collected with the SD-1500 is exploited to infer information on the primary mass of measured UHECRs. This enables the analysis of data up to the highest measured energies of 1.5×10^{20} eV.

The models to describe the signal and time distributions of secondary particles on the ground as a function of the properties of the primary particle are the cornerstones of the universality analysis (see Chapter 2). Together with the universality reconstruction as introduced in Section 3.4, the mass-sensitive variables X_{max} and R_μ (the relative muon content) are reconstructed from data measured only with the SD-1500. Studies on biases and resolutions of the universality reconstruction are discussed in Section 3.4.5. To distinguish the universality reconstruction from the FD results, the universality X_{max} is typically denoted as $X_{\mathrm{max}}^{\mathrm{Univ}}$.

I will concentrate on the discussion of the first two moments of the $X_{\mathrm{max}}^{\mathrm{Univ}}$ distribution, obtained from SD-1500 data. Also, results on the average logarithmic mass $\ln A$ will be given as well as results on R_μ.

As discussed in Section 3.4.5, events with saturated stations are more difficult to reconstruct, resulting in significantly larger resolutions. Additionally, reconstruction biases are present at lower energies. In the following, this subset of events is mostly treated separately, unless stated otherwise.

5.1 Event selection

For this analysis, I am considering data from the period 1. January 2004 until 31. December 2014. The quality selection for SD-1500 data is identical to the one for the analysis of the flux of UHECRs as described in Section 4.1. In addition, golden hybrid events are used for a calibration procedure of the universality method and the study of the precision and resolution of reconstructed quantities. The golden hybrid events were measured simultaneously with the SD and FD. The selection criteria for these events are given in the context of the energy calibration in Section 4.7.2. In particular, the fiducial FoV quality selection is applied to the FD events. This selection ensures that the primary composition is not distorted by the measurement with a limited FoV of the FD telescopes. A total number of 10 856 golden hybrid events pass the general selection criteria.

Except for the requirement that the reconstructed zenith angle is below 50° (to exclude a biased region in the simulations as detailed in Section 3.4.5), the following quality cuts are applied:

- Reconstructed energy above the threshold of full trigger efficiency of the SD-1500 of 3×10^{18} eV,
- Angular difference between reconstructed zenith angles $|\theta_{\mathrm{Univ}} - \theta_{\mathrm{SD}}|$ smaller than 2.5°,
- Range of $X_{\mathrm{max}}^{\mathrm{Univ}}$: $500\,\mathrm{g\,cm^{-2}} < X_{\mathrm{max}}^{\mathrm{Univ}} < 1050\,\mathrm{g\,cm^{-2}}$,
- Range of R_μ: $1.4 < R_\mu < 2.6$.

1498 events are in this energy and zenith angle range; 20 % of these events are saturated. Out of the 1207 remaining non-saturated events, 90 % survive the next quality cuts. However, the quality cuts act more severely on the set of saturated events, of which only 57 % among 291 events are selected. The selection criteria are derived from an analysis of the golden hybrid events. They are necessary to reject failed reconstructions among the current events. Overall, the relative abundance of these problems is less than 10 %. Looking at a histogram of the difference $X_{\mathrm{max}}^{\mathrm{Univ}} - X_{\mathrm{max}}^{\mathrm{FD}}$ in Fig. 5.1a reveals a small additional population with values larger than $200\,\mathrm{g\,cm^{-2}}$. In this representation, the population is hard to distinguish from a tail of the distribution. A better visualization is found later in this section. However, the distribution (representing the detector resolution) is expected to be symmetric (judging from previous analyses and simulations, as studied in Section 3.4.5), which hints at the presence of failed reconstructions. Additionally, the frequency of these problems is higher for saturated events as drawn with the red filled histogram in Fig. 5.1a. Implying a rather conservative cut of $|X_{\mathrm{max}}^{\mathrm{Univ}} - X_{\mathrm{max}}^{\mathrm{FD}}| < 200\,\mathrm{g\,cm^{-2}}$ results in Fig. 5.1b. This histogram of absolute $X_{\mathrm{max}}^{\mathrm{Univ}}$ values illustrates the lower and upper cuts that are chosen. After the selection, the distribution of data ranges from $500\,\mathrm{g\,cm^{-2}}$ to $1050\,\mathrm{g\,cm^{-2}}$. The black filled histogram of all data demonstrates the tail of failed reconstructions that is present before the selection. Additionally, as visible in a histogram of R_μ in Fig. 5.1c, there is a very small

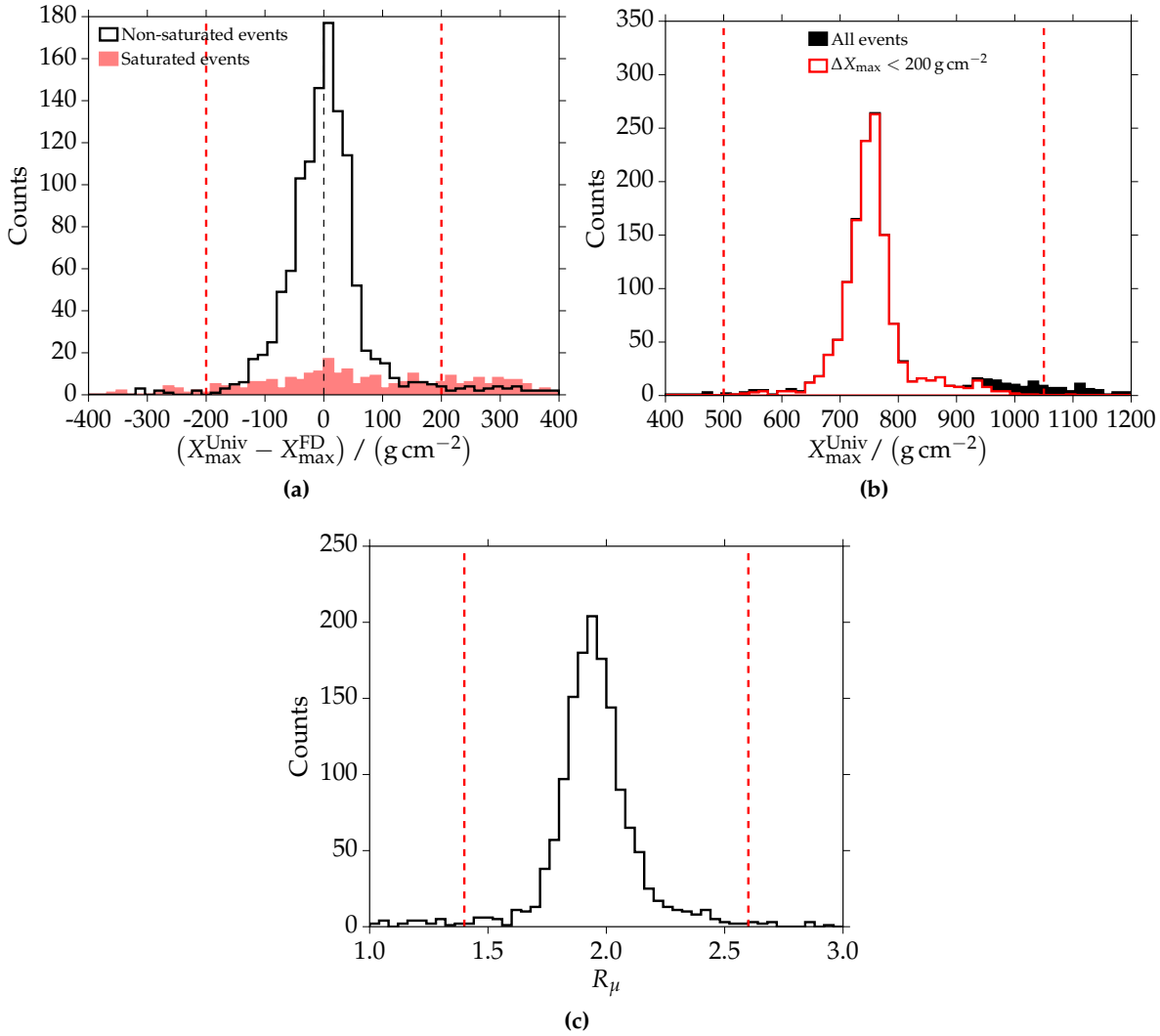


Figure 5.1: Histograms of the quantities $X_{max}^{Univ} - X_{max}^{FD}$, X_{max}^{Univ} and R_μ. Only golden hybrid data within the regions indicated with red dashed lines are selected. Most of the events outside these regions result from failed reconstructions as explained in the text.

number of outliers with respect to the reconstructed R_μ. A range of 1.4 to 2.6 is chosen for the analysis. Among the events that pass all other quality cuts, only seven non-saturated and six saturated events are outside this range.

There is a rather distinct correlation between a wrongly reconstructed X_{max} and an incorrectly reconstructed zenith angle. This correlation is depicted in Fig. 5.2a for QGSJET-II.03 simulations. Most of the problematic events are saturated and have zenith angles above 38°. The main population of well-reconstructed events is located within the white area. The specified cuts reject all events outside this rectangular area. A very similar picture is present for data, as shown in Fig. 5.2b. Also here, most of the failed reconstructions occur for events with large zenith angles or events with low-gain saturated stations. Also, the relative frequency of these problems in data is slightly larger than for simulated events. The analysis of individual events revealed one problem that leads to failed reconstructions: the presence of background signals from one or more accidental muons in the time traces. These are not part of the actual time distribution and can prevent a correct fit of the time shapes, which is the crucial part of the estimation of X_{max}^{Univ} (see Section 3.4). An example

trace is plotted in Fig. 5.2c. The accidental signal is present starting at the time bin 350 (counted in units of 25 ns), and it is clearly separate from the actual signal. The time trace is part of the event with Sd Id 17043675 and a reconstructed energy of 6.6×10^{18} eV. It is the event with the largest value in $|X_{\mathrm{max}}^{\mathrm{Univ}} - X_{\mathrm{max}}^{\mathrm{FD}}|$. Manually excluding the problematic station solves the problem and results in a reconstruction of $X_{\mathrm{max}}^{\mathrm{Univ}} = (741\pm20)\,\mathrm{g\,cm^{-2}}$. This result is very close to the FD measurement of $X_{\mathrm{max}}^{\mathrm{FD}} = (745\pm10)\,\mathrm{g\,cm^{-2}}$. For a station to contribute to the shape fit, the time trace needs to have five bins with signal entries larger than 0.6 VEM/bin. A value of 0.2 VEM/bin was used for simulations, which would explain the presence of failed reconstructions there (accidental muons were not simulated). A sensitivity to fluctuations from thinning could be one possible explanation. Wrongly estimated start times are another source of reconstruction problems. For the data analysis, the threshold criterion to include stations in the shape fit represents a compromise: a lower threshold allows one to use more stations with low signals and thus more information in the fit while there is also a larger frequency of problematic traces that potentially disturb the fit. Due to the steeply falling LDF, stations with small signals are much more frequent, and thus, the probability to have accidental signals in these is much larger. The number of candidate stations increases with increasing zenith angle, and the number of stations with small signals is larger in saturated events. This explains the observed behavior in the population of failed reconstructions. A large signal threshold will exclude most of the problematic stations but will increase the primary energy starting at which events can be reconstructed with the universality reconstruction. At the same time, the resolution of reconstructed quantities will increase significantly. A dynamic rejection of stations and of certain ranges in the time traces would be a valuable addition to the universality reconstruction. It is one possibility to significantly reduce the amount of outliers and also decrease resolutions. This change of the reconstruction algorithm is outside the scope of this thesis.

The presented quality cuts are used for the remainder of this analysis. They allow a rejection of most of the failed reconstructions while they retain the distribution of quantities that we want to reconstruct. This is demonstrated with the distribution of X_{max}. The plots in Fig. 5.3 include profiles of the true X_{max} for different subsets of simulated and real events. Profiles of $X_{\mathrm{max}}^{\mathrm{MC}}$ as a function of energy are shown in Fig. 5.3a. In that plot, the simulated events are divided into non-saturated and saturated events. There is no significant difference between the profiles of these two sets of events. Due to the energy dependence of $X_{\mathrm{max}}^{\mathrm{MC}}$ and the fact that the fraction of saturated events increases with energy, it is better to compare values that are corrected for their intrinsic dependence on energy. This is done by plotting the quantity $X_{\mathrm{max}}^{\mathrm{MC}} - D_{10}\,(\lg(E/\mathrm{eV}) - 19.)$. D_{10} is called the elongation rate and describes the change in the average depth of shower maximum per decade in energy. The numerical value for QGSJET-II.03 simulations is $45.8\,\mathrm{g\,cm^{-2}}$. A plot of $X_{\mathrm{max}}^{\mathrm{MC}}$ corrected for the elongation rate and as a function of energy is given in Fig. 5.3b. The same information is plotted as a function of zenith angle in Fig. 5.3c. In both cases, there is no difference between the distributions of non-saturated and saturated events. The same holds when the subset of events that is selected with the quality cuts specified at the beginning of this section is compared to the rejected events. This is depicted in Fig. 5.3d and Fig. 5.3e. Also here, the profiles are in agreement.

The influence of the selection on the saturation status of events in data is addressed with Fig. 5.3f. Profiles of $X_{\mathrm{max}}^{\mathrm{FD}}$ for the two subsets of golden hybrid events are compared with each other as a function of reconstructed energy. The profiles are in good agreement. These checks verify that the derived quality cuts do not introduce biases in the result. In particular, the plots confirm that results obtained from the subset of non-saturated events are representative of the results obtained from all events.

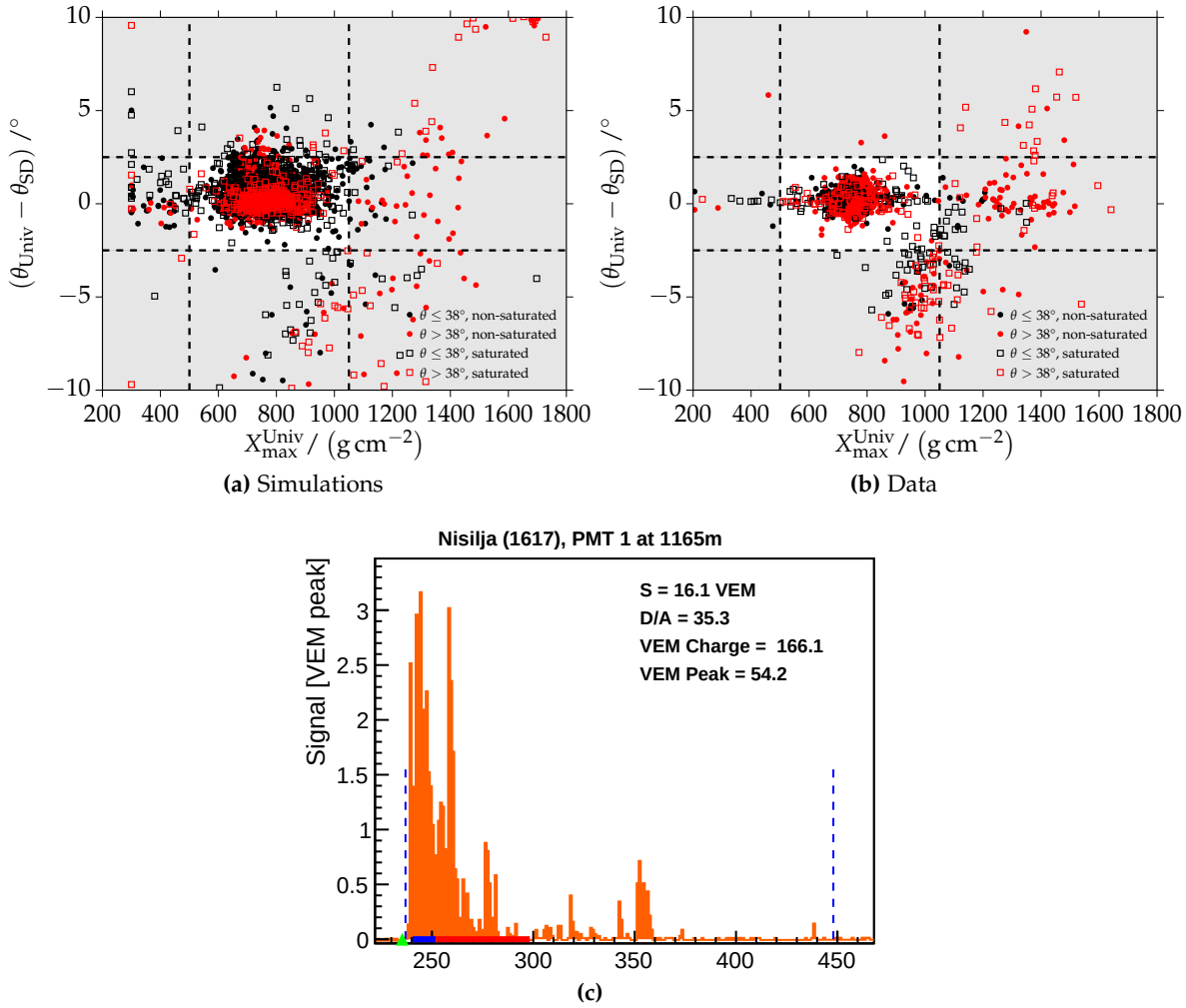


Figure 5.2: (a) Illustration of failed reconstructions in a plane spanned by $X^{\mathrm{Univ}}_{\mathrm{max}}$ and the difference in reconstructed zenith angles. The simulated events are split into two zenith angle ranges and their saturation status. (b) Analogous illustration of failed reconstructions in the set of golden hybrid events. (c) PMT time trace of a station that is part of an event with failed reconstruction (Sd Id: 17043675). Cause for the reconstruction problem is the accidental signal starting in the time bin 350.

5.2 Calibration of the universality reconstruction with data

Using golden hybrid events measured with the SD-1500, the universality reconstruction is calibrated to data. The procedure is to fix all quantities to the measurement from FD, especially the energy and reconstructed X_{max}. Then, using only the signal model and the LDF, the core position and R_μ are fit. In the second step, these quantities are also fixed and the event timing is fit together with Δm_μ. This calibration procedure allows one to parametrize the average $\langle R_\mu \rangle (E, \theta, X_{\mathrm{max}})$ and the time model offset of the muonic components Δm_μ. Results for the current set of golden hybrid events are shown in Fig. 5.4 and Fig. 5.5. The events were selected according to the quality cuts described in Section 4.7.2. Only non-saturated events are used in the parametrization. The dependence of the average R_μ as a function of energy and for different ranges in zenith angle is shown in Fig. 5.4a.

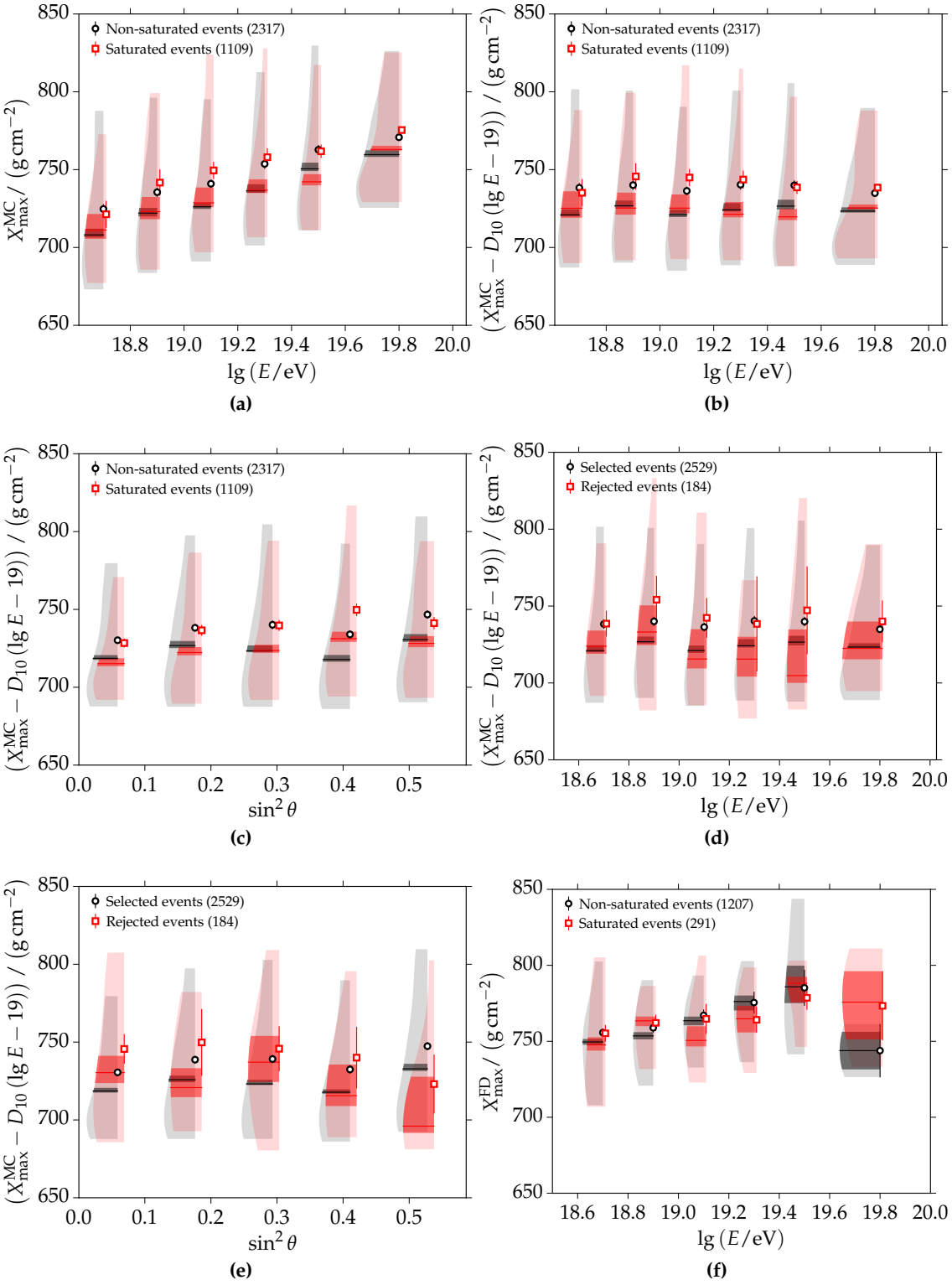


Figure 5.3: Investigating the influence of quality cuts and cuts on saturated events on the $X_{\max}$ distribution as a function of energy and zenith angle. The markers and violins of the profiles of saturated events were slightly shifted to the right for visibility, while their binning is identical to the profiles of non-saturated events.

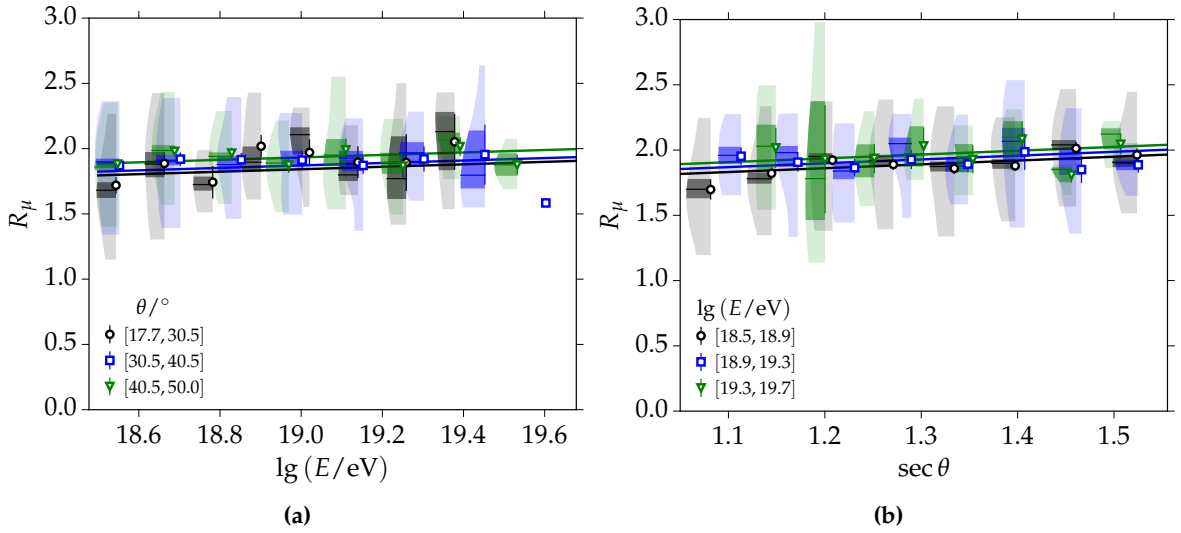


Figure 5.4: Parametrization of the average R_μ as a function of energy and zenith angle. Included in the analysis are golden hybrid events. The R_μ parametrization is derived from a reconstruction in which only R_μ and the core position are fit, the other parameters are fixed to results obtained from the FD reconstruction. In a second step, the time model offset is estimated as described in the text and depicted in Fig. 5.5.

The following model is fit to describe the data:

$$\begin{aligned}\langle R_\mu\rangle(E,\theta) = (2.13\pm0.08) + (0.09\pm0.06)\,(\lg(E/\mathrm{eV}) - 19)\\ + (0.29\pm0.11)\,(\sec\theta - 2).\end{aligned} \tag{5.1}$$

The time model offset Δm_μ is described with an identical functional relation but a different set of parameters:

$$\begin{aligned}\langle \Delta m_\mu\rangle(E,\theta) = (0.04\pm0.05) + (0.03\pm0.04)\,(\lg(E/\mathrm{eV}) - 19)\\ + (0.08\pm0.06)\,(\sec\theta - 2).\end{aligned} \tag{5.2}$$

Both models are obtained with unbinned chi-squared fits to data. The offset is depicted as a function of energy and for different ranges in zenith angle in Fig. 5.5a. There are small dependencies on both energy and zenith angle. It is interesting to note that the offset is slightly negative below 10^{19} eV and slightly positive for most of the energies above. This is in accordance with a change of the mass composition from light to heavy (as detailed in Section 3.4.5).

The correlation of the relative muon content R_μ with $X_{\max}$ is quantified with an analysis of the normalized quantity $R_\mu/\langle R_\mu\rangle(E,\theta)$ as a function of $X_{\max}$. A correlation is expected because proton showers penetrate deeper and have a smaller R_μ compared to iron showers. The correlation is shown in Fig. 5.5c for different ranges in energy. It is described with the following function:

$$\langle R_\mu\rangle(E,\theta,X_{\max}) = \langle R_\mu\rangle(E,\theta)\cdot\frac{1+\frac{0.5}{\pi}\arctan\left(\frac{a\lg(E/\mathrm{eV})-X_{\max}}{40\,\mathrm{g\,cm^{-2}}}\right)}{1+\frac{0.5}{\pi}\arctan\left(\frac{a\lg(E/\mathrm{eV})-\langle X_{\max}\rangle}{40\,\mathrm{g\,cm^{-2}}}\right)}. \tag{5.3}$$

The parameters a and $\langle X_{\max}\rangle$ are estimated with golden hybrid events to $a = (41.1\pm0.4)\,\mathrm{g\,cm^{-2}}$ and $\langle X_{\max}\rangle = (757\pm3)\,\mathrm{g\,cm^{-2}}$. The model is depicted with the red line in Fig. 5.5c. It is drawn as an interpolation of average model values for the measured events. The actual energy dependence is small.

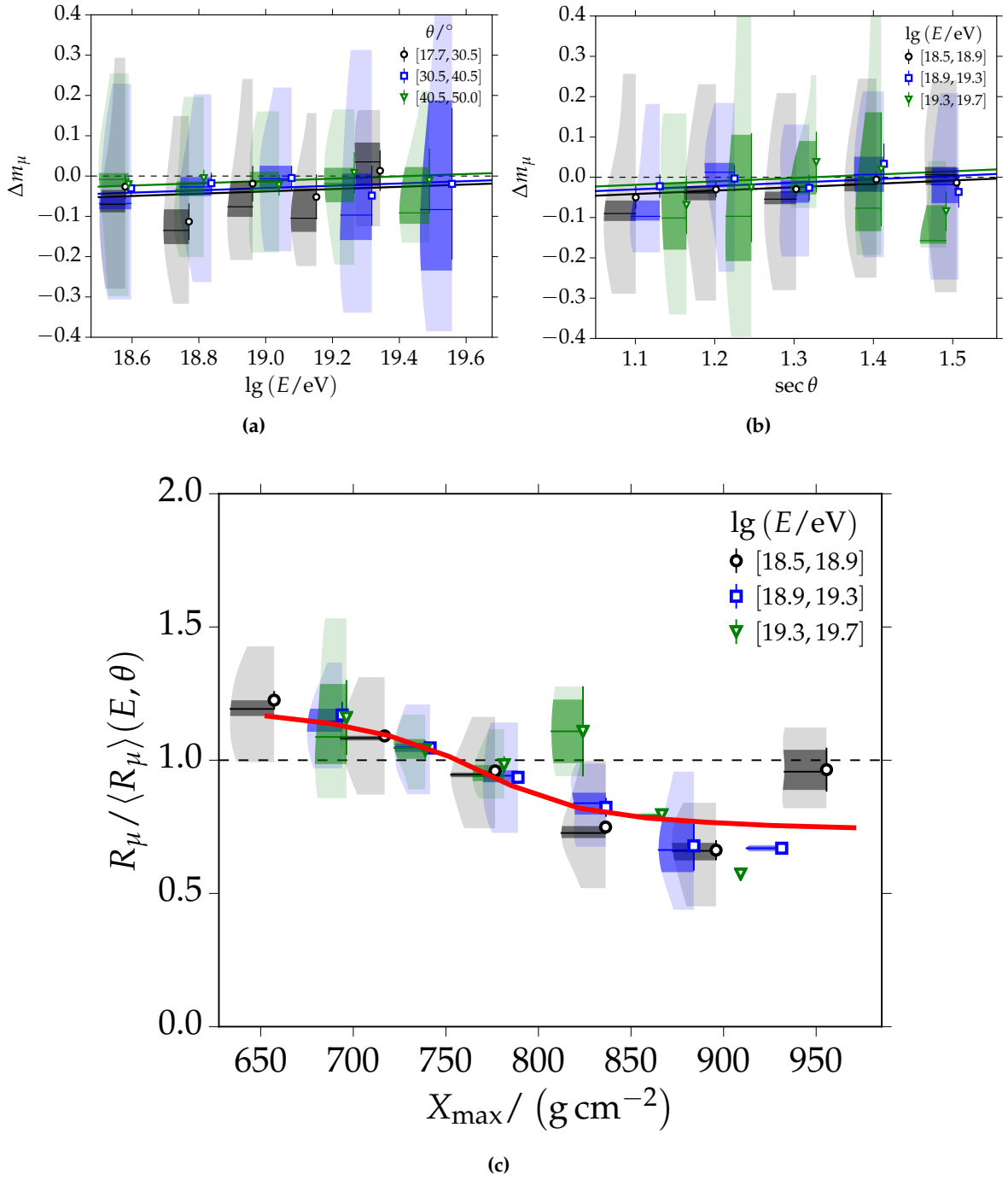


Figure 5.5: (a), (b) Parametrization of the time model offset Δm_μ as obtained from golden hybrid events. (c) Correlation between the relative muon content and the average depth of shower maximum: the normalized R_μ as a function of $X_{\max}$.

5.3 Systematic effects due to the aging of detectors

There are changes in the properties and the detector response of SD stations as they age. In particular, there is a shift in the signal response with time. The reasons for these changes are topic of ongoing studies; first analyses indicate that the reflectivity of the plastic bag that contains the water might change with time, especially due to freezing time periods.

This is discussed in depth in the context of the measurement of the energy spectrum and the CIC method in Section 4.6.

The purpose of the analysis discussed in this section is to quantify the influence of tank aging on the quantities reconstructed with the universality method and to correct for these systematic effects. Two variables that best correlate with the aging are used: the average age of stations within an event, counted in units of years from their date of deployment, and the average AoP, denoted as $\langle \mathrm{AoP} \rangle$. The second quantity describes the ratio of the area to the peak value in the muon calibration histogram (as briefly discussed in Section 1.2.1). The change of this quantity is representative of a change in the detector response to individual particles with time. With increased tank age, signals appear to get more compressed in time, i.e. a smaller AoP. A histogram of $\langle \mathrm{AoP} \rangle$ for golden hybrid events is plotted in Fig. 5.6a. While the mean value in data is close to 3.1, a fixed value of 3.55 was used in the simulation of the continuous library (Appendix E.3). The change of $\langle \mathrm{AoP} \rangle$ with the age of tanks is depicted in Fig. 5.6b. There is a nearly linear decrease with age, independent of the zenith angle or primary energy. An integrated histogram of the number of events as a function of the average tank age and $\langle \mathrm{AoP} \rangle$ is shown in Fig. 5.6c.

To quantify changes in reconstructed quantities with time, profiles of these quantities are first studied as a function of the average tank age for all golden hybrid events. The result for the relative energy difference $E_{\mathrm{Univ}}/E_{\mathrm{FD}} - 1$ is shown in Fig. 5.7a. There is a relatively linear change of 20 % over time. The drift of $X_{\mathrm{max}}^{\mathrm{Univ}} - X_{\mathrm{max}}^{\mathrm{FD}}$ with tank age is plotted in Fig. 5.7b. In that case, there is a change of $-30\,\mathrm{g\,cm^{-2}}$ over time. Note that $X_{\mathrm{max}}^{\mathrm{FD}}$ is stable with time and shows no dependence on the average tank age. This is expected because a change in the signal response of SD stations does not influence the reconstruction of $X_{\mathrm{max}}^{\mathrm{FD}}$. The change in the relative muon content $R_\mu - \langle R_\mu \rangle$ is depicted in Fig. 5.7c. The drift with tank age amounts to 0.1. It should be noted that the energy and X_{max} differences are only unbiased (on average) because of the calibration procedure described in the previous section. The calibration implicitly accounts for differences in the $\langle \mathrm{AoP} \rangle$ between data and simulations.

The three plots are repeated as a function of $\langle \mathrm{AoP} \rangle$ in Fig. 5.8. There are distinct linear dependencies in all three cases. The drifts are parametrized with linear models. The optimal parameters are determined with binned chi-squared fits to the golden hybrid events and stated within each of the plots. No significant dependence on zenith angle or other quantities was found. The models are used to correct SD-1500 and golden hybrid events for the following analyses.

5.4 Analysis method

The universality reconstruction detailed in Section 3.4 is used to obtain estimates of X_{max}, R_μ, the primary energy E and the shower geometry (incoming direction and impact point). The calibration of the method with data is discussed in Section 5.2, and corrections due to the aging of detectors are addressed in the previous section. The iterative method used for the reconstruction is explained in Section 3.4.6. In addition to the study of Monte Carlo events, the performance of this method is studied with golden hybrid events. These events are reconstructed identically to pure SD data; no information from the separate FD reconstruction is used. Afterwards, the reconstructed quantities of both reconstructions are compared with each other. These comparison plots are shown in Fig. 5.9 and Fig. 5.10. The bias in the estimated $X_{\mathrm{max}}^{\mathrm{Univ}}$ of $-5\,\mathrm{g\,cm^{-2}}$, as observed in Monte Carlo simulations (see Section 3.4.6), is already taken into account and corrected for. A scatter plot of E_{Univ} vs. E_{FD} is plotted in Fig. 5.9a. The red line represents the diagonal and not a fit. The relative bias of reconstructed energies is shown as a function of the universality energy in Fig. 5.9b.

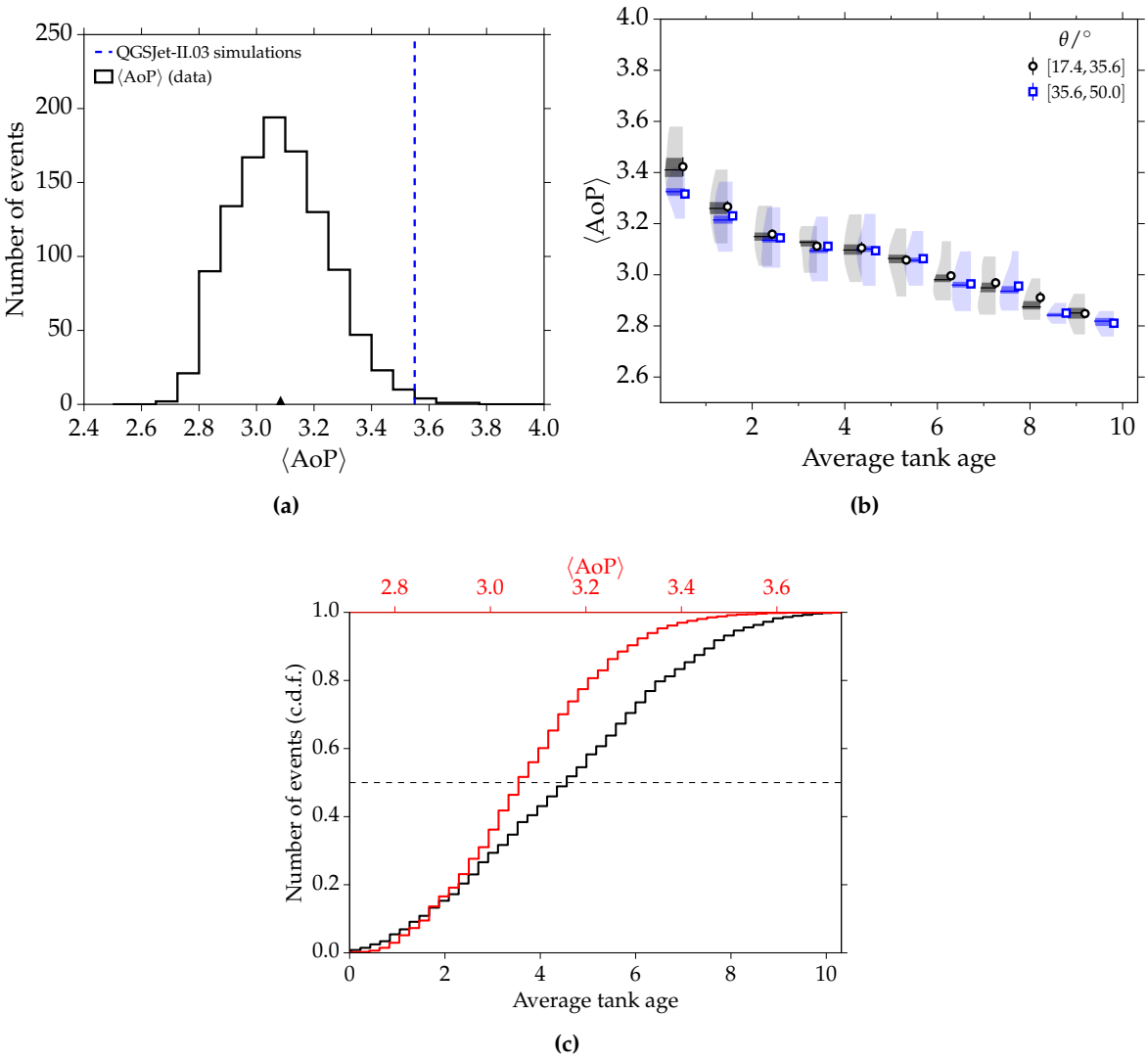


Figure 5.6: (a) Histogram of the average AoP in golden hybrid events. For comparison, the fixed value used in QGSJET-II.03 simulations for this work is shown with a vertical line. (b) The average AoP as a function of the average tank age. The data are divided into two zenith angle ranges. (c) Integrated histogram (c.d.f.) of the number of events as a function of average tank age and average AoP.

In addition, data are divided into different ranges of zenith angle in Fig. 5.9c. Overall, the universality reconstruction yields 4 % larger energies compared to the FD reconstruction. This bias is mostly independent of energy and zenith angle. A comparison of reconstructed zenith angles is shown as a function of the number of candidate stations and for different primary energies in Fig. 5.9d. For a small number of candidate stations, the universality reconstruction yields slightly larger zenith angles. The difference is less than 0.5° and thus within about 1 σ of the statistical uncertainty. The angular reconstruction is unbiased above a number of candidate stations of seven.

The difference in the reconstructed X_{max}: $X_{\mathrm{max}}^{\mathrm{Univ}} - X_{\mathrm{max}}^{\mathrm{FD}}$ is shown as a function of zenith angle in Fig. 5.10a. The drift of values due to the aging of tanks is already accounted for. There is a small increase of the difference with zenith angle. The change of the mean values

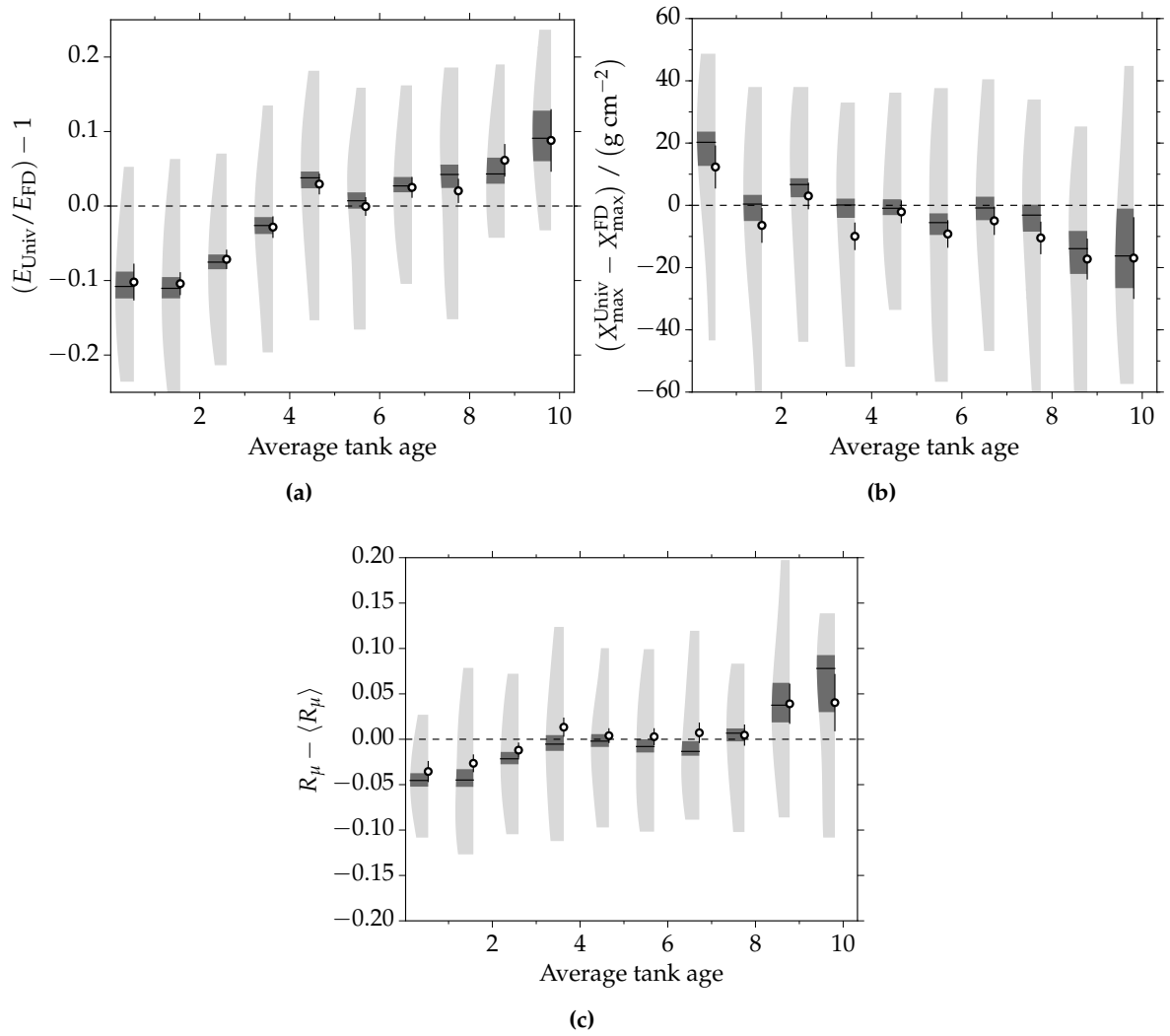


Figure 5.7: The change of quantities with the average tank age: (a) $E_{\mathrm{Univ}}/E_{\mathrm{FD}} - 1$ (b) $X_{\mathrm{max}}^{\mathrm{Univ}} - X_{\mathrm{max}}^{\mathrm{FD}}$ (c) $R_\mu - \langle R_\mu \rangle$.

over the whole angular range is 18 g cm^{-2}. It is parametrized with the following linear model:

$$\Delta X_{\mathrm{max}} (\sec\theta) = (0.3 \pm 1.6) + (35.6 \pm 13.6)\,(\sec\theta - 1.3)\ . \tag{5.4}$$

The best parameters are obtained from a binned chi-squared fit to a mean profile of all non-saturated events. As visible in Fig. 5.10a, saturated events are on average unbiased and do not show a similar trend, but probably due to the very small event statistics. In the following, all events are corrected for the small angular dependence according to the model. A profile of the X_{max} differences as a function of reconstructed energy is shown in Fig. 5.10b. Shown are non-saturated events in black and saturated events in red. Both profiles exhibit no significant deviations from zero. There is an indication of an overestimation of $X_{\mathrm{max}}^{\mathrm{Univ}}$ for saturated events at high energies. This is not conclusive and needs to be studied in the future when greater event statistics are accumulated. After correcting for the residual zenith angle dependence, the profiles are flat as a function of $\sec\theta$ in Fig. 5.10c. There is an indication of an overestimation of $X_{\mathrm{max}}^{\mathrm{Univ}}$ for saturated events at low zenith angles, which is again not conclusive. Detailed residual plots with different ranges of energy or zenith angle are included in Fig. 5.10d and Fig. 5.10e. Overall, all profiles are very close to zero,

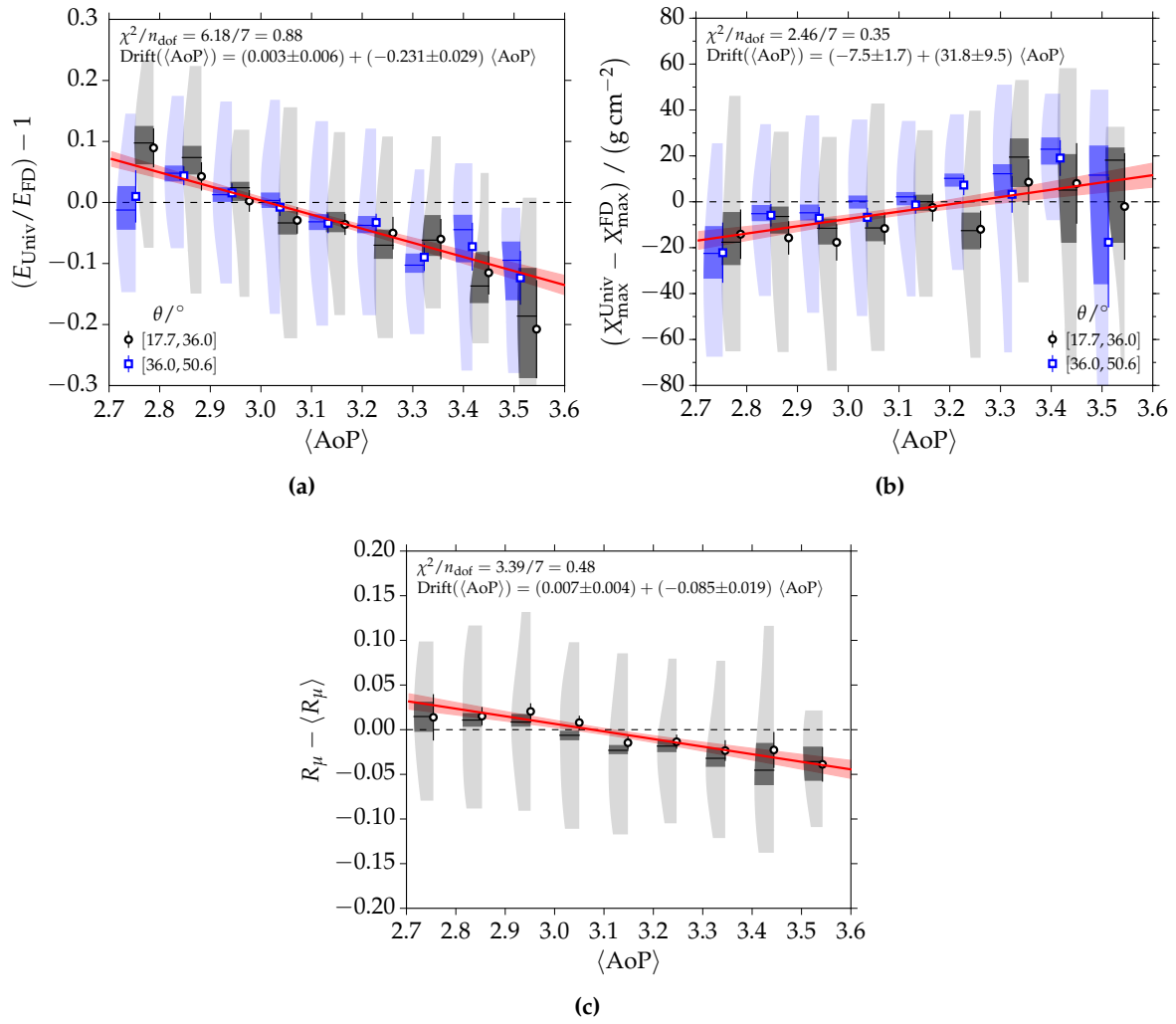


Figure 5.8: The change of quantities with the average AoP: (a) $E_{\mathrm{Univ}}/E_{\mathrm{FD}} - 1$ (b) $X^{\mathrm{Univ}}_{\mathrm{max}} - X^{\mathrm{FD}}_{\mathrm{max}}$ (c) $R_\mu - \langle R_\mu \rangle$. Functions to model the dependencies are indicated with red lines. For completeness, the linear equations resulting from fits are shown in the plots.

confirming the correctness of the universality reconstruction, the calibration procedure and the applied corrections for systematic drifts. Similar plots for the classic universality reconstruction described in Section 3.4.5 are included in Appendix B.4.1.

Histograms of the obtained $X^{\mathrm{Univ}}_{\mathrm{max}}$ values for six different energy ranges are depicted in Fig. 5.11. They include reconstructions of the complete set of SD-1500 data. The binning is fixed to 22 equidistant bins in the X_{max} range 500 g cm^{-2} to 1050 g cm^{-2}. The bin size is thus fixed to 25 g cm^{-2}. These histograms include all events that contribute to the analysis of the first two central moments of the $X^{\mathrm{Univ}}_{\mathrm{max}}$ distributions as detailed in the next sections. Black points and histograms correspond to non-saturated events, while the red markers and dashed lines represent saturated events. It is apparent that the number of saturated events is mostly relevant at the highest energies when about half of the events contain a low-gain saturated station. In all energy ranges, the resolution of saturated events is substantially larger. Reasons for that are discussed in Section 3.4.5. It is apparent that the X_{max} distributions are not symmetric; they have a larger tail towards larger X_{max} values. Results on the average shower maximum are presented in the following section. The average is taken as the mean value of all X_{max} entries within a given energy bin. Calculating median

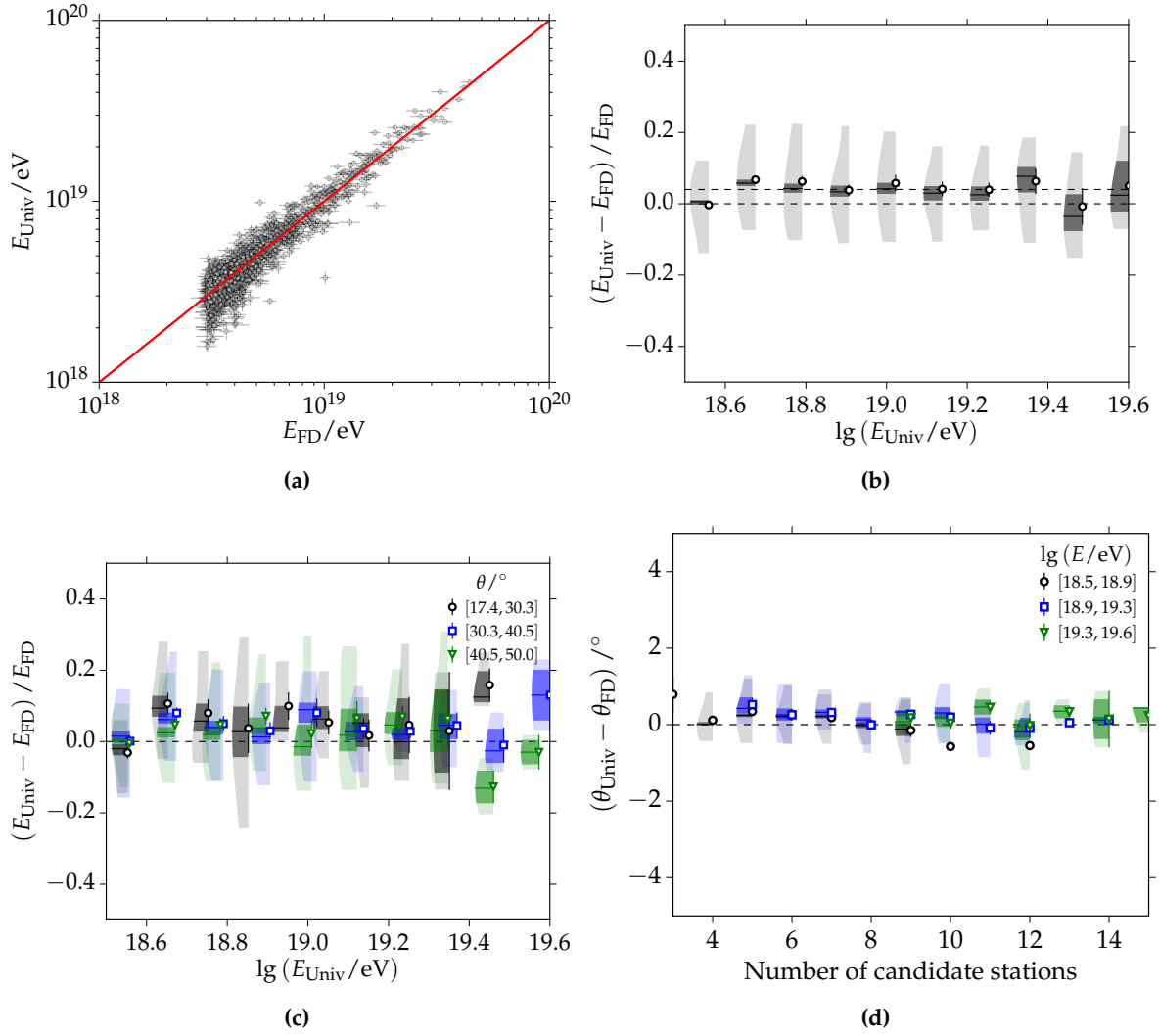


Figure 5.9: A comparison of reconstructed energies and zenith angles using golden hybrid events. The iterative universality fit was used to reconstruct data independently of the FD reconstruction.

values instead of mean values results in differences of up to $5\,\mathrm{g\,cm^{-2}}$ (the median values are always smaller). The difference is independent of energy. The use of mean values is motivated by the fact that also the FD and air shower model results are obtained with mean values, and a comparison between these results is of high interest. The distributions of saturated events show significantly more outliers at large $X_{\mathrm{max}}^{\mathrm{Univ}}$ values. This is due to problems discussed in Section 3.4.5 and Section 3.4.6. Results of these events are therefore shown in separate figures.

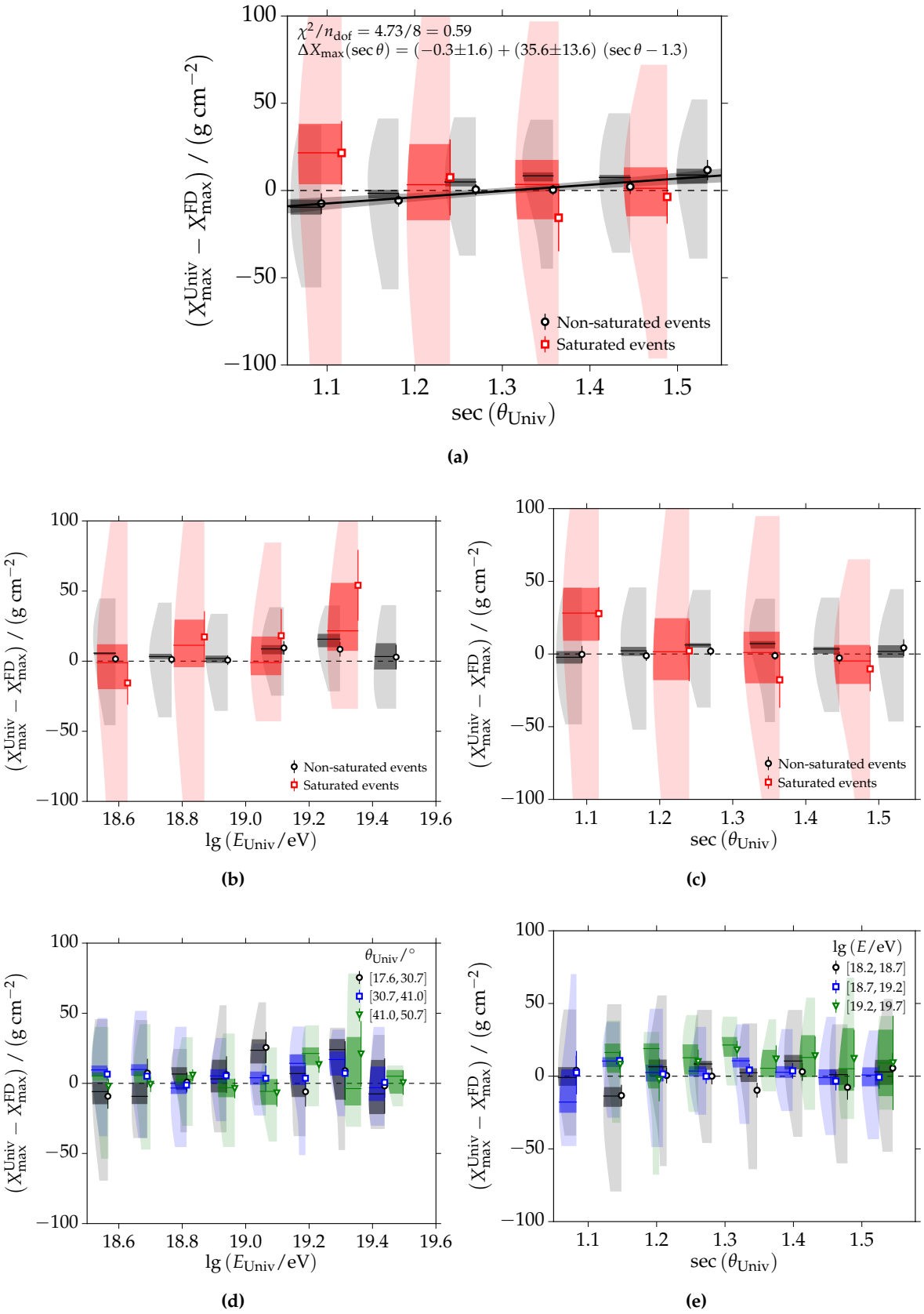


Figure 5.10: Comparison of reconstructed depth of shower maxima using golden hybrid events. The iterative universality method was used to reconstruct data independent of the FD reconstruction.

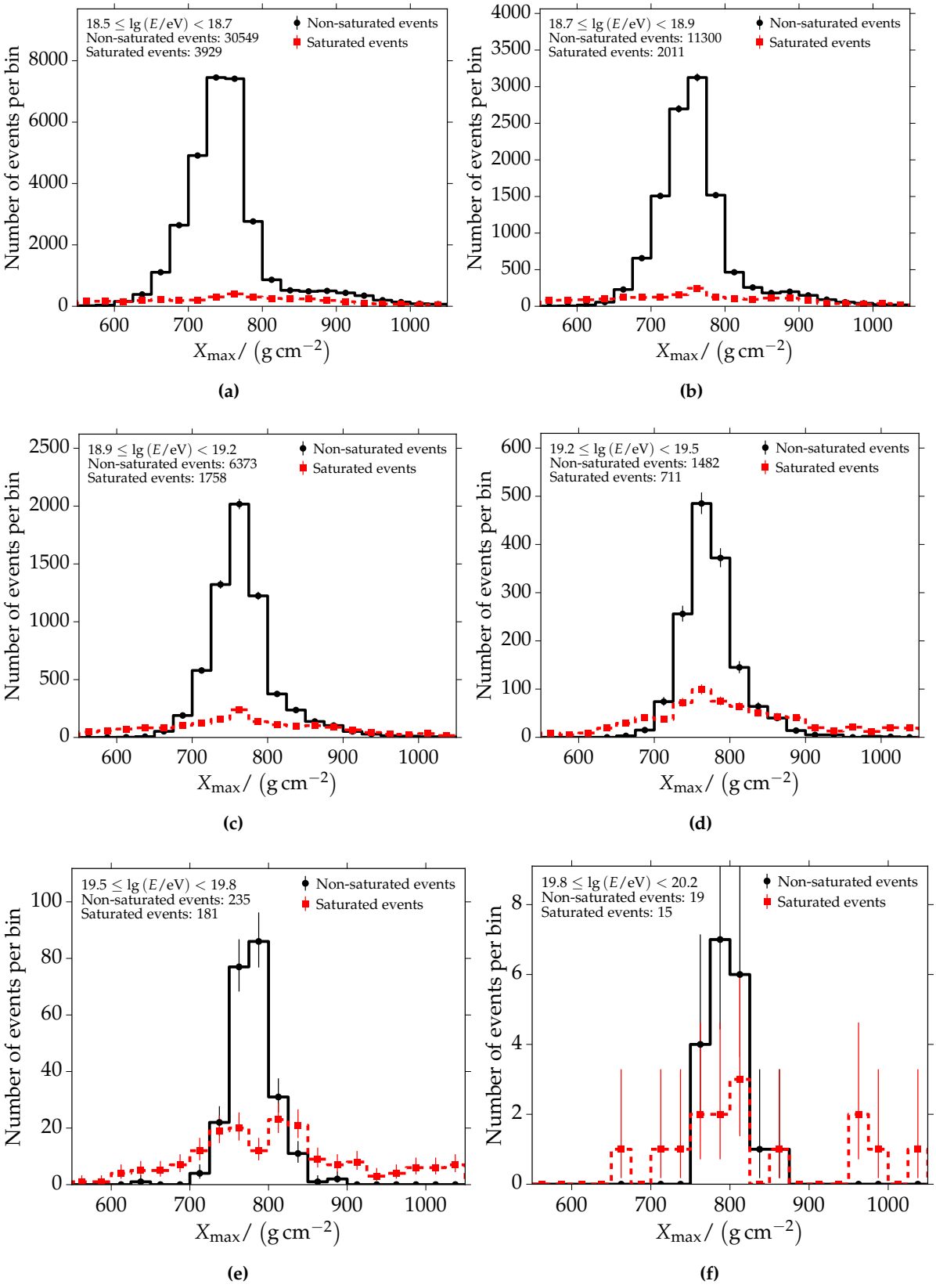


Figure 5.11: Histograms of the reconstructed $X_{\mathrm{max}}^{\mathrm{Univ}}$ using SD-1500 data only. The histograms are split into different ranges of reconstructed energy. The distributions for saturated and non-saturated events are plotted separately as stated in the plots.

5.5 Results on the average depth of shower maximum

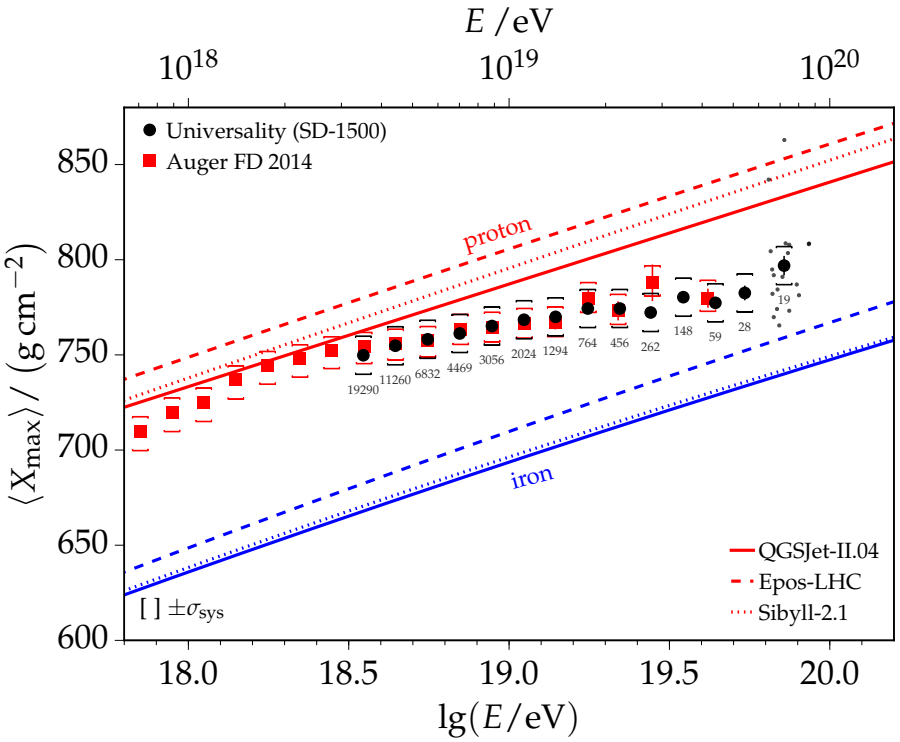
The first central moment of the reconstructed shower maximum ($\langle X_{\mathrm{max}}^{\mathrm{Univ}} \rangle$) is discussed in this section. The complete set of SD-1500 data in the time period 2004 to 2014 is studied; the selection criteria and corrections are applied as discussed in the previous sections. Profiles of the average shower maximum reconstructed with universality are depicted in Fig. 5.12. Only non-saturated events are included in Fig. 5.12a, and all events are shown in Fig. 5.12b. Error bars indicate statistical uncertainties, while the systematic uncertainties are drawn with caps. The statistical uncertainties are obtained from bootstrapping the distribution of measured events and repeatedly calculating the mean. The uncertainty is then calculated from the width of the resulting distribution of mean values. The position of the mean markers on the $\lg E$ axis represents to mean logarithmic energy of all entries within a bin. The total systematic uncertainty of $10\,\mathrm{g\,cm^{-2}}$ comprises:

- $6\,\mathrm{g\,cm^{-2}}$ due to the residual composition dependence of the time model (shown in Fig. 3.13e),
- $5\,\mathrm{g\,cm^{-2}}$ due to tank aging effects and the time calibration (Section 5.3 and Section 5.2),
- $5\,\mathrm{g\,cm^{-2}}$ due to the small influence of quality cuts and the residual zenith angle dependence (see Fig. 5.10a).

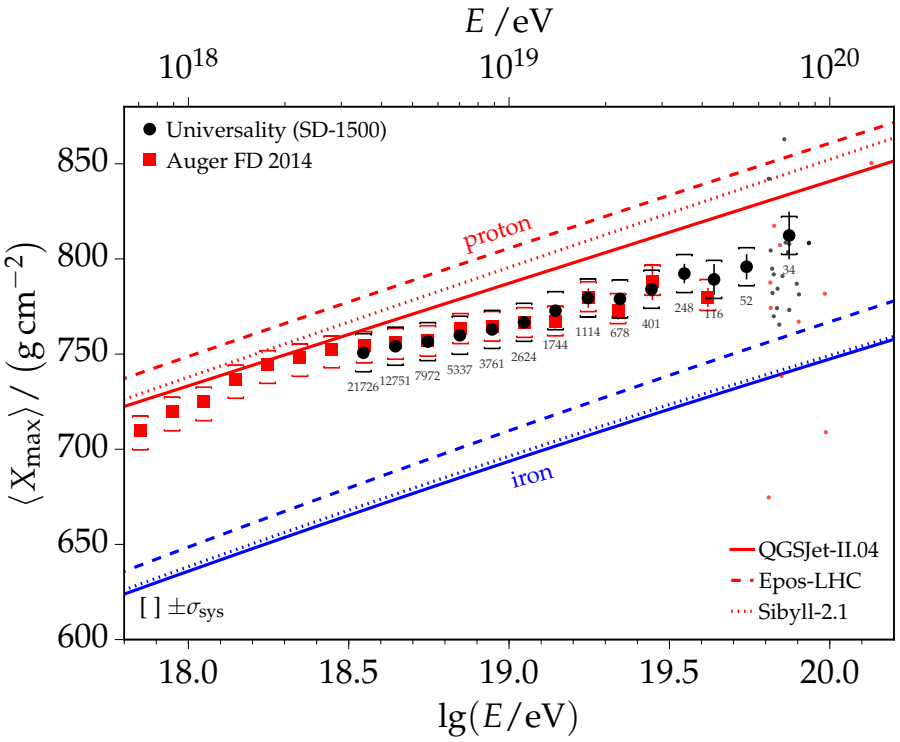
The data are compared to predictions from the contemporary interaction models QGSJet-II.04, Epos-LHC and Sibyll-2.1. The models QGSJet-II.04 and Epos-LHC were tuned with data measured at the LHC[1]. The SD-1500 measurement is compared to the current FD measurement, which is shown with red rectangles [80]. The measurements are compatible within the given uncertainties. Both FD and universality profiles indicate a trend towards heavier composition at the highest energies. However, the last bin in the universality measurement indicates a recovery of light elements, which is not statistically significant at the current level of event statistics. The individual events contributing to the last energy bin are depicted with dots. In Fig. 5.12b, black dots represent individual values from non-saturated events, while red dots indicate saturated events. The total number of events within each bin is printed below the bins. The corresponding violin plots are shown in Fig. 5.13. These include the same data as the previous plots but give additional information. The markers represent the mean values while the small black lines indicate median values of all data points within an energy bin. The shaded regions around mean and median indicate the $1\,\sigma$ spread of the data points. In particular, it is visible that the median profiles for non-saturated and saturated events are in very good agreement, while the mean values for saturated events are pulled to larger values due to the larger fraction of reconstruction outliers that are not completely rejected with the given quality cuts.

Two further important checks are depicted in Fig. 5.14. A comparison to the result without time model calibration is shown in Fig. 5.14a. It is visible that the non-calibrated result deviates more from the FD result, which is used to calibrate the universality reconstruction in Section 5.2. However, the general structure of the profile, especially at the highest energies, remains unchanged. This is an important validation that the observed trend towards lighter elements at the highest energies is not an artifact of the calibration procedure. Another comparison is included in Fig. 5.14b. Here, the blue rectangular profile represents the result obtained with the classic universality reconstruction. In this reconstruction, the energy is fixed to the result from the standard reconstruction, and all other quantities are fit at the same time. The calibration of the average muon content from golden hybrid events

[1] Large Hadron Collider

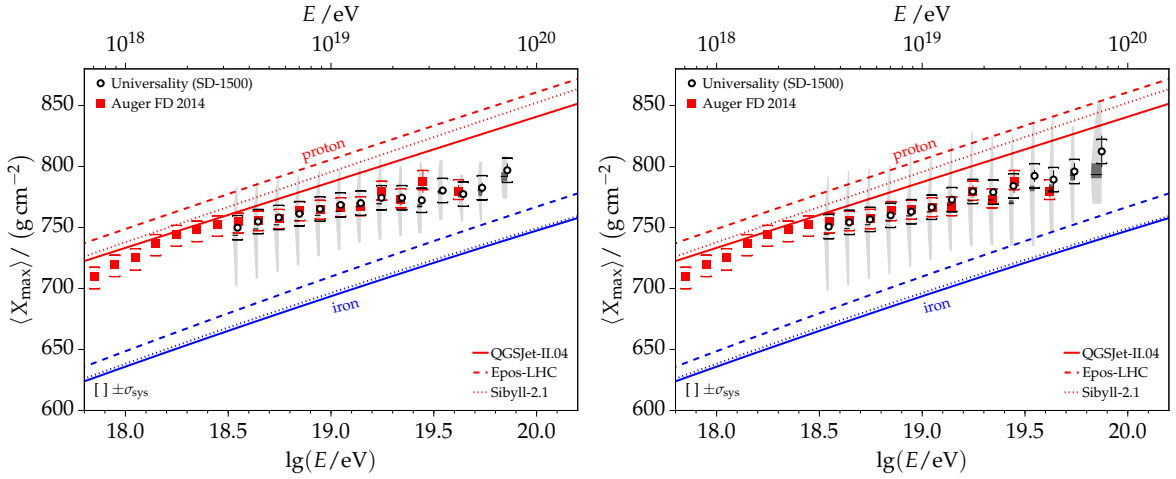


(a) Non-saturated events



(b) All events

Figure 5.12: Profiles of the average reconstructed shower maximum $X_{\mathrm{max}}^{\mathrm{Univ}}$ with statistical and systematic uncertainties. A recent result derived from measurements with the FD of Auger is shown with red rectangular markers. The different lines indicate predictions from simulations using current hadronic interaction models. Events included in (a) non-saturated events and (b) all events.



(a) Non-saturated events **(b)** All events

Figure 5.13: Violin profiles of the average depth of shower maximum from SD-1500 data and reconstructed with shower universality compared to the result from measurements with the FD. The data are identical to Fig. 5.12.

is not used, and no additional constraints enter the fit. Using this reconstruction procedure leads to larger biases below energies of 10^{19} eV and a number of candidate stations of eight (see Section 3.4.5). Therefore, these two cuts were applied. The resulting profile is close to the result using the iterative reconstruction method. In particular, the trend at the highest energies is identical. A fit to the $\langle X_{\mathrm{max}}^{\mathrm{Univ}} \rangle$ mean profile with a linear model was attempted but results in a bad χ^2/n_{dof} of around 3. A much better description with $\chi^2/n_{\mathrm{dof}} = 1.1$ is obtained with a fit to a broken line as depicted in Fig. 5.15. The break point in energy is fit to $\lg(E_0/\mathrm{eV}) = 19.05 \pm 0.09$. It is indicated with the magenta hexagon. Below E_0, the elongation rate $D_{10} = \frac{\mathrm{d}\langle X_{\mathrm{max}} \rangle}{\mathrm{d}\lg E}$ is:

$$D_{10} = (33.8 \pm 2.8)\,\mathrm{g\,cm^{-2}}\,, \tag{5.5}$$

and above E_0:

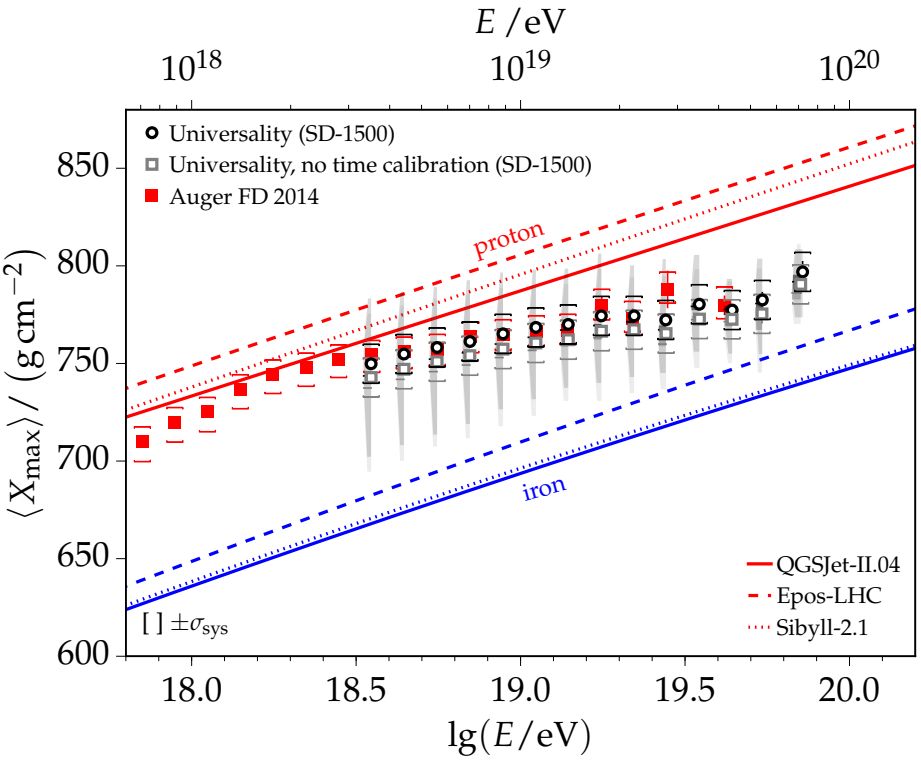
$$D_{10} = (20.8 \pm 3.3)\,\mathrm{g\,cm^{-2}}\,. \tag{5.6}$$

To better visualize the break in the plot, the model prior to the break is continued as a dashed line. The universality result should be compared to the current result derived with FD data:

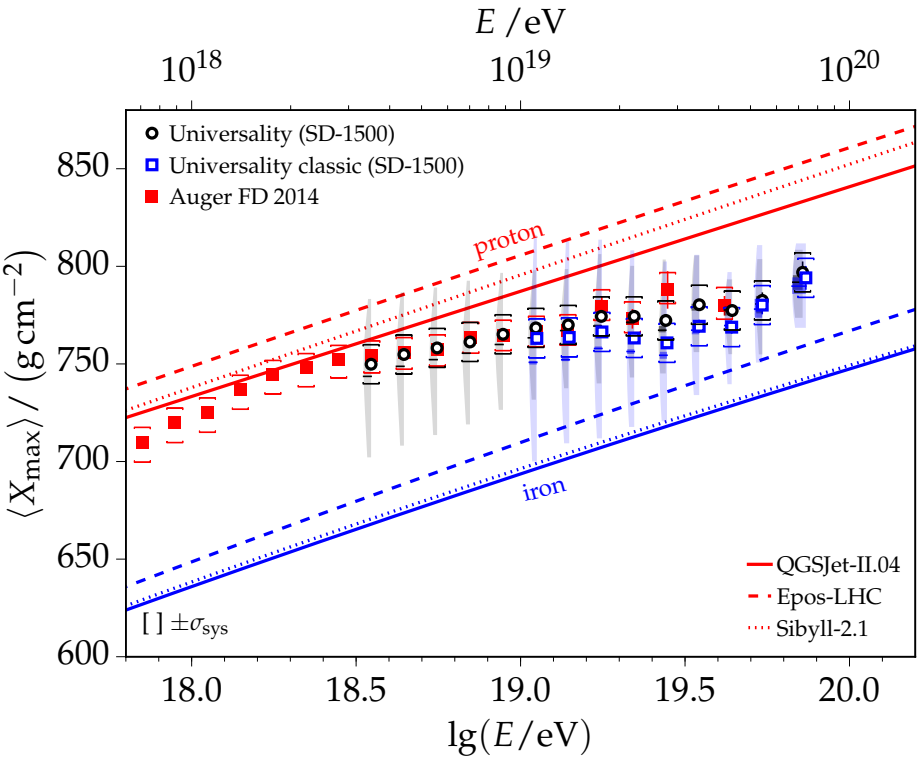
$$D_{10} = 26.4 \pm 2.5(\mathrm{stat})^{+7.0}_{-1.9}(\mathrm{sys})\,\mathrm{g\,cm^{-2}}\,, \tag{5.7}$$

as discussed in [21]. Current air shower simulations predict elongation rates for a constant composition in the range from $54\,\mathrm{g\,cm^{-2}}$ to $64\,\mathrm{g\,cm^{-2}}$. The results of this work indicate that the composition of UHECRs gets heavier in the whole energy range above 3×10^{18} eV. Additionally, there is an increase in the elongation rate above $10^{19.05}$ eV, reflecting a faster change towards a heavier composition. At the highest energies, data hints at the evolution towards slightly lighter elements. The distance of the last data point to the model in units of statistical uncertainty is $1.98\,\sigma$. Also taking the systematic uncertainty into account reduces this number to exactly $1\,\sigma$. That means that, while there is an indication towards lighter elements at the highest energies, this result is not statistically significant with the current event statistics.

A table of the results of the first two central moments of the universality X_{max} together with statistical and systematic uncertainties is included in Appendix C.2.



(a)



(b)

Figure 5.14: (a) Comparison of the result on the average depth of shower maximum to a result obtained without the calibration of the universality time model to golden hybrid data. (b) Comparison to the result obtained with the classic universality reconstruction (only the energy is fixed in the reconstruction).

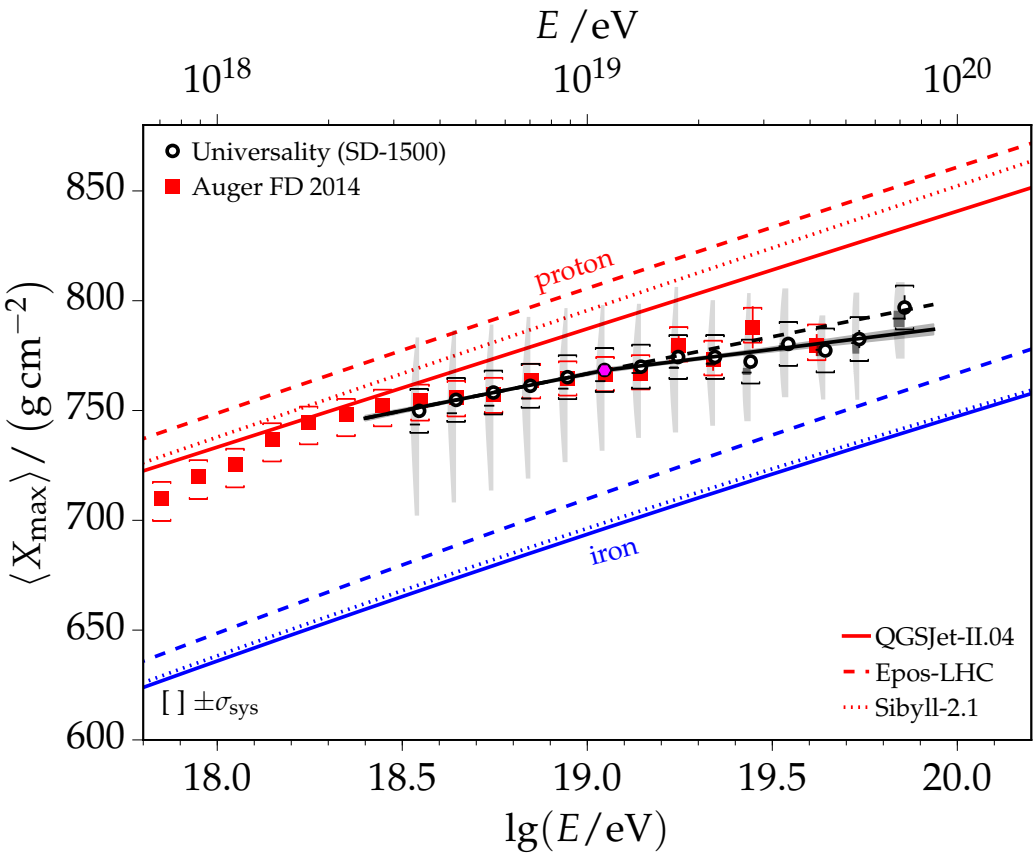


Figure 5.15: Fit of a broken line to the resulting $X_{\mathrm{max}}^{\mathrm{Univ}}$ profile for non-saturated events. Included are events measured with the SD-1500.

5.6 Fluctuations of the depth of shower maximum

The study of the second central moment of the $X_{\mathrm{max}}^{\mathrm{Univ}}$ distribution enables conclusions about the physical fluctuations of the depth of shower maximum. Identically to the study of $\langle X_{\mathrm{max}} \rangle$, the measured fluctuations are compared to predictions for different primary masses using current air shower simulations to draw conclusions about the mass composition. The total resolution is given as the quadratic sum of the physical fluctuations from the shower development and the detector resolution due to sampling fluctuations:

$$\sigma^2_{\mathrm{tot}}(E) = \sigma^2_{\mathrm{phys}}(E) + \sigma^2_{\mathrm{det}}(E)\,. \tag{5.8}$$

Thus, the detector resolution needs to be quantified as detailed in the following section. Due to the significantly increased resolution and systematic uncertainties, saturated events are omitted for this study.

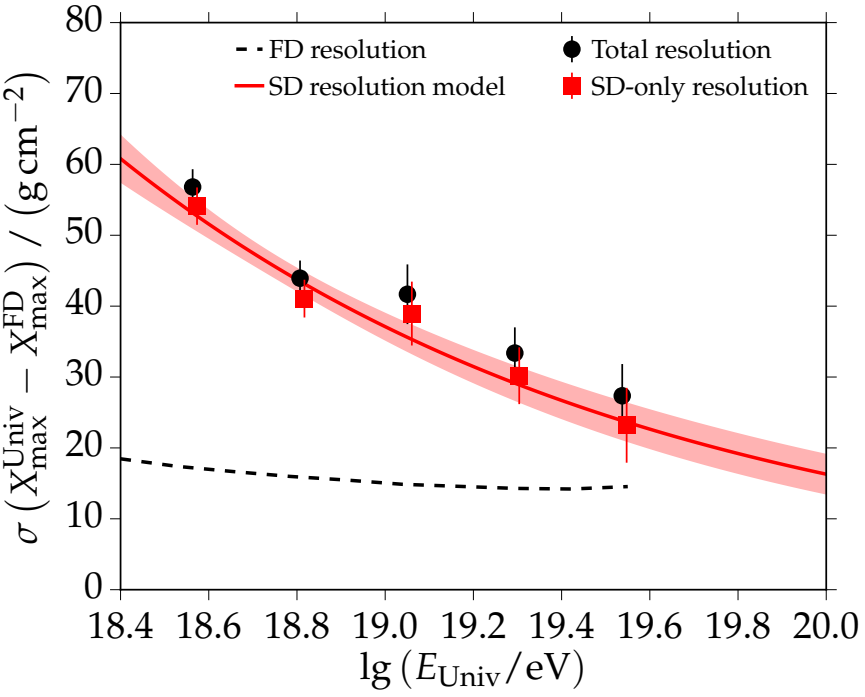
5.6.1 Model of the detector resolution

Sampling fluctuations in the measurement of signals with the SD drive the detector resolution of the universality estimates (can be compared to the signal uncertainty model discussed in Section 3.2.2). The resolution is studied with the set of golden hybrid events discussed in Section 5.4. Results are shown in Fig. 5.16. The resolution is extracted from the energy-dependent width of the distribution of $X_{\mathrm{max}}^{\mathrm{Univ}} - X_{\mathrm{max}}^{\mathrm{FD}}$. The normal standard deviation is used for those calculations. The FD resolution is quadratically subtracted for that comparison. It is detailed in [80] and is indicated with the black dashed line. In Fig. 5.16a, the red rectangular points reflect the resulting resolution of $X_{\mathrm{max}}^{\mathrm{Univ}}$ using the iterative reconstruction method described in Section 3.4.6. The resolution drops from $60\,\mathrm{g\,cm^{-2}}$ at $3{\times}10^{18}\,\mathrm{eV}$ to $40\,\mathrm{g\,cm^{-2}}$ at $10^{19}\,\mathrm{eV}$ and to less than $20\,\mathrm{g\,cm^{-2}}$ at the highest energies. The behavior is parameterized with this exponential function:

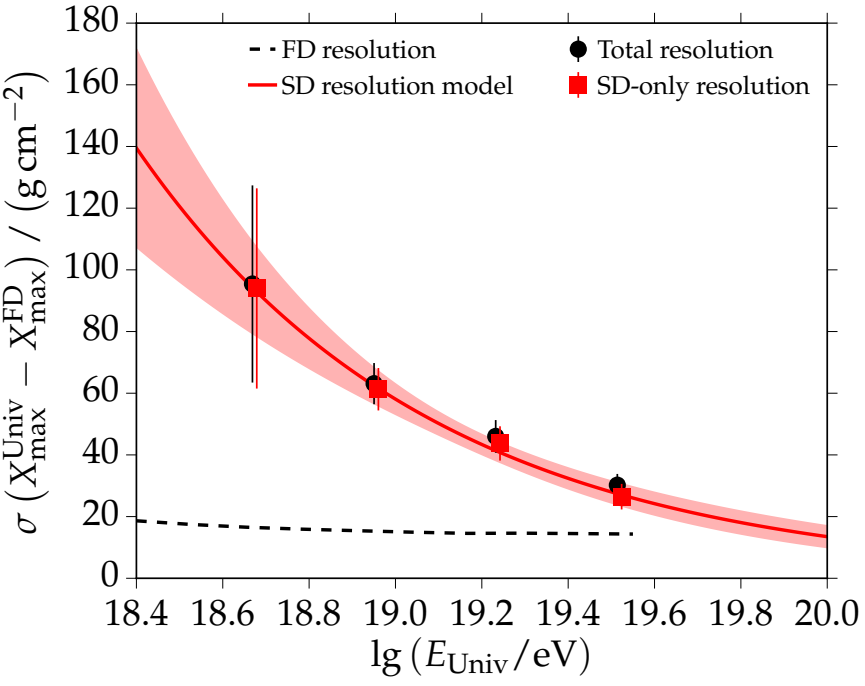
$$\sigma[X_{\mathrm{max}}^{\mathrm{Univ}}](\lg E) = A\,\exp\left(B\,(\lg(E/\mathrm{eV}) - 19)\right). \tag{5.9}$$

The parameters resulting from a binned chi-squared fit with $\chi^2/n_{\mathrm{dof}} = 0.8$ to the golden hybrid data are $A = (37.1{\pm}1.8)\,\mathrm{g\,cm^{-2}}$ and $B = -0.82{\pm}0.13$. The model is very similar to the resolution obtained in simulations (see Fig. 3.20). However, the resolution depends on the mass composition of UHECRs. Using golden hybrid events to derive a model yields an unbiased estimate (in particular due to the quality selection described in Section 5.1). Due to low event statistics, the model needs to be extrapolated to the highest energies. Furthermore, the resolutions extracted from simulations slightly underestimate the truth because true values instead of model values are used in the current iterative reconstruction method. The detector resolution of $X_{\mathrm{max}}^{\mathrm{Univ}}$ obtained with the classic universality reconstruction (see Section 3.4.5) is studied separately. The result is shown in Fig. 5.16b. While the resolution at the highest energies is similar to the one achieved with the iterative reconstruction method, it is roughly doubled at energies below $10^{19}\,\mathrm{eV}$. The significantly larger resolution reflects the difference in the reconstruction types. There are no constraints or iterations in the classic reconstruction, and all quantities are estimated simultaneously. This results in a much larger spread of reconstructed values. Fitting the resolution with Eq. (5.9) results in the parameters $A = (58.1{\pm}4.9)\,\mathrm{g\,cm^{-2}}$ and $B = -1.46{\pm}0.30$. The model and its $1\,\sigma$ uncertainty regime are visualized as a red line and area in Fig. 5.16b.

Next, the second central moment of $X_{\mathrm{max}}^{\mathrm{Univ}}$ as a function of reconstructed energy and obtained with SD-1500 data is corrected for the spread due to the detector resolution using the derived models. The result is depicted in Fig. 5.17. A comparison to current FD data [80] and predictions from simulations is shown. Profiles of the physical fluctuations obtained



(a) Iterative universality reconstruction



(b) Classic universality reconstruction

Figure 5.16: Detector resolution of the measurement of $X_{\mathrm{max}}^{\mathrm{Univ}}$ obtained from reconstructions of golden hybrid events measured with the SD-1500 and FD. (a) iterative reconstruction; (b) classic reconstruction (only the energy is fixed). For visual purposes, different numbers of bins are chosen in the two plots.

with both the classic and iterative reconstructions are depicted in Fig. 5.17a. Due to the requirement that the number of candidate stations needs to exceed seven in case of the classic reconstruction, the result obtained with this reconstruction type is shown only above 10^{19} eV. The result from the classic reconstruction is shown with blue rectangular points. The systematic uncertainties represented with upper and lower brackets stem mostly from the uncertainty in the detector resolution models. Above $10^{19.2}$ eV, it matches very well with the result obtained from FD measurements. Considering the predictions from air shower simulations, there is a distinct evolution towards heavier elements. The last point turns back towards lighter elements. Overall, the profile exhibits the same trends as seen in the average depth of shower maximum in Section 5.5. The result obtained with the iterative universality reconstruction is depicted with black circular markers. It is apparent that the resulting fluctuations are smaller than the ones achieved with FD or the classic reconstruction for all energy bins. At the highest energies, the fluctuations are even in the unphysical range when compared to iron predictions from the current hadronic interaction models. To better understand this, the corrected fluctuations are compared to the uncorrected ones in Fig. 5.17b. Before subtracting the detector resolution, the fluctuations achieved with the iterative method follow the FD measurement. At the same time, the detector resolution obtained with golden hybrid events is similar to the one seen in simulations. These results imply that the intrinsic width of the $X_{\mathrm{max}}^{\mathrm{Univ}}$ distributions is narrowed through the use of constraints and different reconstruction steps in the iterative reconstruction. While X_{max} is not explicitly constrained during the reconstruction, the constraints on the shower geometry and energy act as implicit constraints. The merit of that method is an unbiased result of the mean in the whole energy range. That was the goal behind developing the iterative method. This is usually prevented by the large number of parameters in the universality reconstruction and the low station multiplicity at lower energies. The underestimation of the physical fluctuations can possibly be resolved with less stringent constraints on the shower geometry, for example a constraint on the reconstructed zenith angle of more than 2°.

5.7 Estimation of the logarithmic mass

The superposition model of air showers allows one to formulate linear dependencies between the average depth of shower maximum and the logarithmic mass and energy. These calculations are based on the generalized Heitler model of air showers as discussed in Section 1.1.4. An extended model reads:

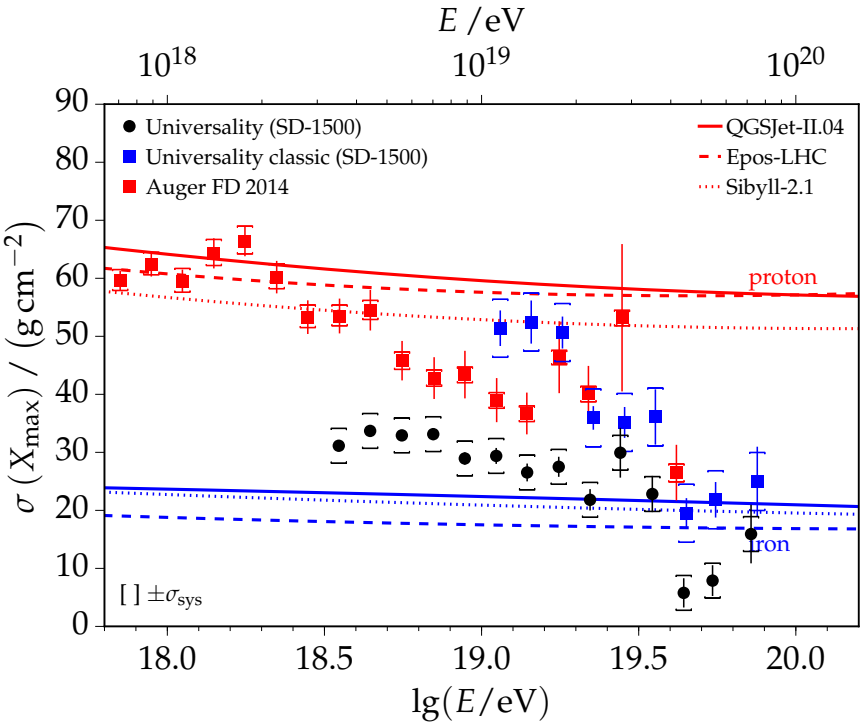
$$\langle X_{\mathrm{max}} \rangle = X_0 + D \lg\left(\frac{E}{E_0 A}\right) + \xi \ln A + \delta \ln A \lg\left(\frac{E}{E_0}\right) . \tag{5.10}$$

This model is generalized to include additional energy and mass dependencies necessary to describe simulations with contemporary hadronic interaction models. More details are given in [28, 186]. X_0 is the mean depth of proton showers at an energy E_0 and D is the elongation rate. The parameters ξ and δ depend on the hadronic interaction model. All parameters are stated in [28, 186]. Eq. (5.10) can be inverted in order to estimate the mean logarithmic mass from the average depth of shower maximum:

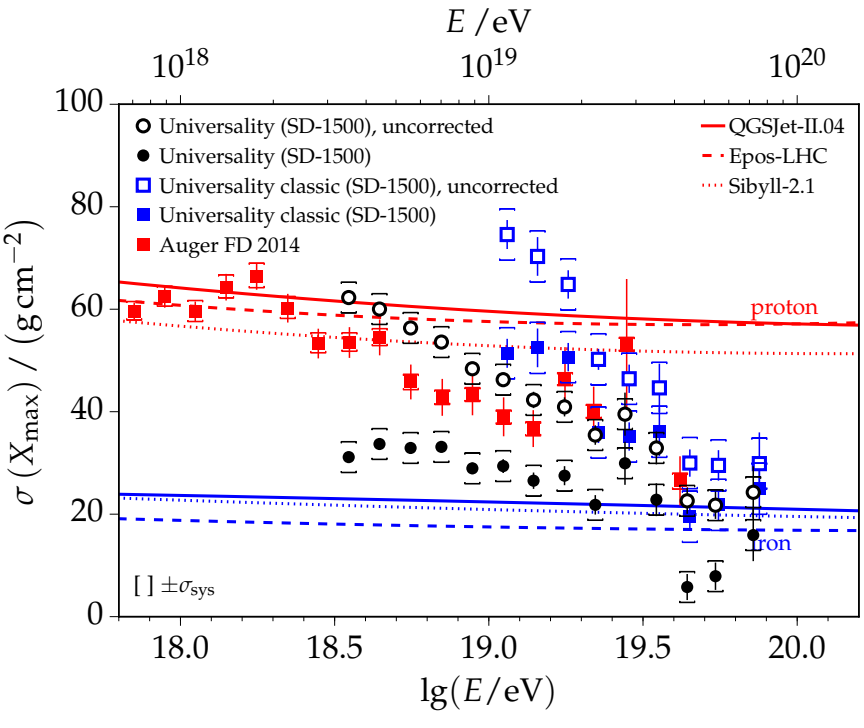
$$\langle \ln A \rangle = \frac{\langle X_{\mathrm{max}} \rangle - \langle X_{\mathrm{max}} \rangle_p}{f_E} , \tag{5.11}$$

with

$$f_E = \xi - \frac{D}{\ln 10} + \delta \lg\left(\frac{E}{E_0}\right) . \tag{5.12}$$

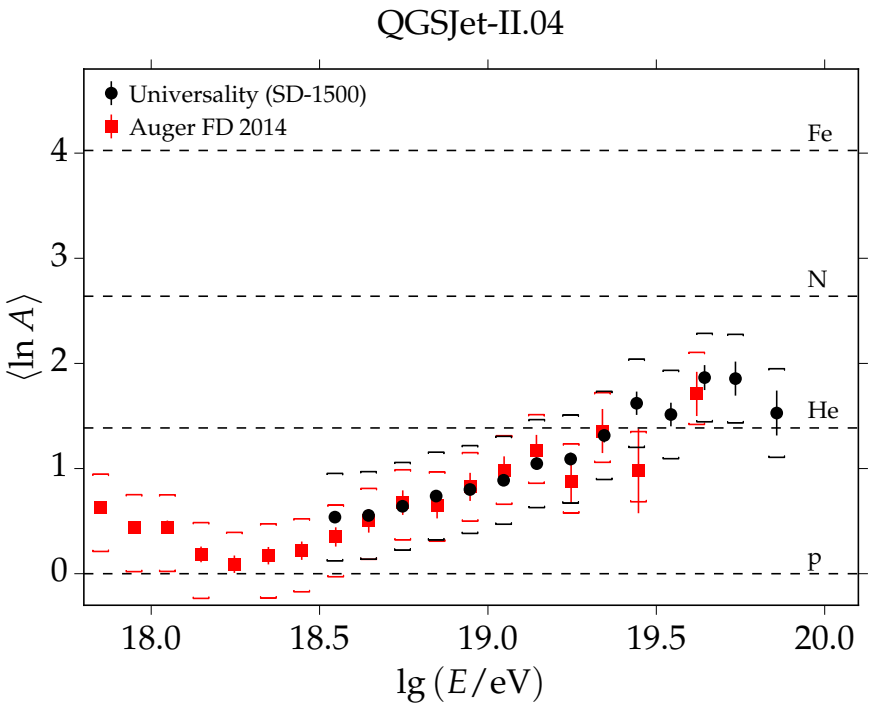


(a)

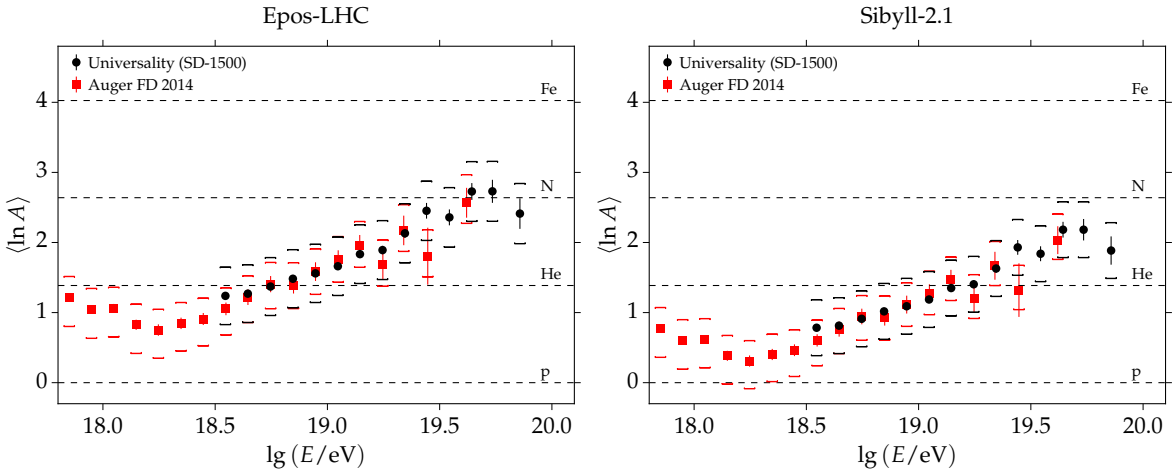


(b)

Figure 5.17: (a) The physical fluctuations of $X_{\rm max}^{\rm Univ}$ derived from SD-1500 data compared to the FD result, and expectations from simulations using current hadronic interaction models. Only non-saturated events are included. (b) A comparison to the uncorrected profiles of the measured fluctuations, prior to subtracting the detector resolution as detailed in the text.



(a) QGSJET-II.04



(b) EPOS-LHC **(c)** SIBYLL-2.1

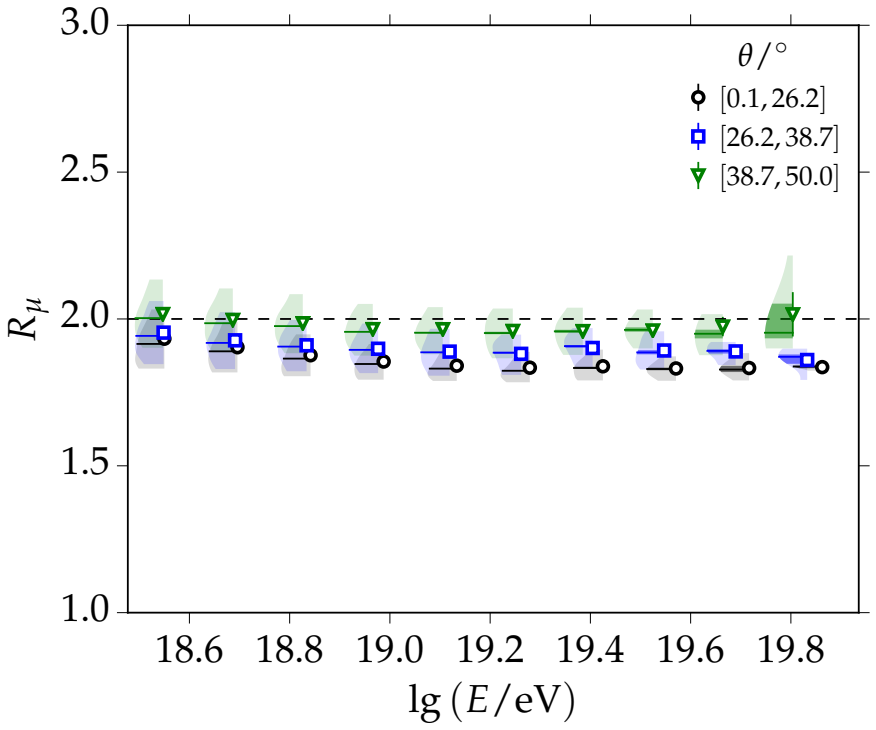
Figure 5.18: Average logarithmic mass derived from the measurement of $\langle X_{\mathrm{max}}^{\mathrm{Univ}} \rangle$. Shown are plots for different hadronic interaction models.

Hereby, $\langle X_{\mathrm{max}} \rangle_p$ is the average depth of shower maximum of proton induced air showers. Results on $\ln A$ from the universality measurement are shown in Fig. 5.18 for the three hadronic interaction models QGSJET-II.04, EPOS-LHC and SIBYLL-2.1. The heaviest composition is predicted by EPOS-LHC. For QGSJET-II.04 and SIBYLL-2.1, the average mass does not increase above nitrogen's mass. Only non-saturated events measured with the SD-1500 and reconstructed with the iterative universality reconstruction are included. The corresponding result on the average depth of shower maximum is discussed in Section 5.5.

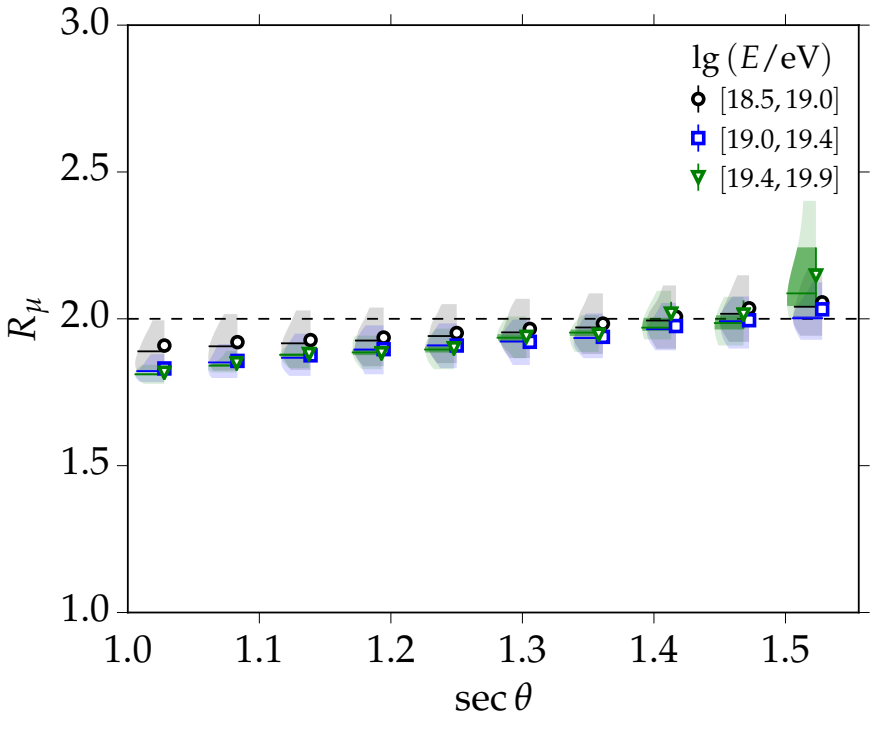
5.8 The relative muon content R_μ

A model of the relative muon content R_μ of air showers was derived in Section 5.2. This model is used in the reconstruction procedure to account for the strong correlation between primary energy and R_μ. Analyzing SD-1500 data, R_μ is refit in a separate reconstruction step, but the result is driven by the input model. The resulting relative muon content is visualized in Fig. 5.19 as a function of energy and zenith angle. The systematic drift due to the aging of the SD detectors as derived in Section 5.3 is accounted for by correcting the values with the fitted model. The reconstruction of data yields an average R_μ of 1.947 ± 0.001 as depicted in Fig. 5.20a, which is well above predictions from simulations. An example for QGSJET-II.03 simulations using the same reconstruction type is shown in Fig. 5.20b. Contemporary hadronic interaction models QGSJET-II.04 and EPOS-LHC predict slightly more muons, but the muon excess in data is still on the order of 30 % to 60 %, depending on the hadronic interaction model. The increase of R_μ at lower energies in Fig. 5.19a is in contradiction to the result found with golden hybrid events in Section 5.2. This is explained by considering the bias of R_μ at lower energies, found in simulations in Fig. 3.19e. The iterative reconstruction was not used for the calibration of golden hybrid events and, as such, these results are not affected by this bias. The increase of R_μ with zenith angle as shown in Fig. 5.19b is compatible with the result obtained with golden hybrid events in Section 5.2. Compared to the evolution of R_μ for a certain composition in simulations, the trend of R_μ in data with energy indicates an evolution towards a heavier composition at higher energies. A more sophisticated study of this evolution and the study of the fluctuations of the muon content is outside the scope of this thesis; it is a very worthwhile task for future analyzes.

In this section, other reconstruction systematics of $X_{\mathrm{max}}^{\mathrm{Univ}}$ as a function of time-related variables and the shower geometry are studied. The influence of tank aging is corrected for as discussed in Section 5.3 and the residual zenith angle dependence is taken into account as demonstrated in Section 5.1. The average reconstructed $X_{\mathrm{max}}^{\mathrm{Univ}}$ is analyzed as a function of time, day of the year, hour of the day and the average age of tanks participating in an event in Fig. 5.21. There is a yearly modulation of $\pm 5\,\mathrm{g\,cm^{-2}}$, which is understood as a change of the atmosphere during the year. A change in the atmospheric parameters influences the longitudinal development of air showers and, thus, the average depth of shower maximum. That could optionally be parametrized and corrected for in future analyses. The change in $X_{\mathrm{max}}^{\mathrm{Univ}}$ with the hour of the day is negligible and on the order of $1\,\mathrm{g\,cm^{-2}}$. The variation with absolute time is below $4\,\mathrm{g\,cm^{-2}}$. This is a residual effect due to the remaining dependence with the average tank age in Fig. 5.21d. Without the correction established in Section 5.3, the change of $X_{\mathrm{max}}^{\mathrm{Univ}}$ with time is on the order of $20\,\mathrm{g\,cm^{-2}}$ and not negligible considering the systematic uncertainty of $10\,\mathrm{g\,cm^{-2}}$. The residual dependence on time is accounted for with an increased systematic uncertainty (see Section 5.5). Dependencies on the shower geometry are depicted in Fig. 5.22. The dependence on zenith angle is shown in Fig. 5.22a. The average X_{max} is plotted as a function of $\sin^2\theta$ and for three different energy thresholds as specified in the legend. Each of the profiles was fit with a constant to quantify the deviation from a constant behavior regarding the χ^2/n_{dof}. There is a significant angular dependence of the low-energy threshold of $10^{18.8}\,\mathrm{eV}$. It could be related to a dependence on the number of candidate stations in the iterative reconstruction method. The profiles for larger energy thresholds are mostly constant with zenith angle, which is the expected behavior. The systematic uncertainty discussed in Section 5.5 is increased by $5\,\mathrm{g\,cm^{-2}}$ to take the residual angular dependence into account. The study of a declination dependence in Fig. 5.22b is of particular importance considering the study of a declination dependence of the flux of UHECRs as detailed in Section 4.9. Again, there is a deviation in the profile of

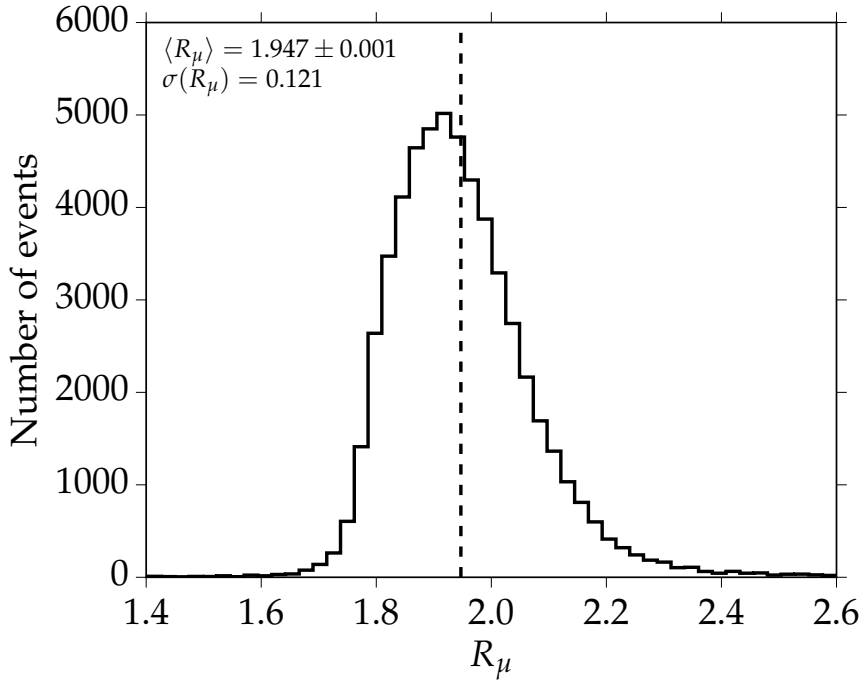


(a)

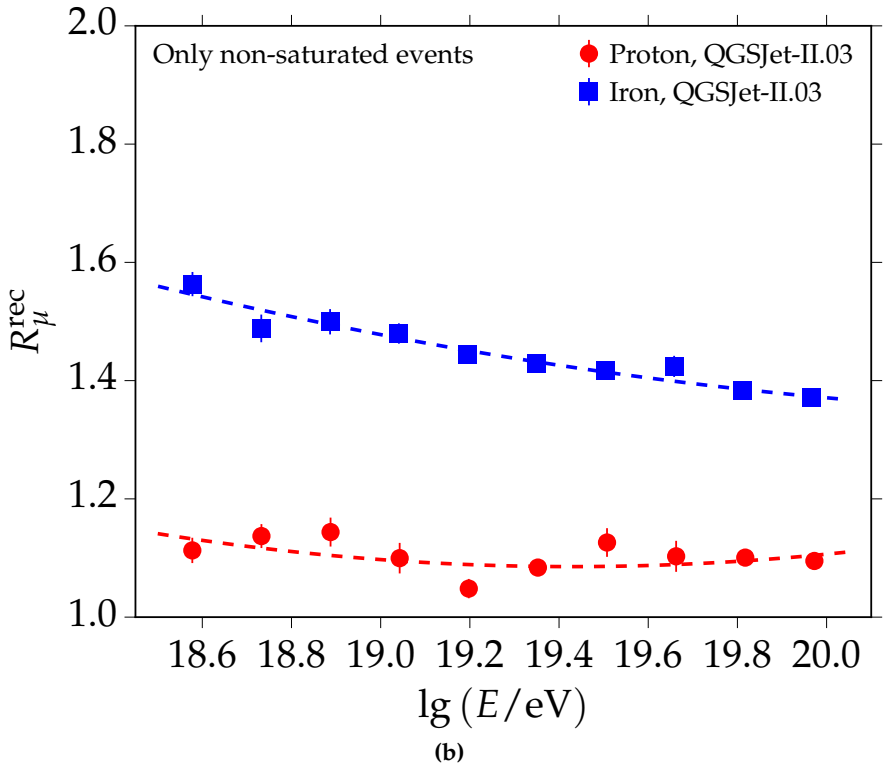


(b)

Figure 5.19: The reconstructed relative muon number R_μ as a function of energy and zenith angle, and for different ranges of reconstructed zenith angle and primary energy. Included are non-saturated events measured with the SD-1500. The iterative universality reconstruction, as developed within this work, was used to reconstruct the events.



(a)



(b)

Figure 5.20: (a) Histogram of all reconstructed R_μ values in SD-1500 data reconstructed with the iterative universality method. (b) Predictions of R_μ for a pure proton and iron composition as a function of energy and from simulations with the hadronic interaction model QGSJET-II.03.

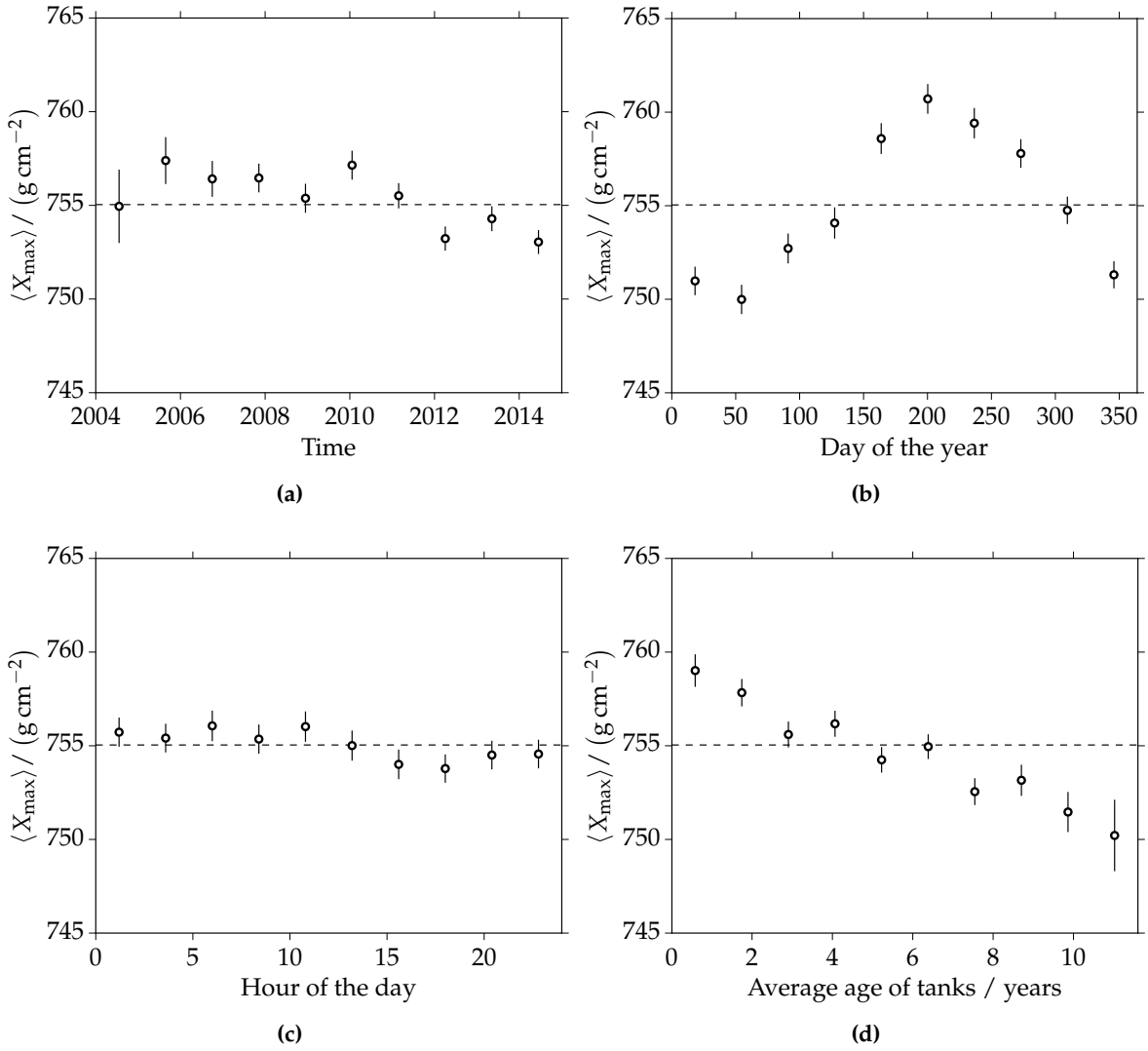


Figure 5.21: Systematic study of the change of the average reconstructed depth of shower maximum with time and the age of tanks. Yearly and daily variations are also shown.

events above $10^{18.8}$ eV. This deviation is completely caused by the relation with zenith angle. Applying larger energy thresholds reveal no significant deviations from a constant behavior. It should be noted though that the reach in declinations on the northern hemisphere is limited to 15° due to an upper zenith angle of 50°. The dependencies on azimuth in Fig. 5.22c are constant for all energy thresholds.

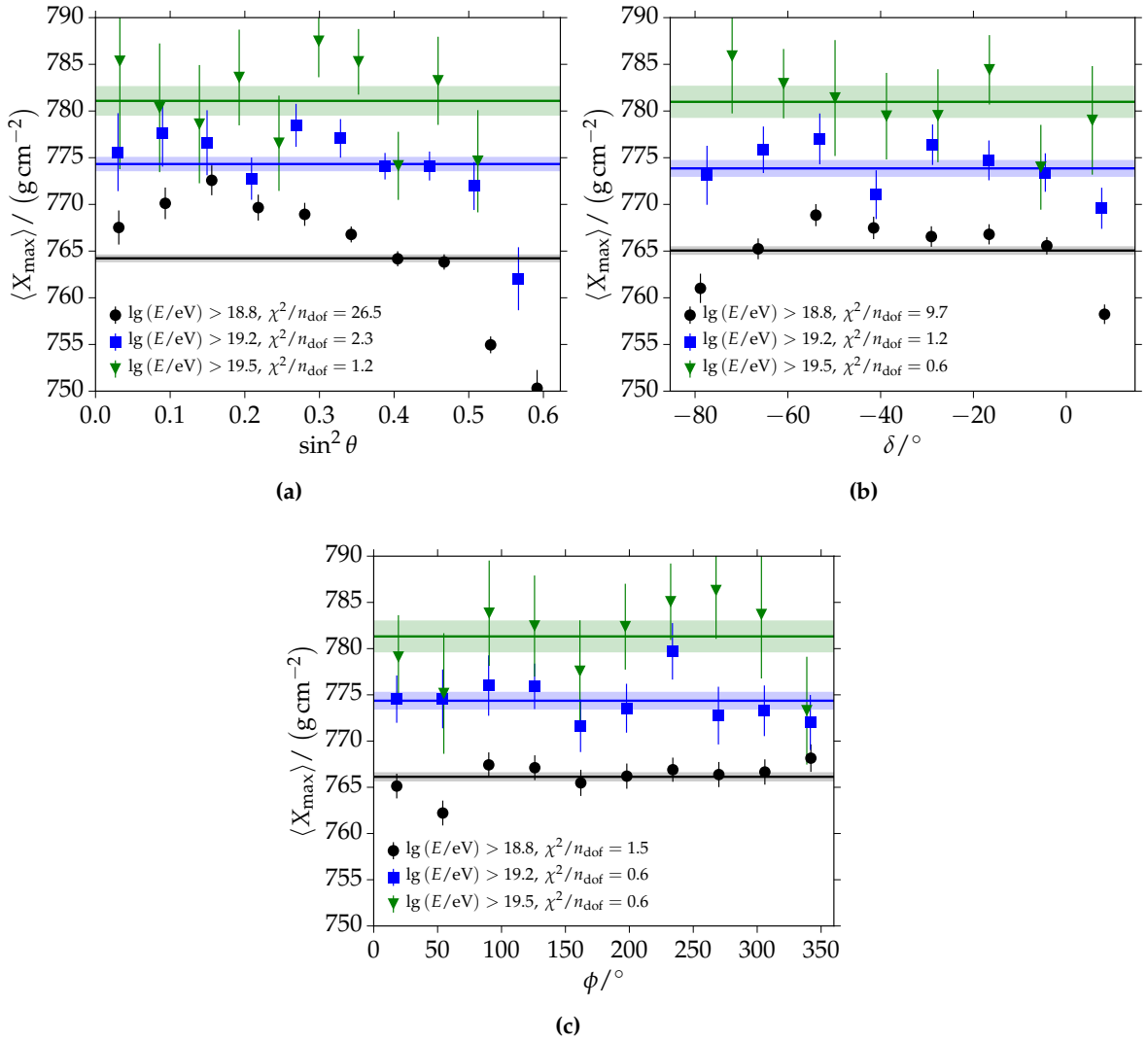


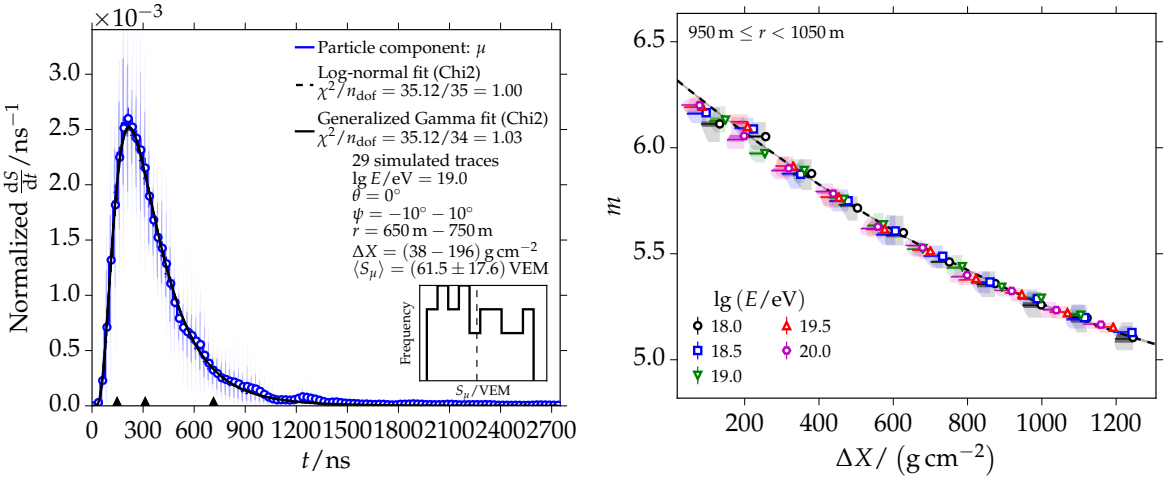
Figure 5.22: Systematic study of the change of the average reconstructed depth of shower maximum with zenith, declination and azimuth angle.

CHAPTER 6

Summary and conclusions

The topic of this thesis is the determination of the mass composition and energy spectrum of ultra-high energy cosmic rays using data measured with the surface detector of the Pierre Auger Observatory. This encompasses the reconstruction and analysis of primary particles with reconstructed energies above 10^{17} eV, as well as the study of simulated air showers and their signal and time patterns in the detectors. A novel reconstruction method based on shower universality was developed and used to derive many of the results in the work, as summarized below. The results produced throughout this work were ingredients to several conference proceedings [23, 86, 187] and will be covered in multiple publications that are currently in progress. The main findings are listed below:

Universality parametrization The paradigm of shower universality allows one to describe the very complex phenomenon of extensive air showers with only a small set of physics parameters, most importantly their energy and the distance to the depth of shower maximum as defined in Fig. A.1b. A major part of this work is the study of the arrival time distributions of different secondary particles of air showers. This was investigated with specific simulations of the WCD surface array of Auger. An application of this method to future detectors - like the upgraded detectors of AugerPrime - is easily possible. The study was performed using simulations of primary particles in the energy range from 10^{17} eV to 10^{20} eV, the primary types proton, carbon, iron and the hadronic interaction models QGSJET-II.03 and EPOS-1.99. Complex analytical time models were derived to describe the arrival time distributions of four distinct particle components. The analysis is based on the fitting of average time distributions in simulations and a subsequent parametrization of the parameters of the fit function as a function of physics parameters. While detailed in Section 2.4, an example of a fit to average traces is shown in Fig. 6.1a; the log-normal distribution was found to give a very good description of the time distributions. An example of the modeling of the mean parameter of the log-normal as a function of ΔX is shown in Fig. 6.1b. The resulting time model is a mixture of physics-motivated and empirical functions. Its correctness was validated with an in-depth analysis of the prediction of individual time distributions and profiles of time quantiles. Dependencies of the time model on the primary energy and geometry were taken into account. Different primary masses and hadronic interaction models cause slight differences in the time model. These systematic effects and their impact on the reconstruction of X_{max} were addressed. An end-to-end validation of the model in the reconstruction of simulated air showers was performed. A completely unbiased estimation of X_{max} is obtained for showers with zenith angles below



(a) **(b)**

Figure 6.1: (a) The average arrival time distribution of muons in a WCD as derived from 29 simulated air showers from proton, carbon and iron primaries with a primary energy of 10^{19} eV. The ranges of other relevant quantities (zenith angle θ, azimuth ψ, radial distance r, distance to the shower maximum ΔX) are stated in the plot. (b) Example of the dependence of the mean parameter of the time model on the distance to the shower maximum ΔX. Shown is an example for muons in WCDs at a distance of 1000 m from the core.

50°. This is a significant step forward with respect to previous works in which there were still substantial biases. The model derived in this work and the reconstruction performance were published in the conference proceeding in [187]. Further publications are in progress.

Universality reconstruction To reconstruct air showers measured with the SD, the universality time model derived in this work was used together with a model of the ground signals. The latter was derived in previous works and validated within this thesis, especially regarding the quality of an extrapolation to energies below 3×10^{18} eV (see Section 2.3). The novel universality reconstruction enables a simultaneous estimation of the primary energy, X_{max}, R_μ, the shower geometry, and the impact position and time of a shower at ground. Due to a large number of parameters and substantial correlations between them, for example energy and muon content, a simultaneous fit of all quantities requires a vast number of stations. Mostly, this is equivalent to a large primary energy. Substantial reconstruction biases manifest themselves at energies below 10^{19} eV. To overcome these issues, a new method of reconstructing air showers in several reconstruction steps and with the use of constraints to previous values was developed. The performance of this iterative reconstruction was studied in detail with simulations in Section 3.4.6. It was shown that all quantities, especially the primary energy and X_{max}, are reconstructed without significant biases. The reconstruction efficiency is 100 %. Outliers in the X_{max} reconstruction were found to have a frequency of less than 10 %. They could be attributed to accidental signal peaks in the time distributions of certain stations participating in the reconstruction.

Method calibration and validation To account for differences in the energy scale, the muon content and the arrival time distributions between data and simulations, a new procedure to calibrate the universality method with golden hybrid events (detected with both SD and FD) was studied. In this procedure, the results of the FD reconstruction are explicitly used. Only the relative muon content, R_μ, and the time model offset, Δm_μ, are

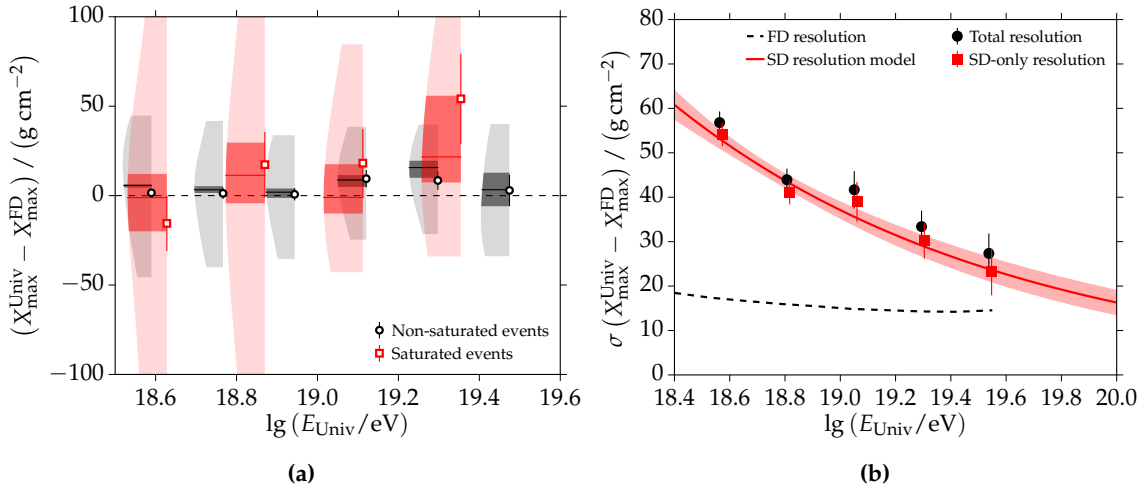


Figure 6.2: (a) The accuracy of the $X_{\max}$ reconstruction as determined with golden hybrid events. (b) The resolution of $X_{\max}^{\mathrm{Univ}}$ for non-saturated events and as a function of primary energy.

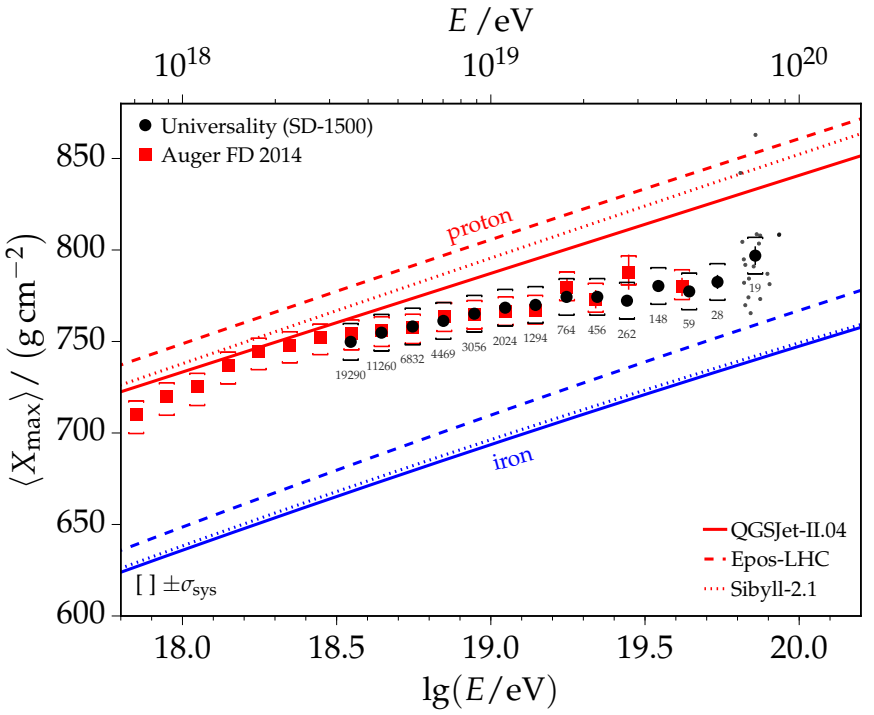
determined. In a first step, event-by-event estimates of these quantities are reconstructed. In a second step, they are parametrized as a function of primary energy and zenith angle. The resulting models are used in the reconstruction of SD data; no information from the FD reconstruction is used. A significant drift of the detector response of SD stations with the time since their deployment was found. The impact of this drift on the reconstructed primary energy, $X_{\max}$ and R_μ was quantified and corrected for. The result is an unbiased reconstruction of $X_{\max}$ when compared to the FD results (visualized for golden hybrid events in Fig. 6.2a). Data are split into non-saturated and saturated events. Both profiles indicate an unbiased estimation of $X_{\max}$ in the whole energy range. The quality cuts and event selection required for obtaining these plots are discussed in Section 5.1. The accompanying resolution in the reconstructed $X_{\max}$ that is achieved with the iterative fit and for the subset of non-saturated events is depicted in Fig. 6.2b. It is below 40 g cm^{-2} for primary energies above 10^{19} eV and reaches less than 20 g cm^{-2} at 10^{20} eV.

Mass composition After calibrating the universality method to data and accounting for the observed time drifts and residual systematic dependencies, the analysis was applied to the full set of data recorded with the SD-1500. The result on the average depth of shower maximum derived in this work is compared with the current result from the FD measurement in Fig. 6.3a. Prediction lines for the elements proton and iron from contemporary hadronic interaction models are included. There is an excellent agreement between the SD and FD results. The general trend in the energy range above $10^{18.5}$ eV is the evolution towards a heavier composition. Surprisingly, the universality results at the highest energies indicate a trend towards lighter elements. While the current significance of this feature is less than $2\,\sigma$, it is of crucial importance to astrophysics and particle astronomy at the very highest energies and needs to be studied further. Using current hadronic interaction models, $\langle X_{\max} \rangle$ was converted to the average logarithmic mass $\langle \ln A \rangle$. Judging from these results, an evolution of the mass composition from proton to helium/nitrogen is favored. The presence of heavier elements is disfavored by the current results. The relative muon content as determined from SD-1500 data is depicted in Fig. 6.3b. On average, R_μ is 1.947±0.001. Employing the energy scale from data, this result implies that the number of muons in simulations is underestimated by 30 % to 60 %, depending on which hadronic interaction

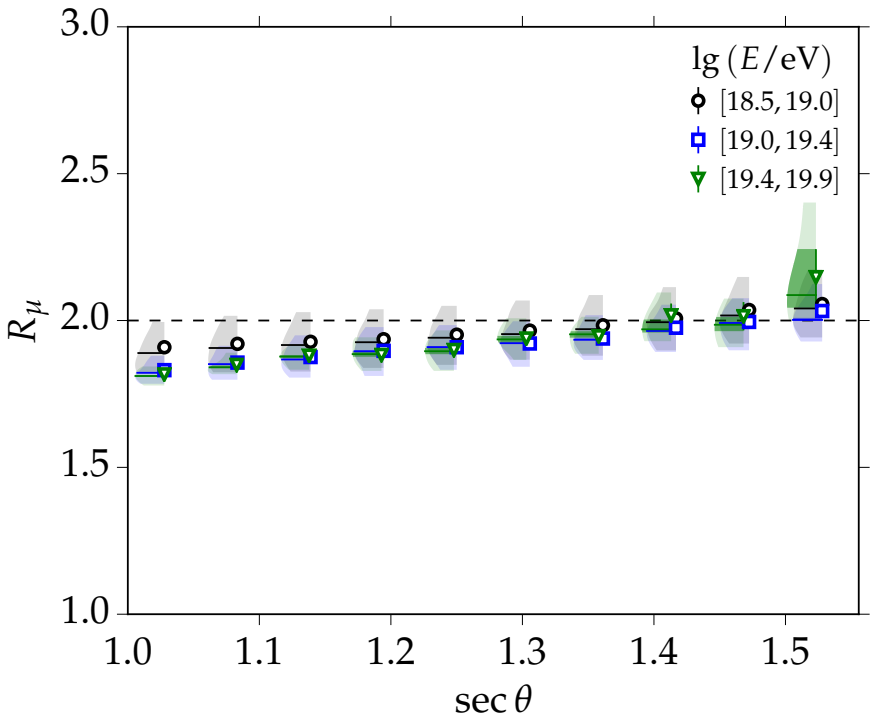
model is used. The evolution of R_μ with the primary energy also supports a trend towards a heavier composition. As concluded in this and other analyses, it is unlikely that the energy scale of data is wrongly determined to compensate for the discrepancy in the muon content. Recent developments in hadronic interaction models might soon shed new light on the puzzle of missing muons.

Energy spectrum A major part of this work was the analysis of the flux of UHECRs (also called the energy spectrum). Together with the universality approach, a refined standard reconstruction was used to obtain these results. This standard approach is based on the empirical description of the overall lateral distribution of secondary particles on the ground. For each event, the distribution is fit with an LDF. The size of this function at an optimal distance to the shower core is a very robust estimate of the primary energy. Within this work, a few advances in the standard reconstruction were obtained. Most of the effort was put into a consistent analysis of both SD-750 and SD-1500 data. Due to the detection efficiency, a measurement with the SD-1500 is only possible above primary energies of 3×10^{18} eV. Using the nested SD-750 array allows us to lower this energy threshold by more than a decade. The correction of the zenith angle dependencies of the shower sizes from SD-750 and SD-1500 with a CIC method was studied in detail. A new method to obtain the attenuation functions was employed. The impact of corrections due to the influence of weather and geomagnetic effects on the shower sizes was also analyzed, resulting in minor corrections. The derived zenith-independent energy estimates for SD-750 and SD-1500 events are called S_{35} and S_{38}. The calibration of both quantities to energies measured with the FD was investigated. A detailed study of the event-by-event energy difference of events measured with both the SD-750 and SD-1500 motivated the development of an improved method to calibrate SD-750 data. The detector resolutions for both measurements were studied in both data and with the use of dedicated simulations. Their impact on the measurement of the flux due to event migration effects was analyzed and corrected for. A new forward-folding method was developed to take into account the respective detector resolutions and efficiencies. The resulting corrected energy spectra were thoroughly studied. A distinct flattening of the flux appears at an energy of $10^{18.72\pm0.01\pm0.01}$ eV. This feature of the spectrum is called ankle. From below to above the ankle, the spectral index changes from $-3.20 \pm 0.01 \pm 0.04$ to $-2.52 \pm 0.03 \pm 0.02$. The spectrum derived from SD-750 data exhibits unexpected and interesting changes in the spectral slope at energies below the ankle. Furthermore, the analysis of this work confirms a strong suppression of the flux above an energy of $10^{19.56\pm0.03\pm0.02}$ eV. Above that energy, the spectral index changes by $-2.6 \pm 0.2 \pm 0.2$. The results are depicted in Fig. 6.4a. A combined SD spectrum is compared to other recent results in Fig. 6.4b. Previous results on the flux of UHECRs derived in this work were used in the conference proceedings [23, 86]. A new analysis was the determination of the energy spectrum from SD-1500 data with shower universality. The result is in perfect agreement with the standard result up to a primary energy of 10^{19} eV (see Fig. 6.4a). A discrepancy at higher energies might have interesting implications for the energy calibration and requires further studies.

Declination dependence of the measured flux Motivated by the large difference between the energy spectra measured by Auger in the southern hemisphere, and by TA in the northern hemisphere, a novel search for a declination dependence of the measured flux of UHECRs was performed. This comprises the study of the experimental exposure as a function of declination and the study of the flux in different declination intervals. No significant dependence of the spectral features on the incoming direction of primary particles was found. However, a distinct excess of events from the south was observed, as indicated



(a)



(b)

Figure 6.3: (a) The average depth of shower maximum derived from SD-1500 data using the shower universality method developed within this work. A comparison to the published FD result is shown. (b) The reconstructed relative muon content in SD-1500 data as a function of zenith angle and for different ranges of primary energy.

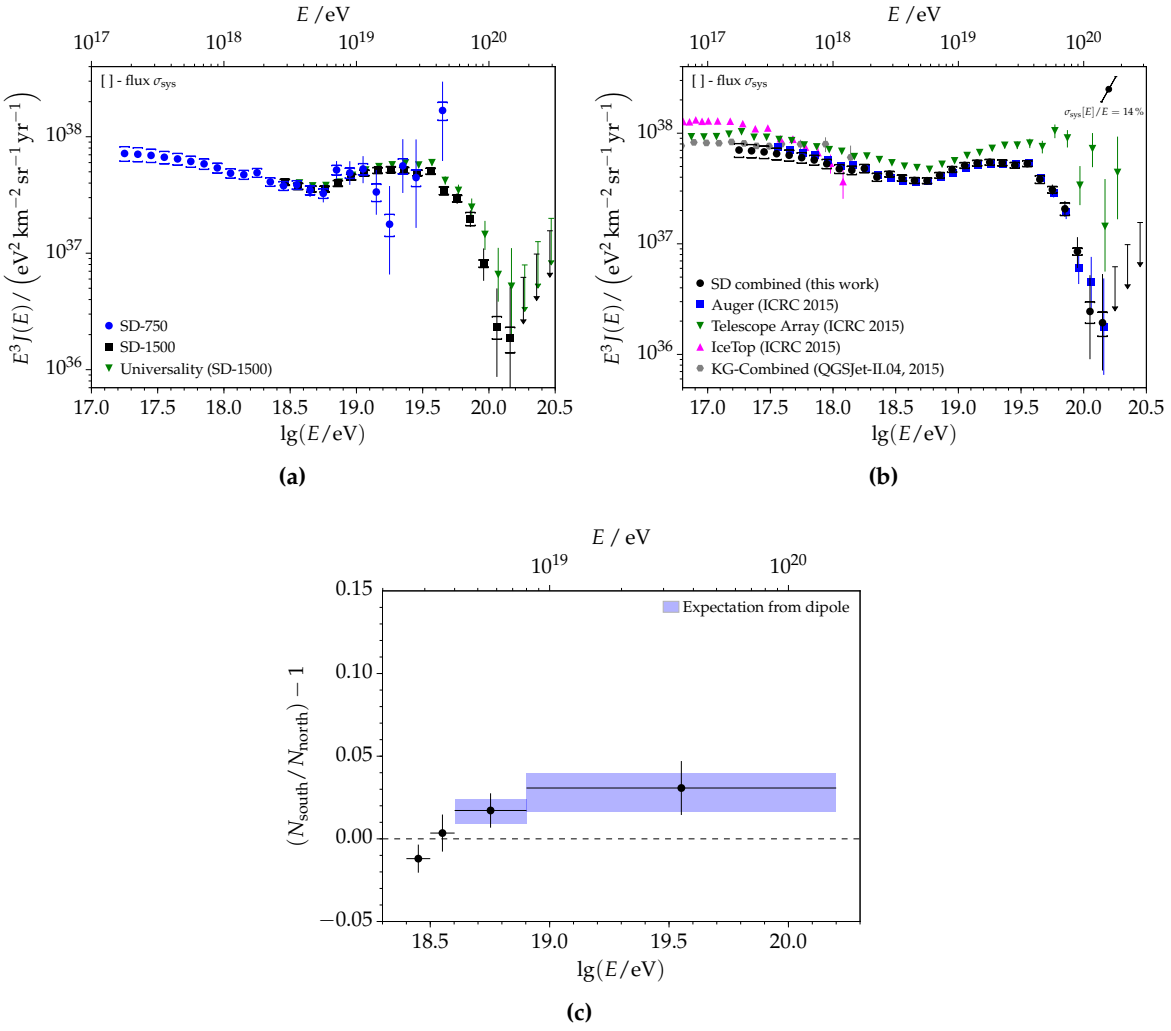


Figure 6.4: (a) The forward-folded energy spectra from data measured with the SD-750 and SD-1500. (b) The combined SD spectrum derived in this work in comparison to other recent measurements. (c) Ratio of the event rates from the northern and southern sky as a function of the primary energy. The results are compared to the expectation from an independent result on the large scale anisotropy, as further detailed in the text.

in Fig. 6.4c. Above $10^{18.8}$ eV, the relative difference in the event rate increases to 3 %, with a statistical significance of $2\,\sigma$. This result confirms and is in agreement with an independent and published study of a large scale anisotropy of events measured with Auger [81]. The expectation from the measured dipole anisotropy is visualized with blue regions in Fig. 6.4c. The results of the declination analysis indicate that the difference in the energy spectra observed by Auger and TA is not caused by an actual difference of the flux from different regions in the sky. However, this is not conclusive because Auger data for events with zenith angles below 60° are limited to a declination range of −90° to 25°. A strong change in the flux for events above a declination of 25° could still explain the difference in the energy spectrum. Differences in the overall energy scales of the experiments are, however, a more likely explanation. A publication of the study of the declination dependence is in progress. A preliminary result obtained in this work shows no dependence of $\langle X_{\mathrm{max}} \rangle$ on declination.

The methods developed throughout this work are essential tools to obtain improved physics results with the upgraded detectors of AugerPrime. Disentangling the muonic and electromagnetic particles with two different detectors will enable a validation of the universality models with data. Furthermore, it will break the degeneracy in the reconstruction of extensive air showers due to the large correlations between energy, muon number and X_{max}. A significant improvement in the resolution of reconstructed quantities is expected.

List of Tables

APPENDIX A

General

A.1 Shower geometry and coordinates

Often used concepts about shower geometry are depicted in Fig. A.1, definition of the shower coordinate system is shown in Fig. A.1a and the definition of the distance ΔX to the shower maximum is plotted in Fig. A.1b.

A.2 General plot comment

For most analyses in this work, data need to be investigated as a function of many distinct variables. If a profile (mean, median,...) of a quantity with respect to another quantity is given, the general rule applies that all data are included unless specified otherwise in the text. For example, if a shower library consists of proton and iron showers and a quantity is plotted as a function of energy and for different ranges of zenith angle, then simulations of both primaries are included unless stated otherwise.

A.3 How to read violin plots

Violin plots are often used within this work to display information and deserve an exhaustive explanation in this section. While the main content of a violin plot should be intuitive to read, some details of the visual presentation might not be immediately clear.

A violin plot is an extended (or I like to say: improved) version of a profile or box plot. It gives the graphical representation of a quantity y as a function of another quantity x. Often x and y are continuous variables and one is interested in seeing the average of y in different (disjunct) ranges of x. That is why one divides the data into i different bins x_i. The typical profile then includes mean and standard deviation of $y_i(x_i)$.

Such a profile is already a significant improvement over a simple scatter plot of all points. The human brain is bad in interpreting scatter plots (also they look ugly, and the plots are huge when saved as a decent vector graphic). However, a scatter plot is optimal in the sense that it contains all the information while an average profile might lack most of the information. Specifically, it lacks information as soon as the data are not distributed according to a perfect Gaussian. One can easily imagine a bimodal distribution with modes symmetrically on either side of a mean value. Then the mean will be in between where there might be no points at all. In that case, the reader would be completely fooled by a

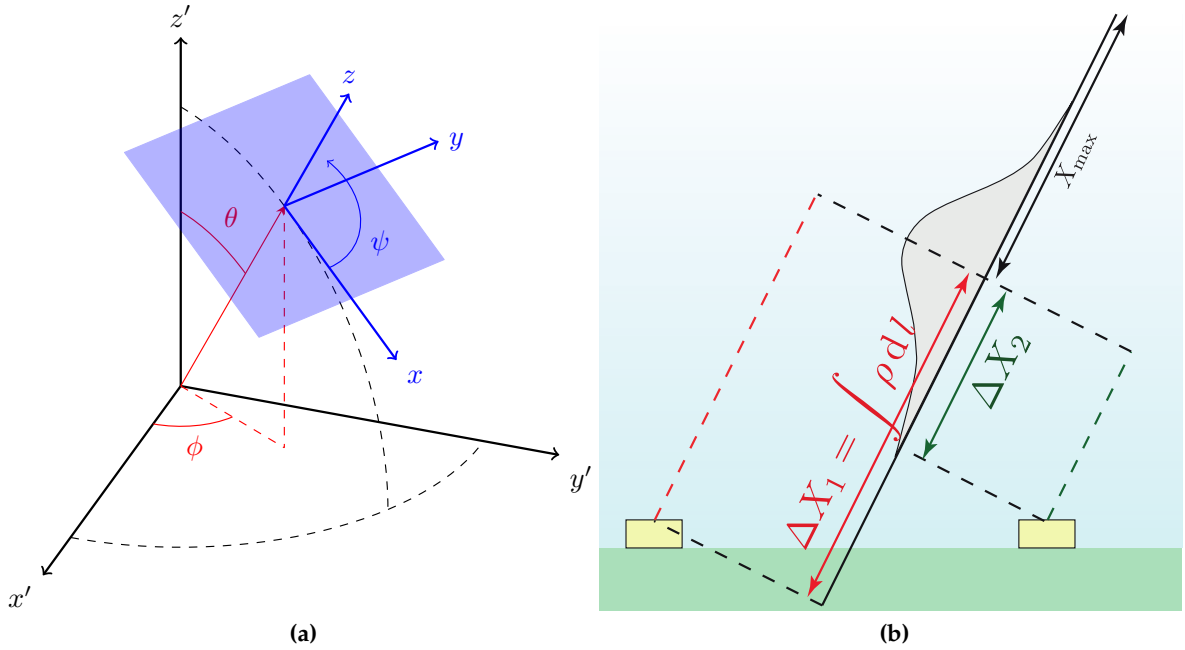
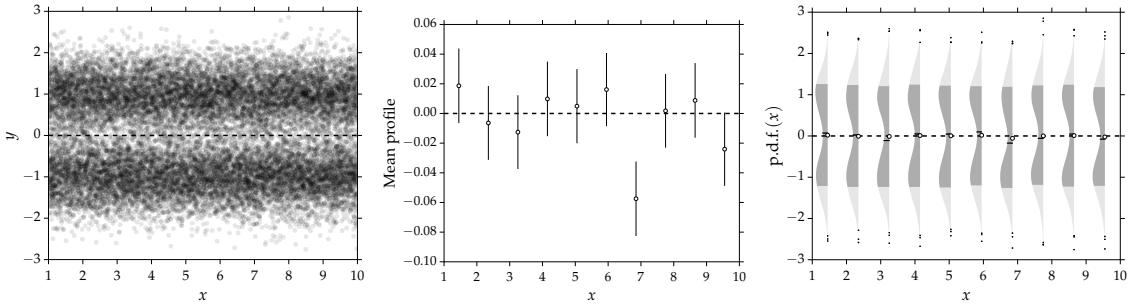


Figure A.1: (a) Visualization of the shower plane coordinate system centered at the shower core. The radial distance to the core is given by $r = \sqrt{x^2 + y^2}$ provided that all the stations are projected into the plane at $z = 0$. The local shower azimuth is denoted as ψ. (b) Schematic of the distance ΔX to X_{max}. It is defined as the atmospheric overburden integrated along the shower axis. In particular, ΔX is different for stations at the same radial but at different azimuthal locations. This is illustrated in the picture with a station in the early region at $\psi = 0$ and a station in the late region at $\psi = 180°$. Early and late part are also called upstream and downstream, respectively. ΔX for the late station is larger than for the early station (from [187]).

profile of mean values because vital information is missing (I am pretty sure that this lead to a crash at NASA at least once). An example is plotted in Fig. A.2. Two Gaussians that are centered at -1 and 1 are superimposed, resulting in a bimodal distribution centered at 0. The structure is visible in the scatter plot, but it is hard to read. The profile does not reveal features of data properly while the violin plot does. Thus, the violin plot extends the common profile by just that missing information, while almost retaining its simplicity.

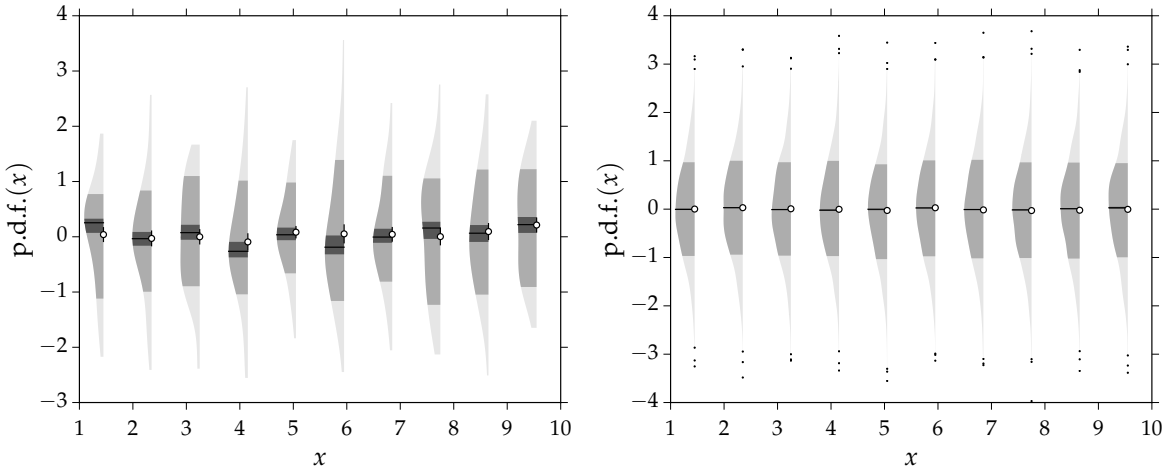
Let me present another example in Fig. A.3. Toy data are linearly distributed in x between 1 and 10 and normally distributed in y with a true mean of 0 and a standard deviation of 1. The sample size is 500 in Fig. A.3a and 20 000 in Fig. A.3b. In each of the ten bins, the distribution of data in y is shown as a shaded area in the background. The magnitude of the violin at a certain point on the y-axis reflects the size of the p.d.f. at this point. The innermost black lines represent median values of y_i for each bin x_i. The dark area around these lines represent asymmetric median $1\,\sigma$ uncertainties. The $1\,\sigma$ asymmetric standard deviations ($\pm 34.1\,\%$) are shown with the dark part of the violin, while the whole violins extend up to $3\,\sigma$ by default ($\pm 49.9\,\%$). The violins are kernel density estimates of the data [188, 189]. Functions from the Python library SCIPY are used [190]. Circles are plotted at the bin centers and give the mean and the standard deviation of the mean as vital information. Also outliers outside of the $3\,\sigma$ quantiles are shown in Fig. A.3b. Due to the small sample size, there are no outliers in Fig. A.3a. In the latter, median and mean values fluctuate much more from the truth though [189].

The whole point of the violin plot is to give the reader a better impression of how data are distributed, which is often crucial, not only because it regularly prevents the



(a) (b) (c)

Figure A.2: A sample of 20 000 points (x, y) with x uniformally distributed between 1 and 10, and y drawn from Gaussians with standard deviation of 0.5 and means at either -1 or 1. Data are shown as (a) scatter plot, (b) profile of mean values and (c) violin plot. The mean value of bin 7 is just a statistical outlier.



(a) (b)

Figure A.3: Violin plots of normally distributed data divided into 10 bins. An explanation is given in the text. (a) Random sample of size 500, i.e. 50 points inside each bin. (b) 2000 points within each bin.

misinterpretation of data during the analysis stage (and I consider it best practice to use it), but because it helps to understand inferences made from data. The actual plot has been heavily optimized according to criteria for excellent visualization of data as given in [191], i.e. reasonable maximization of the data-to-ink ratio. I have found that this ratio can be easily reduced even more but this is paid with worse readability then.

A.4 Comments

A.4.1 Ideas for a future time model

This section contains a few remarks on the idea of a time model that is independent of the detector response as mentioned in Section 2.4.

I judge it as very worthwhile to derive a model for the time-dependent flux of secondary particles in an air shower as a function of all relevant physics and geometry parameters. At

first, it would allow us to apply this model to different detectors (e.g. not only WCDs but scintillators among others). At second, it is a way to learn more and in-depth about particle physics in air showers. A simple detector like the WCD washes out most of the plentiful information before we can even extract it. It would be especially insightful to study which are the most fundamental particle components in the sense that they scale very differently with primary characteristics. Also, the study of different energy and momentum thresholds is of high interest in this context.

However, for this to be applicable, we need to have a model of the detector response to be able to fold it with the time model. The current status is that we are not even close to having such an analytical model. Of course, one can use simulations of the response or even tabulated responses to make it faster [107]. However, if we want to fit measured data to infer properties of the primary particle, the response would have to be calculated in each minimization step. In principle, this could work, but it would be computationally highly challenging and a whole new level of analysis. The other idea would be a hybrid approach: most parts of the model are parametrized before the detector response and the rest afterwards (an extended version of what was done in [97, 100]). Even another approach would be to infer true signals or particle densities from the measured ones and fit a model to these. That requires multiple detectors with different responses to muonic and electromagnetic particles to allow this kind of backward calculation. The upgrade of Auger (see Section 1.2.7) is a step in this direction, but I doubt that it will suffice in this respect. We will probably continue to parametrize models for different detector responses and using these we can already learn a lot from measurements with the upgraded detectors.

One may not forget that whatever we parametrize with simulations we need to trust these simulations. The ultimate goal is to derive models solely from data and for that we need large event statistics with a diversity of detectors, which are differently sensitive to different kind of particles. The future is a challenge.

A.5 Ideas for future improvements and analyses

- Study of the universality time and signal models using simulations with new hadronic interaction models.
- Analysis of the reason for remaining reconstruction biases for events with zenith angles above 50°.
- Refinements to the iterative reconstruction method (station selection, trace scanning) in order to minimize the number of reconstruction outliers.
- Study of universality X_{max} results with SD-750 data.
- Detailed study of the energy spectrum for different mass groups. Results from the shower universality analysis should be used to perform a discrimination on the primary mass.

A.6 Remarks to references

Many references in this work point to Auger internal notes (GAP-notes). The reason is that the related analyses are new and have not been published yet. Still, these references should be given with the remark that GAP-notes are not publicly available. For non-members of the Auger collaboration, the access to specific notes can be requested at markus.roth@kit.edu.

APPENDIX B

Analyses

B.1 Atmospheric effects

Atmospheric parameters like air pressure, temperature and humidity influence the development of EASs. In particular, the relation of atmospheric height to overburden in $\mathrm{g\,cm^{-2}}$ depends on the state of the atmosphere. Secondary particle densities are affected by different parameters in each development stage of the air shower. To account for most of these effects, air shower simulations in this thesis were generated with different monthly atmospheric models, for example, most of the ones described in Appendix E. These models represent the change of the atmosphere over the year and were derived from measurements [44, 143, 192, 193]. Reconstructed quantities derived from simulations thereby include fluctuations due to different atmospheric conditions. This makes simulations comparable to data. The atmospheric height is plotted as a function of grammage in Fig. B.1a. The dashed line at 1400 m represents the average height of detectors at the Auger observatory. The corresponding average atmospheric overburden is $880\,\mathrm{g\,cm^{-2}}$. Functions for different monthly models are shown, but the differences are hardly visible in this representation. The difference of overburden to the average as a function of height is given in Fig. B.1b. Differences due to atmospheric conditions are pronounced in this representation. Differences at the ground level of Auger site are small on the order of $3\,\mathrm{g\,cm^{-2}}$. Maximal differences on the order of $10\,\mathrm{g\,cm^{-2}}$ occur at heights of ~ 10 km. This is typically between the first interaction and the shower maximum, but closer to the latter one. Thus, the variance contributes to the measured fluctuations of the shower maximum. Only functions for vertical showers are shown in the plots in Fig. B.1.

B.2 Air shower signals

In the context of universality analyses, air shower signals are divided into four components as defined in Section 2.2. Additional plots show dependencies of these particle components.

B.3 Universal time model: additional material

This section provides additional material and plots on the time model developed within this work.

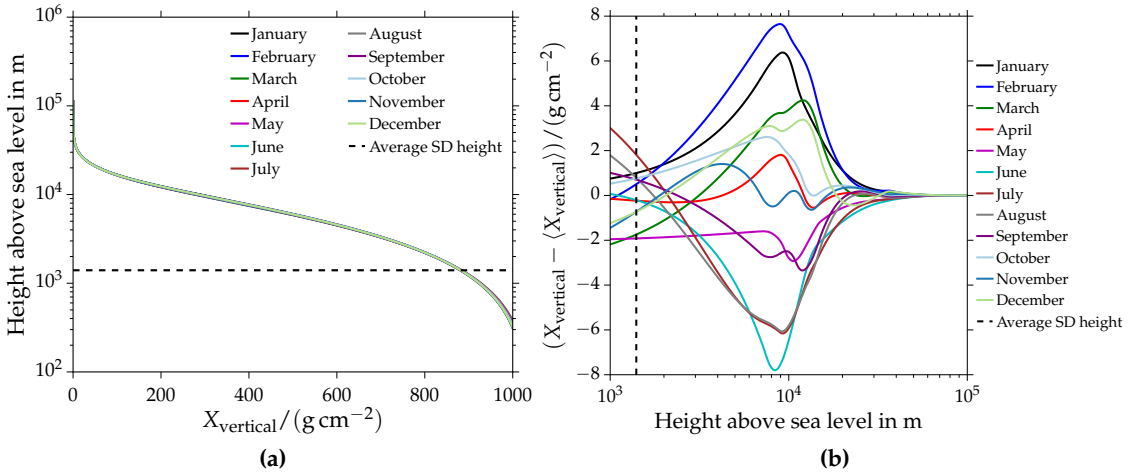


Figure B.1: (a) The average atmospheric height as a function of overburden for different months. (b) The difference of overburdens with respect to the average as a function of height above sea level and for different monthly models.

B.3.1 Dependencies on ΔX

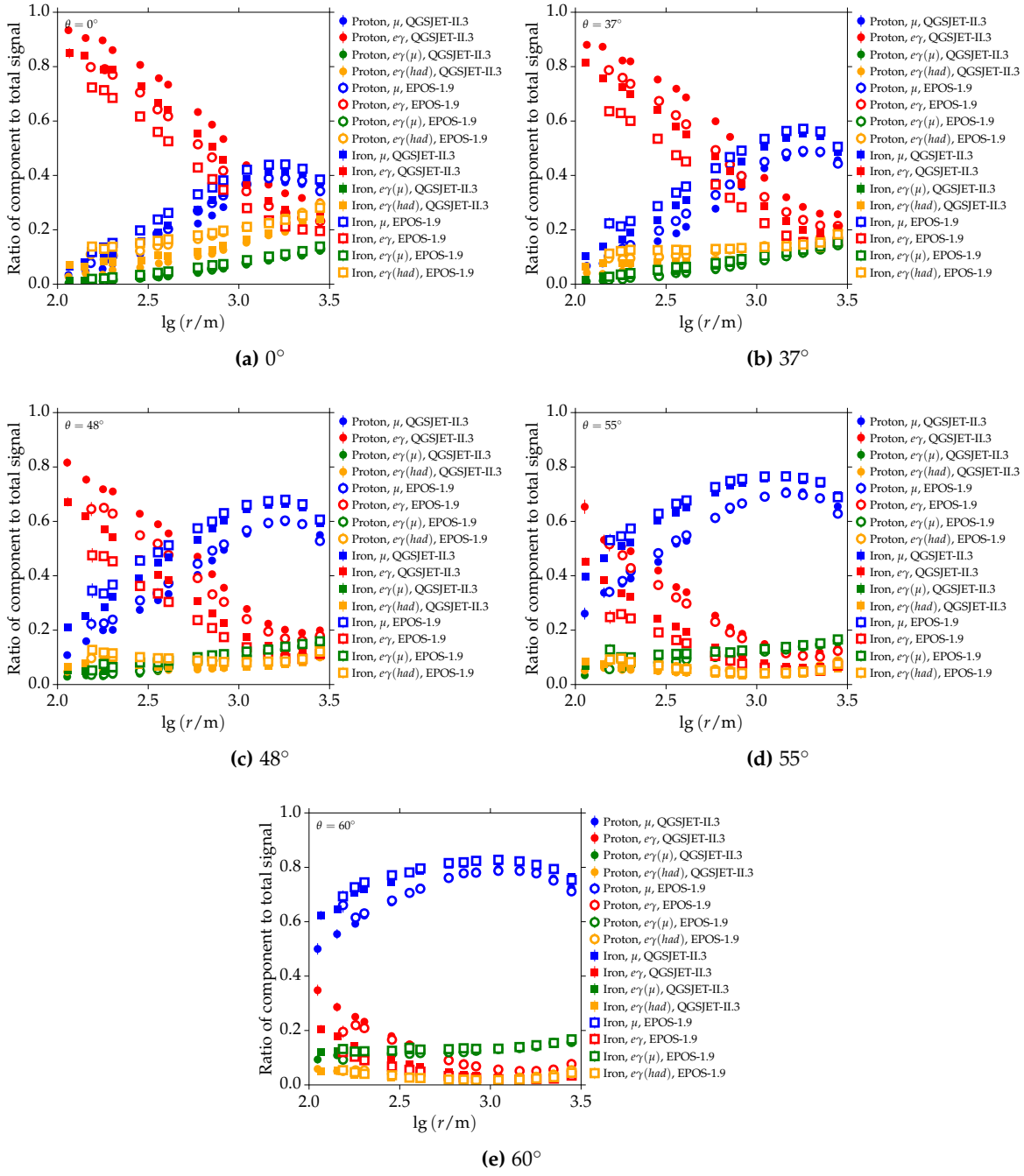
Plots of the main ΔX dependence of the parameters m and s are given for all four particle components. Equivalent plots are included for the average radial distances 500 m, 1000 m and 1500 m. Muons: Fig. B.4 and Fig. B.5, pure electromagnetic component: Fig. B.6 and Fig. B.7, muon decay products: Fig. B.8 and Fig. B.9 and hadron jets: Fig. B.10 and Fig. B.11.

B.3.2 Radial dependencies

Plots of the radial dependence of the parameters of the fit models of m and s are given for all four particle components. The functional form of the model to describe m and s is stated in Eq. (2.18). The parameter for the mean model are listed in Eq. (2.19), while the ones for the model of the width are given in Eq. (2.20). Muons: Fig. B.12 and Fig. B.13, pure electromagnetic component: Fig. B.14 and Fig. B.15, muon decay products: Fig. B.16 and Fig. B.17 and hadron jets: Fig. B.18 and Fig. B.19. It should be noted that the χ^2/n_{dof} values of some fits are bad, because there are cross-correlations between the different parameters. These are not fully taken into account in the segmented fitting procedure (radial dependencies are fit after the other dependencies). This issue can be resolved with a global fit of all parameters. First tests on this did not result in improved parameterizations.

B.3.3 Parameters

Each of the polynomial fits of the plots in the previous section corresponds to a set of parameters. These parameters are given in the tables in this section. The time model parameters for the mean model are given in Table B.1, while the ones for the spread model are stated in Table B.2. They correspond to the parameters of the radial parametrization of the model parameters stated in Eq. (2.18). The quantity $\hat{r}$ used in the polynomial models is defined as $\hat{r} := r/1000\,\text{m}$. To avoid statistical overfitting, the optimal polynomial order is determined with a leave-one-out cross-validation method. The significant digits of the parameters are chosen according to their statistical uncertainties. To increase the readability, the uncertainties are not explicitly given.

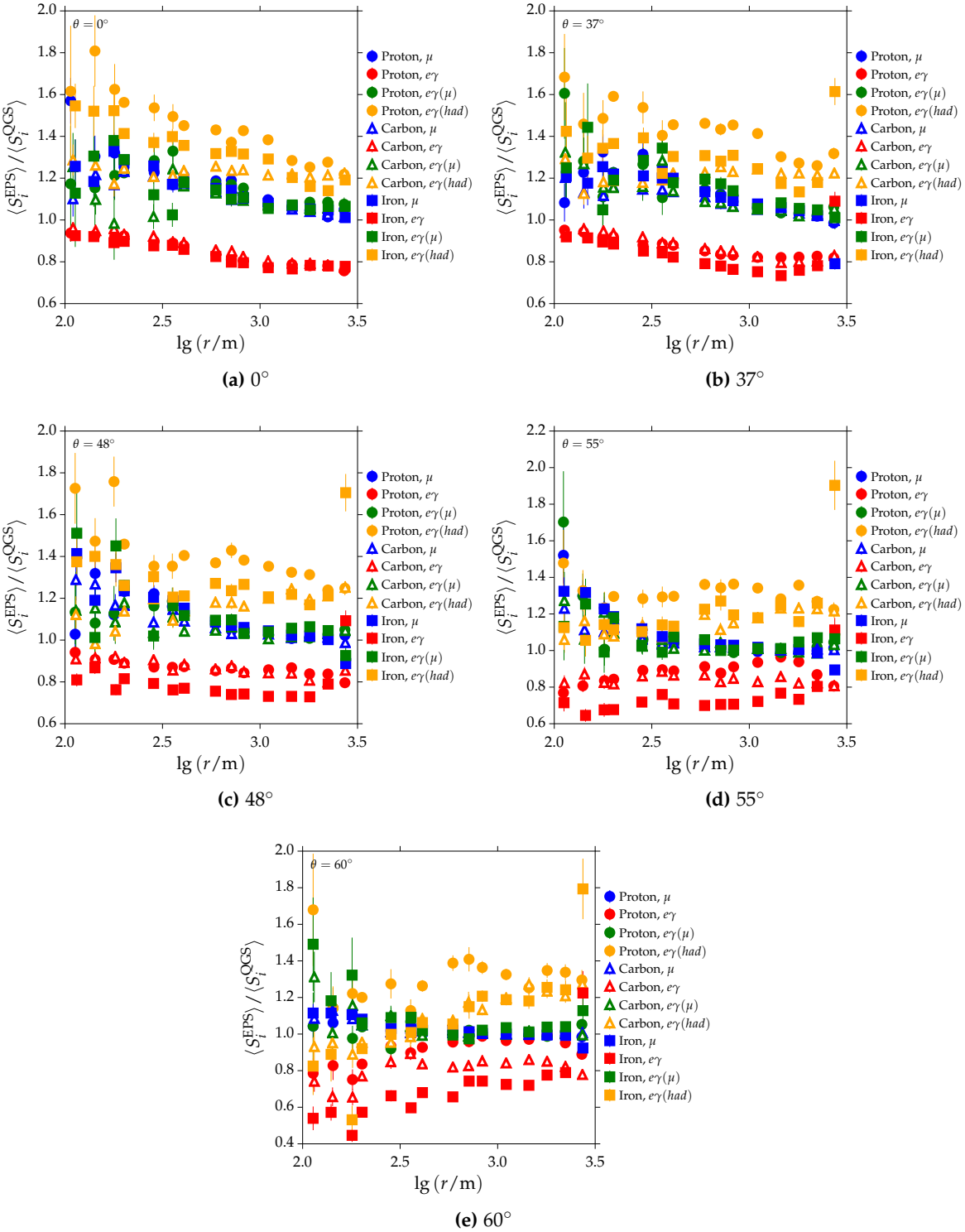


(a) 0° **(b)** 37° **(c)** 48° **(d)** 55° **(e)** 60°

Figure B.2: Ratio of the component signals with respect to the total signal as a function of the logarithmic distance to the core. Each plot corresponds to a subset of events with different zenith angles. Different colors and markers represent different interaction models, particle components and primary species.

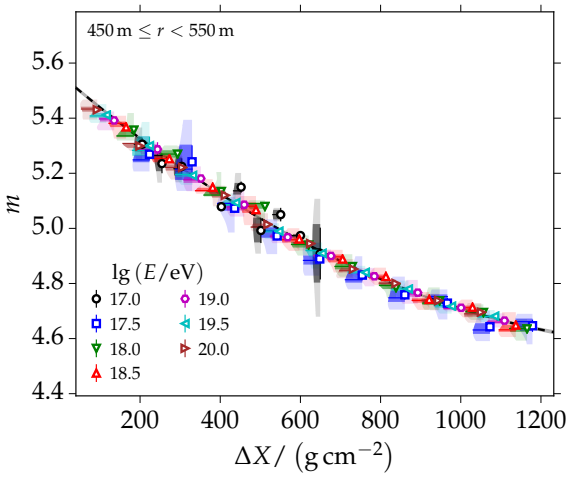
B.3.4 Residuals

Additional residuals of the time model parameters with respect to different quantities and for different particle components are included in this section.

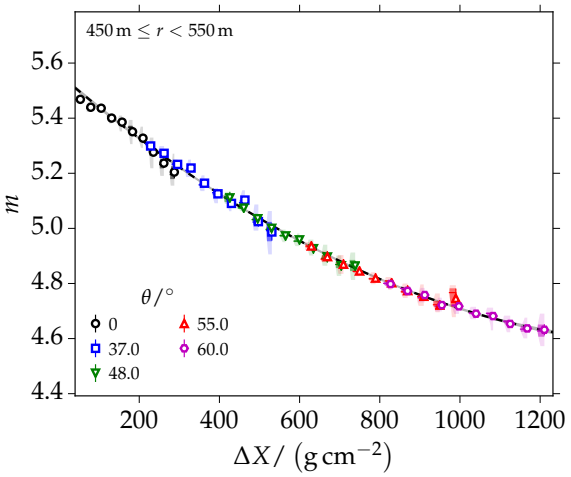


(a) 0° **(b)** 37° **(c)** 48° **(d)** 55° **(e)** 60°

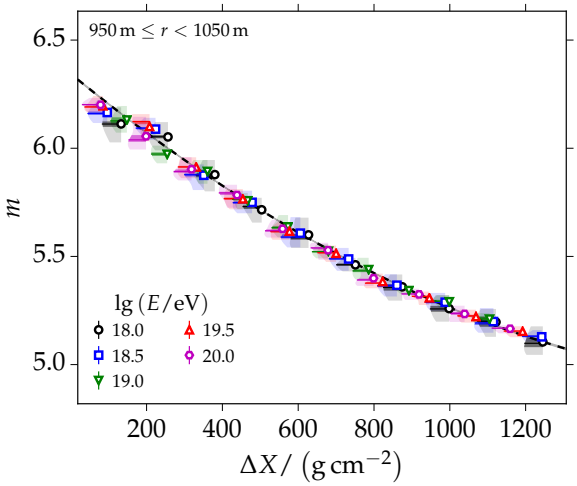
Figure B.3: Ratio of component signals simulated with EPOS-1.99 with respect to the ones simulated with QGSJET-II.03. Different primary species and particle components are indicated with different markers and colors.



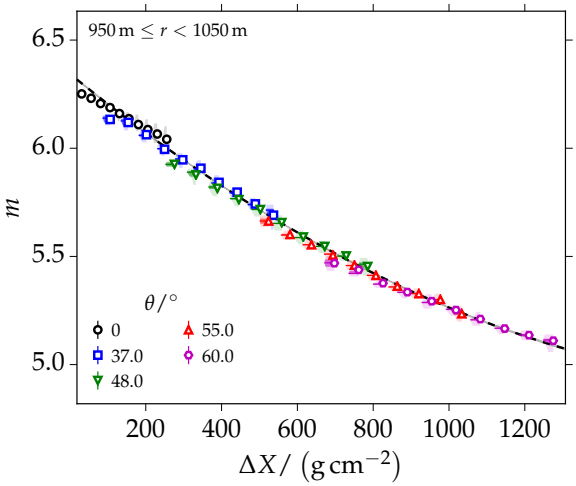
(a) m, 450 m $\leq r <$ 550 m, different energies



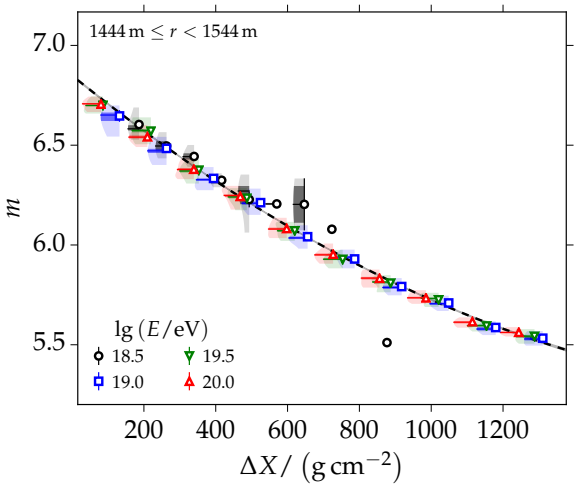
(b) m, 450 m $\leq r <$ 550 m, different zeniths



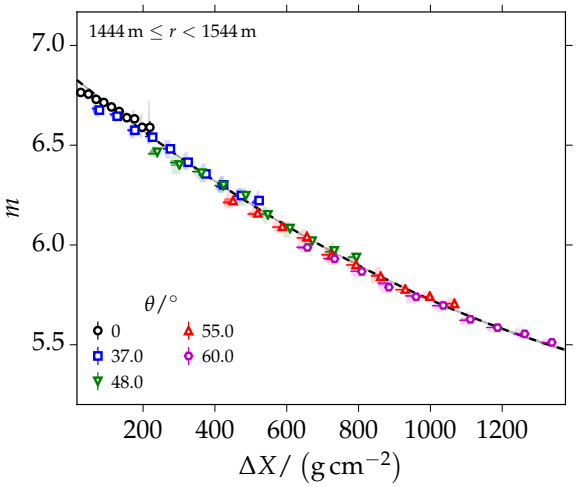
(c) m, 950 m $\leq r <$ 1050 m, different energies



(d) m, 950 m $\leq r <$ 1050 m, different zeniths

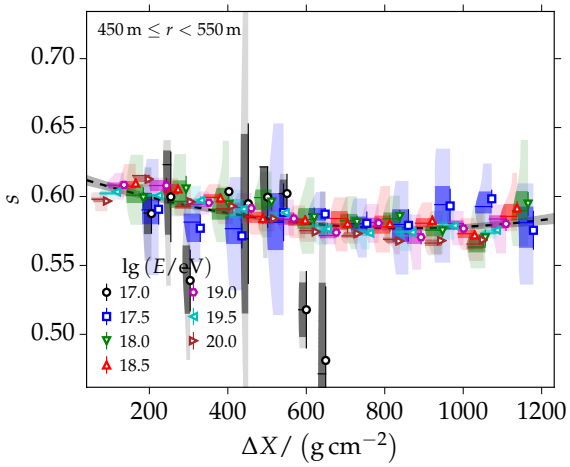


(e) m, 1444 m $\leq r <$ 1544 m, different energies

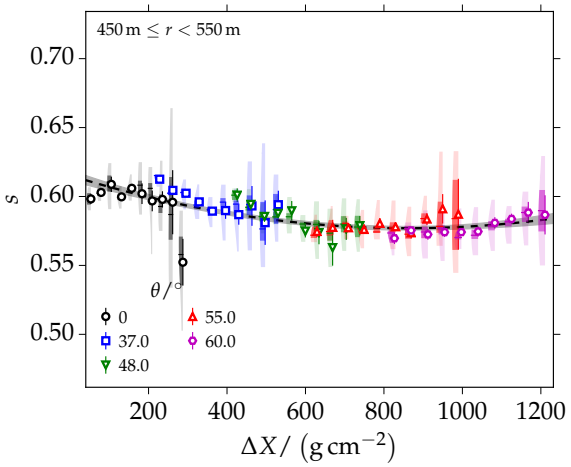


(f) m, 1444 m $\leq r <$ 1544 m, different zeniths

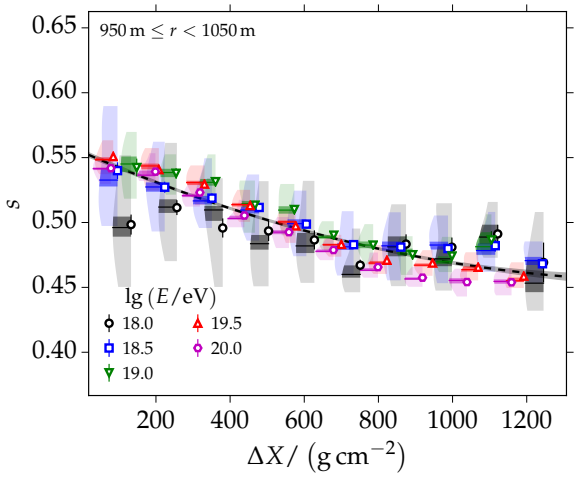
Figure B.4: Time model dependence $m(\Delta X)$ for μ.



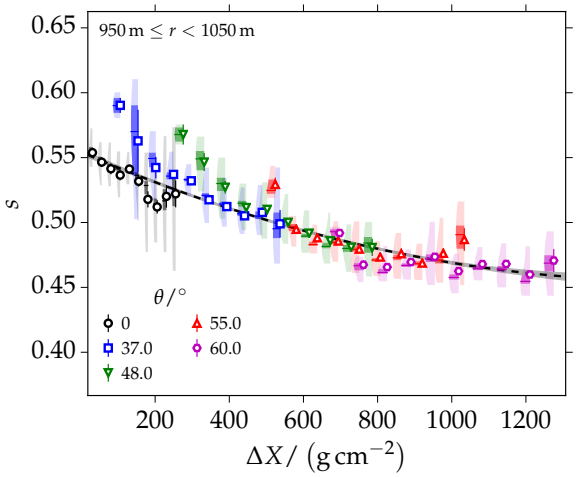
(a) s, 450 m $\leq r <$ 550 m, different energies



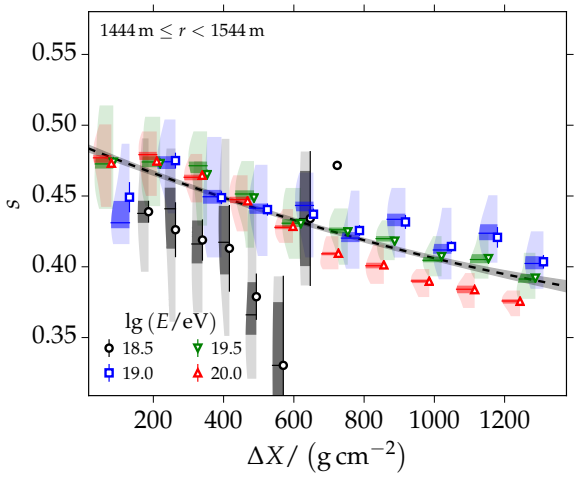
(b) s, 450 m $\leq r <$ 550 m, different zeniths



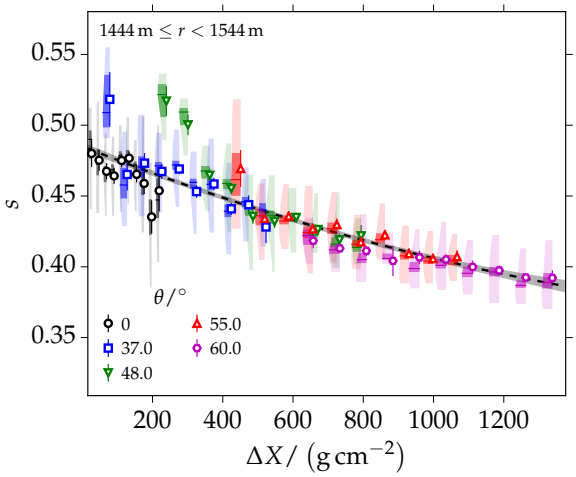
(c) s, 950 m $\leq r <$ 1050 m, different energies



(d) s, 950 m $\leq r <$ 1050 m, different zeniths

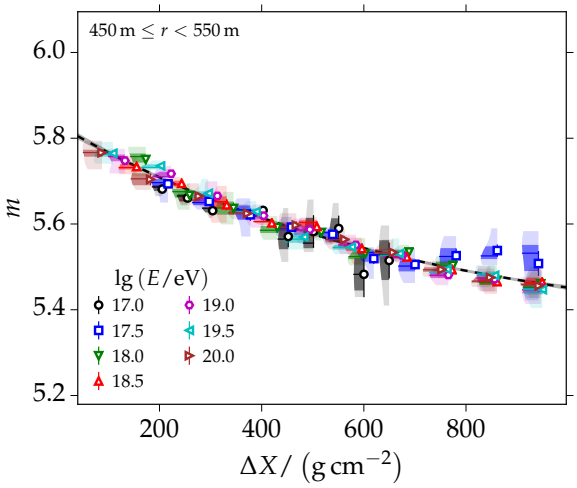


(e) s, 1444 m $\leq r <$ 1544 m, different energies

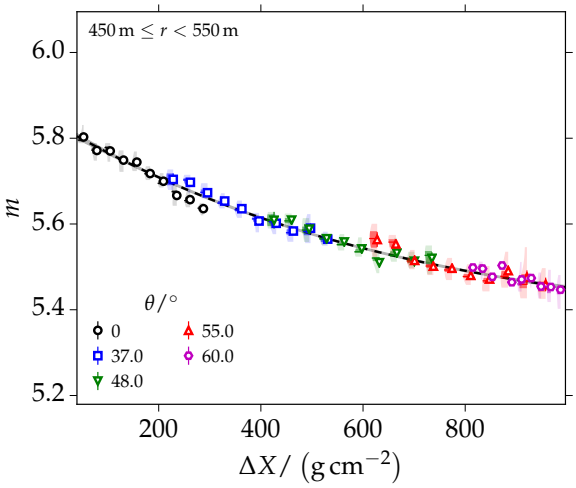


(f) s, 1444 m $\leq r <$ 1544 m, different zeniths

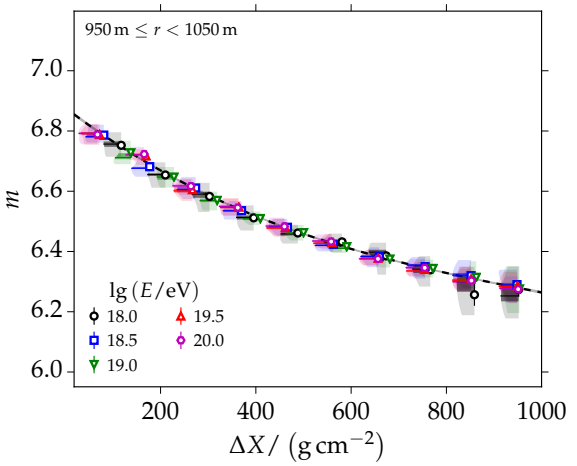
Figure B.5: Time model dependence $s(\Delta X)$ for μ.



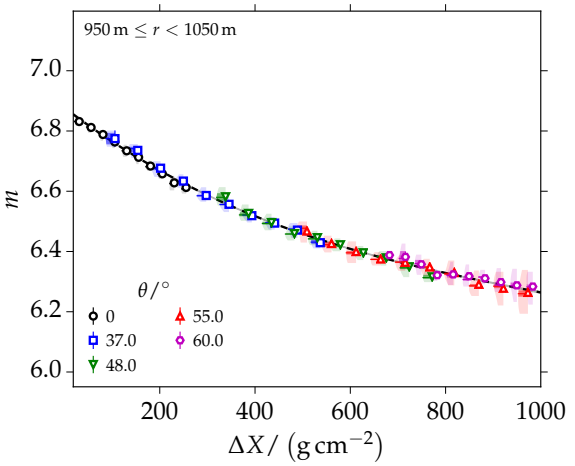
(a) m, 450 m $\leq r <$ 550 m, different energies



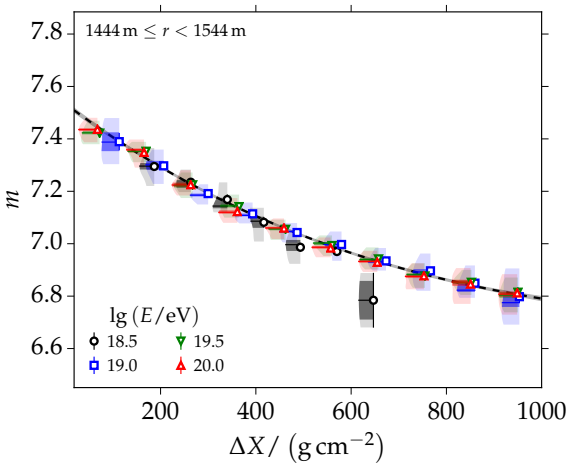
(b) m, 450 m $\leq r <$ 550 m, different zeniths



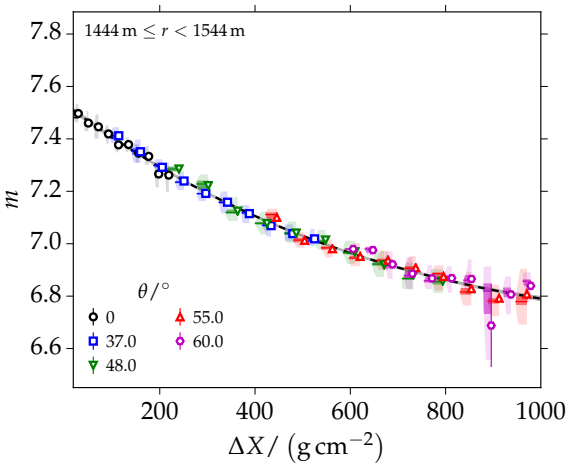
(c) m, 950 m $\leq r <$ 1050 m, different energies



(d) m, 950 m $\leq r <$ 1050 m, different zeniths

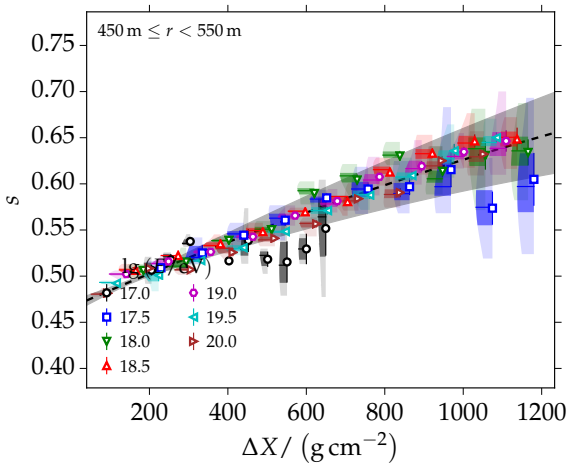


(e) m, 1444 m $\leq r <$ 1544 m, different energies

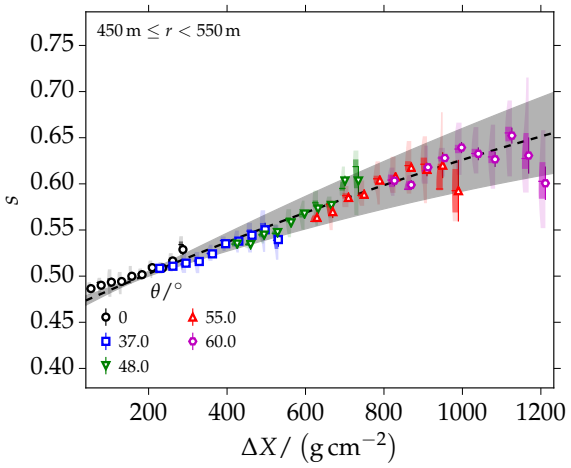


(f) m, 1444 m $\leq r <$ 1544 m, different zeniths

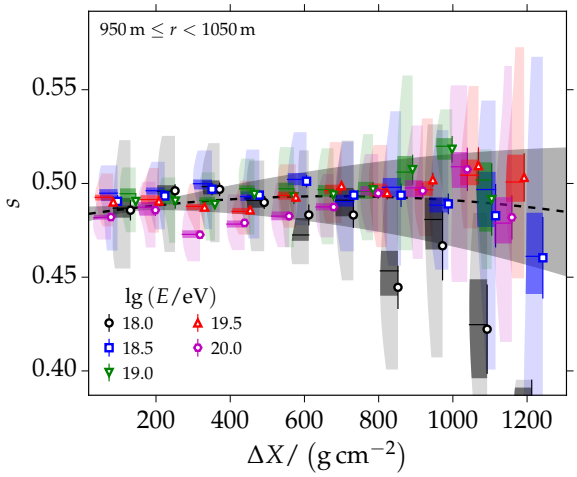
Figure B.6: Time model dependence $m(\Delta X)$ for $e\gamma$.



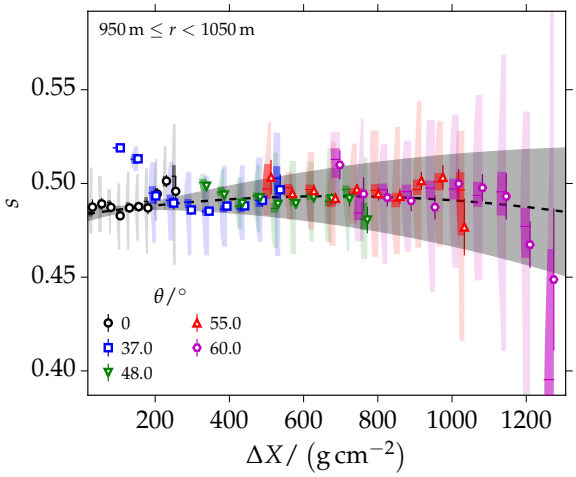
(a) s, 450 m $\leq r <$ 550 m, different energies



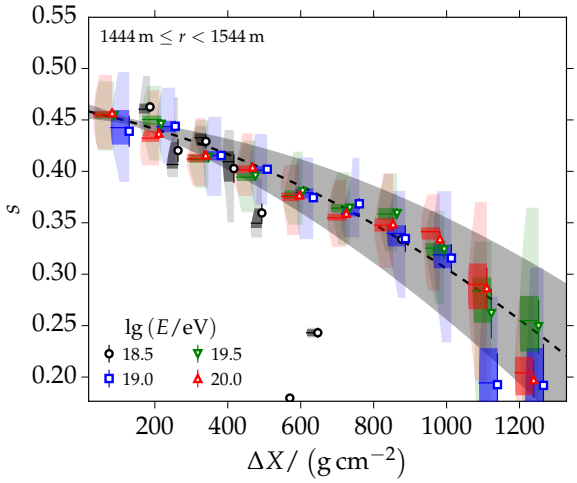
(b) s, 450 m $\leq r <$ 550 m, different zeniths



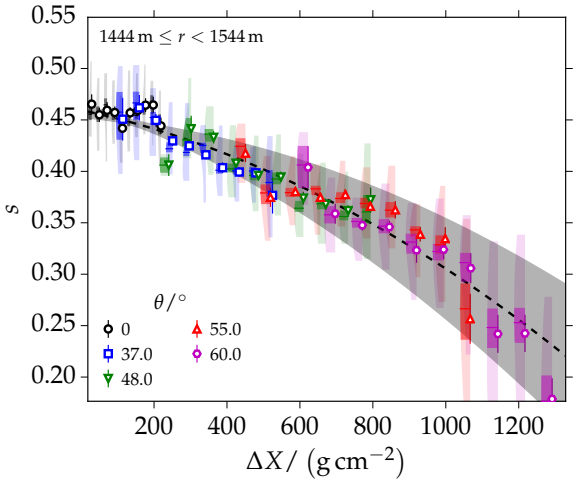
(c) s, 950 m $\leq r <$ 1050 m, different energies



(d) s, 950 m $\leq r <$ 1050 m, different zeniths

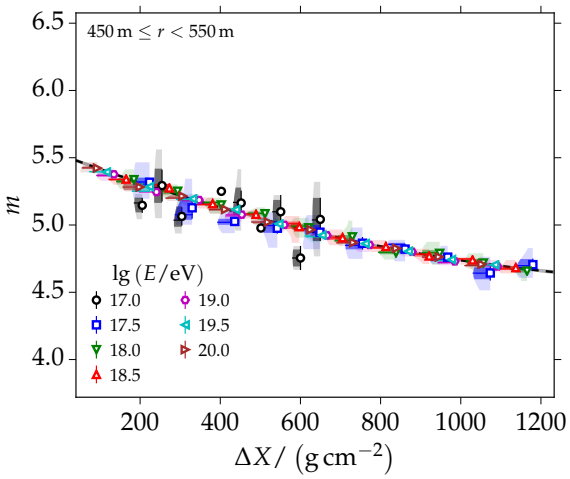


(e) s, 1444 m $\leq r <$ 1544 m, different energies

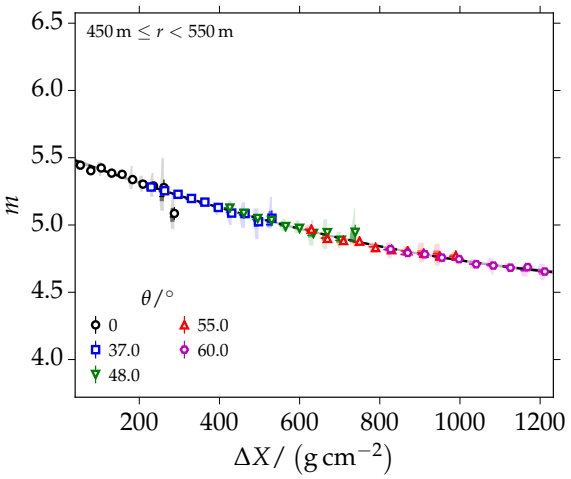


(f) s, 1444 m $\leq r <$ 1544 m, different zeniths

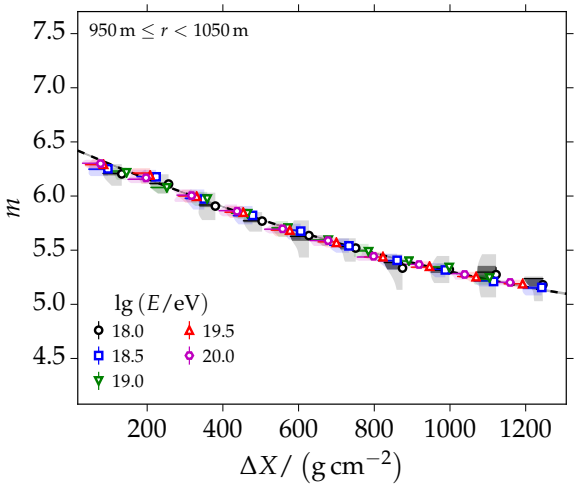
Figure B.7: Time model dependence $s(\Delta X)$ for $e\gamma$.



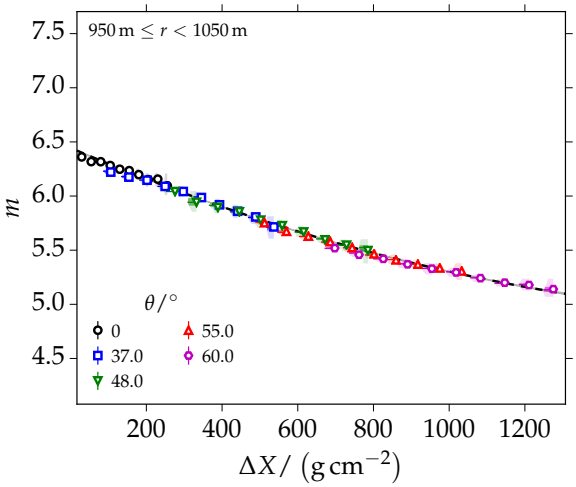
(a) m, 450 m $\leq r <$ 550 m, different energies



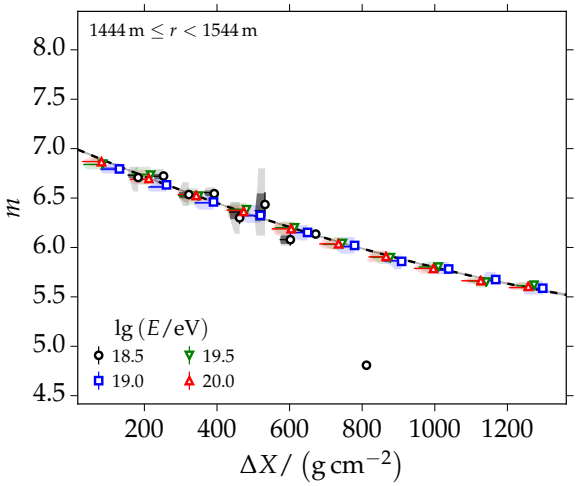
(b) m, 450 m $\leq r <$ 550 m, different zeniths



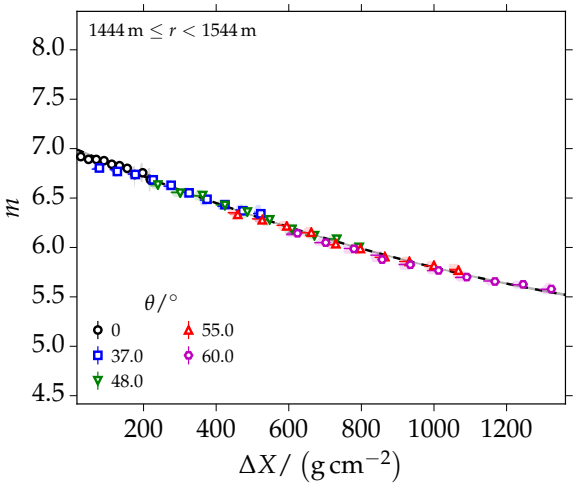
(c) m, 950 m $\leq r <$ 1050 m, different energies



(d) m, 950 m $\leq r <$ 1050 m, different zeniths

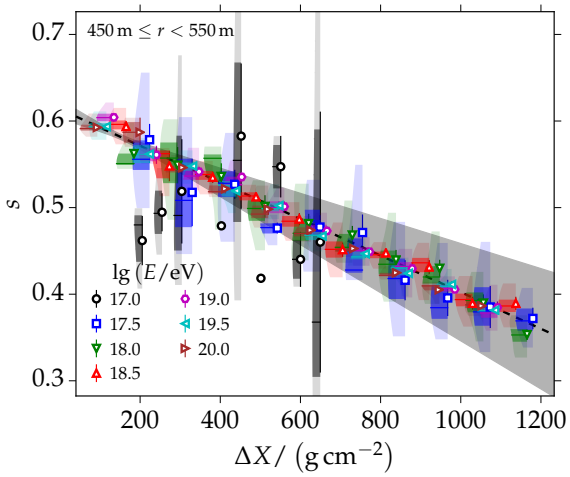


(e) m, 1444 m $\leq r <$ 1544 m, different energies

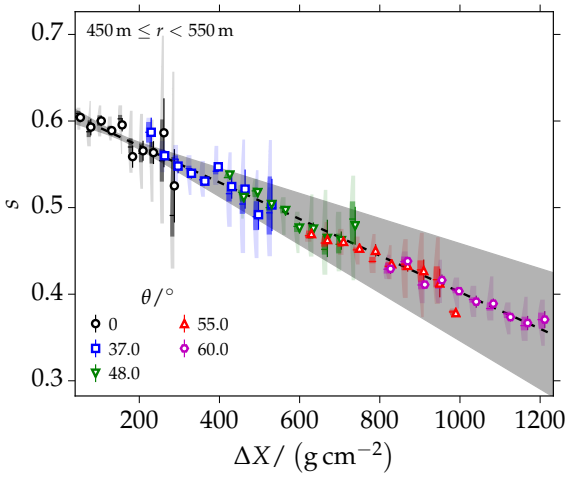


(f) m, 1444 m $\leq r <$ 1544 m, different zeniths

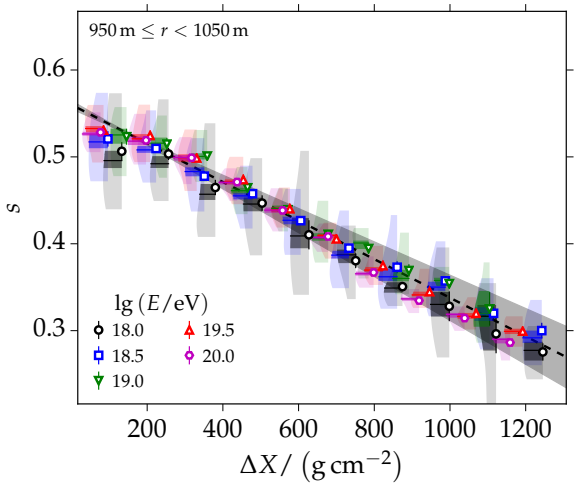
Figure B.8: Time model dependence $m(\Delta X)$ for $e\gamma(\mu)$.



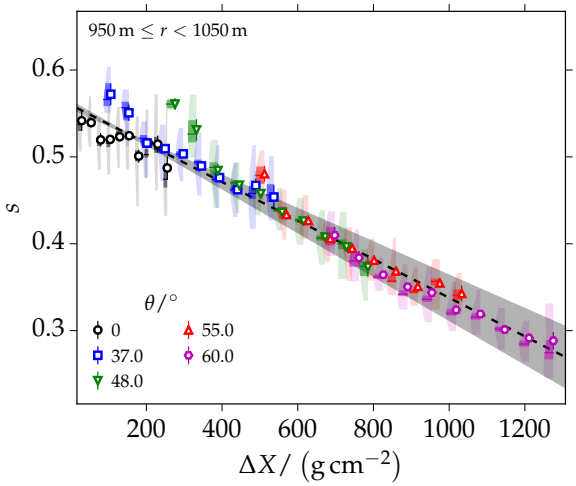
(a) s, 450 m $\leq r <$ 550 m, different energies



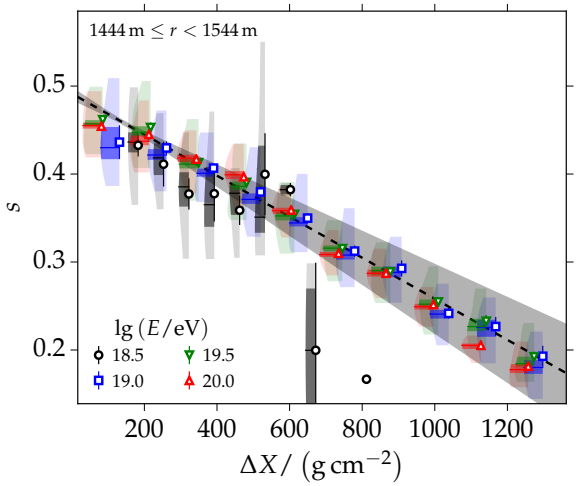
(b) s, 450 m $\leq r <$ 550 m, different zeniths



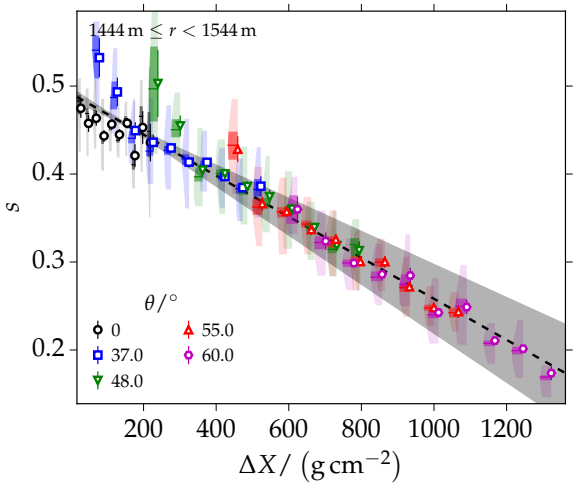
(c) s, 950 m $\leq r <$ 1050 m, different energies



(d) s, 950 m $\leq r <$ 1050 m, different zeniths

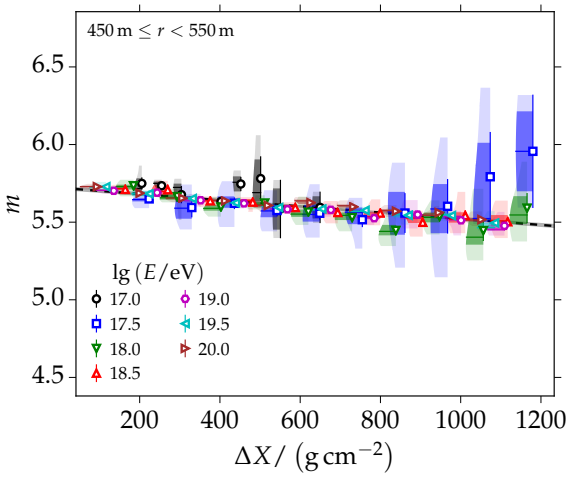


(e) s, 1444 m $\leq r <$ 1544 m, different energies

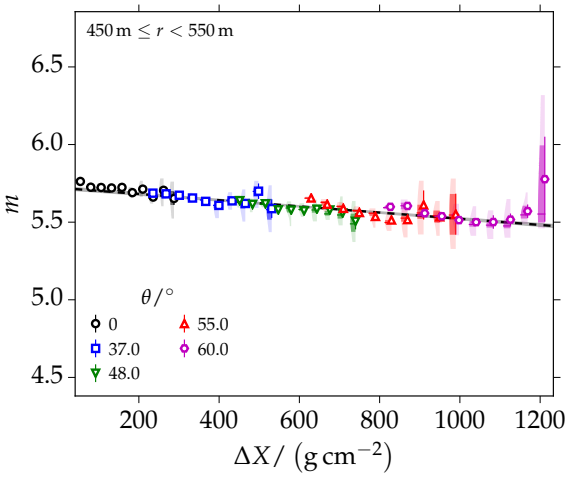


(f) s, 1444 m $\leq r <$ 1544 m, different zeniths

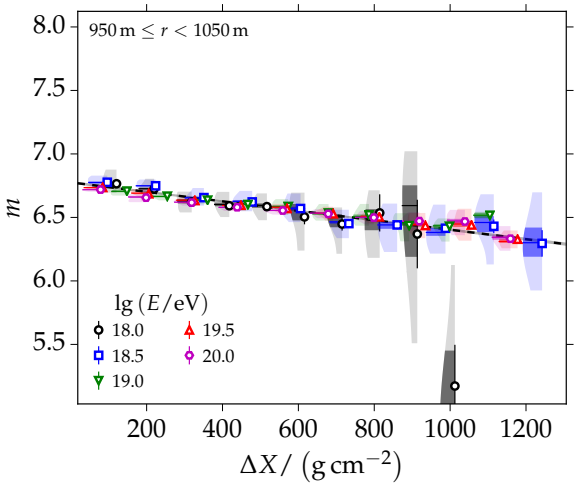
Figure B.9: Time model dependence $s(\Delta X)$ for $e\gamma(\mu)$.



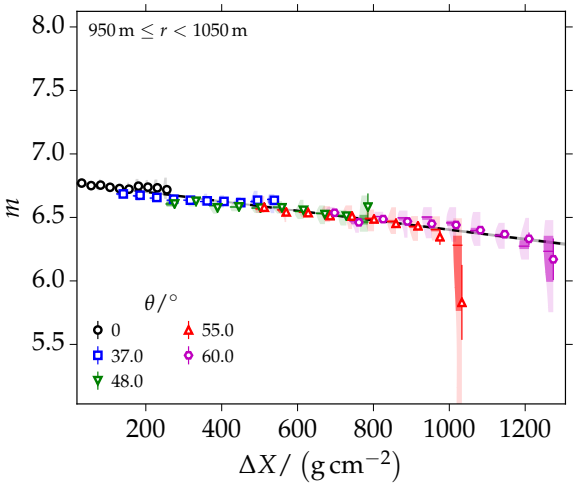
(a) m, 450 m $\leq r <$ 550 m, different energies



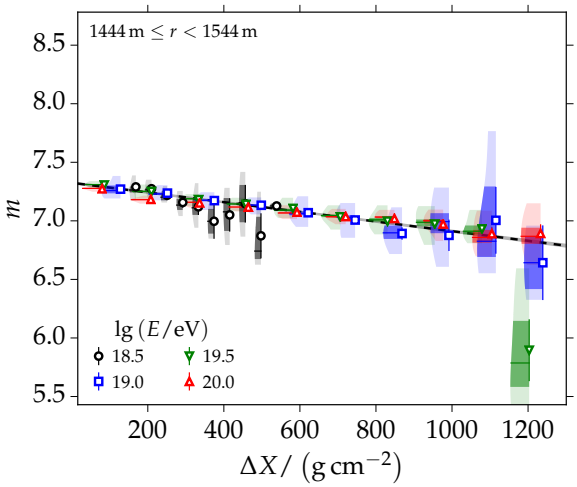
(b) m, 450 m $\leq r <$ 550 m, different zeniths



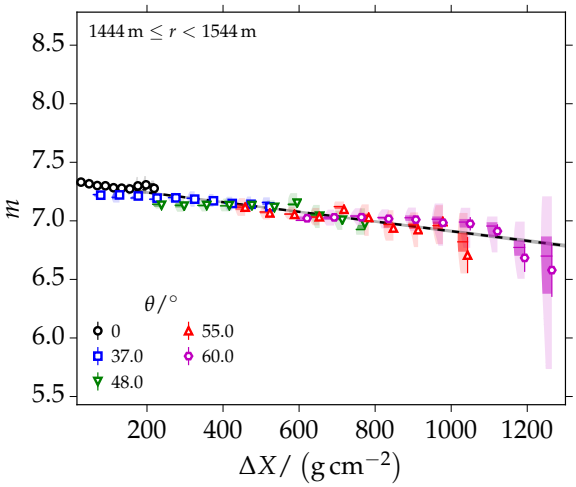
(c) m, 950 m $\leq r <$ 1050 m, different energies



(d) m, 950 m $\leq r <$ 1050 m, different zeniths

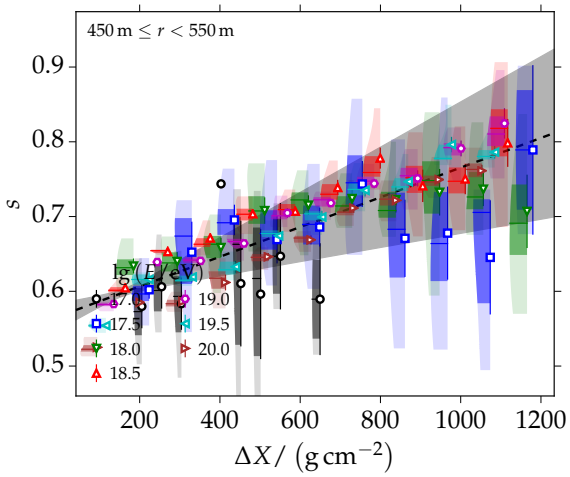


(e) m, 1444 m $\leq r <$ 1544 m, different energies

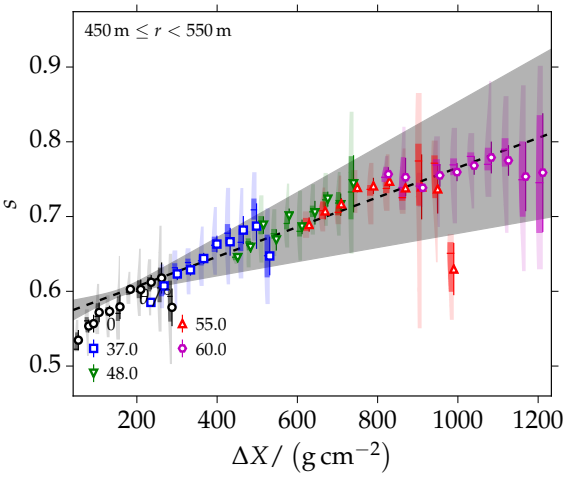


(f) m, 1444 m $\leq r <$ 1544 m, different zeniths

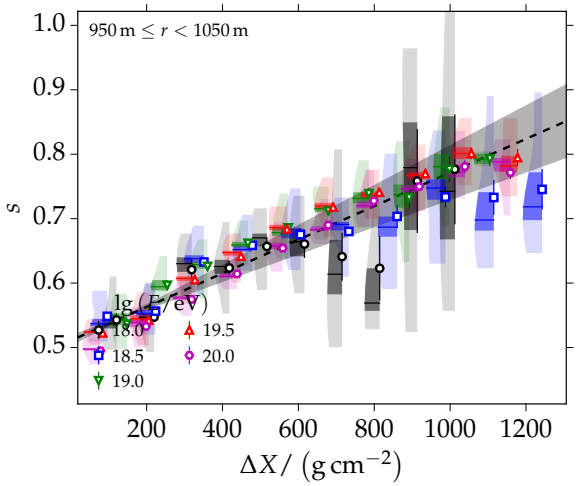
Figure B.10: Time model dependence $m(\Delta X)$ for $e\gamma$(had).



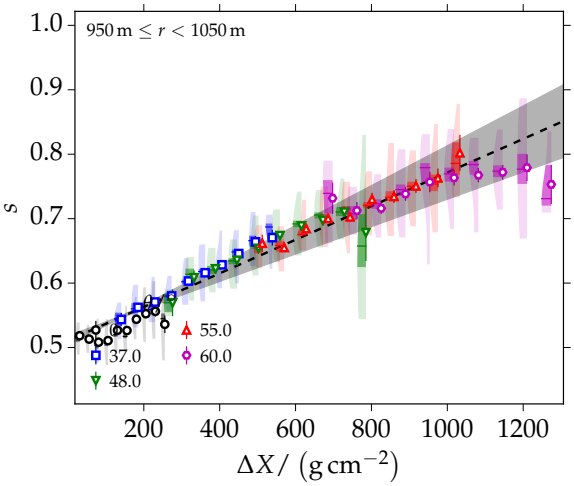
(a) s, 450 m $\leq r <$ 550 m, different energies



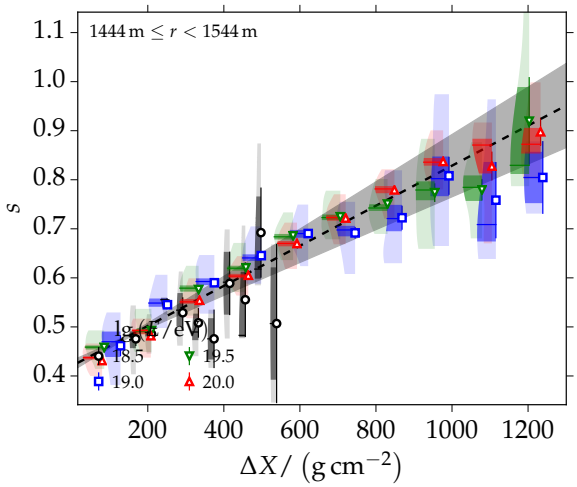
(b) s, 450 m $\leq r <$ 550 m, different zeniths



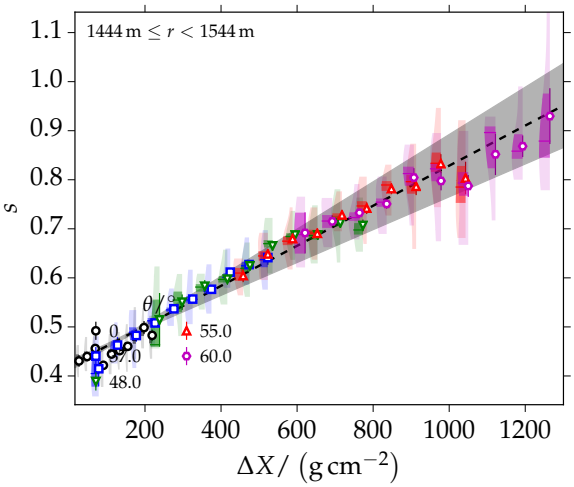
(c) s, 950 m $\leq r <$ 1050 m, different energies



(d) s, 950 m $\leq r <$ 1050 m, different zeniths



(e) s, 1444 m $\leq r <$ 1544 m, different energies



(f) s, 1444 m $\leq r <$ 1544 m, different zeniths

Figure B.11: Time model dependence $s(\Delta X)$ for $e\gamma$(had).

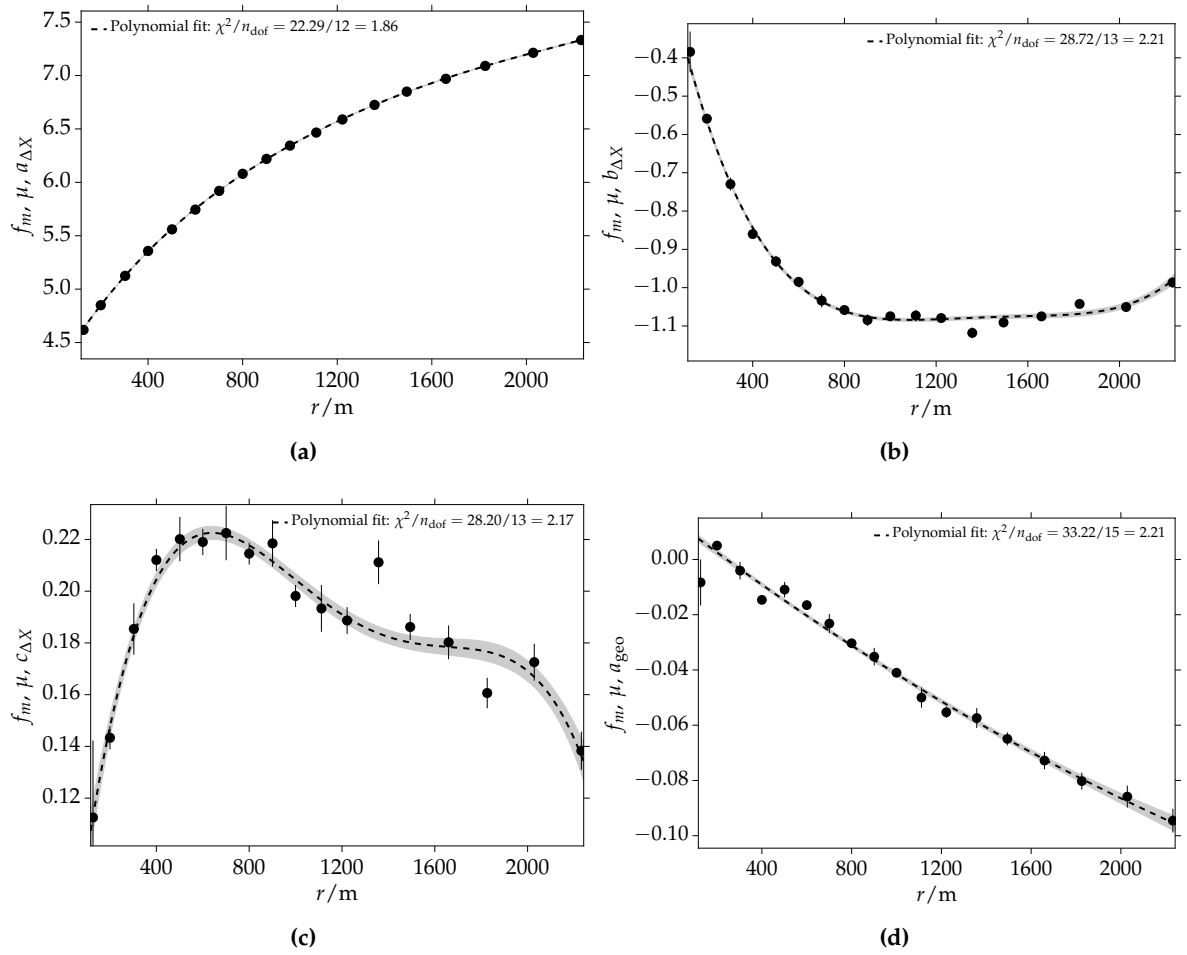


Figure B.12: Fits and profiles of the radial dependencies of the mean parameter m for the muonic component. Different plots correspond to different parameters of the model describing m. Each of the points is the result of a fit according to Eq. (2.18).

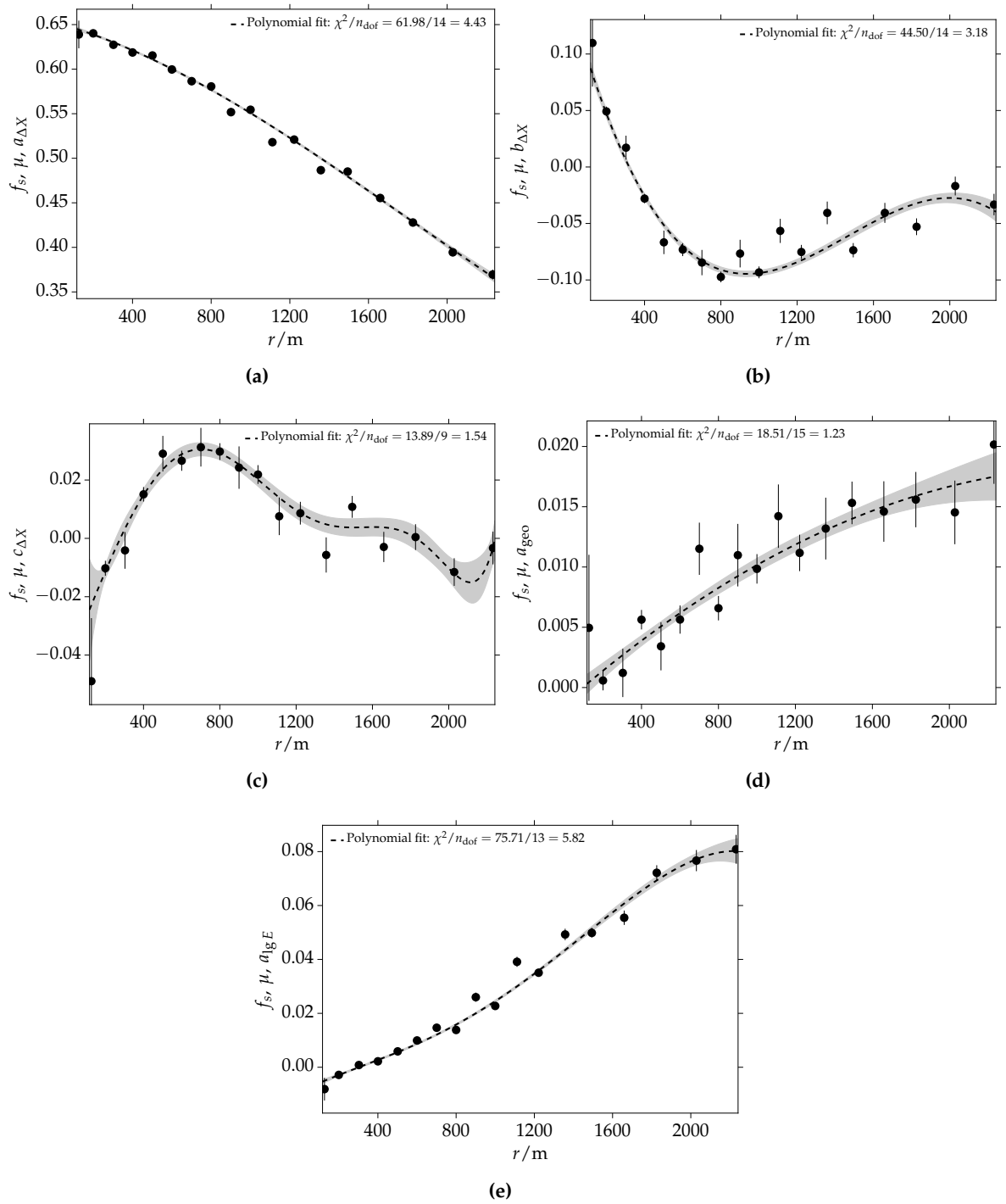


Figure B.13: Fits and profiles of the radial dependencies of the spread parameter s for the muonic component. Different plots correspond to different parameters of the model describing s. Each of the points is the result of a fit according to Eq. (2.18).

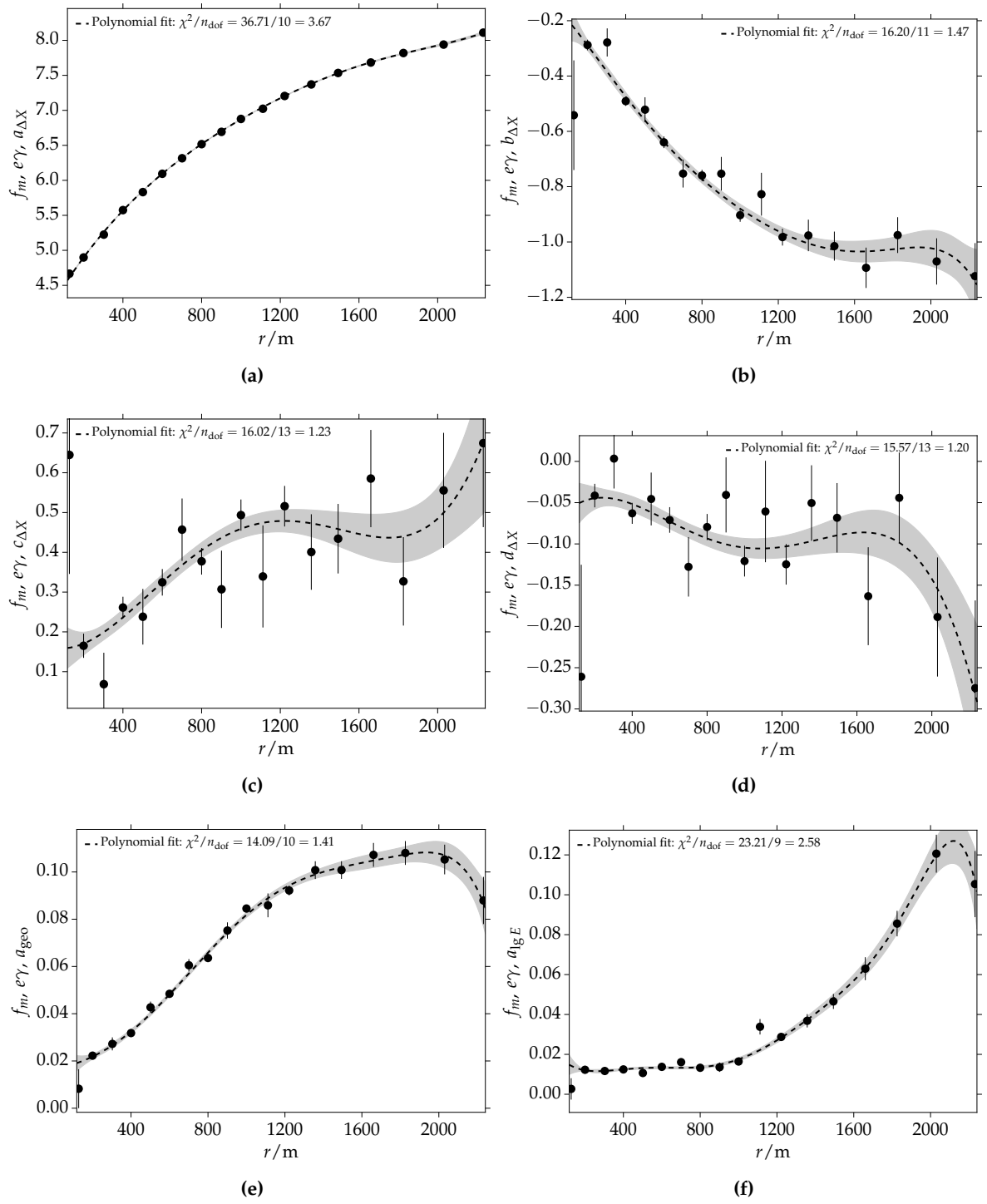


Figure B.14: Fits and profiles of the radial dependencies of the mean parameter m for the electromagnetic component. Different plots correspond to different parameters of the model describing m. Each of the points is the result of a fit according to Eq. (2.18).

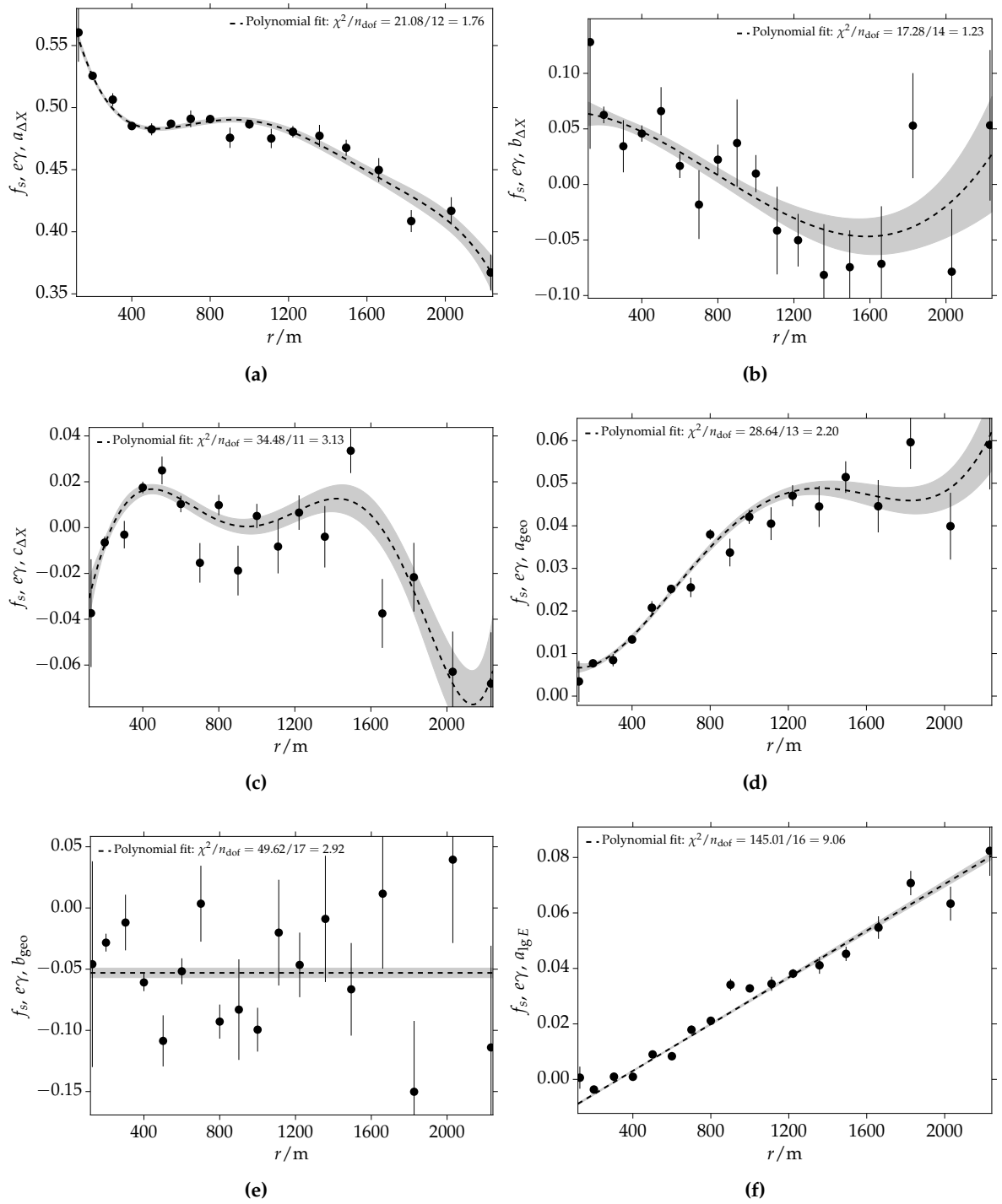


Figure B.15: Fits and profiles of the radial dependencies of the spread parameter s for the electromagnetic component. Different plots correspond to different parameters of the model describing s. Each of the points is the result of a fit according to Eq. (2.18).

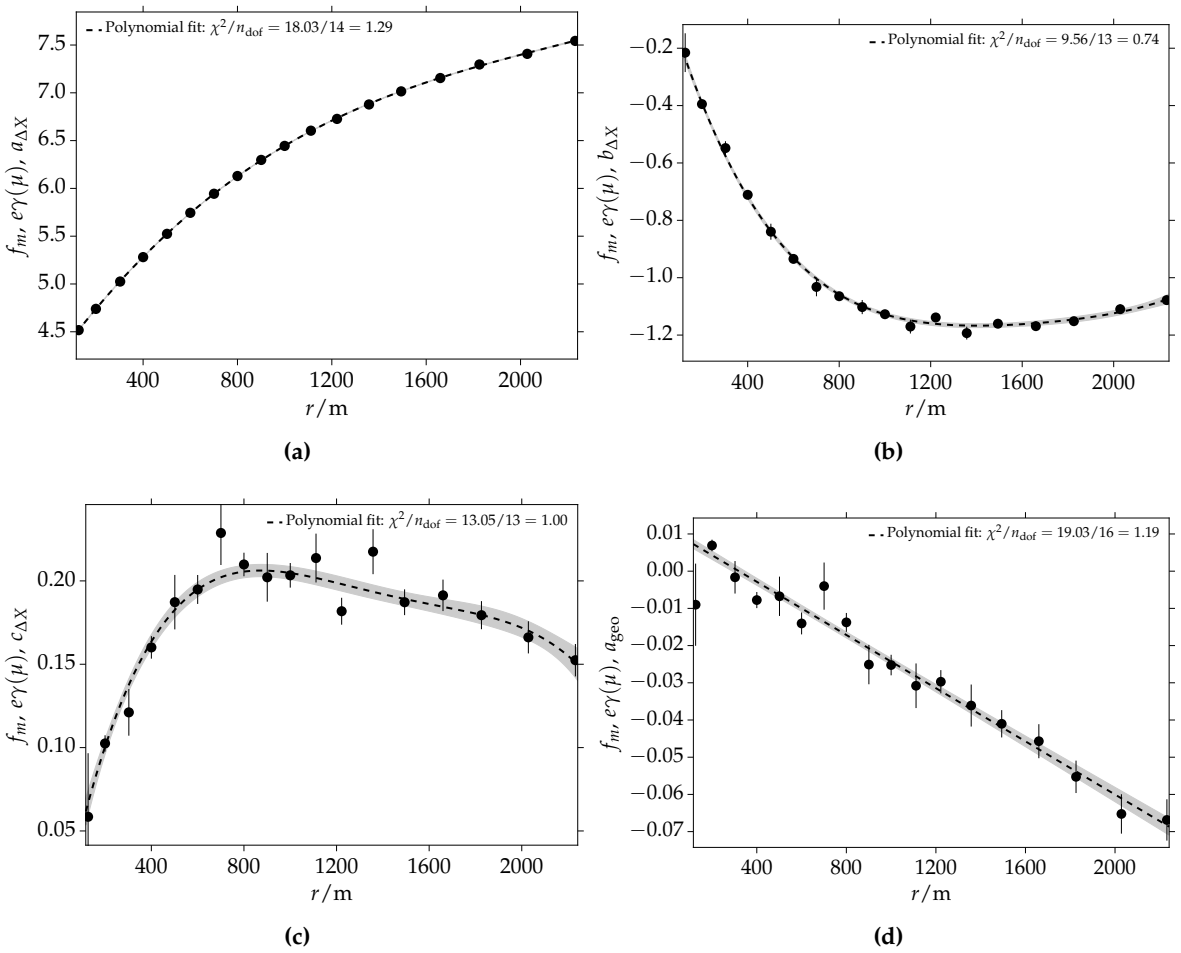


Figure B.16: Fits and profiles of the radial dependencies of the mean parameter m for the muon decay products. Different plots correspond to different parameters of the model describing m. Each of the points is the result of a fit according to Eq. (2.18).

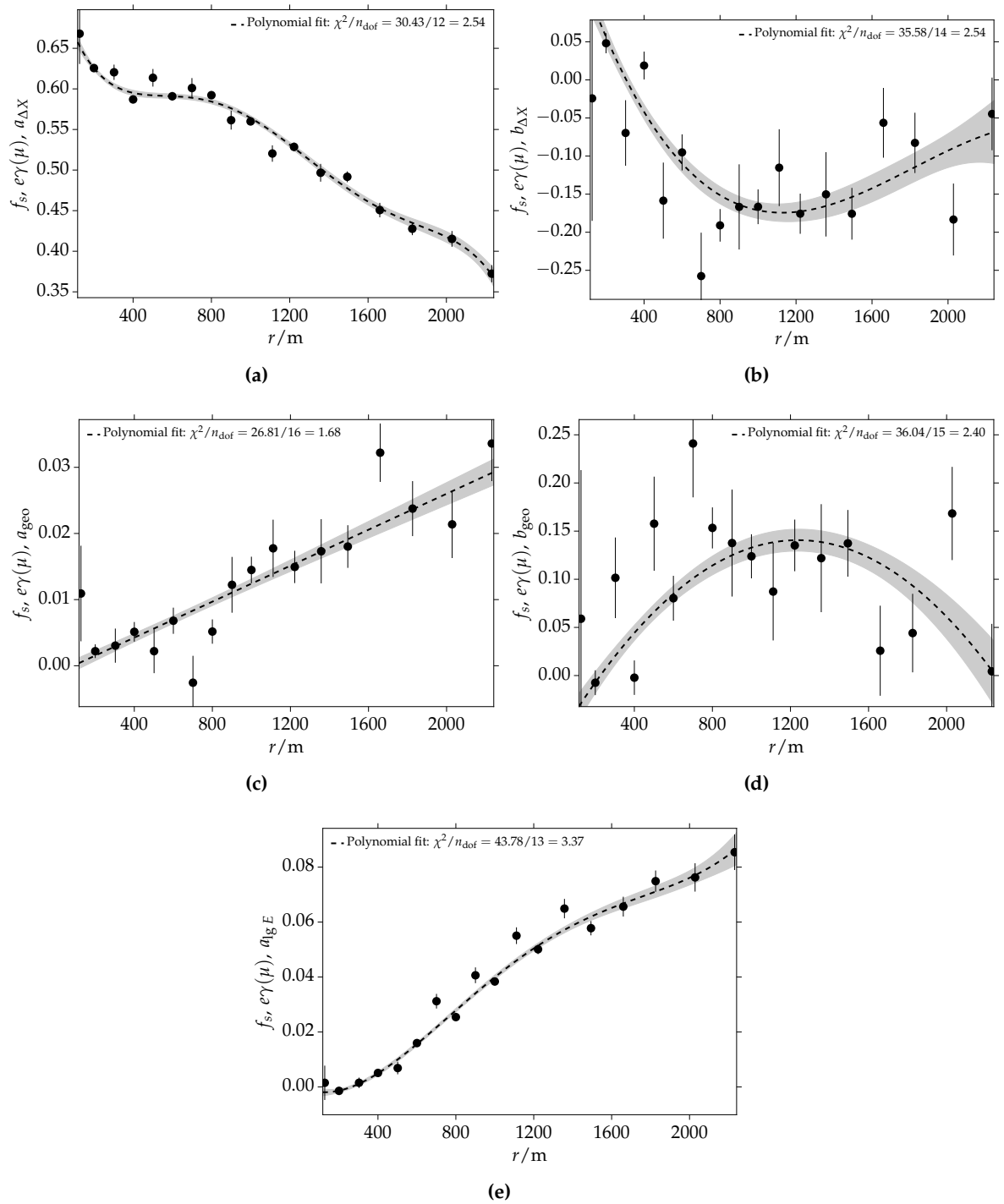


Figure B.17: Fits and profiles of the radial dependencies of the spread parameter s for the muon decay products. Different plots correspond to different parameters of the model describing s. Each of the points is the result of a fit according to Eq. (2.18).

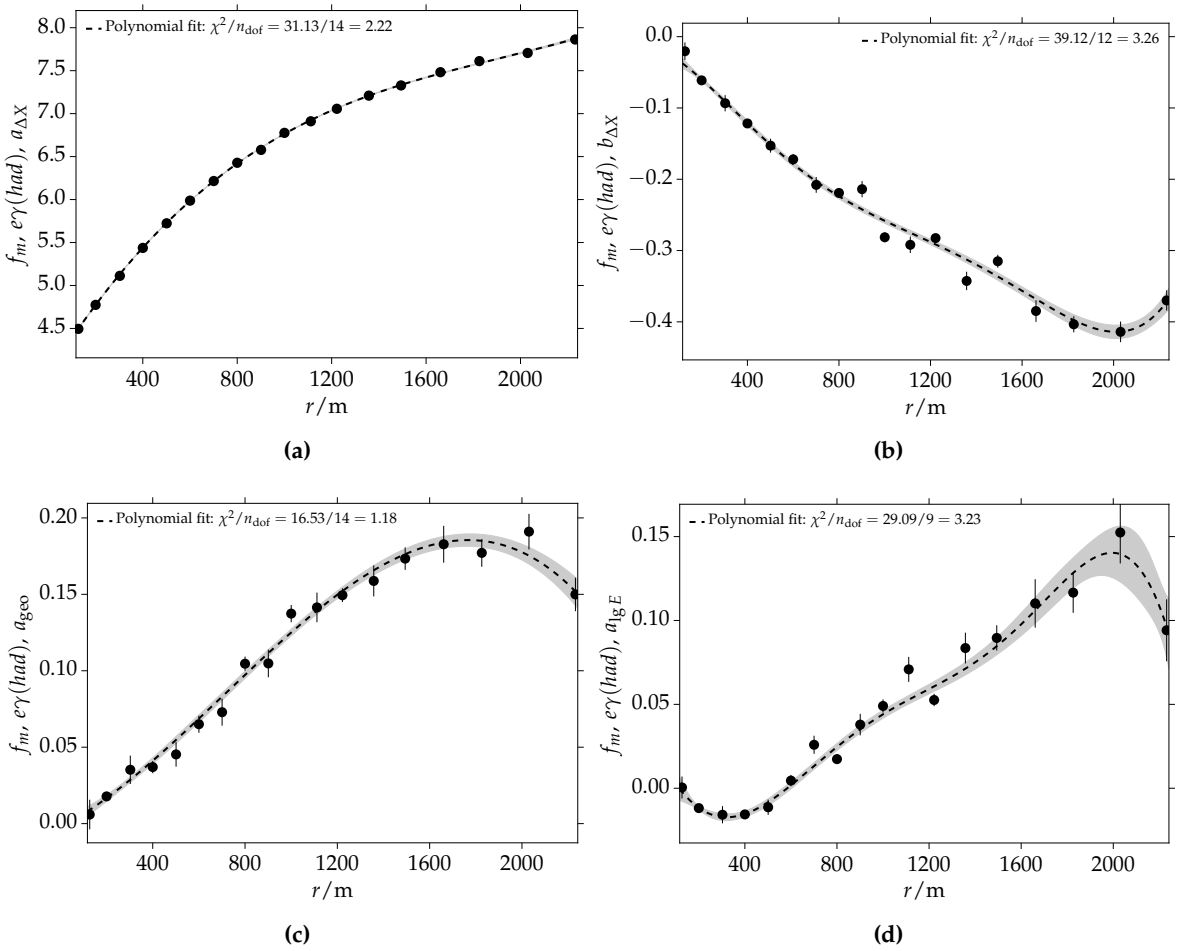


Figure B.18: Fits and profiles of the radial dependencies of the mean parameter m for the hadron jet component. Different plots correspond to different parameters of the model describing m. Each of the points is the result of a fit according to Eq. (2.18).

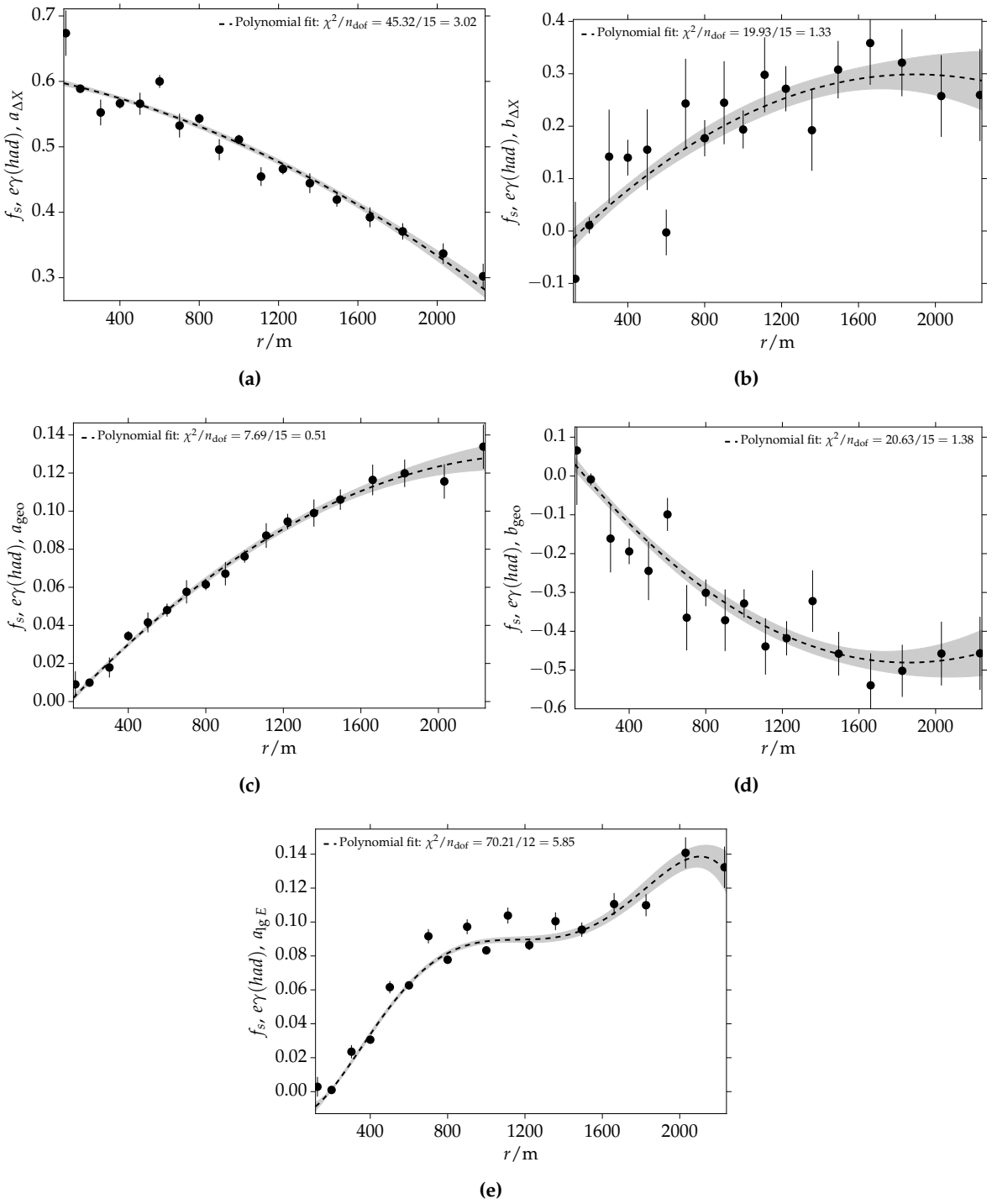


Figure B.19: Fits and profiles of the radial dependencies of the spread parameter s for the hadron jet component. Different plots correspond to different parameters of the model describing s. Each of the points is the result of a fit according to Eq. (2.18).

Table B.1: Table of radial dependencies and parameters for the mean time model. Parameters for all four particle components are given. The quantity $\hat{r}$ used in the polynomial models is defined as $\hat{r} := r/1000\,\mathrm{m}$.

Muons (μ)

$a_{\Delta X}$	$4.20+3.70\,\hat{r}-2.51\,\hat{r}^2+1.30\,\hat{r}^3-0.408\,\hat{r}^4+0.0550\,\hat{r}^5$
$b_{\Delta X}$	$-0.105-2.86\,\hat{r}+3.08\,\hat{r}^2-1.44\,\hat{r}^3+0.251\,\hat{r}^4$
$c_{\Delta X}$	$0.0299+0.781\,\hat{r}-1.07\,\hat{r}^2+0.577\,\hat{r}^3-0.109\,\hat{r}^4$
a_{geo}	$0.0147-0.0617\,\hat{r}+0.005\,58\,\hat{r}^2$

Pure electromagnetic component ($e\gamma$)

$a_{\Delta X}$	$4.05+4.69\,\hat{r}-2.23\,\hat{r}^2-0.689\,\hat{r}^3+2.07\,\hat{r}^4-1.37\,\hat{r}^5+0.378\,\hat{r}^6-0.0363\,\hat{r}^7$
$b_{\Delta X}$	$-0.128-0.609\,\hat{r}-1.50\,\hat{r}^2+3.00\,\hat{r}^3-2.52\,\hat{r}^4+1.05\,\hat{r}^5-0.172\,\hat{r}^6$
$c_{\Delta X}$	$0.170-0.244\,\hat{r}+1.48\,\hat{r}^2-1.25\,\hat{r}^3+0.305\,\hat{r}^4$
$d_{\Delta X}$	$-0.0726+0.270\,\hat{r}-0.769\,\hat{r}^2+0.626\,\hat{r}^3-0.159\,\hat{r}^4$
a_{geo}	$0.0162+0.0208\,\hat{r}+0.0121\,\hat{r}^2+0.149\,\hat{r}^3-0.142\,\hat{r}^4+0.000\,259\,\hat{r}^5+0.0348\,\hat{r}^6-0.008\,99\,\hat{r}^7$
$a_{\lg E}$	$0.0263-0.150\,\hat{r}+0.511\,\hat{r}^2-0.568\,\hat{r}^3-0.327\,\hat{r}^4+1.27\,\hat{r}^5-1.11\,\hat{r}^6+0.414\,\hat{r}^7-0.0582\,\hat{r}^8$

Muon decay products ($e\gamma(\mu)$)

$a_{\Delta X}$	$4.10+3.42\,\hat{r}-1.28\,\hat{r}^2+0.195\,\hat{r}^3$
$b_{\Delta X}$	$0.0880-2.84\,\hat{r}+2.41\,\hat{r}^2-0.922\,\hat{r}^3+0.138\,\hat{r}^4$
$c_{\Delta X}$	$-0.008\,26+0.683\,\hat{r}-0.769\,\hat{r}^2+0.361\,\hat{r}^3-0.0624\,\hat{r}^4$
a_{geo}	$0.0114-0.0357\,\hat{r}$

Hadron jets ($e\gamma(\mathrm{had})$)

$a_{\Delta X}$	$3.97+4.40\,\hat{r}-1.95\,\hat{r}^2+0.339\,\hat{r}^3$
$b_{\Delta X}$	$-0.0154-0.124\,\hat{r}-0.675\,\hat{r}^2+1.01\,\hat{r}^3-0.569\,\hat{r}^4+0.111\,\hat{r}^5$
a_{geo}	$-0.002\,60+0.0800\,\hat{r}+0.0902\,\hat{r}^2-0.0426\,\hat{r}^3$
$a_{\lg E}$	$0.0276-0.304\,\hat{r}+0.561\,\hat{r}^2-0.0691\,\hat{r}^3-0.307\,\hat{r}^4+0.0481\,\hat{r}^5+0.163\,\hat{r}^6-0.0889\,\hat{r}^7+0.0130\,\hat{r}^8$

Table B.2: Table of radial dependencies and parameters for the model of the spread of time distributions. Parameters for all four particle components are given. The quantity $\hat{r}$ used in the polynomial models is defined as $\hat{r} := r/1000\,\mathrm{m}$.

Muons (μ)	
$a_{\Delta X}$	$0.653 - 0.0606\,\hat{r} - 0.0507\,\hat{r}^2 + 0.009\,11\,\hat{r}^3$
$b_{\Delta X}$	$0.154 - 0.629\,\hat{r} + 0.493\,\hat{r}^2 - 0.112\,\hat{r}^3$
$c_{\Delta X}$	$-0.0646 + 0.339\,\hat{r} - 0.413\,\hat{r}^2 + 0.190\,\hat{r}^3 - 0.0305\,\hat{r}^4$
a_{geo}	$-0.001\,31 + 0.0140\,\hat{r} - 0.002\,53\,\hat{r}^2$
$a_{\lg E}$	$-0.009\,17 + 0.0353\,\hat{r} - 0.0264\,\hat{r}^2 + 0.0345\,\hat{r}^3 - 0.009\,70\,\hat{r}^4$

Pure electromagnetic component ($e\gamma$)	
$a_{\Delta X}$	$0.642 - 0.871\,\hat{r} + 1.74\,\hat{r}^2 - 1.54\,\hat{r}^3 + 0.620\,\hat{r}^4 - 0.0939\,\hat{r}^5$
$b_{\Delta X}$	$0.0653 + 0.000\,089\,6\,\hat{r} - 0.134\,\hat{r}^2 + 0.0565\,\hat{r}^3$
$c_{\Delta X}$	$-0.0830 + 0.543\,\hat{r} - 0.896\,\hat{r}^2 + 0.248\,\hat{r}^3 + 0.506\,\hat{r}^4 - 0.399\,\hat{r}^5 + 0.0813\,\hat{r}^6$
a_{geo}	$0.009\,15 - 0.0405\,\hat{r} + 0.181\,\hat{r}^2 - 0.138\,\hat{r}^3 + 0.0313\,\hat{r}^4$
b_{geo}	-0.0530
$a_{\lg E}$	$-0.0138 + 0.0421\,\hat{r}$

Muon decay products ($e\gamma(\mu)$)	
$a_{\Delta X}$	$0.631 - 0.0404\,\hat{r} - 0.0353\,\hat{r}^2$
$b_{\Delta X}$	$0.189 - 0.756\,\hat{r} + 0.483\,\hat{r}^2 - 0.0879\,\hat{r}^3$
a_{geo}	$-0.001\,19 + 0.0136\,\hat{r}$
b_{geo}	$-0.0701 + 0.341\,\hat{r} - 0.137\,\hat{r}^2$
$a_{\lg E}$	$0.000\,051\,1 - 0.0327\,\hat{r} + 0.146\,\hat{r}^2 - 0.0917\,\hat{r}^3 + 0.0181\,\hat{r}^4$

Hadron jets ($e\gamma$(had))	
$a_{\Delta X}$	$0.605 - 0.0631\,\hat{r} - 0.0363\,\hat{r}^2$
$b_{\Delta X}$	$-0.0560 + 0.374\,\hat{r} - 0.0988\,\hat{r}^2$
a_{geo}	$-0.0106 + 0.110\,\hat{r} - 0.0217\,\hat{r}^2$
b_{geo}	$0.101 - 0.626\,\hat{r} + 0.169\,\hat{r}^2$
$a_{\lg E}$	$-0.0140 - 0.0182\,\hat{r} + 0.637\,\hat{r}^2 - 0.909\,\hat{r}^3 + 0.479\,\hat{r}^4 - 0.0861\,\hat{r}^5$

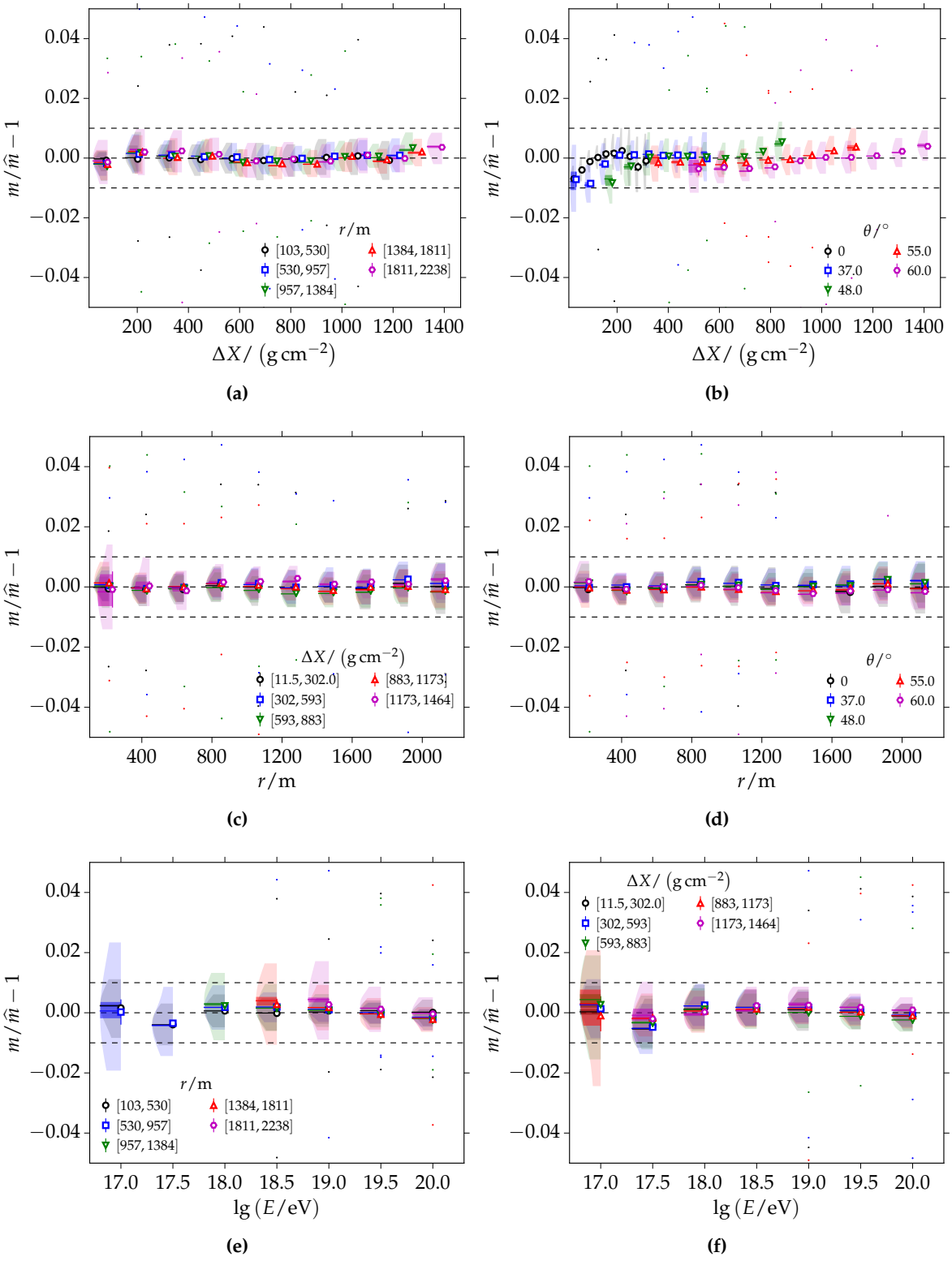


Figure B.20: Residuals of the mean parameter m for the muonic component.

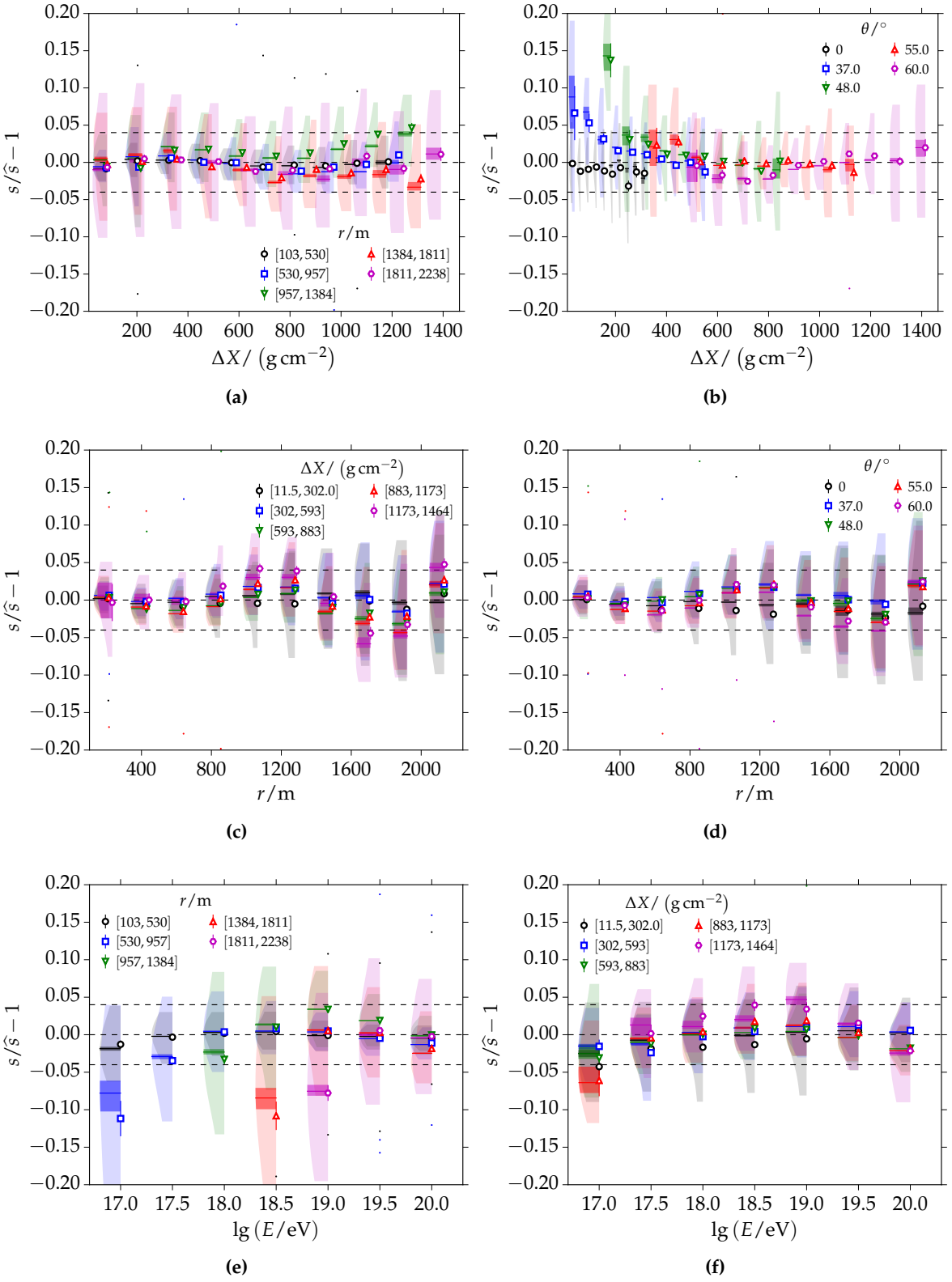


Figure B.21: Residuals of the width parameter s for the muonic component.

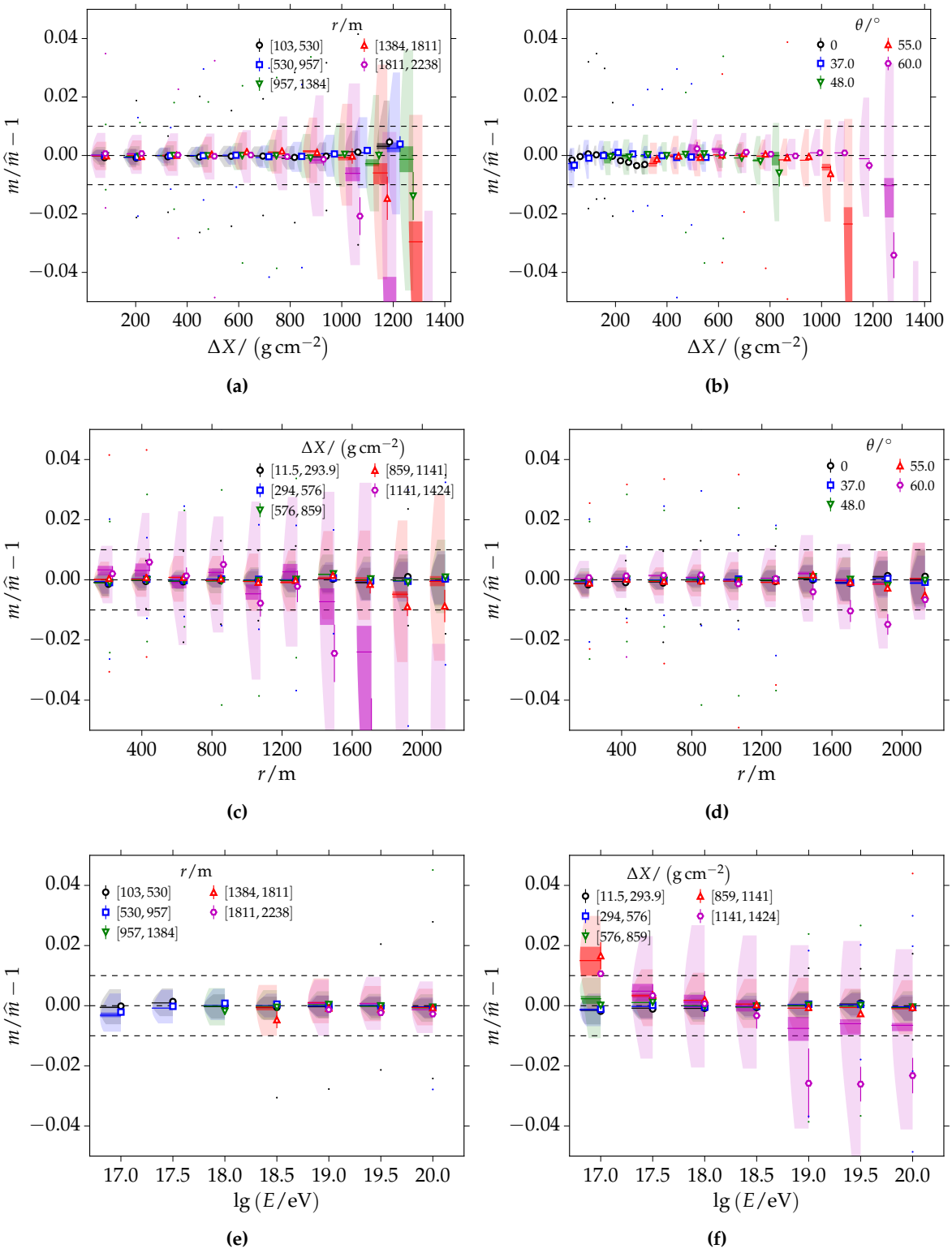


Figure B.22: Residuals of the mean parameter m for the pure electromagnetic component.

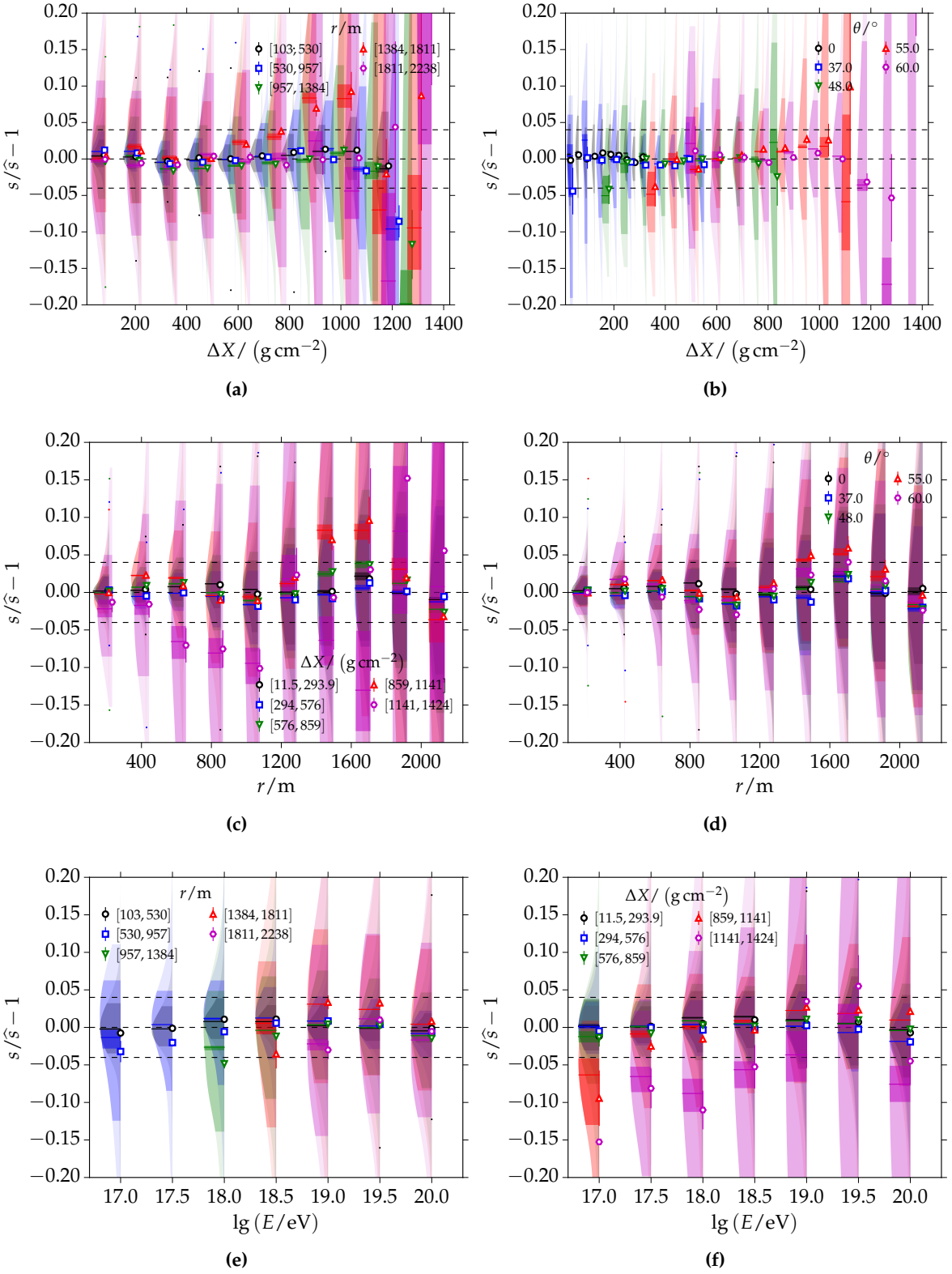


Figure B.23: Residuals of the width parameter s for the pure electromagnetic component.

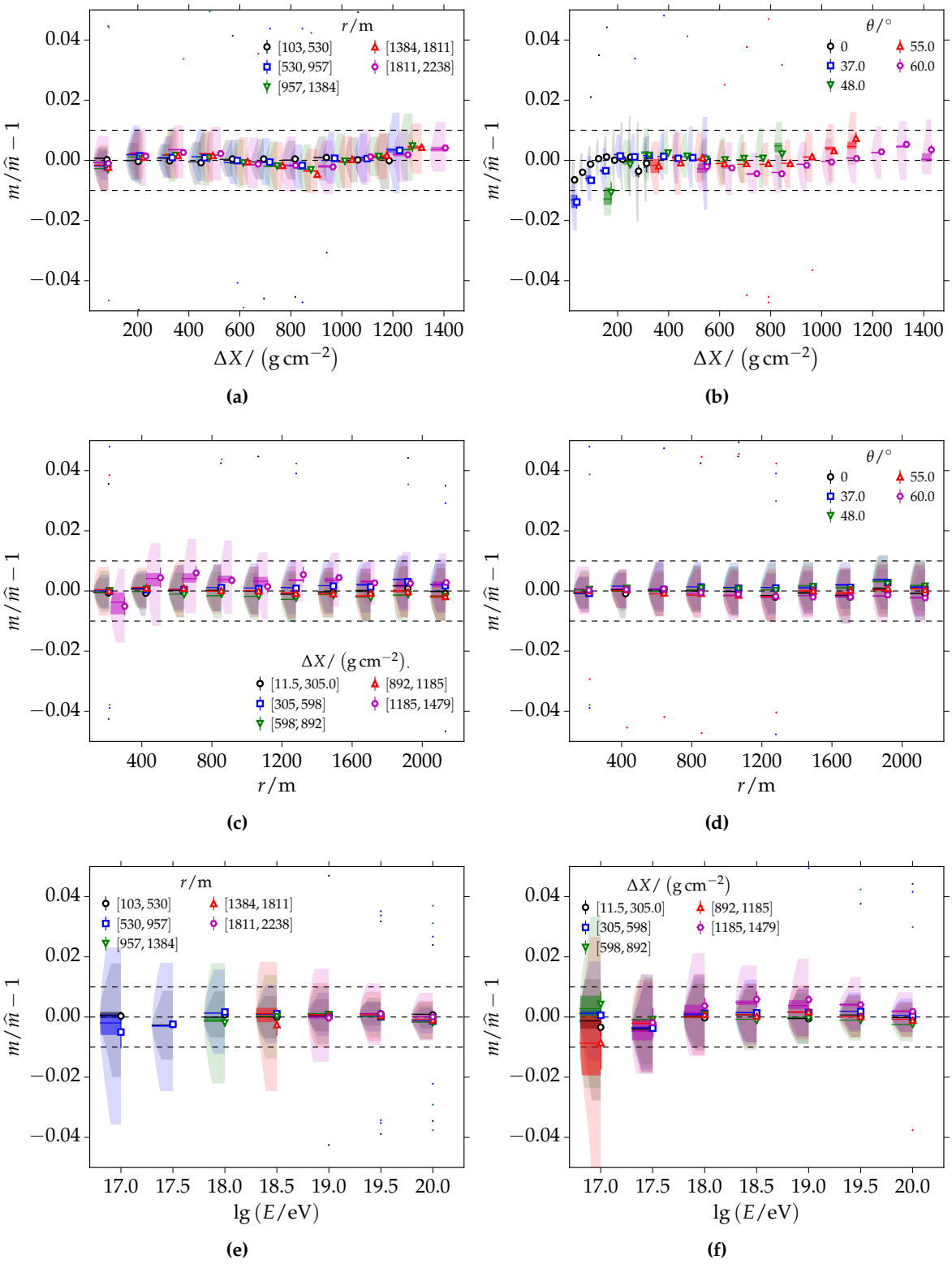


Figure B.24: Residuals of the mean parameter m for the muon decay products.

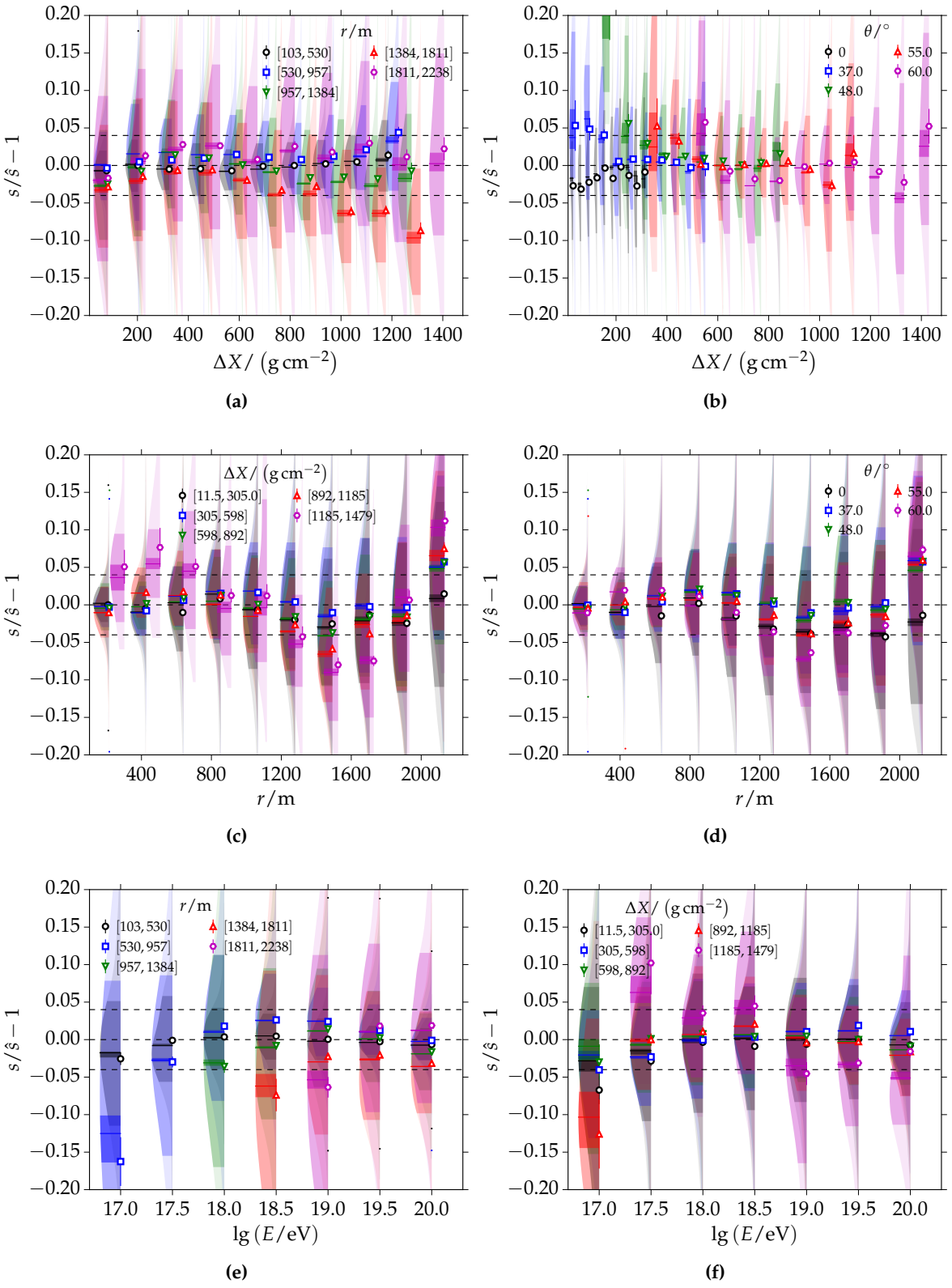


Figure B.25: Residuals of the width parameter s for the muon decay products.

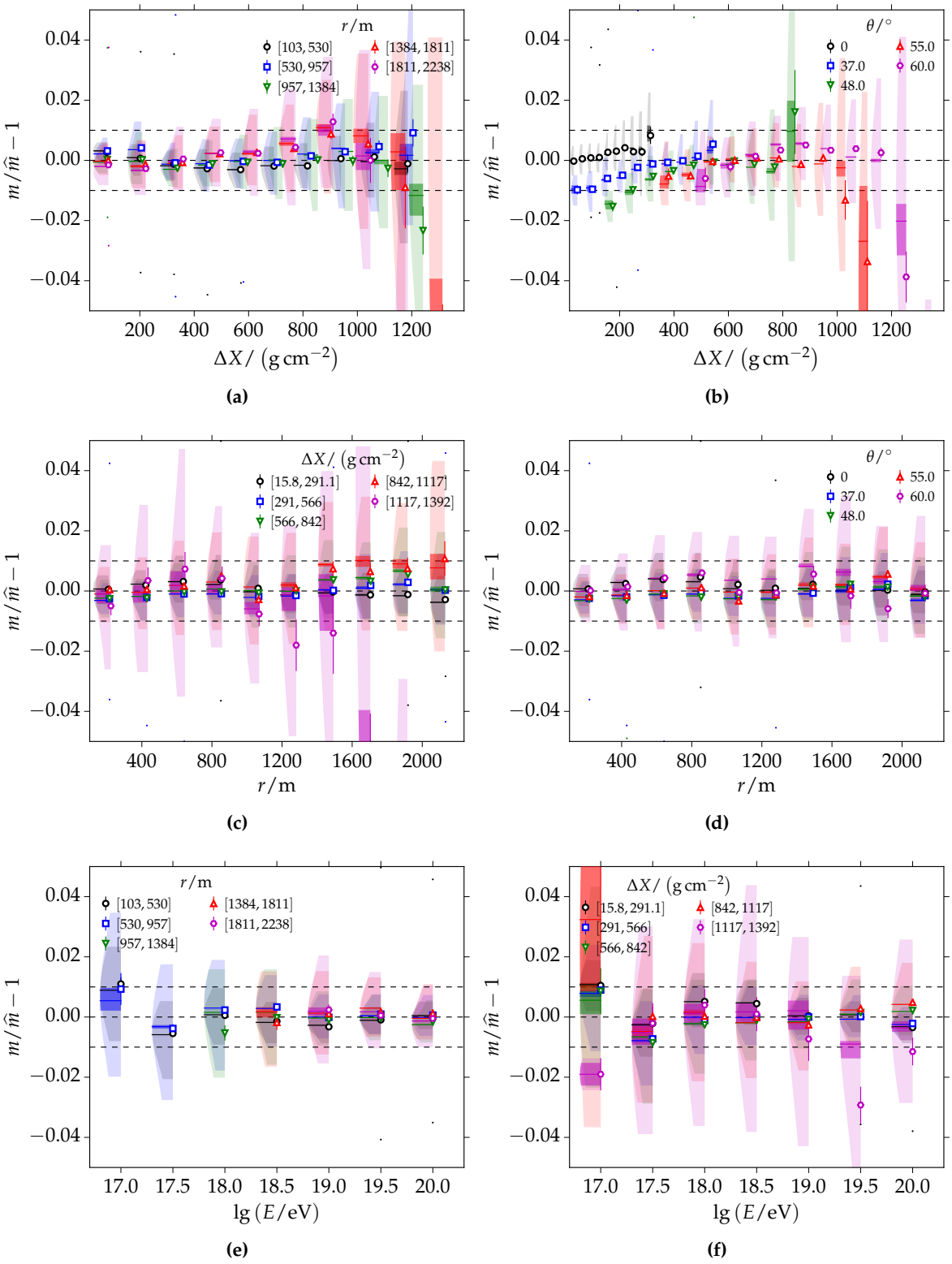


Figure B.26: Residuals of the mean parameter m for the hadron jet component.

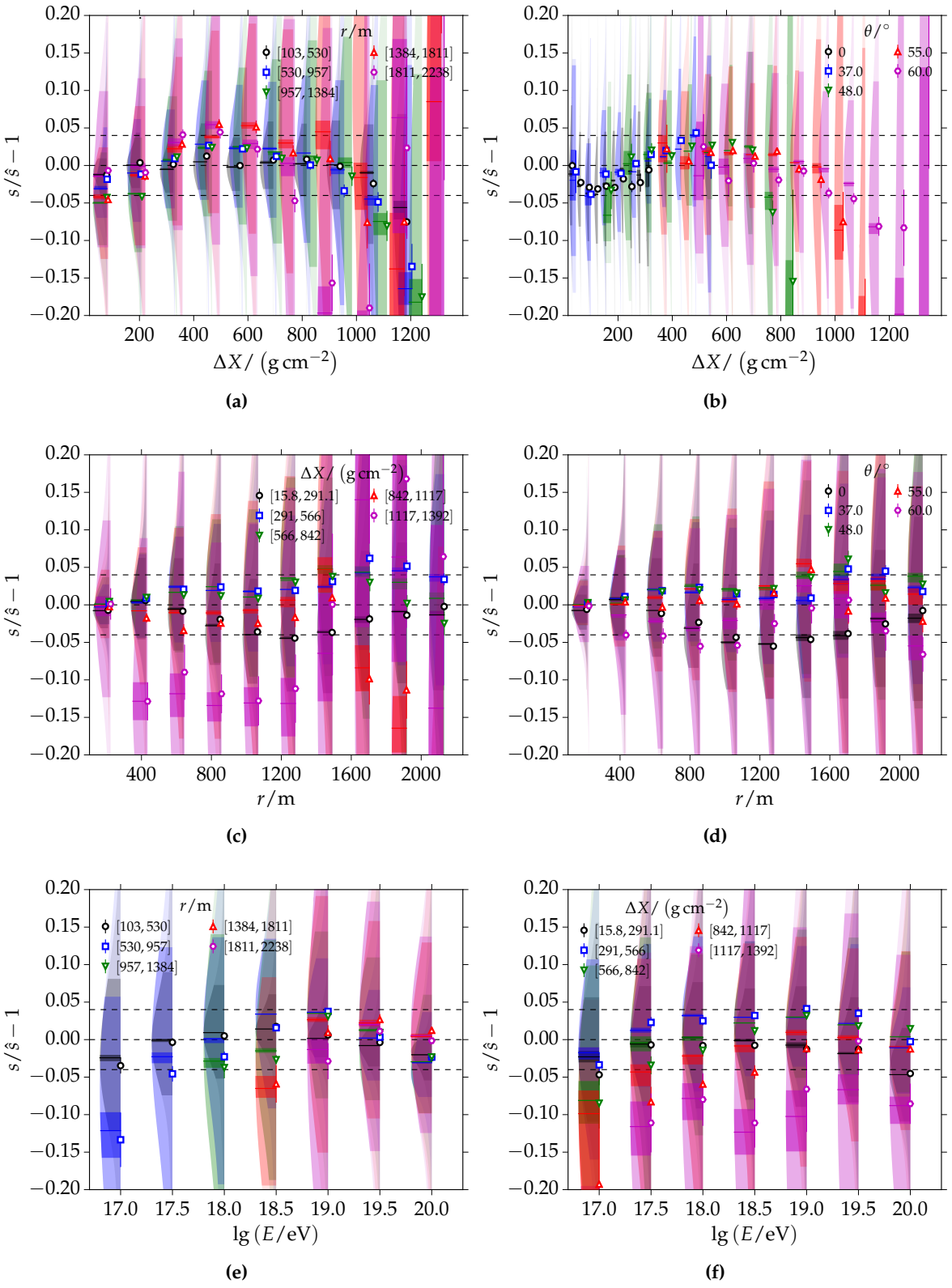


Figure B.27: Residuals of the width parameter s for the hadron jet component.

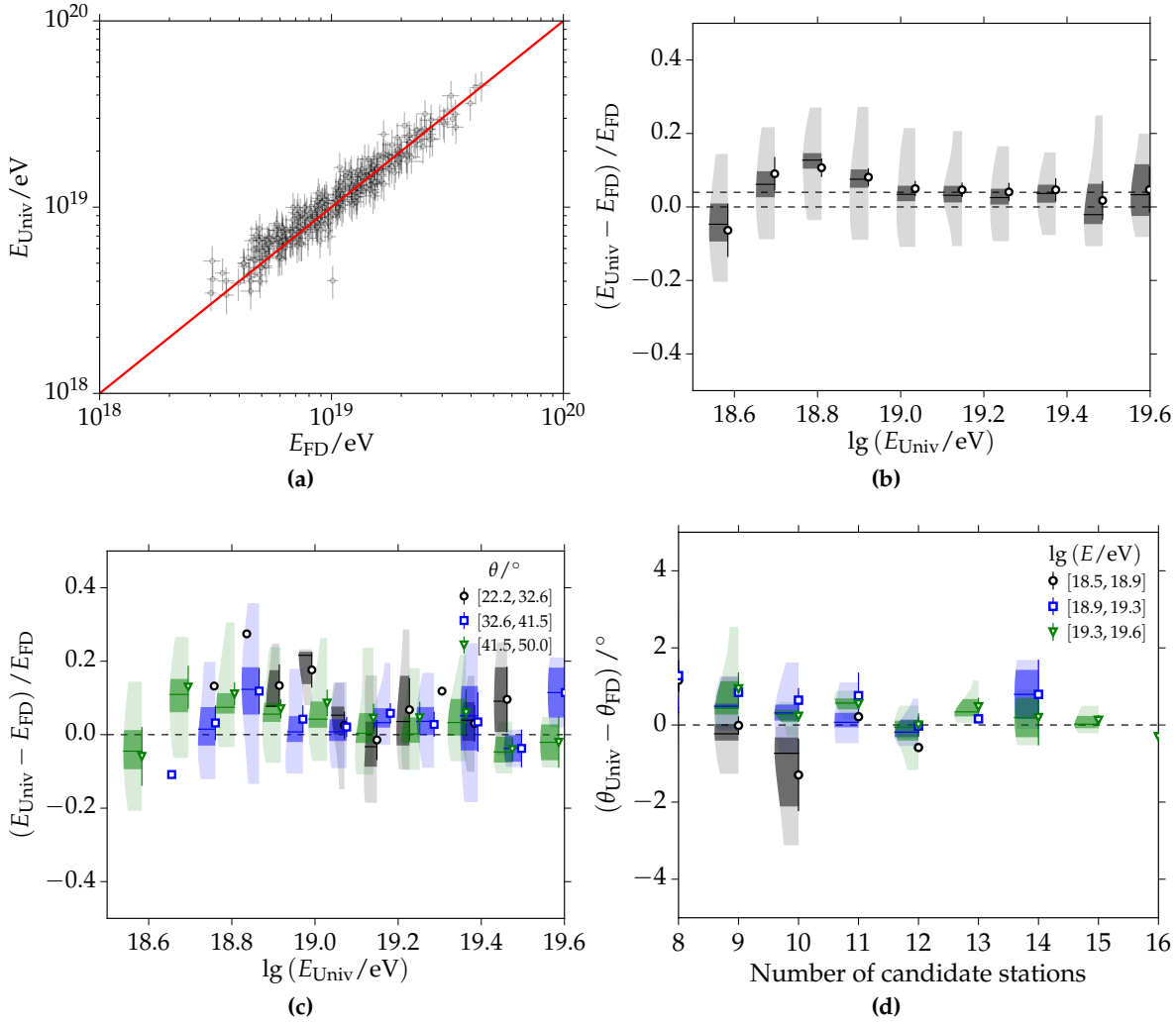


Figure B.28: Comparison of reconstructed energies and zenith angles using golden hybrid events up to December 2014. The classic universality fit was used to reconstruct data independently of the FD reconstruction.

B.4 Universality reconstruction

B.4.1 Validation of the free reconstruction with golden hybrid events

The validation of the iterative universality reconstruction with golden hybrid events is discussed in detail in Section 5.4. As part of Chapter 5, some of the results obtained with the iterative reconstruction are compared to the results obtained with a classic reconstruction type. In that reconstruction, only the energy is fixed to the previous result from the standard reconstruction. All other quantities are fit at the same time without the use of additional models or constraints in the fit. The validation of these results using golden hybrid data is discussed in this section. Due to the large number of parameters in the fit, a minimal number of candidate stations of eight is required (see Section 3.4.5). The resulting residuals between quantities reconstructed with the universality approach and the FD reconstruction are shown in Fig. B.28 and Fig. B.29.

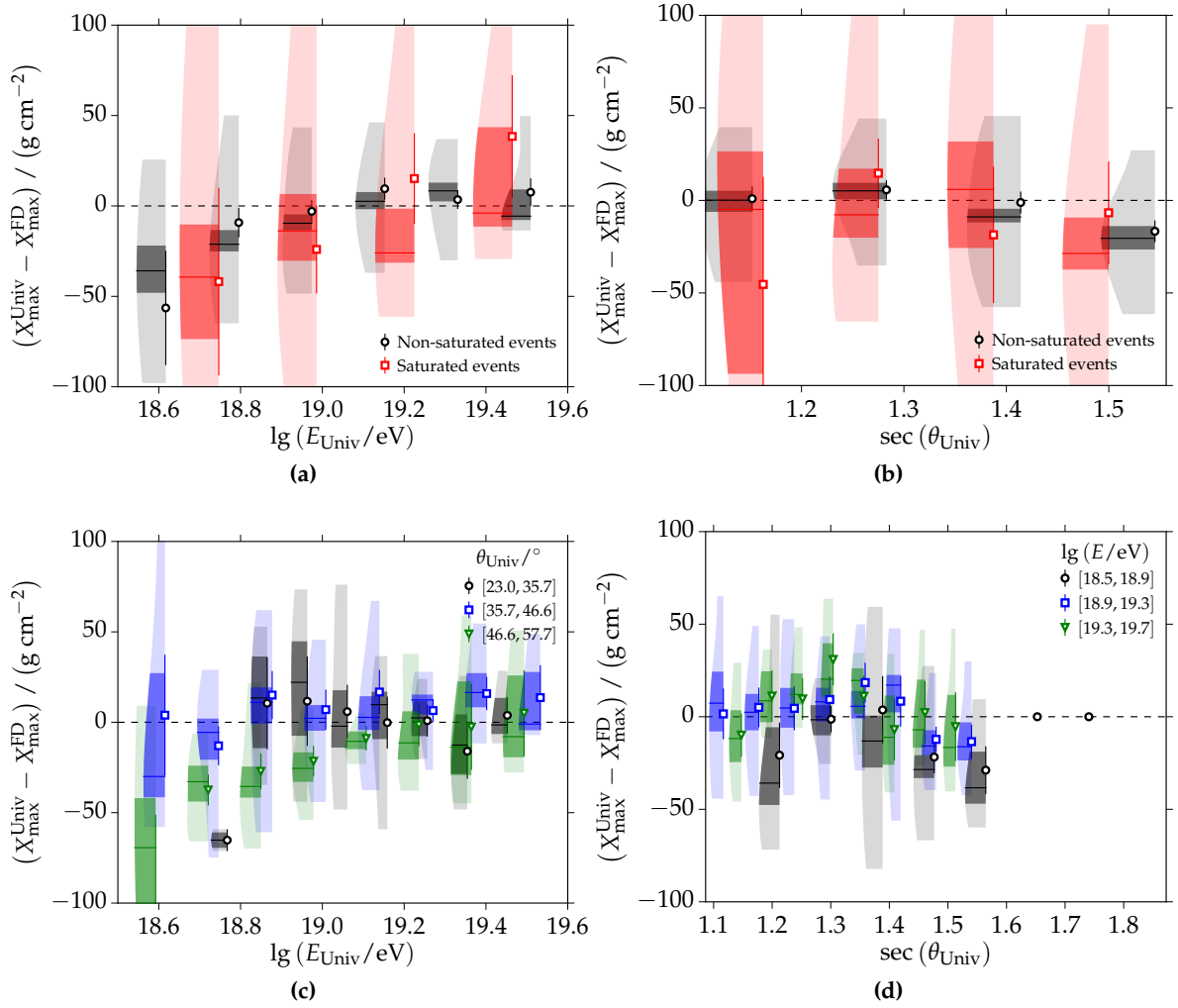


Figure B.29: Comparison of reconstructed depth of shower maxima using golden hybrid events. The classic universality method was used to reconstruct data independent of the FD reconstruction.

B.5 Reconstruction issues

B.5.1 Poisson factor in the classic air shower reconstruction

As described in Section 3.2.2 and Section 3.3.2, a Poisson factor is used to convert signals into effective particle numbers, which are then used in the LDF fit. A problem with the definition of the factor as it is used in Offline was found during the course of this work. The impact of the change on the relevant reconstructed parameters is summarized in the plots included in Fig. B.30. The plots give a comparison of shower sizes before and after the change of the Poisson factor for a selected set of data. The reconstruction was left unaltered except for this change. These plots give changes for the standard SD. On average the shower size S_{1000} shrinks by 0.5 %, while its statistical uncertainty decreases by 10 % on average. Changes are most prominent for the smallest and largest zenith angles due to the functional change operated in the Poisson factor (i.e. no change at $\sim 45°$).

Changes for the shower size S_{450} are very similar but a factor of two larger in case of the sizes. The change in statistical uncertainty is nearly identical. The larger impact for SD-750

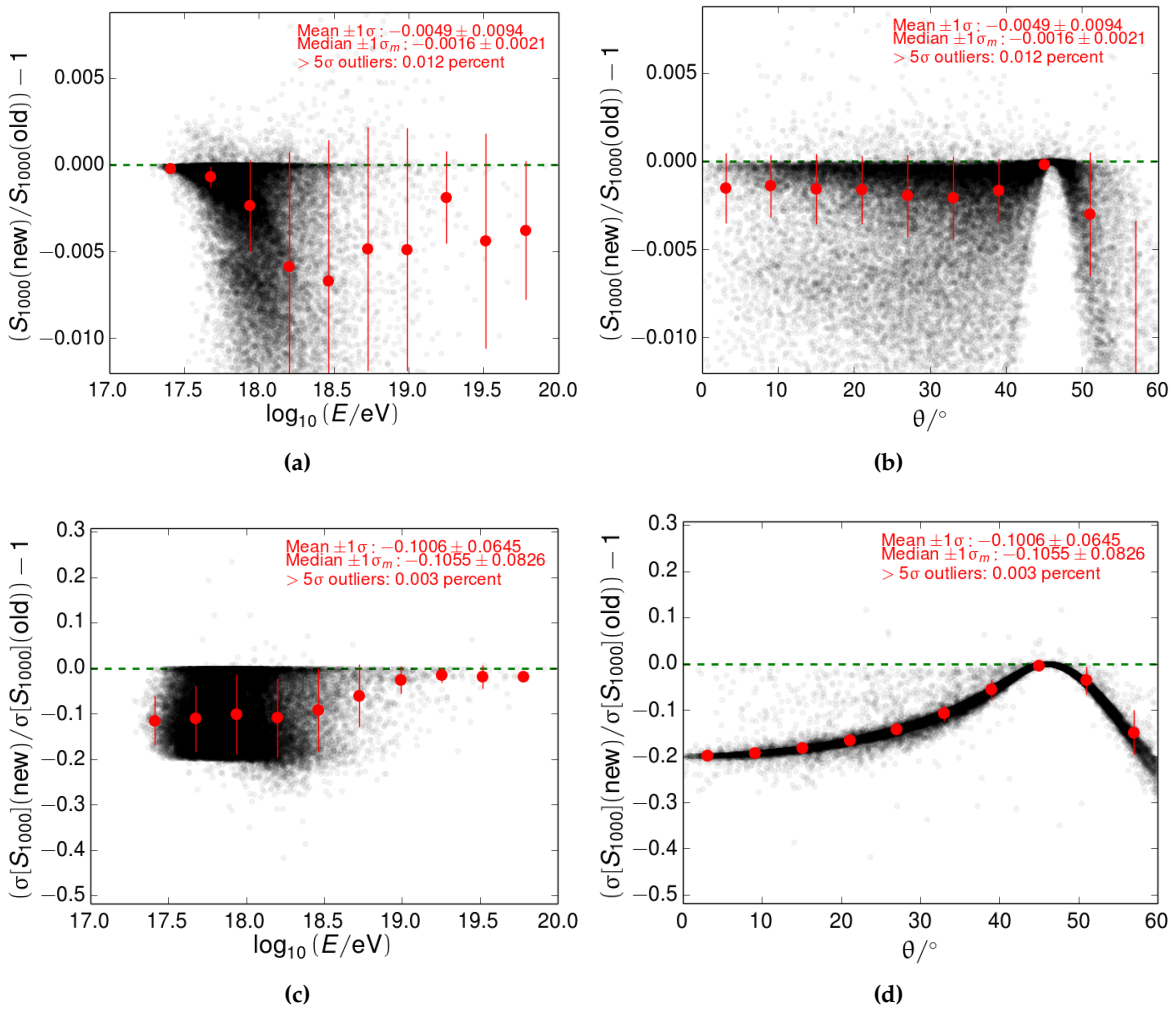


Figure B.30: Change of the shower size S_{1000} and its statistical uncertainty as a function of the reconstructed shower energy and zenith angle. New corresponds to a production with corrected Poisson factor with respect to a broken implementation in Offline prior to February 2014.

events is expected because stations with lower signal have larger weight due to the energy spectrum and the lateral distributions.

B.6 Energy calibration

B.6.1 Composition dependence

As argued in Section 4.7.7, a change in mass composition will lead to a change in the slope of the energy calibration. Not taking this into account might lead to the energy differences shown in Fig. 4.32. This modified energy calibration function is used to describe data better and to test this assumption:

$$S(E) = a \left(\frac{E}{E_0} \right)^{\left(b_0 + b_1 \lg \frac{E}{E_0}\right)} . \tag{B.1}$$

The slope of the function is modified to depend on energy itself. For all the studies and energy comparisons in this section, the energy calibration function for the SD-1500 is fixed to Eq. (4.37).

The energy calibration fit for SD-750 data is repeated using this function (see Section 4.7.5). The result is shown in Fig. B.31a. A non-zero slope of $b_1 = 0.055 \pm 0.020$ is visible in the plot. Using this calibration function, the energy differences between SD-750 and SD-1500 reconstructions are re-evaluated. Compared to the results in Section 4.7.6, the resulting plot Fig. B.31b depicts an average bias of only −2.2 %. That indicates that it might be better to assume an energy-dependent slope for the calibration function. Using only golden hybrid events, no conclusion can be made about which model fits better. A detailed Monte Carlo study should be performed to learn more about the energy calibration under a change in the mass composition. That is outside the scope of this work. There is a distinct energy dependence of the residuals in Fig. B.31b. That is probably due to the lack of events at highest energies.

A combined fit as in Section 4.7.7 was attempted to investigate this in more detail. Again, the modified calibration function is used. The results are shown in Fig. B.32. Now, the energy differences are very small on the order of −1 % as depicted in Fig. B.32b. There is no significant trend with energy because the common events constrain the calibration function at the highest energies. There is no significant difference to the result of a merged fit using the standard calibration function, which leads to a bias of −1.3 % as shown in Fig. 4.33b. Using the extended calibration function leads to various problems. The model in Eq. (B.1) is not easily invertible and the forward-folding and combination analyses need to be adjusted. Therefore, the standard calibration function is used. There is no significant improvement in the description of data with the new function.

B.7 Toy analyses

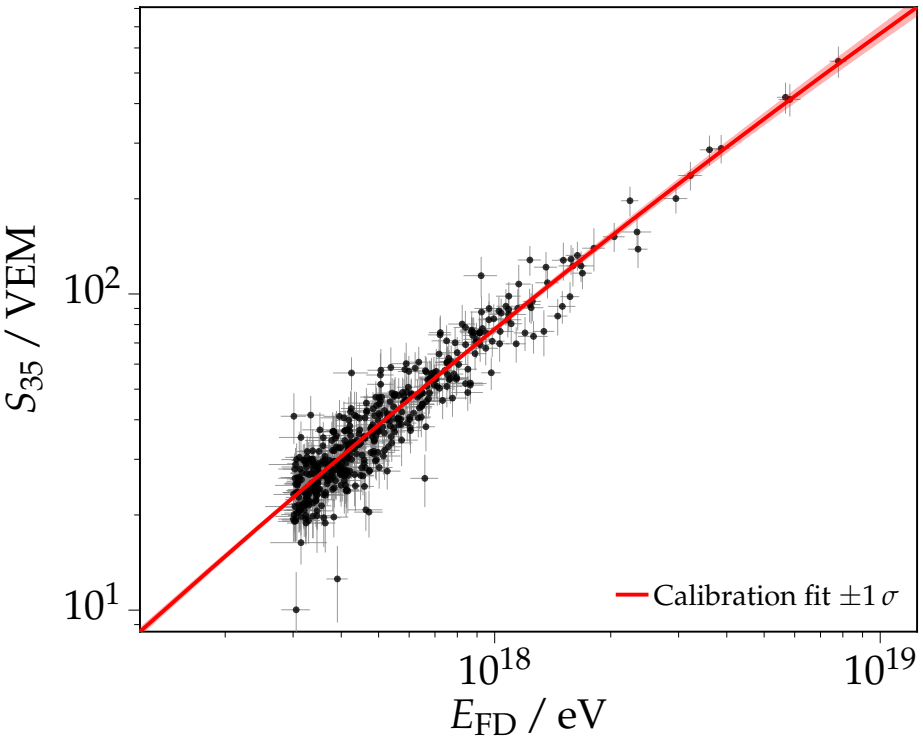
B.7.1 Forward-folding

The forward-folding method is used to correct for migration effects of the measured flux of CRs. These stem from different statistical and systematic sources as discussed in Chapter 4. In the following I will present a toy analysis that demonstrates the validity of the algorithm used in the method. A true spectrum similar to the flux of UHECRs is plugged in and smeared with a toy resolution model similar to the one found for the SD. In addition, a bias in the reconstructed energy is taken into account.

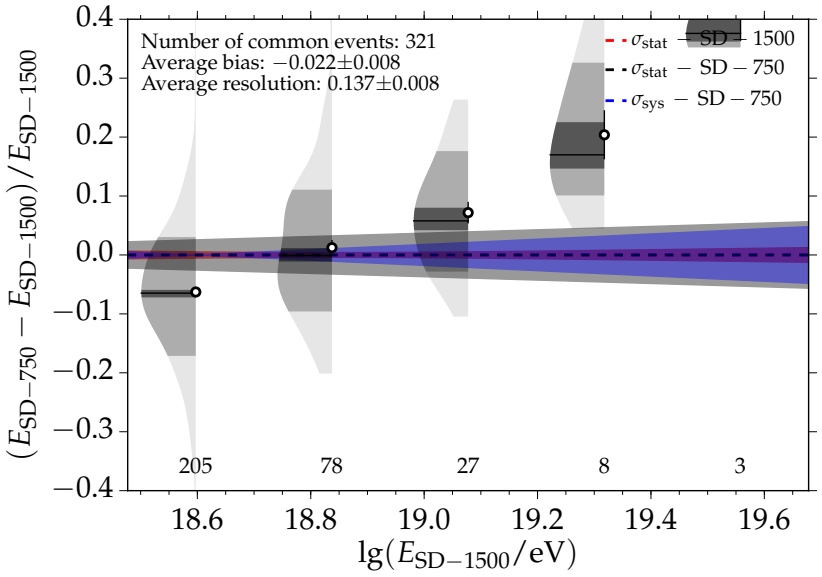
The true model is a broken power-law with instantaneous break at the ankle and a smooth tail suppression. It is described in Appendix B.8.1 with the specific equation Eq. (B.2). The true parameters are chosen similar to what is found for the flux of UHECRs and are stated in Table B.3. The energy resolution model is depicted in Fig. B.33a and a constant energy bias of 5 % is assumed (Fig. B.33b). In reality, this bias could be caused by reconstruction biases in either SD or FD analyses as it is discussed in Chapter 4. The biased energy is 5 % larger than the true energy.

Given the true flux as well as toy resolution and bias models, a numerical integration is performed to obtain the smeared flux. The two fluxes are compared in Fig. B.33c. The ratio of these fluxes gives us the smearing factor as shown in Fig. B.33d. Thus, given the specific energy resolution and bias, the measured flux overestimates the true one by 10 % to 30 %, depending on the energy. Sharp features like the break at the ankle propagate into the factor.

In the forward-folding method, a model to describe the true flux needs to be assumed a priori. In this case I use the true model and plug in the toy resolution and bias models. As

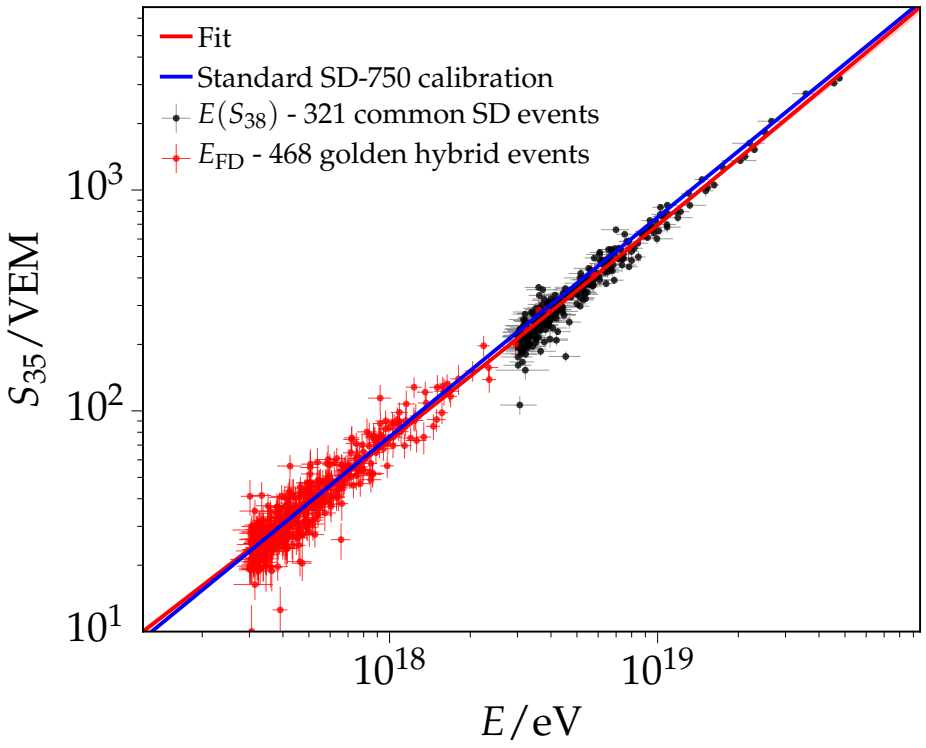


(a) Modified energy calibration fit

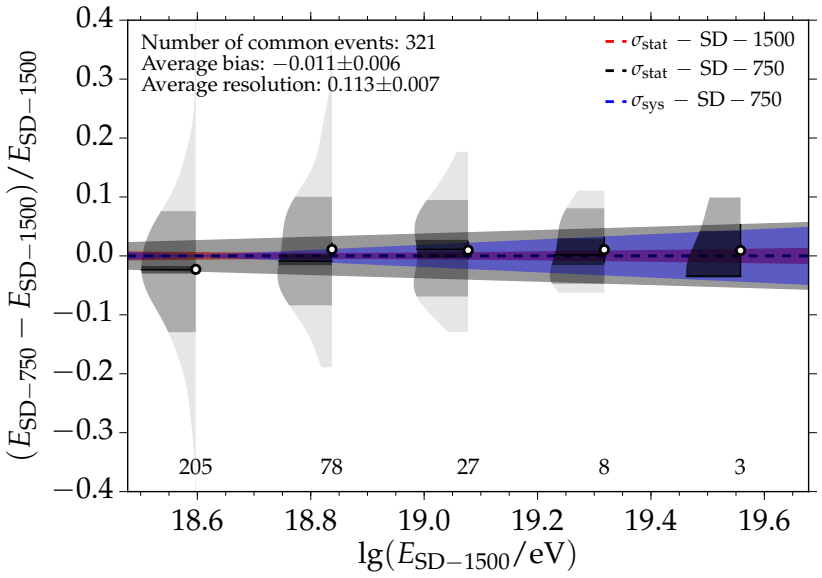


(b) Energy differences

Figure B.31: (a) Resulting energy calibration compared to data. The modified calibration function from Eq. (B.1) is used. (b) Energy differences between SD-750 and SD-1500 reconstructions. The clear trend with energy results from the very low event statistics at highest energies.



(a) Combined energy calibration fit



(b) Energy differences

Figure B.32: (a) Energy calibration fit taking both SD-750 golden hybrid events and events that are common with SD-1500 into account. (b) The energy residuals resulting from this fit.

Table B.3: Overview of the true and fitted parameters for the toy validation of the forward-folding method

	$\lg(E_a/\mathrm{eV})$	γ_1	γ_2	$\lg(E_{1/2}/\mathrm{eV})$	$\lg W_c$
True parameters	18.7	−3.3	−2.7	19.7	0.1
Fitted parameters	18.72±0.03	−3.30±0.01	−2.67±0.08	19.70±0.05	0.08±0.04

a sample spectrum, 89 170 random numbers were drawn from the smeared model. This is a reasonable simulation of a real measurement in terms of statistics. The sample spectrum is depicted in Fig. B.33c together with the underlying models. This spectrum is now used as input in the Poisson maximum-likelihood fit to reconstruct the true spectrum. The method is described in more detail in Section 4.10. The result is shown in Fig. B.33e. The sample spectrum is corrected with the fitted corrections as visualized in Fig. B.33f. The corrected points agree well with the fitted model and the fitted model gives a nice description of the true model. Thus, the true flux was reconstructed correctly within statistical uncertainties. The obtained parameters for this test are compared to the true model parameters in Table B.3. Each reconstructed quantity deviates less than $1\,\sigma$ from the truth. Repeating the test yields very similar results. Tested was a variety of flux models and sample statistics. This example is not an optimal case, often the results are even less biased depending on the shape of the sample flux.

B.8 Energy spectrum

This section contains additional information used in the analysis of the flux of UHECRs.

B.8.1 Flux models

For completeness, the mathematical models to describe the flux of CRs will be stated in this section.

Model with a hard break at the ankle

The model is given by a power law $J(E) \propto E^{-\gamma_1}$ below the ankle E_a and a power law with smooth suppression above:

$$J(E|E > E_a) \propto E^{-\gamma_2} \left[1 + \exp\left(\frac{\lg E - \lg E_{1/2}}{\lg W_c}\right)\right]^{-1}. \tag{B.2}$$

The spectral indices before and after the ankle are γ_1 and γ_2. $\lg W_c$ is the strength of the flux suppression, and $E_{1/2}$ is the energy at which the flux is half of its extraploted value without a suppression.

Model with hard break and smooth suppression

Below the ankle energy E_a, the model is $J(E) = J_0\,(E/E_\mathrm{a})^{-\gamma_1}$. Above the ankle, the model is a power-law with a smooth suppression at the highest energies:

$$J(E) = J_0 \left(\frac{E}{E_\mathrm{a}}\right)^{-\gamma_2} \left[1 + \left(\frac{E_\mathrm{a}}{E_\mathrm{s}}\right)^{\Delta\gamma}\right] \left[1 + \left(\frac{E}{E_\mathrm{s}}\right)^{\Delta\gamma}\right]^{-1}. \tag{B.3}$$

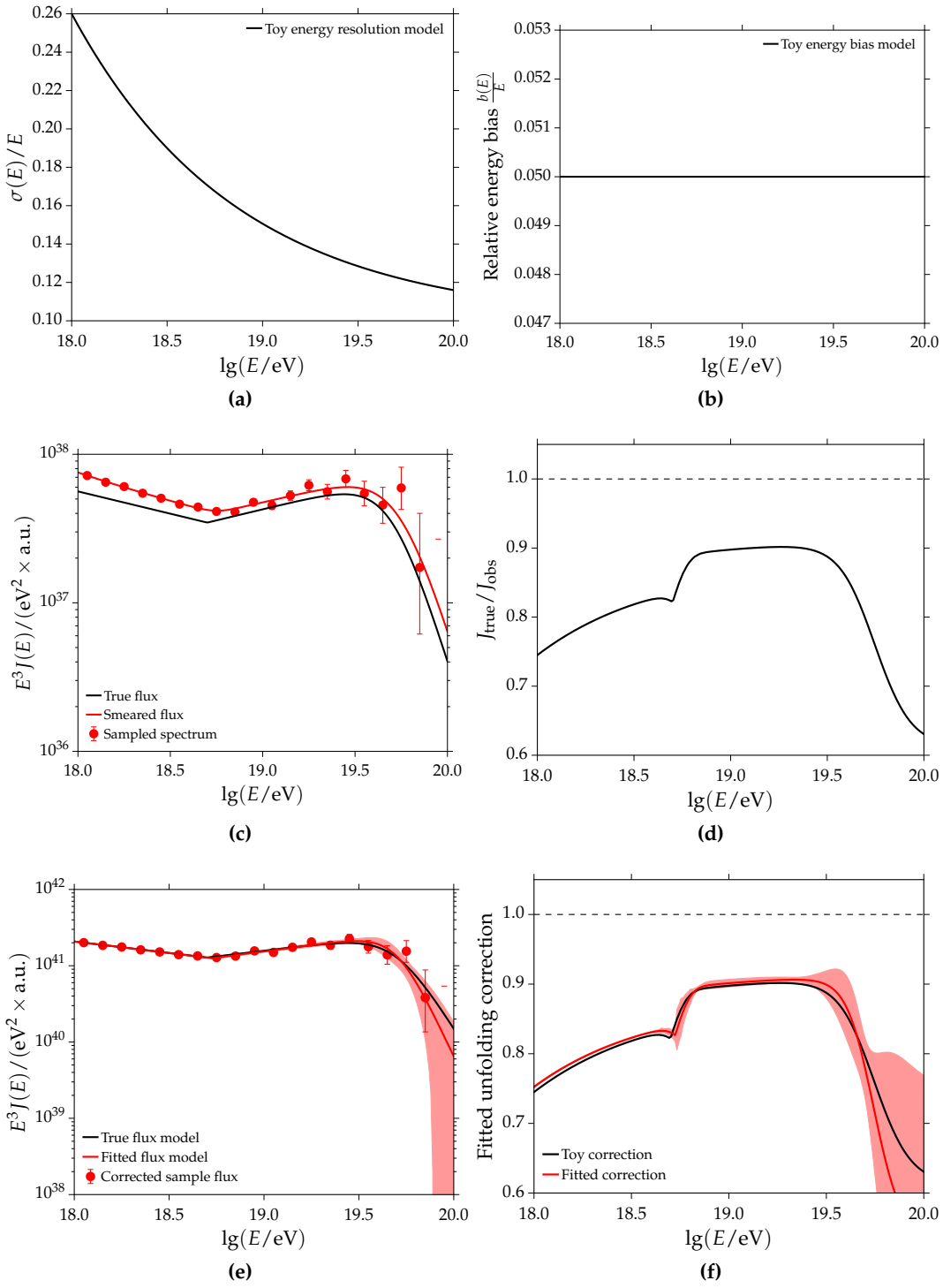
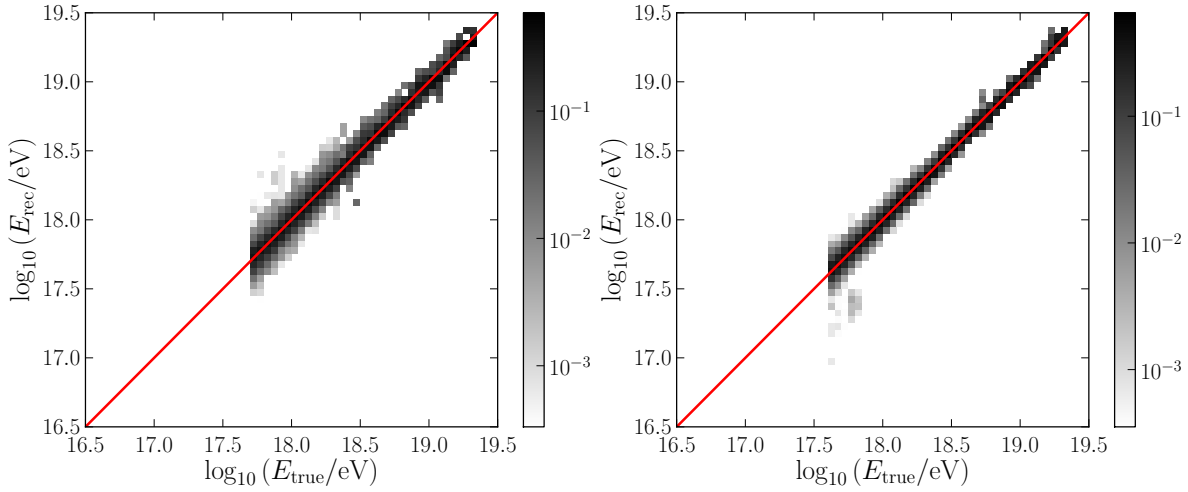


Figure B.33: Validation of the forward-folding method with a toy Monte Carlo.

B.8.2 Migration matrices

Obtained with SD-750 simulations, migration matrices for the two primaries proton and iron are depicted in Fig. B.34. The functional dependencies of the resolutions models determined from simulations and used to forward-fold the measured spectra are stated



(a) Proton **(b)** Iron

Figure B.34: Energy migration matrices obtained from SD-750 simulations. The probability of an event migrating is expressed with the color scale. A darker color resembles an higher probability.

in the following. A model from SD-1500 simulations (QGSJET-II.03) with a 50/50 mix of proton and iron showers:

$$\frac{\sigma[E]}{E} = 0.109 + 0.435\sqrt{\frac{E}{10^{17}\,\mathrm{eV}}}\,. \tag{B.4}$$

A model from SD-750 simulations (QGSJET-II.04) with a 50/50 mix of proton and iron primaries:

$$\frac{\sigma[E]}{E} = 0.078 + 0.165\sqrt{\frac{E}{10^{17}\,\mathrm{eV}}}\,. \tag{B.5}$$

B.8.3 Energy difference between Auger and TA

Assuming a common true flux, Fig. B.35 states the relative difference in the energy scale of Auger and TA as a function of the TA energy. The fitted model results from a combined fit of the Auger and TA data to a common flux model and the assumption of a second-degree polynomial to describe the energy bias. The points represent numerical calculations of the energy difference from the binned fluxes. They agree well with the model.

B.8.4 Distributions of incoming directions

Distributions of incoming directions are given in Fig. B.36 for the inclined measurement and in Fig. B.37 for the SD-750 measurement.

B.8.5 Additional intensity profiles

Intensity profiles as a function of zenith angle are plotted in Fig. B.38 for the different measurements. The intensities as a function of azimuth angle are depicted in Fig. B.39.

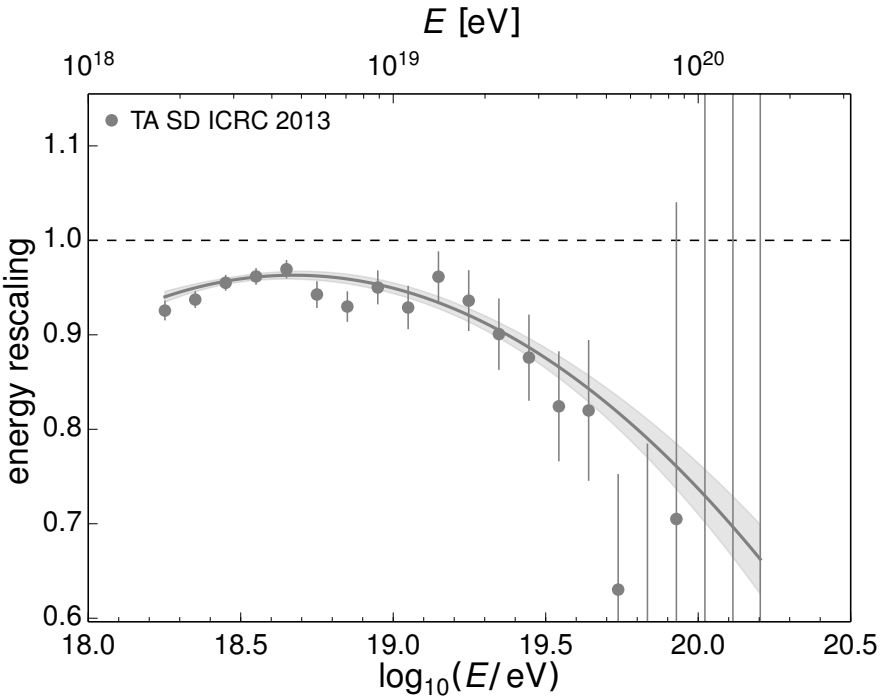


Figure B.35: Function needed to rescale the TA energies with in order to match the corresponding energy spectrum with the one presented by Auger at the ICRC 2013 [23].

B.9 Example events

B.9.1 Event with the highest reconstructed energy

The event with the highest reconstructed energy of 1.39×10^{20} eV is detailed in this section. The information from the Offline EventBrowser is shown in Fig. B.40.

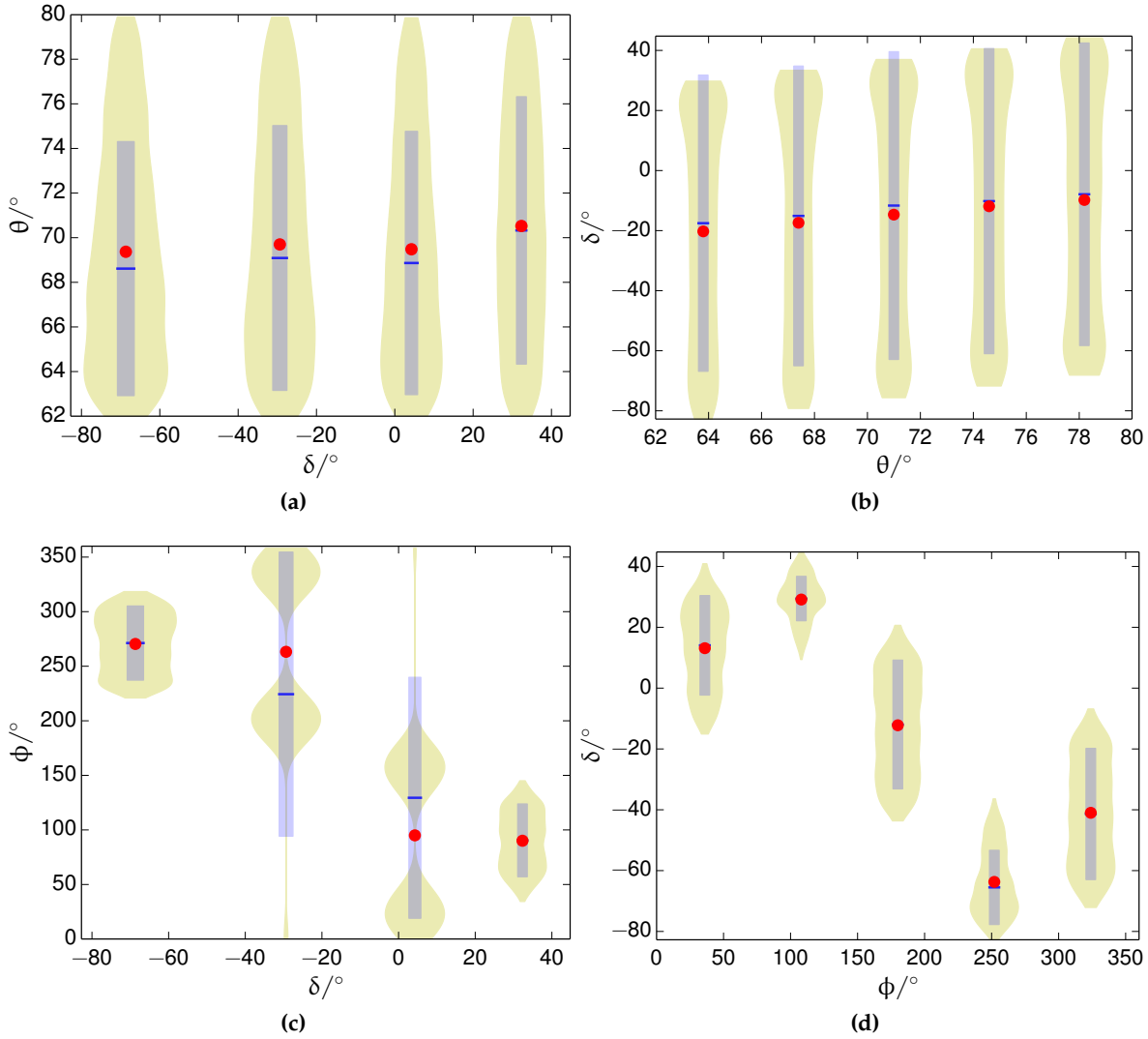


Figure B.36: Distributions of incoming directions for the measurement of inclined showers with the 1500 m array.

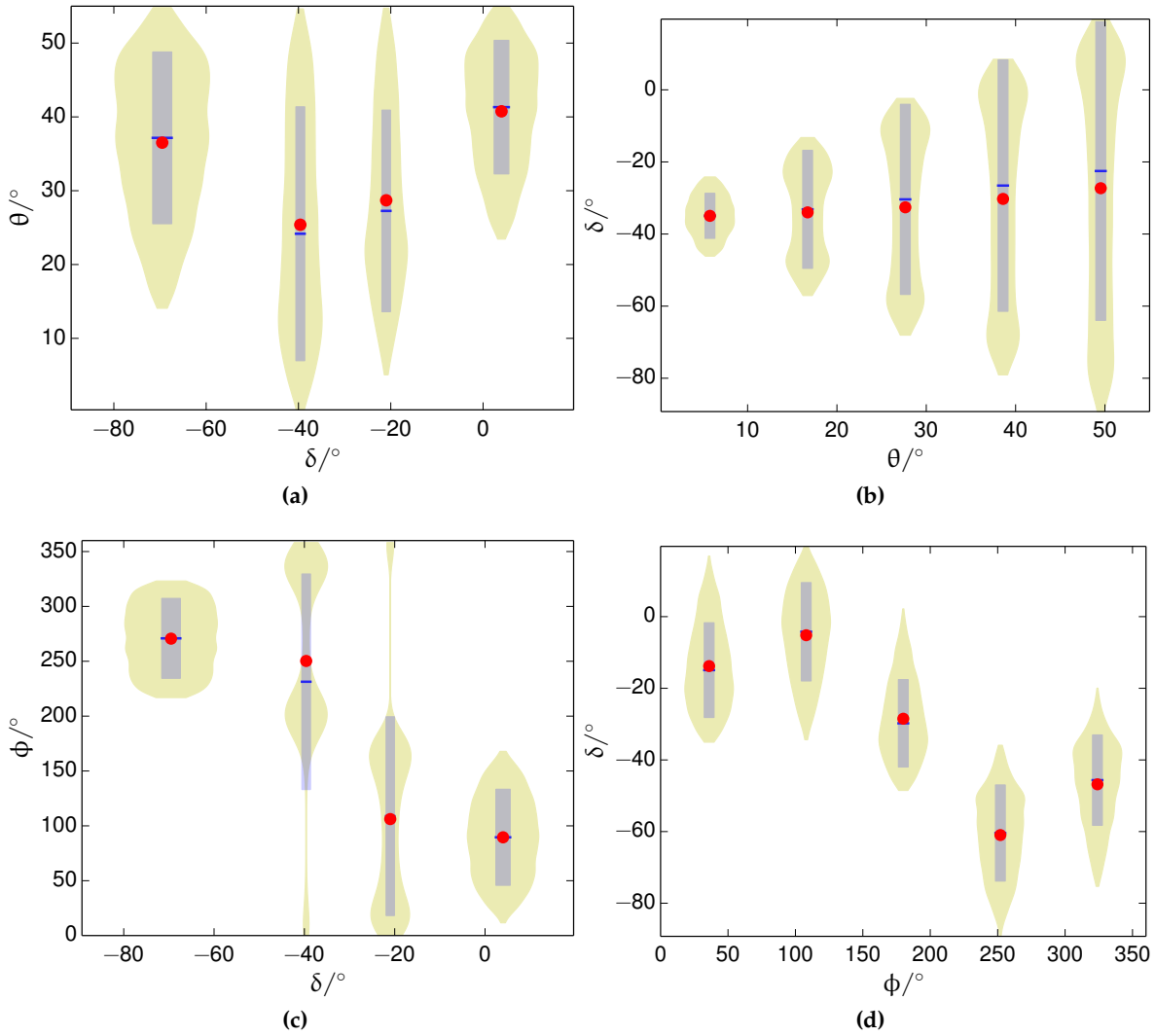
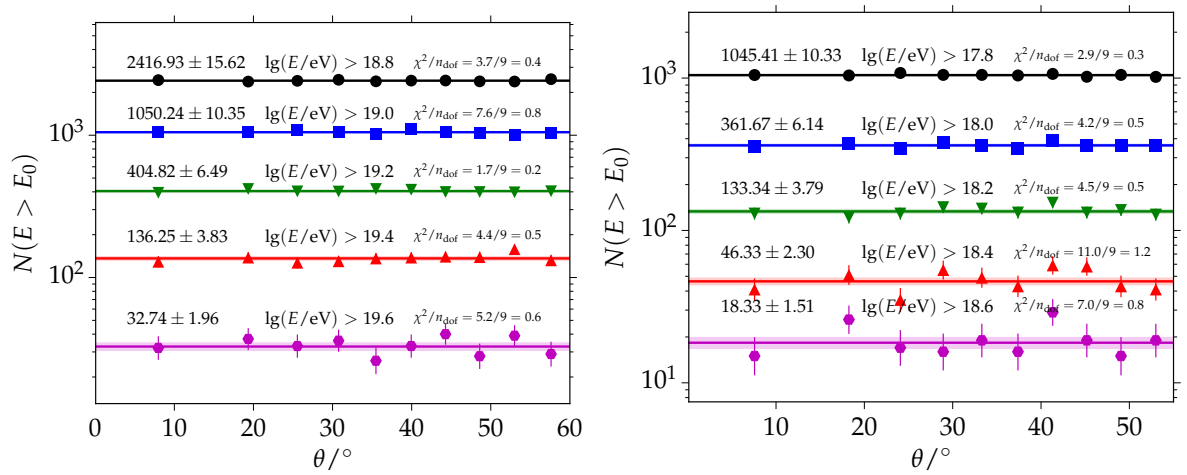


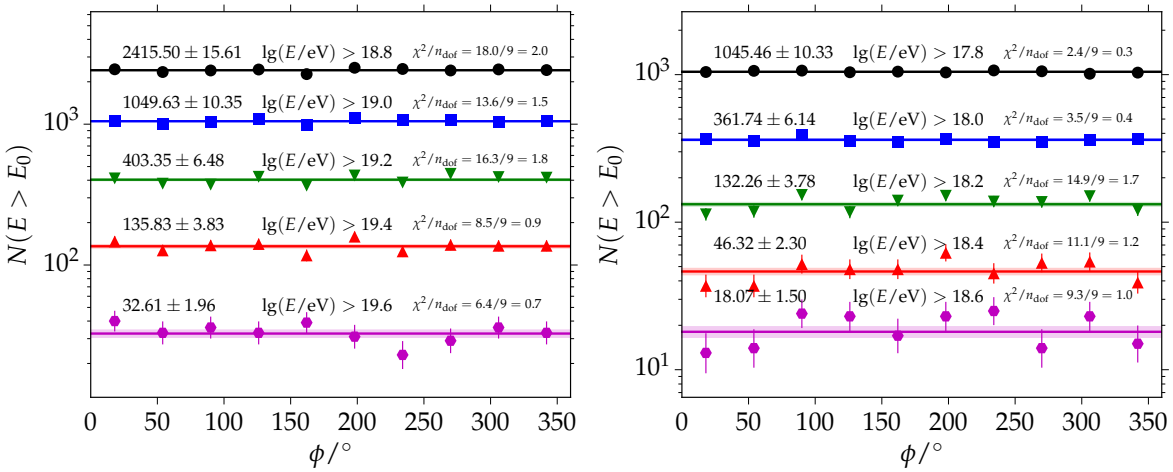
Figure B.37: Distributions of incoming directions for the measurement of vertical showers with the 750 m array.



(a) 1500 m vertical

(b) 750 m vertical

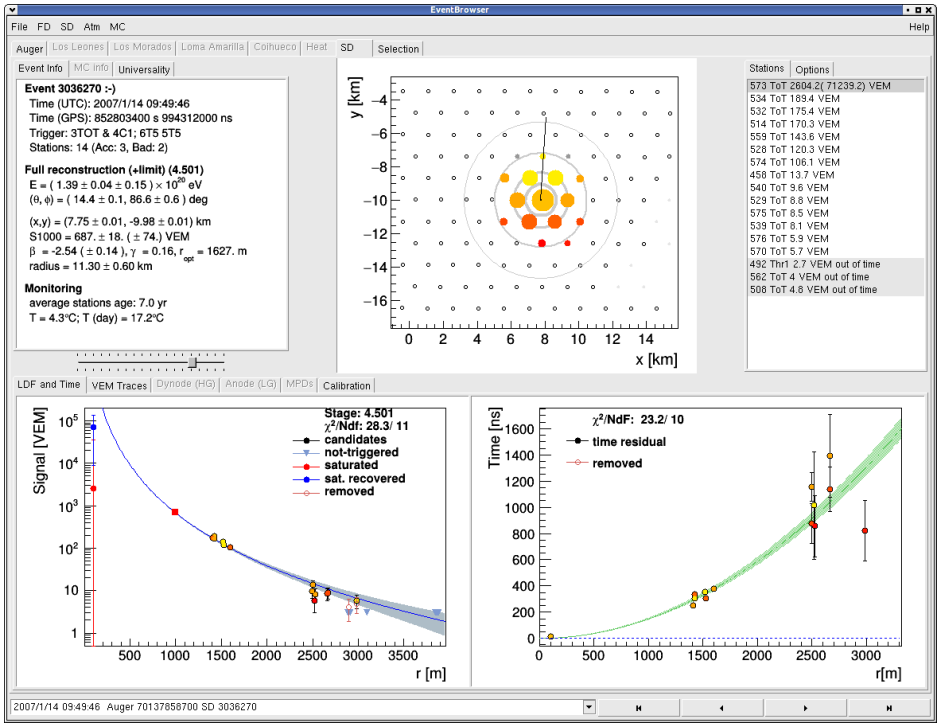
Figure B.38: Intensities as a function of zenith angle extracted from the SD-1500 and SD-750 data.



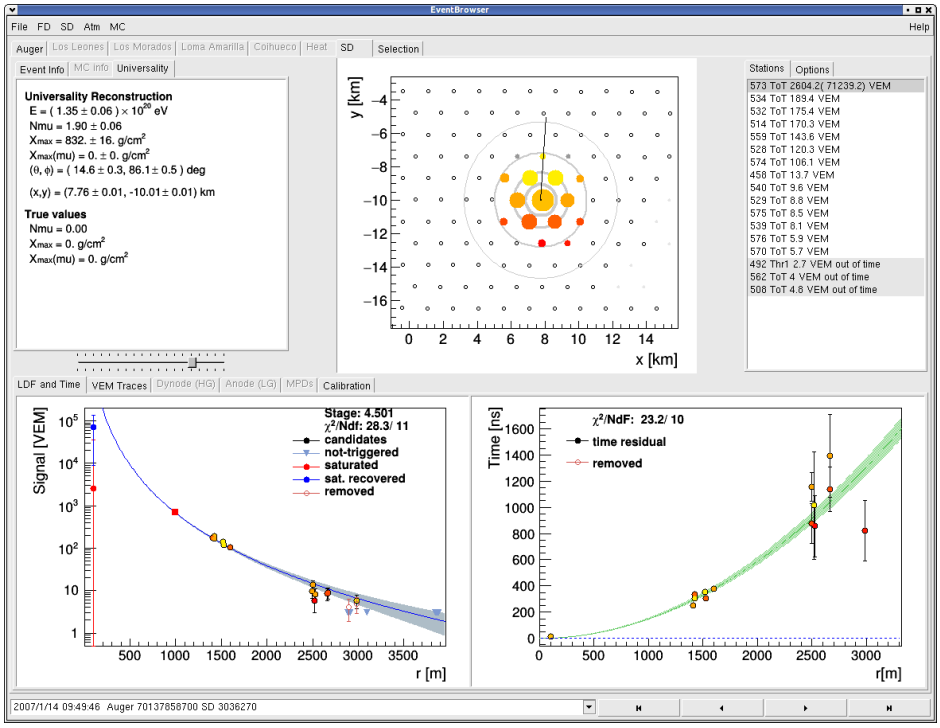
(a) 1500 m vertical

(b) 750 m vertical

Figure B.39: Intensities as a function of azimuth angle extracted from the SD-750 and SD-1500 data.



(a)



(b)

Figure B.40: Screenshots from the Offline EventBrowser for the event with the highest reconstructed energy (Id: .

APPENDIX C

Tables of results

This chapter includes tables of numerical results obtained in this thesis. If numbers are stated with two uncertainties, the use of the following format is implied: number $\pm$ stat.unc. $\pm$ sys.unc..

C.1 The flux of UHECRs

The flux results obtained in this thesis are discussed in Chapter 4. Tables of numerical results from the energy spectra are included in this section.

C.2 The mass composition of UHECRs

Results on the mass composition of UHECRs derived from a shower univerality analysis of SD-1500 data are discussed in detail in Chapter 5. Related tables of results are included in this section.

Table C.1: Results of the flux measurement with the SD-750. The flux is corrected for migration effects due to the finite energy resolution of the detector and due to shower-to-shower fluctuations. All relative uncertainties are stated in percent. The unit of the flux J is $1/(\mathrm{eV\,km^2\,yr\,sr})$.

$\langle \lg\left(\frac{E}{\mathrm{eV}}\right)\rangle$	N	$J\left(\frac{\mathrm{d}N}{\mathrm{d}E\,\mathrm{d}\mathcal{E}}\right)$	$\frac{\sigma^{-}_{\mathrm{stat}}(J)}{J}$	$\frac{\sigma^{+}_{\mathrm{stat}}(J)}{J}$	$\frac{\sigma^{\mathrm{tot}}_{\mathrm{sys}}(J)}{J}$	$\frac{\sigma^{\mathrm{cal(stat)}}_{\mathrm{sys}}(J)}{J}$	$\frac{\sigma^{\mathrm{cal(sys)}}_{\mathrm{sys}}(J)}{J}$	$\frac{\sigma^{\mathrm{add}}_{\mathrm{sys}}(J)}{J}$	Exp. $\mathcal{E}$	$\frac{\sigma_{\mathrm{sys}}(\mathcal{E})}{\mathcal{E}}$	Corr. c	$\frac{\sigma_{\mathrm{stat}}(c)}{c}$	$\frac{\sigma_{\mathrm{sys}}(c)}{c}$
17.55	17 767	1.492×10^{-15}	0.75	0.75	12.09	0.00	11.09	3.50	135.38	3.00	0.931	0.051	1.329
17.65	11 568	7.237×10^{-16}	0.92	0.93	11.59	2.13	10.34	3.50	145.73	3.00	0.940	0.045	1.193
17.75	7227	3.452×10^{-16}	1.17	1.18	10.95	2.13	9.64	3.50	152.86	3.00	0.947	0.040	1.058
17.85	4328	1.649×10^{-16}	1.51	1.53	10.36	2.13	8.99	3.50	153.23	3.00	0.954	0.036	0.922
17.95	2508	7.632×10^{-17}	1.99	2.03	9.79	1.81	8.40	3.50	153.23	3.00	0.959	0.032	0.787
18.05	1417	3.441×10^{-17}	2.64	2.71	9.32	1.65	7.90	3.50	153.23	3.00	0.963	0.029	0.652
18.15	866	1.677×10^{-17}	3.38	3.50	8.92	1.39	7.50	3.50	153.23	3.00	0.967	0.026	0.517
18.25	566	8.736×10^{-18}	4.18	4.36	8.77	1.92	7.21	3.50	153.23	3.00	0.970	0.024	0.382
18.35	297	3.651×10^{-18}	5.77	6.11	8.48	1.01	7.04	3.50	153.23	3.00	0.973	0.022	0.247
18.45	173	1.694×10^{-18}	7.56	8.15	8.85	2.81	7.02	3.50	153.23	3.00	0.976	0.020	0.113
18.55	111	8.649×10^{-19}	9.43	10.36	8.69	1.82	7.13	3.50	153.23	3.00	0.978	0.022	0.021
18.65	63	3.895×10^{-19}	12.50	14.18	9.00	2.15	7.38	3.50	153.23	3.00	0.976	0.807	0.155
18.75	37	1.834×10^{-19}	16.28	19.19	9.48	2.81	7.75	3.50	153.23	3.00	0.986	0.839	0.129
18.85	37	1.462×10^{-19}	16.28	19.19	9.71	2.33	8.22	3.50	153.23	3.00	0.989	0.225	0.025
18.95	22	6.910×10^{-20}	21.06	26.05	10.25	2.59	8.77	3.50	153.23	3.00	0.990	0.213	0.079
19.05	15	3.745×10^{-20}	24.53	28.84	11.22	4.01	9.40	3.50	153.23	3.00	0.991	0.202	0.183
19.15	6	1.191×10^{-20}	36.25	54.67	17.71	13.80	10.09	3.50	153.23	3.00	0.991	0.193	0.287
19.25	2	3.154×10^{-21}	63.00	112.76	21.94	18.51	10.83	3.50	153.23	3.00	0.992	0.185	0.392
19.35	4	5.012×10^{-21}	41.50	69.51	15.12	8.52	11.60	3.50	153.23	3.00	0.992	0.178	0.496
19.45	2	1.991×10^{-21}	63.00	112.76	17.02	10.69	12.40	3.50	153.23	3.00	0.992	0.172	0.601
19.55	0	0.000	0.00	0.00	14.03	0.00	13.23	3.50	153.23	3.00	0.993	0.167	0.706
19.65	3	1.886×10^{-21}	63.16	76.84	18.09	10.35	14.08	3.50	153.23	3.00	0.993	0.162	0.810

Table C.2: Results of the flux measurement with the SD-1500. The flux is corrected for migration effects due to the finite energy resolution of the detector and due to shower-to-shower fluctuations. All relative uncertainties are stated in percent. The unit of the flux J is $1/(\mathrm{eV\,km^2\,yr\,sr})$.

$\langle \lg(\frac{E}{\mathrm{eV}}) \rangle$	N	$J(\frac{\mathrm{d}N}{\mathrm{d}E\,\mathrm{d}\mathcal{E}})$	$\frac{\sigma^-_{\mathrm{stat}}(J)}{J}$	$\frac{\sigma^+_{\mathrm{stat}}(J)}{J}$	$\frac{\sigma^{\mathrm{tot}}_{\mathrm{sys}}(J)}{J}$	$\frac{\sigma^{\mathrm{cal(stat)}}_{\mathrm{sys}}(J)}{J}$	$\frac{\sigma^{\mathrm{add}}_{\mathrm{sys}}(J)}{J}$	Exp. $\mathcal{E}$	$\frac{\sigma_{\mathrm{sys}}(\mathcal{E})}{\mathcal{E}}$	Corr. c	$\frac{\sigma_{\mathrm{stat}}(c)}{c}$	$\frac{\sigma_{\mathrm{sys}}(c)}{c}$
18.45	54 148	1.818×10^{-18}	0.43	0.43	6.31	0.00	3.50	42 052	5.00	0.918	0.069	1.590
18.55	31 573	8.406×10^{-19}	0.56	0.56	6.54	1.85	3.50	42 520	5.00	0.927	0.065	1.449
18.65	18 876	4.001×10^{-19}	0.72	0.73	6.42	1.50	3.50	42 520	5.00	0.929	0.112	1.309
18.75	11 759	1.998×10^{-19}	0.92	0.93	6.28	1.18	3.50	42 520	5.00	0.938	0.382	0.837
18.85	7983	1.121×10^{-19}	1.11	1.13	6.19	0.80	3.50	42 520	5.00	0.976	0.106	0.627
18.95	5680	6.362×10^{-20}	1.32	1.34	6.15	0.63	3.50	42 520	5.00	0.980	0.104	0.444
19.05	3888	3.465×10^{-20}	1.59	1.62	6.14	0.62	3.50	42 520	5.00	0.981	0.079	0.307
19.15	2572	1.823×10^{-20}	1.96	2.00	6.15	0.72	3.50	42 520	5.00	0.982	0.055	0.237
19.25	1652	9.296×10^{-21}	2.45	2.51	6.26	1.36	3.50	42 520	5.00	0.982	0.083	0.265
19.35	1031	4.595×10^{-21}	3.10	3.19	6.39	1.82	3.50	42 520	5.00	0.979	0.161	0.409
19.45	623	2.190×10^{-21}	3.98	4.15	6.18	0.72	3.50	42 520	5.00	0.972	0.216	0.660
19.55	411	1.134×10^{-21}	4.90	5.15	6.43	1.77	3.50	42 520	5.00	0.961	0.220	0.955
19.65	179	3.868×10^{-22}	7.43	8.00	8.50	5.78	3.50	42 520	5.00	0.947	0.343	1.196
19.75	98	1.658×10^{-22}	10.03	11.10	7.24	3.61	3.50	42 520	5.00	0.934	0.630	1.300
19.85	42	5.581×10^{-23}	15.29	17.84	13.57	12.02	3.50	42 520	5.00	0.923	0.900	1.241
19.95	11	1.153×10^{-23}	29.00	34.73	8.06	5.04	3.50	42 520	5.00	0.917	1.075	1.043
20.05	2	1.659×10^{-24}	63.00	112.76	23.41	22.56	3.50	42 520	5.00	0.913	1.158	0.749
20.15	1	6.580×10^{-25}	63.00	175.51	27.22	26.50	3.50	42 520	5.00	0.912	1.182	0.395

Table C.3: Results of the flux measurement with the SD-1500 and the shower universality reconstruction. The flux is not corrected for migration effects. All relative uncertainties are stated in percent. The unit of the flux J is $1/(\mathrm{eV\,km^2\,yr\,sr})$.

$\langle \lg\left(\frac{E}{\mathrm{eV}}\right)\rangle$	N	$J(\frac{\mathrm{d}N}{\mathrm{d}E\,\mathrm{d}\mathcal{E}})$	$\frac{\sigma^{-}_{\mathrm{stat}}(J)}{J}$	$\frac{\sigma^{+}_{\mathrm{stat}}(J)}{J}$	$\frac{\sigma^{\mathrm{tot}}_{\mathrm{sys}}(J)}{J}$	$\frac{\sigma^{\mathrm{add}}_{\mathrm{sys}}(J)}{J}$	Exp. $\mathcal{E}$	$\frac{\sigma_{\mathrm{sys}}(\mathcal{E})}{\mathcal{E}}$
18.55	24 441	8.972×10^{-19}	0.64	0.64	6.10	3.50	33 270	5.00
18.65	14 557	4.245×10^{-19}	0.82	0.83	6.10	3.50	33 270	5.00
18.75	9251	2.143×10^{-19}	1.03	1.04	6.10	3.50	33 270	5.00
18.85	6241	1.148×10^{-19}	1.26	1.27	6.10	3.50	33 270	5.00
18.95	4504	6.582×10^{-20}	1.48	1.50	6.10	3.50	33 270	5.00
19.05	3156	3.664×10^{-20}	1.77	1.80	6.10	3.50	33 270	5.00
19.15	2129	1.963×10^{-20}	2.16	2.20	6.10	3.50	33 270	5.00
19.25	1392	1.020×10^{-20}	2.67	2.74	6.10	3.50	33 270	5.00
19.35	912	5.306×10^{-21}	3.29	3.40	6.10	3.50	33 270	5.00
19.45	552	2.551×10^{-21}	4.23	4.42	6.10	3.50	33 270	5.00
19.55	365	1.340×10^{-21}	5.20	5.48	6.10	3.50	33 270	5.00
19.65	162	4.724×10^{-22}	7.81	8.44	6.10	3.50	33 270	5.00
19.75	84	1.946×10^{-22}	10.83	12.08	6.10	3.50	33 270	5.00
19.85	38	6.991×10^{-23}	16.07	18.90	6.10	3.50	33 270	5.00
19.95	14	2.046×10^{-23}	26.39	30.83	6.10	3.50	33 270	5.00
20.05	4	4.643×10^{-24}	41.50	69.51	6.10	3.50	33 270	5.00
20.15	2	1.844×10^{-24}	63.00	112.76	6.10	3.50	33 270	5.00

Table C.4: Results of the flux measurement with the SD-1500 in the high declination interval 0° to 25°. The flux is not corrected for migration effects. All relative uncertainties are stated in percent. The unit of the flux J is $1/(\mathrm{eV\,km^2\,yr\,sr})$.

$\langle \lg\left(\frac{E}{\mathrm{eV}}\right)\rangle$	N	$J(\frac{\mathrm{d}N}{\mathrm{d}E\,\mathrm{d}\mathcal{E}})$	$\frac{\sigma^{-}_{\mathrm{stat}}(J)}{J}$	$\frac{\sigma^{+}_{\mathrm{stat}}(J)}{J}$	$\frac{\sigma^{\mathrm{tot}}_{\mathrm{sys}}(J)}{J}$	$\frac{\sigma^{\mathrm{add}}_{\mathrm{sys}}(J)}{J}$	Exp. $\mathcal{E}$	$\frac{\sigma_{\mathrm{sys}}(\mathcal{E})}{\mathcal{E}}$
18.45	8257	2.127×10^{-18}	1.09	1.11	6.10	3.50	5969	5.00
18.55	4922	9.873×10^{-19}	1.42	1.44	6.10	3.50	6088	5.00
18.65	2780	4.430×10^{-19}	1.89	1.92	6.10	3.50	6088	5.00
18.75	1703	2.155×10^{-19}	2.41	2.47	6.10	3.50	6088	5.00
18.85	1142	1.148×10^{-19}	2.94	3.03	6.10	3.50	6088	5.00
18.95	817	6.524×10^{-20}	3.48	3.60	6.10	3.50	6088	5.00
19.05	573	3.635×10^{-20}	4.15	4.33	6.10	3.50	6088	5.00
19.15	363	1.829×10^{-20}	5.22	5.50	6.10	3.50	6088	5.00
19.25	237	9.485×10^{-21}	6.46	6.89	6.10	3.50	6088	5.00
19.35	125	3.974×10^{-21}	8.89	9.72	6.10	3.50	6088	5.00
19.45	90	2.273×10^{-21}	10.47	11.63	6.10	3.50	6088	5.00
19.55	64	1.284×10^{-21}	12.40	14.05	6.10	3.50	6088	5.00
19.65	30	4.780×10^{-22}	18.07	21.68	6.10	3.50	6088	5.00
19.75	16	2.025×10^{-22}	22.90	30.01	6.10	3.50	6088	5.00
19.85	5	5.027×10^{-23}	44.90	56.21	6.10	3.50	6088	5.00
19.95	1	7.986×10^{-24}	63.00	175.51	6.10	3.50	6088	5.00
20.05	1	6.343×10^{-24}	63.00	175.51	6.10	3.50	6088	5.00

Table C.5: Results of the flux measurement with the SD-750 and SD-1500. The points result from a combination method described in Section 4.12. The flux is corrected for migration effects due to the finite detector resolution. All relative uncertainties are stated in percent. The unit of the flux J is $1/(\mathrm{eV\,km^2\,yr\,sr})$.

$\langle \lg\left(\frac{E}{\mathrm{eV}}\right)\rangle$	N	$J\left(\frac{\mathrm{d}N}{\mathrm{d}E\,\mathrm{d}\mathcal{E}}\right)$	$\frac{\sigma^{-}_{\mathrm{stat}}(J)}{J}$	$\frac{\sigma^{+}_{\mathrm{stat}}(J)}{J}$	$\frac{\sigma^{\mathrm{tot}}_{\mathrm{sys}}(J)}{J}$
17.25	52 635	1.253×10^{-14}	0.43	0.44	14.29
17.35	39 096	6.180×10^{-15}	0.50	0.51	13.77
17.45	26 938	3.027×10^{-15}	0.61	0.61	13.00
17.55	17 767	1.462×10^{-15}	0.75	0.75	12.23
17.65	11 568	7.088×10^{-16}	0.92	0.93	11.52
17.75	7227	3.380×10^{-16}	1.17	1.18	10.89
17.85	4328	1.614×10^{-16}	1.51	1.53	10.31
17.95	2508	7.471×10^{-17}	1.99	2.03	9.75
18.05	1417	3.368×10^{-17}	2.64	2.71	9.28
18.15	866	1.641×10^{-17}	3.38	3.50	8.90
18.25	566	8.550×10^{-18}	4.18	4.36	8.74
18.35	297	3.573×10^{-18}	5.77	6.11	8.47
18.45	54 321	1.907×10^{-18}	0.43	0.43	6.11
18.55	31 684	8.815×10^{-19}	0.56	0.56	6.40
18.65	18 939	4.198×10^{-19}	0.72	0.73	6.31
18.75	11 796	2.092×10^{-19}	0.92	0.92	6.24
18.85	8020	1.174×10^{-19}	1.11	1.12	6.18
18.95	5702	6.653×10^{-20}	1.32	1.34	6.15
19.05	3903	3.624×10^{-20}	1.59	1.62	6.16
19.15	2578	1.905×10^{-20}	1.96	2.00	6.17
19.25	1654	9.719×10^{-21}	2.44	2.51	6.26
19.35	1035	4.811×10^{-21}	3.09	3.19	6.36
19.45	625	2.294×10^{-21}	3.98	4.15	6.18
19.55	411	1.189×10^{-21}	4.90	5.15	6.33
19.65	182	4.287×10^{-22}	7.38	7.97	8.22
19.75	98	1.736×10^{-22}	10.03	11.10	7.09
19.85	42	5.837×10^{-23}	15.29	17.84	13.08
19.95	11	1.204×10^{-23}	29.00	34.73	7.59
20.05	2	1.732×10^{-24}	63.00	112.76	22.03
20.15	1	6.866×10^{-25}	63.00	175.51	24.88

Table C.6: Table of the first two central moments of the measured X_{max} distributions. The result is obtained with a universality analysis of SD-1500 data. Only non-saturated events are included. The systematic uncertainties are fully correlated between the energy bins.

Range in $\lg(E/\mathrm{eV})$	$\langle \lg(E/\mathrm{eV}) \rangle$	N	$\langle X_{\mathrm{max}} \rangle / \left(\mathrm{g\,cm^{-2}}\right)$	$\sigma(X_{\mathrm{max}}) / \left(\mathrm{g\,cm^{-2}}\right)$
$[18.5, 18.6)$	18.546	19 290	$749.8 \pm 0.5 \pm 10.0$	$31.1 \pm 0.5 \pm 3.0$
$[18.6, 18.7)$	18.645	11 260	$754.8 \pm 0.5 \pm 10.0$	$33.6 \pm 0.7 \pm 3.0$
$[18.7, 18.8)$	18.746	6832	$758.2 \pm 0.7 \pm 10.0$	$32.9 \pm 0.8 \pm 3.0$
$[18.8, 18.9)$	18.846	4469	$761.3 \pm 0.8 \pm 10.0$	$33.1 \pm 0.9 \pm 3.0$
$[18.9, 19.0)$	18.947	3056	$765.2 \pm 0.9 \pm 10.0$	$28.9 \pm 1.2 \pm 3.0$
$[19.0, 19.1)$	19.047	2024	$768.5 \pm 1.0 \pm 10.0$	$29.4 \pm 1.4 \pm 3.0$
$[19.1, 19.2)$	19.145	1294	$770.0 \pm 1.1 \pm 10.0$	$26.5 \pm 1.6 \pm 3.0$
$[19.2, 19.3)$	19.246	764	$774.4 \pm 1.5 \pm 10.0$	$27.5 \pm 1.7 \pm 3.0$
$[19.3, 19.4)$	19.345	456	$774.4 \pm 1.8 \pm 10.0$	$21.8 \pm 2.0 \pm 3.0$
$[19.4, 19.5)$	19.442	262	$772.3 \pm 2.2 \pm 10.0$	$29.9 \pm 4.5 \pm 3.0$
$[19.5, 19.6)$	19.543	148	$780.3 \pm 2.5 \pm 10.0$	$22.8 \pm 3.4 \pm 3.0$
$[19.6, 19.7)$	19.643	59	$777.4 \pm 2.9 \pm 10.0$	$5.8 \pm 2.4 \pm 3.0$
$[19.7, 19.8)$	19.735	28	$782.5 \pm 4.2 \pm 10.0$	$7.9 \pm 2.6 \pm 3.0$
$[19.8, \infty)$	19.858	19	$796.9 \pm 5.4 \pm 10.0$	$15.9 \pm 5.0 \pm 3.0$

Table C.7: Table of the first two central moments of the measured X_{max} distributions. The result is obtained with a universality analysis of SD-1500 data. All events are included. The systematic uncertainties are fully correlated between the energy bins.

Range in $\lg(E/\mathrm{eV})$	$\langle \lg(E/\mathrm{eV}) \rangle$	N	$\langle X_{\mathrm{max}} \rangle / \left(\mathrm{g\,cm^{-2}}\right)$	$\sigma(X_{\mathrm{max}}) / \left(\mathrm{g\,cm^{-2}}\right)$
$[18.5, 18.6)$	18.546	21 714	$763.1 \pm 0.5 \pm 10.0$	$47.6 \pm 0.5 \pm 3.0$
$[18.6, 18.7)$	18.646	12 759	$766.6 \pm 0.6 \pm 10.0$	$51.5 \pm 0.7 \pm 3.0$
$[18.7, 18.8)$	18.747	7980	$769.2 \pm 0.8 \pm 10.0$	$54.0 \pm 0.8 \pm 3.0$
$[18.8, 18.9)$	18.847	5338	$772.8 \pm 1.0 \pm 10.0$	$55.2 \pm 1.1 \pm 3.0$
$[18.9, 19.0)$	18.947	3764	$776.1 \pm 1.0 \pm 10.0$	$55.7 \pm 1.1 \pm 3.0$
$[19.0, 19.1)$	19.047	2626	$779.8 \pm 1.3 \pm 10.0$	$56.2 \pm 1.5 \pm 3.0$
$[19.1, 19.2)$	19.145	1744	$786.9 \pm 1.5 \pm 10.0$	$57.9 \pm 1.9 \pm 3.0$
$[19.2, 19.3)$	19.246	1111	$793.5 \pm 1.8 \pm 10.0$	$57.2 \pm 2.1 \pm 3.0$
$[19.3, 19.4)$	19.346	675	$791.7 \pm 2.6 \pm 10.0$	$63.6 \pm 3.0 \pm 3.0$
$[19.4, 19.5)$	19.444	401	$797.4 \pm 3.6 \pm 10.0$	$70.1 \pm 4.5 \pm 3.0$
$[19.5, 19.6)$	19.547	246	$803.2 \pm 5.2 \pm 10.0$	$70.3 \pm 5.5 \pm 3.0$
$[19.6, 19.7)$	19.639	115	$802.5 \pm 8.0 \pm 10.0$	$73.4 \pm 8.5 \pm 3.0$
$[19.7, 19.8)$	19.741	52	$810.2 \pm 6.8 \pm 10.0$	$45.1 \pm 6.3 \pm 3.0$
$[19.8, \infty)$	19.873	35	$818.0 \pm 16.1 \pm 10.0$	$90.4 \pm 18.5 \pm 3.0$

Table C.8: Table of the first two central moments of the measured X_{max} distributions. The result is obtained with a universality analysis of SD-1500 data. The result is achieved with a less constrained universality reconstruction in which only the energy is fixed and all other quantities are fit simultaneously (classic reconstruction); only non-saturated events are included. The systematic uncertainties are fully correlated between the energy bins.

Range in $\lg(E/\mathrm{eV})$	$\langle \lg(E/\mathrm{eV}) \rangle$	N	$\langle X_{\mathrm{max}} \rangle / \left(\mathrm{g\,cm^{-2}}\right)$	$\sigma(X_{\mathrm{max}}) / \left(\mathrm{g\,cm^{-2}}\right)$
$[19.0, 19.1)$	19.049	1350	$763.0 \pm 2.0 \pm 10.0$	$51.3 \pm 3.0 \pm 5.0$
$[19.1, 19.2)$	19.148	1161	$763.5 \pm 2.0 \pm 10.0$	$52.4 \pm 3.6 \pm 5.0$
$[19.2, 19.3)$	19.247	801	$766.3 \pm 2.3 \pm 10.0$	$50.6 \pm 2.6 \pm 5.0$
$[19.3, 19.4)$	19.346	536	$763.2 \pm 2.3 \pm 10.0$	$35.9 \pm 2.0 \pm 5.0$
$[19.4, 19.5)$	19.445	317	$760.7 \pm 2.7 \pm 10.0$	$35.1 \pm 2.5 \pm 5.0$
$[19.5, 19.6)$	19.543	176	$769.2 \pm 3.1 \pm 10.0$	$36.0 \pm 4.7 \pm 5.0$
$[19.6, 19.7)$	19.642	78	$768.8 \pm 3.3 \pm 10.0$	$19.5 \pm 2.4 \pm 5.0$
$[19.7, 19.8)$	19.735	39	$780.1 \pm 5.1 \pm 10.0$	$21.8 \pm 3.2 \pm 5.0$
$[19.8, \infty)$	19.867	24	$794.1 \pm 5.7 \pm 10.0$	$24.9 \pm 6.0 \pm 5.0$

APPENDIX D

Sequences

D.1 Module sequences and selections

This section includes the most relevant $\overline{\text{Off}}\underline{\text{line}}$ module sequences for the simulation and reconstruction of events, as well as sequences for the selection of reconstructed events.

D.1.1 Module sequences

In the following, relevant $\overline{\text{Off}}\underline{\text{line}}$ module sequences are given as XML listings.

D.1.2 Configuration files for event selections

Configuration files for the selection of ADST events are included in this section.

Table D.1: Cuts for the selection of vertical SD-750 events.

Cut name	Cut value	Meaning
!lightning		reject events caused by lightnings
minRecLevel	3	require reconstructed LDF
maxZenithSD	55	
T4Trigger	2	
T5Trigger	2	6T5 prior
badPeriodsRejectionFromFile		reject events in bad periods
timeInterval	params: 031231 150101	

Listing D.1: "Offline module sequence for the reconstruction of SD-750 and SD-1500 data."

```
<sequenceFile>

  <enableTiming/>

  <moduleControl>

    <loop numTimes="unbounded" pushEventToStack="yes">

      <module> EventFileReaderOG </module>

      <!-- SdCalibratorOG and the SdEventSelectorOG have to be
           outside the loop in order to write only the events
           that pass the bottom up selection -->
      <module> SdQualityCutTaggerOG </module>
      <module> SdPMTQualityCheckerKG </module>
      <module> TriggerTimeCorrection </module>
      <module> SdCalibratorOG </module>
      <module> SdStationPositionCorrection </module>
      <module> SdBadStationRejectorKG </module>
      <module> SdSignalRecoveryKLT </module>
      <module> SdEventSelectorOG </module>

      <!-- SD reconstruction -->
      <try>
        <module> SdPlaneFitOG </module>
        <module> LDFFinderKG </module>
        <module> Risetime1000LLL </module>
        <module> SdEventPosteriorSelectorOG </module>
        <module> EnergyCalculationPG </module>
        <module> UniversalityFitter </module>
      </try>

      <module> RecDataWriterNG </module>
    </loop>

  </moduleControl>

</sequenceFile>
```

Table D.2: Cuts for the selection of vertical SD-1500 events.

Cut name	Cut value	Meaning
!lightning		reject events caused by lightnings
minRecLevel	3	require reconstructed LDF
maxZenithSD	60	
T4Trigger	2	
T5Trigger	2	6T5 prior
badPeriodsRejectionFromFile		reject events in bad periods
timeInterval	params: 031231 150101	

Table D.3: FD quality cuts, usually paired with the SD-1500 quality cuts in Table D.2 for the analysis of golden hybrid events.

Cut name	Cut value	Meaning
hasMieDatabase		Information about atmospheric conditions is available
minLgEnergyFD	1e-20	
skipSaturated		Skip events that saturated the PMT camera
badFDPeriodRejection		Reject events in bad periods
!badPixels	1	Reject events with bad camera pixels
minMeanPixelRMS	17	
maxVAOD	0.1	Vertical aerosol depth below 0.1
LidarCloudRemoval	25	Cloud fraction not too large
MinCloudDepthDistance	params: -50 50	
MaxCloudThickness	100	
xMaxObsInExpectedFOV	params: 40 20	X_{max} reconstructed in expected FoV
xMaxError	40.0	
energyTotError	0.18	
profileChi2Sigma	2.5 -1.1	
maxDepthHole	20.	
maxCoreTankDist	750	
FidFOVICRC13	params: 40 20	Fiducial field of view cut (composition anti-bias cut)

Table D.4: FD quality cuts, usually paired with the SD-750 quality cuts in Table D.1 for the analysis of golden hybrid events.

Cut name	Cut value	Meaning
eyeCut	100000	Select only showers reconstructed with HECO
hasMieDatabase		Information about atmospheric conditions is available
minLgEnergyFD	1e-20	
skipSaturated		Skip events that saturated the PMT camera
badFDPeriodRejection		Reject events in bad periods
!badPixels	1	Reject events with bad camera pixels
minMeanPixelRMS	17	
maxVAOD	0.1	Vertical aerosol depth below 0.1
LidarCloudRemoval	25	Cloud fraction not too large
MinCloudDepthDistance	params: -50 50	
MaxCloudThickness	100	
xMaxObsInExpectedFOV	params: 40 20	$X_{\max}$ reconstructed in expected FoV
xMaxError	40.0	
energyTotError	0.18	
profileChi2Sigma	2.5 -1.1	
maxDepthHole	20.	
maxCoreTankDist	750	
FidFOVICRC13	params: 40 20	Fiducial field of view cut (composition anti-bias cut)

Listing D.2: "Offline module sequence for the reconstruction of golden hybrid events."

```
<sequenceFile>

  <enableTiming/>

  <moduleControl>

    <loop numTimes="unbounded" pushEventToStack="yes">

      <module> EventFileReaderOG </module>

      <!-- SdCalibratorOG and the SdEventSelectorOG have to be
           outside the loop in order to write only the events
           that pass the bottom up selection -->
      <module> SdQualityCutTaggerOG </module>
      <module> SdPMTQualityCheckerKG </module>
      <module> TriggerTimeCorrection </module>
      <module> SdCalibratorOG </module>
      <module> SdStationPositionCorrection </module>
      <module> SdBadStationRejectorKG </module>
      <module> SdSignalRecoveryKLT </module>
      <module> SdEventSelectorOG </module>

      <!-- FD reconstruction (SdCalibratorOG needs to have gotten a
                            shot to set tank times) -->
      <module> FdCalibratorOG </module>
      <module> FdPulseFinderOG </module>
      <module> PixelSelectorOG </module>
      <module> FdSDPFinderOG </module>
      <module> FdAxisFinderOG </module>
      <module> HybridGeometryFinderOG </module>
      <module> FdApertureLightOG </module>
      <module> FdProfileReconstructorKG </module>

      <!-- SD reconstruction -->
      <module> SdPlaneFitOG </module>
      <module> LDFFinderKG </module>
      <module> Risetime1000LLL </module>
      <module> DLECorrectionGG </module>
      <module> MuonProductionDepthFinderGL </module>
      <module> SdEventPosteriorSelectorOG </module>
      <module> EnergyCalculationPG </module>
      <module> UniversalityFitter </module>

      <module> RecDataWriterNG </module>
    </loop>

  </moduleControl>

</sequenceFile>
```

Listing D.3: "Offline bootstrap configuration of the module UniversalityFitter for the iterative reconstruction type."

```
<configLink id="UniversalityFitter">
  <UniversalityFitter>

    <minEnergy unit="EeV"> 0.1 </minEnergy>
    <activeMethod> Karlsruhe </activeMethod>

    <Calibration>
      <!-- Where the calibration constants come from -->
      <!-- [-3] Calib Data [-2] Data [-1] Photon
             [0] QGSJetII-03 [1] EPOS1.99
             [2] QGSJetII-04 [3] EPOS-LHC -->
      <CalibOpt> -3 </CalibOpt>
      <RecMixture> 2 </RecMixture>
    </Calibration>

    <KarlsruheReconstruction>
      <RecType> 2 </RecType>
      <TimeModelVersion> 2 </TimeModelVersion>
      <applyXmaxBiasCorrection> 0 </applyXmaxBiasCorrection>
      <verbosityLevel> 1 </verbosityLevel>
      <IterativeFit> 1 </IterativeFit>
      <Overrides>
        <doStartTimeFit> 0 </doStartTimeFit>
        <doSaturatedStartTimeFit> 0 </doSaturatedStartTimeFit>
      </Overrides>
    </KarlsruheReconstruction>

  </UniversalityFitter>
</configLink>
```

Listing D.4: "Offline bootstrap configuration of the module UniversalityFitter for the classic reconstruction type."

```
<configLink id="UniversalityFitter">
  <UniversalityFitter>

    <minEnergy unit="EeV"> 0.1 </minEnergy>
    <activeMethod> Karlsruhe </activeMethod>

    <Calibration>
      <!-- Where the calibration constants come from -->
      <!-- [-3] Calib Data [-2] Data [-1] Photon
             [0] QGSJetII-03 [1] EPOS1.99
             [2] QGSJetII-04 [3] EPOS-LHC -->
      <CalibOpt> -3 </CalibOpt>
      <RecMixture> 2 </RecMixture>
    </Calibration>

    <KarlsruheReconstruction>
      <RecType> 2 </RecType>
      <TimeModelVersion> 2 </TimeModelVersion>
      <applyXmaxBiasCorrection> 0 </applyXmaxBiasCorrection>
      <verbosityLevel> 1 </verbosityLevel>
      <IterativeFit> 0 </IterativeFit>
      <Overrides>
        <doStartTimeFit> 0 </doStartTimeFit>
        <doSaturatedStartTimeFit> 0 </doSaturatedStartTimeFit>
      </Overrides>
    </KarlsruheReconstruction>

  </UniversalityFitter>
</configLink>
```

Listing D.5: "Offline module sequence for the hybrid detector simulation of the fixed library (see Appendix E.1)."

```
<sequenceFile>

  <enableTiming/>

  <moduleControl>

    <loop numTimes="unbounded" pushEventToStack="yes">

      <module> EventFileReaderOG </module>
      <module> MCShowerCheckerOG </module>

      <loop numTimes="1" pushEventToStack="yes">

        <module> EventGeneratorOG </module>

        <!-- SD simulation part -->
        <loop numTimes="unbounded" pushEventToStack="no">
          <module> CachedShowerRegeneratorOG </module>
          <module> TabulatedTankSimulatorKG </module>
        </loop>

        <module> SdSimulationCalibrationFillerOG </module>

        <loop numTimes="unbounded" pushEventToStack="no">
          <module> SdPMTSimulatorOG </module>
          <module> SdFilterFADCSimulatorMTU </module>
        </loop>

        <module> SdBaselineSimulatorOG </module>
        <module> TankTriggerSimulatorOG </module>
        <module> TankGPSSimulatorOG </module>

        <try>
          <!-- FD simulation part -->
          <module> FdSimEventCheckerOG </module>
          <module> ShowerLightSimulatorKG </module>
          <module> LightAtDiaphragmSimulatorKG </module>
          <module> ShowerPhotonGeneratorOG </module>
          <module> TelescopeSimulatorKG </module>
          <module> FdBackgroundSimulatorOG </module>
          <module> FdElectronicsSimulatorOG </module>
          <module> FdTriggerSimulatorOG </module>
        </try>

        <!-- Trigger and Event builder -->
        <module> CentralTriggerSimulatorXb </module>
        <module> CentralTriggerEventBuilderOG </module>
        <module> EventBuilderOG </module>

        <!-- export simulation in Offline format -->
        <module> EventFileExporterOG </module>

      </loop>
    </loop>

  </moduleControl>

</sequenceFile>
```

Listing D.6: "Offline bootstrap file overrides for the hybrid detector simulations."

```
<configLink id="Atmosphere">
  <AtmosphereInterfaceConfig>
    <ProfileModel> SimShower </ProfileModel>
  </AtmosphereInterfaceConfig>
</configLink>

<configLink id="SdFilterFADCSimulator">
  <SdFilterFADCSimulator>
    <Dynode2AnodeDelay unit="ns"> 0.0 </Dynode2AnodeDelay>
    <StoreFilterSignals> yes </StoreFilterSignals>
    <StoreBaseSignals> yes </StoreBaseSignals>
  </SdFilterFADCSimulator>
</configLink>

<configLink id="TabulatedTankSimulator">
  <TabulatedTankSimulator>
    <signalSeparationMode> Universality </signalSeparationMode>
  </TabulatedTankSimulator>
</configLink>
```

Listing D.7: "Offline module sequence for the reconstruction of SD-750 hybrid simulations."

```
<sequenceFile>

  <enableTiming/>

  <moduleControl>

    <loop numTimes="unbounded" pushEventToStack="yes">

      <module> EventFileReaderOG </module>

      <loop numTimes="1" pushEventToStack="yes">
        <!-- Reconstruction -->
        <module> EventCheckerOG </module>

        <module> SdCalibratorOG </module>

        <!-- Hybrid reconstruction -->
        <try> <!-- limit how far a Continue goes -->
          <module> FdCalibratorOG </module>
          <module> FdEyeMergerKG </module>
          <module> FdPulseFinderOG </module>
          <module> FdSDPFinderOG </module>
          <module> FdAxisFinderOG </module>
          <module> HybridGeometryFinderOG </module>
          <module> HybridGeometryFinderWG </module>
          <module> FdApertureLightKG </module>
          <module> FdEnergyDepositFinderKG </module>
        </try>

        <!-- SD reconstruction -->
        <try> <!-- limit how far a Continue goes -->
          <module> SdEventSelectorOG </module>
          <module> SdMonteCarloEventSelectorOG </module>
          <module> SdPlaneFitOG </module>
          <module> LDFFinderKG </module>
          <module> Risetime1000LLL </module>
          <module> SdEventPosteriorSelectorOG </module>
        </try>

        <!-- export the ADST -->
        <module> RecDataWriterNG </module>

      </loop>
    </loop>

  </moduleControl>

</sequenceFile>
```

APPENDIX E

Monte Carlo air shower libraries

In this appendix, I will give an overview about the Monte Carlo shower libraries used within this work. If not mentioned otherwise, all libraries where created with the CORSIKA code for the simulation of EASs [194–196]. An optimized thinning level of $t = 10^{-6}$ was used [197, 198]. Only a few secondary particles with energies below $\frac{E_0}{t}$ are simulated further. They adopt a weight corresponding to the number of particles of similar type and energy that are dropped. E_0 is the energy of the primary particle. As a compromise between accuracy and time + space requirements, this thinning level became a standard setting for the simulation of EASs at the considered energies [199]. After the shower simulation and prior to the detector simulation, a statistical method called *shower resampling* is used to regain as much of the unthinned shower information is possible. This method is further described in [113, 199, 200]. In these references, it is also discussed that the thinning method, using the optimized thinning value of 10^{-6} up to primary energies of $E = 10^{20}$ eV, does not cause biases in the average signals or time distributions. However, additional thinning fluctuations cannot be avoided and need to be accounted for in analyses sensitive to these structures (in this work, relevant in Section 2.4).
The employed resampling method is implemented in the module CACHED SHOWER REGENERATOR within Offline. Reasonable values for the size of the resampling regions are chosen to be $\pm 0.1 \frac{\mathrm{d}r}{r}$ in radial direction and $\pm 15^\circ$ in azimuthal direction, the polar angle of the shower. A log-normal time smearing of the arrival times of particles is used to avoid large thinning fluctuations in time traces from high energy particles.

E.1 Fixed library

This is the main library I used for the parametrization of the time model in Section 2.4. It consists of simulations of the three primaries proton, carbon ($Z = 6$, $A = 12$) and iron ($Z = 26$, $A = 56$). Hadronic interaction models QGSJET-II.03 and EPOS-1.99 are used. For each mass and interaction model, showers with the following primary energies are available: 10^{17} eV, $10^{17.5}$ eV, 10^{18} eV, $10^{18.5}$ eV, 10^{19} eV, $10^{19.5}$ eV and 10^{20} eV (some showers at $10^{20.5}$ eV exist for testing purposes). Five zenith angles in the interval 0° to 60° and equidistantly distributed in $\sec\theta$ are chosen (0°, 37°, 48°, 55° and 60°). To account for the influence of a varying atmosphere on the shower development, monthly atmospheric models are used within CORSIKA [143, 192, 193]. 10 showers were simulated for each of the mentioned configurations. Due to the statistical nature of particle interactions, these

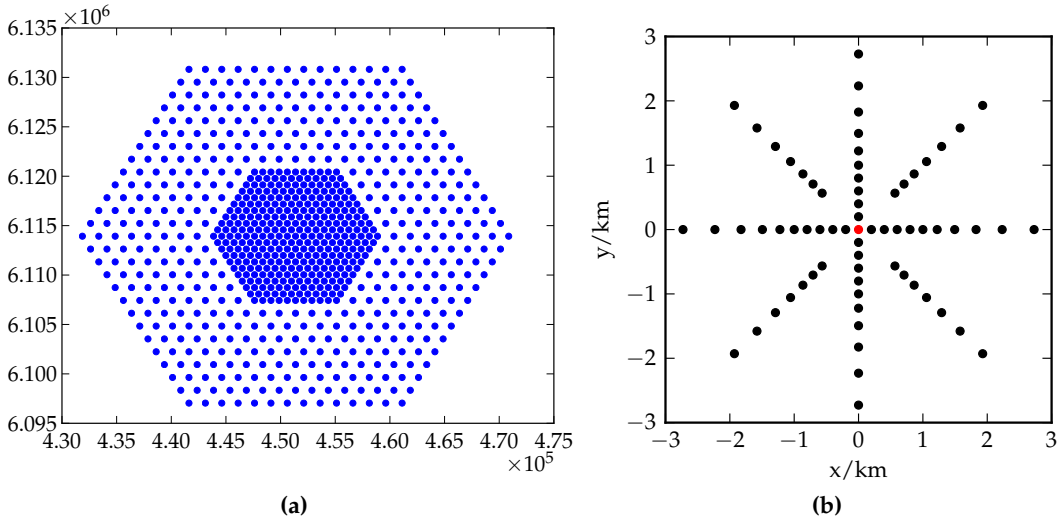


Figure E.1: (a) Schematic of the ideal list of stations used for the detector simulation. Easting and northing are plotted on the x- and y-axis, respectively. The inner part of nested stations represents an ideal SD-750 array, while the outher stations are an ideal standard SD. Station positions are given on ground. (b) Plot of the dense stations used in the simulation. The station positions are given in the shower coordinate system (see Fig. A.1a).

showers are not identical copies of each other. The differences are attributed to shower-to-shower fluctuations. Overall, this library consists of around 25 000 showers (requiring 25 TB of disk capacity). The library was simulated on a large computer cluster over the course of several months.

For each shower there exists an SD detector simulation in the ideal array as illustrated in Fig. E.1a. The core position was generated randomly within the elementary hexagon of the innermost stations. In addition, dense stations as depicted in Fig. E.1b are used. They are placed in rings with radius 200 m, 400 m, 600 m, 800 m, 1000 m, 1222 m, 1494 m, 1826 m, 2232 m and 2728 m centered around the Monte Carlo core. The first 3 rings contain 4 stations each and the rest 8. Detector simulations of FD are also available. The telescope positions were left unchanged with respect to the default ideal telescope list. The SD detector responses were simulated with the TABULATED TANK SIMULATOR, a tabulated version of the full GEANT4 simulation [107]. Differences in simulated photo-electron counts are below the percent level on average. The Offline module sequences used for simulation and reconstruction are listed in Appendix D.1. Particles were separated into components as described in Section 2.2. Simulated time traces are independently available for each of the components.

E.2 New fixed library

This library consists of simulations with the new interaction model QGSJET-II.04 and primaries proton and iron. All remaining parameters are chosen as for the library used in [97].

E.3 Continuous library

A continuous library in the energy range $10^{17.0}$ eV to 10^{20} eV with a spectral slope of E^{-1} is used for various studies, for example in Chapter 3 and Chapter 5. The primary species

proton and iron are available. Zenith angles are distributed according to $\frac{dN}{d\theta} \propto \sin\theta\cos\theta$ as in data. In total, there are around 4000 showers above 10^{19} eV. For most of the analyses in this work, showers in the energy range $10^{18.5}$ eV to 10^{20} eV are used, for example to quantify reconstruction biases and resolutions of the shower universality reconstruction in Section 3.4.5. SD detector simulations and SD-1500 reconstructions exist for each of the simulations.

E.4 New continuous library

A new continuous library for the primaries proton and iron, and in the energy range $10^{16.5}$ eV to $10^{19.5}$ eV was produced to advance studies for the SD-750 array. There are 2500 CORSIKA showers per primary particle and decade in energy. This library was simulated with the new hadronic interaction model QGSJET-II.04. Zenith angles are distributed according to $\frac{dN}{d\theta} \propto \sin\theta\cos\theta$. This distribution is identical to the one in data. An ideal SD-750 array and a dense array are used for the detector simulation. Each shower was simulated ten times with random positions in the SD-750 array. The total number of simulated events is 75 000 per primary particle. Due to a loss of trigger efficiency at the lowest energies, the actual number of triggered events is about 30 % smaller.

APPENDIX F

Mathematical and statistical remarks

F.1 Statistical distributions

This section provides a summary of relevant properties of some probability distribution functions [108] used within this work, for example in the parametrization of a model to describe the time-dependent signal in WCDs. Properties of the mostly used normal distribution are assumed to be well-known and will not be stated in this section. I will use to following terminology for a distribution f of the *variable* x and *parameters* a, b, c:

$$f(x|a,b,c) \tag{F.1}$$

F.1.1 Log-normal distribution

The log-normal distribution describes a random variable X if the logarithm $\log X$ of the variable is normally distributed. Within this work, this distribution is used to describe the arrival time distribution of secondary particles on ground.

The p.d.f. of the asymmetric distribution is given by

$$f(x|m,s) = \frac{1}{x\,s\sqrt{2\pi}}\, e^{-\frac{(\ln x - m)^2}{2s^2}}, \quad x > 0 \tag{F.2}$$

with $m \in \mathbb{R}$ and $s > 0$. The c.d.f. reads

$$F(x|m,s) = \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{\ln x - m}{s\sqrt{2}}\right)\right] = \frac{1}{2}\operatorname{erfc}\left(-\frac{\ln x - m}{s\sqrt{2}}\right) = \Phi\left(\frac{\ln x - m}{s}\right). \tag{F.3}$$

Φ denotes the c.d.f. of the normal distribution, which is directly related to the error-function erf as stated in the equation. In literature, it is common to denote the parameters of the log-normal distribution by μ and σ. I purposely avoid this to emphasize the fact that, unlike in the case of the normal distribution, the parameters are not equal to the first two central moments of the distribution. These are estimated to be

$$\mathrm{E}[X] = e^{m+\frac{1}{2}s^2}, \tag{F.4}$$

$$\mathrm{Var}[X] = (e^{s^2} - 1)e^{2m+s^2} = (e^{s^2} - 1)\ \mathrm{E}[X]^2. \tag{F.5}$$

Thus, it is easy to calculate the parameters from the first central moments and vice-versa. The functional form of the c.d.f. allows for an analytical inversion and an analytical expression of the percent point function or quantile function. This is relevant for convenient calculations of quantiles.

F.1.2 Generalized Gamma distribution

As the name suggests, the generalized gamma distribution results from a generalization of the gamma distribution. The latter has two free parameters and describes, for example, the distribution of time delay between two radioactive decays, both described with exponential distributions. The generalized gamma distribution has three free parameters and combines a family of probability distributions with two free parameters under certain parameter limits. The p.d.f. is given by

$$f(x) = \begin{cases} \frac{|\ell|}{s\,x}\,\frac{1}{\Gamma\left(\frac{1}{\ell^2}\right)}\exp\left[\frac{\ell\,\frac{\ln(x)-m}{s}+\ln\left(\frac{1}{\ell^2}\right)-\exp\left(l\,\frac{\ln(x)-m}{s}\right)}{\ell^2}\right] & ;\ \text{if } \ell \neq 0 \\ \frac{1}{t\,s\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{\ln(x)-m}{s}\right)^2\right) & ;\ \text{if } \ell = 0 \end{cases} \tag{F.6}$$

with time $x > 0$ and the three parameters m, s, ℓ. In the case of $\ell = 0$ the distribution is identical to the log-normal distribution (see Appendix F.1.1). The cumulative distribution function is the integral of the p.d.f.:

$$F(x) = \int_0^\infty f(x)\,\mathrm{d}x = \Gamma\left(\ell^{-2}, \frac{\exp\left(\ell\,(\log x - m)\,/s\right)}{\ell^2}\right) \tag{F.7}$$

This holds for $\ell \neq 0$. For $\ell = 0$ the c.d.f. is identical to the c.d.f. of the log-normal distribution as stated in Eq. (F.3). The moments of the distribution can be derived from the standard representation of this distribution. This exercise has not been explicitly done until now.

F.1.3 Generalized Gumbel distribution

The p.d.f. of the Generalized Gumbel distribution is

$$f(x|m,s,\ell) = \frac{\ell^\ell}{s\,\Gamma(\ell)}\exp\left[-\ell\left(\frac{x-m}{s}+\exp\left(-\frac{x-m}{s}\right)\right)\right] \tag{F.8}$$

for m, ℓ, $s > 0$ and $x \in \mathbb{R}$. The cumulative distribution function $F(x)$ can be found after some reasonable substitutions in the integration:

$$F(x|m,s,\ell) = 1 - \gamma\left(\ell,\, \ell\exp\left(-\frac{x-m}{s}\right)\right). \tag{F.9}$$

$\gamma\,(a,b)$ is the normalized lower incomplete Gamma function. The mean or first central moment of the distribution, given the random variable X, is

$$\mathrm{E}[X] = m + s\left(\ln\ell - \psi^{(0)}\ell\right), \tag{F.10}$$

while the variance is given by

$$\mathrm{Var}[X] = s^2\psi^{(1)}\ell. \tag{F.11}$$

The occurring ψ in both cases are polygamma functions $\psi^{(n)}(x) := \mathrm{d}^{n+1}\ln\Gamma(x)/\,\mathrm{d}^{n+1}x$. The incomplete gamma function γ is invertible with respect to the second argument. This allows an analytical calculation of the inverse c.d.f. and as such of quantiles of the distribution.

F.2 Notes on mathematical calculations

F.2.1 Median absolute deviation

The median absolute deviation MAD is used as a robust estimator of the unbiased standard deviation of a sample [201]. Having calculated the median μ of data, MAD is calculated as the median of the absolute deviations from μ:

$$\text{MAD} = \text{median}_i \left(|x_i - \mu| \right) . \tag{F.12}$$

Because the median describes the 50 % quantile, a scaling factor is needed to estimate the standard deviation of normally distributed data:

$$\sigma = 1.4826 \cdot \text{MAD} . \tag{F.13}$$

The scaling factor is given by $1/\left(\Phi^{-1}(3/4)\right)$, with the inverse of the c.d.f. of the normal distribution Φ^{-1}.
The MAD is a less efficient estimator of the standard deviation for perfectly normally distributed data. This means that the variance of the estimated values is larger. However, for contaminated data with outliers, the MAD is significantly more robust and provides a much more unbiased estimate of the standard deviation of the population.

F.2.2 Calculation of higher statistical moments

For variance models derived within this work, it is important to estimate the variance of the variance or the uncertainty of the uncertainty. Assuming roughly normally distributed data with variance V, fourth central moment m_4 and sample size n, an analytical formula for the estimation of the uncertainty of the uncertainty exists:

$$\sigma[\sigma] = \frac{1}{\sqrt{4V}} \sqrt{\frac{1}{n} \left(m_4 - \frac{n-3}{n-1} V^2 \right)} . \tag{F.14}$$

Another option used within this work is to bootstrap the distributions and the calculation of the standard deviation. The standard deviation of the deviations of the individual bootstrap estimates leads to an equivalent estimate of the uncertainty of the uncertainty.

F.2.3 Unbiased standard deviation

To estimate the standard deviation of a population from a random sample, oftentimes the sample standard deviation defined by $s = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (x_i - \overline{x})^2}$ is used. s^2 is an unbiased estimate of the true variance σ^2 (see e.g. [160]). The square root is a nonlinear function and only linear functions commute with applying the expectation operator. Thus, s is a biased estimate of the standard deviation and it can be shown that it is always an underestimate. Assuming a normal distribution, one can correct for this bias with a function $c_4(n)$ depending only on the sample size n:

$$c_4(n) = \sqrt{\frac{2}{n-1}} \frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{n-1}{2})} . \tag{F.15}$$

Γ is the gamma function. To unbias the standard deviation, the estimate s needs to be divided by $c_4(n)$ such that $\sigma = \text{E}[s]/c_4(n)$ holds. The correction is 20 percent for a sample size of 2 and less than 5 percent for sample sizes greater than 6.

F.2.4 Uncertainty calculation

This paragraph details the calculation of uncertainties or two-sided limits on parameters in a Maximum-likelihood fit. The starting point is the (approx.) knowledge of the optimal parameter vector $\vec{p}_0$ that maximizes the likelihood with a negative log-likelihood value of $-\log L_0$. The $1\,\sigma$ uncertainty interval of a parameter is then constructed by walking up the NLL until reaching the points of $-\log(L) + \frac{1}{2}$. For enough data points in the fit (significantly more than fit parameters), the distribution of $\log L$ converges towards a χ^2-distribution with k degrees of freedom (assuming k fit parameters). The value $1/2$ represents the $\alpha = 0.68$ percentile of $\chi^2_{(1)}$. A parameter interval of $2\,\sigma$ is constructed by walking the $-\log L$ up by 2.

F.2.5 The method of Feldman-Cousins

The method proposed by Feldman and Cousins provides a unified way to construct one or two-sided limits (uncertainties) for estimated parameters. The approach avoids the problem of flip-flopping and proper coverage is ensured [167]. Within this work, the method is mainly used to estimate uncertainties for the flux of UHECRs for bins with small entry numbers below 20.

CHAPTER 7

Acknowledgments

First, I want to express my gratitude towards Prof. Blümer and the Karlsruhe School of Elementary Particle and Astroparticle Physics for giving me the opportunity to study the physics of ultra-high energy cosmic rays at the KIT and work on this Ph.D. thesis within the Pierre Auger Collaboration. I am very thankful for several chances to visit the beautiful Argentina and other countries for workshops and collaboration meetings. My sincerest thanks go to Prof. Weber for accepting to co-referee this thesis at a very late stage of the work. I want to especially thank my advisor Markus Roth for his constant support throughout three and a half very fruitful years. I hope that the further development of the analysis methods will be successful, and that very interesting physics results will be gained. Thanks to all my colleagues at the Campus Nord for their support, many very interesting discussions and lots of nice social events. It was great working with you all. Thanks to Sabine Bucher for perfectly organizing everything. Also, I want to thank my colleagues of the Pierre Auger Collaboration for the good work we did together. I wish everybody a most successful future. The most heartfelt and special thanks go to Ariel Bridgeman for proofreading most of this thesis and providing very helpful feedback. Her support means the most to me. Many thanks to all other proofreaders as well, and to Darko Veberic for convincing me to use the English keyboard layout to write code (and lots of other advice). Finally, I cannot thank my family enough for their great support over all those years. I could not have reached my goals without them.

Bibliography

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

[2] A. D. Angelis, N. Giglietto, S. Stramaglia, Domenico Pacini, the forgotten pioneer of the discovery of cosmic rays, arXiv:1002.2888 (2010). arXiv:arXiv:1002.2888.

[3] J. Blümer, R. Engel, J. R. Hörandel, Cosmic Rays from the Knee to the Highest Energies, arXiv astro-ph/0904.0725 (2009).

[4] P. Auger, P. Ehrenfest, R. Maze, J. Daudin, R. A. Fréon, Extensive cosmic-ray showers, Rev. Mod. Phys. 11 (3-4) (1939) 288–291. doi:10.1103/RevModPhys.11.288.

[5] W. Heitler, The Quantum Theory of Radiation, third ed. Edition, Oxford University Press, Oxford, 1954.

[6] J.Matthews, A Heitler model of extensive air showers, Astroparticle Physics 22 (5-6) (2005) 387 – 397. doi:10.1016/j.astropartphys.2004.09.003.
URL http://www.sciencedirect.com/science/article/pii/S0927650504001598

[7] J. Linsley, Evidence for a primary cosmic-ray particle with energy 10^{20} ev, Phys. Rev. Lett. 10 (4) (1963) 146–148. doi:10.1103/PhysRevLett.10.146.

[8] K. Greisen, End to the cosmic-ray spectrum?, Phys. Rev. Lett. 16 (17) (1966) 748–750. doi:10.1103/PhysRevLett.16.748.

[9] G. T. Zatsepin, V. A. Kuz'min, Upper Limit of the Spectrum of Cosmic Rays, JETP Lett. 4 (1966) 78 – 80.

[10] D. M. Edge *et al.* (Haverah Park Collab.), The cosmic ray spectrum at energies above 10^{17} ev, J. Phys. A 6 (1973) 1612 – 1634.

[11] The High Resolution Fly's Eye Collaboration, Observation of the Ankle and Evidence for a High-Energy Break in the Cosmic Ray Spectrum, Phys.Lett.B619:271-280,2005 (2005). arXiv:arXiv:astro-ph/0501317, doi:10.1016/j.physletb.2005.05.064.

[12] P. Sokolsky for the HiRes Collaboration, Final Results from the High Resolution Fly's Eye (HiRes) Experiment, arXiv:1010.2690 (2010). arXiv:arXiv:1010.2690.

[13] N. Sakaki *et al.* (AGASA Collab.), Cosmic ray energy spectrum above 3×10^{18} eV observed with AGASA, in: Proc. 27th Int. Cosmic Ray Conf., Hamburg, Germany, 2001.

[14] E. Fermi, On the Origin of the Cosmic Radiation, Phys. Rev. 75 (1949) 1169–1174. doi:10.1103/PhysRev.75.1169.
URL http://link.aps.org/doi/10.1103/PhysRev.75.1169

[15] T. Antoni, W. Apel, F. Badea, K. Bekk, K. Bernloehr, H. Bluemer, E. Bollmann, H. Bozdog, I. Brancus, C. Buettner, A. Chilingarian, K. Daumiller, P. Doll, J. Engler, F. Fessler, H. Gils, R. Glasstetter, R. Haeusler, A. Haungs, D. Heck, T. Holst, J. Hoerandel, K.-H. Kampert, J. Kempa, H. Klages, J. Knapp, K. Koehler, G. Maier, H.-J. Mathes, H. Mayer, J. Milke, M. Mueller, J. Oehlschlaeger, M. Petcu, H. Rebel, M. Risse, M. Roth, G. Schatz, J. Scholz, S. Sokhoyan, T. Thouw, H. Ulrich, B. Vulpescu, J. Weber, J. Wentz, J. Wochele, J. Zabierowski, S. Zagromski, Muon density measurements with the KASCADE central detector, Astroparticle Physics 16 (4) (2002) 373 – 386. doi:10.1016/S0927-6505(01)00120-7.
URL http://www.sciencedirect.com/science/article/pii/S0927650501001207

[16] T. Antoni, W. Apel, A. Badea, K. Bekk, A. Bercuci, J. Blümer, H. Bozdog, I. Brancus, A. Chilingarian, K. Daumiller, P. Doll, R. Engel, J. Engler, F. Feßler, H. Gils, R. Glasstetter, A. Haungs, D. Heck, J. Hörandel, K.-H. Kampert, H. Klages, G. Maier, H. Mathes, H. Mayer, J. Milke, M. Müller, R. Obenland, J. Oehlschläger, S. Ostapchenko, M. Petcu, H. Rebel, A. Risse, M. Risse, M. Roth, G. Schatz, H. Schieler, J. Scholz, T. Thouw, H. Ulrich, J. van Buren, A. Vardanyan, A. Weindl, J. Wochele, J. Zabierowski, KASCADE measurements of energy spectra for elemental groups of cosmic rays: Results and open problems, Astroparticle Physics 24 (1-2) (2005) 1 – 25. doi:10.1016/j.astropartphys.2005.04.001.
URL http://www.sciencedirect.com/science/article/pii/S0927650505000691

[17] W. D. Apel, J. C. Arteaga-Velázquez, K. Bekk, M. Bertaina, J. Blümer, H. Bozdog, I. M. Brancus, P. Buchholz, E. Cantoni, A. Chiavassa, F. Cossavella, K. Daumiller, V. de Souza, F. Di Pierro, P. Doll, R. Engel, J. Engler, M. Finger, D. Fuhrmann, P. L. Ghia, H. J. Gils, R. Glasstetter, C. Grupen, A. Haungs, D. Heck, J. R. Hörandel, D. Huber, T. Huege, P. G. Isar, K.-H. Kampert, D. Kang, H. O. Klages, K. Link, P. Łuczak, M. Ludwig, H. J. Mathes, H. J. Mayer, M. Melissas, J. Milke, B. Mitrica, C. Morello, G. Navarra, J. Oehlschläger, S. Ostapchenko, S. Over, N. Palmieri, M. Petcu, T. Pierog, H. Rebel, M. Roth, H. Schieler, F. G. Schröder, O. Sima, G. Toma, G. C. Trinchero, H. Ulrich, A. Weindl, J. Wochele, M. Wommer, J. Zabierowski, KASCADE-Grande Collaboration, Kneelike Structure in the Spectrum of the Heavy Component of Cosmic Rays Observed with KASCADE-Grande, Phys. Rev. Lett. 107 (2011) 171104. doi:10.1103/PhysRevLett.107.171104.
URL http://link.aps.org/doi/10.1103/PhysRevLett.107.171104

[18] V. Berezinsky, Transition from galactic to extragalactic cosmic rays, arXiv:0710.2750v2 (10 2007).
URL http://arxiv.org/abs/0710.2750v2

[19] A. M. Hillas, Cosmic Rays: Recent Progress and some Current Questions, arXiv:astro-ph/0607109 (2006). arXiv:arXiv:astro-ph/0607109.
URL http://arxiv.org/abs/astro-ph/0607109v2

[20] M. Unger, G. R. Farrar, L. A. Anchordoqui, Origin of the ankle in the ultra-high energy cosmic ray spectrum and of the extragalactic protons below it (2015)arXiv:1505.02153.

[21] A. Aab, et al., Pierre Auger Collaboration, Depth of Maximum of Air-Shower Profiles at the Pierre Auger Observatory: Measurements at Energies above $10^{17.8}$ eV, Phys.Rev.D (2014)arXiv:1409.4809.

[22] K.-H. Kampert, P. Tinyakov, Cosmic rays from the ankle to the cutoff, Comptes Rendus Physique 15 (2014) 318–328. arXiv:1405.0575, doi:10.1016/j.crhy.2014.04.006.

[23] Alexander Schulz *et al.* (Pierre Auger Collab.), The measurement of the energy spectrum of cosmic rays above 0.3 EeV with the Pierre Auger Observatory, in: Proc. 33rd ICRC, (2013) 27, arXiv:1307.5059, Rio de Janeiro, Brazil, 2013.

[24] E. Kido, O. Kalashev, Telescope Array Collaboration, Constraining UHECR source models by the TA SD energy spectrum (2013)arXiv:1310.6093.

[25] R. Aloisio, V. Berezinsky, P. Blasi, Ultra high energy cosmic rays: implications of Auger data for source spectra and chemical composition, JCAP 1410 (10) (2014) 020. arXiv:1312.7459, doi:10.1088/1475-7516/2014/10/020.

[26] J. Matthews, A heitler model of extensive air showers, Astropart. Phys. 22 (5-6) (2005) 387–397.

[27] C. Meurer, J. Bluemer, R. Engel, A. Haungs, M. Roth, Muon production in extensive air showers and its relation to hadronic interactions, CZECH.J.PHYS. 56 (2006) A211. URL http://www.citebase.org/abstract?id=oai:arXiv.org:astro-ph/0512536

[28] P. Abreu, et al., Pierre Auger Collaboration, Interpretation of the Depths of Maximum of Extensive Air Showers Measured by the Pierre Auger Observatory, JCAP 1302 (2013) 026. arXiv:1301.6637, doi:10.1088/1475-7516/2013/02/026.

[29] R. Engel, D. Heck, T. Pierog, Extensive air showers and hadronic interactions at high energy, Ann. Rev. Nucl. Part. Sci. 61 (2011) 467–489. doi:10.1146/annurev.nucl.012809.104544.

[30] A. Aab, et al., The Pierre Auger Collaboration, The Pierre Auger Cosmic Ray Observatory (2015)arXiv:1502.01323.

[31] D. Ravignani *et al.* (Pierre Auger Collab.), Measurement of the energy spectrum of cosmic rays above 3×10^{17} eV using the AMIGA 750 m surface detector array of the Pierre Auger Observatory, in: Proc. 33rd Int. Cosmic Ray Conf., Rio de Janeiro, Brazil, 2013.

[32] I. Allekotte *et al.* (Pierre Auger Collab.), The surface detector system of the pierre auger observatory, Nucl. Instrum. Meth. A586 (2007) 409–420, arXiv astro-ph/0712.2832. doi:10.1016/j.nima.2007.12.016.

[33] J. Abraham *et al.* (Pierre Auger Collab.), Trigger and aperture of the surface detector array of the Pierre Auger Observatory, Nucl. Instrum. Meth. A613 (2010) 29–39.

[34] X. Bertou, P. Allison, C. Bonifazi, P. Bauleo, C. M. Grunfeld, M. Aglietta, F. Arneodo, D. Barnhill, J. Beatty, N. Busca, A. Creusot, D. Dornic, A. Etchegoyen, A. Filevitch, P. L. Ghia, I. Lhenry-Yvon, M. C. Medina, E. Moreno, D. Nitz, T. Ohnuki, S. Ranchon, H. Salazar, T. Suomijarvi, D. Supanitsky, A. Tripathi, M. Urban, L. Villasenor, Pierre Auger Collaboration, Calibration of the surface array of the Pierre Auger Observatory, Nucl. Instrum. Meth. A568 (2006) 839–846.

[35] D. Veberič *et al.* (Pierre Auger Collab.), Estimation of Signal in Saturated Stations of the Pierre Auger Surface Detector, in: Proc. 33rd Int. Cosmic Ray Conf., Rio de Janeiro, Brazil, 2013.

[36] M. Aglietta, I. D. Mitri, S. Maglio, S. Maldera, I. C. Maris, D. Martello, G. Navarra, M. Roth, Recovery of saturated signals of the surface detector, Auger internal note GAP-2008-030 (2008).

[37] M. Aglietta, I. De Mitri, S. Maglio, S. Maldera, I. C. Maris, D. Martello, G. Navarra, M. Roth, Recovery of Saturated Signals of the Surface Detector, Auger internal note GAP-2008-30 (2008).

[38] I. C. Mariş, Measurement of the ultra high energy cosmic ray flux using data of the pierre auger observatory, Auger internal note GAP-2008-026 (2008).

[39] D. Mockler, The first measurement of an energy spectrum at 0.1 EeV with the Surface Detector of the Pierre Auger Observatory, Master's thesis, Karlsruhe Institute of Technology, Karlsruhe (10 2014).

[40] J. Abraham *et al.* (Pierre Auger Collab.), The fluorescence detector of the Pierre Auger Observatory, Nucl. Instrum. Meth. A620 (2010) 227 – 251.

[41] M. Kleifges *et al.* (Pierre Auger Collab.), Extension of the Pierre Auger Observatory using high elevation fluorescence telescopes (HEAT), in: Proc. 31st Int. Cosmic Ray Conf., Lódź, Poland, 2009.

[42] J. F. Debatin, Investigation of optical properties of the fluorescence telescopes of the pierre auger observatory, Master's thesis, Karlsruhe Institute of Technology (May 2015).

[43] P. Abreu, et al., Pierre Auger Collaboration, The Rapid Atmospheric Monitoring System of the Pierre Auger Observatory, JINST 7 (2012) P09001. arXiv:1208.1675, doi:10.1088/1748-0221/7/09/P09001.

[44] P. Abreu, et al., Pierre Auger Collaboration, Description of Atmospheric Conditions at the Pierre Auger Observatory using the Global Data Assimilation System (GDAS), Astropart.Phys. 35 (2012) 591–607. arXiv:1201.2276, doi:10.1016/j.astropartphys.2011.12.002.

[45] F. Suarez *et al.* (Pierre Auger Collab.), The AMIGA muon detectors of the Pierre Auger Observatory, in: Proc. 33rd Int. Cosmic Ray Conf., Rio de Janeiro, Brazil, 2013.

[46] H. Falcke, et al., Detection and imaging of atmospheric radio flashes from cosmic ray air showers, Nature 435 (7040) (2005) 313–316. doi:10.1038/nature03614.

[47] S. Buitink, for the LOFAR Collaboration, Contribution no. 579, in: Proc. 33rd Int. Cosmic Ray Conf., Rio de Janeiro, Brazil, 2013.

[48] A. M. van den Berg *et al.* (Pierre Auger Collab.), Radio detection of cosmic rays at the southern auger observatory, in: Proc. 31st Int. Cosmic Ray Conf., Lódź, Poland, 2009.

[49] P. Abreu, et al., Pierre Auger Collaboration, Results of a self-triggered prototype system for radio-detection of extensive air showers at the Pierre Auger Observatory, JINST 7 (2012) P11023. arXiv:1211.0572, doi:10.1088/1748-0221/7/11/P11023.

[50] P. Abreu, et al., Pierre Auger Collaboration, Antennas for the Detection of Radio Emission Pulses from Cosmic-Ray, JINST 7 (2012) P10011. arXiv:1209.3840, doi:10.1088/1748-0221/7/10/P10011.

[51] A. Aab, et al., Pierre Auger Collaboration, Probing the radio emission from air showers with polarization measurements, Phys.Rev. D89 (2014) 052002. arXiv:1402.3677, doi:10.1103/PhysRevD.89.052002.

[52] J. Abraham *et al.* (Pierre Auger Collab.), Pierre Auger Collaboration, Limit on the diffuse flux of ultrahigh energy tau neutrinos with the surface detector of the pierre auger observatory, Phys. Rev. D 79 (10) (2009) 102001. doi:10.1103/PhysRevD.79.102001.

[53] J. Abraham *et al.* (Pierre Auger Collab.), Upper limit on the cosmic-ray photon fraction at eev energies from the pierre auger observatory, Astroparticle Physics 31 (6) (2009) 399 – 406. doi:DOI:10.1016/j.astropartphys.2009.04.003.
URL http://www.sciencedirect.com/science/article/B6TJ1-4W6Y349-1/2/fa58b1b24ace553d6d47e135f06059e1

[54] J. Abraham *et al.* (Pierre Auger Collab.), Measurement of the energy spectrum of cosmic rays above 10^{18} eV using the Pierre Auger Observatory, Physics Letters B 685 (4-5) (2010) 239 – 246. doi:DOI:10.1016/j.physletb.2010.02.013.
URL http://www.sciencedirect.com/science/article/B6TVN-4YC2XC7-3/2/69021ee5b6b08f24e4343776c82fb3ed

[55] J. Abraham *et al.* (Pierre Auger Collab.), Pierre Auger Collaboration, Measurement of the Depth of Maximum of Extensive Air Showers above 10^{18} eV, Phys. Rev. Lett. 104 (9) (2010) 091101. doi:10.1103/PhysRevLett.104.091101.

[56] P. Abreu, et al., Pierre Auger Collaboration, A Search for Ultra-High Energy Neutrinos in Highly Inclined Events at the Pierre Auger Observatory, Phys.Rev. D84 (2011) 122005. arXiv:1202.1493, doi:10.1103/PhysRevD.85.029902,10.1103/PhysRevD.84.122005.

[57] P. Abreu, et al., Pierre Auger Collaboration, The Lateral Trigger Probability function for the ultra-high energy cosmic ray showers detected by the Pierre Auger Observatory, Astropart.Phys. 35 (2011) 266–276. arXiv:1111.6645, doi:10.1016/j.astropartphys.2012.02.005,10.1016/j.astropartphys.2011.08.001.

[58] P. Abreu, et al., Pierre Auger Collaboration, Anisotropy and chemical composition of ultra-high energy cosmic rays using arrival directions measured by the Pierre Auger Observatory, JCAP 1106 (2011) 022. arXiv:1106.3048, doi:10.1088/1475-7516/2011/06/022.

[59] P. Abreu, et al., Pierre Auger Collaboration, The effect of the geomagnetic field on cosmic ray energy estimates and large scale anisotropy searches on data from the Pierre Auger Observatory, JCAP 1111 (2011) 022. arXiv:1111.7122, doi:10.1088/1475-7516/2011/11/022.

[60] P. Abreu, et al., Pierre Auger Collaboration, Advanced functionality for radio analysis in the Offline software framework of the Pierre Auger Observatory, Nucl.Instrum.Meth. A635 (2011) 92–102. arXiv:1101.4473, doi:10.1016/j.nima.2011.01.049.

[61] P. Abreu, et al., Pierre Auger Collaboration, Measurement of the Cosmic Ray Energy Spectrum Using Hybrid Events of the Pierre Auger Observatory, Eur.Phys.J.Plus 127 (2012) 87. arXiv:1208.6574, doi:10.1140/epjp/i2012-12087-9.

[62] P. Abreu, et al., Pierre Auger Observatory Collaboration, A Search for Point Sources of EeV Neutrons, Astrophys.J. 760 (2012) 148. arXiv:1211.4901, doi:10.1088/0004-637X/760/2/148.

[63] P. Abreu, et al., Pierre Auger Collaboration, Large scale distribution of arrival directions of cosmic rays detected above 10^{18} eV at the Pierre Auger Observatory, Astrophys.J.Suppl. 203 (2012) 34. arXiv:1210.3736, doi:10.1088/0067-0049/203/2/34.

[64] P. Abreu, et al., Pierre Auger Collaboration, Search for point-like sources of ultra-high energy neutrinos at the Pierre Auger Observatory and improved limit on the diffuse flux of tau neutrinos, Astrophys.J. 755 (2012) L4. arXiv:1210.3143, doi:10.1088/2041-8205/755/1/L4.

[65] P. Abreu, et al., Pierre Auger Collaboration, Constraints on the origin of cosmic rays above 10^{18} eV from large scale anisotropy searches in data of the Pierre Auger Observatory, Astrophys.J. 762 (2012) L13. arXiv:1212.3083, doi:10.1088/2041-8205/762/1/L13.

[66] P. Abreu, et al., A search for anisotropy in the arrival directions of ultra high energy cosmic rays recorded at the Pierre Auger Observatory, JCAP (2012).

[67] P. Abreu, et al., Pierre Auger Collaboration, Search for signatures of magnetically-induced alignment in the arrival directions measured by the Pierre Auger Observatory, Astropart.Phys. 35 (2012) 354–361. arXiv:1111.2472, doi:10.1016/j.astropartphys.2011.10.004.

[68] P. Abreu, et al., Pierre Auger Collaboration, Ultrahigh Energy Neutrinos at the Pierre Auger Observatory, Adv.High Energy Phys. 2013 (2013) 708680. arXiv:1304.1630, doi:10.1155/2013/708680.

[69] P. Abreu, et al., Pierre Auger Collaboration, Techniques for Measuring Aerosol Attenuation using the Central Laser Facility at the Pierre Auger Observatory, JINST 8 (2013) P04009. arXiv:1303.5576, doi:10.1088/1748-0221/8/04/P04009.

[70] P. Abreu, et al., Pierre Auger, Identifying Clouds over the Pierre Auger Observatory using Infrared Satellite Data, Astropart.Phys. 50-52 (2013) 92–101. arXiv:1310.1641, doi:10.1016/j.astropartphys.2013.09.004.

[71] P. Abreu, et al., Pierre Auger Collaboration, Bounds on the density of sources of ultra-high energy cosmic rays from the Pierre Auger Observatory, JCAP 1305 (05) (2013) 009. arXiv:1305.1576, doi:10.1088/1475-7516/2013/05/009.

[72] A. Aab, et al., Pierre Auger Collaboration, Origin of atmospheric aerosols at the Pierre Auger Observatory using studies of air mass trajectories in South America, Atmos.Res. (2014)arXiv:1405.7551, doi:10.1016/j.atmosres.2014.05.021.

[73] A. Aab, et al., Pierre Auger Collaboration, Muons in air showers at the Pierre Auger Observatory: Mean number in highly inclined events, Phys.Rev.D (2014)arXiv:1408.1421.

[74] A. Aab, et al., Telescope Array Collaboration, Pierre Auger Collaboration, Searches for Large-Scale Anisotropy in the Arrival Directions of Cosmic Rays Detected above Energy of 10^{19} eV at the Pierre Auger Observatory and the Telescope Array, Astrophys.J. 794 (2) (2014) 172. arXiv:1409.3128, doi:10.1088/0004-637X/794/2/172.

[75] A. Aab, et al., Pierre Auger Collaboration, Reconstruction of inclined air showers detected with the Pierre Auger Observatory, JCAP 1408 (08) (2014) 019. arXiv:1407.3214, doi:10.1088/1475-7516/2014/08/019.

[76] A. Aab, et al., Pierre Auger Collaboration, Muons in air showers at the Pierre Auger Observatory: Measurement of atmospheric production depth, Phys.Rev. D90 (2014) 012012. arXiv:1407.5919, doi:10.1103/PhysRevD.90.012012,10.1103/PhysRevD.90.039904.

[77] A. Aab, et al., Pierre Auger Collaboration, Search for patterns by combining cosmic-ray energy and arrival directions at the Pierre Auger Observatory, Eur.Phys.J.C (2014)arXiv:1410.0515.

[78] A. Aab, et al., Pierre Auger Collaboration, A Targeted Search for Point Sources of EeV Neutrons, Astrophys.J. 789 (2014) L34. arXiv:1406.4038, doi:10.1088/2041-8205/789/2/L34.

[79] A. Aab, et al., Pierre Auger Collaboration, A search for point sources of EeV photons, Astrophys.J. 789 (2014) 160. arXiv:1406.2912, doi:10.1088/0004-637X/789/2/160.

[80] A. Aab, et al., Pierre Auger Collaboration, Depth of Maximum of Air-Shower Profiles at the Pierre Auger Observatory: Composition Implications, Phys.Rev.D (2014)arXiv:1409.5083.

[81] A. Aab, et al., Pierre Auger, Large Scale Distribution of Ultra High Energy Cosmic Rays Detected at the Pierre Auger Observatory With Zenith Angles up to 80 degrees, Astrophys. J. 802 (2) (2015) 111. arXiv:1411.6953, doi:10.1088/0004-637X/802/2/111.

[82] A. Aab, et al., Pierre Auger, Searches for Anisotropies in the Arrival Directions of the Highest Energy Cosmic Rays Detected by the Pierre Auger Observatory, Astrophys. J. 804 (1) (2015) 15. arXiv:1411.6111, doi:10.1088/0004-637X/804/1/15.

[83] A. Aab, et al., Pierre Auger, Muons in air showers at the Pierre Auger Observatory: Mean number in highly inclined events, Phys. Rev. D91 (3) (2015) 032003, [Erratum: Phys. Rev.D91,no.5,059901(2015)]. arXiv:1408.1421, doi:10.1103/PhysRevD.91.059901,10.1103/PhysRevD.91.032003.

[84] A. Aab, et al., Pierre Auger, Search for patterns by combining cosmic-ray energy and arrival directions at the Pierre Auger Observatory, Eur. Phys. J. C75 (6) (2015) 269. arXiv:1410.0515, doi:10.1140/epjc/s10052-015-3471-0.

[85] A. Aab, et al., Pierre Auger, Improved limit to the diffuse flux of ultrahigh energy neutrinos from the Pierre Auger Observatory, Phys. Rev. D91 (9) (2015) 092008. arXiv:1504.05397, doi:10.1103/PhysRevD.91.092008.

[86] I. Valino *et al.* (Pierre Auger Collab.), The flux of ultra-high energy cosmic rays after ten years of operation of the Pierre Auger Observatory, in: Proc. 34th Int. Cosmic Ray Conf., The Hague, The Netherlands, 2015.

[87] The Pierre Auger Collaboration, Testing hadronic interactions at ultrahigh energy using air showers measured by the Pierre Auger Observatory, to be published (2015).

[88] P. Lipari, The Concepts of 'Age' and 'Universality' in Cosmic Ray Showers, Phys. Rev. 79 (2008) 063001. arXiv:0809.0190.

[89] A. M. Hillas, Angular and energy distributions of charged particles in electron photon cascades in air, J. Phys. G8 (1982) 1461–1473. doi:10.1088/0305-4616/8/10/016.

[90] M. Giller, A. Kacperczyk, J. Malinowski, W. Tkaczyk, G. Wieczorek, Similarity of extensive air showers with respect to the shower age, J. Phys. G31 (2005) 947–958.

[91] F. Nerling, J. Blümer, R. Engel, M. Risse, Universality of electron distributions in high-energy air showers: Description of cherenkov light production, Astropart. Phys. 24 (2006) 421–437. arXiv:astro-ph/0506729.

[92] P. Lipari, Universality of cosmic ray shower development, Nucl. Phys. Proc. Suppl. 196 (2009) 309–318. doi:10.1016/j.nuclphysbps.2009.09.060.

[93] S. Lafebre, R. Engel, H. Falcke, J. Hörandel, T. Huege, J. Kuijpers, R. Ulrich, Universality of electron-positron distributions in extensive air showers, Astropart. Phys. 31 (2009) 243–254. arXiv:0902.0548.

[94] F. Schmidt, M. Ave, L. Cazon, A. Chou, Applying extensive air shower universality to ground detector data, in: Proc. 30th Int. Cosmic Ray Conf., Pierre Auger Collaboration, Mérida, Mexico, 2007.

[95] M. Ave, R. Engel, J. Gonzalez, D. Heck, T. Pierog, M. Roth, Extensive air shower universality of ground particle distributions, Proc. of 31st Int. Cosmic Ray Conf., Beijing (2011) #1025.

[96] D. Maurel, M. Roth, J. Gonzalez, Universality of the time structure of ground particle distributions and its application to the reconstruction of extensive air showers, in: ICRC Proceedings 2013, 2013.

[97] M. Ave, R. Engel, J. Gonzalez, D.Heck, D. Maurel, T. Pierog, M. Roth, Prediction of the tank response $S(r, \mathrm{DX}, E|\theta, \phi)$ from shower universality, Auger internal note GAP-2011-087 (2011).

[98] M. T. Dova, L. N. Epele, A. G. Mariazzi, The Effect of atmospheric attenuation on inclined cosmic ray air showers, Astropart. Phys. 18 (2003) 351–365. arXiv:astro-ph/0110237, doi:10.1016/S0927-6505(02)00150-0.

[99] M. Ave, N. Busca, L. Cazon, F. Schmidt, T. Yamamoto, Can EPOS Reproduce the Auger SD and Hybrid Data?, Auger internal note GAP-2007-098 (2007).

[100] M. Ave, Ground signal parameterization for water Cherenkov and scintillator detectors and performance of ground arrays that combine both techniques, Auger internal note GAP-2012-141 (2012).

[101] D. Maurel, Mass composition of ultra-high energy cosmic rays based on air shower universality, Dissertation, Karlsruhe Institute of Technology (2013).

[102] D. Maurel, M. Ave, J. Gonzalez, M. Roth, Extended shower universality: A model of the time-dependent signal in water Cerenkov tanks, Auger internal note GAP-2013-022 (2013).

[103] M. Ave, X. Bertou, R. Engel, M. Roth, A model for the time-dependent signal in water Cherenkov detectors, Auger internal note GAP-2013-072 (2013).

[104] F. Schmidt, M. Ave, L. Cazon, A. S. Chou, A Model-Independent Method of Determining Energy Scale and Muon Number in Cosmic Ray Surface Detectors, Astropart. Phys. 29 (2008) 355–365. arXiv:0712.3750.

[105] B. Kégl, D. Veberič, Single Muon Response: Tracklength, Auger internal note GAP-2009-043 (2009).

[106] B. Kégl, R. Bardenet, D. Veberič, Single Muon Response: The signal model, Auger internal note GAP-2010-110 (2010).

[107] M. Ave, R. Engel, J. Gonzalez, D. Maurel, M. Roth, Tabulated response of the Auger surface detector, Auger internal note GAP-2011-086 (2011).

[108] C. Walck, Hand-book on statistical distributions for experimentalists, University of Stockholm, Stockholm, 2007.

[109] NLopt - library for non-linear optimization, http://ab-initio.mit.edu/wiki/.
URL http://ab-initio.mit.edu/wiki/index.php/NLopt_manual

[110] F. James, Minuit Reference Manual, CERN Program Library Office CERN-IT Division CH-1211 Geneva 23 Switzerland (1998).

[111] M. Erfani, M. Risse, M. Settimo, A. Yushkov, Simulation Study of Muons in Air Showers (2), Auger internal note GAP-2014-051 (2014).

[112] D. Garcia-Gamez, L. Molina, The resampling algorithm in Offline: bias in the time of the particles injected in the surface detectors, Auger internal note GAP-2015-006 (2015).

[113] P. Billoir, A sampling procedure to regenerate particles in a ground detector from a 'thinned' air shower simulation output, Astropart. Phys. 30 (2008) 270–285. doi:10.1016/j.astropartphys.2008.10.002.

[114] M. Ave, I. Allekotte, X. Bertou, R. Engel, M. Roth, Preliminary results for the Xmax reconstruction using SD vertical events, Auger internal note GAP-2013-075 (2013).

[115] M. Ave, X. Bertou, R. Engel, D. Maurel, M. Roth, A. Schulz, (Energy,Nmu,Xmax) from the Bariloche reconstruction, Auger internal note GAP-2013-109 (2013).

[116] S. Argirò, S. Barroso, J. Gonzalez, L. Nellen, T. Paul, T. Porter, L. P. Jr., M. Roth, R. Ulrich, D. Veberič, The offline software framework of the Pierre Auger Observatory, Nucl. Instrum. Meth. A 580 (2007) 1485–1496. doi:http://dx.doi.org/10.1016/j.nima.2007.07.010.
URL http://www.sciencedirect.com/science/article/pii/S0168900207014106

[117] P. Billoir, Proposition to improve the local trigger of the Surface Detector for low energy showers, Auger internal note GAP-2009-179 (2009).

[118] P. Billoir, P. Ghia, D. Nitz, R. Sato, First results of the ToTd trigger in the test hexagon, Auger internal note GAP-2011-032 (2011).

[119] L. M. Bueno, P. Billoir, I. C. Maris, Signal variance for the TOTd and MoPS triggers, Auger internal note GAP-2014-035 (2014).

[120] D. Veberič, M. Roth, Offline reference manual: Sd reconstruction, Auger internal note GAP-2005-035, an updated version is shipped with the Offline software distribution (2005).

[121] R. Hiller, M. Roth, An update on the signal accuracy using the infill array, Auger internal note GAP-2012-012 (2012).

[122] M. Horvat, D. Veberič, On shower-front start-time variance, Auger internal note GAP-2007-057 (2007).

[123] I. C. Mariş, M. Roth, T. Schmidt, A Phenomenological Method to Recover the Signal from Saturated Stations, Auger internal note GAP-2006-012 (2006).

[124] K. Kamata, J. Nishimura, The Lateral and the Angular Structure Functions of Electron Showers, Prog. Theoret. Phys. Suppl. 6 (1958) 93–155.

[125] Greisen, Progress in elementary particle and cosmic ray physics, Vol. 3, North-Holland Publishing, Amsterdam, 1956.

[126] D. Newton, J. Knapp, A. A. Watson, The Optimum Distance at which to Determine the Size of a Giant Air Shower, Astropart. Phys. 26 (2007) 414–419.

[127] A. Schulz, Measurement of the Energy Spectrum of Cosmic Rays between 0.1 EeV and 30 EeV with the Infill Extension of the Surface Detector of the Pierre Auger Observatory, Master's thesis, Karlsruhe Institute of Technology (2012).

[128] P. Billoir, A. Letessier-Selvon, C. Macolino, Multi event fit for the LDF, Auger collaboration meeting (2011).

[129] T. Schmidt, I. C. Mariş, M. Roth, Fine tuning of the ldf parameterisation and the influence on s_{1000}, Auger internal note GAP-2007-106 (2007).

[130] M. Ave, Private Communication about updates to the shower universality reconstruction (12 2014).

[131] D. Mockler, private communication (2015).

[132] G. Babu, A note on bootstrapping the variance of sample quantile, Ann. Inst. Statist. Math 38 (Part A) (1986) 439–443.

[133] A. Davison, D. Hinkley, Bootstrap methods and their application, Cambridge University Press, New York, 1997.

[134] H. Dembinski, J. Gonzalez, V. Olmos-Gilbaja, I. Valiño Rielo, M. Roth, T. Schmidt, Reconstruction of very inclined air showers with offline, Auger internal note GAP-2009-008 (2009).

[135] H. Dembinski, Measurement of the Ultra High Energy Cosmic Ray Flux from Data of very inclined showers at the Pierre Auger Observatory, Ph.D. thesis, RWTH Aachen University, Aachen, Germany (2009).

[136] G. Rodriguez *et al.* (Pierre Auger Collab.), Reconstruction of inclined showers at the Pierre Auger Observatory: implications for the muon content, in: Proc. 32nd Int. Cosmic Ray Conf., Beijing, China, 2011.

[137] I. Lhenry-Yvon, P. Ghia, The communications crisis in 2009: proposal for handling data between June and November, Auger internal note GAP-2010-020 (2010).

[138] S. Dagoret-Campagne, The central trigger user guide and reference manual, Auger internal note GAP-2004-020 (2004).

[139] A. Aab, et al., The pierre auger cosmic ray observatory, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment (2015) –doi:http://dx.doi.org/10.1016/j.nima.2015.06.058.
URL http://www.sciencedirect.com/science/article/pii/S0168900215008086

[140] Official web page of the Auger Observer, http://augerobserver.fzk.de/doku.php?id=datatree:root.

[141] Official web page of the Acceptance Group of the Pierre Auger Collaboration, http://ipnwww.in2p3.fr/~augers/AugerProtected/AcceptMain.html.

[142] I. A. Samarai, O. Deligny, I. Lhenry-Yvon, Control of the event rate of the surface detector arrays, Auger internal note GAP-2015-001 (2015).

[143] B. Keilhauer, Investigation of atmospheric effects on the development of extensive air showers and their detection with the pierre auger observatory, Ph.D. thesis, Karlsruhe University, Karlsruhe, Germany (2003).

[144] C. Bleve, Atmospheric effects on extensive air showers observed with the surface detector of the pierre auger observatory, preprint submitted to Elsevier (12 2008).

[145] D. Harari, S. Mollerach, E. Roulet, On the implementation of the energy correction due to weather effects, Auger internal note GAP-2008-175 (2008).

[146] C. Bleve *et al.* (Pierre Auger Collaboration), Weather induced effects on extensive air showers observed with the surface detector of the pierre auger observatory, in: 30th Proc. Int. Cosm. Ray Conf., Merida, Mexico, 2007.

[147] A. Remoto, S. Mollerach, E. Roulet, On the weather correction and energy calibration, Auger internal note GAP-2009-054 (2009).

[148] E. Roulet, I. Allekotte, D. Harari, S. Mollerach, On the Constant Intensity Cut and weather effects, Auger internal note GAP-2009-024 (2009).

[149] S. Mollerach, E. Roulet, O. Taborda, Revisiting the effects of weather on the air shower energy reconstruction, Auger internal note GAP-2014-047 (2014).

[150] S. Mollerach, E. Roulet, O. Taborda, Weather effects on the event rate of the Infill array, Auger internal note GAP-2014-105 (2014).

[151] J. Hersil, I. Escobar, D. Scott, G. Clark, S. Olbert, Observations of Extensive Air Showers near the Maximum of Their Longitudinal Development, Phys. Rev. Lett. 6 (1961) 22–23. doi:10.1103/PhysRevLett.6.22.

[152] T. K. Gaisser, A. M. Hillas, Reliability of the Method of Constant Intensity Cuts for Reconstructing the Average Development of Vertical Showers, in: Proc. 15th Int. Cosm. Ray Conf., Vol. 8, Plovdiv, Bulgaria, 1977, p. 353.

[153] M. Healy, K. Arisaka, D. Barnhill, J. Lee, T. Ohnuki, A. Tripathi, Applying the Constant Intensity Cut to Determine Composition, Energy, and Muon Richness, Auger internal note GAP-2006-020 (2006).

[154] A. Rivera, S. Mollerach, E. Roulet, On the energy dependence of the Constant Intensity Cut, Auger internal note GAP-2009-123 (2009).

[155] H. Dembinski, M. Roth, Constant Intensity Cut method revisited: Uncertainty calculation with the Bootstrap, Auger internal note GAP-2011-074 (2011).

[156] H. Dembinski, M. Roth, Constant intensity cut: Matching the energy scales of vertical and inclined showers, Auger internal note GAP-2012-069 (2012).

[157] D. Veberič, B. Kégl, R. Engel, M. Roth, Constant Intensity Cut: Unbinned Estimation of the Signal Attenuation Function, Auger internal note GAP-2015-065 (2015).

[158] J. P. Roberts, G. R. Farrar, Another look at energy reconstruction systematics, Auger internal note GAP-2012-096 (2012).

[159] A. Schulz, H. Dembinski, M. Roth, Time dependence of the SD energy scale: Analysis with an extended CIC, Auger internal note GAP-2013-039 (2013).

[160] F. James, Statistical Methods in Experimental Physics, 2nd Edition, World Scientific Publishing, Singapore, 2006.

[161] H. Dembinski, T. Hebbeker, T. Schmidt, Determination of the energy calibration function and primary composition of very inclined showers, Auger internal note GAP-2009-048 (2009).

[162] H. Dembinski, M. Roth, A. Schulz, Energy calibration revisited: The simplified likelihood method, Auger internal note GAP-2012-090 (2012).

[163] H. P. Dembinski, B. Kégl, I. C. Mariş, M. Roth, D. Veberič, A likelihood method to cross-calibrate air-shower detectors (2015)arXiv:1503.09027.

[164] S. Y. BenZvi *et al.* (Pierre Auger Collab.), The LIDAR system of the Pierre Auger observatory, Nucl. Instrum. Meth. A574 (2007) 171–184, arXiv astro-ph/0609063. doi:10.1016/j.nima.2007.01.094.

[165] A. Porcelli, Measurement of the depth of shower maximum in the transition region between galactic and extragalactic cosmic rays with the pierre auger observatory, Ph.D. thesis, Karlsruhe Institute of Technology (2014).

[166] L. Cazon, R. Ulrich, The Non-Linearity Between $\langle \ln A \rangle$ and $\langle X_{\max} \rangle$ Induced by the Acceptance of Fluorescence Telescopes, Auger internal note GAP-2012-015 (2012).

[167] G. J. Feldman, R. D. Cousins, Unified approach to the classical statistical analysis of small signals, Phys. Rev. D 57 (1998) 3873 – 3889.

[168] V. Verzi *et al.* (Pierre Auger Collab.), The Energy Scale of the Pierre Auger Observatory, in: Proc. 33rd Int. Cosmic Ray Conf., Rio de Janeiro, Brazil, 2013.

[169] M. Fukushima, Telescope Array, Recent Results from Telescope Array, EPJ Web Conf. 99 (2015) 04004. arXiv:1503.06961, doi:10.1051/epjconf/20159904004.

[170] A. Schulz, H. Dembinski, M. Roth, Search for a declination dependence in the flux of cosmic rays, Auger internal note GAP-2014-013 (2014).

[171] T. Abu-Zayyad et al., The Cosmic Ray Energy Spectrum Observed with the Surface Detector of the Telescope Array Experiment, Astrophysical Journal 768 (2013).

[172] P. Sommers, Cosmic ray anisotropy analysis with a full-sky observatory, Astropart. Phys. 14 (2001) 271–286.

[173] T. Schmidt, I. C. Mariş, M. Roth, LDF systematics in S_{38} from the constant intensity cut method, Auger internal note GAP-2006-070 (2006).

[174] D. Mockler, private communication (2015).

[175] V. Blobel, Unfolding methods in high energy physics experiments, in: Proceedings of the 1984 CERN School of Computing, CERN 85-09 and DESY 84-111, CERN, Switzerland, 1984.

[176] V. Blobel, An unfolding method for high energy physics experiments, Tech. Rep. 02-078, DESY, arXiv hep-ex/0208022v1 (2002).

[177] V. Blobel, The RUN Manual - Regularized Unfolding for High-Energy Physics Experiments, Tech. Rep. TN361, OPAL, DESY, Germany (1996).

[178] H. P. Dembinski, M. Roth, ARU - towards automatic unfolding of detector effects, in: H. B. Prosper, L. Lyons (Eds.), Proceedings of PHYSTAT 2011 (CERN-2011-006), CERN, Geneva, Switzerland, 2011, pp. 285 – 291.

[179] S. G. Johnson, The NLopt nonlinear-optimization package, http://ab-initio.mit.edu/nlopt.

[180] M. Tueros *et al.* (Pierre Auger Collab.), Estimate of the non-calorimetric energy of showers observed with the fluorescence and surface detectors, in: Proc. 33rd Int. Cosmic Ray Conf., Rio de Janeiro, Brazil, 2013.

[181] J. S. Conway, Incorporating Nuisance Parameters in Likelihoods for Multisource Spectra, in: Proceedings, PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding, CERN,Geneva, Switzerland 17-20 January 2011, 2011. arXiv:1103.0354.
URL http://inspirehep.net/record/891252/files/arXiv:1103.0354.pdf

[182] S. Schoo *et al.* (KASCADE-Grande collaboration), The energy spectrum of cosmic rays in the range from 10×10^{14} eV to 10×10^{18} eV, in: Proc. 34th Int. Cosmic Ray Conf., The Hague, The Netherlands, 2015.

[183] K. Rawlins *et al.* (IceCube collaboration), Latest Results on Cosmic Ray Spectrum and Composition from Three Years of IceTop and IceCube, in: Proc. 34th Int. Cosmic Ray Conf., The Hague, The Netherlands, 2015.

[184] R. Aloisio, D. Boncioli, A. F. Grillo, S. Petrera, F. Salamida, SimProp: a Simulation Code for Ultra High Energy Cosmic Ray Propagation, JCAP 1210 (2012) 007. arXiv:1204.2970, doi:10.1088/1475-7516/2012/10/007.

[185] K.-H. Kampert, J. Kulbartz, L. Maccione, N. Nierstenhoefer, P. Schiffer, G. Sigl, A. R. van Vliet, CRPropa 2.0 – a Public Framework for Propagating High Energy Nuclei, Secondary Gamma Rays and Neutrinos, Astropart. Phys. 42 (2013) 41–51. arXiv:1206.3132, doi:10.1016/j.astropartphys.2012.12.001.

[186] Eun-Joo Ahn *et al.* (Pierre Auger Collab.), Inferences about the mass composition of cosmic rays from data on the depth of maximum at the Auger Observatory, in: Proc. 33rd Int. Cosmic Ray Conf., Rio de Janeiro, Brazil, 2013.

[187] M. Ave, M. Roth, A. Schulz, A universal description of temporal and lateral distributions of ground particles in extensive air showers, Proc. of 34th Int. Cosmic Ray Conf., The Hague, The Netherlands (2015).

[188] M. Rosenblatt, Remarks on some nonparametric estimates of a density function, Ann. Math. Statist. 27 (3) (1956) 832–837.

[189] S. Sheather, M. Jones, A reliable data-based bandwidth selection method for kernel density estimation, J. R. Statist. Soc. B 53 (3) (1991) 683 – 690.

[190] Scipy - A Python library for scientific calculations, http://www.scipy.org/.

[191] E. R. Tufte, The Visual Display of Quantitative Information, second edition Edition, Graphics Press LLC, 2007.

[192] B. Keilhauer, J. Blumer, R. Engel, H. Klages, M. Risse, Impact of varying atmospheric profiles on extensive air shower observation: atmospheric density and primary mass reconstruction, Astropart. Phys. 22 (2004) 249–261.

[193] B. Keilhauer, B. Wilczyńska, H. Wilczyński, Molecular atmosphere profiles for Malargüe, Auger internal note GAP-2005-021 (2005).

[194] D. Heck, J. Knapp, J. Capdevielle, G. Schatz, T. Thouw, Corsika: A monte carlo code to simulate extensive air showers, Report FZKA 6019, Karlsruhe (1998).

[195] T. Pierog, R. Engel, D. Heck, S. S. Ostapchenko, K. Werner, Latest Results of Air Shower Simulation Programs CORSIKA and CONEX, in: Proc. 30th Int. Cosmic Ray Conf., Vol. 4, Mérida, Mexico, 2007, pp. 625–628, arXiv astro-ph/0802.1262.

[196] D. Heck, T. Pierog, Extensive air shower simulation with CORSIKA: A user's guide, Institut für Kernphysik, Forschungszentrum Karlsruhe, Karlsruhe, Germany, guide for version 6.6 (2007).

[197] D. Heck, J. Knapp, Upgrade of the monte carlo code corsika to simulate extensive air showers with energies $> 10^{20}$ ev, Report FZKA 6097B, Karlsruhe (1998).

[198] M. Kobal, A thinning method using weight limitation for air-shower simulations, Astropart. Phys. 15 (2001) 259. doi:10.1016/S0927-6505(00)00158-4.

[199] P. Billoir, Does the resampling procedure induce distortions in the fadc traces of the surface detector?, Auger internal note GAP-2005-109 (2005).

[200] P. Billoir, Reconstruction of showers with the ground array: status of the "prototype" program, Auger internal note GAP-2000-025 (2000).

[201] D. C. Hoaglin, F. Mosteller, J. W. Tukey, Understanding Robust and Exploratory Data Analysis, John Wiley and Sons, 1983.

Erklärung

Karlsruhe, den 17.12.2015

Erklärung der selbständigen Anfertigung meiner Dissertationsschrift

Hiermit versichere ich, dass ich die Dissertationsschrift mit dem Titel

Measurement of the Energy Spectrum and Mass Composition of Ultra-high Energy Cosmic Rays

selbständig und ohne unerlaubte fremde Hilfe verfasst habe. Dabei habe ich keine anderen, als die von mir angegebenen Hilfsmittel benutzt.

Alexander Schulz