A New Window to the Universe?

Searching for Ultra-High-Energy Photons at the Pierre Auger Observatory

DISSERTATION

zur Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von
Marcus Niechciol, M.Sc.

eingereicht bei der
Naturwissenschaftlich-Technischen Fakultät
der Universität Siegen

Siegen, Mai 2015

UNIVERSITÄT
SIEGEN

Gutachter der Dissertation:	Prof. Dr. Markus Risse
	Prof. Dr. Claus Grupen
Datum der Disputation:	7. Juli 2015

Gedruckt auf alterungsbeständigem, holz- und säurefreiem Papier.

Zusammenfassung

Gegenstand der vorliegenden Arbeit ist die Suche nach Photonen im EeV-Bereich in Daten des Pierre-Auger-Observatoriums. Photonen sind die wichtigsten Botschafterteilchen bei der Erforschung des Universums. Die Beobachtung von Photonen im EeV-Bereich, deren Existenz von vielen Theorien zur Herkunft von ultrahochenergetischer kosmischer Strahlung vorhergesagt wird, würde nicht nur ein neues Beobachtungsfenster zum Universum öffnen, sondern auch unser generelles Verständnis von kosmischer Strahlung vertiefen. Das Pierre-Auger-Observatorium bietet hervorragende Bedingungen für die Suche nach ultrahochenergetischen Photonen. Um Luftschauer, die von Photonen induziert wurden, vom hadronischen Untergrund zu trennen, wird eine Kombination der Observablen X_{max}, die direkt mit den Fluoreszenzdetektoren gemessen werden kann, sowie der neuartigen Observable F_γ verwendet. Die Observable F_γ basiert auf den Signalen, die in den einzelnen Detektorstationen des Oberflächendetektors gemessen werden. Um beide Observablen zu kombinieren wird eine multivariate Fisher-Diskriminanzanalyse verwendet. Die Anwendung der Methode auf Daten, die zwischen 2004 und 2012 am Pierre-Auger-Observatorium aufgenommen wurden, ergibt ein Luftschauer-Ereignis mit photonartigen Eigenschaften. Es werden Obergrenzen auf den Anteil von Photonen im Gesamtfluss der kosmischen Strahlung im EeV-Bereich bestimmt, die vergleichbar sind mit den strengsten Obergrenzen aus früheren Analysen.

Abstract

The subject of the present thesis is the search for photons in the EeV range in cosmic-ray data collected at the Pierre Auger Observatory. Photons, in general, are the main messenger particles for the exploration of the Universe. An observation of photons in the EeV range, the existence of which is predicted by many current theories about the origin of Ultra-High-Energy Cosmic Rays (UHECRs), would not only open a new observational window to the Universe, but would also further our understanding of UHECRs in general. The Pierre Auger Observatory offers an unprecedented exposure to UHE particles and thus provides a unique opportunity to search for UHE photons. In order to distinguish extensive air showers induced by photons from the hadronic background, a combination of the observable X_{max}, which is directly measured by the Fluorescence Detector (FD), and the novel observable F_γ, which is determined from the signals recorded by the Surface Detector (SD), is employed. Both observables are combined in a multivariate Fisher discriminant analysis. Applying this analysis to data collected at the Pierre Auger Observatory between 2004 and 2012, one event with photon-like characteristics is found. From this result, upper limits on the fraction of photons in the total flux of cosmic rays in the EeV range are determined, which are on par with the most stringent limits set by previous analyses.

Contents

1

Introduction

There is a theory which states that if ever anyone discovers exactly what the Universe is for and why it is here, it will instantly disappear and be replaced by something even more bizarre and inexplicable. There is another theory which states that this has already happened.

(Douglas Adams, The Restaurant at the End of the Universe)

Ever since the dawn of civilization, mankind has been infatuated with the skies, closely observing the motion of the sun, the moon and the stars across the firmament. For thousands of years, these observations have been made with nothing but the naked eye. Nevertheless, the ancient Babylonians, for example, were able to track the cyclic movements of celestial bodies like Mars and Venus with remarkable accuracy. The first major breakthrough in observational astronomy occured at the beginning of the 17th century, when the first optical telescopes had been constructed. With these instruments, it became possible to look even deeper into the skies. Many discoveries were made, which led to a deeper understanding of celestial mechanics—most importantly, perhaps, the realization that the laws of nature governing celestial mechanics are exactly the same laws that underlie terrestrial physics. Still, astronomical observations were limited by the sensitivity of the human eye to the visible light in the wavelength range between $400\,\mathrm{nm}$ and $750\,\mathrm{nm}$ in the electromagnetic spectrum (cf. Fig. 1.1). In the 20th century, however, due to the development of new measurement techniques and the rapid progress in instrumentation, new observational windows to the Universe opened up: in the radio regime at wavelengths in the meter range, the first detection of signals from an astronomical object was made in the 1930s, when radio waves coming from the Milky Way were observed. At smaller wavelengths, corresponding to higher energies than in the visible light, other windows opened up: in the X-ray regime, first measurements of the radiation coming from the Sun were performed in the 1940s using rockets to carry the instruments beyond the atmosphere of the Earth, which would otherwise absorb the X-rays. At even smaller wavelengths, in the γ-ray regime, first measurements were performed in the 1960s using satellites, since also in this regime, the Earth's atmosphere at first limited ground-based observations. These limitations were overcome later with the development of Imaging Air Cherenkov Telescopes (IACTs), which measure the Cherenkov light produced in an extensive air shower that is initiated in the atmosphere by an incoming γ-ray.
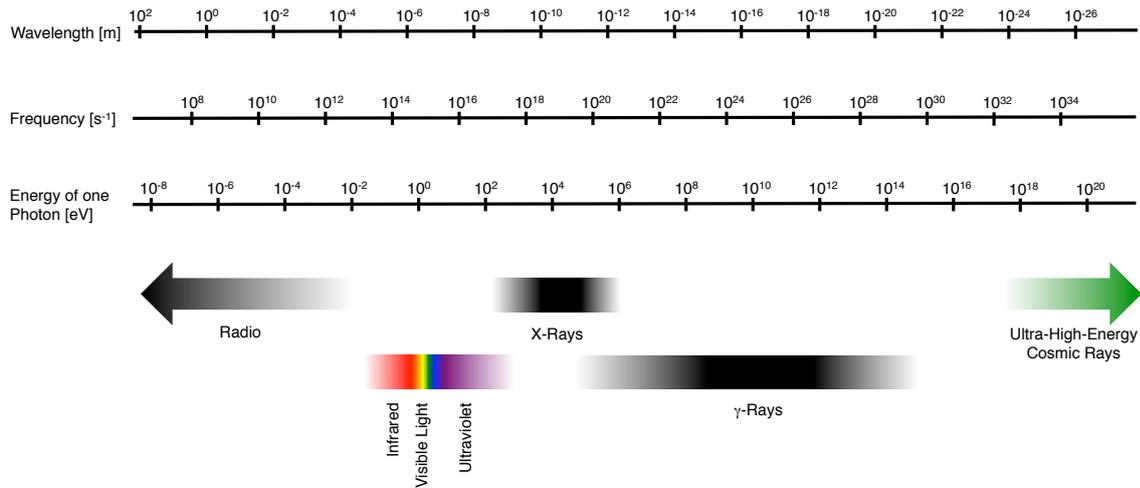
Figure 1.1.: The electromagnetic spectrum and the main wavelength regions typically used today for astronomical studies. Also indicated is the energy region for Ultra-High-Energy Cosmic Rays (UHECRs).

In Fig. 1.2, current instruments used for astronomical observations in the different wavelength regimes are shown. With these instruments, it is possible to survey the Universe with unprecedented accuracy, in an energy range that spans more than 22 orders of magnitude from less than 10^{-8} eV to more than 10^{14} eV. But photons are not the only particles with a cosmic origin reaching the Earth. In 1912, Victor Franz Hess discovered cosmic rays, which consist of charged particles like ionized atomic nuclei, electrons, and positrons. Of particular interest for the current research are Ultra-High-Energy Cosmic Rays (UHECRs) with energies above 10^{18} eV, which can only be studied using the extensive air showers these particles induce in the atmosphere of the Earth. The world's largest air shower experiment is the Pierre Auger Observatory near the town of Malargüe in Argentina. Due to its size, the Pierre Auger Observatory offers an unprecedented exposure for Ultra-High-Energy (UHE) particles. Apart from the study of UHECR, also extensive searches for UHE photons are performed at the Pierre Auger Observatory. These searches are motivated by several theories for the origin of UHECRs, which predict the existence of photons at these energies. The observation of such photons would hence not only open a completely new window to the Universe, but it would also have a large impact on the foundations of particle physics and astrophysics.

In this thesis, an analysis is presented for the search for UHE photons at the Pierre Auger Observatory. This analysis is based on hybrid data, which combine complementary measurements from the Surface Detector (SD) and the Fluorescence Detector (FD) of the Pierre Auger Observatory. The main observables used in the analysis are the atmospheric depth of the shower maximum, X_{max}, which can be measured directly with the FD, and the novel parameter F_γ, which takes into account the signals measured in the detector stations of the SD. Both parameters are combined in a multivariate analysis to extract those events from the data collected at the Pierre Auger Observatory that are most compatible with the hypothesis of a UHE photon as a primary particle. These events are then used to derive upper limits on the fraction of UHE photons in cosmic rays and upper limits on the integral flux of UHE photons impinging on the Earth.

(a)

(b)

(c)

(d)
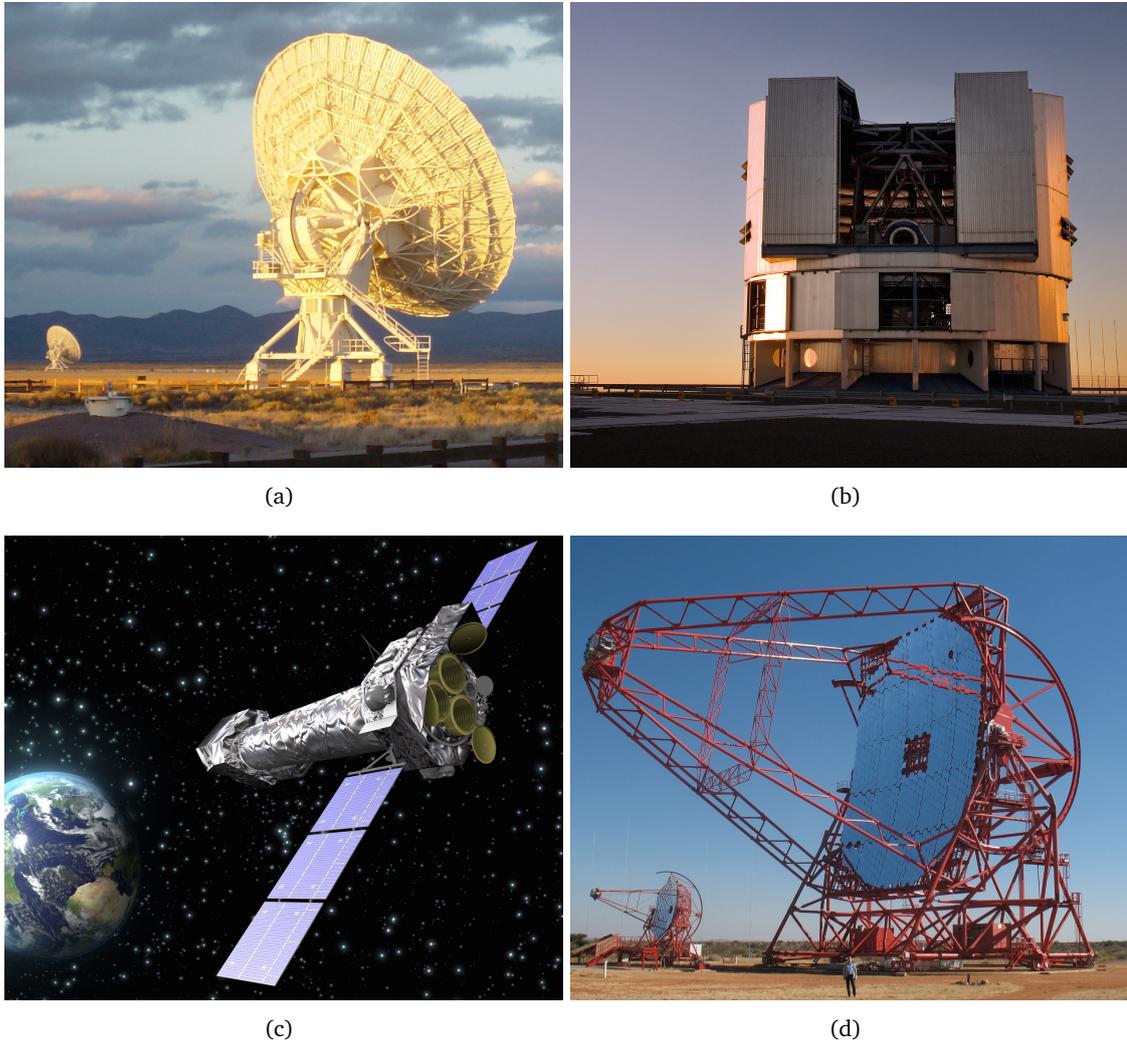
Figure 1.2.: Current instruments used for astronomical observations in the different wavelength regimes (cf. Fig. 1.1). (a) The Very Large Array (VLA) in Socorro, USA, is a radio astronomy observatory consisting of 27 independent dish antennae, each with a diameter of 25 m [Nat04]. (b) The Very Large Telescope (VLT), situated in the Atacama desert in northern Chile, consists of four individual telescopes [Eur10]. Each telescope has a primary mirror with a diameter of 8.2 m. (c) The X-Ray Multi-Mirror Mission - Newton (XMM-Newton), shown here in an artist's impression, is an X-ray observatory orbiting the Earth [Eur08]. (d) The High Energy Stereoscopic System (H.E.S.S.) in the Khomas highland of Namibia is a system of five IACTs used to observe γ-rays with energies up to 100 TeV [Hof12]. Four of the telescopes are equipped with a mirror with a diameter of about 12 m. The fifth telescope (depicted here) has a larger mirror with a diameter of 12 m.

This thesis is structured as follows: an introduction to cosmic-ray physics and a summary of the current experimental results—with focus on UHECRs—are provided in Chap. 2. UHE photons are discussed in detail in Chap. 3, following the path of a UHE photon through the Universe from its production somewhere in the Universe up to its detection on Earth. In Chap. 4, an overview of the Pierre Auger Observatory is given with particulars of the different detector systems and the reconstruction of air shower events from the collected data. The simulations on which the analysis is based as well as the data sample collected at the Pierre Auger Observatory to which the analysis is applied are described in Chap. 5. In Chap. 6, a detailed study of the different observables which are used to differentiate photon-induced air showers from those initiated by protons or heavier nuclei is presented. The two main observables X_{max} and F_γ are first discussed independently and then combined in a multivariate Fisher discriminant analysis. The performance of the analysis is then evaluated in detail. The results of the analysis when it is applied to data are discussed in Chap. 7. The candidate event that best fits the hypothesis of a UHE photon is examined in detail and upper limits on the fraction of UHE photons in cosmic rays and upper limits on the integral flux of UHE photons are derived. Finally, a summary of the main results described in this thesis and an outlook for future analyses are given in Chap. 8.

<div style="text-align: right;">

2

</div>

Cosmic Rays

Contents

Each second, the Earth's atmosphere is hit by countless particles, from neutrinos produced in nuclear fusion processes in the center of our Sun to iron nuclei possibly originating in Active Galactic Nuclei (AGNs) far away from our own galaxy, the Milky Way. Even though all of these particles have a "cosmic" origin, the term "cosmic rays" usually refers to only a fraction of this multitude of particles, namely charged particles like ionized nuclei and electrons or positrons. Photons are chargeless particles and thus not referred to by the term "cosmic rays". Nevertheless, there is a strong connection between photons and charged particles in cosmic rays, especially at the highest energies above the EeV range. Ultra-High-Energy (UHE) photons are discussed in Chap. 3.

In the first section of this chapter (Sec. 2.1), a review of the early history of cosmic-ray physics is presented, covering the period before the actual discovery of cosmic rays in 1912 as well as the early studies of cosmic rays until the 1940s. For a more detailed account of the many small steps leading to the discovery of cosmic rays and the different, sometimes parallel, developments in the early phases of cosmic-ray physics, see e.g. [Fal12] or [Sek85]. Following the historical introduction, the current state of knowledge in cosmic-ray physics is summarized (Sec. 2.2), focussing on Ultra-High-Energy Cosmic Rays (UHECRs) in the energy range above 10^{18} eV. Possible sources and acceleration mechanisms for these UHE particles are discussed, in addition to the energy spectrum observed on Earth, the elemental composition and the anisotropy in the arrival directions. This section concentrates on primary cosmic rays,

i.e. the actual cosmic particles impinging on the Earth. From these primary particles, one has to distinguish those particles that are only created as secondary particles within the Earth's atmosphere. The extensive air showers that are initiated in the atmosphere when a primary cosmic ray interacts with a nucleus from the atmosphere, producing a cascade of secondary particles, are addressed in the last section of this chapter (Sec. 2.3). Special emphasis is given on the development of an extensive air shower in the atmosphere.

2.1. Early History of Cosmic-Ray Physics

On August 7, 1912, the Austrian physicist Victor Franz Hess embarked on the seventh and last of a series of balloon flights to undertake measurements related to the conductivity of air—measurements that would lead to the discovery of cosmic rays and thus mark the dawn of cosmic-ray physics, or in a more general sense even the beginning of the whole field of astroparticle physics.

The problem of the conductivity of air, i.e. the observation that an electroscope—in its simplest form a charged conductor inside an electrically isolated container—will slowly lose its charge over time, was known since the late 18th century, when it was first described by Charles-Augustin de Coulomb [Cou85]. Coulomb attributed this charge loss to dust particles in the surrounding air, but this explanation was not deemed satisfactory, since the charge loss also appeared when the electroscope was well insulated against leaking currents to the outside. This phenomenon puzzled physicists for more than a century, until a better explanation was provided through the discovery of radioactive elements and their accompanying ionizing radiation in 1896 by Henri Becquerel [Bec96]. Julius Elster and Hans Geitel, and independently Charles Thomson Rees Wilson, were the first to relate this ionizing radiation to the problem of the conductivity of air [Els01, Wil01]: small traces of radioactive materials in the surroundings or in the material of the electroscope itself could emit ionizing radiation which produces ions in the air inside the electroscope leading eventually to a slow discharge of the apparatus. In the following years, systematic studies of the rate of ionization were done using electrometers, i.e. gauged electroscopes that can be used to measure the charge loss quantitatively. Soon the consensus was that the conductivity of air was a consequence of radioactive matter in the soil and their accompanying γ-radiation. In 1907, Theodor Wulf presented a new type of electrometer (see Fig. 2.1), which used two tiny strings of quartz instead of a thin gold foil [Wul07]. With this improvement, much more accurate measurements were possible and the Wulf two-string electrometer quickly became the standard apparatus for measuring the ionization rate [Fri11].

Using his electrometer, Wulf performed systematic studies investigating the origin of the γ-radiation in the atmosphere—taking into account also the background radiation from, e.g., the walls of the building—and concluded that the remnant radiation is originating from radioactive elements in the uppermost layer of the Earth [Wul09]. A possible contribution from the atmosphere itself was deemed so small that it is impossible to detect. These conclusions were further supported by measurements done on top of the Eiffel tower in Paris. Here, Wulf found a slight decrease in the radiation, as expected from a source in the ground [Wul10]. However, the decrease was much smaller than expected. At the same time, Domenico Pacini performed measurements of the ionization rate on the surface of the Mediterranean Sea, where radioactive elements in the soil should have no influence. The ionization rate he mea-
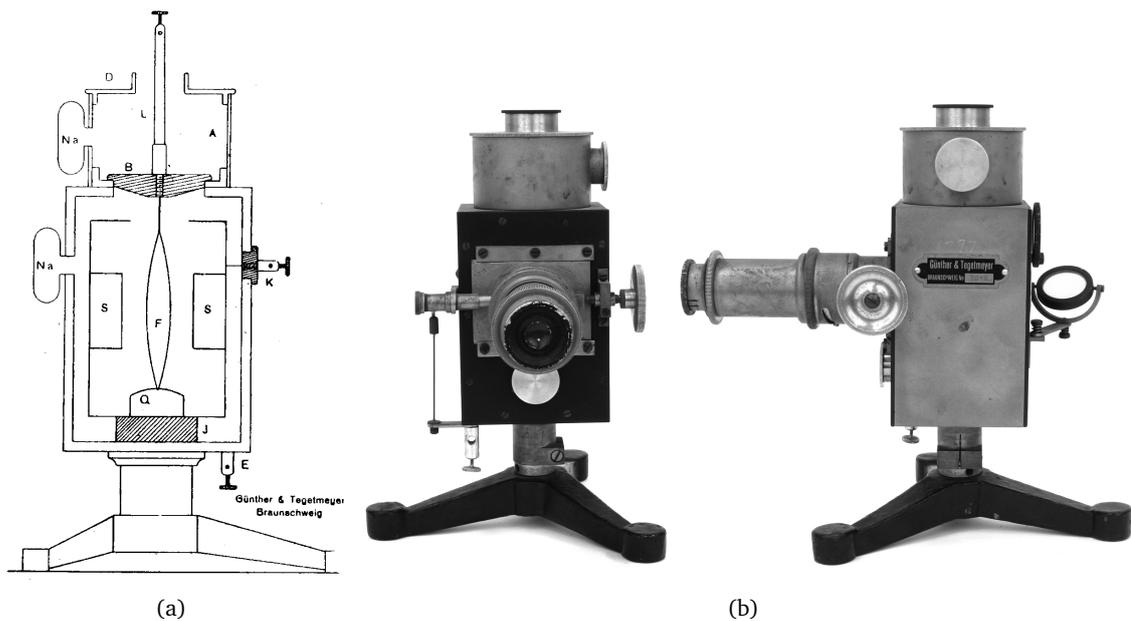
Figure 2.1.: The Wulf two-string electrometer [Fri11]. (a) Schematic depiction. (b) Photograph of the apparatus. This type of electrometer uses two quartz strings, denoted by F in (a), which were glued together at their ends. If a charge is applied, the two strings move in opposite directions due to the repulsive electrostatic force. The measurement of the charge is possible by observing the strings through a microscope. For a more detailed description of the apparatus, see [Fri11].

sured was higher than expected, and so he concluded—in contrast to Wulf—that there should be a significant source of radiation in the atmosphere itself [Pac11].

The first to measure the ionization rate above ground level, i.e. higher in the atmosphere, using a balloon was Franz Linke, already in 1904 [Lin04]. In 1907, Wulf and Albert Gockel started a series of measurements in the Swiss Alps, where they compared the ionization rate at different altitudes on the mountains [Goc08]. They concluded that the ionization rate is not dependent on the altitude and thus, a cosmic radiation only contributes an insignificant fraction to the ionization rate—using the term "cosmic radiation" for the first time. Later, Gockel used balloons to reach higher altitudes and measure the ionization rate up to 4500 m altitude [Goc10]. In these balloon flights, Gockel found a decrease in the ionization rate, but his results were affected by problems with his instruments, which were not fully suited for the temperature and pressure conditions at these heights.

So far, the measurements of Wulf, Gockel, Pacini and others, gave only indications toward an additional cosmic radiation penetrating the Earth's atmosphere. The final discovery of cosmic rays followed in 1912 by Victor Franz Hess (see Fig. 2.2). At Wulf's suggestion, he devised a method to calibrate his two-string electrometers using the γ-radiation from radium sources. Hess was able to increase the sensitivity of the electrometer to a few per mille, instead of a few percent with uncalibrated electrometers [Hes13]. Interpreting the previous results by Wulf and Gockel that the measured rate of ionization does not decrease as significantly as expected, Hess concluded that the reason could be either a new, unknown, type of radiation,

or the absorption of γ-rays in air could be weaker than assumed. In his first experiments, Hess investigated the latter theory using radium sources of known strengths [Hes11]. He found that the absorption was indeed of the expected magnitude, calculated by Arthur Stewart Eve already in 1906 [Eve06]. In 1912, the Imperial Academy of Sciences funded a series of seven balloon flights [Hes12]. The first six flights took place between April and June 1912, all starting from Vienna. To overcome the problems of Gockel and others with their instruments during the flights, Hess used pressure-tight electrometers to measure the γ-radiation. Hess followed a carefully planned measurement program: preceding each flight was a long period of control measurements on the ground to exclude possible radioactive contaminations in the balloon or the ballast sand. The first six flights were all done at lower altitudes up to 2100 m. With the measurements taken during these flights, Hess could confirm the previous results by Gockel and Wulf with high confidence. Some of the flights were also done at night and during a solar eclipse. Hess could find no differences compared to the measurements done during the day. Thus, he concluded that if there is a new kind of radiation from outside the atmosphere, it is not connected to the Sun. The last step was to obtain the final proof for this new radiation during the seventh and last flight, which was taken on August 7, starting in Aussig an der Elbe (today Ústí nad Labem, Czech Republic) and landing in Pieskow, 50 km south-east of Berlin. In contrast to the other flights, this flight was planned to reach higher altitudes. Therefore, Hess used a large hydrogen balloon instead of the smaller coal gas balloons used in Vienna. Reaching a maximum altitude of 5350 m, Hess measured in both of his electrometers an increase of the ionization rate up to a factor of four as compared to the rate on ground. This was the confirmation of the existence of cosmic radiation. However,

Figure 2.2.: Victor Franz Hess in the gondola of a balloon in 1912, surrounded by spectators [Sek85].

Hess' measurements, at first, were not widely accepted. Only after Werner Kolhörster reached even higher altitudes up to 9300 m during his balloon flights and found the same significant increase as Hess, the existence of cosmic rays was generally established [Kol14]. International recognition was eventually given to Hess in 1936 through the Nobel Prize in Physics, which was awarded to him "for the discovery of cosmic radiation" [Nob65].

The next period in cosmic-ray research was devoted to the study of the basic properties of this newly discovered radiation. One important question was the nature of this radiation. Because of their penetrating power, it has always been generally assumed that this radiation is at least similar to the γ-radiation known from radioactive decays. With the invention of new measurement techniques like the Geiger-Müller counter [Gei28], more detailed investigations became possible. Walter Bothe and Kolhörster used two Geiger-Müller counters surrounded by thick layers of absorbant material to measure coincident signals. Using this setup, they showed that cosmic rays must be charged particles, since γ-radiation would not produce coincident signals [Bot29]. Another proof for the particle nature was obtained in the 1930s, when Arthur Holly Compton started a worldwide measurement campaign and found that the intensity of cosmic rays is dependent on the geomagnetic latitude, which again would not be the case for γ-radiation [Com33].

Another important milestone was the discovery of extensive air showers. In 1935, Erich Regener and Georg Pfotzer studied the intensity of cosmic rays in the stratosphere using threefold coincidences [Reg35]. Regener and Pfotzer found, after the well-known increase in intensity with altitude, an unexpected maximum at an altitude of about 14 km. After Homi Jehangir Bhabha and Walter Heitler presented their theory of electromagnetic showers [Bha37], which was based on the ideas of Quantum Electrodynamics (QED), this maximum was interpreted as a result of the multiplication of electrons in the atmosphere. The discovery of extensive air showers is, however, credited to Pierre Auger. In 1939, Auger made coincidence measurements at an altitude of 2350 m at the Jungfraujoch in the Swiss Alps [Aug39]. He measured the coincidence rate as a function of the distance between the single counters up to a distance of 75 m and found that the measured rate significantly exceeds the expected rate of chance coincidences. Thus, he confirmed the existence of extensive air showers of secondary particles induced by primary cosmic-ray particles in the atmosphere.

In retrospect, cosmic-ray physics can also be seen as the birthplace of modern particle physics. With the constant improvements in detection techniques—such as the Wilson cloud chamber, which had been developed in 1911 [Wil11] and afterward successfully used to visualize particle tracks coming from radioactive elements—many new particles were discovered, such as the positron in 1933 [And33], the muon in 1937 [Ned37], and the "strange" kaons in 1947 [Roc47]. All of these discoveries were possible because cosmic rays were the most energetic particles known at that time. Only in the 1950s, man-made particle accelerators took over, and the subsequent discoveries of new particles were made at accelerator-based experiments. Cosmic-ray physics focused on different questions, which are, to a large extent, today still subject of research: the origin of cosmic rays, i.e. their sources, acceleration and propagation mechanisms, their energy spectrum and their composition—albeit today, research is focused on energies several orders of magnitude above what physicists dealt with in the early phases of cosmic-ray physics.

2.2. Primary Cosmic Rays

In the following sections, the current state of knowledge about cosmic rays, in particular UHECRs with energies above 10^{18} eV, is summarized. Although the focus lies on UHECRs, cosmic rays at lower energies are also discussed since at these energies, their properties and their origin are much better understood: partly because their incoming flux is much larger, but also because they can be measured directly, and thus, measurements are not subject to interpretations based on, e.g., extrapolated hadronic interaction models. In addition to the experimental results, astrophysical interpretations of the measurements are discussed.

The experimental results discussed in the following come, to a large extent, from the Pierre Auger Observatory, which is the largest cosmic-ray observatory to date, specifically designed to investigate UHECRs. The Pierre Auger Observatory is discussed in detail in Chap. 4.

2.2.1. Energy Spectrum

The total integrated rate of cosmic-ray particles with energies above the GeV range impinging on the Earth is in the order of $1000 \, \mathrm{m}^{-2} \, \mathrm{sr}^{-1} \, \mathrm{s}^{-1}$. The differential all-particle spectrum is shown in Fig. 2.3. The energy spectrum covers about 11 orders of magnitude in energy, spanning from the GeV range to EeV energies and above. For each decade in energy, the flux drops by about three orders of magnitude, thus spanning more than 30 orders of magnitude—from 1 particle per m^2 per second at 100 GeV down to 1 particle per km^2 per century at 100 EeV. Current man-made particle accelerators, such as the Large Hadron Collider (LHC) in Geneva, Switzerland, can reach energies of 7 TeV per particle [Eva08], which is about seven orders of magnitude below the highest energies measured in cosmic-ray physics. In terms of center-of-mass energy, the LHC can reach 14 TeV in the particle collisions, which corresponds to a cosmic-ray particle with an energy of 100 PeV interacting with a particle at rest in the Earth's atmosphere. Therefore, cosmic rays can reach energy regions that are currently inaccessible in laboratory experiments, in terms of both particle energy and center-of-mass energy.

Even though the energy range covered by the energy spectrum of cosmic rays is quite large, the differential flux above some 10 GeV, where effects of the geomagnetic field and the Sun become negligible, can be approximated by a simple broken power-law:

$$\frac{\mathrm{d}\phi}{\mathrm{d}E} \propto E^{-\gamma}. \tag{2.1}$$

The spectral index γ describes the steepness of the spectrum. In Fig. 2.3, small variations of this steepness are apparent. The visibility of these variations can be enhanced by multiplying the differential flux by a factor $E^{2.6}$, as shown in Fig. 2.4. In particular, three distinctive features become visible: the knee at an energy of about 4×10^{15} eV, where the spectrum gets slightly steeper; the second knee at an energy of about 4×10^{17} eV, where the spectrum once again steepens; and the ankle at an energy of about 4×10^{18} eV, where the spectrum becomes flatter again. In addition, a cut-off above 6×10^{19} eV is visible in Fig. 2.4. In the following, the afore-mentioned features of the energy spectrum will be discussed.

Lower End of the Energy Spectrum

Below some 10 GeV, the flux of primary cosmic rays is strongly influenced by the magnetic field of the Earth and the Sun's activity [Per09], leading to a dependence of the measured flux
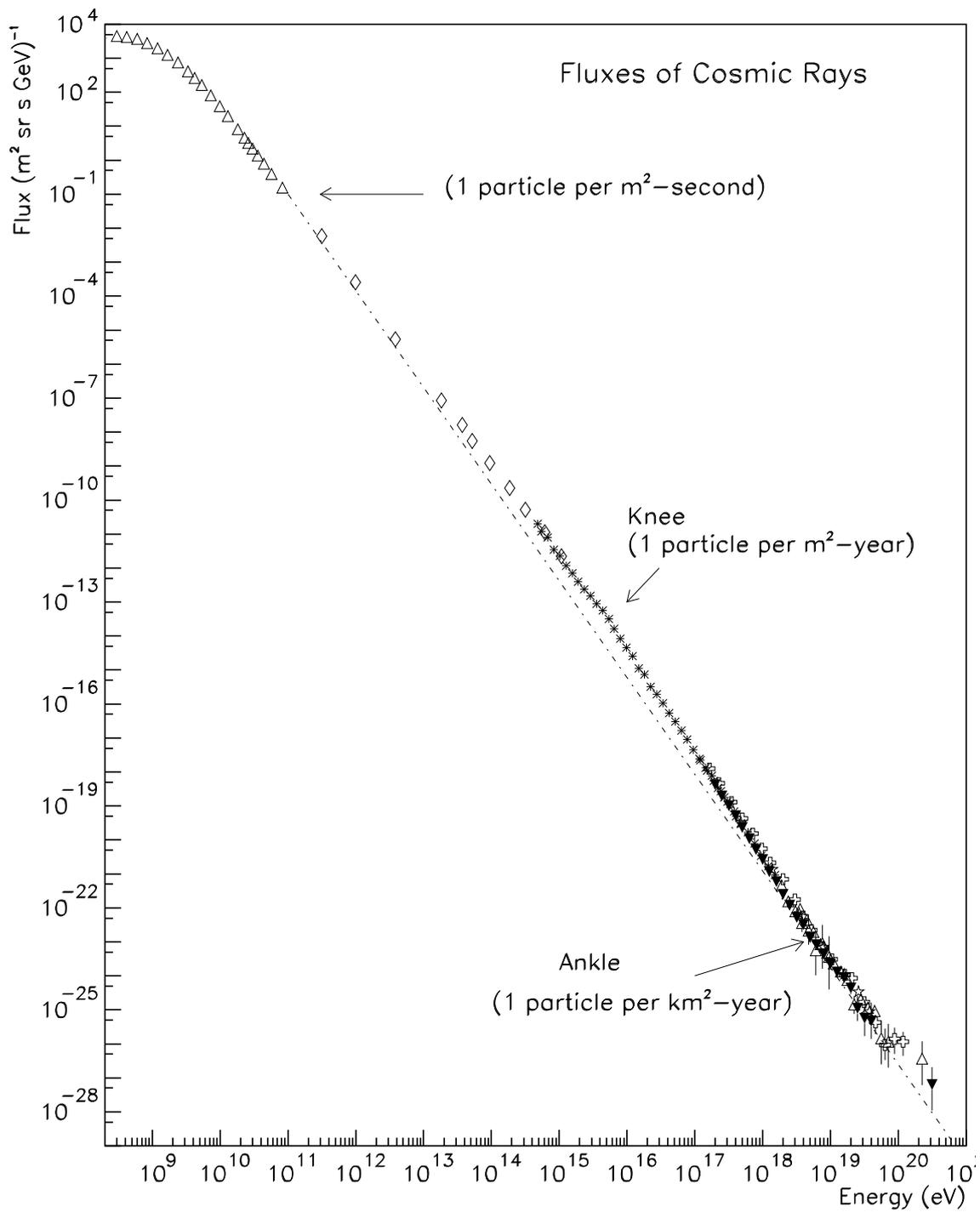
Figure 2.3.: The all-particle energy spectrum of primary cosmic rays as measured by different experiments [Swo01]. The spectrum covers about 11 orders of magnitude in energy and about 32 orders of magnitude in flux.
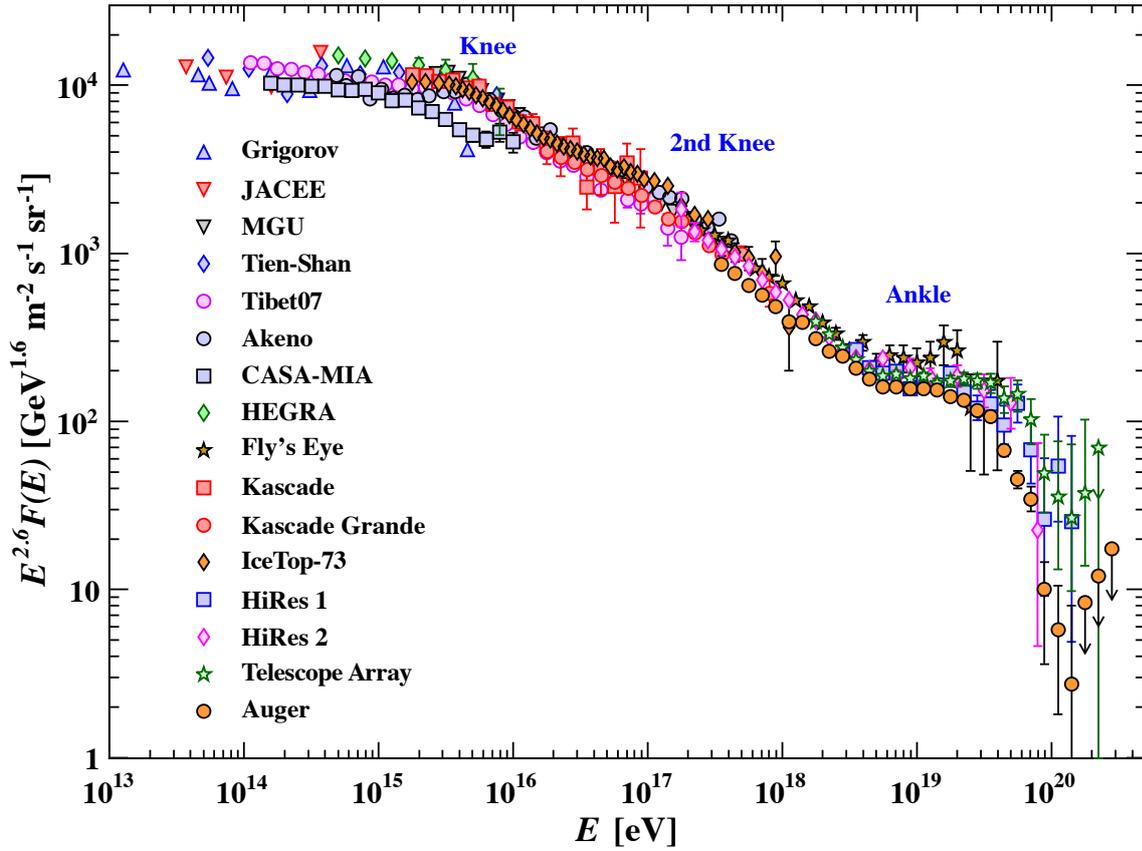
Figure 2.4.: Another representation of the all-particle energy spectrum of primary cosmic rays [Ber12]. Here, the differential flux shown in Fig. 2.3 is multiplied by a factor $E^{2.6}$ to enhance the visibility of the features of the spectrum. The results compiled in this plot come from a variety of experiments, both direct and indirect. In the UHE region, results from the High Resolution Fly's Eye Detector (HiRes) [Bir94, Abb08a], the Telescope Array (TA) [Abu13a] and the Pierre Auger Observatory [Abr08a, Sch13] are included. For a complete list of references, see [Ber12].

in this energy region on both the location of the measurement and the time. The geomagnetic field, which is in first-order approximation a dipole field, leads to a cut-off in the energy spectrum at lower energies, because a charged particle, which is deflected in the dipole field, needs a minimum momentum to penetrate the magnetic field. For a particle with charge $Z \times e$, the cut-off momentum p_{min} is given by

$$p_{min} = 59.6 \, \text{GeV/c} \times Z \times \frac{\cos^4 \lambda}{\left(1 + \sqrt{1 - \sin \zeta \cos^3 \lambda}\right)^2}, \qquad (2.2)$$

where λ denotes the magnetic latitude and ζ the angle of the incoming direction of the particle with respect to the dipole [Per09]. For example, for a proton ($Z = 1$) incident from the vertical ($\sin \zeta = 0$) at the location of the Pierre Auger Observatory in Malargüe, Argentina, ($\lambda = 25.32°$ S) the cut-off momentum according to Eq. 2.2 would be $p_{min} = 9.95 \, \text{GeV/c}$, which corresponds to a kinetic energy of about 9 GeV. However, Eq. 2.2 can only give a general idea of the geomagnetic cut-off. A more precise calculation would have to take the displacement of

the Earth's magnetic dipole from the centre of the Earth into account, as well as higher-order multipole moments in the geomagnetic field [Per09].

In addition to the geomagnetic field, the flux of cosmic rays is also influenced by the Sun's activity through the solar wind, which is a plasma of low-energy particles constantly emitted by the Sun. The intensity of the solar wind is connected to the sunspot activity, which means it follows the same eleven-year cycle. In the magnetic fields associated with the solar wind, the trajectories of cosmic rays are strongly distorted, leading to a modulation of the measured flux on Earth connected to the sunspot activity [Per09].

The Knee

At an energy of about 4×10^{15} eV, the spectral index γ changes from 2.7 to 3.1 [Blü09]. This steepening of the energy spectrum is commonly referred to as the knee. The origin of this feature is not yet clear. Currently, many theories exist to explain the knee. Some of these theories are discussed briefly here. For a more detailed review of the theories aiming at explaining the knee, see [Hör04].

Some theories relates the knee to the acceleration processes of cosmic rays at their sources. The most common models are based on diffusive shock acceleration in Supernova Remnants (SNRs) (see also Secs. 2.2.4). In this scenario, the maximum attainable energy is governed by the diffusion of the accelerated particle out of the acceleration region due to the particle's gyroradius, which is itself depending on the particle's momentum and the reciprocal of the particle's charge Z. The reciprocal dependence on Z leads to a depletion of lighter elements, i.e. elements with smaller Z, at higher energies. The maximum attainable energy in this acceleration process is $\propto Z \times 10^{15}$ eV [Ber96], leading to a charge-dependent drop of the measured flux above this energy. The knee in the all-particle spectrum could then be associated to the cut-off energy of protons ($Z = 1$).

Also a source-related theory is a specific scenario in which a featureless background spectrum is modified by a single, local source [Erl97]. This source could be, for example, the relatively young Monogem Ring SNR [Erl04]. In this scenario, the knee would be formed mainly by elements of the Carbon, Nitrogen, Oxygen (CNO) group, accelerated at the local source. The featureless background could be explained by mechanisms involving pulsars and a modified diffusion model in the Galaxy [Erl97].

Another class of theories relates the knee to the propagation of cosmic rays through the galactic magnetic fields, usually described as a diffusion process. Many different theories exist based on different assumptions on the structure of these fields. In most scenarios a regular magnetic field is superimposed by a random, turbulent component. In one basic model, a global regular field is assumed with the turbulent component existing only in the galactic halo [Ptu93]. In this scenario, the drift of particles in the regular field is responsible for the knee. The drift and diffusion of the particles depend on the energy of the particles and the reciprocal of the charge Z. This leads to a charge-dependent drop of the flux above a threshold energy—and thus a depletion of lighter elements above the knee—as in the source-related scenarios discussed previously.

The models and scenarios discussed so far are all based on astrophysical processes, where the knee is an intrinsic feature of the energy spectrum of primary cosmic rays. In other models, new particle physics processes are considered, which lead to a modification of a featureless primordial spectrum to exhibit the knee feature in measurements. Such models postulate new types of interactions in the development of an extensive air shower in the atmosphere, which transfer energy to new particles that cannot be observed—or have not been observed yet—by current air shower experiments [Hör04]. The energy threshold for these new types of interactions is assumed to be at the knee. Thus, the true energy of a cosmic-ray particle above this threshold energy would be underestimated by air shower measurements, leading to the knee structure in the measured energy spectrum. The new particles needed in these models could be provided by, e.g., extensions of the current Standard Model of Particle Physics, in particular Supersymmetry or Technicolor [Kaz01]. In any case, the new particles should also be observed at the LHC. However, the searches for new physics phenomena at the different experiments at the LHC have yielded neither direct nor indirect evidence of such phenomena up to now, see e.g. [Aad14].

So far, the situation remains unclear on the origin of the knee. Recent results from the Karlsruhe Shower Core and Array Detector-Grande (KASCADE-Grande) experiment on the energy spectra of different elemental groups from 10 PeV to beyond 200 PeV yield a strong indication for a dominance of heavier elements in this energy range [Ape13]. In addition, a knee-like structure has been found in the iron spectrum at around 80 PeV. Together, these results support the models which predict a charge-dependent drop of the flux, with the knee in the all-particle spectrum being caused by the lightest elements. However, some uncertainties remain, especially with the hadronic interaction models used in the data reconstruction [Ape13].

The Second Knee

Two orders of magnitude above the knee, at an energy of about 4×10^{17} eV, a further steepening of the energy spectrum occurs [Blü09]. Here, at the second knee, the spectral index γ changes from 3.1 to about 3.3. It is generally assumed that the second knee is closely connected to the knee. In the phenomenological poly-gonato model—or "many knees model"—the total flux is a result of the superposition of the fluxes of the single elemental components of the charged cosmic rays, with an energy cut-off for each component proportional to the charge Z [Hör03]. In this context, the second knee would be connected to the point where the heaviest elements up to $Z = 92$ no longer contribute significantly to the total flux, while the knee would correspond to the cut-off for the lightest elements, starting with protons with $Z = 1$. Overall, the poly-gonato model describes the measured energy spectra well. However, it relies on several assumptions on the elemental composition of cosmic rays in this energy regime, for example a non-negligible contribution of heavy elements with $Z \geq 28$ [Hör03]. Since the elemental composition at these energies is not yet well known from measurements, it remains difficult to prove these assumptions.

Other theories regarding the origin of the second knee postulate exotic interactions of the cosmic rays with other particles, such as relic electron neutrinos from the Big Bang [Wig00]. The average density of these relic neutrinos in the Universe is estimated to be in the order of $100 \, \text{cm}^{-3}$ with an average kinetic energy in the meV regime, although the density around, e.g., pulsars and AGNs, which are both possible source candidates for cosmic rays, can be several orders of magnitude larger[Wig00]. If the rest mass of the electron neutrino is around

$0.4\,\text{eV}/c^2$—a value which is presently not yet experimentally excluded [Ber12]—primordial α particles in cosmic rays may interact with the relic neutrinos and antineutrinos via

$$\alpha + \nu_e \rightarrow 3p + n + e^-,$$
$$\alpha + \overline{\nu}_e \rightarrow p + 3n + e^+. \tag{2.3}$$

The energy threshold for these processes, i.e. the energies at which a kink in the energy spectrum of cosmic rays should appear, is calculated to be around $3 \times 10^{17}\,\text{eV}$, which is in agreement with the measured value [Wig00].

Similarly, this theory also provides an explanation for the knee, which is interpreted as the energy threshold for protons to interact with relic antineutrinos via inverse β decay [Wig00]:

$$p + \overline{\nu}_e \rightarrow n + e^+. \tag{2.4}$$

However, according to this theory, the knee should be a unique feature of the proton component of cosmic rays, while the second knee should be connected to α particles only, which is in contradiction to the measurements by the KASCADE-Grande experiment discussed before [Ape13].

The Ankle

At an energy of $4 \times 10^{18}\,\text{eV}$, the spectrum flattens again [Blü09], with the spectral index γ changing from 3.3 back to about 2.7. This feature is commonly known as the ankle. As for the knee and the second knee, several theories exist on the origin of the ankle. It is generally assumed that the sources of cosmic rays with energies below the knee and the second knee are located within our galaxy, while cosmic rays at the highest energies have an extragalactic origin. However, it is unclear where exactly the transition from a dominant galactic component of cosmic rays to a dominant extragalactic component occurs. A review of several theories concerning this transition and the experimental constraints on these theories is given in [Alo12].

Traditionally, the ankle is taken directly as a signature of this transition, i.e. the intersection of a flat extra-galactic component, assumed to be purely composed of protons, with a steep galactic component, see e.g. [Hil05]. However, in this model, the ankle model, the spectrum of galactic cosmic rays would have to extend above EeV energies, which is difficult to explain in the scope of most models for the knee discussed previously. Hence, an additional galactic component is assumed dominated by heavy nuclei [Hil05]. This assumption is in contradiction to measurements of the composition in the energy range between 1 and 5 EeV, which indicate a lighter composition [Alo12].

In a different model, the dip model, the transition from galactic to extragalactic cosmic rays is assumed to occur at lower energies, at the second knee, thus being compatible with the estimates of the maximum attainable energy of galactic sources for cosmic rays. In this model, the ankle is caused by energy losses suffered by the cosmic-ray particles during the propagation from their sources to the Earth [Ber06a]. In addition to adiabatic energy losses due to the expansion of the Universe—mainly relevant for distant sources—which affect cosmic rays of all energies and thus do not change the shape of the energy spectrum, an additional source of

energy loss comes from the interactions of cosmic-ray protons with photons from the Cosmic Microwave Background (CMB), creating electron positron pairs:

$$p + \gamma_{\mathrm{CMB}} \rightarrow p + e^- + e^+. \tag{2.5}$$

According to the dip model, this process causes the flattening of the spectrum at the ankle in the form of a dip in the energy range between 1 and 40 EeV [Ber06a]. In this model, a proton-dominated extragalactic component is assumed, taking over from an iron-dominated galactic component at the second knee. This assumption is more compatible with the current measurements than the ankle model. However, at higher energies, the situation is far more unclear. While the measurements by the Telescope Array (TA) show a light composition compatible with this model [Jui12], measurements by the Pierre Auger Observatory indicate a change to a heavier composition [Aab14a], which contradicts the proton-dominated composition required by this model.

The two models discussed so far assume an extragalactic component that is dominated by protons with at most only a small admixture of light nuclei. The mixed composition model is another model that is based on the argument that any acceleration mechanism operating in a gas—e.g. acceleration in AGNs—involves different nuclei and thus the primary flux must have a mixed composition [All05]. In a basic variant of this model, the composition of the extragalactic component is assumed to be close to the composition of the galactic component, i.e. dominated by protons and helium nuclei. At energies above 10 EeV, the composition is predicted to become lighter because heavier nuclei interact in photo-disintegration processes with photons from the Extragalactic Background Light (EBL) [Alo12]. This may change only above 50 EeV, where the proton flux is suppressed due to the predicted Greisen, Zatsepin, Kuz'min (GZK) cut-off, which is discussed in the next section.

Flux Suppression at the Highest Energies

Eventually, above an energy of about 6×10^{19} eV, the flux of primary cosmic rays is strongly suppressed. For some time, it was unclear whether this flux suppression is indeed present in the energy spectrum. On the one hand, measurements from the Fly's Eye experiment indicated a suppression of the flux [Bir94]. On the other hand, the Akeno Giant Air Shower Array (AGASA) recorded a significant number of events with reconstructed energies above 10^{20} eV [Tak98], which seems to contradict the Fly's Eye measurements. More recently, and with higher statistics, the flux suppression has been confirmed by HiRes [Abb08a] and the Pierre Auger Observatory [Abr08a]. The hypothesis that the energy spectrum continues with a constant slope was rejected by the Pierre Auger Observatory with a significance of more than 20σ [Abr10b], hence, the existence of the flux suppression at the highest energies is firmly established today. However, the cause of this suppression is not fully known yet.

The traditional explanation for the flux suppression at the highest energies dates back to the 1960s. In 1965, Arno Penzias and Robert Wilson discovered the Cosmic Microwave Background (CMB) [Pen65], which was immediately interpreted as a relic signature from the Big Bang, filling out the entire Universe [Dic65]. Shortly after the discovery of the CMB, Kenneth Greisen and independently Georgiy Zatsepin and Vadim Kuz'min predicted that the CMB should have an effect on the energy spectrum in the form of a suppression of the flux of

UHECRs due to interactions of cosmic-ray protons with photons from the CMB [Gre66, Zat66]. In these interactions, pions are produced via the $\Delta(1232)^+$ resonance in processes like

$$p + \gamma_{\mathrm{CMB}} \rightarrow \Delta(1232)^+ \rightarrow p + \pi^0,$$
$$p + \gamma_{\mathrm{CMB}} \rightarrow \Delta(1232)^+ \rightarrow n + \pi^+. \tag{2.6}$$

These processes become energetically possible above a threshold energy of about 5×10^{19} eV, leading to a significant energy-loss for UHE protons propagating from their extragalactic sources to the Earth. Hence, if UHE protons with an energy above 10^{20} eV are detected, their sources must lie within the GZK horizon of about 100 Mpc from the Earth [Kac08].

For nuclei, the energy threshold for the GZK process in Eq. 2.6 is higher, but for these particles, also photo-disintegration processes with photons from the CMB as well as from the EBL have to be taken into account:

$$A + \gamma_{\mathrm{CMB,\ EBL}} \rightarrow (A - nN) + nN, \tag{2.7}$$

with A denoting a nucleus with mass number A and N denoting a nucleon. The dominant process is one nucleon emission ($n = 1$) [Alo13a, Alo13b]. These processes also lead to a suppression of the flux of nuclei at the highest energies, although the exact onset of the flux suppression depends on the details of the EBL, which—in contrast to the well-measured CMB—are not known.

Another model relates the flux suppression to the maximum energy to which the sources of UHECRs can accelerate the particles [Alo11]. This model is motivated by measurements from the Pierre Auger Observatory, which indicate a change from a light composition at around 1 EeV to a heavier composition at the highest energies (see also Section 2.2.2) and thus contradict the assumption of a proton-dominated extragalactic component required by the GZK models discussed before. In this model, the maximum energy a cosmic-ray proton can gain is around $4 - 10$ EeV, while the highest energies in the range of $100 - 200$ EeV can be reached only with iron nuclei [Alo11]. A charge-dependent cut-off energy at the source leads to the observed flux suppression without the GZK processes. This also leads to a much smaller flux of UHE neutrinos and photons, which result from the decays of the charged and neutral pions created in the GZK processes. Thus, measuring the flux of UHE photons in conjunction with more detailed measurements of the spectrum and the elemental composition can provide a clue on whether the flux suppression at the highest energies is due to the GZK effect, a maximum attainable energy at the sources or a combination of both.

2.2.2. Composition

Determining the composition of primary cosmic rays is essential for understanding the nature and origin of cosmic rays. However, the need for determining the abundance of single elements and elementary particles in primary cosmic rays extends beyond cosmic ray physics itself. For example, measurements from the Alpha Magnetic Spectrometer (AMS-02) showed an excess of positrons in primary cosmic rays at energies between 0.5 and 350 GeV. This indicates new physics phenomena that may be explained by, e.g., dark matter interactions [Agu13]. Another example, more related to astrophysics, is the measurement of the ^{10}Be/^{9}Be ratio. From measurements of the Isotope Magnet Experiment (ISOMAX), the size of the galactic halo of

the Milky Way can be estimated to be a few kpc [Mol03]. The following sections focus on the elemental composition of cosmic rays, which is discussed at lower energies up to 100 TeV, where direct measurements are still possible, and at the highest energies above the EeV range.

Elemental Composition at Energies up to 100 TeV

Up to about 100 TeV, the composition of primary cosmic rays is well known from balloon-borne and satellite-borne experiments. About 98 % of the primary cosmic rays impinging on the Earth are protons and nuclei, while the remaining 2 % are electrons and positrons. Of the protons and nuclei, 87 % are protons, 12 % are helium nuclei and the remaining 1 % are heavier nuclei [Lon92]. With different abundances, all elements from the periodic table have been found in primary cosmic rays. A compilation of experimental results on the abundance of the different elements in primary cosmic rays at energies of around 1 GeV per nucleon is shown in Fig. 2.5. Overall, the distribution of the relative abundance of elements in cosmic rays largely resembles the one in the Solar System, which indicates that the dominant production process for the elements in cosmic rays is stellar nucleosynthesis [Kla00]. However, for certain elements or elemental groups, the abundance differs, in some cases by several orders of magnitude.

The lightest elements hydrogen ($Z = 1$) and helium ($Z = 2$) are less abundant in cosmic rays than in the Solar System. This can be understood as a consequence of the relatively high ionization energies of these elements [Kla00]. Due to these ionization energies, it can be assumed that a non-negligible fraction of the atoms produced in the sources are not ionized and therefore, as uncharged particles, do not undergo any acceleration process at their source,
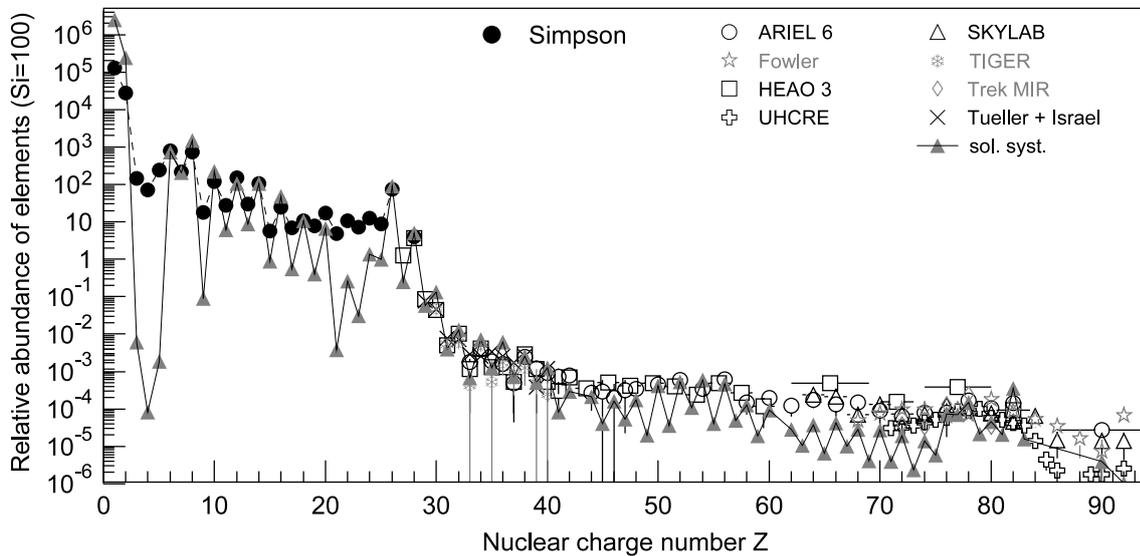
Figure 2.5.: Relative abundances of elements in primary cosmic rays as a function of the nuclear charge Z, at energies around 1 GeV per nucleon [Blü09]. The data points are normalized so that the relative abundance of silicon ($Z = 14$) corresponds to 100 on the ordinate. The data points for $Z \leq 28$ are taken from [Sim83]. For higher charge numbers, results from different experiments are included. For a complete list of references, see [Blü09]. The relative abundance of elements in the Solar System is shown for comparison [Lod03].

leading to a depletion of these elements in the primary cosmic rays observed at Earth.

On the other hand, the light elements lithium ($Z = 3$), beryllium ($Z = 4$), and boron ($Z = 5$) as well as the elements directly below iron ($Z = 26$) and directly below lead ($Z = 82$) are much more abundant in cosmic rays than in the Solar System. They are assumed to be produced in spallation processes of cosmic-ray nuclei during their propagation through the Galaxy [Blü09]. For example, nuclei from the carbon, nitrogen, oxygen (CNO) group fragment into the light elements Li, Be and B. Since the spallation cross sections for these nuclei are known from accelerator-based experiments, information about the propagation path length of cosmic rays though the Galaxy can be gained from measuring the ratio of boron to carbon [Blü09].

Elemental Composition at the Highest Energies

At the highest energies, the knowledge about the elemental composition of cosmic rays is far more limited. Due to the low incoming flux of UHECRs, only indirect measurements at air shower experiments are feasible. Here, the properties of the primary particle have to be inferred from measurements of the extensive air shower. Since hadronic interactions play an important role in the development of an air shower in the atmosphere (see also Sec. 2.3), the interpretation of air shower measurements depends strongly on the knowledge of hadronic interactions at the highest energies. In the UHE region, this knowledge is limited by the lack of data from accelerator-based experiments, hence, hadronic interaction models have to extrapolate from surveyed energy regions, which introduces a significant uncertainty in the interpretation of the measured air shower data. Consequently, a direct reconstruction of the nuclear mass A, or the charge Z, of the primary cosmic-ray particle is very difficult. Nevertheless, determining the elemental composition of UHECRs provides an important measurement to distinguish between the different models and scenarios for the origin of cosmic rays (see also Sec. 2.2.1).

An approach to infer the elemental composition of UHECRs is to exploit the distribution of the atmospheric depth of the shower maximum, X_{max}, which is correlated with the mass number A. Hence, the average depth $\langle X_{max} \rangle$ and the width of the distribution $\sigma(X_{max})$ as a function of energy are sensitive to changes in the overall elemental composition. In Fig. 2.6, experimental results on $\langle X_{max} \rangle$ and $\sigma(X_{max})$ from the Pierre Auger Observatory and HiRes are shown. The experimental results can then be compared with the expectations from air shower simulations for, e.g., a pure proton and a pure iron composition. However, a possible discrepancy exists: while data from the Pierre Auger Observatory indicate a change from a light composition to a heavier composition [Aab14a], the results from HiRes—and the successor experiment TA—are compatible with a proton-dominated composition over the whole energy range [Abb10, Jui12]. Understanding the differences between these experimental results and improving the precision of the composition measurements is one of the major challenges in current cosmic-ray physics.

Neutral Particles at the Highest Energies

Even though the term "cosmic rays" commonly does not include neutral particles like photons, neutrinos and neutrons, they are nevertheless closely linked, especially at the highest energies. Studying these neutral particles in addition to the study of charged cosmic rays

(a)

(b)
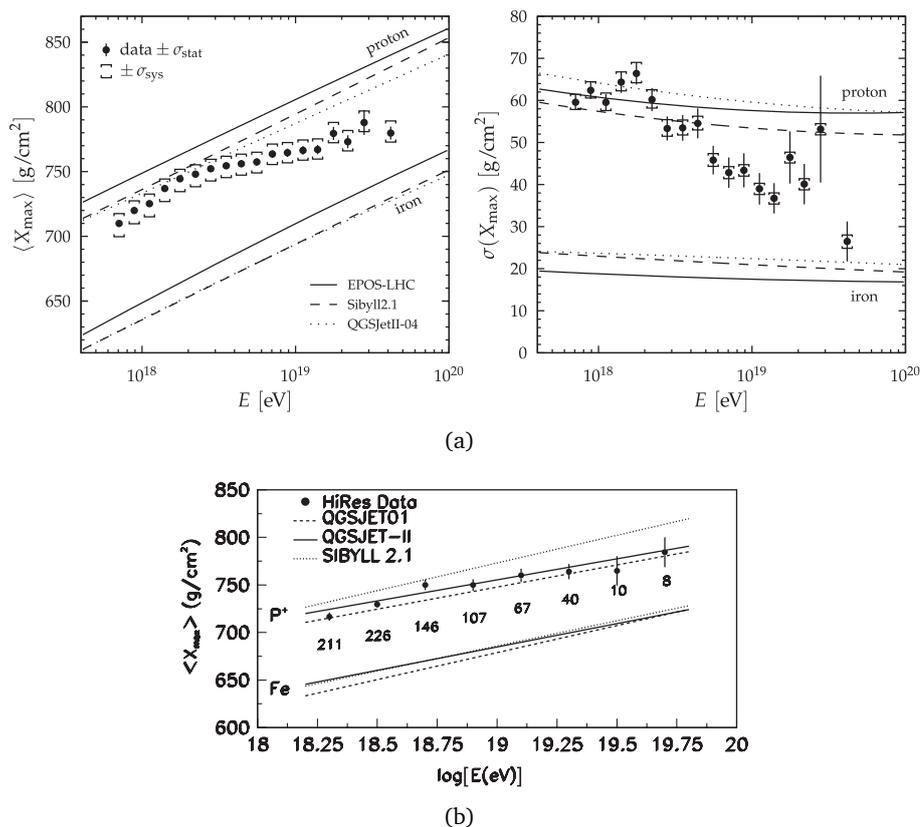
Figure 2.6.: Elemental composition at the highest energies. (a) Measurements of the average depth of the shower maximum, $\langle X_{\mathrm{max}} \rangle$, and the width of the X_{max} distributions, $\sigma(X_{\mathrm{max}})$, from the Pierre Auger Observatory [Aab14a]. (b) Measurements of $\langle X_{\mathrm{max}} \rangle$ from HiRes [Abb10]. The lines shown in the plots represent the expectations derived from MC simulations using different hadronic interaction models for a pure proton and a pure iron composition.

in a multi-messenger approach can reveal complementary information about the origin of UHECRs. In the following, UHE neutrinos and neutrons will be discussed briefly. UHE photons are the topic of Chap. 3.

Neutral particles are not deflected during their propagation through the galactic and extragalactic magnetic fields and thus carry information about the sources. In the atmosphere, they can initiate extensive air showers just like cosmic rays. In the case of neutrons, these showers are indistinguishable from showers induced by protons. At the Pierre Auger Observatory, a search for point sources of UHE neutrons has been performed by looking for an excess of air shower events with energies above the EeV range within the angular resolution of the detectors [Abr12c]. However, no point sources of UHE neutrons have been identified. This constrains models for sources of UHECRs within the Milky Way [Abr12c].

UHE neutrinos can be produced as a consequence of the GZK processes discussed in Sec. 2.2.1, Eq. 2.6, as the decay products of the charged pions. Thus, UHE neutrinos are of particular interest for the study of the origin of the flux suppression at the highest energies. However, UHE

neutrinos are difficult to detect due to their very small interaction cross section. On the other hand, this small cross section leads to a characteristic signature for extensive air showers induced by neutrinos, namely inclined showers that interact deeply in the atmosphere [Abr13]. In the searches for UHE neutrinos at the Pierre Auger Observatory, no neutrino events have been identified so far and upper limits have been placed on the diffuse flux of UHE neutrinos [Abr13]. The highest-energy neutrino events detected at the IceCube Neutrino Observatory at the South Pole have energies around 10^{15} eV [Aar13]. The upper limits on the diffuse flux of neutrinos in the UHE region determined by IceCube are close to the flux predicted by some theoretical models for the origin of UHE neutrinos [Aar13].

2.2.3. Anisotropy in Arrival Directions

Another important puzzle piece for understanding the origin of cosmic rays is the analysis of anisotropies in the arrival directions of cosmic rays at different angular scales. While large-scale anisotropies are connected to the propagation processes of cosmic rays through the Universe, small-scale anisotropies can provide direct hints toward sources or source regions of cosmic rays.

Large-scale Anisotropies

Due to the galactic magnetic fields, cosmic rays propagating through the Milky Way lose virtually all directional information, and their observed flux on Earth is highly isotropic (cf. Eq. 2.8). However, the leakage of cosmic-ray particles from the Galaxy may introduce a certain large-scale anisotropy in the form of a dipole structure spanning the whole sky [Ptu97]. The magnitude of this anisotropy is difficult to predict, as the calculations depend not only on assumptions on the—largely unknown—galactic magnetic fields but also on the concrete model that is used to describe the acceleration and propagation of cosmic rays in the Galaxy (see also Sec. 2.2.4). Some estimates predict anisotropies at the level of a few percent in the PeV to EeV range, see e.g. [Ptu93]. Another effect that causes a large-scale anisotropy in the arrival directions of cosmic rays is the Compton-Getting effect [Com35]. If the Earth moves with respect to the galactic centre, which can also be considered as the center of the rest frame of galactic cosmic rays, a small excess of cosmic ray events is expected in the direction of the motion and a deficit in the opposite direction, which would again lead to a dipole-like structure in the distribution of arrival directions. In the TeV range, large-scale anisotropies have been investigated by, e.g., the Super-Kamioka Nucleon Decay Experiment (Super-Kamiokande) and the Tibet Air Shower Array (Tibet ASγ). The results of these studies imply that the rest-frame of galactic cosmic rays is co-rotating with the Galaxy [Gui07, Ame06]. The results from Tibet ASγ in the PeV range are also compatible with the large-scale anisotropy studies at KASCADE-Grande, which yield no hints for anisotropy at these energies [Ant04a].

At the highest energies, it has been suggested that the Compton-Getting effect becomes more important, as extragalactic cosmic rays are expected to arrive isotropically and thus the movement of the Earth could induce again a dipole structure in the distribution of arrival directions [Kac06]. The orientation and energy dependence of an observed large-scale anisotropy may provide additional information about the transition between galactic and extragalactic cosmic rays. Thus, extensive searches for large-scale anisotropies in the UHE regime have been undertaken by different experiments, including the Pierre Auger Observatory. Although there are some hints indicating a large-scale anisotropy with an amplitude

at the level of the current statistical uncertainties, no significant deviation from isotropy has been found so far [Del13].

Small-scale Anisotropies

Anisotropies in the arrival directions of cosmic rays on a smaller angular scale are of high interest to identify possible source regions or point sources. The identification of individual cosmic ray sources would open up a new field of astronomy, namely charged-particle astronomy. However, as mentioned before, due to the magnetic fields in the Galaxy, the directional information of a cosmic ray event is lost to a large extent. The gyroradius r of a particle with charge number Z and energy E propagating through a magnetic field of strength B can, according to [Blü09], be approximated by

$$r\,[\text{pc}] = 1.08 \frac{E\,[\text{PeV}]}{Z \times B\,[\upmu\text{G}]}.$$ (2.8)

A proton with an energy of 1 PeV moving through a galactic magnetic field with a strength of about 3 μG has, according to Eq. 2.8, a gyroradius of about 0.4 pc. Since the Milky Way has a diameter of about 30 kpc, point sources for charged cosmic rays are not expected to be found in this energy range. However, this changes at the highest energies: for a proton with an energy of 10^{20} eV, the gyroradius exceeds the diameter of the Milky Way and the directional information toward the source is retained at least to some degree. Hence, the identification of source regions or individual sources may become possible at these energies—in particular also since the sources for UHE protons are expected to be located within the GZK horizon, i.e. relatively close to the Earth on cosmological scales. Small-scale anisotropies in the distribution of arrival directions, such as clustering or localized excesses, would be indicative of these source regions.

Another method to identify possible source regions is to investigate correlations of the arrival directions with celestial objects. The Pierre Auger Collaboration reported a correlation of the highest-energy cosmic-ray events with nearby AGNs [Abr07b, Abr10e]. Out of 69 UHECR events with energies larger than 55 EeV detected until December 31, 2009, 29 were found to correlate in their respective arrival directions within an angular window of 3.1° with the positions of AGNs with a redshift smaller than 0.018—corresponding to a distance smaller than 75 Mpc—taken from the 12th edition of the Véron-Cetty & Véron (VCV) catalog [Vér06]. A sky map of the arrival directions of these events and the positions of nearby AGNs from the VCV catalog is shown in Fig. 2.7. Assuming an isotropic flux of UHECR, only 14.5 events would be expected to correlate due to chance coincidences [Abr10e]. However, due to the limited statistics, one cannot conclude that AGNs are the sources of UHECRs. In a repetition of this correlation study by the HiRes Collaboration using the same event and AGN selection, 2 out 13 events were found to be correlated, while under the assumption of an isotropic flux, 3.2 correlated events would be expected [Abb08b]. The differences between these two analyses could be related to the different energy scales of the two experiments, or due to known deficiencies in the VCV catalog of AGNs, which has a different degree of completeness in the two hemispheres [Abr10e].
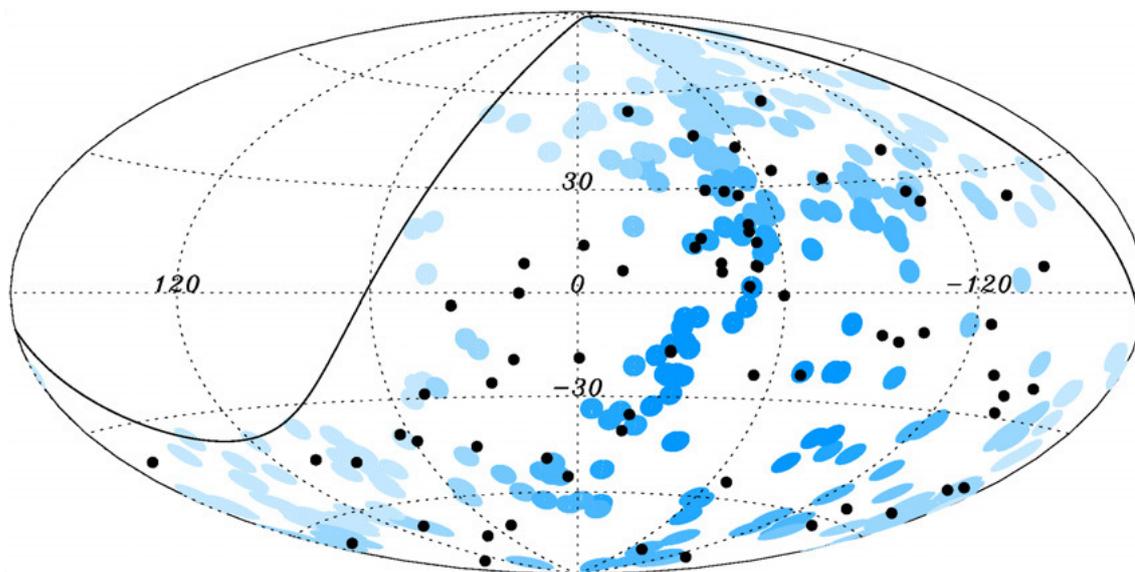
Figure 2.7.: The arrival directions of the highest-energy cosmic-ray events and the locations of nearby AGNs, shown in an Aitoff-Hammer projection of the sky in galactic coordinates [Abr10e]. The arrival directions of the 69 events with $E \geq 55\,\text{EeV}$ detected at the Pierre Auger Observatory until December 31, 2009, are plotted as black dots. The blue circles of radius 3.1° are centered at the positions of the 318 AGNs from the VCV catalog [Vér06] that lie within 75 Mpc from the Earth and are within the field of view of the Pierre Auger Observatory. A darker color of the circles denotes a larger relative exposure. The solid line represents the border of the field of view of the Pierre Auger Observatory for zenith angles below 60°.

2.2.4. Origin of Primary Cosmic Rays

Since the discovery of cosmic rays in 1912, the greatest puzzle has been their origin, i.e. the questions from what sources they originate from and how they acquire their tremendous energies. Consequently, there are many theories aiming at explaining the origin of cosmic rays. However, the difficulty is that the measurements of the different aspects of cosmic rays have to be consistently taken into account, i.e. their energy spectrum, their composition and possible anisotropies in their arrival directions (cf. the previous Sections 2.2.1 to 2.2.3). As discussed in Sec. 2.2.1, it is generally assumed that cosmic rays with energies below the knee and the second knee are of galactic origin, i.e. their sources are located within the Milky Way. At the highest energies, cosmic rays are assumed to be of extragalactic origin. Some astrophysical objects that are possible galactic or extragalactic sources of cosmic rays are discussed at the end of this section. Before the source candidates are discussed, the general mechanisms of how cosmic rays can acquire energies up to 10^{20} eV and beyond are briefly reviewed.

In general, the different theories and models which explain how cosmic rays, in particular UHECRs, gain their energies can follow two different, fundamental approaches. In the more conventional bottom-up approach, low-energy particles are accelerated to high and ultra-high energies. On the other hand, in the more exotic top-down approach, supermassive new particles are postulated which decay directly into the UHECRs that are observed on Earth.

Bottom-Up Models

In the bottom-up approach, the primordial cosmic-ray particles gain energy through acceleration in special astrophysical environments. The details of the acceleration as well as the maximum attainable energy strongly depend on the individual model. In general, the particles can gain energy either in direct acceleration processes, through strong electric fields, or in stochastic acceleration processes, where a particle gains energy step-wise through many interactions with, e.g., shock fronts.

Direct Acceleration   Direct acceleration mechanisms have already been proposed in the 1930s as a possible means for charged particles to attain energies in the GeV range in stellar electric fields [Swa33]. These electric fields are induced by the changing magnetic fields around sun spots on the surface of stars, where potential drops of 10^{10} V can be reached. More recently, similar one-shot mechanisms have been proposed for other astrophysical objects, e.g. rapidly spinning magnetized neutron stars (pulsars) such as the Crab or the Vela pulsars [Che86]. Here, potential drops of up to 10^{21} V may be possible in the magnetosphere of the pulsar [Ven97]. However, since the electric fields are located in regions of high energy densities, the cross sections for energy loss processes are high, reducing the efficiency of such acceleration processes.

Stochastic Acceleration   In stochastic acceleration models, the acceleration takes place not in a one-shot process as in the direct acceleration models discussed before, but rather through a series of interactions, each transferring a small amount of energy to the particle. This idea was first proposed in 1949 by Enrico Fermi, discussing interactions of charged particles with randomly moving, magnetized clouds of interstellar matter [Fer49]. Although an individual particle can gain or lose energy in one interaction with the magnetized cloud, there can be, after numerous interactions, an overall increase in the particle's energy, since interactions in which the particle gains energy are more probable than interactions in which the particle loses energy. Assuming a non-relativistic speed of the cloud, the average energy gain in a single interaction can be calculated to be proportional to the square of the relative velocity β of the magnetized cloud [Fer49]. Thus, this mechanism is also known as second-order Fermi acceleration. For a single magnetized cloud with a fixed β, the resulting energy spectrum of accelerated particles has the form of a power-law spectrum. However, the spectral index depends on β, and thus the energy spectrum observed on Earth (see Sec. 2.2.1) is a superposition of the energy spectra of particles coming from different acceleration sites and not a simple power-law spectrum.

Due to the dependence on β^2 and $\beta \ll 1$, second-order Fermi acceleration is a rather inefficient process. In the late 1970s, a more efficient acceleration process was proposed, involving stochastic acceleration on relativistic shock fronts [Bla78]. The compression of the interstellar medium at the shock front leads to different magnetic field topologies on either side of the shock front. Due to multiple scattering processes, a particle can cross the shock front numerous times and gain energy on each crossing. The fractional energy gain per crossing is directly proportional to β. Hence, this mechanism is also called first-order Fermi acceleration. However, in contrast to second-order Fermi acceleration, the spectral index of the resulting energy spectrum is independent of β. Instead, it depends on the compression ratio at the shock front [Bla78]. The maximum energy that can be transferred to a charged particle through this

mechanism is only limited by the number of acceleration cycles, which in turn only depends on the lifetime of the shock front.

Top-Down Models

Top-down models have been introduced as an alternative to the more conventional bottom-up models. In these models, hypothetical particles, referred to in the following as X particles, decay into UHE particles, circumventing the problem of acceleration to energies above the EeV range. These particles must be super-massive with rest masses $\gg 10^{20}$ eV/c^2, i.e. close to the mass scale of Grand Unified Theories (GUTs). In addition, the density of the X particles in the Universe and their decay rates have to be large enough to account for the observed flux of cosmic rays. The most prominent top-down models involve Topological Defects (TDs), Super-Heavy Dark Matter (SHDM) or the annihilation of relic neutrinos in Z-bursts.

In the case of TDs, it is assumed that they were formed in the very early stages of the Universe during the GUT symmetry breaking phase transition, for example in the form of magnetic monopoles [Hil83] or cosmic strings [Hin95]. In these models, the X particles have a very short lifetime, and thus, they have to be continuously produced in the TDs and emitted from there.

On the other hand, in the SHDM model, the X particles are assumed to be directly produced in the early stages of the Universe [Ber97]. Due to unknown symmetries, they have a lifetime comparable to the age of the Universe and thus constitute a fraction of the cold dark matter distributed in today's Universe. UHECRs originating from the decay of these relic particles would thus trace the cold dark matter distribution, resulting in e.g. an excess of UHECR events from the direction of the galactic halo as a signature of this model. However, data from the Pierre Auger Observatory show no significant excess from this direction [Abr07a].

In the Z-burst model, UHE neutrinos with energies above 10^{21} eV annihilate at the Z boson resonance with relic background neutrinos that are clustered e.g. in the galactic halo, initiating a cascade of particles whose signature might be the UHECRs observed on Earth [Far99]. However, also the UHE neutrinos themselves should be observable. Recent upper limits on the neutrino flux in the energy range above 10^{21} eV strongly constrain this model [Gor04].

Common to all top-down models is the prediction of a large flux of UHE photons as a consequence of the decay of the X particles. However, the current upper limits on the integral flux of UHE photons heavily constrain most top-down models, even excluding some models (see Sec. 3.4). Thus, top-down models are generally disfavored compared to the bottom-up models.

Possible Astrophysical Sources

So far, no astrophysical object has been unambiguously identified as a source for UHECRs. Nevertheless, a variety of source candidates exists. In general, an astrophysical object can be considered a source candidate—i.e. considered to be able to accelerate charged particles up to the highest energies, independent of the concrete acceleration mechanism—if either the magnetic fields at the source are strong enough or the source region itself is large enough to
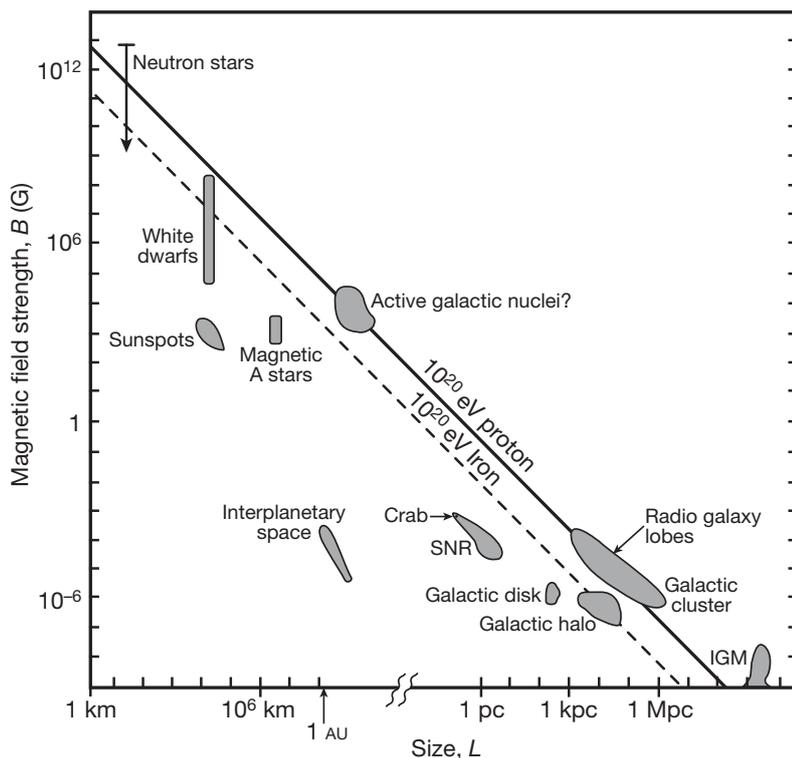
Figure 2.8.: Magnetic field strengths and sizes of the acceleration regions of several astro-
physical objects, shown in a Hillas plot [Bau09]. The solid line and the dashed line are
a graphical representation of Eq. 2.9 for a proton and an iron nucleus, respectively, to be
accelerated to 10^{20} eV, under the assumption of $\beta = 1$.

confine the particle within the acceleration region. This simple criterion has been defined in
1984 by Anthony Hillas [Hil84], and it can be written as

$$B\,[\mu G] \times L\,[\text{kpc}] > E\,[\text{EeV}]\frac{2}{Z\beta}, \tag{2.9}$$

where B denotes the magnetic field strength in the acceleration region and L the size of the
acceleration region. E is the maximum energy of the accelerated particle and Z its charge
number. Finally, β denotes the characteristic velocity of the magnetic scattering centers in
the acceleration region. Fig. 2.8 shows a graphical representation of Eq. 2.9 for various as-
trophysical objects in a Hillas plot. Very few objects satisfy the criterion imposed by Eq. 2.9
for the acceleration of protons to energies of 10^{20} eV. In the following, Active Galactic Nu-
clei (AGNs), pulsars (i.e. fast rotating neutron stars) and the hot spots contained in the radio
lobes of giant radio galaxies are discussed briefly. For a more detailed discussion, see e.g.
[Kot11]. Other promising source candidates for UHECRs, not shown in Fig. 2.8, are Gamma-
Ray Bursts (GRBs), which are also considered. In addition to these source candidates for
UHECRs, Supernova Remnants (SNRs) are discussed, since they are considered an important
source for galactic cosmic rays.

Supernova Remnants (SNRs) SNRs have been considered a source candidate for galactic
cosmic rays since the 1960s, following a simple argument based on the energetics of super-

novae: if only a small fraction of the energy released in supernovae is transferred to cosmic rays, then this could account for the total energy density contained in cosmic rays [Gin64]. In the stellar explosion, matter is ejected into the interstellar gas, forming a shock front at which charged particles can be accelerated in stochastic acceleration processes. The maximum achievable energy for charged particles through first-order Fermi acceleration in SNRs has been calculated to be, in general, in the PeV range [Ber96], although more recent calculations show that iron nuclei can be accelerated up to the EeV range in Type IIb SNRs [Ptu10]. Until recently, a direct evidence for the acceleration of protons in SNRs was lacking. In 2013, the Fermi Large Area Telescope (Fermi LAT) Collaboration identified a characteristic feature in the γ-ray spectra of two different SNRs, which can be related to the decay of neutral pions produced in interactions of the accelerated protons with interstellar matter, thus providing evidence for the acceleration of protons in SNRs [Ack13]. This strongly supports the assumption that SNRs are sources of galactic cosmic rays.

An example for an SNR, the Crab Nebula (NGC 1952), is shown in Fig. 2.9(a). This SNR is particularly interesting since it is one of the brightest persistent sources of γ-rays in the sky and is used as a standard candle in TeV γ-ray astronomy.

Active Galactic Nuclei (AGNs)   AGNs, located in the center of active galaxies, in general consist of an accretion disk around a central super-massive black hole. In radio-loud AGNs, matter is ejected in jets perpendicular to the accretion disk, which terminate in lobes that can be detected in the radio band. Radio-quiet AGNs, on the other hand, generally, do not exhibit

(a)   (b)
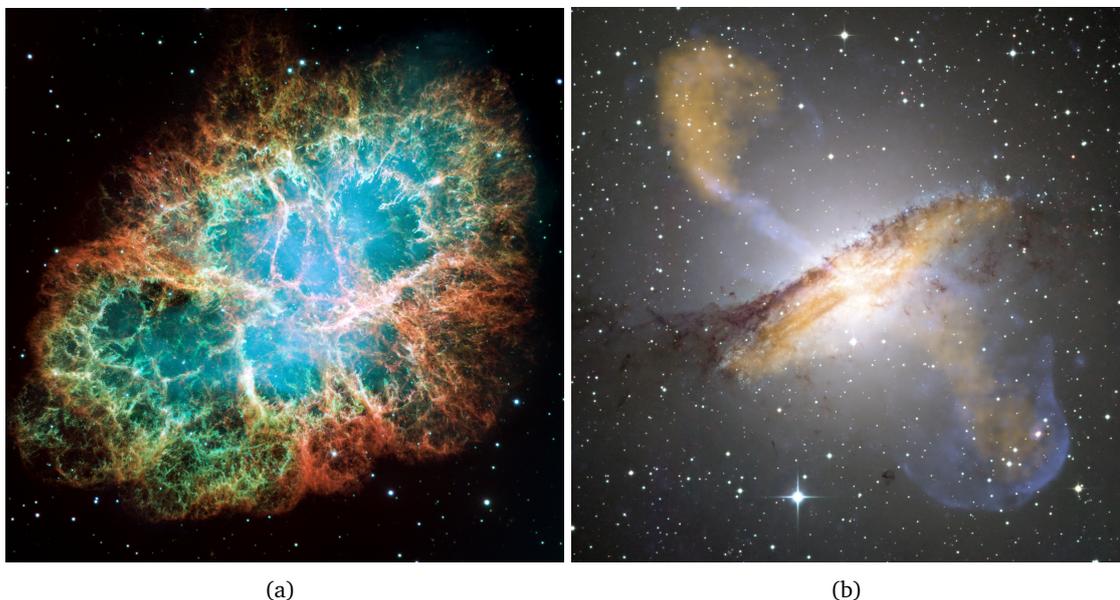
Figure 2.9.: Two examples of possible astrophysical sources of cosmic rays. (a) Composite image of the Crab Nebula (NGC 1952) [Hub05]. The different colors denote different elements in the expanding shell of the SNR. (b) Color composite image of Centaurus A (NGC 5128) [Eur09]. The different colors, in this case, denote the different wavelength bands that have been combined to form this composite image.

prominent radio emission or jets. In principle, both classes of AGNs can accelerate particles to the highest energies. In the central region of an AGN, the magnetic fields are of the order of $300\,\mathrm{G}$, while the size of this region is around $10^{-5}\,\mathrm{pc}$ [Kot11]. Particles with energies up to about $150\,\mathrm{EeV}$ could thus be confined in this region. Charged particles may be accelerated in direct acceleration processes in the electric fields surrounding the magnetosphere of the black hole, where potential drops of the order of $10^{21}\,\mathrm{V}$ may be reached [Bol99]. In radio-loud AGNs, acceleration of charged particles is also possible through first-order Fermi acceleration in the inner jets. Here, energies of the order of $10\,\mathrm{EeV}$ may be reached [Hal97]. However, under realistic conditions, charged particles will lose their energy due to interactions with photons from the dense radiation field around the center of the AGN in addition to the energy losses due to synchrotron radiation and adiabatic losses. Some calculations predict that, under these conditions, only particles with energies below $10^{16}\,\mathrm{eV}$ are emitted from an AGN [Nor95].

In Fig. 2.9(b), Centaurus A (NGC 5128) is shown as an example of an AGN. This AGN is assumed to be one of the most promising source candidates for UHECRs, since it is, at a distance of $3.8\,\mathrm{Mpc}$ [Har10], the closest AGN, well within the GZK horizon. In 2010, the Pierre Auger Collaboration reported a clustering in the distribution of arrival directions of UHECR events in the region around Centaurus A [Abr10e], which seems to corroborate this assumption.

Hot Spots of Radio Galaxies Fanaroff-Riley Class II (FR-II) radio galaxies [Fan74] are among the brightest radio sources in the sky. Even though every FR-II radio galaxy contains an AGN, they can be considered a separate class of possible sources of UHECR, since the acceleration of particles does not take place in the inner regions of the AGN, but rather in hot spots, i.e. regions of intense radiation, in the radio lobes at the end of giant jets emanating from the AGN [Rac93]. It is assumed that these hot spots contain the largest and most powerful shock waves in the Universe, which is supported by measurements of the synchrotron spectra emitted by these galaxies [Bie87]. In the hot spots, particles can be accelerated to energies in the order of $10^{21}\,\mathrm{eV}$ [Rac93]. In contrast to particles accelerated in the inner region of an AGN, particles accelerated in hot spots can leave the acceleration region without large energy losses, since the acceleration regions are located at the edges of the radio lobes, where the radiation field is much less dense than in the inner region of the AGN.

Pulsars Pulsars are formed when the core of a massive star collapses to a neutron star during a supernova. Although pulsars are very compact objects, they are accompanied by very strong magnetic fields (see Fig. 2.8). Hence, they are often considered as potential sites for acceleration of UHECRs. Most models involving pulsars assume direct acceleration processes in the magnetosphere of the neutron star, when the rapidly rotating magnetic fields induce strong electric fields at the surface of the star. The potential drop can be as large as $10^{21}\,\mathrm{V}$ [Ven97]. However, due to synchrotron radiation losses and the subsequent formation of cascades of electron-positron pairs—which influence the electric field—the maximum attainable energy for charged particles accelerated in this environment is in the order of $10^{15}\,\mathrm{eV}$ [Ven97]. On the other hand, some theories predict that young pulsars with strong magnetic fields (above $10^{12}\,\mathrm{G}$) and short spin periods (less than $10\,\mathrm{ms}$) are capable of accelerating iron nuclei from the surface of the pulsar to energies in the order of $10^{20}\,\mathrm{eV}$ through magnetohydrodynamic acceleration in the relativistic winds around the pulsar [Bla00].

Gamma-Ray Bursts (GRBs)   The source candidates discussed so far can all be regarded as continuous emitters of UHECRs. However, there are also models for transient sources such as GRBs. GRBs are highly luminous flashes of γ-rays associated with extremely energetic explosions in distant galaxies, whose exact nature is not yet understood. Provided that the energy released by a GRB in the form of cosmic rays is simular to the energy emitted as γ-rays, GRBs could account for the total cosmic ray flux at the highest energies [Wax95]. The explosion associated with a GRB can lead to multiple shock regions. Different models put the dominant acceleration regions either to inner shocks [Wax95] or outer shocks [Vie95]. Overall, the models predict a maximum energy of the emitted UHECRs in the order of 10^{20} eV. However, it is difficult to obtain direct evidence for GRBs as sources of UHECRs, because they are very rare events (10^{-6} to 10^{-5} per year per galaxy [Pod04]). In addition, since GRBs are transient objects, it is likely that possible UHECRs originating from a GRB arrive at Earth up to 10^7 years after the γ-rays from the GRB [Wax06]. On the other hand, this could be an explanation for the lack of significant correlations of the arrival directions of UHECRs with astrophysical objects (see Sec 2.2.3).

2.3. Extensive Air Showers

Due to the rapidly decreasing incoming flux of primary cosmic rays with energy, direct measurements of these particles become practically impossible above some 100 TeV [Eng12]. Already at the knee, the incoming flux is only about 1 particle per m^2 per year. Current balloon-borne and satellite-borne experiments are limited to an effective area of a few m^2 (see e.g. [Agu13]), so that collecting significant statistics in the energy region above the knee— and in particular at the highest energies—is not possible with these types of experiments on reasonable time scales. However, when the cosmic-ray particle enters the atmosphere, it will eventually interact with an air particle, producing secondary particles. The secondary particles may undergo subsequent interactions with air particles, thus initiating a cascade of particles in the atmosphere. This cascade—the extensive air shower—can be observed using a variety of detection techniques: for example, the Cherenkov or fluorescence light emitted by the secondary particles in the air shower can be measured using Imaging Air Cherenkov Telescopes (IACTs) or fluorescence detectors. Provided the primary cosmic-ray particle initiating the air shower is energetic enough so that the cascade reaches ground level, the particles from the air shower can also be measured by ground-based particle detectors. Since the particles from an extensive air shower are spread out over an area of several square kilometers at the highest energies, the single particle detectors on the ground can also be spread out. Hence, very large areas can be covered. From measurements of the secondary particles from an extensive air shower, the properties of the primary particle inducing the particle cascade in the atmosphere can be deduced.

In the following sections, the phenomenology of extensive air showers will be discussed, in particular the development of a shower in the atmosphere (Sec. 2.3.2). First, a simplified Heitler model is introduced as a toy model for the development of a particle cascade (Sec 2.3.1). The focus of this section lies on extensive air showers that are intiated by protons or nuclei. A separate discussion of extensive air showers induced by UHE photons will follow in Chap. 3.

In the discussion of the development of an extensive air shower in the atmosphere, it is useful to describe the development not as a function of the altitude h, but rather in terms of the atmospheric depth X. X is a measure of the matter traversed by the particle, and it can be calculated from

$$X(h) = \int_h^\infty \rho(h')\,\mathrm{d}h',\tag{2.10}$$

where the integration is carried out over the density profile $\rho(h)$ of the atmosphere effectively from the top of the atmosphere to the altitude h [Sta10]. Eq. 2.10 assumes that the particle is incident from the vertical. For a particle impinging on the atmosphere under a zenith angle θ, the integration has to be carried out along the inclined path of the particle through the atmosphere. For zenith angles below 60°, the curvature of the Earth can be neglected in the flat Earth approximation, and the atmospheric depth traversed by an inclined particle can be approximated as

$$X_{\mathrm{inc}}(h, \theta) = \frac{X(h)}{\cos(\theta)},\tag{2.11}$$

with the vertical atmospheric depth $X(h)$ according to Eq. 2.10 [Sta10].

2.3.1. Simplified Heitler Model

Although an extensive air shower is comprised of billions of particles and involves many different interactions, the key features of the development of an air shower in the atmosphere can be understood through a toy model, proposed by Bhabha and Heitler already in the 1930s [Bha37]. In the following, a simplified Heitler model is discussed [Sta10]. This simplified model describes a particle cascade consisting of identical particles that interact after a fixed interaction length λ. In each interaction, two new particles are created, each carrying half of the original particle's energy. Thus, after every interaction length, the number of particles N is doubled, while the energy per particle E is halved. At an atmospheric depth $X = n\lambda$, the cascade consists of $N = 2^n$ particles, each of which carries an energy $E = E_0/2^n$, with E_0 being the energy of the primary particle initiating the cascade. A schematic view of such a particle cascade is shown in Fig. 2.10.

The particle cascade continues until the energy per particle reaches the critical energy E_c, below which no more interactions are possible. The number of particles does not increase any more, and the particle cascade stops. The maximum number of particles in the cascade is

$$N_{\mathrm{max}} = \frac{E_0}{E_c},\tag{2.12}$$

and the atmospheric depth where N_{max} is reached is given by

$$X_{\mathrm{max}} = \frac{\lambda}{\ln(2)}\,\ln\left(\frac{E_0}{E_c}\right).\tag{2.13}$$

The simplified Heitler model describes the particle cascade only until the maximum number of particles is reached. After the maximum is reached, the particles can, for example, decay, or be absorbed, thus reducing the number of particles, as indicated in Fig. 2.10, but these processes are not in the scope of the simplified Heitler model.

X/λ	N	E/E_0
0	1	1
1	2	1/2
2	4	1/4
3	8	1/8
4	16	1/16
n_{max}	N_{max}	E_c
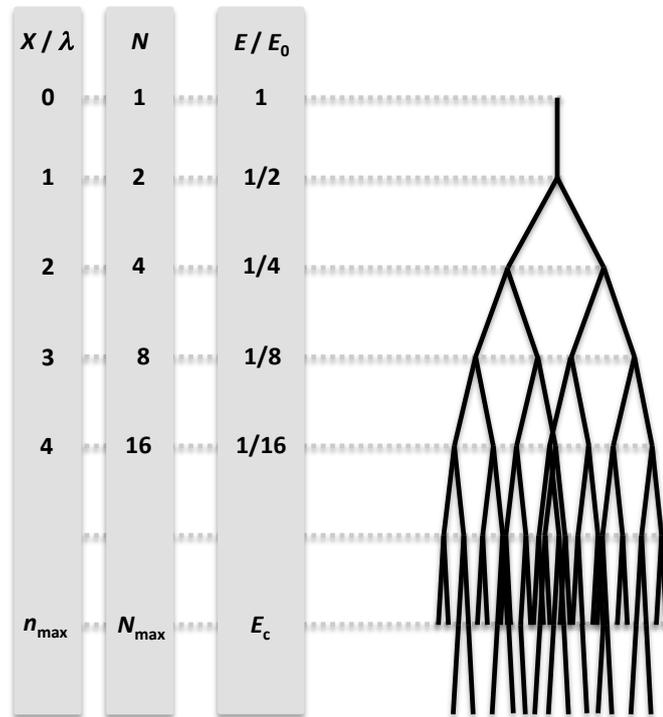
Figure 2.10.: Schematic depiction of a particle cascade in the simplified Heitler model. After each interaction length λ, the number of particles N is doubled, while the energy per particle E is halved. The cascade starts at an atmospheric depth $X = 0$ with one particle with energy E_0, and it continues until the critical energy E_c is reached after n_{max} interaction lengths. At this point, the cascade contains the maximum number of particles.

Even though the Heitler model is very simple, its predictions of the most important features of particle cascades, notably that N_{max} is proportional to E_0 and that X_{max} is proportional to $\ln(E_0)$, are qualitatively in good agreement with detailed MC simulations of particle cascades [Sta10].

2.3.2. Development in the Atmosphere

When a primary cosmic-ray particle, i.e. a proton or a nucleus, enters the atmosphere, the first interaction with a nucleus from the air in the atmosphere—most probably a nitrogen nucleus, since this element is the most abundant at altitudes below 100 km—will take place at an altitude between 15 and 35 km, depending on the mass of the primary cosmic-ray particle, its energy, and the incident angle [Eng12]. The probability P that a primary cosmic-ray particle will interact at a given atmospheric depth X follows an exponential distribution:

$$\frac{\mathrm{d}P}{\mathrm{d}X} = \frac{1}{\lambda_{int}}\, e^{-X/\lambda_{\mathrm{int}}}, \tag{2.14}$$

where λ_{int} denotes the hadronic interaction length, which is inversely proportional to the inclusive cross section for an interaction of a given type of cosmic-ray particle with air [Eng12]. For cosmic-ray protons, the hadronic interaction length is about $60\,\mathrm{g\,cm^{-2}}$ at the knee, decreasing slowly with energy. In the first interaction, mainly pions, kaons and nucleons are
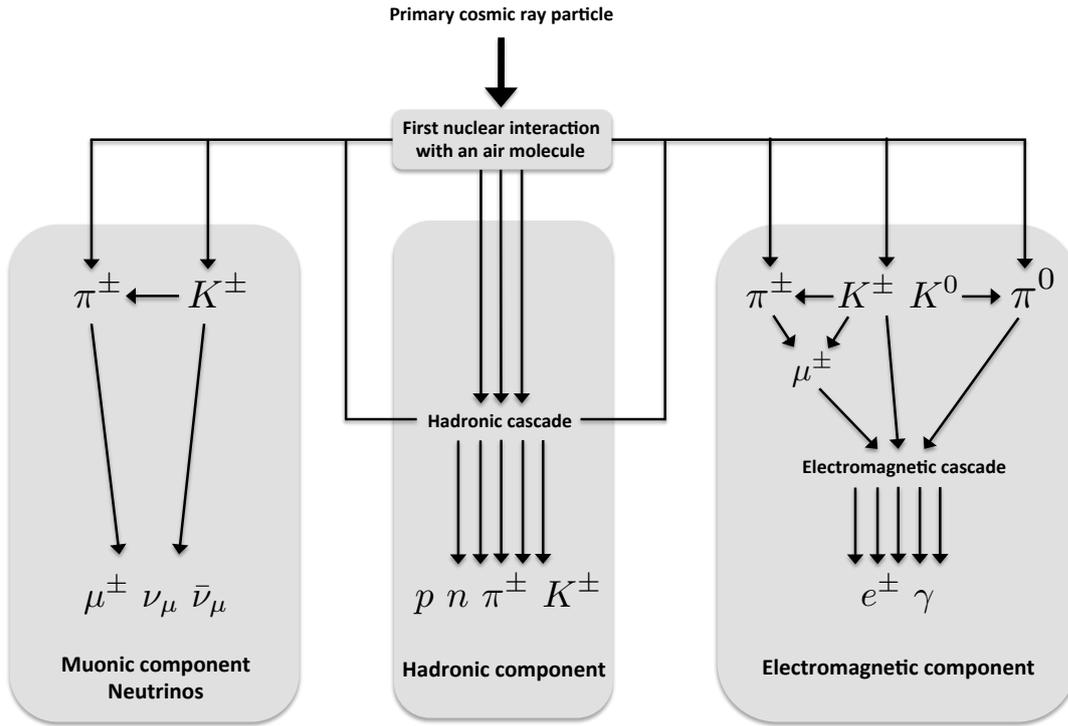
Figure 2.11.: Schematic depiction of the different components of an extensive air shower, including the basic processes and the particles involved in these processes.

produced. These secondary hadrons will also interact with a nucleus from the air, and thus form a hadronic cascade. This cascade forms the hadronic component of an extensive air shower. The hadronic component is crucial for the overall development of the extensive air shower, since it feeds the other components of the shower: the decay of neutral pions into two photons gives rise to electromagnetic cascades which form the electromagnetic component, while the decay of charged hadrons into muons leads to the muonic component. A schematic depiction of the different components of an extensive air shower, including the basic processes and the particles involved in these processes, is shown in Fig. 2.11. In the following, the different components will be briefly discussed.

Hadronic Component

The hadronic particle cascade that forms the hadronic component of an extensive air shower can, in a simplified way, be described by an extended Heitler model (see Sec. 2.3.1) [Mat05]. In this extended Heitler model, only pions are considered as secondary particles from the hadronic interactions, since they are the lightest hadrons and thus the most frequently produced secondary particles. Again, it is assumed that the particles interact with nuclei from the air after a fixed interaction length λ. For pions, the interaction length is $120 \, \mathrm{g \, cm^{-2}}$ at an energy of $100 \, \mathrm{GeV}$, decreasing slowly with energy. In contrast to the simplified Heitler model discussed in Sec. 2.3.1, not only two secondary particles are produced, but n_{charged} charged pions together with $\frac{1}{2} n_{\mathrm{charged}}$ neutral pions. Although n_{charged} is energy-dependent, a constant value of $n_{\mathrm{charged}} = 10$ has been found to be in good agreement with measurements [Mat05]. The secondary particles produced in the first interaction will subsequently interact with nuclei

from the air, if their interaction length λ_{int} is larger than their decay length. Since neutral pions have a very short lifetime of $8.5 \times 10^{-17}\,\mathrm{s}$ [Ber12], it is assumed that they decay instantly into photons, thus initiating the electromagnetic cascades that form the electromagnetic component. The charged pions on the other hand, with their longer lifetime of $2.8 \times 10^{-8}\,\mathrm{s}$ [Ber12], continue through the atmosphere and interact again, producing more secondary particles. The particle cascade continues until the energy per pion is reduced to the critical energy E_c^{π}. At this energy, the decay length of a charged pion exceeds its interaction length, thus the probability for the pion to decay is larger than the probability for a subsequent interaction. In the extended Heitler model, a constant value of $E_c^{\pi} = 20\,\mathrm{GeV}$ is usually adopted [Mat05]. As in the simplified Heitler model, once the critical energy is reached, the particle cascade stops.

While the predictions of the extended Heitler model are in general in good agreement with detailed MC simulations, it should be noted that it neglects some fundamental aspects of the hadronic interactions relevant in the development of a hadronic cascade, for example the inelasticity κ of the interactions [Mat05]. While the model assumes that the energy of the primary particle is distributed equally between all secondary particles, usually a fraction $1 - \kappa$ of the energy is transferred to one leading particle. For a more detailed discussion of this effect, see [Mat05]. In addition, the density profile of the atmosphere is not fully taken into account in this model. Since the atmosphere is very thin at high altitudes, where the first interactions of the particle cascade take place, the interaction length may in fact exceed the decay length of the charged pions, leading to a decay of a part of the charged pions produced int the first interactions [Sta10].

Electromagnetic Component

The electromagnetic component of an extensive air shower is fed from the hadronic component mainly through the decay of neutral pions:

$$\pi^0 \rightarrow \gamma + \gamma. \tag{2.15}$$

The photons from these decays will eventually produce electron-positron pairs:

$$\gamma + N_{\mathrm{Air}} \rightarrow e^- + e^+ + N_{\mathrm{Air}}, \tag{2.16}$$

where the nucleus from the air is needed for reasons of momentum conservation. The electrons and positrons may, in turn, emit photons in bremsstrahlung processes in the Coulomb fields of the nuclei from the air:

$$e^{\pm} + N_{\mathrm{Air}} \rightarrow e^{\pm} + \gamma + N_{\mathrm{Air}}. \tag{2.17}$$

Through these processes, an electromagnetic cascade is formed, which can be described by the Heitler model [Bha37]. The interaction length for electrons is the characteristic radiation length $X_{\mathrm{rad}} = 37\,\mathrm{g\,cm^{-2}}$. For reasons of simplicity, the interaction length for photons is, in the Heitler model, assumed to be the same as for electrons. The critical energy for an electromagnetic cascade is the energy at which energy losses due to collisions of the particles, i.e. Compton scattering for photons and ionization energy losses for electrons, begin to dominate the radiative energy losses. In air, the critical energy is $E_c^{\mathrm{em}} = 85\,\mathrm{MeV}$.

Due to the relatively short interaction length, an electromagnetic cascade is quickly absorbed in the atmosphere. Nevertheless, since the electromagnetic component of an extensive

air shower is, in total, comprised of a superposition of all the electromagnetic cascades originating from the neutral pions that are continuously produced in the hadronic cascade, the electromagnetic component quickly becomes the dominant component of an extensive air shower with respect to the number of particles (see Fig. 2.12).

Muonic Component

In the same way the electromagnetic component of an extensive air shower is fed by the hadronic component through the decays of neutral pions, the muonic component is fed by the hadronic component through the decays of charged mesons, mainly pions and kaons. The most relevant processes are

$$
\begin{aligned}
\pi^\pm &\to \mu^\pm + \nu_\mu(\overline{\nu}_\mu), \\
K^\pm &\to \mu^\pm + \nu_\mu(\overline{\nu}_\mu), \\
K^\pm &\to \pi^0 + \mu^\pm + \nu_\mu(\overline{\nu}_\mu).
\end{aligned}
\tag{2.18}
$$

In addition, up to 10 % of the low-energy muons in an extensive air shower can be produced in photo-nuclear interactions of photons from the electromagnetic component with nuclei from the air [Eng12].

The number of muons in an extensive air shower can be estimated using the extended Heitler model discussed before. Considering a primary proton with energy E_0, the number of charged pions after n interaction lengths is $N_{\text{charged}} = n_{\text{charged}}^n$, while the energy per pion is

$$
E_\pi = \frac{E_0}{\left(\frac{3}{2} n_{\text{charged}}\right)^n},
\tag{2.19}
$$

where also the number of neutral pions has to be taken into account (see above). The particle cascade stops, when the critical energy E_c^π is reached after n_c interaction lengths:

$$
n_c = \frac{\ln\left(\frac{E_0}{E_c^\pi}\right)}{\ln\left(\frac{3}{2}n_{\text{charged}}\right)}.
\tag{2.20}
$$

Assuming that all pions at this stage of the cascade decay into muons, the total number of muons can be calculated with

$$
N_\mu = n_{\text{charged}}^{n_c} = \left(\frac{E_0}{E_c^\pi}\right)^\beta, \qquad \text{with} \quad \beta = \frac{\ln\left(n_{\text{charged}}\right)}{\ln\left(\frac{3}{2}n_{\text{charged}}\right)} \approx 0.85.
\tag{2.21}
$$

Therefore, the number of muons increases with the primary energy [Mat05].

Since muons have a finite lifetime of 2.2 µs [Ber12], a decay via

$$
\mu^\pm \to e^\pm + \nu_e(\overline{\nu}_e) + \overline{\nu}_\mu(\nu_\mu)
\tag{2.22}
$$

is possible for low-energy muons, which are scarcely affected by relativistic time dilation. Higher-energy muons mostly reach the ground, passing the atmosphere nearly undisturbed.

Neutrinos are closely connected to the muonic component. With each muon, a neutrino is produced in turn (see Eq. 2.18), and in the decay of a muon, even two neutrinos are produced (see Eq. 2.22). However, the interaction cross section for neutrinos is very small. Therefore, they play an inferior role in the development of an extensive air shower. Nevertheless, the energy that is carried by neutrinos has to be taken into account when interpreting measurements of an extensive air shower, e.g. in the form of a correction factor.

Overall Longitudinal Development of an Air Shower

In Fig. 2.12(a), the average longitudinal development of the single components of an extensive air shower, i.e. the number of particles as a function of the atmospheric depth along the shower axis—which is defined by the incoming direction of the primary particle—is shown for extensive air showers induced by vertical protons with a primary energy of 10^{19} eV. The distributions have been obtained from detailed MC simulations using Cosmic Ray Simulations for KASCADE (CORSIKA) [Hec98]. The electromagnetic component quickly becomes the dominant shower component, outnumbering hadrons and muons by several orders of magnitude. In an extensive air shower induced by a proton with an energy of 10^{20} eV, a fraction of $90 - 95\%$ of the energy is transferred to the electromagnetic component [Eng12]. Hence, the measurable quantities of an extensive air shower are mainly determined by the electromagnetic component.

The overall longitudinal development of an extensive air shower can be parameterized in the following form [Gai77]:

$$N(X) = N_{\max} \left(\frac{X - X_0}{X_{\max} - \lambda} \right)^{\frac{X_{\max} - \lambda}{\lambda}} \exp \left(-\frac{X - X_0}{\lambda} \right), \qquad (2.23)$$

where the parameter $N_{\max}$ denotes the total number of particles at the shower maximum, which is reached at the atmospheric depth $X_{\max}$. X_0 and λ are the atmospheric depth of the first interaction of the primary particle and the average mean free path of the particles in the cascade, respectively. Eq. 2.23 is commonly known as the Gaisser-Hillas function, and it is used as a standard fit function when reconstructing the longitudinal profile from air shower measurements.

Lateral Distribution of Particles

For ground-based measurements, only the distribution of particles on ground level is of interest. In Fig. 2.12(b), this distribution is shown in terms of the particle density as a function of the distance from the shower core, for an atmospheric depth of $870 \, \mathrm{g \, cm^{-2}}$, which corresponds to the vertical atmospheric depth of the Pierre Auger Observatory. Again, the electromagnetic component outnumbers the other components by several orders of magnitude. In general, all components exhibit a wide lateral distribution, where particles can be measured several kilometers away from the impact point of the shower axis on the ground.

The lateral distribution of the hadronic component is mainly due to the transverse momenta of the secondary hadrons produced in the hadronic cascade [Eng12]. The transverse momentum is $p_\perp = 300 - 400 \, \mathrm{MeV/c}$ for each hadron, almost independent of the energy of the primary particle. Hence, the angle of lower-energy hadrons relative to the shower axis is
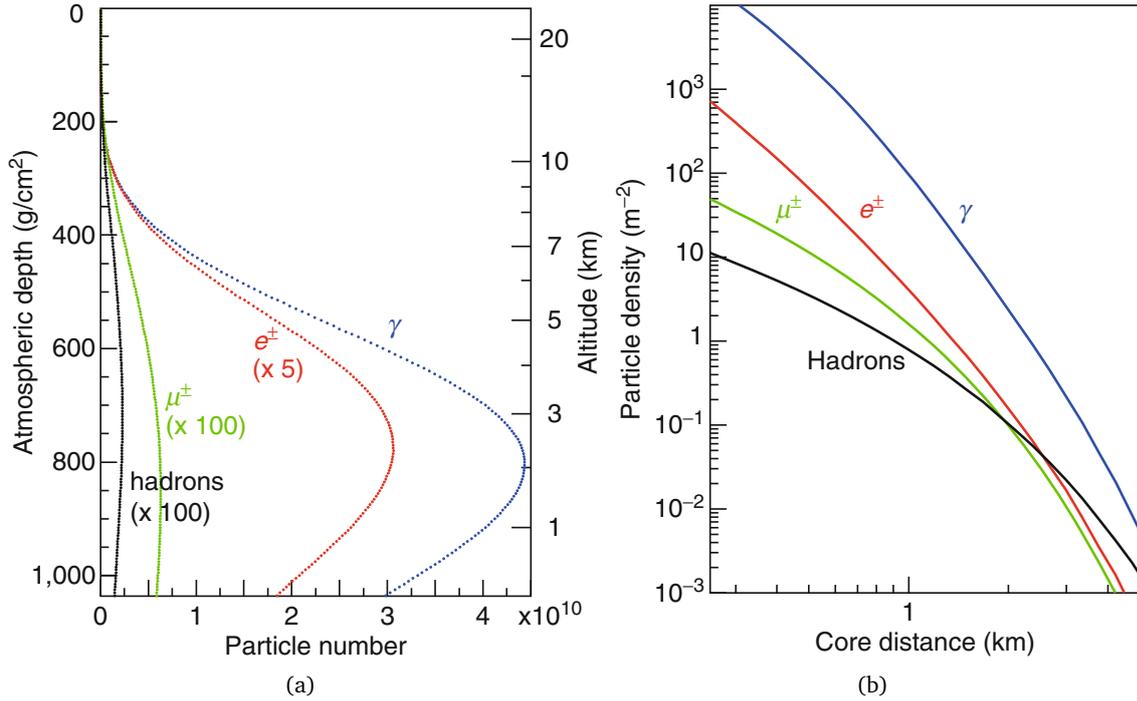
Figure 2.12.: Development of the single components of an extensive air shower in the atmosphere [Eng12]. (a) Number of particles as a function of atmospheric depth. (b) Lateral distribution of particles on ground, calculated for an atmospheric depth of $870\,\mathrm{g\,cm^{-2}}$, which corresponds to the vertical atmospheric depth of the Pierre Auger Observatory. The distributions shown in the plots have been obtained from detailed MC simulations of extensive air showers induced by vertical protons with energies of 10^{19} eV using CORSIKA. The energy thresholds of the simulations were $0.25\,\mathrm{MeV}$ for γ and $e^\pm$, and $0.1\,\mathrm{GeV}$ for $\mu^\pm$ and hadrons.

large, leading to the large spread of the hadronic component. The lateral spread of the electromagnetic component, mainly determined by multiple Coulomb scattering, can be described with a Lateral Distribution Function (LDF) based on the Nishimura, Kamata, Greisen (NKG) approach [Gre56, Kam58]. In this approach, the area density of electrons ρ_e, as a function of the distance r from the shower axis, is given by

$$\rho_e(r) = \frac{N_e}{2\pi r_M^2} \frac{\Gamma(4.5-s)}{\Gamma(s)\,\Gamma(4.5-2s)} \left(\frac{r}{r_M}\right)^{s-2} \left(1+\frac{r}{r_M}\right)^{s-4.5}, \qquad (2.24)$$

where N_e is the total number of electrons in the shower and r_M is the Molière radius, which is a characteristic measure of the lateral spread of electrons in a given material [Eng12]. The parameter s—also known as the shower age—is phenomenologically defined through

$$s = \frac{3X}{X + 2X_{\max}}, \qquad (2.25)$$

where X is the atmospheric depth at which the LDF should be evaluated.

Superposition Model for Nuclei

If the primary cosmic-ray particle is not a proton, but a heavier nucleus with mass number *A*, the extensive air shower initiated by this nucleus can be described using the superposition model. Since the binding energy of a single nucleon in the nucleus is much smaller than the total energy *E* of the nucleus, it can be considered as *A* independent nucleons with energy *E*/*A*, each initiating a single shower [Eng12]. The superposition of these sub-showers equals the total shower as induced by the nucleus. According to the simplified Heitler model and Eq. 2.13, the average atmospheric depth of the superposition of all showers is then given by

$$X_{\text{max}}(E_0, A) \propto \ln\left(\frac{E_0}{A E_c}\right). \tag{2.26}$$

The dependence on *A* implies that, on average, showers initiated by nuclei develop higher in the atmosphere than showers initiated by protons with the same primary energy, leading to a smaller X_{max}. This can be utilized to determine the composition of cosmic rays by measuring the average X_{max} for a given primary energy (see Sec. 2.2.2). In addition, also the number of muons is sensitive to the mass of the primary particle: adding together the muon numbers (Eq. 2.21) of the single sub-showers yields

$$N_\mu(E_0, A) = A\left(\frac{E_0}{A E_c^\pi}\right)^\beta = A^{1-\beta}\left(\frac{E_0}{E_c^\pi}\right)^\beta. \tag{2.27}$$

Eq. 2.27 implies that for example an extensive air shower initiated by an iron nucleus (*A* = 56) contains about a factor of 1.8 more muons than a shower initiated by a proton. Hence, determining the number of muons in an air shower on an event-by-event basis has become one of the main challenges for current and future air shower experiments.

<div style="text-align: right; font-size: 2em; color: #2244aa;">*3*</div>

Ultra-High-Energy Photons

Contents

In Chap. 2, cosmic rays, in particular UHECRs with energies above 10^{18} eV, have been discussed. The following chapter focuses on UHE photons, which—although commonly not referred to by the term cosmic rays—are intimately connected to UHECRs. The search for UHE photons complements measurements of charged cosmic rays and neutrinos at the highest energies toward multi-messenger observations of the Universe. However, an observation of photons in the EeV range would itself open a new window to the Universe, making astronomy at EeV energies possible. Besides the connections to cosmic rays and astroparticle physics in general, which are discussed in this chapter, there are also connections to other, more fundamental areas of physics. For example, observations of UHE photons and the extensive air showers initiated by these particles can serve, through the interactions of the UHE photons with lower-energy photons or nuclei, as probes of aspects of QED and Quantum Chromodynamics (QCD) at the highest energies, which are currently inaccessible at collider experiments.

The structure of this chapter follows the path of an UHE photon through the Universe. In Sec. 3.1, the main production mechanisms for UHE photons are discussed. The interactions of a UHE photon with lower-energy background photons during the propagation through the Universe are examined in Sec. 3.2. When UHE photons eventually reach the Earth, they will initiate extensive air showers in the atmosphere, just like UHECRs. The extensive air showers initiated by UHE photons are briefly reviewed in Sec. 3.3. The differences between showers induced by UHE photons and those initiated by protons or nuclei are discussed. Additional processes relevant for the development of an extensive air shower induced by a UHE photon

in the atmosphere are also considered. The chapter closes with a review of the current status of the search for UHE photons (Sec. 3.4). Current upper limits both for the diffuse and the directional flux of UHE photons are summarized.

3.1. Origin of UHE Photons

The dominant production mechanism of UHE photons is the decay of neutral pions, which are created through some "primary process" [Ris07]:

$$\text{primary process} \to \pi^0(+\pi^\pm + ...) \to \gamma_{\text{UHE}}(+\nu_{\text{UHE}} + ...). \tag{3.1}$$

The production of neutral pions in the primary process can be accompanied by the production of other particles such as charged pions, which can then decay into UHE neutrinos. UHE photons are inherently connected to charged cosmic rays at the highest energies. The exact nature of the primary process depends on the theoretical model that is used to explain the origin of UHECRs (see Sec. 2.2.4).

In the models that follow the bottom-up approach, the dominant primary process is the GZK process discussed in Sec. 2.2.1:

$$p + \gamma_{\text{CMB}} \to \Delta(1232)^+ \to p + \pi^0 \to p + \gamma_{\text{UHE}} + \gamma_{\text{UHE}}. \tag{3.2}$$

The two UHE photons emerging from the decay of the neutral pions produced in the GZK process typically each carry away 10 % of the energy of the primary proton [Ris07]. Since the energy threshold for the GZK process is 5×10^{19} eV, the UHE photons eventually emerging from this process will have a minimum energy of 5×10^{18} eV—but the energy of the photons observed on Earth is lower due to interactions during their propagation to the Earth (see Sec. 3.2). In the case of UHE nuclei instead of protons, the dominant energy loss mechanism is not the GZK process, but rather photo-disintegration (see Sec. 2.2.1). However, UHE protons produced in photo-disintegration processes of UHE nuclei may subsequently undergo the GZK process and eventually produce UHE photons [Hoo11]. To predict the flux of UHE photons from the theoretical models for the acceleration of UHECRs, many assumptions have to be made about the sources themselves, most notably the composition and the energy spectrum of the UHECRs at the source and the distribution of the sources in the Universe. Measurements of the energy spectrum and the composition of UHECRs impinging on the Earth can be used to constrain these calculations, although large uncertainties exist, especially with measurements of the composition (see Sec. 2.2.2). In [Gel08], the flux of UHE photons is calculated for a homogeneous distribution of sources that emit only protons—a scenario that is compatible with composition measurements from HiRes and TA. A scenario that is more compatible with composition measurements from the Pierre Auger Observatory, i.e. compatible with a composition dominated by heavy nuclei at the highest energies, is presented in [Hoo11]. The flux of UHE photons in this scenario is up to two orders of magnitude smaller than the flux calculated in the pure-proton scenario.

For the theoretical models that follow the top-down approach, the primary process is directly related to the decay or the annihilation of the super-massive X particles postulated in most of these models. In the decay or annihilation processes, typically two or more quarks and gluons are produced, which initiate QCD cascades [Alo04]. Eventually, the partons in the

cascades hadronize, and thus the UHECRs are produced along with a large number of pions. In the Z-burst scenario, similar QCD cascades are initiated by quarks and antiquarks from the decay of the Z bosons produced in the resonant interactions of UHE neutrinos with relic background neutrinos. In all top-down models, a large number of neutral pions is expected, which leads to a predicted flux of UHE photons that is, for some models, two orders of magnitude larger than in the bottom-up models discussed before [Gel08]. Also here, some uncertainties exist and assumptions have to be made, e.g., on the density and lifetime of the X particles.

The flux predictions mentioned so far all refer to a diffuse flux of UHE photons, i.e. a flux that is independent of the arrival direction. Some scenarios predict an additional component of the flux from certain directions in the sky. For example, if UHECRs pass through the central region of the Milky Way, they may interact with starlight and infrared photons in processes similar to the GZK process (Eq. 3.2), producing UHE photons [Kus06]. In another scenario, Centaurus A is considered as a source for UHE photons [Kac10]. The energy spectrum of photons emitted from Centaurus A calculated in this scenario is in good agreement with measurements from the High Energy Stereoscopic System (H.E.S.S.) in the TeV range. However, the production mechanisms for UHE photons considered in [Kac10] are based on hadronic processes, and thus require that UHECRs are accelerated in Centaurus A at least to EeV energies.

3.2. Propagation through the Universe

To obtain a prediction for the flux of UHE photons that is observed on Earth, not only the theoretical models for the production of UHE photons have to be taken into account, but also the propagation from the sources or source regions to the Earth. For charged particles, the propagation through the Universe is mainly affected by the galactic and extragalactic magnetic fields the particles traverse on their way from the sources to the Earth. Photons, on the other hand, are neutral particles and are as such not subject to any deflections in magnetic fields. However, like charged particles that interact with photons from the CMB and thus lose energy during propagation, also photons may interact with photons from the different cosmic background radiation fields, producing electron-positron pairs [Ris07]:

$$\gamma_{\mathrm{UHE}} + \gamma_{\mathrm{background}} \rightarrow e^- + e^+. \tag{3.3}$$

The threshold energy E_{thr} for pair production with a background photon of energy ϵ is

$$E_{\mathrm{thr}} = \frac{m_e^2 c^2}{\epsilon} \simeq \frac{2.6 \times 10^{11}\,\mathrm{eV}}{\epsilon\,[\mathrm{eV}]}, \tag{3.4}$$

where m_e denotes the electron mass [Bha00]. Hence, for UHE photons propagating through the Universe, background photons with an energy $\epsilon \lesssim 10^{-6}\,\mathrm{eV}$, corresponding to a frequency of 100 MHz, play an important role. However, the Universal Radio Background (URB) of photons with frequencies in the MHz regime is not well known, mainly because for measurements of the URB, it is very difficult to disentangle the galactic component from the extragalactic component [Bha00]. When the energy of the primary photon becomes smaller, also interactions with the CMB and the Infrared (IR) background become increasingly important. The energy loss lengths of photons due to interactions with the background radiation fields are shown in Fig. 3.1. Typical energy loss lengths for UHE photons range between $7 - 15\,\mathrm{Mpc}$ at
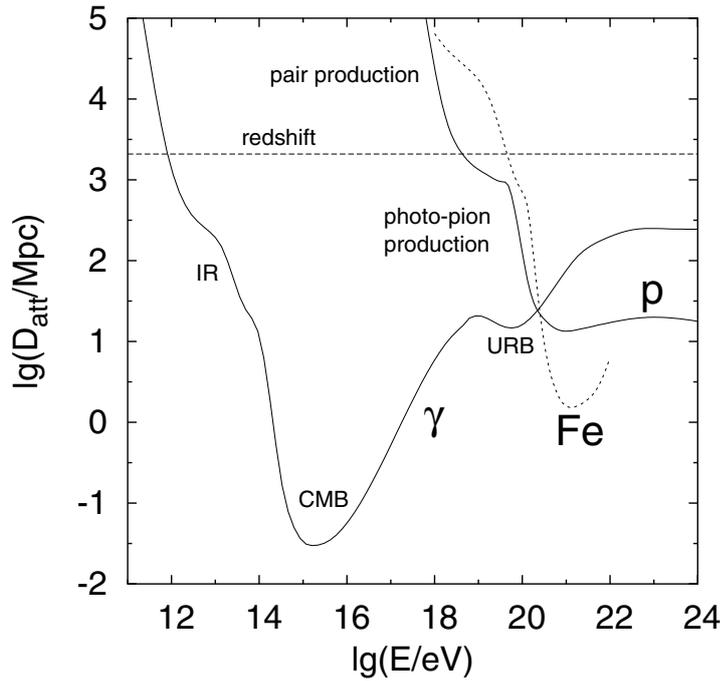
Figure 3.1.: Energy loss length of photons as a function of energy for interactions with the IR
background, the CMB and the URB [Ris07]. For comparison, also the energy loss lengths
of protons and iron nuclei are included. The dashed line gives the energy loss length for
adiabatic energy losses due to the expansion of the Universe.

10^{19} eV and $5 - 20$ Mpc at 10^{20} eV [Ris07]. As a comparison, also the energy loss lengths of
protons and iron nuclei are shown in Fig. 3.1. The dominant processes through which these
particles lose their energy are discussed in Sec. 2.2.1.

The energy distribution between the electron and the positron produced in the pair produc-
tion process according to Eq. 3.3 is not symmetric due to the very high center-of-mass energy.
One of the particles carries away most of the energy of the primary UHE photon [Bha00].
This leading particle can then undergo inverse Compton scattering processes with background
photons:

$$e^{\pm} + \gamma_{\text{background}} \rightarrow e^{\pm} + \gamma_{\text{UHE}}. \tag{3.5}$$

In this process, most of the energy of the electron or positron is transferred to the upscattered
background photon, which can now be considered a UHE photon [Bha00]. Through repeated
cycles of pair production and inverse Compton scattering, an electromagnetic cascade devel-
ops. This cascade stops when the photon energies reach the TeV to GeV range, where the
Universe becomes increasingly transparent to photons, as shown in Fig. 3.1 [Ris07].

In addition, adiabatic energy losses due to the expansion of the Universe have to be taken
into account. The energy loss length for this mechanism can be estimated to be about 4000 Mpc
for the Einstein-de Sitter model of a flat, matter-dominated, Universe, assuming a Hubble con-
stant $H_0 = 75 \, \text{km s}^{-1} \, \text{Mpc}^{-1}$ [Sta10]. As can be seen from Fig. 3.1, adiabatic energy losses are
mainly relevant for photons in the TeV range, as the dominant energy loss mechanism for
UHE photons is due to the interactions with photons from the URB.

3.3. Extensive Air Showers Induced by Photons

Due to the very small incoming flux of UHE photons, direct measurements of these particles are not possible. Similar to charged cosmic rays, only indirect measurements exploiting extensive air showers are feasible. Hence, the differences between air showers initiated by photons and those initiated by protons or nuclei have to be understood in order to distinguish one from another. The development of the air showers can be studied with MC simulations using CORSIKA. In Fig. 3.2, for example, the tracks of the secondary particles from the muonic, electromagnetic and hadronic shower components are shown for three simulated showers initiated by a photon, a proton and an iron nucleus. The energy of the primary particle was in each case 10^{13} eV. The air showers of all three primary particle types develop very strong electromagnetic components. In the case of proton and iron primaries, significant muonic and hadronic components develop as well. These components are mostly absent in air showers initiated by photons, which are almost purely electromagnetic. Although the production of muon pairs is possible in pair production processes similar to Eq. 2.16, such processes are suppressed by a factor $m_e^2/m_\mu^2 = 2.3 \times 10^{-5}$ with respect to the production of electron-positron pairs [Ris07]. Similarly, photonuclear interactions of high-energy photons from the electromagnetic component with nuclei from the air—processes from which secondary hadrons arise—are suppressed by almost two orders of magnitude with respect to pair-production processes in the Coulomb fields of the nuclei [Ris07]. In the following, the differences in the longitudinal and lateral development between extensive air showers initiated by photons and those initiated by hadrons—which have been introduced in Sec. 2.3—are discussed.

3.3.1. Differences in the Longitudinal Development

The average atmospheric depth of the shower maximum, X_{max}, obtained from MC simulations using CORSIKA, for extensive air showers induced by photon, proton and iron primaries as a function of the primary energy is shown in Fig. 3.3. The difference in the average X_{max} between photon and proton primaries is about $60\,\mathrm{g\,cm^{-2}}$ at 10^{16} eV. The difference increases with energy. The difference is even larger in the case of photon and iron primaries. However, it should be noted that for these particles, the choice of the hadronic interaction model used in the MC simulations affects the average X_{max}. Therefore, in Fig. 3.3, three different hadronic interaction models are used for the simulations with protons and iron nuclei as primary particles. The differences between the predictions of the individual models can be as large as $20\,\mathrm{g\,cm^{-2}}$ (see Fig. 3.3). On the other hand, the choice of a specific hadronic interaction model does not influence the results for photon primaries, since the air showers induced by photons are almost purely electromagnetic. The differences in the average X_{max} between different hadronic interaction models are usually less than $5\,\mathrm{g\,cm^{-2}}$ for such showers [Ris07].

The differences in the average X_{max} between air showers induced by photon primaries and those initiated by protons or nuclei are expected, considering the development of a purely electromagnetic air shower in the atmosphere. The multiplicity of the electromagnetic interactions is on average much smaller than the multiplicity of the hadronic interactions (cf. Sec. 2.3.2). Although the smaller multiplicity is partly compensated for by the smaller interaction length of photons and electrons/positrons in air, overall the development of a purely electromagnetic air shower is expected to be delayed with respect to an air shower that is initiated by a proton or a nucleus, resulting in a larger average X_{max} [Set13]. Additional pro-
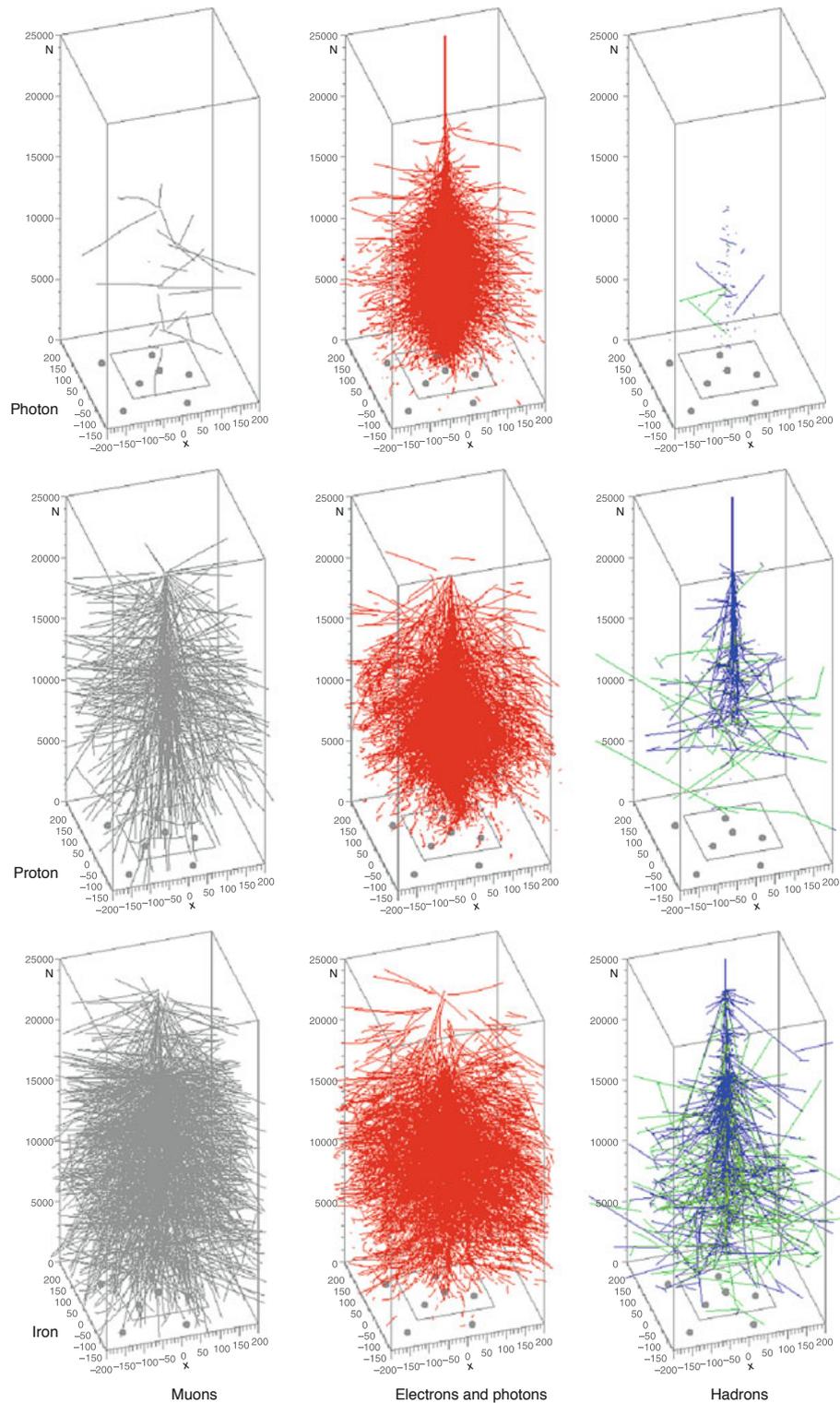
Figure 3.2.: Tracks of the secondary particles from the muonic, electromagnetic and hadronic shower components for three simulated air showers, initiated by a photon, a proton and an iron nucleus, respectively [Eng12]. For the simulations, CORSIKA has been used. The height of each graph corresponds to an altitude of 25 km, while the width corresponds to 400 m.

cesses, which are unique to extensive air showers initiated by UHE photons, further influence the average X_{max}: above some 10^{18} eV, the Landau, Pomeranchuk, Migdal (LPM) effect—i.e. the reduction of the cross sections for pair production and bremsstrahlung at high energies or high matter densities—has to be taken into account, which further delays the shower development, leading to an even larger average X_{max} (see Fig. 3.3). Another effect is the preshower effect, which becomes relevant for primary energies above some 10^{19} eV: a UHE photon may convert into an electron-positron pair in the geomagnetic field above the atmosphere, initiating a particle cascade—the preshower. When the preshower enters the atmosphere, a multitude of electromagnetic air showers is initiated, where the individual primary particles have a lower energy than the initial UHE photon, thus reducing the average X_{max} at these energies (see Fig. 3.3). Both the LPM and the preshower effects are discussed in the following. The use of X_{max} as a observable to distinguish showers induced by photons from those induced by protons in air shower events is discussed in Sec. 6.1.
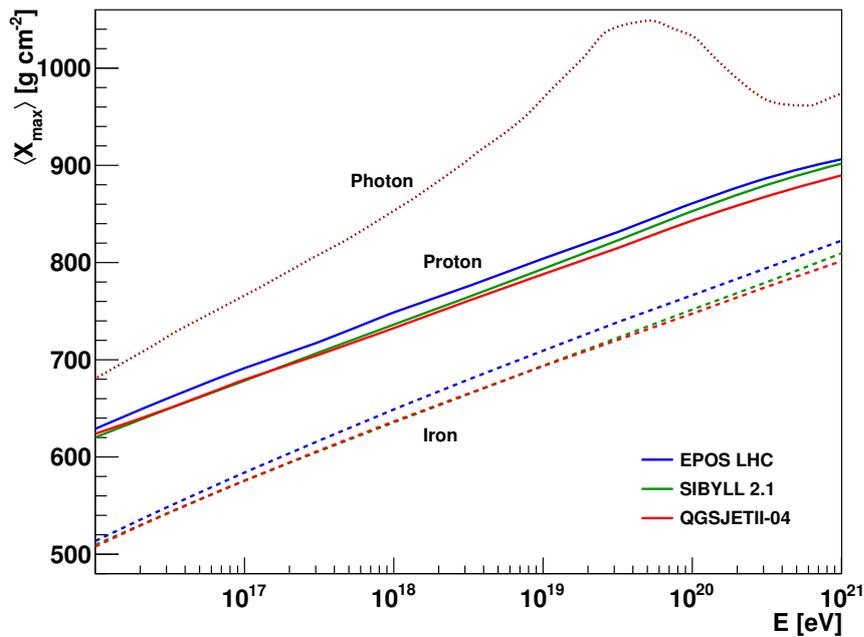
Figure 3.3.: Average atmospheric depth of the shower maximum, $\langle X_{\mathrm{max}} \rangle$, as a function of the primary energy for extensive air showers initiated by photons, protons and iron nuclei as primary particles (adapted from [Set13]). For protons and iron nuclei, three different hadronic interaction models have been used: EPOS LHC [Pie13a], SYBILL 2.1 [Ahn09] and QGSJETII-04 [Ost11]. For photons, the impact of choosing a specific hadronic interaction model is negligible. At the highest energies, additional processes have been taken into account for primary photons, namely the LPM effect and the preshower effect. Due to the preshower effect, $\langle X_{\mathrm{max}} \rangle$ is for energies above about 10^{19} eV not only a function of energy, but also dependent on the location and the incoming direction of the primary photon. For calculating the influence of the preshower effect for this plot, the location of the Pierre Auger Observatory in Malargüe, Argentina, has been used.

LPM Effect

The processes of pair production and bremsstrahlung in matter—which are both crucial to the development of an electromagnetic particle cascade (see Sec. 2.3.2)—are usually described following the formulas derived by Hans Bethe and Walter Heitler in 1934 [Bet34]. However, in 1953, Lev Landau and Isaak Pomeranchuk postulated that these formulas were inapplicable due to interference at high energies or high matter densities [Lan53a, Lan53b]. Later, Arkady Migdal developed a formula to account for this effect [Mig56]. Hence, this effect is commonly known today as the Landau, Pomeranchuk, Migdal (LPM) effect. A review of experimental confirmations of the LPM effect can be found in [Kle99].

In most situations, it is sufficient to assume the pair production and bremsstrahlung processes occur at a single point. However, the Heisenberg uncertainty principle dictates that the processes are spread out over a distance—the formation zone—since the longitudinal momentum transfer between the high-energy particle and the target nucleus is small [Kle99]. During the time the interacting particle traverses the formation zone, multiple scattering on adjacent nuclei in the medium can lead to a reduction of the pair production and bremsstrahlung cross sections due to destructive interference between the single scattering centers. In the case of pair production, the reduced cross section can be approximated by

$$\sigma_{\mathrm{LPM}} = \sigma_{\mathrm{BH}} \sqrt{\frac{E_\gamma E_{\mathrm{LPM}}}{E_e (E_\gamma - E_e)}}, \tag{3.6}$$

where σ_{BH} denotes the cross section calculated from the Bethe-Heitler formulas ($\sigma_{\mathrm{BH}} \approx 0.51\,\mathrm{b}$ in air), and E_γ and E_e are the energies of the incident photon and the electron created in the pair production process [Ris07]. The parameter E_{LPM} can be calculated from

$$E_{\mathrm{LPM}} = \frac{m_e^2 c^3 \alpha X_{\mathrm{rad}}}{4\pi \hbar \rho} \approx 7.7\,\mathrm{TeV\,cm^{-1}}\,\frac{X_{\mathrm{rad}}}{\rho}, \tag{3.7}$$

where X_{rad} and ρ are the radiation length and the density of air, respectively [Ris07]. From Eq. 3.6, it follows that the reduction of the cross section is largest for the creation of an electron-positron pair with the energy equally distributed between the two particles, i.e. if $E_e \approx E_\gamma/2$, while the cross section for a highly asymmetric energy distribution ($E_e/E_\gamma \to 0$ or 1) changes only slightly. Therefore, an asymmetric energy distribution is favored in the pair production process. The cross section for the bremsstrahlung process is suppressed in a similar way [Ris07].

In the case of extensive air showers, the reduction of the cross sections due to the LPM effect leads eventually to a deeper development in the atmosphere and hence a larger X_{max}. Another consequence of the LPM effect is that the fluctuations in the shower development between individual showers with the same energy become larger with increasing energy [Ris07].

Preshower Effect

Protons and nuclei in primary cosmic rays at the highest energies pass the geomagnetic field nearly undisturbed (see Sec. 2.2.1). UHE photons, on the other hand, may convert in the geomagnetic field into an electron-positron pair, which then emits synchrotron radiation, leading to an electromagnetic cascade above the atmosphere, the preshower [Erb66, McB81]. As a
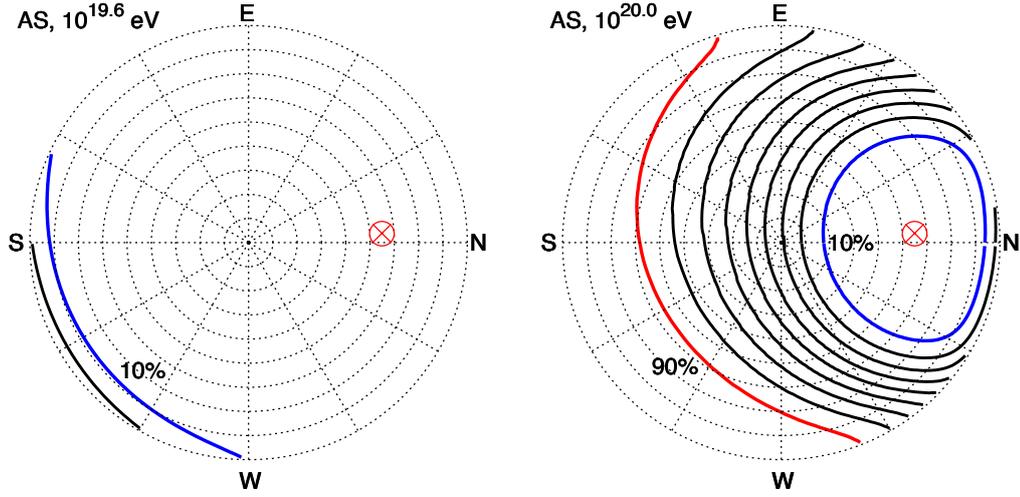
Figure 3.4.: Probabilites P_{conv} for a UHE photon to convert in the geomagnetic field for two different primary energies, calculated for the location of the Pierre Auger Observatory in Malargüe, Argentina [Hom07]. The conversion probabilites are shown as a function of the arrival direction. The contour lines indicate constant conversion probability with a step-size $\Delta P_{\mathrm{conv}} = 10\,\%$. The concentric circles indicate constant zenith angles with a step-size $\Delta \theta = 10°$ with the zenith in the center. The direction of the local magnetic field is illustrated by a red marker.

result, the original, single UHE photon does not reach the top of the atmosphere, but many electrons, positrons and photons, each of which initiates an electromagnetic shower inside the atmosphere. Since the energies of the individual primary particles are smaller than the energy of the initial UHE photon, the individual showers develop higher in the atmosphere, leading to a smaller X_{max}. Since the separation of the preshower particles entering the atmosphere is well below current detector resolutions, both in transverse distance and time, the subsequent showers are observed as a single air shower event [Ris07]. A typical preshower initiated by a photon with an energy of 10^{20} eV starts at an altitude of about 1000 km and enters the atmosphere (at an altitude of about 100 km) with only a few electron-positron pairs with energies around 10^{18} eV and several hundreds of photons [Ris07]. The energies of these photons extend over several decades with a few photons around 10^{19} eV. However, significant deviations from these average values are possible, in particular if the conversion takes place at lower altitudes.

The local differential conversion probability for a photon of energy E depends on the parameter χ, which is given by

$$\chi = \frac{E}{m_e c^2} \frac{B_\perp}{B_c}, \tag{3.8}$$

where $B_\perp$ denotes the component of the local magnetic field perpendicular to the direction of movement of the photon and $B_c = 4.414 \times 10^{13}$ G is a constant [Ris07]. The probability P_{conv} for the photon to convert in the geomagnetic field is calculated by integrating along the trajectory of the photon. Since χ depends on the local magnetic field, P_{conv} strongly depends on the specific trajectory through the geomagnetic field and, thus, on the incoming direction of the primary photon and the location of the experiment. Non-neglible conversion probabilities

($P_{\text{conv}} > 10\%$) can be obtained if the parameter χ exceeds a value of 0.5 along the trajectory [Ris07]. In Fig. 3.4, the conversion probabilites for two primary energies are shown as a function of the incoming direction of the primary photon for the location of the Pierre Auger Observatory in Malargüe, Argentina. At an energy of $10^{19.6}$ eV, the conversion probability is negligible over most of the sky, while at an energy of 10^{20} eV, effectively all photons coming from the southern direction at large zenith angles convert. Hence, the preshower effect plays a role mainly at the highest energies around 10^{20} eV and above.

3.3.2. Differences in the Lateral Distributions

Air showers initiated by photons differ from showers initiated by protons or nuclei not only in the longitudinal development in the atmosphere, but also in the lateral distribution of secondary particles on ground level. Due to the deeper development in the atmosphere and the lack of a significant muonic component, air showers induced by photons exhibit a steeper lateral distribution of secondary particles on ground [Set13]. This can already be inferred from Sec. 2.3.2, Eq. 2.24: for a given energy, the average X_{max} for air showers initiated by photons is larger than for those induced by protons. Hence, the shower age parameter s for a fixed atmospheric depth is smaller, leading to a steeper LDF. The absence of a significant hadronic and muonic component—where the particles can have a large transverse momentum—lead to a smaller spread of the secondary particles on ground as well (cf. Fig. 2.12(b)).

In the analysis that is presented in this work, the differences in the lateral distributions between air showers initiated by photons and those initiated by protons or nuclei are exploited in addition to X_{max} in order to make the best use of the hybrid data from the Pierre Auger Observatory. The observable that is used as a discriminating parameter is discussed in Sec. 6.2.

3.4. Current Upper Limits on the Incoming Photon Flux

So far, photons have been observed up to the 100 TeV range using IACTs [Hol12]. Considering Fig. 3.1, an observation of photons in the PeV range is not expected due to the small energy loss length. Only if their sources were very close, i.e. less than 100 kpc from the Earth, the photons would not cascade down to TeV energies. Measurements from KASCADE-Grande do not yield any evidence for point sources at PeV energies [Ant04b]. The situation changes at higher energies: in the EeV range another window for photon observations opens due to the increasing energy loss length (cf. Fig. 3.1). Hence, extensive searches for UHE photons have been performed by different experiments, see e.g. the review in [Alv13]. However, no unambiguous identification of UHE photons has been reported so far. In the following sections, the current status of the different searches for UHE photons is briefly summarized.

3.4.1. Limits on the Diffuse Photon Flux

The current upper limits on the integral flux of UHE photons, imposed by several experiments, located in both the northern and the southern hemispheres, are shown in Fig. 3.5, compared to the flux predictions by several theoretical models. The upper limits imposed by the Pierre Auger Observatory are the most stringent limits, both in the EeV range, where the limits have been derived using hybrid data [Set11], and above 10^{19} eV, where only data from the surface detectors of the Pierre Auger Observatory has been used [Abr08b]. The current upper limits
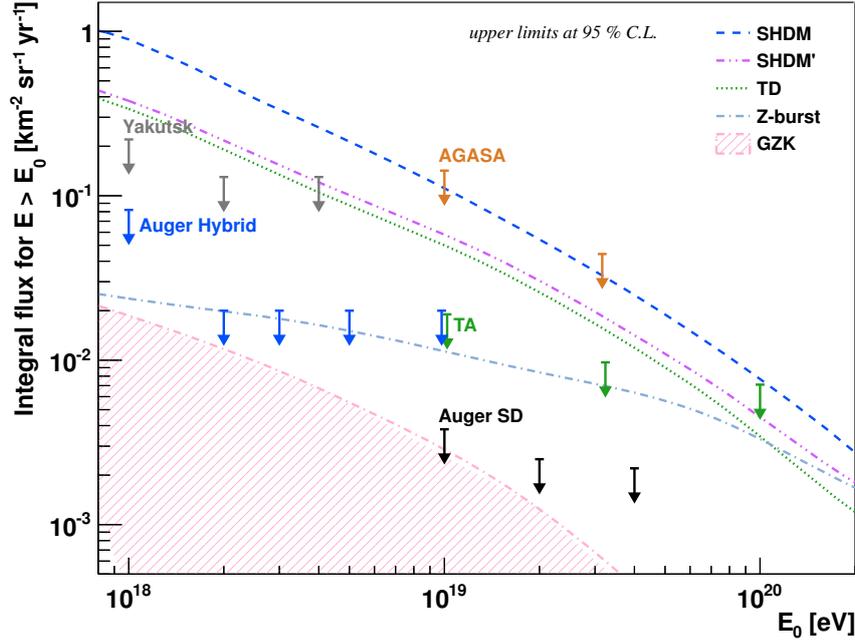
Figure 3.5.: Current upper limits on the integral flux of UHE photons at a confidence level of 95 %, (adapted from [Set13]). Shown are the upper limits determined by the Pierre Auger Observatory (Auger Hybrid [Set11], Auger SD [Abr08b]), TA [Abu13b], the Yakutsk Extensive Air Shower Array [Glu10] and AGASA [Shi02]. The shaded area indicates the flux prediction for UHE photons originating from GZK processes [Gel08]. The lines denote the flux predictions for several top-down models: SHDM, TD, Z-Burst from [Gel08], and SHDM' from [Ell06]. For the upper limits published by the Pierre Auger Observatory, a shift in the overall energy scale between the publication of the limits and the present day has to be taken into account (see Sec. 7.3).

put severe constraints on theoretical models following the top-down approach. The large flux of UHE photons predicted by most of these models is excluded by the current upper limits, so that these models are today generally disfavored with respect to bottom-up models. On the other hand, the predicted flux of UHE photons from GZK processes seems to be within reach within the next years, when larger statistics will be available [Set13]. It should be stressed, however, that the theoretical predictions for the flux of UHE photons from GZK processes strongly depend on specific assumptions about the sources of UHECR (cf. Sec. 3.1). The shaded area indicated in Fig. 3.5 refers to the most optimistic model, in which only protons are accelerated at the sources [Gel08].

Even though no UHE photons have been unambiguously observed yet, already the upper limits on the incoming flux of UHE photons have an impact on our current understanding of fundamental physics. Many theoretical models for quantum gravity predict a breaking of Lorentz invariance, where the magnitude of the effect increases with energy [Gal08]. A consequence of Lorentz invariance violation would be a modification of the dispersion relation for UHE photons. Considering a photon with four-momentum $\left(\omega, \vec{k}\right)$ in the Planck system of units ($c = \hbar = 1$), the dispersion relation at the highest energies can be modified by introduc-

ing Lorentz-violating terms depending on the parameters ξ_n:

$$\omega^2 = k^2 + \xi_n k^2 \left(\frac{k}{M_{\text{Planck}}} \right)^n, \tag{3.9}$$

with $n \geq 1$ and $M_{\text{Planck}} \simeq 10^{19}$ as the Planck mass [Gal08]. Such a modification effectively suppresses pair production processes of UHE photons with lower-energy background photons (Eq. 3.3), enhancing the expected flux of UHE photons, which do not cascade down to lower energies any more. Using current upper limits on the incoming flux of UHE photons and assuming a proton-dominated composition at the highest energies, the upper bounds on the Lorentz-violating parameters ξ_1 and ξ_2 have been improved by seven orders of magnitude, putting tighter constraints on theoretical models involving a breaking of Lorentz invariance [Gal08].

3.4.2. Search for Point Sources of UHE Photons

In addition to the search for a diffuse flux of UHE photons, the Pierre Auger Collaboration performed a search for point sources of UHE photons using hybrid data in the EeV range [Aab14c]. No point source of UHE photons has been detected in this analysis. The sky map of directional upper limits to the flux of UHE photons is shown in Fig. 3.6. The mean value of the flux is $0.035\,\text{km}^{-2}\,\text{yr}^{-1}$, with a maximum of $0.14\,\text{km}^{-2}\,\text{yr}^{-1}$. Under the assumption of an energy spectrum of the UHE photons with a spectral index of 2, these values translate to a mean and a maximum energy flux of $0.06\,\text{eV}\,\text{cm}^{-2}\,\text{s}^{-1}$ and $0.25\,\text{eV}\,\text{cm}^{-2}\,\text{s}^{-1}$, respectively [Aab14c]. Extrapolations from measurements of the energy flux in the TeV range using IACTs predict an energy flux exceeding $1\,\text{eV}\,\text{cm}^{-2}\,\text{s}^{-1}$ for sources such as Centaurus A or the Galactic center region. However, no energy flux that strong has been observed in this analysis [Aab14c]. In the northern hemisphere, the Telescope Array Collaboration has performed a search for steady point-like sources of neutral particles at energies around $10^{18}\,\text{eV}$, which did not yield any evidence for such sources [Abb14].
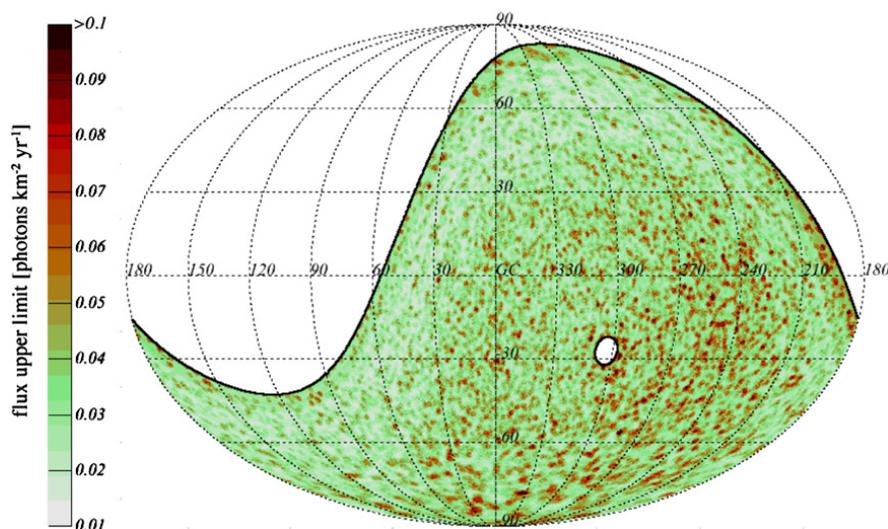
Figure 3.6.: Sky map of directional upper limits to the incoming flux of UHE photons, shown in Galactic coordinates [Aab14c]. The upper limits have been derived at a confidence level of 95 %.

<div align="right">*4*</div>

The Pierre Auger Observatory

Contents

The major challenge in cosmic-ray physics at the highest energies is the rapidly decreasing incoming flux of primary cosmic rays. In the energy range above the ankle, only one particle per km^2 per year reaches the earth (cf. Fig. 2.3). In order to acquire significant statistics in a reasonable amount of time, experiments have to cover very large areas to compensate the low incoming flux. The Pierre Auger Observatory [Pie96] is currently the world's largest cosmic-ray experiment. It is located in Argentina in the province of Mendoza, at the base of the Andes in the Pampa Amarilla, an elevated plain near the town of Malargüe ($35.2°$ S, $69.2°$ W, cf. Fig. 4.1(a)). The average elevation of the detector site is 1400 m above sea level (a.s.l.), which corresponds to an average vertical atmospheric depth of $870\,g\,cm^{-2}$.

According to the orginal design, the Pierre Auger Observatory would have consisted of two independent sites, with one site in the southern and one in northern hemisphere to achieve full sky coverage [Pie96]. While the construction of the southern site near Malargüe has been completed, the planning of the northern site has been ceased due to a lack of funding. In the context of this thesis, the term Pierre Auger Observatory therefore only refers to the southern site.
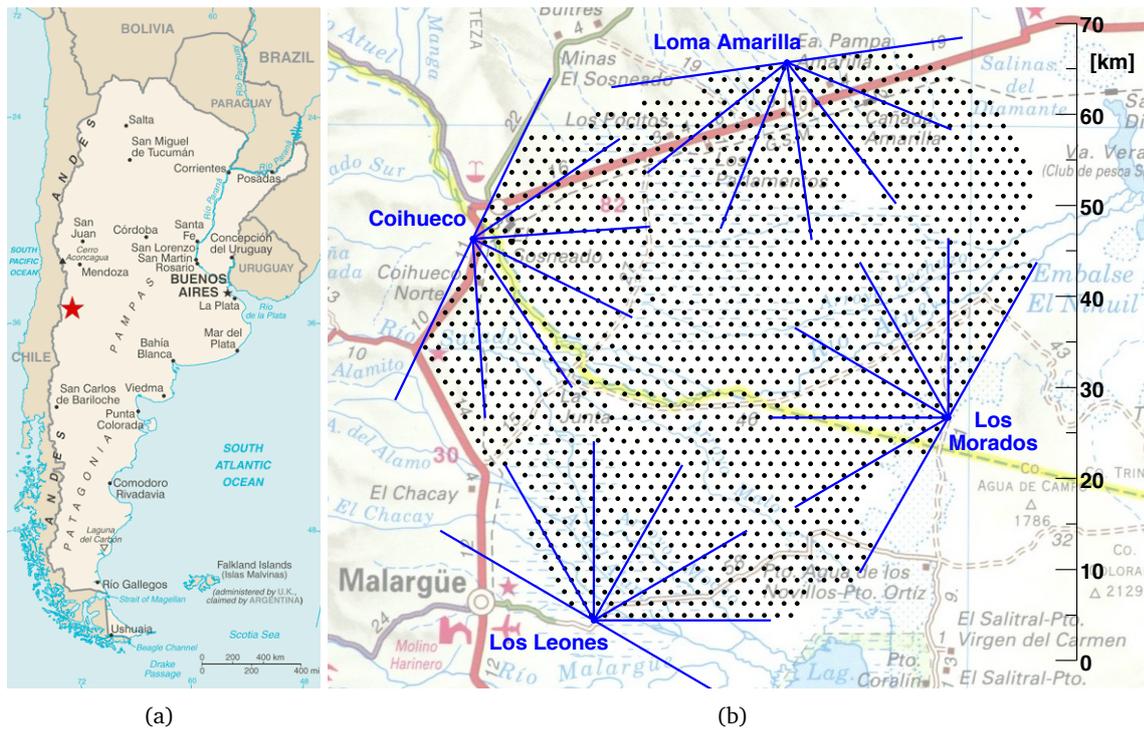
Figure 4.1.: The Pierre Auger Observatory. (a) The location of the Pierre Auger Observatory near Malargüe in Argentina, indicated by a red marker (adapted from [Cen14]). (b) The layout of the detector site of the Pierre Auger Observatory (adapted from [Veb13a]). The more than 1600 detector stations of the SD are marked by black dots. The fields of view of the 24 fluorescence telescopes from the FD are indicated by the blue lines. The telescopes are located in groups of six at the four sites Los Leones, Los Morados, Loma Amarilla and Coihueco at the border of the SD array.

A key feature of the Pierre Auger Observatory is the hybrid concept [Pie96], combining two complementary detection techniques to achieve unprecedented accuracy in the measurements. An array of more than 1600 water Cherenkov detectors forms the Surface Detector (SD), which records the particles from an extensive air shower at ground level. At the same time, the longitudinal development of the air shower is observed using the Fluorescence Detector (FD), which consists of 24 fluorescence telescopes located at the border of the SD array. The layout of the detector site of the Pierre Auger Observatory is shown in Fig. 4.1(b). The detector stations of the SD are arranged in a hexagonal pattern with a distance of 1.5 km between the single detector stations, thus covering an area of more than 3000 km^2 [All08]. The fluorescence telescopes of the FD are located in groups of six at four sites overlooking the SD array [Abr10a].

In the following sections, the detector systems of the Pierre Auger Observatory are discussed in detail: the SD in Sec. 4.1 and the FD in Sec. 4.2. Since the analysis presented in this thesis uses hybrid data, which primarily uses information from the FD, the trigger system of the FD and the detector calibration will be described in more detail. In addition, the monitoring systems for the atmospheric conditions at the experimental site—which are crucial for

the reconstruction of air shower events from data recorded by the FD—are described in this section. The various enhancements of the Pierre Auger Observatory that are currently under construction or under development are summarized in Sec. 4.3. The plans to upgrade the detector systems of the Pierre Auger Observatory in the period from 2015 to 2023 and the physics motivation for these upgrades are briefly outlined in Sec. 4.4. In the last section of this chapter (Sec. 4.5), the reconstruction of air shower events from the data recorded by the SD and the FD is described.

4.1. The Surface Detector

The SD is a ground-based array consisting of more than 1600 completely autonomous water Cherenkov detector stations. The single detector stations can take data with a duty cycle of nearly 100 % [All08]. Hence, the data from the SD constitutes the largest fraction of the data recorded at the Pierre Auger Observatory. The energy threshold for full trigger efficiency for air shower events is mainly determined by the distance between the single detector stations. With a spacing of 1.5 km, the energy threshold for full trigger efficiency is 3×10^{18} eV for zenith angles below 60°, independent of the type of the primary particle initiating the extensive air shower [Abr10d].

A single detector station of the SD is depicted in Fig. 4.2. Each SD station consists of a cylindrical tank with a diameter of 3.6 m [All08]. Each tank is filled with 12,000 l of highly purified water, corresponding to a water depth of 1.2 m inside the tank. The water in the tank is contained within a sealed bag, which has an inner liner with high diffusive reflectivity. The water serves as the detection medium for the air shower particles exploiting the Cherenkov light produced by these particles when they enter the tank. This light is collected by three nine-inch-diameter Photomultiplier Tubes (PMTs), which are optically coupled to the water through three polyethylene windows and placed symmetrically at a distance of 1.2 m from the center at the top of the tank [All08]. The electronic components of the SD station, e.g. the read-out electronics or the PMTs, are powered by a battery system which is charged via solar panels. The communication with the Central Data Acquisition System (CDAS) of the Pierre Auger Observatory is accomplished through a radio link. A Global Positioning System (GPS) system provides precise timing for each station with a timing resolution of 8 ns [Suo09].

When a relativistic charged particle from an extensive air shower enters the tank, it may emit Cherenkov radiation [Che34]. The threshold energy for, e.g., muons, which are the most abundant secondary particles on ground (see Sec. 2.3.2) for this process in water (refractive index $n = 1.33$ [Ber12]) is 54.6 MeV. The Cherenkov radiation is emitted mainly in the forward direction of the charged particle traversing the SD station. Since most particles enter the SD station from above, the Cherenkov radiation needs to be reflected at the inner liner of the tank. The liner has been chosen such that the reflectivity is highly uniform in the Ultraviolet (UV) wavelength regions, where most of the Cherenkov radiation is emitted. The reflected Cherenkov light is then collected by the PMTs, where the signals are amplified with a nominal gain of 2×10^5 [Suo09]. The amplified signals from the PMTs are then filtered and digitized with a frequency of 40 MHz—corresponding to a timing resolution of 25 ns—using Flash Analog-to-Digital Converters (FADCs) [Abr10d]. The digitized data are subsequently stored in ring buffer memories and processed by a Field-Programmable Gate Array (FPGA), through which different local trigger conditions, i.e. on the level of a single
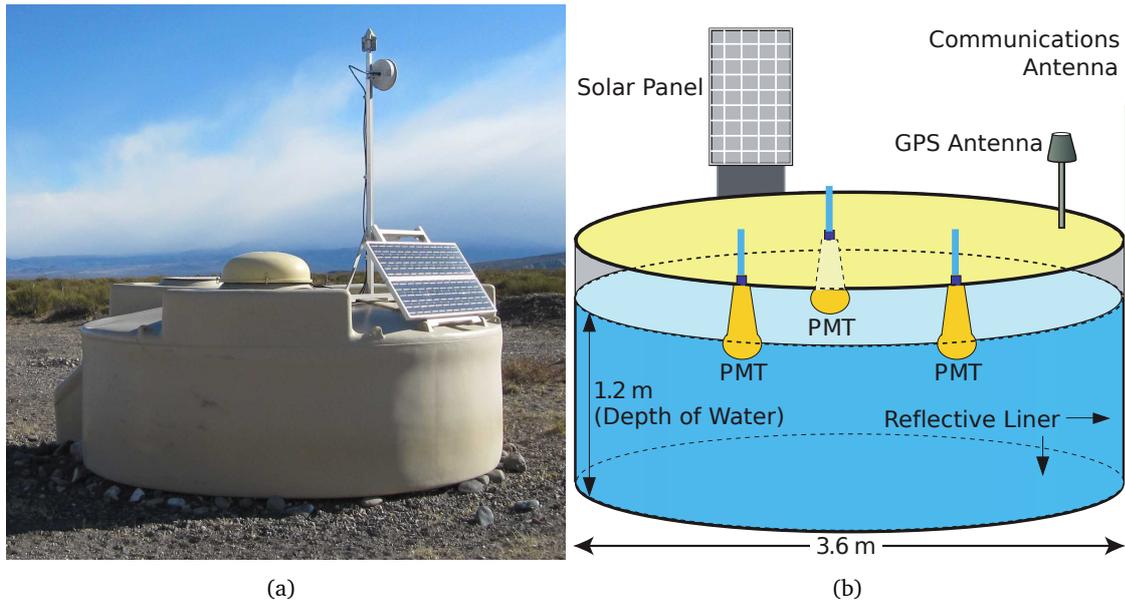
Figure 4.2.: A water Cherenkov detector station from the SD of the Pierre Auger Observatory. (a) Photograph of an SD station in the Pampa Amarilla. (b) Schematic view of an SD station showing the different components (adapted from [Kei04]). Not included are the read-out electronics and the battery system.

detector station, are implemented. In the case of a local trigger, the FADC traces of the single PMTs are stored locally, together with the time stamp of the local trigger. The time stamp and the trigger condition that was fulfilled are then sent to the CDAS, where the third-level trigger (T3) is formed through a combination of the local triggers received from the single detector stations [Abr10d]. If the spatial and temporal distribution of the single triggers fits to a possible air shower event, the T3 trigger is formed and sent back to the locally triggered detector stations. Upon receiving the T3 trigger signal, the locally stored FADC traces are sent via the radio link to the CDAS, where they are permanently stored for further analysis. Since the data recorded by the single stations are only sent to the CDAS if a T3 trigger is received, the load on the wireless communications system, which is the limiting factor, can be reduced to a manageable amount. A full description of the trigger system of the SD can be found in [Abr10d].

The signals recorded in a single detector station are eventually expressed in units of Vertical Equivalent Muon (VEM), which corresponds to the signal measured by a station for a single muon that traverses the station vertically and centrally [Ber06b]. However, the recorded FADC traces also depend on the properties of the detector station, i.e., for example, the exact reflectivity of the inner liner, the gain of the PMTs and the optical coupling of the PMTs to the water. Hence, the signal measured for a single muon varies from station to station. To remove the dependence of the measured signal on the properties of the single detector station, the detectors are calibrated using low-energy background muons. The calibration quantities are constantly monitored and sent to the CDAS, so that also variations of these quantities with time due to changes e.g. in the outside temperature can be compensated. A more detailed account of the different methods used to calibrate the detector stations of the SD can be found in [Ber06b].

4.2. The Fluorescence Detector

The FD is based on the air-fluorescence technique for the detection of extensive air showers. When an extensive air shower passes through the atmosphere, the charged particles contained in the shower may transfer energy to the nitrogen molecules in the air, exciting the molecules. The excited nitrogen molecules can return to their ground states through the emission of fluorescence light with wavelengths from 300 nm to 430 nm, i.e. in the UV regime [Abr10a]. The number of photons emitted through this process per unit of energy deposited in the atmosphere—the fluorescence yield—also depends on the properties of the atmosphere, i.e. pressure, temperature and humidity. The absolute fluorescence yield, in dry air at standard atmospheric conditions (1013 hPa, 293 K), for the 337 nm fluorescence band, which is the dominant wavelength band, is 5.61 ± 0.06 (stat.) ± 0.22 (syst.) photons per MeV [Ave13]. From relative measurements, the fluorescence yield for other wavelength bands and the dependence of the fluorescence yield on the atmospheric conditions can be inferred [Ave07b, Ave08]. The fluorescence light can be observed using optical telescopes such as the FD of the Pierre Auger Observatory. Since the amount of fluorescence light emitted is directly proportional to the energy deposited in the atmosphere and thus the number of charged particles in the extensive air shower, the observation of the fluorescence light provides a measure for the longitudinal shower development in the atmosphere. If the shower fully develops in the atmosphere, the energy of the primary particle initiating the extensive air shower can be inferred from the longitudinal shower profile after a correction for the energy that is transferred to the neutrino component and high energy muons, which do not deposit their energy in the atmosphere and are thus invisible to a fluorescence detector. This correction is in principle dependent on the type of the primary particle and has to be derived from MC simulations (cf. Sec. 4.5.1).

A disadvantage of the air-fluorescence technique is that measurements of the fluorescence light are only possible in clear nights, in which the background light from the moon is dim enough to not overshadow the faint fluorescence light. The FD of the Pierre Auger Observatory operates only in nights where the illuminated fraction of the moon is below 60 % [Abr10a]. An observation period lasts on average 16 days per month, from the waning gibbous moon over the new moon to the waxing gibbous moon. In these nights, the FD is operated from the beginning of the astronomical dusk until the beginning of the astronomical dawn, which amounts to an average observation time of about 10 h per night, ranging from about 14 h in June to 5 h in December [Abr10a].

The FD consists of 24 fluorescence telescopes which are located in groups of six at the four sites Los Leones, Los Morados, Loma Amarilla and Coihueco at the border of the SD array, overlooking the array (cf. Fig. 4.1(b)). As an example, the FD site Los Leones is shown in Fig. 4.3(a). Each FD telescope has a field of view of $30° \times 30°$ in azimuth and elevation, with the center of the field of view at an elevation of 16° from the horizon [Abr10a]. In total, each FD site thus covers 180° in azimuth and 30° in elevation. With this arrangement, a trigger efficiency of 100 % is achieved for primary energies above 10^{19} eV [Abr10a].

The 24 FD telescopes are based on the Schmidt telescope [Sch32] to avoid coma aberrations. The design of an FD telescope is shown in Fig. 4.3(b). The fluorescence light enters the telescope through an aperture system with an optical filter. The filter absorbs visible light to reduce the background noise while transmitting photons in the wavelength range from about

(a)   (b)
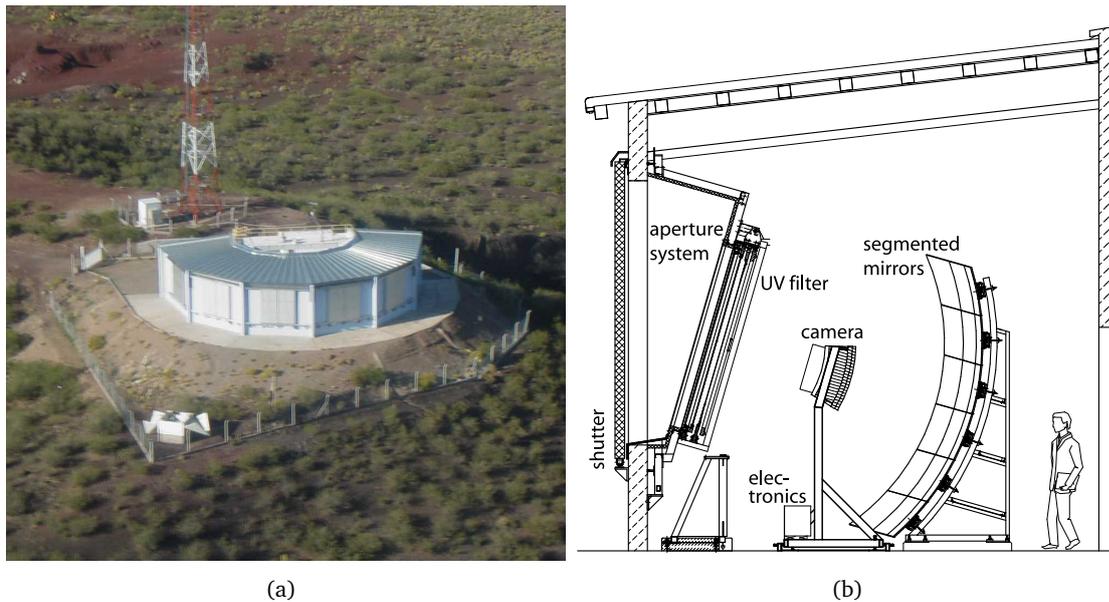
Figure 4.3.: The FD of the Pierre Auger Observatory. (a) Aerial view of the FD site Los
Leones with the shutters closed during the day [Pie06]. (b) Schematic view of an FD
telescope [Abr10a]. The main components of the FD telescope are indicated.

290 nm to about 410 nm. This wavelength range contains all of the relevant wavelength bands
from the fluorescence spectrum of molecular nitrogen [Abr10a]. The diameter of the opening
of the aperture system is 2.2 m, which includes a ring of corrector lenses around the aperture.
The corrector lenses are used to reduce aberrations in the optical system and increasing the
effective aperture area by a factor of two compared to an aperture system without any correct-
ing element. The fluorescence light is collected by a segmented mirror with an area of $13\,m^2$
and a spherical radius of 3.4 m [Abr10a]. In the focal plane of the mirror, a camera consisting
of 440 hexagonal pixels, arranged in a matrix with 22 rows and 20 columns, is mounted.
Each pixel covers a field of view of $1.5° \times 1.5°$. The camera is made from 440 hexagonal PMTs
with additional light collectors between the single PMTs to maximize the light collection and
to provide a sharp transition between the single pixels. The telescopes are housed in clean
climate-controlled buildings to avoid changes in the optical systems due to temperature vari-
ations [Abr10a]. In addition, each telescope has a shutter system that can be closed during
data taking in order to, e.g., prevent damages to the PMTs if the background light gets too
bright.

4.2.1. Data Acquisition and Trigger Systems

The data acquisition and trigger systems of the FD are organized in a hierarchical way, reflect-
ing the physical layout of an FD site [Abr10a]. The 440 PMTs of each of the six fluorescence
telescopes at an FD site are read out by a front-end electronics system consisting of 20 Analog
Boards (ABs) connected to 20 First-Level Trigger (FLT) boards. The signals from the 20 FLT
boards are read out by a single Second-Level Trigger (SLT) board and subsequently sent to a
Personal Computer (PC), where the Third-Level Trigger (TLT) is generated. The data from the
six fluorescence telescopes are then merged and stored for offline analysis.

Each AB receives the signals from a 22-pixel column of the camera, i.e. the signals from 22 PMTs. The main purpose of the ABs is to filter the analog signals coming from the PMTs and to adapt the dynamic range of the signals for the digitization, which is performed on the FLT board [Abr10a]. The PMT signals are processed by an anti-aliasing filter, to match the digitization rate of 10 MHz, and a fourth-order Bessel filter, which has been chosen to minimize the distortion of the signal shapes. After the filtering, the signals are processed by the dynamic range adapter. In the energy regions of interest to the Pierre Auger Observatory, the PMTs of the FD typically record signals in the range from 3 to 10^5 photoelectrons per 100 ns [Abr10a]. This corresponds to a dynamic range of 15 bits. Using the dynamic range compression technique, the dynamic range that has to be covered by the Analog-to-Digital Converters (ADCs) which digitize the signals can be reduced to 12 bits. The details on this technique can be found in [Sch06].

Distinguishing the fluorescence signals of varying intensity from a sizeable and constantly changing background is the main difficulty in performing fluorescence measurements. At the Pierre Auger Observatory, a three-level trigger system is employed. The FLT is a simple threshold trigger. After the signals from an AB have been digitized on the corresponding FLT board with a frequency of 10 MHz using 12-bit ADCs, they are stored in a ring buffer which covers a time period of 100 μs. Then the signals are processed by an FPGA, which generates the FLT [Abr10a]. The threshold of this trigger is continuously adjusted so that the trigger rate is constant at 100 Hz per pixel under varying background conditions [Abr10a].

The trigger signals from the 20 FLT boards are then read out by the SLT board. An air shower event will be observed in the FD telescope as a straight line of triggered pixels, representing the shower axis. The SLT logic, which is implemented in an FPGA, searches for patterns of triggered pixels that match straight line segments and are at least 5 pixels in length [Abr10a]. The five fundamental patterns are shown in Fig. 4.4. Rotations and mirror images of the fundamental patterns are also recognized as straight line segments. However, it is possible that the shower track does not pass through the center of all pixels and therefore some PMTs along the track do not record enough light to pass the FLT. In addition it is possible that a PMT is defective and therefore does not record any signal at all. Hence, the SLT logic requires only four triggering pixels out of the five pixels that form the fundamental patterns [Abr10a]. For each event that passes the SLT, a timestamp is generated by the SLT board. The timing information is, as for the SD, provided by a GPS unit [Abr10a].

After an event has been processed by the FLT and SLT boards and stored in the respective memories, it can be read out by a PC associated to the FD telecope. The TLT is implemented
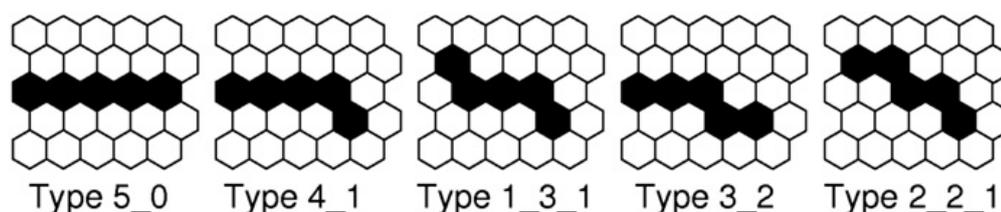
Figure 4.4.: Fundamental patterns that are recognized by the SLT logic as straight line segments [Abr10a].

as a software algorithm running on this PC. The main purpose of the TLT is to filter out noise events due to distant lightning. Since distant lightning can cause large parts of the camera to trigger in bursts of several tens of events per second, it can cause a significant dead time of the detector due to a congestion of the read-out electronics [Abr10a]. The TLT is designed to filter out such lightning events without reading out the full ADC traces, taking into account the time development of the multiplicity of triggered pixels. With this algorithm, approximately 99 % of all lightning events are rejected, effectively removing noise events with more than 25 triggered pixels [Abr10a]. In a second step in the algorithm, noise events with a smaller number of triggered pixels are removed. Such noise events can be caused by, e.g. direct impacts of single muons on the camera and random pixel triggers. In this step, the spatial arrangement of the pixels as well as the peak signal times are used. Hence, the ADC traces have to be read from the memory on the FLT board. In total, about 94 % of all background events are rejected by the TLT, while the fraction of true air shower events rejected by this algorithm is below 0.7 % [Abr10a].

The events that pass the TLT are eventually sent to another PC, where an event builder algorithm merges coincident data from adjacent telescopes to a single event. The events are then stored for offline analysis. In addition, for each TLT, a hybrid trigger is generated and reported to the CDAS, acting as an external trigger to the SD [Abr10a]. The purpose of this external trigger is to record hybrid events in the energy range below 3×10^{18} eV, where the SD array is not fully efficient (cf. Sec. 4.1). In most cases no SD trigger on array level would be generated, since not more than one or two SD stations are triggered. For the hybrid trigger, a preliminary shower direction and ground impact time is calculated with a simple online reconstruction. This information is sent to the CDAS, which requests the data from the SD stations that were triggered close to the calculated time. For each hybrid trigger, only the stations closest to the FD site that generated the external trigger—covering approximately one quarter of the full SD array—are considered [Abr10a]. The data from the FD and the SD are eventually merged offline into hybrid events.

4.2.2. Calibration of the FD

To reconstruct an air shower from the FD measurements, the response of an FD telescope in terms of ADC counts per pixel has to be known for a given flux of photons at the aperture of the telescope. The response of a telescope is influenced by many factors, e.g. the transmittance of the optical filter, the reflectivity of the mirror or the quantum efficiency of the PMTs. At the Pierre Auger Observatory, an end-to-end calibration is employed, where all factors are taken into account simultaneously. Two different calibration systems are used: an absolute calibration using a drum-shaped light source that can be mounted on the outside of the aperture of the FD telescope, and a relative calibration system, where different light sources inside and outside the FD telescope are used.

For the absolute calibration, a portable, drum-shaped light source with a diameter of 2.5 m and a depth of 1.4 m was constructed [Bra04]. The light source can be mounted on the exterior of the telescope aperture (Fig. 4.5(a)) and provides a pulsed flux of photons with known intensity and uniformity to all pixels of the camera. Both the intensity and the uniformity of the flux of photons have been determined in laboratory measurements using calibrated photodiodes [Bra04]. Inside the drum, a pulsed UV Light-Emitting Diode (LED) with a wavelength
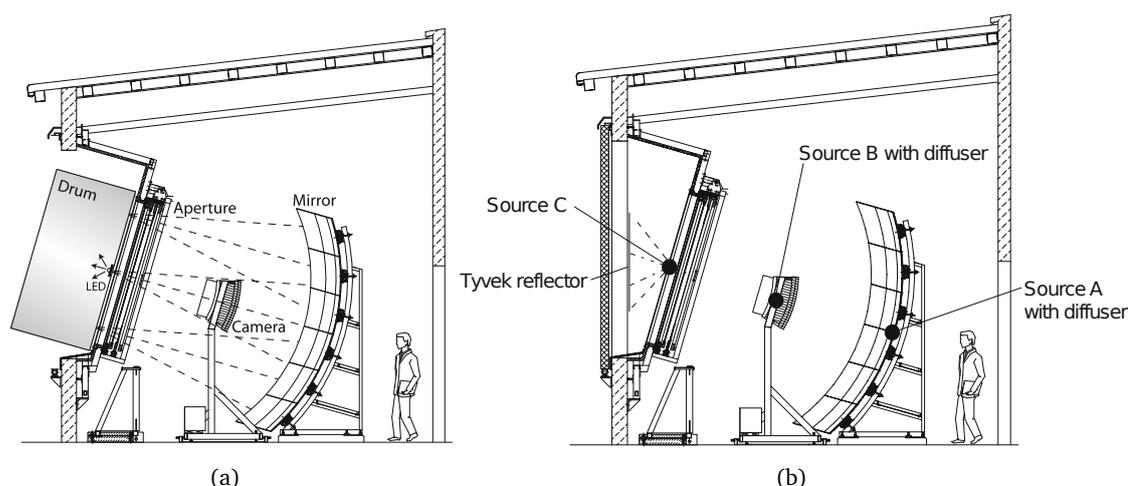
Figure 4.5.: Calibration of the FD telescopes: schematic view of an FD telescope (cf. Fig. 4.3(b)) with the different calibration systems included [Abr10a]. (a) Absolute calibration using a drum-shaped light source mounted on the outside of the aperture. (b) Relative calibration using three different light sources at different positions inside and outside the telescope.

of $(375 \pm 12)\,\mathrm{nm}$ is mounted. To provide the uniform intensity, the interior of the drum has been lined with the same diffusively reflecting material that is also used in the SD stations. The front of the drum facing the telescope aperture is made from a diffusively transmitting sheet of plastic. From the ratio of the measured PMT signals for each pixel to the flux of photons emitted from the light source, the absolute calibration of the FD telescope can be determined. The response of the FD telescopes to other wavelengths is inferred from relative measurements, where the LED is replaced with a xenon flasher and different wavelength filters [Rov09]. The relative FD response together with the laboratory measurements can be used to determine the dependence of the detector response as a function of the wavelength. Finally, the calibration procedure can be cross checked using remote laser shots [Abr10a]. A nitrogen laser with a wavelength of $337\,\mathrm{nm}$ is used to shoot vertical laser pulses in the atmosphere. Since the distance of the laser to the FD telescope is known—chosen such that the scattering due to aerosols can be neglected—the fraction of photons that are scattered in the direction of the aperture can be calculated. Hence, the laser shots provide an independent and redundant end-to-end calibration of the FD telescopes.

The absolute calibration described before is perfomed about once per year for each of the 24 FD telescopes [Abr10a]. In one night, two telescopes can be calibrated. Hence, the absolute calibration procedure for the full FD takes twelve nights. To monitor variations of the detector response in the time period between two successive absolute calibration measurements and to check the overall stability of the FD, relative calibration measurements are automatically performed before and after each night of data taking [Abr10a]. For each camera, three different positions (A, B, and C; cf. Fig. 4.5(b)) are illuminated. With the corresponding measurements of the relative calibration, different parts of the telescope can be monitored. The light source for position A is a pulsed LED with a wavelength of $470\,\mathrm{nm}$. The light from this LED is guided via optical fibers—terminating in a diffusor—to the center of the mirror of each telescope,

from where it directly illuminates the camera [Abr10a]. Hence, the stability of the performance of the PMTs can be monitored. The light sources for the positions B and C are xenon flashers. The optical fiber guiding light from one of the xenon flashers is split near the camera and terminates in diffusors at the edges of the camera (position B), with the light directed to the mirror. The optical fibers from the other xenon flasher are also split, and they terminate in diffusors outside the aperture system of the telescope (position C). From there, the light is reflected on diffusively reflecting sheets on the inside of the shutters into the aperture of the telescope [Abr10a]. Hence, this calibration measurement provides a method to monitor the performance and the stability of the full telescope.

4.2.3. Atmospheric Monitoring

For the determination of the longitudinal shower profile, the number of fluorescence photons produced by the extensive air shower has to be inferred from the number of photons observed at the aperture of the FD telescope. The fluorescence yield depends on the atmospheric state variables at the time and the location of the shower passing through the atmosphere. In addition, the propagation of the fluorescence light from the shower to the FD telescope is influenced by Rayleigh scattering processes with molecules from the air and Mie scattering processes with aerosols in the air, e.g. dust particles or air pollutants. Hence, an extensive program to monitor the atmospheric conditions at the site of the Pierre Auger Observatory has been developed. In Fig. 4.6, the different experimental setups used for the atmospheric monitoring at the detector site are shown. In this section, these setups are briefly discussed. A more detailed description of the atmospheric monitoring program at the Pierre Auger Observatory and the influence of the atmospheric conditions on fluorescence measurements can be found in [Abr10c]. In addition to the regular atmospheric monitoring program, where the atmospheric conditions are determined in fixed time intervals, additional targeted measurements of the atmospheric conditions shortly after the detection of air shower events of special interest—e.g. where the primary energy is very high—have been implemented to improve the resolution of the atmospheric measurements for such events and thus the reliability of the reconstruction of the event [Abr12b].

Monitoring of the Atmospheric State Variables

The atmospheric state variables, i.e. temperature, pressure, and relative humidity, at ground level are monitored using five weather stations. Four of the weather stations are located at the four FD sites and one is situated near the center of the SD array. The weather stations record the atmospheric state variables every five minutes with an accuracy of $0.2 - 0.5\,^\circ\mathrm{C}$ in temperature, $0.2 - 0.5\,\mathrm{hPa}$ in pressure and $2\,\%$ in relative humidity [Abr10c]. The atmospheric profiles, i.e. the dependence of the state variables on the height above the ground, have been measured using balloon-borne radiosondes. The measurement accuracy of the radiosondes is $0.2\,^\circ\mathrm{C}$ in temperature, $0.5 - 1.0\,\mathrm{hPa}$ in pressure and $5\,\%$ in relative humidity [Abr10c]. From 261 radiosonde measurements between August 2002 and December 2008, monthly average models for the atmospheric profiles at the site of the Pierre Auger Observatory have been created [Abr10c]. However, the use of monthly models instead of more frequent balloon flights introduces an additional uncertainty—representing the typical range of conditions observed during the course of one month—on the reconstruction of air shower events from FD measurements. On the other hand, balloon flights are too difficult and expensive to carry out on a daily basis. Hence, the use of the Global Data Assimilation System (GDAS), a global
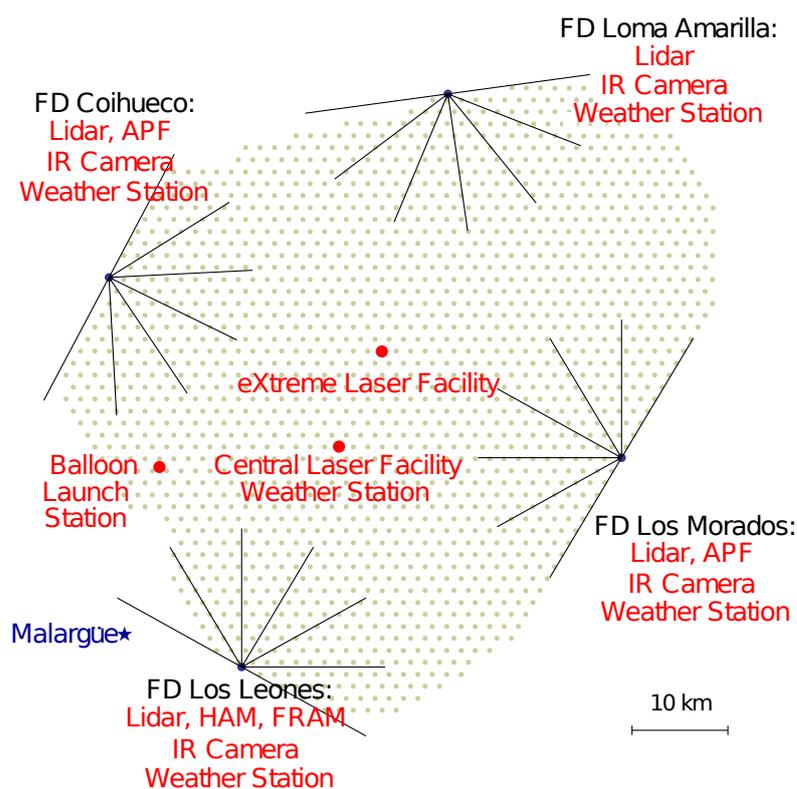
Figure 4.6.: The layout of the detector site of the Pierre Auger Observatory including the different experimental setups to monitor the atmospheric conditions [Abr10c]. The single experimental setups are described in the text.

atmospheric model based on both numerical calculations and meteorological observations, has been evaluated [Abr12a]. The GDAS provides 3-hourly datasets comprising the atmospheric state variables and the atmospheric profiles up to an altitude of about 26 km on a global latitude-longitude grid with a spacing of $1°$. The grid point chosen to represent the detector site of the Pierre Auger Observatory is located at the north-eastern edge of the SD array [Abr12a]. The quality of the GDAS data and the applicability of the data to the reconstruction of extensive air showers measured at the Pierre Auger Observatory has been verified through comparisons with local measurements using the ground-based weather stations and the balloon-borne radiosonde measurements. Thus, the data provided by GDAS are today the basis of the monitoring of the atmospheric state variables [Abr12a].

Aerosol Monitoring

Whereas the influence on the observed fluorescence light due to Rayleigh scattering processes with molecules from the air can be calculated analytically from the atmospheric state variables, no analytic description of the scattering processes with aerosols exists. Hence, the relevant parameters, i.e. the vertical optical depth of aerosols in the air—which is directly related to the attenuation of the fluorescence light due to scattering processes with aerosols—the aerosol phase function—which describes the angular dependence of the scattering process— and the wavelength dependence of the attenuation have to be measured. Several instruments

are used to monitor these quantities.

The vertical optical depth of aerosols is monitored using the Central Laser Facility (CLF) and the eXtreme Laser Facility (XLF), both located near the center of the array (cf. Fig. 4.6), as well as four Light Detection and Ranging (Lidar) systems, located at the single FD sites. The CLF and the XLF provide calibrated laser pulses with a wavelength of 355 nm, which is in the center of the fluorescence spectrum of nitrogen [Fic06]. The nominal energy of 7 mJ per pulse corresponds roughly to the amount of fluorescence light emitted by an extensive air shower initiated by a particle with an energy of 10^{20} eV. During every night of operation of the FD, the lasers fire a set of 50 vertical shots every 15 minutes [Fic06]. The scattered light from the laser shots is then recorded by the FD telescopes. From a comparison with laser shots in clear reference nights, where the influence of aerosols can be neglected, the vertical optical depth of aerosols as a function of the height above ground can be determined on an hourly basis for each FD site, reflecting the fact that the aerosol component in the atmosphere is highly variable with both time and location [Abr10c]. In addition to the CLF and the XLF, four Lidar systems are operated to record the vertical optical depth of aerosols [Ben07a]. Each Lidar consists of a pulsed laser with a wavelength of 351 nm, three parabolic mirrors and three PMTs. The whole system is mounted on a steerable frame. During the nights of operation of the FD, the Lidar systems sweep the sky outside the field of view of the FD telescopes in set, hourly patterns, pulsing the lasers with a frequency of 333 Hz and recording the backscattered light [Ben07a]. Under the assumption of a horizontally uniform distribution of aerosols in the vicinity of each Lidar station, the vertical optical depth of aerosols as a function of the height above ground level can be determined from the differences in the Lidar measurements at different angles [Ben07a]. In general, the measurements of the Lidar system and the measurements using the CLF and the XLF are in good agreement within their respective uncertainties [Abr10c].

The angular dependence of the scattering processes with aerosols can be empirically described by a phase function. At the Pierre Auger Observatory, this phase function is measured by two Aerosol Phase Function Monitor (APF) systems, located at a distance of about 1 km from the FD sites Coihueco and Los Morados [Abr10c]. Each APF consists of a collimated xenon flash lamp with filters to produce shots with wavelengths of 350 nm and 390 nm. During the nights of operation of the FD, an hourly sequence of shots is fired horizontally across the fields of view of the FD telescopes at the respective FD site. These shots are recorded by the FD telescopes and from these measurements, the phase function can be determined [Abr10c].

Finally, the wavelength dependence of the scattering on aerosols is determined using the Horizontal Attenuation Monitor (HAM) [Abr10c]. The HAM uses a high intensity discharge lamp at the FD site Coihueco to provide an intense broad band light source for a Charge-Coupled Device (CCD) camera located at the FD site Los Leones, at a distance of about 45 km. Hence, the total atmospheric attenuation across the detector site can be measured. The wavelength dependence is determined using a filter wheel at the CCD camera to record the source image at five wavelengths between 350 and 550 nm. From the different intensities observed using the different filters, the wavelength dependence of the attenuation due to aerosols can be inferred [Abr10c]. The measurements of the HAM are complemented by the (F/Ph)otometric Robotic Atmospheric Monitor (FRAM), which also primarily determines the wavelength dependence of the attenuation due to aerosols [Ben07b]. FRAM is located

close to the FD site Los Leones. It is a fully automated optical telescope with a diameter of 20 cm, equipped with a CCD camera and a photometer. FRAM automatically observes a set of selected stars and the HAM light source in hourly observation cycles [Ben07b].

Cloud Monitoring

Another important part of the atmospheric monitoring program at the Pierre Auger Observatory is the cloud monitoring. Clouds can distort the observed longitudinal profile of an air shower due to attenuation of the fluorescence light in the cloud or due to scattering of the Cherenkov light also emitted by high-energy particles within the shower into the direction of the FD telescopes—and thus increasing the amount of light observed by the telescope [Abr10c]. However, it is difficult to correct the observed longitudinal profile for the effects induced by clouds. Time periods where parts of the fields of view of the FD telescopes are obscured by clouds are usually removed from the dataset. It is therefore necessary to determine the cloud coverage above the detector site as accurately as possible, so that the events in the selected dataset contain reliable measurements of the longitudinal shower profile. At the Pierre Auger Observatory, the cloud coverage is monitored using different systems. At each FD site, an IR camera is installed, which observes the sky above the detector site during each night of operation of the FD [Chi13]. These cameras record light in the µm wavelength band, which is suitable for distinguishing the warm clouds from the cool background sky. Each camera captures every five minutes a sequence of five images covering the fields of view of the FD telescopes of the respective FD site [Chi13]. In addition, a full-sky sequence of images, covering the entire sky above the FD site, is recorded every fifteen minutes. From the image sequences recorded by the IR camera, the cloud coverage for the direction of each pixel of each FD telescope can be determined [Chi13]. In addition to the IR cameras, the use of satellite data to determine the cloud coverage above the detector site is being evaluated [Chi13]. With both methods, the development of the cloud coverage can be followed throughout the night. However, these methods only provide two-dimensional maps of the cloud coverage for a given direction in the sky and cannot determine the height of the clouds above the ground. This information is provided by the Lidar systems, which can also detect clouds during their hourly sweeps of the sky [Ben07a]. A cloud layer manifests itself in the Lidar measurements as a strong region of backscattered light. From the measurements at different zenith angles and from the arrival times of the backscattered photons at the detector, the height of the clouds above ground level can be inferred. In addition, the Lidar systems also provide the cloud coverage for each pixel on an hourly basis [Ben07a].

4.3. Enhancements of the Observatory

The Pierre Auger Observatory has originally been designed to measure cosmic rays at the highest energies, i.e. above 10^{18} eV. However, also the energy region below 10^{18} eV is of high interest in cosmic-ray physics, because it is generally assumed that in this energy region, the transition from a dominant galactic component of cosmic rays to a dominant extragalactic component occurs (cf. Sec. 2.2.1). Hence, a detailed study of the energy spectrum of primary cosmic rays and their composition in this energy region is needed to distinguish between the different theoretical models describing this transition. This is the motivation for the two low-energy enhancements of the Pierre Auger Observatory, namely Auger Muons and Infill for the Ground Array (AMIGA) and High Elevation Auger Telescopes (HEAT). With these enhance-

ments, the energy range that is observed through SD and FD measurements is extended down to 10^{17} eV. Thus, both the second knee and the ankle are in the observed energy range of the Pierre Auger Observatory. In addition, by extending the observed energy range to lower energies that are also observed by other experiments, such as KASCADE-Grande, a direct comparison of the data in the overlapping region becomes possible. The AMIGA and HEAT enhancements are discussed in Secs. 4.3.1 and 4.3.2, respectively.

Since the detector systems operated at the Pierre Auger Observatory are well understood and yield data of high quality, the detector site also provides an ideal environment for the research and development of new detectors and techniques for measuring extensive air showers. The detection of extensive air showers through radio measurements is researched in the scope of the Auger Engineering Radio Array (AERA) enhancement, which is described in Sec. 4.3.3. In addition, several projects investigate the use of the GHz regime of the electromagnetic spectrum for the detection of extensive air showers. A summary of these activities can be found in [All11].

4.3.1. Auger Muons and Infill for the Ground Array (AMIGA)

The AMIGA enhancement consists of two components, an infill array to the regular SD array and additional underground muon counters, associated to the detector stations of the infill array [Etc07]. For the infill array, an area of $27.3\,\text{km}^2$ near the FD site Coihueco has been equipped with 71 additional SD stations to form a denser grid with a spacing of 750 m between the single detectors (see Fig. 4.7(a)) [Rav13]. The additional detector stations are identical to the ones of the regular SD array with the exception of the communications system, which has been upgraded to increase the bandwidth available for the transmission of the data from both the SD station and the associated muon counter. The trigger system has been adopted from the regular SD array. However, due to the smaller spacing between the single detectors, the threshold for full trigger efficiency is 3×10^{17} eV for zenith angles below 55° [Rav13], which is a full order of magnitude below the threshold for the regular SD array. Since the infill array is overlooked by FD telescopes from the FD site Coihueco site as well as from the HEAT enhancement (see Sec. 4.3.2), hybrid measurements in the energy region below 10^{18} eV are possible. It is also planned to deploy SD stations in an even denser grid with a spacing of 433 m on a portion of the infill array with an area of about $6\,\text{km}^2$ [Etc07]. However, work on this part of the infill array has not been started yet.

The second component of the AMIGA enhancement are muon counters that are placed in the ground alongside the SD stations of the infill array. Since these detectors are buried in a depth of 2.3 m [Mal13], corresponding to a slant depth of $540\,\text{g}\,\text{cm}^{-2}$, the electromagnetic component of an extensive air shower is almost fully absorbed before reaching the muon counters so that essentially only the muonic component is measured, which is strongly correlated with the mass number of the primary particle initiating the air shower (cf. Sec. 2.3.2). Hence, the AMIGA muon counters will provide an opportunity to determine the elemental composition of cosmic rays in the energy range around 10^{18} eV. Each muon counter consists of four modules, with two modules covering an area of $10\,\text{m}^2$ and the remaining two modules covering an area of $5\,\text{m}^2$ each (see Fig. 4.7(b)) [Mal13]. Each module consists of 64 plastic scintillator strips, organized into two groups of 32 strips on either side of a central dome, which houses a 64-pixel PMT and the read out electronics [Etc07]. When a muon passes
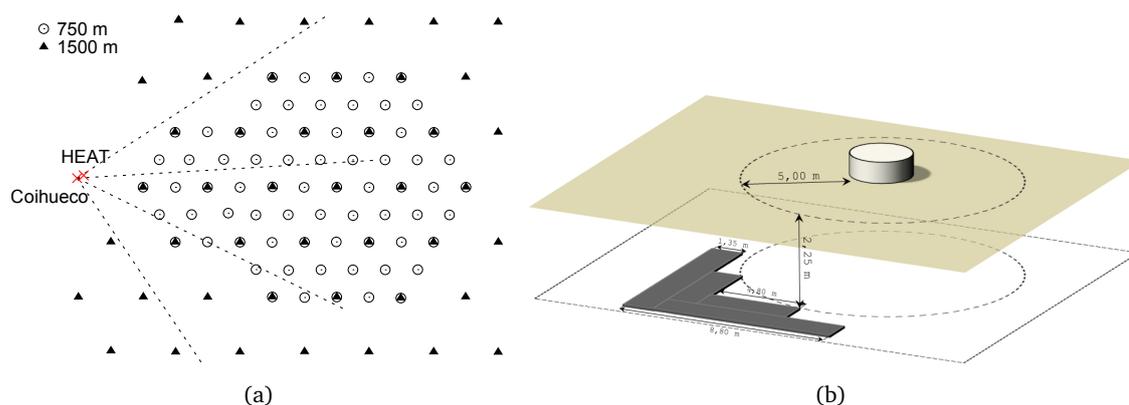
Figure 4.7.: The AMIGA enhancement of the Pierre Auger Observatory. (a) Layout of the infill portion to the regular SD array near the FD site Coihueco (adapted from [Rav13]). SD stations from the regular array with a spacing of 1500 m are marked by triangles, while the additional stations from the infill array with a spacing of 750 m are denoted by circles. Some stations belong to both arrays. In addition, the fields of view of the three FD telescopes overlooking the infill array are indicated. (b) Possible layout of a muon counter buried alongside an SD station [Frö09]. In this layout, the muon counter consists of four independent detector modules.

through one of the scintillator strips, the analog signals from the PMT are digitized by discriminators and sampled by an FPGA with a frequency of 320 MHz. The data are then sent to the associated SD station, which transmits both the SD data and the muon counter data to the CDAS. A counting algorithm is then applied offline to determine the number of muons hitting a given muon counter [Mal13]. To test the baseline design of the muon counters, an engineering array consisting of a single hexagon of SD stations from the infill array has been equipped with muon counters [Mal13]. First data has been recorded with this engineering array in coincidence with the SD to determine the accuracy of the muon counting and develop the final design of the muon counters which is then used to equip all SD stations of the infill array with muon counters.

4.3.2. High Elevation Auger Telescopes (HEAT)

In the scope of the HEAT enhancement, three additional fluorescence telescopes have been installed close to the FD site Coihueco. These telescopes are identical to the regular FD telescopes, except for the ability to tilt the telescopes upward. This enables the HEAT telescopes to record lower-energy showers—which develop in general higher in the atmosphere—and thus lower the threshold for FD observations to below 10^{17} eV [Mat11]. Each HEAT telescope is housed in an individual, pivot-mounted shelter (see Fig. 4.8(a)), which is made out of lightweight insulated walls coupled to a steel structure, which itself rests on a strong steel frame filled with concrete. Each shelter can be tilted upward by 29° within two minutes through a hydraulic system [Mat11]. The schematic layout of a HEAT telescope is shown in Fig. 4.8(b). The optical components of the telescopes are connected to the ground plate to avoid, e.g., wind-induced vibrations and provide a fixed geometry. The mechanical stability of the whole system, including the mirror and the camera, is monitored using distance and inclination sensors [Mat11]. The data acquisition and the trigger system of the HEAT telescopes

(a)   (b)
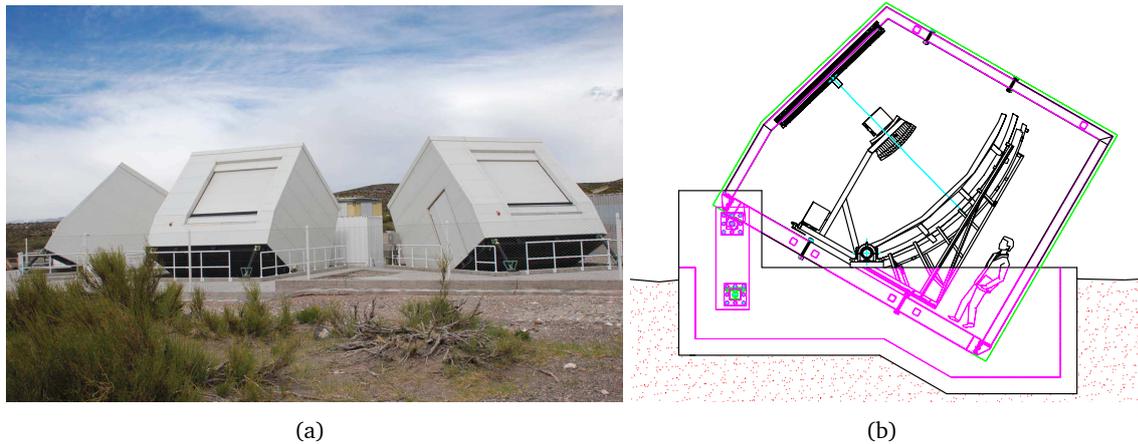
Figure 4.8.: The HEAT enhancement of the Pierre Auger Observatory [Mat11]. (a) Photograph of the three HEAT telescopes in tilted mode with the shutters closed during the day. (b) Schematic view of a HEAT telescope in tilted mode.

are based to the regular FD telescopes, using updated electronics with an increased sampling rate of the digitization of 20 MHz [Mat11].

With the HEAT telescopes, air shower events can be recorded both in the horizontal position as well as in the tilted position. The horizontal position is mainly used for the maintenance of the hardware, and it is also the position in which the absolute calibration of the telescope is performed [Mat11]. In this position, the fields of view of the HEAT telescopes overlap with the telescopes from the FD site Coihueco. Hence, events recorded by telescopes from both sites can be used to check the alignment of the telescopes and provide a cross check for the calibration of the telescopes. In the tilted mode, the FD telescopes from HEAT and the FD site Coihueco cover a combined elevation range from the horizon to 58°. With this extended field of view, the measurement of low-energy showers, which can only be measured if they develop close to the telescope and are thus usually outside the field of view of the regular FD telescopes, becomes possible [Mat11].

4.3.3. Auger Engineering Radio Array (AERA)

When an extensive air shower develops in the atmosphere, the secondary particles contained within the shower not only emit fluorescence or Cherenkov light, but also electromagnetic radiation in the radio frequency band between 10 MHz and 100 MHz. The dominant emission mechanism is geosynchrotron emission, i.e. synchrotron emission of radiation from the electrons and positrons in an air shower in the geomagnetic field [Aab14b]. The short bursts of radio signals emitted by an extensive air shower contain information about the longitudinal development of the shower in the atmosphere. Hence, the radio technique can provide a measurement of the longitudinal shower profile without the limitations of the FD, such as the low duty cycle. Determining the precision of the reconstruction of the properties of the primary cosmic rays with the radio technique and understanding the details of the mechanisms for radio emission are the scientific goals of the AERA enhancement. In addition, AERA serves as

(a) (b)
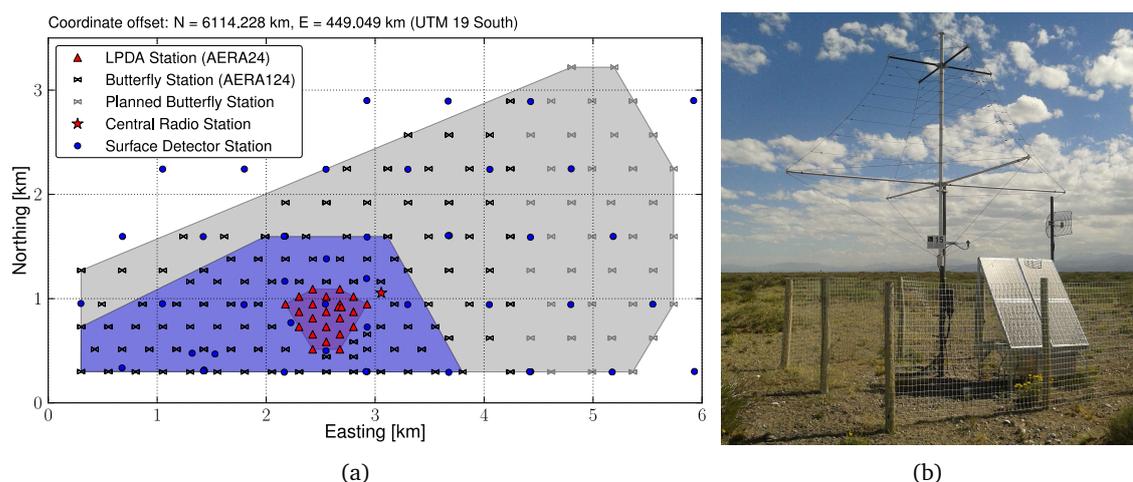
Figure 4.9.: The AERA enhancement of the Pierre Auger Observatory [Wei14]. (a) Layout of the AERA site, which uses two different types of radio detector stations. The AERA site is nested within the infill array. The colored areas highlight the different spacings of the grid of detector stations (150 m in red, 250 m in blue, and 375 m in gray). (b) Photograph of a radio detector station, using the LPDA design. The dish antenna visible on the right hand side is used for the transmission of the data to the central radio station.

a feasibility study for the application of the radio technique on larger scales [Ber09].

AERA is embedded within the SD array of the Pierre Auger Observatory, co-located with the infill array and overlooked by the HEAT telescopes. Currently, 124 radio detector stations of two different types are deployed on hexagonal grids with different spacings, covering an area of about 6 km^2 (see Fig. 4.9(a)) [Wei14]. The dense core of the array consists of 24 detector stations using Logarithmic Periodic Dipole Antennas (LPDAs) (cf. Fig. 4.9(b)) with a distance of 150 m. 100 stations with a different design, using butterfly-shaped antennas, are deployed around the core, on hexagonal grids with spacings of 250 m and 375 m. The central radio station provides a base for the deployment and contains parts of the central data acquisition system of AERA [Wei14].

Except for the design of the antennas, the different detector stations are similar. The antennas of a single station are arranged perpendicular to each other so that polarization sensitive measurements of the transient electric fields are possible in the east-west and the north-south directions. The antennas cover an ultra-broadband frequency range from 30 to 80 MHz [Wei14]. This frequency range is essentially free from continuous man-made noise sources such as radio broadcasts. After amplification and filtering, the signals are digitized with a frequency of 200 MHz using 14-bit ADCs [Wei14]. An FPGA allows for performing complex signal processing, such as fast Fourier transformations, and for implementing flexible trigger algorithms. A multi-threshold time domain trigger is used to identify signals which show the typical radio signature of an air shower event [Wei14]. The data are then transmitted wirelessly to the central radio station. Since the radio stations are designed to work automously like the SD stations, each station is equipped with a solar panel and a battery.

The central data acquisition system of AERA at the central radio station forms an event trigger based on time-coincident trigger signals sent to the central radio station by the single radio stations [Wei14]. A fast geometric reconstruction of the arrival direction of the radio signals is then performed to reject radio signals from known noise sources, which have been identified during the installation of AERA in systematic noise studies [Wei13]. In the last stage of the data acquisition, real air shower events are identified through coincidences with the SD. In addition to the self-trigger, two complementary trigger modes have been implemented: an external trigger from the SD, and an internal trigger based on scintillation detectors located within the AERA array [Wei14].

4.4. Plans to Upgrade the Observatory beyond 2015

In the recent years, the Pierre Auger Observatory has made significant contributions to the field of cosmic-ray physics. Yet, answers to the fundamental questions about the nature and the origin of UHECRs remain elusive. To help providing these answers, several upgrades to the detector systems of the Pierre Auger Observatory are planned for the period between 2015 and 2023 [Pie13b]. The physics goals for these upgrades are threefold: the first goal is establishing the nature of the suppression of the flux of UHECRs above 6×10^{19} eV, i.e. whether the suppression is due to the GZK effect or due to a maximum energy to which the cosmic rays can be accelerated at the sources (cf. Sec. 2.2.1). Determining the contribution of protons to the total flux of UHECRs is the second goal. Measuring the fraction of protons with a sensitivity better than 10 % will not only allow for an evaluation of the physics potential of future experiments in astroparticle phyiscs, but also lead to better theoretical predictions for the fluxes of UHE photons and neutrinos. The third goal is to study in detail the development of extensive air showers induced by UHECRs—in particular hadronic multiparticle production—in order to, e.g., resolve discrepancies between the number of muons that is observed in an air shower and the number of muons that is predicted by MC simulations using current hadronic interaction models. Studying hadronic interactions in the context of UHECRs will allow for an exploration of fundamental particle physics in an energy range that is not accessible at current particle accelerators.

The key to accomplish the physics goals is to determine the elemental composition in the energy region of the flux suppression. Currently, studies of the elemental composition at the Pierre Auger Observatory are mainly done using FD measurements (see Sec. 2.2.2). However, in the energy region of the flux suppression, the statistics that can be obtained with FD measurements is severely limited. Hence, it is planned to upgrade the SD to allow for an event-by-event determination of the electromagnetic and muonic shower components [Pie13b]. Knowing these contributions for an air shower allows for an estimation of the mass of the primary particle, in addition to a reconstruction of the primary energy in a mass-independent way. The separation of the electromagnetic and the muonic air shower components will be achieved by adding a scintillation detector, consisting of a thin scintillator with an area of $2\,\mathrm{m}^2$ and a single PMT, on top of each SD station [Pie13b, Ber14]. Due to the different sensitivities of the different detector systems to the electromagnetic and muonic shower components, a determination of the muonic content of an air shower is possible through a combined analysis of the data from the SD stations and the associated scintillation detectors. In addition to the scintillation detectors, it is foreseen to upgrade the readout electronics of the SD stations, which will provide a three times faster sampling of the signals from the PMTs, a significantly enhanced

dynamic range and the possibility to implement enhanced triggering and monitoring capabilities [Pie13b]. The read-out electronics of the scintillation detectors on top of the SD station will be fully integrated into the upgraded SD station electronics.

Depending on the approval of proposed upgrades, the implementation of the upgrades is expected to be finished by 2017. It is planned to operate the upgraded Obervatory until 2023. This will eventually enlarge the dataset available by 2015 by more than a factor of two [Pie13b]. In particular, it is estimated that about 430 events at the highest energies—i.e. above an energy of $\log_{10}(E\,[\text{eV}]) = 19.7$—will be recorded.

4.5. Reconstruction of Air Shower Events

In the following sections, the reconstruction of air shower events from the raw data recorded by the FD and the SD is described. If an air shower event is recorded simultaneously with both detector systems and passes all trigger levels in both systems, the event can be fully reconstructed from either SD or FD data. These "golden hybrid events" provide an important means to cross check the detector systems and are used to determine the energy scale of the SD by comparisons with the reconstructed energy from the corresponding FD data. At lower energies, i.e. below the threshold for full trigger efficiency of the SD, only one or two SD stations are triggered, which is not enough to fulfill the T3 trigger requirements. However, if the air shower event is recorded simultaneously by the FD, a hybrid trigger is generated, so that the data from the individual SD stations are still stored, even though a full reconstruction of the air shower event from SD data is not possible. In the reconstruction of such hybrid events, the raw data from the FD is complemented by the additional timing information from the SD stations to improve the reconstruction of the geometry of the recorded extensive air shower.

For the analysis presented in this thesis, hybrid data is used. The reconstruction of air shower events from hybrid data is presented in detail in Sec. 4.5.1. In addition, the reconstruction from pure SD data is also briefly summarized (Sec. 4.5.2), since it provides a useful background for the F_γ parameter, which is used in the analysis presented in this thesis to distinguish between UHE photons and protons as primary particles. The basic principles of the event reconstructions are described in the following sections. The actual implementation of these principles in software is done through the Auger Off̄l̲i̲n̲e̲ Software Framework, which is also discussed in Sec. 5.1.2.

4.5.1. Reconstruction from Hybrid Data

The reconstruction of air shower events from hybrid data can be broken down into three stages. In the first stage, the geometry of the air shower is determined, i.e. the impact point on ground and the shower axis. In the second stage, the longitudinal shower profile is calculated from the signals recorded by the individual pixels. In the last step of the reconstruction, the energy of the primary particle is obtained from the longitudinal shower profile.

Geometry Reconstruction

An air shower event is recorded by an FD telescope as a sequence of triggered pixels in the camera (see Fig. 4.10(a)). In the first step of the event reconstruction, the Shower-Detector

(a)   (b)
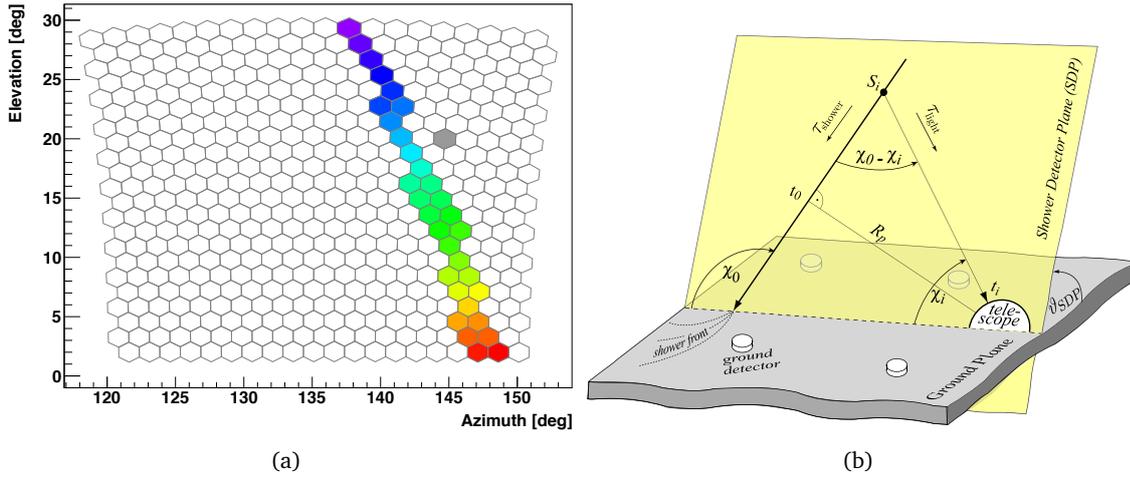
Figure 4.10.: (a) An example sequence of triggered pixels (FD event 1/4933/784, recorded on January 26, 2014, by the FD site Los Leones). The triggered pixels are color-coded according to the arrival time of the signals at the individual pixels, with violet corresponding to early and red corresponding to late. Pixels marked in gray are considered accidental triggers, not fitting to the expected track of an air shower event. (b) Illustration of the geometry used in the reconstruction of an air shower event [Kue08]. The SDP contains both the air shower and the center of the FD site recording the shower. The quantities denoted in the figure are explained in the text.

Plane (SDP) is determined, i.e. the plane which contains the trajectory of the air shower, expressed through the sequence of pixels, and the center of the FD site which recorded the event (cf. Fig. 4.10(b)). The unit normal vector $\vec{n}$ defining the SDP is obtained through a χ^2 minimization, taking into account the pointing directions $\vec{r}_i$ of the triggered pixels [Arg03]:

$$\chi^2_{\mathrm{SDP}} = \sum_i \frac{(\vec{n} \cdot \vec{r}_i)^2 w_i}{\sigma_{\mathrm{SDP}}}, \tag{4.1}$$

where the signals recorded in each individual pixel are used as the weights w_i. The pointing accuracy σ_{SDP} of the SDP reconstruction has been obtained from vertical laser shots, where the geometry of the recorded laser event is known. In the reconstruction procedure, a value of $\sigma_{\mathrm{SDP}} = 0.35°$ is used [Arg03].

Next, the direction of the shower axis within the SDP is determined with the use of the arrival time of the signals at the individual pixels. The arrival times depend on both the propagation time of the shower through the atmosphere and the propagation time of the emitted fluorescence light from the shower to the FD telescope. The propagation time of the shower—which is assumed to propagate at the speed of light in the vacuum c—from a point S_i to the reference point t_0, which corresponds to the point of closest approach of the shower to the FD telescope, can be expressed as

$$\tau_{\mathrm{Shower},\,i} = \frac{R_p}{c \tan(\chi_0 - \chi_i)}, \tag{4.2}$$

where χ_0 and χ_i are the viewing angles of the reference point and the point S_i, respectively, as seen from the FD telescope (cf. Fig. 4.10(b)), and R_p is the distance of closest ap-
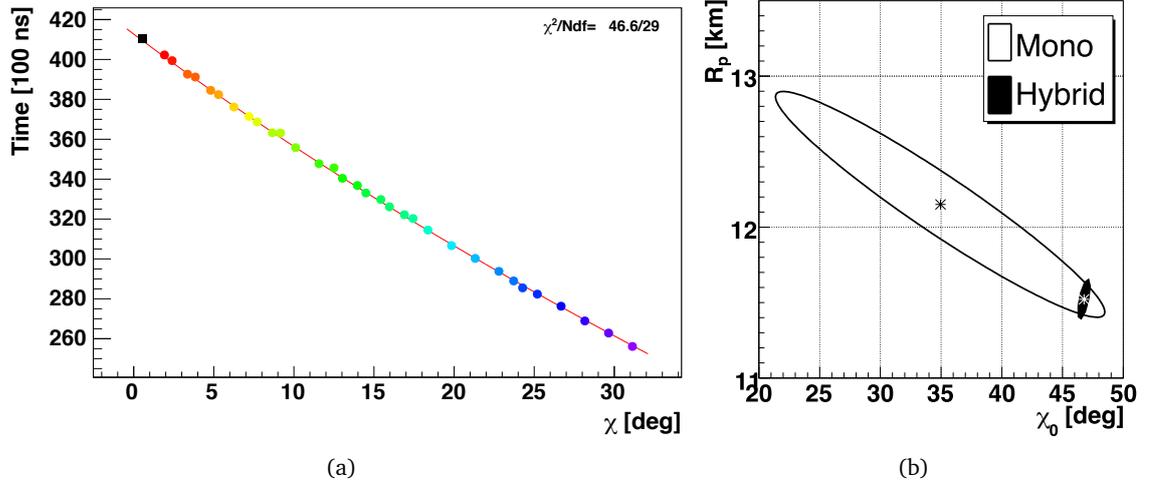
Figure 4.11.: (a) The arrival times of the signals recorded by the single pixels as a function of the viewing angle χ for the example event shown in Fig. 4.10(a). The color code of the data points is the same as in Fig. 4.10(a). In addition to the data points from the FD, the timing information from the SD is marked by the black square. The parameters t_0, R_p and χ_0 from Eq. 4.2 are obtained through a fit to the data points, shown in the figure as a red line. (b) Comparison of the fit results for an FD-only (mono) reconstruction and a hybrid reconstruction for another example event [Mos07]. The shaded areas indicate the 1σ regions around the solutions that minize the χ^2 in the fits, denoted by stars.

proach [Kue08]. Next, the propagation time of the fluorescence light to the FD telescope can be expressed similarly, assuming the light propagates on straight lines with velocity c:

$$\tau_{\text{Light, }i} = \frac{R_p}{c \sin(\chi_0 - \chi_i)}. \tag{4.3}$$

Using Eqs. 4.2 and 4.3 and assuming an instantaneous emission of fluorescence light at the point S_i, the expected arrival time t_i with respect to the reference time t_0 can be calculated [Kue08]:

$$\begin{aligned}
t_i &= t_0 - \tau_{\text{Shower, }i} + \tau_{\text{Light, }i} \\
&= t_0 - \frac{R_p}{c \tan(\chi_0 - \chi_i)} + \frac{R_p}{c \sin(\chi_0 - \chi_i)} \\
&= t_0 + \frac{R_p}{c} \tan\left(\frac{\chi_0 - \chi_i}{2}\right).
\end{aligned} \tag{4.4}$$

The parameters t_0, R_p and χ_0, which fully describe the shower geometry within the SDP, are obtained through a χ^2 minimization:

$$\chi^2_{\text{Time}} = \sum_i \frac{(t_i - t_{\text{meas, }i})^2}{\sigma_i}, \tag{4.5}$$

with the expected arrival times t_i and the measured arrival times $t_{\text{meas, }i}$ with uncertainty σ_i (see Fig. 4.11(a)). So far, only FD data is taken into account in the reconstruction of the geometry of an air shower event. However, the accuracy of the geometry reconstruction is limited,

especially when the measured angular speed $d\chi/dt$ does not change much over the observed track length [Abr10a]. For such events, the fit parameters R_p and χ_0 become strongly correlated and the uncertainties of the fitted parameters become large (cf. Fig. 4.11(b)). The accuracy of the fit can be significantly improved by taking into account additional timing information from the SD. The time at which the shower hits the ground—obtained from the timing information of the SD station with the largest signal—is used as an additional data point in the fit of Eq. 4.4 [Abr10a]. Since this data point is usually far away from the bulk of the data points obtained from the FD, the fit is strongly constrained, and the accuracy in the determination of the fit parameters is improved (cf. Fig. 4.11(b)).

Finally, the impact point of the shower on ground, i.e. the point of intersection of the shower axis with the ground plane, and the incoming direction of the air shower, expressed through the zenith and azimuth angles, can be calculated from the reconstruced shower geometry. Using hybrid data, the impact point can be determined with a resolution of 50 m, while the typical angular resolution in the reconstruction of the arrival direction is 0.6° [Abr10a].

Profile Reconstruction

Once the geometry of the air shower event is known, the longitudinal shower profile can be calculated from the signals measured in the single PMTs, taking also into account the atmospheric conditions. Using the known detector calibration (see Sec. 4.2.2), these signals can be converted into a flux of photons at the aperture of the FD telescope (cf. Fig. 4.12(a)). However, not only fluorescence light is collected by the FD telescope, but also other light, e.g. direct and scattered Cherenkov light produced by high-energy charged particles in the

(a)   (b)
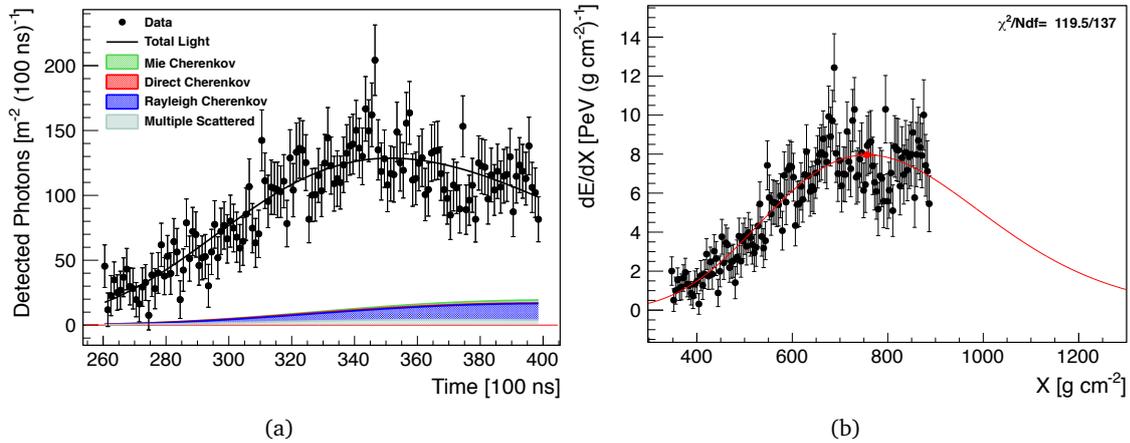
Figure 4.12.: (a) The detected light at the aperture of the FD telescope as a function of time for the example event shown in Fig. 4.10(a). From the data points, the total light profile (black line) is calculated and the different components of the light (colored areas) are disentangled. (b) Longitudinal shower profile in terms of energy loss per unit atmospheric depth, dE/dX, as a function of the atmospheric depth X, reconstructed from the light profile shown in (a). The red line results from a fit of a Gaisser-Hillas function (cf. Eq. 2.23) to the reconstructed profile. The position of the shower maximum, X_{max}, is indicated by a red marker.

extensive air shower. The different components of the measured light distribution have to be disentangled. The method used for disentangling the components is described in [Ung08]. For the fluorescence light component, the number of photons $N_\gamma^{\text{fl.}}$ observed from a certain slant depth X_i is essentially proportional to the energy deposited in the atmosphere by the charged particles from the air shower at this slant depth [Ung08]:

$$N_\gamma^{\text{fl.}}(X_i) = \frac{\mathrm{d}E}{\mathrm{d}X_i}\, Y_i^{\text{fl.}}\, \Delta X_i\, T_i^{\text{atm.}}\, \frac{A}{4\pi r_i^2}\, \epsilon. \tag{4.6}$$

In this equation, $Y_i^{\text{fl.}}$ denotes the fluorescence yield at the given slant depth, ΔX_i is the traversed slant depth in the given observational time window, $T_i^{\text{atm.}}$ is a coefficient that describes the attenuation of the light propagation through the atmosphere due to Rayleigh and Mie scattering processes, $\frac{A}{4\pi r_i^2}$ describes the fraction of the (isotropically emitted) fluorescence light that reaches the aperture of the FD telescope with area A at a distance r_i and ϵ is the light collection efficiency. Thus, once the different components of the light observed at the aperture of an FD telescope are disentangled, the longitudinal shower profile in terms of the energy deposited in the atmosphere—a quantity that is directly proportional to the number of particles in the air shower—as a function of the slant depth can be calculated (cf. Fig. 4.12(b)). However, usually only part of the shower profile is recorded by the FD telescopes. To extrapolate the observed shower profile into the regions outside the field of view of the FD telescopes, a Gaisser-Hillas function following to Eq. 2.23 is fitted to the observed longitudinal profile. Through this fit, the position of the shower maximum, X_{max}, is directly obtained with an uncertainty better than $20\,\text{g}\,\text{cm}^{-2}$ [Ung08].

Energy Reconstruction

In the last stage of the hybrid event reconstruction, the energy E of the primary particle initiating the air shower is determined. First, the longitudinal shower profile is integrated to obtain the calorimetric energy E_{cal}:

$$E_{\text{cal}} = \int_0^\infty \frac{\mathrm{d}E}{\mathrm{d}X}(X)\,\mathrm{d}X. \tag{4.7}$$

E is then obtained from the calorimetric energy by correcting E_{cal} for the "invisible energy" carried away by neutrinos and high-energy muons:

$$E = (1 + f_{\text{inv}})\, E_{\text{cal}}. \tag{4.8}$$

At energies around $10^{19}\,\text{eV}$ and assuming protons or nuclei as primary particles, the correction factor f_{inv} for the invisible energy correction is of the order of $10\,\%$ [Tue13]. For extensive air showers initiated by UHE photons, which do not exhibit a significant neutrino or muon component, the correction is of the order of $1\,\%$ [Pie05].

Systematic uncertainties affecting the energy scale used in the reconstruction of air shower events from hybrid data have been studied in detail in [Ver13]. The total systematic uncertainty on the energy scale is $14\,\%$. Major contributions are given by the calibration of the FD telescopes ($9.9\,\%$), systematic uncertainties in the profile reconstruction ($6.5\,\% - 5.6\,\%$, depending on the energy), systematic uncertainties on the fluorescence yield ($3.6\,\%$) or systematic uncertainties in the determination of the atmospheric conditions ($3.4\,\% - 6.2\,\%$) [Ver13].

4.5.2. Reconstruction from SD Data

Similarly to the reconstruction of air shower events from hybrid data, the reconstruction of an event from SD data can also be broken down into three stages. First, the geometry of the air shower is determined from the timing information and the positions of the single SD stations. In the second step, the lateral profile of the shower is calculated from the signals recorded by each station. In the last stage, the energy of the primary particle initiating the air shower is obtained from the lateral profile. The full reconstruction procedure is described in [Veb13b]. Here, only a summary of the main points of the reconstruction is given.

The signal-weighted barycenter of the triggered SD stations is used as a first estimate for the impact point of the shower on ground. From this first estimate, the direction of the shower axis is calculated from the arrival times of the signals at the single SD stations under the assumption of a planar shower front travelling at the speed of light along the shower axis. The geometry reconstruction is later refined by replacing the planar shower front with a curved shower front as part of an expanding sphere and repeating the procedure [Veb13b].

Once the geometry is known, the lateral distribution of the signals measured in the SD stations can be calculated. The lateral distribution of the signals S, measured at a perpendicular distance r to the shower axis is modeled by an LDF, normalized to the signal measured at a

(a)   (b)
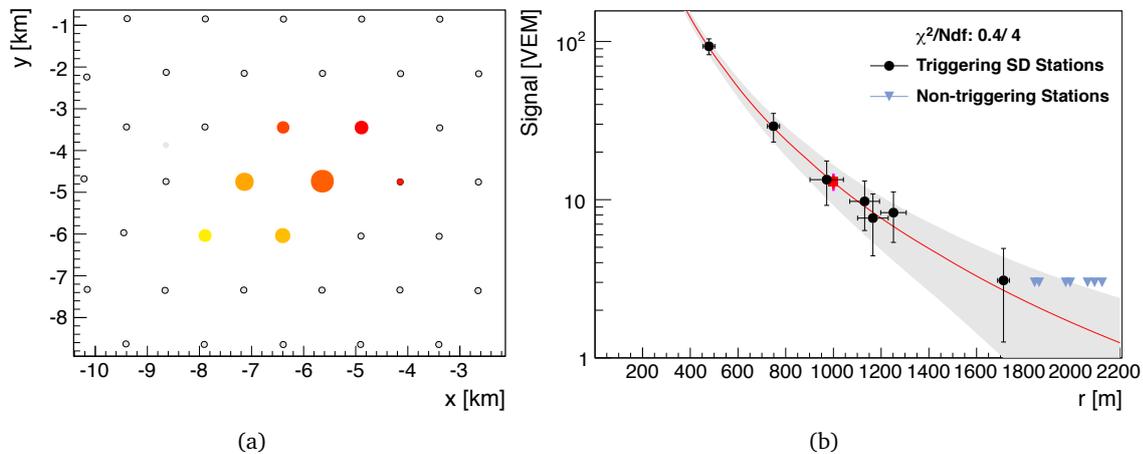
Figure 4.13.: Example of an SD event (SD event 24948716, recorded on January 4, 2014). (a) Footprint of the extensive air shower on the SD array. The triggered stations are color coded according to the timing information of the signals, with yellow corresponding to early and red corresponding to late. The size of the markers corresponds to the signal strength recorded in the SD station. The origin of the coordinate system used in this plot lies in the center of the SD array. (b) The lateral distribution of the signals recorded in the SD stations as a function of the perpendicular distance of the station to the shower axis, r. In addition to the triggering stations (black markers), the positions of close-by non-triggering stations are indicated (blue triangles). The LDF fitted to this distribution is shown as a red line. Also indicated is the S_{1000} obtained from this fit (red marker).

(a)

(b)
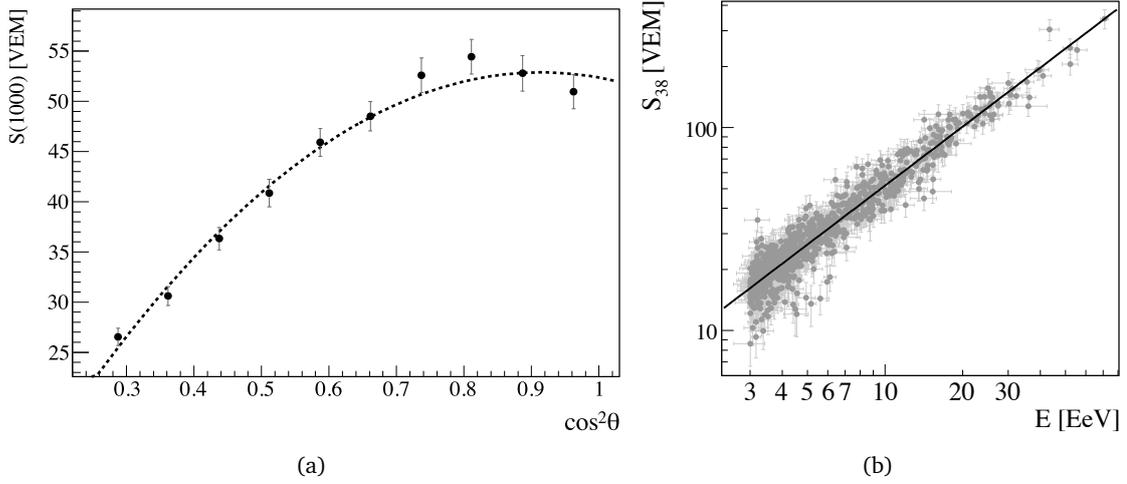
Figure 4.14.: (a) The function $\mathrm{CIC}(\theta)$, obtained from data using the CIC method [Pes11]. A second-order polynomial in $x = \cos^2(\theta) - \cos^2(\bar{\theta})$ with $\bar{\theta} = 38°$ has been fitted to the distribution (dotted line). (b) The relation between S_{38} and E_{FD} for a set of well-measured golden hybrid events [Pes11]. The relation is described using a power-law function according to Eq. 4.12 (solid line).

distance of 1000 m [Veb13b]:

$$S(r) = S_{1000} \times f_{\mathrm{LDF}}(r),$$
$$f_{\mathrm{LDF}}(r) = \left(\frac{r}{1000\,\mathrm{m}}\right)^{\beta} \left(\frac{r + 700\,\mathrm{m}}{1700\,\mathrm{m}}\right)^{\gamma}. \tag{4.9}$$

The functional form of $f_{\mathrm{LDF}}(r)$ given in Eq. 4.9 is commonly known as a modified NKG function (cf. also Sec. 2.3.2). The shape parameters β and γ define the steepness of the LDF. The parameters S_{1000}, β, and γ are obtained through a fit of the LDF to the measured distribution of signals (cf. Fig. 4.13(b)). Due to geometry effects and the attenuation of the secondary particles from the air shower in the atmosphere, the value of S_{1000} decreases with increasing zenith angle for a given primary energy [Pes11]. To remove this dependence, S_{1000} is converted to a zenith-independent quantity, S_{38}:

$$S_{38} = \frac{S_{1000}}{\mathrm{CIC}(\theta)}. \tag{4.10}$$

The function $\mathrm{CIC}(\theta)$ is obtained from data using the CIC method [Her61]. The parameter S_{38} obtained in this way can be regarded as the S_{1000} the shower would have if it impinged on the Earth under the reference zenith angle of $\bar{\theta} = 38°$ [Pes11]. To describe the functional form of $\mathrm{CIC}(\theta)$, a second-order polynomial in $x = \cos^2(\theta) - \cos^2(\bar{\theta})$ has been chosen:

$$\mathrm{CIC}(\theta) = 1 + ax + bx^2, \tag{4.11}$$

where $a = 0.87 \pm 0.04$ and $b = -1.49 \pm 0.20$ [Pes11]. Finally, the energy of the primary particle is calculated from S_{38} using the following relation:

$$E = A \times \left(S_{38}\,[\mathrm{VEM}]\right)^{B}. \tag{4.12}$$

The constants $A = (1.68 \pm 0.05) \times 10^{17}$ eV and $B = 1.035 \pm 0.009$ have been obtained through a comparison of the measured S_{38} and the energy reconstructed from FD data (E_{FD}) for a set of well-measured golden hybrid events (see Fig. 4.14(b)) [Pes11].

5 Simulations and Data

In the search for UHE photons, a detailed knowledge of the development of air showers induced by different primary particle types and the response of the different detector systems to these air showers is necessary. Up to now, no UHE photons have been unambiguously identified in data. Hence, extensive studies of simulated air showers and the corresponding detector response using MC programs are necessary. Using this knowledge, improved analysis techniques can be developed with the aim of better distinguishing air shower events possibly induced by UHE photons from those events that are induced by other particle types. In the following chapter, the simulation samples which are used in the analysis presented in this thesis are discussed in detail. In particular, the settings used to generate the simulation samples are documented in Sec. 5.1. The subset of data collected by the Pierre Auger Observatory to which the analysis is later applied is presented in Sec. 5.2. It should be noted that these samples are "raw" samples, because they contain all triggered events, regardless of the actual quality of the event, for example, how well the longitudinal shower profile can be reconstructed from the recorded signals. Before the samples are used in the analysis, subsamples of well-reconstructed events are extracted from the raw samples. The criteria for the selection of events and their application to the simulation samples and the data sample are described in Sec. 5.3. In the next chapters, the simulation samples are used to optimize an analysis to identify UHE photons (Ch. 6), which is then applied to the data sample (Ch. 7).

5.1. Simulation Samples Used in the Analysis

The generation of the simulation samples is essentially a two-step process: first, extensive air showers are simulated using the MC program CORSIKA, version 7.4000. These simulated showers are subsequently used as an input for the simulation of the detector response, which is based on the Auger $\overline{\text{Off}}\underline{\text{line}}$ Software Framework, version 2.9.1. The CORSIKA part of the simulations, i.e. the basic principles of CORSIKA and the specific settings used to generate the simulation samples discussed in this thesis, is described in Sec. 5.1.1. Similarly, the $\overline{\text{Off}}\underline{\text{line}}$ part is documented in Sec. 5.1.2.

Two simulation samples have been generated: one with photons as primary particles and another sample with protons as primary particles. The energies of the primary particles range from 10^{18} eV to 10^{19} eV. This energy range has been divided into 10 bins with a constant width of 0.1 in terms of $\log_{10}(E\,[\text{eV}])$. In each energy bin and for each primary particle type, 2000 showers have been simulated, i.e. 40 000 showers in total. The size of the data samples has been chosen such that a large range of primary parameters as well as the shower-to-shower fluctuations between showers with the same primary parameters are covered. Each simulated shower was then used five times as input for the $\overline{\text{Off}}\underline{\text{line}}$ part of the simulations to increase the available statistics. Hence, 10 000 events are simulated in each of the 10 energy bins and for each of the two primary particle types, i.e. 200 000 events in total. However, the number of events is reduced in the final sample due to the trigger efficiency of the FD (see also Sec. 5.1.2). Around 10^{19} eV, the trigger efficiency is close to 100 %, independent of the distance of the air shower from the FD telescope. Toward lower energies, the average trigger efficiency decreases, as only showers closer to the FD telescope can be detected [Pet04]. For primary photons, the trigger efficiency is further reduced due to the deeper development of the air shower in the atmosphere: on the one hand, air showers induced by UHE photons may not develop fully within the field of view of the telescopes so that only the part of the

Energy bin	Number of events	
	Photon	Proton
$18.0 \leq \log_{10}(E\,[\text{eV}]) < 18.1$	5412	6229
$18.1 \leq \log_{10}(E\,[\text{eV}]) < 18.2$	5660	6721
$18.2 \leq \log_{10}(E\,[\text{eV}]) < 18.3$	6093	7228
$18.3 \leq \log_{10}(E\,[\text{eV}]) < 18.4$	6826	7757
$18.4 \leq \log_{10}(E\,[\text{eV}]) < 18.5$	7578	8688
$18.5 \leq \log_{10}(E\,[\text{eV}]) < 18.6$	8160	9392
$18.6 \leq \log_{10}(E\,[\text{eV}]) < 18.7$	8712	9750
$18.7 \leq \log_{10}(E\,[\text{eV}]) < 18.8$	9064	9921
$18.8 \leq \log_{10}(E\,[\text{eV}]) < 18.9$	9427	9964
$18.9 \leq \log_{10}(E\,[\text{eV}]) < 19.0$	9551	9987

Table 5.1.: Number of events in the two main simulation samples, subdivided into the single energy bins. The numbers given in the table refer to the raw samples, i.e. after the $\overline{\text{Off}}\underline{\text{line}}$ part of the simulations but before the event selection stage. Hence, the trigger efficiency of the FD is included. Originally, 10 000 events have been simulated in each of the 10 energy bins and for each of the two primary particle types.

longitudinal shower profile before the shower maximum is observed, where less fluorescence light is emitted. On the other hand, when the shower develops deeper in the atmosphere, the fluorescence light has to traverse a larger optical depth on the way to the FD telescope. Hence, a larger fraction of the fluorescence light emitted by a shower induced by a UHE photon is absorbed before it reaches the FD telescope compared to air showers induced by protons, which develop in less dense layers of the atmosphere. The numbers of events in the main simulation samples after the Offline part of the simulation chain are listed in Tab. 5.1.

In addition to the two main samples, several smaller samples have been generated for the purpose of cross checking parts of the analysis and the results. These samples, for which in general the same parameters have been used in the simulations as for the main samples, will be briefly discussed and subsequently evaluated in Sec. 6.4. For some topics discussed in this thesis which are not directly related to the main analysis, other simulation samples have been used, which were not generated with the settings described in this section. For completeness, these samples are briefly discussed later on in the respective sections.

5.1.1. Air Shower Simulations with CORSIKA

Cosmic Ray Simulations for KASCADE (CORSIKA) [Hec98] is a MC program to simulate in detail the development of extensive air showers in the atmosphere. It was developed originally to perform air shower simulations for the KASCADE experiment, but the program has been extended to suit a wider range of applications. Hence, CORSIKA is commonly used today as a standard tool in many air shower experiments. The secondary particles created in an extensive air shower are tracked explicitly until they interact with other particles or decay. These processes are treated according to the current state of knowledge. For particle decays, all decay channels down to a branching ratio of 1 % are taken into account [Hec98]. Different models can be chosen for the treatment of hadronic interactions, i.e. for the calculation of the inelastic hadronic cross-sections, the evaluation of the interactions and the production of further secondary particles in these interactions. For the simulations discussed in this thesis, QGSJETII-04 [Ost11] has been selected to describe hadronic interactions at high energies. The QGSJET model is based on the Reggeon field theory [Gri68], which treats high-energy hadron-hadron collisions essentially as multiple scattering processes involving the exchange of Pomerons, and the quark-gluon string model [Kai82], which supplements the general Reggeon field theory with a specific hadronization model. The most recent version QGSJETII-04 takes into account Pomeron-Pomeron interactions through the resummation of the corresponding enhanced diagrams [Ost11]. In addition, the model has been tuned to current data from the LHC. At lower energies, i.e. below 80 GeV, the MC generator FLUKA 2011.2b.6 [Fer05, Bat07] is used. FLUKA employs a microscopic approach and combines parameterized fits to current experimental data with a variety of theoretical models [Fer05]. The treatment of the electromagnetic component of the extensive air shower is based on the EGS4 code [Nel85], which has been modified and enhanced for use within CORSIKA. EGS4 provides an MC description of all relevant electromagnetic processes, i.e., e.g., for electrons and positrons annihilation, Bhabha and Møller scattering, bremsstrahlung and multiple scattering, as well as Compton scattering, e^-e^+ pair production and photoelectric interactions for photons [Nel85]. In addition to these processes, $\mu^-\mu^+$ pair production and photonuclear interactions have been implemented, as these processes are essential for muon production in extensive air showers initiated by photons [Hec98]. Muon pair production is treated in full analogy to e^-e^+ pair

production. The cross sections for photonuclear interactions are modeled after experimental data [Hec98]. Depending on the energy of the photon, one or more pions are produced in these interactions. For the simulations with UHE photons as primary particles, the LPM effect is taken into account automatically in the modified EGS4 code. Since the energy range considered in the analysis presented in this thesis is below 10^{19} eV, the PRESHOWER option [Hom05] to take into account the preshower effect (cf. Sec. 3.3.1) has not been used.

A CORSIKA simulation is steered via an input file, where the parameters for the simulation are set. An example of an input file used for the simulations that are discussed in this thesis is given in App. A.1. The main steering parameters and the values that were used for the simulations are briefly summarized in this section, for a more detailed explanation of all steering parameters, see [Hec13]. Each input file is used to simulate a single extensive air shower. The type of the primary particle initiating the air shower is set through the PRMPAR parameter. The particle codes most relevant for the simulations discussed here are 1 for photon primaries and 14 for proton primaries. The energy of the primary particle is randomly chosen from the energy range specified through the ERANGE parameter—corresponding to one of the ten energy bins mentioned previously— according to a broken power law with an exponent given by the ESLOPE parameter, which is set to -1 for all simulation runs. The zenith angle is selected at random out of an interval given by the THETAP parameter, in a manner which respects equal fluxes of particles from all solid angle elements of the sky, under the assumption of a horizontal flat detector layout. Similarly, the azimuth angle is chosen randomly from the range given by the PHIP parameter. For the simulations discussed here, zenith angles between $0°$ and $65°$ and azimuth angles from $-180°$ to $180°$ have been considered. The simplified atmosphere used by CORSIKA consists of nitrogen (N_2), oxygen (O_2) and argon (Ar) with volume fractions of 78.1 %, 21.0 %, and 0.9 %, respectively [Hec13]. The density profile of the atmosphere is selected through the ATMOD parameter. For the simulations discussed here, the U.S. standard atmosphere—an idealized global atmosphere model based on measurements made at different mid-latitude geographical locations—as parameterized by John Linsley is used [Hec13]. The geomagnetic field used in the simulation run is controlled by the MAGNET parameter, which has been set to the default values for the location of the Pierre Auger Observatory, i.e. 20.1 µT in the horizontal direction and -14.2 µT in the vertical direction.

In principle, CORSIKA can track all secondary particles from the point where they are created up to the point where they decay or undergo interactions with other particles. However, an extensive air shower induced by an UHE primary particle is comprised of billions of secondary particles. Tracking every single secondary particle is unfeasible in terms of time and computing resources needed to perform the simulations. Therefore, only particles above a given energy threshold are tracked. The thresholds for the different particle types are set through the ECUTS parameter. The thresholds chosen for the simulations discussed here are 0.1 GeV for hadrons (except neutral pions), 0.05 GeV for muons, and 250 keV for electrons, photons and neutral pions. In addition to these energy thresholds, CORSIKA provides a thinning algorithm, where one particle is taken as a representative of a bunch of particles of the same type. This single particle is fully tracked, while the other particles are discarded. The thinning algorithm is controlled by the THIN parameter for electrons and photons and the THINH parameter for muons and hadrons. For the simulations discussed here, optimum 10^{-6} thinning [Kob99, Ris01] is used: the thinning algorithm is used for all secondary particles with an energy below a fraction of 10^{-6} of the primary energy E_0. The maximum weight w_{max}

a thinned particle can get, i.e. essentially the number of particles a single, thinned particle can represent, is set according to

$$w_{\max} = E_0 \,[\text{GeV}] \times 10^{-6}. \tag{5.1}$$

The value set for $w_{\max}$ in the input file refers to the lower bound of the energy range defined through `ERANGE`. Within this energy range, $w_{\max}$ scales linearly with the energy [Hec13]. The `THINH` parameter controls the thinning parameters for hadrons and muons relative to the parameters for electrons and photons. The energy threshold is kept the same, however, the maximum weight of a hadron or a muon is set a factor 100 below the maximum weight of electrons and photons.

Finally, the output of the simulation run is controlled through the `OBSLEV` and `LONGI` parameters. `OBSLEV` defines the observation level, measured above sea level, at which the secondary particles from the extensive air shower are written to a file, which then serves as an input for the simulation of the SD (see Sec. 5.1.2). For the simulations discussed here, an observation level of 1452 m a.s.l. has been chosen. The output of the longitudinal development of the simulated air shower is controlled by the `LONGI` parameter. The number of particles, differentiated by their particle type, is sampled in steps of $5\,\text{g}\,\text{cm}^{-2}$ in vertical atmospheric depth and written to a file, which serves as input for the FD simulation in the next step.

5.1.2. Detector Simulations with the Auger Offline Software Framework

The Auger Offline software framework ($\overline{\text{Off}}\underline{\text{line}}$) [Arg07] provides an infrastructure and tools to support the variety of distinct computational tasks necessary to analyze the data gathered at the Pierre Auger Observatory, in particular the tasks of reconstructing air shower events from the recorded raw data and simulating the detector response to extensive air showers simulated with MC programs such as CORSIKA. The requirements of a large, international collaboration working on an experiment that will take data over decades imposes strong demands on the software framework underlying the analyses [Arg07]. The framework must be flexible and robust enough to support the individual development of algorithms and allow for a comparison of these algorithms. Hence, it is essential that all physics algorithms are exposed, so that a user can easily replace existing algorithms with his own. In addition, the software framework must be extensible to accomodate future upgrades of the detector systems. The framework must also be able to handle a variety of data formats in order to deal with data coming from different sources, such as monitoring data, raw data from the different detector systems or simulated data from different MC programs. Finally, while the underlying framework may use the full power of C++ and object-oriented design, the parts of the code directly accessible by the user should not assume a particularly detailed knowledge of of these topics to facilitate the data analysis.

$\overline{\text{Off}}\underline{\text{line}}$ has been developed according to these requirements. $\overline{\text{Off}}\underline{\text{line}}$ is comprised of three principal components [Arg07]: a collection of processing modules, an event-based data model and a detector description. Most tasks can be factorized into sequences of self-contained processing steps. These steps are implemented in $\overline{\text{Off}}\underline{\text{line}}$ as modules, which inherit a common interface to facilitate exchanging modules and comparing different algorithms for the same processing step. The run-time control over the sequence of modules is afforded through a run controller, which invokes the single modules according to instructions given by an external

XML file. Additional XML files are used to store configurations and and parameters used by the single modules. The different XML files are accessed through a centralized configuration mechanism, which is invoked via a bootstrap file passed to the application at run time. The event-based data model serves as the principal backbone for the communication between the single modules. The overall structure of this model mimics the structure of the detector systems of the Pierre Auger Observatory, with further subdivisions. These subdivisions are, for example, the raw data collected by the detector systems, MC data obtained from simulations, reconstructed quantities and calibration information. During run time, the event is built up dynamically as needed. The single modules access the data contained in the event through a reference to the top of the hierarchy, which is passed to the module interface by the run controller. Finally, the detector description provides a unified interface through which a module can retrieve the detector configuration and the status for a given time as well as the atmospheric conditions. Like the event-based data model, the detector description is structured following the structure of the detector systems, and it provides a set of simple functions to access the data. These functions pass the data requests to a registry of managers, each capable of extracting a particular sort of information from a particular data source. For example, static information about the detector configuration such as the positions of the single SD stations or the orientation of the FD telescopes are stored in XML files, while data that changes with time, such as the atmospheric conditions or the calibration data of the single detectors, are stored in MySQL databases.

In the scope of this thesis, $\overline{\text{Off}}\underline{\text{line}}$ is used to simulate the responses of the FD and the SD to simulated air shower events generated with CORSIKA (see Sec. 5.1.1) and subsequently reconstruct the air shower event from these simulations. The XML file used to specify the sequence of modules can be found in App. A.2. In the following, the main modules involved in the simulation and reconstruction tasks as well as the main parameters used in the configuration of the modules are briefly discussed. For a full description of the modules and their implementation, see, e.g., [Pie14a] and the reference manuals for SD and FD simulation [Bah09, Ass11]. It should be noted that the detector description used in these simulations refers to an ideal detector, which means, for example, that all detector systems are working according to their specifications and the atmosphere is clear and free from aerosols.

The first two modules (`EventFileReaderOG` and `MCShowerCheckerOG`) manage the input files, in this case generated by CORSIKA. The simulated air shower is read in and checked for errors. The `EventGeneratorOG` module then places the simulated shower on the detector array. Different options can be chosen for this module, for example placing the shower according to a list of pre-defined coordinates or randomly over either the whole detector array or sections of it. For the simulations discussed here, a telescope-centered placement has been chosen: the impact point of the simulated air shower on ground is placed randomly inside the sector of a circle centered on telescope 3 of the FD site Los Leones. The radius R_{max} of this circle, which determines the maximum distance of the impact point to the FD telescope, is set energy-dependent according to the average trigger probability, as discussed in the following section. The opening angle of the circle is 30°, so that the full field of view of the FD telescope is covered. Each simulated air shower is used five times to generate an $\overline{\text{Off}}\underline{\text{line}}$ event in order to increase the available statistics. Using the same simulated shower five times is a compromise between reducing the computing resources and time needed and preserving the characteristics of the true distributions of the reconstructed variables.

The next batch of modules deals with the SD part of the simulation: First, in the `Cached-ShowerRegeneratorOG` module, the simulated shower is unthinned using a local sampling method, where a set of unweighted particles is extracted from the thinned particles stored in the CORSIKA output files and then randomly distributed over the walls of the local SD station [Bil00]. Then, the passage of the secondary particles from the air shower through the SD station is simulated with the `G4TankSimulatorOG` module. This module is based on the Geometry and Tracking 4 (Geant4) toolkit [Ago03, All06], which includes a wide range of physics processes relevant for the passage of particles through matter. Of particular interest are the production of Cherenkov light inside the SD station, as well as the transmission and attenuation of the Cherenkov light when traversing the tank and the reflection of the light at the inner liner. At the end of this simulation step, the number of photons impinging on the PMTs is extracted. The next batch of modules simulates the response of the PMTs and the station electronics to these signals. Finally, the SD triggers on station level (see Sec. 4.1) are simulated.

The next set of modules deals with the FD part of the simulation. The production of fluorescence and Cherenkov light within the atmosphere is not taken into account in the CORSIKA simulations. Hence, the `ShowerLightSimulatorKG` module is used, which simulates the emission of Cherenkov and fluorescence photons along the shower axis using the information from the CORSIKA output files. The attenuation of the fluorescence and Cherenkov photons in the atmosphere on the way from the emission point to the aperture of the FD telescopes is then simulated with the `LightAtDiaphragmSimulatorKG` module. Since an ideal detector configuration is used, the attenuation is calculated for ideal atmospheric conditions. The FD telescope itself is simulated in the `TelescopeSimulatorKG`, where the single photons arriving at the aperture are traced through the whole telescope optics. Reflection and refraction at the optical components is taken into account with the transmittance and reflectivity of all materials considered. Finally, the number of photons arriving at a single pixel is extracted. After a simulated background signal has been added, the response of the PMTs and the subsequent electronics is simulated, followed by a simulation of the local FD triggers (see Sec. 4.2.1). After the SD and FD simulations are finished—including the local triggers—the central trigger that is generated at the CDAS is simulated. In the case of a central trigger, the event is put together. The simulated event is then exported in the native $\overline{\text{Off}\underline{\text{line}}}$ file format, and it can then be used as an input for the subsequent reconstruction stage in the same way as the raw data obtained with the real detectors. The modules used for the hybrid reconstruction and the SD-only reconstruction of the simulated air shower events follow the single reconstruction steps that have been already discussed in Secs. 4.5.1 and 4.5.2, hence, they are are not described here any further.

The last module invoked in the module sequence used here is the `RecDataWriterNG` module, which exports the event in the Advanced Data Summary Tree (ADST) file format [Mar10]. The ADST file format is based on the ROOT framework [Bru97]. It contains all high-level quantities that have been reconstructed from both the SD and the FD, as well as the low-level raw data. In addition, the true MC quantities extracted from the CORSIKA output files are stored in an ADST file to allow for a quick and easy comparison of the reconstructed and the MC quantities.

Determining the Maximum Distance R_{max} for the Detector Simulations

As discussed before, the `EventGeneratorKG` module offers several possibilities to determine where the impact point of the simulated air shower is placed on ground. In the eye-centric mode, which is used for the simulations discussed here, the impact point is placed randomly inside the sector of a circle with the opening angle ϕ and the radius R_{max}, centered on a specific telescope. The parameter R_{max} thus determines the maximum distance of a simulated air shower to the observing FD telescope. Since the probability that an air shower will trigger the FD is depending on the energy of the shower and the distance to the FD telescope, the choice of R_{max} for a given simulation run directly influences the efficiency of the simulations, i.e. the ratio of triggered events to the total number of simulated events. In order to make the best use of the available computing resources, this efficiency should be kept high, i.e., regions, where the trigger probability is negligible should be excluded. On the other hand, a bias in the simulation sample has to be avoided, which can be introduced when, e.g., R_{max} is chosen too small and only showers very close to the FD telescope are simulated. Hence, the `EventGeneratorKG` module offers the possibility to set R_{max} automatically as a function of the energy of the simulated air shower that is used as input for the detector simulation [Ung04]:

$$R_{max}^{Offline}(x) = u_1 + x(u_2 + x\,u_3) \tag{5.2}$$

with

$$
\begin{aligned}
x &= \quad \log_{10}(E\,[\text{eV}]), \\
u_1 &= \quad 4.86267 \times 10^5\,\text{m}, \\
u_2 &= -6.72442 \times 10^4\,\text{m}, \\
u_3 &= \quad 2.31169 \times 10^3\,\text{m}.
\end{aligned}
$$

This empirical formula is based on a simulation study of the trigger efficiency of the FD [Pet04]. However, even when using this option, the efficiency of the simulations is about 28 % for photon primaries at 10^{18} eV, increasing to about 56 % at 10^{19} eV. Hence, an updated formula, which increases the simulation efficiency, is presented here.

First, a test sample of simulated air showers has been generated with CORSIKA, using the settings described in Sec. 5.1.1. 1000 Showers have been generated in each of the three energy bins $18.0 \leq \log_{10}(E\,[\text{eV}]) < 18.1$, $18.5 \leq \log_{10}(E\,[\text{eV}]) < 18.6$ and $18.9 \leq \log_{10}(E\,[\text{eV}]) < 19.0$, with photons and protons as primary particles. In total, 6000 showers have been simulated for this study . These showers have been used as input for the detector simulation with $\overline{Off}$line, using the settings documented in Sec. 5.1.2. In particular, each simulated shower has been used five times, and R_{max} has been set energy-dependent according to Eq. 5.2. In Fig. 5.1, the locations of the impact points for these simulations are shown, for both triggered and non-triggered events. As expected, the fraction of triggered events and the average distance of the triggered events to the FD telescope increases with energy. There are no large differences between photon and proton primaries. The new maximum distance R_{max}^{new} is now chosen such that 2 % of the triggered events have a distance larger than R_{max}^{new} (dashed lines in Fig. 5.1). At this limit, the fraction of triggered events increases to about 44 % for photon primaries at 10^{18} eV, and to 72 % at 10^{19} eV. Choosing 2 % as the limit represents a conservative choice. In Fig. 5.2, the R_{max}^{new} determined in this way are shown for the three energy bins and the two primary particle types. The difference between the R_{max}^{new} and $R_{max}^{Offline}$ is around 5 km, while the
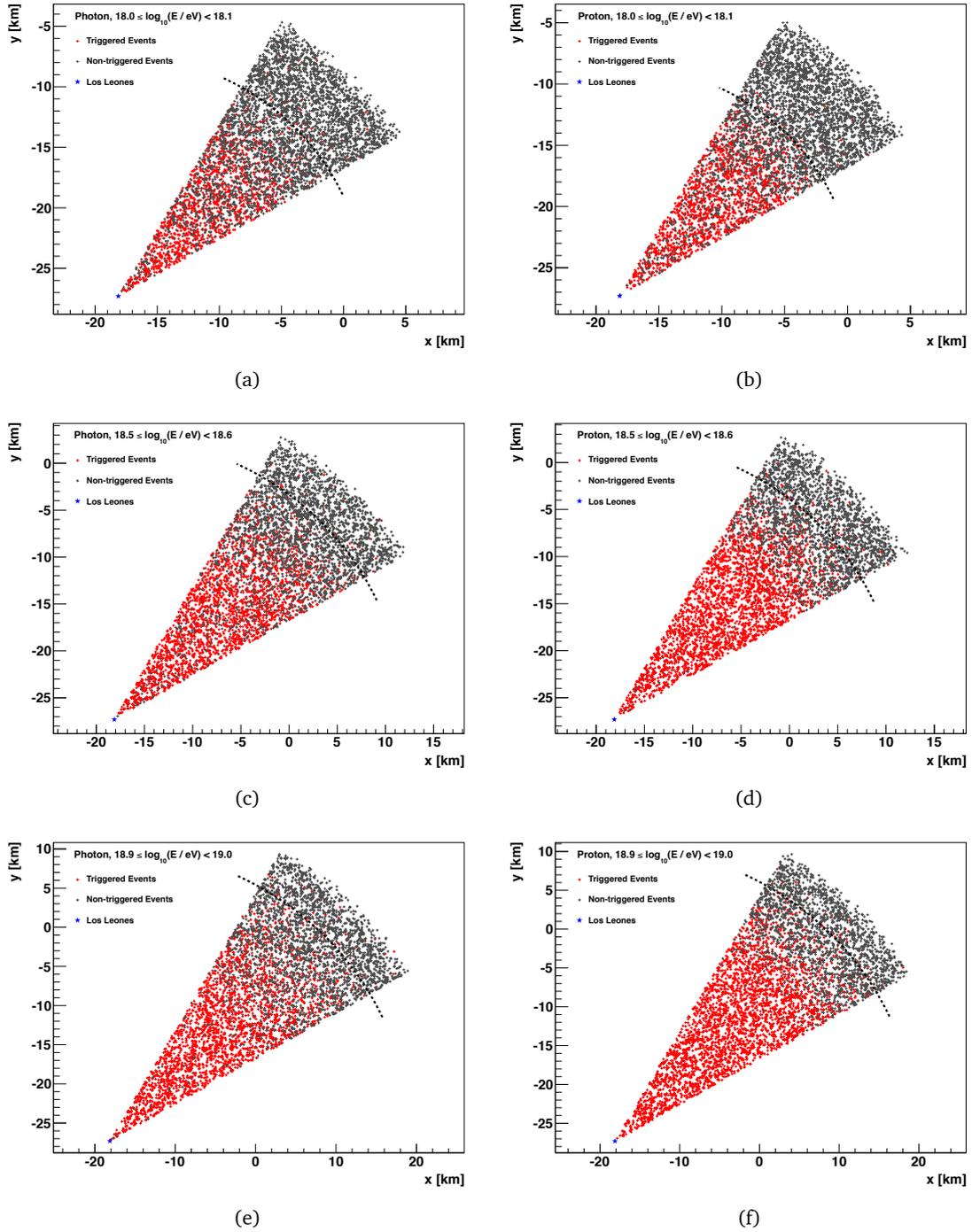
Figure 5.1.: Locations of the simulated impact points for the three energy bins and the two primary particle types. In each energy bin and for each primary particle type, 5000 events have been simulated. Triggered events are shown in red and non-triggered events in gray. The distance $R_{\text{max}}^{\text{new}}$, which has been chosen such that 2% of the triggered events have a distance larger than $R_{\text{max}}^{\text{new}}$, is indicated in each plot by a dashed line.
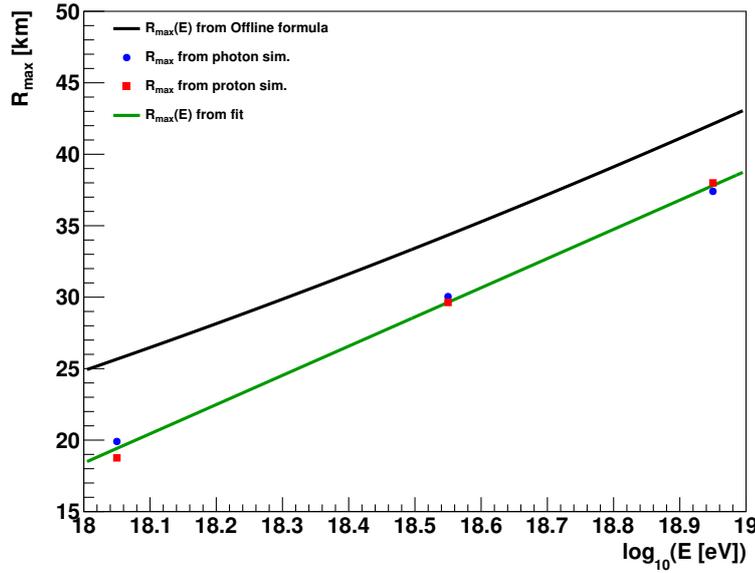
Figure 5.2.: R_{max} as a function of energy. The black line shows $R_{max}(E)$ as determined by the
  Off͟line formula. The R_{max} determined in the study presented here are shown as blue circles
  (photon primaries) and red squares (proton primaries). From a fit to these data points, an
  updated formula for $R_{max}(E)$ has been determined (green line).

difference between R_{max}^{new} for photon and proton primaries is around $1\,km$. In the next step,
a linear function is fitted to the data points—including both the data points for photon and
proton primaries, so that the resulting function gives an average over the two primary particle
types:

$$R_{max}^{new}(x) = v_1 + x\,v_2 \tag{5.3}$$

with

$$
\begin{aligned}
x &= \quad \log_{10}(E\,[\text{eV}]), \\
v_1 &= (-350 \pm 10)\,\text{m}, \\
v_2 &= (20.4 \pm 0.6)\,\text{m}.
\end{aligned}
$$

Eq. 5.3 is then used to determine the values of R_{max} that are set in the XML input files for
the Off͟line simulations. R_{max} is fixed for each energy bin, and the value that is set corresponds
to the center of the respective energy bin.

5.2. Data Sample Used in the Analysis

The data taking at the Pierre Auger Observatory is, to a large extent, an automated process.
The raw data is recorded automatically and distributed from the main computing system sit-
uated at the site of the observatory in Malargüe to several institutions all over the world,
from where they can be accessed by the members of the collaboration. An automated re-
construction of the air shower events from the raw data is provided by the Auger Observer
system [Dem14]. Reconstructed hybrid data from the beginning of data taking at the Pierre
Auger Observatory in 2004 up to the current data taking period is available through the Ob-
server system. It should be noted that the data taking started in 2004 with only a small part

of the SD array and the FD sites Los Leones and Coihueco. The rest of the SD stations and FD sites were gradually installed and included in the automatic data acquisition. However, not the full data period from 2004 up to now can be used in the analysis: for the period before December 2004, no absolute calibration for all of the then-installed FD telescopes is available. For the period after December 2012, the atmospheric databases are not yet officially released at the time of writing of this thesis. These databases have to be filled manually, and this process takes place on a non-regular basis. In the automated event reconstruction, events from these periods are still reconstructed, but this reconstruction is based not on measured atmospheric data and a measured absolute calibration of the telescope, but on tabulated data based on, e.g., monthly atmospheric models [Pie14a]. Hence, events from these periods are not used in the analysis presented in this thesis, effectively reducing the data period that is used to the time between the beginning of December 2004 and the end of December 2012. In total, 3 043 884 hybrid events have been recorded in this period.

5.3. Event Selection

The simulation samples and the data samples described in Secs. 5.1 and 5.2, respectively, are raw samples which contain all triggered events regardless of the quality of the recorded data. In particular, for the data sample this means that also noise events passing all FD trigger levels (see Sec. 4.2.1) are contained in the raw sample, as well as events from periods where the data acquisition cannot be considered reliable, for example due to a missing detector calibration or technical problems with the detector systems. In the simulation samples, an ideal detector configuration is used, where all detector systems are working to their specifications. Nevertheless, in both the raw simulation samples and in the raw data sample, also events where the recorded data are not sufficient to reliably reconstruct the underlying air showers are included. The selection criteria that are used to extract subsamples of well-reconstructed events from the raw samples are documented in detail in Sec. 5.3.1. The selected subsamples from the simulation samples and the data sample, in particular the selection efficiencies when applying the selection criteria to the samples, are described in Secs. 5.3.2 and 5.3.3, respectively.

5.3.1. Description of the Selection Criteria

The event selection can be split up into four levels: the preselection level, the geometry level, the profile level and the atmospheric level. At each level, several selection criteria (cuts) are introduced. At the preselection level, events from time periods with known problems with the detector systems and events where essential data is lacking—i.e. a fully reconstructed hybrid event that combines an FD measurement with SD data—are removed. At the geometry level, several cuts are imposed on the events to remove events without a reliable reconstruction of the event geometry. Similarly, additional cuts at the profile level remove events where the reconstruction of the longitudinal profile is not reliable. At the atmospheric level, two further cuts are applied to the data sample to remove events where the atmospheric conditions might have an influence on the measurement, i.e. clouds obscuring part of the longitudinal profile or too many aerosols in the lower layers of the atmosphere. It should be noted that the set of selection criteria used here is simular to the one used in [Set11].

Preselection Level

The selection criteria that are applied at the preselection level are listed in Tab. 5.2. First, events that were recorded by the HEAT telescopes, which are fully integrated in the automatic data acquisition, are removed (`eyeCut 1111`). In the next step, events from bad data taking periods for the FD are discarded (`badFDPeriodRejection true`). In particular, periods without an absolute calibration for the respective FD telescope are excluded, in addition to periods with known detector problems such as unstable PMTs or erroneous GPS systems. Certain problems with the detector electronics of the FD can be corrected offline, for example an erroneous clock on the FLT board. Events where this correction is missing are discarded (`good10MHzCorrection true`) Since the analysis presented in this thesis is based on both FD and SD data, also bad data taking periods for the SD have to be excluded (`badSDPeriodRejection true`). Between April and November 2009, a problem with the firmware of the communications system between the single detector stations and the CDAS caused frequent disruptions in the data acquisition, so that the data recorded in this period cannot be considered reliable and should thus be excluded from the analysis [Bon10]. In the next step, events that contain one or more pixels that either lack a succesful relative calibration—which is performed before and after each night of data taking—or are saturated during the recording of the event are removed (`badPixelRejection true` and `skipSaturated true`). Events recorded during the closing of the shutters in front of a FD telescope are also discarded (`MeanPixelRMS > 17`). The shutters can close automatically during data acquisition to prevent damages to the detector systems due to, e.g. high winds or nearby lightning. The variance of the signals measured by the single pixels due to background light impinging on the camera is constantly monitored. If the average variance drops below a given value, this indicates closing shutters, and an air shower that crosses the field of view of the respective FD telescope is only recorded partly, making a reliable reconstruction of the event from the recorded data difficult. In the next step, events where no atmospheric monitoring data for the time of the recorded event is available, in particular aerosol data, are excluded (`hasMieDatabase true`), since the atmospheric conditions are an important input for the reconstruction of air shower events from the raw FD data (see Secs. 4.2.3 and 4.5.1).

eyeCut	1111
badFDPeriodRejection	true
good10MHzCorrection	true
badSDPeriodRejection	true
badPixelRejection	true
skipSaturated	true
MeanPixelRMS	> 17
hasMieDatabase	true
FDReconstructionLevel	eHasEnergy
TriggeredSDstations	> 0

Table 5.2.: Selection criteria that are applied at the preselection level. Only the last two criteria are applied to both the simulation samples and the data sample, while the rest of the criteria are only applied to the data sample, since these criteria directly refer to the detector status at a given time, which is not taken into account in the simulations.

The selection criteria described so far apply to the data sample only, since they directly refer to the detector status at the time of the recording of an event which is not relevant for the simulated samples. The last two selection criteria at the preselection level, however, are applied to both the simulation samples and the data sample, since they remove events where necessary data is lacking. In particular, events where a reconstruction of the energy of the primary particle is not possible (`FDReconstructionLevel eHasEnergy`) and events without at least one triggered, non-saturated and non-accidental SD station within 3000 m from the shower axis (`TriggeredSDstations > 0`) are removed.

Geometry Level

After the preselection level, the samples contain only events that are fit for analysis from a technical point of view. At the geometry level, events where the reconstructed event geometry, which is the basis for the reconstruction of the properties of the air shower event, is unreliable are removed. The selection criteria applied at the geometry level are listed in Tab. 5.3. It is required that the reconstruction of the geometry is based on hybrid data, i.e. that the timing information from the SD is taken into account in the determination of the SDP (`NHybridStations > 0`) and that this timing information comes from an SD station that is not further away than 1500 m (`HybridStationAxisDistance ≤ 1500 m`) from the reconstructed shower axis. Furthermore, it is required that the difference between the expected arrival time of the signals at the respective SD station—extrapolated from the timing information from the FD—and the actual arrival time measured by the SD station is below 200 ns (`SD/FD offset < 200 ns`). Events where the χ^2/NDF of the fits to determine the SDP and the geometry of the shower within the SDP is too large are excluded (`SDP fit $\chi^2$/NDF ≤ 7` and `Time fit $\chi^2$/NDF ≤ 8`, cf. Sec. 4.5.1, Eqs. 4.1 and 4.5). Finally, inclined events with a zenith angle larger than 60° are removed (`ZenithAngle ≤ 60°`). For such air showers, the path through the atmosphere becomes so large that the flat Earth approximation cannot be used anymore (see Sec. 2.3). More importantly, most of the secondary particles from the electromagnetic component are absorbed in the atmosphere, leading to a different signature in the detectors compared to less inclined events and thus an erroneous reconstruction, since the reconstruction algorithms are optimized for less inclined air showers.

NHybridStations	> 0
HybridStationAxisDistance	$\leq 1500\,\text{m}$
SD/FD offset	$\leq 200\,\text{ns}$
SDP fit χ^2/NDF	≤ 7
Time fit χ^2/NDF	≤ 8
ZenithAngle	$\leq 60°$

Table 5.3.: Selection criteria that are applied at the geometry level.

Profile Level

At the profile level, events where the longitudinal shower profile has not been reliably reconstructed from the recorded data are removed from the samples. The selection criteria applied at the profile level are listed in Tab. 5.4. The reconstruction of the light profile from

the recorded signals has to take into account contributions from both fluorescence light and Cherenkov light (cf. Fig. 4.12(a)). To ensure a reliable reconstruction of the longitudinal shower profile from the collected light, the contribution from Cherenkov light is limited to 50 % (`CherenkovFraction` $\leq 50\%$). Next, it is required that the reconstructed profile is continuous without large gaps between the data points (`maxDepthHole` $\leq 20\%$). Such gaps can appear, for example, when an air shower is observed by more than one FD telescope, whose fields of view do not exactly overlap. In the next step of the event reconstruction, a Gaisser-Hillas function (Eq. 2.23) is fitted to the reconstructed profile. Events where the χ^2/NDF of the fit is larger than 1.9 are removed to ensure a reliable reconstruction of the longitudinal shower profile (`GH fit` χ^2/NDF ≤ 1.9). Since the analysis presented in this thesis also uses the position of the shower maximum, $X_{\max}$, as a parameter to distinguish air showers induced by UHE photons from those initiated by hadrons, it is required that $X_{\max}$ is observed within the geometrical field of view of the FD telescope that recorded the air shower event (`XmaxInFoV true`). This rather strict requirement is necessary to ensure a reliable determination of $X_{\max}$, although it introduces a bias against showers initiated by UHE photons, which develop deeper in the atmosphere than showers induced by hadrons and thus are more likely to extend below the field of view of the FD telescopes. To reduce the bias introduced by the previous criterion, an energy-dependent fiducial volume cut is applied (`FiducialVolumeCut true`). The fiducial volume cut limits the zenith angle range as well as the shower-telescope distance range to regions where the detector acceptance is similar for primary photons and primary protons [Abr07c, Abr09]:

$$\vartheta > \begin{cases} 35° + 10° \times (\log_{10}(E\,[\text{eV}]) - 19.0) & \text{for } \log_{10}(E\,[\text{eV}]) \leq 19.7, \\ 42° & \text{for } \log_{10}(E\,[\text{eV}]) > 19.7, \end{cases}$$
$$d < \begin{cases} 24\,\text{km} + 12\,\text{km} \times (\log_{10}(E\,[\text{eV}]) - 19.0) & \text{for } \log_{10}(E\,[\text{eV}]) \geq 19.0, \\ 24\,\text{km} + 6\,\text{km} \times (\log_{10}(E\,[\text{eV}]) - 19.0) & \text{for } \log_{10}(E\,[\text{eV}]) < 19.0, \end{cases} \tag{5.4}$$

where ϑ denotes the reconstructed zenith angle and d the distance of the shower core to the FD telescope which recorded the event. Finally, it is required that the calorimetric energy of the primary particle is determined with a relative uncertainty below 20 % (`RelCalEnergyUncertainty` $\leq 20\%$) to have a reliable energy estimate for each event in the samples. The calorimetric energy is used here instead of the fully reconstructed hybrid energy since the missing energy correction needed to determine the hybrid energy would introduce an additional source of uncertainty, which is not relevant to the analysis presented in this thesis at this stage.

CherenkovFraction	$\leq 50\%$
maxDepthHole	$\leq 20\%$
GH fit χ^2/NDF	≤ 1.9
XmaxInFoV	true
FiducialVolumeCut	true
RelCalEnergyUncertainty	$\leq 20\%$

Table 5.4.: Selection criteria that are applied at the profile level.

Atmospheric Level

In the last step of the event selection, events recorded during periods with poor observation conditions due to clouds or a large number of aerosols in the atmosphere are removed. The selection criteria at the atmospheric level, listed in Tab. 5.5, are only applied to data, since in the simulations, the atmosphere is assumed to be free from aerosols and clouds. In data, events where clouds cover more than 25 % of the sky, are removed (`CloudCoverage` $\leq$ 25 %). To determine the cloud coverage, the data from the IR cloud cameras mounted at each FD site are used, since these systems can detect clouds within the fields of view of the FD telescopes. If data from the cloud cameras are not available, data from the Lidar systems are used. If data from neither system are available, the event is discarded. Since the data from the cloud cameras are not yet fully integrated into the ADST files containing the event data, the databases containing the cloud camera data have to be accessed manually for each event. This process is time-consuming, it was decided to move the atmospheric level to the end of the event selection to keep the number of events for which the databases have to be accessed at a minimum. A second selection criterion is imposed on the Vertical Aerosol Optical Depth (VAOD): it is required that the VAOD, integrated from the ground up to a reference altitude of 3 km above ground level is less than 0.1 (`VAOD` $\leq$ `0.1`) to remove periods with poor viewing conditions, where a large number of aerosols accumulates in the lower layers of the atmosphere, leading to a stronger attenuation of the fluorescence light and thus more unreliable measurements.

CloudCoverage	$\leq$ 25 %
maxVAOD	$\leq$ 0.1

Table 5.5.: Selection criteria that are applied at the atmospheric level.

5.3.2. Application to the Simulation Samples

The event selection criteria described in the previous section are now applied to the simulation samples discussed in Sec. 5.1. The number of surviving events after each criterion and the corresponding selection efficiencies are listed in App. B, Tab. B.1. For the two main simulation samples, the selection efficiencies after the preselection, geometry and profile levels as well as the total selection efficiency are shown as a function of energy in Fig. 5.3. When comparing the photon to the proton sample, several differences can be noticed. At the preselection level, the selection efficiency for the photon sample increases with energy, from about 65 % at 10^{18} eV to about 100 % at 10^{19} eV. For the proton sample, the selection efficiency also increases, however, the efficiency at 10^{18} eV is already at about 85 %. This mainly reflects the fact that air showers induced by lower-energy particles have a smaller footprint on ground level and thus trigger, on average, less stations. Toward lower energies, the probability to trigger at least one station decreases, which is reflected in the selection efficiency at the preselection level. For photons, with their steeper lateral distribution of particles at ground level and their lack of muons, the probability to trigger a detector station is again smaller. At the geometry level, the selection efficiency is approximately constant, with little variation between 90 % and 100 % and no significant differences between the primary particles. This behaviour reflects the fact that the reconstruction of the geometry only relies on the timing information and not on the strength of the signals themselves, so that the reconstruction procedure is almost indepen-
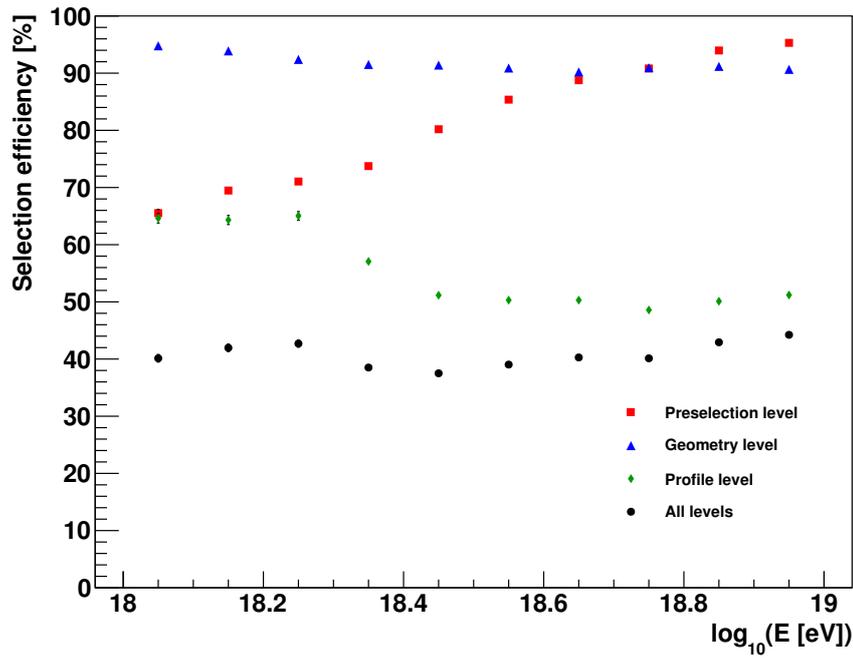
(a)
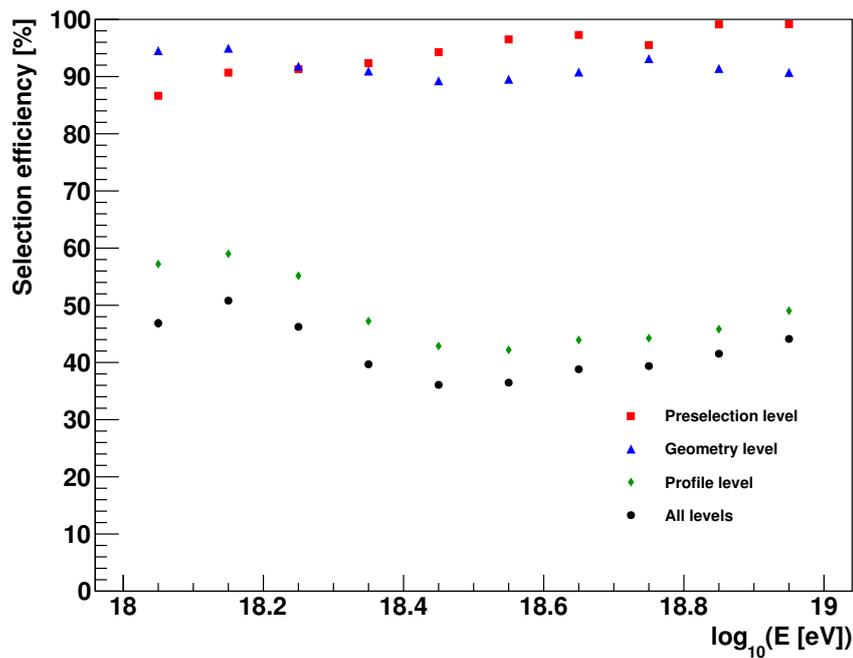
(b)

Figure 5.3.: Application of the event selection to the simulation samples: selection efficiencies after the different selection levels and after all levels for (a) the photon sample and (b) the proton sample. The selection efficiencies for the single selection criteria can be found in App. B. The uncertainties on the selection efficiencies have been determined according to [Pat04]. However, in most cases, the uncertainty is below 1 % and thus not visible in this scale.

dent of the differences between air showers initiated by different primary particle types. At the profile level, larger differences are visible. For the photon sample, the selection efficiency lies between about 65 % at lower energies and about 50 % at higher energies, while fot the proton sample, the selection efficiency is lower, between 60 % and 40 %. Without the fiducial volume cut, it would be expected that the selection efficiency significantly increases with energy for the proton sample: air showers induced by lower-energy protons develop higher in the atmosphere and are thus not fully observed by the FD telescopes—which were optimized for the energy range above 10^{19} eV. Such showers are mainly removed by the requirement that the shower maximum has to be observed within the field of view of the FD telescopes. At higher energies, the average X_{max} gets larger and thus the selection efficiency should increase. However, due to the fiducial volume cut, this increase is cancelled out. For the photon sample, the average X_{max} is larger than for protons so that, at higher energies, the showers are more likely to extend below the field of view of the FD telescopes, so that the selection efficiency of the field of view selection of events decreases. Also here the fiducial volume cut leads to a selection efficiency that is approximately constant with energy. After all selection levels, the selection efficiency for the photon sample is approximately constant around 40 %. This is also the case for the proton sample, albeit with larger deviations especially at lower energies. The lower selection efficiency at the profile level for the proton sample compared to the photon sample effectively cancels out the larger efficiency at the preselection level. This is mainly due to the fiducial volume cut, which has been introduced to remove the bias against photons.

5.3.3. Application to the Data Sample

In Tab. 5.6, the number of events after the single selection criteria have been applied to the data sample described in Sec. 5.2 and the corresponding selection efficiencies are listed. Overall, out of the 3 043 884 events in the raw sample, 84 794 events are selected, which corresponds to about 2.8 % of the total. Most of the events are removed at the preselection level: whereas the more technical selection criteria remove each between 0 % and 16 % of the events, the requirements of a fully reconstructed energy and at least one triggered SD station remove a larger number of events (selection efficiencies of 44.1 % and 69.6 %, respectively). The removed events are mainly noise events due to, e.g., meteorological events like sheet lightning that pass the FD trigger levels and low-energy events where no SD station is triggered. At the geometry level, the selection efficiencies are comparable to the selection efficiencies for the simulation samples, with the exception of the zenith angle selection, where the selection efficiency for the data sample is smaller because the simulations are limited to a maximum zenith angle of 65°, while in the data sample, also events with a larger zenith angle are included. At the profile level, most events are removed due to the requirements that the shower maximum has to be observed within the field of view. It should be noted that the selection efficiency for these criteria is smaller for the data sample than for the simulation samples because no energy selection has been applied. Thus the data sample is dominated by air showers initiated by primary particles with smaller energies—which develop higher in the atmosphere and therefore outside of the field of view of the FD telescopes—due to the steeply falling energy spectrum. About 25 % of the events that pass this requirement are then removed by the fiducial volume cut. Finally, for the 159 093 events that survive the profile level, the cloud coverage is checked using the atmospheric databases. For about half of these events, information from the cloud cameras is available. For another 43 000 events, the cloud information obtained using the Lidar systems can be used. For the rest of the events, nei-

	N	ε
Raw sample	3 043 882	-
eyeCut 1111	2 559 321	84.1 %
badFDPeriodRejection true	2 274 451	88.9 %
good10MHzCorrection true	2 253 531	99.1 %
badSDPeriodRejection true	2 148 950	95.4 %
badPixelRejection true	2 027 192	94.3 %
skipSaturated true	2 020 708	99.7 %
MeanPixelRMS > 17	1 865 774	92.3 %
hasMieDatabase true	1 633 864	87.6 %
FDReconstructionLevel eHasEnergy	720 248	44.1 %
TriggeredSDstations > 0	501 620	69.6 %
Total (preselection level)	501 620	16.5 %
NHybridStations > 0	498 816	99.4 %
HybridStationAxisDistance $\leq$ 1500 m	495 229	99.3 %
SD/FD offset $\leq$ 200 ns	495 216	100.0 %
SDP fit χ^2/NDF $\leq$ 7	493 534	99.7 %
Time fit χ^2/NDF $\leq$ 8	492 827	99.9 %
ZenithAngle $\leq$ 60°	444 411	90.2 %
Total (geometry level)	444 411	88.6 %
CherenkovFraction $\leq$ 50 %	387 485	87.2 %
maxDepthHole $\leq$ 20 %	375 087	96.8 %
GH fit χ^2/NDF $\leq$ 1.9	366 568	97.7 %
XmaxInFoV true	235 233	64.2 %
FiducialVolumeCut true	177 725	75.6 %
RelCalEnergyUncertainty $\leq$ 20 %	159 093	89.5 %
Total (profile level)	159 093	35.8 %
CloudCoverage $\leq$ 25 %	88 336	55.5 %
maxVAOD $\leq$ 0.1	84 794	96.0 %
Total (atmospheric level)	84 794	53.3 %
Total (all levels)	84 794	2.8 %

Table 5.6.: Application of the event selection criteria to the data sample described in Sec. 5.2. Listed are the number of surviving events N after the single criteria and the corresponding selection efficiency ε relative to the number of events before the criterion was applied.

ther data from the cloud cameras nor from the Lidar systems is available and the events are discarded. In total, the requirement that the measured cloud coverage must be below 25 % removes about 45 % of the events surviving the profile level. The additional requirement that the reference VAOD must be less than 0.1 removes an additional 4 %, which indicates that, as expected, the viewing conditions in Malargüe are good for most of the time. For a more detailed study of the effects of the individual selection criteria on the data sample, see [Tho14].

6

Simulation Study of Observables for the Search for UHE Photons

The main challenge in the search for UHE photons lies in distinguishing air shower events that are induced by photons from the background events which are initiated by protons or nuclei. In this chapter, a detailed simulation study of the two main observables used in the analysis presented in this thesis is discussed. The first observable is the position of the shower maximum, X_{max}, which can be measured directly with the FD. The observable X_{max} is complemented by the SD-related observable F_γ, which is based on the fit of a photon-optimized LDF to the signals that are measured in the individual detector stations of the SD. By combining X_{max} and F_γ, the hybrid concept of the Pierre Auger Observatory can be used to its full extent. The combination of both observables is achieved through a Fisher discriminant analysis, which fully exploits the separation power contained in both observables. The Fisher discriminant is then used as an additional observable.

This chapter is structured as follows: first, the observable X_{max} is discussed in Sec. 6.1. In particular, the separation power of this observable between photon and proton primaries is examined. In Sec. 6.2, the observable F_γ is introduced. The combination of both observables in a Fisher discriminant analysis is presented in Sec. 6.3. Using the simulation samples described in Chap. 5, the analysis is optimized for photon/proton separation. Several further studies are described in Sec. 6.4.

6.1. X_{max} as a Discriminating Observable

The atmospheric depth X_{max} at which the longitudinal development of a shower reaches its maximum in terms of the number of secondary particles is an important observable for composition studies at the Pierre Auger Observatory. X_{max} has been used not only for studies of the elemental composition, see e.g. [Aab14a], but also for the determination of upper limits on the diffuse flux of UHE photons, see e.g. [Abr07c, Abr09].

As discussed before, X_{max} is directly correlated with the type of the primary particle initiating the extensive air shower in the atmosphere (see e.g. Eq. 2.26 or Fig. 3.3). In particular, air showers initiated by UHE photons develop, on average, deeper in the atmosphere than those induced by protons or nuclei of the same energy. Furthermore, X_{max} can be measured directly using the air-fluorescence technique. Since the analysis presented in this thesis uses well-reconstructed hybrid events, the observable X_{max}—and thus information on the primary particle type—is available for each event in the data set.

In the following, the separation power of the observable X_{max}, i.e. the capability of the observable to distinguish between air showers induced by UHE photons and those initiated by protons, is examined using the simulation samples described in Sec. 5.1. In Fig. 6.1, the distributions of X_{max} for the photon and the proton samples are shown exemplarily for the first energy bin. The corresponding distributions for the other energy bins are shown in App. C, Fig. C.1. For energies between $10^{18.0}$ eV and $10^{18.1}$ eV, the average X_{max} for photons is about $120 \, \text{g} \, \text{cm}^{-2}$ larger than that for protons. However, the significant spread of the distributions—about $70 \, \text{g} \, \text{cm}^{-2}$ in terms of standard deviation—and the noticeable overlap makes it impossible to distinguish clearly between the different primary particle types on an event-by-event basis in the region between the two distributions. The overlap becomes smaller with increasing energy, hence the overall separation power increases. However, it does not vanish completely. Also in the last energy bin, there is still a region of ambiguity between the two distributions.

In order to quantify the separation power of an observable, the merit factor η is used:

$$\eta = \frac{\left|\mu_\gamma - \mu_p\right|}{\sqrt{\sigma_\gamma^2 + \sigma_p^2}}, \tag{6.1}$$

where μ_γ and μ_p are the mean values of the distributions of the observable for the photon and the proton sample, respectively, while σ_γ and σ_p are the corresponding standard deviations. Qualitatively, the merit factor can be seen as a measure of the degree of overlap between the two distributions. For two distributions that overlap completely, either because their mean values are almost identical or because their widths are very large, the merit factor is close to 0. A large merit factor indicates a small overlap and thus a large separation between the two
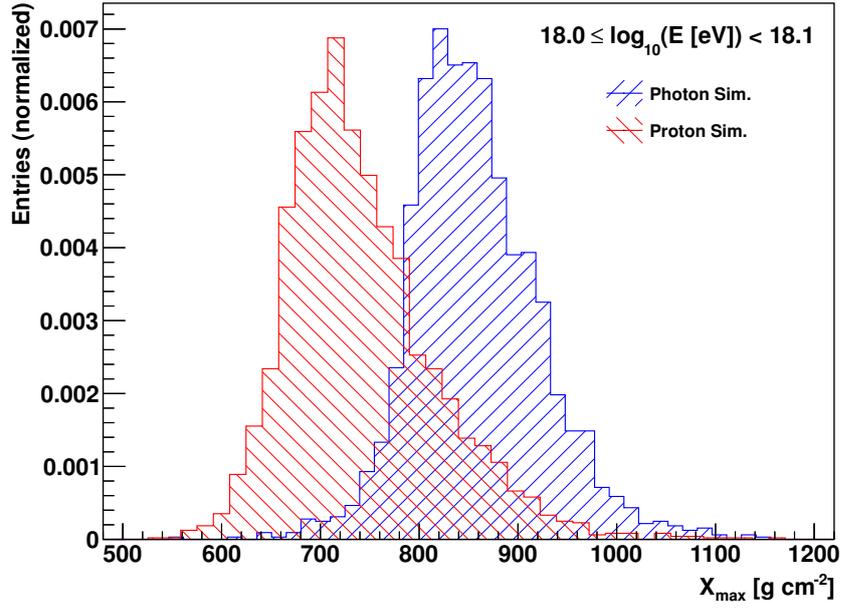
Figure 6.1.: Distributions of X_{max} in the first energy bin ($18.0 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.1$) for the photon sample, shown in blue, and the proton sample, shown in red.

distributions. Usually, an observable is considered "good", if the merit factor is larger than 1. However, it should be noted that in the limit of very asymmetric distributions the merit factor alone is not an optimal measure of the separation power, because it takes into account only the mean values and the widths of the distributions, but not their shapes. Hence, the merit factor is complemented by the background rejection ϱ at a given signal efficiency ε, which does take the shape of the distributions into account. For an observable x with $\mu_\gamma > \mu_p$ and for a given cut value x_{cut}, ϱ and ε are defined as

$$\varrho = \frac{N_p(x < x_{\mathrm{cut}})}{N_{p,\,\mathrm{total}}},$$

$$\varepsilon = \frac{N_\gamma(x \geq x_{\mathrm{cut}})}{N_{\gamma,\,\mathrm{total}}}. \tag{6.2}$$

In this definition, N_p (N_γ) denote the number of events in the proton (photon) sample below (above) the cut value, while $N_{p,\,\mathrm{total}}$ and $N_{\gamma,\,\mathrm{total}}$ describe the total number of events in the samples. For the case $\mu_\gamma < \mu_p$, the larger-than and less-than signs in Eq. 6.2 are reversed. Of particular interest for studying the separation power is the background rejection as a function of the signal efficiency, $\varrho(\varepsilon)$. For each ε, first the corresponding cut value x_{cut} is determined through a scan of the photon distribution. The corresponding cut value is then used to infer the background rejection for this efficiency according to Eq. 6.2. The background rejection at a signal efficiency of 50 % is used as a reference for the separation power. The larger the separation power of the observable, the larger is the background rejection at 50 % signal efficiency.

For the observable X_{max}, the main properties of the distributions—mean values and widths in terms of standard deviation—for the photon and the proton samples for all energy bins are listed in Tab. 6.1. The mean values of the X_{max} distributions increase with energy, for both

Energy Bin	Photon Sample		Proton Sample			
	$\mu\,[\mathrm{g\,cm^{-2}}]$	$\sigma\,[\mathrm{g\,cm^{-2}}]$	$\mu\,[\mathrm{g\,cm^{-2}}]$	$\sigma\,[\mathrm{g\,cm^{-2}}]$	η	$\varrho(\varepsilon=50\,\%)$
$18.0 \le \log_{10}(E\,[\mathrm{eV}]) < 18.1$	857.12	65.78	741.05	75.28	1.16	91.26 %
$18.1 \le \log_{10}(E\,[\mathrm{eV}]) < 18.2$	872.75	69.96	746.33	70.82	1.27	93.91 %
$18.2 \le \log_{10}(E\,[\mathrm{eV}]) < 18.3$	880.61	66.42	746.65	68.90	1.40	94.73 %
$18.3 \le \log_{10}(E\,[\mathrm{eV}]) < 18.4$	893.97	63.01	754.47	65.25	1.54	96.11 %
$18.4 \le \log_{10}(E\,[\mathrm{eV}]) < 18.5$	908.32	68.97	764.90	71.14	1.45	94.32 %
$18.5 \le \log_{10}(E\,[\mathrm{eV}]) < 18.6$	915.37	65.75	765.51	64.05	1.63	96.41 %
$18.6 \le \log_{10}(E\,[\mathrm{eV}]) < 18.7$	926.26	67.01	776.47	70.21	1.54	95.61 %
$18.7 \le \log_{10}(E\,[\mathrm{eV}]) < 18.8$	938.71	68.27	780.91	64.82	1.68	96.70 %
$18.8 \le \log_{10}(E\,[\mathrm{eV}]) < 18.9$	951.42	65.98	784.20	65.17	1.80	97.17 %
$18.9 \le \log_{10}(E\,[\mathrm{eV}]) < 19.0$	961.72	67.05	787.56	62.87	1.89	97.87 %

Table 6.1.: Properties of the X_{max} distributions for the photon and proton samples and the separation power of the observable. Listed are the mean values μ of the distributions and their widths in terms of standard deviation σ. For each energy bin, the merit factor η has been calculated using these parameters. In addition, the background rejection ϱ at a signal efficiency of $\varepsilon = 50\,\%$ was derived from the distributions (cf. App. C, Fig. C.2).
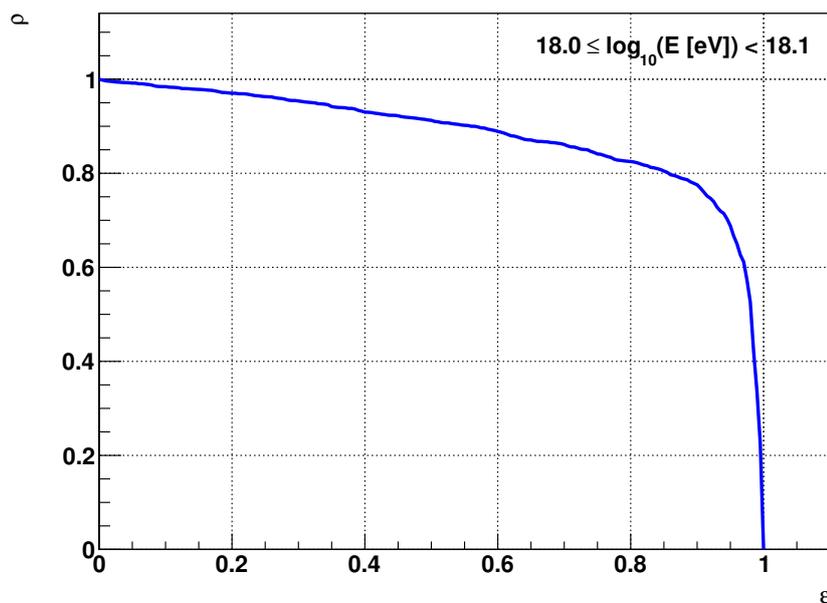
Figure 6.2.: Background rejection ϱ as a function of the signal efficiency ε for the observable X_{max} in the first energy bin ($18.0 \le \log_{10}(E\,[\mathrm{eV}]) < 18.1$).

the photon and the proton sample. Due to the LPM effect, the increase for the photon sample is larger than for the proton sample. The widths of the distributions are approximately constant around $70\,\mathrm{g\,cm^{-2}}$. Using these parameters, the merit factor η was calculated for each energy bin according to Eq. 6.1 and listed in Tab. 6.1. Since the difference between the mean values for the photon and the proton distributions increases with energy and the widths of the distributions remain approximately constant, also η increases with energy. In addition, the background rejection ϱ as a function of the signal efficiency ε was determined for each energy bin. In Fig. 6.2, $\varrho(\epsilon)$ is shown exemplarily for the first energy bin. The corresponding

(a)

(b)
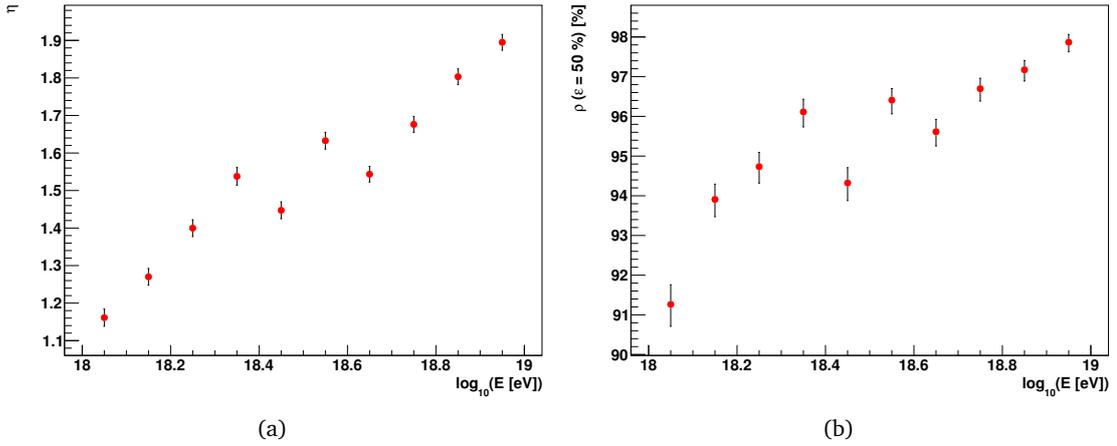
Figure 6.3.: Measures of the separation power of the observable X_{max} as a function of energy. (a) Merit factor η. (b) Background rejection ϱ at a signal efficiency of $\varepsilon = 50\%$. The uncertainties of the individual data points have been estimated using error propagation in the case of η and according to [Pat04] in the case of $\varrho(\varepsilon = 50\%)$.

distributions for the other energy bins are shown in App. C, Fig. C.2. From these distributions, the background rejection at a signal efficiency of 50% was extracted in each energy bin (cf. Tab. 6.1). Both measures of the separation power are plotted as a function of energy in Fig. 6.3. The uncertainties of the individual data points have been estimated in the case of η using the standard errors of the mean values and the standard deviations of the distributions according to error propagation, and in the case of $\varrho(\varepsilon = 50\%)$ according to [Pat04]. Both η and $\varrho(\varepsilon = 50\%)$ increase almost linearly with energy—η from 1.16 around 1 EeV to 1.89 around 10 EeV and $\varrho(\varepsilon = 50\%)$ from about 91% around 1 EeV to about 98% around 10 EeV. The deviations from a linear behaviour can be mainly attributed to statistical fluctuations. Overall, these numbers underline that X_{max} is a good observable suitable for the search for UHE photons.

6.2. A new SD-related Observable: F_γ

The observable X_{max}, as discussed in the previous section, is a pure FD observable. In an analysis based solely on X_{max}, a large part of the data collected at the Pierre Auger Observatory—the data from the SD—is therefore not fully taken into account. However, this data contains valuable information on the type of the primary particle initiating the recorded extensive air shower due to the differences in the lateral development of air showers induced by different primary particle types. The differences in the lateral development are mainly related to the differences in the position of the shower maximum and the number of muons, N_μ (cf. Sec. 3.3.2). In contrast to X_{max}, N_μ cannot be measured directly with the current detector systems of the Pierre Auger Observatory. Hence, observables that indirectly exploit the differences in the number of muons have to be used for composition studies using SD data. Such observables include, for example, the reconstructed radius of curvature of the shower front, R_c, or the risetime of the signals measured in the individual SD stations. Both of these observables make use of the spatial and temporal distribution of the measured SD signals, and both have been used previously for determining upper limits on the incoming flux of UHE photons

using SD data [Abr08b]. However, both observables require a full reconstruction of the air shower event from SD data and are thus not suitable for the energy range between 10^{18} eV and 10^{19} eV, where the majority of the events contain only very few triggered SD stations. In this energy range, other SD-related observables have to be used, such as S_b [Ros11], which makes use of the differences in the shape of the LDFs of air showers initiated by different primary particle types, or the "shape parameter" [Cro03, Aab14c], which exploits the different time structures of the signals recorded in an SD station to distinguish different primary particle types. In order to avoid confusion with the LDF shape parameter β, the "shape parameter" is referred to as ζ in the following. Both S_b and ζ have been used previously in conjunction with X_{max} to determine upper limits on the incoming flux of photons in the EeV range [Set11, Aab14c]. In the analysis presented in this thesis, the new observable F_γ [You11] is used for the first time to search for UHE photons in data. Similar to S_b, F_γ exploits the differences in the LDFs of air showers induced by UHE photons and protons or nuclei. The observable is based on a photon-optimized fit of an LDF to the signals recorded in the individual SD stations. The definition of F_γ is discussed in Sec. 6.2.1. A novel parameterization of the LDF shape parameter β, on which the photon-optimized LDF fit is based, is presented in Sec. 6.2.2, while the LDF fit itself is discussed in detail in Sec. 6.2.3. The performance of this observable in distinguishing air showers induced by UHE photons from those induced by protons or nuclei is examined in Sec. 6.2.4. The separation power of F_γ is then compared to those of S_b and ζ (Sec. 6.2.5). Finally, in Sec. 6.2.6, additional event selection criteria based on F_γ are introduced to further improve the performance of the observable.

6.2.1. Definition of F$_\gamma$

The observable F_γ, as it is used in the context of this thesis, is defined as the ratio of two S_{1000}-like parameters:

$$F_\gamma = \frac{S_{1000|\gamma}}{S_{1000|\mathrm{Hybrid}}}. \tag{6.3}$$

Both $S_{1000|\gamma}$ and $S_{1000|\mathrm{Hybrid}}$ are related to the regular S_{1000} parameter that is used as an energy estimator in the standard reconstruction of air shower events from SD data (see Sec. 4.5.2). The two parameters $S_{1000|\gamma}$ and $S_{1000|\mathrm{Hybrid}}$ are discussed in more detail in the following.

The parameter $S_{1000|\gamma}$ is determined through a fit of an LDF to the signals measured in the individual SD stations. As in the standard SD reconstruction, a modified NKG function is used for the functional form of the LDF (see Eq. 4.9), however, with two important distinctions. Firstly, only one LDF shape parameter (β) is used to determine the steepness of the LDF. Furthermore, β is not a free parameter, but it is parameterized as a function of E_{Hybrid} and θ and fixed in the fit. The parameterization was derived from MC simulations with photons as primary particles, hence, the LDF fit is "photon-optimized". Details on the parameterization are given in Sec. 6.2.2. The LDF that is used in the photon-optimized fit is the following:

$$S(r) = S_{1000|\gamma} \left(\frac{r}{1000\,\mathrm{m}} \right)^{\beta(E_{\mathrm{Hybrid}}, \theta)} \left(\frac{r + 700\,\mathrm{m}}{1700\,\mathrm{m}} \right)^{\beta(E_{\mathrm{Hybrid}}, \theta)}. \tag{6.4}$$

As in the standard SD reconstruction, the LDF is fitted using the maximum-likelihood method for parameter estimation [Cow98]. The likelihood function that is used in the fit is discussed in Sec. 6.2.3, together with details on the fit. In the application of the photon-optimized fit,

the most important difference between this fit and the standard fit in the SD reconstruction is that only one parameter ($S_{1000|\gamma}$) has to be determined through the fit, so that the fit can be applied also to hybrid events where only one SD station was triggered. However, it should be noted that $S_{1000|\gamma}$ cannot be used as an energy estimator like S_{1000} in the standard SD reconstruction since the reconstructed hybrid energy is needed to determine $S_{1000|\gamma}$. Hence, $S_{1000|\gamma}$ denotes the signal that is expected at a distance of 1000 m from the shower axis for a typical photon shower, with given energy and zenith angle, under consideration of the signals measured in the individual SD stations.

The second parameter in Eq. 6.3, $S_{1000|\text{Hybrid}}$, is calculated from the reconstructed energy and zenith angle by inverting the functions used in the standard SD reconstruction for the energy calibration, i.e. for determining the energy of an air shower event from a reconstructed S_{1000} value. The calculation is a two-step process: first, the zenith-angle independent quantity $S_{38|\text{Hybrid}}$ is calculated from the reconstructed energy (E_{Hybrid}) by inverting Eq. 4.12:

$$S_{38|\text{Hybrid}} \, [\text{VEM}] = \sqrt[B]{\frac{E_{\text{Hybrid}}}{A}}. \tag{6.5}$$

In the second step, the dependency on the reconstructed zenith angle θ is introduced through the CIC function (Eq. 4.11) to obtain $S_{1000|\text{Hybrid}}$:

$$S_{1000|\text{Hybrid}} = \text{CIC}(\theta) \times S_{38|\text{Hybrid}}. \tag{6.6}$$

By construction, $S_{1000|\text{Hybrid}}$ thus denotes the average S_{1000} that is expected for an extensive air shower with a given primary energy and a given zenith angle. It should be noted, however, that this refers to the average S_{1000} for showers that are induced by protons and nuclei, as these showers dominate the data set collected by the Pierre Auger Observatory, from which the parameters A and B and the CIC function are determined.

By taking the ratio of the two S_{1000}-like parameters discussed above, the energy dependence introduced by $S_{1000|\gamma}$ is removed and the F_γ parameter can be used as an energy-independent observable for distinguishing air showers induced by photons from those initiated by protons or nuclei. The separation power of the observable will be examined in Sec. 6.2.4.

6.2.2. Determining a Parameterization of the LDF Shape Parameter β for Photons in the EeV Range

The LDF shape parameter β for air showers initiated by photons needs to be parameterized for use in the photon-optimized fit. The parameterization was derived from MC simulations in the energy range $18.0 \leq \log_{10}(E\,[\text{eV}]) < 19.0$. It was already presented in an internal note of the Pierre Auger Collaboration [Nie13] and is discussed in detail in this section.

The MC simulations used to derive the parameterization were performed with CORSIKA, version 6.720, and the Auger Offline Software Framework, version 2.7.8. The settings used to simulate the extensive air showers are the same as described in Sec. 5.1.1, with the exception of the hadronic interaction models. For the simulations described in this section, QGSJETII-03 and FLUKA 2011.b4 are used as the hadronic interaction models at high energies and at low energies, respectively. Two different sets of air showers were simulated with photons as primary particles, one in the energy range $18.0 \leq \log_{10}(E\,[\text{eV}]) < 18.5$ and one in the

subsequent energy range $18.5 \leq \log_{10}(E \, [\text{eV}]) < 19.0$. For the $\overline{\text{Offline}}$ part of the simulations, the settings described in Sec. 5.1.2 are used, albeit with one important modification: instead of the regular SD array with a spacing of 1500 m between the individual SD stations, a much denser array with a spacing of 433 m was used in order to have more triggered stations per event and thus allow for a more precise LDF fit. The additional stations are positioned in between the regular stations, so that one can at a later time easily restrict the analysis to the stations from the regular array, e.g. for cross checks of the parameterization. Each simulated air shower is used five times, with the impact points of the simulated showers randomly distributed over the whole SD array. After the reconstruction stage, the two data sets contain 8375 events ($18.0 \leq \log_{10}(E \, [\text{eV}]) < 18.5$) and 4950 events ($18.5 \leq \log_{10}(E \, [\text{eV}]) < 19.0$), respectively.

Deriving the Parameterization

The process of deriving a parameterization of the LDF shape parameter β, as a function of energy E and zenith angle θ from the simulations described previously is divided into several stages, which are discussed in detail in this section.

It should be noted that for the energy E and zenith angle θ of an event, the MC values are used here instead of the values reconstructed by the $\overline{\text{Offline}}$ algorithms in order to decouple the β parameterization derived here from reconstruction effects and the resulting fluctuations. When using this parameterization for reconstructed real events, it should therefore be ensured that the reconstruction is good enough by applying adequate quality cuts.
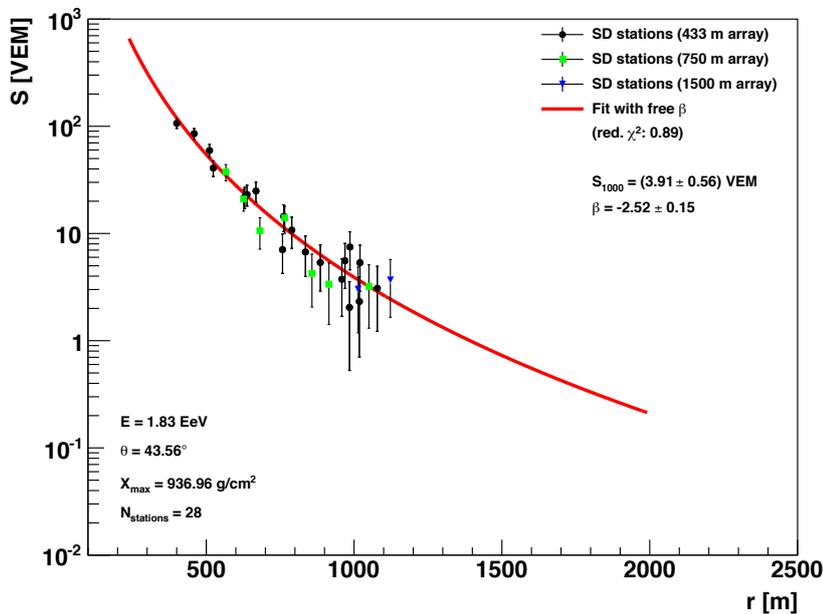
Figure 6.4.: An example of the individual event-by-event LDF fits to the simulated air shower events. A modified NKG function according to Eq. 4.9 with $\gamma = \beta$ has been used, with the two remaining parameters S_{1000} and β unconstrained. 28 SD stations are taken into account in the fit shown here. The quoted energy, zenith angle and X_{max} are the MC values.

In the first stage, an LDF following Eq. 4.9, with $\gamma = \beta$, is fitted to each event individually. In this fit, the two parameters β and S_{1000} are not constrained and only stations that show a non-zero signal are used, i.e. non-triggering stations are ignored. In any case, since a dense array with a spacing of 433 m between the individual detector stations is used, the contribution of non-triggering stations would be negligible. In addition, limits on the minimum and maximum distance of the SD stations to the shower axis are imposed: only SD stations in the range $400\,\text{m} \leq r \leq 4000\,\text{m}$ are used to avoid saturated stations or stations with a very high, upward fluctuating signal that would otherwise dominate the LDF fit. In Fig. 6.4, one example event is shown, including the fitted LDF.

In the next step of this study, both energy bins are divided into five sub-bins each. The width of the sub-bins is 0.1 in $\log_{10}(E\,[\text{eV}])$. The distributions of β as a function of $\sec(\theta) = \frac{1}{\cos(\theta)}$ in each energy sub-bin are then investigated. At this stage, several selection criteria are applied to ensure a good quality of the LDF fit: at least five SD stations have to be used in the LDF fit, none of which showing a saturated signal neither in the high-gain nor in the low-gain channels. In addition, only those events are used where the relative uncertainty of the fitted β is less than 50 %. This selection may lead to a bias in the reduced data set especially at higher zenith angles at low energies, because only upward fluctuating events pass the event selection. To reduce the impact of this bias, the zenith angle range used in the following stages of the study is limited to only those regions where the effect is not significant. Specifically, a cut on $\sec(\theta)$ is introduced at 1.8 for energies $18.0 \leq \log_{10}(E\,[\text{eV}]) < 18.2$, at 1.9 for energies $18.3 \leq \log_{10}(E\,[\text{eV}]) < 18.4$, at 2.0 for energies $18.4 \leq \log_{10}(E\,[\text{eV}]) < 18.5$ and at 2.1 at energies above $\log_{10}(E\,[\text{eV}]) = 18.5$. After the event selection, the final data sets consist of 6788 events ($18.0 \leq \log_{10}(E\,[\text{eV}]) < 18.5$) and 4255 events ($18.5 \leq \log_{10}(E\,[\text{eV}]) < 19.0$), respectively.

As an example, the resulting distribution of $\beta(\sec(\theta))$ for the first energy sub-bin ($18.0 \leq \log_{10}(E\,[\text{eV}]) < 18.1$) is shown in Fig. 6.5. The distributions for the other sub-bins are shown in App. C, Fig. C.3. To model the zenith angle dependence, a simple cubic function is used:

$$\beta(\theta) = a_0 + a_1 \left(\sec(\theta) - 1\right)^3, \tag{6.7}$$

with the parameters a_0 and a_1 describing the y-intercept and the steepness of the function, respectively.

In Fig. 6.6, the fitted functions for the different energy sub-bins are shown together in one plot. A clear ordering of the functions in energy is visible, with the function representing the lowest energies being the steepest and the function representing the highest energies being the flattest. This motivates looking at the parameters a_0 and a_1 as functions of energy. This is shown in Fig. 6.7. To model the energy dependence of the parameters a_0 and a_1, and thus the energy dependence of β, simple functions have been chosen to minimize the final number of parameters: a first-order polynomial for $a_0(E)$ and a second-order polynomial for $a_1(E)$.

Putting everything together, the final parameterization of the LDF shape parameter β as a
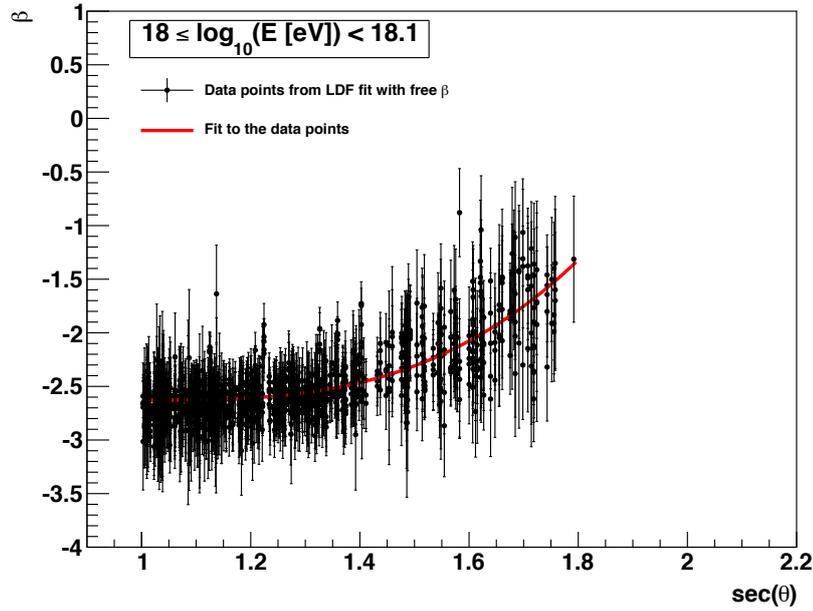
Figure 6.5.: Distribution of the fitted β as a function of $\sec(\theta)$ in the energy range $18.0 \leq \log_{10}(E\,[\text{eV}]) < 18.1$. To reduce the impact of a possible bias because of the event selection applied here, a cut on $\sec(\theta)$ has been introduced. A simple cubic function according to Eq. 6.7 has been fitted to the data points, the result of the fit is shown as a red line.
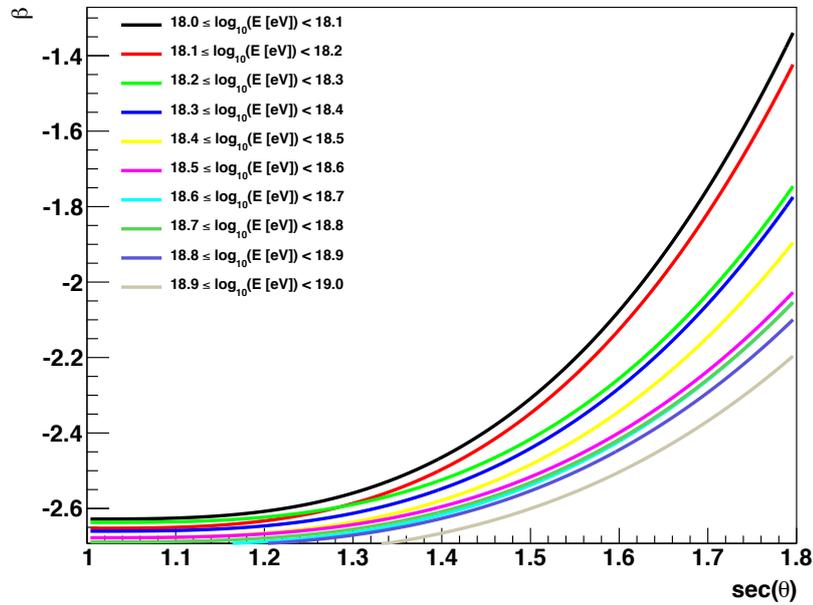
Figure 6.6.: All functions derived from the fits of Eq. 6.7 in the individual energy sub-bins shown in one plot.

(a)   (b)
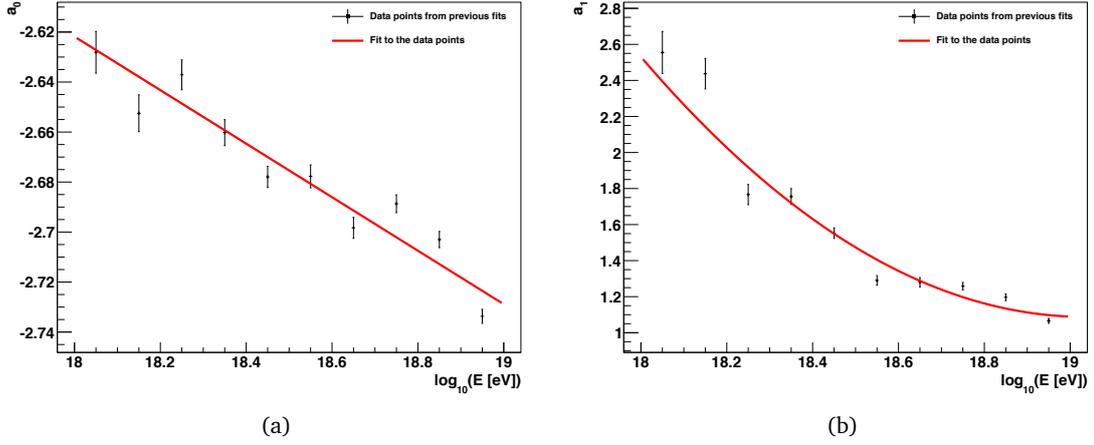
Figure 6.7.: Energy dependence of the parameters a_0 (a) and a_1 (b) from Eq. 6.7. To model the energy dependence, first-order and second-order polynomial functions have been chosen for a_0 and a_1, respectively.

function of energy E and zenith angle θ takes the following form:

$$\beta(E, \theta) = a_0(E) + a_1(E) \left(\sec(\theta) - 1 \right)^3, \quad \text{with}$$

$$
\begin{aligned}
a_0(E) &= b_0 + b_1 \log_{10}(E \,[\text{eV}]), \\
a_1(E) &= c_0 + c_1 \log_{10}(E \,[\text{eV}]) + c_2 \log_{10}(E \,[\text{eV}])^2.
\end{aligned}
\tag{6.8}
$$

The five parameters b_0, b_1, c_0, c_1 and c_2 are determined through the fits shown in Fig. 6.7. The results of the fits are:

$$
\begin{aligned}
b_0 &= -0.695 \pm 0.098, \\
b_1 &= -0.107 \pm 0.005, \\
c_0 &= 506.72 \pm 0.11, \\
c_1 &= -53.159 \pm 0.006, \\
c_2 &= 1.3972 \pm 0.0003.
\end{aligned}
\tag{6.9}
$$

Checking the Parameterization

In this section, the quality of the parameterization is checked and the results are verified. In Fig. 6.8, β as a function of $\sec(\theta)$ is again shown for the first energy sub-bin, this time including the parameterized curve obtained from Eqs. 6.8 and 6.9. Qualitatively, the parameterization describes the overall dependence of β on the zenith angle very well. In Fig. 6.9, the difference between the two curves is shown for all energy sub-bins, taking also into account the zenith angle limits imposed on the individual fits. The discrepancy between the parameterized curve and the fit increases with increasing zenith angle in all sub-bins, but even in the sub-bin with the highest discrepancy, it is still less than 0.15, which is well inside the expected range of variation due to shower-to-shower fluctuations. When applying the parameterization, the zenith angle range above the energy-dependent zenith angle cut introduced in the previous section can only be accessed through extrapolation of the parameterization to higher zenith angles at low energies, as it is shown in Fig. 6.8. For higher energies, such
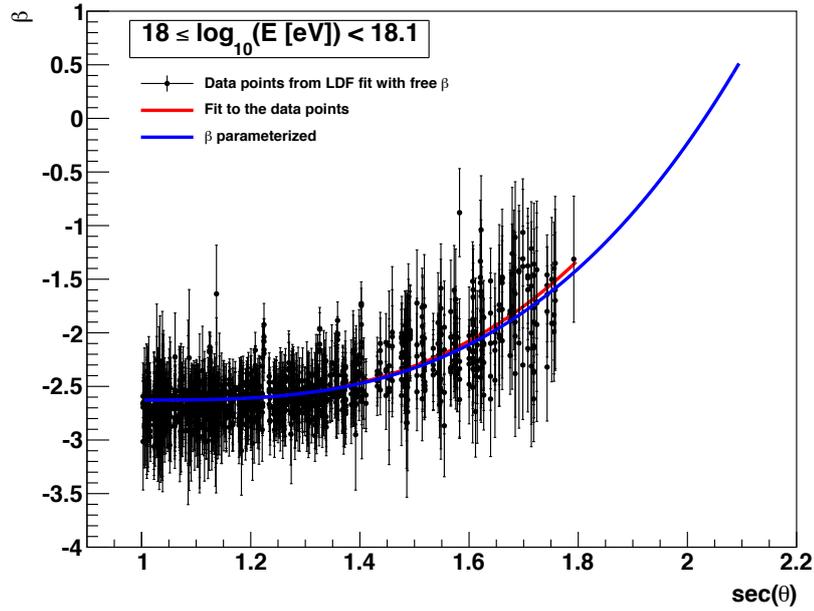
Figure 6.8.: Distribution of the fitted β as a function of $\sec(\theta)$ in the energy range $18.0 \leq \log_{10}(E\,[\text{eV}]) < 18.1$. In addition to the fit of the cubic function according to Eq. 6.7, shown as a red line, the parameterized function obtained from Eq. 6.8 is included here as a blue line.
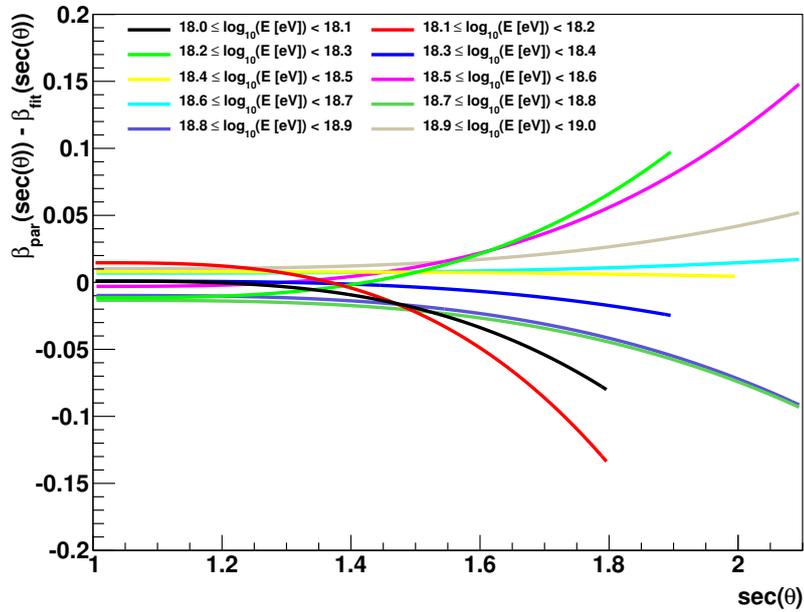
Figure 6.9.: The difference between the parameterized and the fitted curves in the individual energy sub-bins as a function of the zenith angle.

an extrapolation is not necessary since all zenith angles are taken into account in this energy range.

In Fig. 6.10, the difference between the parameterized and the fitted β is shown on an event-by-event basis for the two data sets, divided into three zenith-angle bins. The parameterized β has been calculated using Eqs. 6.8 and 6.9 and the known energy and zenith angle of the event, while the fitted β has been taken from the individual LDF. All distributions exhibit a mean value close to 0. The widths of the distributions increase with increasing zenith angle, which could already be seen in Fig. 6.8, where the shower-to-shower fluctuations increase the spread of the data points when going to higher zenith angles. In addition, a slightly negative skewness of the distributions is visible, which means that the parameterization may slightly overestimate the true $|\beta|$. This skewness is related to the small discrepancy between the fitted and the parameterized curves discussed before. Still, it can be concluded from Fig. 6.10 that the parameterization provides reasonable results in describing the average behaviour of β for photon showers.

An additional crosscheck has been performed by comparing the parameterized β with the result of an LDF fit when only the stations from the regular SD array with a spacing of 1500 m are taken into account. This is shown in Fig. 6.11. As in Fig. 6.10, also these distributions show a mean value around 0, which is expected since there should be no systematic shift in the result of the LDF fit when changing the spacing of the array. The width of these distributions, however, is larger. This is due to the average number of SD stations that are available for the fits. When using the 433 m array, on average 20 stations are triggered per event in the first energy-bin and can thus be used in the LDF fit. This number is reduced to less than two when going to the 1500 m array, hence an unconstrained LDF fit is not even possible at lower energies. In the second energy bin, more stations are triggered, but still the LDF fit is based on much less SD stations when limited to the 1500 m array. This leads to a less precise fit, with a larger spread in the distributions of the fit results.

(a) (b)
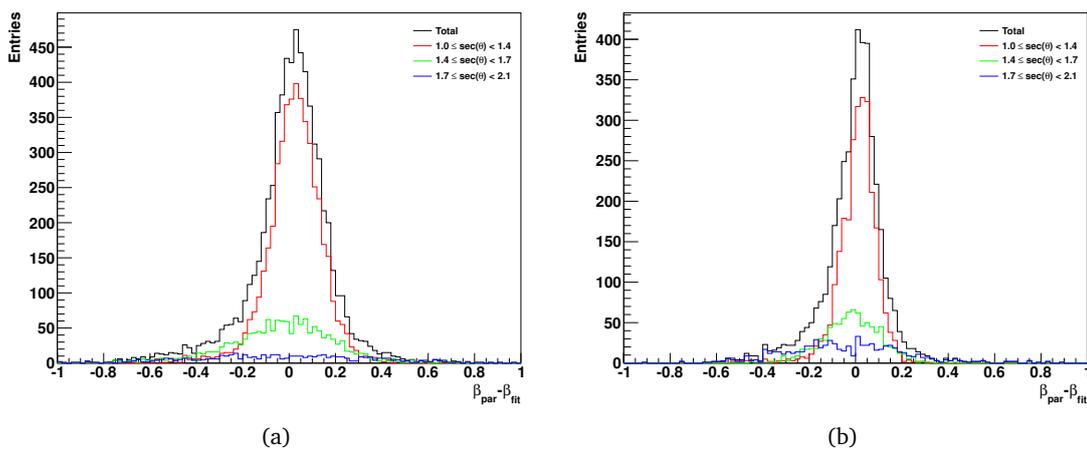
Figure 6.10.: The difference between the parameterized β and the fitted β for the two energy ranges. (a) $18.0 \leq \log_{10}(E\,[\text{eV}]) < 18.5$ and (b) $18.5 \leq \log_{10}(E\,[\text{eV}]) < 19.0$. The distributions have been divided into three zenith angle ranges.

(a)   (b)
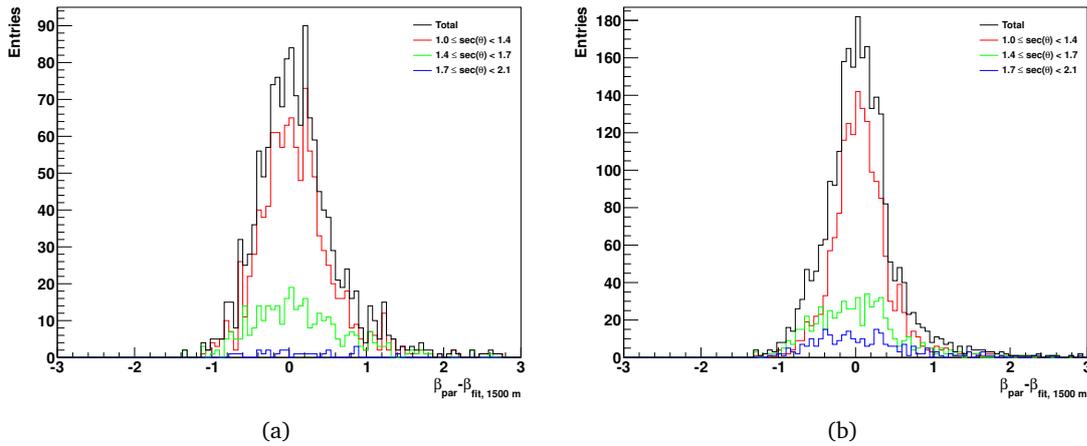
Figure 6.11.: The difference between the parameterized β and the fitted β, only taking into account SD stations from the 1500 m array, for the two energy ranges. (a) $18.0 \leq \log_{10}(E\,[\text{eV}]) < 18.5$ and (b) $18.5 \leq \log_{10}(E\,[\text{eV}]) < 19.0$. The distributions have been divided into three zenith angle ranges.

Overall, it can be stated that results obtained from the parameterization of the LDF shape parameter β are fully in agreement with the expectation. Hence, the parameterization can be used without restrictions in the photon-optimized fit of the LDF for the F_γ observable.

6.2.3. Description of the Photon-Optimized Fit

In the photon-optimized fit of the LDF following Eq. 6.4 to the signals measured in the individual SD stations, the maximum-likelihood method is used to estimate the parameter $S_{1000|\gamma}$. In general, the likelihood function $\mathscr{L}$ is a function of the parameters of a statistical model. The likelihood of a set of input parameters $\vec{p}$ given a set of outcome parameters $\vec{x}$ is equal to the probability P of the observed outcome given those parameters, i.e. $\mathscr{L}(\vec{p}\,|\,\vec{x}) = P(\vec{x}\,|\,\vec{p})$ [Cow98]. The maximum-likelihood method estimates from a given statistical model and a set of outcome parameters the input parameters which maximize the likelihood, i.e. those input parameters that have the highest probability of producing the observed outcome. In the application of the maximum-likelihood method here, the statistical model is the LDF describing the dependence of the measured signals on the perpendicular distance to the shower axis, the input parameters $\vec{p}$ are the parameters of the LDF, i.e. only $S_{1000|\gamma}$ for the photon-optimized LDF fit, and the outcome parameters $\vec{x}$ are the measured signals in the SD stations.

On a technical level, the actual maximization of $\mathscr{L}$, i.e. the minimization of $-\mathscr{L}$, is perfomed using the numerical minimization program MINUIT [Jam75] as implemented in the ROOT analysis framework, version 5.34/03. Within MINUIT, the SIMPLEX and MIGRAD minimization routines are used to determine the minimum of $-\mathscr{L}$.

The Basic Likelihood Function for the Photon-Optimized Fit

The likelihood function $\mathscr{L}$ for the photon-optimized fit is constructed from the measured signals in the individual SD stations in analogy to the likelihood function used in the standard

SD reconstruction [Veb13b]. Three contributions have to be taken into account:

$$\mathscr{L} = \mathscr{L}_{\text{small}} \times \mathscr{L}_{\text{large}} \times \mathscr{L}_{\text{silent}}, \tag{6.10}$$

where $\mathscr{L}_{\text{small}}$ and $\mathscr{L}_{\text{large}}$ denote the contributions from SD stations with small and large signals, respectively, and $\mathscr{L}_{\text{silent}}$ stands for the contribution of non-triggering, or "silent", SD stations. In the standard SD reconstruction, also saturated stations are used in the LDF fit, with the measured saturated signal taken as a lower bound for the real signal in the likelihood function. However, the observable F_γ is intended to be used in the EeV range, where in most events only a single SD station is triggered. To obtain a reliable LDF fit, it is required that all triggered stations taken into account in the fit are not saturated. Also SD stations classified as accidental—i.e. the timing of the signals in the SD station does not fit to the expected shower geometry obtained from the FD—are excluded from the LDF fit. In addition, stations that are part of the infill array as well as stations designated as test tanks [Pie14b] are excluded.

Before the individual contributions to the likelihood function are set up, it is useful to define an effective particle number n as an auxiliary quantity for use in the likelihood function. The effective particle number is directly proportional to the signal S measured in a SD station:

$$n = p(\theta) \times S \, [\text{VEM}], \tag{6.11}$$

where $p(\theta)$ is a proportionality factor called the Poisson factor [Veb13b], which itself is a function of the zenith angle θ of the air shower. The Poisson factor is approximated as

$$p(\theta) = \max\left(1, (0.32 + 0.42/\cos\theta)^{-2}\right), \tag{6.12}$$

independent of the mass and the energy of the primary particle initiating the air shower and the distance of the SD station to the shower axis [Veb13b].

Using the effective particle number, $\mathscr{L}_{\text{small}}$ and $\mathscr{L}_{\text{large}}$ can be set up. In the case of SD stations with a small signal, i.e. in the case of $n \le 30$, the probability that an SD station i at a perpendicular distance r_i to the shower axis measured n_i particles is obtained from Poisson statistics:

$$f_{\text{small}}(n_i, v_i) = \frac{v_i^{n_i} \exp(-v_i)}{n_i!}. \tag{6.13}$$

In the above expression, $v_i = p(\theta) \times S_{\text{LDF}}(r_i)$ denotes the number of particles that are expected from the LDF (Eq. 6.4). In the likelihood function, the contributions of the individual SD stations have to be multiplied:

$$\mathscr{L}_{\text{small}} = \prod_i f_{\text{small}}(n_i, v_i), \tag{6.14}$$

where the index of multiplication i extends over all non-saturated and non-accidental SD stations with $n \le 30$ in the recorded event.

Stations with a large signal, i.e. $n > 30$, are treated in a similar way. However, due to the central limit theorem [Cow98], a Gaussian approximation can be used:

$$f_{\text{large}}(n_j, v_j) = \frac{1}{\sqrt{2\pi}\sigma_j} \exp\left(-\frac{(n_j - v_j)^2}{2\sigma_j^2}\right), \tag{6.15}$$

where $\sigma_j = p(\theta) \times (0.32 + 0.42/\cos\theta) \times \sqrt{S_j}$ [VEM]. Again, the contributions of the individual SD stations have to be multiplied:

$$\mathscr{L}_{\text{large}} = \prod_j f_{\text{large}}(n_j, \nu_j), \qquad (6.16)$$

where the index of multiplication j extends over all non-saturated and non-accidental SD stations with $n > 30$ in the recorded event.

The last contribution to the likelihood function comes from non-triggering SD stations. The information that an active SD station, i.e. a station that is working from a technical point of view and is not taken out of the regular data acquisition process, at a given distance to the shower axis does not show a signal can provide additional information for an LDF fit. For the photon-optimized LDF fit, non-triggering SD stations with a perpendicular distance to the shower axis between 800 m and 1500 m are taken into account. In this range, non-triggering stations have the largest impact on the LDF fit. To distinguish between active SD stations without a signal and those stations that are not working, additional monitoring information is used. The rate at which an SD station reports local triggers to the CDAS is constantly monitored. Since atmospheric muons can also trigger an SD station, the local trigger rate should be approximately constant. If the CDAS does not receive any local triggers from an SD station, then this can be taken as a sign that this station is offline. This monitoring information is automatically compiled on a daily basis into T2Life files [Lhe13].

In general, it can be assumed that more than three muons from an air shower have to traverse an SD station to trigger the station [Veb13b]. Hence, the contribution to the likelihood function from a single non-triggering SD station at a perpendicular distance r_k to the shower axis is obtained by summing up the Poisson probabilities with $n_k \leq 3$:

$$f_{\text{silent}}(\nu_k) = \sum_{l=0}^{3} \frac{\nu_k^l \exp(-\nu_k)}{l!}. \qquad (6.17)$$

Again, the individual contributions are multiplied:

$$\mathscr{L}_{\text{silent}} = \prod_k f_{\text{silent}}(\nu_k). \qquad (6.18)$$

Here, the index of multiplication k extends over all non-triggering and active SD stations with a perpendicular distance to the shower axis between 800 m and 1500 m.

Extending the Likelihood Function

The likelihood function discussed in the previous section is the basis of the LDF fit in the standard SD reconstruction and thus also of the photon-optimized LDF fit discussed here. Before it can be applied to the photon-optimized LDF fit, however, additional terms which are not needed in the standard SD reconstruction have to be added to the likelihood function. In the standard SD reconstruction, the geometry of the air shower event and the shape of the LDF are all determined in the LDF fit, in addition to S_{1000} [Veb13b]. In the photon-optimized fit, the shower geometry, i.e. the direction of the shower axis and the location of the shower core on ground, are taken from the hybrid reconstruction of the air shower event, while the

LDF shape parameter β is determined using the parameterization derived in Sec. 6.2.2. All of these parameters are afflicted with uncertainties, which should be accounted for in the photon-optimized fit by allowing the parameters to fluctuate within their uncertainties. The likelihood function $\mathscr{L}$ is therefore extended by three additional contributions:

$$\mathscr{L} = \mathscr{L}_{\text{small}} \times \mathscr{L}_{\text{large}} \times \mathscr{L}_{\text{silent}} \times \mathscr{L}_{\text{core}} \times \mathscr{L}_{\text{axis}} \times \mathscr{L}_\beta. \tag{6.19}$$

The contribution of the uncertainties in the location of the shower core on ground level is modeled by a two-dimensional Gaussian function:

$$\mathscr{L}_{\text{core}} = \frac{1}{2\pi \Delta x_{\text{rec}}\,[\text{m}]\,\Delta y_{\text{rec}}\,[\text{m}]\sqrt{1-\varrho_{\text{core}}^2}}$$
$$\times \exp\left(-\frac{1}{2(1-\varrho_{\text{core}}^2)}\left[\frac{(x-x_{\text{rec}})^2}{\Delta x_{\text{rec}}^2} + \frac{(y-y_{\text{rec}})^2}{\Delta y_{\text{rec}}^2} - \frac{2\varrho_{\text{core}}\,(x-x_{\text{rec}})\,(y-y_{\text{rec}})}{\Delta x_{\text{rec}}\Delta y_{\text{rec}}}\right]\right),$$
$$\tag{6.20}$$

where the coordinates (x, y) represent the location of the shower core that is varied during the fitting process and is used to determine the perpendicular distance of a given SD station to the shower axis, $(x_{\text{rec}}, y_{\text{rec}})$ denote the location of the shower core obtained from the hybrid reconstruction of the shower geometry with their uncertainties Δx_{rec} and Δy_{rec}, and ϱ_{core} is the correlation coefficient between x_{rec} and y_{rec}.

The contribution of the uncertainties in the direction of the shower axis, which is represented by the zenith angle θ and the azimuth angle ϕ, is treated analogously:

$$\mathscr{L}_{\text{axis}} = \frac{1}{2\pi \Delta \theta_{\text{rec}}\,[\text{rad}]\,\Delta \phi_{\text{rec}}\,[\text{rad}]\sqrt{1-\varrho_{\text{axis}}^2}}$$
$$\times \exp\left(-\frac{1}{2(1-\varrho_{\text{axis}}^2)}\left[\frac{(\theta-\theta_{\text{rec}})^2}{\Delta \theta_{\text{rec}}^2} + \frac{(\phi-\phi_{\text{rec}})^2}{\Delta \phi_{\text{rec}}^2} - \frac{2\varrho_{\text{axis}}\,(\theta-\theta_{\text{rec}})\,(\phi-\phi_{\text{rec}})}{\Delta \theta_{\text{rec}}\Delta \phi_{\text{rec}}}\right]\right).$$
$$\tag{6.21}$$

In this term, θ and ϕ denote the zenith azimuth angles which are varied during the fitting process, while θ_{rec} and ϕ_{rec} denote the reconstructed angles, with their uncertainties $\Delta \theta_{\text{rec}}$ and $\Delta \phi_{\text{rec}}$, and ϱ_{axis} is the correlation coefficient between θ_{rec} and ϕ_{rec}.

The last contribution to the likelihood function comes from taking into account the uncertainty of the LDF shape parameter β. It is modeled using a one-dimensional Gaussian function:

$$\mathscr{L}_\beta = \frac{1}{\sqrt{2\pi}\Delta\beta_{\text{par}}}\exp\left(-\frac{(\beta-\beta_{\text{par}})^2}{2\Delta\beta_{\text{par}}^2}\right), \tag{6.22}$$

where β denotes the LDF shape parameter that is varied during the fitting process and β_{par} the shape parameter that is obtained from the parameterization discussed in Sec. 6.2.2 (Eqs. 6.8 and 6.9). The uncertainty of β_{par}, $\Delta\beta_{\text{par}}$, is determined from the uncertainties of the recon-

structed energy and zenith angle and the uncertainties of the parameters from Eq. 6.9 using error propagation:

$$\Delta\beta_{\text{par}} = \sqrt{\delta b_0^2 + \delta b_1^2 + \delta c_0^2 + \delta c_1^2 + \delta c_2^2 + \delta E^2 + \delta\theta^2}, \quad \text{with}$$

$$
\begin{aligned}
\delta b_0 &= \Delta b_0, \\
\delta b_1 &= \Delta b_1 \, \log_{10}(E\,[\text{eV}]), \\
\delta c_0 &= \Delta c_0 \, (\sec(\theta) - 1)^3, \\
\delta c_1 &= \Delta c_1 \, (\sec(\theta) - 1)^3 \, \log_{10}(E\,[\text{eV}]), \\
\delta c_2 &= \Delta c_2 \, (\sec(\theta) - 1)^3 \, \log_{10}(E\,[\text{eV}])^2, \\
\delta E &= \frac{\Delta E}{E \, \ln(10)} \left(b_1 + c_1 \, (\sec(\theta) - 1)^3 + 2c_2 \, (\sec(\theta) - 1)^3 \right), \\
\delta\theta &= 3a_1 \, (\sec(\theta) - 1)^2 \, \tan(\theta) \sec(\theta) \times \Delta\theta\,[\text{rad}].
\end{aligned}
\tag{6.23}
$$

From the individual contributions discussed previously, the likelihood function is finally constructed. In order to simplify calculating the full likelihood function, the natural logarithm is applied:

$$\ln(\mathscr{L}) = \ln\left(\mathscr{L}_{\text{small}}\right) + \ln\left(\mathscr{L}_{\text{large}}\right) + \ln\left(\mathscr{L}_{\text{silent}}\right) + \ln\left(\mathscr{L}_{\text{core}}\right) + \ln\left(\mathscr{L}_{\text{axis}}\right) + \ln\left(\mathscr{L}_{\beta}\right). \tag{6.24}$$

The minimization routines are then applied to $-\ln(\mathscr{L})$, since the logarithm is a strictly monotonic increasing function, and thus a minimum in $-\ln(\mathscr{L})$ directly translates into a minimum in $-\mathscr{L}$.

6.2.4. Performance of F_γ

The performance of the observable F_γ in distinguishing air shower events induced by photons from those intiated by protons is determined from the simulation samples discussed in Cha. 5. For every event in these samples, it should, in principle, be possible to determine the F_γ observable, since it is required in the prior event selection that each event contains a fully reconstructed hybrid event—i.e. energy and zenith angle are reconstructed from hybrid data. For some events, however, the photon-optimized fit fails due to the minimization routines not converging to a minimum. In most cases, this is due to an unusual station configuration in the reconstructed event, for example, a triggered SD station very far from the shower axis with a number of non-triggering stations closer to the shower axis. However, the number of events with failed LDF fits is less than 0.1 %, independent of the primary particle type (cf. App. B, Tab. B.2). This shows that the photon-optimized fit is robust and only a negligible number of events is lost in the analysis due to technical problems arising from the observable itself.

In Fig. 6.12, the distribution of F_γ is shown in the first energy bin for both the photon and the proton samples. The corresponding distributions in the other energy bins can be found in App. C, Fig. C.5. The average F_γ for photons in this energy range is about 0.4, while the average for protons is about 0.8. In the widths of the distributions similar differences can be found: while the width of the photon distribution is 0.14 in terms of standard deviation, and that of the proton distribution is 0.28. Both distributions exhibit a tail toward larger F_γ values, which is more pronounced in the proton distribution. This tail originates from lower-energy
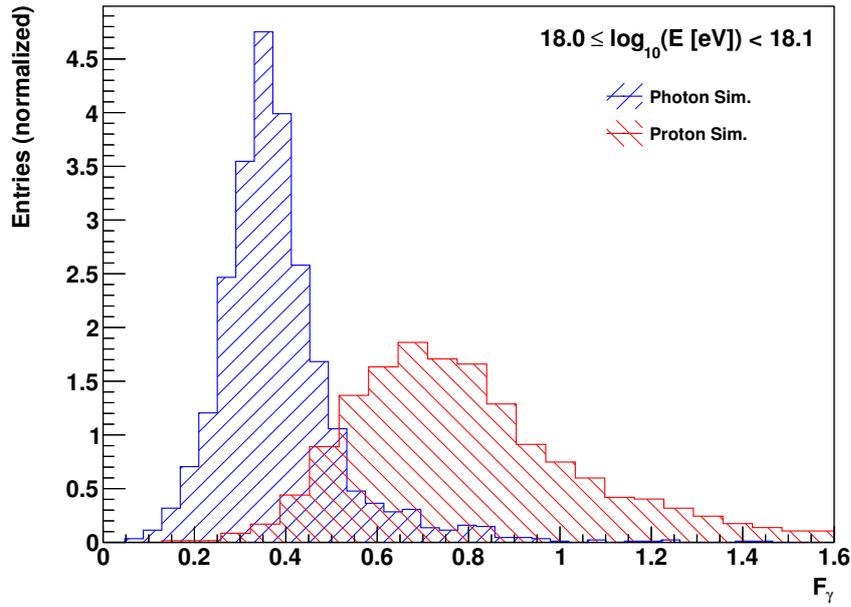
Figure 6.12.: Distributions of the observable F_γ in the first energy bin ($18.0 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.1$) for the photon sample, shown in blue, and the proton sample, shown in red.

events with only a single triggered SD station, where the signal measured in the station fluctuated upward, thereby leading to a larger value in $S_{1000|\gamma}$ in the photon-optimized LDF fit. Downward-fluctuating signals in the SD stations would lead to a similar tail toward smaller F_γ values, however, in most cases such a signal would be below the trigger threshold of the SD station, hence the tail toward smaller F_γ values does not show up in the distributions. The overlap between the two distributions is small and the distributions can be clearly separated. This is quantified through the merit factor, which is 1.36 in this energy bin. In Fig. 6.13, the background rejection is shown as a function of signal efficiency of F_γ, as derived from the distributions shown in Fig. 6.12. The good separation between the two distributions leads to an extended plateau in $\varrho(\varepsilon)$. Only above $\varepsilon = 0.9$, $\varrho(\varepsilon)$ decreases more steeply. At $\varepsilon = 0.5$, the background rejection is still about 99 %.

The properties of the F_γ distributions for the photon and the proton samples as well as the merit factor η and $\varrho(\varepsilon = 50\,\%)$ are listed for all energy bins in Tab. 6.2. With increasing energy, the distributions move further apart—the mean for photons decreases, and the mean for protons increases—while the distributions get smaller in width. This leads to a significant increase in separation power: the merit factor increases almost linearly from 1.34 around 1 EeV to 2.58 around 10 EeV. Similarly, $\varrho(\varepsilon)$ increases from 98.66 % around 1 EeV to 99.89 % around 10 EeV. In Fig. 6.14, η and $\varrho(\varepsilon = 50\,\%)$ for the F_γ observable are shown as a function of energy and compared to the values obtained for the $X_{\max}$ observable. As evident from these results, F_γ performs better than $X_{\max}$ at all energies. In the distribution of the merit factor, the absolute difference between the data points for F_γ and those for $X_{\max}$ even increases with increasing energy. The good performance of the F_γ observable in distinguishing air shower events induced by UHE photons from those initiated by protons can be clearly seen.
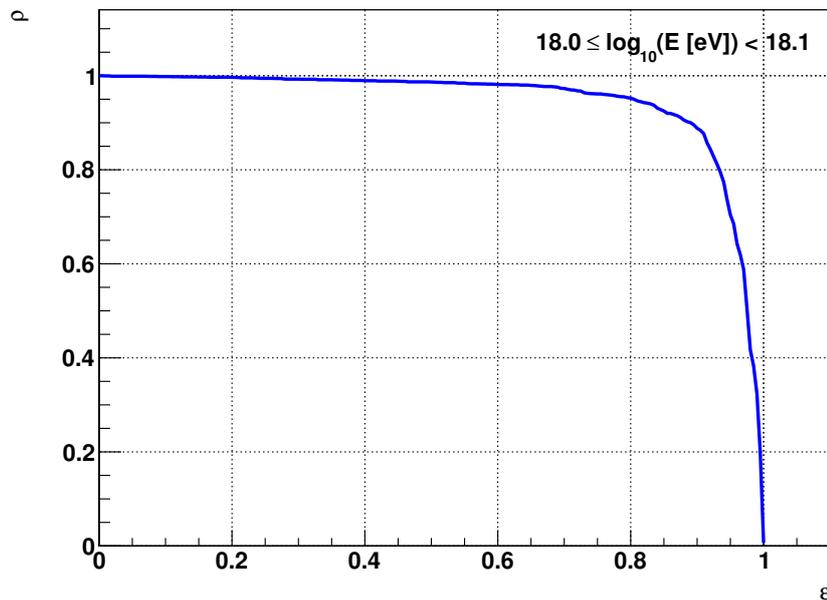
Figure 6.13.: Background rejection ϱ as a function of the signal efficiency ε for the observable F_γ in the first energy bin ($18.0 \leq \log_{10}(E\,[\text{eV}]) < 18.1$).

| Energy Bin | Photon | | Proton | | | |
	μ	σ	μ	σ	η	$\varrho(\varepsilon = 50\%)$
$18.0 \leq \log_{10}(E\,[\text{eV}]) < 18.1$	0.39	0.14	0.81	0.28	1.34	98.66%
$18.1 \leq \log_{10}(E\,[\text{eV}]) < 18.2$	0.37	0.12	0.79	0.26	1.44	98.74%
$18.2 \leq \log_{10}(E\,[\text{eV}]) < 18.3$	0.35	0.11	0.79	0.23	1.69	99.22%
$18.3 \leq \log_{10}(E\,[\text{eV}]) < 18.4$	0.34	0.10	0.78	0.22	1.82	99.35%
$18.4 \leq \log_{10}(E\,[\text{eV}]) < 18.5$	0.34	0.11	0.78	0.20	1.90	99.68%
$18.5 \leq \log_{10}(E\,[\text{eV}]) < 18.6$	0.34	0.11	0.80	0.20	2.08	99.71%
$18.6 \leq \log_{10}(E\,[\text{eV}]) < 18.7$	0.34	0.11	0.81	0.19	2.20	99.87%
$18.7 \leq \log_{10}(E\,[\text{eV}]) < 18.8$	0.34	0.10	0.83	0.18	2.38	99.97%
$18.8 \leq \log_{10}(E\,[\text{eV}]) < 18.9$	0.35	0.10	0.84	0.17	2.48	99.95%
$18.9 \leq \log_{10}(E\,[\text{eV}]) < 19.0$	0.35	0.10	0.86	0.17	2.58	99.89%

Table 6.2.: Properties of the F_γ distributions for the photon and proton samples and the separation power of the observable. Listed are the mean values μ of the distributions and their widths in terms of standard deviation σ. For each energy bin, the merit factor η has been calculated using these parameters. In addition, the background rejection ϱ at a signal efficiency of $\varepsilon = 50\%$ has been derived from the distributions (cf. App. C, Fig. C.5).
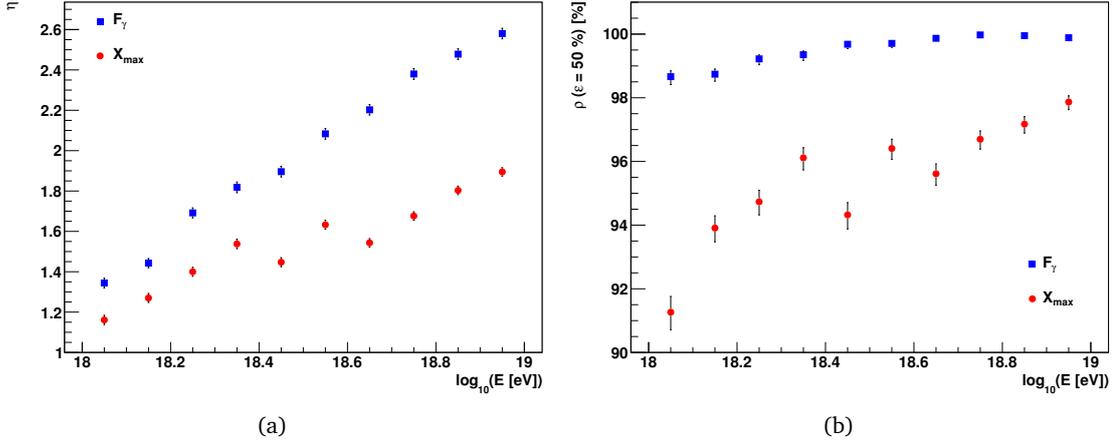
Figure 6.14.: Measures of the separation power of the observable F_γ as a function of energy. (a) Merit factor η. (b) Background rejection ϱ at a signal efficiency of $\varepsilon = 50\%$. The uncertainties of the individual data points have been estimated using error propagation in the case of η and according to [Pat04] in the case of $\varrho(\varepsilon = 50\%)$. In addition to the data points for F_γ, shown as blue squares, the data points for X_{max} (cf. Fig. 6.3) are included as red circles.

6.2.5. Comparison with other SD-related Observables

In this section, the separation power of the observable F_γ is compared to the separation power of other SD-related observables that have been used previously in the search for UHE photons in hybrid data collected at the Pierre Auger Observatory. The observables S_b and ζ are of particular interest here.

The observable S_b exploits the differences in the LDF of air showers induced by photons as compared to the LDF of showers initiated by protons and nuclei due to the differences in the electromagnetic and muonic components. For an air shower event with N triggered SD stations, S_b is defined as:

$$S_b = \sum_{i=1}^{N} \left[S_i \left(\frac{r_i}{1000\,\mathrm{m}} \right)^b \right], \tag{6.25}$$

where S_i and r_i denote the total signal and the perpendicular distance to the shower axis of the individual SD stations. For the choice of the free parameter b, several possibilities exist. In [Set11], $b = 4$ has been used following [Ros10]. However, in a more recent study, it has been argued that the optimum choice for distinguishing air shower events induced by UHE photons from those initiated by protons and nuclei is $b = 3$ [Ros13]. This choice of b has been used in the analysis presented in [Aab14c]. In the following, $b = 4$ is used in order to be consistent with [Set11]. It should be noted that in the application of S_b to data, an additional selection criterion—at least four active SD stations with a perpendicular distance to shower core less than 2000 m—must be applied to exclude events where only an incomplete S_b can be determined and thus prevent a bias in the S_b distributions [Set11].

The spread in arrival times of the secondary particles from an air shower in the SD stations can also be used to distinguish different primary particle types. For showers initiated by UHE

photons, a broader distribution in the arrival times of the secondary particles—and hence a broader shape of the FADC trace in the SD stations—is expected than for showers initiated by protons, mainly because photon-induced showers develop deeper in the atmosphere and contain less muons. The difference in the FADC traces is exploited with the shape parameter ζ, which is defined as the ratio of the early-arriving signal in the SD station with the highest signal to the late-arriving signal in this station [Aab14c]:

$$\zeta = \frac{S_{\text{early}}}{S_{\text{late}}}. \tag{6.26}$$

The early-arriving signal is in turn defined as the integrated signal from the beginning of the FADC trace to the time bin corresponding to $0.6\,\mu s$ on a shifted time scale. This time scale varies with the zenith angle θ and the perpendicular distance r of the SD station to the shower axis:

$$t_i^{\text{scaled}} = t_i \frac{1000\,\text{m}}{r} \times \frac{1}{-0.6 + 1.9\,\cos(\theta)}, \tag{6.27}$$

where t_i^{scaled} and t_i denote the scaled and the real time, respectively, of time bin i in the FADC trace [Aab14c]. Correspondingly, the late-arriving signal is defined as the integrated signal from $0.6\,\mu s$ on the shifted time scale until the end of the FADC trace.

The resulting distributions of S_b and ζ for the photon and the proton samples are shown in Fig. 6.15 for the first energy bin and in App. C, Figs. C.7 and C.9 for the other energy bins. For both observables, the distributions overlap with each other to a large extent. Consequently, the merit factor η is small compared to X_{max} and F_γ: 0.84 in the case of S_b and 0.37 in the case of ζ. The smaller separation power can be seen in $\varrho(\varepsilon)$ (cf. Fig. 6.16). For both observables, $\varrho(\varepsilon)$ already decreases at small ε. The background rejection at a signal efficiency of 50 % is 90.82 % for S_b and 69.22 % for ζ.

The properties of the S_b and ζ distributions for all energy bins are summarized in Tab. 6.3. Due to a technical issue, the FADC traces in the energy bin $18.8 \leq \log_{10}(E\,[\text{eV}]) < 18.9$ could

(a)   (b)
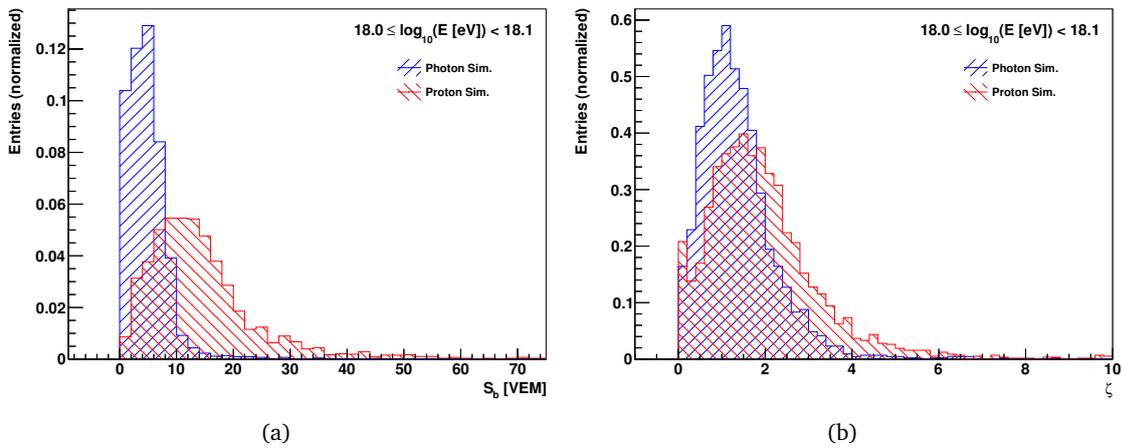
Figure 6.15.: Distributions of (a) S_b and (b) ζ in the first energy bin ($18.0 \leq \log_{10}(E\,[\text{eV}]) < 18.1$) for the photon sample (blue distributions) and the proton sample (red distributions).
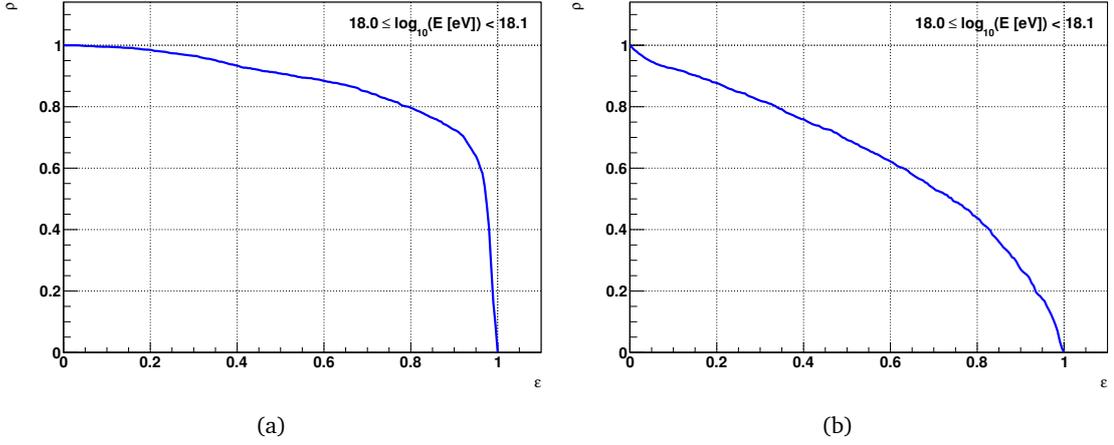
Figure 6.16.: Background rejection ϱ as a function of the signal efficiency ε for (a) S_b and (b) ζ in the first energy bin ($18.0 \leq \log_{10}(E\,[\text{eV}]) < 18.1$).
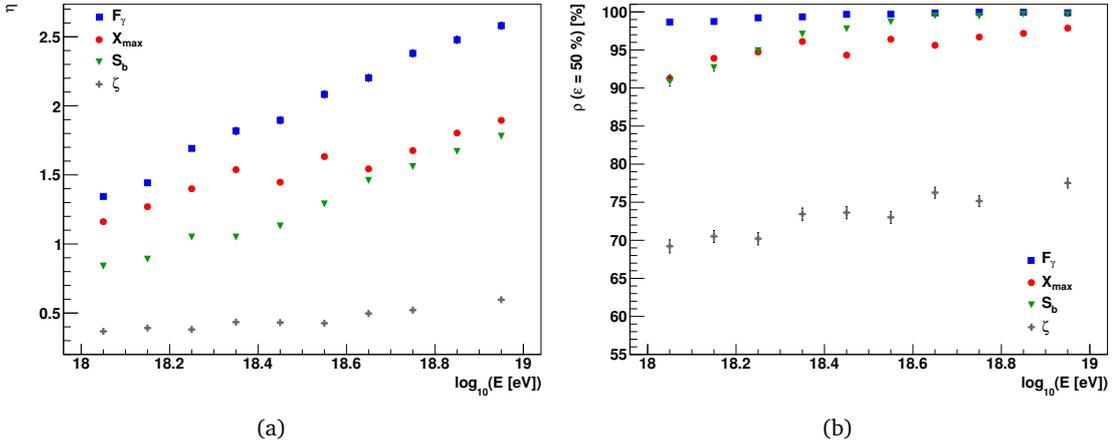
Figure 6.17.: Measures of the separation power of the SD-related observables S_b (green triangles) and ζ (gray crosses) as a function of energy. (a) Merit factor η. (b) Background rejection ϱ at a signal efficiency of $\varepsilon = 50\,\%$. The uncertainties of the individual data points have been estimated using error propagation in the case of η and according to [Pat04] in the case of $\varrho(\varepsilon = 50\,\%)$. In addition, the data points for F_γ and X_{max} are included as blue squares and red circles, respectively.

not be recovered from the ADST files. Therefore, it was not possible to determine the ζ distributions in this energy bin. The calculation of S_b and F_γ was still possible because the total signal measured in the individual SD stations was properly saved in the ADST files. In the other energy bins, no such issues appeared. From the parameters listed in Tab. 6.3, η and $\varrho(\varepsilon = 50\,\%)$ have been determined for the observables S_b and ζ and are shown in Fig. 6.17 as a function of energy. In the case of ζ, η increases only by about 0.2 between 1 EeV and 10 EeV and thus appears approximately constant. For S_b, η increases almost linearly from 0.84 at 1 EeV to 1.78 at 10 EeV. At higher energies, the separation power of S_b is comparable to the separation power of X_{max}, but not as high as the separation power of F_γ. In terms of

Energy Bin	Photon		Proton			
	μ [VEM]	σ [VEM]	μ [VEM]	σ [VEM]	η	$\varrho(\varepsilon = 50\%)$
$18.0 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.1$	4.87	4.23	14.72	10.94	0.84	90.82%
$18.1 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.2$	6.14	5.53	18.77	13.11	0.89	92.68%
$18.2 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.3$	8.00	6.33	25.31	15.22	1.05	94.91%
$18.3 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.4$	9.96	8.20	35.52	22.93	1.05	97.08%
$18.4 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.5$	13.41	12.14	47.11	27.26	1.13	97.77%
$18.5 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.6$	17.67	14.13	66.24	34.80	1.29	98.69%
$18.6 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.7$	22.82	14.49	90.20	43.97	1.46	99.50%
$18.7 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.8$	30.76	18.70	122.19	55.62	1.56	99.44%
$18.8 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.9$	41.70	22.96	167.86	72.15	1.67	99.71%
$18.9 \leq \log_{10}(E\,[\mathrm{eV}]) < 19.0$	56.87	31.41	224.96	88.80	1.78	99.66%

(a)

Energy Bin	Photon		Proton			
	μ	σ	μ	σ	η	$\varrho(\varepsilon = 50\%)$
$18.0 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.1$	1.36	0.87	1.94	1.31	0.37	69.22%
$18.1 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.2$	1.19	0.76	1.75	1.21	0.39	70.51%
$18.2 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.3$	1.14	0.80	1.67	1.13	0.38	70.22%
$18.3 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.4$	1.01	0.68	1.54	1.01	0.43	73.44%
$18.4 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.5$	0.89	0.67	1.39	0.92	0.43	73.64%
$18.5 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.6$	0.85	0.67	1.32	0.89	0.43	73.01%
$18.6 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.7$	0.75	0.57	1.25	0.82	0.50	76.26%
$18.7 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.8$	0.67	0.50	1.13	0.73	0.52	75.16%
$18.8 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.9$	-	-	-	-	-	-
$18.9 \leq \log_{10}(E\,[\mathrm{eV}]) < 19.0$	0.55	0.40	1.00	0.63	0.60	77.53%

(b)

Table 6.3.: Properties of the distributions of (a) S_b and (b) ζ for the photon and proton samples and the separation power of the observable. Listed are the mean values μ of the distributions and their widths in terms of standard deviation σ. For each energy bin, the merit factor η and the background rejection ϱ at a signal efficiency of $\varepsilon = 50\%$ have been calculated using these parameters.

$\varrho(\varepsilon = 50\%)$, the same qualitative result can be found. While the background rejection at a signal efficiency of 50% of S_b is, especially at higher energies, comparable to $X_{\max}$, $\varrho(\varepsilon = 50\%)$ for ζ is significantly smaller.

Summarizing the results of the comparison of the separation power of F_γ and the previously used SD-related observables for photon/proton discrimination in hybrid events, it can be stated that F_γ is a viable alternative to S_b and ζ in the energy range from 1 EeV to 10 EeV.

6.2.6. Introducing additional selection criteria based on F_γ

So far, only the selection criteria discussed in Sec. 5.3 have been applied to the simulation sample. It should be noted that this event selection is not optimized toward any observable in order to prevent a bias in the comparison of the performance of the different observables.

By introducing additional selection criteria optimized for the use of a specific observable, the overall performance of the observable in the separation of photon-induced and proton-induced air shower events can be increased. These additional selection criteria are chosen on the basis of the simulated photon and proton samples. In the case of the observable F_γ, the purpose of the additional selection criteria is to remove events where the quality of the photon-optimized fit of the LDF is not good enough. Such events mainly lie on the tails of the F_γ distribution, hence by removing these events, the width of the distributions is reduced—which in turn reduces the overlap between the photon and the proton distributions. The event selection criteria that are applied in addition to the ones described in Sec. 5.3.1 are listed in Tab. 6.4. First, events where LDF fit failed are removed (`FitFailed false`). This necessary criterion has been mentioned already in Sec. 6.2.4 and has been implicitly included in all plots shown before. In the next step, events where the absolute and the relative uncertainties of the calculated F_γ are too high are discarded (`DeltaFgamma` < 0.35 and `RelFgammaUncertainty` < 0.7). The uncertainty of F_γ is determined using error propagation:

$$\frac{\Delta F_\gamma}{F_\gamma} = \sqrt{\left(\frac{\Delta S_{1000|\gamma}}{S_{1000|\gamma}}\right)^2 + \left(\frac{\Delta S_{1000|\text{Hybrid}}}{S_{1000|\text{Hybrid}}}\right)^2}. \tag{6.28}$$

The uncertainty of $\Delta S_{1000|\gamma}$ is taken from the fit results provided by MINUIT, while the uncertainty of $S_{1000|\text{Hybrid}}$ is determined from the uncertainties on the reconstructed energy E and the reconstructed zenith angle θ as well as the uncertainties of the parameters from Eqs. 4.11 and 6.5 according to:

$$\frac{\Delta S_{1000|\text{Hybrid}}}{S_{1000|\text{Hybrid}}} = \sqrt{\left(\frac{\Delta S_{38|\text{Hybrid}}}{S_{38|\text{Hybrid}}}\right)^2 + \left(\frac{\Delta \text{CIC}(\theta)}{\text{CIC}(\theta)}\right)^2}, \quad \text{with}$$

$$\Delta S_{38|\text{Hybrid}} = \frac{S_{38|\text{Hybrid}}}{b} \sqrt{\left(\frac{\Delta E}{E}\right)^2 + \left(\frac{\Delta a}{a}\right)^2 + \log_{10}(E/a)\left(\frac{\Delta b}{b}\right)^2}, \tag{6.29}$$

$$\Delta \text{CIC}(\theta) = \sqrt{\Delta c^2\, x^2 + \Delta d^2\, x^4 + \Delta x^2 (c + 2dx)^2},$$

$$\Delta x = 2\,\Delta\theta\,[\text{rad}] \times \cos(\theta)\,\sin(\theta).$$

Finally, events where the photon-optimized LDF deviates too far from the signals measured in the individual SD stations are removed through a cut on a χ^2-like quantity (`ChiSquare` < 30):

$$\chi^2 = \sum_i \left(\frac{S_i - S_{\text{LDF}}(r_i)}{\Delta S_i}\right)^2, \tag{6.30}$$

where the S_i denote the signals measured in SD stations that have been taken into account in the photon-optimized LDF fit, $S_{\text{LDF}}(r_i)$ denote the signal that is calculated from the photon-optimized LDF for a station at the respective perpendicular distance to the shower core, and

FitFailed	false
DeltaFgamma	< 0.35
RelFgammaUncertainty	< 0.7
ChiSquare	< 30

Table 6.4.: Addtional selection criteria based on F_γ.

(a)   (b)
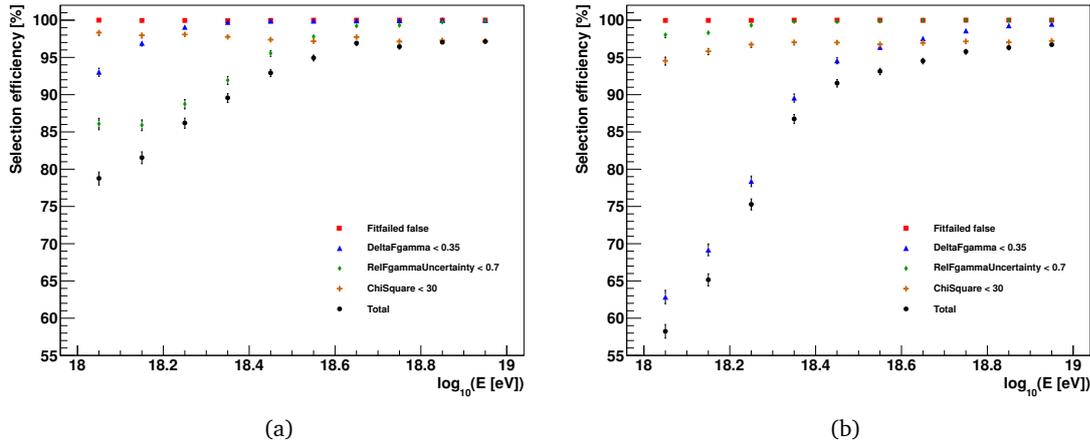
Figure 6.18.: Selection efficiencies in the individual energy bins after the additional selection criteria based on F_γ have been applied for (a) the photon sample and (b) the proton sample. The uncertainties of the selection efficiencies have been determined according to [Pat04].

ΔS_i is the uncertainty of the measured signal. Detailed studies of χ^2 can be found in [Wer13] and [Höf13].

The number of surviving events after each of the additional criteria has been applied to the main simulation samples are listed in App. B, Tab. B.2. The resulting selection efficiencies are shown in Fig. 6.18 for the various energy bins. As mentioned before, the photon-optimized LDF fit fails only for a few events. Consequently, the selection efficiency of the corresponding selection criterion is very close to 100 % for both the photon and the proton sample. For the photon sample, the cut on the relative uncertainty of F_γ is the cut with largest effect on the total selection efficiency, while for the proton sample, it is the cut on the absolute uncertainty of F_γ. This difference comes from the fact that both the average F_γ and the average absolute uncertainty of F_γ for photon-induced events are smaller than for proton-induced events. Hence, the average relative uncertainty of F_γ is larger for photon-induced events. Hence, the cut on the absolute uncertainty of F_γ is more significant for the proton sample, while for the photon sample, the cut on the relative uncertainty of F_γ is more significant. In order to prevent losing too much statistics in the photon sample, the cut on the relative uncertainty on F_γ has been chosen less strict than the cut on the absolute uncertainty of F_γ. The selection efficiency of the cut on the χ^2 quantity is for both samples almost constant with energy at about 98 % for the photon sample and about 95 % for the proton sample. For both samples, the total selection efficiency increases steeply with energy at lower energies, and it remains almost constant at higher energies. This behaviour is rather desirable, since the impact is largest at lower energies, where the available statistics in data is large, while at higher energies, where the available statistics is significantly smaller due to the steeply falling energy spectrum of primary cosmic rays and where the photon-optimized LDF fit is, even without the additional selection criteria, reliable enough, the additional selection criteria remove only few events.

In Fig. 6.19, the distributions of the observable F_γ are shown again for the photon and the proton samples, this time after applying the additional selection criteria. Due to the additional
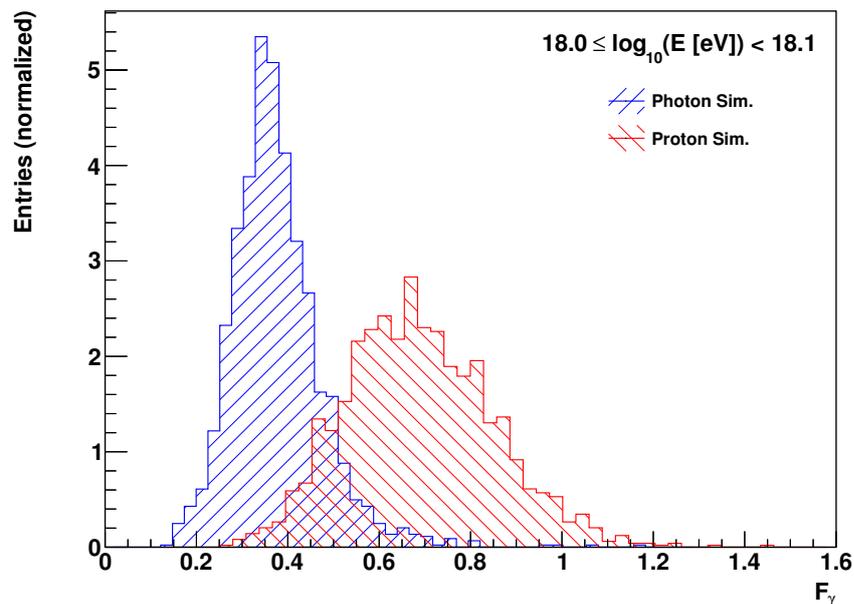
Figure 6.19.: Distributions of the observable F_γ, including the additional selection critereria, in the first energy bin ($18.0 \leq \log_{10}(E\,[\text{eV}]) < 18.1$) for the photon sample, shown in blue, and the proton sample, shown in red.

Figure 6.20.: Background rejection ϱ as a function of the signal efficiency ε for the observable F_γ, including the additional selection criteria, in the first energy bin ($18.0 \leq \log_{10}(E\,[\text{eV}]) < 18.1$).

| | Photon | | Proton | | | |
| | μ | σ | μ | σ | η | $\varrho(\varepsilon = 50\%)$ |
Energy Bin						
$18.0 \leq \log_{10}(E\,[\text{eV}]) < 18.1$	0.38	0.10	0.69	0.16	1.65	98.82 %
$18.1 \leq \log_{10}(E\,[\text{eV}]) < 18.2$	0.38	0.10	0.69	0.17	1.62	98.70 %
$18.2 \leq \log_{10}(E\,[\text{eV}]) < 18.3$	0.37	0.10	0.72	0.17	1.81	99.24 %
$18.3 \leq \log_{10}(E\,[\text{eV}]) < 18.4$	0.35	0.10	0.75	0.18	1.92	99.29 %
$18.4 \leq \log_{10}(E\,[\text{eV}]) < 18.5$	0.35	0.10	0.76	0.18	1.97	99.72 %
$18.5 \leq \log_{10}(E\,[\text{eV}]) < 18.6$	0.34	0.10	0.79	0.18	2.18	99.75 %
$18.6 \leq \log_{10}(E\,[\text{eV}]) < 18.7$	0.34	0.10	0.80	0.18	2.28	99.86 %
$18.7 \leq \log_{10}(E\,[\text{eV}]) < 18.8$	0.34	0.10	0.82	0.17	2.46	99.97 %
$18.8 \leq \log_{10}(E\,[\text{eV}]) < 18.9$	0.35	0.10	0.84	0.17	2.52	99.95 %
$18.9 \leq \log_{10}(E\,[\text{eV}]) < 19.0$	0.35	0.10	0.85	0.16	2.59	99.88 %

Table 6.5.: Properties of the F_γ distributions, after applying the additional selection criteria, for the photon and proton samples and the separation power of the observable. Listed are the mean values μ of the distributions and their widths in terms of standard deviation σ. For each energy bin, the merit factor η has been calculated using these parameters. In addition, the background rejection ϱ at a signal efficiency of $\varepsilon = 50\%$ has been derived from the distributions (cf. App. C, Fig. C.11).

(a)   (b)
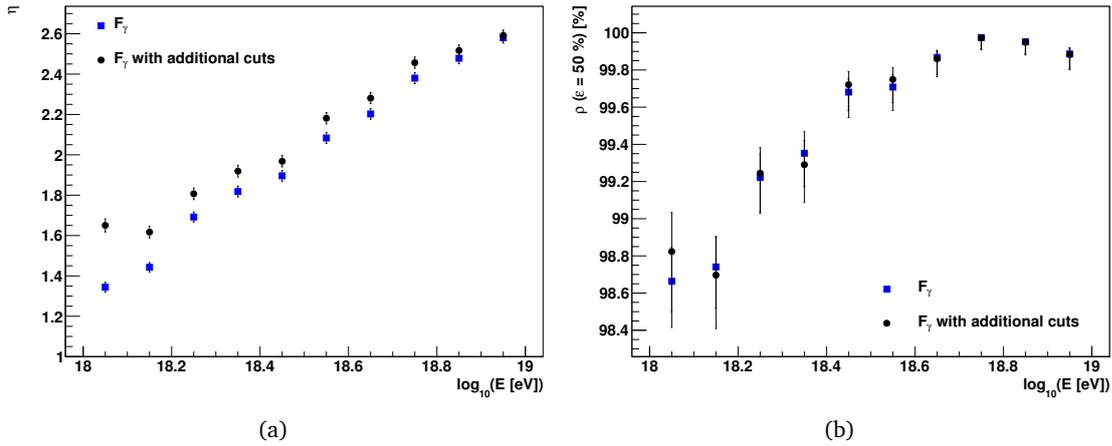
Figure 6.21.: Measures of the separation power of the observable F_γ as a function of energy, without the additonal selection criteria (blue squares) and including the additional selection criteria (black circles): (a) merit factor η. (b) background rejection ϱ at a signal efficiency of $\varepsilon = 50\%$. The uncertainties of the individual data points have been estimated using error propagation in the case of η and according to [Pat04] in the case of $\varrho(\varepsilon = 50\%)$.

selection criteria, the tails of the distributions are reduced. Hence, the width of the distributions in terms of standard deviation becomes smaller compared to the distributions shown in Fig. 6.12. Subsequently, the merit factor increases from 1.34 to 1.65. Also the background rejection at a signal efficiency of 50 % increases slightly (cf. Fig. 6.20), from 98.66 % to 98.82 %. In Tab. 6.5, the properties of the F_γ distributions with the additional selection criteria applied are listed for all energy bins, together with η and $\varrho(\varepsilon = 50\%)$. With increasing energy, the

mean values of the photon and proton distributions move further apart, but the widths of the distributions remain constant (about 0.1 for the photon sample and about 0.17 for the proton sample). As can be expected from the selection efficiencies shown in Fig. 6.18, the impact on the F_γ distributions of the additional selection criteria is largest at lower energies. Here, the increase of the separation power is largest (cf. Tab. 6.2). In Fig. 6.21, η and $\varrho(\varepsilon = 50\%)$ are shown as a function of energy, both with and without the additional selection criteria applied. At all energies, the overall separation power of the observable F_γ is larger with the additional selection criteria applied. At the same time, these additional selection criteria do not affect the distributions of $X_{\max}$ significantly. The mean values and the widths of the distributions only change within the uncertainties expected from the available statistics, hence the separation power of the observable remains unaffected by the additional selection criteria.

6.3. Combining F_γ and X_{max} in a Multivariate Analysis

In the previous sections, the observables $X_{\max}$ and F_γ have been discussed individually and independent from each other. The topic of this section is the classification of events according to their most probable primary particle type using a Multivariate Analysis (MVA) combining both observables.

The two observables F_γ and $X_{\max}$ are shown in a scatter plot in Fig. 6.22. As before, only the first energy bin ($18.0 \leq \log_{10}(E\,[\text{eV}]) < 18.1$) is shown here. The corresponding plots for the other energy bins can be found in App. C, Fig. C.13. Both observables complement each other well. As expected from the individual distributions of the two observables, the distributions in the scatter plots for the photon and the proton samples are quite clearly separated,
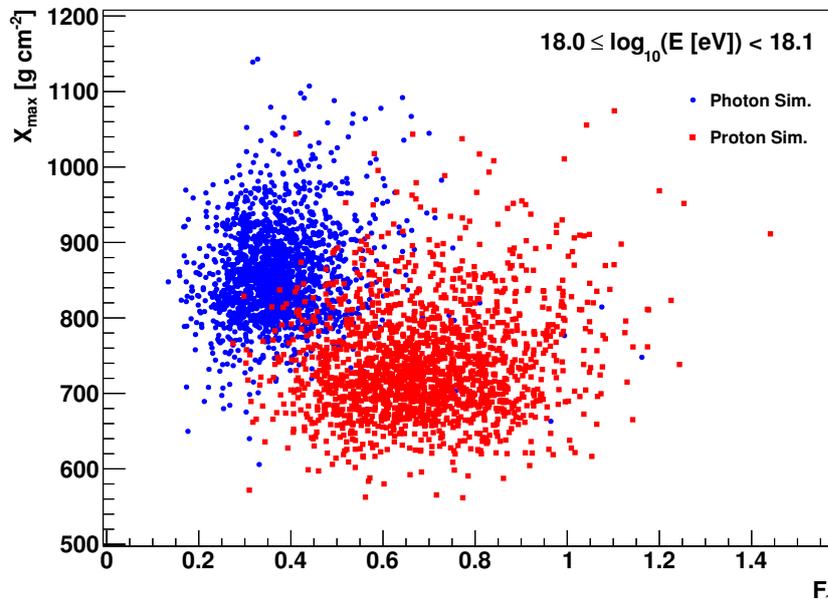
Figure 6.22.: Distribution of the two observables F_γ and $X_{\max}$ shown in a scatter plot for the first energy bin ($18.0 \leq \log_{10}(E\,[\text{eV}]) < 18.1$). The photon sample and the proton sample are shown as blue dots and red squares, respectively.
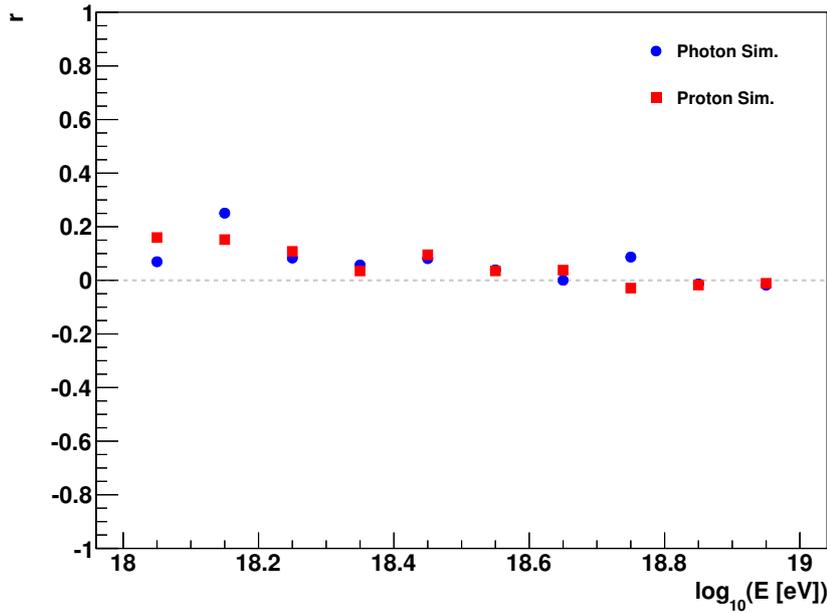
Figure 6.23.: Pearson correlation coefficient r for the two observables F_γ and $X_{\max}$ as a function of energy for the photon sample (blue dots) and the proton sample (red squares).

with the separation increasing with energy. Within the samples, no significant degree of correlation between F_γ and $X_{\max}$ is visible. To quantify the degree of correlation, the Pearson product-moment correlation coefficient is employed [Pea96]. For two observables x and y and a sample of size n, the Pearson correlation coefficient r is defined as

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}, \tag{6.31}$$

where $\bar{x}$ and $\bar{y}$ denote the arithmetic means of the observables. By construction, r can range from -1 to 1, with 1 (-1) denoting full correlation (anticorrelation) and 0 indicating no linear correlations at all. In Fig. 6.23, the Pearson correlation coefficient between F_γ and $X_{\max}$ is shown as a function of energy for the photon and the proton samples. At lower energies, at most a weak correlation is found, while at higher energies the two observables are essentially uncorrelated.

The MVA method used here is the Fisher discriminant analysis [Fis36], which has been used also in [Set11]. In the Fisher discriminant analysis, the two observables F_γ and $X_{\max}$ are combined linearly to obtain the Fisher discriminant t, which is then used as an additional observable for the classification of air shower events into background events, induced by protons or nuclei, and signal events, induced by UHE photons. The analysis presented in [Aab14c] uses tha alternative MVA method of Boosted Decision Trees (BDTs) [Bre84, Sch90], which uses a majority vote on the result of a large number of weighted decision trees for the event classification. However, the classification performance of BDTs is worse than that of a Fisher discriminant analysis for uncorrelated or linearly correlated input variables [Hoe13], which is the case with F_γ and $X_{\max}$ (cf. Fig. 6.23). In addition, BDTs are prone to overtraining issues, where statistical fluctuations in the input samples may affect the BDT training and therefore

the overall performance of the event classification [Hoe13]. The Fisher discriminant analysis is, in general, more robust against overtraining, as well as more transparent [Hoe13].

In Sec. 6.3.1, the Fisher discriminant analysis is introduced and the application to the simulation samples is discussed in detail. As the Fisher discriminant analysis is applied to each energy bin separately, the resulting parameters are parameterized as a function of energy (Sec. 6.3.2). In the next step, a photon candidate cut is defined (Sec. 6.3.3), which is used to finally classify events as background or signal. A conservative approach is chosen here, hence classified signal events are denoted as "photon candidate events" only.

6.3.1. Applying the Fisher Discriminant Analysis

Before the Fisher discriminant analysis is applied to the combination of the two observables F_γ and X_{max}, the mathematical foundations of the Fisher discriminant analysis are discussed briefly, loosely following [Cow98]. The event classification in the Fisher discriminant analysis is performed in a transformed variable space, where the input observables are linearly projected into a one-dimensional space. Considering a sample of N d-dimensional vectors of observables $\vec{x}$, n_1 of which belonging to class C_1 and n_2 belonging to class C_2, the linear projection into the one-dimensional space, i.e. the combination of the d observables x_i into a single observable t, called the Fisher discriminant, is given by

$$t = \sum_{i=1}^{d} f_i x_i = \vec{f}^T \vec{x}, \tag{6.32}$$

where $\vec{f}^T = (f_1, ..., f_d)$ represents the transposed vector of Fisher coefficients. The coefficients are now chosen such that the distributions of the two classes are pushed as far away as possible in the one-dimensional space, while events of the same class are confined in close vicinity to each other. A metric that takes both requirements into account is given by the Fisher criterion:

$$J(\vec{f}) = \frac{\vec{f}^T \mathscr{B} \vec{f}}{\vec{f}^T \mathscr{W} \vec{f}}, \tag{6.33}$$

which is based on the between-class scatter matrix $\mathscr{B}$, which describes the dispersion of events relative to the overall mean values of the observables for the sample, and the within-class scatter matrix $\mathscr{W}$, which similarly describes the dispersion of events relative to the mean values of the observables within the two classes. The components of the two matrices are given by

$$\begin{aligned} \mathscr{B}_{ij} &= (\vec{\mu}_1 - \vec{\mu}_2)_i (\vec{\mu}_1 - \vec{\mu}_2)_j, \\ \mathscr{W}_{ij} &= (V_1 + V_2)_{ij}, \end{aligned} \tag{6.34}$$

where the $\vec{\mu}_k$ and V_k denote the vector of mean values and the covariance matrix for the events in the class C_k. It can be shown that $J(\vec{f})$—and thus the separation between the two classes—is at its maximum, if

$$\vec{f} \propto \mathscr{W}^{-1} (\vec{\mu}_1 - \vec{\mu}_2). \tag{6.35}$$

The Fisher coefficients f_i for determining the Fisher discriminant are therefore only determined up to an arbitrary scaling factor, which has no influence on the separation between the

6. Simulation Study of Observables for the Search for UHE Photons

two classes.

For the analysis presented in this thesis, the implementation of the Fisher discriminant analysis that is provided by the TMVA package [Vos07] version 4.2.0, which is part of the ROOT analysis framework version 5.34/26 , is used. Since only two observables, F_γ and $X_{\max}$, are combined, Eq. 6.32 reduces to

$$t = f_1 \times F_\gamma + f_2 \times X_{\max}, \tag{6.36}$$

with the two Fisher coefficients f_1 and f_2:

$$\begin{aligned}
f_1 &= \alpha \left(\mathscr{W}_{11}^{-1} \left[\mu_\gamma(F_\gamma) - \mu_p(F_\gamma) \right] + \mathscr{W}_{12}^{-1} \left[\mu_\gamma(X_{\max}) - \mu_p(X_{\max}) \right] \right), \\
f_2 &= \alpha \left(\mathscr{W}_{21}^{-1} \left[\mu_\gamma(F_\gamma) - \mu_p(F_\gamma) \right] + \mathscr{W}_{22}^{-1} \left[\mu_\gamma(X_{\max}) - \mu_p(X_{\max}) \right] \right).
\end{aligned} \tag{6.37}$$

Here, $\mu_\gamma(x)$ and $\mu_p(x)$ denote the arithmetic means of the observable x for the photon sample and the proton sample, while the within-class scatter matrix $\mathscr{W}$ is obtained from the covariance matrices of the two observables within the photon and the proton sample according to Eq. 6.34. The scaling factor α is chosen per convention within TMVA as

$$\alpha = \frac{\sqrt{n_\gamma n_p}}{n_\gamma + n_p}, \tag{6.38}$$

where n_γ and n_p denote the number of events in the proton and photon samples, respectively.

The Fisher discriminant analysis is applied to each energy bin separately, using the standard functions implemented in TMVA. In the TMVA training, the photon samples are used as signal and the proton samples are as background. All available events are used to train the MVA, with no additional event weights applied, i.e. the weight of each signal and background event is 1. The resulting Fisher coefficients in each energy bin are listed in Tab. 6.6. To estimate the uncertainty of the Fisher coefficients, a bootstrapping method is employed [Efr79]. In each

| | Fisher coefficients | | | |
Energy Bin	f_1	Δf_1	f_2	Δf_2
$18.0 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.1$	-4.8	0.2	0.0074	0.0003
$18.1 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.2$	-4.9	0.1	0.0082	0.0003
$18.2 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.3$	-5.0	0.1	0.0082	0.0003
$18.3 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.4$	-4.8	0.2	0.0092	0.0003
$18.4 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.5$	-5.0	0.2	0.0083	0.0003
$18.5 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.6$	-5.4	0.1	0.0093	0.0003
$18.6 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.7$	-5.7	0.2	0.0083	0.0003
$18.7 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.8$	-6.3	0.2	0.0092	0.0003
$18.8 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.9$	-6.4	0.2	0.0096	0.0003
$18.9 \leq \log_{10}(E\,[\mathrm{eV}]) < 19.0$	-6.7	0.2	0.0102	0.0003

Table 6.6.: Fisher coefficients determined in the individual energy bins according to Eq. 6.37. The uncertainties of the Fisher coefficients have been obtained using a bootstrapping method (cf. Figs. C.14 and C.15).

(a)   (b)
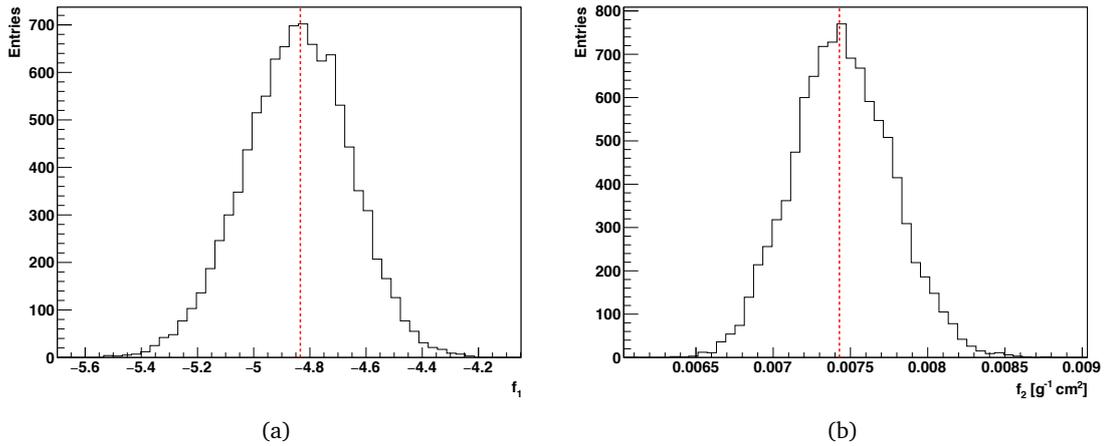
Figure 6.24.: Distributions of the Fisher coefficients (a) f_1 and (b)f_2 obtained through a bootstrapping method in the first energy bin ($18.0 \leq \log_{10}(E\,[\text{eV}]) < 18.1$). The dashed lines indicate the values of f_1 and f_2 that have been determined from the original samples.

energy bin, the photon and proton samples are resampled 10 000 times by randomly drawing, with replacement, events from the original samples. On these bootstrapping samples, the Fisher discriminant analysis is perfomed again, with the same settings, and the Fisher coefficients are extracted. The resulting distributions of the Fisher coefficients obtained from the bootstrapping samples are shown in Fig. 6.24 for the first energy bin and in App. C, Figs. C.14 and C.15, for the other energy bins. The distributions are centered around the values for f_1 and f_2 obtained from the original samples. The standard deviations of the distributions are used as estimates for the uncertainties of the Fisher coefficients.

With the coefficients listed in Tab. 6.6, the Fisher discriminant t is determined for each event in each energy bin according to Eq. 6.36. In Fig. 6.25, the distribution of the Fisher discriminant t is shown for the first energy bin ($18.0 \leq \log_{10}(E\,[\text{eV}]) < 18.1$). The corresponding distributions in the other energy bins can be found in App. C, Fig. C.16. The distributions for the photon and the proton samples separate clearly. The difference between the mean values of the distributions is about 2.4, while the width of the distributions in terms of standard deviation is 0.66 for the photon sample and 0.87 for the proton sample. The latter reflects the larger widths of the underlying distributions of F_γ and X_{max} for the proton sample. The merit factor of the Fisher discriminant as calculated from the mean values and widths of the distributions is 2.19, which is a significant improvement compared to the individual observables. This is corroborated by the background rejection of the observable as a function of the signal efficiency, shown in Fig. 6.26. $\varrho(\varepsilon)$ remains close to 100 % until about $\varepsilon = 80$ %. At a signal efficiency of 50 %, the background rejection is 99.76 %, which is also an significant improvement compared to F_γ and X_{max} taken individually.

In Tab. 6.7, the properties of the distributions of the Fisher discriminant t are summarized for all energy bins, including the merit factor η and the background rejection at a signal efficiency of 50 % as measures of the separation power of the observable. For both the photon and the proton sample, the mean values of the distributions increase, however, the increase is larger for the photon sample, hence also the difference between the mean values for the
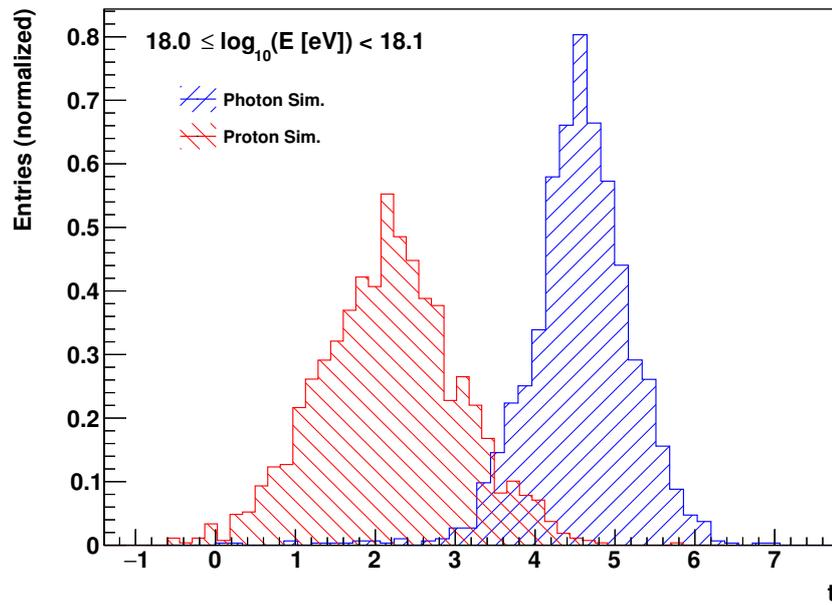
Figure 6.25.: Distributions of the Fisher discriminant t in the first energy bin ($18.0 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.1$) for the photon sample, shown in blue, and the proton sample, shown in red.
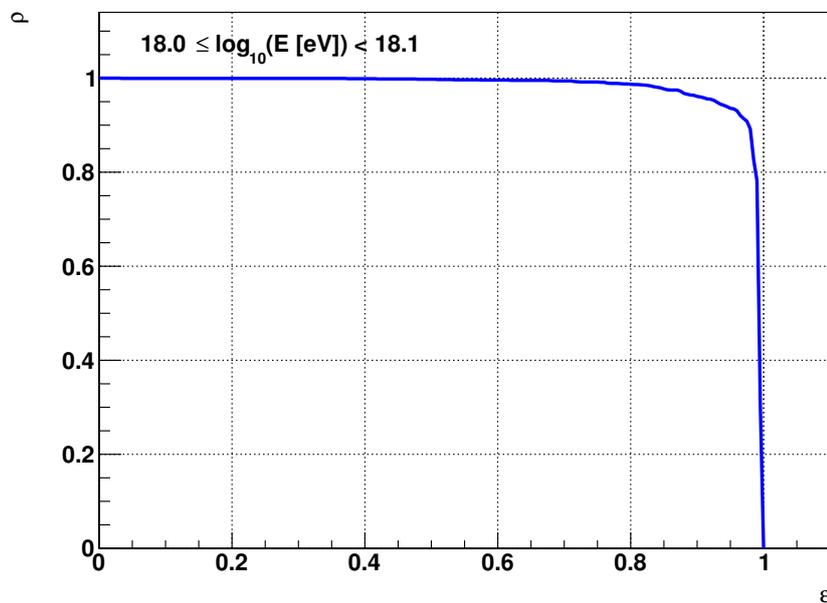
Figure 6.26.: Background rejection ϱ as a function of the signal efficiency ε for the Fisher discriminant t in the first energy bin ($18.0 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.1$).

Energy Bin	Photon		Proton		η	$\varrho(\varepsilon = 50\%)$
	μ	σ	μ	σ		
$18.0 \le \log_{10}(E\,[\text{eV}]) < 18.1$	4.56	0.66	2.17	0.87	2.19	99.76%
$18.1 \le \log_{10}(E\,[\text{eV}]) < 18.2$	5.35	0.66	2.76	0.93	2.28	99.55%
$18.2 \le \log_{10}(E\,[\text{eV}]) < 18.3$	5.45	0.70	2.59	0.97	2.39	99.52%
$18.3 \le \log_{10}(E\,[\text{eV}]) < 18.4$	6.52	0.72	3.33	1.04	2.53	99.85%
$18.4 \le \log_{10}(E\,[\text{eV}]) < 18.5$	5.84	0.74	2.56	1.04	2.56	99.90%
$18.5 \le \log_{10}(E\,[\text{eV}]) < 18.6$	6.67	0.80	2.83	1.13	2.77	99.91%
$18.6 \le \log_{10}(E\,[\text{eV}]) < 18.7$	5.77	0.80	1.88	1.14	2.79	99.89%
$18.7 \le \log_{10}(E\,[\text{eV}]) < 18.8$	6.49	0.83	1.99	1.25	3.00	99.97%
$18.8 \le \log_{10}(E\,[\text{eV}]) < 18.9$	6.89	0.88	2.16	1.26	3.08	99.97%
$18.9 \le \log_{10}(E\,[\text{eV}]) < 19.0$	7.48	0.97	2.38	1.27	3.20	100.00%

Table 6.7.: Properties of the distributions of the Fisher discriminant t for the photon and proton samples and the separation power of the observable. Listed are the mean values μ of the distributions and their widths in terms of standard deviation σ. For each energy bin, the merit factor η has been calculated using these parameters. In addition, the background rejection ϱ at a signal efficiency of $\varepsilon = 50\%$ has been derived from the distributions (cf. App. C, Fig. C.16).

(a)   (b)
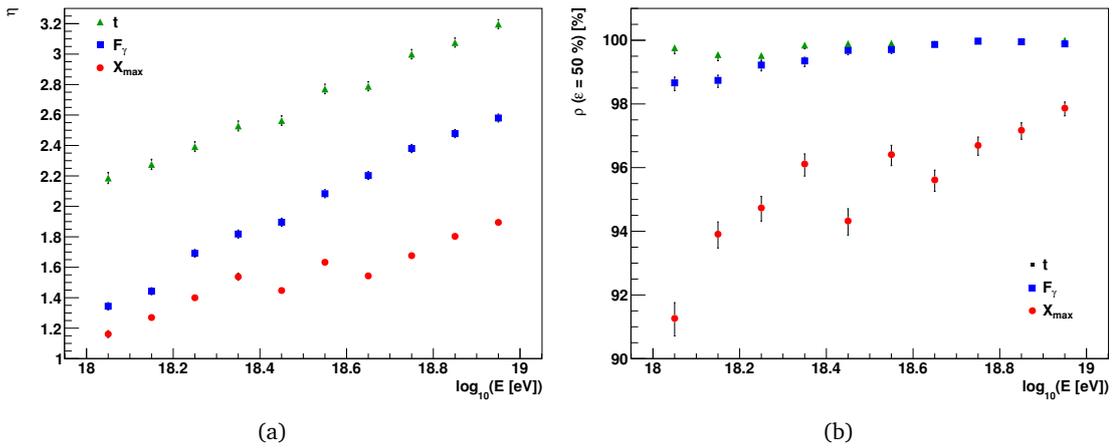
Figure 6.27.: Measures of the separation power of the observable t as a function of energy. (a) Merit factor η, (b) Background rejection ϱ at a signal efficiency of $\varepsilon = 50\%$. The uncertainties of the individual data points have been estimated using error propagation in the case of η and according to [Pat04] in the case of $\varrho(\varepsilon = 50\%)$. In addition to the data points for t, shown as green triangles, the data points for F_γ and X_{max} (cf. Fig. 6.3) are included as blue squares and red circles, respectively.

two samples gets larger. At the same time, the widths of the distributions increase as well— however, not as strongly as the mean values. Consequently, the overall separation power of the observable increases. In Fig. 6.27, η and $\varrho(\varepsilon = 50\%)$ are shown as a function of energy for the Fisher discriminant t as well as the individual observables F_γ and X_{max}. The merit factor of t increases almost linearly from about 2.2 around 1 EeV to about 3.2 around 10 EeV. At all energies, the merit factor of the combined observable t is significantly larger than the

merit factor of the individual observables F_γ and X_{max}. In terms of $\varrho(\varepsilon = 50\%)$, the increase in separation power of the Fisher discriminant compared to the individual observables is most noticeable at lower energies. At higher energies, $\varrho(\varepsilon = 50\%)$ for F_γ is already very close to 100%, but it decreases toward lower energies. For t, $\varrho(\varepsilon = 50\%)$ is above 99% even around 1 EeV.

6.3.2. Parameterization of the Fisher Coefficients as a Function of Energy

In the previous section, the Fisher discriminant t has been introduced as an additional observable combining F_γ and X_{max}. However, t is not a continuous observable, as it is derived from Fisher discriminant analyses performed in discrete energy bins. Hence, the Fisher coefficients f_1 and f_2 are now parameterized as a function of energy to obtain a continuous observable t_{par}, which is defined similarly to Eq. 6.36:

$$t_{par} = f_1(E) \times F_\gamma + f_2(E) \times X_{max}. \tag{6.39}$$

In contrast to the observable t, which is only piecewise defined in the discrete energy bins, t_{par} is defined continuously over the whole energy range.

The Fisher coefficients f_1 and f_2 are shown in Fig. 6.28 as a function of energy. The apparent kink in $f_1(E)$ around $\log_{10}(E\,[\text{eV}]) \approx 18.4$ is a consequence of the within-class scatter matrix $\mathcal{W}$. All other terms used to determine f_1 (cf. Eq. 6.37) show a linear behaviour with energy. From a phenomenological point of view, a change in the behaviour of the input observables is expected in this energy region because of the changes in the efficiencies of the detectors. The energy dependence of the two Fisher coefficients f_1 and f_2 is modeled by linear functions: a piecewise linear function with two sections in the case of f_1 and a single linear function in the case of f_2:

$$f_1(E) = \begin{cases} p_1, & \text{for } \log_{10}(E\,[\text{eV}]) \leq p_0, \\ p_1 + p_2\left(\log_{10}(E\,[\text{eV}]) - p_0\right), & \text{for } \log_{10}(E\,[\text{eV}]) > p_0, \end{cases} \tag{6.40}$$

$$f_2(E) = q_0 + q_1 \log_{10}(E\,[\text{eV}]).$$

(a)   (b)
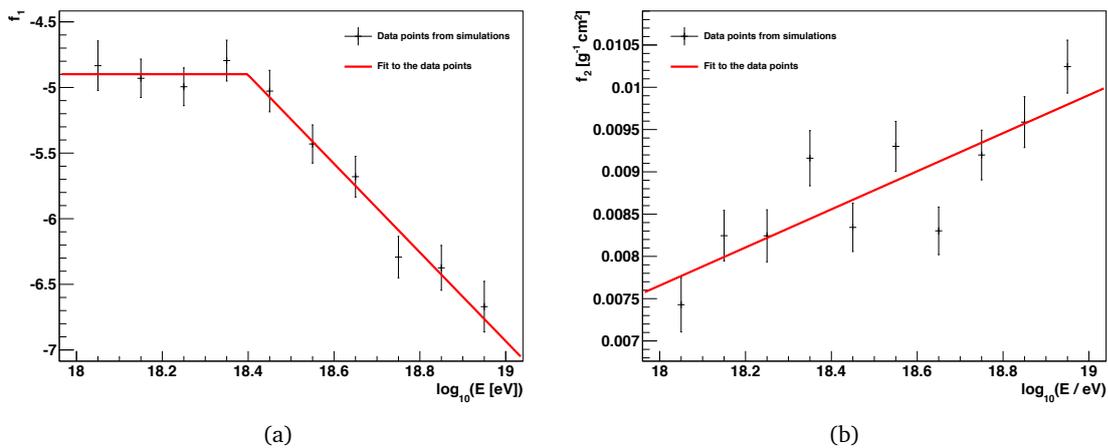
Figure 6.28.: The Fisher coefficients f_1 and f_2 shown as a function of energy (cf. Tab. 6.6). Also shown are the fits to the data points, denoted by the red lines.

These functional forms have been chosen phenomenologically, because they contain only a small number of free parameters. In the case of f_1, the piecewise linear function also exhibits a continuous behaviour even at the transition point, which is described by the parameter p_0. A fit of these functions to the data points shown in Fig. 6.28 yields the following values for the parameters:

$$
\begin{aligned}
p_0 &= 18.40 \pm 0.04, \\
p_1 &= -4.90 \pm 0.08, \\
p_2 &= 3.4 \pm 0.4, \\
q_0 &= (-0.033 \pm 0.006)\,\mathrm{g}^{-1}\,\mathrm{cm}^2, \\
q_1 &= (0.0023 \pm 0.0003)\,\mathrm{g}^{-1}\,\mathrm{cm}^2.
\end{aligned}
\tag{6.41}
$$

In the next step, this parameterization is applied to the main simulation samples to determine whether there is any loss in the separation power of the observable. In each energy bin, the observable t_{par} is determined for each event according to Eq. 6.39, using the parameterization of the Fisher coefficients from Eq. 6.40. As an estimator for the energy of the primary particle in each individual event, the "photon energy" E_γ is used:

$$
E_\gamma = (1 + 1\,\%)\,E_{\mathrm{cal}},
\tag{6.42}
$$

with the calorimetric energy E_{cal} and an invisible energy correction of $1\,\%$ (cf. Sec. 4.5.1). This estimator for the primary energy is used instead of the energy reconstructed by the standard $\overline{\text{Off}}\underline{\text{line}}$ algorithms because it more closely depicts the true primary energy in case of photons as primary particles. The standard reconstruction of the primary energy assumes an invisible energy correction appropriate for air showers initiated by protons and nuclei and hence overestimates the primary energy for photon-induced air showers.

The distributions of t_{par} for the photon and proton samples are shown for the first energy bin ($18.0 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.1$) in Fig. 6.29. The corresponding distributions for the other energy bins can be found in App.C, Fig. C.18. Qualitatively, no significant differences between the distributions of t_{par} and the distributions of t (Fig. 6.25) are visible. The mean values of the distributions increase, for the proton distribution from 4.56 to 4.86 and for the photon distribution from 2.17 to 2.33. The widths of the distributions remain approximately the same: for both distributions, the width increases only by 0.06. The merit factor of t_{par} in this energy bin is 2.16, which is slightly lower than the merit factor of t (2.19). Similarly, $\varrho(\varepsilon = 50\,\%)$ is slightly smaller for t_{par} (99.59 %) than for t (99.76 %). These differences, however, are within the range of statistical fluctuations.The properties of the distributions of t_{par} are listed in Tab. 6.8 for all energy bins. Also listed are η and $\varrho(\varepsilon = 50\,\%)$ for each energy bin. Both measures of separation power are shown in Fig. 6.31 in comparison to η and $\varrho(\varepsilon = 50\,\%)$ for t (cf. Tab.6.7). No significant differences between the two observables are visible. The data points are well compatible within their respective uncertainties. Hence, the differences in the performance of the parameterized observable t_{par} and the original observable t can be seen as negligible. Therefore, t_{par} is used hereafter as the main observable for the discrimination of photon-induced and proton-induced air showers.
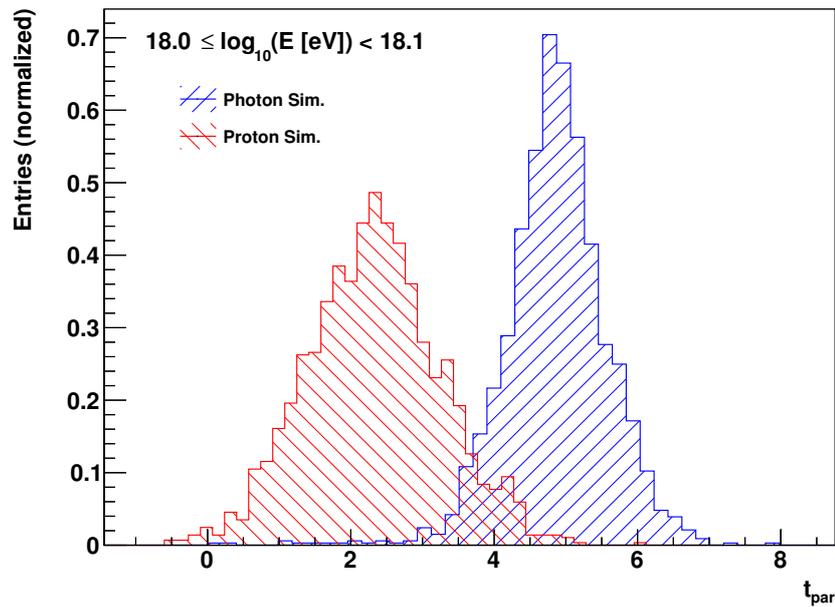
Figure 6.29.: Distributions of the observable t_{par} in the first energy bin ($18.0 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.1$) for the photon sample, shown in blue, and the proton sample, shown in red.
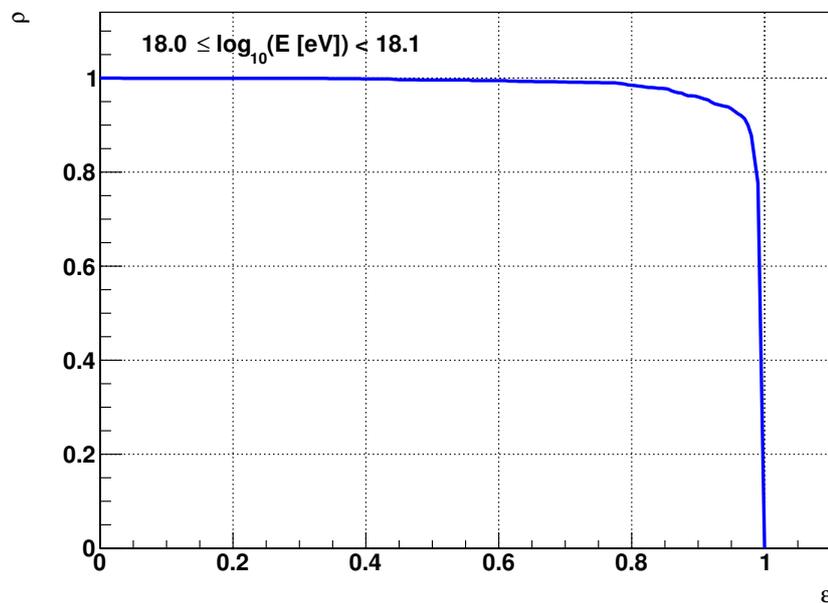
Figure 6.30.: Background rejection ϱ as a function of the signal efficiency ε for the observable t_{par} in the first energy bin ($18.0 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.1$).

	Photon		Proton			
Energy Bin	μ	σ	μ	σ	η	$\varrho(\varepsilon = 50\%)$
$18.0 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.1$	4.86	0.72	2.33	0.93	2.16	99.59%
$18.1 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.2$	5.17	0.69	2.54	0.94	2.25	99.42%
$18.2 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.3$	5.48	0.74	2.58	0.98	2.36	99.52%
$18.3 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.4$	5.86	0.73	2.65	1.05	2.50	99.81%
$18.4 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.5$	6.13	0.78	2.78	1.05	2.57	99.76%
$18.5 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.6$	6.32	0.79	2.58	1.09	2.78	99.94%
$18.6 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.7$	6.52	0.85	2.50	1.15	2.80	99.89%
$18.7 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.8$	6.70	0.84	2.32	1.19	3.01	99.97%
$18.8 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.9$	6.86	0.89	2.15	1.24	3.09	99.97%
$18.9 \leq \log_{10}(E\,[\mathrm{eV}]) < 19.0$	7.02	0.96	1.99	1.24	3.20	100.00%

Table 6.8.: Properties of the distributions of t_{par} for the photon and proton samples and the separation power of the observable. Listed are the mean values μ of the distributions and their widths in terms of standard deviation σ. For each energy bin, the merit factor η has been calculated using these parameters. In addition, the background rejection ϱ at a signal efficiency of $\varepsilon = 50\%$ has been derived from the distributions (cf. App. C, Fig. C.18).

(a)   (b)
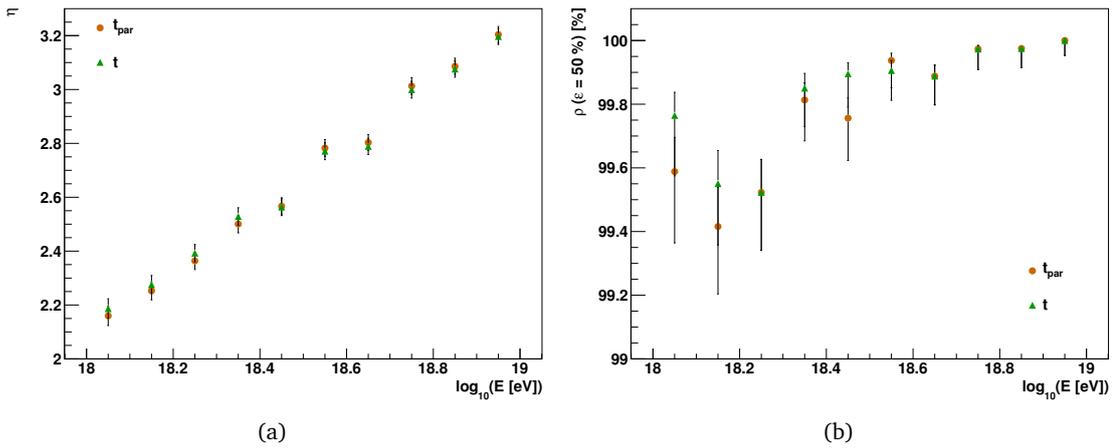
Figure 6.31.: Measures of the separation power of the observable t_{par} as a function of energy. (a) Merit factor η, (b) Background rejection ϱ at a signal efficiency of $\varepsilon = 50\%$. The uncertainties of the individual data points have been estimated using error propagation in the case of η and according to [Pat04] in the case of $\varrho(\varepsilon = 50\%)$. In addition to the data points for t_{par}, shown as orange circles, the data points for t are included as green triangles.

6.3.3. Defining the Photon Candidate Cut

In order to classify air shower events from data—where the primary particle type is necessarily not known—into photon-like and proton-like events using the observables that have been discussed in the preceeding sections, an a priori photon candidate cut is defined in this section. To be consistent with [Set11], the photon candidate cut is based on the median of the photon distribution of the Fisher discriminant, which is known from simulations. In Fig. 6.32,

the median q_{50} of the t_{par} distributions for the photon sample is shown for the individual energy bins. The uncertainties of $q_{50,\gamma}\{t_{par}\}$ have been determined in the individual energy bins using the same bootstrapping method that was already described in Sec. 6.3.1. The resulting bootstrapping distributions can be found in App. C, Fig. C.20. The linear dependence of $q_{50,\gamma}\{t_{par}\}$ on the energy is obvious. The change in the slope around $\log_{10}(E\,[\mathrm{eV}]) \approx 18.4$ is due to the parameterized Fisher coefficient f_1 (cf. Fig. 6.28 (a)). Below this point, $q_{50,\gamma}\{t_{par}\}$ is not constant—like $f_1(E)$—because the distributions of the basic observables F_γ and X_{max} change as well. The energy dependence of $q_{50,\gamma}\{t_{par}\}$ is parameterized by a piecewise linear function with two sections:

$$q_{50,\gamma}\{t_{par}\}(E) = \begin{cases} k_1 + k_2 \log_{10}(E\,[\mathrm{eV}]), & \text{for } \log_{10}(E\,[\mathrm{eV}]) \le k_0, \\ k_1 + k_2 k_0 + k_3\,(\log_{10}(E\,[\mathrm{eV}]) - k_0), & \text{for } \log_{10}(E\,[\mathrm{eV}]) > k_0. \end{cases} \quad (6.43)$$

A fit to the data points shown in Fig. 6.32 yields the following values for the parameters:

$$\begin{aligned} k_0 &= 18.45 \pm 0.01, \\ k_1 &= -53.4 \pm 0.9, \\ k_2 &= 3.23 \pm 0.05, \\ k_3 &= 1.83 \pm 0.05. \end{aligned} \quad (6.44)$$

An air shower event is identified as a photon candidate event, if t_{par} is larger than $q_{50,\gamma}\{t_{par}\}$. Both t_{par} and $q_{50,\gamma}\{t_{par}\}$ are determined for this event using the photon energy E_γ.

In the next step, the performance of this photon candidate cut in identifying photon-induced air shower events and rejecting air shower events initiated by protons is evaluated. In Fig. 6.33, the observable t_{par} is shown versus the energy estimator E_γ in a scatter plot, including all energy bins in the two main simulation samples. The photon candidate cut defined by Eq. 6.43 is also included. For the full photon sample, 14 160 events out of 28 850 are identified as photon candidates. This represents a selection efficiency of 49.1 %. The selection efficiency is not exactly 50 %, which would be expected in principle from the choice of the median of the photon distribution as the candidate cut, because the median is parameterized before it is used. However, the loss in selection efficiency is small because the parameterization describes the energy dependence of the median well. For the proton sample, on the other hand, 37 events out of 30 744 are identified as photon candidates. Hence, a fraction of 0.12 % of all proton-induced air shower events in the energy range between 1 EeV and 10 EeV are misidentified as photon candidates. Among these misidentified events, however, a number of events can be found which are photon-like in one basic observable but not the other. For example, an event can have a very deep X_{max}, which would hint at a primary photon, but at the same time, F_γ can be too large for this event to be considered induced by a photon. Since F_γ and X_{max} are combined into a single observable t_{par}, overall this event can still pass the photon candidate cut. To avoid such misidentifications, the photon candidate cut is extended by requiring that an air shower event not only passes the cut on $q_{50,\gamma}\{t_{par}\}$, but it has to be sufficiently photon-like in both of the basic observables i.e. it has to pass additional cuts on F_γ and X_{max}. These cuts are, analogous to the cut on t_{par}, based on the distributions of the respective observables for the photon sample, however they are not as strict. Instead of the median, the 5-quantile is used for X_{max} and the 95-quantile for F_γ. The two quantiles $q_{5,\gamma}\{X_{max}\}$ and $q_{95,\gamma}\{F_\gamma\}$ are shown for the individual energy bins in Fig. 6.34. The uncertainties of the quantiles have
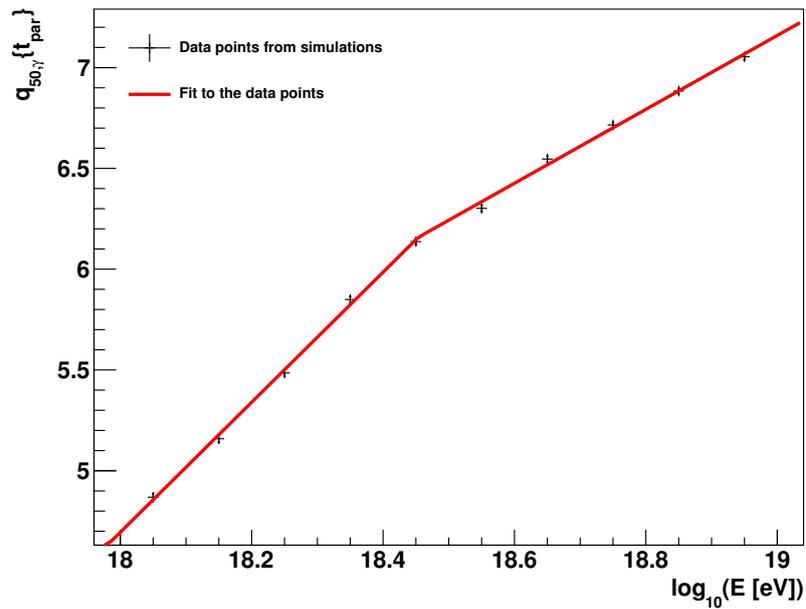
Figure 6.32.: The median q_{50} of the t_{par} distributions for the photon sample as a function of energy. The uncertainties on the data points have been determined using a bootstrapping method, however, they are too small to be visible at this scale. The red line denotes a fit to the data points, which is used as the photon candidate cut.
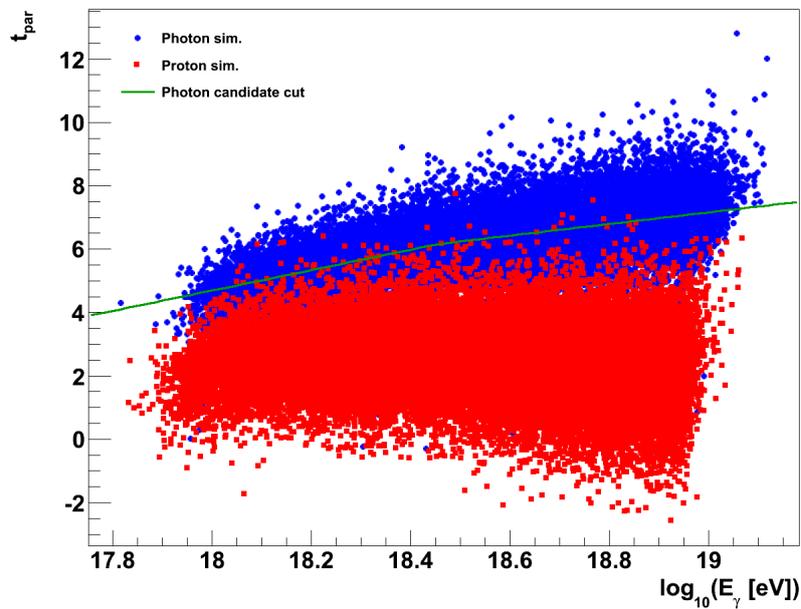
Figure 6.33.: Scatter plot of t_{par} versus E_γ for the photon sample (blue circles) and the proton sample (red squares). The green line denotes the photon candidate cut.

again been determined using the bootstrapping method described before. The resulting distributions, which were used to estimate the uncertainties, can be found in App. C, Figs. C.21 and C.22. The energy dependence of the quantiles is parameterized by simple linear functions, a piecewise linear function with two segments for $q_{95,\gamma}\{F_\gamma\}$ and a single linear function for $q_{5,\gamma}\{X_{\max}\}$:

$$q_{95,\gamma}\{F_\gamma\}(E) = \begin{cases} l_1 + l_2 \log_{10}(E\,[\text{eV}]), & \text{for } \log_{10}(E\,[\text{eV}]) \le l_0, \\ l_1 + l_2 l_0, & \text{for } \log_{10}(E\,[\text{eV}]) > l_0. \end{cases} \tag{6.45}$$

$$q_{5,\gamma}\{X_{\max}\}(E) = m_0 + m_1 \log_{10}(E\,[\text{eV}]).$$

A fit to the data points shown in Fig. 6.34 yields the following values for the parameters:

$$\begin{aligned} l_0 &= 18.58 \pm 0.05, \\ l_1 &= 2.4 \pm 0.3, \\ l_2 &= -0.10 \pm 0.02, \\[4pt] m_0 &= (-1400 \pm 40)\,\text{g cm}^{-2}, \\ m_1 &= (120 \pm 2)\,\text{g cm}^{-2}. \end{aligned} \tag{6.46}$$

Now, an air shower event is identified as as photon candidate event, if it passes the cuts on the individual observables F_γ and $X_{\max}$ in addition to the cut on t_{par}. In contrast to the cut on t_{par}, where it is only required that t_{par} is larger than $q_{50,\gamma}\{t_{\text{par}}\}$, also the uncertainties of F_γ and $X_{\max}$ are taken into account for the photon candidate cut in order to remove events where the uncertainty on the individual observables is too large for the event to be safely considered as a photon candidate event. Hence, it is required that $F_\gamma + \Delta F_\gamma$ is smaller than $q_{95,\gamma}\{F_\gamma\}$ and $X_{\max} - \Delta X_{\max}$ is larger than $q_{5,\gamma}\{X_{\max}\}$. As before, E_γ is used to determine $q_{95,\gamma}\{F_\gamma\}$ and $q_{5,\gamma}\{X_{\max}\}$.

(a)   (b)
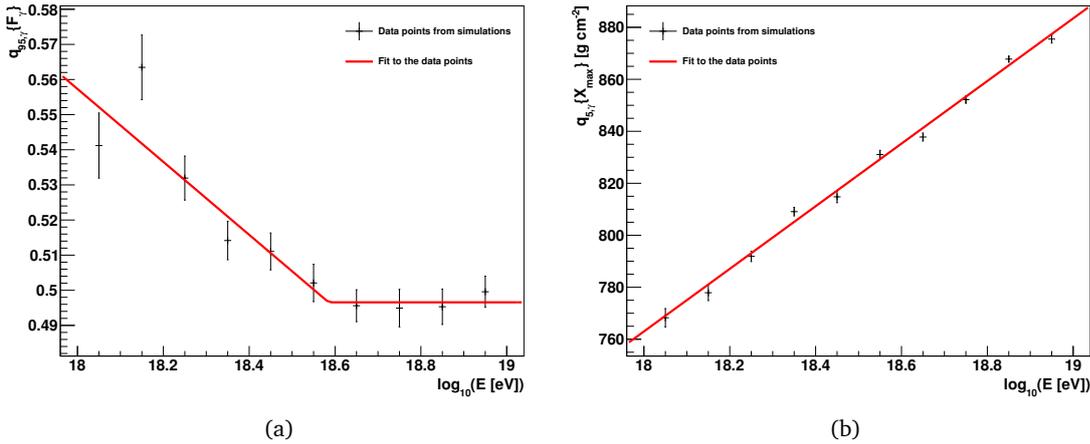
Figure 6.34.: (a) The 95-quantile q_{95} of the F_γ distributions and (b) the 5-quantile q_5 of the $X_{\max}$ distributions for the photon sample as functions of energy. The uncertainties on the data points have been determined using a bootstrapping method. The red lines denote fits to the data points.
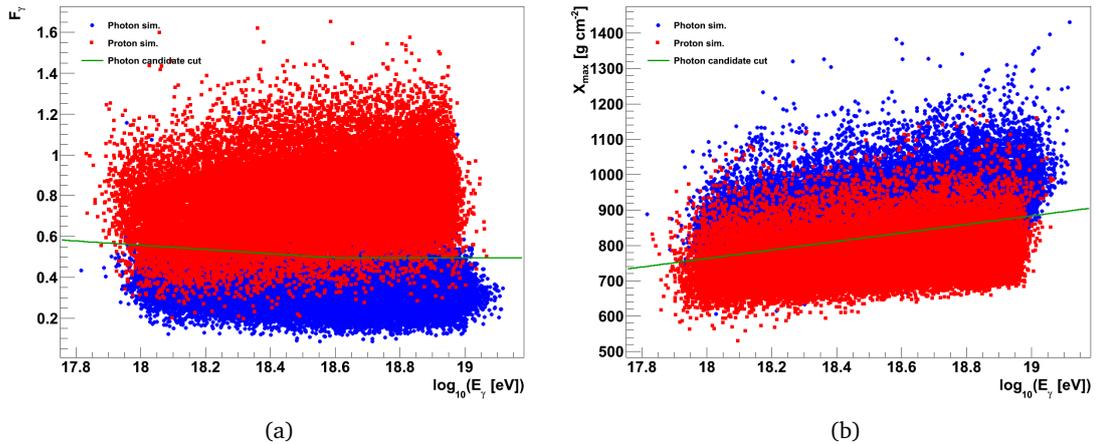
Figure 6.35.: Scatter plots of (a) F_γ versus E_γ and (b) X_{max} versus E_γ, for the photon sample (blue circles) and the proton sample (red squares). The green lines denote the photon candidate cut in each observable.

To visualize these additional cuts, F_γ and X_{max} are shown in Fig. 6.35 as scatter plots versus E_γ. As both cuts have been chosen rather loose, a large number of events from the proton sample pass both cuts individually. Combining the two cuts with the previous cut on t_{par}, however, significantly reduces the number of misidentified events: only 9 out of the 30 744 events from the proton sample, i.e. 0.03 %, pass all three cuts and are thus photon candidate events. This represents a reduction of about 75 % of the number of misidentified proton events compared to the number obtained when using only the cut on t_{par}. This significant improvement in the rejection of proton events comes at the cost of losing a fraction of the photon events: only 11 672 out of the 28 850 events from the photon sample, i.e. 40.5 %, are identified as photon candidate events. About 18 % of the events that were previously identified as photon candidate events, using only the cut on t_{par}, are rejected by the additional cuts. However, this loss in the selection efficiency of photon events is fully justified by the significant increase in the rejection efficiency of proton events.

6.4. Further Studies Using Simulations

In the following sections, several additional studies related to the analysis presented in the preceding sections are discussed. The motivation for these studies lies in the fact that the analysis has been optimized under some specific conditions. In particular, the analysis was optimized for the separation of air showers induced by photons and those initiated by protons—under the assumption of a specific hadronic interaction model in the simulations—in the energy range from 1 EeV to 10 EeV. In the following sections, it is studied whether the performance of the analysis is weaker when these conditions are not met, for example at higher energies or when air showers initiated by nuclei are taken into account. It is also studied whether the choice of a specific hadronic interaction model has an impact on the results of the analysis.

For each of these studies, several smaller simulation samples have been generated. In general, the same settings have been used as for the main samples (cf. Secs. 5.1.1 and 5.1.2),

except when noted. The event selection (see Sec. 5.3.1) remains the same as for the main samples, including the additonal selection criteria listed in Sec. 6.2.6.

6.4.1. Behavior Above 10 EeV

The analysis presented in the preceding sections is optimized for the energy range from $1\,\mathrm{EeV}$ to $10\,\mathrm{EeV}$. However, when upper limits on the integral flux or fraction of UHE photons are to be determined, these limits comprise also the region above $10\,\mathrm{EeV}$ (cf. e.g. Fig. 3.5). The simplest approach would therefore be to extend the energy range in which the analysis is optimized and fully determine the parameterizations, e.g. of the Fisher discriminant, above $10\,\mathrm{EeV}$. However, this approach is not particularly efficient, as the computing resources needed for the simulations significantly increase with energy. At the same time, the number of events that are expected in data in this energy region is comparatively small, due to the steeply falling energy spectrum of primary cosmic rays, about $1\,\%$ of the number of events expected in the energy range from $1\,\mathrm{EeV}$ to $10\,\mathrm{EeV}$. Hence, it was decided not to follow this approach, but rather extrapolate the results obtained in the energy range from $1\,\mathrm{EeV}$ to $10\,\mathrm{EeV}$ to higher energies. In order to facilitate this extrapolation, it was ensured that the functions used in the different parameterizations can be linearly extrapolated. In this section, the performance of the analysis in distinguishing air shower events induced by photons from those events that have been initiated by protons in the energy region above $10\,\mathrm{EeV}$ is examined. For this purpose, several small test samples have been generated, with photons and protons as primary particles, in two energy bins: $19.0 \leq \log_{10}(E\,[\mathrm{eV}]) < 19.1$ and $19.5 \leq \log_{10}(E\,[\mathrm{eV}]) < 19.6$. Above $\log_{10}(E\,[\mathrm{eV}]) = 19.6$, the preshower effect becomes more important (see Sec. 3.3.1), which leads, on average, to a shallower shower development due to the UHE photons that convert in the geomagnetic field. The expected contribution of events above $\log_{10}(E\,[\mathrm{eV}]) = 19.6$ amounts to $0.03\,\%$ of the number of events expected in the energy range from $1\,\mathrm{EeV}$ to $10\,\mathrm{EeV}$. Therefore, a dedicated study of the influence of the preshower effect on air showers with primary energies above $\log_{10}(E\,[\mathrm{eV}]) = 19.6$ has not been done.

After the event selection (cf. App. B, Tabs. B.3 and B.4), 1902 and 1887 events are selected from the photon and the proton samples, respectively, in the energy bin $19.0 \leq \log_{10}(E\,[\mathrm{eV}]) < 19.1$ and 1017 and 1542 events in the energy bin $19.5 \leq \log_{10}(E\,[\mathrm{eV}]) < 19.6$. In the first energy bin, the selection efficiencies are comparable to those of the last energy bin of the main samples. In the second energy bin, the largest effect on the total selection efficiency comes for the photon sample from the requirement that the shower maximum has to be observed within the field of view of the FD telescopes, since the average $X_{\max}$ of photon-induced air showers at these energies is already close to the lower border of the field of view of the FD telescopes. For the proton sample, where the air shower events have on average a shallower longitudinal profile, this requirement discards less events. However, this is compensated by the fiducial volume cut, which was included to ensure that the total selection efficiency is less dependent on the type of the primary particle.

In Fig. 6.36, the results of the analysis when applied to these samples are shown. Qualitatively, the extrapolation of the photon candidate cut in the three observables works well. Of the 1902 events in the photon sample in the first energy bin, 756 are identified as photon candidate events, which corresponds to a photon selection efficiency of $39.7\,\%$. Likewise, in the second energy bin, 513 out of the 1017 events are identified as photon candidate events

(a)

(b)

(c)
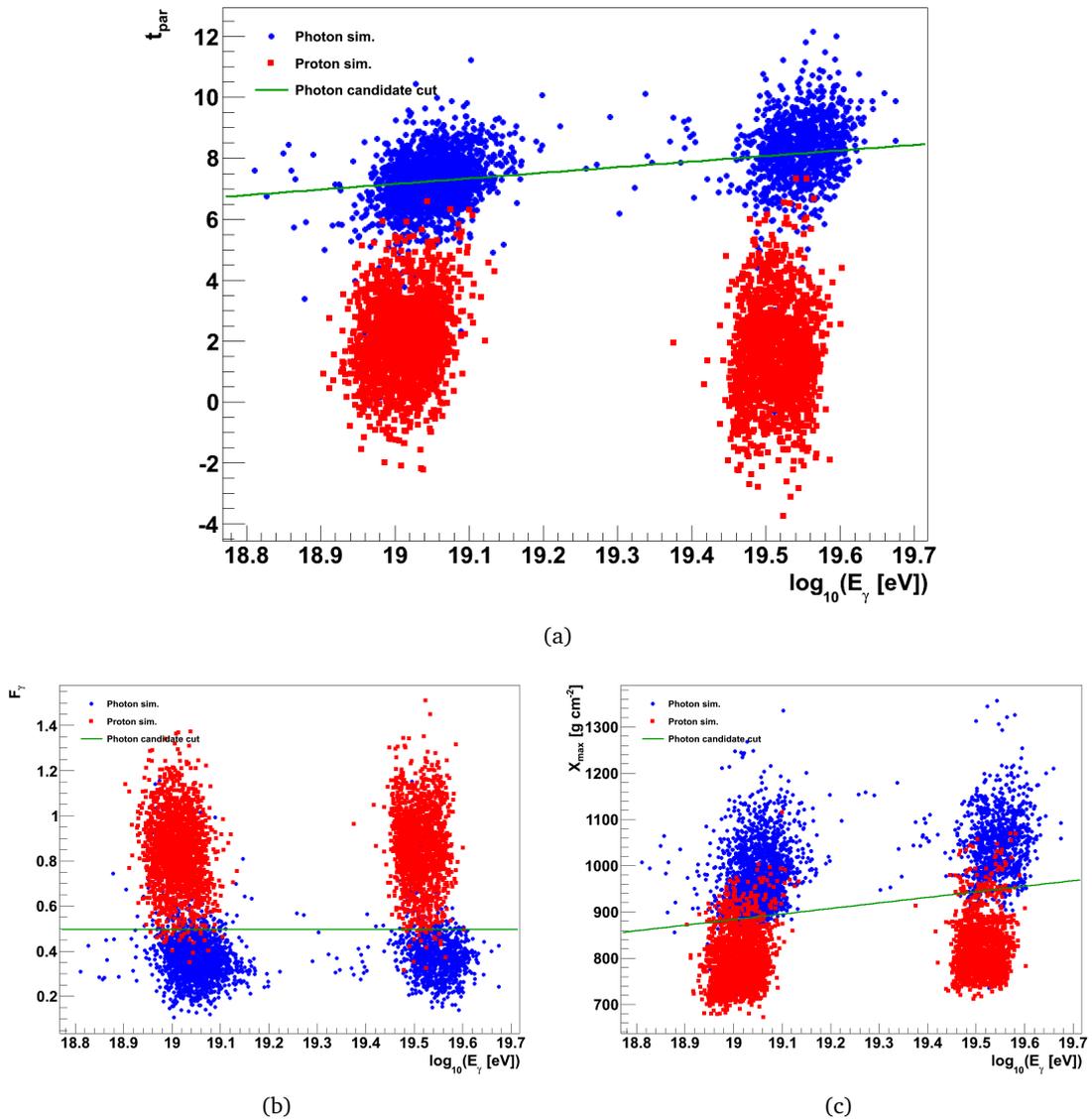
Figure 6.36.: Scatter plots of (a) t_{par} versus E_γ, (b) F_γ versus E_γ and (c) X_{max} versus E_γ, above 10 EeV, for the photon sample (blue circles) and the proton sample (red squares). The green lines denote the extrapolated photon candidate cut in each observable.

(50.4 %). The photon selection efficiency is therefore for the extrapolated analysis almost as good as in the main energy range. In the second energy bin, the selection efficiency is larger than in the main energy range. For the proton sample, no event in both energy bins passes the photon candidate cut in all three observables. Using Poisson statistics, it can therefore be estimated that the fraction of proton-induced air shower events that are misidentified as photon candidate events is, at a confidence level of 68 %, smaller than 0.06 % in the first energy bin and 0.07 % in the second energy bin. These numbers are comparable to the fraction of misidentified events in the main energy range.

Overall, it can be concluded that the extrapolation of the analysis to higher energies, which

is required to determine upper limits on the integral flux or fraction of UHE photons, does not lead to a worse performance in the separation of photon-induced and proton-induced air shower events. In particular, the extrapolation does not significantly increase the risk of misidentified proton-induced air shower events, while even increasing the fraction of photon-induced events that are correctly identified as photon candidate events.

6.4.2. Application to Air Shower Events Initiated by Nuclei

The analysis presented in this thesis has been optimized under the assumption that the background, from which photon-induced air shower events are to be distinguished, is entirely made up of air shower events that have been initiated by protons. It is obvious that this assumption is somewhat simplifying, as measurements of the elemental composition of UHECR indicate a mixed composition, which is dominated by protons as well as by other light nuclei (cf. Sec. 2.2.2). At the same time, it is expected from the superposition model for extensive air showers initiated by nuclei, which has been verified by MC simulations, that the average $X_{\max}$ gets smaller with the nuclear mass of the primary particle, while the average N_μ increases (see Sec. 2.3.2). Hence, the likelihood that an air shower initiated by a nucleus resembles a photon-induced air shower should decrease significantly with increasing nuclear mass. The assumption of a pure proton background represents therefore the worst case, where the maximum number of misidentified background events is expected. In the brief MC study discussed in this section, it is checked whether the fraction of air shower events misidentified as photon candidate events is indeed negligible in the case of primary nuclei, as would be expected from the superposition model.

In this MC study, different nuclei have been used as primary particles: helium (^{4}He, CORSIKA particle code 402), oxygen (^{16}O, CORSIKA particle code 1608), and iron (^{56}Fe, CORSIKA particle code 5626). For each primary particle type, three energy bins have been considered: $18.0 \leq \log_{10}(E\,[\text{eV}]) < 18.1$, $18.5 \leq \log_{10}(E\,[\text{eV}]) < 18.6$, and $18.9 \leq \log_{10}(E\,[\text{eV}]) < 19.0$. The selection efficiencies after each of the selection criteria discussed in Secs. 5.3.1 and 6.2.6 has been applied are listed in App. B, Tabs. B.3 and B.4. In general, the selection efficiencies in these samples are comparable to the efficiencies obtained with the proton sample (see App. B, Tabs. B.1 and B.2). The selection criterion with the largest impact is the fiducial volume cut. In the first energy bin, also the cut on the absolute uncertainty of F_γ reduces the samples considerably, by about 50 %.

In Fig. 6.37, the results of the analysis when applied to these samples are shown as scatter plots of the individual observables t_{par}, F_γ and $X_{\max}$ versus E_γ. Qualitatively, the expectations from the superposition model are confirmed. In the observable F_γ, the distributions shift toward higher values—and thus away from the photon candidate cut in this observable—with increasing primary mass. A similar behaviour is found for the observable $X_{\max}$, where the distributions are shifted toward smaller values, again away from the photon candidate cut. It is also interesting to note that the $X_{\max}$ distributions are more concentrated at higher primary masses, and fluctuations toward a deeper shower development appear less often. In the observable t_{par}, the trend that the distributions move away from the photon candidate cut is even more pronounced. Overall, not a single event in all three samples passes the photon candidate cut in all three observables. Taking the different sample sizes into account, the uncertainties on the fraction of photon candidate events has been calculated according

(a)

(b)

(c)
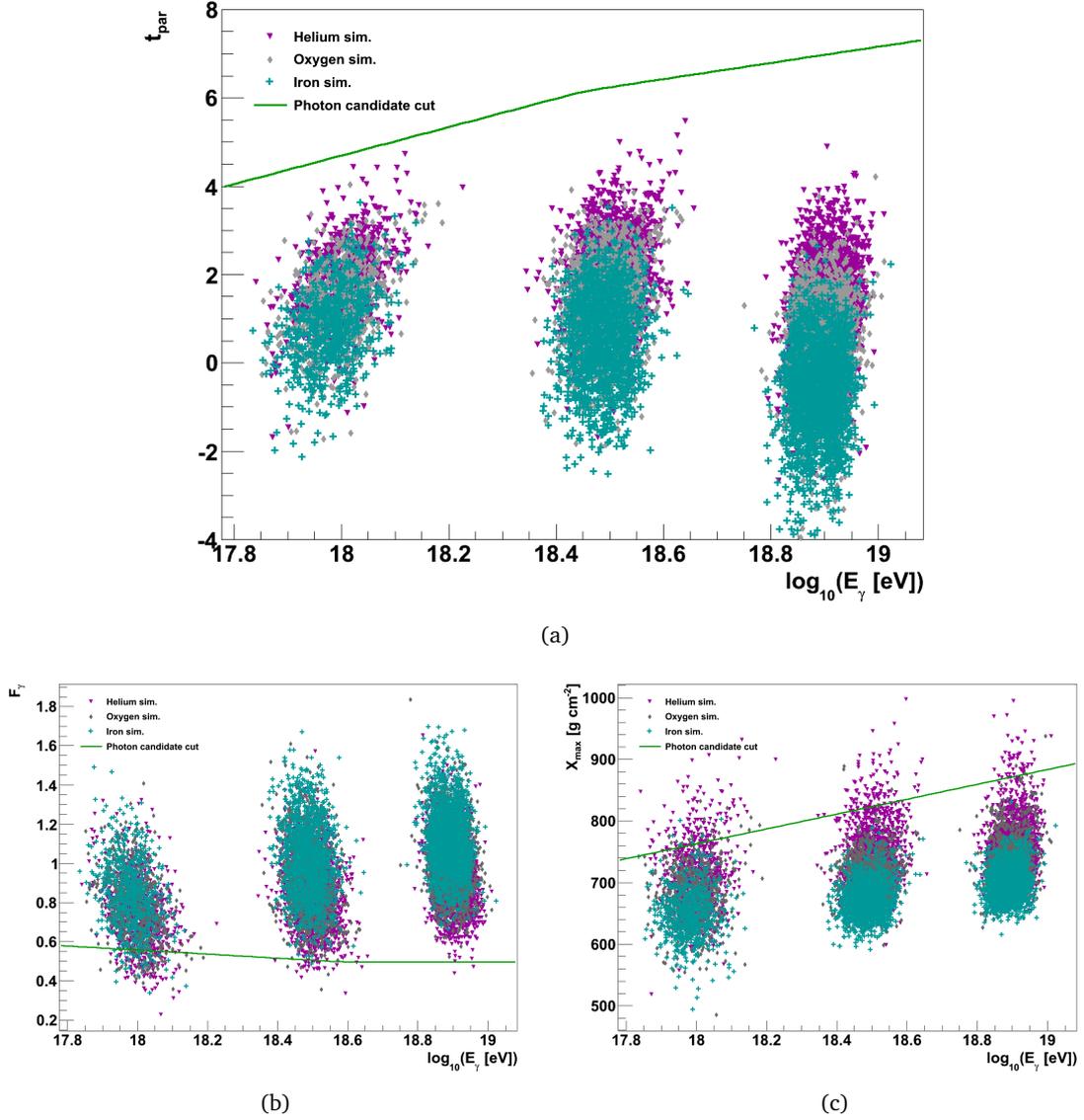
Figure 6.37.: Scatter plots of (a) t_{par} versus E_γ, (b) F_γ versus E_γ and (c) X_{max} versus E_γ, for the helium (violet triangles), oxygen (gray diamonds) and iron samples (turquois crosses), respectively. The green lines denote the photon candidate cut in each observable.

Energy Bin	Helium	Oxygen	Iron
$18.0 \leq \log_{10}(E\,[\text{eV}]) < 18.1$	$\left(0^{+0.25}_{-0}\right)$ %	$\left(0^{+0.26}_{-0}\right)$ %	$\left(0^{+0.33}_{-0}\right)$ %
$18.5 \leq \log_{10}(E\,[\text{eV}]) < 18.6$	$\left(0^{+0.13}_{-0}\right)$ %	$\left(0^{+0.14}_{-0}\right)$ %	$\left(0^{+0.15}_{-0}\right)$ %
$18.9 \leq \log_{10}(E\,[\text{eV}]) < 19.0$	$\left(0^{+0.10}_{-0}\right)$ %	$\left(0^{+0.10}_{-0}\right)$ %	$\left(0^{+0.11}_{-0}\right)$ %

Table 6.9.: Fraction of events in the helium, oxygen and iron samples that pass the photon candidate cut. The asymmetric uncertainties on the fractions have been calculated according to [Pat04].

to [Pat04], cf. Tab. 6.9. Taking also into account the distributions shown in Fig. 6.37 and comparing them to the corresponding distributions for the proton sample (Figs. 6.33 and 6.35), it can be stated that the assumption of a pure proton background in the analysis is justified in the sense that the likelihood of an air shower induced by a nucleus to be misidentified as a photon candidate event can be neglected in comparison to proton-induced air showers.

6.4.3. Robustness Against Exchanging the Hadronic Interaction Model

For the interpretation of the results that are obtained when the analysis presented in this thesis is applied to data, it is important to take into account uncertainties arising from the current lack of knowledge about hadronic interactions at the highest energies. For the simulation samples with which the analysis has been optimized, QGSJETII-04 has been chosen as the hadronic interaction model at high energies. The results obtained so far, in particular the estimation of the fraction of background—i.e. proton-induced—events that are misidentified as photon candidate events, therefore assume that the hadronic component of the air showers in data develops as predicted by the chosen hadronic interaction model. In the brief study discussed in this section, it is checked whether the fraction of misidentified background events changes significantly if another hadronic interaction model is chosen. It should be stressed at this point that the choice of the hadronic interaction model is relevant only for the proton sample, as hadronic interactions play only an inferior role in the development of photon-induced air showers in the atmosphere. Therefore, all results obtained using the photon sample only, for example the photon candidate cut, are not affected by the uncertainties arising from the choice of the hadronic interaction model.

For this brief study, two simulation samples, both with protons as primary particles but using the two models EPOS-LHC and SIBYLL 2.1 to describe hadronic interactions at high energies, have been created. Both models are fully integrated into the version of CORSIKA that has been used for all the simulations discussed in this thesis. For all samples, FLUKA 2011.2b.6 has been used as the hadronic interaction model at low energies. As in the previous study, three energy bins have been considered for both hadronic interaction models: $18.0 \leq \log_{10}(E\,[\text{eV}]) < 18.1$, $18.5 \leq \log_{10}(E\,[\text{eV}]) < 18.6$, and $18.9 \leq \log_{10}(E\,[\text{eV}]) < 19.0$. The selection efficiencies after the individual selection criteria have been applied are listed in App. B, Tabs. B.3 and B.4. The selection efficiencies in the EPOS-LHC and SIBYLL 2.1 samples do not differ significantly from the selection efficiencies in the QGSJETII-04 sample. However, smaller differences in the order of a few percent between the different models exist. Examples for such differences are found at the preselection level—which reflects the trigger probability—and at the profile level, where small differences can be found in the efficiencies due to the requirement that the shower maximum has to be observed within the field of view of the FD and due to the fiducial volume cut. These differences mainly arise from the different predictions of the individual hadronic interaction models for the longitudinal shower development as well as for the muon content.

First, the analysis is applied to the EPOS-LHC sample. The scatter plots of t_{par}, F_γ and $X_{\max}$ versus E_γ are shown in Fig. 6.38. No significant qualitative differences between the EPOS-LHC sample and the main proton sample, for which QGSJETII-04 has been used, are visible. Overall, 2 events out of 4340 in the EPOS-LHC sample are (mis)identified as photon candidate events. The calculation of the corresponding fraction including the asymmetric uncertainties

(a)

(b)

(c)
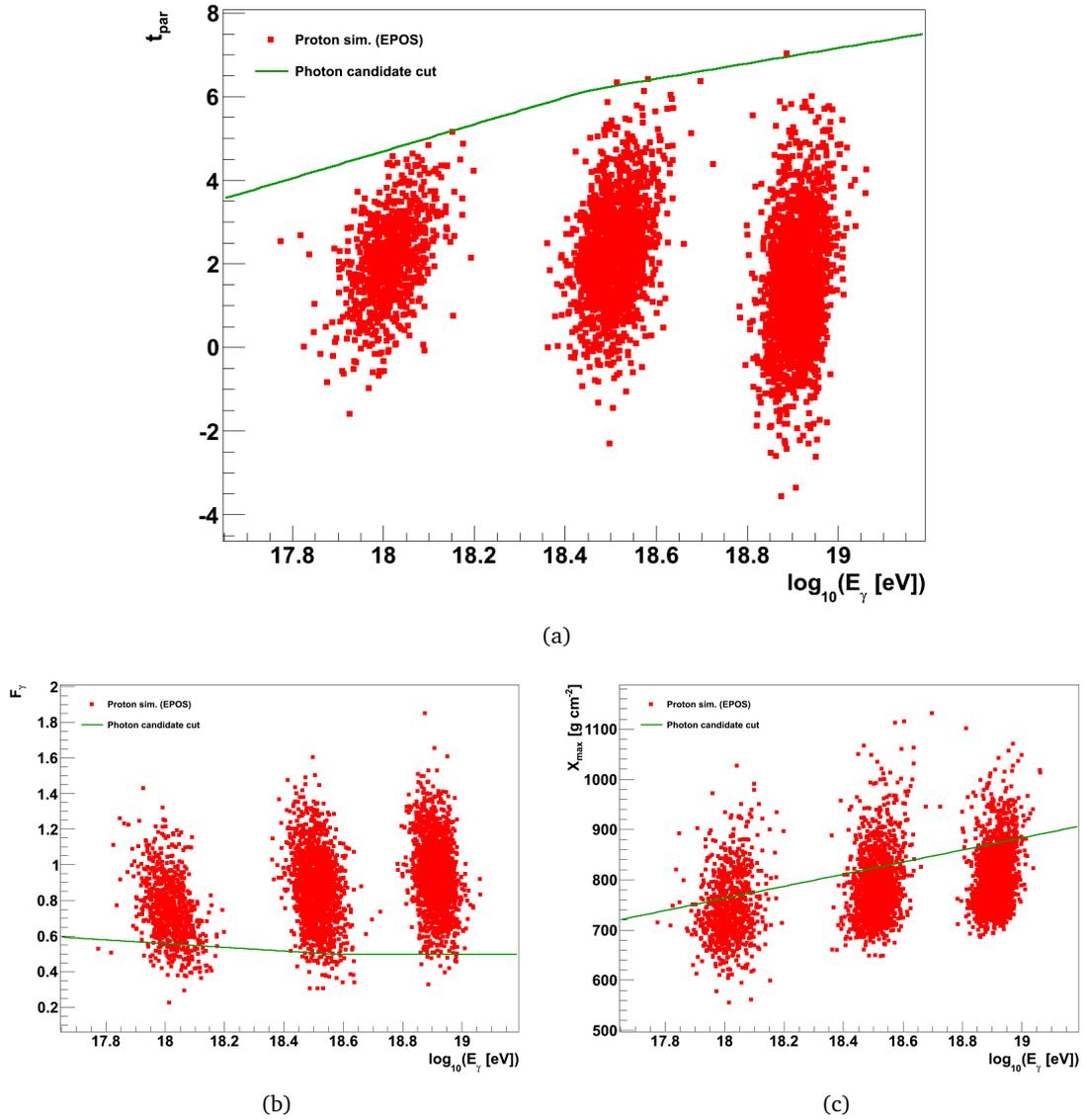
Figure 6.38.: Scatter plots of (a) t_{par} versus E_γ, (b) F_γ versus E_γ and (c) X_{max} versus E_γ, for the EPOS sample. The green lines denote the photon candidate cut in each observable.

according to [Pat04] yields $\left(0.05\,^{+0.06}_{-0.02}\right)$ %. This number agrees with the corresponding fraction for the main proton sample (0.03 %) within one standard deviation. The main reason for these differences can be found in the different predictions for the longitudinal shower development: EPOS-LHC produces, on average, deeper shower profiles, hence proton-induced air showers simulated using this model are more photon-like in this respect compared to showers simulated using QGSJETII-04. However, these differences are small, in the order of $10\,\mathrm{g\,cm^{-2}}$ (see Fig. 3.3), while the predictions for the muon content—which directly affects the signal that is expected in the individual SD stations—are comparable to QGSJETII-04. Hence, the overall difference in the fraction of events identified as photon candidate events between the two models is not large.

(a)

(b)

(c)
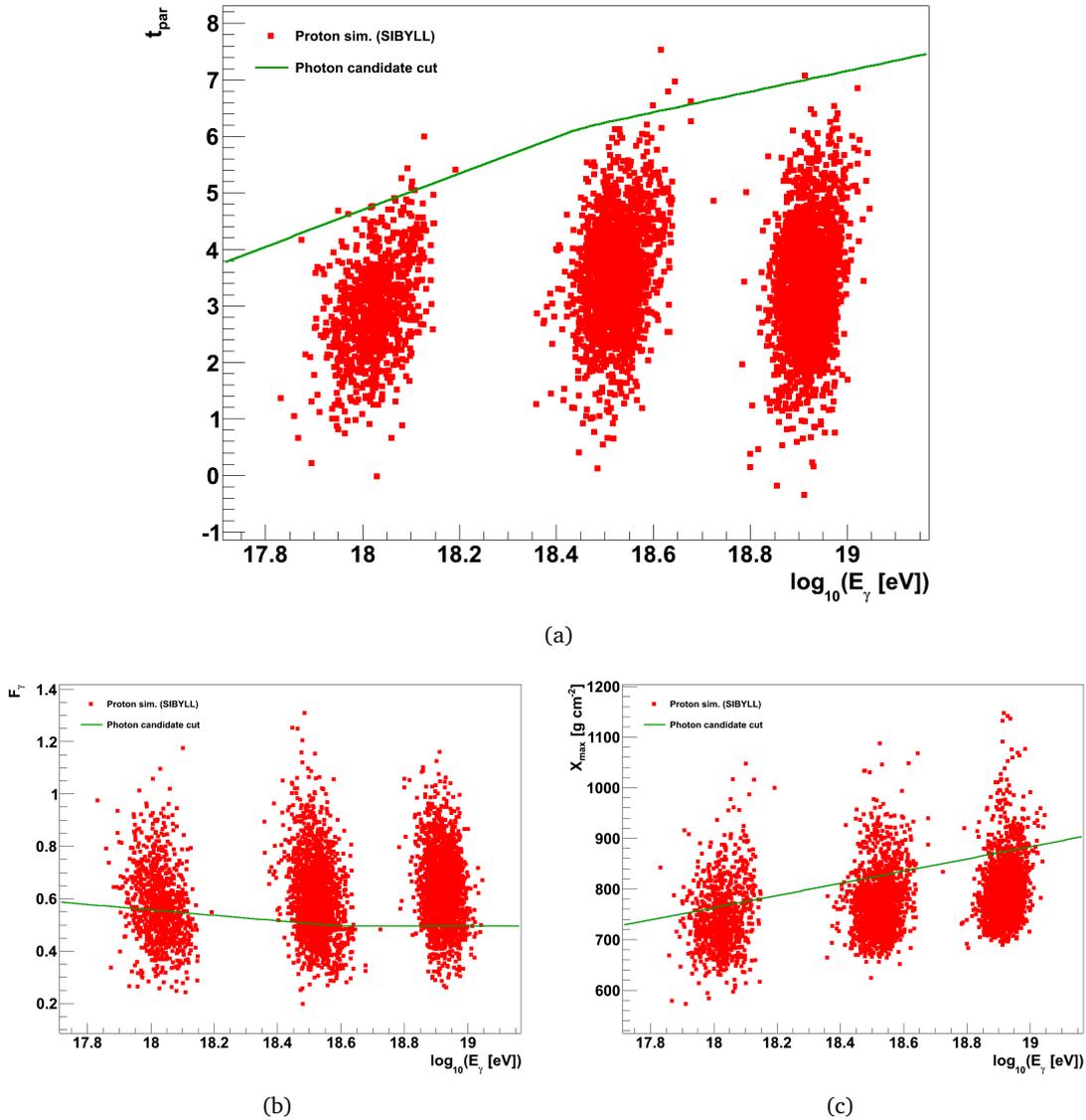
Figure 6.39.: Scatter plots of (a) t_{par} versus E_γ, (b) F_γ versus E_γ and (c) X_{max} versus E_γ, for the SIBYLL sample. The green lines denote the photon candidate cut in each observable.

For the SIBYLL sample, the scatter plots of t_{par}, F_γ and X_{max} versus E_γ are shown in Fig. 6.39. In contrast to the EPOS-LHC sample, larger differences are visible when the scatter plots are compared to the corresponding plots for the main proton sample. While the X_{max} distributions are comparable, the F_γ distributions for the SIBYLL 2.1 sample are, especially in the first energy bin, shifted toward smaller values. This is due to SIBYLL 2.1 predicting a significantly smaller muon content than QGSJETII-04 and EPOS-LHC, which leads to a smaller signal in the SD stations. Overall, the average signal in the SD stations predicted by SIBYLL 2.1 is up to 20 % smaller than what is predicted by QGSJETII-04 [Yus15]. As a consequence, also the t_{par} distributions are shifted toward more photon-like values, and 9 out of 4698 events in the SIBYLL sample pass the photon candidate cut. This represents a fraction of $\left(0.19^{+0.09}_{-0.05}\right)$ %. The difference between this number and the fraction derived from the main proton sample is

more than three standard deviations. The largest fraction of the events identified as photon candidate events are from the first energy bin. In this energy bin, most of the events contain only a single triggered SD station and hence the impact of a smaller signal in this station on the photon-optimized LDF fit is the largest.

A reason for the discrepancies in the results obtained with SIBYLL 2.1 compared to QGSJETII-04 and EPOS-LHC lies in the fact that both QGSJETII-04 and EPOS-LHC have been tuned to current data obtained in measurements at the LHC, while SIBYLL 2.1 predates these measurements. An updated version of SIBYLL, which takes into account the data from current accelerator experiments, including those at the LHC, is currently being worked on [Rie15]. This version is not yet implemented in CORSIKA and can therefore not be used for simulation studies of air showers like the one discussed in this section.

Overall, it can be stated that the impact of chosing a different hadronic interaction model at high energies is small. The differences between the results obtained with the specific models are well understood. Since the impact of exchanging the high-energy model is small, it can be safely assumed that the impact of exchanging the low-energy model is negligible, since the influence of lower-energy hadrons on the overall development of an air shower is much smaller than the influence of the high-energy hadrons and possible differences in the description of these lower-energy particles between the available models are less important.

7

A Search for Photons in the EeV Range Using Hybrid Data

Contents

In the following chapter, the analysis based on the observables discussed in Chap. 6 is applied to real data in order to identify photon-like events in the hybrid data sample collected at the Pierre Auger Observatory. Photon-like events identified with this analysis are denoted as photon candidate events. The application of the analysis to hybrid data is discussed in Sec. 7.1. Out of 6727 events in the data sample, one event passes the photon candidate cut which was defined previously. This event is examined in detail in Sec. 7.2. Special emphasis is given to the FD part of the event—including the atmospheric conditions at the time the event was recorded—since the observable X_{max} is a pure FD observable which can be directly affected by anomalies in the reconstruction. In addition, the arrival direction of the event is investigated, as well as the hypothesis that the event was not induced by a photon, but by a proton or a heavier nucleus. Finally, upper limits based on the results obtained in Sec. 7.1 are determined in Sec. 7.3 on both the fraction of UHE photons in cosmic rays and the differential flux of UHE photons impinging on the Earth. Several systematic effects which may influence the results are discussed.

7.1. Application of the Analysis to Hybrid Data

The data sample to which the analysis described in the previous chapter is now applied has already been discussed in Chap. 5. However, the event selection described in Sec. 5.3.3 does not encompass the additional event selection criteria based on F_γ, which have been defined in Sec. 6.2.6. The selection efficiencies of the additional criteria—when applied to data—are listed in Tab. 7.1. When comparing these selection efficiencies, it should be noted that no energy selection has been applied yet, and therefore the data sample is dominated by lower-energy air shower events with primary energies below 1 EeV. Hence, the selection efficiencies for all selection criteria are smaller than the corresponding efficiencies for the simulation samples, where the primary energy is always above 1 EeV. This is most noticeable when looking at the cut on ΔF_γ. More than 70 % of all events at this stage of the event selection are removed by this cut. In the simulation sample, at most 40 % are removed. The subsequent cuts on the relative uncertainty of F_γ and the χ^2-like quantity each remove about 5 % of the events in the data sample. It is interesting to note that even though the data sample contains mostly lower-energy air shower events—an energy region for which the analysis has not been optimized—the photon-optimized LDF fit fails only for 0.3 % of all events, which again underlines the robustness of the fit procedure.

In the next step, an energy selection is applied to the data sample. As the goal of this analysis is to identify air shower events induced by UHE photons, the energy selection is based on the photon energy E_γ. This is done in order to prevent an overestimation of the true energy in the case of a primary photon, which would occur if the standard missing energy correction was used. It is obvious that this means, at the same time, that the true primary energy is underestimated in the case of a primary proton or nucleus. In view of the goal of the analysis, however, this is not a problem. The energy selection removes all events with $E_\gamma \leq 1$ EeV. This reduces the data sample by about 70 %, which is roughly in agreement with the expectation from an energy spectrum following a broken power law with a spectral index of 3.3. Overall, 6727 events remain for the analysis.

	N	ε
After the event selection from Sec. 5.3	84 794	-
FitFailed false	84 548	99.7 %
DeltaFgamma < 0.35	24 013	28.4 %
RelFgammaUncertainty < 0.7	22 725	94.6 %
ChiSquare < 30	21 644	95.2 %
Total (F_γ level)	21 644	25.5 %
$E_\gamma > 1$ EeV	6727	31.1 %
Total	6727	7.9 %

Table 7.1.: Application of the additional event selection criteria based on F_γ to the data sample described in Sec. 5.2. Listed are the number of surviving events N after applying the criterion and the corresponding selection efficiency ε relative to the number of events before the criterion was applied.
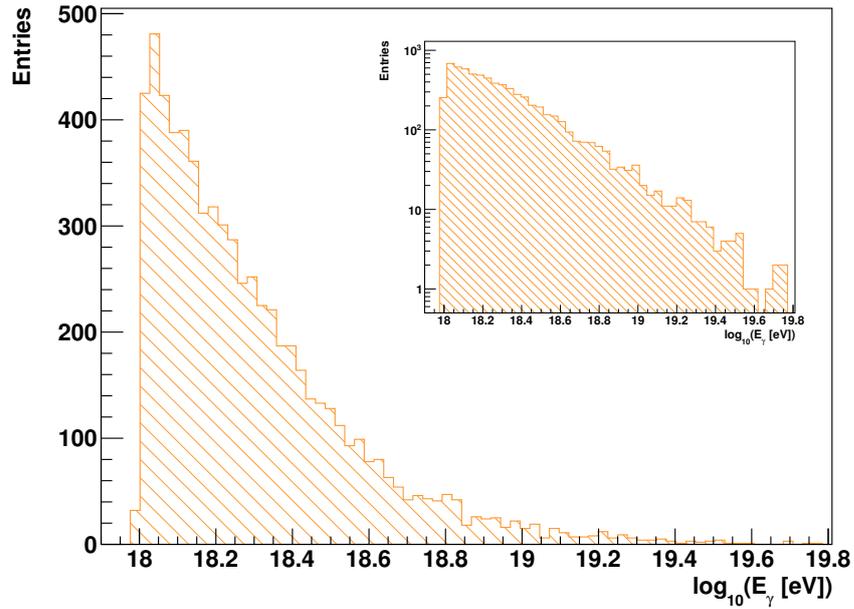
Figure 7.1.: Distribution of the photon energy E_γ for the data sample after the full event selection. The inlay shows the same distribution on a logarithmic scale.

In Fig. 7.1, the distribution of the photon energies E_γ is shown for the data sample after applying the full event selection. The exponential behaviour expected from a broken power-law spectrum is clearly visible. At higher energies, a flattening of the distribution—corresponding to the flattening of the energy spectrum above the ankle—is also visible. The highest-energy event in the sample has an energy around $\log_{10}(E_\gamma\,[\mathrm{eV}]) = 19.8$. It should be noted that for the fully reconstructed hybrid energy the energy distribution would be shifted by about 10 % toward higher values. The shape of the distribution, however, would not be changed.

The distributions related to the geometry of the reconstructed air showers in the data sample are also of interest, in particular the distributions of the zenith and azimuth angles (Fig. 7.2) as well as the distribution of the core positions at ground level (Fig. 7.3). The zenith angle distribution (Fig. 7.2(a)) is mainly affected by the fiducial volume cut, which removes events with small zenith angles below 25°. At these zenith angles, the probability that a photon-induced air shower event is recorded with the shower maximum within the field of view of an FD telescope is smaller than for proton-induced showers, and hence all events with small zenith angles are removed in order to keep the selection efficiencies for photon-induced and proton-induced air showers similar. As the fiducial volume cut is applied in an energy-dependent way (cf. Eq. 5.4), there is no sharp cut-off at small zenith angles. At large zenith angles, the distribution is cut off at 60° due to the corresponding selection criterion. Between the minimum and maximum zenith angles predetermined by the two cuts, the distribution is smooth and roughly flat in $\cos^2(\theta)$. For the distribution of the azimuth angle, no significant deviations from a flat distribution over all angles are expected, as no direct cut on the azimuth angle is applied. An indirect influence on the distribution of the azimuth angle can come from the cut on the contribution of Cherenkov light to the total light collected by an FD telescope, which can remove events that point in the direction of the FD telescopes. However, since the FD telescopes are not all oriented in the same direction, no changes in the over-

(a)　　　　　　　　　　　　　　　　　(b)
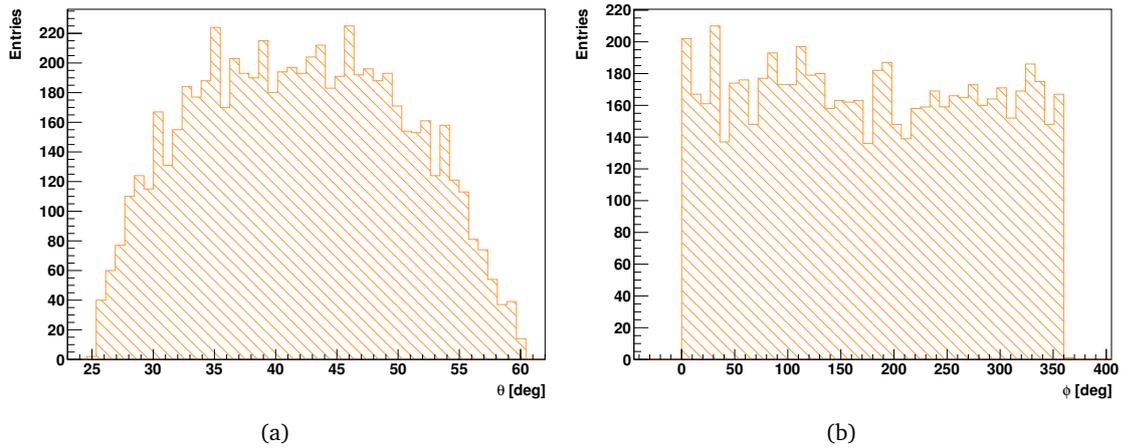
Figure 7.2.: Distributions of (a) the zenith angle θ and (b) the azimuth angle ϕ of the recon-
structed air shower events in the data sample.
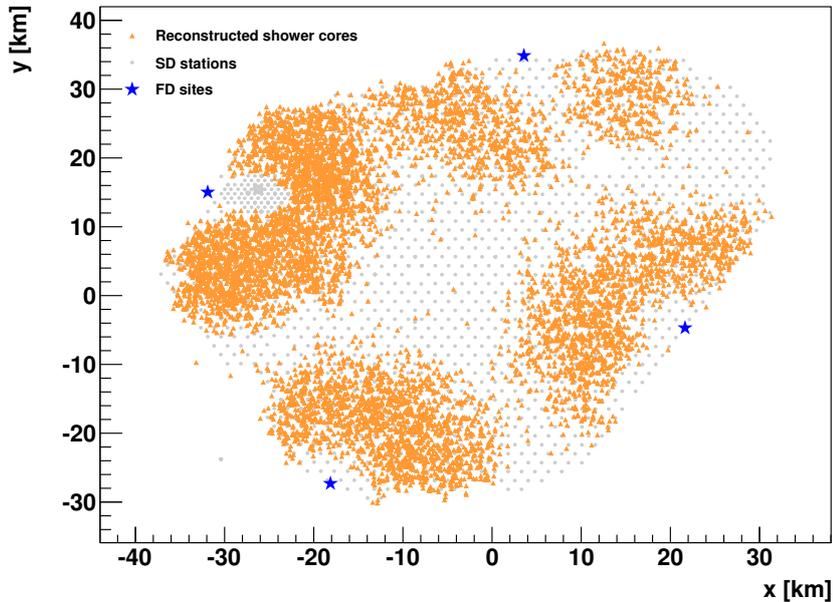
Figure 7.3.: Spatial distribution of the reconstructed position of the shower core at ground
level of the events in the data sample. The positions of the shower cores (orange triangles)
are shown over a map of the positions of the SD stations (gray circles) and the FD sites
(blue stars).

all distribution of the azimuth angle are expected. This is indeed the case for the data sample
(cf. Fig. 7.2(b)), which is, within the observed fluctuations, compatible with a flat distribution.

The reconstructed positions of the shower core on ground level for the events in the data
sample are shown in Fig. 7.3. The positions of the detector stations of the SD array and the
four FD sites at the border of the SD array are also indicated. Several features are visible.
Almost all events are clustered around the FD sites. This is due to the relatively low energies

of the events in the data sample. For an air shower near the center of the SD array to be detected by the FD telescopes, the energy must be at least around 10^{19} eV [Pet04]. The FD sites Coihueco, on the western border of the SD array, and Los Leones, at the southern border close to Malargüe, were the first FD sites to begin taking data, hence there are more events clustered around these two sites than around the eastern site Los Morados and the northern site Loma Amarilla, which started data taking in 2005 and 2007, respectively. In addition to the total data taking time of the FD sites, also the uptime fraction is relevant for the number of events clustering around the individual sites. For Coihueco, also the presence of the infill array has to be taken into account. Events where SD stations from the infill array are triggered are not included in the data sample that is analyzed here, hence there is an uncovered region in the spatial distribution of the core positions around this area. Another uncovered region is visible near Loma Amarilla. The lack of events in this region is due to a misalignment in the mirror of a FD telescope at Loma Amarilla. This misalignment has been corrected in September 2012 [Bel13]. Events recorded by this telescope before that date are excluded from the data set.

Overall, the distributions discussed so far are all in agreement with the expectation. The analysis discussed in Chap. 6 can be applied to the data sample as intended, as there is no indication toward a bias in the data sample that could influence the results of the analysis. The distributions of the two main observables F_γ and X_{max} for the data sample are shown in Fig. 7.4(a) and (b), respectively. The shapes of the distributions for the data sample are comparable to the shapes of the corresponding distributions for the simulation samples. The mean value of the F_γ distribution is at 0.9, which is considerably larger than the mean values of the distributions from the proton sample (cf. Tab. 6.5). This is expected since the data sample contains not only proton-induced air shower events, but also air shower events initiated by heavier nuclei, and such air showers exhibit larger F_γ values (see Sec. 6.4.2). In addition, the data sample covers a wider range of primary energies. As the average F_γ increases with energy for air shower events initiated by protons and nuclei, the mean of the total distribution is also shifted toward higher F_γ values. The width of the distribution for the data sample is larger than that for the pure proton sample: for the data sample, the width is 0.25, while for the proton sample, the width is around 0.17 (cf. Tab. 6.5). For the X_{max} distribution, a similar behaviour could be expected. However, the average X_{max} increases with energy, but decreases with the primary mass. Both effects partly cancel each other out. Hence, the mean and the width of the X_{max} distribution for the data sample are around $750\,\mathrm{g\,cm^{-2}}$ and $65\,\mathrm{g\,cm^{-2}}$, respectively, which is be compatible with the parameters of the X_{max} distributions for the proton sample (see Tab. 6.1). Furthermore, it should also be noted that the X_{max} distributions are affected by the requirement that the shower maximum has to be observed within the field of view and the fiducial volume cut. This removes events with small X_{max}. In addition to the individual distributions of F_γ and X_{max}, a scatter plot of the two observables is shown in Fig. 7.4(c). No correlation between the two observables is visible. The Pearson correlation coefficient r, calculated according to Eq. 6.31, is -0.06, which underlines that there is no correlation between F_γ and X_{max}, as expected from the simulation sample (cf. Fig. 6.23). Of particular interest for this analysis is the upper left region of the scatter plot, where photon-like events are expected which exhibit both large X_{max} and small F_γ values.

The combined observable t_{par} has been calculated for each event according to Eqs. 6.39 and 6.40. The distribution of t_{par} is shown in Fig. 7.4(d). As t_{par} is a linear combination of

(a)

(b)

(c)

(d)
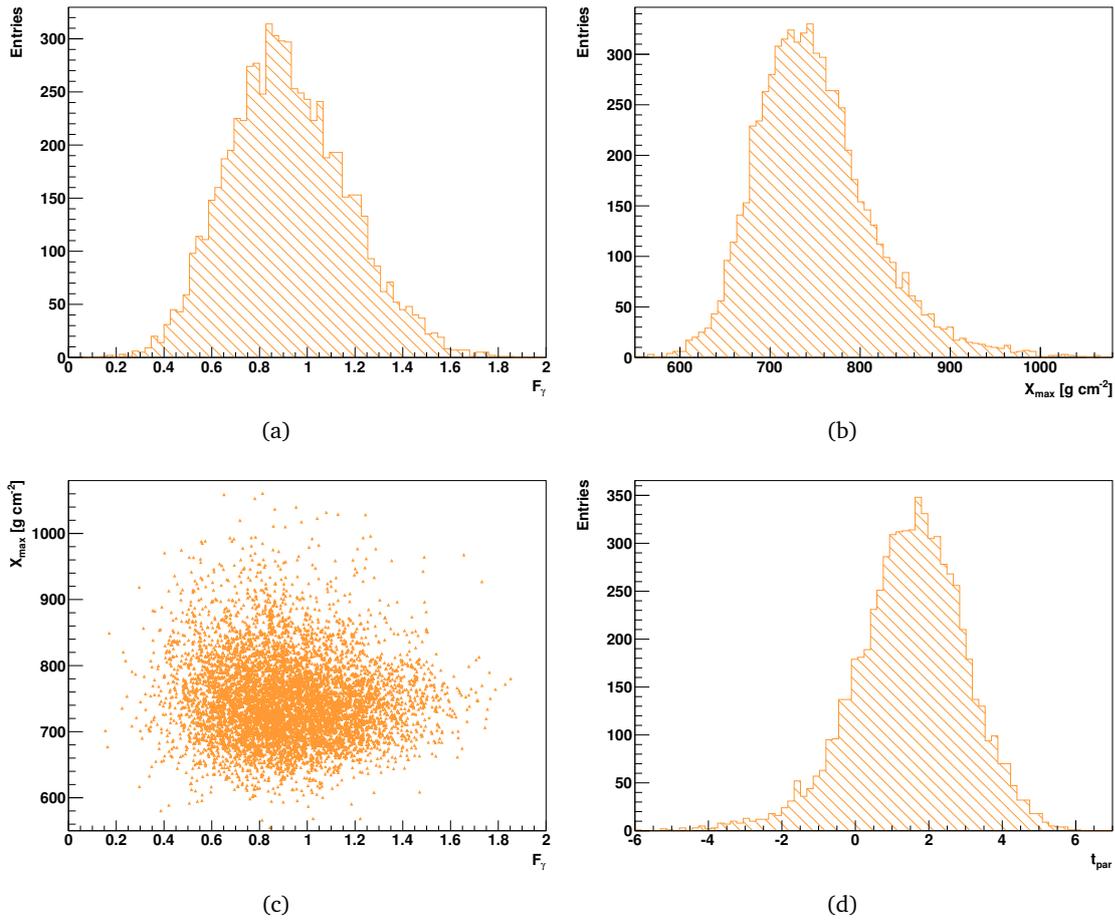
Figure 7.4.: Distributions of the main observables for the data sample: (a) F_γ, (b) X_{max}. In (c), both observables are shown in a scatter plot. The distribution of the combined observable t_{par} is shown in (d).

the basic observables F_γ and X_{max}, features of the corresponding distributions are propagated to the t_{par} distribution. Hence, the mean value of the t_{par} distribution is smaller—i.e. less photon-like—than the mean values of the corresponding distributions for the proton sample (cf. Tab. 6.8). The shape of the distribution is similar to the shape of the distributions for the proton sample.

In the next step, the photon candidate cut is applied to the data sample. The scatter plots of the observables t_{par}, F_γ and X_{max} versus the photon energy E_γ are shown in Fig. 7.5. In the scatter plots, also the photon candidate cut in the individual observables following Eqs. 6.43 and 6.45 is included. In all observables, the data points move, on average, further away from the line indicating the photon candidate cut with increasing energy. This behaviour is also visible, albeit less pronounced, in the proton sample (Figs. 6.33 and 6.35) and, more pronounced, in the helium, oxygen and iron samples (Fig. 6.37). The spread of the data points gets smaller with energy, which is most noticeable in the scatter plot of X_{max} versus E_γ (see Fig. 7.5(c)). This can be interpreted as an effect of the change in the primary mass composition toward higher energies, as the shower-to-shower fluctutions get smaller with increasing mass.
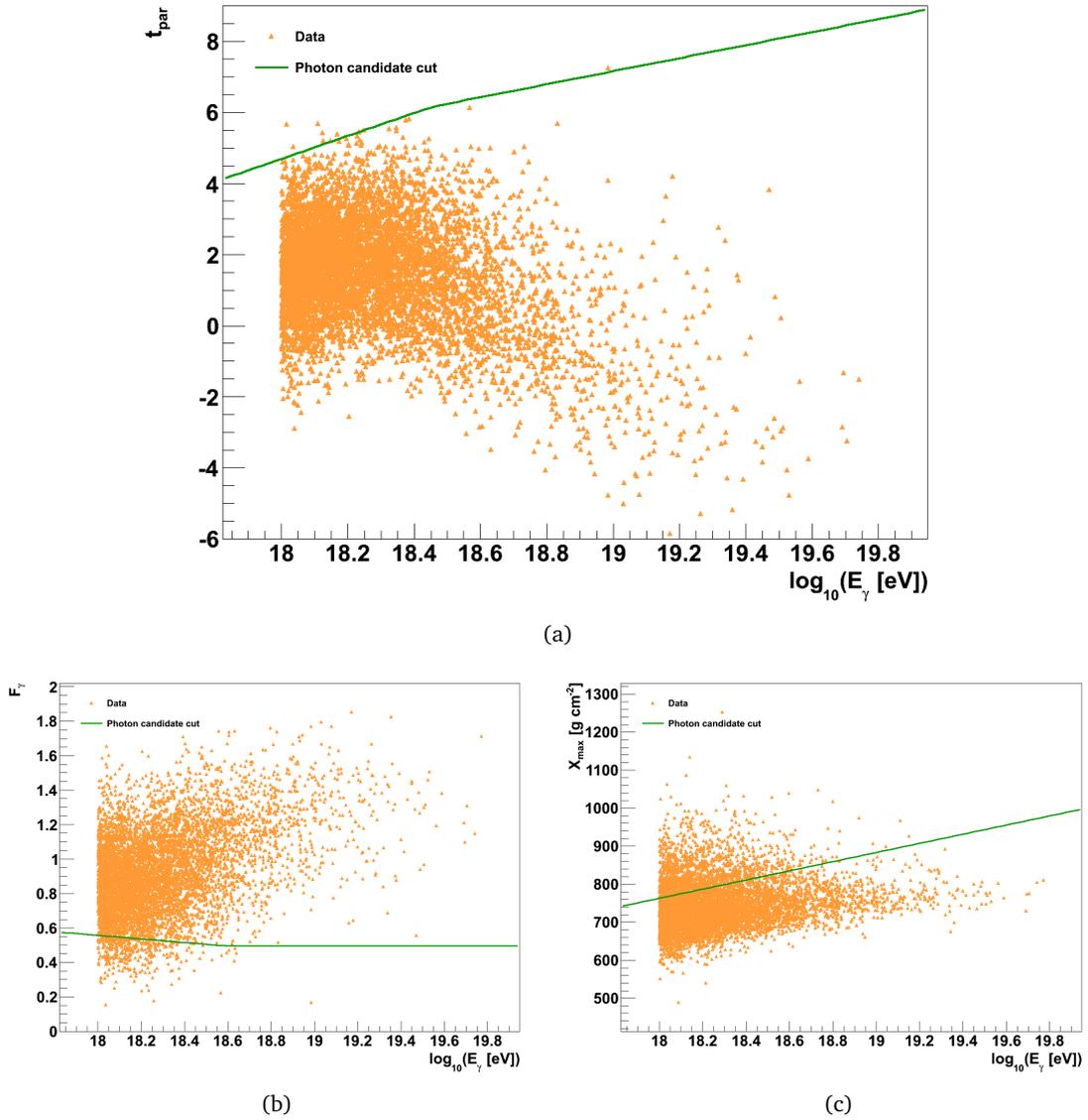
Figure 7.5.: Scatter plots of (a) t_{par} versus E_γ, (b) F_γ versus E_γ and (c) X_{max} versus E_γ, for the data sample. The green lines denote the photon candidate cut in each observable.

Overall, of the 6727 events in the data sample, 1 event passes the photon candidate cut, which corresponds to a fraction of 0.015 %. When only t_{par} is taken into account, 8 events would pass the photon candidate cut. The majority of the events that would be identified as photon candidate events when only t_{par} is taken into account—including events with very large X_{max} values around $1000\,\mathrm{g\,cm^{-2}}$—cannot be reliably considered as photon candidates, since they differ rather much from the photon expectation in one of the basic observables. The one event that passes the photon candidate cut in all three observables was reconstructed with a photon energy of $1.03\,\mathrm{EeV}$. The value of the combined observable t_{par} for this event is 5.66, while the photon candidate cut in this observable at this energy selected all events with $t_{\mathrm{par}} > 4.74$. Also in the two basic observables, this event is sufficiently photon-like: the reconstructed values are $F_\gamma = 0.28 \pm 0.14$ and $X_{\mathrm{max}} = (918 \pm 34)\,\mathrm{g\,cm^{-2}}$, with the photon candidate

cut at this energy selecting events with $F_\gamma + \Delta F_\gamma < 0.56$ and $X_{max} - \Delta X_{max} > 764.64\,\mathrm{g\,cm^{-2}}$.

For the interpretation of this result, it is important to estimate the number of expected background events—i.e. proton-induced air showers and those initiated by nuclei—and compare this number with the observed number of photon candidate events. The expected background is estimated from simulations, which have already been discussed in Chap. 6. Even though these simulations cover only the energy range from 10^{18} eV to 10^{19} eV—with the data sample extending up to $10^{19.8}$ eV—it has been shown that the performance of the analysis at energies beyond 10^{19} eV is comparable to the performance in the EeV range (cf. Sec. 6.4.1). Moreover, only about 2 % of the events in the data sample have energies above 10^{19} eV. Hence, the expected background can be estimated from the simulation samples. Assuming a pure proton background and using QGSJETII-04 as hadronic interaction model, it is expected that 0.026 % of all proton events pass the photon candidate cut. The asymmetric uncertainty of this number, calculated according to [Pat04], is +0.013 % and −0.007 %. For a sample size of 6727 events, this would correspond to $1.8^{+0.9}_{-0.4}$ proton events being misidentified as photon candidate events, which is compatible with the observed number within two standard deviations. A simular result is found when using EPOS-LHC. For this model, 3^{+4}_{-1} events would be expected (see Sec. 6.4.3), which is still compatible with the observed number. For SIBYLL 2.1, on the other hand, a larger deviation is found: 13^{+6}_{-3} photon candidate events would be expected for this model. The observed number of events is more than four standard deviations smaller. However, it has already been discussed in Sec. 6.4.3 that there are known discrepancies between SIBYLL 2.1 and the most recent LHC-tuned models such as QGSJETII-04 and EPOS-LHC. When an LHC-tuned version of SIBYLL becomes available, this study should be repeated to derive a more precise estimate for this model. Independent of the hadronic interaction model the estimates of the expected number of photon candidate events are systematically larger than the observed value, due to the assumption of a pure proton background. Assuming that only 50 % of the background consists of protons and the remainder consists of helium nuclei reduces the expected number of photon candidate events to $0.9^{+3.1}_{-0.4}$. This number has been estimated taking into account the results from Sec. 6.4.2, which are based on QGSJETII-04. For EPOS-LHC and SIBYLL 2.1, similar changes can be expected.

7.2. Study of the Photon Candidate Event

Although the number of photon candidate events is fully compatible with the background expectation and thus the next step would be to derive upper limits on the fraction of photons in UHECRs, it is worthwhile to study the one event that has been identified as a photon candidate event in detail. One aim of this study is to find hints toward the type of the primary particle that initiated the recorded air shower beyond the indications provided by the high-level observables discussed before. At the same time, the quality of the event can be checked in detail at all stages of the event reconstruction.

The candidate event has been recorded on the 15th of August 2007, at 01:04:05 UTC. In addition to telescope 1 of the FD site Coihueco, three SD stations have been triggered. From the data recorded by these stations, a full SD reconstruction of the air shower event was possible, hence the event is a golden hybrid event. The internal identifiers of the individual FD and SD events are 4/2268/1852 and 3837310, respectively, while the merged event which

contains both parts has the identifier auger_200722602073.

The air shower was recorded close to the western border of the SD array, about 10 km south of Coihueco. In the site coordinate system, the position of the shower core—taken from the hybrid geometry reconstruction—is $(x, y) = (-34.11 \pm 0.04, 4.5 \pm 0.04)$ km. The reconstructed direction of the shower axis, represented by the zenith angle θ and the azimuth angle ϕ, is given by $(\theta, \phi) = (35.9 \pm 0.7, 180.5 \pm 0.6)°$. In Fig. 7.6, the recorded event is visualized in a three-dimensional drawing.

The reconstructed primary energy is comparatively low, close to the lower bound of the energy range considered in this analysis. The calorimetric energy measured by the FD is (1.02 ± 0.14) EeV. The fully reconstructed hybrid energy, including the standard correction for the missing energy, is (1.20 ± 0.16) EeV. The photon energy, i.e. the reconstructed energy including a missing energy correction suitable for primary photons, is (1.03 ± 0.14) EeV. The primary energy which was reconstructed from the SD data is (0.96 ± 0.18) EeV. The small discrepancy between the fully reconstructed hybrid energy and the SD energy can be an indication that the air shower was indeed induced by a photon, for which such a discrepancy would be expected. However, due to the large uncertainties, the reconstructed hybrid and SD energies are still compatible.
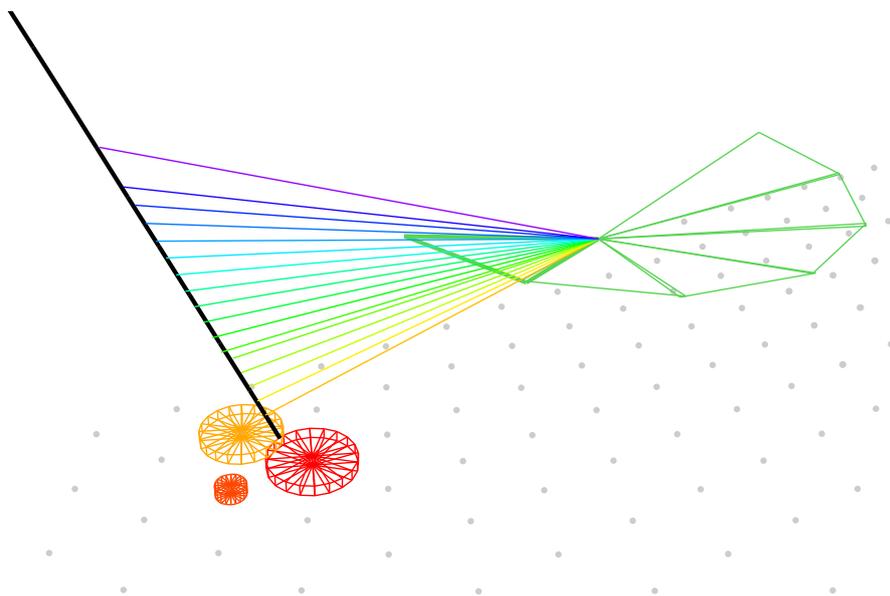
Figure 7.6.: Schematic, three-dimensional drawing of the photon candidate event. The event has the internal identifier auger_200722602073. SD stations that were not triggered are represented by gray dots, while triggered stations are represented by colored cylinders. The size and the color of the cylinders indicate the signal size and the signal timing, respectively. The orientations of fields of view of the six FD telescopes at Coihueco are sketched as well. The black line represents the reconstructed shower axis. The colored lines between the shower axis and the FD telescope indicate the viewing directions of the triggered pixels in the telescope, with the color representing the signal time.

In the following sections, the FD and SD parts of the event will be examined in detail. In addition, dedicated studies of the arrival direction of the air shower and the possible primary particle type are presented.

7.2.1. FD Part of the Event

Before the FD measurements are discussed, the atmospheric conditions at the time of recording of the photon candidate event are reviewed. A detailed knowledge of the atmospheric conditions is important for the correct interpretation of the FD measurements, for example to decide whether a measured longitudinal profile is distorted due to a cloud passing through the field of view of the FD telescope observing the event, or due to a large number of aerosols in the atmosphere, or whether the general viewing conditions were too poor to enable reliable FD measurements.

The atmospheric state variables—temperature and pressure—as measured by the weather stations at the detector site, were 271.5 K and 860 hPa, respectively. These values are close to the averages for the month of August at the detector site [Abr10c]. Therefore, it can be assumed that the model used to determine the atmospheric profile from the weather station measurements (cf. Sec. 4.2.3) provides a reliable estimate of the true atmospheric profile.

During the night of data taking in which the photon candidate event was recorded, the Lidar system at the FD site Coihueco was operating normally and performing the standard sequence of measurements. The evaluated Lidar measurements at the time of the recording of the event are shown in Fig. 7.7. The optical depth of aerosols was measured about 15 min

(a)   (b)
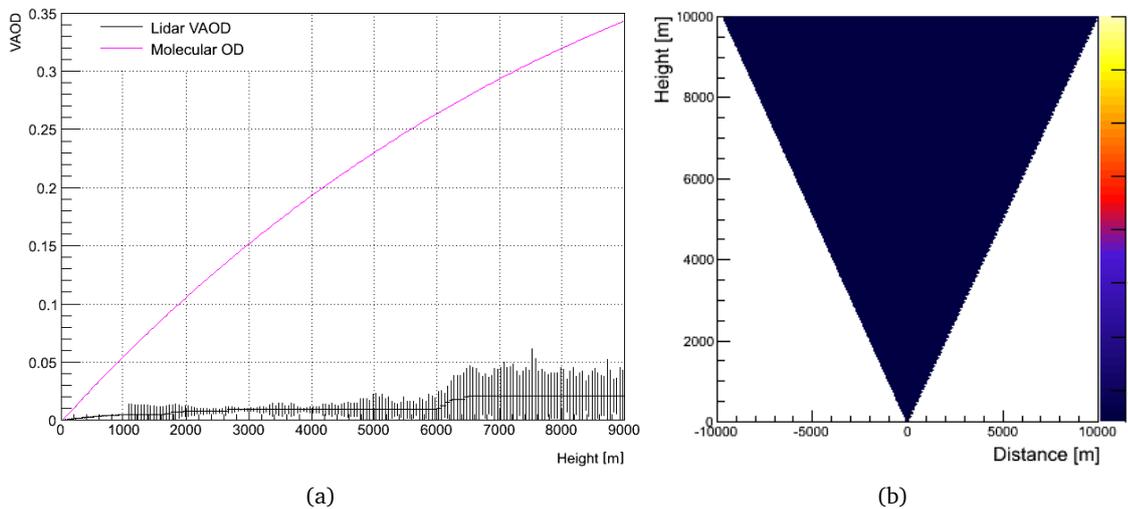
Figure 7.7.: Evaluated measurements from the Lidar at the FD site Coihueco, taken from the measurement runs 16998 and 16999 [Ton15]. (a) Reconstructed optical depth of aerosols as a function of the height above the ground. In addition to the measured optical depth of aerosols, the estimated optical depth of molecules in the atmosphere is shown as a pink line. (b) Measurement of the cloud coverage above the FD site. The color scale indicates the amount of backscattered light, which is an indicator for clouds in the field of view of the Lidar.

prior to the photon candidate event. The measured VAOD as a function of the height above ground is shown in Fig. 7.7(a). Compared to the optical depth of molecules in the atmosphere, the optical depth of aerosols can be neglected. At the reference height of 3 km, the VAOD is significantly smaller than 0.1, which was the cut value used in the event selection. The same conditions are found when measurements from the other Lidar sites are considered. In Fig. 7.7(b), another Lidar scan of the sky above the FD site is shown. This scan was started about 2 min prior to the photon candidate event, and it required about 10 min to complete. Clouds above the detector site would appear in this plot as a region of strong backscattered light. Up to an altitude of 10 km, no clouds are visible. The lowermost, very thin layer of clouds is found at a height of about 14 km. At this height, clouds have no influence on the FD measurements. For the photon candidate event, in particular, the first signals from the air shower were recorded from a height of about 5 km. The shower maximum is located at a height of about 1 km. At these heights, the Lidar measurements do not indicate any clouds.

As an alternative to the Lidar measurement of the cloud coverage, the continuous scans of the sky provided by the IR cloud cameras are typically used. However, at the time the photon candidate event was recorded, the cloud camera at Coihueco could not be used due to technical problems [Gru14]. Another cross check of the cloud coverage is provided by satellite data. In Fig. 7.8, a map of the cloud coverage over the full detector site, estimated using IR data from the Geostationary Operational Environmental Satellite (GOES) system [Chi13], is
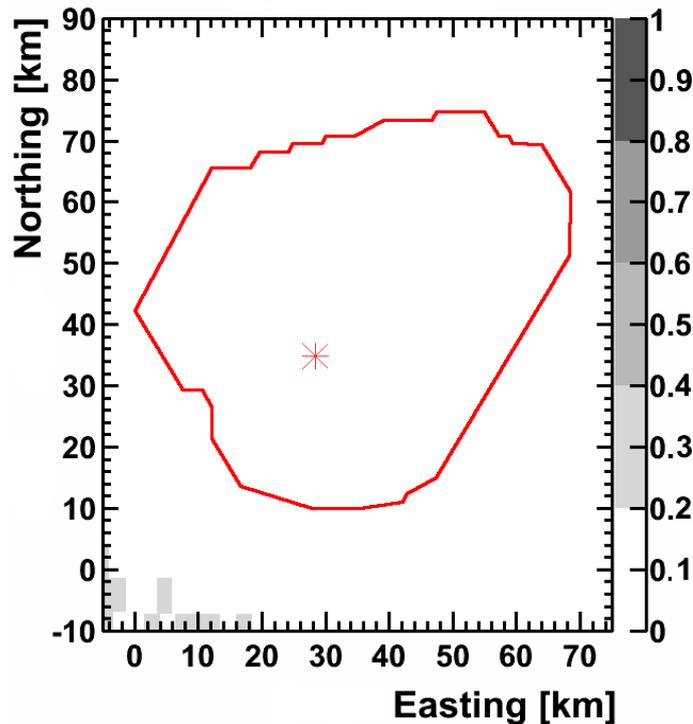
Figure 7.8.: Map of the cloud coverage over the detector site of the Pierre Auger Observatory, based on IR data from the GOES system [Chi15]. The SD array is outlined in the plot. As an additional reference point, the location of the CLF is indicated as well. The GOES-12 satellite passed over the detector site about 4 min after the photon candidate event was recorded.

shown. The GOES data give no indications for clouds above the detector site at the time of the photon candidate event.

Overall, the measurements of the atmospheric properties suggest that there were very good conditions for FD measurements at the time the photon candidate event was recorded. It can therefore be safely assumed that the reconstruction of the air shower event from the FD measurements was not subject to any significant negative atmospheric effects.

The different stages of the reconstruction of the air shower event from the FD measurements are shown in Fig. 7.9. The distribution of triggered pixels in the FD telescope is shown in Fig. 7.9(a). The triggered pixels form a near-straight track in the camera. The χ^2/NDF of the fit of the SDP to this track (cf. Eq. 4.1) is 15.35/15, which indicates a reliable estimation of the SDP. From the time distibution shown in Fig. 7.9(b), the geometry of the event within the SDP is determined. In addition to the timing information of the individual pixels, the timing information of the SD station 627 (Cacho) is used. The offset between the expected and the observed signal timing at the SD station is about 5 ns. The χ^2/NDF of the fit of a time function (cf. Eq. 4.4) to determine the shower geometry is, 33.40/14 , which is still acceptable given the good agreement with the overall distribution of the data points.

In the next step of the event reconstruction, the number of photons at the aperture of the telescope is determined from the signals in the individual pixels. The number of photons as a function of time is shown in Fig. 7.9(c). The contribution of Cherenkov light is indicated as well. Since the air shower is viewed mostly from the side, the contribution of direct Cherenkov light—which is emitted preferentially in forward-direction—to the total light profile is negligible. The contribution of scattered Cherenkov light is also small compared to the contribution of fluorescence light. The fraction of Cherenkov light in the total light is estimated to be around 7 %. From the reconstructed fluorescence light, the longitudinal shower profile in terms of the energy deposited in the atmosphere as a function of the slant depth is determined (Fig. 7.9(d)). The χ^2/NDF of the fit of a Gaisser-Hillas function (Eq. 2.23) to the profile is around 100.22/106. Although the observed longitudinal profile is rather flat, overall the profile is well-reconstructed. For Fig. 7.9(e), the data points from Fig. 7.9(d) have been binned into 20 bins of equal width to visually enhance the shape of the measured profile. The shower maximum $X_{\max}$, as determined from the fit in Fig. 7.9(d), is located at a slant depth of $(918 \pm 34)\,\mathrm{g\,cm^{-2}}$, which is well inside the field of view of the FD telescope. The reconstructed value of $X_{\max}$ is larger than the average $X_{\max}$ expected from MC simulations for photon-induced air showers at these energies (cf. Fig. 6.1). Compared to the expected distribution of proton-induced showers around 1 EeV, it is found that about 2 % of the simulated events exhibit $X_{\max}$ values larger than the observed one. Hence, although the measured depth of the shower maximum indicates that this event could very well be induced by a photon, the hypothesis of a primary proton cannot be excluded (see also the dedicated study discussed in Sec. 7.2.4).

Summarizing the discussion of the FD part of the photon candidate event, it can be stated that the event was, from the point of view of the FD, well-measured and well-reconstructed. The reconstructed quantities and observables, in particular $X_{\max}$, can be assumed to be reliable estimates of the true quantities.

(a)

(b)

(c)

(d)

(e)
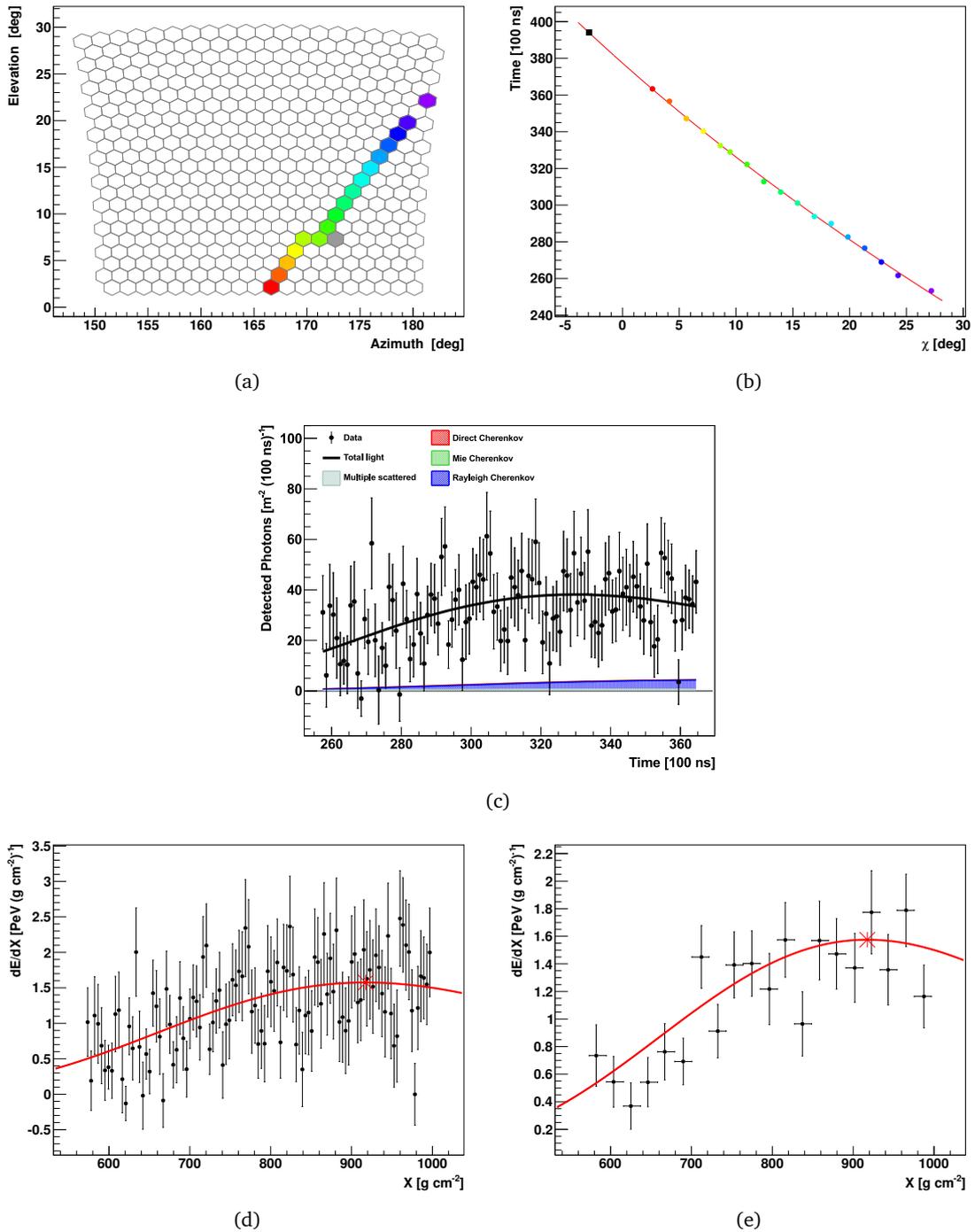
Figure 7.9.: Reconstruction of the photon candidate event from the FD part of the recorded data. (a) Sequence of triggered pixels, color-coded according to the arrival time of the signals at the individual pixels. (b) The arrival times of the signals recorded by the single pixels and the additional SD station as a function of the viewing angle χ. (c) The detected light at the aperture of the FD telescope as a function of time, including the total light profile and the disentangled components of the light profile. (d) The longitudinal shower profile in terms of energy loss per unit atmospheric depth, dE/dX, as a function of the atmospheric depth X. In (e), the longitudinal shower profile is shown again with the data points binned in 20 bins of equal width to visually enhance the shape of the profile. For more detailed descriptions of the individual plots, see the corresponding plots in Sec. 4.5.1.

7.2.2. SD Part of the Event

The photon candidate is a golden hybrid event, therefore a full reconstruction of the air shower event is also available from SD data. This reconstruction is briefly discussed in the following paragraphs. The footprint of the air shower on the SD array in terms of triggered stations is shown in Fig. 7.10(a). Six SD stations have been triggered in total. However, three of these triggered stations are considered accidental, and hence are not shown in Fig. 7.10(a). Even though the energy of the air shower event is rather low, it is possible for air showers at these energies to trigger enough stations for a full reconstruction of the event from SD data. From studies of the SD trigger efficiency, it is known that about 60% of all air showers with primary energies around 1 EeV produce an SD trigger on array level [Abr10d]. Even though the trigger efficiency for air shower induced by primary photons is smaller—mostly due to their lack of a significant muon component—it is still not unlikely for a photon-induced air shower with a primary energy around 1 EeV to trigger more than three SD stations [Set08].

From the spatial and temporal distribution of the triggered stations, the position of the shower core has been determined to be $(x, y) = (-34.12 \pm 0.04, 4.29 \pm 0.14)$ km. The reconstructed direction of the shower axis is given by $(\theta, \phi) = (34.0 \pm 0.5, 178 \pm 2)°$. The geometries of the air shower reconstructed from hybrid data and SD data, respectively, are compatible within their uncertainties.

In the next step of the event reconstruction, an LDF following Eq. 4.9 is fitted to the lateral distribution of the signals (Fig. 7.10(b)). Qualitatively, the fit is acceptable. However, the quality of the fit cannot be reliably quantified with χ^2/NDF, because the fit is based on a maximum-likelihood method [Veb13b]. The S_{1000} determined by the LDF fit is (5.16 ± 1.18) VEM. From the reconstructed S_{1000}, the primary energy was reconstructed using the energy calibration of the SD to be (0.96 ± 0.18) EeV. It should be noted at this point, that the energy calibration itself is based on higher-energy data, above 3 EeV and thus has to be extrapolated for the application to lower energies such as the one observed here, which may introduce additional uncertainties.

So far, only higher-level data, such as the total signal measured in the SD stations, has been discussed. Lower-level data, for example the FADC traces recorded by the SD stations, also provide useful information. The time structure of the traces contains additional information which may give indications toward the probable primary particle type. Air showers initiated by photons lead to an FADC trace that is more spread out in time than the traces typically expected from air showers initiated by protons and nuclei. This is mainly due to the different muon content and the differences in the longitudinal shower development [Abr08b]. The information contained in the FADC traces can be accessed in different ways. In [Abr08b], the risetime $t_{1/2}$ of the signals within the individual SD stations, interpolated at a perpendicular distance of 1000 m, is used. For this observable, however, at least four SD stations are needed. Another way to exploit the FADC traces is the shape parameter ξ, which has been discussed already in Sec. 6.2.5. In Fig. 7.10(c), the FADC traces recorded by the SD station Cacho, which recorded the largest signal in this event, are shown. The average risetime of the signals is about 120 ns, and the shape parameter, determined according to Eq. 6.26, is 1.73. However, comparing this value to the expected distributions of ζ for air showers around 1 EeV (cf. Fig. 6.15(b)), yields no indication toward any primary particle type, as the separation power of the observable ζ around 1 EeV is rather small.

(a)

(b)

(c)
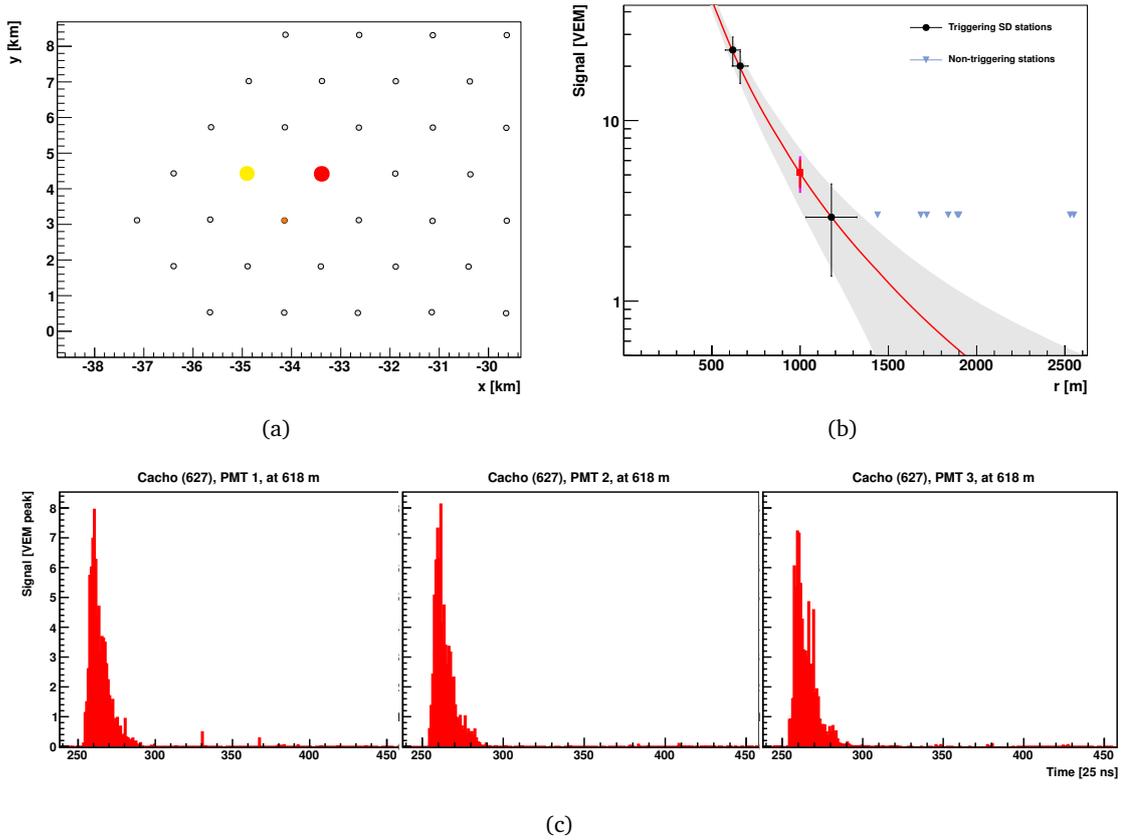
Figure 7.10.: SD part of the photon candidate event: (a) Footprint of the extensive air shower measured by the SD array, with the color and the size of the triggered SD stations indicating the timing and the size of the signals, respectively. (b) The lateral distribution of the signals recorded in the SD stations as a function of the perpendicular distance of the station to the shower axis, r. The red line indicates the fit of an LDF to the data points. (c) FADC traces recorded by the three PMTs in the SD station 627 (Cacho). For more detailed descriptions of (a) and (b), see the corresponding plots in Sec. 4.5.2.

To conclude the discussion of the SD part of the photon candidate event also the observable S_b is evaluated. From the lateral distribution of the signals, S_b is determined according to Eq. 6.25 with $b = 4$. The calculated value of 18.3 VEM is larger than the average S_b expected for photon-induced showers at these energies, which is around 5 VEM (see Fig. 6.15(a)). Only about 1 % of the simulated events exhibit S_b values larger than the observed one. The observed value is more compatible with the expectation for proton-induced showers, where the average S_b is around 15 VEM.

Photon-optimized LDF Fit for F_γ

In this section, the determination of the observable F_γ for the photon candidate event is briefly reviewed. Of the three triggered SD stations, only one could be used for the photon-optimized LDF fit. The other two stations are designated test stations [Pie14b]. Even though these stations were not removed from the automatic data acquisition, it was decided, prior
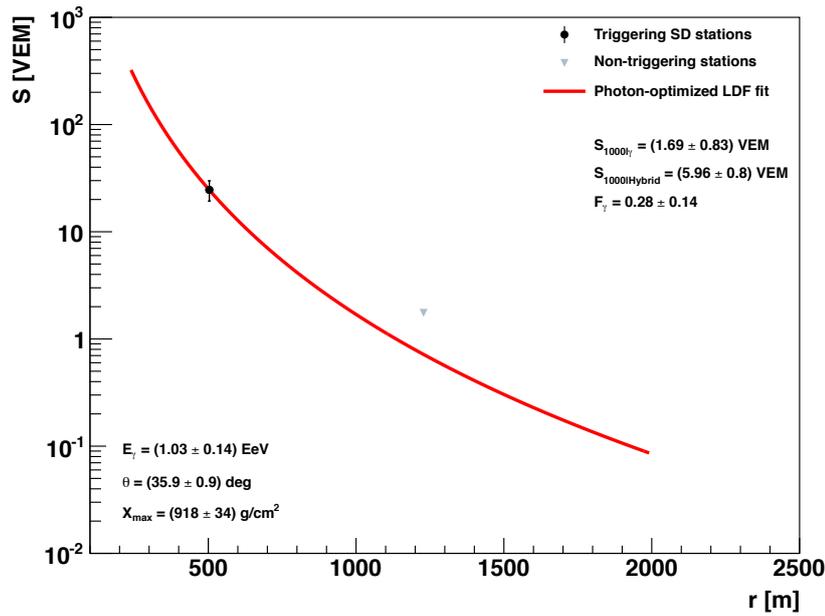
Figure 7.11.: Photon-optimized LDF fit to the data recorded by the SD for the photon candidate event.

to the application of the analysis to real data, to exclude data from test stations from the photon-optimized LDF fit. In addition to the remaining SD station, also the information from one non-triggering station was used in the fit. The result of the photon-optimized fit is shown in Fig. 7.11. It has been checked that the signals measured in the excluded test stations is compatible with the shown fit within the uncertainties of the signals. The $S_{1000|\gamma}$ obtained from the fit, (1.69 ± 0.83) VEM, is significantly smaller than the expected S_{1000} ($S_{1000|\text{Hybrid}}$) at this energy and zenith angle, which is (5.96 ± 0.8) VEM. It is interesting to note that the result of the LDF fit in the standard SD reconstruction is more compatible with $S_{1000|\text{Hybrid}}$ than the result of the photon-optimized fit, although the results of the two fits are based on different assumptions and are therefore not directly comparable. The observable F_{γ} is determined from the two S_{1000} values. The observed value $F_{\gamma} = 0.28 \pm 0.14$ suggests that the photon candidate event was initiated by a primary photon. Compared to the expectation from the MC simulations shown in Fig. 6.19, it is found that only 0.3 % of all simulated events with protons as primary particles exhibit a smaller F_{γ} value, and about 80 % of all simulated photon-induced events exhibit a larger value of F_{γ}. As with the observable X_{max}, this is not conclusive evidence for the hypothesis that the photon candidate event was indeed initiated by a photon.

7.2.3. Study of the Arrival Direction

In the following section, the arrival direction of the photon candidate event is studied in detail and compared to the known positions of astrophysical objects. If a correlation is found between the reconstructed arrival direction of the photon candidate event and the position of an astrophysical object which could be the source of UHE photons, this may serve as an additional hint that the photon candidate event was induced by a photon. If no such correlation is found, then this is not an indication that the primary particle was not a photon, since

the most probable production mechanism for UHE photons is the GZK mechanism, which can take place anywhere in the Universe.

Using the known position and orientation of the Earth at the time the photon candidate event was recorded and the reconstructed shower geometry, the arrival direction of the event in celestial coordinates can be determined. In galactic coordinates, which are oriented such that the primary direction is aligned with the direction of the center of the Milky Way and the fundamental plane is in the galactic plane, the reconstructed coordinates of the arrival direction of the photon candidate event are $(l, b) = (337.0 \pm 0.5, 25.4 \pm 0.6)°$. In the following, this arrival direction is compared to the known positions of astrophysical objects from a number of catalogs. The choice of the catalogs is mainly motivated by current theories for the origin of UHE photons and UHECRs in general.

The first set of astrophysical objects is a compilation of different catalogs, which were already used in a targeted search for point sources of UHE neutrons using data from the Pierre Auger Observatory [Aab14d]. The use of the same catalogs for a targeted search for UHE photons, which is based on the analysis described in Sec. 3.4.2, is currently being evaluated as well [Kue14]. The set consists of 366 objects in nine subsets of different classes, including millisecond and γ-ray pulsars, low-mass and high-mass X-ray binary systems, different classes of TeV-γ-ray sources identified by H.E.S.S. as well as microquasars and magnetars. Common to all subsets is that only objects with a declination below 20° in equatorial coordinates—which corresponds to the total field of view of the Pierre Auger Observatory for zenith angles below 60°—are considered. For a detailed description of the different subsets and the selection process, see [Sal13]. In Fig. 7.12, the positions of the objects in the different subsets are shown in a Mollweide projection of the sky in galactic coordinates. This map and the subsequent maps have been created using the Auger Coverage & Anisotropy Toolkit [Ham10] version 3.0, which is based on the Hierarchical Equal Area isoLatitude Pixelization (HEALPix) package [Gór05]. The selected objects shown in Fig. 7.12 mainly populate the region around the galactic center and along the galactic plane. For each object in the different subsets, the angular distance ρ between the position of the object and the arrival direction of the photon candidate event is calculated according to spherical trigonometry. For most objects, the angular distance is large. The five objects with the smallest angular distance and their locations are listed in Tab. 7.2. The closest object is 4U 1456-32, also known as Cen X-4, which is a binary system visible in the X-ray part of the electromagnetic spectrum. The angular distance of this object is about 4.5°. This distance is too large for the object to be reliably considered a source candidate for the photon candidate event, even when the uncertainties of the reconstructed arrival direction are taken into account.

The second set of astrophysical objects consists of AGNs from the 13th edition of the VCV catalog [Vér10]. The inclusion of this set is motivated by theories which extrapolate the flux of photons in the TeV range emitted by AGNs such as Centaurus A into the EeV regime [Kac10]. AGNs have also been studied in the context of UHECR (see Sec. 2.2.3), although it should be pointed out that these studies were carried out at higher energies, above 50 EeV. For this study, the same AGN selection as in [Abr07b] is used, i.e. only AGN with a redshift smaller than 0.024 are considered. In addition, the same declination cut as before is used. In total, 579 AGNs are considered. The distribution of the selected AGNs over the sky is shown in Fig. 7.13. As before, the five closest AGNs in terms of the angular distance are listed in Tab. 7.3. The
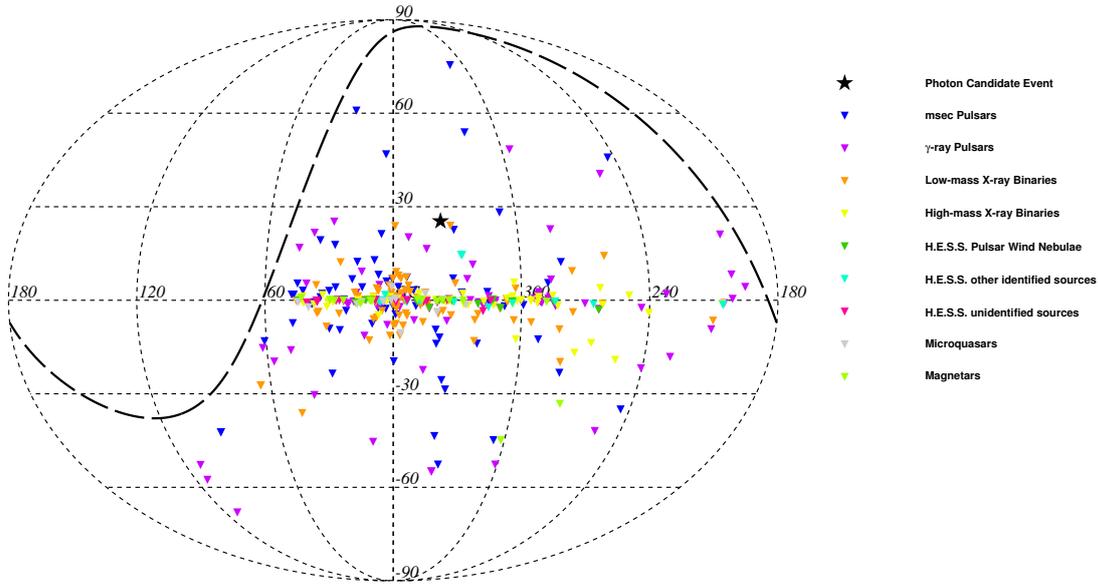
Figure 7.12.: The arrival direction of the photon candidate event (black star) and the locations of astrophysical objects from different classes (colored triangles), shown in a Mollweide projection of the sky in galactic coordinates. For a detailed description of the different classes of astrophysical objects, see [Sal13]. The uncertainty of the arrival direction of the photon candidate event is about 0.5° degree in both galactic longitude and latitude and therefore not visible in this scale.

Name	Subset	ρ [°]
4U 1456-32	low-mass X-ray binary system	4.54
J1455-3300	msec pulsar	6.36
J1600.7-3053	γ-ray pulsar	11.13
J1504-418	H.E.S.S. other	13.79
J1502-421	H.E.S.S. other	14.15

Table 7.2.: The five astrophysical objects from Fig. 7.12 which are closest to the arrival direction of the photon candidate event in terms of angular distance. ρ is the angular distance between the location of the object and the arrival direction.

closest AGN, at an angular distance of about 3.6°, is PKS 1521-30, a type-I Seyfert galaxy. However, also for this object, the angular distance is too large for the object to be considered a source candidate. For Centaurus A, which is the closest AGN from the Earth and therefore treated separately here, the angular distance is about 20°.

To conclude the discussion of the arrival direction of the photon candidate event, a set of blazars is considered. Although blazars are a subclass of AGNs, they are treated separately here, motivated by theories which postulate a conversion of photons into axions in the magnetic fields around various astrophysical objects, including blazars [Fai11]. Although such theories are currently highly hypothetical since the existence of axions has not been exper-
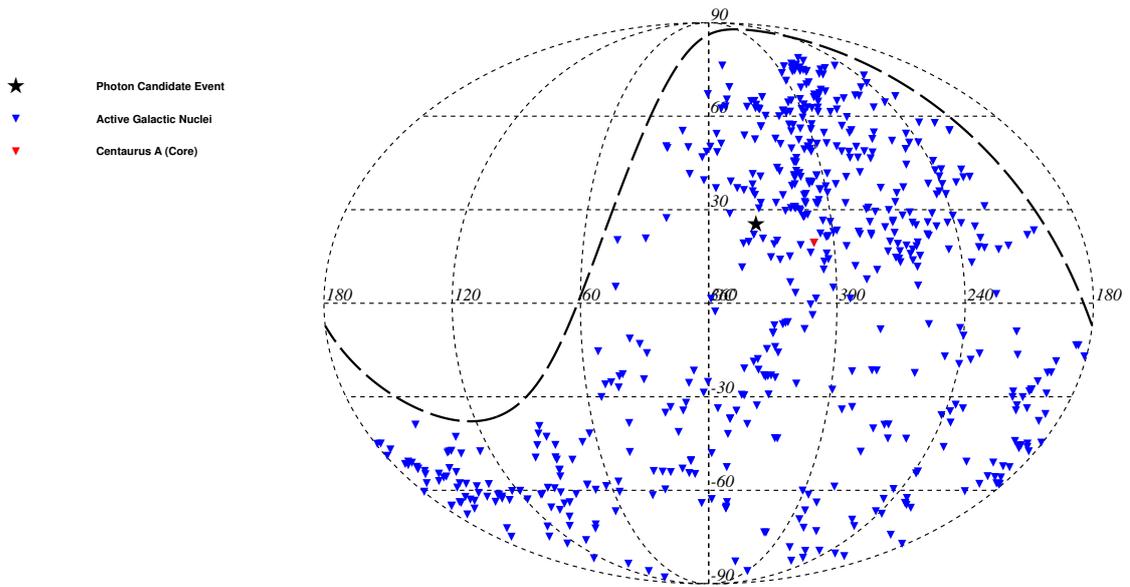
Figure 7.13.: The arrival direction of the photon candidate event (black star) and the locations of nearby AGN from the 13th edition of the VCV catalog (blue triangles), shown in a Mollweide projection of the sky in galactic coordinates. Centaurus A, which is the closest AGN from the Earth, is marked in red. The uncertainty of the arrival direction of the photon candidate event is too small to be visible in this scale.

Name	ρ [°]
PKS 1521-30	3.58
MCG -03.35.014	5.31
NVSS J15123-3333	5.47
NGC 5968	6.71
ESO 512-G20	7.32
Centaurus A	19.42

Table 7.3.: The five AGN from Fig. 7.13 which are closest to the arrival direction of the photon candidate event in terms of the angular distance. ρ is the angular distance between the location of the object and the arrival direction. In addition to the five closest AGN, Centaurus A is included in this list.

imentally proven, it is still worthwhile to investigate blazars as potential source candidates, in particular in view of claims of a correlation between the arrival directions of UHECRs and the positions of γ-ray-loud blazars [Gor02]. In this study, blazars from the 5th edition of the Roma BZCAT [Mas09] have been selected by the same criteria as previously applied. In total, 387 blazars are considered here. As in the case of the other sets, the distribution of the blazars over the sky is shown in Fig. 7.14, and the five closest blazars are listed in Tab. 7.4. With an angular distance of about 2.6°, the blazar 5BZBJ1522-2730 is not only the closest blazar to the arrival direction of the photon candidate, but also overall the closest astrophysical object considered in the study that is discussed in this section. However, the angular distance be-
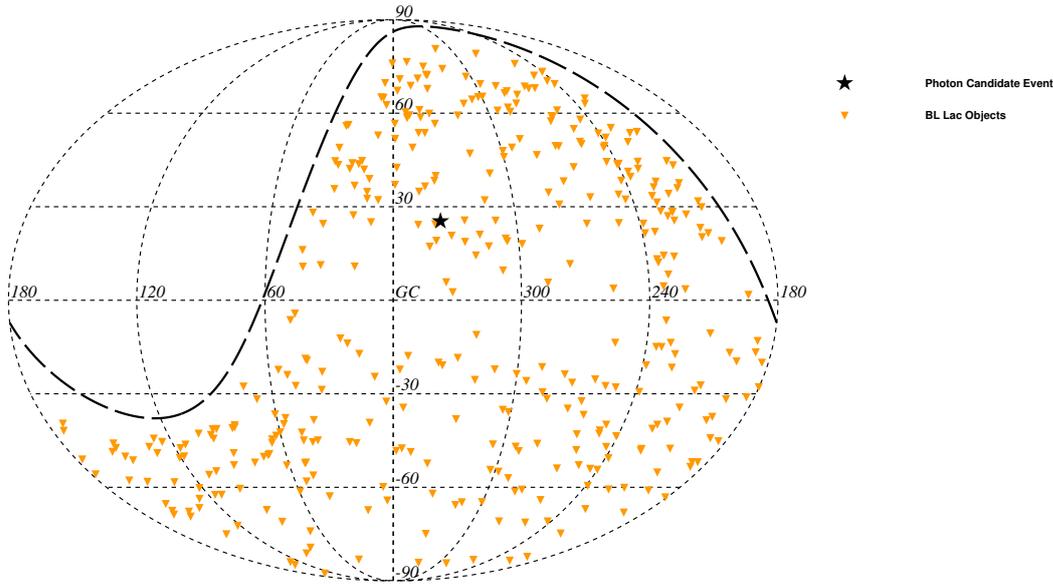
Figure 7.14.: The arrival direction of the photon candidate event (black star) and the locations of nearby blazars from the 5th edition of the Roma BZCAT catalog (orange triangles), shown in a Mollweide projection of the sky in galactic coordinates. The uncertainty of the arrival direction of the photon candidate event is too small to be visible in this scale.

Name	ρ [°]
5BZBJ1522-2730	2.57
5BZBJ1505-3432	6.59
5BZBJ1536-3151	6.70
5BZBJ1553-3118	9.72
5BZBJ1548-2251	9.95

Table 7.4.: The five blazars from Fig. 7.14 which are closest to the arrival direction of the photon candidate event in terms of the angular distance. ρ is the angular distance between the location of the object and the arrival direction.

tween the location of the blazar and the arrival direction of the photon candidate event is still too large for the object to be considered a source candidate.

Overall, the study presented in this section yields no indication of a known astrophysical object, of the types considered, as the source of the photon candidate event. The closest objects in the catalogs considered here have angular distances of a few degrees. For a photon directly emitted from an astrophysical object, a stronger correlation, i.e. a smaller angular distance, is be expected. However, the possibility of a UHE photon originating from the GZK effect and then impinging on the Earth cannot be tested with a study like the one shown here.

7.2.4. Possible Primary Particle Type

In Sec. 7.1, it was argued that the number of photon candidates obtained by applying the analysis to data is fully compatible with the expectation from a background comprised of air shower events initiated by protons and nuclei. However, this is no evidence toward a certain primary particle type, considering the total statistics, which is comparatively small. In this section, the hypothesis that the photon candidate event was not induced by a photon, but by a proton or a heavier nucleus, is investigated. Two different approaches are used here. To test whether the photon candidate event could be initiated by a proton, a full MC approach is used, whereas for nuclei, a simpler ansatz based on the expected distributions of the shower maximum X_{max} for a given primary energy and mass is employed. Both approaches will be discussed in the following sections.

Dedicated Proton Simulations

The study discussed in this section is based on dedicated MC simulations of proton-induced air showers. For these simulations, the reconstructed quantities of the photon candidate event, i.e. the reconstructed energy, zenith and azimuth angles, and core position have been used as input. To account for the uncertainties of the reconstructed quantities, a Gaussian smearing is applied to the input values for the simulations. The sequence used for these simulations is essentially the same as described in Sec. 5.1. Instead of the ideal detector used for the main simulation samples, the status of the detector systems, including the atmospheric conditions, at the time the photon candidate event was recorded is reproduced from the official databases.

In total, 2500 extensive air showers have been simulated using CORSIKA. In contrast to the main simulation sample, each CORSIKA shower is used only once as input for the $\overline{\text{Off}\underline{\text{line}}}$ part of the simulation sequence, since the core position of the simulated shower within the detector array is fixed, and therefore one degree of freedom is removed from the overall sample. If the showers were still used five times each, this could lead to a bias in the results since the impact of random shower-to-shower fluctuations can be overestimated in this case. Of the 2500 simulated events, 1516 were found to trigger the FD. This number is in agreement with the expected trigger efficiency of the FD around 1 EeV considering the distance between the core position and the FD telescope [Pet04]. Of the 1516 triggered events, 932 pass the event selection stage. Almost all of the events removed in the event selection were discarded by the cut on the uncertainty of F_γ. This is mostly due to the fact that the photon-optimized LDF fit is very sensitive to fluctuations in the signals recorded in the SD stations, in particular if there are only one or two triggered stations in the event. In Fig. 7.15, the distributions of the observables F_γ and X_{max} for the dedicated proton simulation sample are shown in comparison to the values of the respective observables from the photon candidate event. The observed values of F_γ and X_{max} are located in the tails of these distributions. About 3 % of the events in the proton sample exhibit an F_γ value smaller than the one observed in the photon candidate event. At the 1σ level, i.e. if the uncertainty on the observed value is taken into account, up to 20 % of the events exhibit a smaller value. For X_{max}, the corresponding numbers are about 1 % and up to 3 %. Although the average values for proton-induced showers with the given energy and core position of the distributions are rather far from the observed values, the hypothesis that the photon candidate event was initiated by a proton can not be excluded. This is even more apparent when the two observables are plotted against each other (Fig. 7.15(c)). Also here, the bulk of the simulated events is located rather far from the photon candidate event.

(a)

(b)

(c)
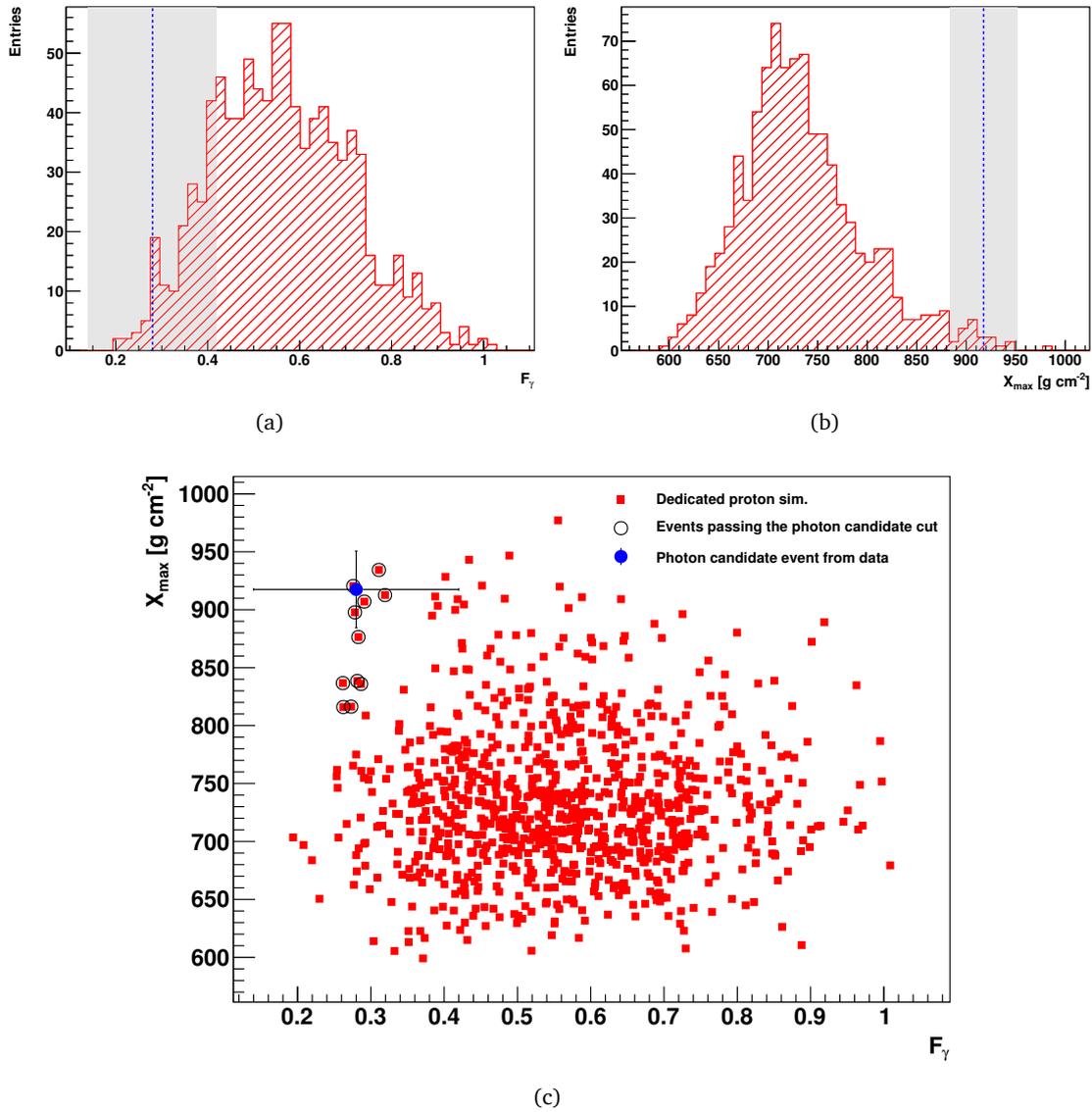
Figure 7.15.: Distributions of the observables (a) F_γ and (b) X_{max} for the sample of dedicated proton simulations. The values of the observables from the photon candidate event are indicated by the dashed line, with the shaded area denoting the uncertainties of the reconstructed values. In (c), both observables are shown in a scatter plot. The photon candidate event is indicated by a blue marker. Simulated events passing the photon candidate cut are marked with open circles.

However, a few simulated events do cluster around the candidate event. If the full analysis is applied to the dedicated sample, it is found that all of the simulated events clustered around the candidate event pass the photon candidate cut. In addition, a few events with smaller X_{max}, but comparable F_γ values also pass the candidate cut. In total, 11 events from the proton sample discussed here would pass the photon candidate cut. This corresponds to a fraction of 1.2 %, with uncertainties of +0.5 % and −0.3 %. From this result, the hypothesis that the photon candidate event was induced by proton can neither be confirmed nor excluded.

Drawing on the results obtained in Sec. 6.4.3, the impact of exchanging the hadronic interaction model for the dedicated proton simulations can also be estimated. Compared to QGSJETII-04, EPOS-LHC on average produces deeper, i.e. more photon-like, longitudinal shower profiles. As in Sec. 6.4.3, this would lead to more proton events passing the photon candidate cut. The same overall result is found for SIBYLL 2.1. The smaller average muon content of air showers simulated with this model leads to a larger number of proton events passing the photon candidate cut compared to events simulated with QGSJETII-04.

Nuclei as Potential Primary Particles

In this section, it is tested whether the photon candidate event could be initiated by a helium nucleus or a heavier nucleus. In contrast to the previous section, where a full MC approach was employed, here an analytical ansatz is used. This ansatz is based on the parameterization of the distribution of the shower maximum X_{max} as a function of the primary energy E and the primary mass A [Dom13]. This parameterization uses the generalized Gumbel function

$$\mathcal{G}_{\text{PDF}}(X_{\text{max}} \,|\, E, A) = \frac{1}{\sigma} \frac{\lambda^{\lambda}}{\Gamma(\lambda)} \, e^{-\lambda z - \lambda e^{-z}}, \qquad \text{with } z = \frac{X_{\text{max}} - \mu}{\sigma}. \tag{7.1}$$

The parameters μ, σ and λ are functions of E and A. In addition, these parameters depend on the hadronic interaction model. A full description of the parameterization and its dependencies is given in [Dom13]. In Fig. 7.16, the probability density functions according to Eq. 7.1 for a primary energy of 1.20 EeV, corresponding to the reconstruced energy of the photon candidate event, and different nuclear masses are shown. With increasing mass, both the width of the distributions and the average X_{max} decrease, as expected from the superposition model (cf. Sec. 2.3.2). For X_{max} values above $800\,\text{g}\,\text{cm}^{-2}$, the helium probability dominates among all other probabilities.

In the following, the approach used in [Hei14] is adopted. From the parameterized Gumbel function, the probability for a single primary particle with energy E and mass A to produce an
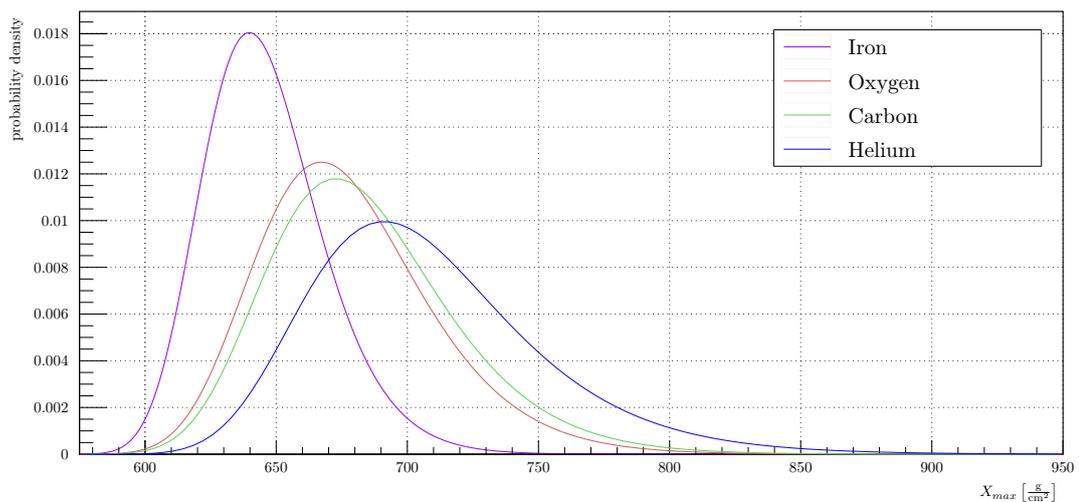
Figure 7.16.: Parameterized probability distributions for X_{max} obtained from the generalized Gumbel function (Eq. 7.1) for an energy of 1.2 EeV and different primary masses [Hei15].

air shower with $X_{\text{max}} \geq X_{\text{max}}^{\text{obs}}$ is determined through integration:

$$P(X_{\text{max}} \geq X_{\text{max}}^{\text{obs}} \,|\, E, A) = \int_{X_{\text{max}}^{\text{obs}}}^{\infty} \mathcal{G}_{\text{PDF}}(X_{\text{max}} \,|\, E, A) \, dX_{\text{max}}. \qquad (7.2)$$

Using Eq. 7.2, the probabilities for a primary particle of the same energy as in the reconstructed photon candidate event to produce an air shower with an X_{max} larger than the observed value of $917.5\,\text{g cm}^{-2}$ are determined under the assumptions of different hadronic interaction models (cf. Tab. 7.5). The probabilities for heavier nuclei, beginning with lithium, are very small and thus negilible compared to the probabilities for helium. Assuming QGSJETII-04 as the hadronic interaction model, the probability for a helium-induced air shower at this energy to exhibit a larger value of X_{max} is 0.084 %, and thus also negligible. These numbers corroborate the results obtained in Sec. 6.4.2, where it was also found that events induced by helium nuclei or heavier nuclei have a very small probability of passing the photon candidate cut.

The same approach can also be used to test the hypothesis of protons as primary particles. Assuming a primary proton with an energy of 1.20 EeV, Eq. 7.2 yields a probability of around 2 % [Hei15], with differences in the order of 0.5 % between the different hadronic interaction models. This number is comparable to that obtained from the integration of the X_{max} distribution shown in Fig. 6.2, which was based on a full MC approach, and can hence serve as an additional cross check.

Overall, the brief study discussed in this section indicates that the primary particle which induced the photon candidate event was not a helium nucleus or nucleus with heavier mass. The probability for a helium nucleus with an energy of 1.20 EeV to produce an X_{max} value equal to or larger than the one observed in the photon candidate event is more than an order of magnitude below the corresponding probability for a primary proton at this energy. For heavier nuclei, even smaller probabilities are found. On the other hand, since the calculated probability for a proton is 2 %, the hypothesis that the photon candidate event was induced by a proton cannot be excluded using the analytical X_{max}-based approach.

| Model | $P(X_{\text{max}} \geq X_{\text{max}}^{\text{obs}} \,|\, E, A)$ [%] | | | | |
	He ($A = 4$)	Li ($A = 7$)	C ($A = 12$)	O ($A = 16$)	Fe ($A = 56$)
QGSJETII-04	0.084	0.019	0.0031	0.0009	6.36e-08
SYBILL 2.1	0.122	0.031	0.0061	0.0022	3.37e-06
EPOS-LHC	0.075	0.031	0.0008	0.0001	1.19e-11

Table 7.5.: Probabilities $P(X_{\text{max}} \geq X_{\text{max}}^{\text{obs}} \,|\, E, A)$ obtained from the integration of the parameterized X_{max} distributions for a primary energy of 1.20 EeV and an observed X_{max} value of $917.5\,\text{g cm}^{-2}$, corresponding to the reconstructed quantities from the photon candidate event [Hei15]. Different hadronic interaction models have been used.

7.3. Determining Upper Limits in the EeV Range

In Sec. 7.1, the number of air shower events with photon-like characteristics according to the observables used in this analysis has been determined the hybrid data recorded by the Pierre Auger Observatory. It has also been estimated that the observed number of photon candidate events is compatible with the expectation from a pure hadronic background. In order to obtain from the observed number of photon candidate events the fraction of UHE photons in the total flux of cosmic rays in a statistically sound way, an exact knowledge of the expected background is neccessary. The exact determination of the hadronic background is difficult at best, mainly due to the uncertainties in the hadronic interaction models at the highest energies. To circumvent the need for the detailed knowledge of the hadronic background, upper limits on the fraction of UHE photons in the total flux of cosmic rays are determined, under the conservative assumption that the observed number of photon candidate events is only due to background events. In the following sections, the calculation of the upper limits on the photon fraction is discussed in detail. In addition, upper limits on the integral flux of photons impinging on the Earth are determined using the upper limits on the photon fraction and the all-particle energy spectrum measured by the Pierre Auger Observatory. The section closes with a discussion of the impact of various systematic uncertainties on the results obtained in this thesis.

7.3.1. Upper Limits on the Photon Fraction

The upper limit on the fraction f_γ of photons in the total flux of cosmic rays above a given energy threshold E_0 and for a given confidence level α is given by:

$$f_\gamma^\alpha(E > E_0) = \frac{N_\gamma^\alpha(E > E_0)}{N_{\text{total}}(E > E_0)}, \tag{7.3}$$

where $N_\gamma^\alpha(E > E_0)$ denotes the upper limit on the number of photon events above the threshold energy, determined at the confidence level α, while $N_{\text{total}}(E > E_0)$ is the total number of events in the data sample above the threshold energy. In order to compare the upper limits obtained here with previous results, a confidence level of 95 % is chosen. As the true type of the primary particle for a given event in the data sample is not known, the missing energy correction for photons is applied to all events in the data sample in order to obtain the number of photon candidate events and the total number of events above the given threshold, i.e. $N_\gamma^\alpha(E > E_0)$ in Eq. 7.3 is replaced by $N_\gamma^\alpha(E_\gamma > E_0)$ and $N_{\text{total}}(E > E_0)$ is replaced by $N_{\text{total}}(E_\gamma > E_0)$. If the larger missing energy correction for hadrons was used instead, this would in turn lead to a larger total number of events above the threshold. Hence, using the smaller correction for photons leads to a larger, more conservative, upper limit on the fraction.

The upper limit on the number of photon events is determined from the observed number of photon candidate events $N_{\gamma,\text{obs}}^{95}(E_\gamma > E_0)$, again at a confidence level of 95 %, after correcting this number for the acceptance of the dector and the efficiency of the photon candidate cut. Since the acceptances of the detector systems are different for photon-induced and proton-induced air showers, the recorded data sample is biased toward proton-induced air showers. In the event selection, the fiducial volume cut has been introduced to counteract the different acceptances, however, this cut cannot fully remove the bias. Hence, an additional acceptance correction factor κ is used here for the calculation of the upper limits. The acceptance cor-

rection is based on the relative acceptances $\mathscr{A}$ for photon-induced air showers and showers initiated by protons or nuclei:

$$\kappa(E_0) = \frac{\min\left[\mathscr{A}_\gamma(E)\,|\,E > E_0\right]}{\max\left[\mathscr{A}_{\mathrm{p/nucl}}(E)\,|\,E > E_0\right]}.$$ (7.4)

The acceptance correction is derived in detail in App. D. At this point, a less conservative approach is used by limiting the energy range considered for the relative acceptances also toward higher energies:

$$\kappa(E_0) = \frac{\min\left[\mathscr{A}_\gamma(E)\,|\,E_0 < E < E_{\max}\right]}{\max\left[\mathscr{A}_{\mathrm{p/nucl}}(E)\,|\,E_0 < E < E_{\max}\right]}.$$ (7.5)

The maximum energy $E_{\max}$ is based on the energy distribution found in the data sample (Fig. 7.1). From this distribution, $E_{\max}$ is determined for each threshold energy E_0 such that the fraction of events with an energy between E_0 and $E_{\max}$ compared to the total number of events with energies above E_0 is 80%. The acceptances for photon-induced and proton-induced air showers, as obtained from the MC simulations discussed in Chap. 5, are shown in Fig. 7.17(a) as a function of energy. The acceptances are given as relative acceptances, i.e. normalized such that the acceptance for proton-induced air showers at $\log_{10}(E\,[\mathrm{eV}]) = 18.5$ is equal to 1. For comparison, the acceptances for iron-induced air showers are shown as well. At all energies, the acceptance for iron-induced air showers is smaller than the acceptance for proton-induced showers. Hence, only proton-induced showers are relevant for Eq. 7.5. In order to obtain the relative acceptances as continuous functions of energy, linear functions are used:

$$\begin{aligned} \mathscr{A}_{\gamma,\mathrm{rel}}(E) &= d_0 + d_1 \log_{10}(E\,[\mathrm{eV}]), \\ \mathscr{A}_{\mathrm{p,rel}}(E) &= e_0 + e_1 \log_{10}(E\,[\mathrm{eV}]). \end{aligned}$$ (7.6)

A fit to the data points shown in Fig. 7.17(a) yields the following values for the parameters:

$$\begin{aligned} d_0 &= 14.8 \pm 0.5, \\ d_1 &= 0.85 \pm 0.02, \\ e_0 &= 15.1 \pm 0.5, \\ e_1 &= 0.87 \pm 0.03. \end{aligned}$$ (7.7)

The second correction to the observed number of photon candidate events accounts for the selection efficiency of the photon candidate cut. In Sec. 6.3.3, only the overall selection efficiency, averaged over all energies, has been discussed. For the calculation of upper limits at different energy thresholds, however, the selection efficiency ϵ_{cand} of the photon candidate cut as a function of energy is needed. In Fig. 7.17(b), the selection efficiency for photons, obtained from the main simulation sample, is shown. As before, the energy dependence is modeled by a linear function:

$$\epsilon_{\mathrm{cand}}(E) = g_0 + g_1 \log_{10}(E\,[\mathrm{eV}]).$$ (7.8)

The fit of this function to the data points results in the following values for the parameters:

$$\begin{aligned} g_0 &= -1.2 \pm 0.2, \\ g_1 &= 0.084 \pm 0.010. \end{aligned}$$ (7.9)
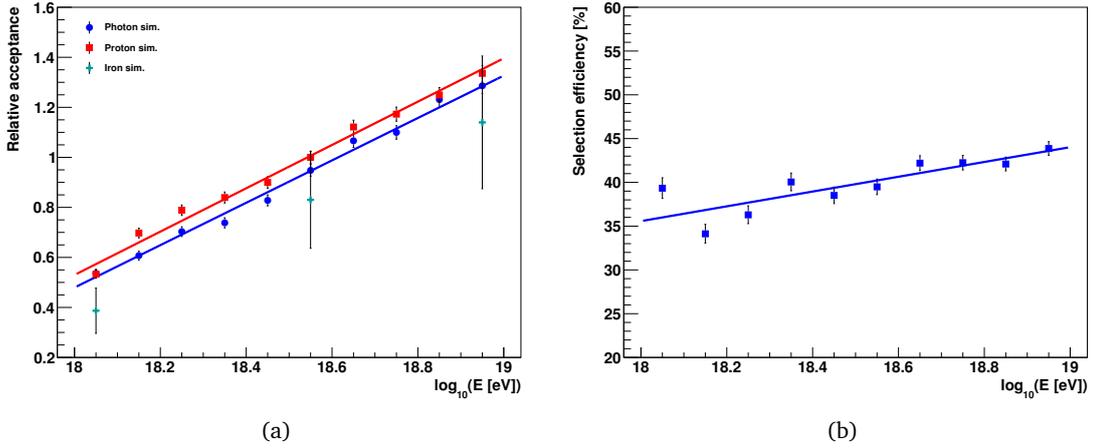
Figure 7.17.: (a) The relative acceptance $\mathscr{A}_{\mathrm{rel}}$ as a function of energy for the photon (blue circles), proton (red squares) and iron (turquoise crosses) samples, respectively. The acceptances have been normalized so that the relative acceptance for the proton sample at $\log_{10}(E[\mathrm{eV}]) = 18.5$ is equal to 1. The colored lines denote the fits to the data points according to Eq. 7.6. (b) The selection efficiency of the photon candidate cut for the photon sample as a function of energy. The line denotes a fit to the data points according to Eq. 7.8.

For calculating the upper limits, the minimum of $\epsilon_{\mathrm{cand}}(E)$ in the range above E_0 is used in order to maximize the upper limits. Since $\epsilon_{\mathrm{cand}}(E)$ is strictly monotonic increasing, this means that the minimum of the function will always be at the energy threshold.

The observed number of photon candidate events $N_{\gamma,\mathrm{obs}}^{95}(E_\gamma > E_0)$ at a confidence level of 95 % is determined from the number of photon candidates identified by the analysis presented in Chap. 6 according to Poisson statistics. The number ν of photon candidate events is expected to fluctuate according to a Poisson distribution [Poi37] with mean μ:

$$P(\nu \,|\, \mu) = \frac{\mu^\nu \, e^{-\mu}}{\nu!}. \tag{7.10}$$

In order to impose an upper limit on ν at a confidence level α, it is required that only for a fraction $1 - \alpha$ of identical measurements, a number of photon candidate events smaller or equal to the observed number of photon candidate events n is observed. Hence:

$$1 - \alpha = \sum_{\nu=0}^{n} P(\nu \,|\, \mu) = \sum_{\nu=0}^{N} \frac{\mu^\nu \, e^{-\mu}}{\nu!}. \tag{7.11}$$

The value of μ obtained by this equation is then the upper limit on ν at the given confidence level, i.e. in this case $\mu = N_{\gamma,\mathrm{obs}}^{95}(E_\gamma > E_0)$ with $n = N_{\gamma,\mathrm{obs}}(E_\gamma > E_0)$ as the number of photon candidate events observed in the analysis above a given energy threshold. It should be pointed out here again that no background is subtracted from $N_{\gamma,\mathrm{obs}}(E_\gamma > E_0)$ prior to calculating $N_{\gamma,\mathrm{obs}}^{95}(E_\gamma > E_0)$.

Putting everything together, Eq. 7.3 becomes

$$f_\gamma^{95}(E > E_0) = \frac{N_{\gamma,\mathrm{obs}}^{95}(E_\gamma > E_0)\, \frac{1}{\epsilon_{\mathrm{cand}}(E_0)}\, \frac{1}{\kappa(E_0)}}{N_{\mathrm{total}}(E_\gamma > E_0)}, \tag{7.12}$$

In this work, upper limits on the photon flux at threshold energies of 1, 2, 3, 4 and 5 EeV are determined. The maximum energies E_{max} corresponding to these thresholds are 2.9, 5.2, 7.6, 10.0 and 12.2 EeV. It has been decided to not determine upper limits at higher energies, because this would mean that the results obtained so far have to be extrapolated significantly into an energy range for which the analysis has not been optimized. In Sec. 6.4.2 it has been shown that this extrapolation does not lead to a larger number of background events being identified as photon candidate events. For the calculation of upper limits on the photon fraction, on the other hand, additional parameterizations for κ and ϵ are used. These functions cannot be reliably extrapolated to much higher energies. For example, to calculate an upper limit above 10 EeV, an extrapolation of the relative acceptances shown in Fig. 7.17(a) would yield steadily increasing values. However, from the simulation samples used in Sec. 6.4.1, it is known that the overall acceptance in fact decreases due to the fiducial volume cut. Hence, κ is not estimated correctly for such thresholds and an extrapolation to much larger energies should be avoided. For each of the energy thresholds listed previously, the corresponding upper limit on the fraction of photons in the total flux of primary cosmic rays is calculated according to Eq. 7.12. The individual quantities needed in Eq. 7.12 for the different thresholds and the calculated upper limits are listed in in Tab. 7.6.

The upper limits derived in this work are now compared to previous upper limits which were determined from the data recorded at the Pierre Auger Observatory and the expecta-

E_0 [EeV]	$N_{\gamma,\text{obs}}(E_\gamma > E_0)$	$N^{95}_{\gamma,\text{obs}}(E_\gamma > E_0)$	$N_{\text{total}}(E_\gamma > E_0)$	$\epsilon_{\text{cand}}(E_0)$	$\kappa(E_0)$	$f^{95}_\gamma(E > E_0)\,[\%]$
1	1	4.74	6727	0.36	0.51	0.39
2	0	3.00	2575	0.38	0.64	0.48
3	0	3.00	1302	0.40	0.68	0.85
4	0	3.00	771	0.41	0.71	1.35
5	0	3.00	537	0.41	0.73	1.85

Table 7.6.: Upper limits f^{95}_γ for different energy thresholds E_0 and the individual quantities needed to calculate the upper limits according to Eq. 7.12.

	$f^{95}_\gamma(E > E_0)\,[\%]$	
E_0 [EeV]	[Abr09]	[Set11]
1.16	-	0.43
2.3	3.8	0.47
3.45	2.4	1.09
5.7	3.5	2.76
11.2	11.7	9.45

Table 7.7.: Upper limits on the fraction of UHE photons in the total flux of cosmic rays above the EeV range derived in previous publications based on data recorded at the Pierre Auger Observatory. The correction of the overall energy scale used in the event reconstruction between the publication of the limits and the present day has been taken into account here by shifting the energy threshold E_0. The upper limits on the photon flux from [Set11] have been translated into fraction limits using the energy spectrum from [Sch13].
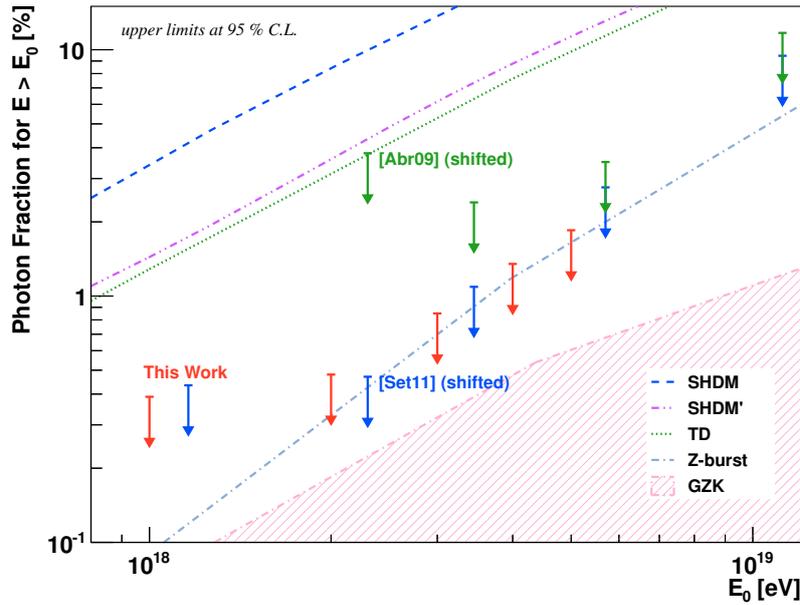
Figure 7.18.: The upper limits on the fraction of UHE photons in the total flux of cosmic rays, at a confidence level of 95 %, obtained in this work compared to previous results and the predictions of current theoretical models. The upper limits from [Abr09] and [Set11] have been shifted to higher energies to take into account the correction of the overall energy scale (cf. Tab. 7.7). For the theoretical models, the predictions for the integral flux of photons (see Fig. 3.5) have been translated into fractions using the energy spectrum as measured by the Pierre Auger Observatory. For a complete list of references, see Fig. 3.5.

tions from current theoretical models. Before the limits can be compared, however, it needs to be taken into account that the overall energy scale used in the reconstruction of events from the data was corrected between the publication of the limits and the present day. The overall change ranges from +16 % at 1 EeV to +12 % at 10 EeV [Ver13]. This change in the energy scale means that the upper limits from previous publications have to be interpreted as referring to a higher energy threshold, because a quoted threshold of 1 EeV is in fact a threshold of 1.16 EeV. In the following, the upper limits from previous publications are shifted accordingly. The upper limits quoted in [Abr09] were obtained using only the observable X_{max}. In [Set11], upper limits on the integral flux of photons impinging on the Earth were determined using a Fisher disciminant analysis combining X_{max} and S_b. The upper limits on the photon flux from [Set11] were translated into fraction limits using the all-particle energy spectrum measured by the Pierre Auger Observatory as derived in [Sch13] (cf. Sec. 7.3.2). In Tab. 7.7, the shifted upper limits on the photon fraction from [Abr09] and [Set11] are listed. These limits are shown in comparison to the upper limits derived in this work and the predictions of current theoretical models in Fig. 7.18. The upper limits determined in this work are up to an order of magnitude below the limits derived in [Abr09]. The improvement is mainly due to the increase in separation power provided by the additional observable F_γ. Compared to [Set11], a slight improvement at higher energies above 5 EeV can be found. At lower energies, the upper limits obtained in this work are on par with [Set11]. Although the analysis presented in this thesis is based on about 2 years worth of additional data compared to [Set11], there are no significant differences between the results of the analyses, mostly due to the more conser-

vative approach used in the calculation of the upper limits in this work (see also the following section). Overall, the upper limits obtained here corroborate previous results in putting severe constraints on theoretical models following the top-down approach, in particular those involving SHDM and TD. For the Z-burst scenario, the limits are still compatible with the predicted photon fractions. Currently, the limits are also above the region of fractions expected from GZK processes. However, it can be assumed that this region can be reached within the next few years with an increased data sample.

7.3.2. Upper Limits on the Integral Photon Flux

Upper limits on the integral flux of photons impinging on the Earth with energies above a given threshold can be determined in two ways, either directly or indirectly. In the direct approach, used in [Set11] and [Abr08b], the number of photon candidate events above a given threshold is divided by the exposure of the detector for photon-induced air showers. In the case of data from the SD, the integrated exposure can be calculated geometrically. For hybrid data, the exposure cannot be calculated analytically, but it has to be inferred from MC simulations [Abr11]. For these simulations, the full detector status of both the FD, including the atmospheric conditions, and the SD has to be reproduced for any time in the period where the data considered in the analysis was recorded. In addition, the full energy range has to be covered up to the highest energies with sufficient statistics. This approach is very demanding in terms of computing resources needed for the reliable calculation of the hybrid exposure, and it is therefore not followed here.

In the indirect ansatz, the upper limits f_γ^{95} on the fraction are translated into upper limits Φ_γ^{95} on the integral flux by multiplying the fraction limits with the integral flux of cosmic rays obtained from the all-particle energy spectrum $J(E)$:

$$\Phi_\gamma^{95}(E > E_0) = f_\gamma^{95}(E > E_0) \times \int_{E_0}^{\infty} J(E)\, \mathrm{d}E. \tag{7.13}$$

In the following, the energy spectrum determined according to [Sch13] is used. The energy spectrum parameterized therein by a single power law below the ankle and a power law with a smooth suppression above the ankle:

$$J(E) \propto \begin{cases} E^{-\gamma_1}, & \text{for } E \leq E_a, \\ E^{-\gamma_2} \times \left[1 + \exp\left(\frac{\log_{10}(E\,[\mathrm{eV}]) - \log_{10}(E_{1/2}\,[\mathrm{eV}])}{\log_{10}(W_c)} \right) \right]^{-1} & \text{for } E > E_a. \end{cases} \tag{7.14}$$

In this parameterization, γ_1 and γ_2 are the spectral indices below and above the ankle, respectively. The ankle itself is located at the energy E_a. $E_{1/2}$ is the energy at which the differential flux has dropped to half of its peak value before the suppression, the steepness of which is described by W_c. The values for the different parameters given in [Sch13] are

$$\begin{aligned} \gamma_1 &= 3.23, \\ \gamma_2 &= 2.63, \\ \log_{10}(E_a\,[\mathrm{eV}]) &= 18.72, \\ \log_{10}(E_{1/2}\,[\mathrm{eV}]) &= 19.63, \\ \log_{10}(W_c) &= 0.15. \end{aligned} \tag{7.15}$$

The overall normalization of the energy spectrum is given by $\log_{10}\left(J(E = 10^{18.5}\,\text{eV})\right) = -17.85$ [Sch14]. The differential flux of cosmic rays impinging on the Earth as a function of energy as calculated from Eq. 7.14 is shown in Fig. 7.19.

In order to calculate the upper limits on the integral flux of photons, the energy spectrum is integrated above the different energy thresholds E_0. The integrated fluxes as well as the corresponding upper limits on the integral flux of photons above this energy threshold are listed in Tab. 7.8. These limits are shown in comparison to previous results and the predictions of current theoretical models in Fig. 7.20. The upper limits determined in [Set11] were, taking into account the correction of the energy scale, $0.082\,\text{km}^{-2}\,\text{sr}^{-1}\,\text{yr}^{-1}$ above $1.16\,\text{EeV}$ and $0.02\,\text{km}^{-2}\,\text{sr}^{-1}\,\text{yr}^{-1}$ above 2.3, 3.45 and 5.7 EeV. When comparing the limits derived in this work to the upper limits from [Set11], the differences in the methods to calculate the upper limits have to be kept in mind. The direct ansatz to determine the flux limits from [Set11] is based solely on the exposure of the hybrid detector for photon-induced air showers, whereas
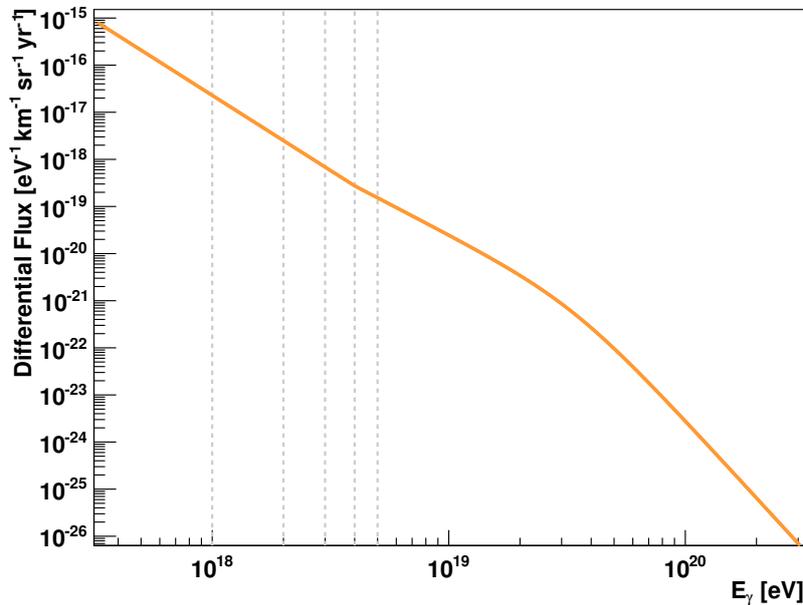
Figure 7.19.: The differential flux of cosmic rays, obtained from Eq. 7.14, as a function of E. The energy thresholds E_0 for which the upper limits are derived are marked by the dashed lines.

E_0 [EeV]	$\int_{E_0}^{\infty} J(E)\,\text{d}E\,[\text{km}^{-2}\,\text{sr}^{-1}\,\text{yr}^{-1}]$	$\Phi_\gamma^{95}(E > E_0)\,[\text{km}^{-2}\,\text{sr}^{-1}\,\text{yr}^{-1}]$
1	26.22	0.101
2	5.73	0.027
3	2.43	0.021
4	1.37	0.019
5	0.91	0.017

Table 7.8.: Upper limits Φ_γ^{95} for different energy thresholds E_0 and the integrated fluxes obtained from the energy spectrum above the given thresholds.
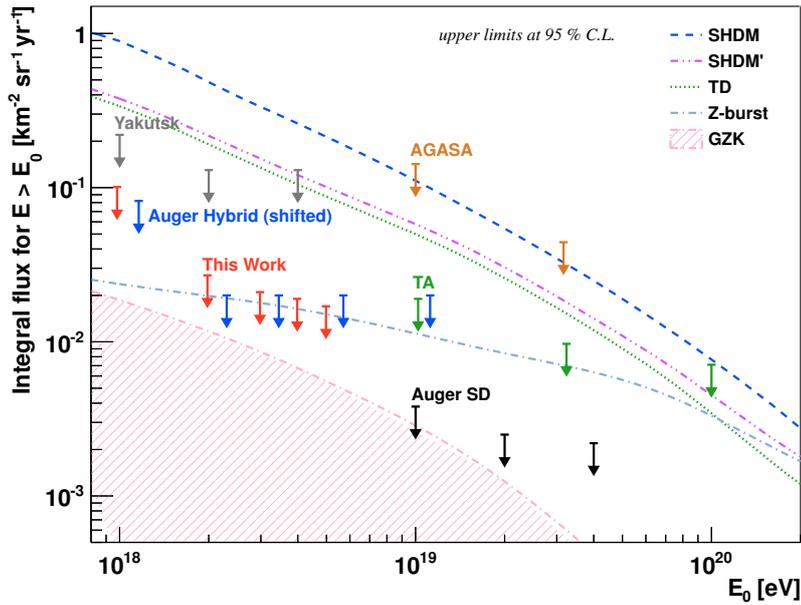
Figure 7.20.: The upper limits on the integral flux of UHE photons, at a confidence level of 95 %, obtained in this work compared to previous results and the predictions of several theoretical models. The upper limits published by the Pierre Auger Observatory have been shifted to higher energies to take into account the correction of the overall energy scale. For a complete list of references, see Fig. 3.5.

in the indirect ansatz, also the acceptances for proton-induced showers have to be taken into account, which in a conservative estimation leads to higher upper limits. Nevertheless, the upper limits on the integral flux of photons impinging on the Earth are comparable to the upper limits determined in [Set11], which are the most stringent limits determined so far in the EeV range.

7.3.3. Discussion of Systematic Effects

To conclude this chapter, the robustness of the results against different sources of systematic effects is discussed. The systematic uncertainty in the determination of the observable X_{max} is smaller than $10\,\mathrm{g\,cm^{-2}}$ for all energies [Aab14a]. The main contributions originate from the detector calibration ($3\,\mathrm{g\,cm^{-2}}$), the reconstruction of the longitudinal shower profile from the measured light (up to $8\,\mathrm{g\,cm^{-2}}$, depending on the energy) and the atmospheric monitoring (up to $8\,\mathrm{g\,cm^{-2}}$, depending on the energy). At the highest energies, the uncertainties in the atmospheric monitoring dominate the systematic uncertainties in the determination of X_{max} [Aab14a]. When the reconstructed X_{max} values for the data sample are increased (reduced) by $\Delta X_{\mathrm{max}}^{\mathrm{sys}} = 10\,\mathrm{g\,cm^{-2}}$ while keeping the simulation samples unchanged—which is justified since the simulations are based on an ideal detector; hence the uncertainties of, e.g., the atmospheric monitoring play no role in the reconstruction of the simulated data—then the number of photon candidate events changes by $+1$ (±0) between 1 and 2 EeV. Consequently, the upper limit on the photon fraction above 1 EeV changes by $+0.13\,\%$ ($\pm0\,\%$). The limits above the other thresholds remain unaffected.

Another possible source of systematic uncertainty originates from the LDF fit used to determine F_γ. An uncertainty of about 6% can be taken an estimate for the systematic uncertainty of F_γ, when it assumed that the systematic uncertainties in the photon-optimized LDF fit are similar to the ones in the standard SD reconstruction [Ave07a]. When the reconstructed F_γ values for the data sample are increased (reduced) by $\Delta F_\gamma^{\mathrm{sys}}/F_\gamma = 6\%$, again keeping the simulation samples unchanged, then the number of photon candidate events changes by ± 0 $(+1)$ between 1 and 2 EeV. This changes the upper limit on the photon fraction above 1 EeV by $\pm 0\%$ $(+0.13\%)$. As before, the limits above the other thresholds are not changed.

On the simulation side, the extrapolation of the photonuclear cross section into the UHE regime leads to a systematic uncertainty in the predictions of MC simulations for the average longitudinal shower profiles of photon-induced air showers. This uncertainty can be estimated to be below $10\,\mathrm{g\,cm^{-2}}$ [Ris06]. While this uncertainty does not affect the number of photon candidates obtained from data, it does have an impact on the correction that is applied to this number for the determination of the upper limits on the photon fraction, because the correction has been determined using the simulated photon sample. Uncertainties in the predictions for X_{max} for the photon sample lead to uncertainties in the selection efficiency of the photon candidate cut. When all X_{max} values in the photon sample are increased (reduced) by $\Delta X_{\mathrm{max}}^{\mathrm{sys},\gamma} = 10\,\mathrm{g\,cm^{-2}}$, then the selection efficiency changes by about $+5\%$ (-5%), independent of the energy. The resulting relative effect on the upper limits on the photon fraction is about -11% $(+15\%)$.

Another systematic uncertainty of air shower simulations in general comes from the choice of the hadronic interaction model. However, only simulations of air showers initiated by protons or nuclei are noticeably affected by such uncertainties. Photon-induced air showers do not develop a significant hadronic component, hence uncertainties arising from the treatment of this component can be neglected in this case. The uncertainties in the modelling of proton- and nucleon-induced air showers, i.e. the background events for the analysis presented in this work, would would have a large impact if the exact fraction of photons in cosmic rays was to be determined. However, for the conservative approach of determining upper limits on the fraction, the uncertainties in the modelling of the background have no influence on the results.

The overall energy scale to which the upper limits derived before refer to has a total systematic uncertainty of 14% [Ver13]. The main contributions to the systematic uncertainty of the energy scale originate from the atmospheric monitoring (up to 6.2%, depending on the energy), the calibration of the FD (9.9%), the uncertainties in the energy reconstruction from FD data (up to 6.5%, depending on the energy) and from the overall stability of the energy scale (5%) [Ver13]. Taking into account the total systematic uncertainty of the energy scale, i.e. shifting all energy values upward or downward, the numerical values of the upper limits above a given threshold energy E_0 can therefore be interpreted as referring to an effective threshold E_0', with $E_0' = k_{\mathrm{eff}} \times E_0$, where $k_{\mathrm{eff}} \in [0.86, 1.14]$ describes the systematic shift in the energy scale.

For the indirect calculation of the upper limits on the integral flux of photons, also the systematic uncertainties in the determination of the energy spectrum are relevant. For the energy spectrum presented in [Sch13], sources of systematic uncertainties include, for example, the determination of the exposure of the hybrid detector and the reconstruction of the

energy of the primary particle from hybrid data. Varying the parameters of the energy spectrum (Eq. 7.15) within the systematic uncertainties quoted in [Sch13] changes the integrated flux above a given threshold by less than 5 %. The impact of the aforementioned systematic uncertainties on the calculated upper limits on the integral photon flux is thus also less than 5 %.

8

Summary and Outlook

The subject of this thesis is the search for photons in the EeV range in cosmic-ray data collected at the Pierre Auger Observatory. Photons, in general, are the main messenger particles for the exploration of the Universe, with observational windows ranging from less than 10^{-8} eV to more than 10^{14} eV. But photons are not the only particles of cosmic origin impinging on the Earth. There are also cosmic rays—which are the most energetic particles known so far, with energies of more than 10^{20} eV—and by studying them, a deeper understanding of the Universe can be achieved. At the highest energies, cosmic rays and photons are intimately connected, as many of the current theories for the origin of Ultra-High-Energy Cosmic Rays (UHECRs) also predict the existence of photons in this energy range. An observation of such photons would therefore not only open a new observational window to the Universe, but it would also help in finding answers to the most fundamental questions about the origin of the UHECRs. At the Pierre Auger Observatory, the extensive air showers initiated by Ultra-High-Energy (UHE) particles, when they interact with particles from the Earth's atmosphere, are detected using a hybrid approach, combining the Fluorescence Detector (FD), which uses the air-fluorescence technique, with the Surface Detector (SD), a ground array of particle detectors. Due to its size, the Pierre Auger Observatory offers an unprecedented exposure to UHE particles and thus provides a unique opportunity to search for UHE photons.

The main challenge in the search for UHE photons is distinguishing air showers induced by photons from the hadronic background which consists of air showers initiated by protons and nuclei. In this thesis, a combination of the atmospheric depth of the shower maximum, $X_{\max}$, which is an FD observable, and the novel SD-related observable F_γ was used to fully take advantage of the capabilities of the hybrid approach. The observable F_γ is based on the photon-optimized fit of a Lateral Distribution Function (LDF) to the signals recorded in the individual detector stations of the SD array. In a simulation study, it was shown that the performance of F_γ in separating photon-induced air showers from the hadronic background is better than the performance of other SD-related observables currently used in the search for UHE photons. The two observables F_γ and $X_{\max}$ are combined in a multivariate Fisher discriminant analysis. The overall separation power of the Fisher discriminant is significantly larger than that of the individual observables. Around 1 EeV, the merit factor η of the Fisher discriminant is 2.2, and it increases to 3.2 around 10 EeV. Based on Monte Carlo (MC) simulations, several criteria have been developed to identify photon-like air shower events in data using the afore-

mentioned observables.

The analysis was then applied to hybrid data recorded at the Pierre Auger Observatory in the period between December 2004 and December 2012. Out of 6727 events which were analyzed here, one air shower event was found with photon-like characteristics. This number is in agreement with the number of background events from a pure hadronic background, which was estimated from MC simulations. A detailed examination of the photon candidate showed that the event was well-measured and well-reconstructed. A dedicated study to investigate the primary particle type gave inconclusive results, the hypothesis that the photon candidate event was not induced by a photon, but by a proton, can hence not be excluded. From the results of the analysis, upper limits on the fraction of photons in the total flux of cosmic rays above the EeV range and upper limits on the integral flux of photons above this energy range have been determined, both of which are summarized in Tab. 8.1. The upper limits obtained with the novel observable F_γ are on par with the most stringent limits from previous analyses, even improving them above 5 EeV. These results illustrate well the capabilities of a combination of F_γ and X_{max} thus underline the advantages of the use of this observable in the search for UHE photons.

In the near future, it will be possible to further improve the limits obtained from this analysis using the addtional data collected at the Pierre Auger Observatory. If the size of the data sample is doubled, which should be possible within the next five years, then the upper limits will be improved by more than a factor of two, assuming the number of photon candidate events scales with the total statistics. In a more optimistic scenario, when it is assumed that the overall performance of the analysis in separating photon-induced air showers from the hadronic background can be optimized, the improvement of the upper limits is more than a factor of three above 1 EeV.

Future photon searches at the Pierre Auger Observatory in general will profit vastly from the planned upgrades of the detector systems. A better measurement of the muon content of the extensive air showers, which will be provided by the additional scintillators on top of the SD stations, combined with the larger statistics of the SD compared to hybrid measurements, will give a significant boost to the sensitivity of future photon searches. A combined analysis using X_{max} and the muon number as observables should give the best possible performance in separating photon-induced air showers from the hadronic background. Already now, new algorithms have been implemented in the trigger system of the SD in order to increase the

E_0 [EeV]	$f_\gamma^{95}(E > E_0)$ [%]	$\Phi_\gamma^{95}(E > E_0)$ [$\mathrm{km^{-2}\,sr^{-1}\,yr^{-1}}$]
1	0.39	0.101
2	0.48	0.027
3	0.85	0.021
4	1.35	0.019
5	1.85	0.017

Table 8.1.: Upper limits f_γ^{95} and Φ_γ^{95} on the photon fraction and the integral photon flux, respectively, which were derived in this work for different energy thresholds E_0. The upper limits have been determined at a confidence level of 95 %.

sensitivity to UHE photons. Overall, it is expected that in the near future, the upper limits on the integral flux of UHE photons can be improved by up to one order of magnitude, assuming no photon-induced air shower is unambiguously identified. The increase in sensitivity puts the expected range of photon fluxes from more optimistic scenarios involving Greisen, Zatsepin, Kuz'min (GZK) processes in reach. Through the combination of the results from the search for UHE photons with those from complementary searches for UHE neutrinos in a multi-messenger approach, it will be possible to provide strong constraints on current models for the origin of UHECRs.

So far, the photon window to the Universe at EeV energies remains closed. However, it is just a matter of time until the existence of UHE photons is either unambiguously confirmed or excluded, and the Pierre Auger Observatory with its unique combination of powerful detector systems will play a large role in this undertaking. And even if the search for UHE photons remains unsuccessful, and the window is sealed shut, the knowledge gained from these searches will be crucial to provide answers to the most interesting questions about the nature and the origin of the highest energy particles observed so far.

<div style="background-color: #b8bcdf; padding: 2em; text-align: right;">

Appendices

</div>

Contents

A. Input Files for the Simulations

A.1. CORSIKA Input File

In this section, an example of a CORSIKA input file used for the air shower simulations that are discussed in this thesis is given. The main steering parameters are discussed in Sec. 5.1.1. For a more detailed explanation of all steering parameters, see [Hec13].

```
1   RUNNR  1
2   EVTNR  1
3   NSHOW  1
4
5   SEED     86738337 0   0
6   SEED     340023802  0   0
7   SEED     887696731  0   0
8   SEED     239309143  0   0
9   SEED     215086971  0   0
10
11  PRMPAR  1
12  ESLOPE  −1
13  ERANGE  1.0e9  1.258925412e9
14  THETAP  0.0 65.0
15  PHIP    −180 180
16
17  OBSLEV  145200
18  MAGNET  20.1  −14.2
19  ATMOD   1
20  ARRANG  0.
21
```

```
22 QGSJET   T   0
23 QGSSIG   T
24 HADFLG   0   0   0   0   0   2
25 ELMFLG   T   T
26 STEPFC   1.0
27 RADNKG   200.E2
28 LONGI    T   5.   T   T
29 MUMULT   T
30
31 THIN     1e−6 1000 1.5e4
32 THINH    1.00 1.0e2
33 ECUTS    0.1   0.05 2.5e−4 2.5e−4
34
35 MUADDI   T
36 FLUDBG   T
37 MAXPRT   1
38 ECTMAP   10000
39 PAROUT   T   T
40 DATBAS   T
41 DEBUG    F   6   F   1000000
42
43 DIRECT   /home/niechciol/AirShowerSimulations/Fgamma/Photon/180−181/job1/
44 DATDIR   /home/niechciol/auger/corsika−74000/run/
45 USER     niechciol
46 HOST     slc1
47
48 EXIT
```

A.2. Off̄l̲i̲n̲e Input File

In this section, the XML file used for steering the simulation and reconstruction of air shower events with Off̄l̲i̲n̲e is listed. In the steering file, the sequence of modules which are used during the run time is specified. The main modules and parameters are discussed briefly in Sec. 5.1.2. For a more detailed description of the modules and their implementation, see [Pie14a].

```
1  <sequenceFile>
2
3    <moduleControl>
4
5      <loop numTimes="1" pushEventToStack="yes">
6
7        <module> EventFileReaderOG </module>
8        <module> MCShowerCheckerOG </module>
9
10       <loop numTimes="5" pushEventToStack="yes">
11
12         <module> EventGeneratorOG </module>
13
14         <!−− SD simulation part −−>
15         <loop numTimes="unbounded" pushEventToStack="no">
16           <module> CachedShowerRegeneratorOG </module>
17           <module> G4TankSimulatorOG </module>
18         </loop>
19
20         <try>  <!−− catch triggerless events for RecData∗ −−>
21
22           <module> SdSimulationCalibrationFillerOG </module>
23           <module> SdPMTSimulatorOG </module>
24           <module> SdFilterFADCSimulatorMTU </module>
25           <module> SdBaselineSimulatorOG </module>
26           <module> TankTriggerSimulatorOG </module>
27           <module> TankGPSSimulatorOG </module>
```

```
28
29              <!-- FD simulation part -->
30              <module> FdSimEventCheckerOG </module>
31              <module> ShowerLightSimulatorKG </module>
32              <module> LightAtDiaphragmSimulatorKG </module>
33              <module> ShowerPhotonGeneratorOG </module>
34              <module> TelescopeSimulatorKG </module>
35              <module> FdBackgroundSimulatorOG </module>
36              <module> FdElectronicsSimulatorOG </module>
37              <module> FdTriggerSimulatorOG </module>
38
39              <!-- Trigger and Event builder -->
40              <module> CentralTriggerSimulatorXb </module>
41              <module> CentralTriggerEventBuilderOG </module>
42              <module> EventBuilderOG </module>
43
44              <!-- export simulation in Offline format -->
45              <module> EventFileExporterOG </module>
46
47              <!-- Reconstruction -->
48              <module> EventCheckerOG </module>
49
50              <try> <!-- run RecData*, even if checker, calibrator send Continue -->
51
52                <module> SdCalibratorOG </module>
53
54                <!-- Hybrid reconstruction -->
55                <try>  <!-- limit how far a Continue goes -->
56                  <module> FdCalibratorOG </module>
57                  <module> FdPulseFinderOG </module>
58                  <module> PixelSelectorOG </module>
59                  <module> FdSDPFinderOG </module>
60                  <module> FdAxisFinderOG </module>
61                  <module> HybridGeometryFinderOG </module>
62                  <module> FdApertureLightOG </module>
63                  <module> FdProfileReconstructorKG </module>
64                </try>
65
66                <!-- SD reconstruction -->
67                <try>   <!-- limit how far a Continue goes -->
68                  <module> SdEventSelectorOG </module>
69                  <module> SdMonteCarloEventSelectorOG </module>
70                  <module> SdPlaneFitOG </module>
71                  <module> LDFFinderKG </module>
72                  <module> Risetime1000LLL </module>
73                  <module> SdEventPosteriorSelectorOG </module>
74                </try>
75
76              </try> <!-- catch calibrator Continues -->
77            </try> <!-- catch trigger Continues -->
78
79            <!-- export the ADST -->
80            <module> RecDataWriterNG </module>
81
82          </loop>
83        </loop>
84
85    </moduleControl>
86
87 </sequenceFile>
```

B. Detailed Statistics for the Event Selection

Top half

	\(18.0 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.1\)				\(18.1 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.2\)				\(18.2 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.3\)				\(18.3 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.4\)				\(18.4 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.5\)			
	Photon		Proton		Photon		Proton		Photon		Proton		Photon		Proton		Photon		Proton	
	N	ε	N	ε	N	ε	N	ε	N	ε	N	ε	N	ε	N	ε	N	ε	N	ε
Raw samples	5412	-	6229	-	5660	-	6721	-	6093	-	7228	-	6826	-	7757	-	7578	-	8688	-
FDReconstructionLevel eHasEnergy	4560	84.3%	5660	90.9%	4794	84.7%	6219	92.5%	4962	81.4%	6657	92.1%	5471	80.1%	7205	92.9%	6390	84.3%	8208	94.5%
TriggeredSDstations > 0	3546	77.8%	5387	95.2%	3932	82.0%	6095	98.0%	4329	87.2%	6598	99.1%	5034	92.0%	7164	99.4%	6077	95.1%	8191	99.8%
Total (preselection level)	3546	65.5%	5387	86.5%	3932	69.5%	6095	90.7%	4329	71.0%	6598	91.3%	5034	73.7%	7164	92.4%	6077	80.2%	8191	94.3%
NHybridStations > 0	3544	99.9%	5387	100.0%	3929	99.9%	6095	100.0%	4324	99.9%	6596	100.0%	5032	100.0%	7160	99.9%	6072	99.9%	8191	100.0%
HybridStationAxisDistance ≤ 1500 m	3542	99.9%	5380	99.9%	3924	99.9%	6087	99.9%	4315	99.8%	6592	99.9%	5022	99.8%	7152	99.9%	6069	100.0%	8186	99.9%
SD/FD offset ≤ 200 ns	3542	100.0%	5380	100.0%	3924	100.0%	6087	100.0%	4315	100.0%	6592	100.0%	5022	100.0%	7150	100.0%	6069	100.0%	8185	100.0%
SDP fit χ²/NDF ≤ 7	3541	100.0%	5374	99.9%	3916	99.8%	6080	99.9%	4310	99.9%	6579	99.8%	5015	99.9%	7142	99.9%	6065	99.9%	8173	99.9%
Time fit χ²/NDF ≤ 8	3533	99.8%	5367	99.9%	3903	99.7%	6068	99.8%	4293	99.6%	6556	99.7%	4997	99.6%	7112	99.6%	6046	99.7%	8142	99.6%
ZenithAngle	3361	95.1%	5102	95.1%	3692	94.6%	5788	95.4%	4000	93.2%	6058	92.4%	4607	92.2%	6517	91.6%	5555	91.9%	7312	89.8%
Total (geometry level)	3361	94.8%	5102	94.7%	3692	93.9%	5788	95.0%	4000	92.4%	6058	91.8%	4607	91.5%	6517	91.0%	5555	91.4%	7312	89.3%
CherenkovFraction ≤ 50%	3304	98.3%	5008	98.2%	3657	99.1%	5691	98.3%	3972	99.3%	5975	98.6%	4582	99.5%	6454	99.0%	5506	99.1%	7251	99.2%
maxDepthHole ≤ 20%	3249	98.3%	4959	99.0%	3608	98.7%	5624	98.8%	3919	98.7%	5899	98.7%	4518	98.6%	6396	99.1%	5437	98.7%	7182	99.1%
GH fit χ²/NDF ≤ 1.9	3234	99.5%	4952	99.9%	3593	99.6%	5608	99.7%	3905	99.6%	5883	99.7%	4501	99.6%	6376	99.7%	5409	99.5%	7162	99.7%
XmaxInFoV true	2621	81.0%	4200	84.8%	2939	81.8%	4954	88.3%	3286	84.1%	5207	88.5%	3714	82.5%	5725	89.8%	4458	82.4%	6588	92.0%
FiducialVolumeCut true	2283	87.1%	3069	73.1%	2470	84.0%	3533	71.3%	2682	81.6%	3431	65.9%	2696	72.6%	3142	54.9%	2936	65.9%	3180	48.3%
RelCalEnergyUncertainty ≤ 20%	2172	95.1%	2919	95.1%	2375	96.2%	3415	96.7%	2602	97.0%	3342	97.4%	2629	97.5%	3087	98.2%	2842	96.8%	3135	98.6%
Total (profile level)	2172	64.6%	2919	57.2%	2375	64.3%	3415	59.0%	2602	65.1%	3342	55.2%	2629	57.1%	3087	47.4%	2842	51.2%	3135	42.9%
Total (all levels)	2172	40.1%	2919	46.9%	2375	42.0%	3415	50.8%	2602	42.7%	3342	46.2%	2629	38.5%	3087	39.8%	2842	37.5%	3135	36.1%

Bottom half

	\(18.5 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.6\)				\(18.6 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.7\)				\(18.7 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.8\)				\(18.8 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.9\)				\(18.9 \leq \log_{10}(E\,[\mathrm{eV}]) < 19.0\)			
	Photon		Proton		Photon		Proton		Photon		Proton		Photon		Proton		Photon		Proton	
	N	ε	N	ε	N	ε	N	ε	N	ε	N	ε	N	ε	N	ε	N	ε	N	ε
Raw samples	8160	-	9392	-	8712	-	9750	-	9064	-	9921	-	9427	-	9964	-	9551	-	9987	-
FDReconstructionLevel eHasEnergy	7193	88.1%	9073	96.6%	7883	90.5%	9492	97.4%	8326	91.9%	9752	98.3%	8920	94.6%	9880	99.2%	9136	95.7%	9906	99.2%
TriggeredSDstations > 0	6966	96.8%	9064	99.9%	7733	98.1%	9485	99.9%	8232	98.9%	9746	99.9%	8858	99.3%	9877	100.0%	9102	99.6%	9902	100.0%
Total (preselection level)	6966	85.4%	9064	96.5%	7733	88.8%	9485	97.3%	8232	90.8%	9746	98.2%	8858	94.0%	9877	99.1%	9102	95.3%	9902	99.1%
NHybridStations > 0	6965	100.0%	9060	100.0%	7732	100.0%	9483	100.0%	8226	99.9%	9743	100.0%	8853	99.9%	9877	100.0%	9096	99.9%	9899	100.0%
HybridStationAxisDistance ≤ 1500 m	6959	99.9%	9054	99.9%	7728	99.9%	9483	100.0%	8226	100.0%	9743	100.0%	8849	100.0%	9875	100.0%	9094	100.0%	9898	100.0%
SD/FD offset ≤ 200 ns	6957	100.0%	9054	100.0%	7728	100.0%	9483	100.0%	8225	100.0%	9743	100.0%	8845	100.0%	9872	100.0%	9091	100.0%	9895	100.0%
SDP fit χ²/NDF ≤ 7	6953	99.9%	9039	99.8%	7719	99.9%	9469	99.9%	8215	99.9%	9732	99.9%	8839	99.9%	9859	99.9%	9080	99.9%	9884	99.9%
Time fit χ²/NDF ≤ 8	6929	99.7%	9017	99.8%	7690	99.6%	9444	99.7%	8191	99.7%	9722	99.9%	8813	99.7%	9854	99.9%	9070	99.9%	9881	100.0%
ZenithAngle ≤ 60°	6332	91.4%	8116	90.0%	6976	90.7%	8612	91.2%	7487	91.4%	8827	90.8%	8077	91.6%	9030	91.6%	8252	91.0%	8984	90.9%
Total (geometry level)	6332	90.9%	8116	89.5%	6976	90.2%	8612	90.8%	7487	90.9%	8827	90.6%	8077	91.2%	9030	91.4%	8252	90.7%	8984	90.7%
CherenkovFraction ≤ 50%	6315	99.7%	8070	99.4%	6953	99.7%	8572	99.5%	7465	99.7%	8802	99.7%	8055	99.7%	9001	99.7%	8230	99.7%	8950	99.6%
maxDepthHole ≤ 20%	6245	98.9%	8007	99.2%	6864	98.7%	8504	99.2%	7378	98.8%	8735	99.2%	7945	98.6%	8930	99.2%	8132	98.8%	8872	99.1%
GH fit χ²/NDF ≤ 1.9	6221	99.6%	7986	99.7%	6821	99.4%	8479	99.7%	7334	99.4%	8701	99.6%	7902	99.5%	8897	99.6%	8082	99.4%	8833	99.6%
XmaxInFoV true	4946	79.5%	7373	92.3%	5153	75.5%	7848	92.6%	5164	70.4%	7988	91.8%	5293	67.0%	8152	91.6%	5167	63.9%	8057	91.2%
FiducialVolumeCut true	3265	66.0%	3444	46.7%	3556	69.0%	3795	48.4%	3675	71.2%	3911	49.0%	4080	77.1%	4139	50.8%	4236	82.0%	4407	54.7%
RelCalEnergyUncertainty ≤ 20%	3185	97.5%	3425	99.4%	3509	98.7%	3784	99.7%	3637	99.0%	3906	99.9%	4047	99.2%	4138	100.0%	4224	99.7%	4406	100.0%
Total (profile level)	3185	50.3%	3425	42.2%	3509	50.3%	3784	43.9%	3637	48.6%	3906	44.3%	4047	50.1%	4138	45.8%	4224	51.2%	4406	49.0%
Total (all levels)	3185	39.0%	3425	36.5%	3509	40.3%	3784	38.8%	3637	40.1%	3906	39.4%	4047	42.9%	4138	41.5%	4224	44.2%	4406	44.1%

Table B.1.: Application of the event selection criteria to the main simulation samples described in Sec. 5.1. Listed are the number of surviving events *N* after applying the criterion and the corresponding selection efficiency *ε* relative to the number of events before the criterion was applied.

| | $18.0 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.1$ | | | | $18.1 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.2$ | | | | $18.2 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.3$ | | | | $18.3 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.4$ | | | | $18.4 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.5$ | | | |
| | Photon | | Proton | | Photon | | Proton | | Photon | | Proton | | Photon | | Proton | | Photon | | Proton | |
	N	ε	N	ε	N	ε	N	ε	N	ε	N	ε	N	ε	N	ε	N	ε	N	ε
After the event selection from Sec. 5.3	2172	-	2919	-	2375	-	3415	-	2602	-	3342	-	2629	-	3087	-	2842	-	3135	-
FitFailed false	2172	100,0%	2918	100,0%	2374	100,0%	3414	100,0%	2601	100,0%	3341	100,0%	2627	99,9%	3087	100,0%	2841	100,0%	3135	100,0%
DeltaFgamma < 0.35	2021	93,0%	1834	62,9%	2301	96,9%	2362	69,2%	2577	99,1%	2619	78,4%	2620	99,7%	2765	89,6%	2838	99,9%	2965	94,6%
RelFgammaUncertainty < 0.7	1740	86,1%	1798	98,0%	1977	85,9%	2322	98,3%	2287	88,7%	2601	99,3%	2409	91,9%	2760	99,8%	2712	95,6%	2959	99,8%
ChiSquare < 30	1711	98,3%	1700	94,5%	1937	98,0%	2225	95,8%	2243	98,1%	2516	96,7%	2355	97,8%	2678	97,0%	2641	97,4%	2870	97,0%
Total	1711	78,8%	1700	58,2%	1937	81,6%	2225	65,2%	2243	86,2%	2516	75,3%	2355	89,6%	2678	86,8%	2641	92,9%	2870	91,5%

| | $18.5 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.6$ | | | | $18.6 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.7$ | | | | $18.7 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.8$ | | | | $18.8 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.9$ | | | | $18.9 \leq \log_{10}(E\,[\mathrm{eV}]) < 19.0$ | | | |
| | Photon | | Proton | | Photon | | Proton | | Photon | | Proton | | Photon | | Proton | | Photon | | Proton | |
	N	ε	N	ε	N	ε	N	ε	N	ε	N	ε	N	ε	N	ε	N	ε	N	ε
After the event selection from Sec. 5.3	3185	-	3425	-	3509	-	3784	-	3637	-	3906	-	4047	-	4138	-	4224	-	4406	-
FitFailed false	3185	100,0%	3425	100,0%	3509	100,0%	3783	100,0%	3637	100,0%	3906	100,0%	4047	100,0%	4138	100,0%	4224	100,0%	4406	100,0%
DeltaFgamma < 0.35	3182	99,9%	3301	96,4%	3507	99,9%	3690	97,5%	3636	100,0%	3851	98,6%	4047	100,0%	4108	99,3%	4224	100,0%	4383	99,5%
RelFgammaUncertainty < 0.7	3112	97,8%	3297	99,9%	3480	99,2%	3690	100,0%	3611	99,3%	3851	100,0%	4037	99,8%	4108	100,0%	4221	99,9%	4383	100,0%
ChiSquare < 30	3024	97,2%	3190	96,8%	3401	97,7%	3577	96,9%	3508	97,1%	3741	97,1%	3927	97,3%	3986	97,0%	4103	97,2%	4261	97,2%
Total	3024	94,9%	3190	93,1%	3401	96,9%	3577	94,5%	3508	96,5%	3741	95,8%	3927	97,0%	3986	96,3%	4103	97,1%	4261	96,7%

Table B.2.: Application of the additional event selection criteria based on F_γ to the main simulation samples. Listed are the number of surviving events N after applying the criterion and the corresponding selection efficiency ε relative to the number of events before the criterion was applied.

Top table

	$19.0 \leq \log_{10}(E\,[\text{eV}]) < 19.1$				$19.5 \leq \log_{10}(E\,[\text{eV}]) < 19.6$				$18.0 \leq \log_{10}(E\,[\text{eV}]) < 18.1$				$18.5 \leq \log_{10}(E\,[\text{eV}]) < 18.6$				$18.9 \leq \log_{10}(E\,[\text{eV}]) < 19.0$			
	Photon		Proton		Photon		Proton		Proton (EPOS)		Proton (SIBYLL)		Proton (EPOS)		Proton (SIBYLL)		Proton (EPOS)		Proton (SIBYLL)	
	N	ε	N	ε	N	ε	N	ε	N	ε	N	ε	N	ε	N	ε	N	ε	N	ε
Raw samples	4804	-	4998	-	4866	-	4997	-	3169	-	2945	-	4661	-	4737	-	4997	-	5000	-
FDReconstructionLevel eHasEnergy	4620	96.2%	4970	99.4%	4688	96.3%	4876	97.6%	2941	92.8%	2640	89.6%	4479	96.1%	4562	96.3%	4951	99.1%	4959	99.2%
TriggeredSDstations > 0	4604	99.7%	4969	100.0%	4674	99.7%	4867	99.8%	2844	96.7%	2437	92.3%	4476	99.9%	4550	99.7%	4949	100.0%	4956	99.9%
Total (preselection level)	4604	95.8%	4969	99.4%	4674	96.1%	4867	97.4%	2844	89.7%	2437	82.8%	4476	96.0%	4550	96.1%	4949	99.0%	4956	99.1%
NHybridStations > 0	4597	99.8%	4966	99.9%	4608	98.6%	4825	99.1%	2843	100.0%	2437	100.0%	4473	99.9%	4550	100.0%	4948	100.0%	4953	99.9%
HybridStationAxisDistance ≤ 1500 m	4596	100.0%	4962	99.9%	4591	99.6%	4784	99.2%	2843	100.0%	2437	100.0%	4472	100.0%	4549	100.0%	4946	100.0%	4952	100.0%
SD/FD offset ≤ 200 ns	4594	100.0%	4960	100.0%	4561	99.3%	4761	99.5%	2843	100.0%	2437	100.0%	4472	100.0%	4548	100.0%	4943	99.9%	4951	100.0%
SDP fit χ^2/NDF ≤ 7	4590	99.9%	4954	99.9%	4544	99.6%	4721	99.2%	2835	99.7%	2427	99.6%	4466	99.8%	4546	99.8%	4937	99.9%	4946	99.9%
Time fit χ^2/NDF ≤ 8	4585	99.9%	4952	100.0%	4540	99.9%	4721	100.0%	2831	99.9%	2427	100.0%	4455	99.8%	4537	99.8%	4937	100.0%	4945	100.0%
ZenithAngle ≤ 60°	4116	89.8%	4533	91.5%	4089	90.1%	4365	92.5%	2683	94.8%	2302	94.8%	4080	91.6%	4162	91.7%	4492	91.0%	4517	91.3%
Total (geometry level)	4116	89.4%	4533	91.2%	4089	87.5%	4365	89.7%	2683	94.3%	2302	94.5%	4080	91.2%	4162	91.5%	4492	90.8%	4517	91.1%
CherenkovFraction ≤ 50%	4108	99.8%	4521	99.7%	4080	99.8%	4349	99.6%	2648	98.7%	2259	98.1%	4054	99.4%	4140	99.5%	4475	99.6%	4507	99.8%
maxDepthHole ≤ 20%	4053	98.7%	4462	98.7%	3936	96.5%	4220	97.0%	2616	98.8%	2235	98.9%	4020	99.2%	4094	98.9%	4436	99.1%	4467	99.1%
GH fit χ^2/NDF ≤ 1.9	4021	99.2%	4440	99.5%	3723	94.6%	3993	94.6%	2611	99.8%	2230	99.8%	4015	99.9%	4085	99.8%	4412	99.5%	4452	99.7%
XmaxInFoV true	2343	58.3%	3993	89.9%	1402	37.7%	3388	84.8%	2215	84.8%	1886	84.6%	3730	92.9%	3784	92.6%	3920	88.8%	4013	90.1%
FiducialVolumeCut true	1982	84.6%	1949	48.8%	1268	90.4%	1638	48.3%	1643	74.2%	1395	74.0%	1764	47.3%	1824	48.2%	2009	51.3%	2110	52.6%
RelCalEnergyUncertainty ≤ 20%	1974	99.6%	1948	99.9%	1268	100.0%	1638	100.0%	1573	95.7%	1349	96.7%	1755	99.5%	1816	99.6%	2009	100.0%	2110	100.0%
Total (profile level)	1974	48.0%	1948	43.0%	1268	31.0%	1638	37.5%	1573	58.6%	1349	58.6%	1755	43.0%	1816	43.6%	2009	44.7%	2110	46.7%
Total (all levels)	1974	41.1%	1948	39.0%	1268	26.1%	1638	32.8%	1573	49.6%	1349	45.8%	1755	37.7%	1816	38.3%	2009	40.2%	2110	42.2%

Bottom table

	$18.0 \leq \log_{10}(E\,[\text{eV}]) < 18.1$						$18.5 \leq \log_{10}(E\,[\text{eV}]) < 18.6$						$18.9 \leq \log_{10}(E\,[\text{eV}]) < 19.0$					
	Helium		Oxygen		Iron		Helium		Oxygen		Iron		Helium		Oxygen		Iron	
	N	ε	N	ε	N	ε	N	ε	N	ε	N	ε	N	ε	N	ε	N	ε
Raw samples	3106	-	3224	-	3136	-	4767	-	4802	-	4828	-	4999	-	5000	-	5000	-
FDReconstructionLevel eHasEnergy	2836	91.3%	2964	91.9%	2875	91.7%	4617	96.9%	4648	96.8%	4707	97.5%	4969	99.4%	4973	99.4%	4971	99.4%
TriggeredSDstations > 0	2760	97.3%	2872	96.9%	2816	97.9%	4612	99.9%	4644	99.9%	4706	100.0%	4966	99.9%	4971	100.0%	4969	100.0%
Total (preselection level)	2760	88.9%	2872	89.1%	2816	89.8%	4612	96.7%	4644	96.7%	4706	97.5%	4966	99.3%	4971	99.4%	4969	99.4%
NHybridStations > 0	2759	100.0%	2872	100.0%	2816	100.0%	4612	100.0%	4642	100.0%	4705	100.0%	4964	100.0%	4970	100.0%	4968	100.0%
HybridStationAxisDistance ≤ 1500 m	2756	99.8%	2866	99.8%	2814	99.9%	4609	100.0%	4641	100.0%	4705	100.0%	4964	100.0%	4969	100.0%	4966	100.0%
SD/FD offset ≤ 200 ns	2756	100.0%	2866	100.0%	2814	100.0%	4609	100.0%	4641	100.0%	4704	100.0%	4963	100.0%	4968	100.0%	4964	100.0%
SDP fit χ^2/NDF ≤ 7	2751	99.8%	2859	99.8%	2809	99.8%	4601	99.8%	4638	99.9%	4696	99.8%	4955	99.8%	4965	99.9%	4952	99.8%
Time fit χ^2/NDF ≤ 8	2746	99.8%	2853	99.8%	2807	99.9%	4586	99.7%	4627	99.8%	4686	99.8%	4955	100.0%	4965	100.0%	4952	100.0%
ZenithAngle ≤ 60°	2633	95.9%	2773	97.2%	2708	96.5%	4175	91.0%	4235	91.5%	4332	92.4%	4499	90.8%	4572	92.1%	4507	91.0%
Total (geometry level)	2633	95.4%	2773	96.6%	2708	96.2%	4175	90.5%	4235	91.2%	4332	92.1%	4499	90.6%	4572	92.0%	4507	90.7%
CherenkovFraction ≤ 50%	2573	97.7%	2715	97.9%	2649	97.8%	4141	99.2%	4204	99.3%	4294	99.1%	4486	99.7%	4557	99.7%	4488	99.6%
maxDepthHole ≤ 20%	2546	99.0%	2685	98.9%	2619	98.9%	4102	99.1%	4164	99.0%	4255	99.1%	4457	99.4%	4526	99.3%	4447	99.1%
GH fit χ^2/NDF ≤ 1.9	2537	99.6%	2683	99.9%	2619	100.0%	4088	99.7%	4158	99.9%	4246	99.8%	4441	99.6%	4507	99.6%	4423	99.5%
XmaxInFoV true	2178	85.8%	2256	84.1%	2156	82.3%	3830	93.7%	3917	94.2%	3984	93.8%	4154	93.5%	4336	96.2%	4186	94.6%
FiducialVolumeCut true	1675	76.9%	1664	73.8%	1606	74.5%	1718	44.9%	1638	41.8%	1591	39.9%	2021	48.7%	2159	49.8%	1932	46.2%
RelCalEnergyUncertainty ≤ 20%	1593	95.1%	1607	96.6%	1536	95.6%	1714	99.8%	1637	99.9%	1586	99.7%	2020	99.9%	2159	100.0%	1932	100.0%
Total (profile level)	1593	60.5%	1607	58.0%	1536	56.7%	1714	41.1%	1637	38.7%	1586	36.6%	2020	44.9%	2159	47.2%	1932	42.9%
Total (all levels)	1593	51.3%	1607	49.8%	1536	49.0%	1714	36.0%	1637	34.1%	1586	32.9%	2020	40.4%	2159	43.2%	1932	38.6%

Table B.3.: Application of the event selection criteria to the additional simulation samples described in Sec. 5.1. Listed are the number of surviving events N after applying the criterion and the corresponding selection efficiency ε relative to the number of events before the criterion was applied.

Table B.4.: Application of the additional event selection criteria based on F_γ to the additional simulation samples. Listed are the number of surviving events N after applying the criterion and the corresponding selection efficiency ε relative to the number of events before the criterion was applied.

| | $19.0 \leq \log_{10}(E\,[\mathrm{eV}]) < 19.1$ | | | | $19.5 \leq \log_{10}(E\,[\mathrm{eV}]) < 19.6$ | | | |
| | Photon | | Proton | | Photon | | Proton | |
	N	ε	N	ε	N	ε	N	ε
After the event selection from Sec. 5.3	1974	-	1948	-	1268	-	1638	-
FitFailed false	1974	100.0%	1948	100.0%	1268	100.0%	1638	100.0%
DeltaFgamma < 0.35	1974	100.0%	1939	99.5%	1268	100.0%	1638	100.0%
RelFgammaUncertainty < 0.7	1973	99.9%	1939	100.0%	1268	100.0%	1638	100.0%
ChiSquare < 30	1902	96.4%	1887	97.3%	1017	80.2%	1542	94.1%
Total	1902	96.4%	1887	96.9%	1017	80.2%	1542	94.1%

| | $18.0 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.1$ | | | | $18.5 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.6$ | | | | $18.9 \leq \log_{10}(E\,[\mathrm{eV}]) < 19.0$ | | | |
| | Proton (EPOS) | | Proton (SIBYLL) | | Proton (EPOS) | | Proton (SIBYLL) | | Proton (EPOS) | | Proton (SIBYLL) | |
	N	ε	N	ε	N	ε	N	ε	N	ε	N	ε
After the event selection from Sec. 5.3	1573	-	1349	-	1755	-	1816	-	2009	-	2110	-
FitFailed false	1573	100.0%	1349	100.0%	1755	100.0%	1816	100.0%	2009	100.0%	2110	100.0%
DeltaFgamma < 0.35	903	57.4%	978	72.5%	1627	92.7%	1797	99.0%	1991	99.1%	2109	100.0%
RelFgammaUncertainty < 0.7	884	97.9%	928	94.9%	1626	99.9%	1794	99.8%	1991	100.0%	2109	100.0%
ChiSquare < 30	831	94.0%	893	96.2%	1572	96.7%	1746	97.3%	1937	97.3%	2059	97.6%
Total	831	52.8%	893	66.2%	1572	89.6%	1746	96.1%	1937	96.4%	2059	97.6%

| | $18.0 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.1$ | | | | | | $18.5 \leq \log_{10}(E\,[\mathrm{eV}]) < 18.6$ | | | | | | $18.9 \leq \log_{10}(E\,[\mathrm{eV}]) < 19.0$ | | | | | |
| | Helium | | Oxygen | | Iron | | Helium | | Oxygen | | Iron | | Helium | | Oxygen | | Iron | |
	N	ε	N	ε	N	ε	N	ε	N	ε	N	ε	N	ε	N	ε	N	ε
After the event selection from Sec. 5.3	1593	-	1607	-	1536	-	1714	-	1637	-	1586	-	2020	-	2159	-	1932	-
FitFailed false	1588	99.7%	1605	99.9%	1533	99.8%	1714	100.0%	1636	99.9%	1586	100.0%	2020	100.0%	2159	100.0%	1932	100.0%
DeltaFgamma < 0.35	900	56.7%	822	51.2%	703	45.9%	1633	95.3%	1489	91.0%	1379	86.9%	1998	98.9%	2126	98.5%	1868	96.7%
RelFgammaUncertainty < 0.7	886	98.4%	813	98.9%	695	98.9%	1633	100.0%	1489	100.0%	1379	100.0%	1998	100.0%	2126	100.0%	1868	100.0%
ChiSquare < 30	809	91.3%	760	93.5%	617	88.8%	1576	96.5%	1424	95.6%	1324	96.0%	1945	97.3%	2058	96.8%	1818	97.3%
Total	809	50.8%	760	47.3%	617	40.2%	1576	91.9%	1424	87.0%	1324	83.5%	1945	96.3%	2058	95.3%	1818	94.1%

C. Additional Plots for Chapter 6
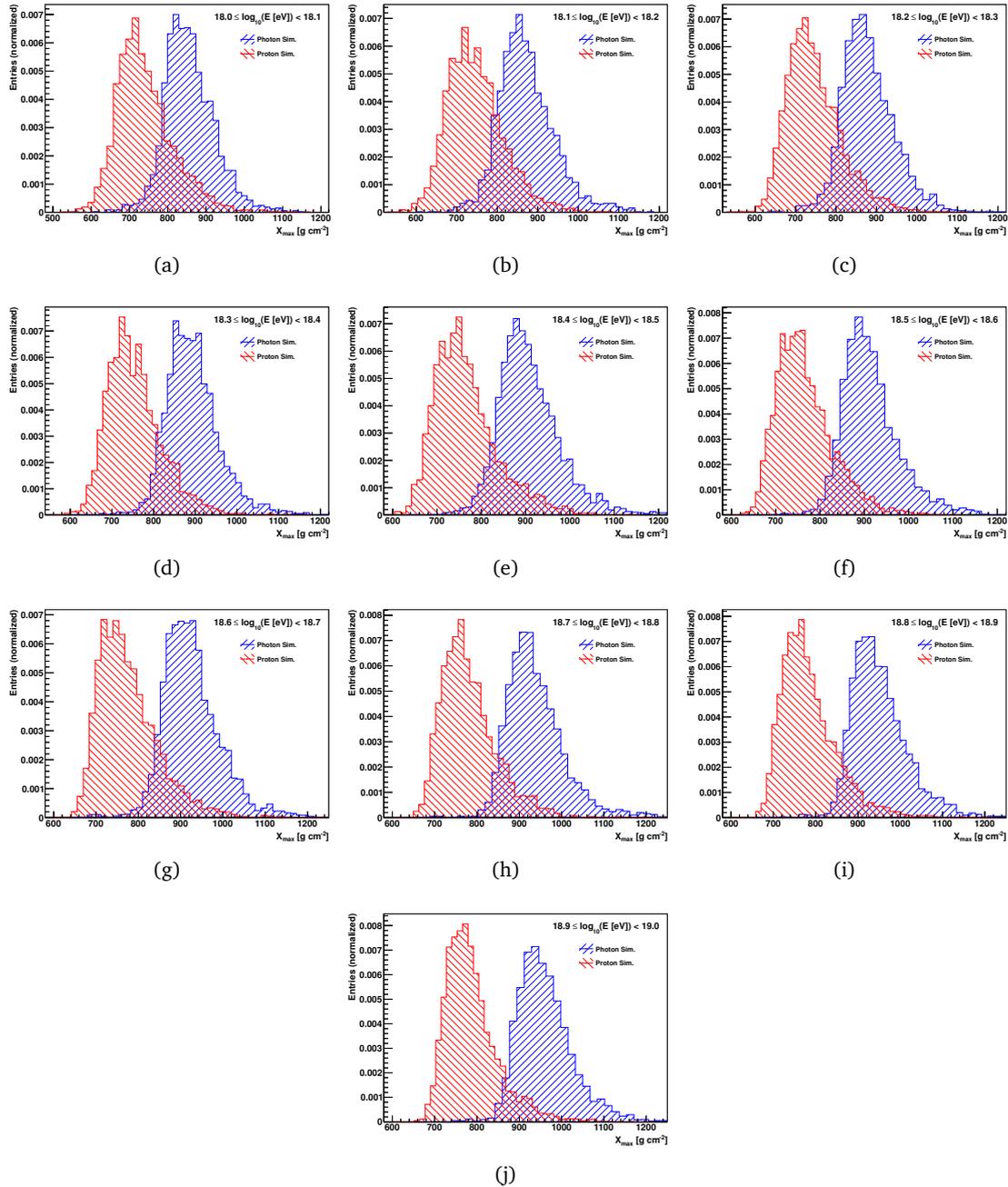
Figure C.1.: Distributions of the observable $X_{\max}$ in the individual energy bins for the photon sample, shown in blue, and the proton sample, shown in red.

Figure C.2.: Background rejection ϱ as a function of the signal efficiency ε for the observable X_{max} in the individual energy bins.
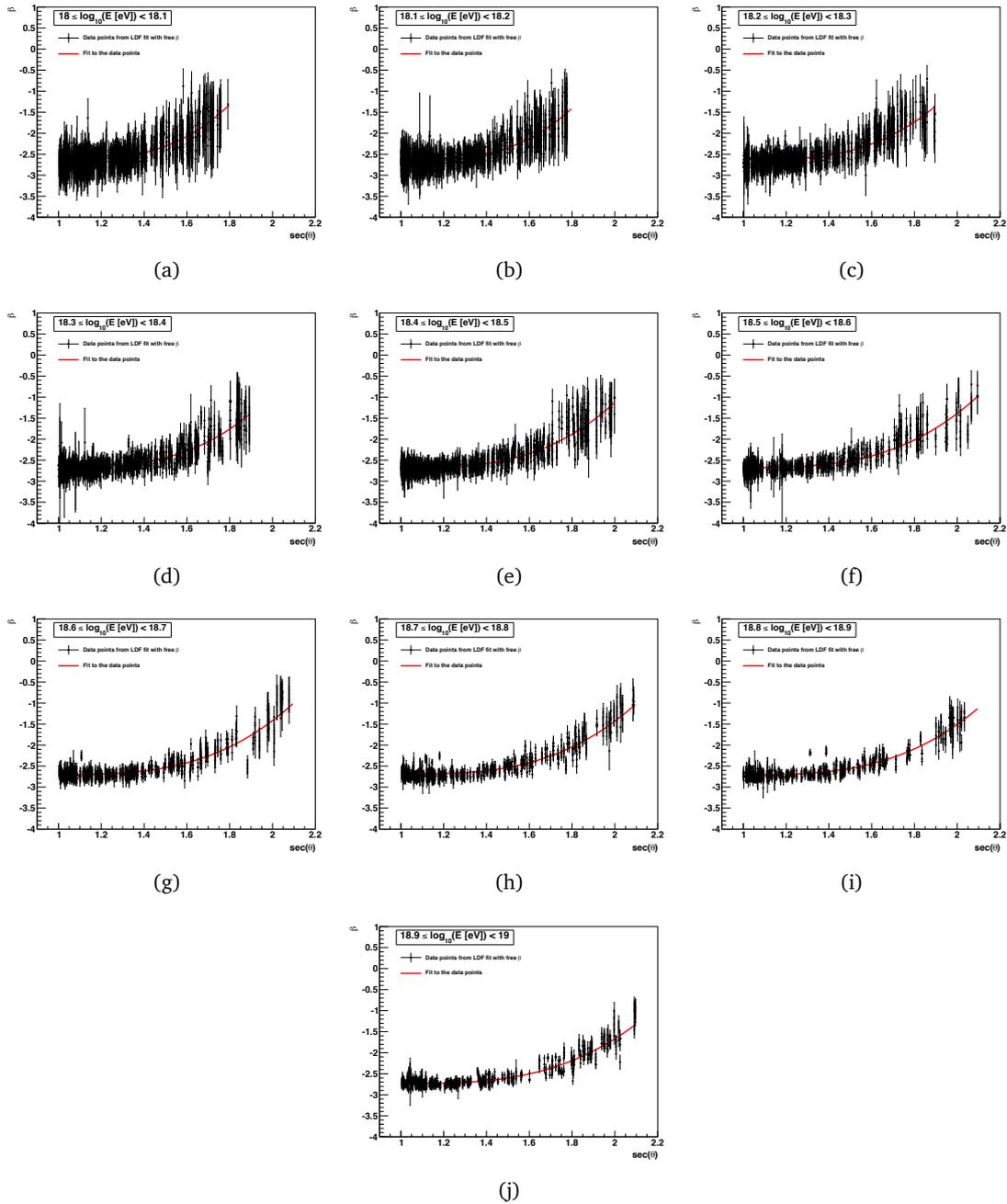
Figure C.3.: Distributions of the fitted β as a function of $\sec(\theta)$ in the different energy ranges. A simple cubic function according to Eq. 6.7 has been fitted to the data points in each plot, the results of the fits are shown as red lines.
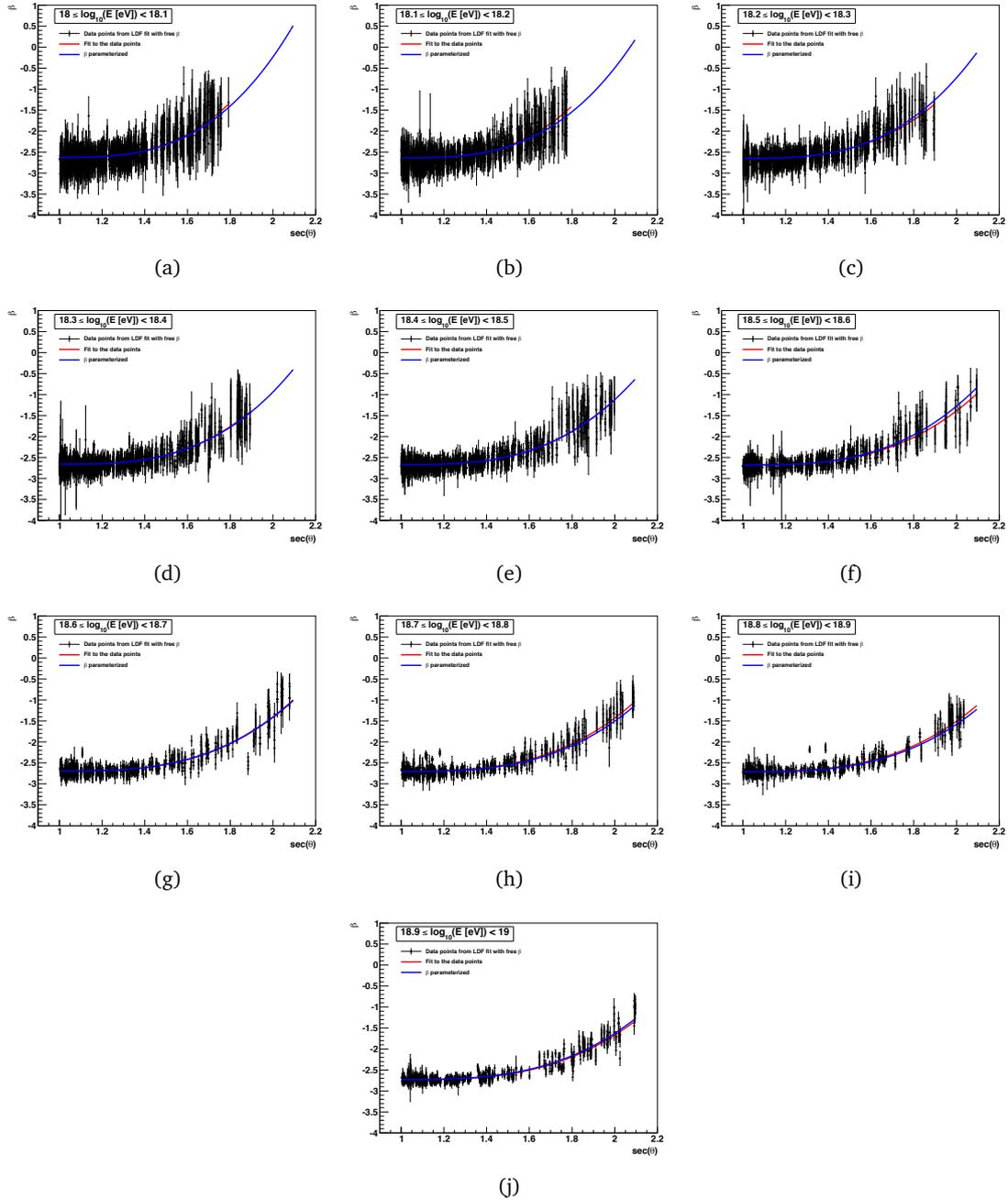
Figure C.4.: Distributions of the fitted β as a function of $\sec(\theta)$ in the different energy ranges. In addition to the fits of the cubic function according to Eq. 6.7, shown as red lines, the parameterized functions obtained from Eq. 6.8 are included here as blue lines.
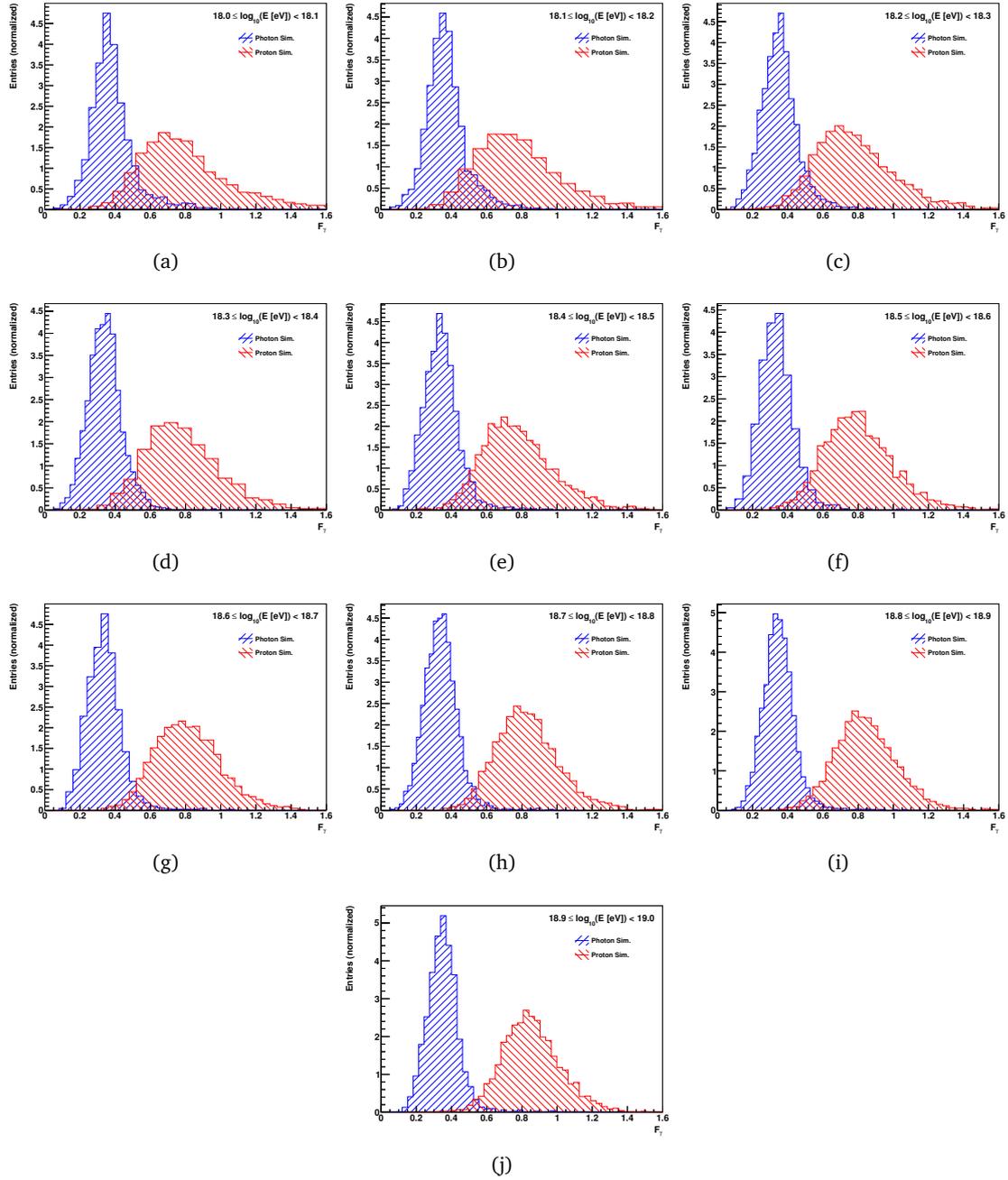
Figure C.5.: Distributions of the observable F_γ in the individual energy bins for the photon sample, shown in blue, and the proton sample, shown in red.
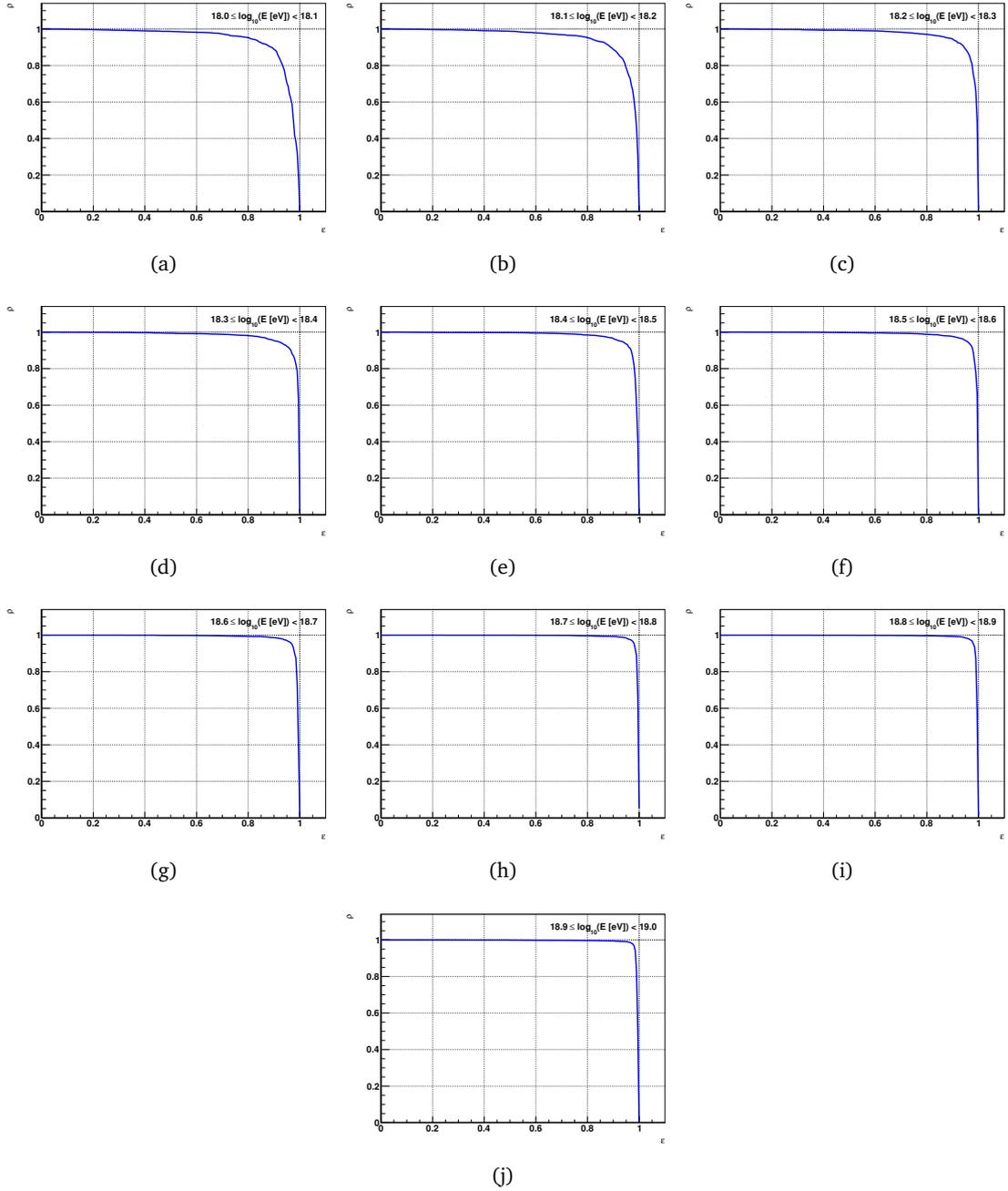
Figure C.6.: Background rejection ϱ as a function of the signal efficiency ε for the observable F_γ in the individual energy bins.
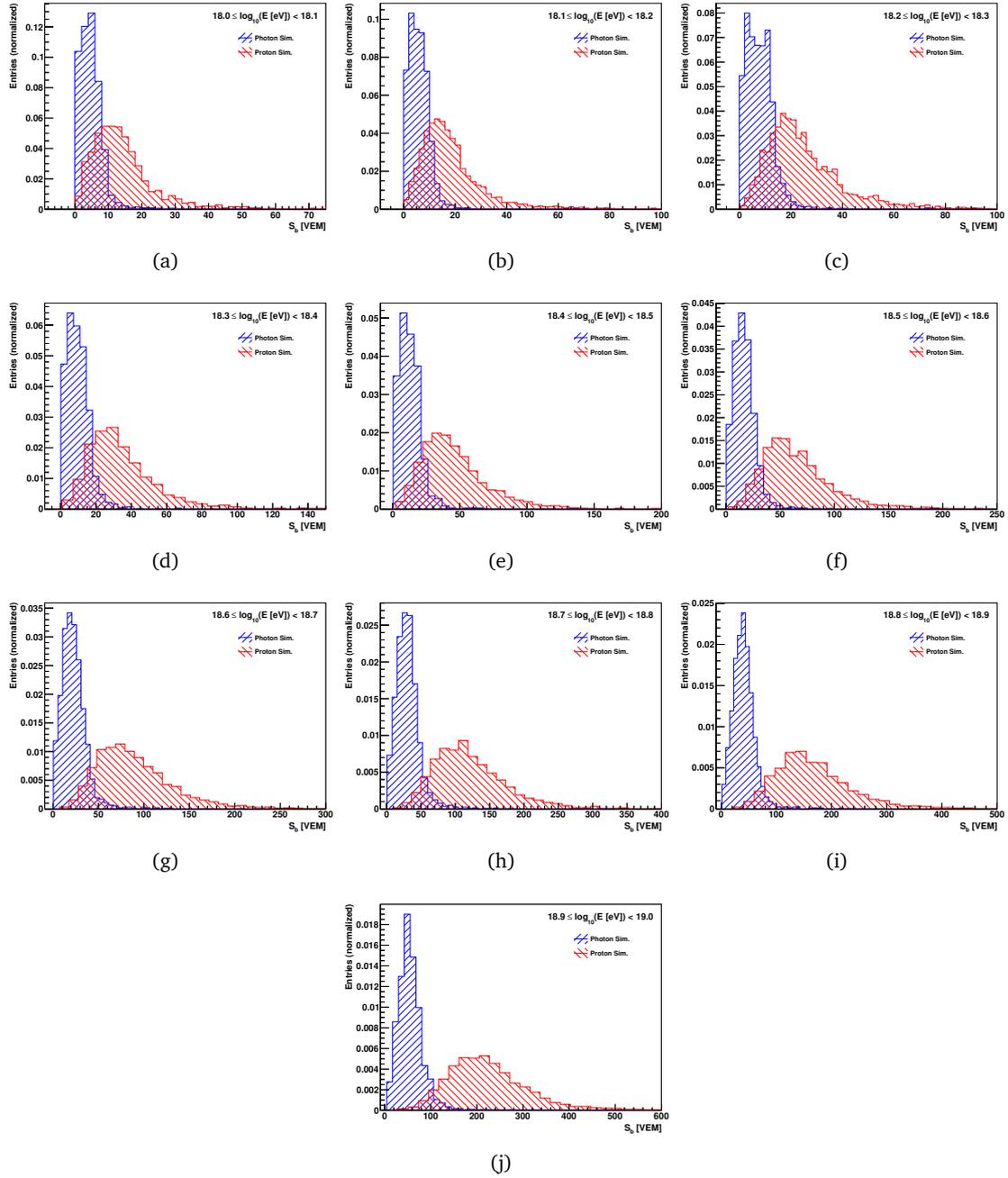
Figure C.7.: Distributions of the observable S_b in the individual energy bins for the photon sample, shown in blue, and the proton sample, shown in red.
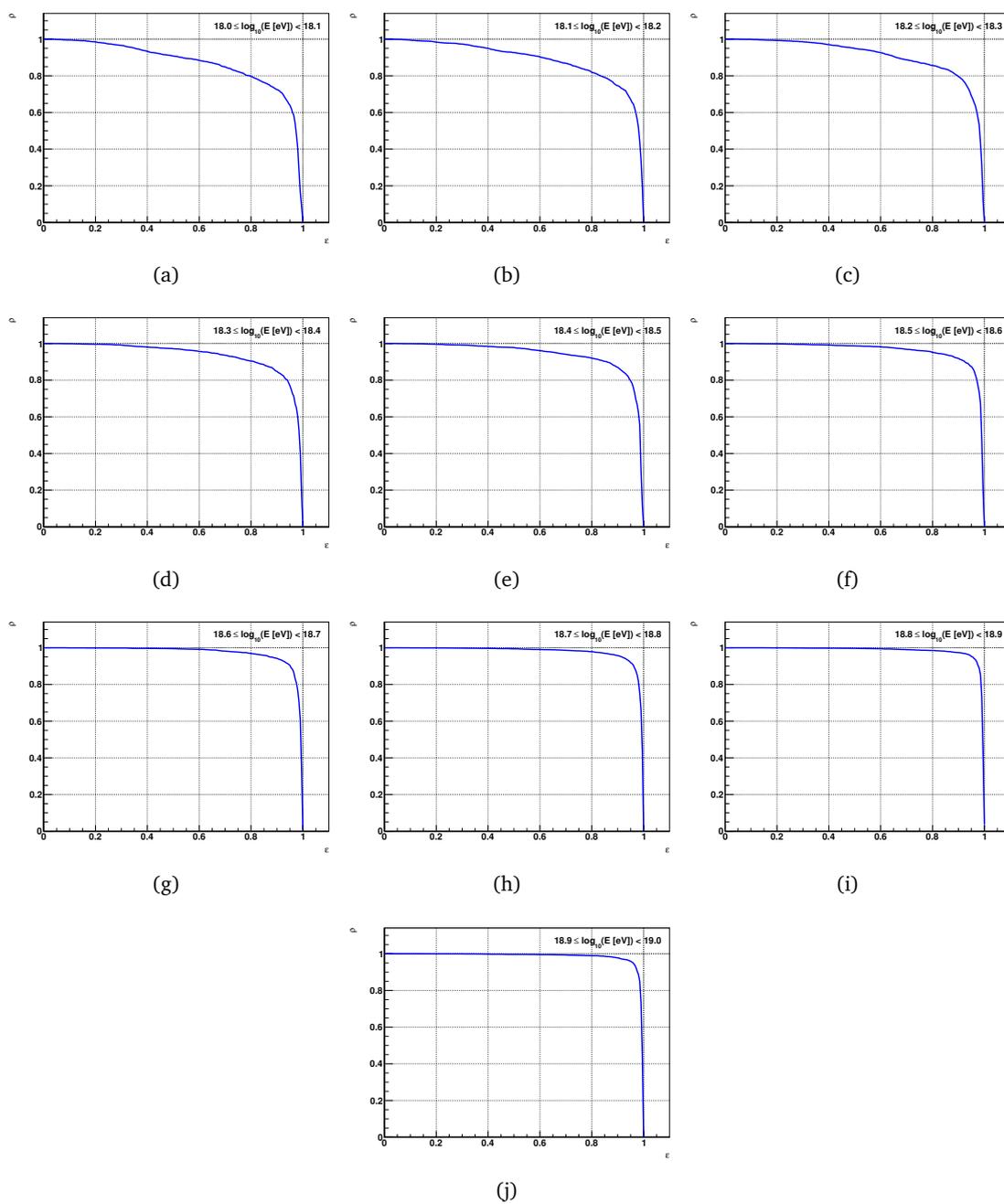
Figure C.8.: Background rejection ϱ as a function of the signal efficiency ε for the observable S_b in the individual energy bins.
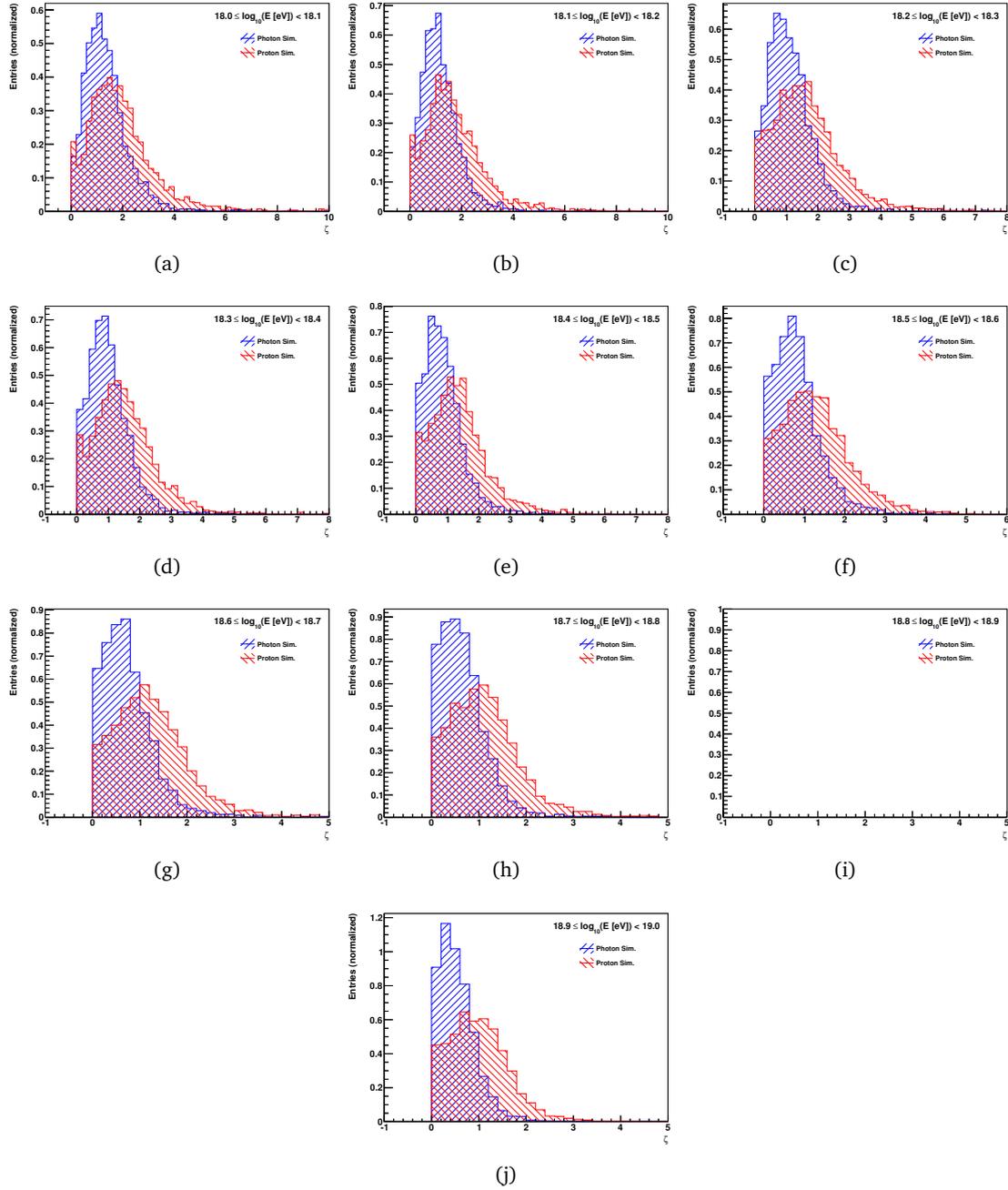
Figure C.9.: Distributions of the observable ζ in the individual energy bins for the photon sample, shown in blue, and the proton sample, shown in red.
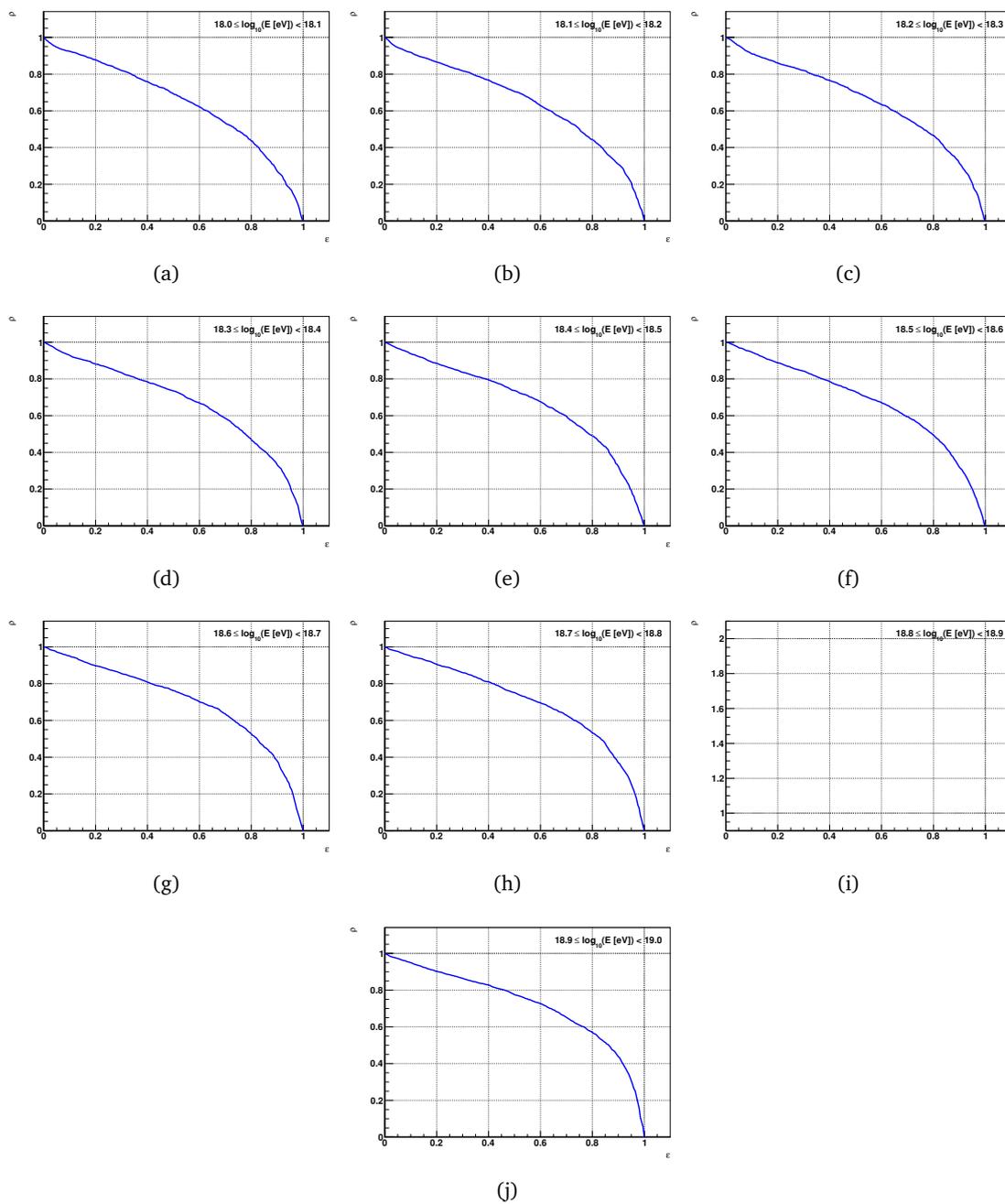
Figure C.10.: Background rejection ϱ as a function of the signal efficiency ε for the observable ζ in the individual energy bins.
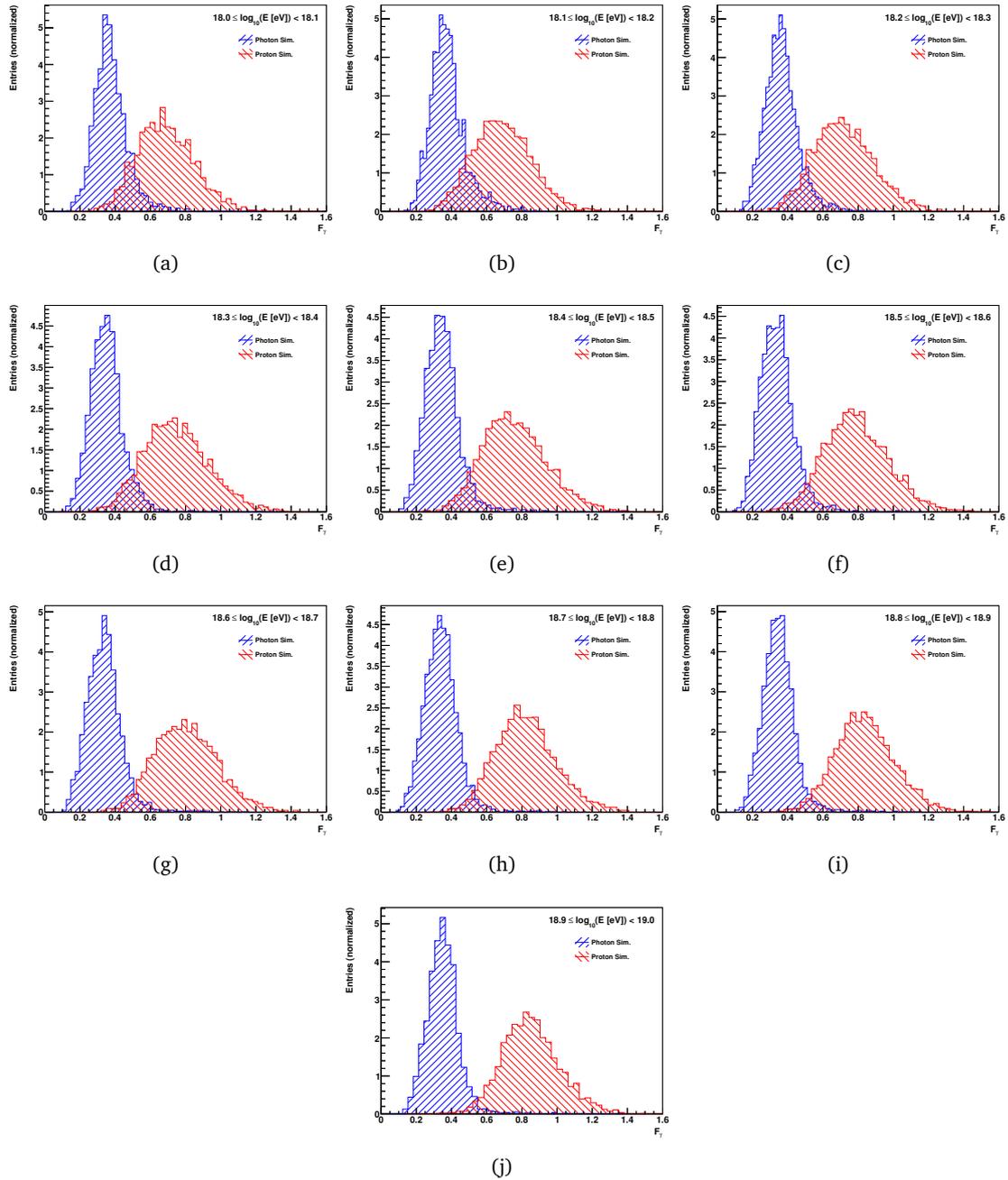
Figure C.11.: Distributions of the observable F_γ, including the additional selection criteria, in the individual energy bins for the photon sample, shown in blue, and the proton sample, shown in red.
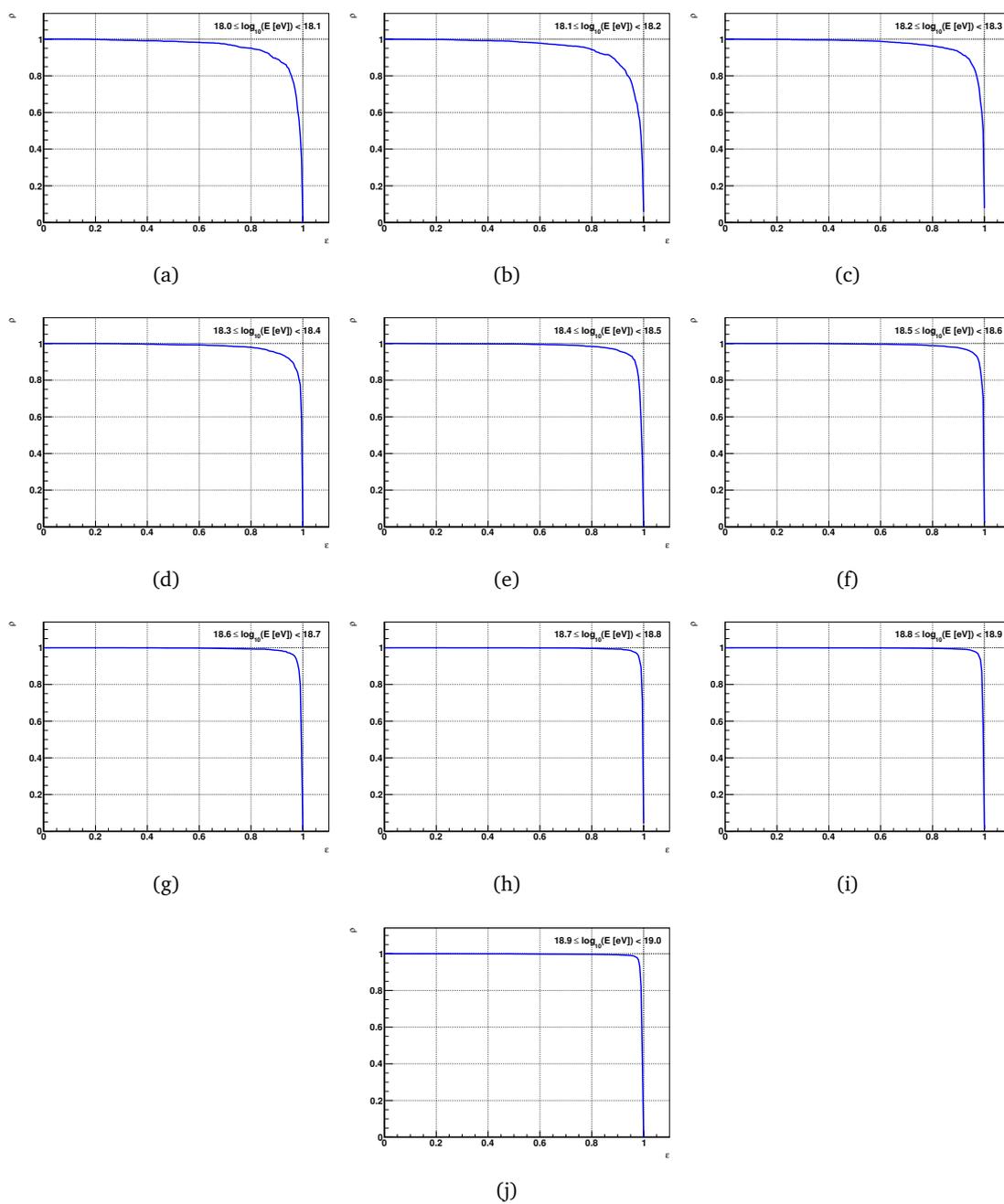
Figure C.12.: Background rejection ϱ as a function of the signal efficiency ε for the observable F_γ, including the additional selection criteria, in the individual energy bins.
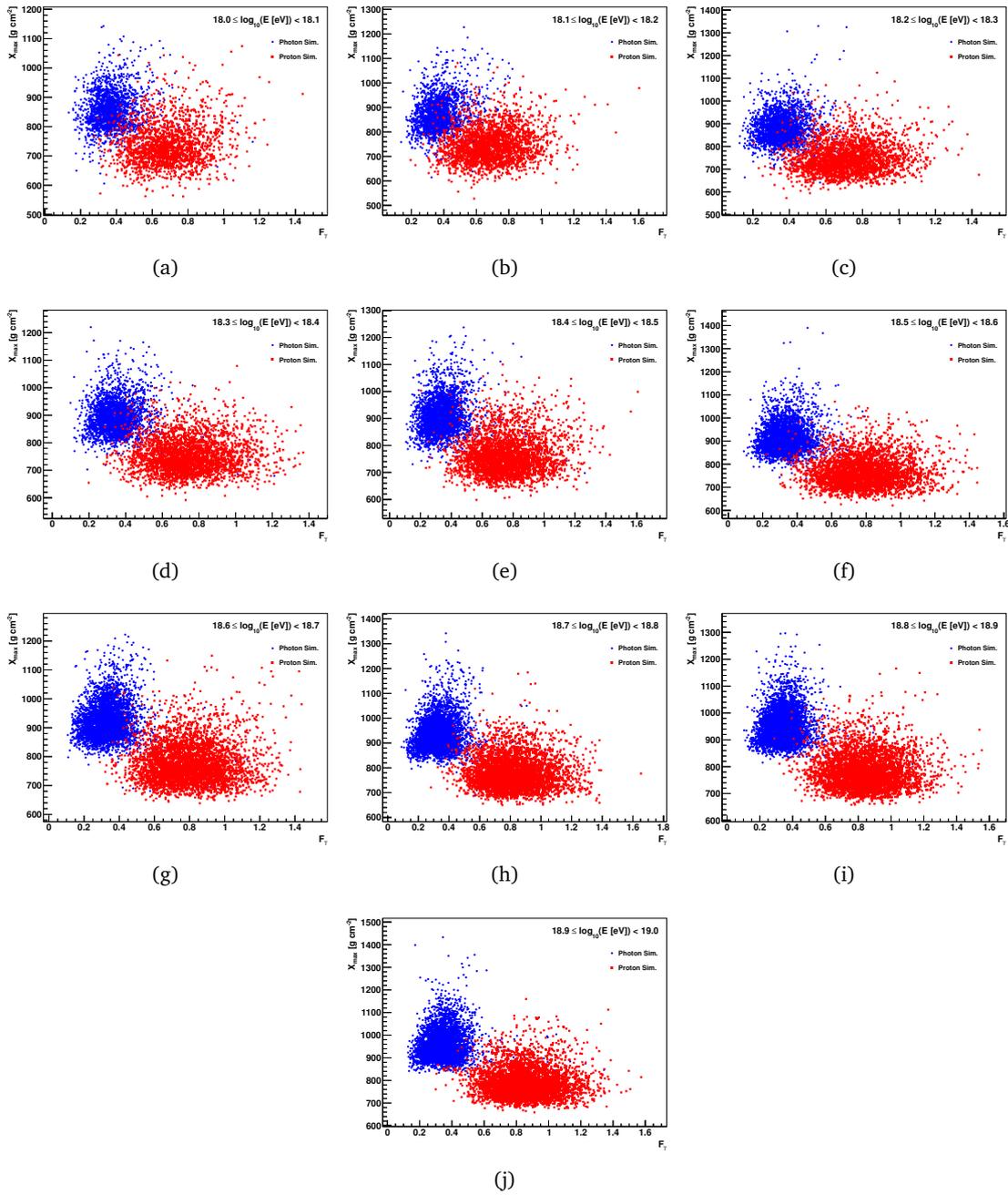
Figure C.13.: Distribution of the two observables F_γ and $X_{\max}$, shown in a scatter plot, in the individual energy bins. The photon sample and the proton sample are shown as blue dots and red squares, respectively.
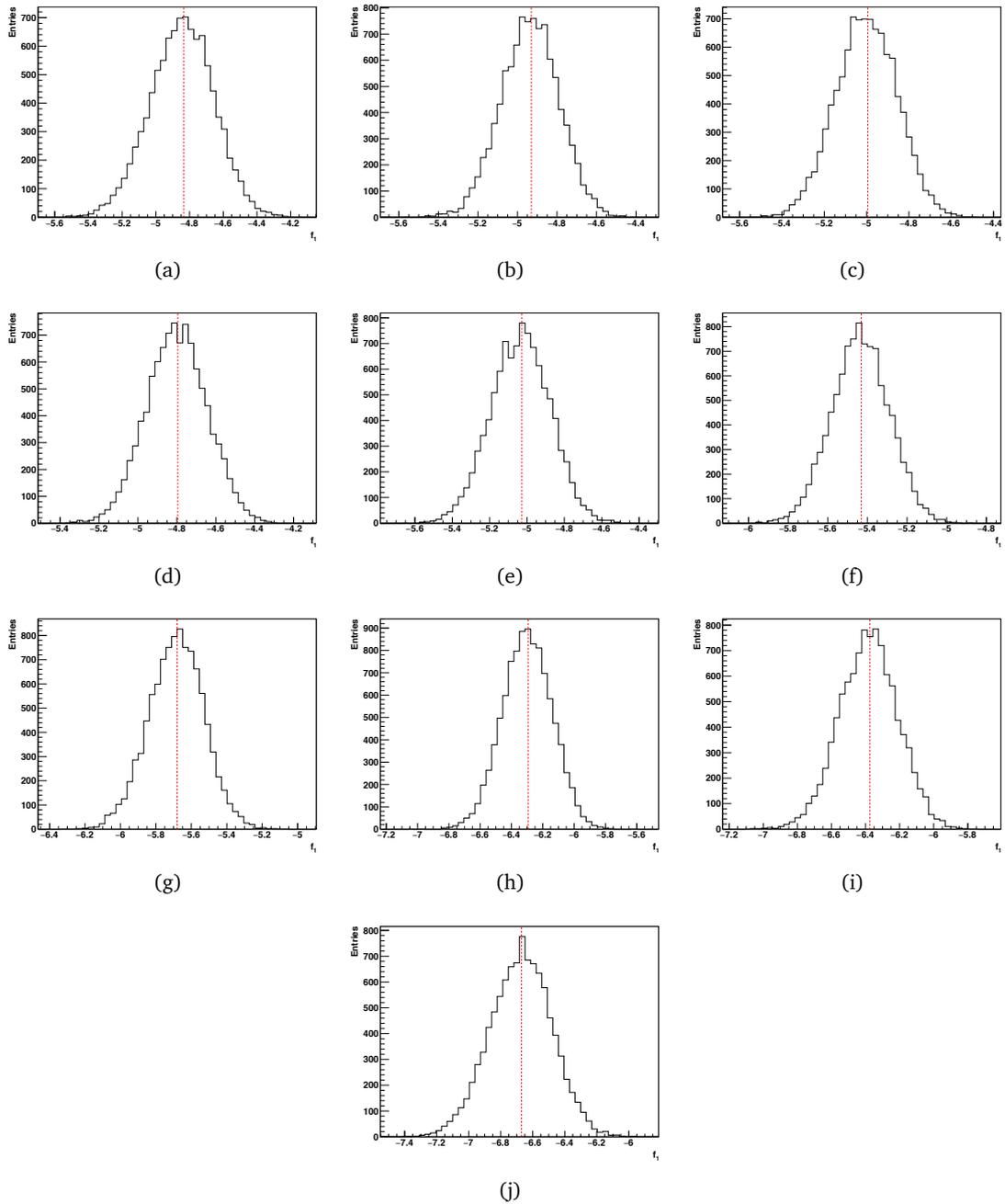
Figure C.14.: Distributions of the Fisher coefficient f_1 obtained through a bootstrapping method in the individual energy bins. The dashed lines indicate the values of f_1 that have been determined for the original samples
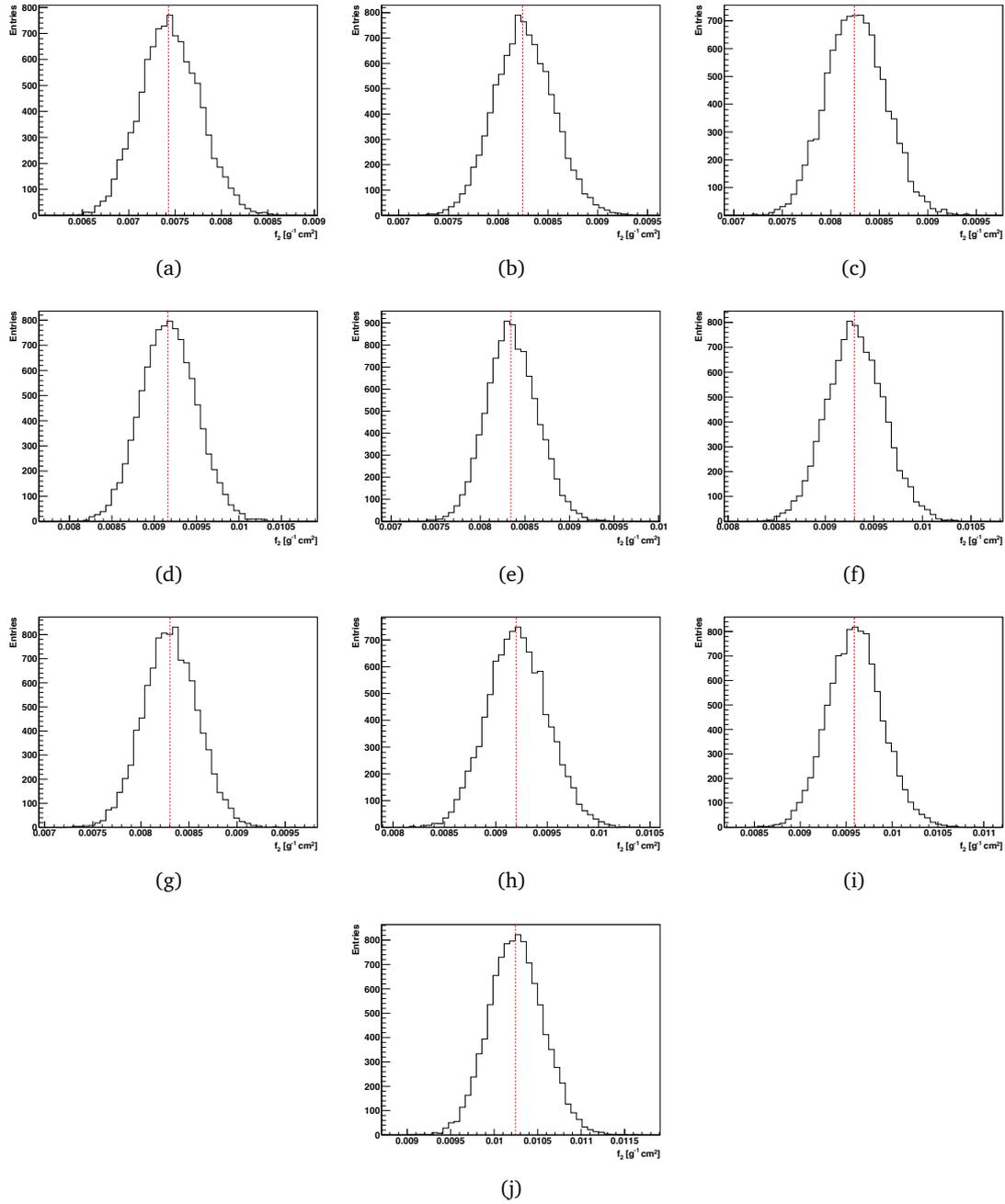
Figure C.15.: Distributions of the Fisher coefficient f_2 obtained through a bootstrapping method in the individual energy bins. The dashed lines indicate the values of f_2 that have been determined for the original samples
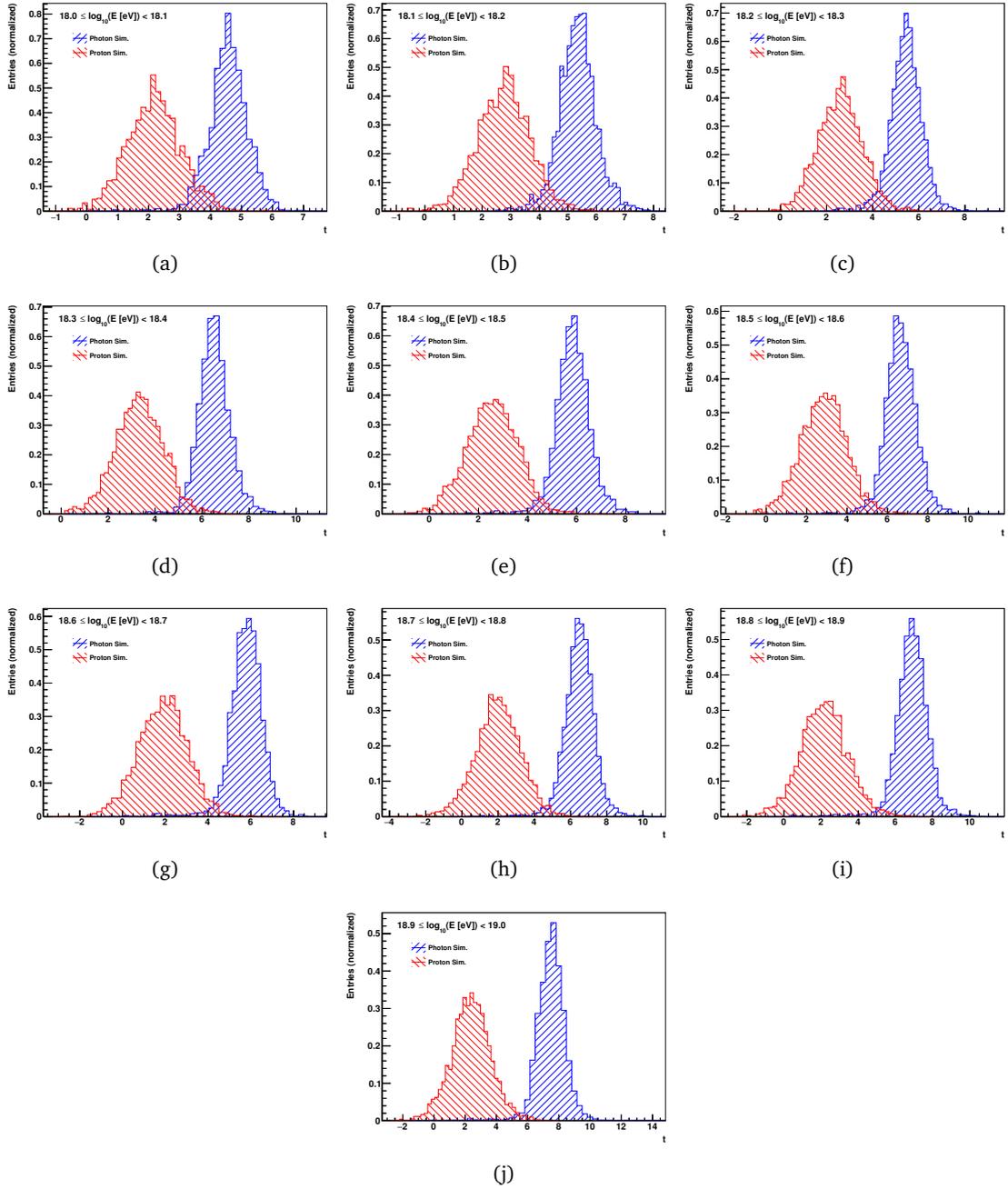
Figure C.16.: Distributions of the observable t in the individual energy bins for the photon sample, shown in blue, and the proton sample, shown in red.
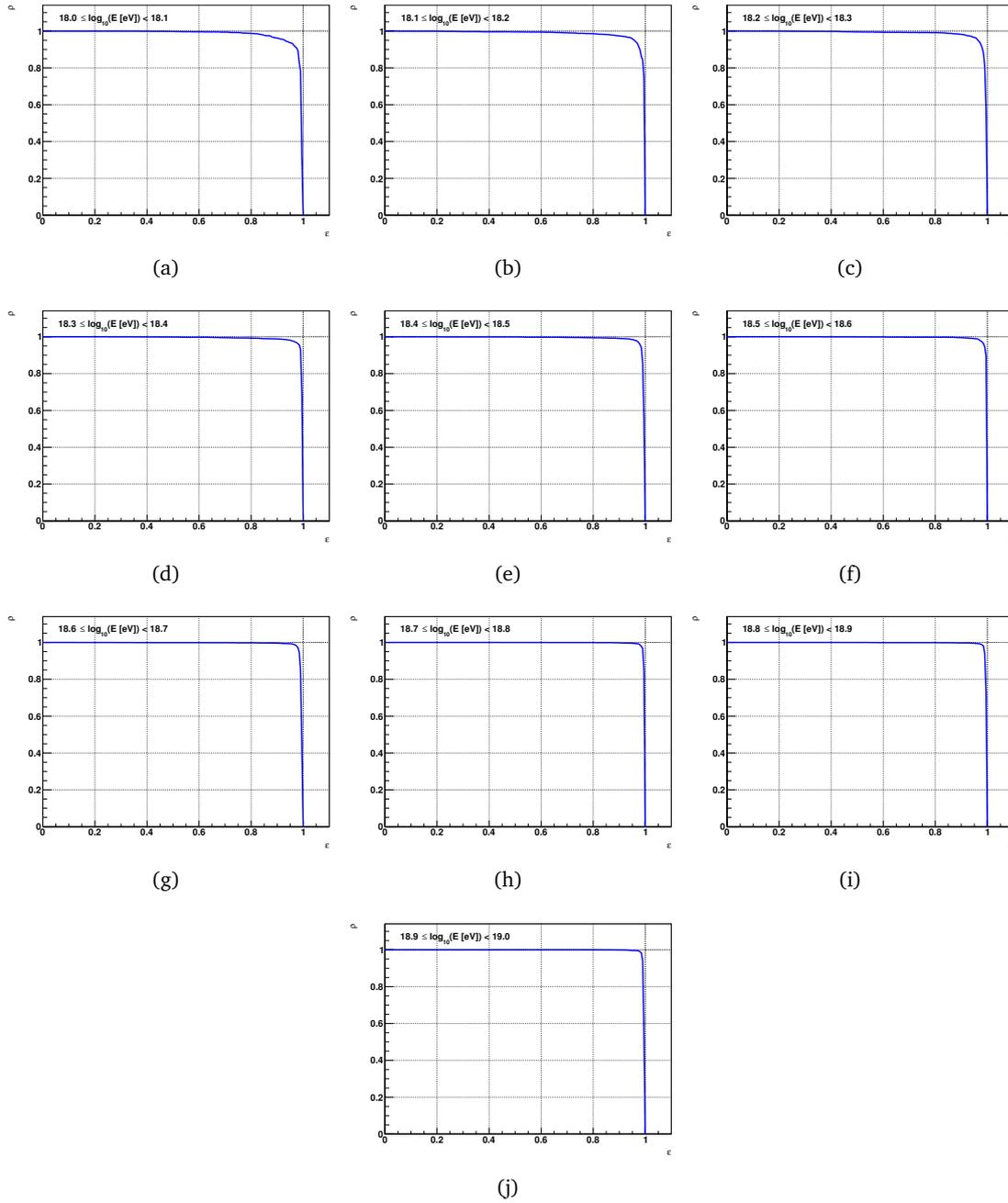
Figure C.17.: Background rejection ϱ as a function of the signal efficiency ε for the observable t in the individual energy bins.
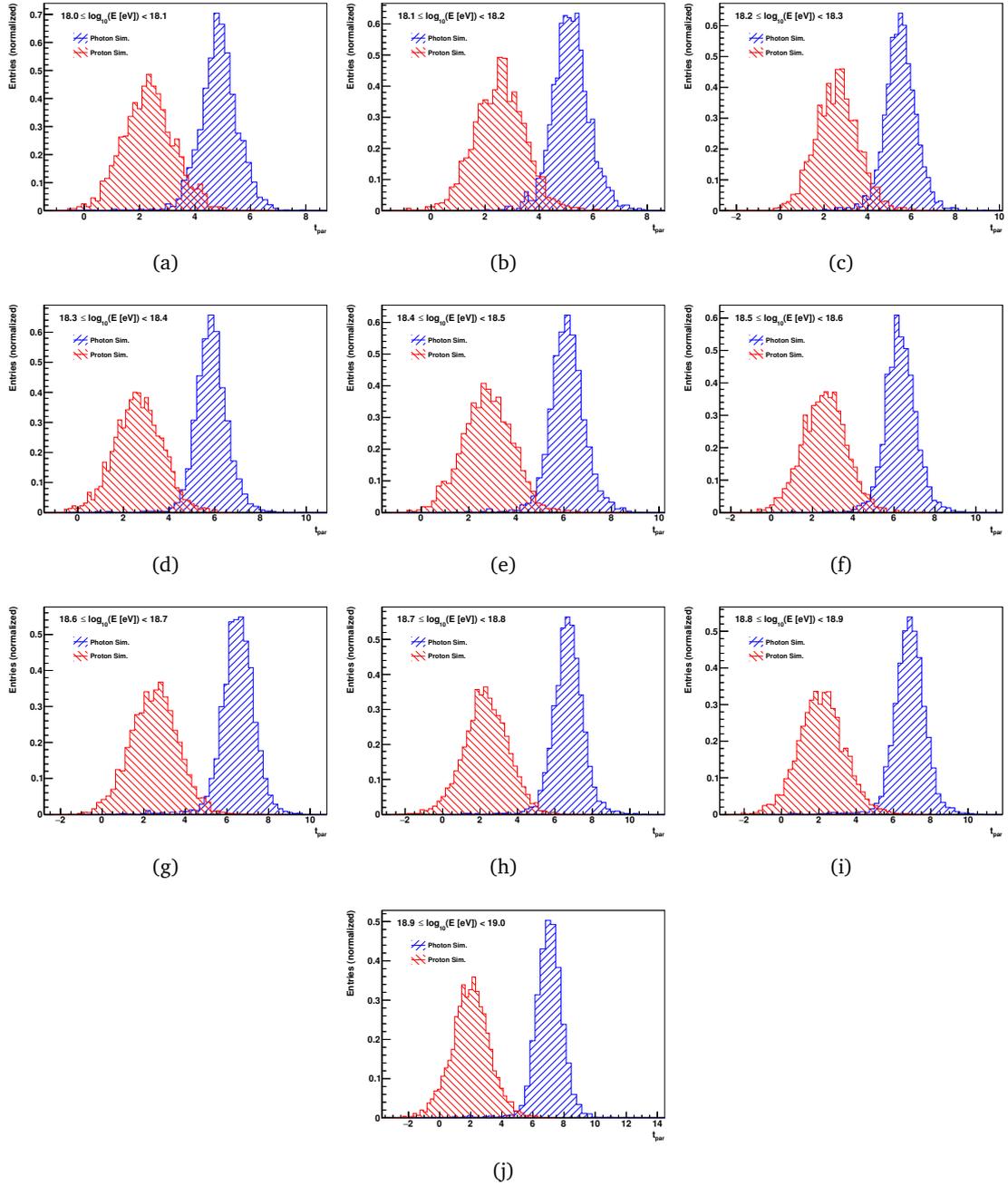
Figure C.18.: Distributions of the observable t_{par} in the individual energy bins for the photon sample, shown in blue, and the proton sample, shown in red.
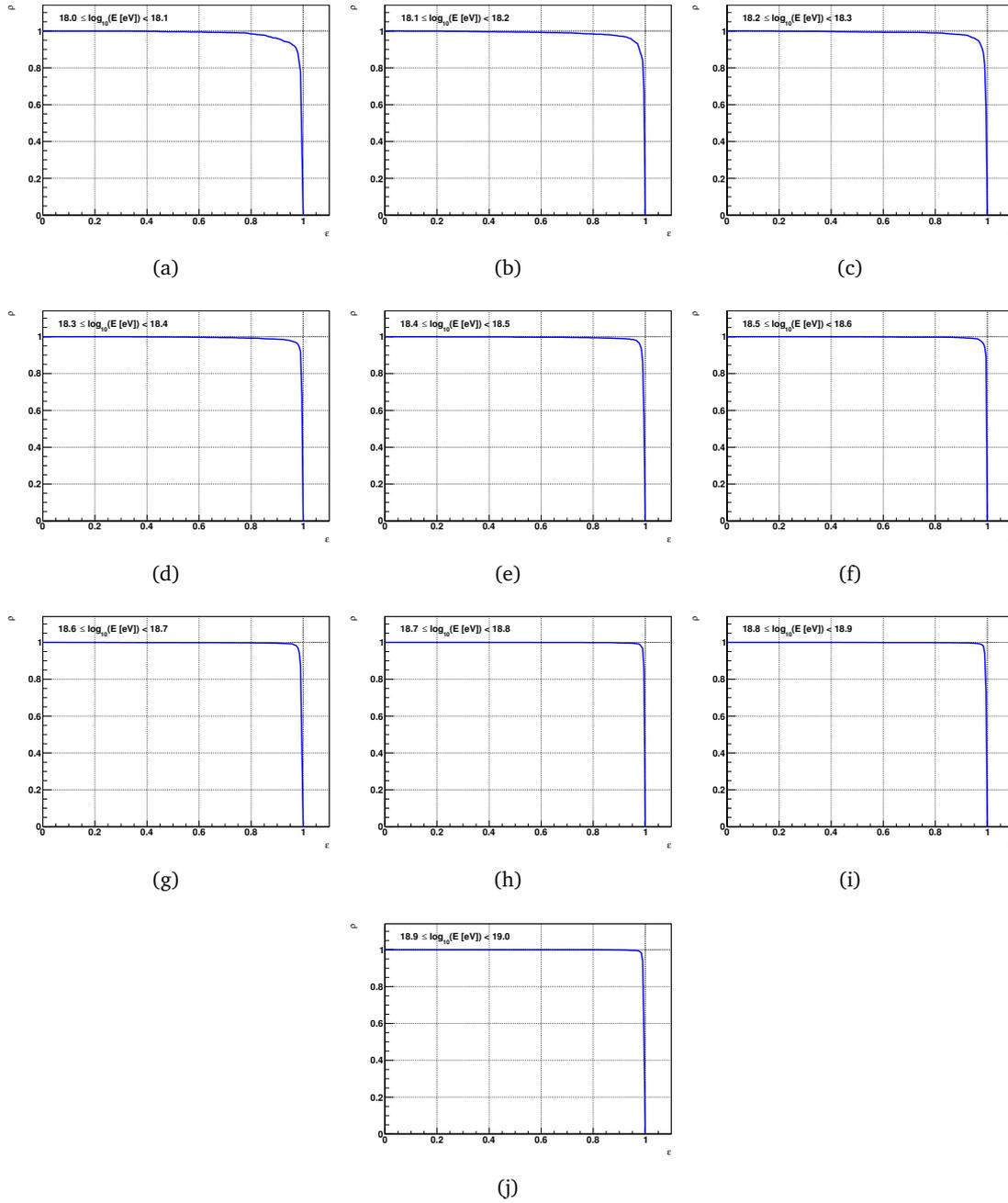
Figure C.19.: Background rejection ϱ as a function of the signal efficiency ε for the observable t_{par} in the individual energy bins.
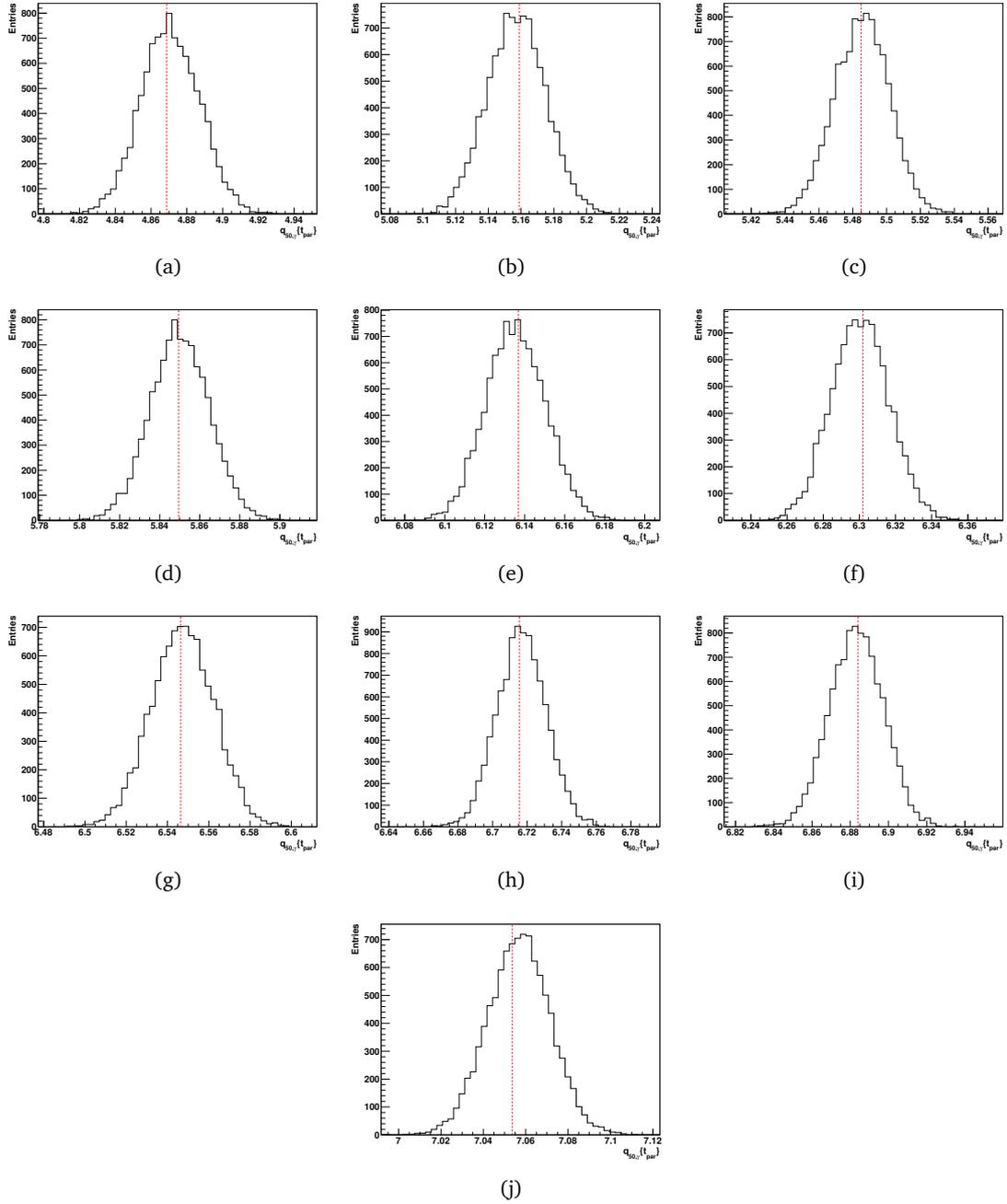
Figure C.20.: Distributions of the median q_{50} of the t_{par} distributions for the photon sample obtained through a bootstrapping method in the individual energy bins. The dashed lines indicate the values of q_{50} that have been determined for the original samples
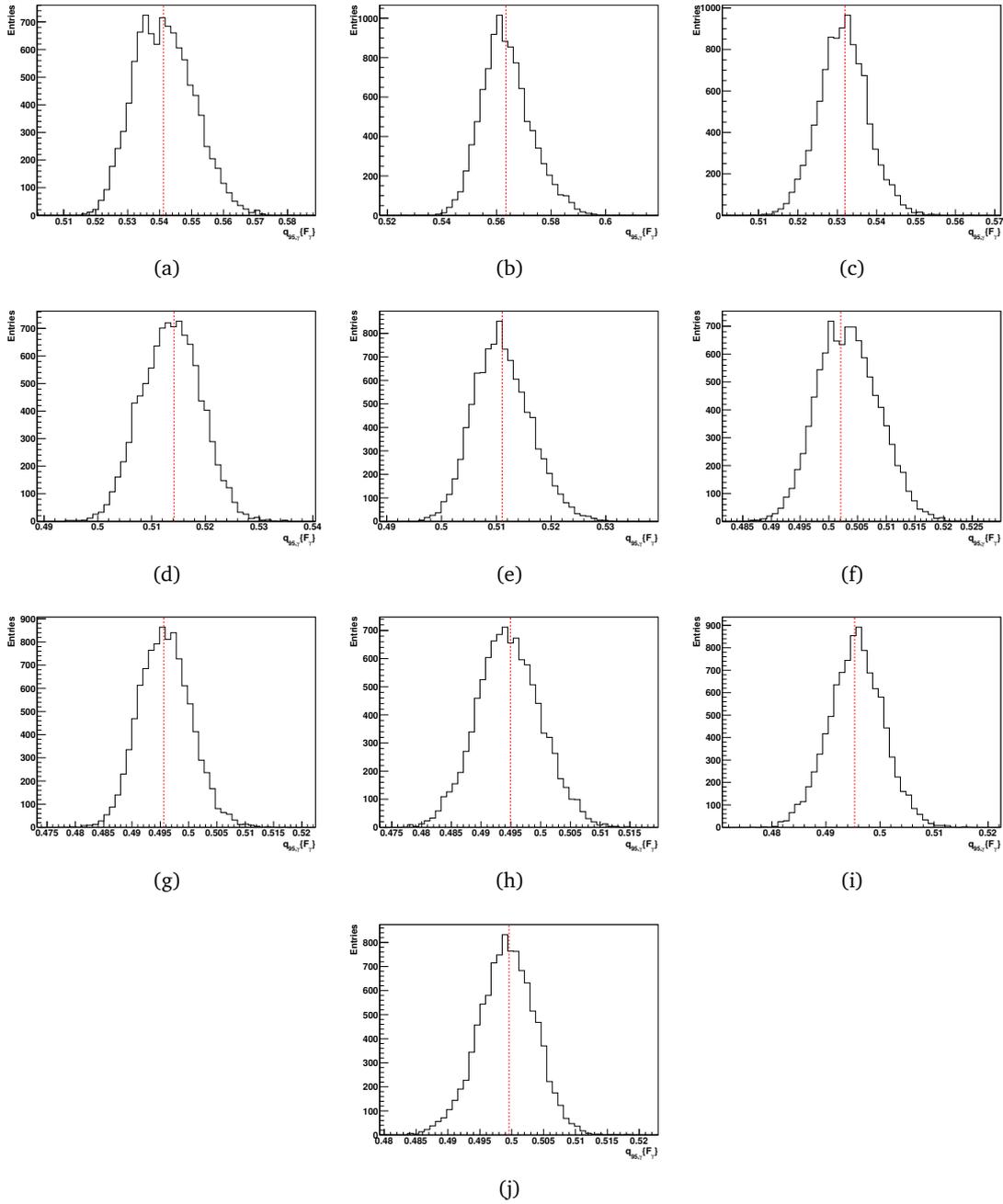
Figure C.21.: Distributions of the 95-quantile q_{95} of the F_γ distributions for the photon sample obtained through a bootstrapping method in the individual energy bins. The dashed lines indicate the values of q_{95} that have been determined for the original samples
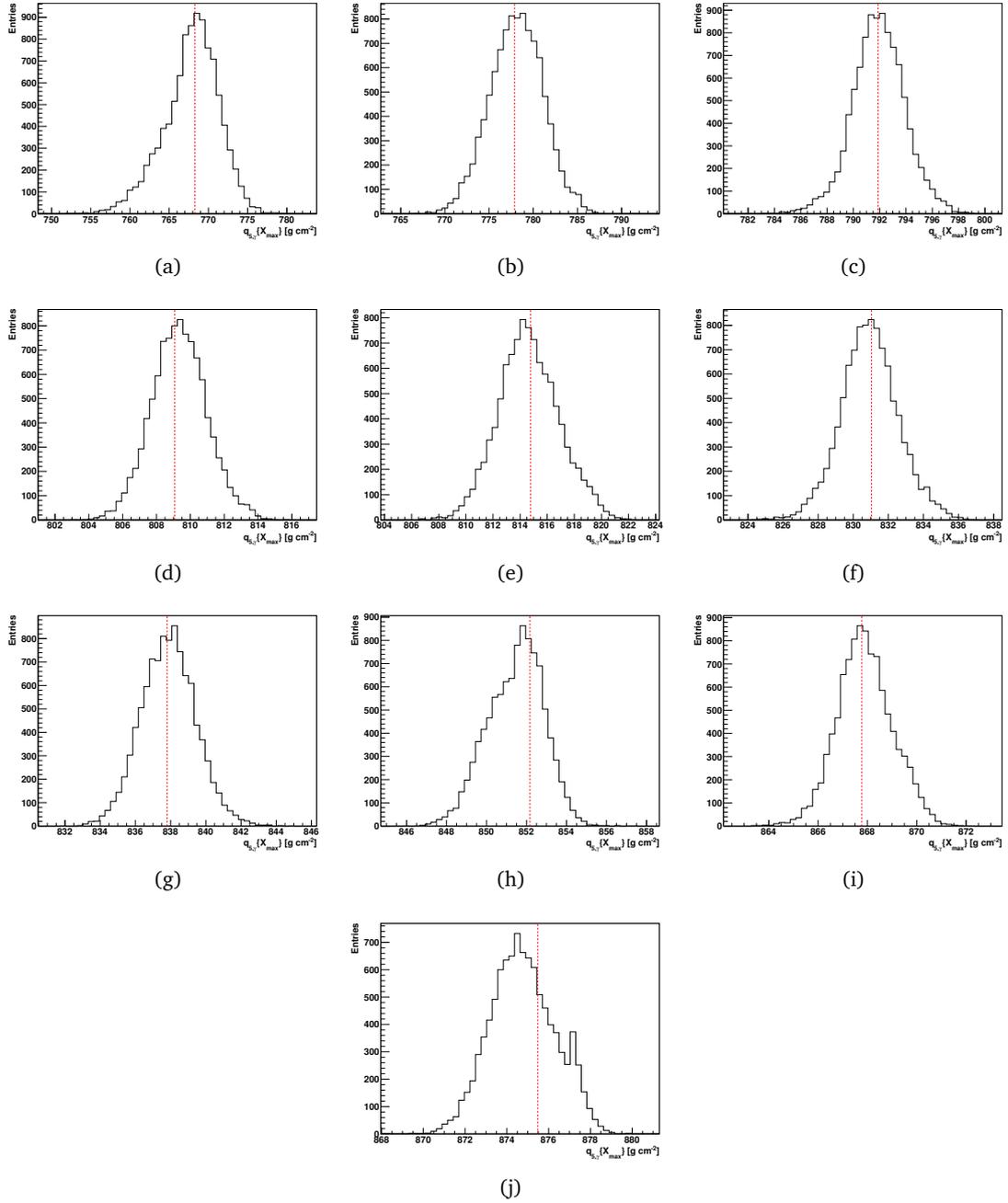
Figure C.22.: Distributions of the 5-quantile q_5 of the X_{max} distributions for the photon sample obtained through a bootstrapping method in the individual energy bins. The dashed lines indicate the values of q_5 that have been determined for the original samples

D. Acceptance Correction

The acceptance correction κ was introduced in Sec. 7.3.1 to account for the different acceptances of the detector for photon-induced and proton-induced air showers. This correction is used for the conservative determination of the upper limits on the photon fraction. In this section, κ is deduced for the calculation of upper limits from data. The derivation of κ given here summarizes the detailed calculation in [Abr09].

The fraction f_γ of photons in the total flux of cosmic rays above a given energy threshold can be written as

$$f_\gamma(E > E_0) = \frac{\int_{E_0}^{\infty} \varphi_\gamma(E)\, dE}{\int_{E_0}^{\infty} \varphi_\gamma(E)\, dE + \sum_i \int_{E_0}^{\infty} \varphi_i(E)\, dE} \tag{D.1}$$

where $\varphi_\gamma(E)$ denotes the differential flux of primary photons impinging on the Earth and the $\varphi_i(E)$, with $i = $ p, He, ..., represent the differential fluxes of primary protons and nuclei. The fraction that is measured by an experiment can be written in a similar way, however, the different acceptances $\mathscr{A}$ of the detectors to the different primary particle types have to be taken into account. In addition, the effective energy threshold E_i is dependent on the type of the primary particle:

$$f_\gamma^{\mathrm{meas}}(E > E_0) = \frac{\int_{E_0}^{\infty} \mathscr{A}_\gamma(E)\, \varphi_\gamma(E)\, dE}{\int_{E_0}^{\infty} \mathscr{A}_\gamma(E)\, \varphi_\gamma(E)\, dE + \sum_i \int_{E_i}^{\infty} \mathscr{A}_i(E)\, \varphi_i(E)\, dE}. \tag{D.2}$$

Hence, if an upper limit on the photon fraction is to be derived from data, the differences in the acceptance and the energy threshold have to be properly accounted for. From simulations, the relation between E_i and E_0 can be inferred for the case of an experiment based on the air-fluorescence technique:

$$E_i = E_0 \frac{m_i}{m_\gamma}, \tag{D.3}$$

where m_i and $m_\gamma < m_i$ denote the missing energy corrections needed for air showers induced by protons or nuclei and photons, respectively. Replacing E_i by E_0 and introducing the minimum acceptance for photon-induced air showers $A_\gamma^{\mathrm{min}} = \min\left[\mathscr{A}_\gamma(E)\,|\,E > E_0\right]$ yields

$$f_\gamma^{\mathrm{meas}}(E > E_0) > \frac{\int_{E_0}^{\infty} \mathscr{A}_\gamma^{\mathrm{min}}\, \varphi_\gamma(E)\, dE}{\int_{E_0}^{\infty} \mathscr{A}_\gamma^{\mathrm{min}}\, \varphi_\gamma(E)\, dE + \sum_i \int_{E_0}^{\infty} \mathscr{A}_i(E)\, \varphi_i(E)\, dE}. \tag{D.4}$$

Now, the acceptance ratio $\kappa_i(E) = \mathscr{A}_\gamma^{\mathrm{min}}/\mathscr{A}_i(E)$ is introduced:

$$f_\gamma^{\mathrm{meas}}(E > E_0) > \frac{\int_{E_0}^{\infty} \mathscr{A}_\gamma^{\mathrm{min}}\, \varphi_\gamma(E)\, dE}{\int_{E_0}^{\infty} \mathscr{A}_\gamma^{\mathrm{min}}\, \varphi_\gamma(E)\, dE + \sum_i \int_{E_0}^{\infty} \mathscr{A}_\gamma^{\mathrm{min}}/\kappa_i(E)\, \varphi_i(E)\, dE}. \tag{D.5}$$

Cancelling out $\mathscr{A}_\gamma^{\mathrm{min}}$ and introducing the minimum acceptance ratio $\kappa = \min\left[\kappa_i(E)\,|\,E > E_0\right]$ with $\kappa < 1$ gives

$$f_\gamma^{\mathrm{meas}}(E > E_0) > \kappa \frac{\int_{E_0}^{\infty} \varphi_\gamma(E)\, dE}{\int_{E_0}^{\infty} \varphi_\gamma(E)\, dE + \sum_i \int_{E_0}^{\infty} \varphi_i(E)\, dE} = \kappa\, f_\gamma(E > E_0). \tag{D.6}$$

Therefore, an upper limit f_γ^{ul} to the photon fraction can be conservatively calculated from data as

$$f_\gamma^{\mathrm{ul}} = f_\gamma^{\mathrm{ul,\, meas}}/\kappa > f_\gamma^{\mathrm{meas}}/\kappa > f_\gamma. \tag{D.7}$$

It should be pointed out that the upper limit obtained in this way does not require any assumptions on the differentual fluxes of the different components φ_i of the total flux of cosmic rays.

Bibliography

[Aab14a] ALEXANDER AAB ET AL. (PIERRE AUGER COLLABORATION), *Depth of maximum of air-shower profiles at the Pierre Auger Observatory. I. Measurements at energies above $10^{17.8}$ eV*, Physical Review D **90** (2014), p. 122005.

[Aab14b] ALEXANDER AAB ET AL. (PIERRE AUGER COLLABORATION), *Probing the radio emission from air showers with polarization measurements*, Physical Review D **89** (2014), p. 052002.

[Aab14c] ALEXANDER AAB ET AL. (PIERRE AUGER COLLABORATION), *A search for point sources of EeV photons*, The Astrophysical Journal **789** (2014), pp. 160–171.

[Aab14d] ALEXANDER AAB ET AL. (PIERRE AUGER COLLABORATION), *A targeted search for point sources of EeV neutrons*, The Astrophysical Journal Letters **789** (2014), pp. L34–L41.

[Aad14] GEORGES AAD ET AL. (ATLAS COLLABORATION), *A general search for new phenomena with the ATLAS detector in pp collisions at $\sqrt{s} = 8$ TeV*, ATLAS-CONF-2014-006, CERN, Geneva, 2014.

[Aar13] MARK G. AARTSEN ET AL. (ICECUBE COLLABORATION), *Probing the origin of cosmic rays with extremely high energy neutrinos using the IceCube Observatory*, Physical Review D **88** (2013), p. 112008.

[Abb08a] RASHA U. ABBASI ET AL. (HIRES COLLABORATION), *First Observation of the Greisen-Zatsepin-Kuzmin Suppression*, Physical Review Letters **100** (2008), p. 101101.

[Abb08b] RASHA U. ABBASI ET AL. (HIRES COLLABORATION), *Search for correlations between HiRes stereo events and active galactic nuclei*, Astroparticle Physics **30** (2008), pp. 175–179.

[Abb10] RASHA U. ABBASI ET AL. (HIRES COLLABORATION), *Indications of Proton-Dominated Cosmic-Ray Composition above 1.6 EeV*, Physical Review Letters **104** (2010), p. 161101.

[Abb14] RASHA U. ABBASI ET AL. (TELESCOPE ARRAY COLLABORATION), *A Northern Sky Survey for Point-Like Sources of EeV Neutral Particles with the Telescope Array Experiment*, Manuscript Submitted for Publication, Preprint Available at arXiv:astro-ph/1407.6145, 2014.

[Abr07a] JORGE ABRAHAM ET AL. (PIERRE AUGER COLLABORATION), *Anisotropy studies around the galactic centre at EeV energies with the Auger Observatory*, Astroparticle Physics **27** (2007), pp. 244–253.

[Abr07b] JORGE ABRAHAM ET AL. (PIERRE AUGER COLLABORATION), *Correlation of the Highest-Energy Cosmic Rays with Nearby Extragalactic Objects*, Science **318** (2007), pp. 938–943.

[Abr07c] JORGE ABRAHAM ET AL. (PIERRE AUGER COLLABORATION), *An upper limit to the photon fraction in cosmic rays above 10^{19} eV from the Pierre Auger Observatory*, Astroparticle Physics **27** (2007), pp. 155–168.

[Abr08a]   Jorge Abraham et al. (Pierre Auger Collaboration), *Observation of the Suppression of the Flux of Cosmic Rays above 4×10^{19} eV*, Physical Review Letters **101** (2008), p. 061101.

[Abr08b]   Jorge Abraham et al. (Pierre Auger Collaboration), *Upper limit on the cosmic-ray photon flux above 10^{19} eV using the surface detector of the Pierre Auger Observatory*, Astroparticle Physics **29** (2008), pp. 243–256.

[Abr09]   Jorge Abraham et al. (Pierre Auger Collaboration), *Upper limit on the cosmic-ray photon fraction at EeV energies from the Pierre Auger Observatory*, Astroparticle Physics **31** (2009), pp. 399–406.

[Abr10a]   Jorge Abraham et al. (Pierre Auger Collaboration), *The fluorescence detector of the Pierre Auger Observatory*, Nuclear Instruments and Methods in Physics Research A **620** (2010), pp. 227–251.

[Abr10b]   Jorge Abraham et al. (Pierre Auger Collaboration), *Measurement of the energy spectrum of cosmic rays above 10^{18} eV using the Pierre Auger Observatory*, Physics Letters B **685** (2010), pp. 239–246.

[Abr10c]   Jorge Abraham et al. (Pierre Auger Collaboration), *A study of the effect of molecular and aerosol conditions in the atmosphere on air fluorescence measurements at the Pierre Auger Observatory*, Astroparticle Physics **33** (2010), pp. 108–129.

[Abr10d]   Jorge Abraham et al. (Pierre Auger Collaboration), *Trigger and aperture of the surface detector array of the Pierre Auger Observatory*, Nuclear Instruments and Methods in Physics Research A **613** (2010), pp. 29–39.

[Abr10e]   Pedro Abreu et al. (Pierre Auger Collaboration), *Update on the correlation of the highest energy cosmic rays with nearby extragalactic matter*, Astroparticle Physics **34** (2010), pp. 314–326.

[Abr11]   Pedro Abreu et al. (Pierre Auger Collaboration), *The exposure of the hybrid detector of the Pierre Auger Observatory*, Astroparticle Physics **34** (2011), pp. 368–381.

[Abr12a]   Pedro Abreu et al. (Pierre Auger Collaboration), *Description of atmospheric conditions at the Pierre Auger Observatory using the Global Data Assimilation System (GDAS)*, Astroparticle Physics **35** (2012), pp. 591–607.

[Abr12b]   Pedro Abreu et al. (Pierre Auger Collaboration), *The rapid atmospheric monitoring system of the Pierre Auger Observatory*, Journal of Instrumentation **7** (2012), p. P09001.

[Abr12c]   Pedro Abreu et al. (Pierre Auger Collaboration), *A Search for Point Sources of EeV Neutrons*, The Astrophysical Journal **760** (2012), pp. 148–158.

[Abr13]   Pedro Abreu et al. (Pierre Auger Collaboration), *Ultrahigh Energy Neutrinos at the Pierre Auger Observatory*, Advances in High Energy Physics **2013** (2013), p. 708680.

[Abu13a]   Tareq Abu-Zayyad et al. (Telescope Array Collaboration), *The Cosmic-Ray Energy Spectrum Observed with the Surface Detector of the Telescope Array Experiment*, The Astrophysical Journal Letters **768** (2013), pp. L1–L5.

[Abu13b]   Tareq Abu-Zayyad et al. (Telescope Array Collaboration), *Upper limit on the flux of photon with energies above 10^{19} eV using the Telescope Array surface detector*, Physical Review D **88** (2013), p. 112005.

[Ack13]   Markus Ackermann et al. (Fermi LAT Collaboration), *Detection of the Characteristic Pion-Decay Signature in Supernova Remnants*, Science **339** (2013), pp. 807–811.

[Ago03]   Stefano Agostinelli et al. (Geant4 Collaboration), *Geant4—a simulation toolkit*, Nuclear Instruments and Methods in Physics Research A **506** (2003), pp. 250–303.

[Agu13]   Manuel Aguilar et al. (AMS Collaboration), *First Result from the Alpha Magnetic Spectrometer on the International Space Station: Precision Measurement of the Positron Fraction in Primary Cosmic Rays of 0.5 - 340 GeV*, Physical Review Letters **110** (2013), p. 141102.

[**Ahn09**] Eun-Joo Ahn et al., *Cosmic ray interaction event generator SIBYLL 2.1*, Physical Review D **80** (2009), p. 094003.

[**All05**] Denis Allard et al., *UHE nuclei propagation and the interpretation of the ankle in the cosmic-ray spectrum*, Astronomy & Astrophysics **443** (2005), pp. L29–L32.

[**All06**] John Allison et al., *Geant4 Developments and Applications*, IEEE Transactions on Nuclear Science **53** (2006), pp. 270–278.

[**All08**] Ingomar Allekotte et al. for the Pierre Auger Collaboration, *The surface detector system of the Pierre Auger Observatory*, Nuclear Instruments and Methods in Physics Research A **586** (2008), pp. 409–420.

[**All11**] Patrick S. Allison for the Pierre Auger Collaboration, *Microwave detection of cosmic ray showers at the Pierre Auger Observatory*, Proceedings of the 32nd International Cosmic Ray Conference (ICRC 2011), Beijing, China, 2011.

[**Alo04**] Roberto Aloisio, Veniamin Berezinsky, and Michael Kachelriess, *Fragmentation functions in supersymmetric QCD and ultrahigh energy cosmic ray spectra produced in top-down models*, Physical Review D **69** (2004), p. 094023.

[**Alo11**] Roberto Aloisio, Veniamin Berezinsky, and Askhat Gazizov, *Ultra high energy cosmic rays: The disappointing model*, Astroparticle Physics **34** (2011), pp. 620–626.

[**Alo12**] Roberto Aloisio, Veniamin Berezinsky, and Askhat Gazizov, *Transition from galactic to extragalactic cosmic rays*, Astroparticle Physics **39-40** (2012), pp. 129–143.

[**Alo13a**] Roberto Aloisio, Veniamin Berezinsky, and Svetlana Grigorieva, *Analytic calculations of the spectra of ultra high energy cosmic ray nuclei. I. The case of CMB radiation*, Astroparticle Physics **41** (2013), pp. 73–93.

[**Alo13b**] Roberto Aloisio, Veniamin Berezinsky, and Svetlana Grigorieva, *Analytic calculations of the spectra of ultra high energy cosmic ray nuclei. II. The general case of background radiation*, Astroparticle Physics **41** (2013), pp. 94–107.

[**Alv13**] Jaime Alvarez-Muñiz, Markus Risse for the Pierre Auger Collaboration, Grigory I. Rubtsov, and Benjamin T. Stokes for the Telescope Array Collaboration, *Review of the Multimessenger Working Group at UHECR-2012*, EPJ Web of Conferences **53** (2013), p. 01009.

[**Ame06**] Michihiro Amenomori et al. (Tibet ASγ Collaboration), *Anisotropy and Corotation of Galactic Cosmic Rays*, Science **314** (2006), pp. 439–443.

[**And33**] Carl D. Anderson, *The positive electron*, The Physical Review **43** (1933), pp. 491–494.

[**Ant04a**] Torsten Antoni et al. (KASCADE Collaboration), *Large-scale cosmic-ray anisotropy with KASCADE*, The Astrophysical Journal **604** (2004), pp. 687–692.

[**Ant04b**] Torsten Antoni et al. (KASCADE Collaboration), *Search for cosmic-ray point sources with KASCADE*, The Astrophysical Journal **608** (2004), pp. 865–871.

[**Ape13**] Wolf D. Apel et al. (KASCADE-Grande Collaboration), *KASCADE-Grande measurements of energy spectra for elemental groups of cosmic rays*, Astroparticle Physics **47** (2013), pp. 54–66.

[**Arg03**] Stefano Argirò, *FdSDPFinder.cc*, 2003, Documentation of the Auger Offline Software Framework, version 2.9.1.

[**Arg07**] Stefano Argiró et al., *The offline software framework of the Pierre Auger Observatory*, Nuclear Instruments and Methods in Physics Research A **580** (2007), pp. 1485–1496.

[**Ass11**] Pedro Assis et al., *Offline Reference Manual FD Simulation*, 2011, Revision 19813, based on the internal note GAP-2008-014.

[**Aug39**] Pierre Auger, Roland Maze, Paul Ehrenfest Jr., and André Fréon, *Les grandes gerbes de rayons cosmiques*, Journal de Physique et Le Radium **10** (1939), pp. 39–48.

[Ave07a] MAXIMO AVE FOR THE PIERRE AUGER COLLABORATION, *Reconstruction accuracy of the surface detector array of the Pierre Auger Observatory*, Proceedings of the 30th International Cosmic Ray Conference (ICRC 2007), Mérida, México, 2007.

[Ave07b] MAXIMO AVE ET AL. (AIRFLY COLLABORATION), *Measurement of the pressure dependence of air fluorescence emission induced by electrons*, Astroparticle Physics **28** (2007), pp. 41–57.

[Ave08] MAXIMO AVE ET AL. (AIRFLY COLLABORATION), *Temperature and humidity dependence of air fluorescence yield measured by AIRFLY*, Nuclear Instruments and Methods in Physics Research A **597** (2008), pp. 50–54.

[Ave13] MAXIMO AVE ET AL. (AIRFLY COLLABORATION), *Precise measurement of the absolute fluorescence yield of the 337 nm band in atmospheric gases*, Astroparticle Physics **42** (2013), pp. 90–102.

[Bah09] JULIO L. BAHILO, THOMAS PAUL, TROY PORTER, AND DARKO VEBERIČ FOR THE $\overline{\text{OFF}}\underline{\text{LINE}}$ TEAM, $\overline{\text{Off}}\underline{\text{line}}$ *Reference Manual SD Simulation*, 2009, Revision 14717.

[Bat07] GIUSEPPE BATTISTONI ET AL., *The FLUKA code: description and benchmarking*, AIP Conference Proceedings **896** (2007), p. 31.

[Bau09] PABLO M. BAULEO AND JULIO RODRÍGUEZ MARTINO, *The dawn of the particle astronomy era in ultra-high-energy cosmic rays*, Nature **458** (2009), pp. 847–851.

[Bec96] HENRI BECQUEREL, *Sur les radiations invisibles émises par les corps phosphorescents*, Comptes rendus hebdomadaires des séances de l'Académie des Sciences **122** (1896), pp. 501–503.

[Bel13] JOSÉ BELLIDO, BRUCE DAWSON, AND STEVEN SAFFI, *Fixing the Systematics in the Hybrid Reconstruction of the Shower Axis in Loma Amarilla*, 2013, internal note GAP-2013-010.

[Ben07a] SEGEV Y. BENZVI ET AL., *The Lidar system of the Pierre Auger Observatory*, Nuclear Instruments and Methods in Physics Research A **574** (2007), pp. 171–184.

[Ben07b] SEGEV Y. BENZVI ET AL. FOR THE PIERRE AUGER COLLABORATION, *New method for atmospheric calibration at the Pierre Auger Observatory using FRAM, a robotic astronomical telescope*, Proceedings of the 30th International Cosmic Ray Conference (ICRC 2007), Mérida, México, 2007.

[Ber96] EVGENY G. BEREZHKO, *Maximum energy of cosmic rays accelerated by supernova shocks*, Astroparticle Physics **5** (1996), pp. 367–378.

[Ber97] VENIAMIN BEREZINSKY, MICHAEL KACHELRIESS, AND ALEXANDER VILENKIN, *Ultrahigh Energy Cosmic Rays without Greisen-Zatsepin-Kuzmin Cutoff*, Physical Review Letters **79** (1997), pp. 4302–4305.

[Ber06a] VENIAMIN BEREZINSKY, *On transiton from galactic to extragalactic cosmic rays*, Journal of Physics: Conference Series **47** (2006), pp. 142–153.

[Ber06b] XAVIER BERTOU ET AL. FOR THE PIERRE AUGER COLLABORATION, *Calibration of the surface array of the Pierre Auger Observatory*, Nuclear Instruments and Methods in Physics Research A **568** (2006), pp. 839–846.

[Ber09] ADRIAAN M. VAN DEN BERG FOR THE PIERRE AUGER COLLABORATION, *Radio detection of cosmic rays at the southern Auger Observatory*, Proceedings of the 31st International Cosmic Ray Conference (ICRC 2009), Łódź, Poland, 2009.

[Ber12] JÜRG BERINGER ET AL. (PARTICLE DATA GROUP), *Review of Particle Physics*, Physical Review D **86** (2012), p. 010001, and the 2013 partial update for the 2014 edition.

[Ber14] ADRIAAN M. VAN DEN BERG ET AL., *Final recommendations of the Upgrade Committee*, 2014, internal document of the Pierre Auger Collaboration.

[Bet34] HANS BETHE AND WALTER HEITLER, *On the Stopping of Fast Particles and on the Creation of Positive Electrons*, Proceedings of the Royal Society London A **146** (1934), pp. 83–112.

[Bha37]   Homi J. Bhabha and Walter Heitler, *The Passage of Fast Electrons and the Theory of Cosmic Showers*, Proceedings of the Royal Society of London A **159** (1937), pp. 432–458.

[Bha00]   Pijushpani Bhattacharjee and Günter Sigl, *Origin and propagation of extremely high-energy cosmic rays*, Physics Reports **327** (2000), pp. 109–247.

[Bie87]   Peter L. Biermann and Peter A. Strittmatter, *Synchrotron emission from shock waves in active galactic nuclei*, The Astrophysical Journal **322** (1987), pp. 643–649.

[Bil00]   Pierre Billoir, *Reconstruction of showers with the Ground Array: status of the "prototype" program*, 2000, internal note GAP-2000-025.

[Bir94]   David J. Bird et al., *The Cosmic ray energy spectrum observed by the Fly's Eye*, The Astrophysical Journal **424** (1994), pp. 491–502.

[Bla78]   Roger D. Blandford and Jeremiah P. Ostriker, *Particle acceleration by astrophysical shocks*, The Astrophysical Journal Letters **221** (1978), pp. L29–L32.

[Bla00]   Pasquale Blasi, Richard I. Epstein, and Angela V. Olinto, *Ultra-high-energy cosmic rays from young neutron star winds*, The Astrophysical Journal Letters **533** (2000), pp. L123–L126.

[Blü09]   see e.g. Johannes Blümer, Ralph Engel, and Jörg R. Hörandel, *Cosmic Rays from the Knee to the Highest Energies*, Progress in Particle and Nuclear Physics **63** (2009), pp. 293–338.

[Bol99]   Elihu Boldt and Pranab Ghosh, *Cosmic rays from remnants of quasars?*, Monthly Notices of the Royal Astronomical Society **307** (1999), pp. 491–494.

[Bon10]   Carla Bonifazi, Piera Luisa Ghia, and Isabelle Lhenry-Yvon, *On the unstable period of data-taking in 2009*, 2010, internal note GAP-2010-002.

[Bot29]   Walter Bothe and Werner Kolhörster, *Das Wesen der Höhenstrahlung*, Zeitschrift für Physik **56** (1929), pp. 751–777.

[Bra04]   Jeff T. Brack, Rishi Meyhandan, Gertjan Hofman, and James Matthews, *Absolute photometric calibration of large aperture optical systems*, Astroparticle Physics **20** (2004), pp. 653–659.

[Bre84]   Leo Breiman, Jerome H. Friedman, Richard A. Olshen, and Charles J. Stone, *Classification and regression trees*, Wadsworth International Group, Belmont, 1984.

[Bru97]   Rene Brun and Fons Rademakers, *ROOT - An object oriented data analysis framework*, Nuclear Instruments and Methods in Physics Research A **389** (1997), pp. 81–86.

[Cen14]   Central Intelligence Agency, *The World Factbook 2013-14: Argentina*, 2014, https://www.cia.gov/library/publications/the-world-factbook/index.html.

[Che34]   Pavel A. Cherenkov, *Vidimoe svechenie chistyh zhidkostej pod dejstviem γ-radiacii*, Doklady Akademii Nauk SSSR **2** (1934), pp. 451–454.

[Che86]   Kwong S. Cheng, Cheng Ho, and Malvin Ruderman, *Energetic radiation from rapidly spinning pulsars. II. Vela and Crab*, The Astrophysical Journal **300** (1986), pp. 522–539.

[Chi13]   Johana Chirinos for the Pierre Auger Collaboration, *Cloud Monitoring at the Pierre Auger Observatory*, Proceedings of the 33rd International Cosmic Ray Conference (ICRC 2013), Rio de Janeiro, Brazil, 2013.

[Chi15]   Johana Chirinos et al. (Pierre Auger Collaboration), *Atmospheric Studies for the Pierre Auger Observatory at MTU*, 2015, http://befnet2.auger.mtu.edu.

[Com33]   Arthur H. Compton, *A Geographic Study of Cosmic Rays*, The Physical Review **43** (1933), pp. 387–403.

[Com35]   Arthur H. Compton and Ivan A. Getting, *An Apparent Effect of Galactic Rotation on the Intensity of Cosmic Rays*, The Physical Review **47** (1935), pp. 817–821.

[Cou85]   CHARLES-AUGUSTIN DE COULOMB, *Troisième Mémoire sur l'Electricité et le Magnétisme*, Mémoires de l'Académie Royale des Sciences (1785), pp. 612–638.

[Cow98]   SEE E.G. GLEN COWAN, *Statistical data analysis*, Oxford University Press, Oxford, 1998.

[Cro03]   JAMES W. CRONIN, *Particle discrimination using the FADC traces from the Auger Observatory surface detectors*, 2003, internal note GAP-2003-076.

[Del13]   OLIVIER DELIGNY FOR THE PIERRE AUGER COLLABORATION, *Large-Scale Distribution of Arrival Directions of Cosmic Rays Detected at the Pierre Auger Observatory Above 10 PeV*, Proceedings of the 2nd Cosmic Ray Anisotropy Workshop, Madison, USA, 2013.

[Dem14]   HANS DEMBINSKI ET AL. (THE OBSERVER TEAM), *The Auger Observer*, 2014, http://augerobserver.fzk.de/.

[Dic65]   ROBERT H. DICKE, PHILLIP J. E. PEEBLES, PETER G. ROLL, AND DAVID T. WILKINSON, *Cosmic black-body radiation*, The Astrophysical Journal Letters **142** (1965), pp. L414–L419.

[Dom13]   MANLIO DE DOMENICO, MARIANGELA SETTIMO, SIMONE RIGGI, AND ERIC BERTIN, *Reinterpreting the development of extensive air showers initiated by nuclei and photons*, Journal of Cosmology and Astroparticle Physics **2013** (2013), p. 050.

[Efr79]   BRADLEY EFRON, *Bootstrap methods: another look at the jackknife*, The Annals of Statistics **7** (1979), pp. 1–26.

[Ell06]   JOHN ELLIS, VAN E. MAYES, AND DIMITRI V. NANOPOULOS, *Ultrahigh-energy cosmic rays particle spectra from crypton decays*, Physical Review D **74** (2006), p. 115003.

[Els01]   JULIUS ELSTER AND HANS GEITEL, *Weitere Versuche über die Elektrizitätszerstreuung in abgeschlossenen Luftmengen*, Physikalische Zeitschrift **2** (1901), pp. 560–563.

[Eng12]   SEE E.G. RALPH ENGEL, *Indirect Detection of Cosmic Rays*, in: Claus Grupen and Irène Buvat (eds.), Handbook of Particle Detection and Imaging, Springer-Verlag, Berlin, 2012.

[Erb66]   THOMAS ERBER, *High-Energy Electromagnetic Conversion Processes in Intense Magnetic Fields*, Reviews of Modern Physics **38** (1966), pp. 626–659.

[Erl97]   ANATOLY D. ERLYKIN AND ARNOLD W. WOLFENDALE, *A single source of cosmic rays in the range 10^{15} - 10^{16} eV*, Journal of Physics G: Nuclear and Particle Physics **23** (1997), pp. 979–989.

[Erl04]   ANATOLY D. ERLYKIN AND ARNOLD W. WOLFENDALE, *Cosmic rays and the Monogem supernova remnant*, Astroparticle Physics **22** (2004), pp. 47–63.

[Etc07]   ALBERTO ETCHEGOYEN FOR THE PIERRE AUGER COLLABORATION, *AMIGA, Auger Muons and Infill for the Ground Array*, Proceedings of the 30th International Cosmic Ray Conference (ICRC 2007), Mérida, México, 2007.

[Eur08]   EUROPEAN SPACE AGENCY, *XMM-Newton Spacecraft*, 2008, http://sci.esa.int/jump.cfm?oid=42885, Full Image Credit: ESA.

[Eur09]   EUROPEAN SOUTHERN OBSERVATORY, *Centaurus A**, 2009, Photo Release eso0903, Full Image Credit: ESO/WFI (Optical), MPIfR/ESO/APEX/A.Weiss et al. (Submillimetre), NASA/CXC/CfA/R.Kraft et al. (X-ray).

[Eur10]   EUROPEAN SOUTHERN OBSERVATORY, *VLT Unit Telescope at Paranal*, 2010, Photo Id ib-paranal09, Full Image Credit: Iztok Boncina/ESO.

[Eva08]   LYNDON EVANS AND PHILIP BRYANT (EDS.), *LHC machine*, Journal of Instrumentation **3** (2008), p. S08001.

[Eve06]   ARTHUR S. EVE, *On the Radioactive Matter in the Earth and the Atmosphere*, Philosophical Magazine London **12** (1906), pp. 189–200.

[Fai11]   MALCOLM FAIRBAIRN, TIMUR RASHBA, AND SERGEY V. TROITSKY, *Photon-axion mixing and ultra-high energy cosmic rays from BL Lac type objects: Shining light through the Universe*, Phys. Rev. D **84** (2011), p. 125019.

[Fal12]   BRIGITTE FALKENBURG AND WOLFGANG RHODE (EDS.), *From Ultra Rays to Astroparticles*, Springer Netherlands, Dordrecht, 2012.

[**Fan74**] BERNARD L. FANAROFF AND JULIA M. RILEY, *The morphology of extragalactic radio sources of high and low luminosity*, Monthly Notices of the Royal Astronomical Society **167** (1974), pp. 31–35.

[**Far99**] DANIELE FARGION, BARBARA MELE, AND ANDREA SALIS, *Ultra-High-Energy Neutrino Scattering onto Relic Light Neutrinos in the Galactic Halo as a Possible Source of the Highest Energy Extragalactic Cosmic Rays*, The Astrophysical Journal **517** (1999), pp. 725–733.

[**Fer49**] ENRICO FERMI, *On the Origin of the Cosmic Radiation*, The Physical Review **75** (1949), pp. 1169–1174.

[**Fer05**] ALFREDO FERRARI, PAOLA R. SALA, ALBERTO FASSÓ, AND JOHANNES RANFT, *FLUKA: a multiparticle transport code*, CERN-2005-010, CERN, Geneva, 2005.

[**Fic06**] BRIAN FICK ET AL., *The Central Laser Facility at the Pierre Auger Observatory*, Journal of Instrumentation **1** (2006), p. P11003.

[**Fis36**] RONALD A. FISHER, *The use of multiple measurements in taxonomic problems*, Annals of Eugenics **7** (1936), pp. 179–188.

[**Fri11**] RUDOLF G. A. FRICKE, *Günther & Tegetmeyer 1901-1958 — Instrumente für die Wissenschaft aus Braunschweig*, AF Verlag, Wolfenbüttel, 2011.

[**Frö09**] UWE FRÖHLICH, *Charakterisierung der Szintillatoren und der Ausleseelektronik des AMIGA-Myonsystems*, Master's Thesis, University of Siegen, Siegen, Germany, 2009.

[**Gai77**] THOMAS K. GAISSER AND ANTHONY M. HILLAS, *Reliability of the method of constant intensity cuts for reconstructing the average development of vertical showers*, Proceedings of the 15th International Cosmic Ray Conference (ICRC 1977), Plovdiv, Bulgaria, 1977.

[**Gal08**] MATTEO GALAVERNI AND GÜNTER SIGL, *Lorentz Violation for Photons and Ultrahigh-Energy Cosmic Rays*, Physical Review Letters **100** (2008), p. 021102.

[**Gei28**] HANS GEIGER AND WALTHER MÜLLER, *Elektronenzählrohr zur Messung schwächster Aktivitäten*, Die Naturwissenschaften **16** (1928), pp. 617–618.

[**Gel08**] GRACIELA GELMINI, OLEG E. KALASHEV, AND DMITRI V. SEMIKOZ, *GZK Photons as Ultra-High-Energy Cosmic Rays*, Journal of Experimental and Theoretical Physics **106** (2008), pp. 1061–1082.

[**Gin64**] VITALY L. GINZBURG AND SERGEI I. SYROVATSKII, *The Origin of Cosmic Rays*, Pergamon Press Ltd., Oxford, 1964.

[**Glu10**] ALEXANDER V. GLUSHKOV ET AL. (YAKUTSK EAS ARRAY), *Constraints on the flux of primary cosmic-ray photons at energies $E > 10^{18}$ eV from Yakutsk muon data*, Physical Review D **82** (2010), p. 041101(R).

[**Goc08**] ALBERT GOCKEL AND THEODOR WULF, *Beobachtungen über die Radioaktivität der Atmosphäre im Hochgebirge*, Physikalische Zeitschrift **9** (1908), pp. 907–911.

[**Goc10**] ALBERT GOCKEL, *Luftelektrische Beobachtungen bei einer Ballonfahrt*, Physikalische Zeitschrift **11** (1910), pp. 280–282.

[**Gor02**] DMITRY S. GORBUNOV, PETER G. TINYAKOV, IGOR I. TKACHEV, AND SERGEY V. TROITSKY, *Evidence for a Connection between the γ-Ray and the Highest Energy Cosmic-Ray Emissions by BL Lacertae Objects*, The Astrophysical Journal Letters **577** (2002), pp. L93–L96.

[**Gor04**] PETER W. GORHAM ET AL., *Experimental Limit on the Cosmic Diffuse Ultra-high Energy Neutrino Flux*, Physical Review Letters **93** (2004), p. 041101.

[**Gór05**] KRZYSZTOF M. GÓRSKI ET AL., *HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere*, The Astrophysical Journal **622** (2005), pp. 759–771, http://sourceforge.net/projects/healpix/.

[**Gre56**] KENNETH GREISEN, *The extensive air showers*, in: John G. Wilson (ed.), Progress in Cosmic Ray Physics, Volume III, North-Holland Publishing Company, Amsterdam, 1956.

[**Gre66**] KENNETH GREISEN, *End to the cosmic-ray spectrum?*, Physical Review Letters **16** (1966), pp. 748–750.

[Gri68] Vladimir N. Gribov, *A Reggeon diagram technique*, Journal of Experimental and Theoretical Physics **26** (1968), pp. 414–423.

[Gru14] Trent Grubb, *Personal Communication*, October 2014.

[Gui07] Gene Guillian et al. (Super-Kamiokande Collaboration), *Observation of the anisotropy of 10 TeV primary cosmic ray nuclei flux with the Super-Kamiokande-I detector*, Physical Review D **75** (2007), p. 062003.

[Hal97] Francis Halzen and Enrique Zas, *Neutrino fluxes from active galaxies: a model-independent estimate*, The Astrophysical Journal **488** (1997), pp. 669–674.

[Ham10] Jean-Christophe Hamilton, Benoît Revenu, and Benjamin Rouillé d'Orfeuil, *Coverage & Anisotropy Toolkit (v3.0)*, 2010, http://apcauger.in2p3.fr/Protected/Toolkit/.

[Har10] Gretchen L. H. Harris, Marina Rejkuba, and William E. Harris, *The Distance to NGC 5128 (Centaurus A)*, Publications of the Astronomical Society of Australia **27** (2010), pp. 457–462.

[Hec98] Dieter Heck et al., *CORSIKA: A Monte Carlo Code to Simulate Extensive Air Showers*, Forschungszentrum Karlsruhe, Wissenschaftliche Berichte, FZKA 6019, 1998.

[Hec13] Dieter Heck and Tanguy Pierog, *Extensive Air Shower Simulation with CORSIKA: A User's Guide (Version 74xxx from September 3, 2013)*, Institut für Kernphysik, Karlsruher Institut für Technologie (KIT), 2013.

[Hei14] Philipp Heimann, Markus Risse, Marek Werner, and Alexey Yushkov, *Looking deep: a way to test the existence of primary protons*, 2014, internal note GAP-2014-117.

[Hei15] Philipp Heimann, *Personal Communication*, April 2015.

[Her61] Juan Hersil et al., *Observations of extensive air showers near the maximum of their longitudinal development*, Physical Review Letters **6** (1961), pp. 22–23.

[Hes11] Victor F. Hess, *Über die Absorption der γ-Strahlen in der Atmosphäre*, Physikalische Zeitschrift **12** (1911), pp. 998–1001.

[Hes12] Victor F. Hess, *Über Beobachtungen der durchdringenden Strahlung bei sieben Freiballonfahrten*, Physikalische Zeitschrift **13** (1912), pp. 1084–1091.

[Hes13] Victor F. Hess, *Über Neuerungen und Erfahrungen an den Radiummessungen nach der γ-Strahlenmethode*, Physikalische Zeitschrift **14** (1913), pp. 1135–1141.

[Hil83] Christopher T. Hill, *Monopolonium*, Nuclear Physics B **224** (1983), pp. 469–490.

[Hil84] Anthony M. Hillas, *The origin of ultra-high-energy cosmic rays*, Annual Review of Astronomy and Astrophysics **22** (1984), pp. 425–444.

[Hil05] Anthony M. Hillas, *Can diffusive shock acceleration in supernova remnants account for high-energy galactic cosmic rays?*, Journal of Physics G: Nuclear and Particle Physics **31** (2005), pp. R95–R131.

[Hin95] Mark B. Hindmarsh and Thomas W. B. Kibble, *Cosmic strings*, Reports on Progress in Physics **58** (1995), pp. 477–562.

[Hoe13] Andreas Hoecker et al., *TMVA 4 - Toolkit for Multivariate Data Analysis with ROOT: Users Guide*, CERN, Geneva, 2013.

[Hof12] Werner Hofmann et al. (H.E.S.S. Collaboration), *Largest ever Cherenkov telescope sees first light*, 2012, Press Release, Full Image Credit: H.E.S.S. Collaboration, Stefan Schwarzburg.

[Höf13] Felix Höfer, *Untersuchung der Güte des Lateralverteilungsfits zur Optimierung der Suche nach ultrahochenergetischen Photonen*, Bachelor's Thesis, University of Siegen, Germany, August 2013.

[Hol12] see e.g. Jamie Holder, *TeV gamma-ray astronomy: A summary*, Astroparticle Physics **39-40** (2012), pp. 61–75.

[Hom05] Piotr Homola et al., *Simulation of ultra-high energy photon propagation in the geomagnetic field*, Computer Physics Communications **173** (2005), p. 71.

[**Hom07**] PIOTR HOMOLA ET AL., *Characteristics of geomagnetic cascading of ultra-high energy photons at the southern and northern sites of the Pierre Auger Observatory*, Astroparticle Physics **27** (2007), pp. 174–184.

[**Hoo11**] DAN HOOPER, ANDREW M. TAYLOR, AND SUBIR SARKAR, *Cosmogenic photons as a test of ultra-high energy cosmic ray composition*, Astroparticle Physics **34** (2011), pp. 340–343.

[**Hör03**] JÖRG R. HÖRANDEL, *On the knee in the energy spectrum of cosmic rays*, Astroparticle Physics **19** (2003), pp. 193–220.

[**Hör04**] JÖRG R. HÖRANDEL, *Models of the knee in the energy spectrum of cosmic rays*, Astroparticle Physics **21** (2004), pp. 241–265.

[**Hub05**] HUBBLE SPACE TELESCOPE, *A Giant Hubble Mosaic of the Crab Nebula*, 2005, News Release STScI-2005-37, Full Image Credit: NASA, ESA, J. Hester and A. Loll (Arizona State University).

[**Jam75**] FRED JAMES AND MATTS ROOS, *Minuit - a system for function minimization and analysis of the parameter errors and correlations*, Computer Physics Communications **10** (1975), pp. 343–367.

[**Jui12**] CHARLES C. H. JUI FOR THE TELESCOPE ARRAY COLLABORATION, *Cosmic Ray in the Northern Hemisphere: Results from the Telescope Array Experiment*, Journal of Physics: Conference Series **404** (2012), p. 012037.

[**Kac06**] MICHAEL KACHELRIESS AND PASQUALE D. SERPICO, *The Compton-Getting effect on ultra-high energy cosmic rays of cosmological origin*, Physics Letters B **640** (2006), pp. 225–229.

[**Kac08**] MICHAEL KACHELRIESS, ETIENNE PARIZOT, AND DMITRI V. SEMIKOZ, *The GZK Horizon and Constraints on the Cosmic Ray Source Spectrum from Observations in the GZK Regime*, Journal of Experimental and Theoretical Physics Letters **88** (2008), pp. 553–557.

[**Kac10**] MICHAEL KACHELRIESS, SERGEY OSTAPCHENKO, AND RICARD TOMÀS, *TeV Gamma Rays from Ultrahigh Energy Cosmic Ray Interactions in the Cores of Active Galactic Nuclei: Lessons from Centaurus A*, Publications of the Astronomical Society of Australia **27** (2010), pp. 482–489.

[**Kai82**] ALEXEY B. KAIDALOV AND KAREN A. TER-MARTIROSYAN, *Pomeron as quark-gluon strings and multiple hadron production at SPS-Collider energies*, Physics Letters B **117** (1982), pp. 247–251.

[**Kam58**] KOICHI KAMATA AND JUN NISHIMURA, *The Lateral and the Angular Structure Functions of Electron Showers*, Progress of Theoretical Physics Supplements **6** (1958), pp. 93–155.

[**Kaz01**] DEMOSTHENES KAZANAS AND ARGYRIS NICOLAIDIS, *Cosmic ray "knee": A herald of new physics?*, Proceedings of the 27th International Cosmic Ray Conference (ICRC 2001), Hamburg, Germany, 2001.

[**Kei04**] BIANCA G. KEILHAUER, *Investigation of Atmospheric Effects on the Development of Extensive Air Showers and their Detection with the Pierre Auger Observatory*, Ph.D. thesis, University of Karlsruhe (TH) and Forschungszentrum Karlsruhe, Germany, 2004.

[**Kla00**] SEE E.G. HANS V. KLAPDOR-KLEINGROTHAUS AND KAI ZUBER, *Particle Astrophysics*, revised edition, Institute of Physics Publishing, Bristol and Philadelphia, 2000.

[**Kle99**] SPENCER KLEIN, *Suppression of bremsstrahlung and pair production due to environmental factors*, Reviews of Modern Physics **71** (1999), pp. 1501–1538.

[**Kob99**] MATJAŽ KOBAL, ANDREJ FILIPČIČ, AND DANILO ZAVRTANIK FOR THE PIERRE AUGER COLLABORATION, *Thinning of High-Energy Cosmic-Ray Air-Showers*, Proceedings of the 26th International Cosmic Ray Conference (ICRC 1999), Salt Lake City, USA, 1999.

[**Kol14**] WERNER KOLHÖRSTER, *Messungen der durchdringenden Strahlung bis in Höhen von 9300 m*, Verhandlungen der Deutschen Physikalischen Gesellschaft **16** (1914), pp. 719–721.

[Kot11]   Kumiko Kotera and Angela V. Olinto, *The Astrophysics of Ultrahigh Energy Cosmic Rays*, Annual Review of Astronomy and Astrophysics **49** (2011), pp. 119–153.

[Kue08]   Daniel Kuempel, Karl-Heinz Kampert, and Markus Risse, *Geometry reconstruction of fluorescence detectors revisited*, Astroparticle Physics **30** (2008), pp. 167–174.

[Kue14]   Daniel Kuempel, *Targeted search for photons*, June 2014, talk at the Pierre Auger Collaboration Analysis Meeting in Golden, USA.

[Kus06]   Alexander Kusenko, Jason Schissel, and Floyd W. Stecker, *Interactions of ultrahigh-energy cosmic rays with photons in the galactic center*, Astroparticle Physics **25** (2006), pp. 242–245.

[Lan53a]   Lev Landau and Isaak Pomeranchuk, *Predely primenimosti teorii tormoznogo izluchenija jelektronov i obrazovanija par pri bol'shih jenergijah*, Doklady Akademii Nauk SSSR **92** (1953), pp. 535–536.

[Lan53b]   Lev Landau and Isaak Pomeranchuk, *Jelektronno-lavinnye processy pri sverhvysokih energijah*, Doklady Akademii Nauk SSSR **92** (1953), pp. 735–738.

[Lhe13]   Isabelle Lhenry-Yvon, *Personal Communication*, July 2013.

[Lin04]   Franz Linke, *Luftelektrische Messungen bei zwölf Ballonfahrten*, Abhandlungen der Gesellschaft der Wissenschaften in Göttingen, mathematisch-physikalische Klasse **3** (1904), pp. 1–90.

[Lod03]   Katharina Lodders, *Solar System Abundances and Condensation Temperatures of the Elements*, The Astrophysical Journal **591** (2003), pp. 1220–1247.

[Lon92]   see e.g. Malcolm S. Longair, *High Energy Astrophysics, Volume 1: Particles, photons and their detection*, second edition, Cambridge University Press, Cambridge, 1992.

[Mal13]   Simone Maldera for the Pierre Auger Collaboration, *The AMIGA enhancement of the Pierre Auger Observatory*, Proceedings of the 14th ICATPP Conference on Astroparticle, Particle, Space Physics and Detectors for Physics Applications, Como, Italy, 2013.

[Mar10]   Ioana C. Mariş, Fabian Schüssler, Ralf Ulrich, and Michael Unger, *ADST and Event-Browser Reference Manual: Data Summary Trees and Shower Visualization for Reconstructed and Simulated Auger Events*, 2010, Revision 18212, based on the internal note GAP-2006-081.

[Mas09]   Enrico Massaro et al., *Roma-BZCAT: a multifrequency catalogue of blazars*, Astronomy & Astrophysics **495** (2009), pp. 691–696, and the update for the 5th edition available at http://arxiv.org/abs/1502.07755.

[Mat05]   James Matthews, *A Heitler model of extensive air showers*, Astroparticle Physics **22** (2005), pp. 387–397.

[Mat11]   Hermann-Josef Mathes for the Pierre Auger Collaboration, *The HEAT Telescopes of the Pierre Auger Observatory Status and First Data*, Proceedings of the 32nd International Cosmic Ray Conference (ICRC 2011), Beijing, China, 2011.

[McB81]   Brian McBreen and Colin J. Lambert, *Interactions of high-energy ($E > 5 \times 10^{19}$ eV) photons in the Earth's magnetic field*, Physical Review D **24** (1981), pp. 2536–2538.

[Mig56]   Arkady B. Migdal, *Bremsstrahlung and Pair Production in Condensed Media at High Energies*, The Physical Review **103** (1956), pp. 1811–1820.

[Mol03]   Alexander Molnar and Manfred Simon, *A new Thought on the Energy Dependence of the $^{10}Be/^{9}Be$ Ratio*, Proceedings of the 28th International Cosmic Ray Conference (ICRC 2003), Tsukuba, Japan, 2003.

[Mos07]   Miguel Mostafá for the Pierre Auger Collaboration, *Hybrid Activities of the Pierre Auger Observatory*, Nuclear Physics B (Proceedings Supplement) **165** (2007), pp. 50–58.

[Nat04]   National Radio Astronomy Observatory, *VLA at Sunset 10/27/2004*, 2004, http://images.nrao.edu/Telescopes/VLA/488, Full Image Credit: Image Courtesy of NRAO/AUI and Photo by Amy DuVall.

[Ned37]   SETH H. NEDDERMEYER AND CARL D. ANDERSON, *Note on the Nature of Cosmic-Ray Particles*, The Physical Review **51** (1937), pp. 884–886.

[Nel85]   WALTER R. NELSON, HIDEO HIRAYAMA, AND DAVID W. O. ROGERS, *The EGS4 code system*, SLAC-R-265, SLAC, Stanford, USA, 1985.

[Nie13]   MARCUS NIECHCIOL, MARKUS RISSE, AND MARIANGELA SETTIMO, *A simple parameterization of the LDF shape parameter β for photons*, 2013, internal note GAP-2013-011.

[Nob65]   THE NOBEL FOUNDATION, *Nobel Lectures, Physics 1922-1941*, Elsevier Publishing Company, Amsterdam, 1965.

[Nor95]   COLIN A. NORMAN, DONALD B. MELROSE, AND ABRAHAM ACHTERBERG, *The origin of cosmic rays above $10^{18.5}$ eV*, The Astrophysical Journal **454** (1995), pp. 60–68.

[Ost11]   SERGEY OSTAPCHENKO, *Monte Carlo treatment of hadronic interactions in enhanced Pomeron scheme: QGSJET-II model*, Physical Review D **83** (2011), p. 014018.

[Pac11]   DOMENICO PACINI, *La radiation pénétrante sur la mer*, Le Radium **8** (1911), pp. 307–312.

[Pat04]   MARC PATERNO, *Calculating Efficiencies and Their Uncertainties*, FERMILAB-TM-2286-CD, Fermilab, Batavia, USA, 2004, for additional information, see also Thomas Ullrich and Zhangbu Xu: Treatment of Errors in Efficiency Calculations, arXiv:physics/0701199.

[Pea96]   KARL PEARSON, *Mathematical Contributions to the Theory of Evolution. III. Regression, Heredity, and Panmixia*, Philosophical Transactions of the Royal Society of London A **187** (1896), pp. 253–318.

[Pen65]   ARNO A. PENZIAS AND ROBERT W. WILSON, *A measurement of excess antenna temperature at 4080 Mc/s*, The Astrophysical Journal Letters **142** (1965), pp. L419–L421.

[Per09]   SEE E.G. DONALD H. PERKINS, *Particle Astrophysics*, second edition, Oxford University Press, Oxford, 2009.

[Pes11]   ROBERTO PESCE FOR THE PIERRE AUGER COLLABORATION, *Energy calibration of data recorded with the surface detectors of the Pierre Auger Observatory: an update*, Proceedings of the 32nd International Cosmic Ray Conference (ICRC 2011), Beijing, China, 2011.

[Pet04]   SERGIO PETRERA, *Towards a full simulation of the FD Aperture*, 2004, internal note GAP-2004-015.

[Pie96]   PIERRE AUGER COLLABORATION, *The Pierre Auger Observatory Design Report*, FERMILAB-PUB-96/024, Fermi National Accelerator Laboratory, Batavia, USA, 1996.

[Pie05]   TANGUY PIEROG ET AL., *Dependence of the longitudinal shower profile on the characteristics of hadronic multiparticle production*, Proceedings of the 29th International Cosmic Ray Conference (ICRC 2005), Pune, India, 2005.

[Pie06]   PIERRE AUGER COLLABORATION, *Pierre Auger Observatory Public Website, Image Gallery, Aerial photos of the Observatory*, 2006, http://www.auger.org/media/.

[Pie13a]   TANGUY PIEROG ET AL., *EPOS LHC: test of collective hadronization with LHC data*, DESY-13-125, DESY, Hamburg, 2013.

[Pie13b]   PIERRE AUGER COLLABORATION, *Plans for a Proposal to Upgrade the Pierre Auger Observatory*, 2013, document from October 1, 2013.

[Pie14a]   PIERRE AUGER COLLABORATION, *Documentation of the Auger $\overline{Off\underline{line}}$ Software Framework*, 2014, https://www.auger.unam.mx/AugerWiki/OfflineSoftware.

[Pie14b]   PIERRE AUGER COLLABORATION, SD TRIGGER AND ACCEPTANCE GROUP, *Tanks coordinates and instruction for use*, 2014, http://ipnwww.in2p3.fr/~augers/AugerProtected/AcceptTanks.html.

[Pod04]   PHILIPP PODSIADLOWSKI ET AL., *The rates of hypernovae and gamma-ray bursts: implications for their progenitors*, The Astrophysical Journal Letters **607** (2004), pp. L17–L20.

[Poi37]   SIMÉON D. POISSON, *Recherches sur la probabilité des jugements en matière criminelle et en matière civile*, Bachelier, Paris, 1837.

[Ptu93] Vladimir S. Ptuskin et al., *Diffusion and drift of very high energy cosmic rays in galactic magnetic fields*, Astronomy & Astrophysics **268** (1993), pp. 726–735.

[Ptu97] Vladimir S. Ptuskin, *Transport of high energy cosmic rays*, Advances in Space Research **19** (1997), pp. 697–705.

[Ptu10] Vladimir S. Ptuskin, Vladimir N. Zirakashvili, and Eun-Suk Seo, *Spectrum of galactic cosmic rays accelerated in supernova remnants*, The Astrophysical Journal **718** (2010), pp. 31–36.

[Rac93] Jörg P. Rachen and Peter L. Biermann, *Extragalactic ultra-high energy cosmic rays, I. Contribution from hot spots in FR-II radio galaxies*, Astronomy & Astrophysics **272** (1993), pp. 161–175.

[Rav13] Diego Ravignani for the Pierre Auger Collaboration, *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*, Proceedings of the 33rd International Cosmic Ray Conference (ICRC 2013), Rio de Janeiro, Brazil, 2013.

[Reg35] Erich Regener and Georg Pfotzer, *Vertical Intensity of Cosmic Rays by Threefold Coincidences in the Stratosphere*, Nature **136** (1935), pp. 718–719.

[Rie15] Felix Riehn and ohers, *Charm production in SIBYLL*, manuscript submitted to EPJ Web of Conferences, preprint available under arXiv:1502.06353, 2015.

[Ris01] Markus Risse, Dieter Heck, Johannes Knapp, and Sergey S. Ostapchenko, *EAS Simulations at Auger Energies with CORSIKA*, Proceedings of the 27th International Cosmic Ray Conference (ICRC 2001), Hamburg, Germany, 2001.

[Ris06] Markus Risse et al., *Photon air showers at ultra-high energy and the photonuclear cross-section*, Czechoslovak Journal of Physics **56** (2006), pp. A327–A336.

[Ris07] Markus Risse and Piotr Homola, *Search for ultra-high energy photons using air showers*, Modern Physics Letters A **22** (2007), pp. 749–766.

[Roc47] George D. Rochester and Clifford C. Butler, *Evidence for the existence of new unstable elementary particles*, Nature **160** (1947), pp. 855–857.

[Ros10] Germán Ros et al., *S_b for photon-hadron discrimination*, 2010, internal note GAP-2010-052.

[Ros11] Germán Ros et al., *A new composition-sensitive parameter for ultra-high energy cosmic rays*, Astroparticle Physics **35** (2011), pp. 140–151.

[Ros13] Germán Ros et al., *Improving photon-hadron discrimination based on cosmic ray surface detector data*, Astroparticle Physics **47** (2013), pp. 10–17.

[Rov09] Adrián C. Rovero et al. for the Pierre Auger Collaboration, *Multi-wavelength calibration procedure for the Pierre Auger Observatory Fluorescence Detectors*, Astroparticle Physics **31** (2009), pp. 305–311.

[Sal13] Francisco Salesa Greus, *On the selection of the target sets for the targeted search for point sources of EeV neutrons paper*, 2013, internal note GAP-2013-117.

[Sch32] Bernhard Schmidt, *Ein lichtstarkes komafreies Spiegelsystem*, Mitteilungen der Hamburger Sternwarte in Bergedorf **7** (1932), pp. 15–17.

[Sch90] Robert E. Schapire, *The strength of weak learnability*, Machine Learning **5** (1990), pp. 197–227.

[Sch06] Viviana Scherini for the Pierre Auger Collaboration, *Performance of the Pierre Auger Fluorescence Detector at the highest energies*, Proceedings of the 20th European Cosmic Ray Symposium (ECRS 2006), Lisbon, Portugal, 2006.

[Sch13] Alexander Schulz for the Pierre Auger Collaboration, *Measurement of the energy spectrum of cosmic rays above 3×10^{17} eV with the Pierre Auger Observatory*, Proceedings of the 33rd International Cosmic Ray Conference (ICRC 2013), Rio de Janeiro, Brazil, 2013.

[Sch14] ALEXANDER SCHULZ, *Personal Communication*, September 2014.

[Sek85] YATARO SEKIDO AND HARRY ELLIOT (EDS.), *Early history of cosmic ray studies*, D. Reidel Publishing Company, Dordrecht, 1985.

[Set08] MARIANGELA SETTIMO ET AL., *Lateral Trigger Probability functions for photon primary at low energies*, 2008, internal note GAP-2008-009.

[Set11] MARIANGELA SETTIMO FOR THE PIERRE AUGER COLLABORATION, *An update on a search for ultra-high energy photons using the Pierre Auger Observatory*, Proceedings of the 32nd International Cosmic Ray Conference (ICRC 2011), Beijing, China, 2011.

[Set13] MARIANGELA SETTIMO FOR THE PIERRE AUGER COLLABORATION, *Search for ultra-High Energy Photons with the Pierre Auger Observatory*, Proceedings of the International Conference on the Structure and the Interactions of the Photon (Photon 2013), Paris, France, 2013.

[Shi02] KENJI SHINOZAKI ET AL., *Upper limit on gamma-ray flux above 10^{19} eV estimated by the Akeno Giant Air Shower Experiment*, The Astrophysical Journal Letters **517** (2002), pp. L117–L120.

[Sim83] JOHN A. SIMPSON, *Elemental and Isotopic Composition of the Galactic Cosmic Rays*, Annual Review of Nuclear and Particle Science **33** (1983), pp. 323–381.

[Sta10] SEE E.G. TODOR STANEV, *High Energy Cosmic Rays*, second edition, Springer-Verlag, Berlin, 2010.

[Suo09] TIINA SUOMIJÄRVI FOR THE PIERRE AUGER COLLABORATION, *Performance and operation of the Surface Detector of the Pierre Auger Observatory*, Proceedings of the 31st International Cosmic Ray Conference (ICRC 2009), Łódź, Poland, 2009.

[Swa33] WILLIAM F. G. SWANN, *A Mechanism of Acquirement of Cosmic-Ray Energies by Electrons*, The Physical Review **43** (1933), pp. 217–220.

[Swo01] SIMON P. SWORDY, *The energy spectra and anisotropies of cosmic rays*, Space Science Reviews **99** (2001), pp. 85–94.

[Tak98] MASAHIRO TAKEDA ET AL., *Extension of the Cosmic-Ray Energy Spectrum beyond the Predicted Greisen-Zatsepin-Kuz'min Cutoff*, Physical Review Letters **81** (1998), pp. 1163–1166.

[Tho14] ERIK THORN, *Studies on the Event Selection for the Search for ultra-high energetic Photons at the Pierre Auger Observatory*, Bachelor's Thesis, University of Siegen, Germany, May 2014.

[Ton15] AURELIO TONACHINI ET AL. (PIERRE AUGER COLLABORATION), *LIDAR Information Database*, 2015, http://www.auger.to.infn.it/lidar/.

[Tue13] MATIAS J. TUEROS FOR THE PIERRE AUGER COLLABORATION, *Estimate of the non-calorimetric energy of showers observed with the fluorescence and surface detectors of the Pierre Auger Observatory*, Proceedings of the 33rd International Cosmic Ray Conference (ICRC 2013), Rio de Janeiro, Brazil, 2013.

[Ung04] MICHAEL UNGER, *EventGenerator.cc*, 2004, Documentation of the Auger Off̅line Software Framework, version 2.9.1.

[Ung08] MICHAEL UNGER ET AL., *Reconstruction of longitudinal profiles of ultra-high energy cosmic ray showers from fluorescence and Cherenkov light measurements*, Nuclear Instruments and Methods in Physics Research A **588** (2008), pp. 433–441.

[Veb13a] DARKO VEBERIČ, *Maps of the Pierre Auger Observatory*, 2013, https://web.ikp.kit.edu/darko/auger/auger-array/.

[Veb13b] DARKO VEBERIČ, MARKUS ROTH, AND HANS DEMBINSKI FOR THE OFF̅LINE TEAM, *Off̅line Reference Manual SD Reconstruction*, 2013, document version 0.8, based on the internal note GAP-2005-035.

[Ven97] Aparna Venkatesan, M. Coleman Miller, and Angela V. Olinto, *Constraints on the production of ultra-high-energy cosmic rays by isolated neutron stars*, The Astrophysical Journal **484** (1997), pp. 323–328.

[Vér06] Marie-Paule Véron-Cetty and Philippe Véron, *A catalogue of quasars and active nuclei: 12th edition*, Astronomy & Astrophysics **455** (2006), pp. 773–777.

[Vér10] Marie-Paule Véron-Cetty and Philippe Véron, *A catalogue of quasars and active nuclei: 13th edition*, Astronomy & Astrophysics **518** (2010), p. A10.

[Ver13] Valerio Verzi for the Pierre Auger Collaboration, *The Energy Scale of the Pierre Auger Observatory*, Proceedings of the 33rd International Cosmic Ray Conference (ICRC 2013), Rio de Janeiro, Brazil, 2013.

[Vie95] Mario Vietri, *The acceleration of ultra-high-energy cosmic rays in gamma-ray bursts*, The Astrophysical Journal **453** (1995), pp. 883–889.

[Vos07] Helge Voss, Andreas Hoecker, Jörg Stelzer, and Fredrik Tegenfeldt, *TMVA, the Toolkit for Multivariate Data Analysis with ROOT*, Proceedings of Science **ACAT** (2007), p. 040.

[Wax95] Eli Waxman, *Cosmological Gamma-Ray Bursts and the Highest Energy Cosmic Rays*, Physical Review Letters **75** (1995), pp. 386–389.

[Wax06] Eli Waxman, *Gamma-ray bursts: Potential sources of ultra high energy cosmic-rays*, Nuclear Physics B (Proceedings Supplement) **151** (2006), pp. 46–53.

[Wei13] Klaus Weidenhaupt for the Pierre Auger Collaboration, *The Auger Engineering Radio Array*, Acta Polytechnica (Supplement) **53** (2013), pp. 825–828.

[Wei14] Klaus Weidenhaupt for the Pierre Auger Collaboration, *Radio Detection of Cosmic Rays at the Auger Engineering Radio Array*, Proceedings of the 3rd International Conference on Technology and Instrumentation in Particle Physics (TIPP 2014), Amsterdam, the Netherlands, 2014.

[Wer13] Marek Werner, *Suche nach neuen Observablen zur Identifizierung primärer Photonen in der kosmischen Strahlung*, Bachelor's Thesis, University of Siegen, Germany, March 2013.

[Wig00] Richard Wigmans, *On Big Bang Relics, the Neutrino Mass and the Spectrum of Cosmic Rays*, Nuclear Physics B (Proceedings Supplement) **85** (2000), pp. 305–310.

[Wil01] Charles T. R. Wilson, *On the Ionisation of Atmospheric Air*, Proceedings of the Royal Society of London A **68** (1901), pp. 151–161.

[Wil11] Charles T. R. Wilson, *On a Method of making Visible the Paths of Ionising Particles through a Gas*, Proceedings of the Royal Society of London A **85** (1911), pp. 285–288.

[Wul07] Theodor Wulf, *Ein neues Elektrometer für statische Ladungen*, Physikalische Zeitschrift **8** (1907), pp. 246–248.

[Wul09] Theodor Wulf, *Über den Ursprung der in der Atmosphäre vorhandenen γ-Strahlung*, Physikalische Zeitschrift **10** (1909), pp. 997–1003.

[Wul10] Theodor Wulf, *Beobachtungen über Strahlung hoher Durchdringungsfähigkeit auf dem Eiffelturm*, Physikalische Zeitschrift **11** (1910), pp. 811–813.

[You11] Patrick Younk, *SD Composition Indicator for Hybrid Events*, 2011, internal note GAP-2011-095.

[Yus15] Alexey Yushkov, *Personal Communication*, April 2015.

[Zat66] Georgiy T. Zatsepin and Vadim A. Kuz'min, *Upper limit of the spectrum of cosmic rays*, Journal of Experimental and Theoretical Physics Letters **4** (1966), pp. 78–80.

List of Acronyms and Abbreviations

AB	Analog Board
ADC	Analog-to-Digital Converter
ADST	Advanced Data Summary Tree
AERA	Auger Engineering Radio Array
AGASA	Akeno Giant Air Shower Array
AGN	Active Galactic Nucleus
AMIGA	Auger Muons and Infill for the Ground Array
AMS-02	Alpha Magnetic Spectrometer
APF	Aerosol Phase Function Monitor
a.s.l.	above sea level
BDT	Boosted Decision Tree
CCD	Charge-Coupled Device
CDAS	Central Data Acquisition System
CIC	Constant Intensity Cut
CLF	Central Laser Facility
CMB	Cosmic Microwave Background
CNO	Carbon, Nitrogen, Oxygen
CORSIKA	Cosmic Ray Simulations for KASCADE
EBL	Extragalactic Background Light
EGS	Electron Gamma Shower
EPOS	Energy conserving quantum mechanical multiple scattering approach, based on Partons (parton ladders) Off-shell remnants, and Spliting of parton ladders
FADC	Flash Analog-to-Digital Converter
FD	Fluorescence Detector
Fermi LAT	Fermi Large Area Telescope
FLT	First-Level Trigger
FLUKA	Fluktuierende Kaskade

FPGA	Field-Programmable Gate Array
FR-II	Fanaroff-Riley Class II
FRAM	(F/Ph)otometric Robotic Atmospheric Monitor
GDAS	Global Data Assimilation System
Geant4	Geometry and Tracking 4
GOES	Geostationary Operational Environmental Satellite
GPS	Global Positioning System
GRB	Gamma-Ray Burst
GUT	Grand Unified Theory
GZK	Greisen, Zatsepin, Kuz'min
HAM	Horizontal Attenuation Monitor
HEALPix	Hierarchical Equal Area isoLatitude Pixelization
HEAT	High Elevation Auger Telescopes
H.E.S.S.	High Energy Stereoscopic System
HiRes	High Resolution Fly's Eye Detector
IACT	Imaging Air Cherenkov Telescope
IR	Infrared
ISOMAX	Isotope Magnet Experiment
KASCADE	Karlsruhe Shower Core and Array Detector
LDF	Lateral Distribution Function
LED	Light-Emitting Diode
LHC	Large Hadron Collider
Lidar	Light Detection and Ranging
LPDA	Logarithmic Periodic Dipole Antenna
LPM	Landau, Pomeranchuk, Migdal
MC	Monte Carlo
MVA	Multivariate Analysis
NDF	Number of Degrees of Freedom
NGC	New General Catalogue of Nebulae and Clusters of Stars
NKG	Nishimura, Kamata, Greisen
PC	Personal Computer
PMT	Photomultiplier Tube
QCD	Quantum Chromodynamics
QED	Quantum Electrodynamics
QGSJET	Quark Gluon String Jet
SD	Surface Detector
SDP	Shower-Detector Plane
SHDM	Super-Heavy Dark Matter
SLT	Second-Level Trigger
SNR	Supernova Remnant
SQL	Structured Query Language
Super-Kamiokande	Super-Kamioka Nucleon Decay Experiment
TA	Telescope Array

TD	Topological Defect
Tibet ASγ	Tibet Air Shower Array
TLT	Third-Level Trigger
UHE	Ultra-High-Energy
UHECR	Ultra-High-Energy Cosmic Ray
URB	Universal Radio Background
UTC	Coordinated Universal Time
UV	Ultraviolet
VAOD	Vertical Aerosol Optical Depth
VCV	Véron-Cetty & Véron
VEM	Vertical Equivalent Muon
VLA	Very Large Array
VLT	Very Large Telescope
XLF	eXtreme Laser Facility
XML	Extensible Markup Language
XMM-Newton	X-Ray Multi-Mirror Mission - Newton

Danksagung

An dieser Stelle möchte ich die Gelegenheit nutzen und den vielen Personen danken, ohne deren Unterstützung diese Arbeit kaum möglich gewesen wäre.

Zuerst möchte ich meinem Doktorvater Prof. Dr. Markus Risse danken: Dafür, dass ich die letzten vier Jahre in diesem spannenden Teilgebiet der Physik arbeiten durfte. Für die großartige Betreuung, die ständige Unterstützung und die motivierenden Worte, wenn es mal nicht so lief wie geplant. Für die vielen Diskussionen, aus denen ich so viel mitnehmen konnte. Und nicht zuletzt auch für die Bereitschaft, meine vielen Reisen zu Collaboration Meetings, Konferenzen und Schulen zu finanzieren.

Einen besonderen Dank möchte ich auch Prof. Dr. Claus Grupen aussprechen, der die oftmals undankbare Aufgabe des Zweitgutachters übernommen hat, sowie Prof. Dr. Karl-Heinz Kampert und Prof. Dr. Thomas Mannel für die Bereitschaft in der Promotionskommission mitzuwirken.

Ein großer Dank gebührt auch Dr. Mariangela Settimo für die Co-Betreuung dieser Arbeit in der Anfangsphase, für die vielen Anregungen, wie man die Analyse noch weiter verbessern könnte, und vor allem für die geduldige Beantwortung meiner vielen Fragen, selbst nachdem sie Siegen in Richtung Paris verlassen hatte. Erwähnt werden sollte auch Dr. Patrick Younk, der während seiner Zeit in Siegen die Observable F_γ „erfunden" und damit den Anstoß für diese Arbeit gegeben hat.

Nicht unterschätzt werden sollte der Beitrag, den Dr. Hendrik Czirr, Dr. Uwe Fröhlich, Dr. Bakul Gaur und Dr. Alexey Yushkov an dieser Arbeit hatten. Vielen Dank für das Korrekturlesen des gesamten Textes, vor allem in den letzten, besonders stressigen Wochen kurz vor der Abgabe der Arbeit, und für die vielen hilfreichen Kommentare, die dabei geholfen haben, meine Gedanken in eine lesbare Fassung zu bringen.

Bei der gesamten Arbeitsgruppe Experimentelle (Astro-)Teilchenphysik an der Universiät Siegen möchte ich mich bedanken für die angenehme Arbeitsatmosphäre und dafür, dass jeder einzelne bei Fragen und Problemen immer mit Rat und Tat bereitstand. Besonders hervorheben möchte ich, zusätzlich zu den zuvor Genannten, M.Sc. Alexander Aab, Dr. Thomas Bäcker, B.Sc. Christian Dehn, M.Sc. RWTH Philipp Heimann, M.Sc. Bertha Heimel, B.Sc. Frederik Lauber, M.Sc. Amir Noori Shirazi, Dr. Michael Pontz, B.Sc. Sebastian Sonntag sowie Dr. Andreas