Search for Ultra-High-Energy Photons with the Pierre Auger Observatory

Dissertation
zur Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften
der
Fakultät Mathematik und Naturwissenschaften
Fachbereich Physik
der Bergischen Universität Wuppertal

BERGISCHE
UNIVERSITÄT
WUPPERTAL

vorgelegt von

Philipp Papenbreer

Wuppertal, den 2.10.2020

The PhD thesis can be quoted as follows:

urn:nbn:de:hbz:468-20201203-114928-2
[http://nbn-resolving.de/urn/resolver.pl?urn=urn%3Anbn%3Ade%3Ahbz%3A468-20201203-114928-2]

DOI: 10.25926/h3hs-t241
[https://doi.org/10.25926/h3hs-t241]

Contents

Abstract

In this thesis, a search for photons at ultra-high energies is presented. A discovery of these particles would greatly advance the understanding of the origin of ultra-high energy cosmic rays and would have multiple implications for high-energy particle physics. Upper limits to the flux of these particles, on the other hand, would constrain the parameter space of astrophysical scenarios concerning the primary composition at the cosmic ray sources as well as certain models of the origin of cosmic rays. In this work, 13.5 years of data measured with the Surface Detector of the Pierre Auger Observatory have been analyzed. In preparation for this, many improvements to the quality of the measurement data were made. Since multiple photon-like events were identified by the analysis, the compatibility of these event candidates with a photon as well as a background hypothesis has been tested. Taking the conservative approach of assuming the background hypothesis, the world's strongest limits to the diffuse photon flux and fraction have been set.

Zusammenfassung

Die vorliegende Arbeit befasst sich mit der Suche nach Photonen in der ultrahochenergetischen kosmischen Strahlung. Eine Entdeckung dieser Teilchen würde wertvolle Hinweise über den Ursprung der kosmischen Strahlung liefern und böte darüber hinaus vielfältige Implikationen für die Teilchenphysik bei hohen Energien. Obergrenzen für den Fluss dieser Teilchen begrenzen den Parameterraum, sowohl von astrophysikalischen Szenarien bezüglich der Komposition an ihren Quellen, als auch bestimmter Modelle für den Ursprung der kosmischen Strahlung. 13.5 Jahre Messdaten, die mit dem Oberflächendetektor des Pierre-Auger-Observatoriums aufgenommen wurden, sind in dieser Arbeit analysiert worden. In Vorbereitung dazu sind einige Verbesserungen der Messdatenqualität unternommen worden. Da in der Analyse mehrere photonartige Ereignisse gefunden wurden, wurde die Kompatibilität dieser Kandidatenereignisse mit einer Photonhypothese und einer Untergrundhypothese untersucht. Mit dem konservativen Ansatz einer Untergrundhypothese sind die weltweit strengsten Obergrenzen auf den Anteil von Photonen an der kosmischen Strahlung und auf den diffusen Photonenfluss gesetzt worden.

Abbreviations

X_{max} atmospheric depth of the shower maximum. 5, 6, 13, 17, 19, 22, 36, 71–73, 96, 97, 100, 102, 107, 112, 114

ADST Advanced Data Summary Tree. 35, 70

AGN Active Galactic Nuclei. 7

C converted. 68, 70, 71, 112, 115, 116, 118, 119

C+UC converted and unconverted. 116

CMB Cosmic Microwave Background. 4, 9

CORSIKA COsmic Ray SImulations for KAscade. 21–23, 36, 45, 67, 68, 72, 77, 81, 109, 114, 137

CR Cosmic Ray. 3, 4, 6–8, 10, 20, 120, 121

DSA Diffusive Shock Acceleration. 8

EBL Extra-Galactic Background Light. 9

FADC Flash Analog to Digital Converter. 27, 28, 30, 31, 36, 37, 40, 41, 43, 44, 46, 49–51, 55–57, 59–64, 69, 93, 97, 143

FD Fluorescence Detector. 6, 25, 26, 33, 35, 96, 97

GRB Gamma-Ray Bursts. 7

GZK Greisen-Zatsepin-Kuzmin. 1, 2, 4, 6, 9, 10, 20, 120, 121

HG high-gain. 27, 61, 70, 93, 144

ICRC International Cosmic Ray Conference. 2, 67, 88, 95, 118

KS Kolmogorov-Smirnov. 50

LDF Lateral Distribution Function. 33, 34, 36–38, 42, 77, 81, 84, 96, 97, 114

LG low-gain. 27, 37, 40, 61, 70, 93, 144, 145

LPM Landau-Pomeranchuk-Migdal. 14–17

MC Monte Carlo. 52, 67, 70–72, 77, 78, 80, 81, 84, 85, 87, 89, 92, 109, 112, 114, 115, 118, 152

MCMC Markov Chain Monte Carlo. 53, 54

MoPS Multiplicity of Positive Steps. 29–31, 43–47, 49, 50, 52–54, 61

NKG Nishimura-Kamata-Greisen. 33, 34, 36, 37, 42

PAO Pierre Auger Observatory. 2, 5, 6, 10, 15, 26, 27, 67, 122, 124

PCA Principal Component Analysis. 2, 67–70, 79, 81, 84, 85, 87–89, 91–94, 96, 109, 111, 112, 117, 120, 149

PE trace Photoelectron Trace. 48–51

PMT Photomultiplier tube. 26–30, 36, 37, 40, 46–50, 55–65, 93, 96, 97, 123

RMS Root Mean Square. 59, 62

S1000 Signal at 1000m distance from the shower axis. 34, 71, 72, 75, 77, 78, 80, 81, 118, 153

SD Surface Detector. 2, 5, 6, 25–29, 32–34, 36, 37, 41, 43–46, 54, 55, 63, 67–70, 72, 78, 79, 81, 84, 123, 124

SHDM Super-Heavy Dark Matter. 9

Th Threshold. 29

TOT Time Over Threshold. 29–31

ToTd Time over Threshold deconvoluted. 29, 30, 43–46, 52, 54, 59, 60

UC unconverted. 68, 70–72, 74, 77, 81, 84, 112, 115, 116, 118

UHECR ultra-high-energy cosmic rays. 1, 2, 9, 10, 18

VEM Vertical Equivalent Muon. 27, 28, 40, 56, 96, 97

List of Figures

List of Tables

Chapter 1

Introduction

The study of ultra-high-energy cosmic ray (UHECR) physics deals with the interactions of particles at highest energies, i.e. above 10^{18} eV. These particles naturally arrive at Earth after being accelerated in close proximity to some of the most extreme objects in the universe like active galactic nuclei or (multiple) supernova explosions. Naturally, the potential of discovery through the measurements of these particles is considerable as their energies are unmatched by any particles scientists can accelerate on Earth. However, making measurements at these energies is difficult and requires large detectors and complex analysis procedures. This is largely a consequence of the low flux of UHECR, which is of the order of a few particles per square kilometer per year. Upon entering Earth's atmosphere, cosmic ray particles interact and then distribute their energy to a multitude of secondary particles creating a so-called extensive air shower. These secondary particles propagate through the atmosphere nearly at the speed of light, spreading the primary particle's energy over an area of several square kilometers. UHECR experiments are designed to measure these secondary particles and to reconstruct the energy, arrival direction, particle type, and further properties of the primary particle through the use of complex algorithms.

The search for ultra-high-energy photons, i.e. photons with energies above 10^{18} eV, is a topic of current research with multiple implications for physics. In the past, upper limits to the photon primary cosmic ray fluxes have almost ruled out many models for the origin of UHECR that are not based on acceleration mechanisms[1] but on the decay of super-heavy particles, since the flux levels predicted by these models have not been observed. A measurement of UHECR photons would provide new information on the origin of cosmic rays and the physical processes which occur during their propagation to Earth. In particular, since photons are unaffected by magnetic fields, their arrival directions would point directly towards their points of origin.

A flux suppression has been observed in the all-particle energy spectrum of cosmic rays above 5×10^{19} eV[1]. Currently, this is theorized to either be attributable to the maximum energy of the cosmic ray acceleration mechanisms being reached or to be caused by a photo-pion production process, the GZK effect, which would create ultra-high-energy photons. Recent results of the Pierre Auger Collaboration favor the former cause of the flux suppression, however the Telescope Array Collaboration observes a proton dominated composition at highest energies which suggests the latter. An obser-

[1] so-called Top-Down models

vation of a primary photon flux could provide an independent confirmation of the GZK process. Furthermore, an observation of just a small flux of primary photons could be used to greatly decrease the uncertainties of photon-air cross sections at highest energies.

The Pierre Auger Observatory is the world's largest detector for cosmic ray studies and thus naturally well suited for studies at highest energies, i.e. lowest fluxes. In this work, 13.5 years of measurement data taken with the Surface Detector (SD) of the Pierre Auger Observatory has been used to search for primary photons in the UHECR flux. The main goal of this study was to find a flux of primary photons, and, if none were found, to provide new and stringent upper limits to the cosmic ray photon flux.

The main part of this thesis begins with Chapter 2, which provides a brief introduction to cosmic ray physics, focused particularly on the aspects important to the search for primary photons which is presented in this thesis. Chapter 3 introduces the Pierre Auger Observatory, again with a strong focus on the components that are used in the analyses of this work.

The description of the analysis begins with the trigger studies presented in Chapter 4. A procedure is introduced which can improve the trigger algorithms that select air shower signals from the background data in the SD stations of the observatory. Since photon studies are limited by the small number of participating detector stations, an improvement to these triggers can greatly increase the efficiencies of photon analyses.

In Chapter 5, new algorithms are introduced which improve the data quality of the SD: some high-level analyses, like the photon analysis presented in this work, rely on very clean detector signals and signal shapes. However, in a large and complex experiment, like the SD, having a small fraction of the detector components produce odd signals due to problems with the detector electronics is unavoidable. To counter this, in this work automated filtering mechanisms which find these problematic components and remove their data from the analysis were developed and implemented in the data reconstruction procedure.

The search for ultra-high-energy air showers produced by primary photons is presented in Chapter 6. A burn-sample consisting of 2% of the available data was used in combination with photon air shower simulations to prepare a Principal Component Analysis which was then applied to the full data set. New limits to the diffuse photon flux were then derived which have been presented at the International Cosmic Ray Conference 2019 [2] and are being prepared for a journal publication by the Pierre Auger Collaboration. The primary nature of the candidate events in the lowest energy bin has been studied to check compatibility with a proton primary background hypothesis.

In Chapter 7, a summary of the results of this thesis is given and an outlook for possible future works based on the performed studies is presented.

Chapter 2

Cosmic Rays

In this Chapter a summarizing introduction to cosmic ray physics will be given, with a special focus on photon primary air showers. Cosmic rays and the main properties of their flux are discussed in Sec. 2.1. Cosmic ray sources and propagation are discussed in Sec. 2.2 and Sec. 2.3. Air showers in general are introduced in Sec. 2.4, while photon primary air showers are discussed in more detail in Sec. 2.5. Since a primary goal of this thesis is to search for a flux of ultra-high-energy photons, the published limits from the most current ultra-high energy photon searches are discussed in Sec. 2.6. Finally, in Sec. 2.7, the methods and tools for simulating air showers are presented.

2.1 Introduction

The atmosphere of Earth is exposed to a flux of high energy particles called Cosmic Rays (CR). These particles were discovered in the year 1912 when Victor Hess measured the ionization profile of Earth's atmosphere. He correctly reasoned that the increase in ionization with altitude is likely to be explained by a new type of radiation which does not originate on Earth [3]. In 1936, he was awarded the Nobel Prize in Physics for this discovery [4].

Many insights concerning the field of astroparticle physics have been gained since its emergence at the time, yet the primary nature and origin of cosmic rays is still a subject of ongoing research. The current status of this research will be briefly summarized in this Chapter. Special attention will be paid to the most important aspects which concern an experimental search for cosmic ray photons at the highest energies.

2.1.1 Energy Spectrum

The flux of CR, $J(E)$, for a given energy, E, closely follows a power law of the form $J(E) \propto E^{-\gamma}$ indicating their non-thermal origin. This spectrum has been measured over 11 orders of magnitude in energy and 32 orders of magnitude in flux. The combined spectrum measurements of multiple CR experiments are shown in Fig. 2.1. In order to make spectral features more noticeable, a $E^{2.7}$ scaling has been applied to the measured fluxes. Within this huge energy range there are a few notable points where the spectral index changes quickly. Between 10^{15} eV and 10^{16} eV, the spectral index, γ, changes from about 2.7 to 3.1, the so-called *knee* of the spectrum [5, 6].

This knee can likely be attributed either to a rigidity dependent maximum energy of acceleration processes, or to diffusive propagation of CR within the galaxy with a rigidity dependent escape probability [7]. The cutoff position in both cases depends on the charge of the primary particle and therefore its type. The knee, at about 4×10^{15} eV, has been shown to be caused by a drop off of the lightest component of the composition (protons) [8]. Correspondingly, a second knee-like structure has been detected at about 8×10^{16} eV where the knee of iron primaries would be expected if the spectral feature was caused by a rigidity dependent effect. This indicates a transition from a lighter to a heavier composition between both knee-like features [9]. Furthermore, it should be noted that most popular models for the knee-like structures indicate a transition of galactic to extragalactic origin of the CR primary particles which begins in this energy region.

At about 3×10^{18} eV, the energy spectrum flattens with γ going from ≈ 3.2 to ≈ 2.7. This transition, called the *ankle*, could either be explained by a new population of high-energy particles beginning to dominate at these energies or by a pile-up due to the suppression of the flux at higher energies. The possible additional population is assumed to be due to extragalactic CR sources becoming predominant in this energy region [10]. The model of a CR flux suppression at higher energies is based on the assumption of energy losses due to interactions with the Cosmic Microwave Background (CMB) [11], the so-called GZK effect (see Sec. 2.3), or photo disintegration of heavy primaries. The GZK effect requires a light CR composition at the highest energies and is thus challenged by the recent results of the Pierre Auger Collaboration (see Sec. 2.1.2). Furthermore, measurements of the equatorial dipole of CR fluxes support the model of a transition from galactic to extragalactic CR sources in the region between the second knee and the ankle [12].

The spectrum becomes much steeper in the *cutoff* energy region above the ankle. This feature could either be created by energy losses in high-energy interactions of protons and heavier nuclei with the astrophysical photon background or be caused by the maximum energies reached in CR acceleration sources. Recent measurements have shown that in this region between the ankle and the cutoff, a broken power law is not sufficient to describe the data and more complex models might be needed [13].

2.1.2 Composition

At CR energies on the GeV scale, the chemical composition is well known to be close to the abundances in the solar system with the exception of a few elements which are over-abundant in CR. These elements (Li, Be, B), are only created in spallation interactions of heavier CR (C, N, O) with the interstellar media [24] and otherwise would be almost entirely absent from the solar system.

At higher energies, where the flux of CRs is so low that they can only be measured with large experiments that rely on indirect measurement techniques (see Chapter 3), the composition is much more difficult to determine. A fit to the cosmic ray composition above 100 GeV with four mass groups, from hydrogen to iron [25], is shown in Fig. 2.2. For most of the spectrum, the flux is dominated by light elements (H, He) with notable

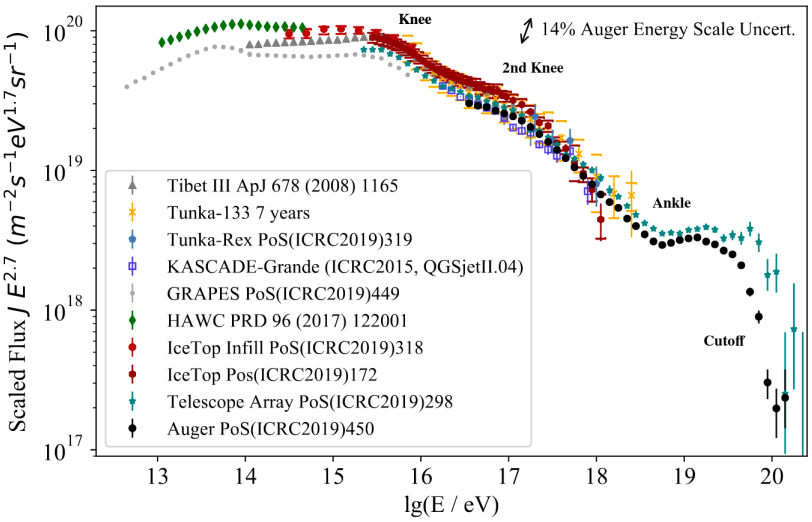


Figure 2.1: Cosmic ray all-particle energy spectrum multiplied by $E^{2.7}$ to make features more visible. The effect of the uncertainty on the absolute energy is shown for the Pierre Auger Observatory data. Plot modified from [14] (Data from [13, 15, 16, 17, 18, 19, 20, 21, 22, 23]).

exceptions around the second knee and at the highest energies. For energies above 10^{17} eV, the primary particle composition is mostly studied using the depth of the air shower maximum, X_{max}, in the atmosphere [26, 27, 28] which is highly correlated with the logarithmic primary mass [29]:

$$\langle X_{\mathrm{max}} \rangle = \alpha(\ln E - \langle \ln A \rangle) + \beta, \tag{2.1}$$

where $\langle \ln A \rangle$ is the mean logarithmic primary mass and α and β are parameters depending on hadronic interactions.

With the air fluorescence technique, X_{max} can be measured directly. However, at highest energies where the statistics of fluorescence measurements become low, a study with the Surface Detector of the Pierre Auger Observatory was performed using the Δ-technique [27], which is also used in this work (see 3.5.2). The measurements of Δ_{Leeds} have been used to estimate X_{max} from Surface Detector data and thus infer the primary particle composition. The analysis results are shown in Fig. 2.3.

In the ankle region of 3×10^{18} eV, a mostly protonic composition is indicated while towards the highest energies the composition gets heavier. Furthermore, by measuring the spread of the X_{max} distribution, it has been shown that the composition becomes less mixed towards these highest energies favoring a rather low component of light elements [28]. While the X_{max} measurements are compatible with a similar study from the Telescope Array Collaboration (TA) [30], the results of TA do not exclude a purely light composition beyond the ankle. Recent studies combining the spectrum and composition

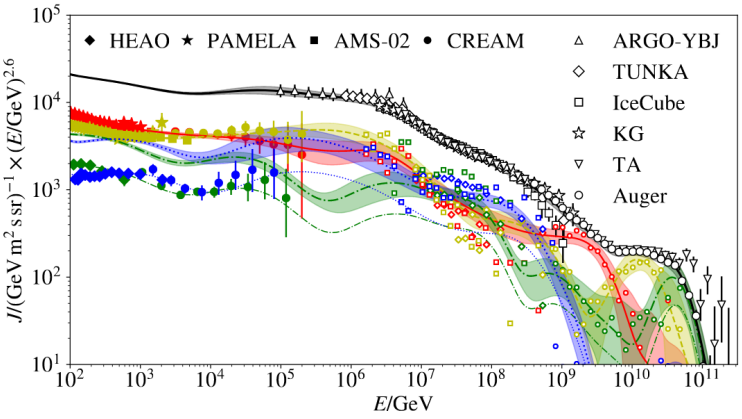


Figure 2.2: Data-driven model of the CR flux of four different mass groups using a Global Spline Fit: protons (red), helium (yellow), oxygen (green), and iron (blue). The all-particle flux is shown in black [25].

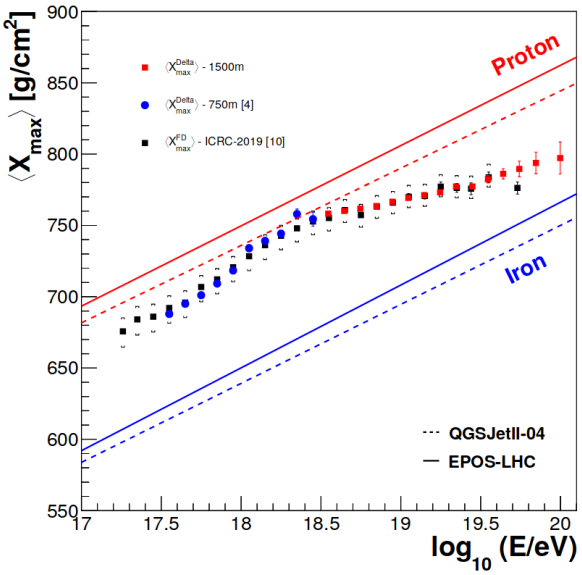


Figure 2.3: Composition measurements at highest energies from the Pierre Auger Observatory. SD measurements of Δ_{Leeds} have been used to calculate the mean X_{max} for different energy bins which is shown by the red and blue points. The results are compared to direct X_{max} measurements from the FD in black. The red and blue lines indicate X_{max} values for proton and iron primaries inferred from different hadronic interaction models [27].

measurements of multiple experiments favor a sub-dominant protonic component even at the highest energies and allow more than 10 % of CRs above 5×10^{19} eV to be protons with all hadronic interaction models [31].

The fraction of proton cosmic rays at the sources has a direct influence on the expected photon flux from the GZK effect (see Sec. 2.3): a heavier composition at the

sources would cause a decrease of the photon flux which aims to be measured in this work. Strong limits on the photon flux would constrain the parameter space of astrophysical scenarios concerning the composition at the CR sources.

2.2 Sources

One of the fundamental open questions about cosmic rays is which astrophysical objects are able to accelerate particles to the high energies observed and what are the processes involved in their acceleration. It is currently mostly thought that gradual acceleration mechanisms take place with a process called Fermi-acceleration being the most prominent (see Sec. 2.2.1). For this type of acceleration to occur, the particles must be confined within the acceleration region; i.e. the Larmor radius of the CRs must be smaller than the acceleration region's size L. This defines a maximum energy for a source class [32] of

$$E/(10^{15}\text{ eV}) < B[\mu\text{G}] \cdot L[\text{pc}] \cdot Z \cdot \beta. \tag{2.2}$$

Here, B is the magnetic field strength in the acceleration region, Z is the CR charge number, and β is the relativistic velocity of the scattering centers in the acceleration region, which are vital to the acceleration process (see Sec. 2.2.1).

In Fig. 2.4, candidates for astrophysical sources are shown plotted by their respective size and magnetic field strength. This so-called Hillas plot illustrates that the number of astrophysical objects which are able to accelerate cosmic rays to the highest energies is limited with Active Galactic Nuclei (AGN), Gamma-Ray Bursts (GRB), neutron stars, and galaxy clusters (via galaxy merger shocks) being the most prominent examples.

2.2.1 Fermi Acceleration

Charged particles can be accelerated to high energies by scattering off moving magnetized gas clouds. Basically, each gas cloud works as a moving *mirror* reflecting the particle and accelerating it in the case of a head-on collision while decelerating it in the case of a tail-on collision. Since head-on collisions are more likely for a moving particle than tail-on collisions, statistically an overall energy gain is expected [34], which depends on the relativistic velocity of the gas clouds, β, as

$$\frac{\delta E}{E} \propto \beta^2, \tag{2.3}$$

where δE is the energy gain of a particle with energy E. This acceleration model, which is called *second order Fermi acceleration* because it is proportional to the square of the cloud velocity, is slow because $\beta \ll 1$.

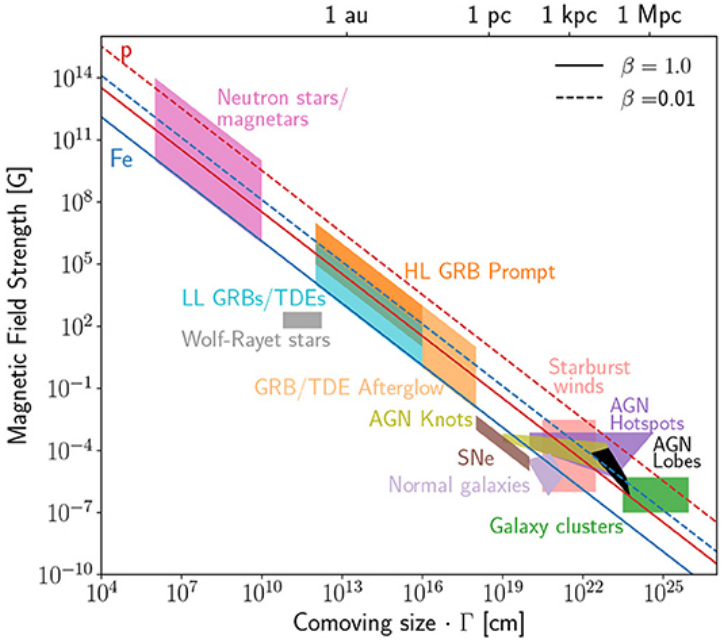


Figure 2.4: Different source classes as a function of their magnetic field strengths, B, and characteristic acceleration region sizes, R (the Hillas diagram). The solid (dashed) lines indicate the minimum values of the product BR to accelerate protons (red) and iron (blue) to an energy of 10^{20} eV for high (low) shock velocities β. The characteristic acceleration region size, R, is equal to the comoving size of the source times the Lorentz factor of the flow, Γ [33].

Later this mechanism was modified and a more efficient way to accelerate CRs to the highest energies was found [35] for the kind of shock waves that occur near supernovae. This mechanism is called Diffusive Shock Acceleration (DSA): A relativistic particle crossing the supernova shock front will, in its own rest frame, observe an approaching ionized gas on both sides of the shock and thus will be magnetically reflected each time it crosses the shock front. This process is capable of accelerating particles proportional to the gas velocity and hence is called *first order Fermi acceleration*:

$$\frac{\delta E}{E} \propto \beta. \tag{2.4}$$

Interestingly, the energy spectrum produced in this process is naturally a power law with a spectral index of -2.

2.2.2 Top-Down Models

In addition to acceleration models, there are several top-down models which try to explain part of the cosmic ray energy spectrum as the result of the decay of super-heavy particles. In Super-Heavy Dark Matter (SHDM) models, these are metastable particles from the early universe that contribute to cold dark matter and thus are located in the halos of galaxies [36, 37]. Another potential source, topological defects [38], are a general consequence of Grand Unified Theories [39]. They are relics from the early universe and could produce high energy particles. However, since their densities are typically very low, they either produce low fluxes or large anisotropy. Alternatively, Z-burst models [40] predict interactions of high energy neutrinos with background neutrinos in a resonant annihilation process creating Z-bosons, which then decay into hadrons and leptons. All top-down models predict high fluxes of cosmic ray primary photons and neutrinos and are mostly incompatible with the already established upper limits.

2.3 Propagation at the Highest Energies

Ultra-high-energy cosmic rays are expected to have an extragalactic origin [41]. On their path to Earth, they are mostly affected by magnetic fields and the astrophysical photon background [42] which consists mainly of the Cosmic Microwave Background (CMB) and the Extra-Galactic Background Light (EBL). Magnetic fields strongly bend the trajectories of charged, low-energy particles rendering it impossible to trace them back to their sources through their arrival directions. At the highest energies, however, the angular deflections from the sources are often as low as just a few degrees [43].

UHECR particles interact with the CMB and EBL in pair- and photo-pion production processes. Nuclei heavier than protons additionally suffer photo-disintegration [42]. Most important for this work is the photo-pion production which occurs when protons meet the blueshifted CMB, called the GZK process. It was named after Greisen, Zatsepin, and Kuzmin who predicted a flux-suppression at the highest energies due to this process as early as 1966 [44, 45]. A high-energy proton interacts with a CMB photon in one of the two following processes:

$$p + \gamma \rightarrow \Delta^+ \rightarrow p + \pi^0 \tag{2.5}$$

$$p + \gamma \rightarrow \Delta^+ \rightarrow n + \pi^+. \tag{2.6}$$

The minimum proton energy threshold for photo-pion production to occur on the 2.7 K CMB is about $10^{19.5}$ eV. Due to the GZK process, a proton with sufficiently high energy will lose about 20 % of its energy every about 5-10 Mpc until it falls below the threshold energy. For this reason, the maximum distance a proton can travel while maintaining an energy above the threshold is about 100 Mpc [46], the GZK horizon. Any protons observed with higher energies must therefore have been accelerated in a nearby source [47]. The neutral pions created in the GZK process will decay to high-energy photons, which travel in a straight path afterwards.

The AugerPrime upgrade [48] of the Pierre Auger Observatory which is currently underway is aimed at addressing the open question of whether the observed cutoff of UHECR fluxes at the highest energies is caused by a maximum energy of the CR sources, the photodisintegration of heavy primaries, or by the GZK effect. The most recent composition results of the Pierre Auger Collaboration favor a flux suppression caused by the maximum energy of the sources [49], while the Telescope Array collaboration observes a proton dominated composition at the highest energies which favors the GZK scenario [50]. Since ultra-high-energy photons are a key signature of the GZK effect, their observation would be an independent confirmation of this process [42].

UHECR photons, being neutral particles, are not deflected by magnetic fields and thus could be traced back to their sources. However, interactions with background photons can result in pair production before the photons reach Earth. The high-energy electrons created quickly lose energy by inverse Compton scattering or synchrotron radiation [51]. For this reason, higher photon fractions are only expected in scenarios with original sources close to the observer (i.e. only a few Mpc).

2.4 Air Showers

The rarity of UHECR, with fluxes of about one particle per square kilometer per year or less, makes their direct observation almost impossible. Fortunately, when UHECR enter Earth's atmosphere, their energy is distributed to a multitude of newly created particles scattered over square kilometers, called an extensive air shower. In the following Sections, the most important processes taking place in both hadronic, and photon primary air showers are summarized. Hadronic air showers consist of two main elements: the hadronic cascade and the electromagnetic cascade. Both cascades are commonly described by the Heitler model [52] with some modifications needed for the hadronic cascade [53].

2.4.1 Hadronic Cascades

According to the modified Heitler model illustrated in Fig. 2.5, when a hadron primary cosmic ray strikes an air molecule all three types of pions are produced in equal fractions. While neutral pions (π^0) almost immediately decay into two photons, the charged pions (π^+, π^-) have time dilated decay times large enough to survive long enough to further interact with air molecules. These charged pion interactions take place roughly once per hadronic interaction length λ_{int}, which is about $120\,\mathrm{g/cm^2}$.

With each step of interactions, the average energy of the charged pions decreases until they reach a critical energy $\xi_C^\pi \approx 20\,\mathrm{GeV}$. At this energy, the relativistic boost is no longer sufficient for them to interact before decaying to muons. Therefore, this critical energy is defined as the energy for which the hadronic interaction length of a charged pion is equal to its decay length, $\lambda_{\mathrm{dec}} = \rho\gamma c\tau$, where γ is the Lorentz boost, τ is the charged pion lifetime, and $\rho = n\cos(\theta)\lambda_{\mathrm{int}}/h_0$ is the density of the atmosphere

depending on the scale height h_0 the number n of interactions which took place, and the shower incident angle θ [26].

The hadronic interactions are characterized by their average multiplicity, N. This means that after n interactions there are N^n pions and the Lorentz boost of the pions is $\gamma = E_0/(N^n m_\pi c^2)$. It follows that the number of interactions n_C needed to reach the critical energy increases with the energy of the primary particle, E_0 [26]:

$$\gamma = \frac{E_0}{N^{n_C} m_\pi c^2} = \frac{\lambda_{\text{dec}} h_0}{n_c \lambda_{\text{int}} c\tau \cos(\theta)}, \tag{2.7}$$

$$\overset{\text{at the critical energy}}{\Rightarrow} \frac{n_c}{N^{n_c}} = \frac{h_0 m_\pi c^2}{E_0 c\tau \cos(\theta)}, \tag{2.8}$$

and therefore

$$n_C = -\frac{W_{-1}\left(-\frac{h_0}{c\tau}\frac{m_\pi c^2}{E_0}\frac{\ln N}{\cos(\theta)}\right)}{\ln N}, \tag{2.9}$$

where W_{-1} is the lower branch of the Lambert-W function [54].

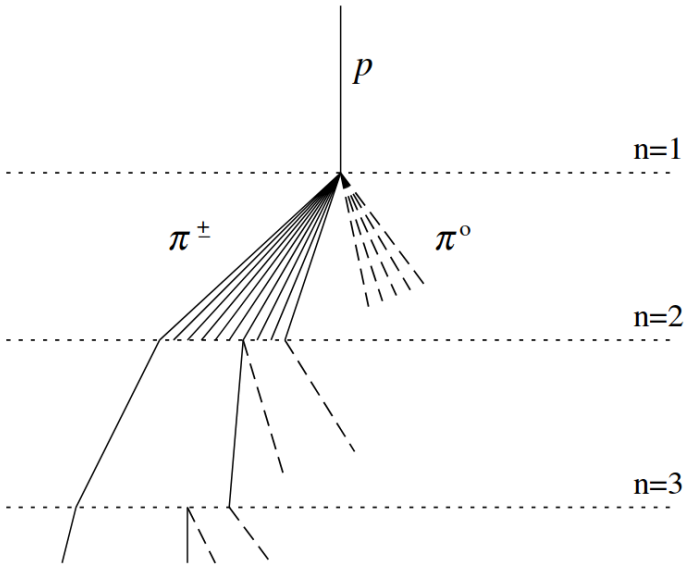


Figure 2.5: Sketch of the Heitler model of the hadronic cascade: a primary proton interacts after one interaction length creating N pions, 2/3 of them being charged. While the neutral pions decay, thus starting electromagnetic cascades, the charged pions interact one interaction length later each creating N new pions [53].

2.4.2 Electromagnetic Cascades

High-energy photons, like those created in the decay of π^0 produced in hadronic cascades, undergo pair production in interactions with the Coulomb fields of atmospheric

nuclei. The generated electrons and positrons then repeatedly produce high-energy photons in bremsstrahlung emissions. The combination of these two-body splittings forms the electromagnetic cascade of an air shower, which consists of a multitude of photons, electrons, and positrons, some of which reach ground. In the Heitler model of electromagnetic cascades [52], this doubling of particle counts is assumed to take place once per splitting length, λ_r, which is about $25\,\mathrm{g/cm^2}$ [55]. A sketch of this process is shown in Fig. 2.6. In the Heitler model, the number of particles increases until the average energy of the photons and electrons reaches the critical energy of $\xi_C^e = 85\,\mathrm{MeV}$ at which the energy loss by collisions surpasses the radiative energy loss. The number of interactions n_C until this happens is about

$$n_C = \log_2(E_0/\xi_C^e), \tag{2.10}$$

with a primary particle energy of E_0. This point of an air shower is called the electromagnetic shower maximum and contains a total number of 2^{n_C} electromagnetic particles. The atmospheric depth, X_max^γ, of the shower maximum logarithmically increases with the primary energy as

$$X_\mathrm{max}^\gamma = n_C \lambda_\mathrm{r} = \lambda_\mathrm{r}\ \log_2(E_0/\xi_C^e). \tag{2.11}$$

This increase is described by the elongation rate

$$\Lambda^\gamma = \frac{dX_\mathrm{max}^\gamma}{d\ \log_{10}(E_0)}, \tag{2.12}$$

with a typical value of $84.4\,\mathrm{g/cm^2}$ per order of magnitude in primary particle energy [55].

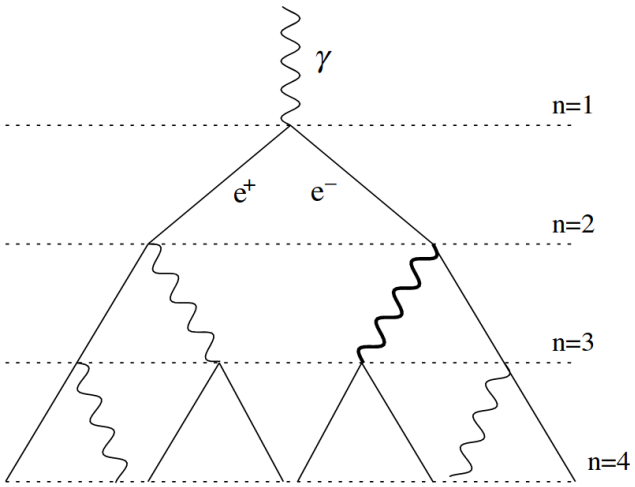


Figure 2.6: Sketch of the Heitler model of the electromagnetic cascade: after one electromagnetic interaction length, a photon creates an e^+/e^- pair via pair production. One interaction length later, each of these particles emit a high-energy photon due to bremsstrahlung. Thus, the total number of electromagnetic particles doubles once per interaction length [53].

2.4.3 Hadronic Primary Air Showers

Air showers with hadronic primaries consist of both hadronic and electromagnetic cascades with relative sizes depending on the primary particle energy and type. An air shower initiated by a particle heavier than a proton with a mass number A, can be approximated as A simultaneous proton air showers each with energy E/A (the superposition model). In this model, the atmospheric depth of the shower maximum of nuclei $X^A_{\max}$ is [26]

$$X^A_{\max}(E) = X^p_{\max}(E/A) = c + D_p \ln(E/A), \tag{2.13}$$

with the parameters c and D_p depending on characteristics of hadronic interactions. Since lighter nuclei have deeper shower maxima (assuming the same primary energy), the measured position of the shower maximum is an indicator of the primary particle type.

Once the critical energy is reached, the total number of muons in an air shower can be approximated by the number of created charged pions and therefore for proton primary air showers is given by [26]

$$N^p_\mu \approx \left(\frac{E_0}{\xi^\pi_C}\right)^\beta, \tag{2.14}$$

with

$$\beta = \frac{\ln(\frac{2}{3}N)}{\ln N}, \tag{2.15}$$

where the factor $\frac{2}{3}$ accounts for the fraction of charged pions. This results in a range of β between 0.88 and 0.92. Applying the superposition model, it becomes clear that air showers with heavier primaries should be expected to have a higher number of muons:

$$N^A_\mu \approx A\left(\frac{E_0/A}{\xi^\pi_C}\right)^\beta = N^p_\mu A^{1-\beta} \tag{2.16}$$

The energy of the electromagnetic cascade, E_{em}, only weakly depends on the primary particle mass A, and therefore often is used to measure the primary particle energy:

$$E_{\text{em}} = E_0\left(1 - \left(\frac{2}{3}\right)^{n_c}\right). \tag{2.17}$$

2.5 Photon Primary Air Showers

Photon primary air showers, in contrast to hadronic air showers, have a very strong electromagnetic and a rather small hadronic shower component which is only created by high energy photonuclear interactions early in the shower. Also, the comparatively slow development of the electromagnetic cascade is further slowed down due to the LPM effect (see below). At higher energies, photons have a small probability of interacting before they reach the atmosphere causing a so-called preshower. These effects, which are of particular importance for identifying photon induced air showers, will be summarized in the following Sections.

2.5.1 LPM Effect

Discovered in 1953, the Landau-Pomeranchuk-Migdal (LPM) effect describes a strong suppression of the bremsstrahlung and pair production splitting processes at high energies [56, 57]. This effect becomes increasingly important for an electromagnetic cascade with mean particle energies above 10^{15} eV. At low energies, the splitting rate for bremsstrahlung, Γ, can be well approximated as

$$\Gamma \approx nv\sigma, \tag{2.18}$$

where n is the density of the medium, v is the relative velocity, and σ is the splitting cross section [58]. The time these splitting processes take increases with energy. This can be understood as a consequence of the uncertainty principle since the low momentum transfer at high energies must correspond to a large interaction distance. At very high energies (about 10^{15} eV depending on the medium density [59]), the typical time needed for the splitting process becomes longer than the average time between collisions with the medium. Consequently, the multiple coulomb scattering that occurs between each splitting process causes a decoherence that strongly suppresses the splitting rate [58]. This in turn causes a significant increase in the shower length [60] and thus the atmospheric depth of the shower maximum and its event-to-event fluctuation [51].

2.5.2 Preshowers

Photon primary cosmic rays with energies of at least 10^{19} eV can potentially interact with the geomagnetic field of Earth and be converted into an electron-positron pair before their first interaction in the atmosphere [61, 62]. The probability of this increases significantly with shower energy and depends on the strength of the magnetic field component perpendicular to the movement direction of the primary photon (see Fig. 2.7). The electrons created in this process can emit high-energy photons due to synchrotron radiation which is again depending on the electron energy and the perpendicular component of the magnetic field. As a result of these two processes, at high energies electromagnetic cascades can be produced before a photon enters the Earth's atmosphere. These cascades are called preshowers and result in air showers initiated by several less-energetic electromagnetic particles instead of a single high-energy photon.

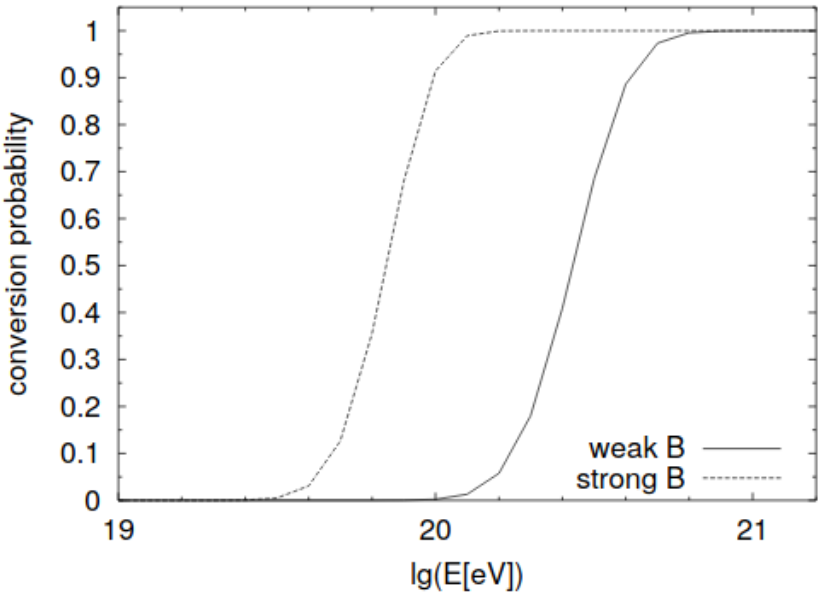


Figure 2.7: Calculation of the preshower probability as a function of the primary energy for two different arrival directions (with respect to the geomagnetic field). The calculations have been done for the Pierre Auger Observatory [63].

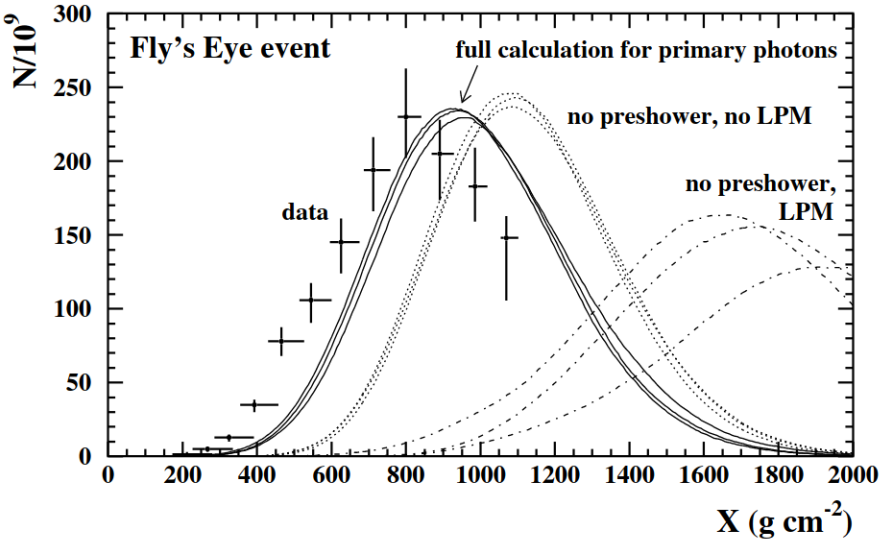


Figure 2.8: Simulated profiles of a photon primary air shower compared to data of a hadronic primary cosmic ray event at very high energy. The changes in the profile due to the preshower and LPM effects are shown by the various lines. While the LPM effect delays the shower development, the preshower effect results in an earlier development in the atmosphere [51].

Typical preshowers for photon energies of 10^{20} eV would start at about 1000 km above ground. These then would enter the atmosphere ($\sim$ 100 km altitude) with just one to a few electron-positron pairs and several hundred photons with a wide spread in energy [51]. The shower-to-shower fluctuations of air showers started by a preshower are lower than for other photon air showers since they begin with a multitude of particles. The spread of these highly boosted preshower particles however is low enough for them to be measured as a single air shower [51]. Since the large number of particles in the beginning of a preshowered photon air shower results in an earlier development of the cascade, the separation of a photonic air shower from a hadronic one is significantly more difficult if it preshowered (see Fig. 2.8).

2.5.3 Photonuclear Interactions

Air showers with photon primaries have a hadronic shower component due to photonuclear interactions at high energies. While the cross section of pair production decreases with growing energy due to the LPM effect (see Sec. 2.5.1), the cross section of interactions of photons with target nucleons increases. Usually, these interactions take place when the high-energy photon fluctuates into a quark-antiquark pair ($q\bar{q}$) enabling hadronic interactions in the nucleus [64]. At very large energies, direct interactions are possible as well. For most target materials, the cross section of these interactions even surpasses that of pair production if the energy is above 10^{20} eV [57, 64]. Already at a few tenths of EeV, electromagnetic air showers can develop significant hadronic components [64].

The actual cross section of photonuclear interactions at energies above the TeV scale has not been measured directly and therefore is estimated using extrapolations over several orders of magnitude in energy [63] (see Fig. 2.9). Because of this lack of data, the uncertainties of the estimated photon-air cross sections could be greatly reduced with the observation of a small number of photon induced air showers [51].

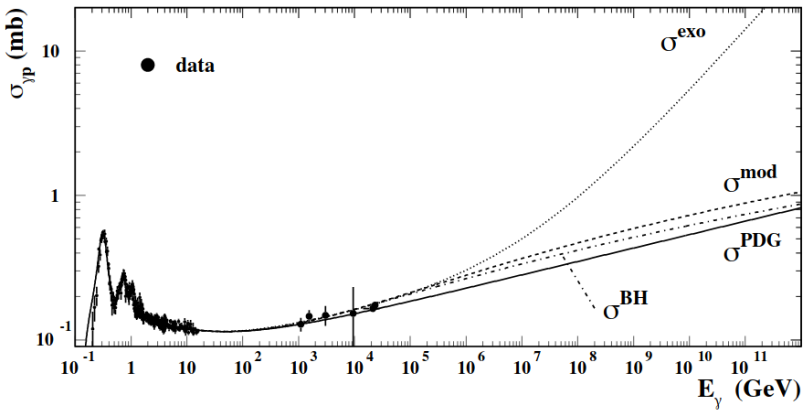


Figure 2.9: Photonuclear cross section as a function of the photon energy [63]. Created using data from [65], extrapolations from [65] and [66], parametrizations from [67] and [68], and a fit from [69].

2.5.4 Properties of Photon Induced Air Showers

At low primary particle energies, photon induced showers are shorter than hadronic showers [57], but since the increase in length of hadronic air showers with energy is slow, already at energies above 10^{14} eV photon induced air showers have a deeper $X_{\max}$ than their hadronic counterparts. This faster increase of $X_{\max}$ is strongly enhanced by the LPM effect but can be mostly mitigated at highest energies if the air shower begins with a preshower (see Fig. 2.10).

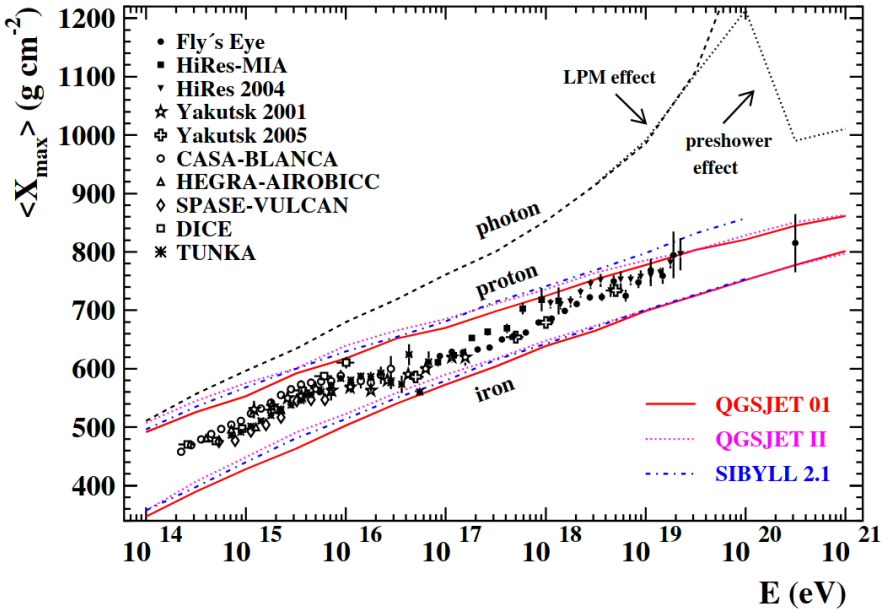


Figure 2.10: Simulation of the average $X_{\max}$ as a function of the primary particle energy for photons, protons, and iron, compared to data of various experiments [51] (data and simulation references in [70])

The small hadronic shower component results in a low number of muons being produced from the decay of charged pions. Additionally, the creation of muons from muon pair production in the electromagnetic cascade is also suppressed by a factor of $(m_e/m_\mu)^2 \approx 2.3 \cdot 10^{-5}$ [71]. The resulting distributions of muon numbers for different atmospheric depths are shown in Fig. 2.11.

The lack of muons in photon air showers causes a fast decrease of the measured signals at ground at larger distances from the shower core, since the measured signal at these larger distances is usually dominated by muons [73, 74]. This effect is further enhanced by the late development of high-energy photon air showers. These characteristics together result in photon primaries having late particle arrival times with respect to a planar shower front (see Fig. 2.12) approximation at large distances from the shower axis [72]. It is these distinct features which are used in the algorithms, presented in Sec. 3.5, which aim at separating photonic from hadronic air showers.

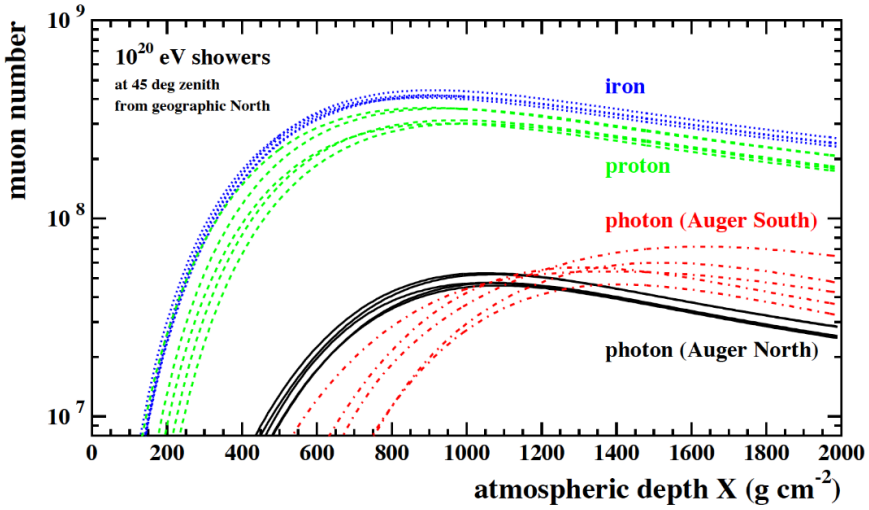


Figure 2.11: Muon number profiles for different primary particle types [72]

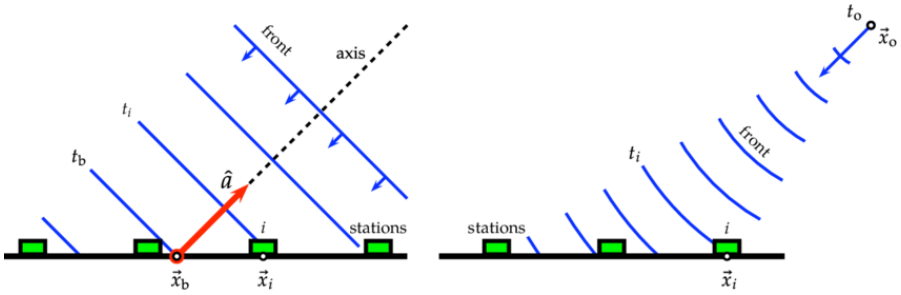


Figure 2.12: Left: Schematic view of a plane shower front approximation. Right: Schematic view of a spherical shower front [75].

2.6 Current Upper Limits on the Photon Flux

Limits on the diffuse photon flux have been provided by multiple experiments. A goal of this work is to either find a flux of UHECR photons or to set new upper limits on the diffuse photon flux at highest energies. In this Section, the most recent published upper limits on the cosmic ray photon flux at high energies are presented (Figs. 2.13 and 2.14).

At energies of around 10^{17} eV, upper limits on the diffuse photon flux were set by a recent analysis of data taken with the EAS-MSU experiment between 1984-1990 [76, 77]. The data was taken using 76 charged-particle detectors, arranged over a total area of about 0.5 km^2 coupled with a central underground muon detector. In this analysis, each event that did not trigger the muon detector was considered a photon candidate event. The hadronic background was estimated using simulated proton air showers assuming a proton fraction of 46 ± 6 % which was then subtracted from the number of selected candidates.

The KASCADE-Grande collaboration published upper limits on the diffuse gamma ray flux at energies between 10^{16} eV and $10^{17.5}$ eV using their 37 scintillation detector stations forming an instrumented area of 0.49 km^2 [78]. Information about the muon content of the observed air showers was retrieved from muon detectors located in a subregion of the detector array which was required to have participated in each analyzed shower. Photon candidates were selected by comparing the measured electromagnetic and muonic signals and no background subtraction was applied [78].

More recently, measurements using the 1660 water Cherenkov detector stations of the Pierre Auger Observatory combined with data from the 27 surrounding fluorescence telescopes was used to search for photons with energies above 1 EeV [74]. A variable sensitive to the radius of curvature of the shower front was used which was developed for the low number of surface detector stations triggered at these energies. This variable was combined with measurements of X_{max} from the fluorescence telescopes in a boosted decision tree (BDT) analysis. The training of the BDTs was done using simulations of proton and photon induced air showers and the photon candidate cut was fixed to a 50 % photon selection efficiency.

A similar analysis was done using the denser 750 m spacing sub-array of the Pierre Auger Observatory combined with the low-energy fluorescence detector HEAT to search for a photon flux in an energy range of 0.2 to 1 EeV [2]. In this case, the training of the BDTs was performed using a 5 % data burn sample.

Recent photon searches by the Telescope Array (TA) collaboration use BDTs to combine 16 composition sensitive variables using the 507 scintillation detector stations of the TA surface detector array [79]. Some of these parameters are derived from the radius of curvature and the total shower size. The cut position was optimized for the hypothesis of all events being protons. Simulated proton and photon air showers were used for the training of the BDTs.

In 2008, early before the hybrid detector analysis [74], the Pierre Auger Collaboration published limits on the diffuse photon flux using the water Cherenkov detector array [80]. Due to the larger exposure without the need of fluorescence detector measurements, flux limits at the highest energies (above 10^{19} eV) were set. This was done by combining measurements of the curvature of the shower front with reconstructed signal risetimes at 1000 m from the shower core in a Principal Component Analysis.

In addition to these diffuse photon flux searches, some analyses aim for the detection of photon sources by examining the arrival directions of photon-like events (e.g. [81]) or try to find a photon flux by searching for angular correlations with a stacked catalog of possible high-energy photon sources (e.g. [82]).

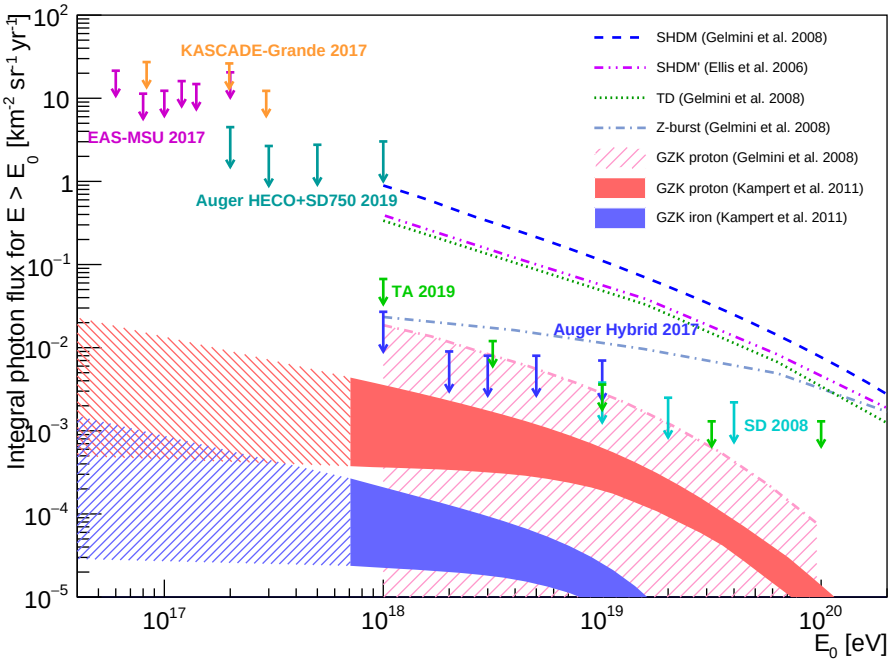


Figure 2.13: Upper limits on the integral photon fluxes as a function of the energy threshold. Results of various experiments at 95 % CL [74, 79, 80] and 90 % CL [77] [78] are shown and compared to model predictions [83, 84, 85].

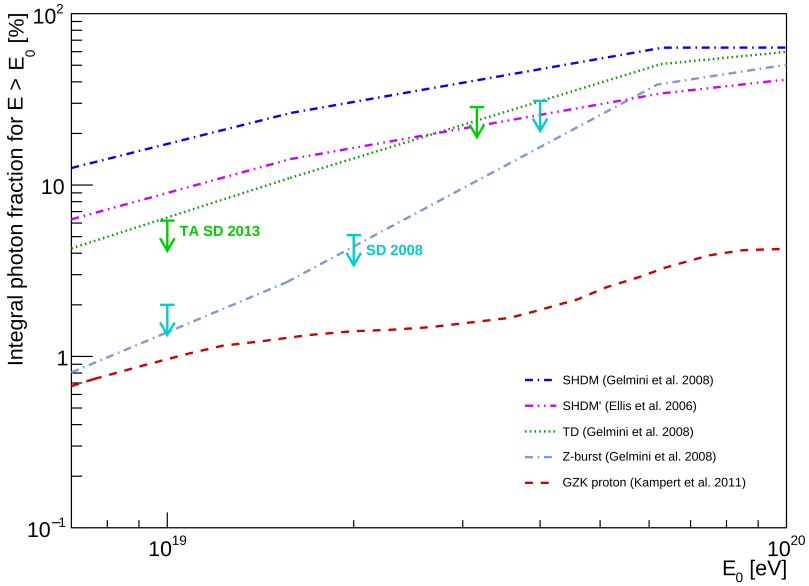


Figure 2.14: Limits on the integral CR photon fraction [80, 86] as a function of the energy threshold in comparison to GZK expectations as well as a selection of top-down models [83, 84, 85].

2.7 Air Shower Simulation

Since the development of high-energy air showers in the atmosphere is a large and complex process, models are used to simulate air showers at different energies, geometries, and primary particle types. To do so, many interdependent sub-processes must be accounted for. Most of these are well understood but are difficult to account for all at once as they can interact. However, some processes are not well understood and can only be approximated with theory and extrapolations. The largest difficulty is posed by soft hadronic interactions at the highest energies as no data from man-made accelerators is available at these energies and the uncertainties in any available data result in model uncertainties on the order of 20 % [70].

The extended air shower simulation code CORSIKA [87, 88] can accommodate different hadronic interaction models and can be used to create simulated air shower data for arbitrary geometries and primary particle types. It operates by creating an internal stack of particles whose interactions are individually processed: first, the point of the next interaction (or decay) of a particle is generated and any continuous effects are accounted for on its way to this point. These continuous effects include multiple scattering, energy loss processes, deflections due to the magnetic field, and Cherenkov light emission. In the next step, if the particle has not yet passed the defined observation level, either a particle interaction is processed using the appropriate interaction model (for electromagnetic, low-energy hadronic, or high-energy hadronic processes) or the particle decay is simulated. Cuts are then applied to the generated secondary particles, and the remaining particles are then placed in the particle stack to be processed in turn. These steps are then repeated until all particles on the stack are cut or have passed the observation level.

In this work, air shower simulations created with CORSIKA are used to optimize the photon analysis (see Sec. 6.5.1) and then later to test the proton hypothesis of the candidate events (see Sec. 6.5.6). For surface detector studies, the most important output of the simulation are the details of the particles which reach the observation level.

It is worth noting that all state-of-the-art hadronic interaction models generate fewer muons than are observed in real cosmic ray data [14]. However, the effect of this uncertainty in high-energy air shower simulations is strongly mitigated when simulating photon air showers since they mostly consist of electromagnetic particles [51]. For this reason, the limits to the photon flux set in this work were produced using only photon air shower simulations and real data.

2.7.1 High-Energy Hadronic Interaction Models

The largest uncertainties of extended air shower simulations are introduced by high-energy hadronic interactions. Two of the most successful models used for this type of interaction in air shower simulations are EPOS LHC [89] and QGSJETII-04 [90]. Both of these models calculate cross sections using Gribov Regge theory [91] to build

a Pomeron[1] from the interaction components and then apply corrections afterwards. The main difference between these interaction models is that EPOS LHC includes an interference term which also takes energy sharing into account [92]. For this reason, all air shower simulations used in this work have been created using EPOS LHC. In practice, EPOS LHC predicts about $20\,\mathrm{g\,cm^{-2}}$ deeper X_{max} values than QGSJETII-04 and about 10 % larger muon numbers at ground level [93]. A set of typical images of simulated photon, proton, and iron primary air showers produced with CORSIKA is shown in Fig. 2.15.

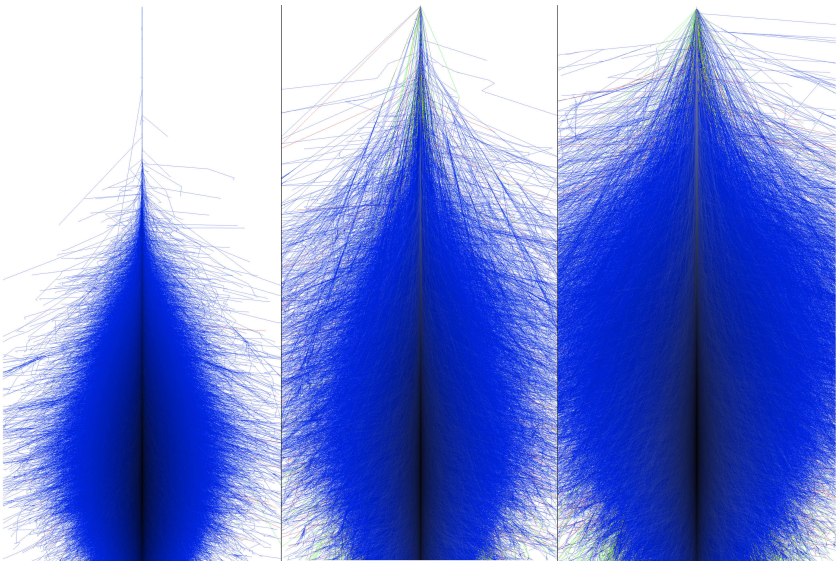


Figure 2.15: Typical representations of simulated air showers caused by a primary photon (left), proton (middle), and iron nucleus (right), modified from [94]. All simulations were started with a first interaction height of 30 km and with a primary particle energy of 10^{15} eV. The vertical range is 30.1 km, the horizontal range is 10 km. Only muons (green) and hadrons (red) with $E > 100$ MeV, and electromagnetic particles (electrons, positrons, and photons in blue) with $E > 100$ keV are shown.

2.7.2 Preshower Simulation

Since the simulation of air showers usually starts when a cosmic ray particle enters Earth's atmosphere, a special treatment for photon primary air showers which contain a preshower is needed. This is done through an additional program, *PRESHOWER* [61, 95], which can be combined with CORSIKA to simulate photon-induced particle cascades outside the atmosphere. The calculation of a possible conversion of the primary photon to an electron-positron pair begins at a distance of 5 Earth radii and is performed

[1] Pomerons are quasi-particles that effectively account for the transferred momentum while carrying the quantum numbers of the vacuum.

in step sizes which have been optimized for both accurate calculations and fast evaluation times [96]. If a preshower occurs, the information of all resulting particles which enter the atmosphere is passed to CORSIKA and the point of first interaction is set to a negative value in atmospheric depth.

2.7.3 The Thinning Procedure

Ultra-high-energy air showers contain a huge number of secondary particles. For example at a primary energy of $10^{19}\,\mathrm{eV}$, on the order of 10^{12} particles are created. It is not currently possible to simulate all of these particles fast enough (and with reasonable amounts of storage space) to build a set of simulations with large enough statistics for air shower analyses. For this reason, the thinning technique has been developed [97]. When simulating an air shower using the thinning procedure, instead of following every generated particle in the cascade, only a subset of the shower particles is chosen to be tracked. These particles are then assigned a weight which is inverse to the probability for the respective particle to be selected for the tracked subsample. In this way, all estimators based on sums over particles can be obtained without biases [98]. Since the largest statistical fluctuations in air-showers are sourced from the first few interaction processes, the thinning procedure is only applied to particles below a threshold energy, $E_{\mathrm{threshold}}$, which depends on the user selected thinning level, ϵ_{th}, which is defined as

$$\epsilon_{\mathrm{th}} = E_{\mathrm{threshold}}/E_0, \tag{2.19}$$

where E_0 is the energy of the primary particle.

When simulating the detector response to a thinned air shower, the potentially large number of particles hitting each detector station needs to be recovered from the list of particles and weights. This is done through a resampling procedure [98]: for each detector station a sampling region is defined that is large enough to contain sufficient particle statistics in the thinned particle subset but small enough to represent the same position in the air shower (see Fig. 2.16). Each shower particle from the thinned set which hits this *thinning region* is replicated a number of times chosen according to the weight of the original particle. The number of these artificial particles hitting the detector station is then determined randomly by sampling from a Poisson distribution. In Fig. 2.17 a sketch of the resampling procedure is shown.

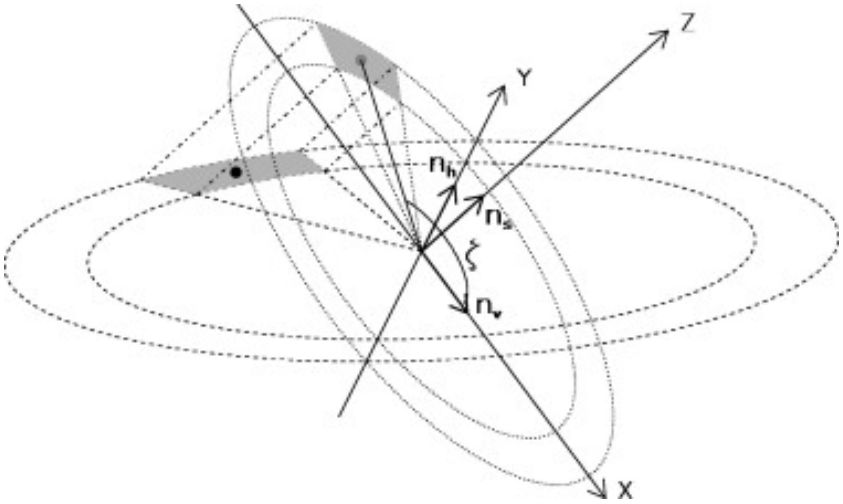


Figure 2.16: Definition of the thinning region around a detector station (gray area, black dot). The regions' limits are derived from a ring segment around the shower axis projected to the ground [98].

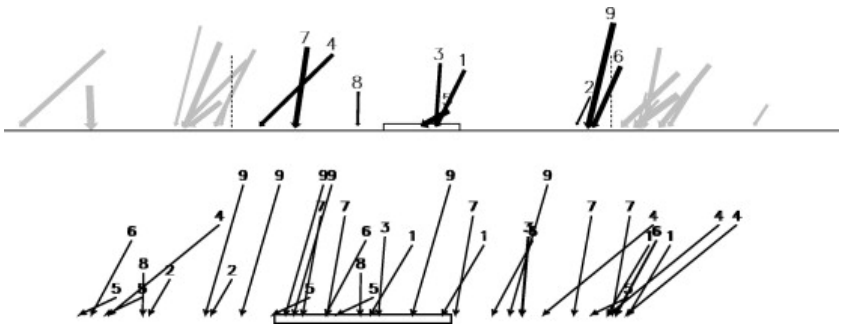


Figure 2.17: Sketch of the resampling procedure used in the thinning technique:
Top: Each particle hitting the sampling region is represented by an arrow (arrow length proportional to energy, width proportional to particle weight).
Bottom: The particles from above are multiplied according to their weight and placed in random positions in the sampling region [98].

Chapter 3

The Pierre Auger Observatory

The Pierre Auger Observatory was specifically designed to address questions on the origin of cosmic rays at the highest energies [99]. Due to the extremely low flux of cosmic rays at these energies, a large detector was required to gather enough statistics to make measurements in this energy region. It was designed as a hybrid detector combining a Surface Detector (SD) with a Fluorescence Detector (FD) in the Argentinian Pampa Amarilla. With this design, air showers are observed in two complementary ways, combining the strengths of both methods and providing a valuable measurement redundancy [99].

The SD consists of 1660 detector stations which are used to measure the secondary particles of the air shower which reach the ground level. The photon search presented in Chapter 6, as well as the trigger studied presented in Chapter 4, use the measurements made by this detector. For that reason, a detailed description of the SD is provided in Sec. 3.2 while a much more brief description of the FD will be given immediately below. The reconstruction procedures used for events measured with the SD are presented in Sec. 3.3. The software framework $\overline{\text{Off}}\underline{\text{line}}$, which is used for these reconstruction procedures, is introduced in Sec. 3.4. Finally, photon sensitive event variables based on measurements of the SD are presented in Sec. 3.5.

3.1 The Fluorescence Detector

The Fluorescence Detector is used to measure the longitudinal development of extensive air showers, while they are passing through Earth's atmosphere [100]. The charged particle component of the air shower excites nitrogen molecules on its way, which then emit fluorescence light in the UV range. This fluorescence light is measured by 27 telescopes, each covering a field of view of $30° \times 30°$ with 440 pixels that are digitized every 100 ns.

The number of emitted fluorescence photons in an air shower is proportional to the electromagnetic energy loss in the atmosphere and can thus be used to measure the total energy deposit of the charged shower particles [100]. Since this energy deposit represents about 90 % of the primary particle energy, it provides a nearly calorimetric measurement of the air shower energy. The direct measurement of the shower profile provided by the FD makes very accurate measurements of the shower maximum possible which is invaluable for composition measurements (see Sec. 2.1.2).

3.2 The Surface Detector

The Surface Detector (SD) array of the Pierre Auger Observatory consists of 1660 autonomously working water Cherenkov detectors [101] placed on a triangular grid with a spacing of 1.5 km. This gives the SD a total instrumented area of about 3000 km^2 [102] (Fig. 3.1).

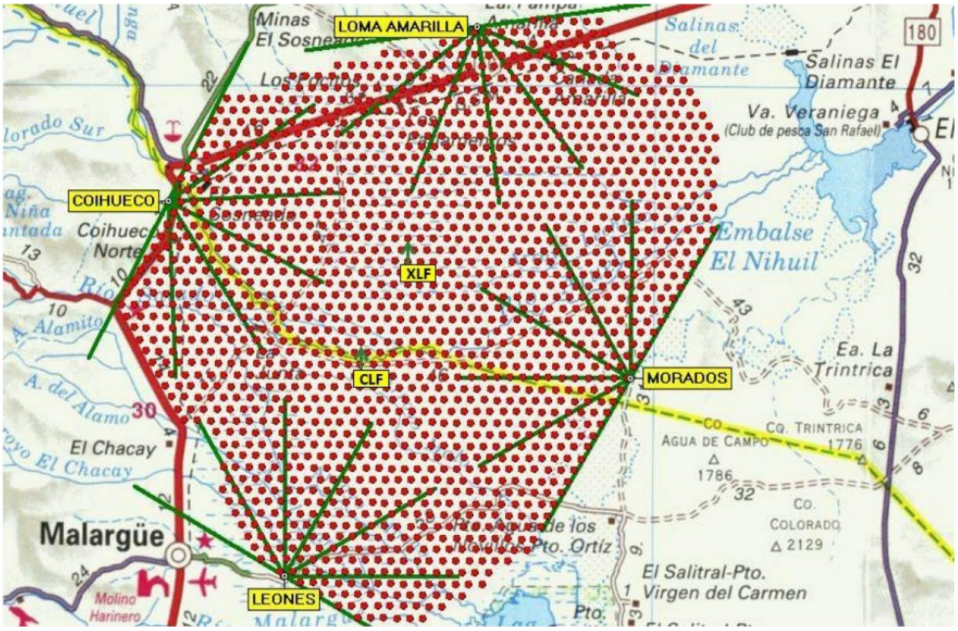


Figure 3.1: Map of the Pierre Auger Observatory. Each dot represents a Surface Detector station. The green lines indicate the field of view of the Fluorescence Detector telescopes [99].

3.2.1 Surface Detector Stations

Each water Cherenkov detector station (SD station) consists of a cylindrical weather- and lightproof polyethylene tank with a diameter of 3.6 m and a height of about 1.6 m [99]. This tank contains a sealed liner filled with 12 000 l of ultra-pure water. The liner provides an excellent diffuse reflectivity for the Cherenkov light produced by relativistic charged particles traversing the water [103], and provides an additional layer of light shielding from the outside. The Cherenkov light produced in the water is measured using three Photomultiplier tubes (PMTs) each with a 9 inch diameter. These are equally distributed on the surface of the liner at a radial distance of 1.2 m from the central tank axis and face downward into the water through windows in the liner surface.

Since every detector station is required to operate autonomously, i.e. without exterior cabling, each station is equipped with a solar panel and two batteries for power supply, a GPS receiver for accurate timing, and a radio transceiver to transmit measurement data to the main campus of the observatory. The outside of an SD station, sitting on site in the Argentinian Pampa, is shown in Fig. 3.2.

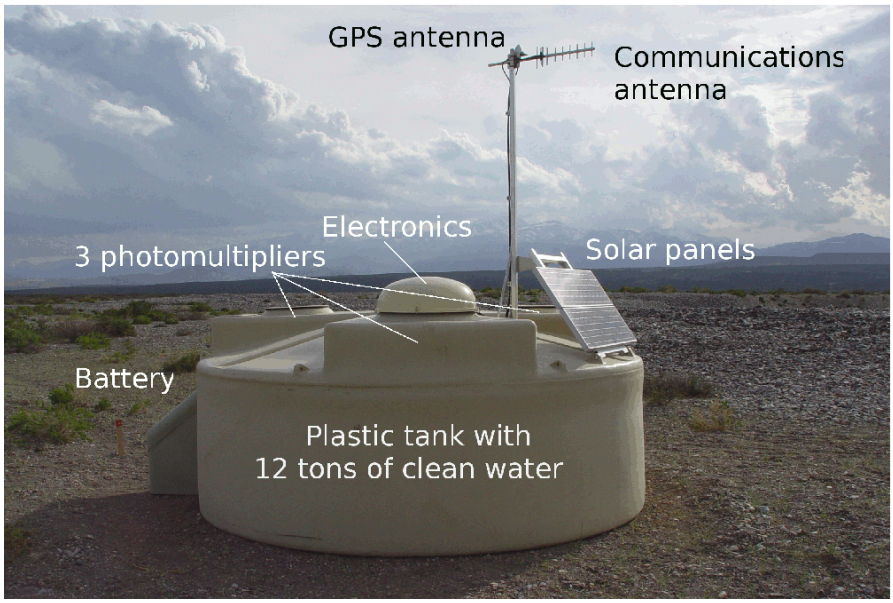


Figure 3.2: Picture of an SD station in the Argentinian Pampa Amarilla showing the main detector components [104].

Each PMT provides two signals with differing amplification factors. The low-gain signal (LG) is directly taken from the anode of the PMT while the high-gain signal (HG) is taken from the last dynode and amplified by an additional nominal factor of 32 [102]. This provides a large dynamic range for the digitization of signals both close to the shower core (with about 1000 particles/µs) and far from it (about 1 particle/µs). Both of these signals are digitized using a 40 MHz, 10-bit Flash Analog to Digital Converter (FADC) resulting in a time resolution of 25 ns. These FADCs are currently being replaced as part of the ongoing "AugerPrime" upgrade of the Pierre Auger Observatory [48]. The new FADCs have a sampling rate of 120 MHz and will thus increase the time resolution to 8.3 ns. Furthermore, on top the detector stations will be equipped with an additional scintillator unit, which will, in combination with the water Cherenkov detector, allow for a better discrimination of primary particle types.

3.2.2 Calibration of the Surface Detector

The Surface Detector PMT signals are recorded by 10-bit FADCs, covering an input range of 0-2 V. In order to provide a common reference level between all detector stations and to facilitate comparisons with detector simulations, each station regularly performs an automatic self-calibration. This calibration is used to convert the FADC output to VEM (or Q_{VEM}), a unit equal to the average charge measured from a muon vertically traversing the detector (a Vertical Equivalent Muon) [105]. Additionally, the equivalent

amplitude, I_{VEM}, of a VEM-particle is determined to provide the reference pulse height needed for the triggering algorithms (see Sec. 3.2.3).

The continuous on-line calibration makes use of atmospheric muons which pass through each detector at a rate of about 2500 Hz. Each time a coincidence of signals with at least 5 FADC counts above the baseline level occurs in all working PMTs, these signals are recorded and used to fill charge- and pulse height histograms (see Fig. 3.3). Both of

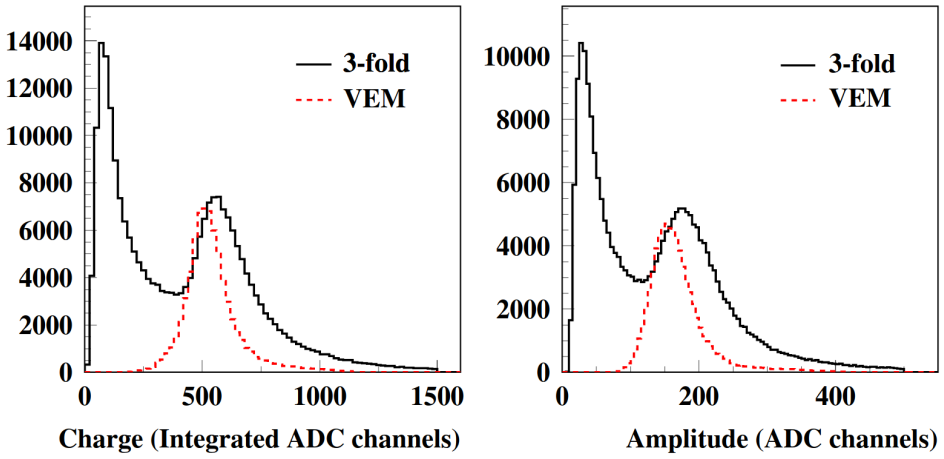


Figure 3.3: Charge (left) and pulse height (right) histograms of background measurements in an SD station. The second peaks are created by vertical muons and used to calculate $Q_{\mathrm{VEM}}^{\mathrm{peak}}$ and $I_{\mathrm{VEM}}^{\mathrm{peak}}$ [105]. The dashed line shows data triggered by an external muon telescope to select vertical and central muons only.

these calibration histograms contain two distinct peaks. The first peak is created by the distribution of low-energy particles accidentally satisfying the trigger requirements [105]. The second peak is due to muons traversing the detector and is therefore the one used for the calibration.

The muon peak of the charge histogram, $Q_{\mathrm{VEM}}^{\mathrm{peak}}$, can be directly converted to the charge of a vertical-equivalent muon, Q_{VEM}, by applying a factor of 1.09 VEM/$\mathrm{Q}_{\mathrm{VEM}}^{\mathrm{peak}}$, which is derived from the mean trajectory through a geometrical model of a detector station [106]. The value of $I_{\mathrm{VEM}}^{\mathrm{peak}}$ is not determined online since this would require large dead times of the detector to remove the baseline. Instead, the trigger levels rely on an estimate of $I_{\mathrm{VEM}}^{\mathrm{peak}}$, which is defined by tuning its value until the calibration trigger rate is 70 Hz [105]. The individual gains of the 3 PMTs are regularly updated so that $I_{\mathrm{VEM}}^{\mathrm{peak}}$ is at 50 FADC counts.

3.2.3 Surface Detector Triggers

The recordable event rate of each Surface Detector station is strongly limited by the requirement that all data must fit within the bandwidth of the wireless communication system, which must transmit to receivers at up to 40 km distance with a power consumption below 1 W [102]. For this reason, efficient filtering mechanisms must be applied to reduce the event rate from single station events, but retain as much usable air shower data as possible. The constraints of the wireless communication system require an event rate reduction on station level from the 3 kHz background to a rate of less than one transmitted event per hour. This reduction is mostly realized by the first two levels of the hierarchical SD trigger system: T1 and T2.

T1 and T2 triggers

There are four algorithms applied to the SD data which can produce a T1 trigger: Threshold (Th), Time Over Threshold (TOT), Time over Threshold deconvoluted (ToTd), and Multiplicity of Positive Steps (MoPS), all of which are described below. The difference between T1 and T2 trigger levels only applies to the Th-trigger, which requires a higher threshold to form a T2 trigger than for a T1. For all other algorithms, the requirements of T1 and T2 are identical.

Th Trigger The threshold trigger (Th) requires all PMTs to measure a signal above 1.75 $I_{\mathrm{VEM}}^{\mathrm{peak}}$ for a T1 trigger and above 3.2 $I_{\mathrm{VEM}}^{\mathrm{peak}}$ for a T2 trigger [102, 107]. In the case of only one or two of the PMTs working, these thresholds are increased, according to Table 3.1, to reduce the effect of random coincidences [107].

# **Working PMTs**	**Th-T2**/$[I_{\mathrm{VEM}}^{\mathrm{peak}}]$	**Th-T1**/$[I_{\mathrm{VEM}}^{\mathrm{peak}}]$
3	3.20	1.75
2	3.60	2.00
1	5.00	2.85

Table 3.1: T1 and T2 threshold trigger settings for different numbers of working PMTs [107]

TOT Trigger The Time Over Threshold (TOT) algorithm is targeted at triggering detector stations with a signal that is significantly spread in time. It requires 13 bins (N bins) to have signals exceeding a threshold of 0.2 $I_{\mathrm{VEM}}^{\mathrm{peak}}$ (Th) within a time window of 120 bins[1] (*window* bins) in at least two of the three PMTs [102]. If only one PMT is working, the algorithm is applied to this PMT alone.

[1] 120 bins = 3 µs

ToTd Trigger The ToTd trigger is a refinement of the TOT trigger which makes use of the predictable time delay which comes from light reflecting on the liner surface (see Sec. 3.2.1). Basically, the Cherenkov light produced by a particle going through the detector occurs in a time-frame which is smaller than the time resolution of the FADCs. However, since a considerable portion of this light undergoes multiple scatterings on the liner surface before hitting the sensitive area of a PMT, this part of the signal is delayed.

Assuming that each Cherenkov photon produced within the detector volume has an approximately constant probability to propagate towards one of the PMTs after each reflection on the liner surface, the signal structure of a particle going through the detector should have an exponential tail. The decay time, τ, of this tail mostly depends on the detector geometry and has been shown to be about $\tau \approx 67 - 70\,\text{ns}$, which is well above the time resolution of the FADCs [107, 108]. To prevent TOT trigger conditions from being met by a small number of particles with large signals and thus long tails, the ToTd algorithm applies a deconvolution to the FADC traces to remove the tails leaving only the peaks. This deconvolution works by reducing the signal in each FADC bin by a constant fraction of the signal in the preceding bin and then applying a normalization term [108]:

$$d_i = \frac{(a_i - f a_{i-1})}{(1-f)} \quad \text{with} \quad f = \exp(-\Delta t/\tau), \tag{3.1}$$

where d_i is the deconvoluted signal in bin i, Δt is the binning size (25 ns or 8.3 ns, see Sec. 3.2.1), and a_i and a_{i-1} are the original signals in bins i and $i-1$, respectively. Afterwards, the triggering conditions of the normal TOT algorithm are applied to the deconvoluted trace with the same thresholds set as shown in the TOT Section above.

MoPS Trigger Like the TOT and ToTd algorithms, the MoPS trigger aims at selecting signals produced by a series of low-energy particles. This is done by selecting FADC traces that contain a certain number of *positive steps*, M, above a threshold within a moving time *window* [109]. The algorithm is based on the assumption that, typically, each of these steps corresponds to the arrival of a new particle in the detector. In contrast to the other trigger algorithms, the MoPS trigger is completely independent of $I_{\text{VEM}}^{\text{peak}}$.

In the MoPS trigger, a *positive step* is defined as a sequence of bins in which the FADC trace increases (see Fig. 3.4). It has a minimum step size, $y_{\min}$, to avoid statistical fluctuations and a maximum step size, $y_{\max}$, to avoid muon-like signals [107]. Also, a veto is included which prevents the counting of additional steps due to fluctuations in the tails of large peaks. The number of vetoed bins after a step depends on the step size Δy [107]:

$$\text{Veto} = \text{floor}(\log_2(\Delta y) + 1 - \text{OFS}), \tag{3.2}$$

where the offset, OFS, is a free parameter used to adjust the veto sizes. In former versions of the MoPS implementation, the veto was set with the paramenter N_{veto} to $\text{Veto} = \text{floor}(\log_2(\Delta y) - 4 + N_{\text{veto}})$ and was only applied when $\Delta y > 16$ [109]. For a MoPS trigger, the trigger condition must be satisfied for at least two of the three PMTs of a station. The standard settings of this trigger algorithm are given in Table 3.2.

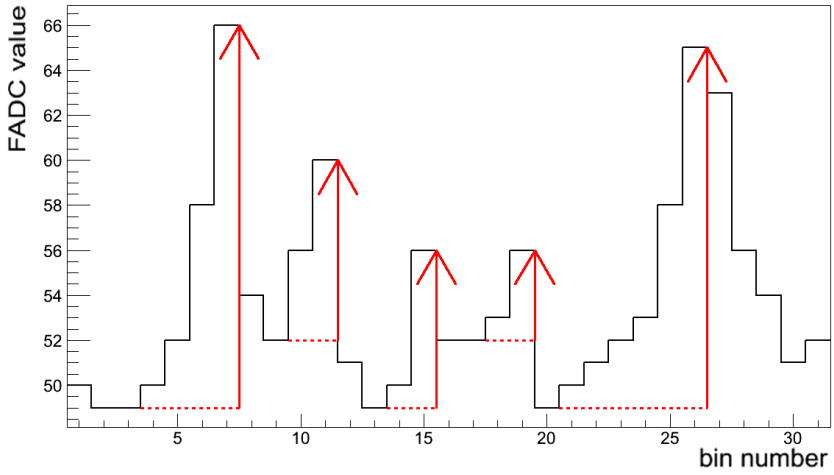


Figure 3.4: Sketch of *positive steps* in the MoPS algorithm: One step (corresponding to one red arrow) is an accumulation of consecutive increases of the FADC trace with a minimum and a maximum size.

Parameter	**value**
y_{min}	3
y_{max}	31
M	4
OFS	3
window	120

Table 3.2: Settings of the MoPS trigger algorithm [107]

T3 triggers

T3 triggers are centrally formed with only the spatial and timing information of T2 triggered stations. To do this, all clusters of T2 signals within a time window of $\pm 25\,\mu s$ are examined for spatial coincidences [99]. When at least three closely clustered stations trigger with at least two TOT triggers or when at least four closely clustered stations with any T2 trigger are found, a T3 is formed. Upon the formation of a T3 trigger, the FADC traces are sent to the central data acquisition, where two additional off-line triggers are implemented [102].

T4 triggers

The T4 trigger is used to test if the signal timings can be fit to a plane shower front moving with the speed of light and is therefore a physics trigger. The trigger conditions make sure that 99 % of the stations containing a physical signal from the shower are kept [99] while keeping the number of random coincidences to less than one per day over the full detector array [102].

T5 triggers

To provide a set of air shower measurements without the possibility of important parts of the shower missing, e.g. because the shower arrived close to the border of the array, the fiducial trigger, T5, was introduced. It requires all six detector stations adjacent to the station with the highest signal to be in operation at the time of the event and is therefore also called 6T5. Some studies only require four or five working adjacent stations, which are called 4T5 or 5T5 respectively [99].

3.2.4 Aperture and Exposure of the Surface Detector

The total aperture of the Surface Detector array depends on the number of hexagons of active detector stations at each moment of data taking. Each of these hexagons consists of one central and six surrounding detector stations, and has an elemental cell with a size of $A_{\text{cell}} = 1.95\,\text{km}^2$ [102]. A sketch of an elemental cell within a detector station hexagon is shown in Fig. 3.5. The aperture per hexagon, a_{cell}, is obtained by integrating

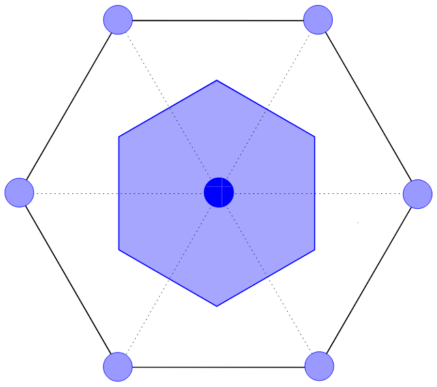


Figure 3.5: Sketch of the elemental cell (blue hexagon) used to calculate the aperture of the SD. Each cell has one central and six surrounding stations [102].

the effective elemental cell area for air showers with a zenith angle of θ over the used solid angle range:[2]

$$a_{\text{cell}} = \int A_{\text{cell}} \cos(\theta)\, d\Omega. \tag{3.3}$$

For all air showers with zenith angles between 0° and 60°, assuming an isotropic flux, the aperture per cell is

$$a_{\text{cell}}(0° - 60°) = 4.59\,\text{km}^2\,\text{sr}. \tag{3.4}$$

For the zenith angle range 30°−60°, used in the SD photon search presented in Chapter 6, the aperture per cell a_{cell} is

$$a_{\text{cell}}(30° - 60°) = \frac{2}{3} a_{\text{cell}}(0° - 60°) = 3.06\,\text{km}^2\,\text{sr}. \tag{3.5}$$

[2] At primary energies below 3 EeV, additional trigger effects need to be taken into account because the 1500 m array has not yet reached full trigger efficiency.

The number of active hexagons, N_{cell}, is monitored second-by-second and can therefore be used for an accurate determination of the SD exposure, A, which is calculated by integrating the aperture over the total time of data taking:

$$A = \int N_{\text{cell}} \cdot a_{\text{cell}} \, dt. \tag{3.6}$$

Additionally, time periods where the detector array was not working properly are labeled as *bad periods* and are removed from the aperture integration.

3.3 Air Shower Reconstruction

The timing and signal sizes from the triggered Surface Detector stations can be used for the reconstruction of air shower geometries and primary particle energies. The energy reconstruction for hadronic air showers uses a calibration derived from hybrid measurements with the FD. However, this method is not applicable to the reconstruction of the energies of photon primary air showers. For this reason, a dedicated photon energy reconstruction has been developed. The geometric reconstruction, however, is the same for both hadron and photon primaries.

3.3.1 Angular Reconstruction

The air shower geometry is derived from a fit of the SD station signal times to a shower front moving with the speed of light. If the number of triggered detector stations is sufficient, a spherical shower front is used instead of a simple planar front [99]. The center of this sphere, interpreted as a virtual point of shower origin, is used together with the shower impact point on the ground to derive the arrival direction of the air shower, with an angular resolution better than 1.6° for events with more than three stations, and better than 0.9° for events with more than six stations [110].

3.3.2 Lateral Distribution Function (LDF)

The impact point of the air shower on the ground is obtained from a fit of the Lateral Distribution Function (LDF) [99]. This function describes the measured signals, S, in the individual detector stations as a function of their distance, r, from the shower axis. It is commonly parameterized using a modified Nishimura-Kamata-Greisen (NKG) function [111, 112]:

$$S(r) = S(r_{\text{opt}}) \left(\frac{r}{r_{\text{opt}}} \right)^{\beta} \left(\frac{r + r_1}{r_{\text{opt}} + r_1} \right)^{\beta+\gamma}. \tag{3.7}$$

The parameter $S(r_{\text{opt}})$ is an estimator of the shower size. It is given by the signal at the distance r_{opt} from the shower axis which is optimized for an accurate shower size determination and depends on the detector geometry [113]. For the SD, r_1 is fixed

at $r_1 = 700\,\mathrm{m}$ and $r_{\mathrm{opt}} = 1000\,\mathrm{m}$ [99]. The parameters β and γ are obtained from a parametrization depending on $S(1000\,\mathrm{m})$ (in short S1000) and θ. Thus, S1000 and the impact point of the shower on the ground are the only remaining free parameters and can be obtained from the fit. An exemplary LDF is shown in Fig. 3.6.

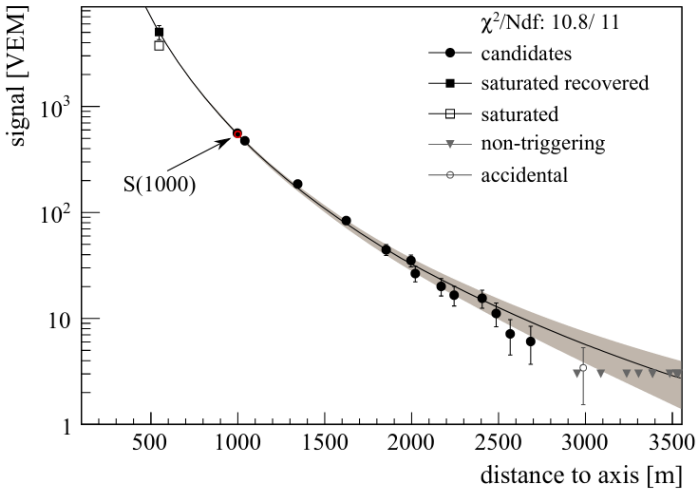


Figure 3.6: Lateral Distribution Function (LDF) from an SD event and its NKG type parametrization [99]

3.3.3 Energy Reconstruction

The standard energy reconstruction relies on an empirically determined relation between S1000 and the shower energy. The first step of this reconstruction is to correct the shower size for its zenith angle dependence. This is done by calculating the equivalent shower size if it had arrived at the median zenith angle of 38°, S_{38}. The conversion is well known and is parameterized with a Constant Intensity Cut (CIC) method[3] using a third degree polynomial, $f_{\mathrm{CIC}}(\theta)$ [115, 116]:

$$S_{38} = \frac{\mathrm{S1000}}{f_{\mathrm{CIC}}(\theta)}, \tag{3.8}$$

$$f_{\mathrm{CIC}}(x) = 1 + ax + bx^2 + cx^3, \text{with} \tag{3.9}$$

$$a = 0.980 \pm 0.004, \tag{3.10}$$

$$b = -1.68 \pm 0.01, \text{and} \tag{3.11}$$

$$c = -1.30 \pm 0.45. \tag{3.12}$$

The relation between S_{38} and the shower energy has been studied using hybrid measurements which make use of the very accurate energy measurement of the Fluorescence Detector. Thus, the shower energy can be estimated from S_{38} with a statistical uncertainty

[3]The energy reconstruction procedures presented here have been used for the starting values of the iterative photon energy reconstruction in Chapter 6. A newer version of the hadronic energy reconstruction is described in [114].

of less than about 16 % and a systematic uncertainty (dominated by the uncertainty in the absolute calibration of the FD) of about 14 % [99, 116]:

$$E_{\mathrm{SD}} = A \left(\frac{S_{38}}{\mathrm{VEM}} \right)^{B}, \text{with} \tag{3.13}$$

$$A = (1.90 \pm 0.05) \cdot 10^{17}\,\mathrm{eV} \text{ and} \tag{3.14}$$

$$B = 1.025 \pm 0.007. \tag{3.15}$$

Due to the unique properties of photon air showers (see Sec. 2.5), the hadronic energy reconstruction is not applicable to photon primaries, and needs to be adapted to account for their different shower development [117]. The procedure used by earlier analyses is based on the concept of shower universality [118], which states that most main observables of air showers (e.g. energy spectra of secondaries, angular distributions of electromagnetic particles [119, 120], and the energy deposit close to the shower core [121]) can be well described as functions of primary energy, shower maximum, normalization of the muon content, and atmospheric depth only. This principle was used to develop an iterative procedure to evaluate the energy of photon-induced air showers in [117] which was refined in [122]. In [123], the parametrizations were updated and corrected for apparent deviations from universality. This iterative photon energy reconstruction procedure will be revisited and explained in more detail in Sec. 6.4.3.

3.4 The $\overline{\text{Off}}\underline{\text{line}}$ Framework

The simulation and reconstruction procedures used in analyses for the Pierre Auger Collaboration make use of the internally developed $\overline{\text{Off}}\underline{\text{line}}$ software framework [124, 125]. The main components of $\overline{\text{Off}}\underline{\text{line}}$ are an event-based[4] data structure, a time-dependent detector description and a collection of *modules* containing physics-related algorithms [125]. The reconstruction of air showers is executed by running modules specified in a *module sequence*: at first, raw data is read in, then simulation and reconstruction algorithms are applied, and finally the reconstructed data is written out. Each module, as well as the module sequence itself, can be configured using XML files. For this purpose, a set of standard module sequences and configurations is provided for most types of analyses depending on the detector components used and air shower types.

The reconstructed shower data is stored in Advanced Data Summary Tree (ADST)-files. ADST is a standalone package based on the ROOT [126] toolkit. It provides the *EventBrowser*, a graphical display for reconstructed event properties, and allows for high-level analyses to be performed on ADST-files directly without the need to modify the $\overline{\text{Off}}\underline{\text{line}}$ source code and perform a full reconstruction of the raw data. Standard analyses are usually implemented in $\overline{\text{Off}}\underline{\text{line}}$ for ease of reproduction, updating, and for availability to the collaboration.

[4] An event corresponds to one air shower.

3.4.1 Simulated Air Showers in Offline

In addition to the reconstruction of measured air showers, the Offline framework can be used to process air shower simulations like those created with CORSIKA (see Sec. 2.7). In this case, instead of raw detector data, simulated air shower data is read in and the detector response to the air shower is simulated. At first, the impact point of the simulated air shower is placed somewhere in the Surface Detector array as the simulated shower is generic in the location with respect to the detector stations [124]. Then, the number, energy, and type of shower particles entering each detector station are determined and their energy loss and Cherenkov light emission is simulated. Then, a simulation of the PMT and detector electronics response follows to create simulated FADC-traces, which are used to determine the trigger response to each detector station and to the full event. Once the detector response to the event has completely been simulated, the same event reconstruction procedures can be applied that are used for real measured data.

3.5 Photon-Hadron Separation

The characteristic properties of photon induced air showers (see Sec. 2.5) have been used to develop observables which are suitable to distinguish photon primary air showers from those of hadronic origin. The main difficulty in detecting photon air showers is the much larger abundance of proton primary showers that can be hard to distinguish from showers of photonic origin in case of larger shower-to-shower fluctuations. In analyses using the fluorescence technique, searches for high-energy photon air showers usually make use of their late development and therefore large $X_{\max}$ values. Since no direct measurement of the shower maximum is possible using a surface detector only, the observables used are based on the lateral and time distributions of secondary particle signals. In [123] it has been shown that, so far, the strongest combination of observables for the SD is the risetime parameter Δ_{Leeds} combined with the LDF-based parameter R_{NKG}. These two parameters are explained in more detail below. This combination works particularly well since it is a combination of two methods of photon-hadron separation which rely on different air shower properties and thus have a comparatively low correlation.

3.5.1 LDF Method

The lower muonic content of photon air showers, as compared to those of hadronic origin, results in them having a steeper decrease of the measured signal with larger distances from the shower core. Several methods to quantify this effect have been developed and are compared in [127]. The parameter used in this work (Chapter 6) is called R_{NKG} and was chosen for its high selection efficiency ($\approx 97\,\%$). It is based on the fact that the NKG-function used to describe the LDF of hadronic air showers overestimates the signals of photon showers at large distances from the shower core. Both, the steep LDF of a simulated photon event, and the NKG-function used in the calculation of R_{NKG},

can be seen in the top of Fig. 3.7. R_{NKG} is calculated by averaging over the relative deviations of the station signals from the NKG parameterized LDF:

$$R_{\mathrm{NKG}} = \frac{1}{N} \sum_{i=1}^{N} \frac{S_i}{LDF(x_i)}, \tag{3.16}$$

where S_i is the measured signal in station i, $LDF(x_i)$ is the signal predicted by the LDF-function fit during the SD reconstruction (see Eq. 3.7) evaluated at the position of station i, and N is the total number of participating detector stations. All station signals are not equally well suited to a robust calculation of R_{NKG}. Therefore, detector stations must meet the following requirements:

- The station distance from the shower axis is larger than 1000 m.
- The LG-signal does not exceed the range of the FADC, i.e. is not saturated.
- No problematic PMTs are used (see Chapter 5).

As long as at least one station meets these requirements, R_{NKG} can be calculated and further used.

3.5.2 Delta-Risetime Method

The late development of photonic air showers in the atmosphere results in a lower average production height of the measured secondary particles. For geometric reasons, particles with the same impact point arrive more delayed with respect to a planar shower front if they were generated close to the ground (see Fig. 3.8). Similarly, a batch of particles produced over a certain path length will be more spread in time if the production was closer to ground [80]. For these reasons, the time spread of shower particles entering a detector station will typically be considerably larger in case of a photon air shower as compared to a hadron primary air shower.

The time spread of particle arrivals in a detector station is quantified using a station risetime parameter $t_{1/2}$ [128]. It is defined as the time it takes to go from 10 % to 50 % of the integrated signal in the averaged calibrated FADC trace of all working PMTs in a detector station (see Fig. 3.9). An asymmetry correction is applied to the station risetimes [123, 129, 130] before using them to calculate their risetime parameter. This correction accounts for the fact that, in case of an inclined air shower, detector stations at the same distance from the shower axis will be hit by the shower front at different shower ages depending on their azimuth angle with respect to the shower axis.

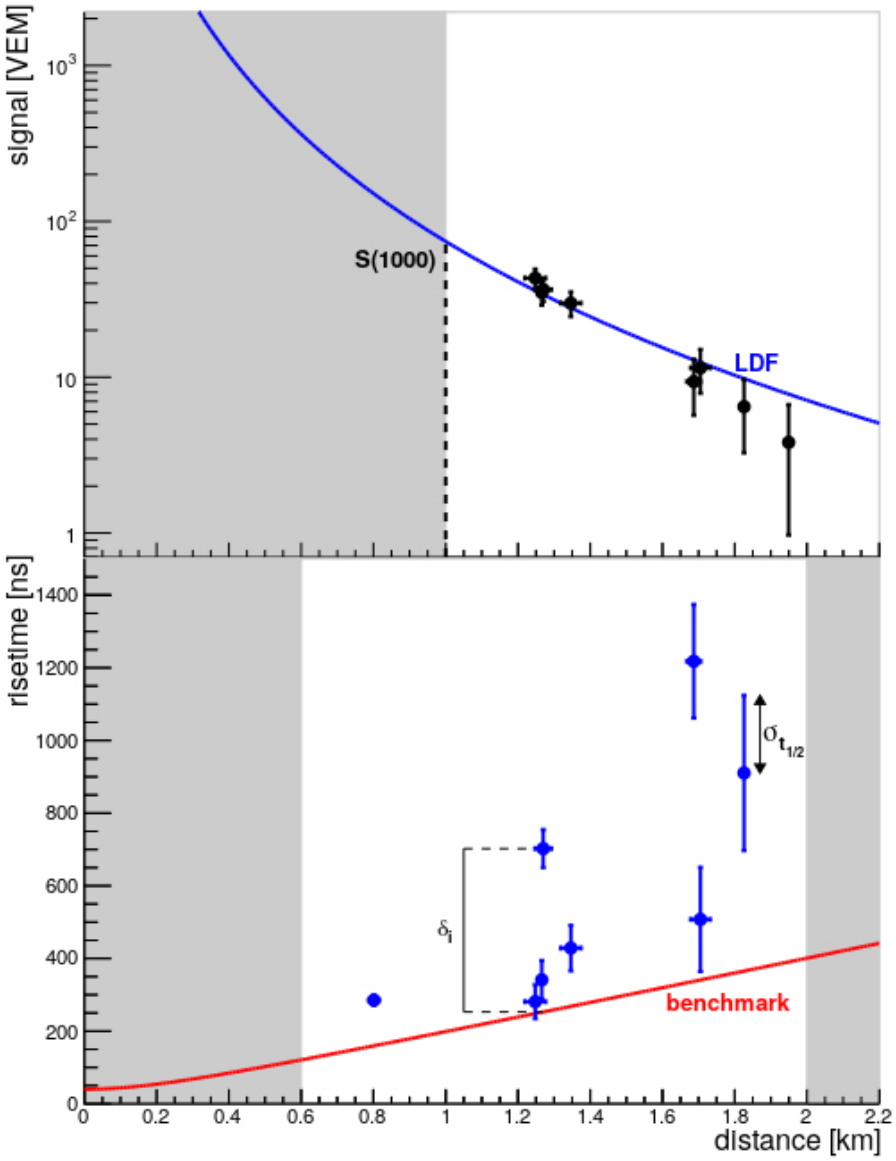


Figure 3.7: Example event to illustrate the determination of R_{NKG} and Δ_{Leeds}. Distance regions not used in the calculation are shaded gray.
Top: R_{NKG} is calculated by determining the average relative signal deviation from the parameterized LDF function for all detector stations at distances above 1000 m from the shower axis.
Bottom: Δ_{Leeds} is calculated by determining the average deviation, δ_i, of the station risetime from the benchmark in units of the risetime uncertainty, $\sigma_{t_{1/2}}$. Only detector stations between 600 m and 2000 m from the shower axis are used.

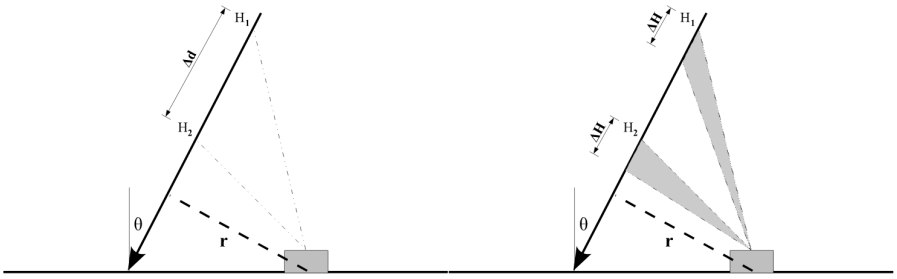


Figure 3.8: Left: Sketch of the shower geometry causing later particle arrival times for a low production height, H_2, compared to a larger production height, H_1. Right: Sketch of the shower geometry which causes larger signal time spreads in the case of a late developing shower: a batch of particles created over the distance ΔH will arrive at the detector with a larger spread in time if it was created deeper in the atmosphere due to the larger path length differences [80].

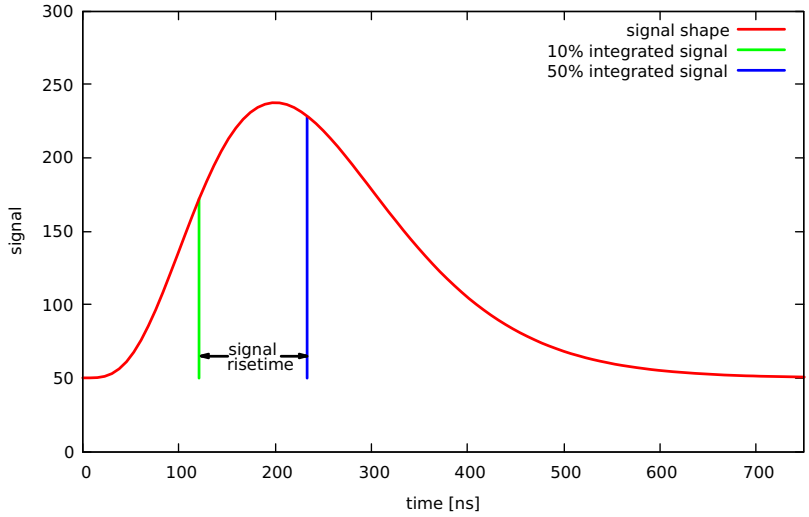


Figure 3.9: Sketch of the signal risetime $t_{1/2}$: the risetime is defined as the time span between 10 % and 50 % integrated signal in a detector station.

The Δ_{Leeds} parameter is a well-established photon sensitive event variable, which combines the individual station risetimes with a parametrization of the average risetime (and its standard deviation) taken from data [123, 129, 130, 131, 132]. The parametrization is called the *benchmark* $t_{1/2}^{\text{bench}}$ and is given as a function of core distance, zenith angle, and energy bin . It is used as a reference value to calculate a deviation, δ_i, of each individual detector station risetime, $t_{1/2}^i$, from its expected value:

$$\delta_i = t_{1/2}^i - t_{1/2}^{\text{bench}}. \tag{3.17}$$

Δ_{Leeds} is calculated by averaging over these N deviations, weighted by the parameterized risetime standard deviations $\sigma^i_{t_{1/2}}$ (see Fig. 3.7, bottom):

$$\Delta_{\text{Leeds}} = \frac{1}{N} \sum_{i=1}^{N} \frac{\delta_i}{\sigma^i_{t_{1/2}}}. \tag{3.18}$$

In Sec. 6.4.1, the most recent parametrizations of the risetime benchmark are compared for applicability in the photon search in Chapter 6. Like in the calculation of R_{NKG}, stations must meet a minimum set of requirements to ensure that the resulting δ_i value is robust. These requirements for a detector station to contribute to Δ_{Leeds} with the chosen benchmark are [123]:

- The asymmetry-corrected risetime is $> 40\,\text{ns}$.
- The signal is above 6 VEM.
- The station's distance from the shower axis is between 600 m and 2000 m.
- The LG-signal does not exceed the range of the FADC, i.e. is not saturated.
- No problematic PMTs are used (see Chapter 5).

Additionally, at least four detector stations must contribute to Δ_{Leeds} for an event to be used in the analysis of Chapter 6.

3.5.3 Other Photon Separation Parameters

A multitude of photon sensitive Surface Detector variables have been used in the past. A short summary of a selection of these parameters is given below.

Radius of Curvature Compared to hadronic air showers, photon air showers tend to have smaller radii of curvature. Since high-energy photon air showers develop deeper in the atmosphere, the virtual point of shower origin for the secondary particles is closer to ground and therefore the path length of secondaries arriving in a detector station increases quickly with distance to the shower axis.

To obtain the radius of curvature R as an event parameter, the trigger times, t_i, are fit to a spherical model by minimizing

$$\chi^2 = \sum_{i=1}^{N} \frac{[c(t_i - t_0) - [R\vec{a} - \vec{x}_i]]^2}{c^2\sigma_t^2}, \tag{3.19}$$

where c is the speed of light, t_0 is the arrival time of the shower in the center of curvature, $\vec{a}$ is the unit vector along the shower axis, $\vec{x}_i$ are the locations of the detector stations relative to the shower core, and σ_t is the uncertainty of the shower arrival time [80]. To reduce the contributions of signals not related to the considered air shower, a software filter is applied while calculating each t_i. The radius of curvature method was combined with RT1000 (see below) in the last published search for a diffuse photon flux with the Surface Detector of the Pierre Auger Observatory [80].

Entity Method In the entity method not only the station risetimes are evaluated, but rather the complete FADC-traces are compared to a parametrization derived from data [133]. This is done in a procedure which consists of five main steps:

1. The mean FADC-traces and uncertainties in data are parameterized depending on the event geometry.

2. For each detector station in an event, the FADC-trace is compared to the parameterized mean trace resulting in a $\mathrm{Log}_{10}(\chi^2/\mathrm{ndf})$-value.

3. The probability density functions of the $\mathrm{Log}_{10}(\chi^2/\mathrm{ndf})$-values are calculated for photon simulations and data.

4. The $\mathrm{Log}_{10}(\chi^2/\mathrm{ndf})$ values are converted to probabilities of a detector station to belong to a data or a simulated photon event, then the probability ratio is calculated.

5. The logarithmic probability ratios are averaged to form an event parameter.

RT1000 The risetime at 1000 m (RT1000) is a parameter which combines the risetimes of detector stations to an event parameter without the use of a benchmark function. Instead, a second degree polynomial is fit to the risetime as a function of the distance to the core. This function is then evaluated at 1000 m from the shower core to get the event parameter [134]:

$$t_{1/2}(r) = 40\,\mathrm{ns} + ar + br^2. \tag{3.20}$$

In [135], the method was revisited and the uncertainty estimation was improved by including a correlation term.

S_b The S_b parameter is based on the total signal in each detector station. It is defined as

$$S_b = \sum_{i=1}^{N} \left[S_i \cdot \left(\frac{r_i}{r_0} \right)^b \right], \tag{3.21}$$

where N is the number of candidate stations, S_i is an individual station signal, r_i is the station's distance from the shower axis, and $r_0 = 1000\,\mathrm{m}$ is a reference distance [136]. The photon sensitivity of S_b is based on the steeper decrease in measured signals with distance from the shower core in case of photon air showers as compared to normal events. The parameter is formed by comparing the signal in each detector station to a reference LDF $(r_i/r_0)^{-b}$. The discrimination power of S_b for photons is maximized at $b = 4$, thus using S_4 [137]. The parameter R_{NKG} was developed from S_b by using the standard LDF (i.e. the NKG-function) instead of the reference LDF and dividing the result by the number of stations N.

Number of Candidate Stations Due to the lower number of muons in photon induced air showers and the steeper LDF, a photon induced shower will typically have a smaller footprint and therefore a lower number of triggered detector stations than a hadronic air shower of the same energy and geometry. For this reason, the number of triggered detector stations is a very simple parameter that can be used in combination with other photon-sensitive variables (e.g. in [79]).

Chapter 4

Trigger Improvements

The ongoing upgrade of the Surface Detector, including an increase of the FADC sampling rate from 40 MHz to 120 MHz (see Sec. 3.2.1), requires a reevaluation of the station-based trigger settings.[1] Since the information available on the inner structure of the SD traces will be increased by reducing the size of the time bins from 25 ns to 8.33 ns, it stands to reason that a better understanding of the measured particles will be possible. This should not only improve the composition sensitivity of the Surface Detector by increasing the time resolution of signal shape sensitive observables, but should also allow for an improvement to the data selection. Ideally, these resolution improvements can also increase the efficiency of the trigger algorithms (see Sec. 3.2.3) without causing an increase of the trigger rates; i.e. selecting a larger number of extensive air showers while simultaneously reducing the trigger rates from the random muon background.

All trigger algorithms currently implemented on the Surface Detector stations depend on a variety of trigger parameters that have been optimized to maximize the trigger efficiency. The communication systems of the SD stations, however, limit the maximum possible trigger rates. The goal of the trigger analysis presented in this Chapter is to develop new trigger parameters which are adapted to the higher sampling rate and specifically have a higher efficiency for triggering on photon-induced air showers.

Due to their small muonic shower component, photon primary air showers have an overall lower trigger efficiency than hadron-induced air showers, and thus should particularly profit from adapted trigger settings (see Sec. 2.5). These new trigger settings could be implemented in addition to the established ones or even replace them if they prove to perform better overall. Due to their inherent focus on the electromagnetic shower component, and thus their natural suitability for photon searches, the ToTd and MoPS trigger algorithms (see Sec. 3.2.3) were chosen to be optimized for triggering on photon air showers.

[1] Currently, the data taken with 120 MHz is sampled down to 40 MHz in a compatibility mode for the triggering.

4.1 Current Trigger Performance

For hadronic primaries, the SD reaches full efficiency at an energy of about $10^{18.5}$ eV. Photon air showers, on the other hand, contain a very small number of muons, and therefore only reach 100 % trigger efficiency at about 10^{19} eV [102, 138]. Generally, this comparatively low trigger efficiency for photon primaries is improved when using the new trigger algorithms ToTd and MoPS. Their usage increases the overall number of triggered air showers and the number of available detector stations far from the shower core. At photon energies between $10^{18.5}$ eV and 10^{19} eV, with the new triggers, the maximum distance at which a station has a not negligible trigger probability increases from ~ 2.5 km to ~ 3.0 km [138]. The energy at which the SD reaches full efficiency is reduced to $10^{18.3}$ eV for hadrons and to $10^{18.5}$ eV for photons with the ToTd and MoPS triggers.

4.2 Trigger Improvement Goals

An increased sampling rate of 120 MHz can potentially increase the efficiency of both the ToTd and MoPS trigger algorithms. Since ToTd is essentially a *time over threshold* trigger, with a higher time resolution the start and stop times of signal windows can be determined more accurately. Furthermore, since the signal in the FADCs is averaged over each time bin, low-energy particles with signals shorter than the time binning are easier to detect if they are measured with a higher time resolution. For this reason, at 120 MHz, low-energy particles can contribute to the ToTd trigger that would otherwise not be detected due to a stronger time averaging of a lower sampling rate. Similarly, the improvement in the detection of low-energy particles can increase the MoPS trigger accuracy as these particles result in larger step sizes.

An increase of the T2 efficiencies would directly affect the number of detector stations per air shower available for analysis. This is beneficial because many photon sensitive observables require a minimum number of active detector stations. The parameter Δ_{Leeds} (see Sec. 3.5.2), for example, requires four active detector stations in a certain distance range from the shower axis. Obviously, an increase of the number of triggered detector stations would directly increase the number of qualifying air showers and would therefore increase the overall efficiency of the analysis.

The primary goal of the trigger studies is to maximize the efficiency of the T2 triggers, while making sure that the amount of data taken stays within the bandwidth limitations of the communication system. This is achieved by setting the current background trigger rates of the ToTd and MoPS trigger algorithms as the maximum rates of background triggers for the examined sets of trigger parameters. For this reason, an increase in efficiency, as compared to the former trigger settings, is only possible with a simultaneous increase of the data selection purity.

4.3 Efficiency Simulation

The efficiency of the ToTd and MoPS trigger algorithms for each parameter set was evaluated using Offline. To do this, detector simulations and reconstructions (see Sec. 3.4.1) were performed on test sets of simulated low-energy photon air showers and the resulting trigger efficiency was calculated. These sets of air shower simulations were created with primary energies of $10^{18.5}$ eV using CORSIKA (see Sec. 2.7). This energy level will result in a mixture of events that will and won't be triggered in the simulation sample for each trigger algorithm applied alone.

In order to measure the T2 trigger efficiency with large statistics and for various detector station positions relative to the shower core, the detector response was simulated with 600 virtual SD stations. Since the air showers were simulated using the thinning technique (see Sec. 2.7.3), the virtual detector stations had to be placed with a dependence on the shower geometry to ensure that only one detector station is used per thinning region (see Fig. 4.1). The detector simulation was performed using an adapted version of Offline modified to use a 120 MHz sampling frequency. In these simulations,

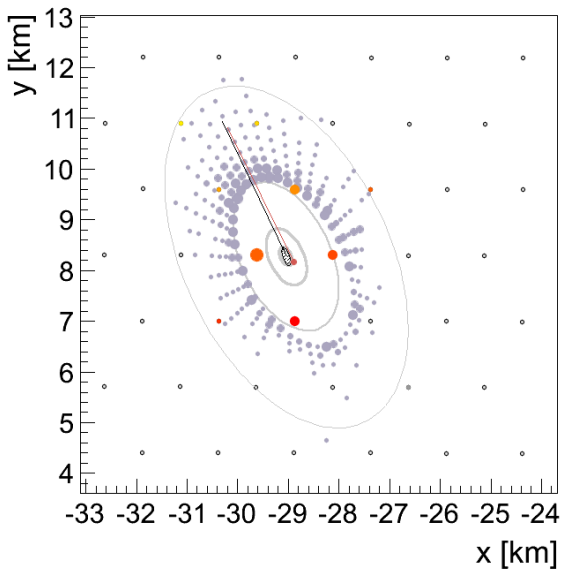


Figure 4.1: Map of the SD station positions relative to the shower core of a simulated photon air shower. Grey dots mark the positions of virtual detector stations included to increase simulation statistics. The red and orange dots indicate the positions of regular detector stations participating in the event. The black and red lines indicate the simulated and reconstructed shower direction.

the percentage of SD stations which fulfill the T2 trigger conditions was used to quantify the T2 efficiency for each set of ToTd and MoPS trigger parameters tested.

4.4 Purity Simulation

Since no data taken with the upgraded detector electronics was available at the time of this analysis, simulated single muons were used to study the background effects on the T2 trigger rates. In nature, so-called *random muons* are generated from very low-energy cosmic ray air showers which are well below the energy threshold of the SD. Because of the high flux of cosmic rays at these low energies, the random muons constitute the main background source for accidental T2 triggers. To simulate this background, single muons with energies of 2 GeV are injected from a hemisphere closely surrounding a modeled SD station in the *ParticleInjector* module of Offline. These simulated muons undergo the same Geant4-based detector and PMT simulation procedures used for the standard detector simulations in Offline. For each injected muon, the FADC traces of all three PMT signals are built with sampling frequencies which have been increased from 40 MHz to 120 MHz. Finally, the trigger criteria of the ToTd and MoPS trigger algorithms are applied to these simulated FADC traces. Because a large number of trigger settings needed to be evaluated, an external program was used to analyze the traces to save computing time. Similar to the estimation of the trigger efficiencies above, the percentage of single muon traces creating T2 triggers was used as a measure of the trigger purity for each set of trigger parameters. However in this case, a high trigger efficiency for single muons indicates a low trigger purity and is therefore undesired.

4.5 Afterpeak Feature

During the simulations for the trigger purity, it was found that the FADC traces of single muons sampled with 120 MHz contain distinct peaks in the signal tail (see Fig. 4.2).

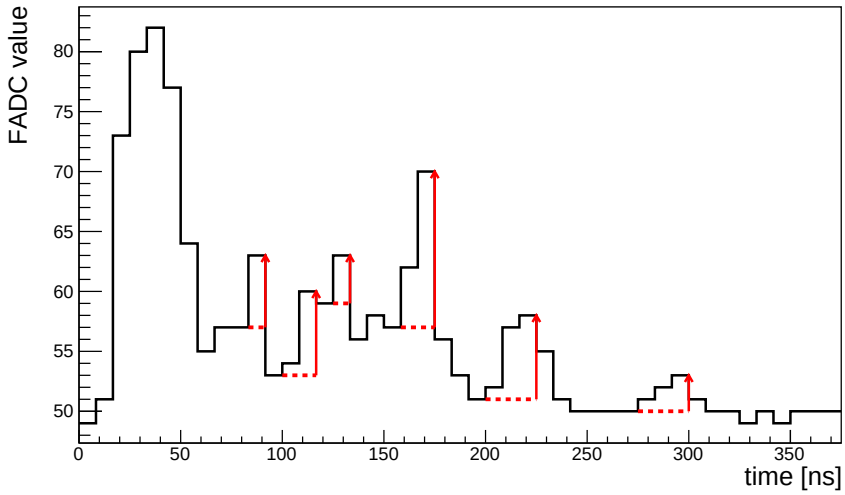


Figure 4.2: Afterpeak structure of a single muon simulated with 120 MHz. Red arrows indicate steps found by the MoPS algorithm.

This *afterpeak* feature was never visible with the original 40 MHz sampling rate and was shown to disappear with the stronger time averaging of 25 ns bins (see Fig. 4.3). Additionally, in order to identify the origin of the afterpeak feature, tests were developed to study whether the afterpeaks were results of formerly unobserved physical processes, a bug in the simulation reconstruction process, or were due to the larger relative uncertainties in the smaller 8.33 ns time bins. These tests are described in the following two Sections.

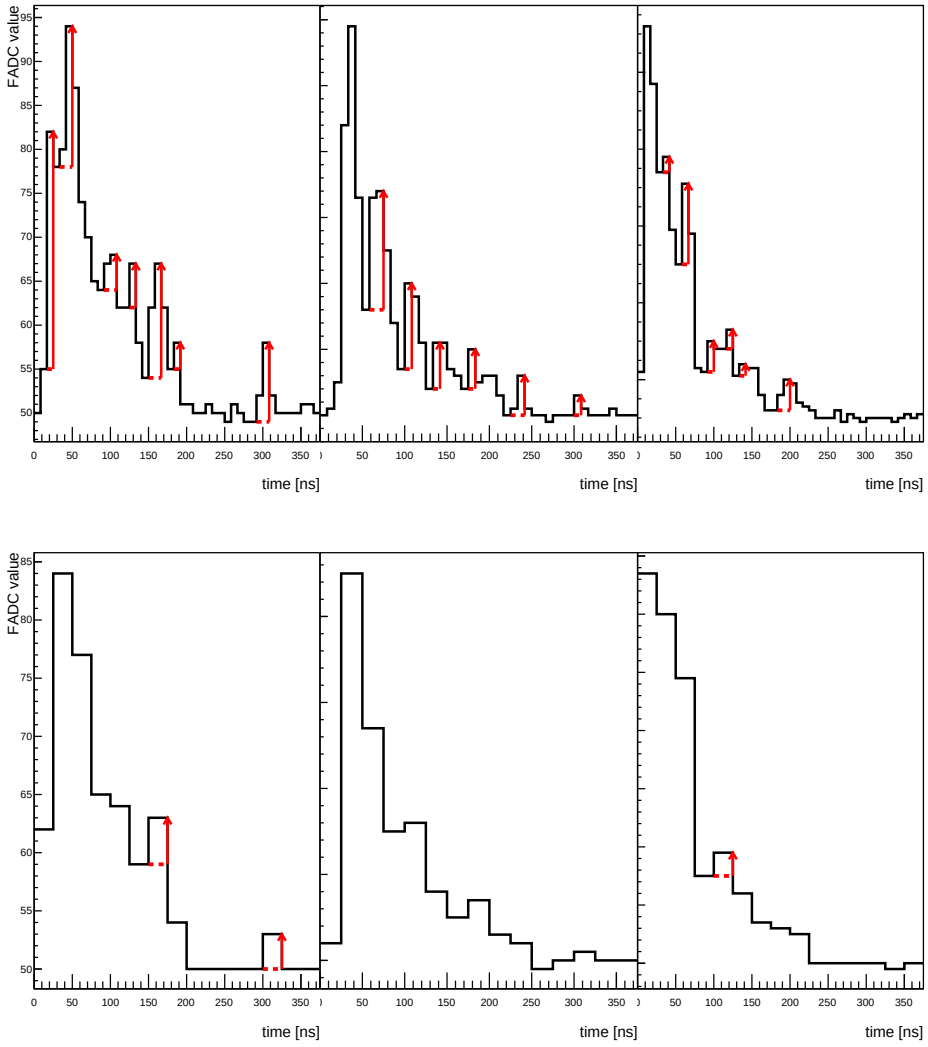
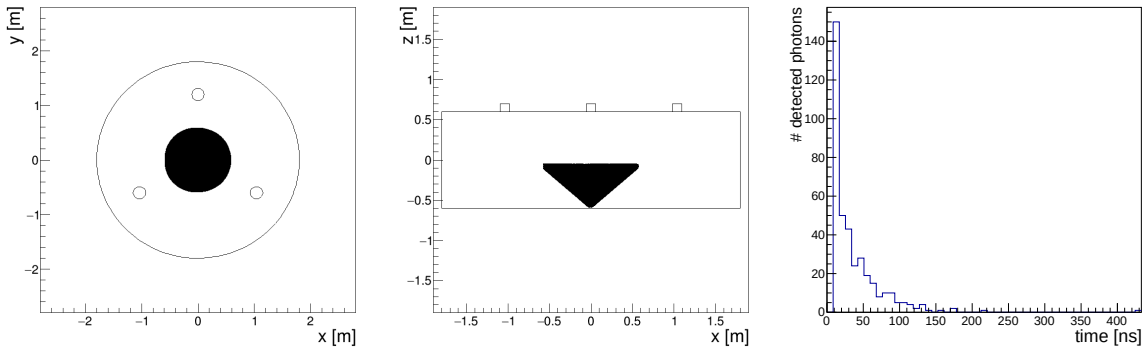


Figure 4.3: Top: 120 MHz simulations of the three PMT traces in a detector station hit by a single muon.
Bottom: Rebinning of the PMT traces above to 40 MHz. The afterpeak features of the 120 MHz simulation vanish in the averaging of the 40 MHz time bins. Red arrows indicate steps found by the MoPS algorithm.

4.5.1 Physical Processes and Detector Geometry

The possibility of physical processes -like the creation of delta electrons, pair production, or muon decay to create additional particles, which then cause the afterpeak feature- was examined by repeating the trigger purity simulation process while disabling physical processes in the Geant4 simulations. Since disabling all interaction processes has no significant effect on the afterpeaks, physical processes were excluded as a cause. Also, by confirming that direct photon hits on the PMTs result in clean, featureless peaks, the PMTs themselves were ruled out as causing the additional peaks. The effects of the specular reflectance of the liner surface was tested by artificially setting all simulated reflections to 100 % diffusivity. This also does not remove the afterpeaking. The possibility of the tank geometry generating signal accumulations after single, twofold, threefold etc. reflections was studied using a standalone simulation. In this simulation, light was propagated through a cylinder and was either diffusely reflected or absorbed on the surface (see Fig. 4.4). The resulting structure in the photon arrival time distribution at the PMTs caused by reflections already vanishes with a 8.33 ns time binning, ruling it out as a possible cause.



(a) Top view. (b) Side view. (c) Number of detected photons per time bin.

Figure 4.4: (a) Top and (b) side view of the detector station geometry in the standalone simulation. The circles on top of the station indicate the PMT positions. The black dots indicate the positions of photons emitted along a moving point. These photons then move along a cone from their points of origin. (c): The distribution of photon arrival times in the PMTs from the standalone geometry simulation.

4.5.2 Photoelectron Traces and Random Fluctuations

The Photoelectron Trace (PE trace) containing the time distribution of photoelectrons generated from a photon hitting one of the PMTs is stored in the Offline software after the Geant4 simulation for the in-detector particle propagation. Figure 4.5 shows

these PE traces for all three PMTs of a simulated detector station with the resulting simulated 120 MHz FADC traces superimposed. The PE traces show that each afterpeak

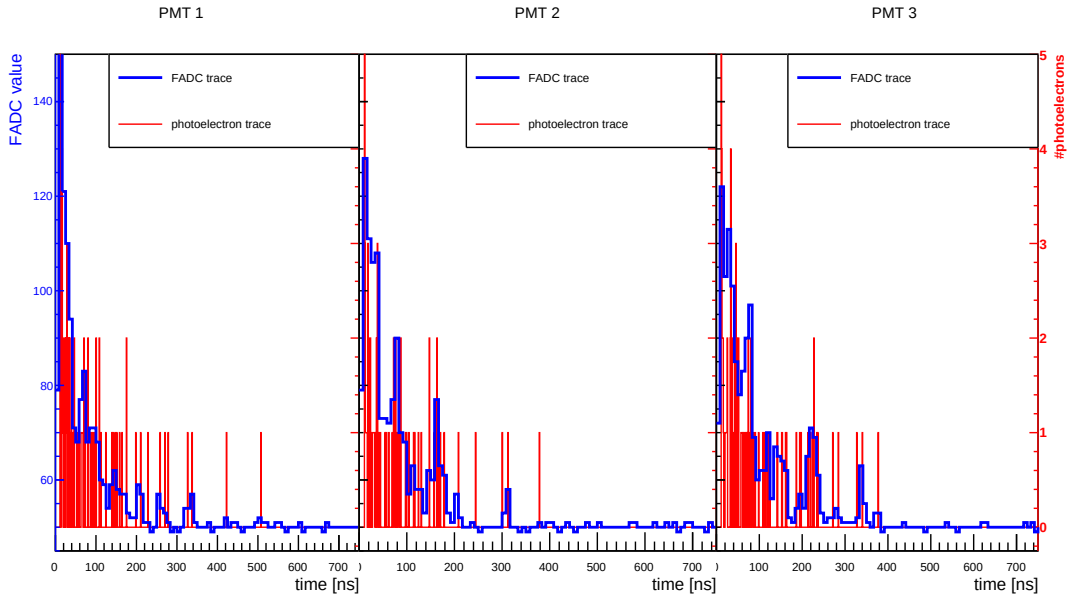


Figure 4.5: Red: Time distributions of generated photoelectrons from a photon hitting one of the three PMTs taken from the Geant4 detector simulation.
Blue: Resulting FADC traces by resampling the PE traces with 120 MHz.

in the FADC traces corresponds to a time frame with a high density of photoelectrons. Therefore, the afterpeaks do not seem to be created anywhere in the simulation process after Geant4. Furthermore, since the same PE traces are used for 40 MHz and 120 MHz simulations, it can be concluded that the feature also affects 40 MHz FADC traces but results in much smaller positive steps which are most of the time not recognized by the MoPS algorithm.

The creation of afterpeaks from the time distribution of photoelectrons was studied by stacking the PE traces of 100 muons vertically injected to a detector station centre. The result is shown in Fig. 4.6. In the first 30 ns there are features indicating the first reflections from the liner surfaces. Besides these, the structure is remarkably featureless and follows the expected exponential light decay. Since the observed afterpeaks in the FADC traces mostly appear later than these very first reflections, the features must be created randomly at varying time positions and thus vanish in the averaging process of the PE trace stacking.

Since the afterpeaks were shown to originate completely from the photoelectron traces, only two scenarios for their creation remain: either there is an effect causing a clustering of photoelectron times at varying time positions in the traces, or the afterpeaks, though quite distinct, have to be created by random fluctuations of the small number of photoelectrons per FADC time bin. In order to test these hypotheses, the mean PE trace (see Fig. 4.6) was used as a probability density function of photoelectron

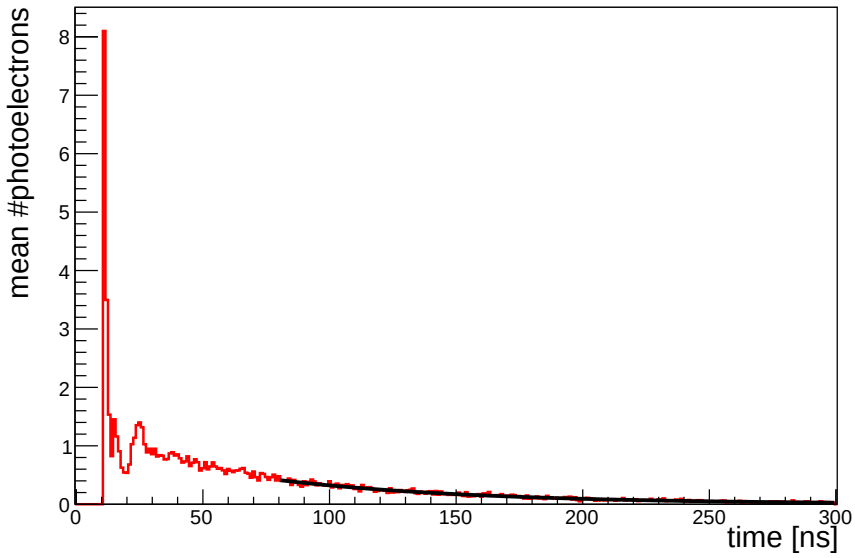


Figure 4.6: Mean Photoelectron Traces of 100 muons injected vertically into the centre of a simulated detector station. The peaks in the beginning of the trace indicate light arrival times after the first reflections on the liner. An exponential fit to the data beginning at 80 ns is shown in black.

arrival times in the PMTs. Artificial PE traces that could be compared to those simulated in Offline were built using these random photoelectron times. The number of photoelectrons forming an artificial photoelectron trace was chosen randomly from a Poisson distribution with the mean set to the average number of photoelectrons per trace (≈ 96) in the Offline simulations.

The comparison between the Offline simulated and the artificial PE traces was done using the distributions of time differences from one photoelectron to the next one in the same trace (see Fig. 4.7). Both distributions are in very good agreement, with a Kolmogorov-Smirnov (KS) distance of 0.0026 and a KS probability of 0.99998. This result indicates that there is no additional effect causing a clustering of photoelectron arrival times in Offline and that the afterpeaks are a feature of random fluctuations in the photoelectron traces.

In a final test, the likelihood of the artificial PE traces creating FADC traces with as many peaks as the Offline PE traces was examined. The number of peaks was evaluated by rebinning all PE traces down to 8 ns time bins and applying the MoPS algorithm. The rebinned artificial PE traces contained 2521 while the rebinned Offline traces only contained 2392 detected peaks. The observed excess of peaks in the rebinned artificial traces was found to be mostly due to peaks which were rather late in the traces. Therefore, it is likely that they are caused by positive fluctuations in the tail of the photoelectron time distribution (Fig. 4.6) which was used to create the artificial PE traces. With a new batch of artificial PE traces, created by replacing the tail (beginning

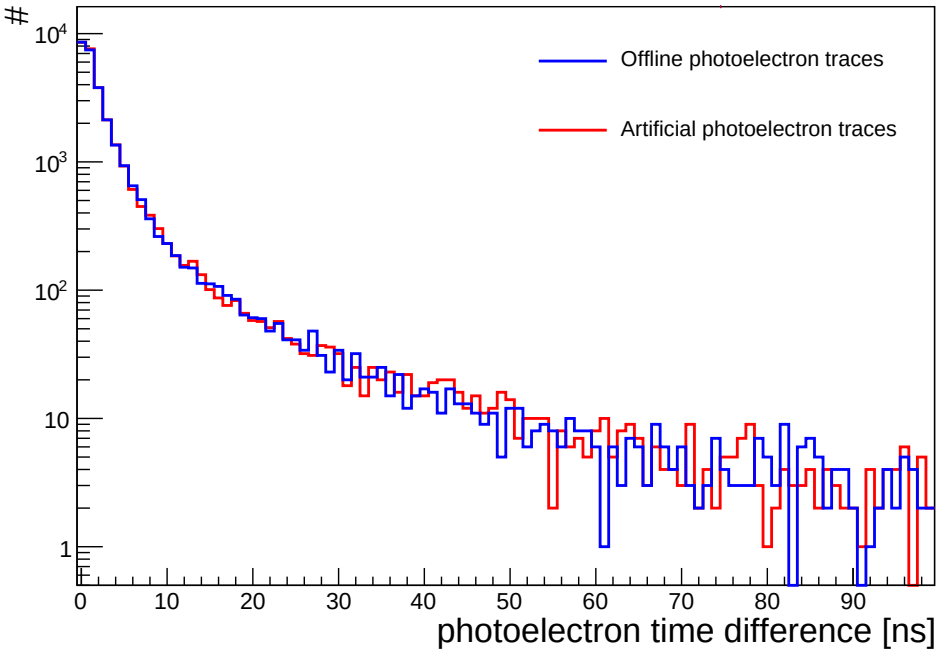


Figure 4.7: Distribution of time differences from one photoelectron to the next one in the same FADC trace for $\overline{\text{Off}}\underline{\text{line}}$ simulated Photoelectron Traces (blue) and artificial Photoelectron Traces with randomized photoelectron times (red)

at 80 ns) of the photoelectron time distribution with an exponential fit, the number of peaks in the rebinned $\overline{\text{Off}}\underline{\text{line}}$ traces could be reproduced very accurately, with the new rebinned artificial traces containing 2388 peaks compared to the 2392 peaks seen in the rebinned $\overline{\text{Off}}\underline{\text{line}}$ traces. This confirms that the afterpeak feature is very likely a result of random fluctuations in the arrival times of photoelectrons and thus of the larger relative uncertainties in the smaller time bins obtained with the 120 MHz sampling rate.

Outside of the simulation studies presented here, the expectation of considerably larger peak numbers with higher sampling rates was confirmed using data from the *AugerNorth* detector [139]. Instead of a 120 MHz rate, they have a sampling rate of 100 MHz (see Fig. C.1), but the effect was nonetheless clearly visible. These AugerNorth detectors had to be used as data taken with the new electronics was not yet available at the time of this analysis.

4.5.3 Results of the Afterpeak Study

Simulated FADC traces of single muons with 120 MHz sampling rate contain significantly more peaks than FADC traces with the former sampling rate of 40 MHz. This increase can be attributed to fluctuations in the small number of photoelectrons in each time bin at 120 MHz sampling rate. For this reason, it is expected that all data taken with the higher sampling rate will contain more peaks than previously expected which makes

the optimization of the MoPS algorithm necessary and more challenging. This is mostly because at 120 MHz, a large number of small peaks does not necessarily indicate a signal predominantly generated by the electromagnetic shower component, but can also be created by peaks in the tails of muon signals.

4.6 Parameter Space Studies

The goal for a new set of trigger parameters is to find trigger settings that result in a higher trigger efficiency for photon air showers without significantly increasing the T2 trigger rates (see Sec. 4.2). Specifically this means that:

- The trigger efficiency for single muons is not higher than that of the currently implemented trigger settings of the ToTd/MoPS trigger algorithms.
- The trigger efficiency for photon primary air showers is as high as possible, but at least as high as with the standard trigger settings at the 40 MHz sampling rate.

Since the ToTd and MoPS algorithms depend on four and five main trigger parameters respectively, there are huge parameter spaces of possible trigger settings[2] for both that need to be searched for optimal settings. As a first step toward exploring this parameter space, for both trigger algorithms, rough multidimensional parameter scans were performed, starting with the parameter sets established for the 40 MHz sampling rate and changing each parameter with constant step sizes.[3] In these scans, parameter sets for the MoPS algorithm were found that triggered on the simulated photon air showers with an up to 8 % higher efficiency than on proton air showers of the same energy and zenith angle. This is remarkable since, at identical MC energies, proton air showers are usually much easier to trigger on due to their larger signals. The parameter sets that preferred photon air showers typically had large N_{veto} values and small *window* values. Large values of N_{veto} can help to reduce the number of peaks counted from a single muon in 120 MHz traces. Low *window* values, on the other hand, can remove late muons from the trigger windows. Furthermore, MoPS parameter sets were also found that fulfilled the T2 rate and purity requirements stated above. The estimated efficiencies and purities of one of these parameter sets, combined with the deviations from changing one of the parameters, are shown in Fig. 4.8. The shown parameter set seems to be more efficient in triggering photons than the established trigger settings and simultaneously has a lower chance to be triggered by single muons in the data traces.

For many parameter sets, including the one depicted, the *background trigger* fraction is zero which indicates a perfectly clean data set. For those parameter sets, the MoPS algorithm does not trigger on single muons at all. Since a perfectly clean data set is unrealistic and since the effect of detector noise on the trigger algorithms could not be studied, real measurements of the background and noise sampled with 120 MHz were needed for reasonably reliable results.

[2] For ToTd the parameters are Th, N_{bins}, *window*, and τ; for MoPS the parameters are y_{min}, y_{max}, N_{veto}, M, and *window*; see Sec. 3.2.3.

[3] Parameters corresponding to a certain amount of time were adapted to the new bin sizes.

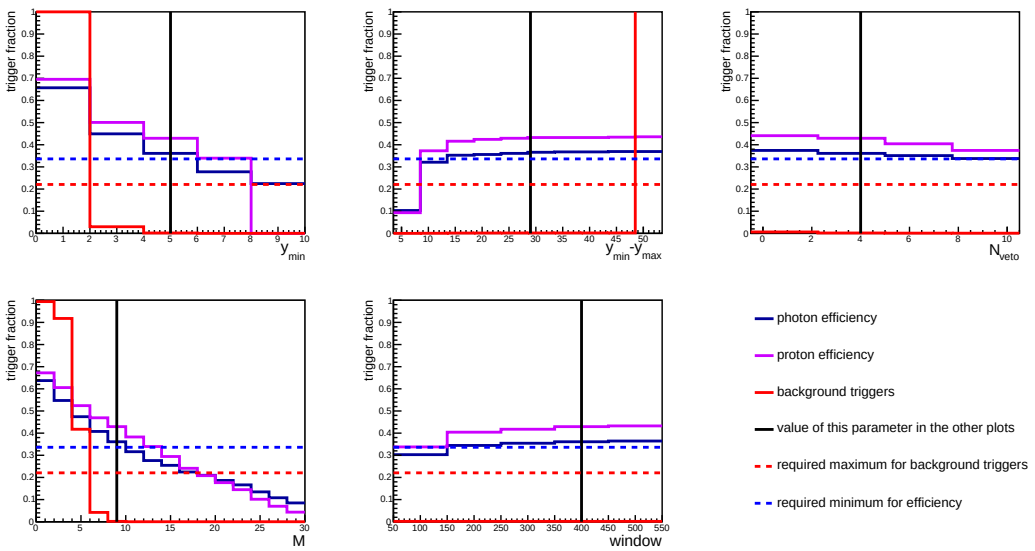


Figure 4.8: Trigger fractions (as defined in Sections 4.3 and 4.4) vs the MoPS parameters. Each image shows the consequences of changing a single trigger parameter on multiple trigger probabilities. For each plot, the value of the other parameters is fixed to the value indicated by the black vertical line. The trigger efficiency for photons (dark blue) is required to be above the blue dashed line indicating the photon trigger efficiency with established trigger settings at 40 MHz. The fraction of single muons triggering the algorithm (solid red) is required to be lower than with the established trigger settings (dashed red). The trigger efficiency for protons is shown in purple.

4.7 Optimization Using Markov Chain Monte Carlo

To increase the efficiency, setting one parameter to a more strict value requires at least one of the other parameters to be released. Some trigger parameters, however, can only be changed coarsely (e.g. the number of MoPS steps), while other parameters can be changed finely (e.g. *window* sizes). The smallest possible change of a coarse parameter might require a big change of a fine parameter to result in a reasonable parameter set. For this reason, the parameter spaces of both trigger algorithms were expected to contain many local minima. Since an automated optimization of the trigger parameters using common minimizers was expected to easily get stuck in a local minimum, the Markov Chain Monte Carlo (MCMC) method was chosen to perform a global optimization. The MCMC method is based on a Markov chain of small and random parameter changes; i.e. every change of the parameters (called a step) only depends on the results of the similar parameter set tested immediately before. If the parameter set after this random step has a better result (e.g. a higher efficiency and a purity still below threshold) the step is automatically accepted. If the step results in a worse parameter set it is still accepted with a probability taken from a Boltzmann distribution depending on how much worse it performs. For this reason the algorithm is able to step out of local minima. If a step

is not accepted, the last parameter set is used again to create the next random step. The convergence of the algorithm is enforced by gradually increasing the temperature of the Boltzmann distribution making jumps to other parameter regions increasingly less likely. Thus, after a while the algorithm stays in the same minimum and starts searching for its exact position. However, to ensure that the global minimum is found, the MCMC process is performed multiple times which makes it possible to estimate the likelihood to be stuck in a local minimum.

A Markov Chain Monte Carlo has been prepared and tested for the optimization of the ToTd and MoPS trigger parameters. The application on the trigger parameter spaces proved to be quite time consuming since every step of the optimizing includes the testing of that parameter set on a large number of simulated detector traces. The MCMC optimizing was not fully performed since a realistic background model for the purity estimation was not yet available at the time of this analysis.

4.8 Outlook

A solid base for an in-depth optimization of the T2 trigger algorithms has been created with the presented studies on trigger parameters for the MoPS and ToTd algorithms. This is especially true with respect to triggering on photon primary air showers. However, since the trigger purity analysis based on single muons yielded unrealistic results, the global optimization could not be fully applied and only preliminary results could be obtained. The afterpeak feature of the 120 MHz data traces makes finding improvements to the current parameter settings more difficult than anticipated, but not impossible. For the MoPS algorithm, parameter sets with increased efficiencies for photons have been found with reduced trigger windows but increased veto values. Since by now real background data sampled with 120 MHz is available, it would be possible to conclude this analysis. The analysis could also significantly profit from large statistics of simulated photon and proton air showers. The simulation set created for the SD photon analysis in Chapter 6 could be used for this purpose. However, due to time constraints further work on this trigger analysis is left to future research and is beyond the scope of this thesis.

Chapter 5

Improvement of the PMT Signal Quality

Searches for rare events, like the SD photon search presented in Chapter 6, depend on starting with a very clean data set in order to identify candidate events. Earlier studies searching for a primary photon flux have shown that the expected number of photon candidate events, as well as the proton background, are on the order of a few events. For this reason, in case no photon flux is detected, a single fake photon candidate caused by detector issues will have a large effect on the resulting flux limits. For example above 40 EeV, where no proton background is expected, a single (fake) candidate would increase the fractional limit from 1.6 % to 2.6 % if a zenith angle range corrected exposure of 39 000 $\mathrm{km}^2\,\mathrm{sr}\,\mathrm{yr}$ is assumed.

Unlike many other analyses which only use the total signal in each detector station, the methods used to search for photon air showers rely on the time structure of the signals within the PMTs as well. This is mostly through utilizing the signal risetime (see Sec. 3.5.2). As such, special care has to be taken since some detector components might have disturbed signal shapes even when they produce a reliable total signal and thus are fine for use in most analyses, but will cause problems in a photon analysis.

Prior to this work, the elimination of these malfunctioning, so-called *Bad PMTs*, from the data sets of sensitive analyses was done using a list of PMTs that was populated manually. In order to have a more reliable elimination procedure, algorithms were developed that can be applied to the FADC traces to find problematic PMTs automatically. Each affected FADC trace found is then flagged in the reconstruction software to prevent it from being used in shape sensitive analyses like the SD photon and neutrino searches. A further advantage of these automated methods is that the PMT flagging can be done on an event-by-event basis instead of flagging a whole time period, which could needlessly veto many events that are not affected. To do this, identifying procedures for typical problems observed so far, the event pathologies presented below, were required.

5.1 Event Pathologies

During the search for PMTs with problematic signal shapes, three main pathologies were identified. Each of these usually affects a PMT for only a certain period of time, often with the effects of the malfunction increasing during the problematic time period (see Fig. 5.6). Nevertheless, even during the most severely affected parts of these time periods, events with usable signals can be found. For example, sometimes the problems with the signals of a PMT appear seemingly at random or sometimes the problems only significantly affect events for which a large signal is recorded in the respective detector station [140].

5.1.1 PMTs with Afterpulses

The FADC traces of PMTs with afterpulses contain additional signal relatively late in the trace which is not found in the FADC traces of the other PMTs in the detector station during the event. An example of this behavior can be clearly seen in Fig. 5.1. For short signals these afterpulses are well separated from the real signal but for signals

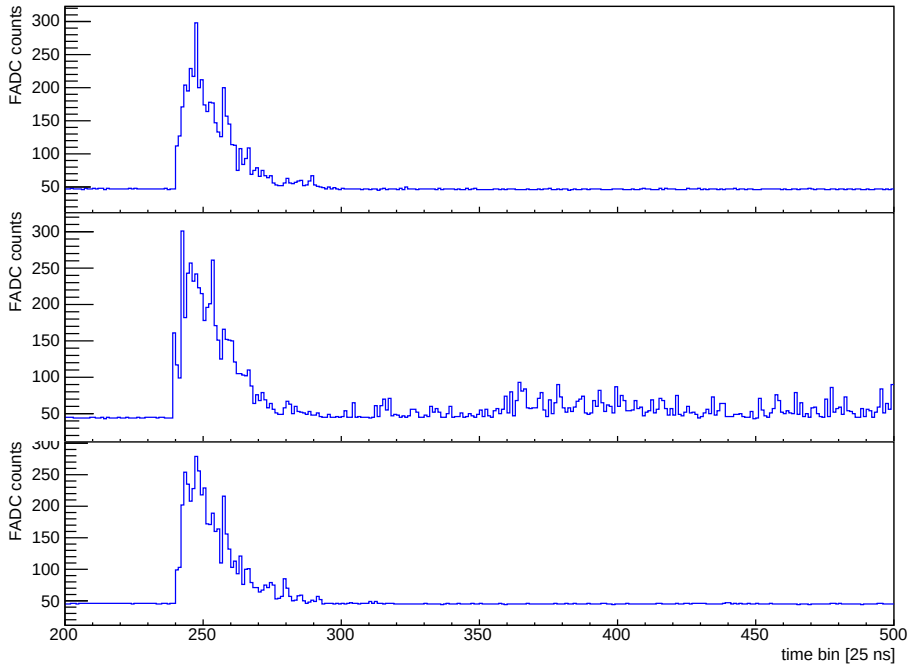


Figure 5.1: FADC traces for the three PMTs of a detector station. The middle PMT shows a strong afterpulse feature.

larger than $\sim 20\,\mathrm{VEM}$, it can overlap with the real signal and would be considered as part of the signal in the event reconstruction. The size of the afterpulses is proportional to the real signal. For this reason, the afterpulse pathology is most severe in FADC traces with a large signal.

5.1.2 PMTs with Oscillating Baselines

The second type of PMT pathology is defined by an oscillation pattern in the baseline. While FADC traces with a weak oscillation can still be usable, some PMTs have oscillations strong enough that they need to be rejected for further analysis. The oscillations are most often caused by the PMT electronics since the feature usually ceases after exchanging them [140]. Another possible cause of these oscillation patterns are lightning events but those are considered a separate issue. Excluding lightning events, oscillation PMTs with three varying sub-features have been identified:

- PMTs with regularly oscillating baselines.
- PMTs with an oscillation which includes a modulation.
- PMTs with an oscillation within the signal.

Even though the origins of these problems are likely different, all of them can be identified by the single procedure outlined in Sec. 5.2.2. The FADC traces of a detector station affected by an oscillating baseline feature are shown in Fig. 5.2.

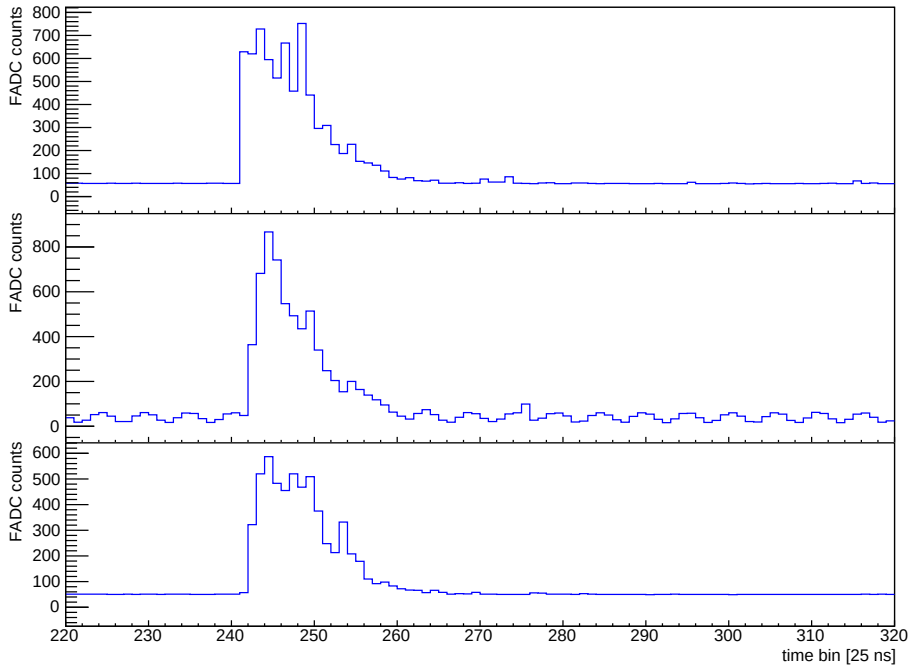


Figure 5.2: FADC traces for the three PMTs of a detector station. The middle PMT shows an oscillation pattern.

5.1.3 PMTs with Large Decay Times

The third type of PMT pathology is defined by traces with a very slow decrease of the signal following each signal peak. An example of this feature is shown in Fig. 5.3. Similar to the afterpulse feature, the impact of this problem is stronger for events with a large signal. The decay time feature can have a very large effect on signal risetimes which is why its identification is especially crucial to the photon analysis.

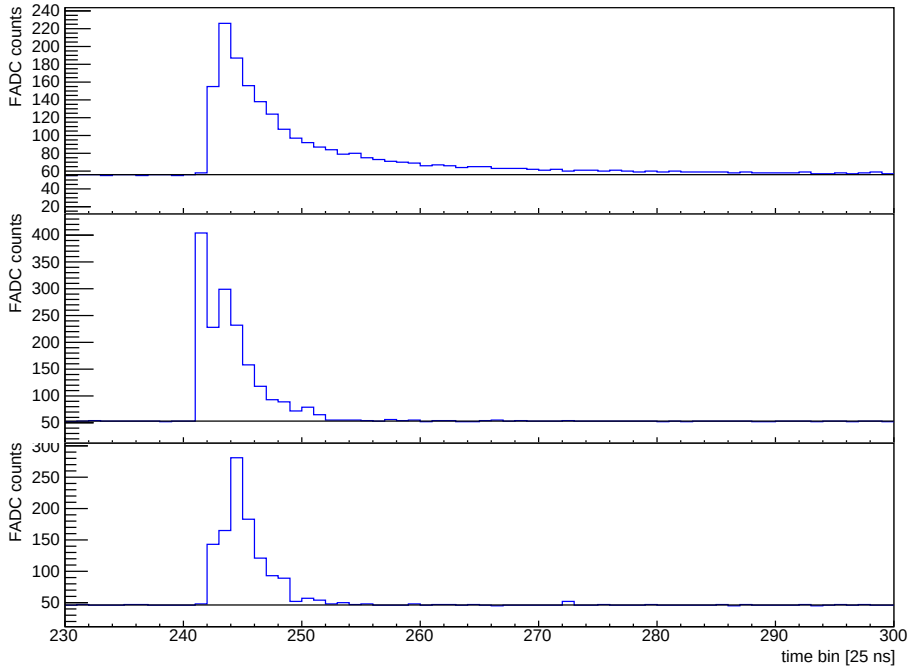


Figure 5.3: FADC traces for the three PMTs of a detector station. The top PMT shows a very large decay time.

5.2 Filtering Procedures

Filtering procedures, which can be applied on an event-by-event basis, have been developed, tested, and implemented for all the above pathologies. These procedures have become the standard method for data cleaning in neutrino and photon searches. The individual filtering mechanisms are described below.

5.2.1 Filter for Afterpulses

The identification of afterpulse events is based on the calculation of a *late signal* in the FADC traces. This late signal is defined as the integrated signal between bins 350 and 600. To make sure no *direct light*[1] is included, the contents of the bin with the highest signal and its two neighboring bins are removed. The filtering mechanism then applies the following three criteria to the late signal:

- The late signal must be at least 5 % of the total signal in the FADC trace.
- The late signal must be at least 7 times larger than the late signal in any of the other working PMTs.
- The late signal is larger than 1 VEM.

All of these criteria have been optimized to keep as many usable traces as possible while ensuring a very clean data set. Due to the second criterion, this filtering mechanism can only be applied to detector stations with at least two working PMTs. This requirement is common in these filters as the comparison to the other two PMTs in each station has proven to be very efficient in separating PMT problems from other fine, but unusual signal forms.

5.2.2 Filter for Oscillating Baselines

The filter dedicated to identifying oscillating baselines uses the RMS of the baseline, which is calculated during the calibration process of each detector station. Two cuts on this RMS value already existed: a cut requiring the RMS to be below 1.75 FADC counts on a daily level and a cut requiring less than 12 FADC counts on the event level. The event level cut turned out to be too lenient, while the daily cut failed to filter out most problematic events since the issue does not necessarily affect a large enough number of events in a given time period to raise the RMS for the full day [140]. Because of these features, after testing the effects of different cut thresholds, a new event level cut at 1.5 FADC counts was introduced.

5.2.3 Filter for Large Decay Times

To find PMTs with a large decay time, a filtering mechanism has been developed that uses both the calibrated FADC trace and its baseline. This filtering mechanism was inspired by the ToTd triggering mechanism (see Sec. 3.2.3). It is based on the fact that large decay time FADC traces will contain many more *negative steps* (i.e. decreases from one bin to the following bin) than FADC traces of regular events. Since regular FADC traces contain many negative steps as well (due to light decay in the water tanks, see Sec. 3.2.1), a deconvolution with an exponential, which removes all tails with the

[1] Direct light is a term for large signals in one or two bins of a single PMT due to Cherenkov light in the tank traveling directly to the PMT instead of after scattering off of the liner.

standard decay time of $\sim 70\,\text{ns}$ is performed. This deconvolution is identical to the one used in the ToTd algorithm described in Sec. 3.2.3, which was

$$d_i = \frac{(a_i - f a_{i-1})}{(1-f)} \quad \text{with} \quad f = \exp(-\Delta t/\tau), \tag{5.1}$$

where d_i indicates the signal in the i-th bin of the deconvoluted trace, a_i is the (baseline subtracted) signal in the i-th bin of the regular trace, Δt is the bin width, and τ is the time constant set to 70 ns.

After deconvolution, the resulting trace can be interpreted as a time distribution of particles in the detector station [108]. An example of the effects of this deconvolution on a normal, as well as a large decay time FADC trace, is shown in Fig. 5.4. The deconvolution procedure creates traces with hardly any tails for normal PMTs but still leaves shortened tails for large decay time PMTs. This difference is used to improve the discrimination of the next step of the filter algorithm.

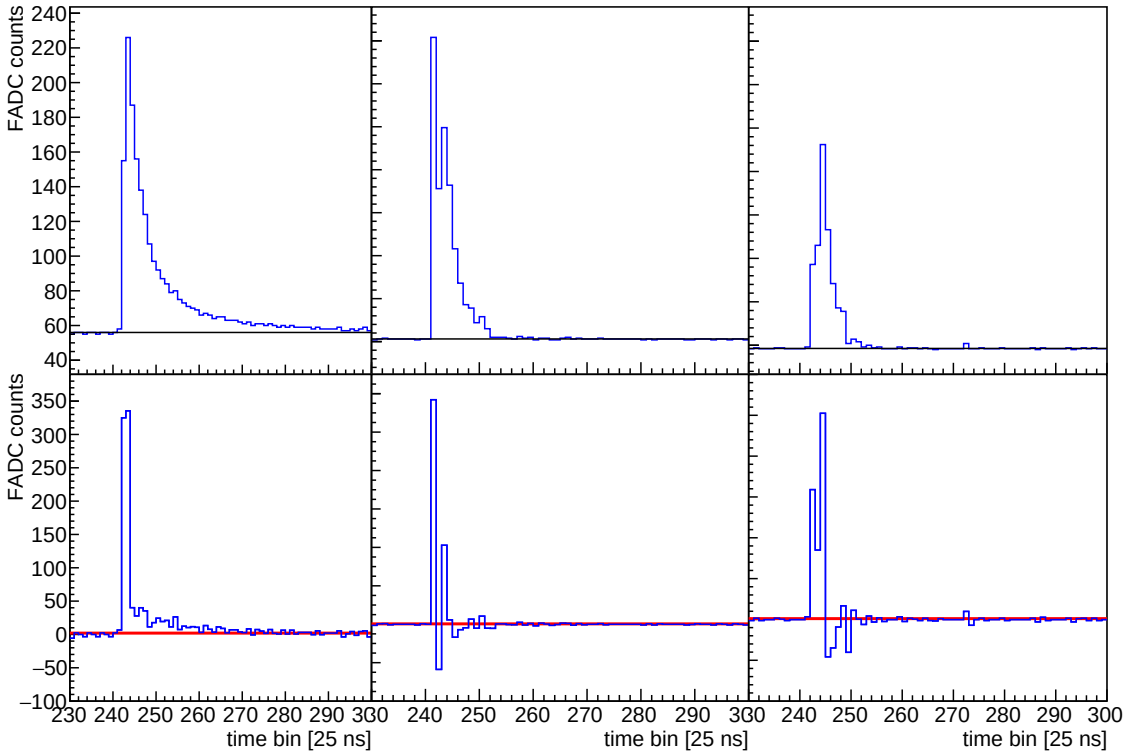


Figure 5.4: Top: Signal traces in a detector station with a decay time feature in the left PMT
Bottom: Deconvolution of the signal traces above: while the tails in the middle and right PMTs vanish, the tail in the left PMT only gets shorter. The red line indicates the minimum signal for a negative step to be counted in the large decay time filtering mechanism.

Next, the remaining tails of the problematic PMTs are quantified by counting the number of negative steps remaining in the deconvoluted traces. Here, a negative step is only counted if there is a decrease of the signal from one bin to the next bin in both the deconvoluted trace and the original trace:[2]

$$d_i > d_{i+i} \wedge a_i > a_{i+1} \wedge d_{i+1} > 2 \text{ FADC counts.} \tag{5.2}$$

Furthermore, in order to avoid counting baseline fluctuations, a negative step is only found if the second (lower) bin of the step is at least 2 FADC counts above the baseline as illustrated in Fig. 5.4.

An imbalance, Δ_i, in the number of negative steps between each PMT in a detector station is calculated:

$$\Delta_i = \frac{N_i - \frac{N_j + N_k}{2}}{\sigma_N}. \tag{5.3}$$

In this Equation, N_i denotes the number of negative steps in the i-th PMT and σ_N is the uncertainty of the mean for the N_i values of all working PMTs.
The final criteria for issuing a large decay time flag in the i-th PMT are:

- $N_i \geq 16$.
- $\Delta_i \geq 8.5$.

In $\overline{\text{Off}}\underline{\text{line}}$, the individual flags for each channel (HG and LG) were produced using the other FADC traces from the same gain. This becomes important if the event reconstruction uses, for example, the HG channel of one PMT and the LG channels of the other PMTs.

5.3 Performance of the Identification Procedures

The effects of the filtering mechanisms were evaluated by examining their results on already known Bad PMTs in their affected time windows. Additionally, the number of traces which were identified where none should be expected was considered. A summary of the results is found below.

5.3.1 Performance of the Afterpulse Filter

About 0.8 % of all reconstructed FADC traces were flagged by the afterpulse filter resulting in a very high detection efficiency of the afterpulse feature. In fact, if the signal was above 20 VEM, every single trace in a set of known malfunctioning PMTs was detected in the affected time interval. Furthermore, all malfunctioning PMTs with a signal above 5 VEM were also detected if a subset of particularly strongly affected PMTs was considered [140].

[2] Note: in this algorithm, a step has the size of a single bin. This is in contrast to the MoPS algorithm, where a step usually consists of multiple bins.

5.3.2 Performance of the Oscillation Filter

The newly implemented filtering mechanism for oscillating baselines flagged only about 0.1 % of all PMT traces. Among the PMTs with oscillating baselines, many FADC traces with small baseline fluctuations (resulting in a small baseline RMS) can be found which are still usable and therefore do not need to be cut from data. Strongly affected traces, on the other hand, are often already flagged by the long decay time and afterpulse filters.

PMTs which were wrongly identified by this filter were often found to have a large amount of negative signal due to misreconstructions of the PMT baseline. These PMTs are technically fine, but the misreconstructed baseline will of course cause problems in calculations of other parameters (e.g. the signal risetime) as well. Therefore, their elimination from shape sensitive analyses is considered a positive additional feature of the oscillation filter mechanism.

5.3.3 Performance of the Large Decay Time Filter

While most Bad PMT signatures are quite striking by eye, with the new large decay time filter even rather subtle examples of affected FADC traces could be found an example of which can be seen in Fig. 5.5. However, on a test sample of well known large decay time

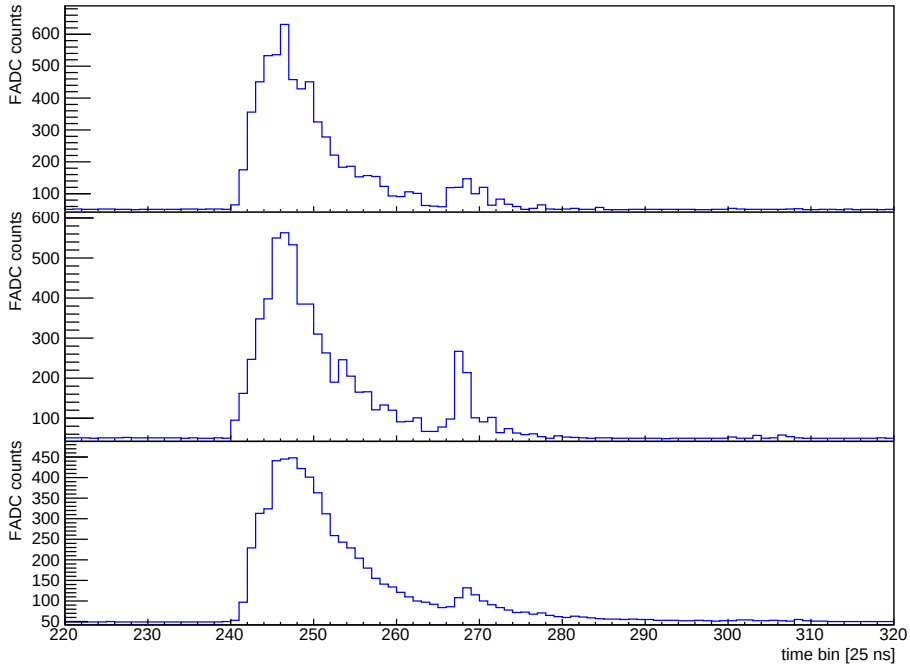


Figure 5.5: FADC traces for the three PMTs of a detector station. The bottom PMT shows a subtle large decay time feature.

PMTs in the affected time periods only about 25 % of the FADC traces were flagged. This is a net positive, because even during these times most events are not significantly affected by the decay time feature and therefore do not need to be cut from further

analysis. For this reason, many not affected FADC traces could be recovered for shape-sensitive analyses by the introduction of the new event-by-event filtering mechanisms. Additionally, many raining PMTs[3], with FADC traces that were affected by the raining feature, were found. The removal of these strongly raining PMTs is a secondary beneficial feature of the decay time filter.

Overall, 0.3 % of all PMT traces were flagged by this filtering procedure. The time distribution of the identified FADC traces for data until December 2015 is shown in Fig. 5.6. The horizontal lines show problematic PMTs. Some of them have strong features and were already known (e.g. station 602 and 762). Many of the other, weaker lines denote newly found PMTs (e.g. station 194 and 984). Overall, 63 formerly unknown problematic PMTs have been identified. As mentioned before, not all events during an affected time period are problematic. However, sometimes a feature disappears at some point only to reappear later in the data. With the new filtering algorithms these changes over time are dealt with automatically.

5.4 Summary of the Filtering Results and Expected Effects on Photon Analyses

Automated filters for all three problematic PMT pathologies have been developed and extensive tests have been performed. Since the former treatment of Bad PMTs involved flagging time periods, and therefore many unaffected FADC traces, a direct test of the algorithms by a one-to-one comparison of results was not possible. For this reason, all filtering mechanisms were tested by examining the detection rates of well known affected PMTs and time periods. Additionally, many events which were not filtered out in those time frames were screened by eye for problematic signatures.

Checks on PMTs which received a flag without being affected by one of the features were also done. These indicate that the oscillation filter and the large decay time filter are both also sensitive to misreconstructions of the PMT baseline. In practice, cutting these rare events from shape sensitive analyses is a welcome side effect since an incorrect baseline estimation affects shape sensitive observables as well. The overall success in finding new methods to automatically detect problematic PMTs made manually updating the Bad PMT listings in $\overline{\text{Off}}\underline{\text{line}}$ obsolete. Furthermore, the new procedure eliminates the possibility of newly appearing Bad PMTs affecting sensitive analyses before they are found.

For photon searches, a reliable detection of the afterpulse and especially the large decay time features is crucial since they directly increase signal risetimes and thus can easily cause fake photon candidates. Fortunately, the filters for these pathologies proved to be very reliable and many formerly unknown problematic PMTs have been identified. For the SD photon search in Chapter 6, the effect that unflagged events have on the signal risetime was evaluated. This was done by examining a set of events which contain

[3] A *raining PMT* is a PMT with a feature causing a drop of $I_{\text{VEM}}^{\text{peak}}$ (see Sec. 3.2.2) reoccurring on the order of tens of minutes.

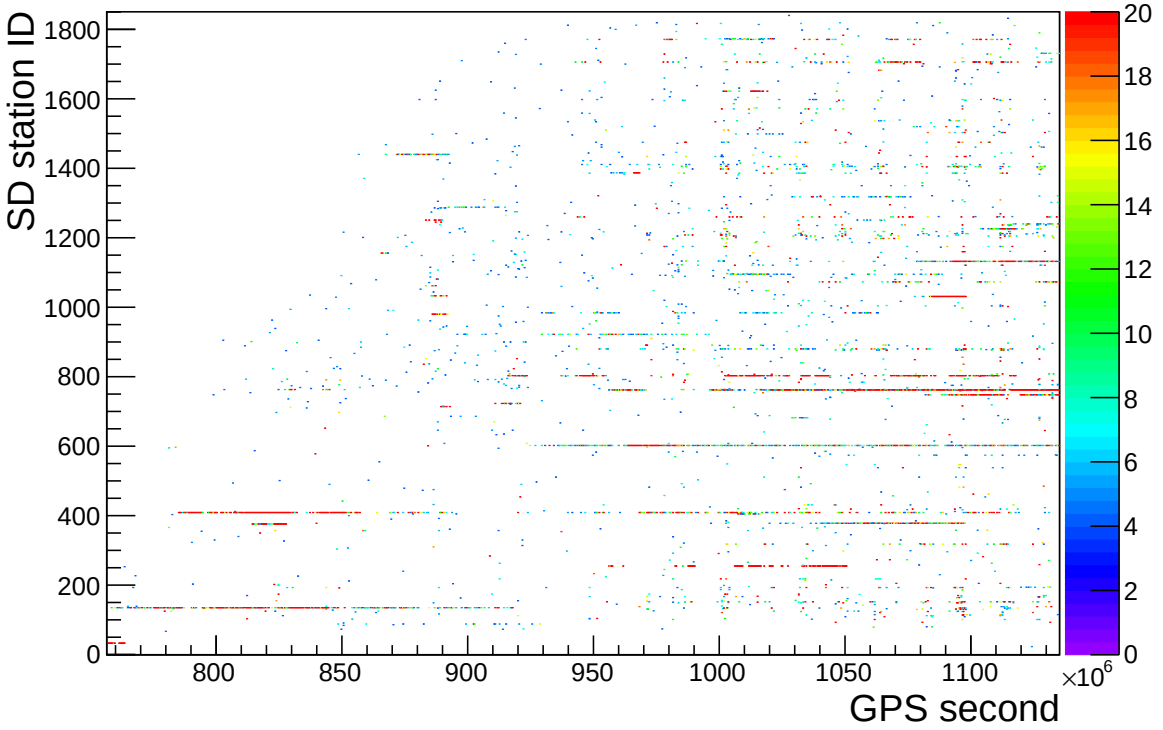


Figure 5.6: 2D histogram of event times and SD station IDs of PMTs identified with the new filtering algorithm for large decay times applied to data measured between 2004 and 2015. The color scale indicates the number of identified traces in the according time bin. Horizontal lines correspond to problematic PMTs, often with strong decay time problems. The start and end of their problematic time periods are visible even though not every FADC trace in this time period must be affected. Additionally, the low number of randomly flagged PMTs is visible in the small number of isolated hits on otherwise non-problematic PMTs.

a PMT which is currently in a period where it is experiencing a decay time feature that was not flagged by one of the algorithms. For each of these events, the risetime of the affected PMT has been compared to the risetimes of the other intact PMTs. The results of this comparison are visualized in Figs. 5.7 and 5.8. Even though the affected PMTs on average had slightly higher risetimes ($\approx 10\,\%$), this effect is considered to be small enough to be unproblematic for the photon search. This is because neither the relative strength nor count of long risetime outliers significantly increased among the affected PMTs. Also, an examination of FADC traces with a visible decay time feature that were not flagged by the algorithm revealed that the signal shapes of these traces are not much different from those in the other PMTs. For this reason, these events are considered unproblematic for usage in shape sensitive studies. An example of this type of event is shown in Fig. 5.9.

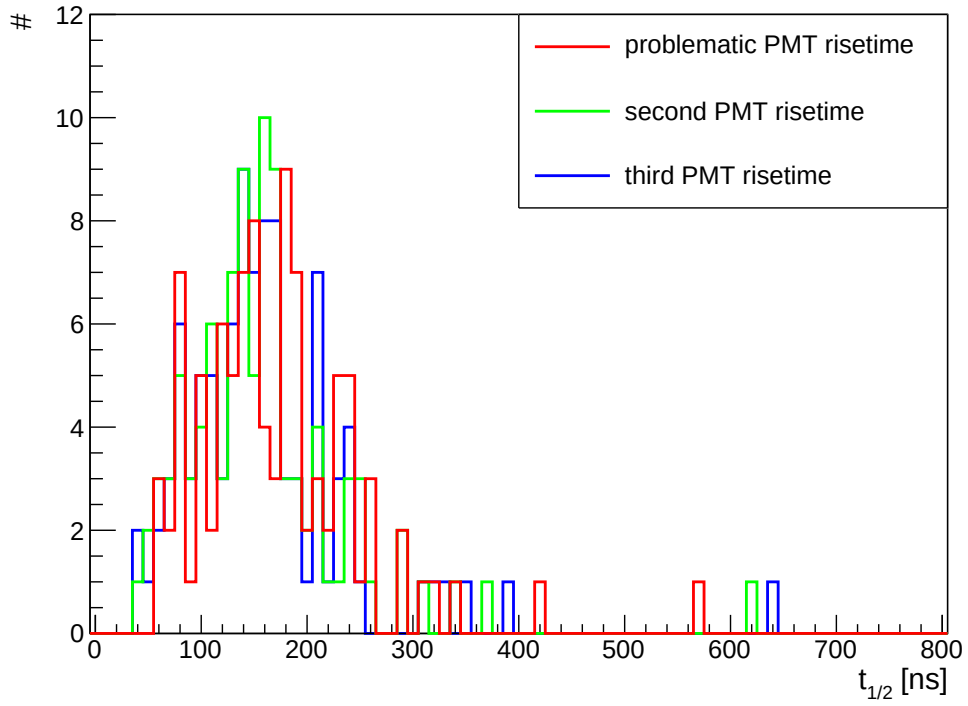


Figure 5.7: Individual risetimes of the three PMTs in decay time problem affected stations and time periods for events that have not been flagged by the filtering mechanism. The risetime of the FADC trace with a possibly undetected decay time problem is shown in red while the other PMTs' risetimes are shown in green and blue.

5.5 Implementation in the $\overline{\text{Off}}\underline{\text{line}}$ Framework

The filters for malfunctioning hardware components were included in a newly introduced $\overline{\text{Off}}\underline{\text{line}}$ module called *SdPMTSignalShapeQualityChecker*, and the parameters used in the algorithms have been internally published in [140]. Since the effects of the studied pathologies on the total station signals are usually small, this module is intended to be primarily used for signal shape sensitive studies.[4] The flags used to ensure the removal of affected PMTs, or PMT anodes, have already been established and are used e.g. by the *SdCalibrator* module[5] making the usage of the new $\overline{\text{Off}}\underline{\text{line}}$ module as simple as possible.

[4]The SdPMTSignalShapeQualityChecker module is included in module sequences using the *SdPreSelectionShapeSensitive* standard sequence.

[5]The PMT flags can be accessed using "→IsTubeOk()" and "→IsLowGainOk()" on the PMT calibration data of the affected event.

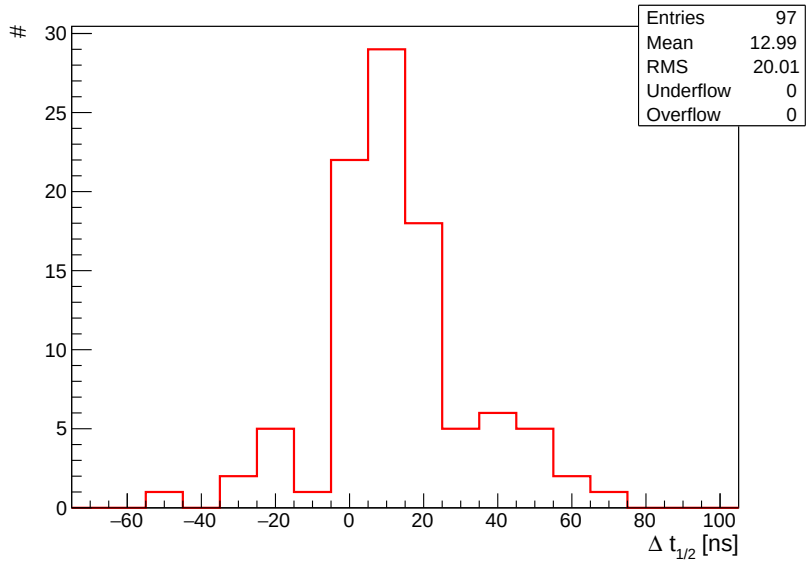


Figure 5.8: Risetime differences in decay time problem affected stations and time periods for events that have not been flagged by the filtering mechanism: for each of these events the difference of the affected PMT's risetime to the mean of the other two PMTs' risetimes is shown.

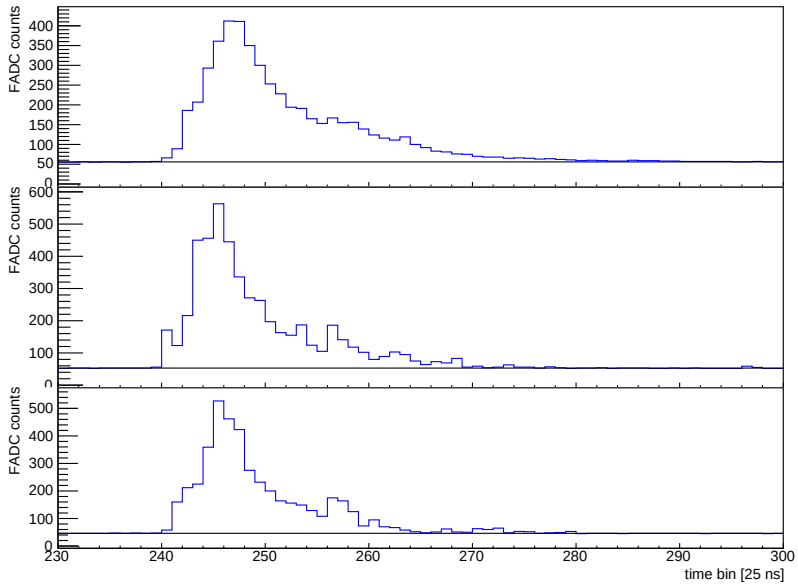


Figure 5.9: Typical example of a decay time problem affected FADC trace not found by the filtering algorithm: even though the top PMT is clearly affected by a large decay time feature it did not receive a decay time flag. Since the shape of the signal is only slightly affected, it is considered to be still usable for shape sensitive analyses.

Chapter 6

Search for Ultra-High-Energy Photons

The main goal of the analysis presented in this thesis was to search for a diffuse flux of primary photons at energies above 10^{19} eV using the Surface Detector of the Pierre Auger Observatory. This search has been performed using the data period spanning from Jan 1st 2004 to June 30th 2018 and a set of simulated photon air showers created with CORSIKA using the interaction models EPOS LHC and FLUKA (see Sec. 2.7). Before using the data in the photon search, it was cleaned using the filtering mechanisms presented in Chapter 5. This version of the analysis has already been presented at the International Cosmic Ray Conference (ICRC) 2019 [2], however, a more detailed description of the procedure used to search for primary photons and eventually to set upper limits on the diffuse photon flux and fraction is presented below. In addition, multiple improvements to the photon search are presented which have not been used in [2] but are applied for the publication by the Pierre Auger Collaboration.

6.1 Principal Component Analysis (PCA)

The search for a diffuse flux of primary photons was performed with a Principal Component Analysis (PCA) which used models of the expected signal and background. In this process, signal and background are unseparated and treated equally. This is an important feature of the PCA for the SD photon search since it is not known if the experimental data contains photon events. Because of this, a clear separation of signal (photon MC) and background (data) is not possible.

The PCA is based on finding the axis in the two-dimensional input parameter space along which the variance of the data/simulation combined data set, is maximized. This axis is consequently called the PC axis and the projection onto this axis is called the PC value. The data is then sorted by the PC values and signal candidate events are chosen by setting a PC value threshold called the candidate cut. While a strict candidate cut will result in a smaller number of background candidates, it will also result in a smaller analysis efficiency since not all signal events will exceed the threshold. In this work, the background has been modeled using 2 % of the available data in the PCA calculation. By using this *burn sample*, no simulations of hadronic air showers (with the included hadronic interaction uncertainties, see Sec. 2.7.1) are needed for the photon

search. To ensure a blind analysis, the burn sample used for the calculation of the PCA has been excluded from the search for a photon flux.

The PCA was performed in three integrated energy bins: $E > 10\,\mathrm{EeV}$, $E > 20\,\mathrm{EeV}$, and $E > 40\,\mathrm{EeV}$. No upper energy limit was set since no cosmic ray photons have been found in previous searches (see Sec. 2.6) and the parameter space was therefore kept as large as possible. Since a flux of primary photons could not be definitely shown, for each energy bin an upper limit on the diffuse photon flux and fraction has been set. The individual steps of the Principal Component Analysis are presented in the following Sections.

6.2 Photon Simulations

The model for the photon signal used in the PCA calculation (Sec. 6.5.1) was created with a dedicated set of simulated photon air showers. These simulations were generated with CORSIKA (Sec. 2.7) using *PRESHOWER* (Sec. 2.7.2) to get a realistic sub-sample of air showers with a first interaction in the geomagnetic field. Simulated events containing a preshower will subsequently be called converted (C) air showers while those without a preshower will be called unconverted (UC) air showers. The simulation set was created with the following specifications:

Number of air shower simulations	30000
High-energy interaction model	EPOS LHC
Low-energy interaction model	FLUKA
Zenith angle range	0°-60°
Minimum energy	$10^{18.5}\,\mathrm{eV}$
Energy spectrum	E^{-1} (reweighted to E^{-2} in the analysis)
Thinning level ϵ_{th}	10^{-6}

Table 6.1: Specifications of the photon air shower simulation set

An example CORSIKA steering card can be found in Appendix A. The detector simulation and reconstruction of the simulated air showers was performed using $\overline{\text{Off}}\underline{\text{line}}$ (see Sec. 3.4, trunk version *v3r0p0*). The module sequence used is included in Appendix B.1. The calculation of photon sensitive event parameters, as well as the reconstruction of the photon energy, has been performed using the SdCompositionParameters module (see Sec. 6.4.2). To increase the statistics of the simulation set, each CORSIKA simulated air shower was thrown five times each with different azimuth angles and positions in the SD array. For a correct calculation of the trigger efficiency, the simulation feature which removes detector stations very close to the shower axis from analyses which used thinned air shower simulations had to be taken into account. This removal is done because at distances close to the shower core the resampling procedure of shower particles becomes highly susceptible to statistical fluctuations (see Sec. 2.7.3). In this work, simulated

events with detector stations closer than 300 m to the shower core were resimulated using adjusted thinning settings.[1]

6.3 Data Set and Reconstruction

The PCA has been applied to Surface Detector data taken between Jan 1st 2004 and June 30th 2018. The measured air showers were reconstructed using Offline (version *v3r0p0*) with the module sequence shown in Appendix B.2. The main differences from the standard SD reconstruction module sequence are the usage of the SdPMTSignalShapeQualityChecker (see Sec. 5.5), to ensure a reliably high data quality of the FADC signal shapes, and the usage of the SdCompositionParameters module, for photon sensitive event parameters and its photon energy reconstruction (see Sec. 6.4.2). These modules are included as they are specifically designed for signal shape sensitive analyses and are therefore not needed in studies which rely only on total station signals.

6.4 Preparation of Shower Dependent Photon Variables

The discrimination between photon-induced air showers and hadronic background is done using the variables Δ_{Leeds} (or simply Δ), and R_{NKG}, both presented in Sec. 3.5. The photon search presented in this work has been made available for review, future photon searches, and multi-messenger analyses by including the calculation of all used parameters in the SdCompositionParameters module in Offline[2] (see Sec. 6.4.2).

6.4.1 Choice of a Photon Benchmark

The Δ_{Leeds} parameter evaluates the photon-likeness of an air shower by comparing the risetime of each detector station (meeting certain quality criteria) to a parametrization of the average risetime found in the data that we refer to as *benchmark* (see Sec. 3.5.2). The value of this risetime benchmark depends on the air shower energy, its zenith angle, and the distance of the detector station from the shower axis. Two independently derived photon benchmarks have been created for this purpose, one by Krohm [123] (called *B1*) and one other by Sanchez-Lucas [130] (called *B2*).

The main differences between these parametrizations are summarized in Table 6.2. Mainly, using B2 would allow to use very steep air showers and include detector stations that are closer to the shower core, but these are gained at the cost of losing zenith angles above 46°. The two benchmarks cover different parameter spaces since B2 was created for hadronic primary mass separation, while B1 was created specifically for the

[1] To calculate the trigger and reconstruction efficiency, the trigger simulations were performed without the removal of detector stations close to the shower axis. The resimulations were performed with Offline version *r31509*.

[2] version *r29341* or greater

	Benchmark B1 Krohm 2017	Benchmark B2 Sanchez-Lucas 2016
Distance range	600 m − 2000 m	300 m − 2000 m
Zenith angle range	30° − 60°	0° − 46°
	common LG/HG benchmark	individual LG/HG benchmarks

Table 6.2: Differences of the risetime benchmarks from [123] and [130]

search for cosmic ray photons. In the parameter space covered by both benchmarks, the differences are mostly well below the $1\,\sigma$ level (see Fig. C.2, C.3), especially when the LG benchmark from B2 is used. Large differences only occur in the parameter space not covered by B2 which includes an undefined region where B2 can not be evaluated. Since good agreement between the two models was found, it was decided that the more extensively tested B1 benchmark optimized for photon searches should be used and air showers between 46° and 60° should be included in the data set.

6.4.2 Implementation in the $\overline{\text{Off}}\underline{\text{line}}$ Framework

The calculation of all photon parameters used was performed in the $\overline{\text{Off}}\underline{\text{line}}$ framework via the newly implemented module *SdCompositionParameters*.[3] During the analysis, the photon parameters are saved to the ADSTs in a parameter storage class introduced for SD related variables. All code for the photon variable calculation has been validated by comparing the resulting photon parameter distributions to the calculations in [123]. The accuracy of the photon energy reconstruction derived in [123], when applied to the new photon simulations, is shown in Fig. 6.1.

Since the parameters used in the photon energy reconstruction were created using the hadronic interaction model QGSJETII.03 (EPOS LHC was used for the new simulations), the relative differences between the photon energy reconstructions and the MC truth were studied (both for C and UC photons). Considering the resolution of the energy reconstruction (see Sec. 6.4.3), there is a good agreement between the reconstructed energy and the MC energy with an average deviation of around 5 %. For this reason, this energy reconstruction was used in [2]. The energies of C air showers are clearly underestimated, however the PCA was performed only using UC photons. For this reason, the underestimation of C photon energies can only cause a decrease of the photon search efficiency.

[3] In former studies, all photon parameters had been calculated after the $\overline{\text{Off}}\underline{\text{line}}$ reconstruction in standalone programs using ADSTs.

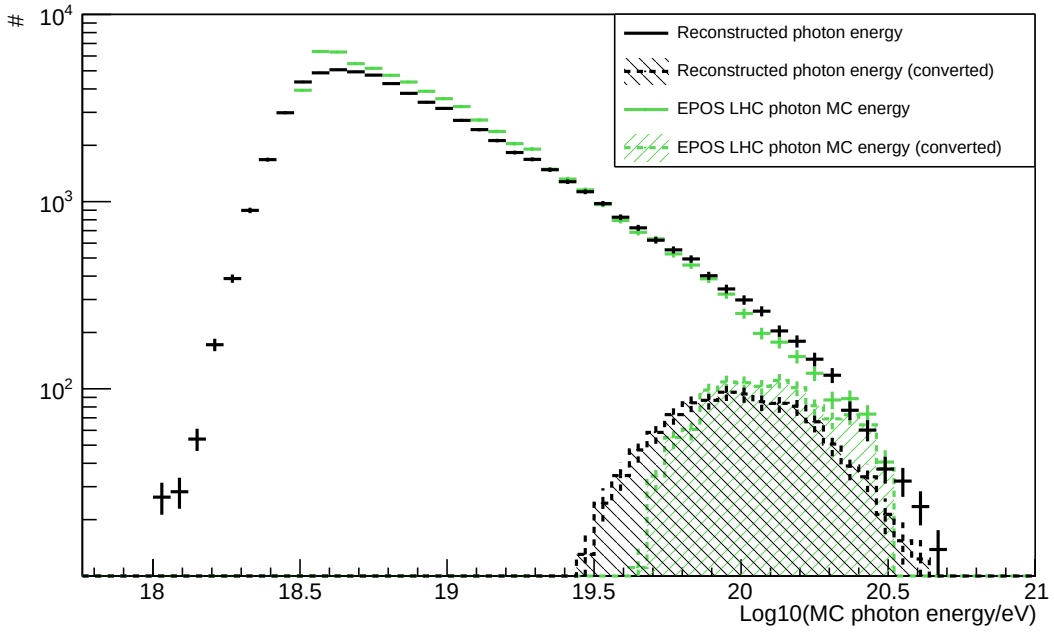


Figure 6.1: Distributions of reconstructed photon energies (black) and MC energies (green) of C and UC photon air showers.

6.4.3 Photon Energy Reconstruction

The energy reconstruction procedure used in [2] was revised using the larger statistics of the EPOS LHC photon MC set. Two methods to determine the photon energy were tested. Both methods only depend on the zenith angle, θ, and the signal at 1000 m from the shower axis, S1000. The two methods are:

- A refinement of the established *iterative procedure* using the new photon MC set.
- A *tabular energy reconstruction* utilizing the large MC statistics to assign a photon energy based on bins in S1000 and θ.

Iterative Procedure The photon energy reconstruction developed in [117] uses an iterative procedure consisting of the following steps:

1. A first estimate of the shower energy is set to $E_\gamma = 2 \cdot E_{SD}$ (for E_{SD} see Sec. 3.3.3).
2. The position of the shower maximum $X_{\max}$ is then calculated according to the parameterized elongation rate (Equation 6.1) using the current E_γ energy estimate.
3. Finally, a new energy estimate E_γ is calculated from $X_{\max}$ according to the parameterized profile (Equation 6.5).

Steps 2 and 3 are repeated until the energy converges with a relative tolerance of 10^{-5}. To ensure a good reconstruction, if the procedure finds the shower maximum to be more than $50\,\mathrm{g\,cm^{-2}}$ below ground, it is cut due to bias issues. This restriction reduces the efficiency for more vertical air showers especially at the highest energies (see Sec. 6.6.1).

The elongation rate used in step 2 of the iteration was fit with a second degree polynomial (Fig. 6.2):

$$X_{\max} = q_0 + q_1 \cdot \log_{10}(E_\gamma/\mathrm{EeV}) + q_2 \cdot \log_{10}^2(E_\gamma/\mathrm{EeV}), \tag{6.1}$$

with a $\chi^2/\mathrm{ndf} = 4.646/7$ resulting in the parameters

$$q_0 = (867.7 \pm 6.1)\,\mathrm{g\,cm^{-2}}, \tag{6.2}$$
$$q_1 = (72.34 \pm 10.49)\,\mathrm{g\,cm^{-2}}, \tag{6.3}$$
$$q_2 = (33.4 \pm 4.2)\,\mathrm{g\,cm^{-2}}. \tag{6.4}$$

The data points consist of the reconstructed UC photon MC events with simulated energies between $10^{18.5}\,\mathrm{eV}$ and $10^{20.5}\,\mathrm{eV}$, zenith angles between 30° and 60°, and $X_{\max}$ values of at least $50\,\mathrm{g\,cm^{-2}}$ above ground.[4] The shower profile used in step 3 of the iteration is modeled as

$$E_\gamma = (\mathrm{S1000}/G(X_{\mathrm{ground}} - X_{\max}))^{1/\alpha}\,. \tag{6.5}$$

Here, $G(X_{\mathrm{ground}} - X_{\max})$ is a Gaisser-Hillas function

$$G(X_{\mathrm{ground}} - X_{\max}) = p_0 \cdot \left(1 + \frac{X_{\mathrm{ground}} - X_{\max} - p_2}{p_1}\right)^{p_1/\Lambda} \cdot e^{-(X_{\mathrm{ground}} - X_{\max} - p_2)/\Lambda}, \tag{6.6}$$

with Λ fixed at $100\,\mathrm{g\,cm^{-2}}$ and p_0, p_1, and p_2 obtained from the fit shown in Fig. 6.3:

$$p_0 = (2.847 \pm 0.005)\,\mathrm{VEM}, \tag{6.7}$$
$$p_1 = (726.5 \pm 5.7)\,\mathrm{g\,cm^{-2}}, \tag{6.8}$$
$$p_2 = (172.1 \pm 0.8)\,\mathrm{g\,cm^{-2}}. \tag{6.9}$$

This parametrization of the shower profile has been obtained with the same cuts that were used for the elongation rate except the MC energy threshold, which has been increased to that of the SD photon search ($10^{19}\,\mathrm{eV}$).

The parameter α quantifies the energy dependence of S1000 [123] as

$$\log_{10}\left(\frac{\mathrm{S1000}}{\mathrm{VEM}}\right) = \alpha\ \log_{10}\left(\frac{E_\gamma}{\mathrm{EeV}}\right) + \beta, \tag{6.10}$$

where $\beta = \log_{10}(G(X_{\mathrm{ground}} - X_{\max}))$ contains the shower profile. Since α is also a function of $\Delta X_{\mathrm{gr}} = X_{\mathrm{ground}} - X_{\max}$, the simulation data has been divided into 10 bins of ΔX_{gr},

[4] In contrast to the description of the simulation set in 6.2, here only one Offline reconstruction has been used for each CORSIKA simulated air shower.

and Equation 6.10 has been fit to each (Fig. 6.4). The resulting values of α have been plotted as a function of ΔX_{gr} in Fig. 6.5. A linear model, as well as a quadratic model, have been fit to the data. Since the quadratic model did not significantly improve the energy reconstruction accuracy, the parameters presented in the following are derived from the linear model of α.

$$\alpha = \alpha_0 + \alpha_1 \cdot \Delta X_{\mathrm{gr}} + \alpha_2 \cdot \Delta X_{\mathrm{gr}}^2, \tag{6.11}$$

with α_0 and α_1 obtained from the fit and α_2 fixed to zero for the linear model:

$$\alpha_0 = 0.841 \pm 0.011, \tag{6.12}$$
$$\alpha_1 = (1.77 \pm 0.31) \times 10^{-4}\,\mathrm{g}^{-1}\,\mathrm{cm}^2, \tag{6.13}$$
$$\alpha_2 = 0. \tag{6.14}$$

The energy dependent accuracy of the iterative photon energy reconstruction method is shown in Fig. 6.6. Unfortunately, it has a bias towards underestimating photon energies which increases with energy climbing to about $-20\,\%$ at 100 EeV. This bias is caused by the required cutting of events where X_{max} occurs more than $50\,\mathrm{g\,cm}^{-2}$ below ground. This is because events with underestimated X_{max} will be generally less likely to be cut, but also will on average be reconstructed with lower energies.

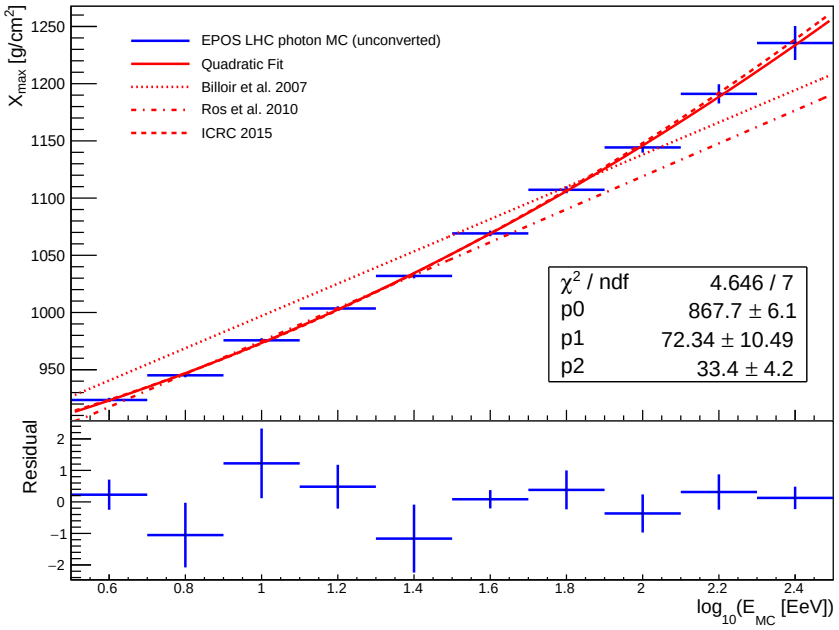


Figure 6.2: Parameterized elongation rate of UC photon MC air showers. The fit of a second degree polynomial is compared to the results of [117], [122], and [123].

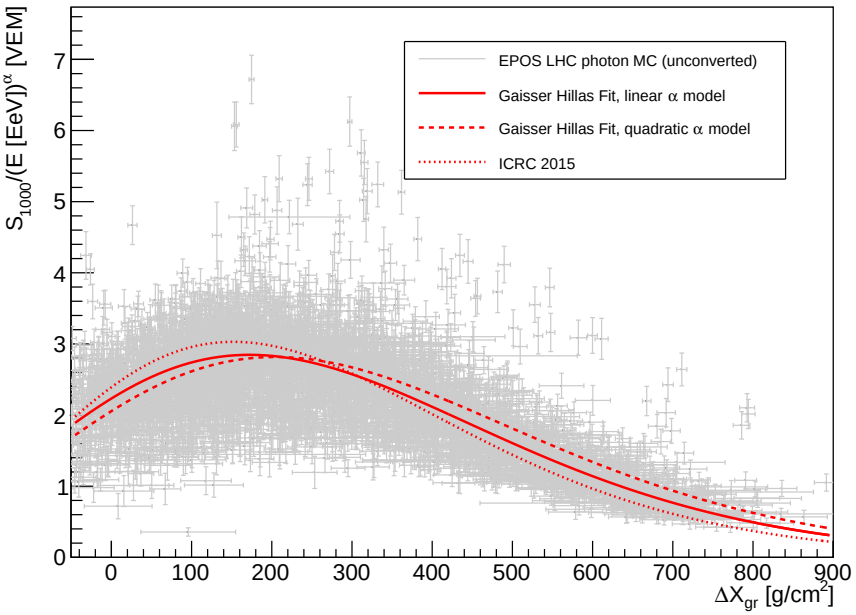


Figure 6.3: Fit of a Gaisser-Hillas function to the universal shower profile for photon primary air showers. Two fits are shown for the linear and the quadratic model of the energy correction factor α. For comparison, the fit results from [123] have been added. The photon MC data points (gray) refer to the linear model of α.

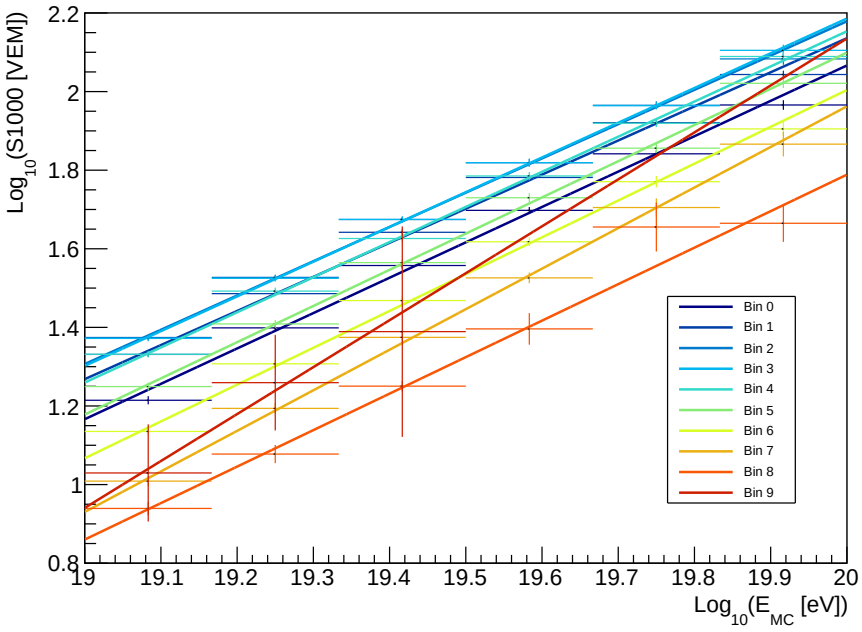


Figure 6.4: Energy dependence of S1000 in 10 equally sized bins of $\Delta X_{\mathrm{gr}} = X_{\mathrm{ground}} - X_{\mathrm{max}}$ between $-50\,\mathrm{g\,cm^{-2}}$ (start of Bin 0) and $900\,\mathrm{g\,cm^{-2}}$ (end of Bin 9).

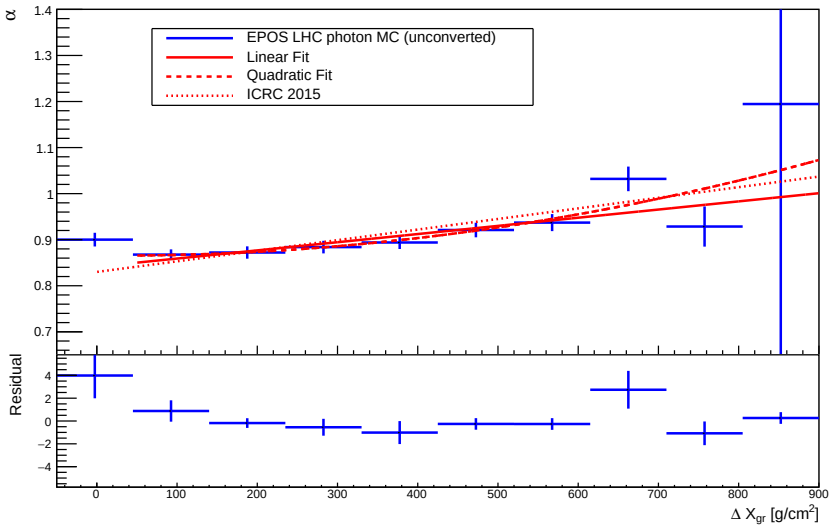


Figure 6.5: Fit of the correction factor α (see Eq. 6.10) with a linear model compared to a quadratic model and the fit from [123].

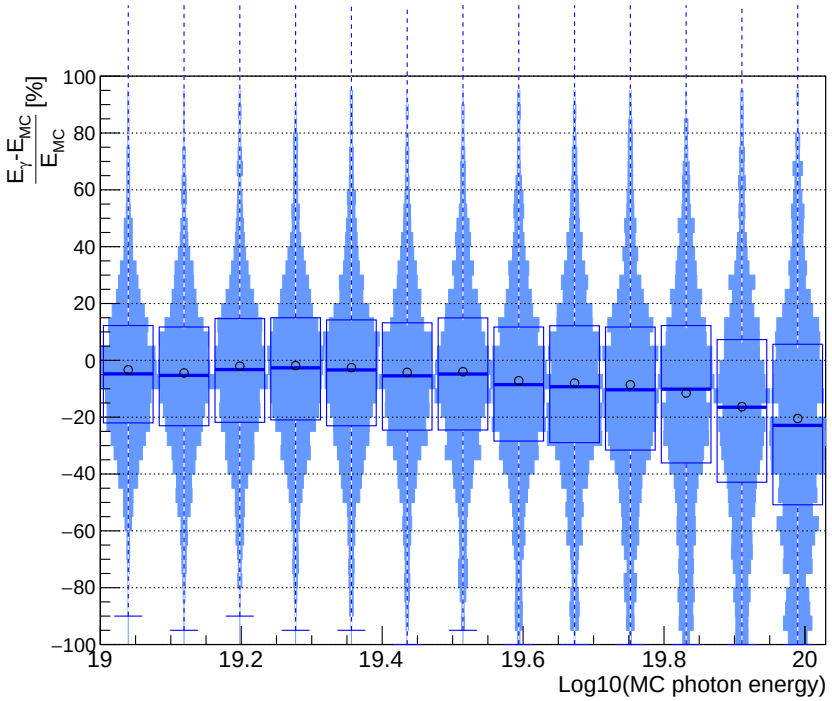


Figure 6.6: Relative energy reconstruction accuracy with the iterative method. Circular markers indicate mean values, dark blue lines median values. The blue box indicates the part between the 25th and the 75th percentile. The iterative photon energy reconstruction has an increasing bias towards low energies caused by the selective cutting of events with overestimated X_{max}.

Tabular Energy Reconstruction The iterative energy reconstruction is a complex procedure relying on the averaged elongation rate. Since it depends on two observables only, S1000 and the zenith angle, an alternative method has been developed by simply tabulating the mean photon MC energies in bins of S1000 and zenith angle. This method profits from the large statistics of the EPOS LHC photon MC set. Only unconverted simulated photon air showers with a 6T5 trigger and both a reconstructed LDF and shower axis are used to populate the table.[5] The resulting energy reconstruction table,

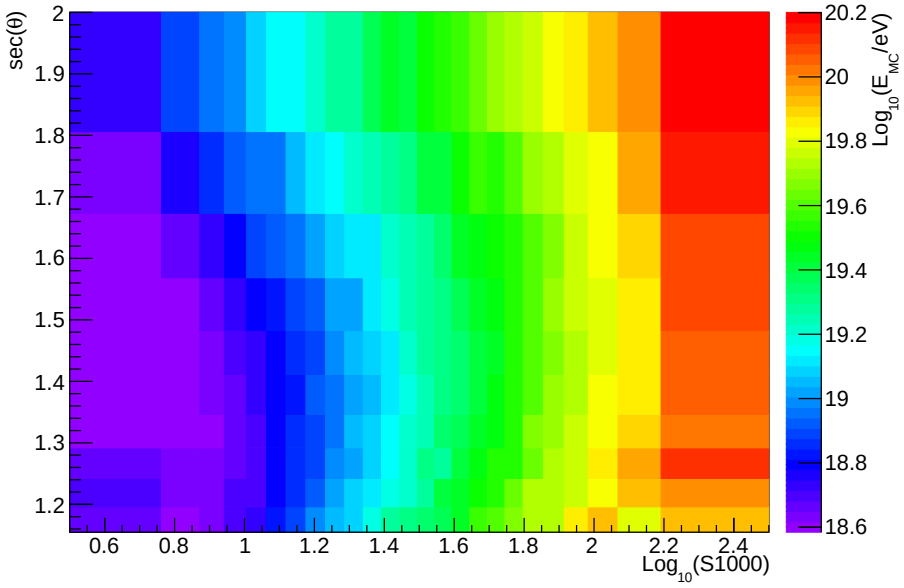


Figure 6.7: Table to convert the measured S1000 and zenith angle to the mean logarithmic MC energies. The binning was chosen for roughly equal statistics in each bin.

represented by a two-dimensional histogram, is shown in Fig. 6.7. On both axes, the binning has been chosen to result in roughly equal statistics in each bin.

The accuracy of the energy reconstruction using the tabular approach is shown in Fig. 6.8. Similar to the iterative energy reconstruction procedure, there is an increasing bias towards low reconstructed energies for high MC energies. However, here this is caused by the reconstruction procedure having been optimized to have no bias as a function of S1000. Particularly large values of S1000 are always more likely to be caused by a lower energy air shower with a comparatively large measured signal than vice versa. For this reason, at high energies a shower of energy E which would on average result in a measured S1000 signal of $S1000_E$, will always have a higher energy than the average energy of all showers which have resulted in that $S1000_E$ value. In other words: the

[5] As in the iterative procedure, only one Offline reconstruction has been used for each CORSIKA simulated air shower.

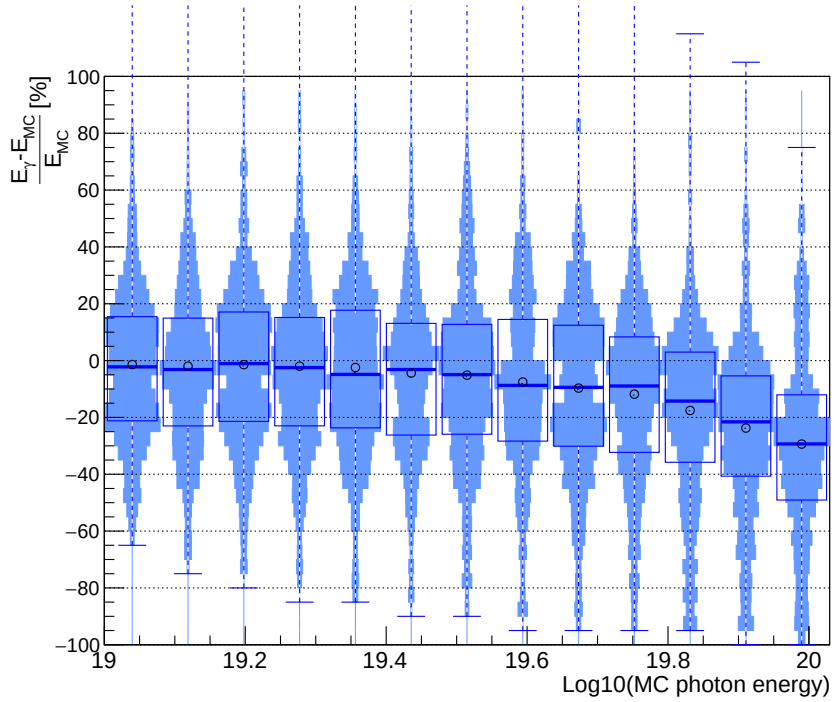


Figure 6.8: Relative energy reconstruction accuracy with the tabular method. Circular markers indicate mean values, dark blue lines median values. The blue box indicates the part between the 25th and the 75th percentile. The increasing bias towards high energies is caused by the reconstruction being designed to have no bias as a function of S1000.

energy of showers with a *mean signal* of $S1000_E$ is higher than the *mean energy* of showers with the *signal* $S1000_E$.

Comparison of Energy Reconstruction Accuracies The relative reconstruction bias of the photon energy is given for both the iterative and the tabular methods. In Fig. 6.9, both methods show an increasing bias to lower reconstructed energies at high MC energies, reaching up to $-30\,\%$ at 100 EeV. While at high energies this bias is larger with the tabular energy reconstruction method, at energies around 10^{19} eV (i.e. for most air showers used in the SD photon search) it is considerably smaller. As noted before, the tabular energy reconstruction was designed to provide a good estimate of the shower energy for each measured value of S1000. For this reason, even though there is a bias of the reconstructed energy as a function of the MC energy, there is (by design) no absolute bias and only a small relative bias of the reconstructed energies as a function of S1000 (see Fig. 6.10). Interestingly, the iterative energy reconstruction method has an increasing positive energy reconstruction bias as a function of S1000 for large measured signals (due to the difference between mean energy and mean signal, see above).

Since no cut on air showers with $X_{\max}$ below ground is required with the tabular energy reconstruction method, the analysis efficiency is greater than that of the iterative

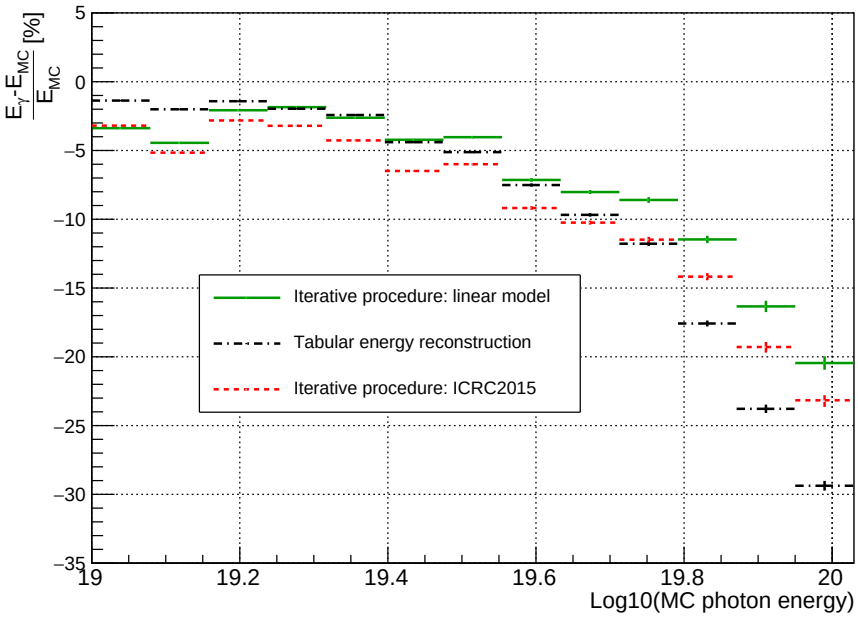


Figure 6.9: Comparison of the relative energy reconstruction bias with the iterative energy reconstruction (green), the tabular energy reconstruction (black dash-dotted), and the version of the iterative energy reconstruction from [123] (red dashed) used in [2].

energy reconstruction. This increase is on a level of 3 % to 11 % (see Fig. C.6). Furthermore, these simulated showers that are not rejected by the tabular method show no significant difference in the PCA distribution (see Fig. C.7).

The most important step of the SD photon search with reliance on the energy reconstruction is the assignment of air showers to an integrated energy bin. For this reason, the accuracy of this assignment has been compared using the different reconstruction methods. The results of this comparison are shown in Fig. 6.11. At energies slightly below the threshold of a new energy bin, misassignments to a too high energy bin become more likely (because only a small misreconstruction is needed). The other way around, slightly above the threshold, misassignments to a too low energy threshold become more likely. Correct assignments are most likely far away from the bin thresholds. While the iterative energy reconstruction has a lower fraction of air showers below 10^{19} eV misassigned to the first integrated energy bin, the tabular energy reconstruction performs the best for most of the 10^{19} eV to $10^{20.5}$ eV energy range. Because of this, its larger efficiency, and the substantially lower complicacy, the tabular method is the favored energy reconstruction method for photons in the Pierre Auger Collaboration.

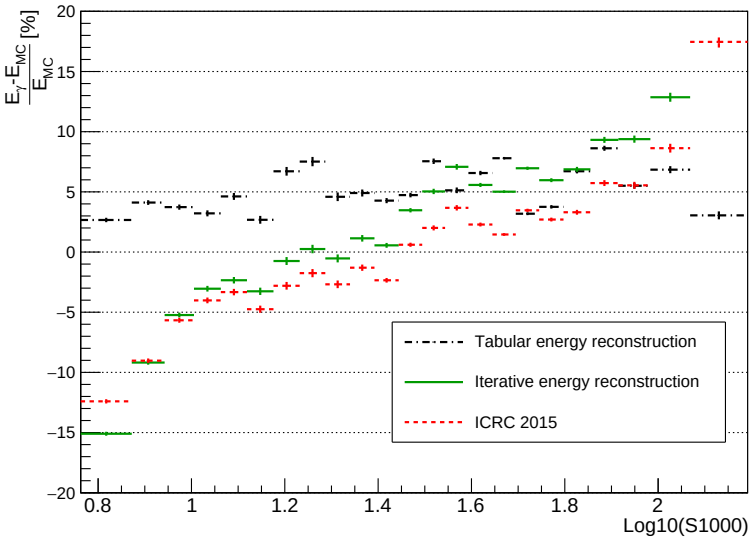


Figure 6.10: Relative energy reconstruction bias in bins of S1000. The small constant bias towards overestimating the energy with the tabular method is caused by the mean energy resulting in no absolute (but therefore a relative) bias: since $E_\gamma - E_{MC}$ has a weighted average of 0, a bias in $(E_\gamma - E_{MC})/E_{MC}$ occurs.

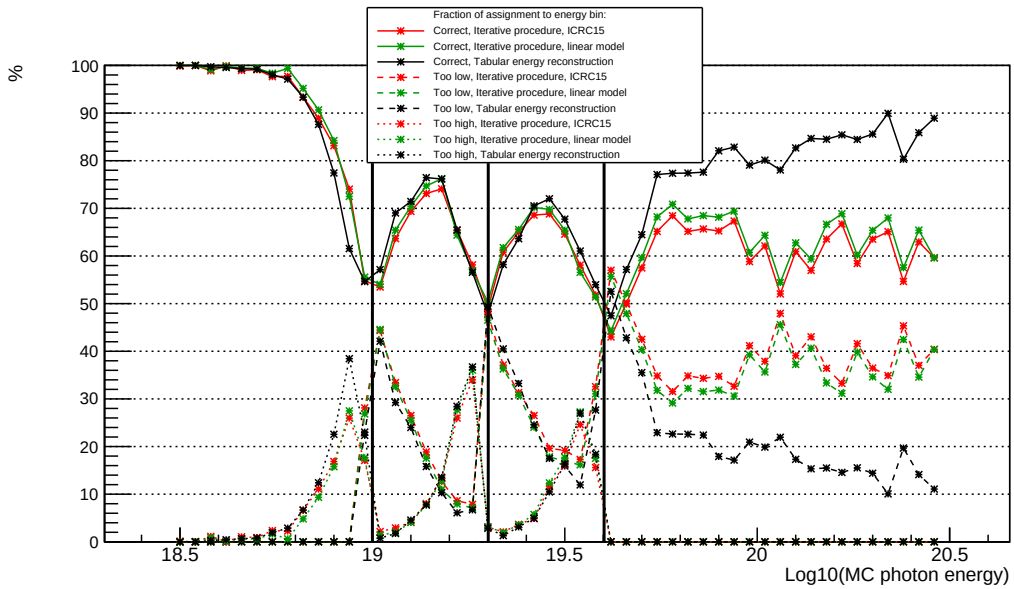


Figure 6.11: Percentages of simulated air showers assigned to the correct lowest energy bin as a function of the MC energy (solid lines). The dashed (dotted) lines show fractions of events misassigned to a lowest bin of a too low (too high) energy threshold. Vertical lines indicate the boundaries of the integrated energy bins.

6.4.4 Linear Transformation of the Photon Variables

The Principal Component Analysis (PCA) used in this SD photon search is based on the LDF parameter

$$\mathrm{L_{LDF}} = \mathrm{Log_{10}}(R_{\mathrm{NKG}}), \tag{6.15}$$

and the risetime parameter Δ_{Leeds}. Both parameters have been linearly transformed to be centered around the mean values of UC photons with a standard deviation of 1:

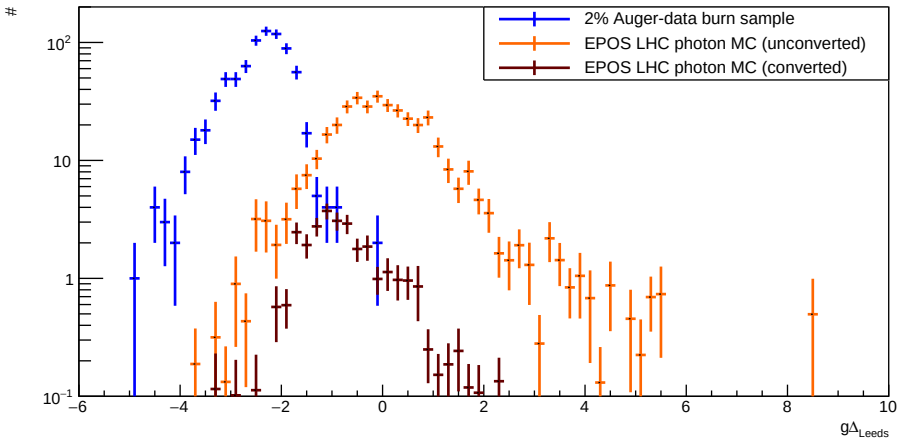
$$\mathrm{g}\Delta_{\mathrm{Leeds}} = \frac{\Delta_{\mathrm{Leeds}} - \overline{\Delta^{\gamma}_{\mathrm{Leeds}}}(E_{\gamma,\mathrm{rec}}, \theta_{\mathrm{rec}})}{\sigma^{\gamma}_{\Delta_{\mathrm{Leeds}}}(E_{\gamma,\mathrm{rec}}, \theta_{\mathrm{rec}})}, \tag{6.16}$$

$$\mathrm{gL_{LDF}} = \frac{\mathrm{L_{LDF}} - \overline{\mathrm{L^{\gamma}_{LDF}}}(E_{\gamma,\mathrm{rec}}, \theta_{\mathrm{rec}})}{\sigma^{\gamma}_{L_{\mathrm{LDF}}}(E_{\gamma,\mathrm{rec}}, \theta_{\mathrm{rec}})}. \tag{6.17}$$

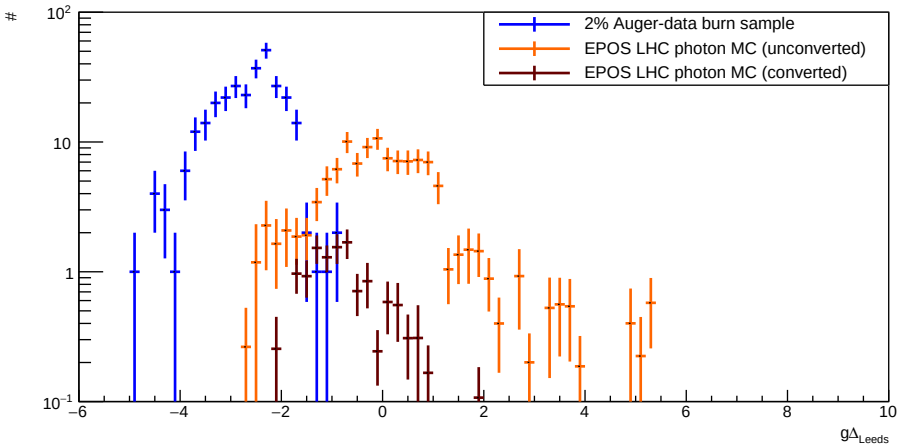
Since both parameters have significant energy and zenith angle dependence (see Figs. C.8 and C.9), this transformation has been performed individually for each bin of reconstructed energy and zenith angle. In this way, an energy and zenith angle independent photon median is constructed and the separation power of the variables is optimized for the full parameter space. For the results presented in [2], the mean and standard deviations derived in [123] were used, resulting in the separation between photon MC and burn sample data shown in Figs. 6.12 and 6.13 for each integrated energy bin.

The new EPOS LHC photon simulation set with its increased statistics has been used to renew the renormalization. Unlike [123], here only the observables Δ_{Leeds} and $\mathrm{L_{LDF}}$ are used and the common selection is applied (i.e. the quality criteria of both variables) during the determination of the renormalization. Furthermore, the new parametrization was created only with the zenith angle range of the SD photon search ($30^\circ < \theta < 60^\circ$) and therefore has a binning which is substantially different from that in [123]. The results of this process are shown in Fig. C.10.

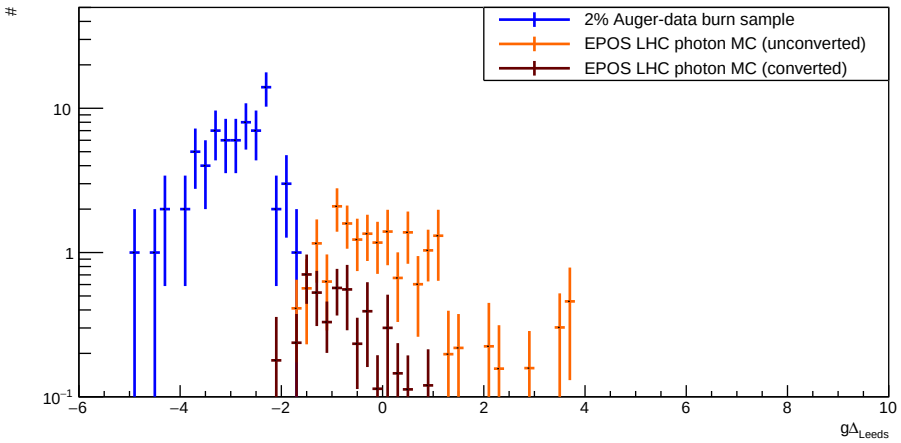
To reduce the effects of the energy reconstruction uncertainties on the linear transformation, a second parametrization based directly on S1000, instead of the reconstructed photon energy, has been created. This parametrization is shown in Fig. C.11. It is worthwhile to mention that one simulated CORSIKA air shower has been removed from the parametrization by hand because it contained a very exotic signal structure which resulted in huge risetimes.



(a) $E > 10\,\mathrm{EeV}$

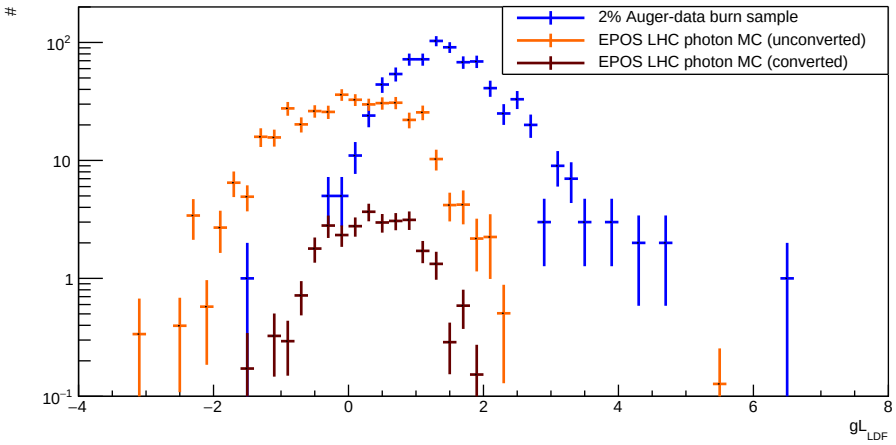


(b) $E > 20\,\mathrm{EeV}$

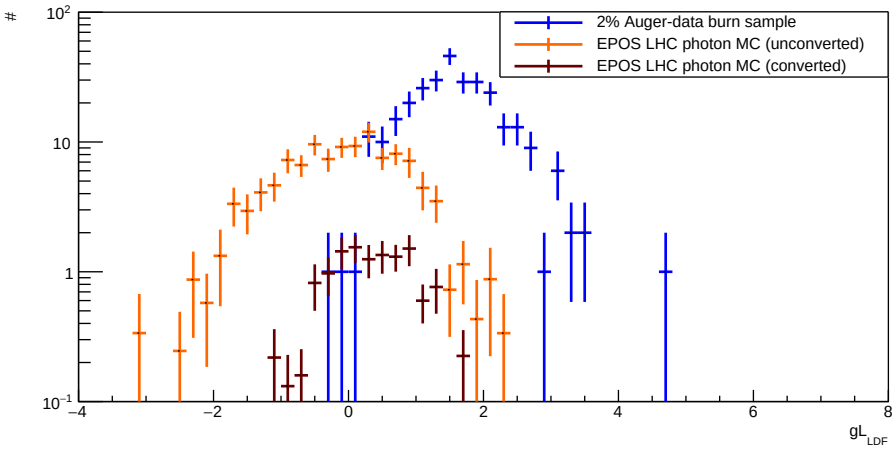


(c) $E > 40\,\mathrm{EeV}$

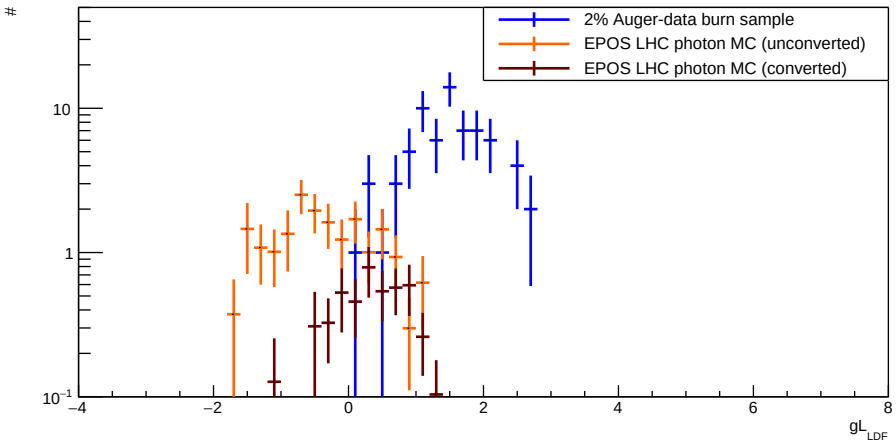
Figure 6.12: Distribution of $\mathrm{g}\Delta_{\mathrm{Leeds}}$ for burn sample data (blue), the reduced set of unconverted photon MC (see Sec. 6.5.1, orange), and converted photon MC (dark red) for the three integrated energy bins



(a) $E > 10\,\mathrm{EeV}$



(b) $E > 20\,\mathrm{EeV}$



(c) $E > 40\,\mathrm{EeV}$

Figure 6.13: Distribution of $\mathrm{gL_{LDF}}$ for burn sample data (blue), the reduced set of unconverted photon MC (see Sec. 6.5.1, orange), and converted photon MC (dark red) for the three integrated energy bins

6.5 Multivariate Photon Search

In this Section, the determination of the PC axis and the results of the PCA are presented. The PC axis itself was determined using a 2 % burn sample and a set of simulated photon air showers. Subsequently, the PCA was applied to the SD data taken between Jan 1st 2004 and June 30th 2018. To ensure the search for photons is unbiased, once this *unblinding* had been performed no further changes were made to the analysis. Afterwards, photon candidates were identified in the data. These event candidates, their possible primary nature, as well as tests to see if their quantity can be explained by a proton background, will also be discussed.

6.5.1 PCA Determination

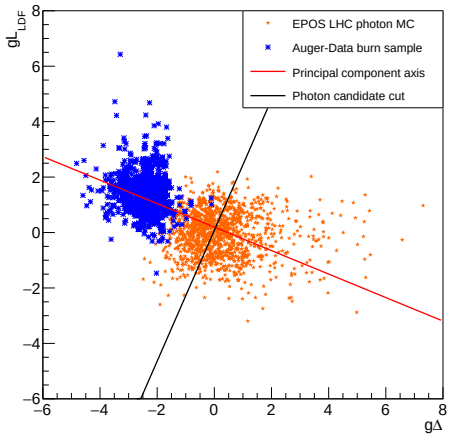
As stated above, the PC axis was calculated using a 2 % burn sample containing 662 events which were then excluded from further analysis. A sub-sample of these events had already been *burned* in previous photon searches [123]. These were chosen randomly and identified by a list of event IDs. For this analysis, this subset has been expanded with new events measured between May 15th 2013 and June 30th 2018 with $30^\circ < \theta < 60^\circ$, a reconstructed LDF, a 6T5 trigger, and a requirement of

$$\mathrm{SDID}\%50 = 0, \tag{6.18}$$

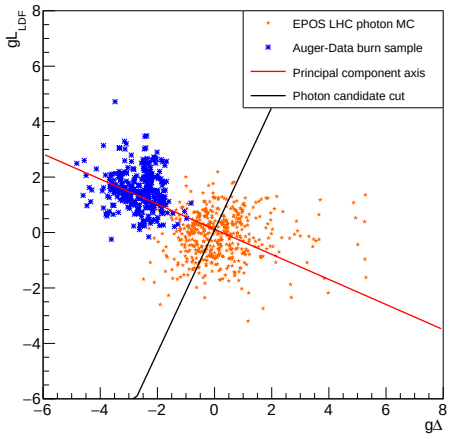
where SDID is the Id of the SD event. This results in a selection of the 2 % of the data which has the disadvantage of not being completely random, but the advantage of being easy to update and reproduce.

As described in the previous sections, the photon signal was modeled using simulated (UC) photon air showers (see Sec. 6.2). Even though large statistics were available, only small sets of simulations were chosen to represent the photons in the PCA. This reduction of the photon MC statistics was necessary to avoid the PC axes from being dominated by the random fluctuations in photon simulations instead of focusing on the separation of the MC signal and the burn sample background. For this reason, the sizes of the reduced MC sets were chosen depending on the size of the burn sample in each energy bin, with a ratio between photon MC and burn sample size of $\sim 5/3$. The calculations of the PC axis and the photon median were then performed for each integrated energy bin individually in order to adapt to the energy dependence of the photon sensitive parameters.

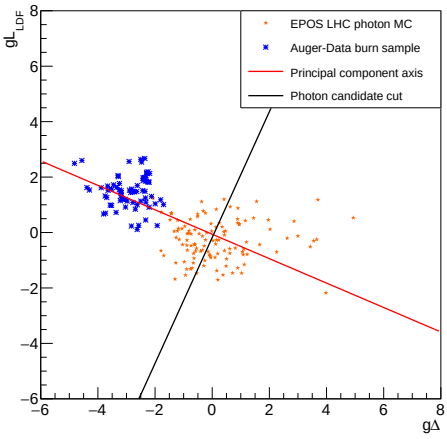
In Fig. 6.14, the burn sample, photon MC, and PC axes for all three bins of the reconstructed photon energy are shown in the PCA parameter space. Perpendicular to the PC axis, the position of the photon candidate cut is shown which was set to the median of the UC photon MC distribution. This, somewhat arbitrary, cut was used in order to stay consistent with former publications [2, 74, 80]. Also, by choosing the photon median, the linearly transformed variables $\mathrm{g}\Delta_{\mathrm{Leeds}}$ and $\mathrm{gL}_{\mathrm{LDF}}$ can be interpreted as a measure of how significantly an event differs from the expected photon response in the detector.



(a) $E > 10\,\mathrm{EeV}$



(b) $E > 20\,\mathrm{EeV}$



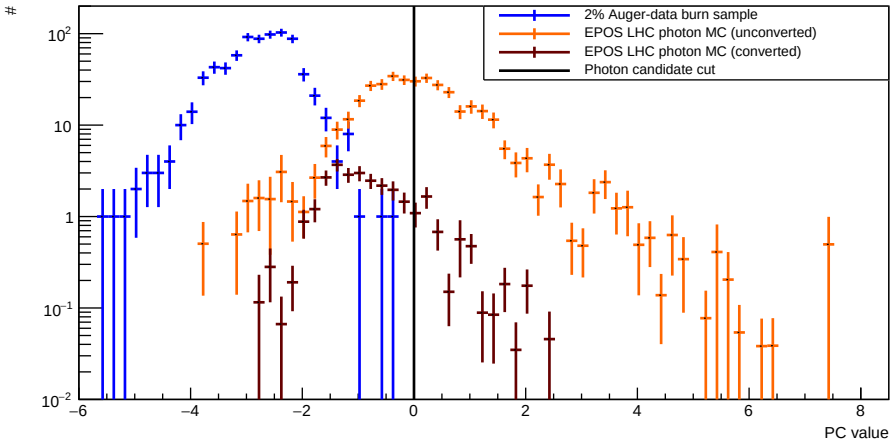
(c) $E > 40\,\mathrm{EeV}$

Figure 6.14: PCA of burn sample data (blue) and the reduced photon MC set (orange) using the parameters $\mathrm{g}\Delta_{\mathrm{Leeds}}$ and $\mathrm{gL}_{\mathrm{LDF}}$

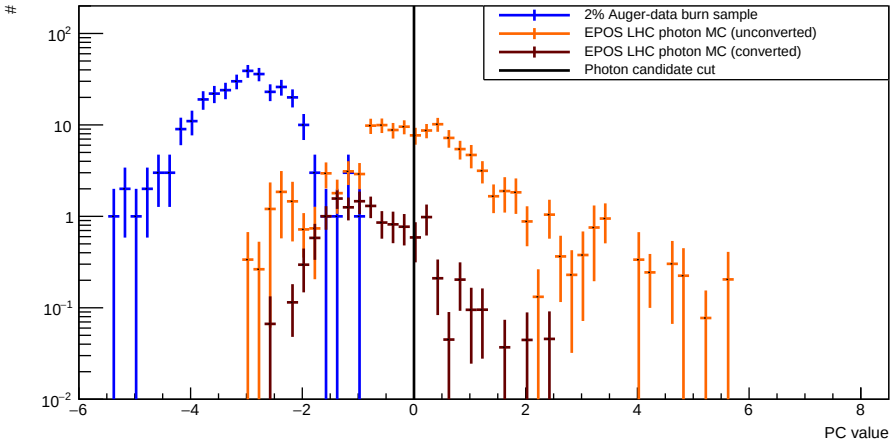
The projections of burn sample data and photon MC to the PC axis are shown in Fig. 6.15. From here on, a shift has been applied to the PC values so that the median of the photon distribution is at the origin. The individual angles of the PC axis and PC values of the photon median for all three integrated energy bins are given in Table 6.3.

	$E > 10\,\mathrm{EeV}$	$E > 20\,\mathrm{EeV}$	$E > 40\,\mathrm{EeV}$
Photon median	0.88	0.99	1.12
PCA angle	$-23.0°$	$-24.3°$	$-23.8°$

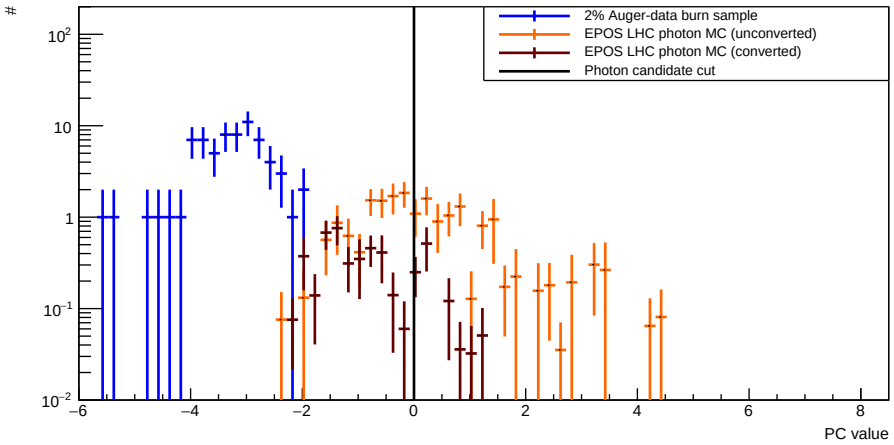
Table 6.3: PCA angles and photon medians for the three integrated energy bins



(a) $E > 10\,\mathrm{EeV}$



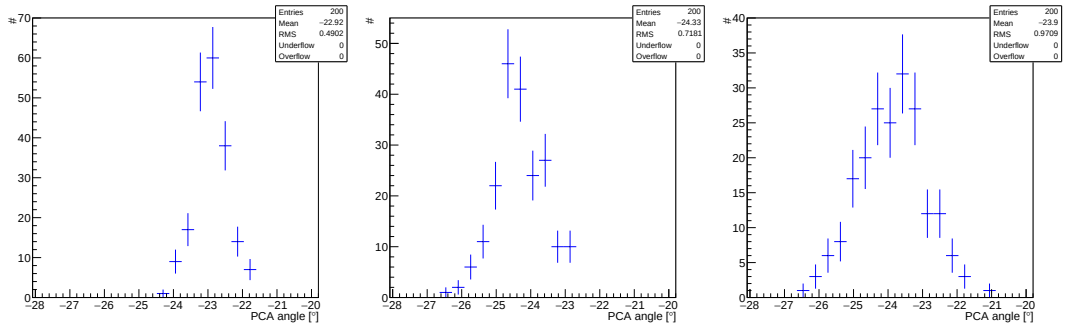
(b) $E > 20\,\mathrm{EeV}$



(c) $E > 40\,\mathrm{EeV}$

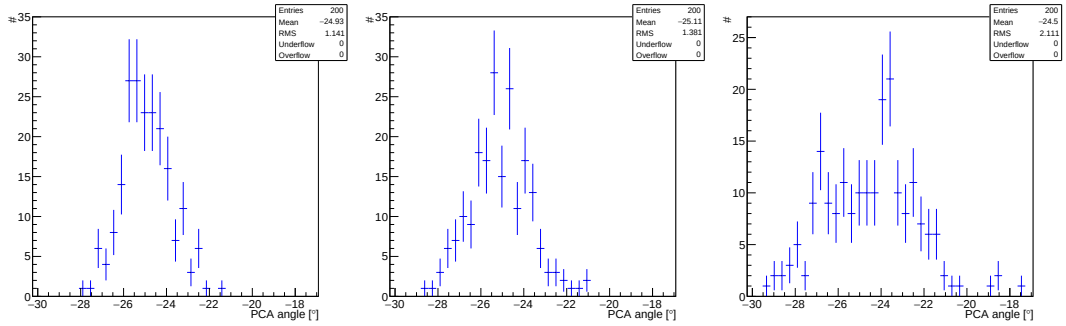
Figure 6.15: Separation of photon MC and burn sample data projected to the PC axis for the three integrated energy bins. The PC value was shifted to have the photon median at zero.

Since the angle of the PC axis has a strong influence on the analysis results, the significance of its positioning was studied. Whether the analysis would profit from an increase of the burn sample size was tested by bootstrapping on the burn sample data. 200 full sized subsamples of the burn sample data have been created using sampling with replacement. For each of the subsamples, the resulting PCA angle has bin calculated. Additionally, the statistical uncertainty due to the size of the MC sample was determined by carrying out the PCA on 200 equally sized subsamples from the much larger sample of available photon MC air shower simulations. The results of these tests are shown in Figs. 6.16 and 6.17. Since the standard deviations of both distributions are on the level of only about 1° in the first energy bin, the burn sample (and therefore also the MC) statistics have not been increased in order to keep as much data for the photon search as possible.



(a) $E > 10\,\mathrm{EeV}$ (b) $E > 20\,\mathrm{EeV}$ (c) $E > 40\,\mathrm{EeV}$

Figure 6.16: Distributions of PCA angles in a bootstrapping of 200 full-sized burn sample sets. The samples were created from the same data using sampling with replacement.



(a) $E > 10\,\mathrm{EeV}$ (b) $E > 20\,\mathrm{EeV}$ (c) $E > 40\,\mathrm{EeV}$

Figure 6.17: Distributions of PCA angles in a bootstrapping with 200 equally sized randomly chosen MC sets.

6.5.2 Position of the Candidate Cut

The choice to set the candidate cut at the position of the photon median is an established, but essentially arbitrary, choice [2, 74, 80]. Following up on the ICRC2019 analysis presented in this work, a refinement of the process of determining the cut position was developed. The new process was specifically designed to increase the power of the analysis to set upper limits. To study the effects of the positioning of this cut, the renormalization of the PCA variables based on the reconstructed photon energy and the zenith angle presented in Sec. 6.4.4 and shown in Fig. C.10 is used.

In order to estimate the number of expected background candidate events in the full data sample, the tail of the burn sample PC distribution for the $E > 10\,\mathrm{EeV}$ bin was fit with an exponential function (Fig. 6.18). By integrating this exponential tail

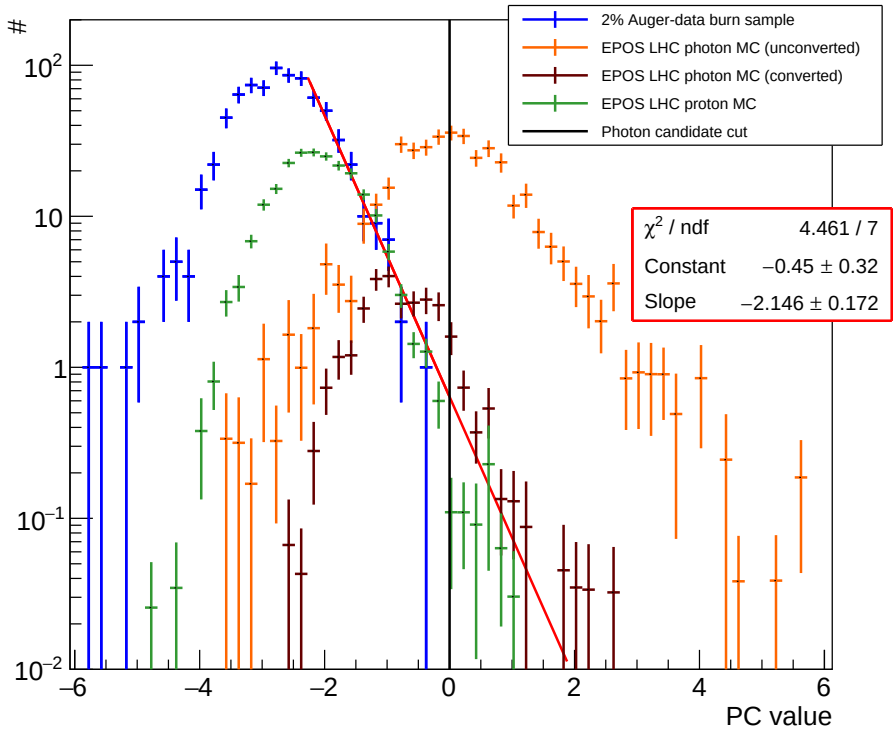


Figure 6.18: Exponential fit (red) to the tail of the burn sample PC distribution (blue). For comparison, the distribution of simulated proton air showers, scaled by an arbitrary factor to compare the distribution tails, and reweighted to a spectrum of E^{-3}, is shown in green.

from the cut position to infinity and scaling the result to the statistics of the full data sample, a prediction of the number of background candidate events as a function of the cut position (Fig. 6.19a) can be obtained. The resulting number of 'Feldman-Cousin candidates' [141] is shown in Fig. 6.19b.

While this candidate background is exponentially suppressed with larger cut values, the analysis efficiency is reduced significantly as well. For this reason, the candidate cut efficiency has been studied as a function of the candidate cut position with the methods presented in Sec. 6.6 with results shown in Fig. 6.19c. The strength of an upper limit on the photon flux, assuming no photons in the data sample and a background following the exponential shape of the burn sample, can be estimated by

$$\Phi_\gamma^{95\mathrm{CL}} \propto \frac{N_{\mathrm{cand}}^{\mathrm{FC}}}{\epsilon_{\text{candidate cut}}}, \tag{6.19}$$

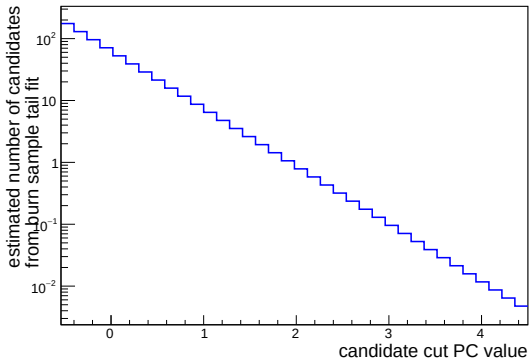
where $N_{\mathrm{cand}}^{\mathrm{FC}}$ is the number of Feldman-Cousin candidates at 95 % confidence level and $\epsilon_{\text{candidate cut}}$ is the candidate cut efficiency. The cut value which minimizes this term (see Fig. 6.19d) denotes a natural position to set the candidate cut which takes the background expectation from the burn sample into account. An additional advantage of this method of deriving the cut position is that the available exposure is also automatically taken into account.[6] Even though stronger limits were expected with this optimized cut position, in order to optimize the analysis to find a photon flux and to stay consistent with former publications, the cut at the photon median has been used in the photon search presented in this work.

6.5.3 PCA Results

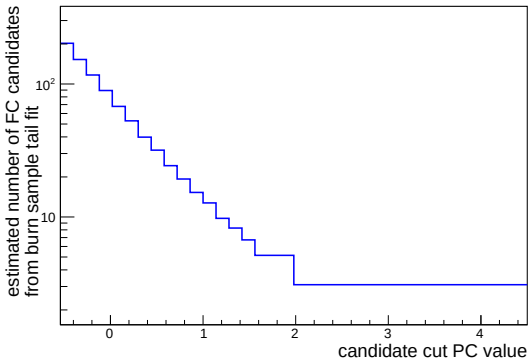
The result of the PCA on the full data set in the parameter space of the PCA variables is shown in Fig. 6.20. The region to the right of the cut axis (large $g\Delta_{\mathrm{Leeds}}$ and low gL_{LDF} values) in the $E > 10\,\mathrm{EeV}$ energy bin contains 12 candidate events. This is significantly larger than the 7.1 events expected from upscaling the number of candidate events found in [142] to the new exposure in this energy bin. One of these events has a $g\Delta$ value of ≈ 4.38, which is far larger than the rest of the data, and even quite high compared to most of the photon MC. This event was later shown to be a fake candidate (see Sec. 6.5.4) and was therefore removed from the analysis and the following plots. Two of the remaining 11 photon candidates have energies above 20 EeV but below 40 EeV and thus appear in the second energy bin as well.

All photon candidate events, except for the mentioned fake candidate, are positioned very close to the candidate cut axis. Therefore, slight changes of the cut position would have a large effect on the analysis results. In Fig. 6.21, the projection of the PCA result to the PC axis is shown. In the $E > 10\,\mathrm{EeV}$ and $E > 20\,\mathrm{EeV}$ energy bins, the photon candidate events seem to be contained in a roughly exponential tail of the PC distribution of the data. The tails of these PC distributions will be discussed in more detail in Sec. 6.5.6.

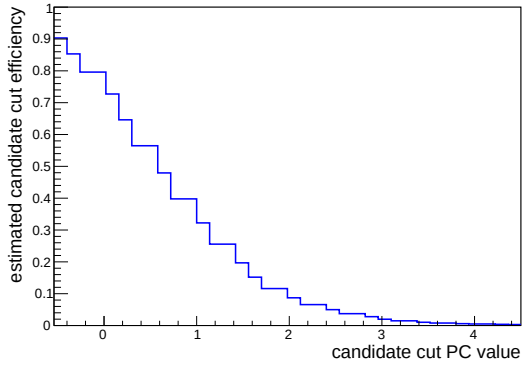
[6] With larger exposure the background expectation increases while the candidate cut efficiency remains unchanged. For this reason, larger exposures result in a more strict candidate cut, in contrast to a fixed cut at the photon median.



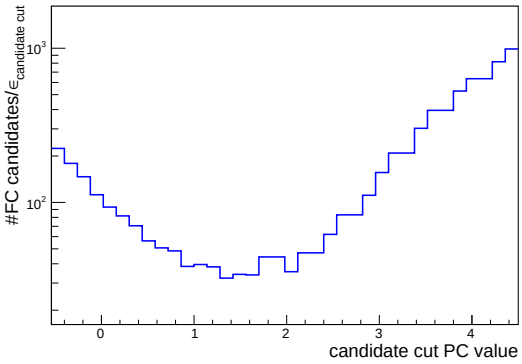
(a) Expected number of background candidate events as a function of the candidate cut position. The background has been modeled using an exponential fit to the burn sample data.



(b) Expected number of Feldman-Cousin candidates as a function of the candidate cut position.

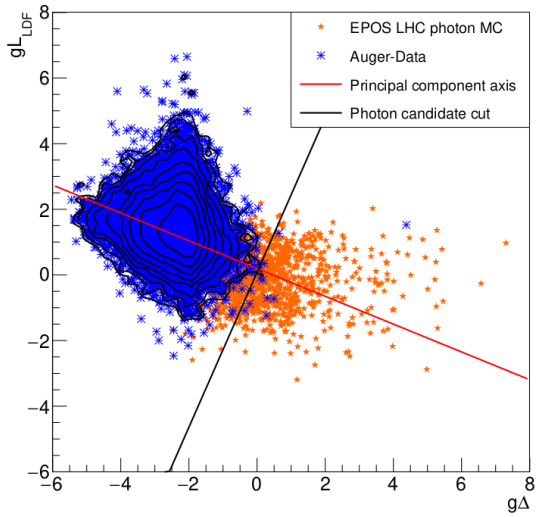


(c) Efficiency of the photon candidate cut as a function of the candidate cut position.

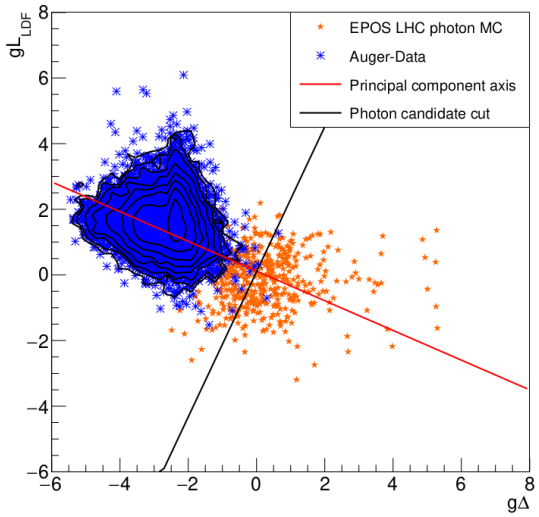


(d) Number of Feldman-Cousin candidates divided by candidate cut efficiency. This gives an estimator for the candidate cut position resulting in the strongest upper limits on the photon flux assuming no photon flux and an exponential shape of the hadronic background.

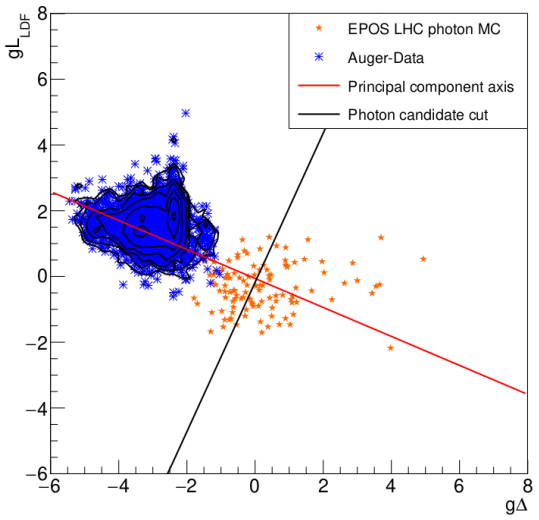
Figure 6.19: Effects of the candidate cut position on (a)/(b) the number of expected background candidates, (c) the analysis efficiency, and (d) the strength of the upper limit (assuming no photon flux).



(a) $E > 10$ EeV,
11 photon candidates + 1 fake candidate

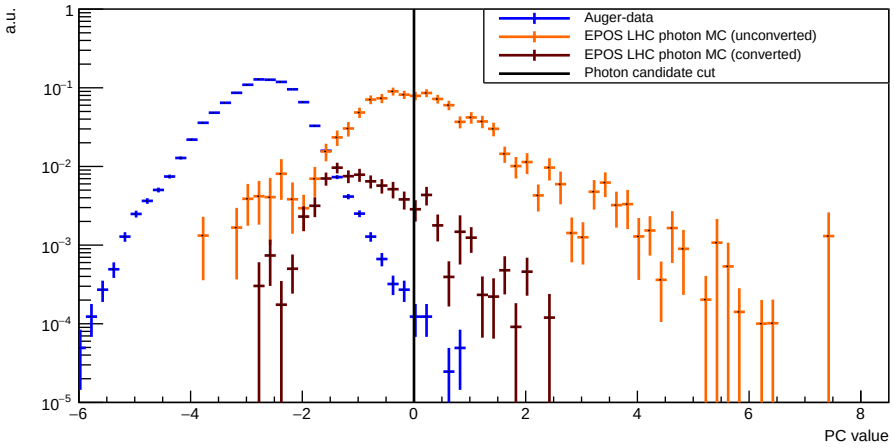


(b) $E > 20$ EeV,
2 photon candidates

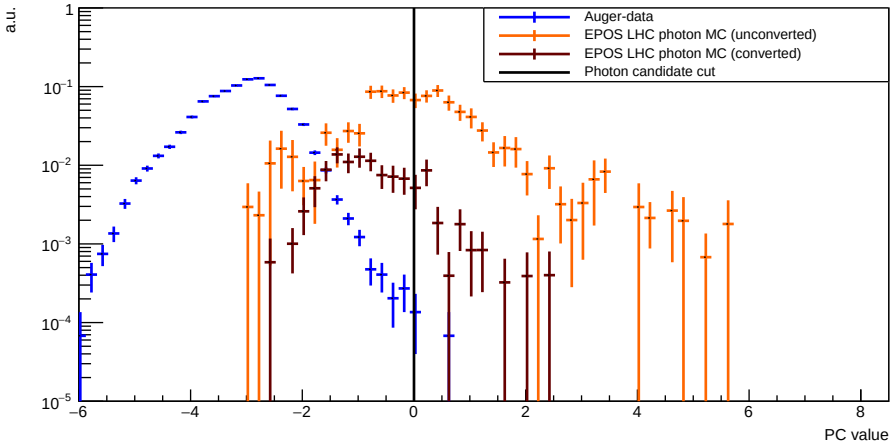


(c) $E > 40$ EeV,
0 photon candidates

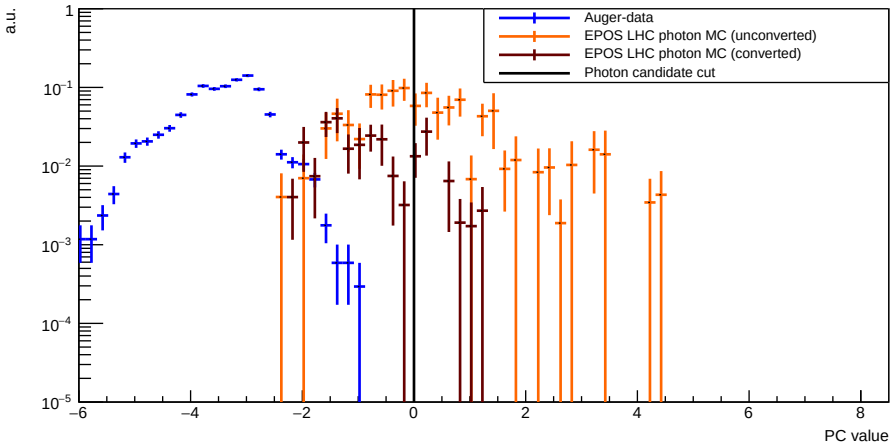
Figure 6.20: PCA results of the full Auger SD data set from Jan 1st 2004 to June 30th 2018 (blue) and the reduced photon MC set (orange). All data events to the right of the cut axis are photon candidates. One fake candidate is discussed in Sec. 6.5.4.



(a) $E > 10\,\mathrm{EeV}$, 11 photon candidates



(b) $E > 20\,\mathrm{EeV}$, 2 photon candidates



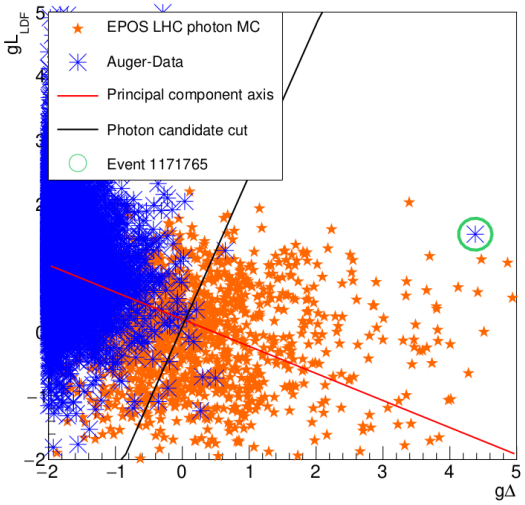
(c) $E > 40\,\mathrm{EeV}$, 0 photon candidates

Figure 6.21: Projection of the PCA results to the PC axis for the three integrated energy bins. The fake candidate was removed from the $E > 10\,\mathrm{EeV}$ plot. The shown statistics of MC photons were reduced for the calculation of the PCA.

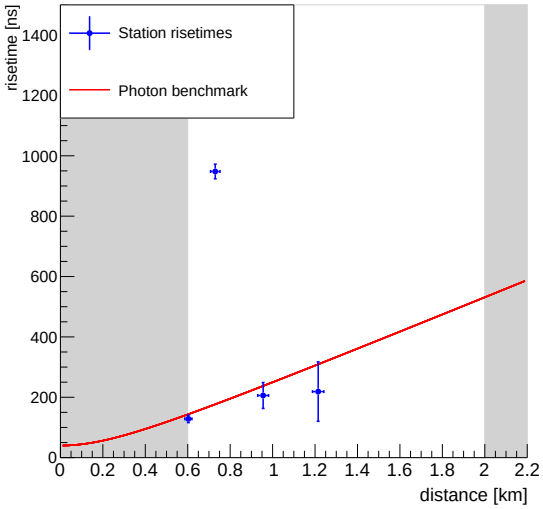
6.5.4 Candidate Discussion

The unblinding of the analysis revealed one candidate event (event 1171765) very far from the PC axis with a PC value of 3.46. This large value was driven by the very large $g\Delta_{\text{Leeds}}$ of 4.38, which is unusually high even for photons, while the gL_{LDF} value of 1.52 is rather typical for hadrons. In Fig. 6.22a, the position of this event in the PCA parameter space is shown. The value of $g\Delta_{\text{Leeds}}$ was entirely caused by a single detector station with a huge risetime value. All of the other detector stations involved had risetimes below the benchmark (see Fig. 6.22b). The HG channels of all three PMTs in that particular detector station were saturated during the event. The LG signal traces are shown in Fig. 6.22c. PMT1 and PMT3 have normal signal forms, however PMT2 has a completely empty signal trace. The risetime calculation of this empty FADC trace had a result large enough to create a fake candidate event alone. Due to a technical problem, the empty trace was not removed by the SdCalibrator which is the module responsible for catching such traces. Removing this problematic trace results in a $g\Delta_{\text{Leeds}}$ value of -3.29 for the event and consequently a PC value of -2.72 which is clearly below the candidate cut. For this reason, event 1171765 has been removed from the analysis. After finding this, all FADC traces of the other photon candidate events were hand examined for Bad PMT signatures. In these candidates, no problematic signatures were found indicating a successful removal of the trace pathologies described in Chapter 5.

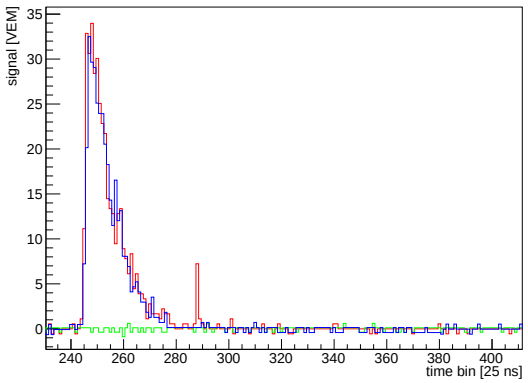
Candidate event no. 8742491 contains a detector station with a measured signal which is split into two separate parts (see Fig. 6.23). A similarly split signal is only apparent in one of the four other participating detector stations (see Fig. C.12). In the reconstruction used in this work, the second part of the signal was interpreted as the signal start. In Fig. 6.24, a comparison of the timing fits for this event with an older version of the reconstruction software is shown. The time fit which includes the first part of the signal is considerably more successful with a χ^2/NDF of smaller than 0.1/1 compared to 41.7/1 in the original reconstruction. A comparison of the reconstruction results shows that the zenith angle changed from the 42° found with the older reconstruction to 55° when the new reconstruction is used. This change in geometry is important because it also caused an increase of the reconstructed photon energy. In this event, the detector station with the split signal was not used for the calculation of Δ_{Leeds}, but was the only station used for the calculation of the L_{LDF} parameter. Even though including the first part of the PMT traces of this station would move the L_{LDF} parameter further to the data region, event no. 8742491 was kept as a photon candidate since the large PC value was mostly driven by the large risetimes found in the four other detector stations.



(a) Position of event 1171765 in the PCA parameter space. This event is far from the PC axis due to a huge $g\Delta_{\text{Leeds}}$ even though gL_{LDF} is in the typical region for hadrons.



(b) Individual station risetimes compared to the benchmark function. A single detector station has a huge risetime. All other detector stations have risetimes below the benchmark.



(c) PMT traces of detector station 348 in event 1171765. The second PMT has an empty signal trace which was not removed from the analysis.

Figure 6.22: Main observables of event 1171765.

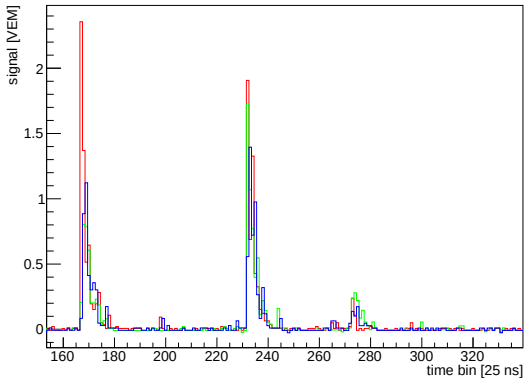
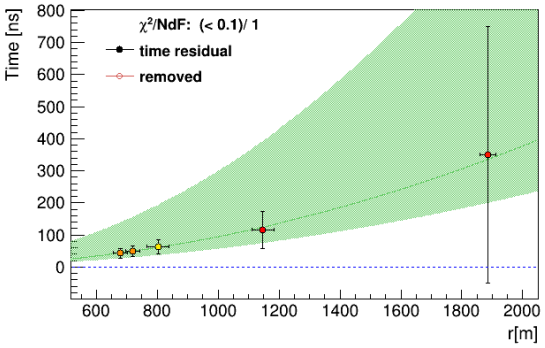
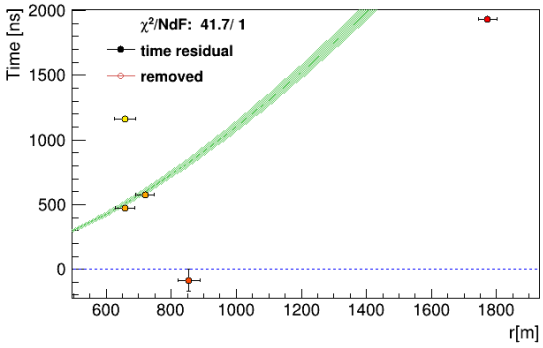


Figure 6.23: Signal traces of all three PMTs in detector station 737 for event 8742491. The signal is split into two main parts.



(a) Time fit for event 8742491 with the ICRC2013 reconstruction



(b) Time fit for event 8742491 with the newer version of the reconstruction software used in this work

Figure 6.24: Comparison of time fits for event 8742491

6.5.5 Photon Candidate Events

Among the 11 candidate events found in the PCA, three events (46285053, 9701198, and 15797618) were also observed by the Fluorescence Detector. One of these events (15797618) was in the field of view of two FD telescopes simultaneously. Additionally, three of the photon candidate events (9701198, 10759292, and 15797618) had been identified as photon candidates in [142] which included two of the FD events. The only remaining candidate event of that analysis (7543164), was below the candidate cut in this work with a PC value of -0.14. No common candidate events with the entity method based analysis (see Sec. 3.5.3) from [133] have been found. Measurement details of all candidate events are shown in Figs. 6.25 to 6.35.

Candidate event **8742491** has the largest PC value (0.77) and the second largest $g\Delta_{\text{Leeds}}$ value (0.50) of all candidate events. The large $g\Delta_{\text{Leeds}}$ value is caused by four stations with risetimes well above the risetime benchmark making it a very promising candidate. This event has an ambiguous signal start time in station 737, which was discussed in Sec. 6.5.4.

Event no. **9099370** has a $g\Delta_{\text{Leeds}}$ value largely driven by a single station with a very high risetime. The individual PMT traces of this station are shown in C.13 and appear to contain a multitude of particles with widely spread arrival times as would be expected from a photonic air shower.

For event no. **9701198** an FD measurement is available. This event was measured on the border of the Surface Detector array. It has four stations with risetimes well above the risetime benchmark and an FD-measured X_{max} value of $(1158 \pm 43)\,\text{g}\,\text{cm}^{-2}$. This X_{max} is above the values expected for protons and even very high for photons but was not in the field of view of the telescope (see Fig. C.21). The large $g\Delta_{\text{Leeds}}$ value is caused by four detector stations with very large risetimes, while the LDF is compatible with the expectation for hadrons.

Event **10759292** is very close to the PC axis and contains four detector stations contributing to $g\Delta_{\text{Leeds}}$. Two of these stations have risetimes close to the risetime benchmark while the two other stations have very large risetimes causing the large $g\Delta_{\text{Leeds}}$ value. The VEM traces of the detector stations with large risetimes are shown in images C.14 and C.15.

Event no. **15797618** is the only candidate event with measurements of two FD telescopes. It has a reconstructed photon energy of 24.3 EeV and thus is a candidate in the first two integrated energy bins. All four stations used for the Δ_{Leeds} parameter have risetimes above the benchmark with one of them being particularly high possibly caused by multiple late muons. The corresponding VEM trace is shown in C.16. Both fluorescence telescope measurements indicate a very high X_{max} value ($\approx 1250\,\text{g}\,\text{cm}^{-2}$) which is larger than expected for a hadronic event and even unusually high for a photon. The profiles measured by the FD are shown in C.22 and C.23. In the profile measured by the FD telescope Loma Amarilla, data is missing close to the shower maximum maybe due to a heavy cloud. The measurement from Los Morados confirms the deep shower maximum but has a much larger measurement uncertainty.

Event **23952703** is the first candidate event not covered in the data period of [123] and the last candidate event covered in the data period of [133]. It includes four detector stations with signals below the LDF function and five stations with risetimes significantly above the benchmark.

Event **32654533** includes measurements of two detector stations that are significantly below the LDF fit and four detector stations used for the Δ_{Leeds} parameter. One of these stations has a very large risetime and seems to contain a multitude of low energy particles. The VEM traces of this station are shown in C.17.

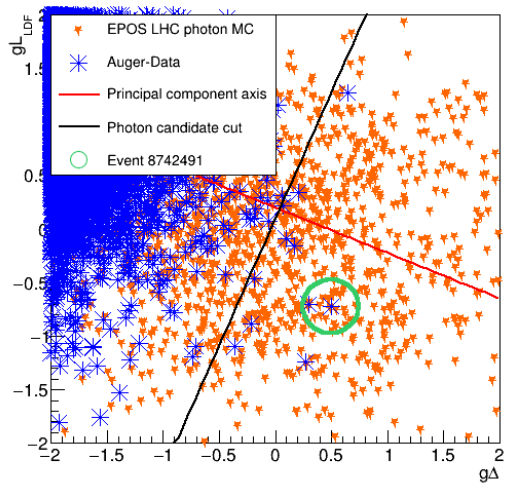
Event **40457464** has the second lowest $\mathrm{gL_{LDF}}$ parameter of all candidate events due to one station with a signal significantly lower than expected from the LDF. Furthermore, all four stations used to calculate the $\mathrm{g}\Delta_{\mathrm{Leeds}}$ parameter have risetimes above the benchmark.

Event no. **43993111** has the third largest PC value of all photon candidate events. It is driven by two detector stations with signals well below the LDF and four out of five detector stations with risetimes much larger than the benchmark. The risetime of the only station compatible with the risetime benchmark is probably affected by a wrong starting position due to a small peak in the FADC trace of a single PMT (see Fig. C.18).

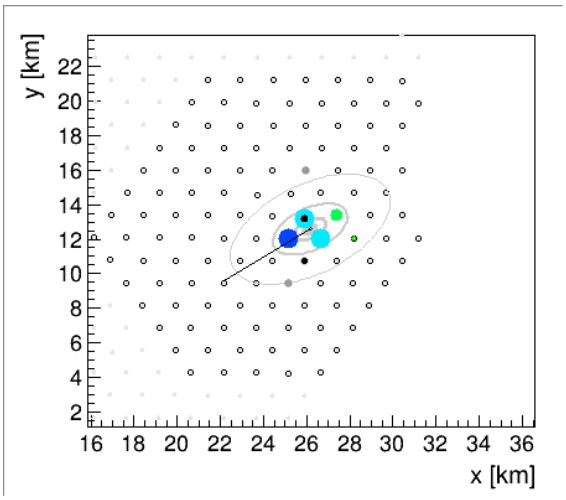
Event no. **46285053** is located at the edge of the Surface Detector array. It has the second largest PC value of all candidate events and the lowest $\mathrm{gL_{LDF}}$ value. The low $\mathrm{gL_{LDF}}$ is caused by two stations with signals far below the LDF fit. Their VEM traces are shown in C.19 and C.20. All four stations used to calculate the risetime parameter are significantly above the risetime benchmark. A measurement of this event from the FD telescope Loma Amarilla is available. The shower maximum was not in the field of view of the telescope (see Fig. C.24). The measured X_{max} is lower than the average X_{max} of photons but compatible with the expectation of protons at about 20 EeV suggesting this event might be caused by a proton.

Event **48726514** has the lowest PC value of all photon candidate events and is just above the candidate cut. It contains two stations below but still compatible with the LDF parametrization of the signal and four stations with risetimes significantly above the risetime benchmark.

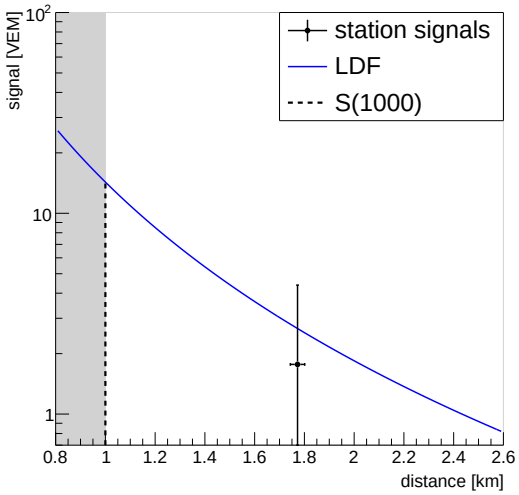
SD event ID	8742491
Date	Nov 22nd, 2009
Photon energy	12.2 EeV
Hadronic energy	(4.78 ± 0.39) EeV
Zenith angle	55.0°
Azimuth angle	216.6°
$g\Delta_{\text{Leeds}}$, gL_{LDF}	0.50, −0.73
PC value ($E > 10$ EeV)	0.77



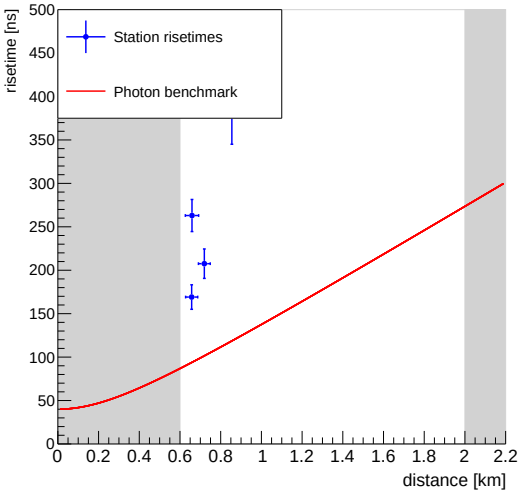
(a) Position of the event in the PCA parameter space



(b) Position of the event on the grid of the Surface Detector



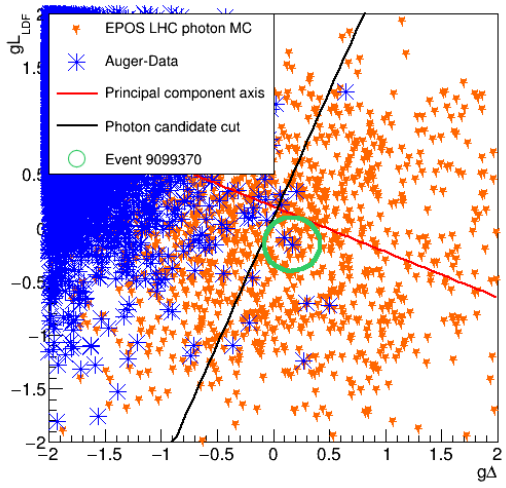
(c) Individual detector station signals compared to the LDF function



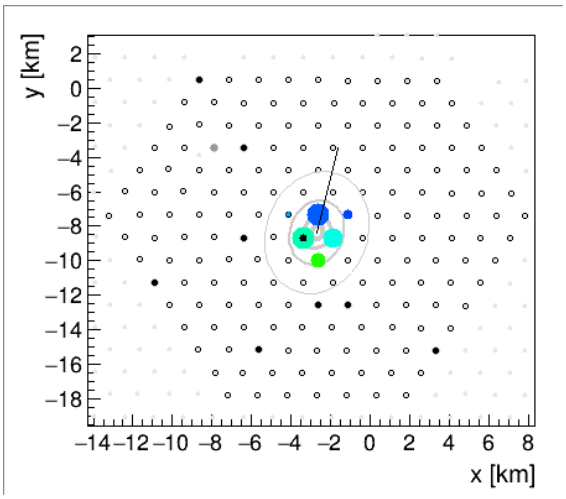
(d) Individual station risetimes compared to the benchmark function

Figure 6.25: Main observables of event 8742491

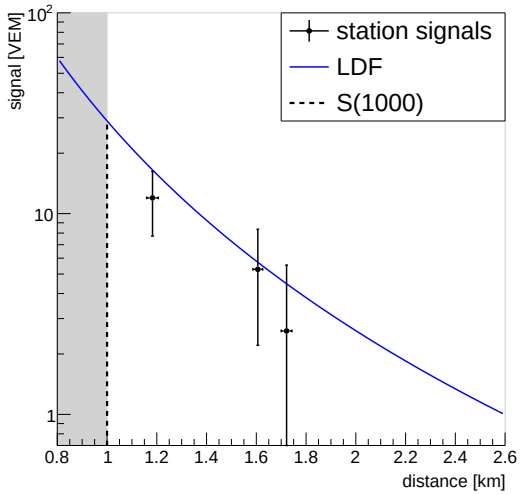
SD event ID	9099370
Date	Feb 1st, 2010
Photon energy	13.6 EeV
Hadronic energy	(6.67 ± 0.42) EeV
Zenith angle	43.4°
Azimuth angle	77.5°
$g\Delta_{Leeds}$, gL_{LDF}	0.17, −0.15
PC value ($E > 10$ EeV)	0.24



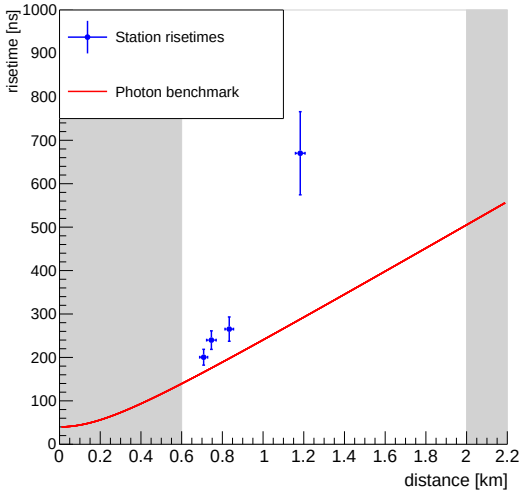
(a) Position of the event in the PCA parameter space



(b) Position of the event on the grid of the Surface Detector



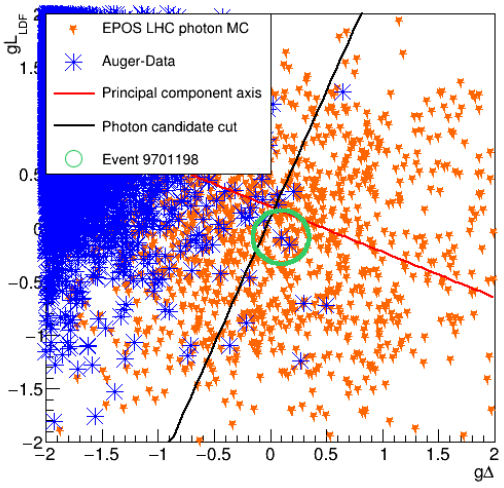
(c) Individual detector station signals compared to the LDF function



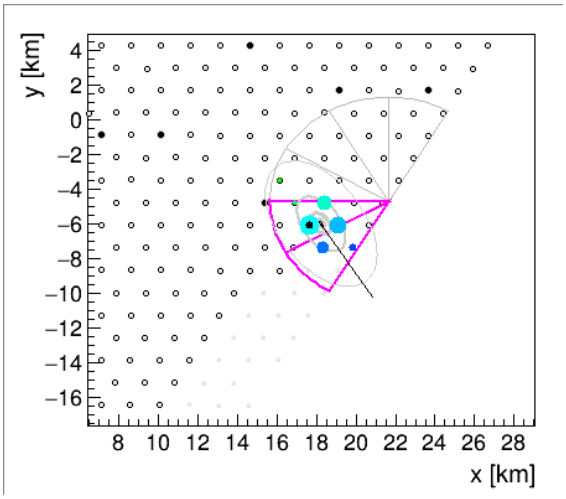
(d) Individual station risetimes compared to the benchmark function

Figure 6.26: Main observables of event 9099370

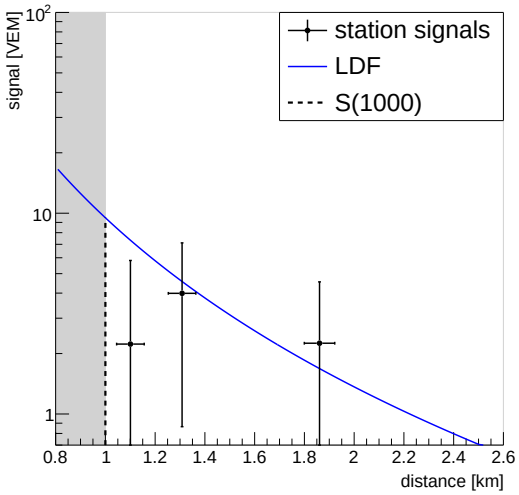
SD event ID	9701198
Date	June 8th, 2010
Photon energy	11.1 EeV
Hadronic energy	(3.33 ± 0.36) EeV
FD energy	(2.19 ± 0.24) EeV
Zenith angle	56.4°
Azimuth angle	301.5°
$X_{\max}$	(1158 ± 43) $\frac{\mathrm{g}}{\mathrm{cm}^2}$
$\mathrm{g}\Delta_{\mathrm{Leeds}}$, $\mathrm{gL}_{\mathrm{LDF}}$	0.10, −0.09
PC value ($E > 10$ EeV)	0.15



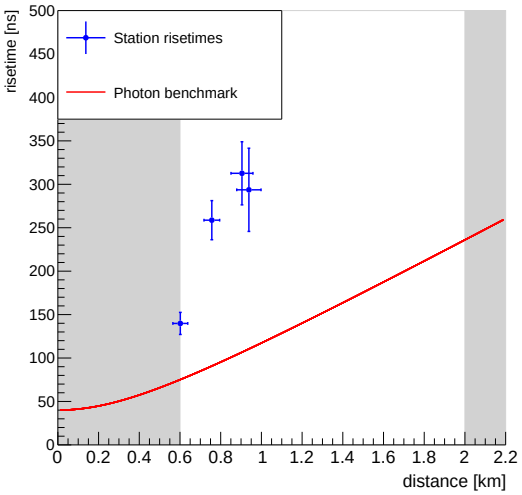
(a) Position of the event in the PCA parameter space



(b) Position of the event on the grid of the Surface Detector



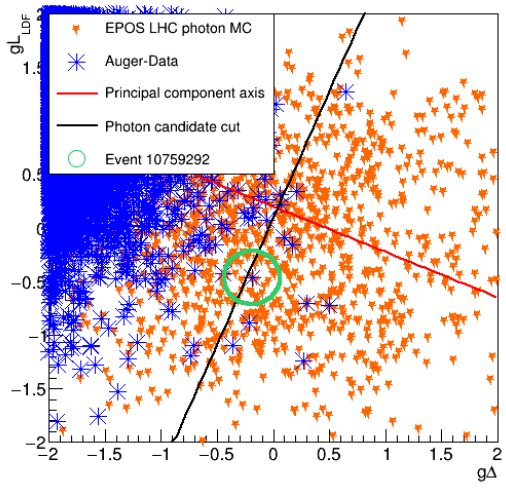
(c) Individual detector station signals compared to the LDF function



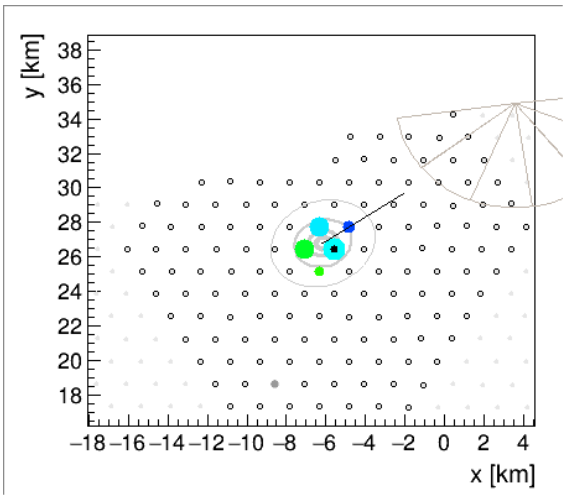
(d) Individual station risetimes compared to the benchmark function

Figure 6.27: Main observables of event 9701198

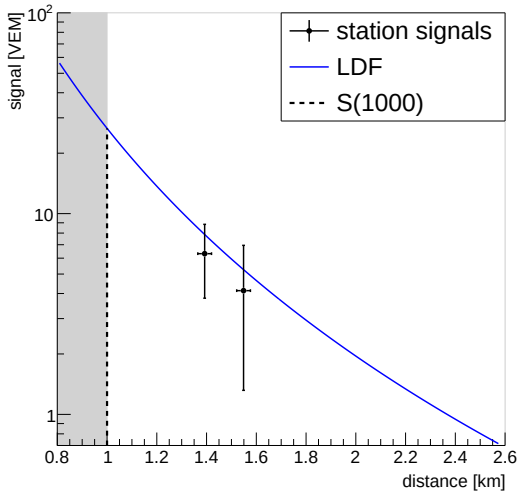
SD event ID	10759292
Date	Dec 12th, 2010
Photon energy	16.0 EeV
Hadronic energy	(5.01 ± 0.33) EeV
Zenith angle	30.9°
Azimuth angle	35.4°
$g\Delta_{\text{Leeds}}$, gL_{LDF}	-0.19, -0.47
PC value ($E > 10$ EeV)	0.03



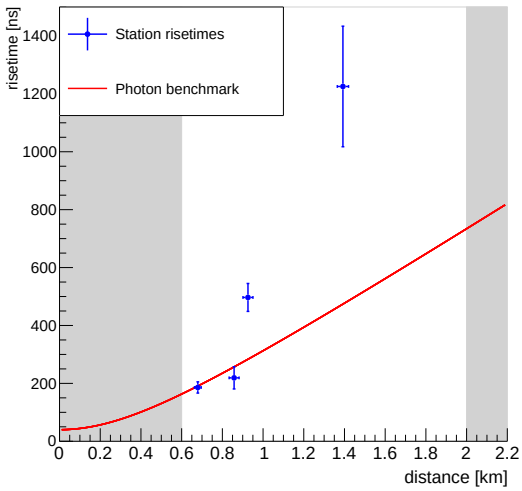
(a) Position of the event in the PCA parameter space



(b) Position of the event on the grid of the Surface Detector



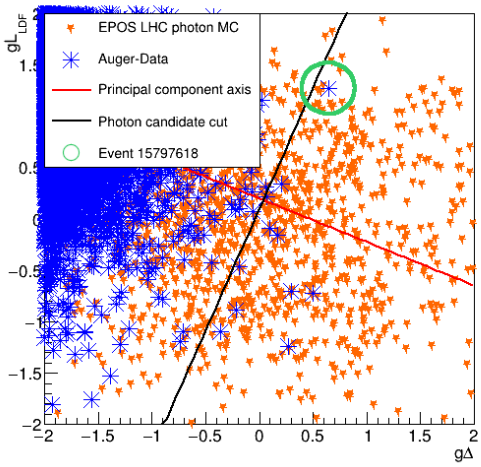
(c) Individual detector station signals compared to the LDF function



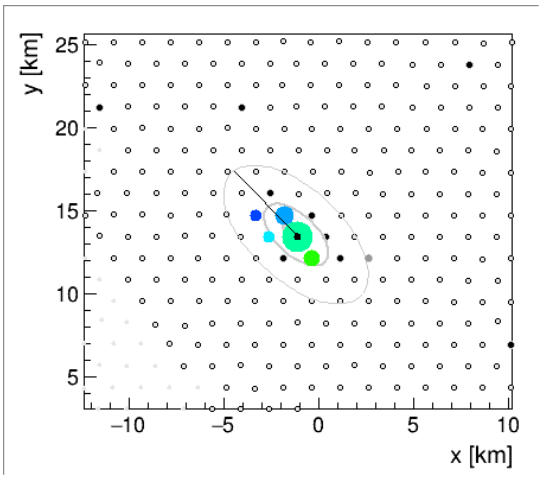
(d) Individual station risetimes compared to the benchmark function

Figure 6.28: Main observables of event 10759292

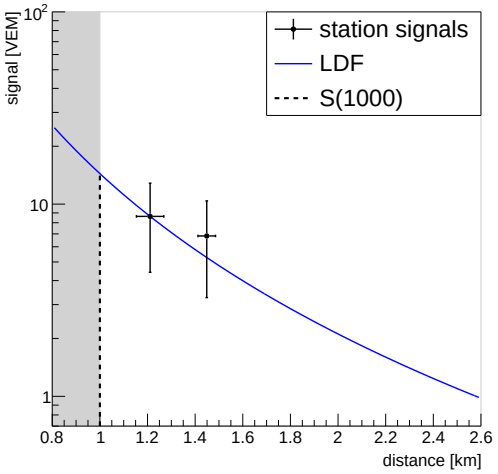
SD event ID	15797618
Date	July 23rd, 2012
Photon energy	24.3 EeV
Hadronic energy	(5.96 ± 0.73) EeV
FD energy (Telescope 1)	(2.93 ± 0.45) EeV
FD energy (Telescope 2)	(4.06 ± 0.15) EeV
Zenith angle	59.3°
Azimuth angle	130.5°
X_{max} (Telescope 1)	(1240 ± 80) $\frac{\mathrm{g}}{\mathrm{cm}^2}$
X_{max} (Telescope 2)	(1260 ± 35) $\frac{\mathrm{g}}{\mathrm{cm}^2}$
$\mathrm{g}\Delta_{\mathrm{Leeds}}$, $\mathrm{gL}_{\mathrm{LDF}}$	0.64, 1.27
PC value ($E > 10$ EeV)	0.12
PC value ($E > 20$ EeV)	0.10



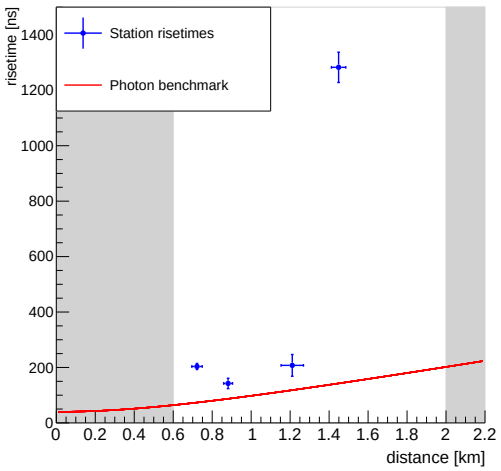
(a) Position of the event in the PCA parameter space



(b) Position of the event on the grid of the Surface Detector



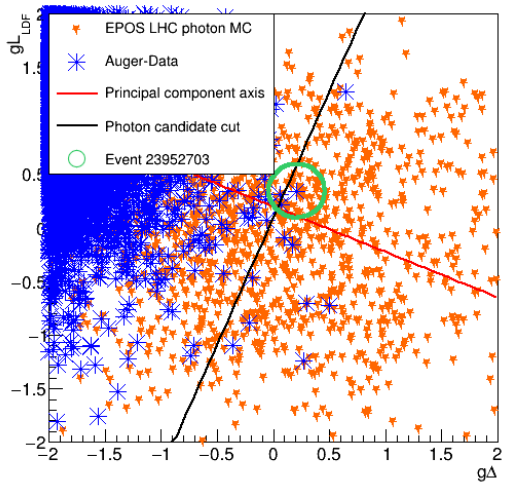
(c) Individual detector station signals compared to the LDF function



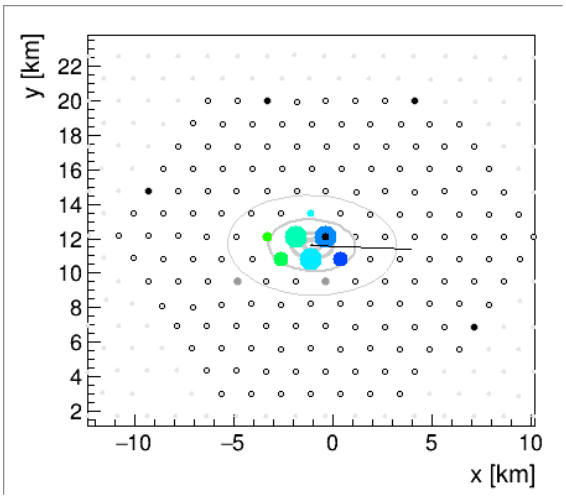
(d) Individual station risetimes compared to the benchmark function

Figure 6.29: Main observables of event 15797618

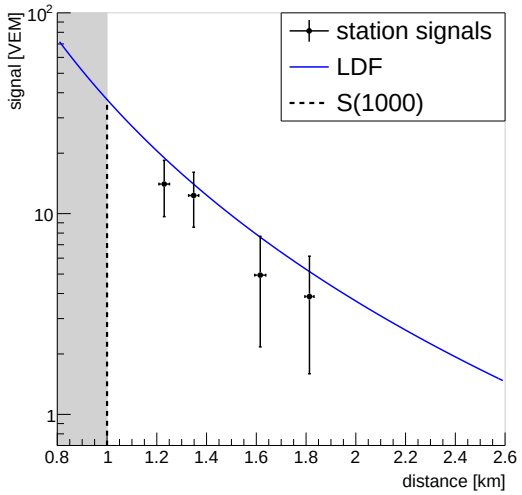
SD event ID	23952703
Date	Nov 4th, 2013
Photon energy	19.2 EeV
Hadronic energy	(9.70 ± 0.54) EeV
Zenith angle	47.6°
Azimuth angle	357.0°
$g\Delta_{Leeds}$, gL_{LDF}	0.21, 0.34
PC value ($E > 10$ EeV)	0.09



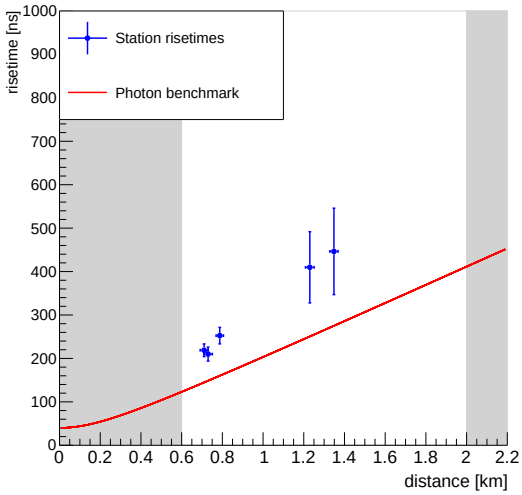
(a) Position of the event in the PCA parameter space



(b) Position of the event on the grid of the Surface Detector



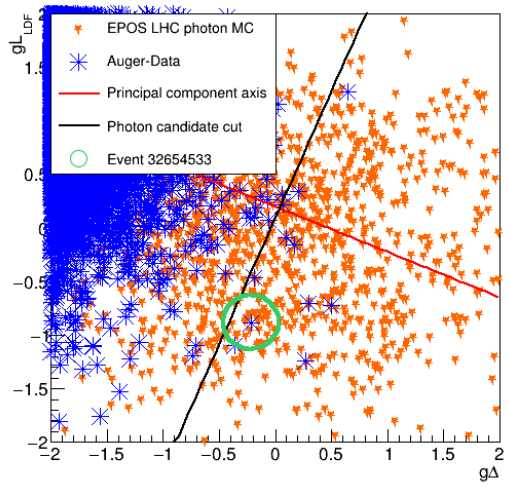
(c) Individual detector station signals compared to the LDF function



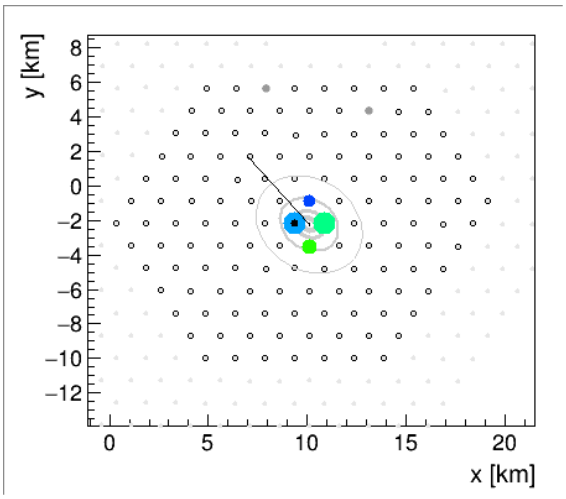
(d) Individual station risetimes compared to the benchmark function

Figure 6.30: Main observables of event 23952703

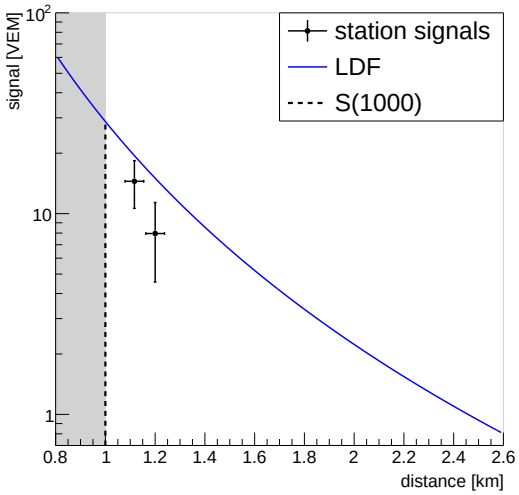
SD event ID	32654533
Date	May 2nd, 2015
Photon energy	14.7 EeV
Hadronic energy	(5.73 ± 0.32) EeV
Zenith angle	35.8°
Azimuth angle	128.5°
$g\Delta_{\text{Leeds}}$, gL_{LDF}	-0.22, -0.88
PC value ($E > 10$ EeV)	0.17



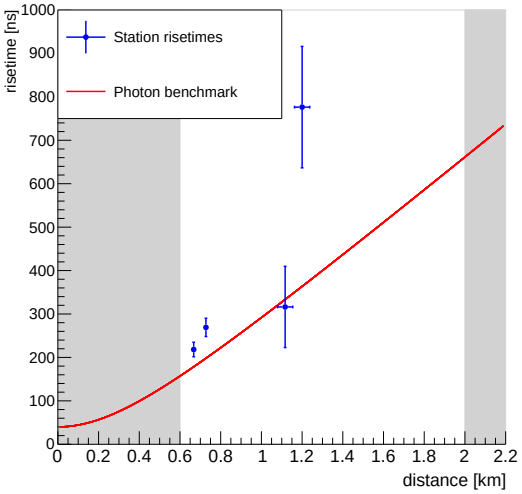
(a) Position of the event in the PCA parameter space



(b) Position of the event on the grid of the Surface Detector



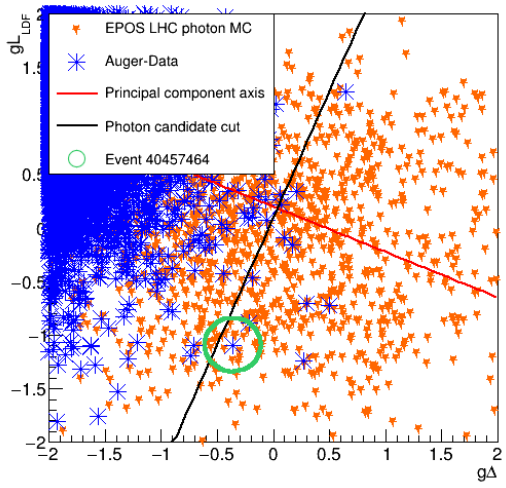
(c) Individual detector station signals compared to the LDF function



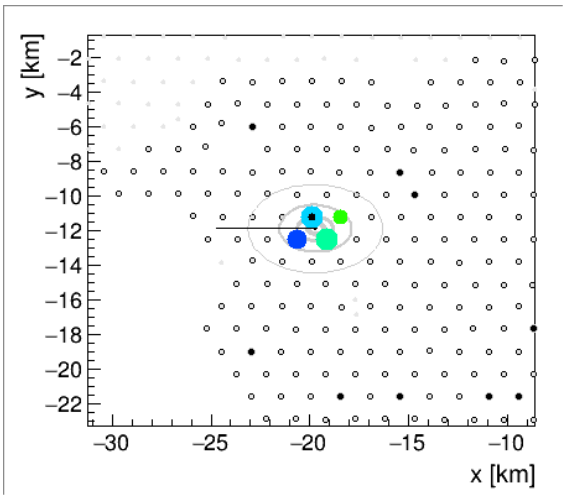
(d) Individual station risetimes compared to the benchmark function

Figure 6.31: Main observables of event 32654533

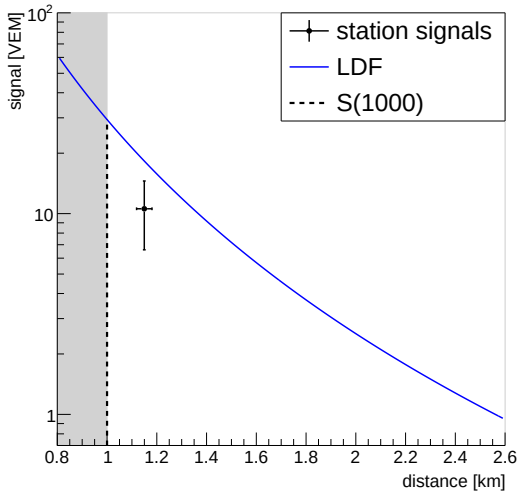
SD event ID	40457464
Date	Nov 25th, 2016
Photon energy	13.7 EeV
Hadronic energy	(6.48 ± 0.38) EeV
Zenith angle	41.4°
Azimuth angle	179.7°
$g\Delta_{\text{Leeds}}$, gL_{LDF}	-0.36, -1.09
PC value ($E > 10$ EeV)	0.12



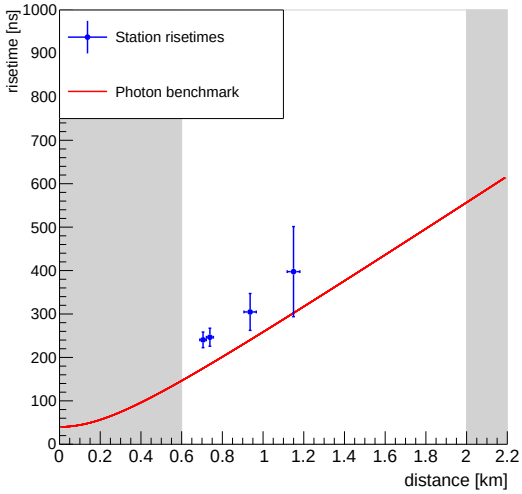
(a) Position of the event in the PCA parameter space



(b) Position of the event on the grid of the Surface Detector



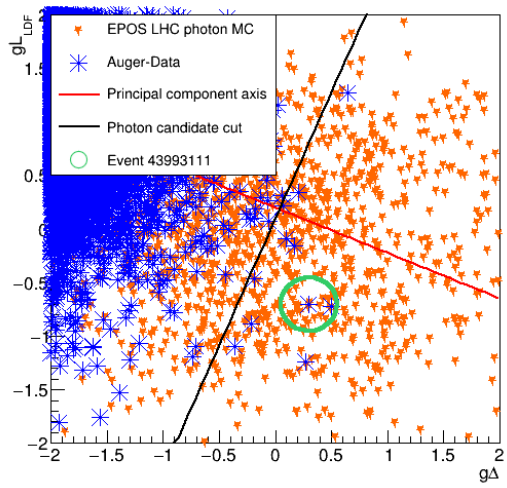
(c) Individual detector station signals compared to the LDF function



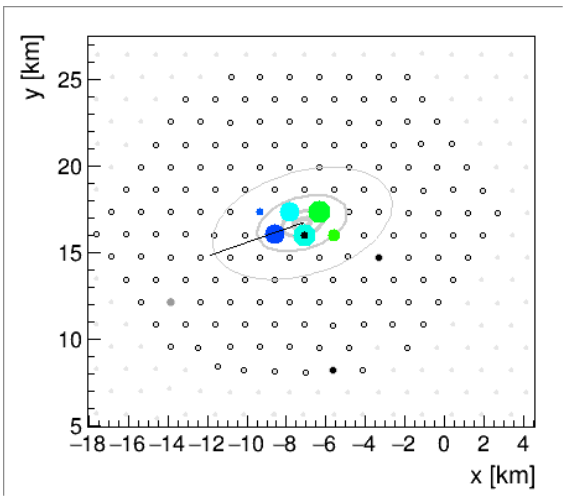
(d) Individual station risetimes compared to the benchmark function

Figure 6.32: Main observables of event 40457464

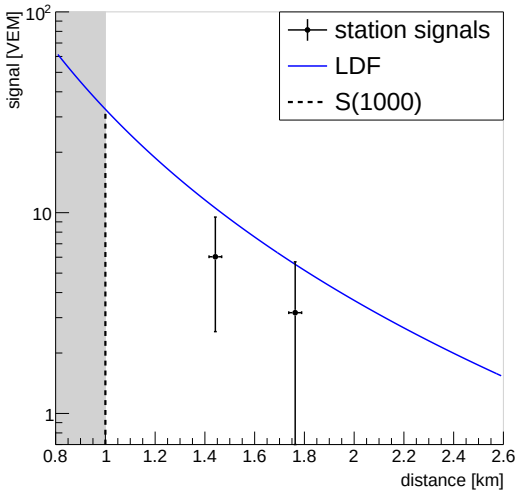
SD event ID	43993111
Date	Aug 10th, 2017
Photon energy	20.1 EeV
Hadronic energy	(9.89 ± 0.58) EeV
Zenith angle	51.8°
Azimuth angle	201.8°
$g\Delta_{\text{Leeds}}$, gL_{LDF}	0.30, −0.71
PC value ($E > 10$ EeV)	0.58
PC value ($E > 20$ EeV)	0.60



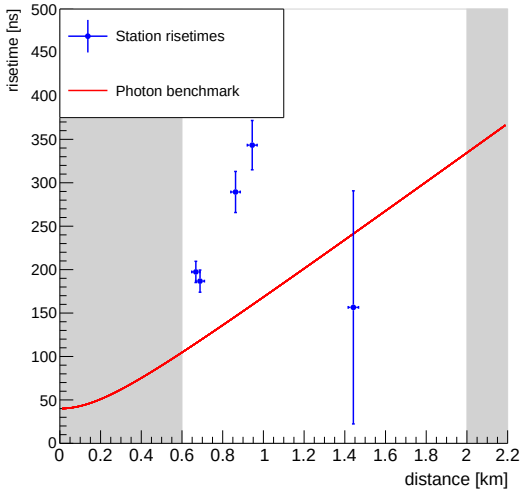
(a) Position of the event in the PCA parameter space



(b) Position of the event on the grid of the Surface Detector



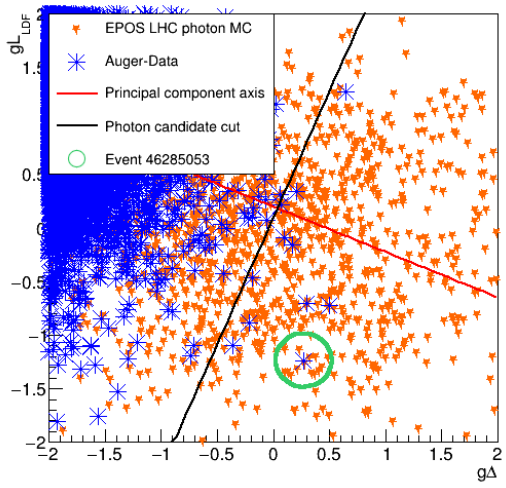
(c) Individual detector station signals compared to the LDF function



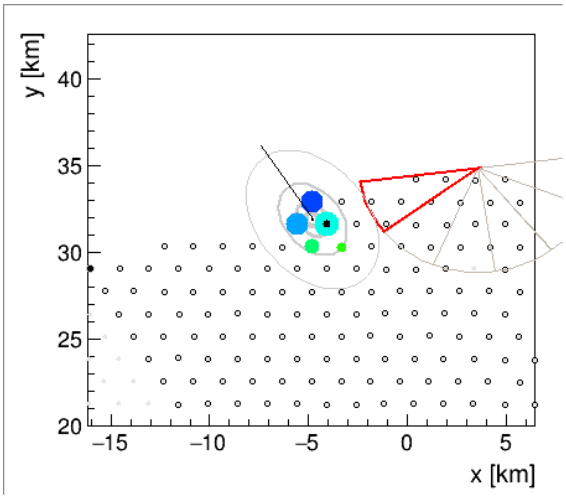
(d) Individual station risetimes compared to the benchmark function

Figure 6.33: Main observables of event 43993111

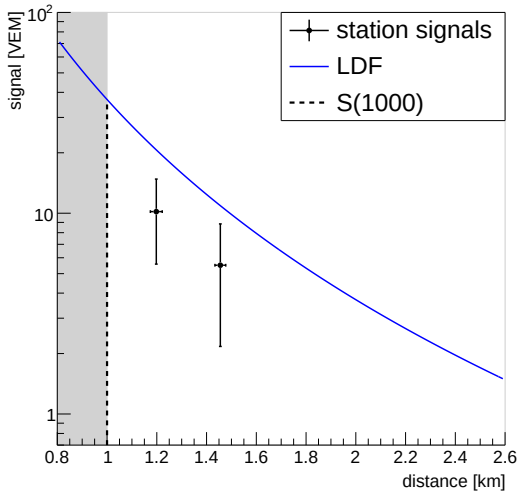
SD event ID	46285053
Date	Jan 20th, 2018
Photon energy	19.6 EeV
Hadronic energy	(9.83 ± 0.57) EeV
FD energy	(12.9 ± 2.6) EeV
Zenith angle	48.1°
Azimuth angle	121.6°
$X_{\max}$	(820 ± 75) $\frac{\mathrm{g}}{\mathrm{cm}^2}$
$\mathrm{g}\Delta_{\mathrm{Leeds}}$, $\mathrm{gL}_{\mathrm{LDF}}$	0.27, −1.24
PC value ($E > 10$ EeV)	0.76



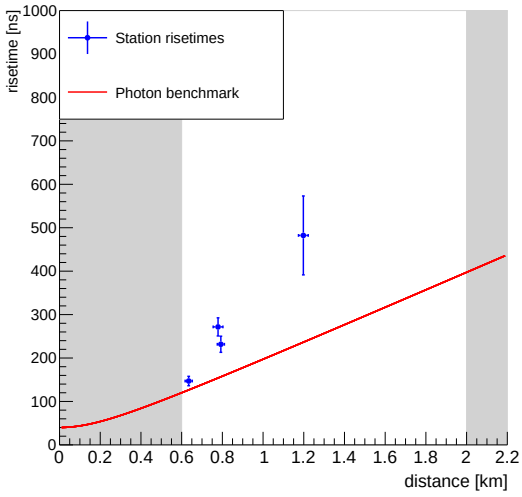
(a) Position of the event in the PCA parameter space



(b) Position of the event on the grid of the Surface Detector



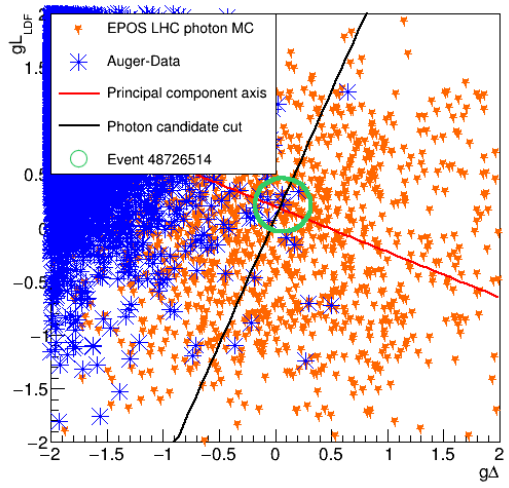
(c) Individual detector station signals compared to the LDF function



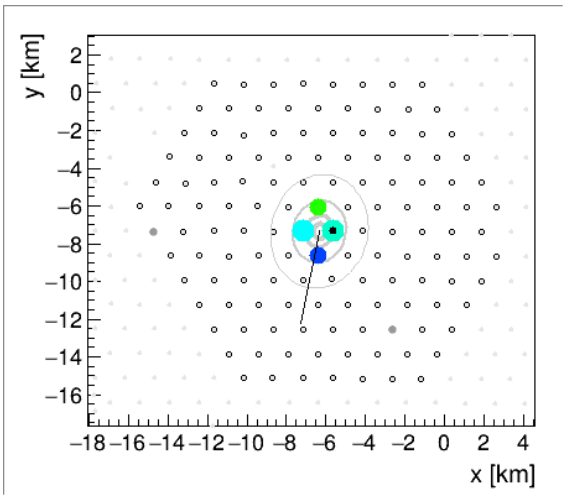
(d) Individual station risetimes compared to the benchmark function

Figure 6.34: Main observables of event 46285053

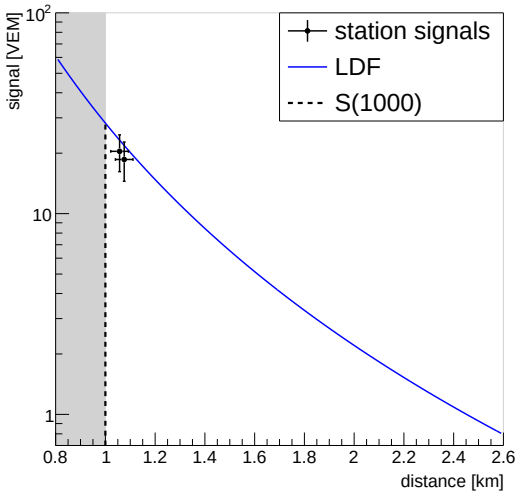
SD event ID	48726514
Date	June 26th, 2018
Photon energy	14.4 EeV
Hadronic energy	(5.65 ± 0.35) EeV
Zenith angle	36.1°
Azimuth angle	259.3°
$g\Delta_{\text{Leeds}}$, gL_{LDF}	0.07, 0.21
PC value ($E > 10$ EeV)	0.01



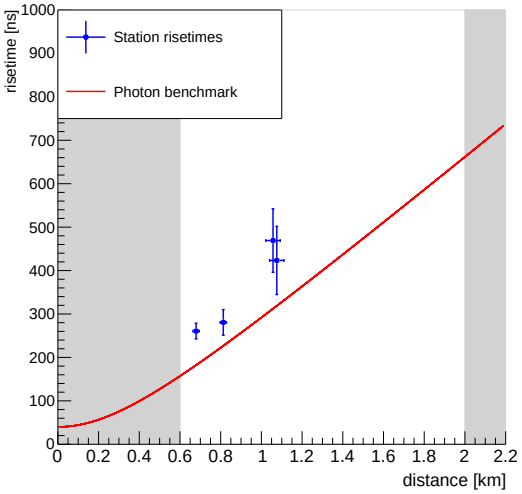
(a) Position of the event in the PCA parameter space



(b) Position of the event on the grid of the Surface Detector



(c) Individual detector station signals compared to the LDF function



(d) Individual station risetimes compared to the benchmark function

Figure 6.35: Main observables of event 48726514

6.5.6 Tests of the Primary Nature of the Candidate Events

The large number of candidate events in the $E > 10\,\mathrm{EeV}$ energy bin may be hinting at the measurement of a photon signal at these energies. On the other hand, the candidate events could be a feature of the tail of the hadronic PC distribution (most likely created by protons). To test the latter hypothesis, the tail of the data PC distribution was fitted with an exponential function. The best fit is shown in Fig. 6.36. If this exponential is integrated from the position of the photon candidate cut to infinity, an expectation of 8.33 candidates is found. The probability to obtain at least the observed 11 candidate events assuming a Poisson distribution with $\lambda = 8.33$ is about 22 %. Thus, the candidate events might be part of an exponential tail of the data PC distribution but could also hint at an excess at large PC values. However, it is important to note that there is no physical reason for the tail to be exponential after the bin-wise renormalization of the photon variables, and therefore this conclusion has to be taken with care.

To test whether the candidate number and distribution can be explained by a low flux of photon air showers, a second fit was performed. For this fit the PC distribution of the photon MC was multiplied by a fitted scaling factor then added to an exponential. The results of this fit are shown in Fig. 6.37. The best fit was obtained using a scaling factor of 7.35 which is equivalent to an integral of 7.35 photons. This fit including a photon component had a χ^2/ndf of 9.5 which is worse than the $\chi^2/\mathrm{ndf} = 8.8$ of a pure exponential fit over the same range. This result might hint at a small photon component in the data, but still depends on the assumption of an exponential model of the PC distribution tail.

To test the probability of the candidate events being due to proton air showers, a set of 6961 CORSIKA proton air showers with energies between $10^{18.5}\,\mathrm{eV}$ and $10^{20.5}\,\mathrm{eV}$ was used. These were generated with an energy spectrum of E^{-1} (then reweighted to E^{-3}) and were thrown into the detector simulations five times each. This resulted in a data set of 34805 proton air showers which created 20 photon candidate events. While most of the proton MC is similarly distributed in the PCA parameter space as the experimental data, the proton candidate events tend to have larger $\mathrm{gL_{LDF}}$ values than the candidate events from the data (see Fig. 6.38). Assuming a proton fraction of $F_p = 10\,\%$ at $10^{19}\,\mathrm{eV}$ [143, 144], it is possible to make a prediction of the number of background candidates using the weighted fraction of proton simulations which passed the candidate cut ($\epsilon^p_{\text{candidate cut}} = 0.29\,\%$)[7]. Applied to the size $N_{\text{data}} = 40570$ of the data sample past all cuts except the candidate cut, $N_{\text{data}} \cdot F_p \cdot \epsilon^p_{\text{candidate cut}} = 11.8$ photon candidates are expected. This prediction is astoundingly close to the observed number of 11 photon candidates. More candidates, however, are expected from simulated proton events than from real air showers since simulated hadronic air showers on average contain lower muon numbers than their real counterparts (see Sec. 2.7).

[7] $\epsilon^p_{\text{candidate cut}} = 0.29\,\%$ is the candidate cut efficiency of the simulated proton events, as introduced later on in Sec. 6.6.3.

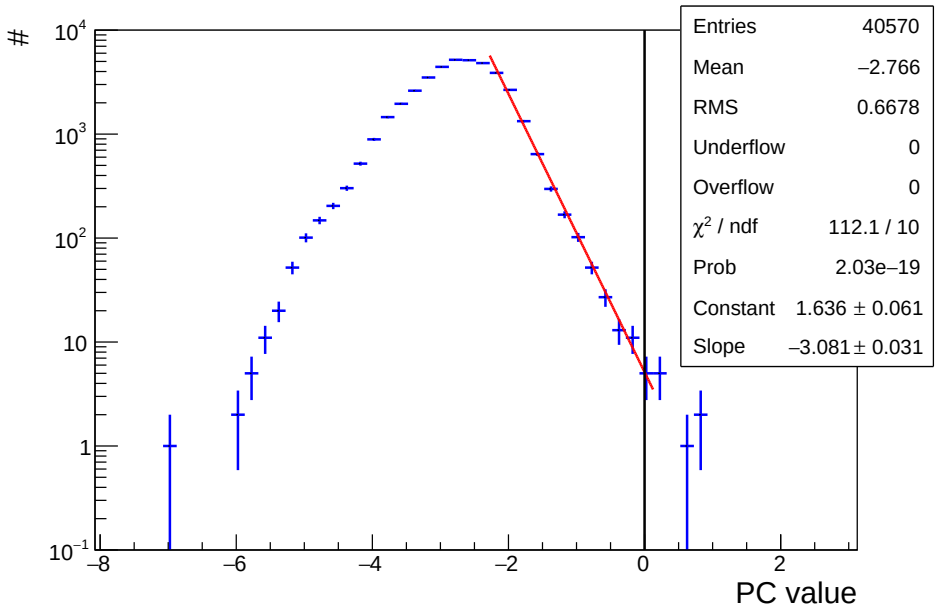


Figure 6.36: Exponential fit (red) to the tail of the data PC distribution (blue). The position of the photon candidate cut is shown in black.

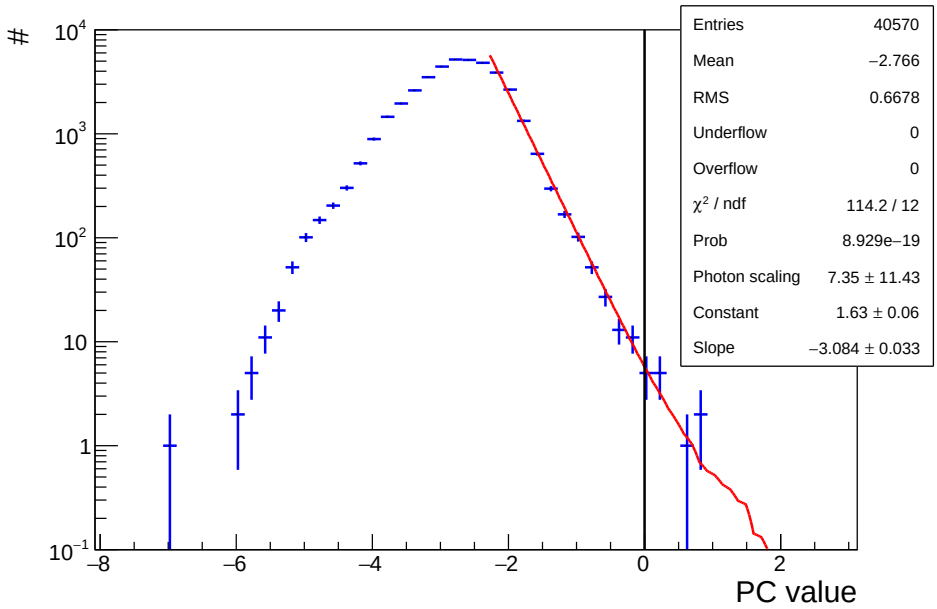


Figure 6.37: Fit of the scaled photon MC PC distribution + exponential to the PC distribution of data. The features at the end of the fit function are caused by the uneven photon MC PC distribution. The photon distribution was scaled to an integral of 7.35 photons.

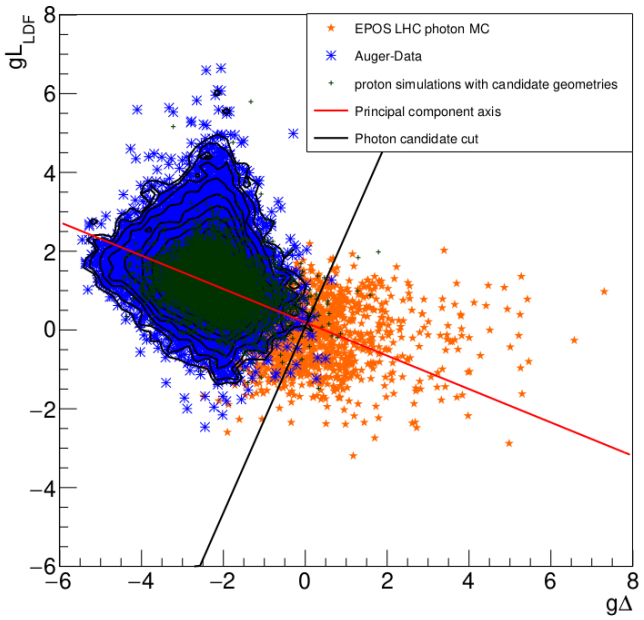


Figure 6.38: PCA results for E > 10 EeV of the burn sample data (blue), the reduced photon MC set (orange), and proton simulations with energies between $10^{18.5}$ eV and $10^{20.5}$ EeV (dark green). The proton events beyond the candidate cut on average have larger gL_{LDF} values than the photon candidates.

Primary proton air showers are particularly likely to create candidate events if one of the leading particles of the first hadronic interaction is a π^0 with a large elasticity, which, with its fast decay to two photons, results in a large electromagnetic shower component. The calculations from [133] were used to estimate the number of protons with high elasticity leading π^0. At a proton fraction of 10 % [143, 144], with EPOS LHC (QGSJETII-04), 8.1 (1.6) events after all cuts except the candidate cut, are expected to have a π^0 carrying at least 50 % of the interaction energy. 10 % of these events will have elasticities of at least 80 % and therefore be nearly indistinguishable from photon air showers. These numbers of course are highly dependent on the proton fraction above $E > 10^{19}$ eV and naturally increase with fractions larger than 10 %.

The tests performed in this Section indicate that the large number of candidates in the $E > 10$ EeV energy bin might hint at a photonic component of the flux but can also possibly be explained by the fluctuations of proton air showers. Because of this, the strength of upper limits on the photon flux in the $E > 10$ EeV energy bin are currently not limited by exposure, but rather by the rejection of background events. Since the real primary nature of the event candidates remains unknown, a conservative upper limit on the primary photon flux has been set taking the photon candidate events into account.

6.5.7 Arrival Directions of the Candidate Events

Photons, unlike nuclei, are not deflected in magnetic fields and point back towards their sources. The hypothesis of the candidates being photons would be strengthened if they pointed back to a common source. To test this, the arrival directions of the 11 photon candidate events have been plotted in galactic coordinates in Fig. 6.39. The background heat map of this Figure shows the arrival directions of all events which passed every cut except the candidate cut. No clustering of the candidate event arrival directions compared to the aperture of the analysis is apparent. In particular, there are no events with arrival directions coinciding within their arrival direction uncertainties. The points of origin of the candidate events have also been tested against the background heat map with a Kolmogorov-Smirnov test resulting in a KS distance of 0.29 and a p value of 0.32 confirming no strong divergence from the exposure.

6.6 Efficiency Calculation

The signal efficiency of the analysis was studied by applying the PCA to photon MC air showers and calculating the (weighted) fraction which are selected as candidate events. In a conservative approach, the combined set of C and UC photon air shower simulations have been used. All efficiency calculations were done assuming a spectral index of $\alpha = -2$ for the photon flux. The analysis contains three main steps which affect the efficiency:

- The air shower has to trigger the detector and be reconstructed successfully (including a successful photon energy reconstruction with X_{ground}-$X_{\text{max}} > -50\,\text{g}\,\text{cm}^{-2}$).
- The quality cuts for the PCA variables have to be met.
- The PC value of the air shower must meet or exceed the candidate cut.

For each of these steps an individual efficiency was calculated:

$$\epsilon = \frac{\sum_i w_{i+}}{\sum_i w_i}, \tag{6.20}$$

where w_{i+} are the MC weights of the selected events and w_i are the MC weights of all events prior to the selection. The statistical uncertainties of these efficiencies were calculated using the methods of [145] and the derivation from [146]:

$$\delta\epsilon = \sqrt{\frac{(\sum_i w_{i+})^2 \sum_i w_{i-}^2 + (\sum_i w_{i-})^2 \sum_i w_{i+}^2}{(\sum_i w_i)^4}}, \tag{6.21}$$

where w_{i-} are the MC weights of all events which were not selected.

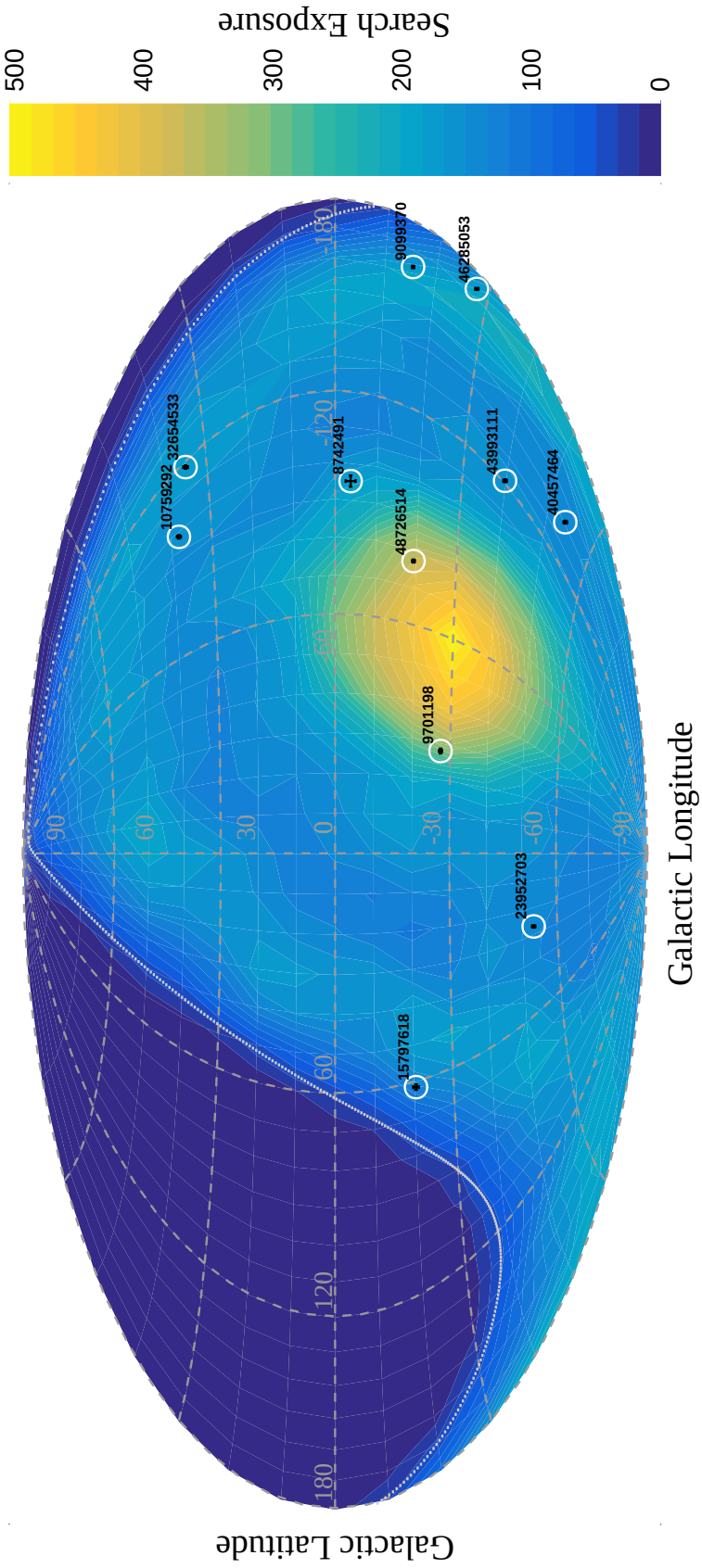


Figure 6.39: Arrival directions of the photon candidate events in galactic coordinates. The background shows a heat map of all data events that passed the selection prior to the candidate cut.

6.6.1 Trigger and Reconstruction Efficiency

The *trigger and reconstruction efficiency* describes the probability that a photon air shower has:

- A 6T5 trigger (see Sec. 3.2.3).
- A reconstructed shower axis and LDF (see Sec. 3.3).
- A successfully reconstructed photon energy above the energy threshold.

To increase the statistics for this calculation, all five simulation/reconstructions of each CORSIKA photon air shower have been used. The trigger and reconstruction efficiencies as a function of the threshold energy are shown in Fig. 6.40. The focus in this Section is put on the reconstruction efficiencies since the trigger efficiency of events at these energies is above 99 %. The reconstruction efficiencies, on the other hand, have a trend of decreasing with MC energy starting at 90 % at 10 EeV and decreasing to 42 % above 100 EeV.

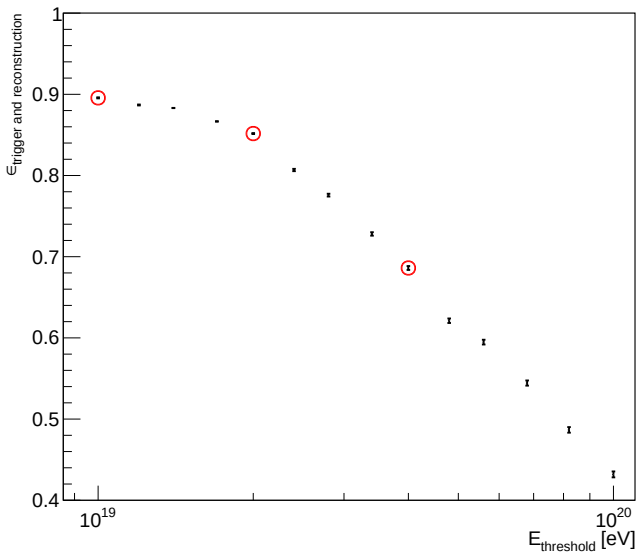


Figure 6.40: Trigger and reconstruction efficiency derived from photon MC for variable threshold MC energies. The red circles indicate the efficiencies of the three energy bins used in the SD photon search.

At lower energies, the effect which causes most air showers to not be successfully reconstructed is the requirement of X_{max} to be not more than $50\,\text{g}\,\text{cm}^{-2}$ below ground required for a successful iterative energy reconstruction (see Sec. 6.4.3). This requirement additionally introduces a zenith angle dependence to the reconstruction efficiency since steep air showers travel a shorter distance through the atmosphere before reaching ground. In Fig. 6.41, this relation between zenith angle and reconstruction efficiency is

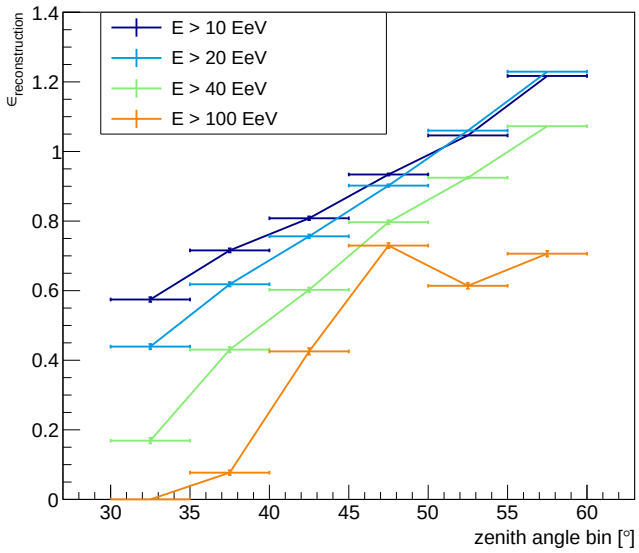
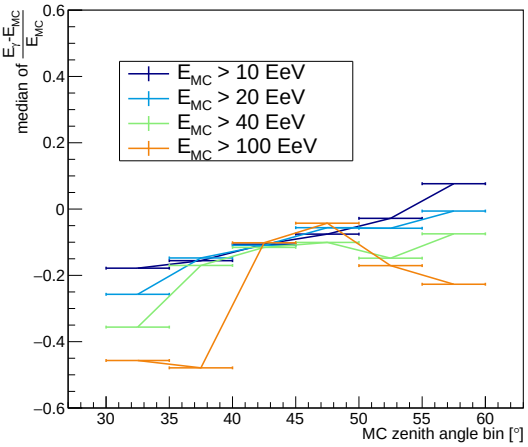


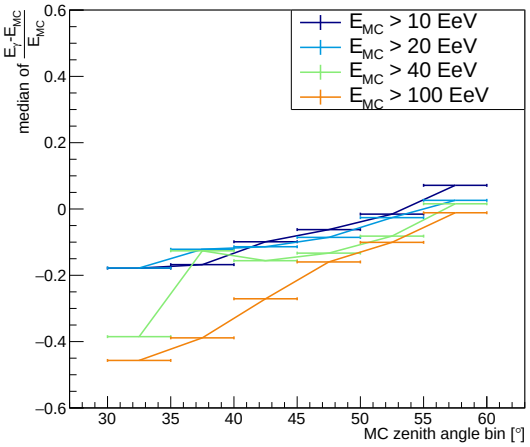
Figure 6.41: Reconstruction efficiency derived from photon MC as a function of the zenith angle at different threshold energies.

shown for different energy thresholds. Shallow, low energetic air showers are considerably more likely to be successfully reconstructed than steep, high energetic ones. In the lowest zenith angle/highest energy bin ($30° < \theta < 35°, E_{\text{MC}} > 100\,\text{EeV}$), none of the 3352 6T5 triggered simulated air showers were successfully reconstructed. The reconstruction efficiencies of the three lower energy bins interestingly increase to values above one at large zenith angles, while the efficiency in the 100 EeV bin starts decreasing at zenith angles greater than 45°. The efficiencies higher than one are caused by a migration from lower energies and the fact that the efficiency is calculated as the number of air showers with a reconstructed energy above the threshold compared to the number of simulated air showers with MC energies above the threshold. The flattening of the 100 EeV curve is caused by an effect of the energy reconstruction having been optimized for UC photons, but being applied to a mixture of both C and UC photons in the efficiency calculations.

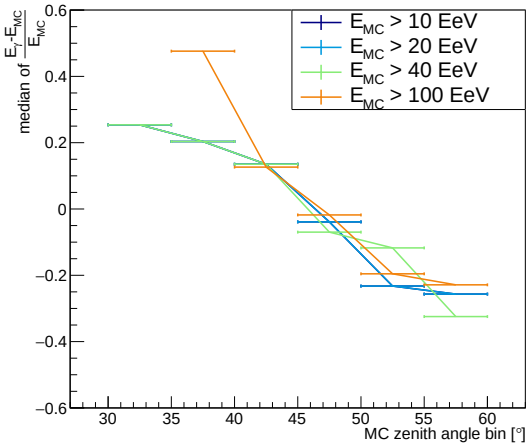
In Fig. 6.42a, the medians of relative differences between reconstructed photon energies and MC energies are shown. At low energies and large zenith angles the photon energy is overestimated for most simulated air showers causing the aforementioned migration effect. Additionally, there is a clear trend of a lower reconstructed energy for high MC energies and low zenith angles (see Sec. 6.4.3). In the 100 EeV graph, it can be seen that the negative energy reconstruction biases decrease in the 30° to 45° region and then again increase near 60°. The results of Fig. 6.42a are split into UC and C photon air showers and plotted separately in Figs. 6.42b and 6.42c. The graphs for UC photons indicate that the iterative energy reconstruction has a stronger negative bias for high energies and low zenith angles (see Fig. 6.6) causing air showers with MC energies above threshold to have their energies reconstructed below threshold. For the C photons the trend in zenith angles is completely reversed with larger zenith angles being system-



(a) C+UC photons



(b) UC photons



(c) C photons

Figure 6.42: Median of the relative difference between reconstructed photon energy and MC energy as a function of the zenith angle. The individual graphs correspond to different threshold energies.

atically reconstructed at lower primary energies. In the low energy threshold and low zenith angle region, the graphs are overlapping since all events have energies exceeding 40 EeV.

The changes in the trend of the 100 EeV graphs in Figs. 6.41 and 6.42a are caused by the C photon simulations dominating the large zenith angle region and UC photon simulations dominating the low zenith angle regions. In Fig. C.5, the separation between UC air showers with mostly overestimated energies and C air showers with mostly underestimated energies is shown for the high energy, large zenith angle region. The reconstruction efficiency is further reduced by C air showers with low zenith angles that have an estimated $X_{\max}$ below ground even though the MC $X_{\max}$ is above it.

6.6.2 Parameter Selection Efficiency

The *parameter selection efficiency* describes the combined efficiency of all quality cuts required by the photon sensitive variables used in the PCA. This includes:

- The Δ_{Leeds} parameter quality flag.
- The L_{LDF} parameter quality flag.
- The no lightning flag.
- A maximum reconstructed photon energy of $10^{20.5}$ eV.

The Δ_{Leeds} and L_{LDF} parameter quality flags are described in more detail in Sec. 3.5. The maximum energy cut at $10^{20.5}$ eV was introduced for technical reasons since the renormalization of the PCA variables (based on photon energy and zenith angle) has been performed up to this energy (see Sec. 6.4.4). In Fig. 6.43, the energy dependent parameter selection efficiencies are shown. They are most prominently influenced by

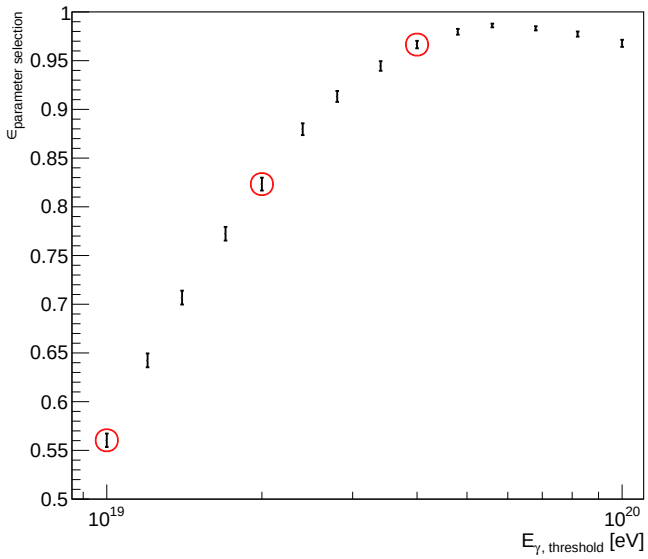


Figure 6.43: Parameter selection efficiency for variable reconstructed photon energy thresholds derived from photon MC. The red circles indicate the efficiencies of the three energy bins used in the SD photon search.

the Δ_{Leeds} parameter quality flag since it requires that each air shower has at least four stations located between 600 m and 2000 m from the shower axis. Since the number of triggered stations in an air shower is strongly correlated to the primary energy, the parameter selection efficiency increases with energy. The efficiency slightly decreases only at the highest energies, above 50 EeV, due to the maximum reconstructed energy cut. At these energies, the photon analysis could gain in efficiency if larger statistics of simulated air showers at high energies are produced. Since the parameter selection

efficiency is still above 95 %, an increase of the amount of high energy photon MC was not considered necessary for the ICRC2019 analysis. In an upcoming analysis, a new parametrization based on S1000 and θ (see Sec. 6.4.4) made the cut on $E_{\mathrm{MC}} < 10^{20.5}\,\mathrm{eV}$ obsolete.

6.6.3 Candidate Cut Efficiency

The probability of a simulated photon air shower, which has met all other requirements, having a PC value which exceeds the candidate cut is called the *candidate cut efficiency*. Since the cut position has been chosen to be at the median of the unconverted photon distribution, the efficiencies in all energy bins have values of roughly 50 %. In Fig. 6.44,

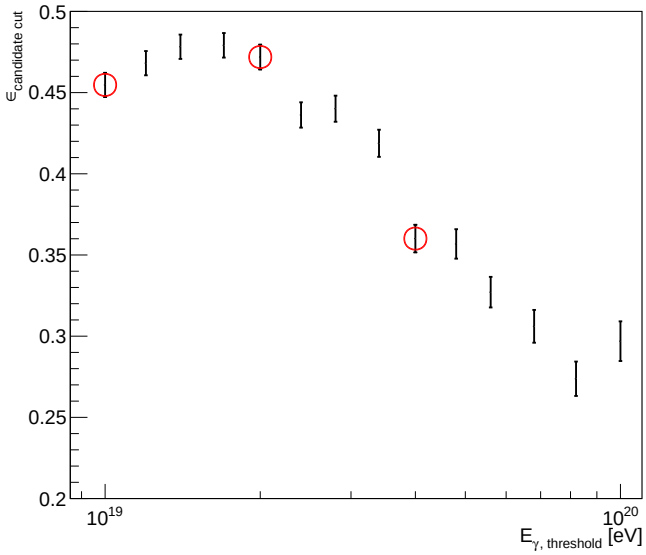


Figure 6.44: Candidate cut efficiency derived from photon MC for variable reconstructed photon energy thresholds. The red circles indicate the efficiencies of the three energy bins used in the SD photon search.

the energy dependence of the candidate cut efficiency is shown. At high energies, towards 100 EeV, the candidate cut efficiency drops significantly falling to values below 30 %.

The falloff at high energies is caused by the conservative approach of using a mixed set of C and UC photon MC for the efficiency studies. Since the analysis was targeted at finding UC photon air showers, the photon median was determined using a set of UC photons only. The large fraction of C air showers in the highest energy bins therefore leads to lower candidate cut efficiencies since converted air showers are significantly more similar to the hadronic background.

6.6.4 Total Efficiency

Since all components of the efficiency are strongly energy dependent, the total efficiency of the analysis has an energy dependence as well, which is clearly shown in Fig. 6.45.

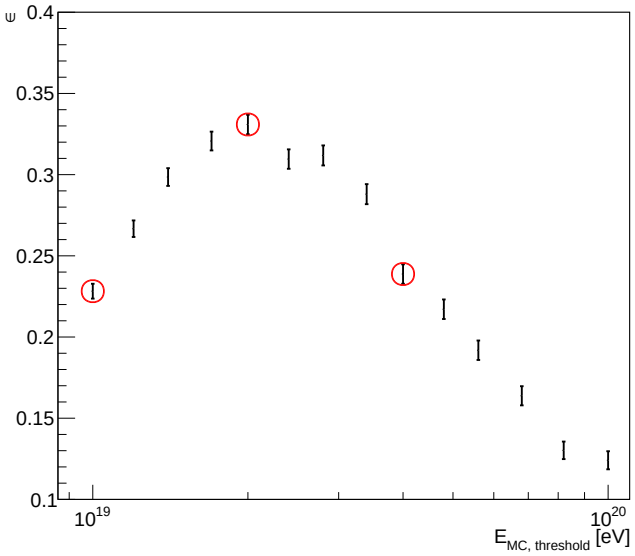


Figure 6.45: Combined efficiency derived from photon MC for variable reconstructed photon energy thresholds. The red circles indicate the efficiencies of the three energy bins used in the SD photon search.

The differing trends result in an increase of the efficiency from 10 EeV to a maximum at about 20 EeV and subsequently a decrease towards the maximum examined energy threshold of 100 EeV. The increase at low energies is driven by the increasing parameter selection efficiency as more detector stations become available to the analysis at higher energies. At energies above 20 EeV, the decrease of the reconstruction and candidate cut efficiencies due to a combination of the X_{max} below ground cut, underestimated energies, and larger fractions of C air showers becomes prevalent.

6.7 Limit on the Diffuse Photon Flux and Fraction

The 11 photon candidate events selected by the Principal Component Analysis were used to calculate the maximum number of candidate events at 95 % confidence level using the Feldman-Cousin method [141]. With this number of Feldman-Cousin candidates, upper limits on the diffuse photon flux, Φ_γ, have been set:

$$\Phi_\gamma^{95\mathrm{CL}}(E_\gamma > E_0) = \frac{N_{\mathrm{cand}}^{\mathrm{FC}}}{\epsilon \cdot A}. \tag{6.22}$$

Here, $N_{\mathrm{cand}}^{\mathrm{FC}}$ is the number of Feldman-Cousin candidates, ϵ is the total efficiency in the integrated energy bin ($E_\gamma > E_0$) and A is the exposure of the detector in the used time period. The exposure has been calculated using the procedure presented in Sec. 3.2.4 considering a zenith angle range of $30° - 60°$ and the data fraction not used for the calculation of the PC axis (98%). The results of the limit calculation are shown in Table 6.4. In Fig. 6.46 a comparison to earlier experimental results is presented.

Data period: Jan 1st 2004 - June 30th 2018			
Exposure: $A = 59698.1$ km^2 sr yr			
Exposure corrected for zenith angle range and burn sample fraction: $A_{\mathrm{corr}} = 39002.8$ km^2 sr yr			
	$E > 10$ EeV	$E > 20$ EeV	$E > 40$ EeV
Burn sample size	768	290	69
$\epsilon_{\text{trigger \& reconstruction}}$ [%]	89.56 ± 0.06	85.17 ± 0.06	68.61 ± 0.24
$\epsilon_{\text{parameter selection}}$ [%]	56.04 ± 0.69	82.33 ± 0.66	96.65 ± 0.38
$\epsilon_{\text{candidate cut}}$ [%]	45.47 ± 0.74	47.19 ± 0.77	36.01 ± 0.85
ϵ [%]	22.82 ± 0.46	33.09 ± 0.60	23.88 ± 0.60
PCA angle	$-23.0°$	$-24.3°$	$-23.8°$
Photon median (PC value)	0.88 ± 0.04	0.99 ± 0.07	1.12 ± 0.14
Number of photon candidate events N_{cand}	11	2	0
Flux limit $\Phi_\gamma^{95\mathrm{CL}}/10^{-3}[\frac{1}{\mathrm{km^2\ sr\ yr}}]$	2.17	0.52	0.33
Fraction limit $F_\gamma^{95\mathrm{CL}}$ [%]	0.52	0.49	1.59

Table 6.4: PCA and limit calculation results for all three integrated energy bins

Currently, as a result of this analysis, the world's strongest upper limits on the diffuse CR photon flux at ultra-high energies have been set. At $E > 10$ EeV, the improvement compared to earlier published upper limits is comparatively low since the results are limited by the hadronic background. This background could not be rejected using high-energy hadronic air shower simulations as this would introduce large uncertainties from the uncertainties in hadronic interaction models at high energies. The measured flux limits are clearly below the levels predicted by top-down models and are entering the regions predicted from the GZK process.

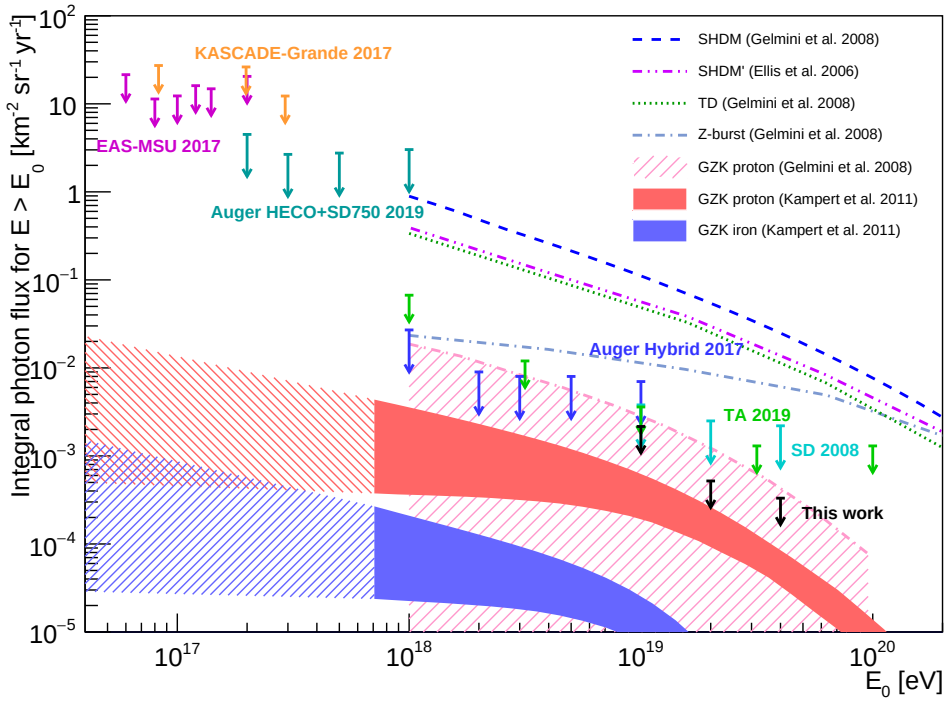


Figure 6.46: Integral Feldman-Cousin upper limits on the primary photon flux at a 95 % confidence level as compared to model predictions [83, 84, 85] and other experimental limits at 95 % CL [74, 79, 80] and 90 % CL [77, 78].

In addition to the flux limits, upper limits on the primary photon fraction have also been set using the cosmic ray energy spectrum measurement from [13] (see Table 6.4, Fig. 6.47). Again, the upper limits strongly disfavor the predictions of top-down models. Furthermore, the predictions from the GZK process are disfavored in the case of a purely protonic CR composition above 10 EeV.

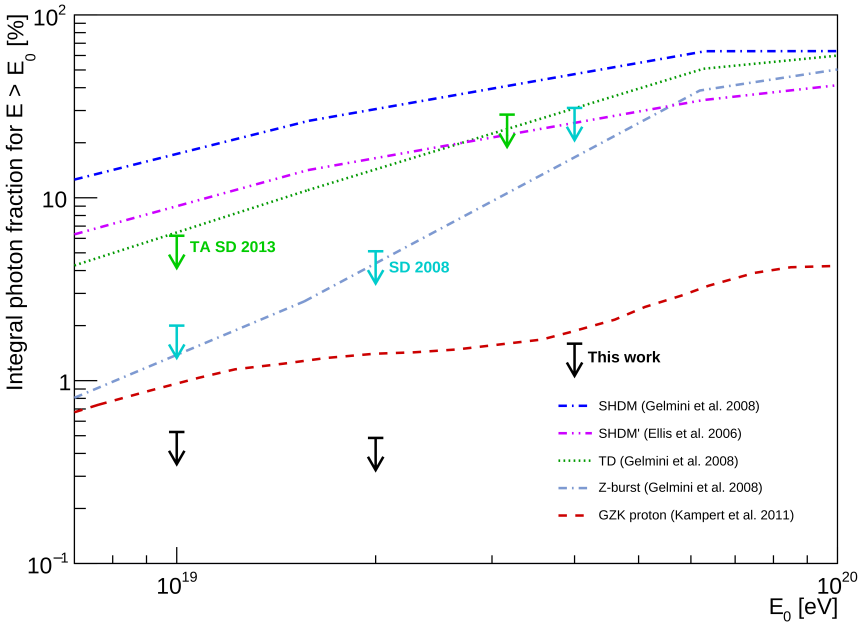


Figure 6.47: Feldman-Cousin upper limits on the primary photon fraction at a 95% confidence level as compared to other experimental limits [80, 86] and model predictions [83, 84, 85]. The fractions have been calculated based on the cosmic ray energy spectrum measurement from [13] which uses data from the Pierre Auger Observatory taken until 2019.

Chapter 7

Summary and Outlook

In this thesis, a search for ultra-high-energy photons with energies above 10^{19} eV is presented which resulted in an improvement of the world's strongest upper limits on both the diffuse photon flux and fraction at these energies. In order to achieve this goal, multiple improvements to the reconstruction procedures of the Pierre Auger Collaboration have been made. In Chapter 4, a procedure developed to improve the station based trigger mechanisms of the Surface Detector was described. First results from these trigger studies have shown that improvements of the triggering mechanisms, especially for photon air showers, can be achieved. The final determination of the trigger parameters is now possible as first measurement data have recently been taken. In Chapter 5, newly developed data cleaning procedures were presented which have been shown to work very reliably. These procedures significantly improve upon the bias prone by-eye scanning of events formerly needed to remove faulty PMT signals from the data. As a result of this work, the search for problematic detector components, which formerly had to be updated regularly, is now fully automated. In Chapter 6, upper limits to the diffuse photon flux have been set using 13.5 years of measurement data combined with photon air shower simulations using up-to-date hadronic interaction models (EPOS LHC and FLUKA). These limits are the main result of the thesis. The compatibility of the surprisingly large number of identified photon candidates in the lowest integrated energy bin ($E > 10\,\mathrm{EeV}$) has been tested against a photon as well as a proton background hypothesis.

The upper limits on the diffuse photon flux and fraction at 95% CL are summarized in Table 7.1. These are the world's best limits on the ultra-high-energy photon flux and fraction and represent a major improvement over previously published limits. These limits are strict enough to present a challenge to many theoretical models that predict a flux of ultra-high-energy photons e.g. as decay products.

	$E > 10$ EeV	$E > 20$ EeV	$E > 40$ EeV
Flux limit $\Phi_\gamma^{95\mathrm{CL}}/10^{-3}[\frac{1}{\mathrm{km}^2\ \mathrm{sr\ yr}}]$	2.17	0.52	0.33
Fraction limit $F_\gamma^{95\mathrm{CL}}$ [%]	0.52	0.49	1.59

Table 7.1: Upper limits on the diffuse photon flux and fraction at 95% CL

Outside of this thesis, the results of the trigger study (Chapter 4) as well as the automated data quality improvement procedures (Chapter 5) have been published as internal collaboration notes [140, 147]. The results of the photon search have been presented at the International Cosmic Ray Conference (ICRC) 2019 [2, 148], and have already been used to constrain the mass and lifetime of super-heavy dark matter particles [149]. A collaboration publication with updated upper limits to the diffuse photon flux and fraction is in preparation and well advanced.

Both, the automated data cleaning procedures and the calculation of the event variables used in the photon search have been implemented in the $\overline{\text{Off}}\underline{\text{line}}$ framework and are now easily available to the collaboration for further studies. The data cleaning procedures have become a part of all shape-sensitive SD studies and have already been used in published searches for a neutrino flux by the Pierre Auger Collaboration [150]. The calculation of the event variables for the photon search are currently being used for follow-up studies [151] combining the photon search in a multi-messenger approach with gravitational wave measurements from the LIGO and Virgo detectors [152].

Several parts of the analysis offer the potential for further improvement to the presented upper limits on the diffuse photon flux and fraction. This is especially true at the highest energies, the analysis of which has been shown to be limited by the photon energy reconstruction which relies on a requirement that X_{max} be less than $50\,\text{g}\,\text{cm}^{-2}$ below ground. With the tabular method introduced in Sec. 6.4.3, a part of this parameter space can be recovered. Furthermore, the limits could be improved by adapting the photon candidate cut as presented in Sec. 6.5.2. The possibility of the photon event candidates to be, at least partly, created by proton air showers with leading π^0 is currently under extensive research. From the experimental point of view, the ongoing AugerPrime upgrade (see Sec. 3.2.1) of the Pierre Auger Observatory is especially promising for photon studies. Most importantly, the newly installed scintillation detectors will allow a better differentiation between the muonic and electromagnetic signals in the Surface Detector stations and thus greatly increase the power to discriminate between photons and hadrons.

Bibliography

[1] Alexander Aab et al. Features of the Energy Spectrum of Cosmic Rays above $2.5 10^{18}$ eV Using the Pierre Auger Observatory. *Phys. Rev. Lett.*, 125(12):121106, 2020.

[2] Julian Rautenberg. Limits on ultra-high energy photons with the Pierre Auger Observatory. *PoS*, ICRC2019:398, 2020.

[3] Victor Hess. On the observations of the penetrating radiation during seven balloon flights, 2018.

[4] Victor F. Hess - Biographical. NobelPrize.org. Nobel Media AB 2020. `https://www.nobelprize.org/prizes/physics/1936/hess/biographical/`. (accessed February 7th, 2020).

[5] G. Khristiansen G. Kulikov. On the size spectrum of extensive air showers. *Sov.Phys.JETP,8,441*, 1959.

[6] B. Peters. Primary cosmic radiation and extensive air showers. *Nuovo Cim.,22,800*, 1961.

[7] Jorg R. Hoerandel. Models of the knee in the energy spectrum of cosmic rays. *Astropart. Phys.*, 21:241–265, 2004.

[8] W. D. Apel et al. KASCADE-Grande measurements of energy spectra for elemental groups of cosmic rays. *Astropart. Phys.*, 47:54–66, 2013.

[9] W. D. Apel et al. Kneelike Structure in the Spectrum of the Heavy Component of Cosmic Rays Observed with KASCADE-Grande. *Phys. Rev. Lett.*, 107:171104, Oct 2011.

[10] D. J. Bird et al. The Cosmic ray energy spectrum observed by the Fly's Eye. *Astrophys. J.*, 424:491–502, 1994.

[11] Veniamin Berezinsky, Askhat Gazizov, and Svetlana Grigorieva. On astrophysical solution to ultrahigh energy cosmic rays. *Phys. Rev. D*, 74:043005, Aug 2006.

[12] Lorenzo Caccianiga. Anisotropies of the Highest Energy Cosmic-ray Events Recorded by the Pierre Auger Observatory in 15 years of Operation. *PoS*, ICRC2019:206, 2020.

[13] Valerio Verzi. Measurement of the energy spectrum of ultra-high energy cosmic rays using the Pierre Auger Observatory. *PoS*, ICRC2019:450, 2020.

[14] Frank G. Schröder. News from Cosmic Ray Air Showers (ICRC 2019 – Cosmic Ray Indirect Rapport). *PoS*, ICRC2019:030, 2020.

[15] M. Amenomori et al. The All-particle spectrum of primary cosmic rays in the wide energy range from 10^{14} eV to 10^{17} eV observed with the Tibet-III air-shower array. *Astrophys. J.*, 678:1165–1179, 2008.

[16] D. Kostunin et al. Seven years of Tunka-Rex operation. *PoS*, ICRC2019:319, 2020.

[17] N.M. Budnev et al. The primary cosmic-ray energy spectrum measured with the Tunka-133 array. *Astropart. Phys.*, 117:102406, 2020.

[18] Mario Bertaina et al. KASCADE-Grande energy spectrum of cosmic rays interpreted with post-LHC hadronic interaction models. *PoS*, ICRC2015:359, 2016.

[19] Fahim Varsi et al. Energy spectrum and composition measurements of cosmic rays from GRAPES-3 experiment. *PoS*, ICRC2019:449, 2020.

[20] R. Alfaro et al. All-particle cosmic ray energy spectrum measured by the HAWC experiment from 10 to 500 TeV. *Phys. Rev.*, D96(12):122001, 2017.

[21] Ramesh Koirala and Thomas K. Gaisser. Low Energy Cosmic Ray Spectrum from 250 TeV to 10 PeV using IceTop. *PoS*, ICRC2019:318, 2020.

[22] Karen Andeen and Matthias Plum. Cosmic Ray Spectrum and Composition from PeV to EeV from the IceCube Neutrino Observatory. *PoS*, ICRC2019:172, 2020.

[23] Dmitri Ivanov. Energy Spectrum Measured by the Telescope Array. *PoS*, ICRC2019:298, 2020.

[24] J. George, K. Lave, M. Wiedenbeck, W. Binns, A. Cummings, A. Davis, G. de Nolfo, Paul Hink, M. Israel, R. Leske, Richard Mewaldt, L. Scott, E. Stone, T. Rosenvinge, and Nathan Yanasak. Elemental composition and energy spectra of galactic cosmic rays during solar cycle 23. *The Astrophysical Journal*, 698:1666, 06 2009.

[25] Hans Peter Dembinski, Ralph Engel, Anatoli Fedynitch, Thomas Gaisser, Felix Riehn, and Todor Stanev. Data-driven model of the cosmic-ray flux and mass composition from 10 GeV to 10^{11} GeV. *PoS*, ICRC2017:533, 2018. [35,533(2017)].

[26] Karl-Heinz Kampert and Michael Unger. Measurements of the cosmic ray composition with air shower experiments. *Astroparticle Physics*, 35(10):660–678, May 2012.

[27] Carlos J. Todero Peixoto. Estimating the Depth of Shower Maximum using the Surface Detectors of the Pierre Auger Observatory. *PoS*, ICRC2019:440, 2020.

[28] Alexey Yushkov. Mass Composition of Cosmic Rays with Energies above $10^{17.2}$ eV from the Hybrid Data of the Pierre Auger Observatory. *PoS*, ICRC2019:482, 2020.

[29] J. Abraham et al. Measurement of the Depth of Maximum of Extensive Air Showers above 10^{18} eV. *Phys. Rev. D*, 104:091101, 2010.

[30] William Hanlon. Telescope Array 10 Year Composition. *PoS*, ICRC2019:280, 2020.

[31] Marco Stein Muzio, Michael Unger, and Glennys R. Farrar. Constraints on UHECR sources and their environments, from fitting UHECR spectrum and composition, and neutrinos and gammas. *PoS*, ICRC2019:364, 2020.

[32] A. M. Hillas. The origin of ultra-high-energy cosmic rays. *Annual Review of Astronomy and Astrophysics*, 22(1):425–444, 1984.

[33] Rafael Alves Batista et al. Open Questions in Cosmic-Ray Research at Ultrahigh Energies. *Front. Astron. Space Sci.*, 6:23, 2019.

[34] Enrico Fermi. On the origin of the cosmic radiation. *Phys. Rev.*, 75:1169–1174, Apr 1949.

[35] R. D. Blandford and J. P. Ostriker. Particle acceleration by astrophysical shocks. *ApJ*, 221:L29–L32, Apr 1978.

[36] V. Berezinsky, M. Kachelriess, and A. Vilenkin. Ultrahigh-energy cosmic rays without GZK cutoff. *Phys. Rev. Lett.*, 79:4302–4305, 1997.

[37] V. A. Kuzmin and V. A. Rubakov. Ultrahigh-energy cosmic rays: A Window to postinflationary reheating epoch of the universe? *Phys. Atom. Nucl.*, 61:1028, 1998.

[38] A. Vilenkin and E. P. S. Shellard. *Cosmic Strings and Other Topological Defects*. Cambridge University Press, 2000.

[39] M. Kachelriess. The rise and fall of top-down models as main UHECR sources. In *20th Rencontres de Blois on Challenges in Particle Astrophysics*, pages 215–224, 2008.

[40] T. Weiler. Resonant absorption of cosmic-ray neutrinos by the relic-neutrino background. *Phys. Rev. Lett.*, 49:234–237, Jul 1982.

[41] R. Aloisio, V. Berezinsky, and A. Gazizov. Transition from galactic to extragalactic cosmic rays. *Astropart. Phys.*, 39-40:129–143, 2012.

[42] Roberto Aloisio. Acceleration and propagation of ultra high energy cosmic rays. *PTEP*, 2017(12):12A102, 2017.

[43] Glennys R. Farrar and Michael S. Sutherland. Deflections of UHECRs in the Galactic magnetic field. *JCAP*, 05:004, 2019.

[44] Kenneth Greisen. End to the cosmic-ray spectrum? *Phys. Rev. Lett.*, 16:748–750, Apr 1966.

[45] G. T. Zatsepin and V. A. Kuzmin. Upper limit of the spectrum of cosmic rays. *JETP Lett.*, 4:78–80, 1966. [Pisma Zh. Eksp. Teor. Fiz.4,114(1966)].

[46] Christopher Heiter, Daniel Kuempel, David Walz, and Martin Erdmann. Production and propagation of ultra-high energy photons using CRPropa 3. *Astropart. Phys.*, 102:39–50, 2018.

[47] D Harari, S Mollerach, and E Roulet. On the ultrahigh energy cosmic ray horizon. *Journal of Cosmology and Astroparticle Physics*, 2006(11):012–012, nov 2006.

[48] Alexander Aab et al. The Pierre Auger Observatory Upgrade - Preliminary Design Report. *arXiv e-prints*, page arXiv:1604.03637, April 2016.

[49] Alexander Aab et al. Combined fit of spectrum and composition data as measured by the Pierre Auger Observatory. *JCAP*, 1704(04):038, 2017. [Erratum: JCAP1803,no.03,E02(2018)].

[50] Peter Tinyakov. Latest results from the telescope array. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 742:29 – 34, 2014. 4th Roma International Conference on Astroparticle Physics.

[51] Markus Risse and Piotr Homola. Search for ultrahigh energy photons using air showers. *Mod. Phys. Lett.*, A22:749–766, 2007.

[52] W. Heitler. *The quantum theory of radiation*, volume 5 of *International Series of Monographs on Physics*. Oxford University Press, Oxford, 1936.

[53] J. Matthews. A heitler model of extensive air showers. *Astroparticle Physics*, 22(5):387 – 397, 2005.

[54] Darko Veberic. Having Fun with Lambert W(x) Function. *arXiv e-prints*, page arXiv:1003.1628, March 2010.

[55] Jorg R. Hoerandel. Cosmic rays from the knee to the second knee: 10^4 to 10^{18} eV. *Mod. Phys. Lett.*, A22:1533–1552, 2007. [,63(2006)].

[56] Lev Davidovich Landau and I I Pomeranchuk. The limits of applicability of the theory of Bremsstrahlung by electrons and of the creation of pairs at large energies. *Dokl. Akad. Nauk SSSR*, 92:535, 1953.

[57] Lisa Gerhardt and Spencer R. Klein. Electron and photon interactions in the regime of strong Landau-Pomeranchuk-Migdal suppression. *Phys. Rev. D*, 82:074017, Oct 2010.

[58] Peter Arnold and Shahin Iqbal. The LPM effect in sequential bremsstrahlung. *JHEP*, 04:070, 2015. [Erratum: JHEP09,072(2016)].

[59] Todor Stanev, Ch. Vankov, R. E. Streitmatter, R. W. Ellsworth, and Theodore Bowen. Development of ultrahigh-energy electromagnetic cascades in water and lead including the Landau-Pomeranchuk-Migdal effect. *Phys. Rev. D*, 25:1291–1304, Mar 1982.

[60] J. Alvarez-Muniz and E. Zas. EeV hadronic showers in ice: The LPM effect. In *26th International Cosmic Ray Conference*, volume 1, page 506, 1999.

[61] Piotr Homola, Mathias Risse, D. Góra, D. Heck, Hans Klages, J. Pekala, B. Wilczynska, and H. Wilczynski. Simulation of ultra-high energy photon showers with PRESHOWER. *Nuclear Physics B-proceedings Supplements - NUCL PHYS B-PROC SUPPL*, 151:119–120, 01 2006.

[62] Thomas Erber. High-energy electromagnetic conversion processes in intense magnetic fields. *Rev. Mod. Phys.*, 38:626–659, Oct 1966.

[63] M. Risse, P. Homola, R. Engel, D. Góra, D. Heck, J. Pekala, B. Wilczyńska, and H. Wilczyński. Photon air showers at ultra-high energy and the photonuclear cross-section. *Czechoslovak Journal of Physics*, 56(1):A327–A336, Sep 2006.

[64] Spencer R. Klein. e+ e- pair production from 10-GeV to 10-ZeV. *Radiat. Phys. Chem.*, 75:696–711, 2006.

[65] S. Eidelman, K.G. Hayes, K.A. Olive, et al. Review of Particle Physics. *Physics Letters B*, 592:1+, 2004.

[66] J. R. Cudell, V. V. Ezhela, P. Gauron, K. Kang, Yu. V. Kuyanov, S. B. Lugovsky, B. Nicolescu, and N. P. Tkachenko. Hadronic scattering amplitudes: Medium-energy constraints on asymptotic behavior. *Phys. Rev. D*, 65:074024, Mar 2002.

[67] A. Donnachie and P. V. Landshoff. New data and the hard Pomeron. *Phys. Lett.*, B518:63–71, 2001.

[68] Leonid B. Bezrukov and E. V. Bugaev. Nucleon Shadowing Effects in Photon Nucleus Interaction. (In Russian). *Yad. Fiz.*, 33:1195–1207, 1981. [Sov. J. Nucl. Phys.33,635(1981)].

[69] M. M. Block and F. Halzen. Evidence for the saturation of the froissart bound. *Phys. Rev. D*, 70:091901, Nov 2004.

[70] J. Knapp, D. Heck, S. J. Sciutto, M. T. Dova, and M. Risse. Extensive air shower simulations at the highest energies. *Astropart. Phys.*, 19:77–99, 2003.

[71] H Burkhardt, S R Kelner, and R P Kokoulin. Monte Carlo Generator for Muon Pair Production. Technical Report CERN-SL-2002-016-AP. CLIC-Note-511, CERN, Geneva, May 2002.

[72] Scherini V. Search for ultra-high energy photons at the Pierre Auger Observatory. *EPJ Web of Conferences*, 53, 06 2013.

[73] F. Ashton, J. Fatemi, H. Nejabat, A. Nasri, W. S. Rada, E. Shaat, A. C. Smith, T. R. Stewart, M. G. Thompson, M. W. Treasure, and I. A. Ward. The Lateral Distribution of Muons in Extensive Air Showers at Sea Level. In *International Cosmic Ray Conference*, volume 11, page 400, January 1977.

[74] Alexander Aab et al. Search for photons with energies above 10^{18} eV using the hybrid detector of the Pierre Auger Observatory. *JCAP*, 1704(04):009, 2017.

[75] A. Aab et al. Reconstruction of Events Recorded with the Surface Detector of the Pierre Auger Observatory. *Journal of Instrumentation*, 15, 2020.

[76] S. N. Vernov, G. B. Khristiansen, V. B. Atrashkevich, G. V. Bogoslovskii, V. I. Boitsov, O. V. Vedeneev, G. V. Kulikov, A. P. Lebedev, S. I. Matsenov, and V. I. Melnikov. The new installation at Moscow State University for the study of extensive air showers with energies up to 10^{18} eV. *Akademiia Nauk SSSR Izvestiia Seriia Fizicheskaia*, 44:537–543, March 1980.

[77] Yu. A. Fomin, N. N. Kalmykov, I. S. Karpikov, G. V. Kulikov, M. Yu Kuznetsov, G. I. Rubtsov, V. P. Sulakov, and S. V. Troitsky. Constraints on the flux of $\sim (10^{16} - 10^{17.5})$ eV cosmic photons from the EAS-MSU muon data. *Phys. Rev.*, D95(12):123011, 2017.

[78] W. D. Apel et al. KASCADE-Grande limits on the isotropic diffuse gamma-ray flux between 100 TeV and 1 EeV. *The Astrophysical Journal*, 848(1):1, oct 2017.

[79] Mikhail Kuznetsov, Oleg Kalashev, and Grigory Rubtsov. Telescope Array Search for EeV Photons. *PoS*, ICRC2019:326, 2020.

[80] J. Abraham et al. Upper limit on the cosmic-ray photon flux above 10^{19} eV using the surface detector of the Pierre Auger Observatory. *Astropart. Phys.*, 29:243–256, 2008.

[81] Alexander Aab et al. A search for point sources of EeV photons. *Astrophys. J.*, 789(2):160, 2014.

[82] Alexander Aab et al. A targeted search for point sources of EeV photons with the Pierre Auger Observatory. *Astrophys. J.*, 837(2):L25, 2017.

[83] G. B. Gelmini, O. E. Kalashev, and D. V. Semikoz. GZK photons as ultra-high-energy cosmic rays. *Soviet Journal of Experimental and Theoretical Physics*, 106(6):1061–1082, Jun 2008.

[84] John Ellis, V. E. Mayes, and D. V. Nanopoulos. Ultrahigh-energy cosmic rays particle spectra from crypton decays. *Physical Review D*, 74(11):115003, Dec 2006.

[85] Biswajit Sarkar, Karl-Heinz Kampert, and Joerg Kulbartz. Ultra-High Energy Photon and Neutrino Fluxes in Realistic Astrophysical Scenarios. In *32nd International Cosmic Ray Conference*, volume 2, page 198, 2011.

[86] T. Abu-Zayyad et al. Upper limit on the flux of photons with energies above 10^{19} eV using the Telescope Array surface detector. *Phys. Rev. D*, 88(11):112005, 2013.

[87] D. Heck. The air shower simulation program CORSIKA. In *Simulation and analysis methods for large neutrino telescopes. Proceedings, Workshop, Zeuthen, Germany, July 6-9, 1998*, pages 228–237, 1998.

[88] D. Heck, J. N. Capdevielle, G. Schatz, T. Thouw, et al. CORSIKA: A Monte Carlo Code to Simulate Extensive Air Showers, Technical Report FZKA 6019, Forschungszentrum Karlsruhe, 1998.

[89] T. Pierog, Iu. Karpenko, J. M. Katzy, E. Yatsenko, and K. Werner. EPOS LHC: Test of collective hadronization with data measured at the CERN Large Hadron Collider. *Phys. Rev.*, C92(3):034906, 2015.

[90] Sergey Ostapchenko. Monte Carlo treatment of hadronic interactions in enhanced Pomeron scheme: I. QGSJET-II model. *Phys. Rev.*, D83:014018, 2011.

[91] Vladimir Gribov. *Frontmatter*, pages i–iv. Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology. Cambridge University Press, 2008.

[92] H. J. Drescher, M. Hladik, S. Ostapchenko, T. Pierog, and K. Werner. Parton based Gribov-Regge theory. *Phys. Rept.*, 350:93–289, 2001.

[93] Tanguy Pierog. Air Shower Simulation with a New Generation of post-LHC Hadronic Interaction Models in CORSIKA. *PoS*, ICRC2017:1100, 2018.

[94] F. Schmidt. Corsika shower images. `https://www-zeuthen.desy.de/~jknapp/fs/showerimages.html`. 2020 (accessed January 8th, 2020).

[95] Piotr Homola, D. Gora, D. Heck, H. Klages, J. Pekala, M. Risse, B. Wilczynska, and H. Wilczynski. Simulation of ultrahigh energy photon propagation in the geomagnetic field. *Comput. Phys. Commun.*, 173:71–90, 2005.

[96] Piotr Homola, Ralph Engel, Agata Pysz, and Henryk Wilczyński. Simulation of ultra-high energy photon propagation with PRESHOWER 2.0. *Comput. Phys. Commun.*, 184:1468–1475, 2013.

[97] A.M. Hillas. Shower simulation: lessons from MOCCA. *Nuclear Physics B - Proceedings Supplements*, 52(3):29 – 42, 1997.

[98] Pierre Billoir. A sampling procedure to regenerate particles in a ground detector from a "thinned" air shower simulation output. *Astroparticle Physics*, 30(5):270 – 285, 2008.

[99] Alexander Aab et al. The Pierre Auger Cosmic Ray Observatory. *Nucl. Instrum. Meth.*, A798:172–213, 2015.

[100] J. Abraham et al. The Fluorescence Detector of the Pierre Auger Observatory. *Nucl. Instrum. Meth.*, A620:227–251, 2010.

[101] I. Allekotte et al. The Surface Detector System of the Pierre Auger Observatory. *Nucl. Instrum. Meth.*, A586:409–420, 2008.

[102] J. Abraham et al. Trigger and Aperture of the Surface Detector Array of the Pierre Auger Observatory. *Nucl. Instrum. Meth.*, A613:29–39, 2010.

[103] A. Filevich, P. Bauleo, H. Bianchi, J. Rodríguez Martino, and G. Torlasco. Spectral-directional reflectivity of tyvek immersed in water. *Nucl. Instrum. Meth*, A423(1):108 – 118, 1999.

[104] R. Smida. Observation of Ultra-high Energy Cosmic Rays. In *29th International Conference on Physics in Collision (PIC 2009) Kobe, Japan, August 30-September 2, 2009*, 2010.

[105] Xavier Bertou et al. Calibration of the surface array of the Pierre Auger Observatory. *Nucl. Instrum. Meth.*, A568:839–846, 2006.

[106] A. Etchegoyen, P. Bauleo, Xavier Bertou, C. B. Bonifazi, A. Filevich, Maria Clementina Medina, Diego G. Melo, Adrian C. Rovero, A. D. Supanitsky, and A. Tamashiro. Muon-track studies in a water Cherenkov detector. *Nucl. Instrum. Meth.*, A545:602–612, 2005.

[107] Alan Coleman. The New Trigger Settings. *Auger internal note*, GAP2018–001, 2018.

[108] Pierre Billoir. Proposition to improve the local trigger of the Surface Detector for low energy showers. *Auger internal note*, GAP2009–179, 2009.

[109] Pierre Billoir. New proposal to improve the local trigger of the Surface Detector. *Auger internal note*, GAP2011–089, 2011.

[110] C. Bonifazi and Pierre Auger Collaboration. The angular resolution of the Pierre Auger Observatory. *Nuclear Physics B Proceedings Supplements*, 190:20–25, May 2009.

[111] Koichi Kamata and Jun Nishimura. The Lateral and the Angular Structure Functions of Electron Showers. *Progress of Theoretical Physics Supplement*, 6:93–155, 02 1958.

[112] K. Greisen. Cosmic ray showers. *Ann. Rev. Nucl. Part. Sci.*, 10:63–108, 1960.

[113] David Newton, J. Knapp, and A. A. Watson. The Optimum Distance at which to Determine the Size of a Giant Air Shower. *Astropart. Phys.*, 26:414–419, 2007.

[114] A. Aab et al. Measurement of the cosmic-ray energy spectrum above 2.5×10^{18} eV using the Pierre Auger Observatory. *Phys. Rev. D*, 102:062005, Sep 2020.

[115] J. Hersil, I. Escobar, D. Scott, G. Clark, and S. Olbert. Observations of extensive air showers near the maximum of their longitudinal development. *Phys. Rev. Lett.*, 6:22–23, Jan 1961.

[116] Alexander Schulz. The measurement of the energy spectrum of cosmic rays above 3×10^{17} eV with the Pierre Auger Observatory. In *Proceedings, 33rd International Cosmic Ray Conference (ICRC2013): Rio de Janeiro, Brazil, July 2-9, 2013*, page 0769, 2013.

[117] Pierre Billoir, Cecile Roucelle, and Jean-Christophe Hamilton. Evaluation of the Primary Energy of UHE Photon-induced Atmospheric Showers from Ground Array Measurements. *arXiv e-prints*, pages astro–ph/0701583, January 2007.

[118] Aaron S. Chou, Katsushi Arisaka, Maximo David Ave Pernas, David Barnhill, Pierre Billoir, Arun Tripathi, and Tokonatsu Yamamoto. An universal description of the particle flux distributions in extended air showers. In *Proceedings, 29th International Cosmic Ray Conference (ICRC 2005): Pune, India, August 3-11, 2005*, 2005.

[119] Frank Nerling, J. Bluemer, R. Engel, and M. Risse. Universality of electron distributions in high-energy air showers: Description of Cherenkov light production. *Astropart. Phys.*, 24:421–437, 2006.

[120] M. Giller, H. Stojek, and G. Wieczorek. Extensive air shower characteristics as functions of shower age. *Int. J. Mod. Phys.*, A20:6821–6824, 2005.

[121] Dariusz Gora, R. Engel, D. Heck, P. Homola, H. Klages, J. Pekala, M. Risse, B. Wilczynska, and H. Wilczynski. Universal lateral distribution of energy deposit in air showers and its application to shower reconstruction. *Astropart. Phys.*, 24:484–494, 2006.

[122] Germán Ros, Gustavo A. Medina-Tanco, Daniel Supanitsky, Luis del Peral, and María Dolores Rodríguez-Frías. Reconstruction of the primary energy of UHE photon-induced showers from Auger SD. *Auger internal note*, GAP2010–047, 2010.

[123] Nicole Krohm. Search for Ultra-High Energy Photons with the Surface Detector of the Pierre Auger Observatory. PhD thesis, Bergische Universität Wuppertal, 2017, http://elpub.bib.uni-wuppertal.de/servlets/DocumentServlet?id=7390.

[124] S. Argiro, S. L. C. Barroso, J. Gonzalez, L. Nellen, Thomas Cantzon Paul, T. A. Porter, L. Prado, Jr., M. Roth, R. Ulrich, and D. Veberic. The Offline Software Framework of the Pierre Auger Observatory. *Nucl. Instrum. Meth.*, A580:1485–1496, 2007.

[125] Javier G. Gonzalez. The Offline Software of the Pierre Auger Observatory: Lessons Learned. *arXiv e-prints*, page arXiv:1208.2154, Aug 2012.

[126] R. Brun and F. Rademakers. ROOT: An object oriented data analysis framework. *Nucl. Instrum. Meth.*, A389:81–86, 1997. https://root.cern.ch.

[127] Nicole Krohm, Carla Bleve, Karl-Heinz Kampert, and Julian Rautenberg. Looking for an LDF Parameter for the SD Photon Search. *Auger internal note*, GAP2012–059, 2012.

[128] A A Watson and J G Wilson. Fluctuation studies of large air showers: the composition of primary cosmic ray particles of energy $E_p \sim 10^{18}$ eV. *Journal of Physics A: Mathematical, Nuclear and General*, 7(10):1199–1212, jul 1974.

[129] Heather Louise Cook. Limit to the Ultra-High Energy Gamma-Ray Flux Using Data from the Surface Detectors of the Pierre Auger Observatory. Master thesis, University of Leeds, 2012, http://etheses.whiterose.ac.uk/id/eprint/3863.

[130] Patricia Sánchez Lucas. The $\langle\Delta\rangle$ Method: An Estimator for the Mass Composition of Ultra-High-Energy Cosmic Rays. PhD thesis, Universidad de Granada, 2016, https://hera.ugr.es/tesisugr/26357422.pdf.

[131] Christopher Wileman. The Spread in the Arrival Times of Particles in Air-Showers for Photon and Anisotropy Searches above 10 EeV. PhD thesis, University of Leeds, 2008, Auger internal note, GAP2008–160.

[132] Benjamin Edward Smith. The Mass Composition of Cosmic Rays Above 1 EeV Inferred Using the Spread in Arrival Times of Air Shower Particles. PhD thesis, University of Leeds, 2008, Auger internal note, GAP2008–161.

[133] Lu Lu. A Search for Photons of Energy above 6×10^{18} eV Using Data from the Water-Cherenkov Detectors of the Pierre Auger Observatory. PhD thesis, University of Leeds, 2014, http://etheses.whiterose.ac.uk/7354/.

[134] David Scott Barnhill. Composition Analysis of Ultrahigh Energy Cosmic Rays Using the Pierre Auger Observatory Surface Detector. PhD thesis, University of California, 2005, Auger internal note, GAP2005–082.

[135] Cecilia Jarne, Hernán Wahlberg, and María Teresa Dova. Risetime at 1000 m revisited, 2013. Auger internal note, GAP2013–010.

[136] G. Ros, A. D. Supanitsky, G. A. Medina-Tanco, L. del Peral, J. C. D'Olivo, M. D. Rodriguez-Frias, A. D. Supanitsky, G. A. Medina-Tanco, J. C. D'Olivo, and M. D. Rodriguez Frias. A new composition-sensitive parameter for Ultra-High Energy Cosmic Rays. *Astropart. Phys.*, 35:140–151, 2011.

[137] Germán Ros, Gustavo A. Medina-Tanco, Daniel Supanitsky, Luis del Peral, and María Dolores Rodríguez-Frías. S_b for photon-hadron discrimination, 2010. Auger internal note, GAP2010–052.

[138] M. Settimo, P. Billoir, Mariş, and L. Molina Bueno. Trigger probability for single stations and air-showers with the TOTd and MoPS algorithms, 2013. Auger internal note, GAP2013–114.

[139] The Telescope Array Collaboration and The Pierre Auger Collaboration. Pierre Auger Observatory and Telescope Array: Joint Contributions to the 35th International Cosmic Ray Conference (ICRC 2017). 2018.

[140] Isabelle Lhenry-Yvon and Philipp Papenbreer. Bad PMTs : the third levels of quality cuts, tuned for PMT traces analysis., 2018. Auger internal note, GAP2018–029.

[141] Gary J. Feldman and Robert D. Cousins. A Unified approach to the classical statistical analysis of small signals. *Phys. Rev. D*, 57:3873–3889, 1998.

[142] Carla Bleve. Updates on the neutrino and photon limits from the Pierre Auger Observatory. In *Proceedings, 34th International Cosmic Ray Conference (ICRC 2015): The Hague, The Netherlands, July 30-August 6, 2015*, page 1103, 08 2016.

[143] A. Aab et al. Depth of maximum of air-shower profiles at the Pierre Auger Observatory. II. Composition implications. *Phys. Rev. D*, 90(12):122006, 2014.

[144] Marco Stein Muzio, Michael Unger, and Glennys R. Farrar. Progress towards characterizing ultrahigh energy cosmic ray sources. *Physical Review D*, 100(10), Nov 2019.

[145] Diego Casadei. Estimating the selection efficiency. *JINST*, 7:P08021, 2012.

[146] Benno List. Statistical error of efficiency determination from weighted events, 1999. updated 2003, `www.desy.de/~blist/notes/effic.ps.gz`.

[147] Philipp Papenbreer. Afterpulse structure in single muon PMT traces simulated at 120 MHz, 2015. Auger internal note, GAP2015–010.

[148] Antonella Castellina. Highlights from the Pierre Auger Observatory (ICRC2019). *PoS*, ICRC2019:004, 2020.

[149] Olivier Deligny. Hunting super-heavy dark matter with ultra-high energy photons. In *Snowmass 2021, Primary frontiers:CF1, CF7(2020)*, 2020.

[150] Alexander Aab et al. Probing the origin of ultra-high-energy cosmic rays with neutrinos in the EeV energy range using the Pierre Auger Observatory. *Journal of Cosmology and Astroparticle Physics*, 2019(10):022–022, Oct 2019.

[151] Philip Ruehl, Marcus Niechciol, and Markus Risse. A follow-up search for UHE photons after gravitational wave events observed during the first two observation runs of the Advanced LIGO and Virgo detectors, 2019. Auger internal note, GAP2019–043.

[152] B. P. Abbott et al. GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs. *Phys. Rev.*, X9(3):031040, 2019.

Appendix A

CORSIKA Simulation Settings

For the photon air shower simulations used in chapter 6, the simulation program CORSIKA has been used. An example of a steering card used in the simulations is shown below.

```
DATBAS T
DIRECT 'data/'
ECTMAP 250000
ECUTS 0.1 0.05 0.00025 0.00025
ELMFLG T T
ERANGE 3162277660.16838 316227766016.838
ESLOPE -1
EVTNR 1
GCOORD -69.585 -35.463 2013 1 0
HOST cluster
LONGI T 5 T T
MAGNET 19.812 -14.3187
MAXPRT 1
MUADDI T
MUMULT T
NSHOW 1
OBSLEV 145200
PAROUT T T
PHIP -180 180
PRMPAR 1
RADNKG 500000
RUNNR 1
SEED 997 0 0
STEPFC 1
THETAP 0 70
THIN 1e-06 3162.27766016838 30000
THINH 1 100
USER kuempel
EPOS T 0
EPOPAR input ./epos/epos.param
EPOPAR fname inics ./epos/epos.inics
EPOPAR fname iniev ./epos/epos.iniev
EPOPAR fname initl ./epos/epos.initl
EPOPAR fname inirj ./epos/epos.inirj
EPOPAR fname inihy ./epos/epos.ini1b
EPOPAR fname check none
EPOPAR fname histo none
EPOPAR fname data none
EPOPAR fname copy none
SEED 998 0 0
SEED 999 0 0
EXIT
```

Appendix B

Offline module sequences

The simulation and reconstruction procedures in Offline are controlled by *module sequence* XML-files. The sequences used in the photon analysis of this work are shown below.

B.1 Module Sequence for Photon Simulations

```
<!-- A sequence for an SD only reconstruction -->

<sequenceFile
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:noNamespaceSchemaLocation='/lustre/gridsoft/auger/Offline/rev30478/share/auger-offline
  /config/ModuleSequence.xsd'>

  <enableTiming/>

  <moduleControl>

    <loop numTimes="1" pushEventToStack="yes">

      <module> EventFileReaderOG </module>

      <!-- increase numTimes if you want to throw the shower
           into the array more than once -->
      <loop numTimes="5" pushEventToStack="yes">

        <module> EventGeneratorOG </module>

        <!-- simulation of muon background -->
        <module> SdAccidentalInjectorKG </module>
        <module> TabulatedTankSimulatorKG </module>

        <!-- simulate shower particles in batches
             to limit memory consumption          -->
        <loop numTimes="unbounded" pushEventToStack="no">
          <module> CachedShowerRegeneratorOG </module>
          <module> G4StationSimulatorOG          </module>
        </loop>
        <module> ClearParticleLists              </module>
        <module> SdSimulationCalibrationFillerOG </module>
        <module> SdPMTSimulatorOG                </module>
        <module> SdFilterFADCSimulatorMTU        </module>
        <module> SdBaselineSimulatorOG           </module>
        <module> TankTriggerSimulatorOG          </module>
        <module> TankGPSSimulatorOG              </module>
        <module> CentralTriggerSimulatorXb       </module>
        <module> CentralTriggerEventBuilderOG    </module>
```

```
        <module> EventBuilderOG                  </module>
        <module> EventCheckerOG                  </module>
        <module> SdPMTQualityCheckerKG           </module>
        <module> SdCalibratorOG                  </module>
        <module> SdSignalRecoveryKLT             </module>
        <module> SdMonteCarloEventSelectorOG     </module>
        <module> SdEventSelectorOG               </module>
        <module> SdPlaneFitOG                    </module>
        <module> LDFFinderKG                     </module>
        <module> DLECorrectionWG                 </module>
        <module> SdCompositionParameters         </module>
        <module> SdEventPosteriorSelectorOG      </module>
        <module> Risetime1000LLL                 </module>

        <!-- export in Offline format -->
        <module> EventFileExporterOG             </module>

        <!-- export the ADST -->
        <module> RecDataWriterNG                 </module>

      </loop>

    </loop>

  </moduleControl>

</sequenceFile>
```

B.2 Module Sequence for Data

```
<!-- A sequence for an SD only reconstruction -->

<!DOCTYPE sequenceFile [
  <!ENTITY % sd SYSTEM "/lustre/papenbreer/Offline_svn/install.20180130/share/auger-offline
  /config/standardSdSequences.dtd">
  %sd;
] >

<sequenceFile
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:noNamespaceSchemaLocation='/lustre/papenbreer/Offline_svn//install.20180130/share/auger-offline
  /config/ModuleSequence.xsd'>

  <enableTiming/>
  <moduleControl>

    <loop numTimes="unbounded" pushEventToStack="yes">

      <module> EventFileReaderOG </module>
      <module> EventCheckerOG </module>
      <module> SdPMTQualityCheckerKG </module>
      <module> TriggerTimeCorrection </module>
      <module> SdCalibratorOG </module>
      <module> SdPMTSignalShapeQualityChecker </module>
      <module> SdStationPositionCorrection </module>
      <module> SdBadStationRejectorKG </module>
      <module> SdSignalRecoveryKLT </module>
      <module> SdEventSelectorOG </module>
      <module> SdPlaneFitOG </module>
      <module> LDFFinderKG </module>
      <module> DLECorrectionWG </module>
      <module> SdCompositionParameters </module>
      <module> SdEventPosteriorSelectorOG </module>
      <module> RecDataWriterNG </module>

    </loop>

  </moduleControl>

</sequenceFile>
```

Appendix C

Additional Plots

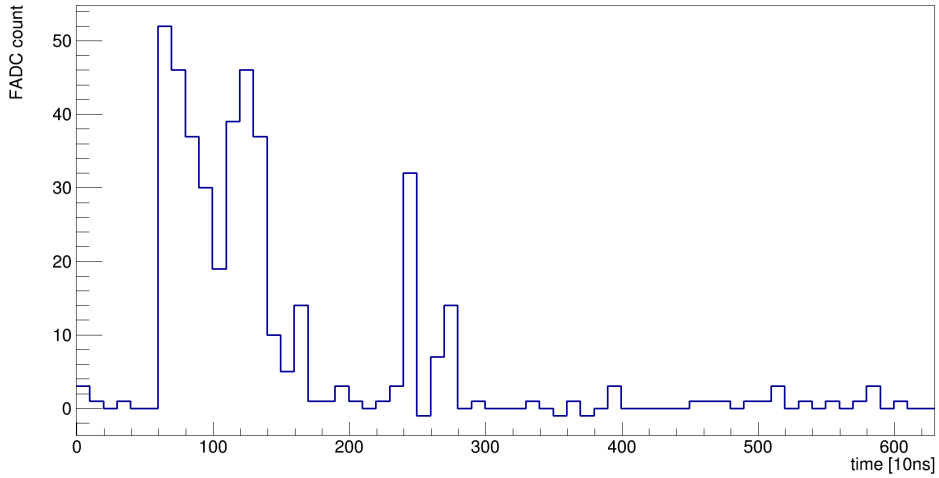


Figure C.1: FADC trace of a single muon measured with 100 MHz sampling rate at the AugerNorth detector station in the Telescope Array. Due to the large sampling rate a single muon causes multiple distinct peaks in the FADC trace.

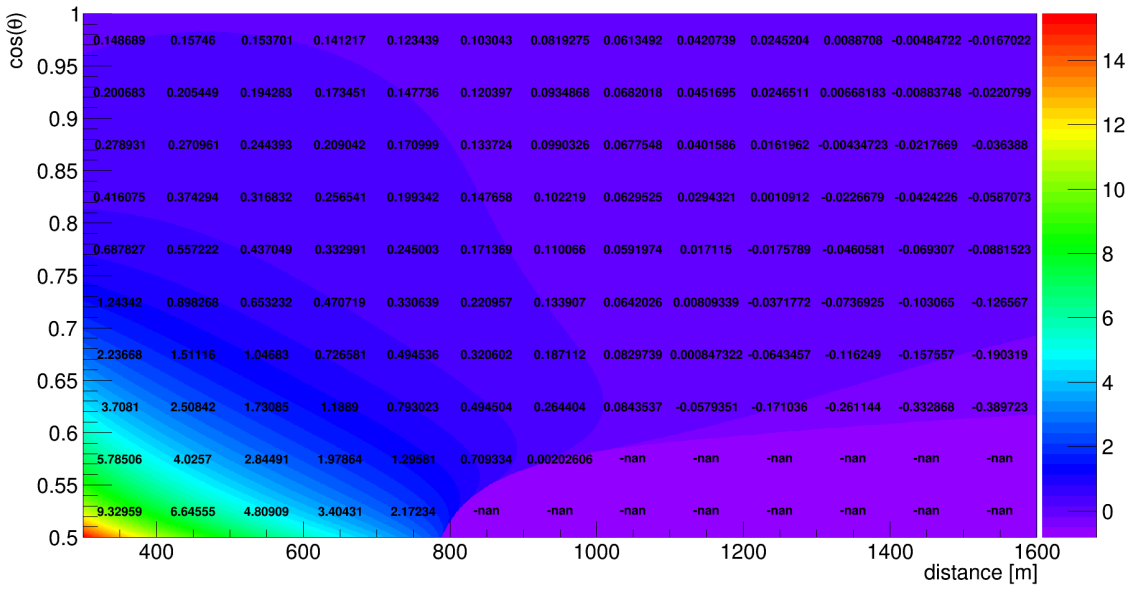


Figure C.2: Differences of the risetime benchmark derived in [123] to the HG benchmark from [130]: in each bin the difference of both benchmarks divided by the benchmark uncertainty from [123] is shown $\left(t^1_{\frac{1}{2}} - t^2_{\frac{1}{2}}\right) / \sigma t^1_{\frac{1}{2}}$. The region below $\cos(\theta) = 0.69$ was not parameterized in [130] and contains a large undefined region. Besides this region, most residuals are well below the 1 σ level showing a good agreement of the two benchmarks in the commonly analyzed parameter space. The agreement is even better when comparing the benchmark from [123] to the LG parametrization from [130] (see C.3).

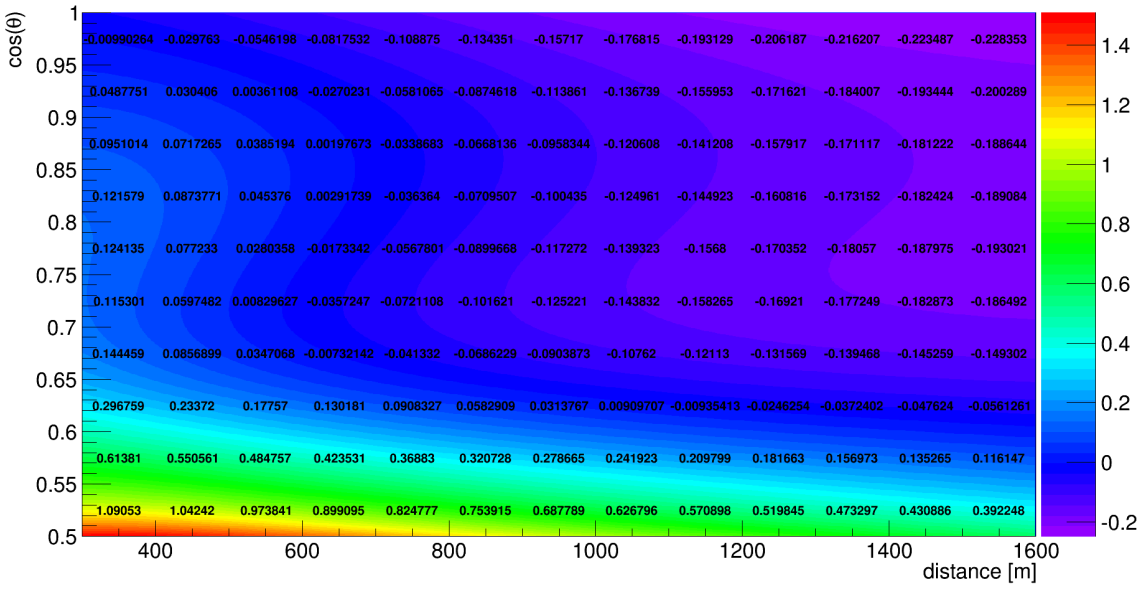


Figure C.3: Differences of the risetime benchmark derived in [123] to the LG benchmark from [130]: in each bin the difference of both benchmarks divided by the benchmark uncertainty from [123] is shown $\left(t^1_{\frac{1}{2}} - t^2_{\frac{1}{2}}\right) / \sigma t^1_{\frac{1}{2}}$.

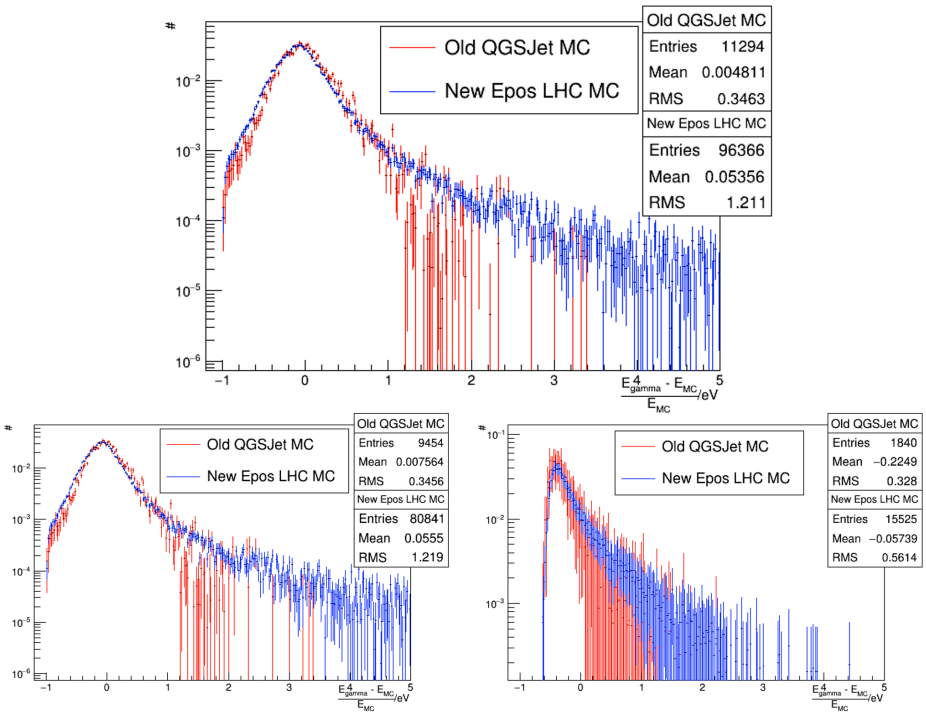


Figure C.4: Top: Relative differences of reconstructed photon energies for EPOS LHC (red) and QGSJetII.03 (blue).
Bottom left: As above using only unconverted MC photons.
Bottom right: As above using only converted MC photons.

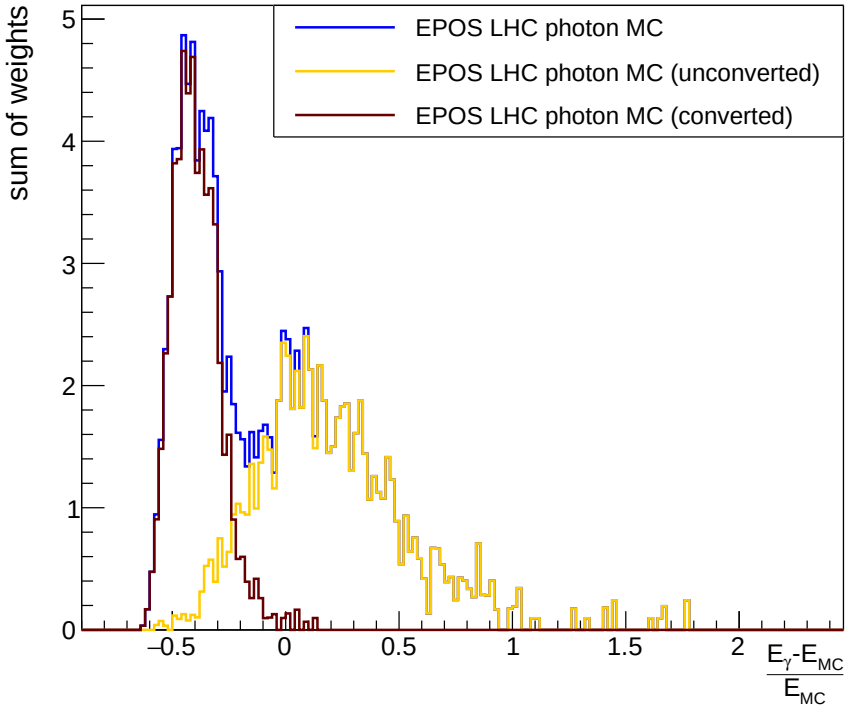


Figure C.5: Photon energy reconstruction accuracy for $E > 100\,\mathrm{EeV}$ and $55° < \theta < 60°$. The energies of converted air showers are mostly underestimated while the energies of unconverted air showers are mostly overestimated.

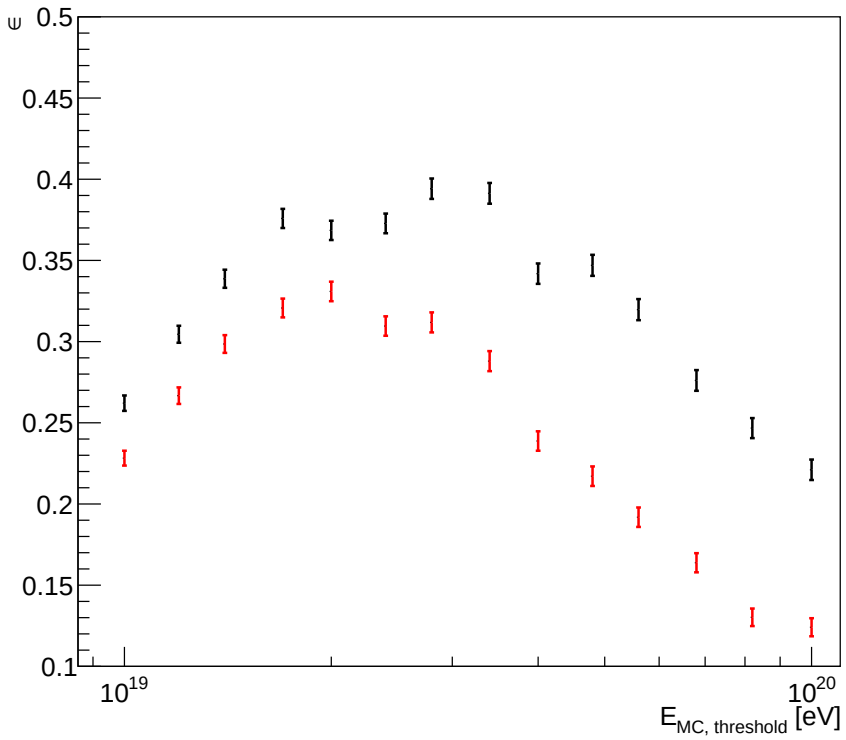


Figure C.6: Efficiency of the photon analysis with the iterative energy reconstruction (red) and the tabular energy reconstruction (black) as a function of the simulated photon energy. The lower efficiencies with the iterative method are caused by the cut on X_{max} not more than $50\,\mathrm{g\,cm^{-2}}$ below ground.

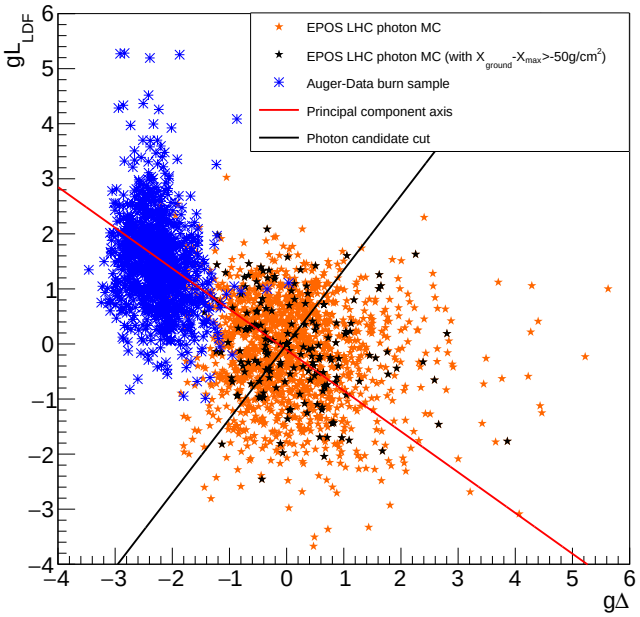
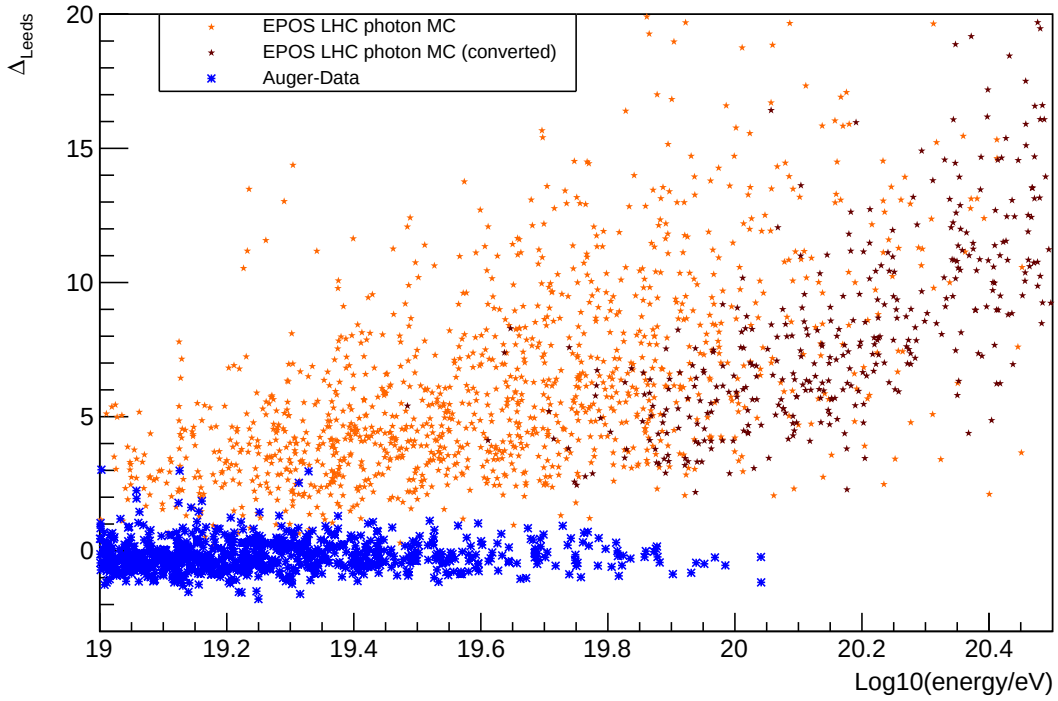
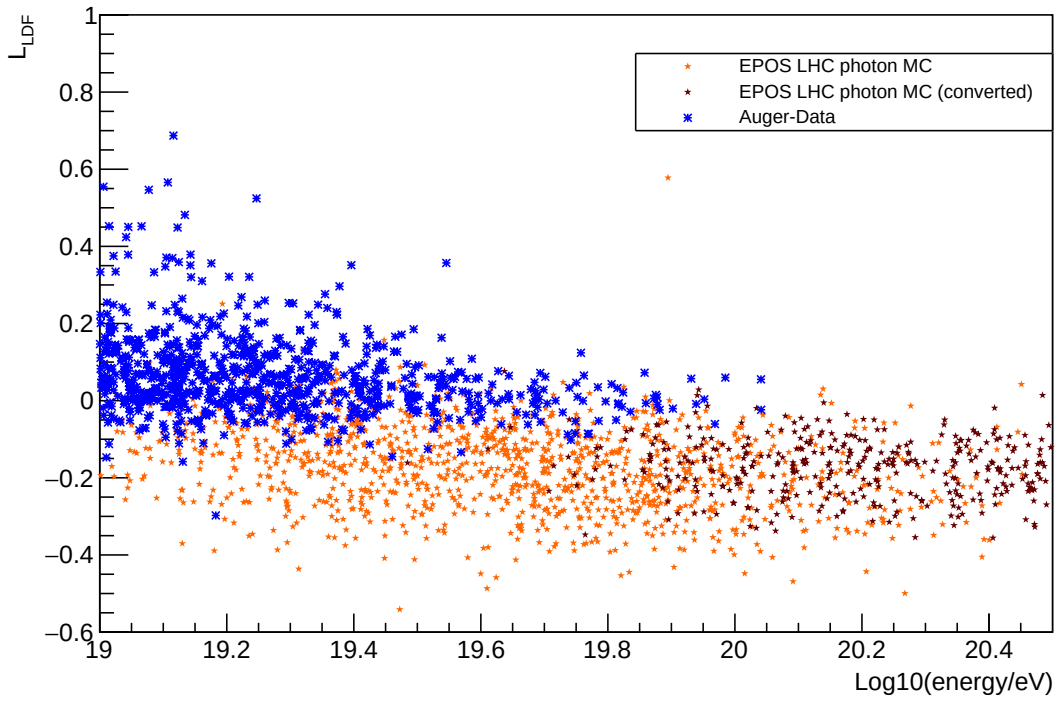


Figure C.7: Distribution of events with $X_{\max}$ more than $50\,\mathrm{g\,cm^{-2}}$ below ground (black) in the PCA parameter space. For this plot the parametrization from Fig. C.11 has been used.

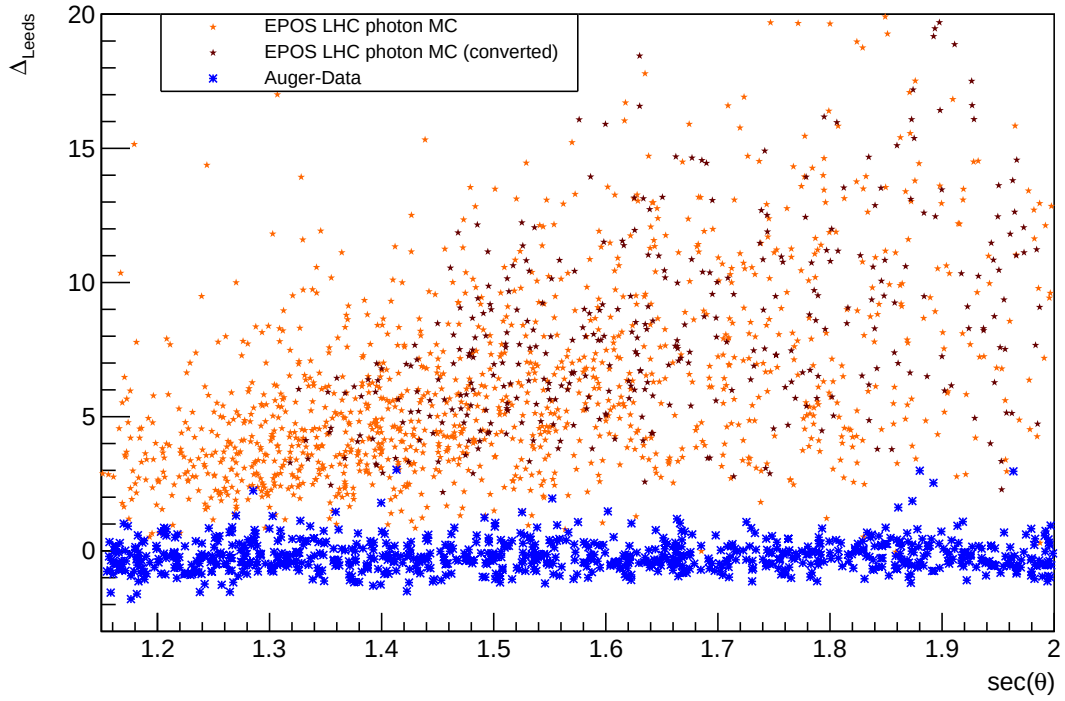


(a) Δ_{Leeds}

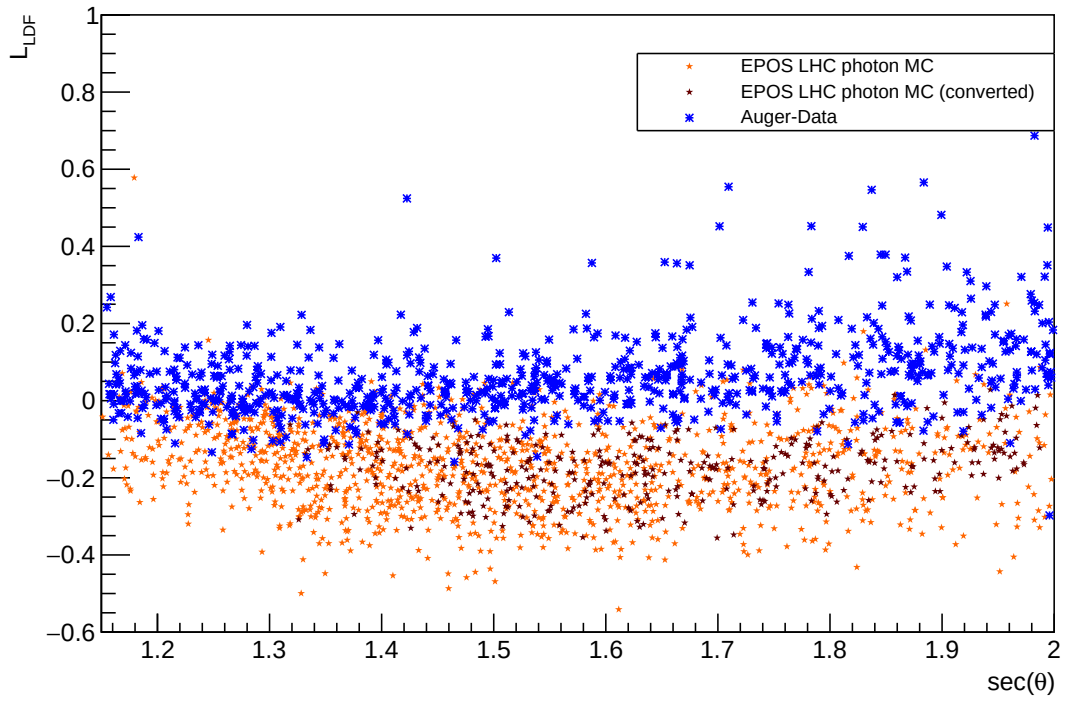


(b) L_{LDF}

Figure C.8: Energy dependent separation of burn sample data and photon MC

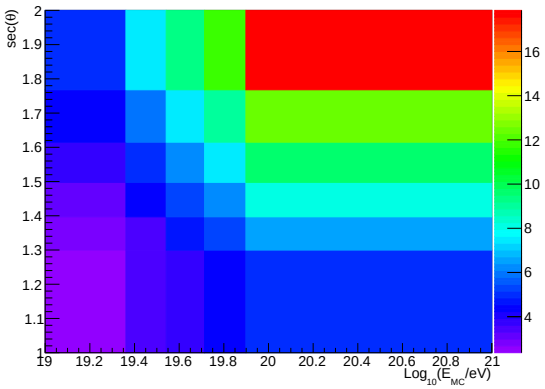


(a) Δ_{Leeds}

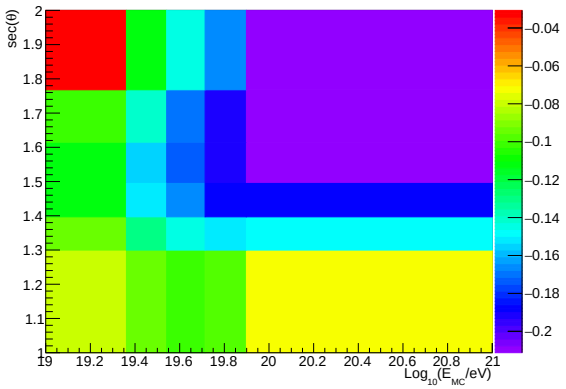


(b) L_{LDF}

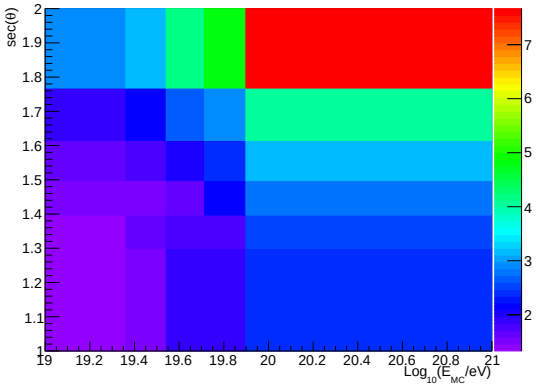
Figure C.9: Zenith angle dependent separation of burn sample data and photon MC



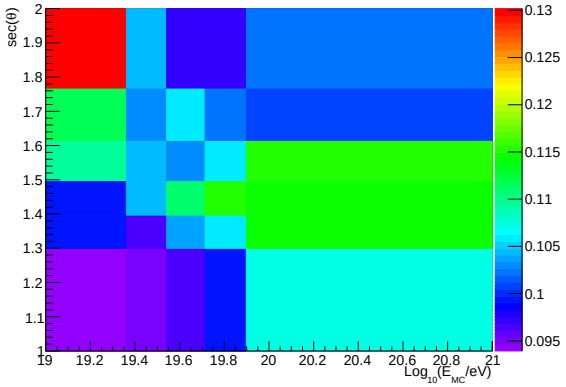
(a) Mean values of Δ_{Leeds}



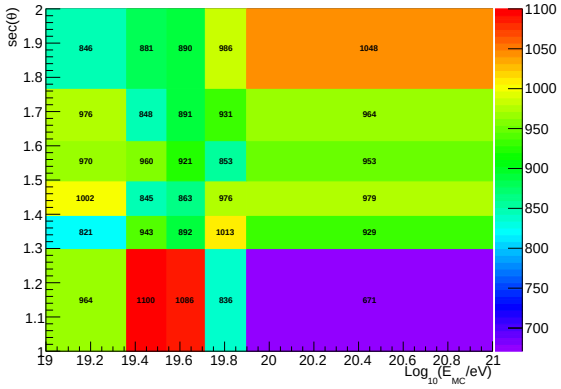
(b) Mean values of $\mathrm{L}_{\mathrm{LDF}}$



(c) σ values of Δ_{Leeds}

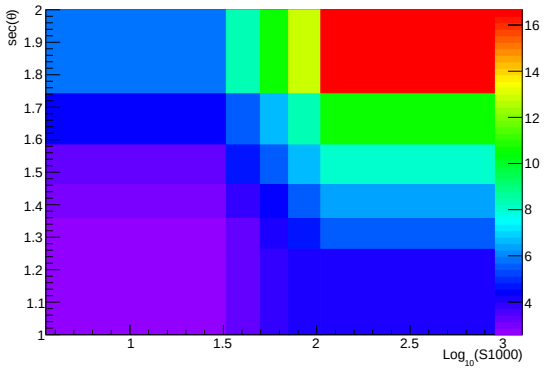


(d) σ values of $\mathrm{L}_{\mathrm{LDF}}$

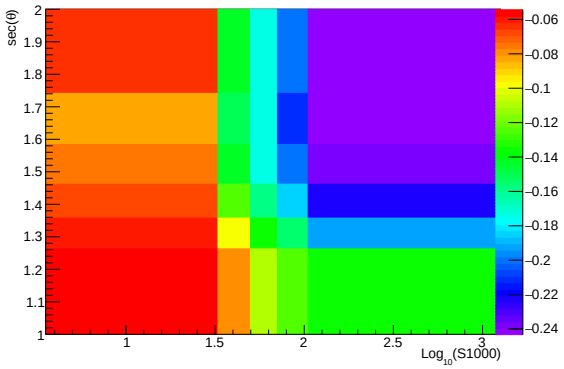


(e) Number of entries per bin in E_{MC} and θ

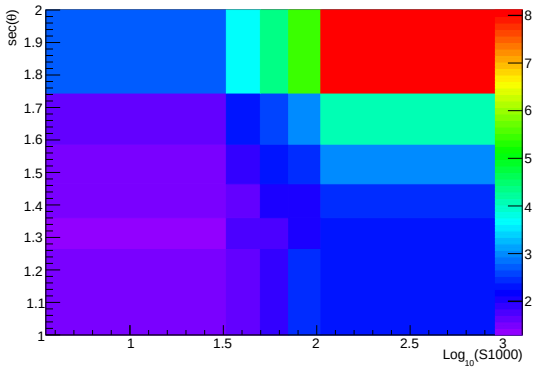
Figure C.10: Parametrization histograms for $\mathrm{g}\Delta_{\mathrm{Leeds}}$ and $\mathrm{gL}_{\mathrm{LDF}}$ as a function of MC energy and $\sec(\theta)$. The binning was chosen for roughly equal statistics in each bin.



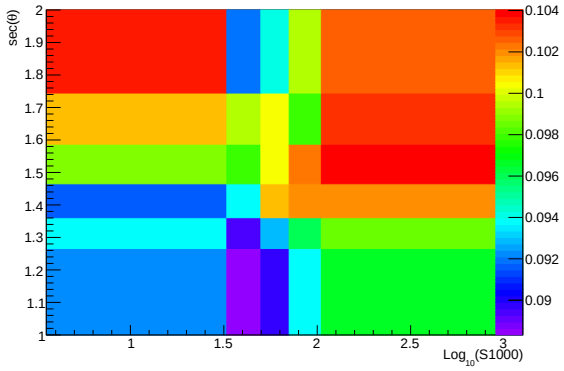
(a) Mean values of Δ_{Leeds}



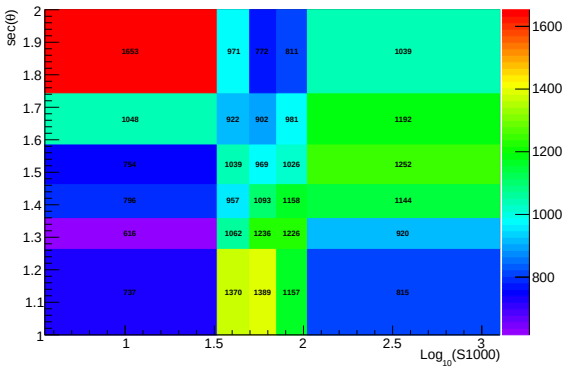
(b) Mean values of $\mathrm{L}_{\mathrm{LDF}}$



(c) σ values of Δ_{Leeds}



(d) σ values of $\mathrm{L}_{\mathrm{LDF}}$



(e) Number of entries per bin in S1000 and θ

Figure C.11: Parametrization histograms for $\mathrm{g}\Delta_{\mathrm{Leeds}}$ and $\mathrm{gL}_{\mathrm{LDF}}$ as a function of S1000 and θ. The binning was chosen for roughly equal statistics in each bin.

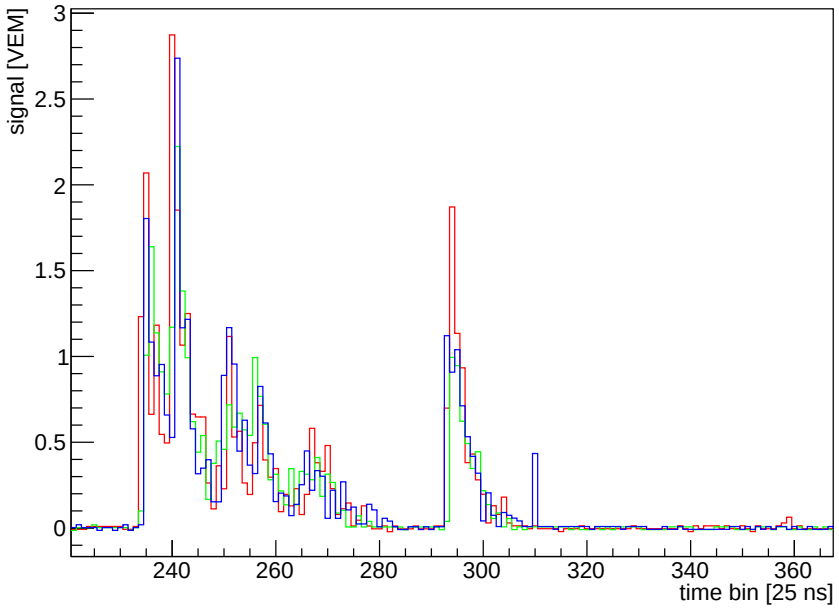


Figure C.12: PMT traces of event 8742491 at station 1681 with a split signal.

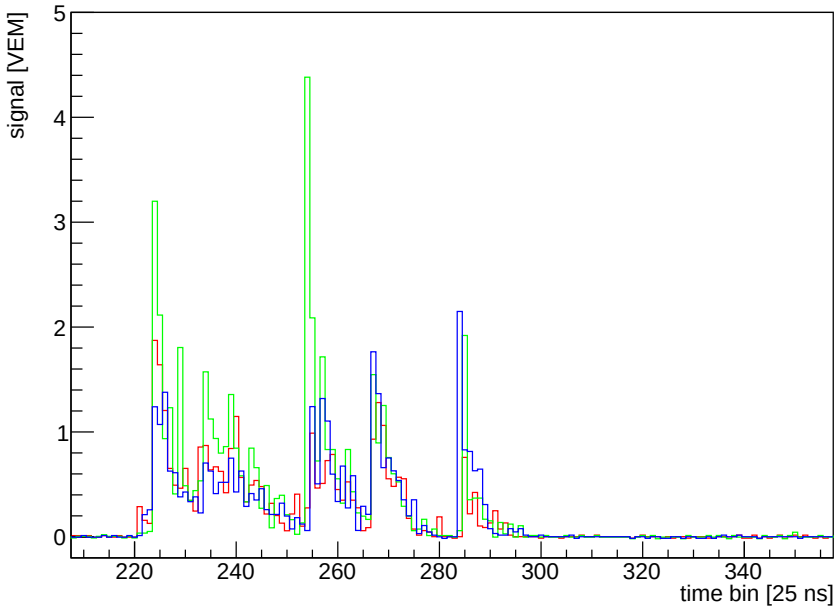


Figure C.13: PMT traces of event 9099370 at station 223 with a large risetime.

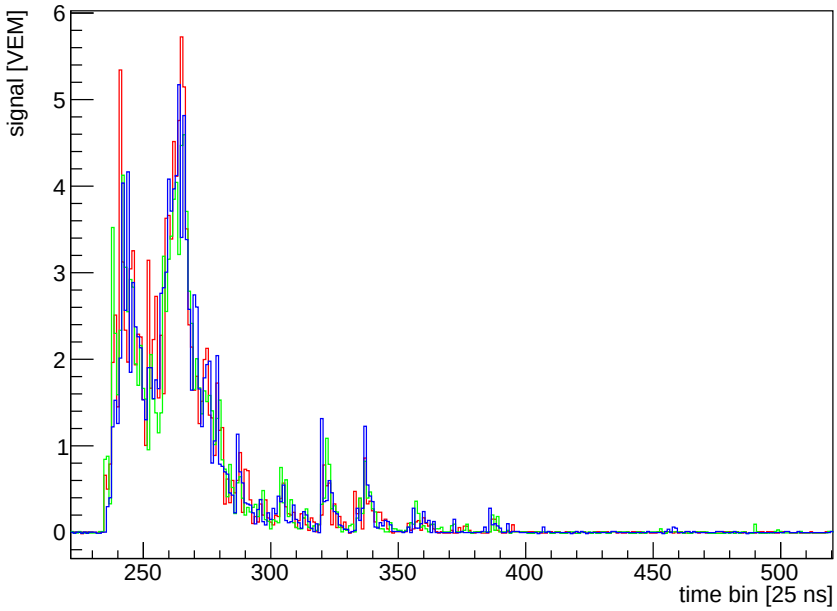


Figure C.14: PMT traces of event 10759292 at station 975 with a large risetime.

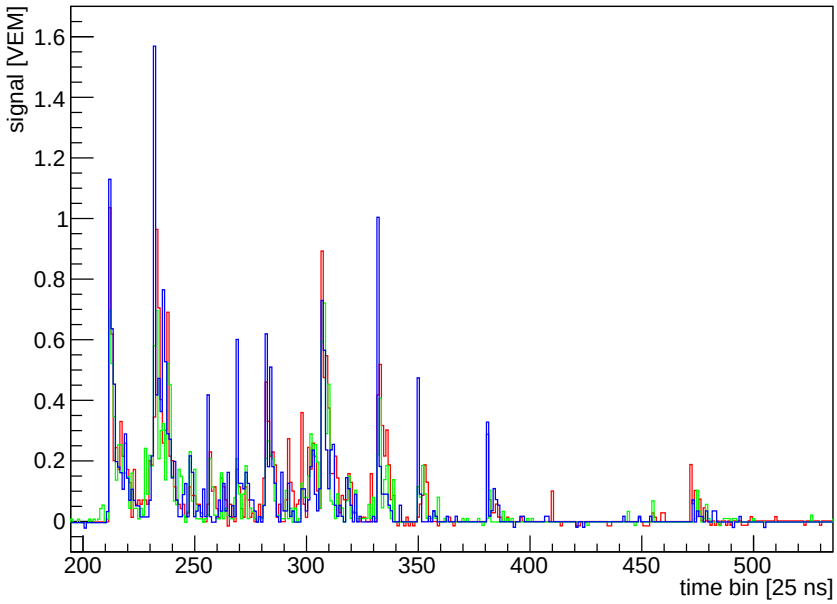


Figure C.15: PMT traces of event 10759292 at station 1486 with a large risetime.

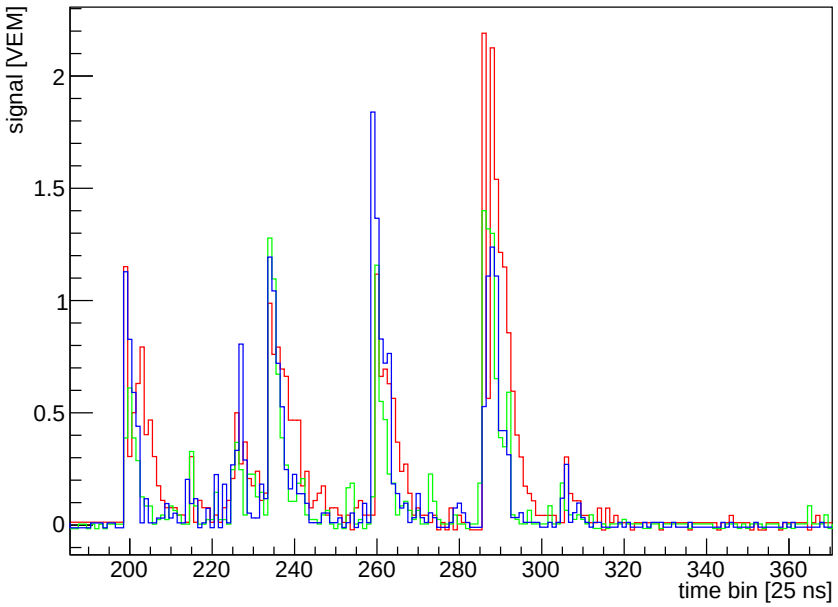


Figure C.16: PMT traces of event 15797618 at station 852 with a large risetime possibly caused by multiple late muons.

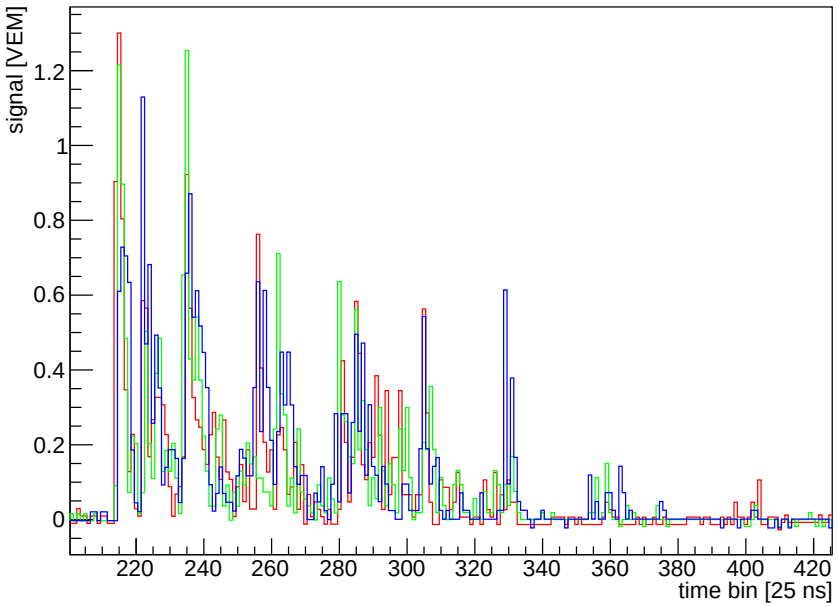


Figure C.17: PMT traces of event 32654533 at station 709 with a large risetime possibly caused by a multitude of low energy particles.

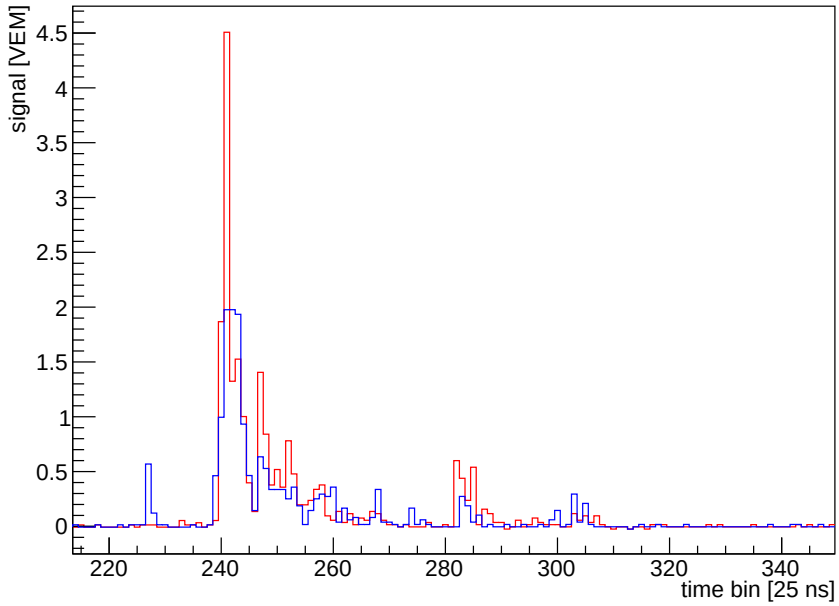


Figure C.18: PMT traces of event 43993111 at station 1216. The signal start time has been determined based on PMT3 (blue) and was set to bin 228.

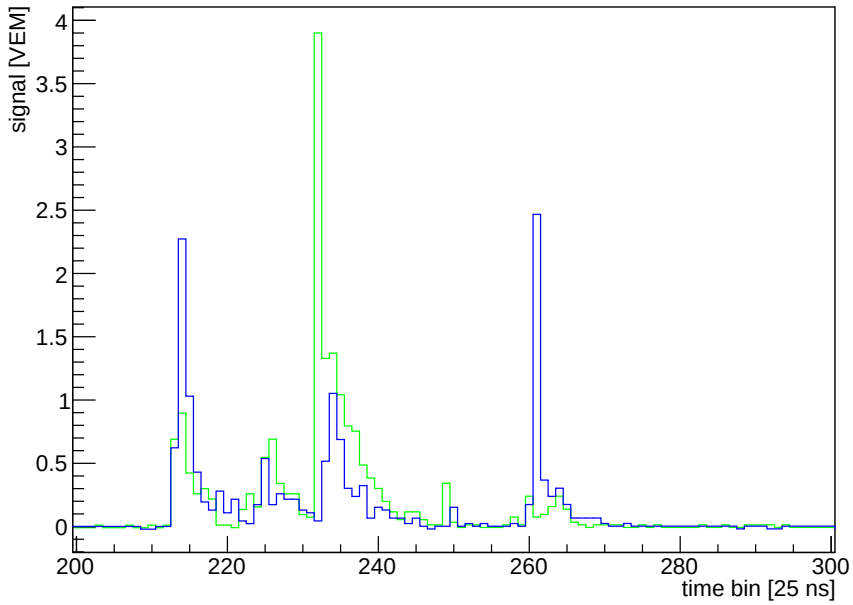


Figure C.19: PMT traces of event 46285053 at station 1545.

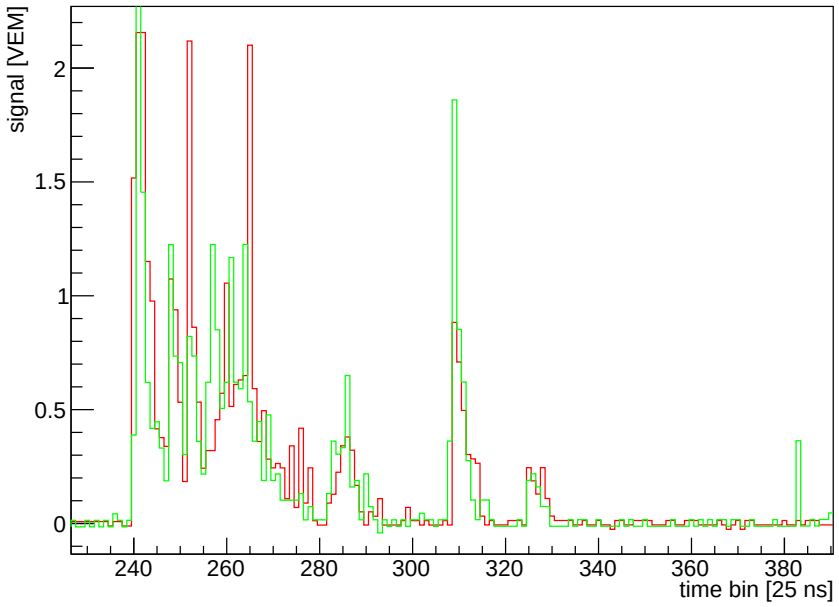


Figure C.20: PMT traces of event 46285053 at station 1546.

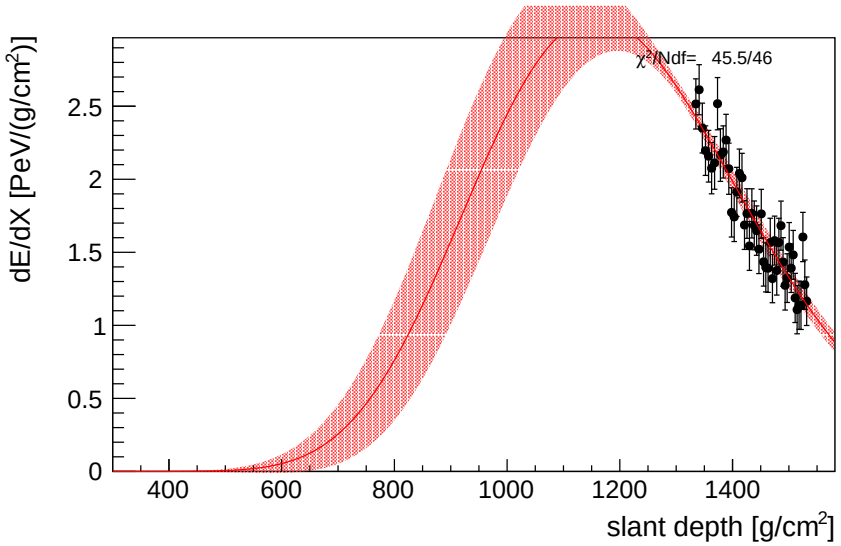


Figure C.21: Profile of event 9701198 measured by the FD telescope Los Morados. The reconstruction of $X_{\max}$ has a large uncertainty since it is not in the field of view of the telescope.

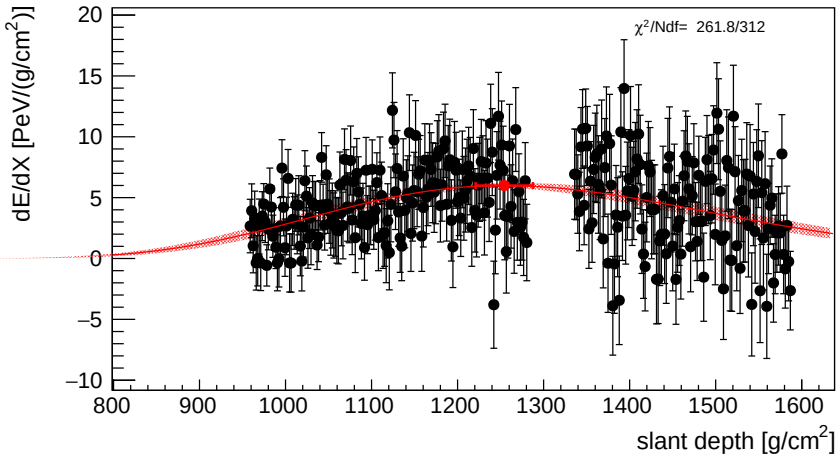


Figure C.22: Profile of event 15797618 measured by the FD telescope Loma Amarilla.

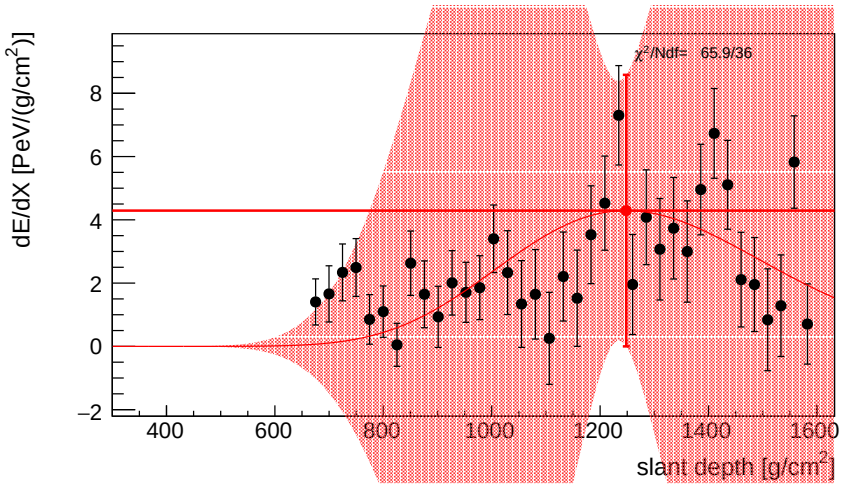


Figure C.23: Profile of event 15797618 measured by the FD telescope Los Morados.

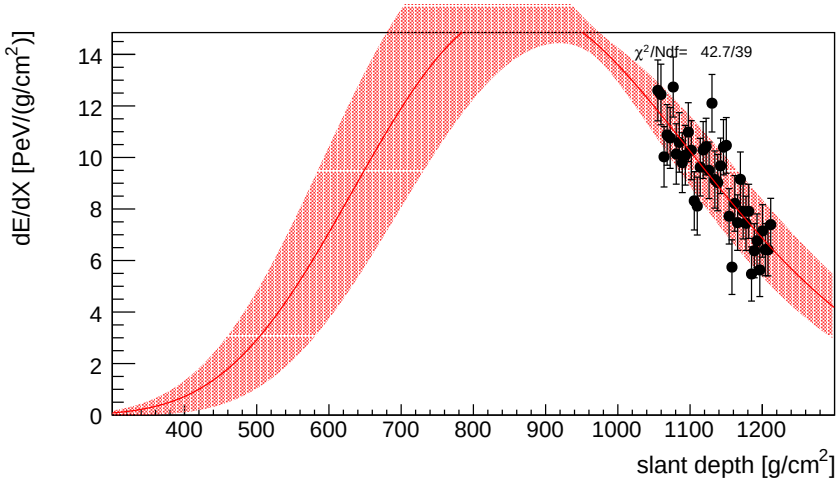


Figure C.24: Profile of event 46285053 measured by the FD telescope Loma Amarilla.

Acknowledgment

This work would not have been possible without the help and support of many people to whom I owe a debt of gratitude.

- My first thanks goes to Prof. Dr. Karl-Heinz Kampert. My initial decision to work in the field of astroparticle physics is based on his excellent lectures and was never regretted. I am grateful that I was allowed to work on such an exciting topic and to him for the mentoring of my thesis.

- Likewise, I want to thank Corinne Bérat for agreeing to be the second reviewer of this work and for all the helpful input for the photon paper.

- I would like to thank Julian Rautenberg for countless hours of discussions on the analyses presented in this work, for his invaluable expertise and feedback. I am very thankful for the proofreading of my thesis and his continuous support.

- My thanks goes to the Editiorial Board of the photon paper including Julian Rautenberg, Corinne Bérat, Carla Bleve, Olivier Deligny, Marcus Niechciol, Alexey Yushkov, Nicole Krohm and Daniel Kümpel. It was a pleasure to have all these fruitful discussions with so many experts on cosmic ray photons.

- I would also like to thank Lorenzo Perrone for the very warm welcome at the Università del Salento and Isabelle Lhenry-Yvon for visiting me there and introducing me to the BadPMT studies.

- In Wuppertal, a big additional thanks goes to Eric Mayotte for additional proofreading and many thoughtful insights pertaining to my work.

- I want to thank my friends Raphael Kleinemühl and Max J. Ried for reading my thesis and the nice time Raphael and I had together as fellow physics students and roommates.

- Finally, I am immensely thankful to my sister and parents who have always been there for me and will always support me. My wholehearted gratitude goes to Tani for her patience and support, especially in the final months of writing this thesis.

Eidesstattliche Erklärung

Hiermit erkläre ich, die vorliegende Arbeit ohne fremde Hilfe und nur unter Verwendung der angegebenen Hilfsmittel und Quellen, selbstständig verfasst zu haben. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten oder nicht veröffentlichten Arbeiten anderer entnommen sind, habe ich kenntlich gemacht.
Die Dissertation hat weder in der gegenwärtigen, noch in einer anderen Fassung, schon einem anderen Fachbereich einer wissenschaftlichen Hochschule vorgelegen.

Wuppertal, den 3.12.2020

_____
Philipp Papenbreer