UNIVERSITY OF NOVA GORICA
GRADUATE SCHOOL

APPLICATION OF MACHINE LEARNING TECHNIQUES FOR COSMIC RAY EVENT CLASSIFICATION AND IMPLEMENTATION OF A REAL-TIME ULTRA-HIGH ENERGY PHOTON SEARCH WITH THE SURFACE DETECTOR OF THE PIERRE AUGER OBSERVATORY

DISSERTATION

Lukas Zehrer

Advisor: prof. dr. Serguei Vorobiov

Nova Gorica, 2021

UNIVERZA V NOVI GORICI
FAKULTETA ZA PODIPLOMSKI ŠTUDIJ

RAZVOJ METOD STROJNEGA UČENJA ZA IDENTIFIKACIJO KOZMIČNIH DELCEV EKSTREMNIH ENERGIJ TER NJIHOVA IMPLEMENTACIJA PRI ISKANJU FOTONOV EKSTREMNIH ENERGIJ S POVRŠINSKIMI DETEKTORJI OBSERVATORIJA PIERRE AUGER

DISERTACIJA

Lukas Zehrer

Svetovalec: prof. dr. Serguei Vorobiov

Nova Gorica, 2021

UNIVERSITY OF NOVA GORICA
GRADUATE SCHOOL

Lukas Zehrer, *Application of machine learning techniques for cosmic ray event classification and implementation of a real-time ultra-high energy photon search with the surface detector of the Pierre Auger Observatory,* Dissertation, (2021)

Meinen tiefsten Dank und meine Wertschätzung für ihre Ermutigungen und für all die bedingungslose Unterstützung in diesen sehr intensiven Studienjahren, sowie während meines gesamten Lebens, möchte ich meinen Eltern Ingrid und Heinrich und meinem Bruder Valentin aussprechen. Ohne euch hätte ich das nicht geschafft, besonderen Dank!

Povzetek

Kozmični delci so bili odkriti pred več kot desetletjem, vendar do danes še vedno ne poznamo vseh njihovih lastnosti. Pri prehodu skozi Zemljino atmosfero sprožijo kozmični delci obširne sekundarne plazove atmosferskih delcev, ki jih lahko zaznamo z zbirko detektorjev na površju Zemlje. Eden izmed večjih observatorijev za opazovanje kozmičnih delcev je Observatorij Pierre Auger, ki omogoča tudi detekcijo kozmičnih delcev ekstremnih energij.

Izvor redkih kozmičnih delcev z energijami nad 10^{18} eV (UHECR) trenutno še ni poznan, pogosto opazovanje UHECR fotonov pa bi močno pripomoglo k našemu razumevanju kozmičnih delcev. Hkrati bi detekcije takšnih dogodkov omogočile nov način opazovanja vesolja ter morebiti razkrile fizikalne pojave onkraj standardnega modela.

V tem delu predstavimo razvoj orodja za posredovanje podatkov o možnih detekcijah UHECR dogodkov med observatorijem Pierre Auger in globalnim omrežjem observatorijev AMON (Astrophysical Multimessenger Observatory Network) v realnem času. Orodje klasificira dogodke detektirane v ustreznem območju kot fotone ekstremnih energij z uporabo metod strojnega učenja in kombinacijo opazovalnih parametrov pridobljenih z observatorijem Pierre Auger ter multivariabilno analizo.

Takšen način klasifikacije omogoča neposredno vključitev podatkov observatorija Pierre Auger v svetovne študije na področju večglasniške astrofizike. S hitrim poročanjem o možni detekciji UHECR delcev omrežju AMON pridobijo observatoriji znotraj omrežja možnost, da lahko najbolj zanimivim dogodkom z dodatnimi opazovanji sledijo v živo ali pa jih kasneje preučijo na podlagi arhivskih podatkov.

Ključne besede: ključne besede
PACS: 96.50.S-, 96.50.sd, 07.05.Kf

Abstract

Despite their discovery already more than a century ago, Cosmic Rays (CRs) still did not divulge all their properties yet. Theories about the origin of ultra-high energy (UHE, $> 10^{18}$ eV) CRs predict accompanying primary photons. The existence of UHE photons can be investigated with the world's largest ground-based experiment for detection of CR-induced extensive air showers (EAS), the Pierre Auger Observatory, which offers an unprecedented exposure to rare UHE cosmic particles.

The discovery of photons in the UHE regime would open a new observational window to the Universe, improve our understanding of the origin of CRs, and potentially uncloak new physics beyond the standard model.

The novelty of the presented work is the development of a "real-time" photon candidate event stream to a global network of observatories, the Astrophysical Multimessenger Observatory Network (AMON). The stream classifies CR events observed by the Auger surface detector (SD) array as regards their probability to be photon nominees, by feeding to advanced machine learning (ML) methods observational air shower parameters of individual CR events combined in a multivariate analysis (MVA).

The described straightforward classification procedure further increases the Pierre Auger Observatory's endeavour to contribute to the global effort of multi-messenger (MM) studies of the highest energy astrophysical phenomena, by supplying AMON partner observatories the possibility to follow-up detected UHE events, live or in their archival data.

Keywords: astroparticle physics, ultra-high energy cosmic rays, ultra-high energy photons, extensive air showers, Pierre Auger Observatory, multi-messenger, AMON, machine learning, multivariate analysis
PACS: 96.50.S-, 96.50.sd, 07.05.Kf

Contents

Povzetek . . . . . **i**

Abstract . . . . . **iii**

Contents . . . . . **v**

1. Introduction **1**

2. Ultra-high energy cosmic rays **5**
2.1. Early history of cosmic rays 5
2.2. Direct cosmic ray measurements 9
2.3. Energy spectrum 11
2.4. Origin of ultra-high energy cosmic rays 16
2.4.1. Acceleration models (Bottom-up models) 17
2.4.1.1. Fermi acceleration 19
2.4.2. Non-acceleration models (Top-down models) 21
2.5. Potential sources of UHECRs 23
2.6. Propagation of Cosmic Rays 31
2.6.1. Deflection in magnetic fields 31
2.6.2. Energy loss processes 34
2.7. Mass composition 39
2.8. Anisotropy 42
2.9. Practical application of cosmic rays 45

3. Ultra-high Energy Photons **47**
3.1. Experimental state of the art 47
3.2. Production of cosmogenic UHE photons 48
3.3. Propagation of UHE photons 50
3.4. Testing Lorentz Invariance 52
3.5. Multi-messenger observations 54
3.5.1. Neutrino searches 55
3.5.2. Galactic Neutrons 57
3.5.3. Electrons and positrons $e^{\pm}$ 58
3.5.4. Global networks of experiments and observatories 58
3.5.4.1. The Astrophysical Multimessenger Observatory Network (AMON) 58

4. Extensive Air Showers (EASs) **61**
4.1. Air shower development 62
4.1.1. Electromagnetic cascade & Heitler model 64
4.1.2. Hadronic component & superposition model 65
4.1.3. Muonic component 66
4.1.4. Longitudinal & lateral development of air showers 67
4.2. Photon induced extensive air showers 68
4.2.1. Landau-Pomeranchuk-Migdal effect 69
4.2.2. Preshower effect 69
4.2.3. Electromagnetic UHE shower properties 70

5. The Pierre Auger Observatory **73**
5.1. The fluorescence detector 73
5.2. Surface detector 75
5.2.1. Calibration 76
5.2.2. The triggers 77
5.2.3. Aperture and Exposure of the Surface Detector 80
5.2.4. Data reconstruction 81
5.2.4.1. Shower geometry 82
5.2.4.2. Lateral Distribution Function 86
5.2.4.3. SD Energy Reconstruction 87
5.3. The Offline framework 88
5.4. Upgrade of the observatory *AugerPrime* 89

6. Applying machine learning for UHE photon search **91**
6.1. Air shower simulations 91
6.1.1. Utilized CORSIKA air shower libraries 92
6.2. Detector response 93
6.3. Data reconstruction 93
6.4. Experimental data - SD burn sample 95
6.5. Machine learning techniques in data analysis 96
6.5.1. Approaches and techniques in machine learning 96
6.5.1.1. Principal component analysis 98
6.5.1.2. (Boosted) Decision Trees 99
6.5.1.3. Linear discriminant analysis (LDA) 104
6.5.1.4. Support vector machines 105
6.5.1.5. Multilayer perceptrons / Neural Networks 106
6.6. Analysis results 110
6.6.1. TMVA workflow 110
6.6.2. Application of ML techniques for photon candidate event selection 110
6.6.3. MVA with two observables 114
6.6.3.1. **A**: Recalculated rise-time & Curvature 114
6.6.3.2. **B**: Δ_R rise-time & Curvature 114
6.6.3.3. ROC curves: case **A** vs. **B** 115

6.6.4. MVA with three observables . . . 117
6.6.4.1. **C**: recalculated rise-time & Curvature & NSt . . . 117
6.6.4.2. **D**: Δ_R rise-time & Curvature & NSt . . . 117
6.6.4.3. ROC curves: case **C** vs. **D** . . . 118
6.6.5. MVA with four observables . . . 120
6.6.5.1. Case **A**: rise-time recalculated + curvature + NSt + θ . . . 122
6.6.5.2. Case **B**: Δ_R + curvature + NSt + θ . . . 122
6.6.5.3. ROC curves: case **A** vs. **B** . . . 122
6.6.6. Application to the SD burn data . . . 124
6.6.6.1. Comparison against mixed proton-iron background . . . 134

7. Conclusions . . . **137**

Bibliography . . . **139**

List of Figures . . . **172**

Appendix . . . **181**
A.1. Further plots . . . 182
A.1.1. Benchmark function fits . . . 182
A.1.2. LDF β . . . 183
A.1.3. $S(1000)$ . . . 183
A.1.4. S_{38} . . . 184
A.1.5. $S(1000)$ vs. S_{38} . . . 184
A.2. MVA distributions and cut efficiencies for tested methods on SD burn sample . . . 185
A.3. ReLU activating function in MLPBFGS for MVA . . . 187
A.4. PCA transformed MVA input variables . . . 188
A.5. Comparison of methods . . . 189
B.1. $\overline{\text{Off}}\underline{\text{line}}$ module sequences . . . 191
B.1.1. Simulations with $\overline{\text{Off}}\underline{\text{line}}$ version: ape-0.99-v3r3p4 . . . 191
B.1.2. Simulations with $\overline{\text{Off}}\underline{\text{line}}$ trunk version 32493 . . . 192
B.2. Proposal for the technical implementation . . . 193
B.3. Settings of tested ML methods . . . 195

Acronyms . . . **197**

1. Introduction

The whole Universe with all its galaxies and thereby also our home galaxy, the Milky Way, with our Solar system are permanently exposed to a flux of highly energetic particles: the cosmic rays (CRs). After a century of research since their discovery [1], much more is known about CRs today, and their properties have been measured over a notably large energy range starting from MeV region to ultra–high energies (UHEs) $> 10^{18}$ eV, up to a few 10^{20} eV, but at the highest energies still several major questions remain unanswered:

- What are the physical processes that produce CRs at the highest energies?
- Which astrophysical objects are capable to serve as ultra–high energy cosmic ray (UHECR) sources?
- What is the maximal energy that can be reached at such cosmic accelerators? What are the energy limiting factors in the involved acceleration mechanisms?
- Of which chemical elements precisely are UHECRs composed? Is our understanding of CRs biased, due to the unknown properties of hadronic interactions at these energies, inaccessible at man-made accelerators?
- At which energies does the transition from Galactic to extragalactic CR sources take place?

Measurements of CRs at UHEs present experimental difficulties due to the very low CR flux (above 10^{20} eV, only about 1 CR per km^2 per century arrives at the Earth). UHECRs can therefore be studied only indirectly, using the extensive air showers (EASs) induced by the primary CRs in the atmosphere, and with the help of detectors on the ground. The world's largest experiment to measure such air showers is the Pierre Auger Observatory. The detector is located near the town of Malargüe in Argentina and, since starting collecting CR data in 2004, it helped enormously to increase our understanding of the properties of the highest-energy CRs. Due to its large instrumented detector area ($\sim$ 3000 km^2), the Pierre Auger Observatory offers an unprecedented exposure for UHE particles. The combination of various detector systems, principally of a large array of water Cherenkov detectors (WCDs) and fluorescence telescopes, in a hybrid approach enhances the data precision and enables a better control over systematic uncertainties for CR measurements.

The ongoing upgrade of the Observatory, *AugerPrime*, will supply it with new complementary detector types, which will boost the Auger UHECR mass composition sensitivity and in this way will further contribute to revealing properties of the highest-energy particles.

Additional important capabilities of the Pierre Auger Observatory to tackle the questions about the origin of extreme energy CRs arise from its sensitivity to UHE photons, neutrinos and neutrons. Searches for these messengers of the Nature's fundamental forces at the highest energies are motivated by theoretical scenarios describing the origin of UHECRs. In particular, during CR acceleration and propagation, hadronic interactions occur where photons and neutrinos are produced. As neutral messengers they provide information about CR source direction on the sky, and the physical conditions at the CR production site, while the charged CRs suffer from loss of directional information by deflection at intergalactic and galactic magnetic fields (MFs). Neither neutrinos nor photons have been observed at UHEs so far, which resulted in the corresponding stringent flux limits which constrained many models of UHECR origin.

The presented work focuses on implementing CR data analyses sensitive to UHE photons, with an aim to contribute to multi–messenger (MM) astrophysical studies performed by the Pierre Auger Observatory. Astrophysical photons have not yet been observed above a few PeV energy [2]. An observation of UHE photons – predicted by CR models – would hence not only open a completely new observational window to the Universe, but it would also have large impact on particle physics and astrophysics.

In this thesis a new approach for a "real-time" photon candidate data stream and a corresponding trigger from the Pierre Auger Observatory to the Astrophysical Multimessenger Observatory Network (AMON), aiming for MM identification of astrophysical sources of events, is presented. In order to reach this goal, measurements from the surface detector (SD) array of the Pierre Auger Observatory are analysed and a classification of the primary particles as UHE photon candidates is performed using machine learning (ML) techniques. The CR event classification is realized via the combination of four observables of the recorded SD events in a multivariate analysis (MVA). The selected photon candidate events will be then sent to the observatories participating in AMON, which can conduct follow-up studies of the detected UHE events.

The structure of this thesis is as follows: chapter 2 introduces CR research and describes its current experimental results, including its MM aspects, with a focus on UHEs. In the following chapter 3, UHE photons are discussed in detail, from the plausible scenarios of their origin to their propagation through the Universe till their detection on Earth. A description of EAS and air-shower components, and of the properties of UHE photon-induced particle cascades is given in chapter 4. An overview of the Pierre Auger Obser-

vatory and its several detector components and performances, followed with a short summary of the upgrades of the observatory is given in chapter 5. Simulations of air showers, on which the analysis is based, as well as the data sample collected at the Pierre Auger Observatory, to which the analysis is applied, and a detailed study of observables used for photon discrimination are presented in chapter 6. The implemented MVA, the tested ML techniques, the analysis performances, the results of the MVA application to the CR data, and the proposal for the technical implementation of a UHE photon candidate event data stream to AMON are described there in detail. A summary of the main results of the presented thesis work and an outlook for possible future extensions of the proposed data stream are given in chapter 7.

2. Ultra-high energy cosmic rays

A short survey of the history of CRs, including their discovery and the past milestones of the cosmic-ray research, as well as the experiments studying the lower part of the CR energy spectrum via direct detection is given as an introduction of this chapter. For completeness a few words are given on the lower regime of the CR energy spectrum before the focus of this chapter is changed on the high end of the observed energy spectrum of CRs, providing a description of the main acceleration mechanisms, some of the potential UHECR source classes and the propagation of CRs.

2.1. Early history of cosmic rays

The last of a series of balloon flights on the 7^{th} of August 1912, dedicated to measurements related to the conductivity of the air, by the Austrian physicist Victor Franz Hess [1], led to the discovery of CRs, which marked the dawn of astroparticle physics and was acknowledged by awarding Hess with the Nobel Prize in physics in 1936 [3].

To describe where the name "cosmic rays" comes from, what led to the CR discovery, to understand how science came to the picture that we have about CRs today, and why cosmic-ray physics is interpreted as a building block of modern particle physics, some CR research milestones are presented here in a brief chronological manner.

Compared to the state of present-day knowledge very little was known about radiation in general in the end of the 19^{th} and the beginning of the 20^{th} century. The reason for the natural conductivity of the air was accepted not to arise from spontaneous disintegration of the air molecules due to thermal excitation, but being induced by some unknown radiation. A global hunt for the origins of this ionizing radiation was launched. However, first hypotheses of the possibility that this ionization is due to radiation coming from extraterrestrial sources by Marie Curie [4] and later by Wilson [5] were abandoned. Elster and Geitel [6] and independently Wilson [5] were the first to conclude that some type of radiation is constantly ionizing the air surrounding us. Various experiments showed the dependency of the conductivity of air on atmospheric conditions, changing with altitude - more intense over land than over sea, concluding that the radioactive emissions accumulating in the atmosphere and the radioactive substances contained in the Earth's

soil/crust are the main source of ionization of the air. Their findings led to the commonly accepted assumptions that the main source of the air ionization is coming from α-, β- and γ-radiation from radioactive substances in the soil. Since α- and β- radiation is rapidly absorbed in the air, the more penetrating γ-radiation was considered as the most important cause of the ionization. Several measurements to investigate how the speed of ionization changes with height were made possible by the invention of the electrometer, an improved electroscope which was a more sensitive and portable device, by Wulf [7], who performed measurements on top of the Eiffel tower (a height of 330 m), and thereby noticed that the speed of ionization decreases with height, but much more slowly than expected [8]. First results of measurements during balloon flights by Linke [9] were unfortunately interpreted in a wrong way. In subsequent measurements by Gockel and Wulf [10] in the Swiss Alps at different altitudes it has been concluded that the ionization rate is not dependent on the altitude, and that some type of radiation from outer space – using for the first time the term "cosmic radiation" – should only contribute insignificantly to the ionization rate. The nowadays used, but misleading term "rays" originates from these times, when for quite a long time "cosmic rays" were assumed to consist only of photons, in analogy with the gamma-rays – the strongest source of ionization known at that time. In contrast to Wulf, further systematic studies of Pacini [11] on the shoreline, at sea and even/especially underwater, where radioactive elements in the soil should have no influence, found, that the air ionization rate is higher than expected, and therefore concluded, that there should be a significant source of radiation in the atmosphere itself. Proclaimed uncertainties in the observations, technical problems with instruments, false interpretations of the measurement results prevented scientists so far to make definitive conclusions about "cosmic radiation".

Initially intended to determine the absorption factor of γ-radiation in the air, Victor Franz Hess undertook a series of several balloon flights in the years between 1911 and 1913. The first flights were conducted at lower altitudes up to 2100 m starting from Vienna. In the famous flight (with a maximum altitude of 5350 m) on August 7^{th} 1912, starting in Aussig an der Elbe (today Ústí nad Labem, Czech Republic) and landing in Pieskow (50 km south-east of Berlin), whereat Hess used more sensitive electrometers inside a hermetic vessel, overcame the problems of Gockel and others with their instruments, recorded in very precise detail the weather conditions and established the cosmic origin of the ionizing radiation with high confidence [1, 12, 13]. To exclude the sun as the source, Hess undertook flights not only during the day, but under different atmospheric conditions, during the night and at a time of a solar eclipse, where he did not find any difference in the measured ionization rate. The observed level of radiation decreased up to an altitude of about 1 km, but at higher altitudes an increase of the ionization rate up to a factor of four as compared to the rate on the ground was detected. Hess made

the assumption about the need of a new source of radiation of extraterrestrial origin, thus correctly concluding about the existence of cosmic radiation.

It took many years before this discovery and its interpretation became commonly accepted. Also the WWI stopped further scientific investigations. The confirmation of Hess' discovery, and the establishment of the existence of CRs was delivered by Werner Kolhörster, who reached even higher altitudes up to 9300 m during his balloon flights [14], where he found the same significant increase of radiation level as Hess.

Moreover, the scepticism of the Nobel prize winner Robert Millikan towards this discovery of extraterrestrial radiation unfortunately had strong influence on the opinions of many physicists at that time. After several years of efforts to disprove the results by Hess and Kolhörster, it was Millikan himself together with Cameron [15] who, via their experiments on lakes at high altitudes, confirmed Hess' and Kolhörster's findings. Millikan however still believed that his measurements proved that the observed radiation was due to high energy photons, and thereby introduced the name "cosmic rays", which is used all over the world to date since 1926.

Naturally arising and always accompanying a discovery of a phenomenon the questions of its nature and origin initiated the hunt for the properties of CRs. Jacob Clay was the first who found evidence that CRs are charged particles [16] in 1927, since his findings of an increase of the cosmic radiation level with the latitude indicated that the primary CRs are deflected by the geomagnetic field. New measurement techniques allowed to finally disprove the assumption that the cosmic rays were γ-radiation. Experiments (with results published in 1929) using Geiger-Muller counters demonstrated that CRs must carry electric charge and be of corpuscular nature, since coincidence signals in a setup shielded by an absorber material were produced [17].

Based on these findings, Bruno Rossi predicted the "*east-west effect*" in 1930 [18], which was later detected by three independent experiments in 1933 and 1934 [19–21]. This effect results in a difference between the intensities of CRs arriving from the east and the west that depends upon the charge and the energy of the primary charged particles. In fact, it was found that the CR intensity is greater from the west, thus proving that most primaries are positive.

A further milestone was unleashed in 1933 with the discovery of the dependence of the intensity of CRs on the geomagnetic latitude by Compton [22], which further disproved the assumptions for γ-radiation origin of cosmic rays.

Cosmic-ray physics can also be seen as the birthplace of modern particle physics, as became clear after the discovery of the first antiparticle, the positron. It was predicted by Dirac [23] based on quantum electrodynamics (QED), and discovered by Anderson in CR interactions in 1933 [24].

Hess eventually received international recognition for his discovery in 1936 through the Nobel Prize in Physics, which was awarded to him "for the dis-

covery of cosmic radiation" [3], and shared with Anderson awarded "for his discovery of the positron", see figure 2.1.

Figure 2.1.: Nobel prize ceremony 1936. Front row from left to right: Prof. Otto Loewi, Sir Henry Dale (both laureates in medicine), Prof. Peter Debye (laureate in chemistry), Dr. C.D. Anderson, Prof. V.F. Hess. Figure taken from [25].

The next milestone in the cosmic ray studies was the discovery of EASs via coincidence measurements in 1939, which is accredited to Pierre Auger [26]. At an altitude of 2350 m at the Jungfraujoch in the Swiss Alps, Auger measured the rate of coincidence as a function of the distance between single counters (up to a distance of 75 m), whereat he measured that the rate significantly exceeds the expected rate of chance coincidences. This was an indication that primary cosmic-ray particles induce cascades of secondary particles in the atmosphere. But it should be mentioned that already in 1935 Erich Regener and Georg Pfotzer used threefold coincidences to study the intensity of cosmic rays in the stratosphere [27] and reported an unexpected maximum at an altitude of about 14 km. Also Rossi in 1934 reported coincidences, near-simultaneous discharges of two widely separated Geiger counters, at a rate larger than the expected accidental rate [28].

In 1937 Homi Jehangir Bhabha and Walter Heitler published, based on the ideas of QED, their theory of electromagnetic showers [29], what led to the interpretation of the observed maximum from Regener and Pfotzer as the

result of the multiplication of electrons in the atmosphere, or the cascade production of gamma rays and electron-positron pairs. More on the Heitler model is given in section 4.1.1.

In the ongoing effort of cosmic-ray research many other new particles beside the positron were discovered, as there were the muons in 1937 [30]. After the pause in the cosmic-ray research forced by WWII, the "strange" kaons were discovered in CR-induced air showers in 1947 [31], followed by the charged and neutral pions, and hyperons, long before the technical level allowed for the high-energy particle research with man-made accelerators.

Concluding results of cosmic-ray research in the years from 1930 to 1948, primary CRs were found to consist of mostly protons, with approximately 10% of the primaries being helium nuclei (α-particles) and 1% heavier nuclei of elements such as carbon, iron, and lead. The secondary radiation produced in the atmosphere and reaching the ground was found to be composed mainly of electrons, photons and muons [32, 33].

It should be mentioned at this point that the term "cosmic rays" refers only to charged primary particles like ionized nuclei and electrons or positrons. Photons as well as neutrinos are chargeless particles and thus not referred to as "cosmic rays". Nevertheless, there is a strong connection between photons and neutrinos with charged particles in CRs. A focus on photons especially at the highest energies above the EeV (1 EeV = 10^{18} eV) range – called UHEs – will be given in chapter 3.

In order to explain the observed high energies of cosmic rays, in 1949 Enrico Fermi introduced a mechanism to describe the acceleration of charged particles in astrophysical plasmas (more in sec. 2.4.1.1) [34]. This so-called *Fermi acceleration* employing astrophysical shock waves is nowadays considered to be the main mechanism involved in the acceleration of CRs to UHEs.

Soon after the discovery of the cosmic microwave background (CMB) radiation, Zatsepin and Kuzmin, and independently Greisen derived in 1966 that CRs protons with energies above $\sim$ 60 EeV would strongly interact with the CMB photons, what effectively reduces the CR energy and restricts their path lengths [35, 36]. The expected cut-off in the energy spectrum of UHECRs resulting from this effect is called the Greisen–Zatsepin–Kuzmin (GZK) effect or cut-off (see more in section 2.6.2).

2.2. Direct cosmic ray measurements

The energy spectrum of CRs ranges from a few 10^8 eV to at least a few 10^{20} eV, see fig. 2.2. At the lower part of the spectrum CRs can be investigated directly thanks to experiments launched on air balloons like BESS [38] and CREAM [39], in satellites like PAMELA [40] and on the International Space Station (ISS) with the AMS-02 [41] (further details can be found in [42] and references therein). The direct detection experiments make use of the high flux of CRs

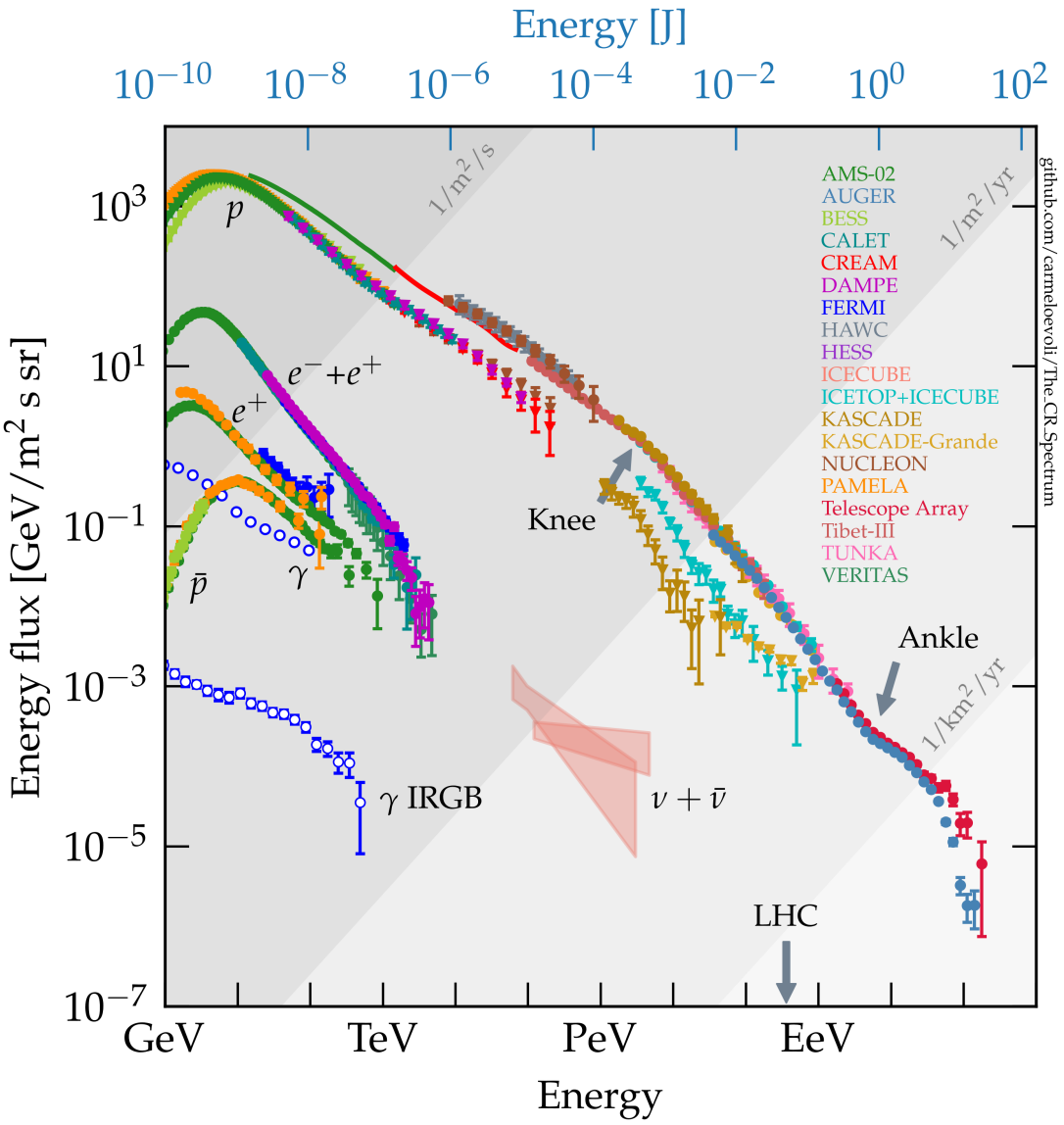


Figure 2.2.: Compilation of the full range CR energy spectrum measured by several experiments (after 2000). Figure taken from [37].

(compared to the extremely low flux at the high end of the energy spectrum) and thereby have direct access to properties of the primary particles, since they are measured before any particle interaction in the atmosphere. Historically, the direct experiments on balloons and rockets identified the first point sources of high energy cosmic electromagnetic (EM) radiation as for example the Crab Nebula, and the binary star systems Vela X-1 and Cygnus X-3 [43].The direct CR measurements can be performed up to a few PeV energies, where the flux of CRs is of the order of one particle per square meter per year.

Below the GeV-range the cosmic ray flux is suppressed, as the particles from outside the solar system are deflected by a magnetic field generated by

solar activity, as well as by the geomagnetic field. Thus up to GeV our Sun is the main contributor to the low energy cosmic ray flux, emitting electrons, protons and a small fraction of heavier nuclei [44]. Furthermore, the measured flux of CRs on Earth is modulated, because it is connected to the sunspot activity, which is in turn connected to the intensity of the solar wind and thereby following the same eleven-year cycle. Since the arrival directions of CRs at energies above GeV do not show any correlation with the sun, they evidently originate from outside the solar system.

2.3. Energy spectrum

Already in 1963 the Volcano ranch experiment, near Albuquerque New Mexico, reported the detection of an event with an energy of $\sim 10^{20}$ eV [45], but it took till the 1990s for the existence of CRs at these energies to be confirmed by experiments such as the Haverah Park array [46] and AGASA [47].

In 1991 the Fly's Eye observatory in Utah detected the most energetic particle that has ever been measured with a reconstructed energy of $(3.2 \pm 0.9) \cdot 10^{20}$ eV [48], puzzling the astroparticle-physics community. Since a strong suppression of the flux of CRs is expected above $\sim 6 \cdot 10^{19}$ eV due to the aforementioned GZK effect, see sec. 2.6.2, detections of particles with energies above the suppression would henceforth suggest nearby UHECR source(s) or require new physics - see sec. 3.1.

Previous UHECR detectors lacked both statistics and precision in measurements of the very rare cosmic ray events at these extreme energies. This aspired the development of the world's largest CR experiment, the Pierre Auger Observatory, which is described in detail in chapter 5. Now an overview of the present knowledge of the energy spectrum, arrival directions and mass composition of the UHECRs is given.

The number N of CRs in a certain energy band is represented by the energy spectrum, which can be approximated by an inverse power law in energy (E), thereby clearly indicating a non-thermal origin of CRs, and is represented via a differential flux given by

$$\frac{\mathrm{d}N}{\mathrm{d}E} \propto E^{-\gamma}\,, \tag{2.1}$$

where γ is the so-called spectral index, which indicates the steepness of the decrease of the measured CR flux with increasing CR energy. The energy spectrum is shown in figure 2.3 as a combination of the measurements of multiple experiments specialized to investigate CRs, where a scaling with $E^{2.7}$ has been applied in order to accentuate the specific spectral features.

Over the broad range in energy shown in fig. 2.3 most of the spectrum is rather featureless and can be well described by a power law (eq. (2.1)) with a constant spectral index γ. Changes in the spectral index γ at the certain points called the "knee", "2$^{\text{nd}}$ knee", "ankle" and the "cut-off" are interpreted to

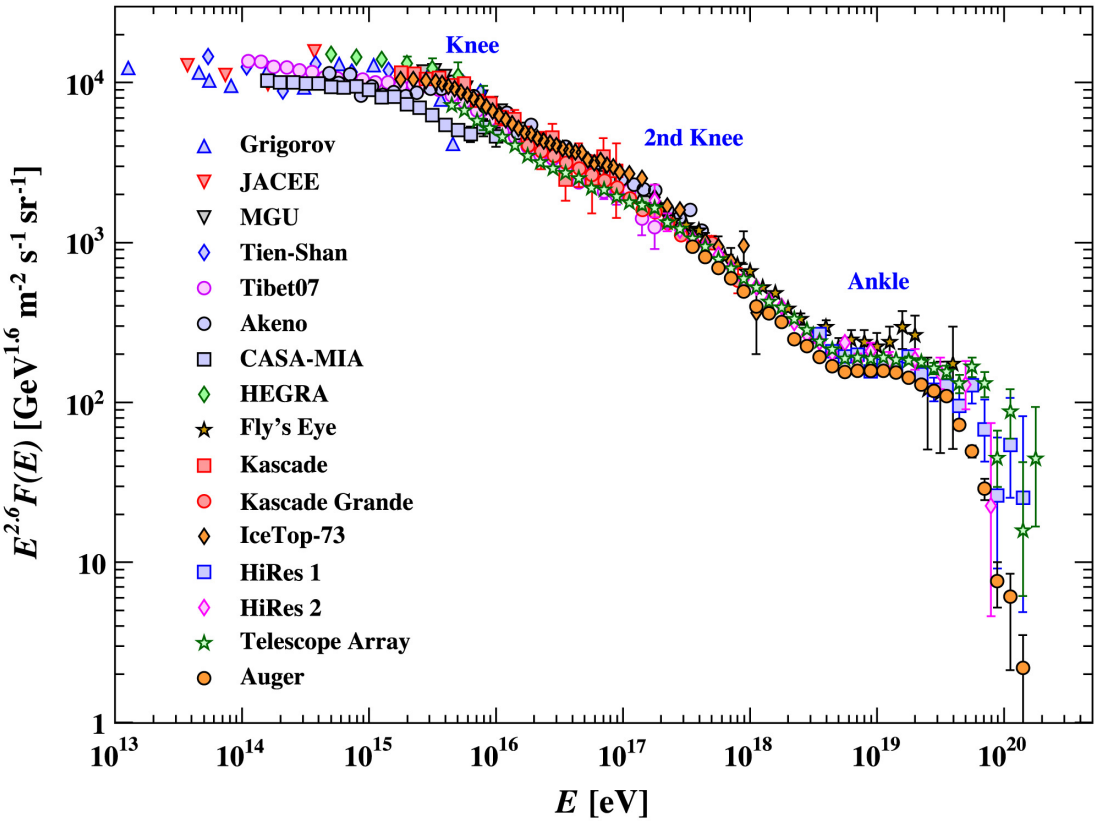


Figure 2.3.: Differential energy spectrum of charged cosmic rays, measured from 10^{13} eV to few 10^{20} eV by cosmic ray observatories. The spectrum has been multiplied by $E^{2.6}$ in order to put more emphasis on the features of the steep spectrum which are otherwise difficult to discern. These features are: the "Knee" at around $4 \cdot 10^{15}$ eV, the "2nd Knee" at 10^{17} eV and the "Ankle" at $4 \cdot 10^{18}$ eV. Figure taken from [42].

arise due to different astrophysical phenomena, as due to the various types of sources of CRs, the mechanisms of production, interactions and propagation through the interstellar medium and to the chemical composition of CRs [49].

Knee

The spectrum follows a power-law (eq. (2.1)) with $\gamma = 2.7$ till the "knee" at $\sim 4 \cdot 10^{15}$ eV, where a drop off of the lightest component of the composition (protons) has been shown for example by the KASCADE-Grande experiment [50]. The steepening to a $\gamma = 3.1$ that occurs above, can be explained as to arise from a rigidity-dependent energy cutoff of CR nuclei accelerated at sources, due to limits of magnetic confinement, the (EM) rigidity $R = |\vec{p}|/eZ$ [51], where $|\vec{p}|$ is the absolute value of the momentum, e the elementary charge, and Z the nuclear charge number, e.g. $Z = 1$ for protons and $Z = 26$ for iron nuclei. This results in the maximum energy of nuclei proportional to

their charge; that is, in the cutoff rigidity (energy/charge) the same for all nuclear species, and therefore in a limit on the acceleration power of protons by galactic sources as for example supernovae [52]. Another contributing factor could be due to a diffusion and drift effect, which causes CRs to escape the galaxy above a certain energy threshold [53]. Nevertheless, the origin of this spectral feature is not yet fully clear and several models exist to explain the knee [54].

2nd knee

A weaker second knee has been found at $\sim 4 \cdot 10^{17}$ eV [55] and can be seen to arise as a logical extension of the scenario presented for the first knee, leading to a spectral index of $\gamma = 3.3$. Between these two knee-like features composition studies have hinted toward a transition from lighter to heavier nuclei [56] and would correspond to the rigidity-dependent acceleration limit of galactic supernovae for the heaviest CRs [57]. A phenomenological model considers the total flux to be the result of the superposition of the fluxes of the single elemental components of CRs, called the "many knees model", with an energy cut-off for each component proportional to the charge Z [58]. However, this model relies on several assumptions on the chemical composition of CRs in this energy regime which is not yet well measured.

Furthermore, there are models [55, 59, 60] interpreting the second knee as the transition region from galactic to extragalactic origin of the primary CRs.

The CR dipole phase and amplitude as a function of energy (shown in figure 2.4), as well as the evolution of the CR mass composition with energy provide hints on the transition [61]. The change of the phase of the dipole with energy (fig. 2.4) from pointing towards the Galactic Center (GC) (-90 deg $= 270$ deg = 18h00m) at 1–10 PeV range to completely away from the GC at the highest energies is an expressive representation of the transition from galactic to extragalactic origin of the primary CRs. Except for the transition region between a few PeV and EeV, a statistically significant dipole anisotropy (be it weak anisotropy amplitudes of the order of a permill or lower) has been observed at almost all energies [62].

The dipole component of the anisotropy in the TeV to PeV range might be explained by the movement of the solar system in the Milky Way – the *Compton-Getting effect* [63, 64] – though does this hypothesis need to be studied with more accurate measurements of the energy evolution of the anisotropy [65].

The ankle region

The next prominent spectral feature, the ankle, occurs at $\sim 4 \cdot 10^{18}$ eV, where a flattening of the spectrum is observed with γ changing to ~ 2.7. At such high CR energies, an extragalactic component starts to take over in the total flux

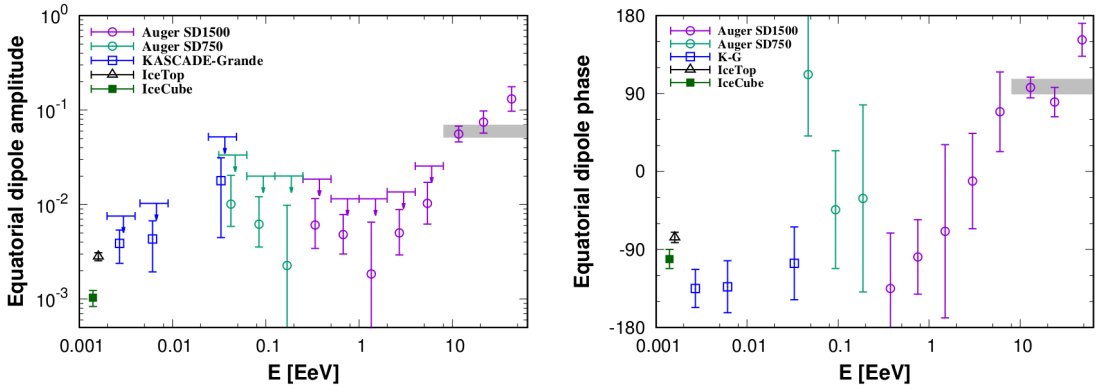


Figure 2.4.: Amplitude (left) and phase (right) of the equatorial dipole amplitude. Plot taken from [62].

[48, 66]. Several models concerning the galactic-extragalactic transition and the relevant experimental constraints are reviewed in [60]. An indicator of the commonly believed extragalactic origin of CRs with energies above the ankle is that no correlations with the galactic plane or the GC have been found to date. A recent discovery by the Pierre Auger Observatory of a dipole structure of the large-scale angular distribution of the arrival directions of UHECRs (see figure 2.5) pointing about 125° away from the GC firmly established the extragalactic UHECR origin [62, 67].

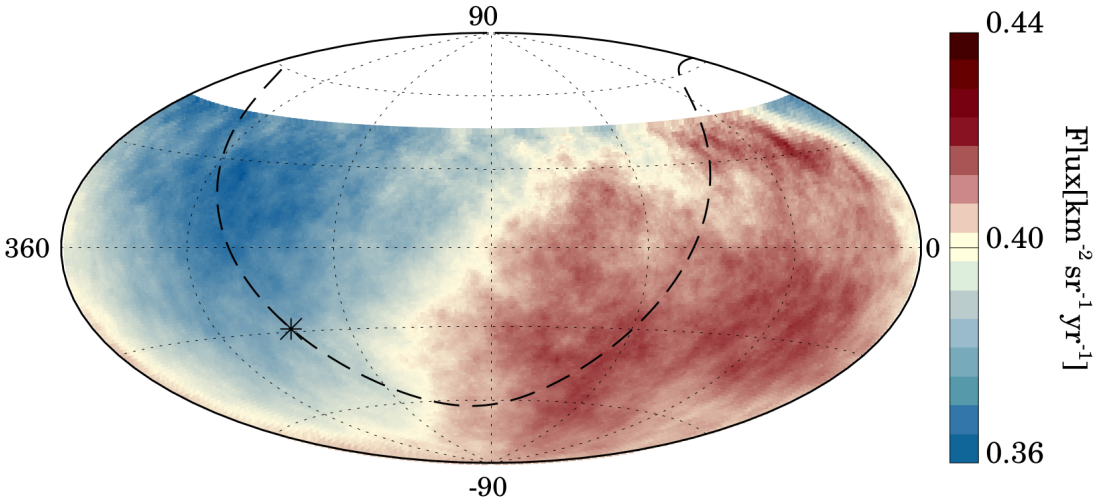


Figure 2.5.: The flux of CRs above 8 EeV in equatorial coordinates, where the dashed line represents the galactic plane and the star indicates the GC. Plot taken from [62].

Historically, the ankle has been interpreted in the framework of the so called "dip"-model. It assumes a proton-dominated extragalactic component and the transition from galactic to extragalactic CRs occurring at energies of the second knee, thus being compatible with the estimates of the maximum at-

tainable energy of galactic sources for CRs [55, 68]. In this model two main processes cause the energy losses of CRs [69]. The first process of energy loss can be explained via adiabatic energy losses (see sec. 2.6.2) due to the expansion of the Universe, affecting CRs of all energies, therefore not changing the shape of the energy spectrum, and is more relevant for distant sources. A second process of energy loss, responsible for the ankle formation, would be a pair-production interaction of the UHE protons with CMB photons [59], creating electron-positron pairs:

$$p + \gamma_{CMB} \rightarrow p + e^- + e^+ , \tag{2.2}$$

(or more general any nucleon with photons of the cosmic background radiation (CBR), see sec. 2.6.2) and causing the flattening of the spectrum at the ankle in the form of a dip (more on energy loss process is described in sec. 2.6.2). However, UHECR mass composition measurement results by the Pierre Auger Observatory indicate a mixed CR composition at the ankle [70, 71], thus disfavouring the dip-model as a sole explanation for the ankle origin..

Another concept proposed to explain the ankle involves photo-disintegration of CR nuclei in the environment surrounding the source, which acts like a "high-pass" filter on UHECRs, where only the most energetic CRs can escape their environment [72]. For a range of source conditions, such mechanism leads to a light extragalactic component near and below the ankle, which evolves to intermediate mass composition above the ankle, in agreement with the measurements of the evolution of the shower maximum depths with energy [61].

The increased CR event statistics, together with the high measurement precision of the Pierre Auger Observatory allowed to reveal that the spectrum above the ankle has a shape more complex than a the simple broken power law that was used until recently to describe it. A power-law is not anymore considered an appropriate description at all the cosmic-ray energy spectrum above the ankle. A new feature has to be included in the UHECR energy spectrum description between the ankle and the cut-off [73]. The reported measurement of the energy spectrum of CRs above $2.5 \cdot 10^{18}$ eV by the Pierre Auger Collaboration, based on 215,030 events, with most recent values for the spectral index and features of the spectrum in agreement with models of energy-dependent mass composition can be seen in figure 2.6 [49].

Cut-off

A strong suppression of the flux of CRs has been observed starting from $\sim 5 \cdot 10^{19}$ eV. The question arises whether the flux suppression is caused by the GZK effect, or if the capabilities of acceleration of sources are limited (see the next section, sec. 2.4.1), which by chance manifests in similar spectral features as predicted for the GZK effect (see e.g. [74]).

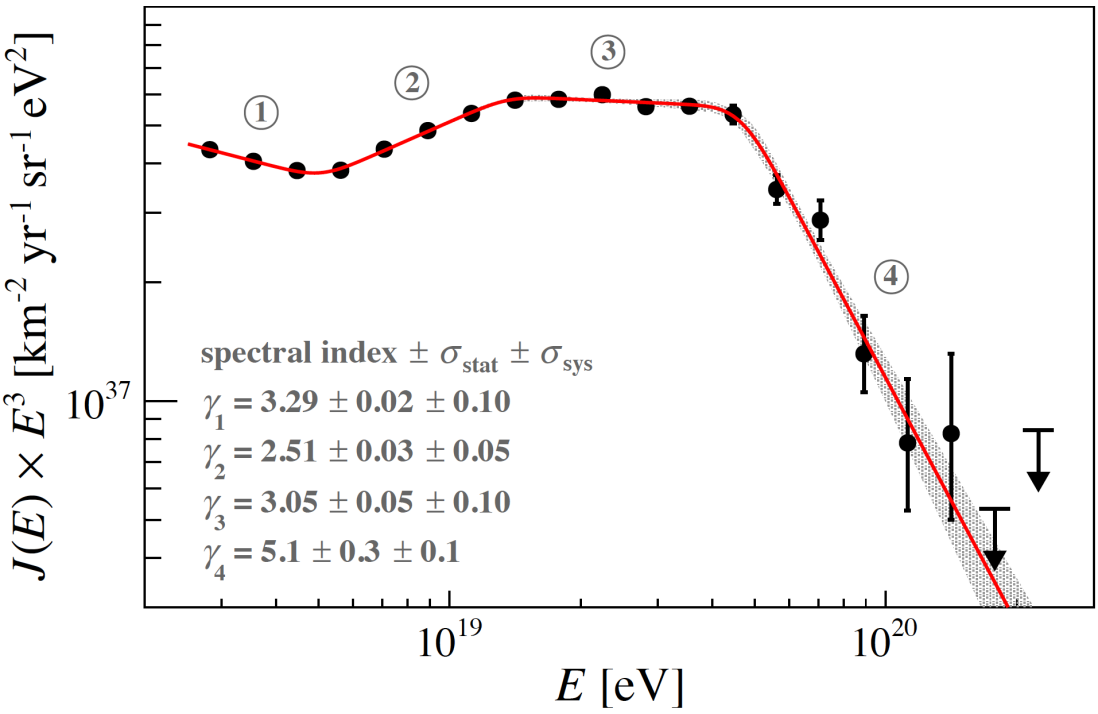


Figure 2.6.: The energy spectrum measured at the Pierre Auger Observatory, scaled with E^3 and fitted with a sequence of four power laws represented by the red line. The spectral index γ_i of the power law corresponds to the energy intervals identified with the numbers ($i = 1, 2, 3, 4$), the shaded band represents the statistical uncertainty of the fit, and the upper limits are at the 90% confidence level (CL). Plot taken from [49].

Measurements of the energy spectra with two observatories with largest exposures for UHECR event detection, the Pierre Auger Observatory in the southern hemisphere and the Telescope Array (TA) in the northern hemisphere, are in agreement with each other when taking into account statistical, systematic, and energy scale uncertainties [75], see fig. 2.7. Some tension occurs only at the very highest energies above the cut-off between the flux measured by the two experiments and can be explained only partly by the observation of different parts of the sky [76].

2.4. Origin of ultra-high energy cosmic rays

The mechanisms capable of generating the highest observed energies are still not entirely clear. Astrophysical objects serving as possible candidate sources are explained in the next section 2.5, while here the underlying physical mechanisms are briefly discussed.

To explain how the charged particles can reach UHEs, two radically different categories of UHECR origin scenarios have been proposed: astrophysical acceleration models in which particles acquire energy, and therefore also

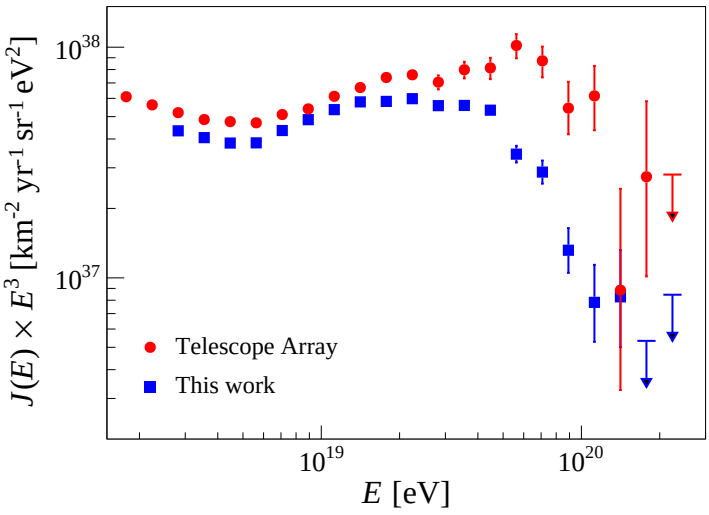


Figure 2.7.: Comparison between the spectrum derived with the Pierre Auger Observatory in blue and the one derived at the Telescope Array, scaled with E^3. The blue graph referred to by the words "this work" in the legend refers to the article [75] from which the plot is taken.

called *bottom-up* models, and the more exotic *top-down* models. All these approaches have of course to be consistent with the measured CR properties, i.e. their energy spectrum (sec. 2.3), mass composition (sec. 2.7) and possible anisotropies (2.8) in their arrival directions. For example, the measured energy spectrum of CRs reveals a non-thermal process that would require e.g. an active acceleration mechanism in order to maintain a power law spectral shape over several decades in energy [77].

Another aspect to keep in mind is the particles' voyage through the Universe. Once produced at their sites, CRs have to travel through the Universe, where they will interact with the surrounding environment – especially photon fields – before reaching detectors on our planet. As a consequence, during the UHECR propagation the composition and the shape of the spectrum as well as the CR arrival directions - due to the deflections of charged particles in cosmic MFs - will get modified.

2.4.1. Acceleration models (Bottom-up models)

The bottom-up models have in common that CRs undergo acceleration at astrophysical environments with magnetized plasmas, which require for their detailed description the complex magneto-hydro-dynamic (MHD) processes. These models can in general be differentiated into two main categories:

- **"One-shot" models**
  Scenarios where CRs are directly, in one step, accelerated to the highest

energies may be realized in jets – relativistic outflows of ionised matter (or "plasma outflows") transporting energy and momentum from "power engines", e.g. from supermassive black holes (SMBHs) in the centres of active galactic nuclei (AGNs), to remote locations, – or in rapidly spinning magnetized neutron stars (pulsars) such as the Crab, see fig. 2.12, or the Vela pulsars [78]. While the obvious advantage of such scenarios is the fast acceleration processes, the high density environment in the vicinity of the sources will lead to important energy losses. Another reason why this mechanism is not widely favoured is that it cannot explain the observed power law spectrum. A representative of this class of models, a relativistic blast wave scenario, is dubbed "espresso" acceleration [79, 80].

- **Stochastic Acceleration**
  In another approach, named after Enrico Fermi who was the first to propose it in 1949, and referred to as *Fermi acceleration* mechanisms, CRs are slowly and iteratively accelerated through multiple stochastic collisions in shock fronts with moving magnetized plasma. Thereby CRs need to be kept confined in the acceleration zone – called Fermi-acceleration zone – to reach UHEs. The elaborated versions of this scenario successfully explain the power law shape of the CR energy spectrum and, due to the ubiquity of astrophysical shocks in the Universe, are considered as the prime mechanism contributing to the UHECR flux. The Fermi acceleration mechanisms are described in more detail in the following section. However, these mechanisms are slower as compared to the electric field acceleration, and it is hard to keep the particles confined within the acceleration site.

  An upper limit of the energy to which sources can accelerate particles in this stochastic approaches can be estimated by evaluating the probability of the particles to escape from the acceleration region, as was done for the first time by A. M. Hillas [81]. Particles undergoing this type of acceleration are hard to keep confined in the acceleration site, if the Larmor radius r_L (or also called gyroradius, cf. eq. (2.7)) of a relativistic particle of charge Ze is of the order of the acceleration site's characteristic size $R = l \cdot \Gamma$, with l the comoving size of the source, and Γ the Lorentz factor of the motion. The resulting maximum energy E_{max} is related to the magnetic field strength B in the source by the commonly referred "Hillas condition"

$$E_{max} \approx \beta_s c \cdot Ze \cdot B \cdot R \,, \tag{2.3}$$

  where $\beta_s c$ is the relativistic velocity of the scattering centres in the acceleration region, which are vital to the acceleration process.

  Classes of objects capable to accelerate CRs to the highest energies in terms of the product of $R \cdot B$ are shown in figure 2.8. Such type of plots is called a Hillas diagram after A.M. Hillas who introduced them. The

solid diagonal lines show the minimum product of $R \cdot B$, which is required to accelerate protons (red) and iron nuclei (blue) to 10^{20} eV by a fast shock. The categories of objects to the left of the lines do not satisfy the Hillas criterion (eq. (2.3)). As shown with the dashed diagonal lines, the required product of $R \cdot B$ is higher for slower shocks ($\beta_s = 0.01$ is used for illustration). Normal galaxies, supernovae, and Wolf-Rayet stars (stars that drive strong magnetized winds) do not satisfy the confinement condition, while other source classes in fig. 2.8 do with AGNs, gamma-ray bursts (GRBs), neutron stars, and galaxy clusters (via galaxy merger shocks) being the most prominent UHECR source candidates. The largest Lorentz factor values Γ are observed in sources that are described later in more detail (see sec. 2.5): $\Gamma \sim 10 - 50$ in AGN jets and $\Gamma \sim 10 - 1,000$ in GRBs [82].

Another argument for extragalactic origin of UHECR sources arises from eq. (2.3): even if Galactic accelerators could reach EeV energies, most of the produced highest energy particles would escape the Galaxy, since the maximum energy for a proton that can be confined in the Galaxy (with $B \approx 0.3$ μG, $R < 1$ kpc and $\beta_s = 1$) is $E_{max,Gal} < 0.3$ EeV.

In addition to Fermi acceleration scenarios, several other mechanisms (fast acceleration models) should be mentioned, as for example fast magnetic reconnection in the vicinity of black holes (BHs) and relativistic jets [83, 84], inductive electric fields in magnetized, relativistic sheared outflows [85], and proton acceleration via unipolar induction in an aligned pulsar magnetosphere [86], yet have these scenarios not delivered sufficient observational proof.

2.4.1.1. Fermi acceleration

In general one has to differentiate between two sets of models, depending on the relativistic velocity of the scattering centres, β.

Second-order Fermi acceleration

Historically introduced by Fermi in his original paper [34] and less efficient is the nowadays called second-order Fermi acceleration. In this scenario charged particles are accelerated through elastic scattering with moving magnetized gas clouds. The particles are reflected on the clouds since each gas cloud works as a moving mirror and accelerates the particle in the case of a head-on collision while decelerating it in the case of a tail-on collision. As Fermi derived, statistically an overall energy gain is expected, since head-on collisions are more likely than tail-on collisions, depending on β via

$$\frac{\delta E}{E} \propto \beta^2 , \tag{2.4}$$

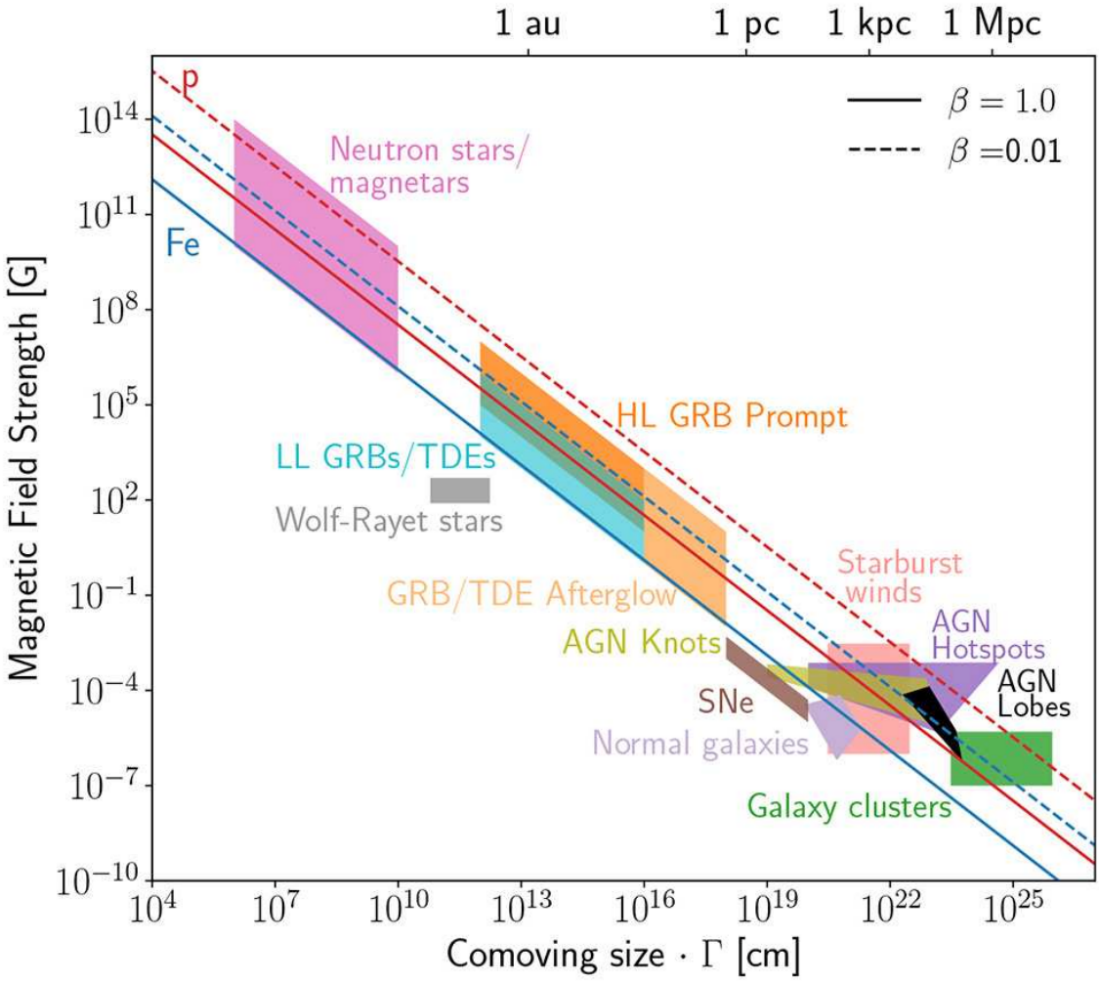


Figure 2.8.: Hillas diagram, representing source classes shown as function of their characteristic size and magnetic field strength (quoted values are in the comoving frame of the source). The diagonal lines indicate the limit beyond which confinement of protons (red) and iron (blue) nuclei with energy 10^{20} eV is possible. Two different shock velocities β (in units of c) are indicated as solid and dashed lines. Figure taken from [82].

where δE is the energy gain of a particle with energy E. The self-explaining attribute "second-order" originates from this dependence, and since $\beta \ll 1$ this mechanism is very slow. Later this acceleration mechanism has been shown to produce a thermal spectrum [87, 88].

First-order Fermi acceleration

Blandford and Ostriker [89] modified the previous mechanism, enabling a more efficient way to accelerate CRs to the highest energies, by taking as scattering centres the shock waves that occur near supernovae. This mechanism, referred to as diffusive shock acceleration (DSA), can accelerate particles pro-

portionally to the gas velocity and is hence called (in comparison to eq. 2.4) first-order Fermi acceleration

$$\frac{\delta E}{E} \propto \beta \,. \tag{2.5}$$

If a relativistic particle crosses the supernova shock front, it will scatter on the approaching ionized gas on both sides of the shock, and therefore it will get magnetically reflected each time it crosses the shock front. In addition to being more efficient than the second-order process, an interesting feature of this mechanism is the naturally produced power law with a spectral index of -2 in the energy spectrum [90]. The first-order Fermi mechanism is considered as the main acceleration mechanism for UHECRs.

2.4.2. Non-acceleration models (Top-down models)

In order to explain the CR measurement results by the AGASA experiment [91] and in particular the observation of CR events above 10^{20} eV, several top-down models have been introduced, some conform with, others beyond the Standard Model of particle physics (SM). Further intention behind these models is to evade difficulties with acceleration mechanisms for particles reaching energies in the EeV range. These types of models imply hypothetical and super-massive particles, referred to in the following as X particles, which subsequently decay into particles with UHEs. The rest masses m_X of those particles are thus required to be $> 10^{20}$ eV/c^2, i.e. close to the mass scale of Grand Unified Theories (GUTs). Furthermore, to be compatible with the observed flux of CRs, the density of those X particles in the Universe as well as their decay rates are required to be large enough. As a direct consequence of the decay of the X particles, ineluctably all top-down models predict a large flux of UHE photons accompanying the UHECRs. This property serves as an accessible scope to test those exotic models.

With the increase and the quality improvement of experimental data over time, upper limits on the integral flux of UHE photons already excluded a big part of the parameter space of some of these models. The current upper limits heavily constrain top-down models (see sec. 3.1), leading to a general disfavour of these models compared to the astrophysical acceleration scenarios.

Here, most prominent top-down models, which have not yet been ruled out (entirely) by experiments, like Topological Defects (TDs), the annihilation of relic neutrinos in Z-bursts, or Super-Heavy Dark Matter (SHDM), are briefly introduced.

Topological Defects

Models of TDs assume their formation very early in the history of the Universe via the phase transition of GUT symmetry breaking. Examples of pos-

sible realization are cosmic strings [92] or magnetic monopoles [93] with very short lifetime, which are therefore required to be continuously produced in the TDs. The decay channels are quarks and leptons, where after the hadronization of the quarks, hadrons and leptons decay into photons, leptons and nucleons. These scenarios can result in several photons emitted per nucleon (depending on the decay model).

Z-burst model

This model's name reveals the involvement of the Z boson, which arises as a resonance in the interactions of UHE ν's with relic background neutrinos (clustered e.g. in the galactic halo), and then decays into a fermion/anti-fermion pair [94]:

$$\nu_{UHE} + \nu_{rb} \rightarrow Z \rightarrow f\bar{f} \,. \tag{2.6}$$

The resulting fermion/anti-fermion pairs can be hadrons, quarks, neutrinos or charged leptons (at different branching ratios) having energies in the UHE range. Also the later (sec. 2.6.2) mentioned energy limit E_{Hor} is below the energy of primary protons ($\sim 10^{24}$ eV) in the most popular Z-burst model [95]. Along with the UHECR production, UHE ν's and photons are generated too, whereat upper limits on the fluxes of neutrinos and photons at the highest energies 10^{21} eV strongly constrain this model [96, 97] (see also sec. 3.1 and 3.5.1).

Decay of heavy particles (Dark matter)

The models that consider SHDM are based on super-heavy metastable particles with lifetimes comparable to the age of the Universe, which would have been directly produced in the inflation epoch of the Universe. A hypothesized candidate could be the crypton, which is thought to exist in a hidden sector of a superstring-derived flipped SU(5) GUT model [98]. These particles are assumed to compose part of the cold dark matter (CDM) and in particular part of the dark halo of our Galaxy. Thereby these models avoid the GZK cut-off, since the produced flux of UHECRs originates close enough to the Earth. Decays or annihilations of those particles into UHECRs [99–102] would serve as a tracer of CDM, with predicted signatures such as a dominant photon fraction above the ankle region and an excess of UHECR events from the direction of the galactic halo (resulting in a large-scale anisotropy). The parameter space of these models is therefore almost totally excluded by the stringent upper photon limits and the large scale anisotropy measurement results by the Pierre Auger Observatory, see [61, 103] and sec. 3.1.

2.5. Potential sources of UHECRs

With solid theoretical consideration of mechanisms capable of accelerating charged particles up to the highest energies in the Universe, several astrophysical objects can be taken into account as potential UHECR source candidates while fulfilling certain requirements, however no object has been identified as source for UHECRs hitherto.

As already mentioned in sec. 2.4.1, several conditions must be met for sources to accelerate CRs to UHEs. The constraints that need to be satisfied are:

- **Time-scale constraint**:
  The ceasing of the acceleration of the particles occurs once the dynamical time-scale of the system becomes comparable to the acceleration time.

- **Cooling constraint**:
  The energy gain ceases when the particle can no longer be confined within the acceleration region (escape), or when various cooling processes (energy losses) are fast enough to inhibit further acceleration. The dominant energy losses for protons during acceleration are in most of the models due to the synchrotron radiation, as e.g. in [104]. Other processes, such as the photo-pion production or the curvature radiation, may also contribute to the energy losses.

As shown in sec. 2.4.1, both constraints may be met in the most powerful sources and some candidates have already been introduced in figure 2.8. Since injection mechanisms and/or outflow flux varies for different types of sources, and due to the lack of observational data on the inner mechanisms of the sources, each type of source candidate has to be studied individually. Only a few types of astrophysical objects are known to meet the criteria, with the most prominent ones briefly introduced below. Other astrophysical site cases are considered in detail in [81, 105]. To summarize the constraints:
Geometry: accelerated particle should be kept inside the source while being accelerated.
Power: the source must possess the required amount of energy to transfer to the particles.
Radiation losses: the energy lost by a particle as radiation in the accelerating field should not exceed the energy gain.
Interaction losses: the energy lost by interactions with other particles must be less than the energy gain.
Emissivity: the total number and power of sources must explain the observed UHECR flux.

Additional important criteria constraining the source models include requirements on the density of the sources and the produced UHECR flux to match the observed spectrum and composition on Earth. This in turn implies

constraints on the distance to the UHECR production sites, taking into account the energy losses during the propagation of the UHECRs to the Earth (see next sec. 2.6).

Supernova Remnants

Remarkable astrophysical objects, Supernova Remnants (SNRs) as sources for galactic CRs were discussed in 1934 by Baade and Zwicky [106].

Nowadays SNRs are widely considered to have a dominant contribution to CRs at least up to PeV energies. PeV energies are assumed to be reached in young SNRs before the phase in which its shell is swept off (called Sedov phase) and for very dense media and high masses. The ejected matter of a stellar explosion forms a shock front at which charged particles in the interstellar gas can be accelerated in stochastic acceleration processes through first-order Fermi acceleration to maximum energies in the PeV range [68]. Other calculations, have shown the possibility for iron nuclei to be accelerated even up to the EeV range in Type IIb SNRs [107]. From experimental side, the detection of the characteristic pion-decay feature in the γ-ray spectra of two SNRs, IC 443 and W44, with the Fermi LAT provided direct evidence that CR protons are accelerated in SNRs [108, 109]. Another experiment which helped to constrain acceleration mechanisms [110] and had important contributions to prove CR (re)acceleration in SNRs via the detected spectrum of the γ-ray emission is the AGILE (satellite) [111, 112]. AGILE in combination with data from H.E.S.S. also showed strong support for a hadronic model for the γ-ray production in W28 [113].

As an example of a SNR, more exactly a shell-type SNR, which are supposedly CR "PeVatrons", Cassiopeia A (Cas A) is shown in figure 2.9.

Active Galactic Nuclei

A galaxy is called an "active galaxy", if it hosts an AGN in its center - an active SMBH. Being the most luminous persistent sources of EM radiation in the Universe, the non-stellar radiation from an AGN is theorized to result from the accretion of matter onto the SMBH. Highly relativistic and confined jets of particles are the driver of particle acceleration and a common feature of these objects.

A $\sim 3\sigma$ correlation between the arrival directions of some 20 UHECR events above $5.6 \cdot 10^{19}$ eV with the positions of AGNs in the Veron-Cetty and Veron catalogue (with distances < 75 Mpc) have been announced by the Pierre Auger Collaboration [115, 116]. Nevertheless, this correlation with AGNs does not necessarily mean that they are the true sources, since any other type of objects following large-scale matter distribution, such as GRBs or clusters of galaxies, may also be a plausible UHECR source candidate. Furthermore, the high correlation level reported in [115] has probably been a result of a statistical

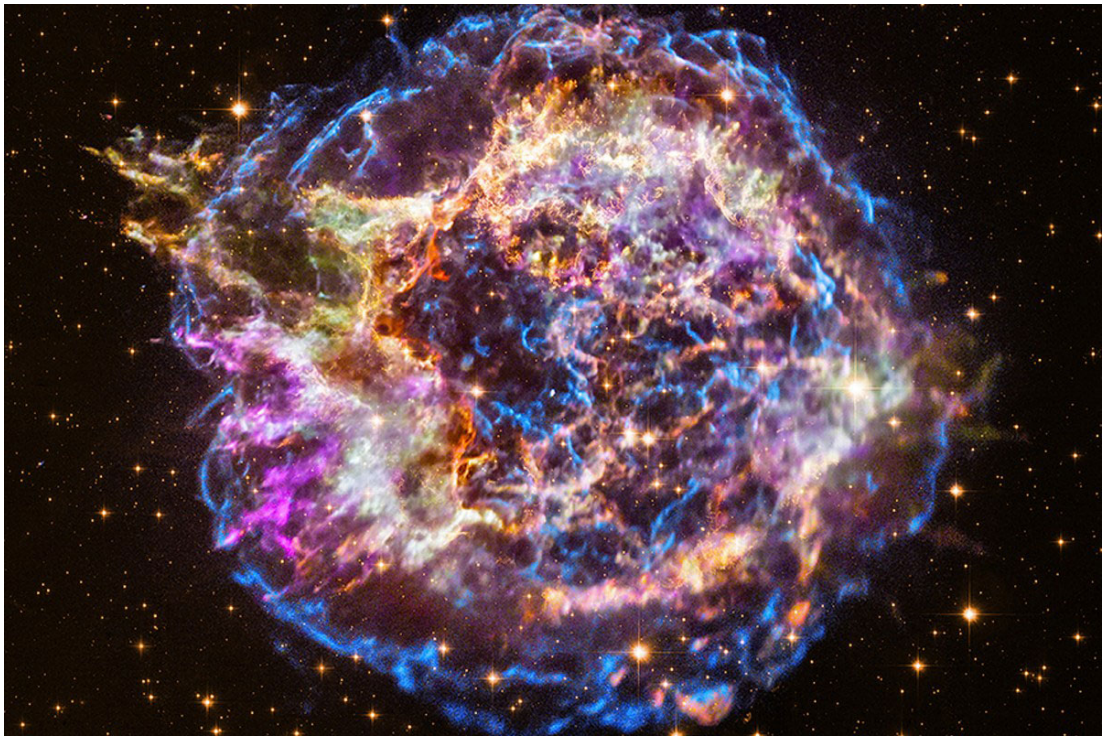


Figure 2.9.: This is a mosaic image of Cas A, composed of X-ray data from NASA's Chandra X-ray Observatory mission, infrared data from the Spitzer Space Telescope, and optical data from other telescopes [114].

fluctuation in the arrival direction distribution of UHECR events. There were also additional uncertainties concerning the inhomogeneity of the utilized AGN catalogue. The correlation with another selection of AGNs, gamma-ray-selected AGNs, has been recently confirmed at somewhat lower statistical level [117, 118]. However, as also pointed in [118], there is still a lack in understanding of the impact of bulk magnetic deflections, which renders any conclusions of a preferred UHECR source type at this stage preliminary.

Furthermore could particles be accelerated up to UHEs by flares that are generated when a star approaches a BH, gets disrupted by it and a jetted tidal disruption event (TDE) occurs [119].

The majority of AGNs are considered to be of the radio-quiet type, with no strong jets and thus unlikely as sources for UHECRs due to the severe radiative losses expected near the nucleus, and the absence of non-thermal emission components. In contrast roughly 10% of AGNs are known to be radio-loud, ejecting powerful, relativistic jets away from the nucleus [120] and are described in the next section.

Radio Galaxies

Although radio galaxies (RGs) are home to AGNs in their centres, they are considered a separate class of possible sources of particles with the highest energies.

This classification is due to the fact that the acceleration of particles reaching energies up to 10^{21} eV supposedly does not take place in the inner regions of the AGN, but rather in hot spots, i.e. regions of intense radiation, in the radio-lobes at the end of giant (spanning > 100 kpc), relativistic, collimated jets emanating from the AGN [121]. Historically RGs have been divided into Fanaroff-Riley (FR) I and II subclasses, depending on their morphology in the radio band, with FR-I jets being more diffuse and weaker. Recently, a new class of individually fainter, but much more numerous FR-0 RGs has been proposed [122].

Acceleration of particles to highest energies is achieved in FR class II via diffusive acceleration or shear acceleration (due to a velocity gradient) [123] and by termination shocks in first order Fermi acceleration [121].

Since these long jets reach out into the Interstellar Medium (ISM), where they create large MFs and the radiation field is much less dense than in the inner region of the AGN, energy losses are lower compared to the acceleration in jets, and no adiabatic deceleration is expected.

Especially hot spots of radio-loud, FR-II RGs [124], based on early observations of the synchrotron spectra emitted by these galaxies [125], are known to create the largest and most powerful shock waves in the Universe.

The well known candidates for this class of sources include Virgo A (M87, NGC 4486) and Centaurus A (NGC 5128, Cen A, shown in figure 2.10) – one of the closest RGs to Earth, which implies that emitted UHECRs might suffer smaller magnetic deflections. Indeed, the Pierre Auger Collaboration reported an excess of events in the direction of Cen A [126].

A recent work suggests that FR-0 RGs could contribute to the overall UHECR flux, being much more numerous than the more energetic FR-I and FR-II RGs [127].

Radio-loud galaxies inside the detectable volume (< 10 Mpc) represent interesting targets for the UHE photon (see next chapter) search with the Pierre Auger Observatory.

Figure 2.10.: This image of Centaurus A is composed of images obtained with three instruments, operating at very different wavelengths and thereby revealing the lobes and jets emanating from the active galaxy's central black hole: in orange the 870-micron submillimetre data, from LABOCA on APEX, in blue X-ray data from the Chandra X-ray Observatory, and visible light data from the WFI on the MPG/ESO 2.2 m telescope show the background stars and the Galaxy's characteristic dust lane in "true colour".

Pulsars

Another type of astrophysical object considered as potential source for CRs reaching UHE are highly magnetized rotating compact stars, such as neutron stars that emit beams of EM radiation out of magnetic poles [128].

The discovery of the first pulsar in 1967 at the radioastronomy laboratory in Cambridge, England was awarded to Antony Hewish with the Nobel Prize in 1974 "for his decisive role in the discovery of pulsars", together with Sir Martin Ryle "for his observations and inventions, in particular of the aperture synthesis technique" [129].

Merging processes of binary neutron star (BNS) systems produce highly turbulent MFs and potentially result in the birth of magnetars – young neutron stars hosting a particularly strong magnetic field, and able to accelerate charged particles to UHEs (see [130–132] and references therein). To be mentioned for the sake of completeness, other mergers of binary systems (BSs), also the merging of BSs of stellar mass black holes (BHs) are speculated – under appropriate boundary conditions – to be capable of accelerating CRs to high energies, but with less efficiency than BNS mergers [133]. A schematic representation of compact BS mergers and their outcome as a pulsar, an accreting BH with jets, or a BH can be seen in figure 2.11.

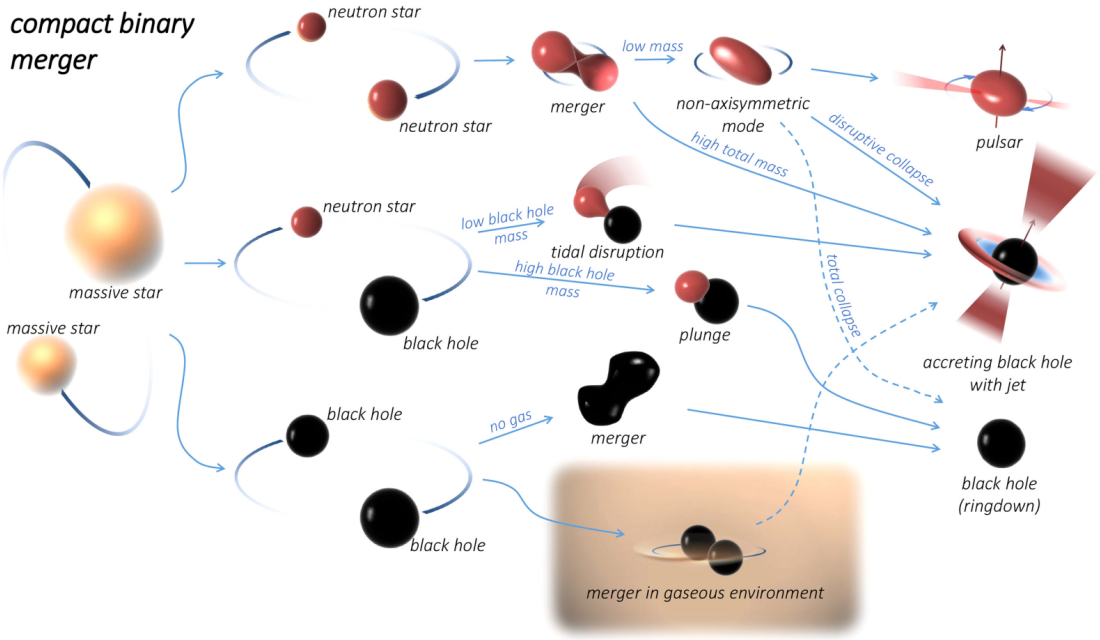


Figure 2.11.: Schematic depiction of compact BS mergers. Figure taken from [134].

The first binary pulsar ever discovered in 1974, the Hulse–Taylor binary, is a BS system composed of a neutron star and a pulsar (known as PSR B1913+16, PSR J1915+1606 or PSR 1913+16), used to indirectly confirm the existence of gravitational radiation. Russell Alan Hulse and Joseph Hooton Taylor, Jr. were awarded the 1993 Nobel Prize in Physics "for the discovery of a new type of pulsar, a discovery that has opened up new possibilities for the study of gravitation" [135].

Even more exotic, and the only known double pulsar is PSR J0737–3039, which consists of two neutron stars emitting electromagnetic waves in the radio wavelength in a relativistic BS [136].

A direct observation of the ultimate merging process of a BNS system has finally been achieved with the detection of a gravitational wave signal in 2017 by the Advanced LIGO and Virgo detectors [137].

As a classical pulsar example, the Crab Pulsar (PSR B0531+21) is shown in figure 2.12, which is accredited to be the first identified source of PeV electrons [138].

Figure 2.12.: The pulsar in the center of the Crab Nebula. Optical data from the Hubble Space Telescope (in red) are combined with X-ray images from Chandra X-ray Observatory (in blue).

Gamma-Ray Bursts

Belonging to transient source classes, in contrast to above mentioned continuous emitters of UHECRs, are flashes of gamma rays associated with the most

energetic explosions in the Universe, and only observed in distant galaxies hitherto, so-called GRBs. With durations of the bursts lasting from several milliseconds to several hours [139], multiple shock regions can occur, where different models put the dominant acceleration regions either to inner shocks [104] or outer shocks [140].

Long GRBs (LGRBs) are supposed to be the product of the collapse of massive stars, and are therefore considered as tracers of star formation, where a relatively recent work has provided more insights on the nature of LGRBs and also on the star formation process in the early universe based on comparisons of the observed abundance ratios in the interstellar medium with detailed chemical evolution models to constrain the nature and the age of LGRBs host galaxies [141].

Furthermore could GRBs account for the total flux of CRs at the highest energies, since the released energy in the form of CRs is similar to the energy emitted as γ-rays [104]. But since these explosions are extremely rare (a few per galaxy per million years [142]), it is difficult to accredit GRBs as direct source of UHECRs. These properties of transient character, together with the possibility of a delay of emitted UHECRs from a GRB arriving up to 10^7 years later at Earth than its emitted γ-rays [143], could in principle explain the lack of significant correlations of the arrival directions of UHECRs with astrophysical objects.

The most prominent model to describe GRBs is the so-called fireball model [144, 145], which enables protons to be accelerated up to the highest energies [146].

Clusters of Galaxies

Known as the largest gravitationally bound systems in the universe, clusters of galaxies, composed of stars, hot gas and dark matter, are, according to the currently favoured picture of hierarchical structure formation in the CDM cosmology, the latest objects to form. They are assumed to be surrounded by strong accretion shocks as a consequence of continuing infall of Dark Matter (DM) and baryonic gas [147], with these shocks proposed as sources of UHECRs [148]. Cosmological hydrodynamic simulations consistent with the observed CR flux suggest that protons can be accelerated up to $\sim 6 \cdot 10^{19}$ eV [149] and nuclei with higher Z having correspondingly shorter acceleration time so that Fe may be even accelerated up to 10^{20} eV [150, 151]. Thereby the escape of CRs from clusters may be possible via energy-dependent escape upstream of the shock, or by diffusion in directions away from the filaments, or perhaps by advection in partial outflows during merging events [120].

2.6. Propagation of Cosmic Rays

Cosmic-ray particles are assumed to propagate from their site of creation in a diffusive process through galaxies and intergalactic medium, being deflected many times by the randomly oriented MFs, interacting with radiation and undergoing various other processes of energy loss. The identification of extrasolar sources of CRs is complicated, since electrically charged particles are being deflected by MFs and do thus not point back to their site of production or acceleration. As a consequence of the widely accepted extragalactic origin of UHECRs (at least above the ankle of the cosmic-ray spectrum), the measured cosmic-ray spectrum on Earth has to be shaped by the effect of the propagation of the particles in the extragalactic medium.

Since evidence suggests that CRs at UHEs are created outside the Galaxy but within $50 - 100$ Mpc, the gyro-radius for a particle in the galactic MF (GMF) with $E \gtrsim 10^{19}$ eV is larger than the Galaxy, and for sources of UHECRs within the Galaxy, the arrival direction would point back to the source. This follows from the expression for the gyroradius (also called Larmor radius) r of a particle with charge number Z and energy E propagating through a magnetic field of strength B [55]:

$$r[\mathrm{pc}] \;=\; 1.08\,\frac{E\,[\mathrm{PeV}]}{Z\cdot B\,[\mu\mathrm{G}]}\;. \tag{2.7}$$

Extensive simulations [152, 153] of the propagation of CR nuclei from the local supercluster (LSC) of galaxies have shown that for a LSC field of 10^{-7} G CRs with energies $< 10^{19}$ eV undergo a random walk as they travel from source to observer, while those $> 10^{20}$ eV follow a relatively straight path (further discussion in sec. 2.8). Also in [154] it is shown that at the highest energies, the angular deflections of CRs from the sources could be as low as just a few degrees.

This section summarizes the processes influencing UHE nuclei on their propagation between their creation and their final detection on Earth, while a more detailed description can be found e.g. in [155, 156], and an overview of UHE photon propagation is presented in section 3.3.

2.6.1. Deflection in magnetic fields

Another fascinating and challenging quest in modern astrophysics and directly connected to cosmic-ray research is the origin of galactic and extragalactic MFs. Beside detections of MFs in galaxies of all types and in galaxy clusters (with appropriate observations), there is mounting evidence that these fields exist in galaxies at cosmological redshifts (denoted with z), where they are characterized by a modest strength of $10^{-7} - 10^{-5}$ G and huge spatial scale of $\sim$ Mpc.

Estimates of the average strength of extragalactic MFs (EGMFs) are of the order of 10^{-9} G (1 nG) [157]. Even larger MFs have been observed in clusters of galaxies, but already fields of nG strength would affect the propagation of UHECRs.

In our Galaxy a MF, besides its large scale component oriented along the spiral arms, is assumed to be randomly oriented, having an strength of $B \sim 3\ \mu$G and extending to the galactic halo as well. A residence time in the Galaxy for particles with GeV energies has been found to be of about $15 \cdot 10^6$ years [158], and with the ISOMAX detector the scale height of the halo has been estimated via the ^{10}Be/^{9}Be-ratio to be a few kpc [159].

MFs as observed in disk galaxies, are generally assumed to be amplified and maintained by an $\alpha\omega$-dynamo [155], wherein the combination of helical turbulences and differential rotation is continuously regenerating the field. On the contrary appears the MF in non-rotating or slowly rotating systems (such as elliptical galaxies and clusters) to have a characteristic coherence scale much smaller than the size of the system itself and may be generated by a local, turbulent dynamo. In the absence of rapid rotation, the field does not organize on large scales. Nevertheless must the dynamo paradigm be considered incomplete, since it does not explain the origin of the initial fields that act as seeds for subsequent dynamo action. Moreover, the standard $\alpha\omega$ dynamo is not capable to explain the existence of MFs in elliptical galaxies and clusters.

The combination of radio and X-ray measurements led to the discovery of MFs along stacked Cosmic Filaments (CFs) [160] which connect the largest gravitationally bound structures in the Universe, galaxy groups and clusters.

Cosmological MFs, including those that existed prior to the epoch of galaxy formation (called primordial), could be imagined as a field that is essentially uniform across our Hubble volume. At present it is not known whether cosmological MFs exist, and a detection of sufficiently strong cosmological fields would not just provide tremendous support to the primordial field hypothesis but furthermore open a new observational window to the early Universe and in addition – since very weak cosmological fields can act as seeds for the galactic dynamo – help complete the dynamo paradigm [155].

Further details about MFs in galaxies, clusters, superclusters, and beyond are described in [155, 161].

The four most common methods used to study astrophysical MFs are briefly introduced below and a more thorough discussion of observational techniques can be found for example in [156].

Polarization of Optical Starlight

Polarized light from stars can reveal the presence of large-scale MFs in our Galaxy and those nearby. To mention just a few milestones, it was after first observations of polarized starlight by Hiltner [162, 163] and Hall [164] in 1949,

that the existence of a GMF as a means of confining CRs was proposed by Alfvén [165] and Fermi [166] in 1949 too. Later, in 1951, a connection between polarized starlight and a GMF was made by Davis and Greenstein [167], suggesting that elongated dust grains would have a preferred orientation in a magnetic field. However, polarization of optical starlight has limited value as a probe of EGMFs, since this method is not sensitive enough even for distant stars in the Milky Way [156, 168]. As with polarimetry of the synchrotron radiation, and in contrast to Faraday rotation measure (FRM) imaging, optical polarimetry does not convey information about the field sign [169].

Synchrotron Emission

This radiation, which is produced by relativistic electrons spiralling along magnetic field lines, can be used to study MFs in astrophysical sources ranging from compact objects like pulsars (sec. 2.5) to superclusters.

One of the two primary estimates for the strength of MFs in galaxies and clusters is the degree of polarization, which is an important indicator of the field's uniformity and structure, while the other, the total synchrotron emission depends on the energy distribution of electrons, $n_e(E)$. For the latter a commonly used class of models, assumed to be valid over some range in energy, is based on a power-law distribution

$$n_e(E)dE \;=\; n_{e0}\left(\frac{E}{E_0}\right)^{-\gamma} dE \tag{2.8}$$

with the constant $n_{e0} \equiv n_e(E_0)$ setting the normalization and a spectral index $\gamma \simeq 2.6 - 3.0$ typical for spiral galaxies [155].

Faraday rotation

Astrophysical measurement of MFs can be estimated from rotation measures (therefore often called Faraday rotation measures (FRMs)) given a knowledge of the electron number density. When EM waves propagate through a region of both MF and free electrons they experience Faraday rotation wherein left and right-circular polarization states travel with different phase velocities.

Radio pulsars are well suited sources where the dispersion caused by these electrons results in a time delay between pulses received at different wavelengths, which can be measured in terms of the electron column density, or dispersion measure. A combination of measurements of the dispersion measure and the rotation measure therefore yields the weighted mean of the magnetic field along the line of sight, and can be achieved also in other astrophysical objects, if their dispersion measure can be estimated based on reasonable guesses about the propagation path length and typical electron densities.

Furthermore, in the case of extended sources, will Faraday rotation lead to a decrease in the polarization, called Faraday depolarization, since the com-

bined signal from waves originating in different regions of the source will experience different amounts of Faraday rotation and thereby leading to a spread in polarization directions. Faraday depolarization can be a used to measure MFs in the foreground of a source of polarized synchrotron emission.

FRMs helped to establish that the MF lines follow the spiral arms of the Galaxy. The largest CR magnetic deflections occur, when charged CRs propagate close to the GC. Apart from the GC, the average deflection angle for 10^{20} eV protons is between 3.1° and 4.5° in different galactic field models and for $2 \cdot 10^{18}$ eV protons, it is in the range $17.7° - 25.9°$ [170].

Zeeman Splitting

Beside that Zeeman splitting is providing the most direct method available for observing astrophysical MFs, it is also of historical importance since it led to the first known example of extraterrestrial MFs via the discovery of MFs in sunspots [171].

While synchrotron emission and Faraday rotation probe the line-of-sight MF, Zeeman splitting is sensitive to the regular MF at the source, but is unfortunately extremely difficult to observe. The two most common spectral lines in Zeeman-effect observations are the 21 cm line (for neutral hydrogen (H)) and the 18 cm OH line (for molecular clouds). Detections have been restricted to regions of low temperature and high MF and in this way Zeeman effect measurements within the Galaxy helped to provide information on the MF in star forming regions and near the GC.

2.6.2. Energy loss processes

A very important aspect in the propagation studies of UHECRs is the modelling of their energy losses and the change of mass composition due to various interactions. In the extragalactic medium, except maybe in the immediate vicinity of the UHECR accelerators, mainly interactions with photon backgrounds are relevant, where protons and nuclei interact mainly with CMB photons which represent the densest photon background, as well as with infrared (IR), optical and ultra–violet (UV) photons, categorized together as CBR.

All the UHECR production models discussed in sec. 2.4.1 need to consider various energy losses, already on site of the source, which might vary for different CR nuclear species.

To mention just a few mechanisms, potential energy losses are due to collisional and inverse Compton scattering (ICS) in a radiation field. An energetic particle will also loose energy radiatively while moving through magnetic and electric fields which confine and accelerate this particle, radiative synchrotron losses, pair production (Bethe-Heitler processes [172], also see 4.2.1)

and Bremsstrahlung (more in sec. 4.1.1) [86, 95]. Also photohadronic interactions could impact the particle motion and contribute to energy losses, via GZK-like processes, already at the acceleration site. For example as shown in [95] will a proton with energy larger than $\sim 7 \cdot 10^{22}$ eV lose its energy through the interaction with the 2.7 K CMB photons right at the acceleration site. Furthermore is the ultimate upper threshold on the energy of electromagnetically accelerated protons in [95] derived to be $E_{Hor} \sim 4 \cdot 10^{23}$ eV, and the size of the accelerator would exceed the size of the Universe for energies above.

While here only an overview of the main energy loss processes during UHECR propagation is presented, a more detailed description of the interactions on the CRs' voyage through the Universe can be found e.g. in [173].

Three of the below discussed processes, namely pair production (PP), pion photoproduction and Photodisintegration (PD), all produce high energy photons. The fraction of the energy transferred to EM channels in general increases with the CR energy. For iron nuclei with energies between 10^{18} eV and 10^{21} eV, this fraction ranges from a few 10^{-6} to a few 10^{-1} [173].

Photo-meson production and GZK-effect

Photo-meson production is the most common interaction, and for the most prominent channel often called photo-pion production. This interaction leads to one of the most important features of the extragalactic propagation of UHECRs consisting in a cut-off in the energy spectrum above a few 10^{19} eV due to interactions of the charged CR particles with photons of the CBR [35, 36]. This has also triggered a long series of studies on the extragalactic propagation of UHECRs including the production of secondary neutrinos and photons.

For protons with energies above a few 10^{19} eV in their rest frames encountered thermal photons are highly blue-shifted and therefore the energy of the CMB photons γ_{CMB} is sufficient to excite baryon resonances that subsequently leads to pion production. This is the so called GZK cut-off, resulting as the threshold of pion production in the interaction of cosmic-ray protons with CMB photons. At a few tens of EeV the cross section for this process is dominant, leading to the so-called GZK-suppression of the UHECR flux via the $\Delta^+(1232)$ resonance:

$$p + \gamma_{\text{CMB}} \quad \rightarrow \quad \Delta^+(1232) \quad \rightarrow \quad n + \pi^+ \tag{2.9}$$

$$\rightarrow \quad p + \pi^0 \quad , \tag{2.10}$$

with n the neutron and p a proton (with branching ratios of about 2/3 for eq. (2.10) and 1/3 for eq. (2.9), which are determined by isospin and are resulting from the relevant Clebsch-Gordon coefficients [174, 175]). Generally speaking, there are other "GZK-type processes" with nucleons N

$$N + \gamma_{\text{CBR}} \quad \rightarrow \quad N + \pi \tag{2.11}$$

But also other baryon resonances can occur with increasing energy:

$$p + \gamma_{\mathrm{CBR}} \quad \rightarrow \quad \Delta^{++} + \pi^{-} \quad \rightarrow \quad p + \pi^{+} + \pi^{-} \,, \tag{2.12}$$

where Δ^{++} denotes e.g. the $\Delta(1620)$ or $\Delta(1700)$ resonance. The corresponding threshold energy E_{th}, for the nucleon to produce pions assuming head-on collision can be found to be

$$E_{th}^{N,\pi} \;=\; \frac{m_\pi c^4 \,(2m_N \,+\, m_\pi)}{4\varepsilon} \;\simeq\; 6.8 \cdot 10^{19}\ \mathrm{eV} \,, \tag{2.13}$$

where ε is the energy of a CMB photon $\varepsilon \simeq 10^{-3}$ eV, m_π the pion mass and m_N the nucleon mass.

This GZK process will affect a proton with sufficiently high energy, leading to losses of $\sim$ 20% of its energy every about 5-10 Mpc until ultimately its energy falls below $E_{th}^{N,\pi}$, resulting in the maximum distance a proton can travel (maintaining an energy above the threshold) of $\sim$ 100 Mpc [173], which is called the *GZK horizon*. Thus, as stated earlier, any protons observed with higher energies must have been accelerated in a more nearby source [176].

The neutral pions created in process eq. (2.10) will further decay to two high energy (HE) photons (due to the law of conservation of energy and momentum),

$$\pi^0 \quad \rightarrow \quad \gamma_{\mathrm{HE}} \,+\, \gamma_{\mathrm{HE}} \,, \tag{2.14}$$

travelling a straight path afterwards (more in sec. 3.2). Photo-meson production is the main production channel of UHE photons and neutrinos by hadronic CRs [173] and therefore would the observation of UHE photons be a key signature and an independent confirmation of the GZK effect [74, 177]. Results of the Pierre Auger Collaboration about composition favour a flux suppression caused by the maximum energy achievable by the sources [70], while the TA collaboration observed a proton dominated composition at the highest energies which in turn favours the GZK scenario [178].

It should be noted that with the CMB being represented by a Planck distribution even lower GZK effect threshold energies are possible. Also other extragalactic background lights, e.g. infrared or radio background are affecting the propagation of UHECRs as can be seen in figure 2.14 a) for protons and b) for iron.

Pair-production (Bethe–Heitler)

Another important process of energy loss of CRs is the $e^{\pm}$ pair production (PP) by a nucleus X with mass number A and atomic number Z on a photon (see fig. 2.13), which becomes important at energies below the GZK-suppression and is also known as Bethe-Heitler process:

$$^{A}_{Z}X + \gamma_{\mathrm{CBR}} \quad \rightarrow \quad ^{A}_{Z}X + e^{+} + e^{-} \,. \tag{2.15}$$

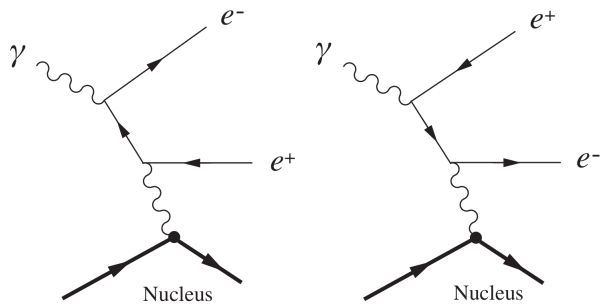


Figure 2.13.: Schematic depiction in Feynman diagrams of the PP process in lowest order. Figure taken from [179].

In Feynman diagrams such as shown on fig. 2.13, time runs from left to right, while a thin solid line with its arrow pointing to the right (left) indicates an electron (positron). Compared to the photo-pion production, the threshold energy here is two orders of magnitude smaller,

$$E_{th}^{N,e^{\pm}} = \frac{m_e c^4 \left(m_X + m_e\right)}{\varepsilon} \simeq 4.8 \cdot 10^{17} A \text{ eV} , \tag{2.16}$$

where m_e the electron/positron mass. The produced $e^{\pm}$ pair can subsequently interact via ICS to produce high-energy secondary photons and thus PP is an indirect photon production channel and is especially important for producing secondary photons below PeV energies.

Photodisintegration

In the process called Photodisintegration (PD) a UHE nucleus (not a proton) interacts inelastically with a CMB photon, leading to partial fragmentation of the nucleus, potentially leaving the residual nucleus in an excited state, which leads to the emission of one or more HE photons.

However, due to the sparsity of the available PD measurements there are large uncertainties in modelling of this process which is especially important for describing nuclei propagation at ultra-high energies. The relevant energy losses vary with the CR species and are given for comparison in figures 2.14 a) for proton and b) for iron. The total energy loss length χ_{loss} of different species is shown in figure 2.15.

In the example of protons first above $\sim 5 \cdot 10^{17}$ eV pair-production starts taking over the energy loss, while above $\sim 7 \cdot 10^{19}$ eV photo-pion production becomes dominant, so that above 10^{20} eV χ_{loss} falls to less than 100 Mpc, which strongly limits the propagation of UHE protons.

Heavier nuclei experience various PD channels, in which the CR nucleus loses some of its nucleons, showing threshold energies varying with the mass of the nucleus. The evolution of the contribution to the energy loss of PD and PP is shown in fig. 2.14 b).

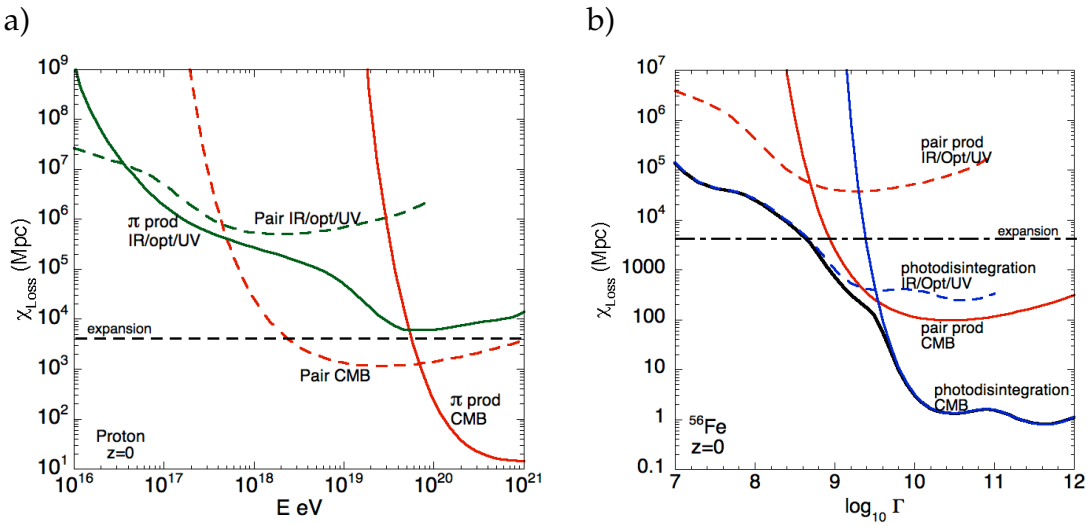


Figure 2.14.: (In the nucleus rest frame) **a)** Contributions of different energy loss processes to the energy evolution of the energy loss length χ_{loss} of protons. **b)** Contributions of PP and PD to the total attenuation length of iron nuclei with evolution of the logarithm of the Lorentz factor on the abscissa. Both figures taken from [180].

As can be seen in figure 2.14 b) for nuclei, the dominant energy loss processes above 10^{19} eV during intergalactic propagation are PD and PP interactions with photons of the CMB [181].

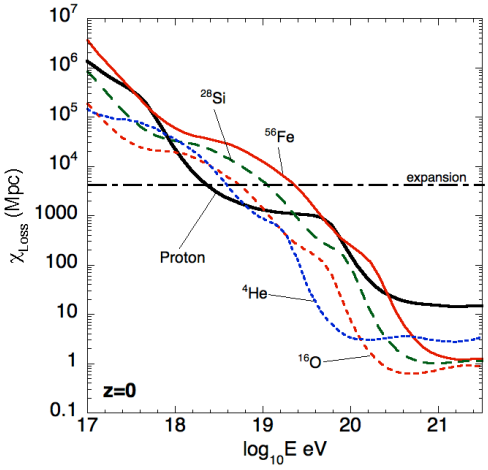


Figure 2.15.: Energy loss lengths χ_{loss} of different species. Figure taken from [180].

Adiabatic energy loss

This process of adiabatic fractional energy loss is the least important mechanism described herein, but it dominates near and below the pair-production threshold, and is redshifting due to the expansion of the universe. It is described as

$$-\frac{1}{E}\left(\frac{dE}{dt}\right)_{adiabatic} = H_0\,, \tag{2.17}$$

whit H_0 the present Hubble constant. Figures 2.14 a) and b) and 2.15 show estimates of this effect on the energy loss denoted as "expansion". All other loss processes are negligible for the propagation, although are important in very dense central CR source regions (e.g. at the site of acceleration).

2.7. Mass composition

The determination of the exact chemical composition of CRs up to the highest energies is a further important aspect in answering the open questions about their nature and origin.

Up to $\sim$ 100 TeV the abundance of individual elements has been measured directly with detectors above the atmosphere, and all the elements of the periodic table have been found in cosmic rays [55]. Furthermore have antiparticles (anti-protons and positrons) also been found in CRs, but are extremely rare compared to the particle fluxes. However, their abundance is well consistent with the expectations of anti-particles generated in collisions of CRs with interstellar dust and thereby ruling against the hypothesis of possible large amounts of anti-matter in the observable universe [43]. Differences to the composition of stellar material in our solar system are relatively small, as shown in figure 2.16 as function of the nuclear charge number. However,

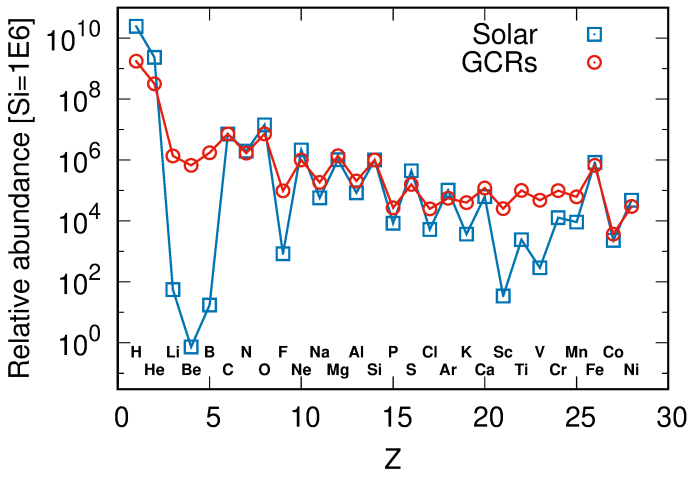


Figure 2.16.: The relative abundances of elements in the solar system and in low energy galactic CRs (up to nickel (Ni)), normalized to a value of 10^6 for silicon (Si). Figure taken from [116].

some important differences are the overabundance of hydrogen (H) and helium (He) for solar system, while lithium (Li), beryllium (Be) and boron (B) are overabundant in CRs. The cosmic-ray abundance of iron (Fe) agrees quite well with solar system composition, but there is an excess of elements slightly lighter than Fe. Beside revealing information about the acceleration of CRs

this difference can be understood assuming an initially equal composition as of the solar matter at the CR origin. During their propagation through the interstellar medium, CRs interact with gas and dust particles, which results in heavier nuclei spallating into lighter nuclei via collisions with the elements of the CNO-group (carbon (C), nitrogen (N), oxygen (O)). Moreover might the local under-abundance of H and He reflect the chemical composition of their respective sources, and their large ionization energy pose an obstacle for their injection into the initial acceleration mechanism [116]. Knowledge about the spallation cross sections can be obtained in collider experiments, and can be used to determine the amount of matter traversed by CRs between sources and Earth.

Being composed of predominantly H at the lowest energies at around 1 GeV, the CR composition turns heavier with increasing energy, showing an hardening of the spectrum at around 100 GeV, which is more pronounced for heavier nuclei, resulting in He being more abundant than H at around 1 PeV and beyond the knee even heavier elements in the mass range of Fe start taking over the CR spectrum [116], see fig. 2.17.

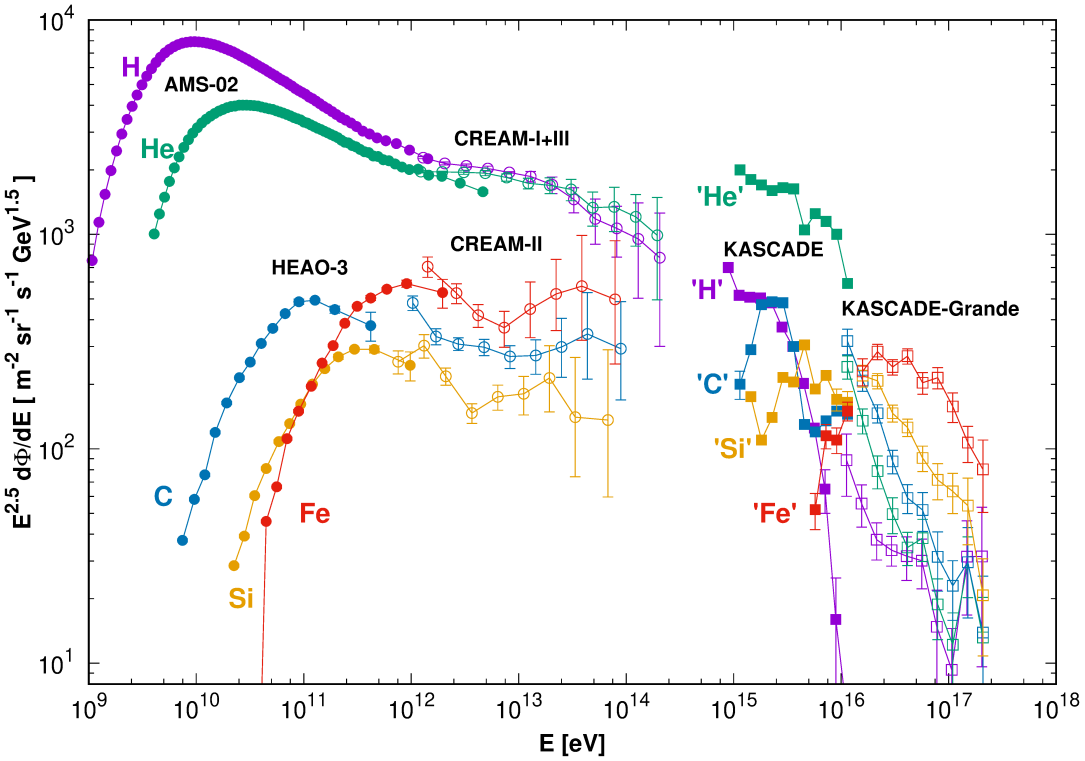


Figure 2.17.: The CR spectra for individual primary particle types as measured by various experiments for energies up to and above the knee, including the uncertainties of the hadronic interaction models that were used to perform the simulations on which the analysis is based. Figure taken from [116].

Due to the significantly decreasing statistics with increasing energy, in-

direct detection via air shower experiments with large apertures are so far used at $\gtrsim$ 100 TeV for an analysis of the average mass composition with progress towards event-by-event analyses. Since this highest energy range (above 10^{17} eV) is not yet accessible by man-made accelerator or collider experiments, air shower physics (more details in Chap. 4) must rely on simulations of different CR nuclei primaries. Beside the strong dependence on extrapolations of hadronic interaction properties in air shower simulations, additionally there is a large variance between the hadronic interaction models and thereby relevant systematic uncertainties in the experimental mass composition studies of UHECRs. Current hadronic interactions models lead to a deficit in the number of muons N_μ in the simulated showers compared to data [182], with inconsistencies occurring for energies $> 10^{16}$ eV. For the data measured with the Pierre Auger Observatory, the larger average N_μ in data than in simulations would correspond to a very heavy composition above 10^{19} eV, containing elements heavier than iron, which is in disagreement with other mass-composition studies. Experimental upgrades, as the *AugerPrime* (c.f. sec. 5.4), together with more refined analyses are expected to further reduce the uncertainties of the mass-composition related experimental results.

At energies approaching the ankle a trend towards a lighter composition, and above the ankle towards a heavier composition has been shown by recent results from the Pierre Auger Observatory, as shown in figure 2.18 [71].

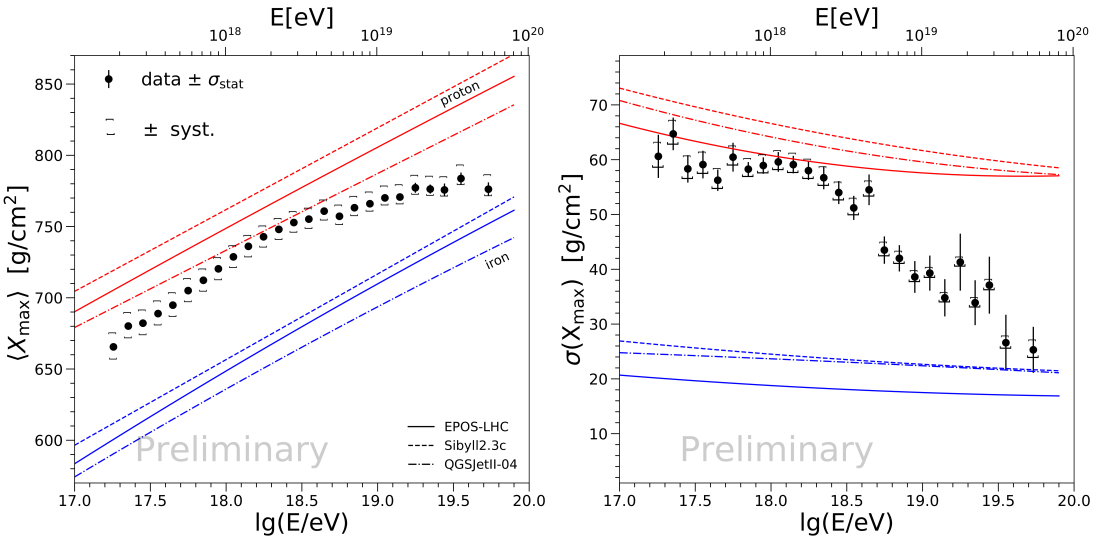


Figure 2.18.: The Pierre Auger Observatory's measurements of $\langle X_{\max} \rangle$ (**left**) and $\sigma(X_{\max})$ (**right**) compared to the predictions for proton and iron nuclei of three different hadronic interaction models. The hadronic interaction models QGSJETII-04 [183], EPOS-LHC (V3400) [184], and Sibyll 2.3c [185] are commonly used to simulate air showers and are all based on the simple parton model associated with the Gribov-Regge multiple scattering approach, further explained in their respective references as well as in [186]. Figure taken from [71].

The first and second moments of the distribution of the depth of the air shower maximum X_{max} are used to constrain the UHECR mass composition (see fig. 2.18). This observable, which represents the atmospheric depth of the maximum energy deposit by shower particles, and measured as a atmospheric column density in units of $\mathrm{g/cm^2}$, is explained in more detail in ch. 4 while introducing the Heitler model.

From the analysis of X_{max}, it follows that the mean mass of the UHECRs is getting lighter up to $10^{18.3}$ eV and is becoming heavier afterwards. The spread of the masses $\sigma(\ln A)$ is becoming smaller up to the ankle ($E \sim 10^{18.7}$ eV) and more constant at higher energies. The spread of the masses near the ankle is compatible with $0.85 < \sigma(\ln A) < 1.6$. For energies below the ankle ($10^{18.5} - 10^{18.7}$ eV), the data allows to exclude pure and proton-helium mixed compositions.

2.8. Anisotropy

A complementary information about the search for anisotropies in the arrival direction of CRs on different angular scales can contribute to the understanding of the cosmic-ray origin, in particular to the identification of source regions or individual sources. Small-scale anisotropies would directly hint towards cosmic-ray sources. Furthermore are galaxies in our local neighbourhood distributed inhomogeneously. This should lead to an observation of the large angular scale anisotropy in the CR arrival direction distribution, if the CR sources are distributed similarly to the matter in the Universe. One significant aspect to be taken into account in anisotropy studies is that during CR propagation in galactic and extragalactic MF, anisotropies are reduced, depending on the mass composition and energy of the CR particles.

At lower energies, below PeV the already mentioned dipolar anisotropy could possibly be due to Compton-Getting effect (cf. page 13), but the much smaller observed amplitude of the dipolar component indicates that the cosmic rays actually co-rotate with the local stars [187] and the orbital motion of the Earth around the Sun leads also to a dipolar anisotropy in the CR flux which is a function of the sidereal time, and can then be observed when analyzed in the solar frequency, as reported in [188].

For energies above the knee more precise anisotropy measurements would be important to better understand the origin of the knee and of the second knee, especially if they could be performed separately for the different mass groups.

The results of the Pierre Auger Observatory [62], cf. figures 2.5 and 2.4 together with the values obtained at few PeV energies by the IceCube [189], IceTop [190], and KASCADE-Grande [191] (also in fig. 2.4) experiments suggest that the transition between a predominantly Galactic and an extragalactic origin for the dipolar anisotropies is taking place somewhere around 1 EeV.

Moreover, may (though at smaller amplitude than observed at the highest energies) anisotropies on large angular scales, such as a dipole or a quadrupole, be present at all CR energies. Possibilities for the origin of the large-scale anisotropy lie in the anisotropic distribution of the CR sources themselves, or may originate from the characteristics of the CR propagation (quasi-rectilinear or diffusive), e.g. resulting from individual extragalactic CR sources in case the MFs are strong enough so that the propagation is diffusive. Large-scale anisotropies could also be produced at lower energies by CRs from Galactic sources, as they escape from the Galaxy [62].

Since the predictions of the UHECRs deflections in magnetic fields are uncertain, a high statistics of events is essential when searching for anisotropies in a sky map. Regardless the origin of the suppression in their flux and their nature, UHECRs are expected to come from sources relatively close to the Earth, where the galaxies are distributed non uniformly. On small angular scales anisotropy signals may occur, even if low-charge particles were to contribute only a fraction ($\sim 10\%$) of the primary CRs, as the gathered statistics increases with time.

As discussed already in the context of the CR gyroradius, see eq. (2.7), the identification of small-scale anisotropies in the distribution of arrival directions, such as clustering or localized excesses, would be indicative of the CR source regions and may become possible at energies $> 10^{19}$ eV, especially if the sources for UHE protons are located within the GZK horizon, i.e. relatively close to the Earth on cosmological scales.

The arrival distribution of CRs at the highest energies has shown to be remarkably isotropic, at odds with the original expectations which assumed only a few sources of CR protons contributing at these energies. Furthermore the lack of small scale anisotropies, with just some hints of localized flux excesses at intermediate angular scales, implies the significant effect of the Galactic and/or extragalactic MFs on UHECRs trajectories, as expected in mixed composition scenarios [116]. Thereby the arrival direction distribution of heavy CR nuclei up to the highest energies would be significantly spread out by EGMFs with amplitudes in the nG range, even for nearby extragalactic sources, leading to a washing out of the small scale anisotropies while still leading to significant large scale anisotropies.

Additional CR propagation aspects influencing the magnetic deflections of the particles shall be briefly discussed. As the direction of the particles will be changed after travelling a field coherence distance l_c, the CRs will perform a random walk through the different field domains leading to a diffusive spatial propagation, if the Larmor radius r_L is smaller than the coherence length l_c. In the case of isotropic diffusion a scalar diffusion coefficient D is sufficient for describing the increase of the mean square distance of the particles from the original point with time. For further characterization of particle diffusion, in Galactic (G) and extragalactic (xG) MF (B), the critical energy E_c is defined

via $r_L(E_c) = l_c$, hence $E_c = ZeBl_c$, and thus ([116])

$$E_c^G \simeq 80\,Z\,\frac{B}{3\,\mu\mathrm{G}}\,\frac{l_c}{30\,\mathrm{pc}}\,\mathrm{PeV} \quad , \quad E_c^{xG} \simeq 0.9\,Z\,\frac{B}{\mathrm{nG}}\,\frac{l_c}{\mathrm{Mpc}}\,\mathrm{EeV} \quad . \tag{2.18}$$

E_c separates the regimes of resonant diffusion ([192], deflections due to interactions with the MF) at lower energies, in which particles have large deflections caused by MF modes with scales comparable to the Larmor radius, and the non-resonant regime at $E > E_c$ in which the deflections after traversing a distance l_c are small.

From the already mentioned (fig. 2.5) observation of a significant dipole, combined with the lack of significant anisotropies at small angular scales, non-negligible effects on UHECR trajectories can be deduced. In [161] the root mean square (rms) deflection of a particle of charge Z and energy E coming from a distance L through a homogeneous turbulent MF with rms amplitude B and coherence length l_c has been shown to be

$$\delta_{\mathrm{rms}} \simeq 30^\circ \left(\frac{B}{\mathrm{nG}}\right)\left(\frac{4\,Z\,\mathrm{EeV}}{E}\right)\sqrt{\frac{l_c}{\mathrm{Mpc}}}\sqrt{\frac{L}{10\,\mathrm{Mpc}}}\,. \tag{2.19}$$

As an example for CRs in the non-resonant regime ($E > E_c$, $L \gg l_c$), plugging into eq. (2.19) values for oxygen nuclei with 30 EeV coming from a distance $L \simeq 10$ Mpc results in deflections by about 30° for an EGMF B of 1 nG, which is consistent with the bounds from CBR and Faraday rotation measures.

From these considerations and reasonable values for the extragalactic magnetic field in the LSC, $Bl_c \sim \mathrm{nG}\sqrt{Mpc}$ it follows that:

- protons may then propagate diffusively up to EeV energies from the closest sources, and up to higher energies for sources farther away, and
- the propagation of heavier nuclei will have similar behaviour, shifted up in energy by a factor Z, i.e. for the same rigidities.

Following the deflections in the intergalactic space, additional deflections due to the GMF (see sec. 2.6) will further affect the observed images of the CR sources.

To wrap this up, the general result of the diffusion effects is a suppression of the CR flux at low rigidities, since particles are not able to arrive to the observer from distant sources and take a much longer time than in the case of rectilinear propagation to arrive from the nearby ones – which is known as the "magnetic horizon" effect. Although this effect limits the distance from which particles from a given source can arrive, it can be shown with the help of the propagation theorem [193] that even at energies for which far away sources do not contribute, the observed energy spectrum will not change significantly. Only when the CR flux from the nearest sources gets suppressed, the overall spectrum will be modified, as long as the nearby sources lie at a distance

from the observer smaller than the diffusion length. It follows that as long as the distance to the nearest sources is smaller than other relevant length scales (diffusion length and the energy loss length χ_{loss}, see sec. 2.6), the total CR flux from sources of similar strength will be the same as that expected ignoring MF effects and for a continuous distribution of sources.

In the case of the dominant energy losses coming from the adiabatic expansion of the Universe, i.e. for $E/Z < 1$ EeV (see fig. 2.15), the spectrum will be suppressed as long as the propagation time is larger than the Hubble time, or for energies smaller than E_H given by

$$E_H \simeq 0.4\, Z \frac{B}{\mathrm{nG}} \sqrt{\frac{l_c}{\mathrm{Mpc}}} \frac{d_s}{40\,\mathrm{Mpc}}\,\mathrm{EeV}\,, \tag{2.20}$$

with d_s being the average distance between UHECR sources (for $E > E_c^{xG}$) [116]. The magnetic horizon effect is expected to strongly suppress the flux below this energy, getting more intense for decreasing energies, resulting in a hardening of the CR spectrum.

Further CR anisotropy-related details and reviews of observations can be found e.g. in [116, 194].

2.9. Practical application of cosmic rays

After the immense importance of CRs for fundamental research was pointed out, some comments on practical applications of cosmic-ray physics have to be added here for the sake of completion.

Since CRs can cause genetic defects, they can potentially be helpful or harmful. On the positive side, they can be understood as a driver of evolution. On the other hand, organisms travelling through space are unlikely to survive cosmic irradiation, since CRs are a potentially harmful sterilizer. The CR studies are thus vital for the radiation protection of the people travelling on spaceships and airplanes. The space weather impacts the flux of low energy charged particles at Earth, as e.g. manifested by the Northern Lights, which has to be accounted to protect the people, electric circuits, etc. during geomagnetic storms.

As described, they are a tracer of cosmological events, and even to be a source of energy they can be accredited as, for example, some UV energy can be captured by solar cells. Via the use of muon detectors CRs serve as an altimeter and very interesting is also the possibility to use CRs as a tomography probe.

First used in 1955 by Eric George to measure the depth of the overburden of a tunnel in Australia [195], muon tomography can be used to create three-dimensional images of volumes. Depending on the atomic number of the material, muons are scattered and absorbed in characteristic ways when

interacting with a dense material (for example uranium). Since cosmic muons are uniformly available across the Earth's surface, muon tomography is particularly desirable. Muons can penetrate far deeper into dense materials than conventional imaging methods, including X-rays. Impressive is the possibility in detecting nuclear material in road transport vehicles and cargo containers (nuclear proliferation) [196], as was also recently reported to be able to deliver high-resolution 3D images of a small lead block inside a large sensing area, to stop dangerous nuclear materials from being transported illegally [197]. Furthermore, to monitor potential underground sites used for carbon sequestration, to predict volcanic eruptions via imaging of magma chambers [198], for exploration of Mars geology [199], or to get a realistic estimate of the extent of the damage in the Fukushima Daiichi power plant at its very core, and to obtain knowledge of the location of the molten fuel [200, 201]. Via isotopes measured in the sand of rivers, which are formed by CRs, geologists quantified the erosion of the Swiss Alps and showed for the first time that the mountains are being lifted faster than they are being lowered through erosion and are thus growing even higher [202]. Moreover, muon analysis allows scientists to look deep inside ancient monuments without drilling holes or causing other damage to the precious structures. A particularly remarkable recent application of the muon tomography is the detection of a void in the Great Pyramid at Giza in Egypt [203].

3. Ultra-high Energy Photons

As a constant companion of CRs, cosmic photons cover a wide range of energies, from the radio regime in MHz, over the microwave, IR, optical light, UV and X-ray regime up to γ-rays. The highest photon energy observed is about 1.4 PeV, as recently detected from one of 12 Galactic PeVatron sources recorded by China's Large High Altitude Air Shower Observatory (LHAASO) [2].

During CR acceleration and propagation, photons and neutrinos are produced in hadronic interactions of CRs. As (electrical) neutral messengers they are thus not referred to as "cosmic rays", and provide information about CR source direction on the sky, and the physical conditions at the CR production site, while the (charged) CRs suffer from loss of directional information by deflecting at intergalactic and galactic MFs. A diffuse photon flux is produced in the Galaxy by the interaction of GeV CRs with the interstellar medium. At extreme energies, such secondary photons and neutrinos from UHECR interactions are called *cosmogenic*, referring to the production by the decay of primary CR particles, as e.g. the pions from the GZK-process.

Apart from their great value for astrophysics, observations of UHE photons would grant access to fundamental physics, as a direct accessible probe of aspects of QED and Quantum Chromodynamics (QCD) at the highest energies yet not accessible with man made experiments.

3.1. Experimental state of the art

A non-observation of photon-induced air showers above few PeV so far allowed to constrain the overall flux of primary photons and was used to derive upper limits on the diffuse flux of cosmic photons with energies above $6 \cdot 10^{16}$ eV – as shown in figure 3.1 – which are restricting and even excluding several UHECR production models as discussed in sec. 2.4.

With large-scale observatories such as the Pierre Auger Observatory or TA, the CR data of unprecedented quality and quantity is being accumulated and flux limits already constrain optimistic predictions of GZK-type scenarios. However, these predictions are strongly dependent on the chemical composition at the highest energies. The Pierre Auger Observatory is the most sensitive air-shower detector for primary photons with energies above several 10^{16} eV. In addition to the diffuse photon flux searches also analyses aiming for point sources of photons or searching for angular correlations with a stacked catalogue of possible HE photon sources are conducted [205, 206].

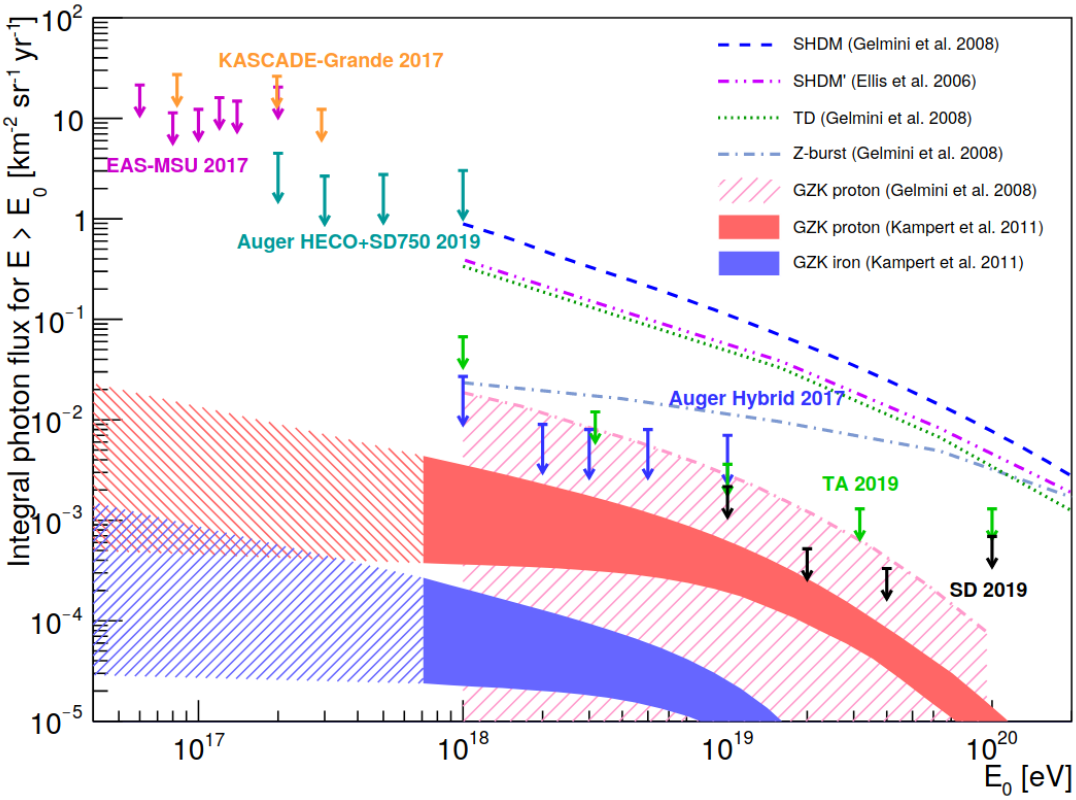


Figure 3.1.: Upper limits on the integrated photon flux at 90% CL by KASCADE-Grande and EAS-MSU and 95% CL by TA and the Pierre Auger Observatory compared to model predictions [204].

3.2. Production of cosmogenic UHE photons

Thermal emission from hot celestial objects is not sufficient to create photons above the GeV energy range, nominating UHE photons as probe of the non-thermal Universe where other, non-thermal mechanisms allow the concentration of large amounts of energy into a single particle.

From all the previous mentioned processes that can create particles of the highest energies, the subsequent decay of a neutral pion from the channel of photo-meson production (sec. 2.6.2) is known to be by far the most dominant in producing photons with energies above 10^{18} eV [173].

Either from the GZK-type processes of a UHECR nucleus N interacting with photons of the low-energy CBR (eq. (2.11), or more specifically a proton p with the CMB eq. (2.9) and (2.10) or eq. 2.12)), or from proton-proton interactions the produced pion π will further decay. The decay of the pions is one of the most efficient sources of neutral particles (photons γ and neutrinos ν) with UHE in conventional astrophysical bottom-up scenarios:

a) b)

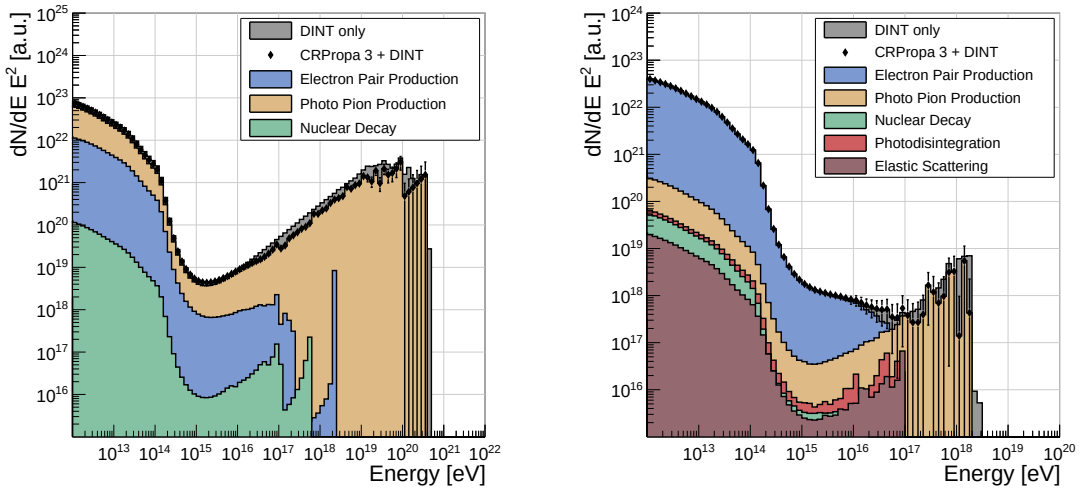


Figure 3.2.: Photon flux, scaled with E^2 **a)** from protons and **b)** from iron. Black dots show the total photon flux. Both figures taken from [173].

$$\pi^0 \rightarrow \gamma_{\mathrm{UHE}} + \gamma_{\mathrm{UHE}} \tag{3.1}$$

$$\pi^+ \rightarrow \mu^+ + \nu_{\mu,UHE} \tag{3.2}$$

with the γ_{UHE} receiving about a tenth [207] of the energy of the primary UHE protons (and less for heavier nuclei). This can be seen in figure 3.2 in comparison of a primary proton in a) vs. primary iron in b) emitted with a very hard power-law spectrum $dN/dE \propto E^{-1}$ between 1 and 1000 EeV from uniformly distributed sources at 3 – 1000 Mpc distance (to reach higher simulated photon statistics in all production channels). While additional photon production channels like PD [208], elastic scattering and radiative decay are negligible for cosmic-ray nuclei, they can give rise to a relevant contribution of UHE photons.

In b) photo-pion production marks the dominant contribution to the photon flux at highest energies, in the iron scenario, whereas electron PP is most relevant for lower energies, and elastic scattering, PD and nuclear γ-decay are of subdominant contribution.

In comparison in a), the proton scenario, the contribution of photo-pion production is dominant throughout the considered energy range, as well as nuclear decay contribution originating from the β-decay of neutrons produced via photo-pion production, while the subdominant contributions are from electron PP and nuclear decay.

Production of photons at the highest energies can be further achieved via *elastic scattering*, where a background photon effectively receives a double

Lorentz boost

$$ {}^{A}_{Z}X + \gamma_{\mathrm{CBR}} \rightarrow {}^{A}_{Z}X + \gamma_{\mathrm{UHE}} \,, \tag{3.3} $$

and no additional particles are produced.[1]

Since CRs can also contain radioactive nuclides, one or more photons can be emitted in the nuclear relaxation process following a *radioactive decay*

$$ {}^{A}_{Z}X^{*} \rightarrow {}^{A}_{Z}X + n\gamma \,, \tag{3.4} $$

whereat the number and energy of the emitted photons depends on the preceding decay, as well as on nuclear level transitions inside the excited nucleus, potentially resulting in an observable photon energy in the sub-EeV range (for particles with a Lorentz factor of 10^9).

To mention only briefly, for the sake of completeness, photons in the UHE range have theoretically a possibility of production via the already earlier introduced top-down models (sec. 2.4.2), due to the decay or the annihilation of the super-massive X particles (in most of these models). The production of quarks and gluons in these decay or annihilation processes can initiate QCD cascades, in which subsequently partons hadronize, producing UHECRs together with a large number of pions [209]. Also in the Z-burst model can the decay of the Z boson lead to quarks and antiquarks initiating QCD cascades. What all these top-down models have in common is the prediction of a large fraction of neutral pions and hence photons, up to two magnitudes more than in bottom-up models [97]. However, the parameter space for UHE photons from top-down models has been already severely restricted by experiments, as can be seen in figure 3.1.

3.3. Propagation of UHE photons

UHE photons as well as the CRs experience interaction processes during their propagation before reaching the Earth. The most relevant processes in this case is the $e^{\pm}$ PP from interactions with photons of the low-energy CBR:

$$ \gamma_{\mathrm{UHE}} + \gamma_{\mathrm{CBR}} \rightarrow e^{+} + e^{-} \,, \tag{3.5} $$

as compared to eq. (2.15) for UHECRs, and the process of double pair production (DPP):

$$ \gamma_{\mathrm{UHE}} + \gamma_{\mathrm{CBR}} \rightarrow e^{+} + e^{-} + e^{+} + e^{-} \,. \tag{3.6} $$

[1] A similar photon production channel particularly important below PeV energies originates via the ICS of low energy background photons by HE electrons either from PP (Bethe–Heitler, sec. 2.6.2), or from cosmic synchrotron emission, e.g. in pulsars and pulsar-driven nebulas.

The thereby created secondary HE $e^{\pm}$ quickly lose their energy via ICS or synchrotron radiation [207], forming an EM cascade. Since energy distribution between the electron and the positron produced in the process (3.5) is not symmetric (because of the very high center-of-mass energy), one of those particles obtains most of the energy of the primary UHE photon and can in turn again up-scatter a background photon to UHEs via ICS [210]

$$e^{\pm} + \gamma_{\mathrm{CBR}} \rightarrow e^{\pm} + \gamma_{\mathrm{UHE}} \,. \tag{3.7}$$

This interplay of PP and ICS continues and produces EM cascade until the photons energies are down to GeV energies.

These interactions with photons of the IR background, CMB and universal radio background (URB) strongly limit the propagation of UHE photons, as represented via the interaction lengths λ as a function of the energy in figure 3.3.

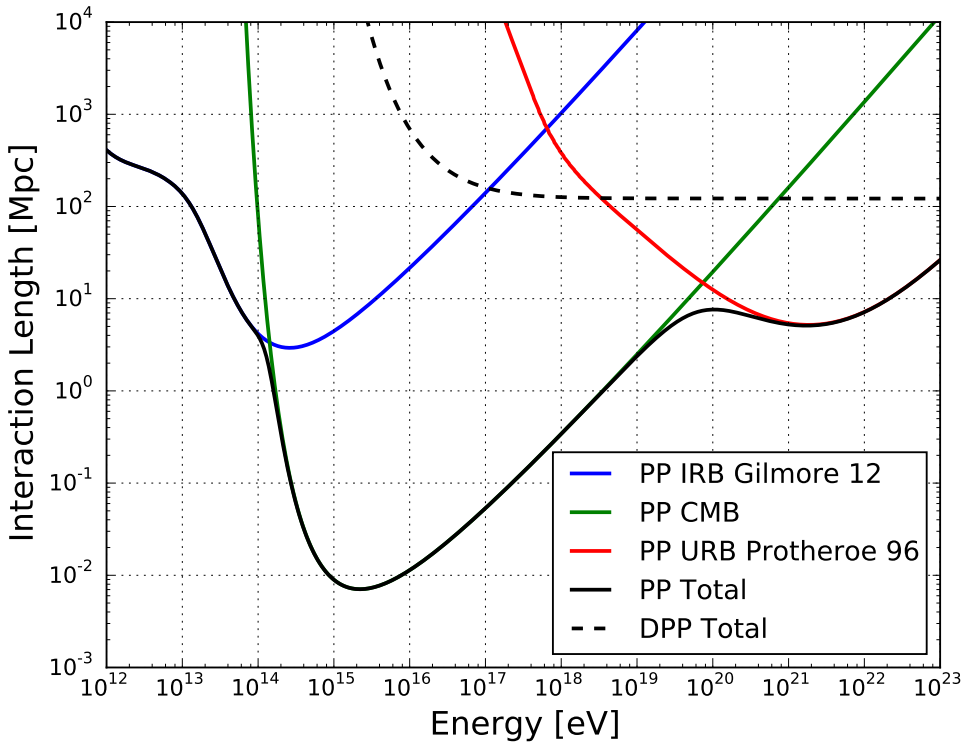


Figure 3.3.: Interaction lengths for cosmic-ray photons interacting with cosmic photon backgrounds CBR. The relevant processes are PP (PP, black solid line) and double PP (DPP, black dashed line). Figure taken from [173].

Finally, as in the case of UHECR primaries, adiabatic energy losses due to the expansion of the Universe have to be considered. With an energy loss length for this mechanism of $\sim$ 4000 Mpc [211] (for the Einstein-de Sitter model of a flat, matter-dominated, Universe, and a Hubble constant $H_0 =$ 75 km s^{-1} Mpc^{-1}) adiabatic energy losses are mainly relevant for photons

with energies below the TeV range.

Taking all processes into account, UHE photons could reach the Earth only from nearby sources, at distances of a few Mpc, as shown in fig. 3.3.

3.4. Testing Lorentz Invariance

Contrary to neutrinos, photons undergo interactions with the photons of the CBR, inducing electromagnetic cascades, and are therefore sensitive to the extragalactic environment. As a consequence, propagation effects make photons an excellent probe of new physics scenarios as e.g. violation of Lorentz invariance or photon-axion conversion.

The decades old quest for a quantum theory of gravitation called Quantum Gravity (QG) has posed a challenge to theoretical physics and remains far from understood. At the *Planck scale* M_{Pl} which is an energy scale around $1.22 \cdot 10^{19}$ GeV, quantum effects of gravity become strong and descriptions of sub-atomic particle interactions in terms of quantum field theory (QFT) break down and become inadequate, due to the impact of the apparent non-renormalizability of gravity within current theories.

One example of such a fundamental symmetry that is believed to hold exactly in nature and could be broken by QG is CPT invariance, requiring physics to be unchanged under the combination of charge conjugation (C), parity inversion (P) and time reversal (T).

The connection of Lorentz symmetry to CPT symmetry arises from QFT, in which Lorentz invariance is one of the hypotheses of the "CPT theorem". While the CPT breaking in any unitary, local, relativistic point-particle field theory implies violation of Lorentz invariance, this is not holding vice versa, as it is possible to violate Lorentz invariance while keeping CPT exact.

QG phenomenology via the violation of discrete symmetries [212] with e.g. the discovery of Lorentz invariance violation (LIV) would be an important signal of physics beyond the SM plus General Relativity (GR).

For yielding a glimpse of QG it is interesting to study both the theory and the phenomenology of LIV, which can be investigated in the hadron sector via modified UHECR propagation and subluminal LIV on the photon sector via modified propagation of GZK photons. Since effects of LIV are expected to be suppressed at low energies and for short travel distances [213], the most energetic known particles in the Universe, UHECRs with their extragalactic origin as well as UHE photons are therefore a unique possibility for testing it. Furthermore are in several Lorentz violating tests the relevant particles photons and electrons, making Lorentz violating QED the appropriate theory.

While also in this case the non-observation of LIV signals results in restrictive limits on the LIV coefficient [214], LIV in the photon sector is limited by the upper limits in the photon flux.

The aforementioned modifications were described in a phenomenological approach in [213, 215], summarizing LIV effects in a perturbative modification

of particle's dispersion relation for fermions:

$$E^2 \;=\; p^2 \,+\, m^2 \,+\, f(E,p;\mu;M)\;, \tag{3.8}$$

with $c = 1$; E and p are the particle energy and momentum, respectively; μ is a particle-physics mass-scale (associated with a symmetry breaking) and M denoting the relevant QG scale assuming $M \sim M_{Pl}$. The function $f(E,p;\mu;M)$ can thereby be expanded in powers of the momentum and the lowest order LIV terms (p, p^2 and p^3) have primarily been considered in [213].

And accordingly for photons in the case of Effective Field Theory (EFT), the modification via non-renormalizable, Planck suppressed LIV operators of the effective Lagrangian of QED leads to the following HE modified dispersion relations (MDR) [216]:

$$\omega^2 \;=\; k^2 \,+\, \frac{\xi^{(n)}\, k^n}{M^{n-2}}\;, \tag{3.9}$$

with the momentum k and ω the energy of the HE photon, $n = 3$ for dimension 5 operators and $n = 4$ for dimension 6 ones, and the constants $\xi^{(n)}$ indicating the strength of the LIV. In this framework of LIV in EFT also processes which are forbidden in Lorentz invariant physics are allowed and can effectively dump the photon flux: photon decay in vacuum and photon splitting ($\gamma \to N\gamma$) [216].

Since the modification of the dispersion relation leads to a change in the kinematics of the interactions of the particles, causing modifications in the extragalactic propagation of the UHECRs (and UHE photons), LIV effects are expected to be visible in the measured spectrum and composition.

As described in the previous section, highest energy photons interact with the CBR via PP, what strongly suppresses the expected flux of GZK photons, but due to LIVs an increase in the mean free path is obtained, leading to photons travelling farther and consequently an enhanced flux of GZK photons is expected.

As described in [215, 217] there may not be a GZK cut-off after all, arguing that already small departures from Lorentz invariance have effects that increase rapidly with energy, possibly kinematically preventing cosmic-ray nucleons from undergoing inelastic collisions with CBR photons resulting in the elimination of the cut-off and leading to a deeply cosmological origin of UHECRs.

An example for preserving CPT and gauge invariance but breaking Lorentz invariance (via the addition of a single term to the Lagrange density [218]) was presented in [219], where the impact of two Lorentz-violating processes (photon decay into an electron-positron pair and modified neutral-pion decay into two photons) was investigated, based on measurable modifications in the development of EASs due to an isotropic LIV in the photon sector [220].

The described theoretical possibilities serve as the driving motivation for this work, which aims to contribute to the Pierre Auger Observatory search

for the expected UHE photons.

What seemed too far out of reach just a while ago is today already feasible: existing experiments are able to strongly constrain Planck-suppressed effects motivated by QG scenarios. This can be achieved, because even tiny violations of a fundamental symmetry such as Lorentz invariance can lead to detectable effects at energies well below the Planck scale, as i.e. it is just a matter of time for the Pierre Auger Observatory to observe expected photons with energies above 10^{19} eV and hopefully signatures of LIV and in this way further restrict the allowed region of parameter space for LIV QED.

3.5. Multi-messenger observations

Multi–messenger (MM) astrophysics is the combination of the messengers of the Nature's four fundamental forces, i.e. cosmic rays, photons, neutrinos and gravitational waves (GWs), in correlated and coordinated studies of the most energetic phenomena in the Universe.

In a chronological order, photons, the force carrier of the EM force, has been the only source of information about extraterrestrial physics - i.e. in optical light - for a long time and is still the richest source for astronomical observations. The accessible EM spectrum of astrophysical sources and vastly broadened with the observations of radio waves, X-rays and gamma rays and the development of dedicated instruments. Multiwavelength observations of distinct astrophysical objects maximized the accumulated information about and boosted the understanding of the physics of these objects. The first detection of cosmic very–high energy (VHE) γ-rays from the Crab Nebula was reported by the Whipple imaging atmospheric Cherenkov telescope (IACT) in the 1980s [221]. Gammas at lower MeV–GeV gamma-ray energies have been studied in great detail with Fermi Gamma-ray Space Telescope [222, 223], at VHEs with arrays of IACTs, such as H.E.S.S. [224], MAGIC [225] and VERITAS [226], and with dedicated high-altitude EAS detector arrays such as HAWC [227].

The earliest known messengers taking part in the strong interaction at the VHEs, CRs, were detected in 1912 by Victor Hess with his balloon experiments, as described previously in chapters 1 and 2.

First detections of the messengers that undergo the weak interaction, neutrinos ν, arriving from extrasolar sources at high energies were announced in 1987, in the first ever MM observation of coincident neutrinos and photons from supernova SN1987A [228, 229]. Recently, the IceCube detector [230] reported evidence for extraterrestrial neutrinos in the 30 TeV-2 PeV energy range [231] with a diffuse flux significantly above the expected atmospheric neutrino background. A notable astrophysical MM observation of IceCube, Fermi-LAT, MAGIC, AGILE, HAWC, H.E.S.S. (to name just a few of the already mentioned) collaborations is the detection of a flare of high energy

photons in coincidence with a high energy muon neutrino event (IceCube-170922A) from the direction of a distant blazar TXS-0506+056 in 2017 [232, 233]. Conclusions suggest that blazars are a source of (a substantial fraction of) the cosmic neutrino flux, and a potential source of VHE CRs.

Disturbances in the curvature of spacetime, generated by accelerated masses, that propagate as waves outward from their source at the speed of light, the gravitational waves (GWs) are the wave solutions of the linearized weak-field equations, the final formulation of the field equations of GR, as proposed by Henri Poincaré in 1905 [234] and predicted by Albert Einstein in 1916 [235, 236]. In 2016 the first GW signals from the merging of a BS of BHs, GW150914, were reported [237]. This detection of GWs from the BSs merger has not only marked the birth of GW astronomy, but enabled to combine the messengers of all four of nature's fundamental forces.

One advantage of neutral particles compared to CRs is their immunity to MF during their propagation through the Universe and thereby deliver information about their sources. When arriving at Earth they can initiate EAS in the atmosphere just like CRs and for example in the case of neutrons, these showers are indistinguishable from showers induced by protons.

Designed for the main purpose to answer the big questions about UHECRs the Pierre Auger Observatory serves as a well suited detector to contribute via several messengers to MM astrophysics with unprecedented accuracy and statistics. In the following, the Pierre Auger Observatory's capabilities to investigate other messengers, like neutrinos and galactic neutrons, in a way complementary to other observatories is addressed.

3.5.1. Neutrino searches

Concerning their physical properties, neutrinos are the ideal messenger to study the HE Universe, and searches at UHEs for these particles have been constantly undertaken by the Pierre Auger Observatory [238, 239].

Neutrinos can travel in straight paths through cosmological distances unlike charged particles, which are deflected by cosmic MFs, or photons which are absorbed. They are a clear and direct signature of hadronic acceleration or interaction, and are capable of going through large matter depths without absorption. Beside all of that they reveal information about composition of CRs, the local environment at the sources, and their evolution with redshift [240].

Observatories designed specifically for HE astrophysics with ν are experiments such as IceCube [241], Antares [242] and KM3NeT [243]. The data of observations of these astrophysical HE ν can then be combined with the data of the Pierre Auger Observatory in a MM approach.

Due to the very weak interaction of ν with matter, the discrimination from other primary particles is based on the fact that neutrinos can potentially interact at any point in the atmosphere for any zenith angle, which means

they are, in contrast to all other particles, able to interact very deep in the atmosphere. This offers the opportunity to search for neutrino primaries in very inclined showers since the EASs from hadronic primaries are reduced to mostly HE muons well before reaching the Earth's surface at the large zenith angles and thus lowering background for neutrino searches.

With the Pierre Auger Observatory searches for neutrino candidate events are performed separately for Earth-skimming (ES) – in a zenith angle range $90^\circ \leq \theta \leq 95^\circ$ – and down-going low (DGL) – $60^\circ \leq \theta \leq 75^\circ$ – and down-going high (DGH) – $75^\circ \leq \theta \leq 90^\circ$ – neutrino-induced showers [244, 245]. Thereby can the shower in the DGL or DGH channel be induced by any neutrino flavours, while in the ES channel only via tau neutrinos (ν_{tau}). These ν_{tau} are traversing the Earth's crust, with some ν_{tau} interacting close to the surface after skimming through the ground, producing a tau lepton that will start an upward-going shower above the ground.

It is possible with the Pierre Auger Observatory to detect UHE neutrinos from a large fraction of the sky, from very close to the South Celestial Pole to declination values up to $\delta \sim 60^\circ$. For a steady source of UHE ν in this range, there constantly is a time window during a sidereal day in which the source is in the field of view (FOV) of the ES, DGH or DGL channels.

In the search of downward going events the discrimination power is optimised with the use of the Fisher's linear discriminant method (cf. sec. 6.5.1.3), using a combination of several observables to optimise the separation between the background hadronic showers. For the ES analysis one variable is enough to clearly discriminate, since there is no background in this channel. Since the target mass provided by the Earth is large compared to the atmosphere and also because the threshold energy is low, the ES is the most effective channel for neutrino detection.

The unmatched sensitivity of the SD of the Pierre Auger Observatory to potential sources of EeV neutrinos in the northern celestial hemisphere arises from the fact that this region of the sky cannot be exploited in the EeV energy range by IceCube, due to the opacity of the Earth to neutrinos arriving from those directions when seen from the South Pole.

Despite its very good sensitivity in all ν channels, no UHE ν has been identified by the Pierre Auger Observatory to date, which allowed to derive strong limits on the diffuse flux of UHE neutrinos, comparable to the one achieved by the IceCube experiment, as can be seen in figure 3.4a), and upper limits on flux from point-like sources of neutrinos, as can be seen in fig. 3.4 b). UHE neutrinos thus provide an independent probe of UHECR source properties and of the origin of the UHECR flux suppression at the highest energies. From the limit on diffuse fluxes it can be seen that the Auger data are disfavouring UHECR scenarios with proton-dominated UHECR composition and strong evolution of the sources with redshift. The scenarios of a mixed or heavy composition with weak evolution are more likely, which implies a suppressed neutrino flux.

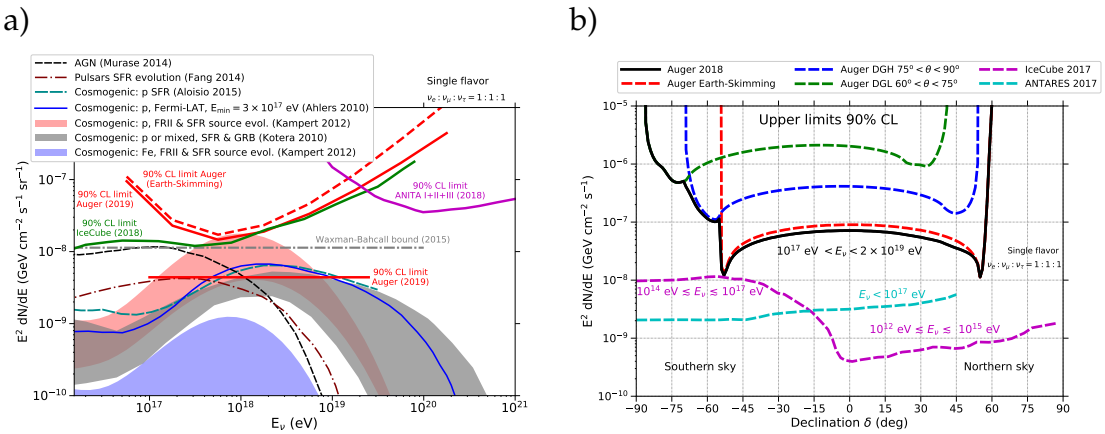


Figure 3.4.: a) Upper limit on diffuse fluxes, compared to the differential limits obtained by other experiments and astrophysical and cosmogenic neutrino models. In red are shown the integral and differential upper limit for the normalisation constant assuming a E^{-2} energy flux for a single flavour for all the three neutrino channels, and the dashed red line is the ES channel alone, which contributes the strongest to the limit [239]. **b)** Auger 90% CL upper limit on the fluxes of point-like sources as a function of equatorial declination obtained from the non-observation of neutrinos. Note the different energy ranges where the limits of each observatory apply. See [238] and references therein.

Neutrinos from point-like sources across the sky can be detected with the Pierre Auger Observatory's SD array, with peak sensitivities at declination around $-53°$ and $+55°$. Since no neutrinos have been identified, upper limits on the neutrino flux from point-like steady sources have been calculated as a function of declination as shown in fig. 3.4 b) together with IceCube and ANTARES upper limits at lower energies.

3.5.2. Galactic Neutrons

EASs induced by neutrons cannot be distinguished from the ones induced by CRs on the basis of the shower development. Nevertheless, is it in principle possible to determine sources of neutrons by identifying an event excess from given directions, or by exploiting potential correlations in time and direction, due to the non-deflected paths of neutrons. With a mean travel distance for relativistic neutrons with energies E_n of 9.2 kpc $\cdot E_n$/EeV, the distance from Earth to the Galactic center of about 8.3 kpc, and the radius of the Galaxy approximately 15 kpc, neutrons at UHEs can reach Earth from the entire Galaxy but not from much further away [246].

Temporal coincidence with lower-energy EM counterparts of UHE neutrons from Galactic sources [247], and UHE gamma rays from nearby extragalactic sources (within $\sim$10 Mpc) may both be exploited for the identification of

UHECR accelerators [248].

Since neutrons are, as well as photons, produced in photo-hadronic interactions, photon sources can be chosen as potential neutron sources. Anyhow, also in the case of neutrons no significant excess of a neutron flux has been found in the Pierre Auger Observatory's searches from any class of candidate sources, but strong upper limits at 95% CL on the energy flux in neutrons have been deduced, thereby setting strong constraints on UHE proton production in our galaxy [249, 250].

3.5.3. Electrons and positrons $e^{\pm}$

Beside being accelerated at astrophysical sources, electrons and positrons are created during the propagation of (mostly galactic) CRs in interactions with the interstellar matter. First high statistics measurement results on the positron fraction in CRs were published by PAMELA [251] and later confirmed by Fermi-LAT [252] and AMS-02 [41].

While no additional primary component from astrophysical sources (e.g. pulsars or dark matter) is required due to the consistency of the measurements of AMS-02 (in the range of 0.5 – 350 GeV) with a secondary origin of the positrons [253], results from PAMELA are well described by only plain diffusion or low re-acceleration models, without any need for a charge-asymmetric extra component, and an increase of the positron fraction up to 250 GeV is confirmed by all three experiments.

3.5.4. Global networks of experiments and observatories

By virtue of MM detection capabilities, unrivaled at the extreme energies, the Pierre Auger Collaboration is contributing to global networks of observatories which coordinate numerous observational facilities worldwide and in space, such as the Deeper, Wider, Faster (DWF) program initiated to study fast transients (millisecond-to-hours duration) [254] and the AMON [255, 256]. In the cooperation with such MM networks, the Pierre Auger Observatory can follow up and/or send a trigger for potentially interesting astrophysical events.

As the goal of this work is the development of a photon candidate data stream to AMON, a summary about AMON is given in the following.

3.5.4.1. The Astrophysical Multimessenger Observatory Network (AMON)

AMON provides a framework for cross-correlation analyses of the HE astrophysical signals across all four astrophysical messengers: photons, CRs, neutrinos, and GWs [255]. It is a cyber-infrastructure and a network of astrophysical detectors and observatories [257], created for the purpose of enabling near real-time coincidence searches [258], and is structured as an open and extensible network, with an Memorandum of Understanding (MoU) allowing for

straight-forward incorporation of new facilities [256]. This greatly increases the speed and effectiveness of reaction to a trigger, compared with the large number of traditional individual observatory-to-observatory connections.

As a virtual system it collects, analyses and distributes MM data under control of its participating observatories in real time for follow-up observations, where individual facilities participating in AMON can be characterized as triggering facilities, follow-up facilities, or both. The latter is the case for the Pierre Auger Observatory, which uses the excellent angular resolution and the particle identification capabilities of the observatory. Triggering facilities search in a wide FOV or monitor known sources for transient behaviour and are typically sensitive to one or more messenger type. The participating collaborations decide about all decisions concerning data analyses and alert distribution, and the shared data through the network remains the property of the originating collaborations.

AMON accepts, stores and analyses MM data of sub- and above threshold events, and distributes electronic alerts for follow-up observations. The main strength of AMON arises from combining and analysing sub-threshold data from different facilities (which cannot generally be used stand-alone to identify astrophysical sources) and thereby enhance the SNR. The definition of a "*sub-threshold*" event is left by AMON to the individual experiments.

AMON's analyses algorithms can identify statistically significant coincidence candidates of MM events, leading to the distribution of alerts used by partner observatories for real-time follow-up to potentially identify and confirm the authenticity and leverage e.g. the GRB coordinates network/transient astronomy network (GCN/TAN) [259, 260], among others, to promulgate its alerts.

The three main goals of AMON are defined as [256, 261] to:

- Receive events and broadcast them through the GCN/TAN to the astronomical community for follow-up.
- Store events into a database to perform archival coincidence searches.
- Perform coincidence searches of sub-threshold events of various observatories in real-time, and distribute prompt alerts to follow-up observatories.

As a central hub connected to individual observatories, the servers of AMON are hosted by the Institute for CyberScience at the Pennsylvania State University. The system has been constructed using the Python programming language, with the heart of the AMON software project called *AmonPy*, and has been under development since 2014. The intended growth through new observatories joining, subsequently creating new data streams and requiring new monitoring tools is explicitly supported.

The Pierre Auger Observatory and AMON

As a follow-up and triggering partner, the Pierre Auger Observatory is at present the only AMON partner sensitive to photons, neutrinos and hadrons at UHEs.

A MoU with AMON was signed in 2013 and since 2015 the Pierre Auger Observatory is already sending events to AMON, which is shown in figure 3.5 as the rates of events per day in comparison to other partner experiments. Thereby the Pierre Auger Observatory transmits the arrival direction and the

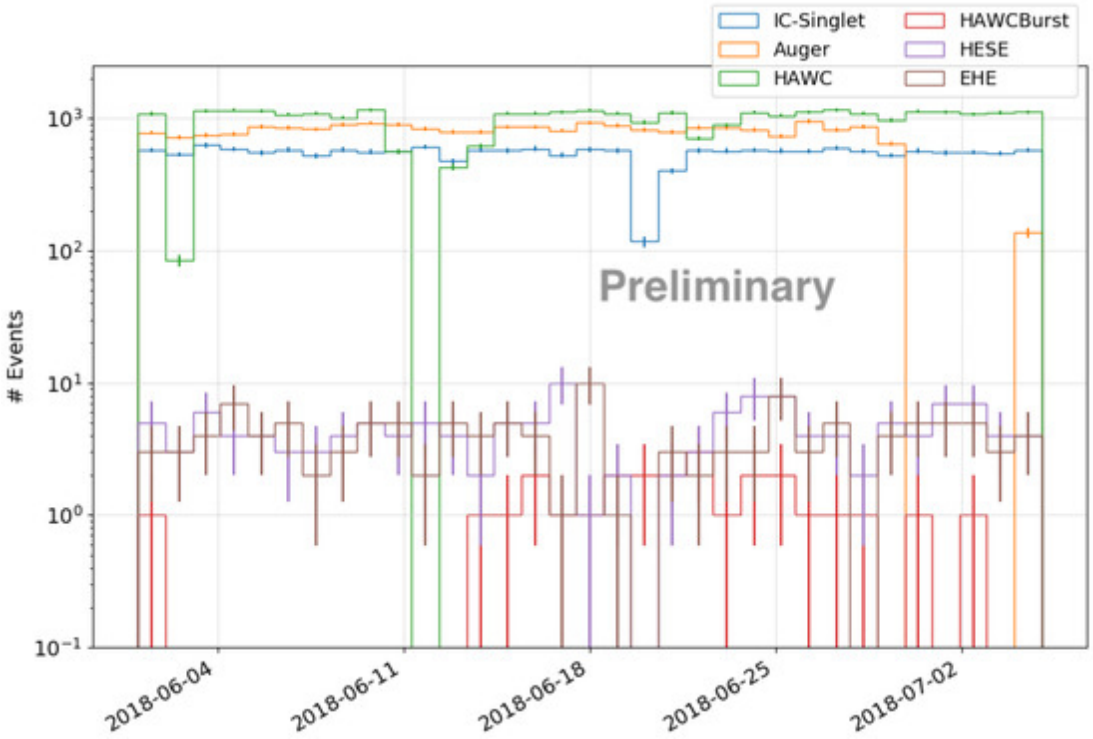


Figure 3.5.: Rates of selected streams received by AMON. The plot shows the number of events per day received during the month of June 2018. Figure taken from [257].

energy of high quality CR events to AMON in real-time [246], with energy above 3 EeV and zenith angle $\theta < 60°$. To complement the sent information with a statement about the nature of primary CR particles is one of the motivations for this work.

The Pierre Auger Observatory's capacities as a MM observatory and a contributor to global networks are expected to further increase with the ongoing development of analyses, as well as after the completion of the currently ongoing detector upgrade (*AugerPrime* see sec. 5.4).

4. Extensive Air Showers (EASs)

Before reaching the Earth's atmosphere, all CRs arriving at Earth first enter the geomagnetic field, where charged particles with energy $E < 1$ GeV can get magnetically trapped and reflected between the Earth's magnetic poles multiple times before even reaching the top of the atmosphere. Above that energy CR deflections by the MF leads to a latitude-dependent apparent anisotropy in arrival directions. A superb manifestation of the trapped low energy particles (mostly electrons and protons) are Northern Lights, while some CRs can escape back into outer space.

Above some 100 TeV, in the energy region approaching the knee, the incoming flux of primary CRs decreases rapidly with energy, reaching only $\sim$ 1 particle per m^2 per year and together with limitations of the effective detector area of balloon-borne and satellite-borne experiments of a few m^2, renders direct measurements impossible for collecting significant statistics on reasonable time scales.

However, as known since the 1930s, when CRs enter the atmosphere, they interact with atmospheric molecules (mostly N_2, O_2, Argon (Ar)), thereby producing secondary particles, which subsequently undergo further interactions in the atmosphere, thus initiating a cascade of nuclear and electromagnetic interactions that develops to an extensive air shower (EAS) of millions of secondary particles in the atmosphere. The thereby created number of secondary particles at first multiplies, till it reaches a maximum before it attenuates more and more as particles fall below the threshold for further particle production. In this way, the Earth's atmosphere serves as part of the detector, compatible to a large calorimeter.

To study EASs at sea level, at different elevations (on mountains), or even beneath the Earth, precise knowledge about the atmosphere, nuclear (especially proton-air) and photo-nuclear cross-sections and the geomagnetic field is required. A variety of detection techniques for these secondaries and their by-products has been developed to deduce the properties of the primary CR particle, as there are for example:

- imaging atmospheric Cherenkov telescopes (IACTs) measuring the Cherenkov light from air showers above few 10 GeV and allowing to select rare gamma-rays from astrophysical sources on the largely exceeding background from primary CR nuclei, based on the shape and orientation of the Cherenkov light images from charged shower particles,

- fluorescence light isotropically emitted by the secondary particles in the

air shower, in order to characterize the EAS's longitudinal development,

- arrays of ground-based particle detectors, which sample the lateral particle densities by measuring the shower front (both amplitude and time) of signals created in individual detectors by shower particles; the energy threshold for these techniques can be as low as in the TeV range for the detectors located at high mountain altitudes (e.g. HAWC); the individual detectors in arrays have to be spread over larger areas when aiming to detect higher CR energies, which implies the more sparse deployment of detectors.

Further detection techniques are worth to be mentioned, that potentially can be used as an independent technique to detect EASs or by complementing surface detector arrays or fluorescence telescopes, including

- **radar** reflections of the ionization columns produced by EASs [262];
- detection of **microwave** molecular bremsstrahlung radiation (MBR) originating from free electron collisions with neutrals in the tenuous plasma left after the passage of the shower [263];
- **acoustic** detection of EASs, already developed in the 1950s by Askaryan [264], which is based on the reconstruction of characteristic sound pulses that are generated by particle cascades in water or ice, where the energy deposition of cascade particles is connected to a local heating accompanied by an expansion of the medium [265];
- detection of **radio** pulses emitted from EASs, which are produced by several mechanisms, as e.g. from coherent synchrotron emission by the electron and positron pairs propagating in the Earth's magnetic field [266].

4.1. Air shower development

Following the same principles as for the cascades studied in particle physics, the development of an EAS can be described via simplified models used to describe EM and hadronic showers, which are able to describe the shower development.

A simplified sketch of cascade processes initiated by a primary particle is shown in figure 4.1. After the first interaction of a primary CR in the Earth's atmosphere, the leading nucleon propagates close to the trajectory of the primary particle (a "shower axis" in CR jargon) further developing hadronic cascade produced neutral and charged pions and kaons contribute respectively to EM cascade and to muon and neutrino production, so the total shower can be seen consisting of four components:

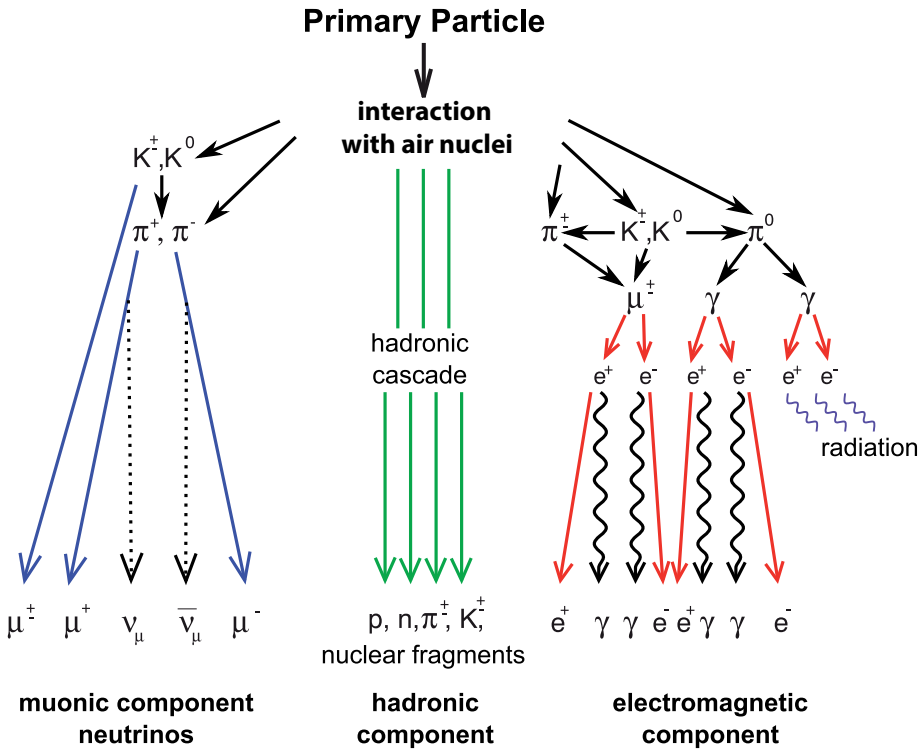


Figure 4.1.: Schematic view of an EAS, with the hadron, muon (and neutrino), and EM components. Figure taken from [267].

- EM component
  - ◇ Pure EM (Bremsstrahlung, pair production)
  - ◇ EM from μ - contributions from pion cascade
  - ◇ EM from hadronic jets
- Muons
- Hadrons (nucleons, $\pi^{\pm,0}$, $K^{\pm,0}$)
- Neutrinos

In case of an EM primary, whereat the first interaction initiates an EM cascade, the main components are EM particles with an additional small fraction of photo-nuclear interactions resulting in a small muon fraction.

While the hadronic component is typically contained within a few tens of meters from the shower axis, periphery of the shower is dominated by the EM and muonic component. The development of these latter two components proceeds in a different way: while the EM component propagates "diffusively" (due to the multiple Coulomb scattering of electrons), muons propagate radially from their parent mesons. The shower front can therefore thought to be a superposition of a more extended in time, later arriving EM

component and a temporally thin first arriving muonic component. Due to the suppressed muonic component in case of photon induced EASs, information of the primary particle is imprinted in the slope of the lateral distribution, shower curvature and structure of the shower front, as will be described later in sec. 5.2.4.

4.1.1. Electromagnetic cascade & Heitler model

The almost immediate electroweak decay of neutral pions $\pi^0 \rightarrow \gamma + \gamma$ triggering the EM cascade is based on two processes: pair production and bremsstrahlung ($e^{\pm} \rightarrow e^{\pm} + \gamma$). The latter is a deceleration radiation, where the moving particle loses kinetic energy, which is converted into radiation (i.e. photons) by the deceleration of a charged particle deflected by another charged particle, typically an electron by an atomic nucleus ($e^- + (Z, A) \rightarrow e^- + \gamma + (Z, A)$), as depicted in the dominant Feynman diagram in figure 4.2.

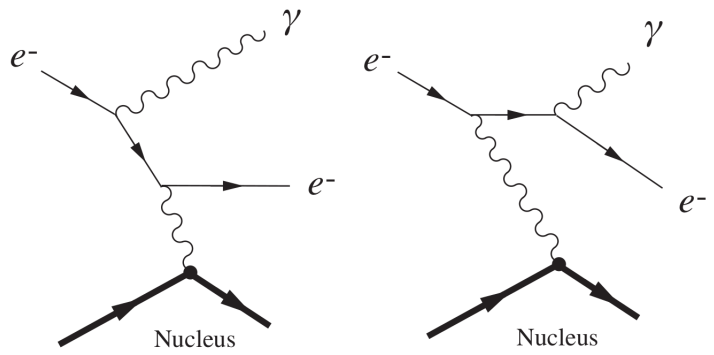


Figure 4.2.: Schematic depiction of the dominant Feynman diagrams for the bremsstrahlung process for an electron in the EM field of a nucleus. Figure taken from [179].

The energy loss due to bremsstrahlung is described as

$$-\left(\frac{dE}{dx}\right)_{\text{rad}} = \kappa_{\text{brems}} \frac{E}{A} \left[Z^2 \left(\ln\left(\frac{184.15}{Z^{1/3}}\right) - f(Z) \right) + Z \ln\left(\frac{1194}{Z^{2/3}}\right) \right], \quad (4.1)$$

where $\kappa_{\text{brems}} = 4\alpha N_A r_e^2 = 1.396 \cdot 10^{-7}\ \text{cm}^2\ \text{g}^{-1}$ is a collection of constants (as r_e the classical electron radius and N_A Avogadro's number), and $f(Z)$ a function of Z [175].

Although being comprised of billions of particles and involving many different interactions, the key features of the longitudinal development of the EM cascade of an air shower in the atmosphere can approximately be described by a simple toy model, the *Heitler model* [29, 268] and even a simplified Heitler model as discussed in [211].

The model is based on the assumption that per interaction two particles are created, each carrying half of the original particle's energy and at an

atmospheric depth called $X_{\max}$ the number of shower particles is maximal $N_{\max} = E_0/\varepsilon_c$ and ionization processes become dominant.

For an EM cascade (following [269]) with an electron radiating a single photon after travelling one splitting length $d = X_0 \ln 2$ and X_0 the radiation length in the medium (here air), one finds, that a distance of $x = nX_0 \ln 2$ is reached after n splitting lengths with a total shower size (electrons and photons) $N = 2^n = e^{x/X_0}$. From plugging the equations into each other one obtains:

$$X_{\max}^{EM} = X_0 \ln\left(\frac{E_0}{\varepsilon_c^{EM}}\right) , \tag{4.2}$$

where E_0 the primary energy and $\varepsilon_c^{EM} \approx 85$ MeV the critical energy below which the ionization processes are dominant. Despite of its simplicity, the Heitler model predicts the most important features of particle cascades, notably that $N_{\max}$ is proportional to E_0 and that $X_{\max}$ is proportional to $\ln(E_0)$, and is in good agreement with detailed Monte Carlo (MC) simulations of particle cascades [211].

4.1.2. Hadronic component & superposition model

A further generalization of the Heitler model for hadronic showers has been introduced by Matthews in [269], which together with [270] serves as description herein. It is assumed, that about 2/3 of the primary energy is converted into the productions of charged pions, and subsequently in the pion cascade at each step 1/3 of the energy is transferred to the electromagnetic cascade. The other 1/3 is transferred into neutral pions, which quickly decay into photons that further initiate secondary EM showers. The charged pions continue interacting and finally **d**ecouple from the shower by decays into muons and neutrinos after n_d interactions, where N_{ch} **ch**arged pions have been created with an energy per $\pi^{\pm}$, $\varepsilon_d^{\pi} \approx 10$ GeV [269] (in comparison to the previous section 4.1.1)

$$N_{\pi}^{\pm} = N_{ch}|_{n_d} = \left(\frac{E_0}{\varepsilon_d^{\pi}}\right)^{\beta} , \tag{4.3}$$

where the number of secondary particles is proportional to the primary energy E_0, and β assumed to be $\ln N_{ch}/\ln(3/2N_{ch}) \approx 0.85$ for $N_{ch} = 10$ with an estimated [269] energy-independent multiplicity [271, 272]. The energy thereby transferred in the EM channel is

$$E_{EM} \simeq E_0 \cdot \left(1 - \left(\frac{2}{3}\right)^{n_d}\right) , \tag{4.4}$$

and hence E_{EM} is a good measure of the primary energy, where $n_d = 6$ can be approximated [269, 273]. Furthermore can the *depth of the shower maximum* for a primary proton be approximated with elasticity κ_{ela} and the interaction

length of the proton λ_p, if the total multiplicity of hadrons produced in the main interaction is N and the average hadron energy is E_0/N

$$\langle X^p_{\max} \rangle \simeq \lambda_p + X_0 \ln \left(\frac{\kappa_{\text{ela}} E_0}{2N\varepsilon_c^{EM}} \right) , \tag{4.5}$$

with the factor 2 taking into account the decay of neutral pions into two photons, thereby correctly estimating the correlation $X_{\max} \propto \ln(E_0)$.

For heavier nuclei the superposition model is a good approximation, where the shower is initiated by a nucleus of atomic number A is treated as the superposition of A showers initiated by protons with energies E_0/A. With the elongation rate D, which quantifies the change of the average shower maximum depth per logarithm in energy,

$$D = \frac{\mathrm{d}X_{\max}}{\mathrm{d}\ln E} , \tag{4.6}$$

such (semi-)superposition model results in logarithmic growth with energy of $X_{\max}$ for nuclei as well, and the following dependence with respect to protons:

$$\langle X^A_{\max} \rangle \simeq c + D_p \cdot \ln(E/A) = \langle X^p_{\max} \rangle - D_p \cdot \ln A , \tag{4.7}$$

with $D_p = \mathrm{d}\langle X^p_{\max} \rangle / \mathrm{d}\ln E$ approximately constant (due to an approximately logarithmic decrease of λ_p with energy in hadronic interaction models) and depending on the properties of the first interaction and on the multiplicity.

Continuing for a mixed compositions of nuclei of mass A_i at fixed energy E, with fractions f_i and $\langle X_{\max} \rangle = \sum_i f_i \langle X_{\max} \rangle_i$ and $\langle \ln A \rangle = \sum_i f_i \langle \ln A \rangle_i$, the mean logarithmic mass is

$$\langle \ln A \rangle = \frac{\langle X^p_{\max} \rangle - \langle X_{\max} \rangle}{D_p} , \tag{4.8}$$

explicitly demonstrating the relation of $\langle X_{\max} \rangle$ to the *average logarithmic mass* of the cosmic ray composition. With many assumptions this model predicts correctly amongst others that the number of muons increases with the primary energy according to a power law; that the maximum development depth of a hadronic shower is shallower than for an EM shower; also, the heavier the nucleus, the larger the muon fraction.

4.1.3. Muonic component

The muonic component of an EAS has the origin similar to the one of the EM component, which is fed by the hadronic component through the decays of neutral pions, with the difference, that it originates from the decays of charged

mesons, mainly pions and kaons, with the most relevant processes being:

$$\begin{aligned} \pi^{\pm} &\rightarrow = \mu^{\pm} + \nu_{\mu}(\bar{\nu}_{\mu}) \\ K^{\pm} &\rightarrow = \mu^{\pm} + \nu_{\mu}(\bar{\nu}_{\mu}) \\ K^{\pm} &\rightarrow = \pi^{0} + \mu^{\pm} + \nu_{\mu}(\bar{\nu}_{\mu}) \, . \end{aligned} \tag{4.9}$$

In addition, a fraction of up to 10% of the (low-energy) muons in EASs is produced in photo-nuclear interactions of photons from the EM component with nuclei from the air [270].

The total number of muons in EASs can be approximated by the number of created charged pions once the critical energy (ε_c^{π}) is reached and for proton primary air showers by [269]

$$N_{\mu}^{p} \approx \left(\frac{E_0}{\varepsilon_c^{\pi}} \right)^{\beta} , \tag{4.10}$$

with

$$\beta = \frac{\ln(\frac{2}{3}N)}{\ln N} \tag{4.11}$$

in the range $\beta = 0.88 - 0.92$, thereby the number of muons increases with the energy of the primary. The factor $2/3$ accounts for the fraction of charged pions. For the air showers induced by heavier primaries, the superposition model predicts:

$$N_{\mu}^{A} \approx A \left(\frac{E_0/A}{\varepsilon_c^{\pi}} \right)^{\beta} = N_{\mu}^{p} \, A^{1-\beta} , \tag{4.12}$$

and consequently a higher number of muons.

With their finite lifetime of 2.2 μs [175] low-energy muons further decay via

$$\mu \rightarrow e^{\pm} + \nu_e(\bar{\nu}_e) + \bar{\nu}_{\mu}(\nu_{\mu}) , \tag{4.13}$$

while high energy muons are affected by relativistic time dilation and therefore mostly reach the ground, passing the atmosphere nearly undisturbed.

4.1.4. Longitudinal & lateral development of air showers

In the previous sections it was pointed out that there are several channels contributing to the EM component, which overall quickly becomes the dominant shower component, outnumbering hadrons and muons by several orders of magnitude, as can be also seen in the lateral and longitudinal distribution of particles in figure 4.3.

As an example, in an EAS induced by a proton with an energy of 10^{20} eV, a fraction of $90 - 95\%$ of the energy is transferred into the EM channel [274], so the measurable quantities of the EAS are mainly determined by the EM component.

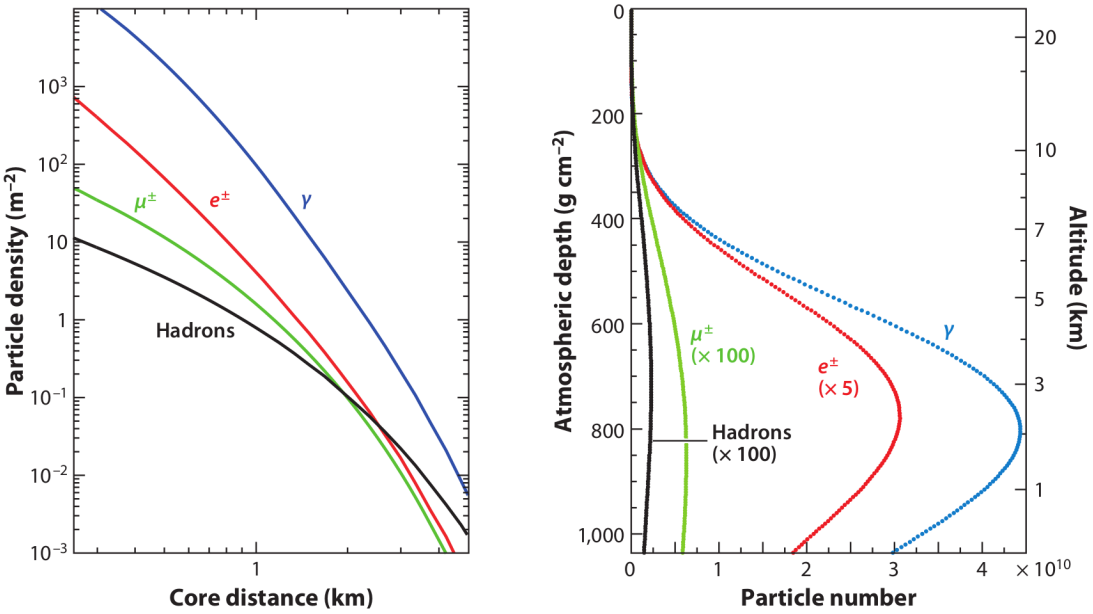


Figure 4.3.: Average (left) lateral and (right) longitudinal shower profiles for vertical, proton-induced showers at 10^{19} eV, with energy thresholds in the simulation of 0.25 MeV for γ and $e^{\pm}$, and 0.1 GeV for muons and hadrons. Figure taken from [274].

A function used to parametrize the number of particles $N(X)$ as a function of the slant traversed atmospheric depth X, the *longitudinal profile*, is the so-called Gaisser-Hillas function [275]:

$$N(X) = N_{\max} \left(\frac{X - X_0}{X_{\max} - X_0} \right)^{\frac{X_{\max} - X_0}{\lambda}} \cdot e^{\left(\frac{X_{\max} - X}{\lambda}\right)}, \tag{4.14}$$

where $N_{\max}$ is the maximum number of particles observed at depth $X_{\max}$, and X_0 and λ are primary mass and energy dependent parameters.

4.2. Photon induced extensive air showers

As for extreme energy CRs, the measurement of UHE photons is only feasible via indirect measurements exploiting EASs, using the same detection techniques as for hadronic showers. However, there are some specific features for photon-induced EASs that will be explained here.

The difference in phenomenology between γ induced showers and hadron-induced showers arises from the almost purely electromagnetic nature of the former ones, with a rather small hadronic shower component (just created by high energy photo-nuclear interactions early in the shower). The two types of air showers can thus be differentiated from each other (see figure 4.5). Furthermore, the relatively slow development of the EM cascade is addition-

ally slowed down due to the Landau-Pomeranchuk-Migdal (LPM) effect. At energies above $10^{19.5}$ eV photons have a probability of interacting in the terrestrial magnetosphere before they reach the atmosphere causing a so-called *preshower*.

4.2.1. Landau-Pomeranchuk-Migdal effect

The formulas for bremsstrahlung and pair creation in matter formulated by Hans Bethe and Walter Heitler [172] (Bethe-Heitler formula, eq. (2.15)) were shown to be inapplicable at high energy or high matter density in the early 1950s by Lev Landau and Isaak Pomeranchuk [276, 277], based on which Arkady Migdal developed a description which accounted for the effect of high energies or high matter densities [278], nowadays called LPM effect.

The Bethe-Heitler (BH) cross-section σ_{BH} for pair production by photons decreases due to destructive interference from several scattering centres and very small longitudinal momentum transfers arising due to ultra-relativistic EM interactions.

At energies below 10^{15} eV, the splitting rate for bremsstrahlung, Γ, can be approximated as $\Gamma \approx nv\sigma$, with n the density of the medium, v the relative velocity, and σ the splitting cross section [279]. The time these splitting processes take and the importance of this effect increase with energies, where for EM cascades above 10^{15} eV multiple Coulomb scattering that occurs between each splitting process causes a decoherence (dependent on the density of the medium [280]) which strongly suppresses the splitting rate [279], thereby subsequently significantly increases the shower length [281] and the atmospheric depth of the shower maximum (X_{max}). An event-to-event (X_{max}) fluctuation is increased as well, since for a shower that has a first interaction high in the atmosphere the LPM effect is less pronounced, leading to a faster shower development [207].

4.2.2. Preshower effect

In contrast to UHECRs which pass the geomagnetic field nearly undisturbed a primary photon with energies above 10^{19} eV may convert in the geomagnetic field into an $e^{\pm}$ pair, which subsequently emits synchrotron radiation, leading to an EM cascade above the atmosphere, the preshower [282, 283]. Upon their entry into the atmosphere, a multitude of electromagnetic sub-air showers is produced, with the individual primary particles having a lower energy than the initial UHE photon, and the pre-showers developing higher in the atmosphere, thereby reducing the average X_{max}. The preshower probability increases significantly with energy and depends on the strength of the magnetic field component perpendicular to the momentum of the primary photon. At a primary photon energy of 10^{20} eV the preshower would start at

about 1000 km above the ground, producing a few $e^{\pm}$ pairs and several hundred photons with a wide spread in energy entering the atmosphere [207].

In [284] it was reported that (1) via combining fluorescence techniques (to observe longitudinal profile of EAS) with detection by surface detectors (to observe the muon content) of the Pierre Auger Observatory the identification of unconverted UHE photons and of EAS produced by the preshower effect is possible, that (2) a strong directional dependence of the UHE photon first conversion would be present, and that (3) the detector location itself had a great impact on the conversion probability, as e.g. a larger transition region would be observed closer to the poles when simulating for a northern site of the Pierre Auger Observatory, what is by virtue of the larger strength of the local magnetic field.

4.2.3. Electromagnetic UHE shower properties

Although muon pair production is possible (compare with PP processes eq. (3.5) and (2.15)), such processes are suppressed by a factor $m_e^2/m_\mu^2 = 2.3 \cdot 10^{-5}$ with respect to the production of $e^{\pm}$ pairs [207]. Another process which is suppressed by almost two orders of magnitude, is the interaction of HE photons from the EM component with nuclei of the air, from which potentially secondary hadrons arise [207].

This muon supression further leads to differences in the lateral distribution of secondary particles on ground level of photon induces EASs. They are characterized by the deeper development in the atmosphere and the lack of a significant muonic component, while showing a steeper lateral distribution of secondary particles on the ground [285].

Another imprint of the insignificant hadronic and muonic component of photon showers is a smaller spread of the secondary particles on ground, which is further enhanced by the late development of EAS from HE photons. The combination of these characteristics results in photon-induced events having late particle arrival times with respect to a planar shower front (see figure 4.4) approximation at large distances from the shower axis [284]. This makes these showers distinguishable via algorithms, presented in chapter 6, aiming to separate photons from hadronic events.

Furthermore can both, the LPM and the preshower effect, be present and influence each other, e.g. will a preshowered event be less affected by the LPM effect since the energies of each sub-shower are lower [287].

While the LPM effect leads to a decrease of the cross section of PP with energy, the cross section of interactions of photons with target nucleons increases. The γ-nucleon interactions occur if the high-energy photon produces a quark-antiquark pair $q\bar{q}$, enabling hadronic interactions in the nucleus and resulting in a hadronic component of the shower initiated by a primary photon [288]. At the highest energies above 10^{20} eV even direct interactions are possible, with the cross section of these interactions even surpassing that of

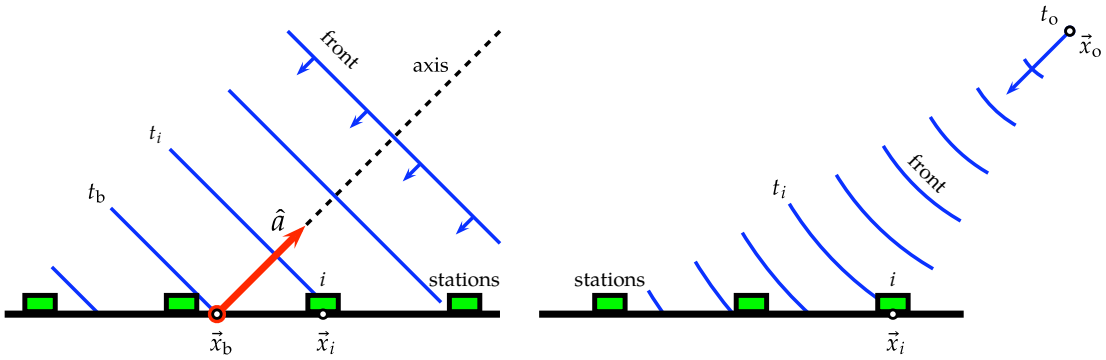


Figure 4.4.: Schematic view of (left) a plane shower front approximation, (right) a spherical shower front, where the front reaches the station i at position $\vec{x}_i$ at time t_i. Figure taken from [286].

PP for most target materials [288], and therefore EM EASs possibly develop significant hadronic components already at a few tenths of EeV.

Above a few TeV the actual cross section of photonuclear interactions is inaccessible for direct measurements and must therefore be extrapolated over several orders of magnitude in energy. The observation of only a small number of photon induced EAS at the highest energies would already greatly reduce the uncertainties of the estimated photon-air cross sections [207].

In fig. 4.5 the tracks of the secondary particles from the muonic, EM and hadronic shower components are shown for three simulated showers initiated by a photon, a proton and an iron nucleus with and energy of the primary particle in each case of 10^{13} eV using Cosmic Ray Simulations for Kascade (CORSIKA) [289, 290]. It can be seen that in all three cases a very strong EM component develops, while only in the case of proton and iron primaries, significant muonic and hadronic components are present.

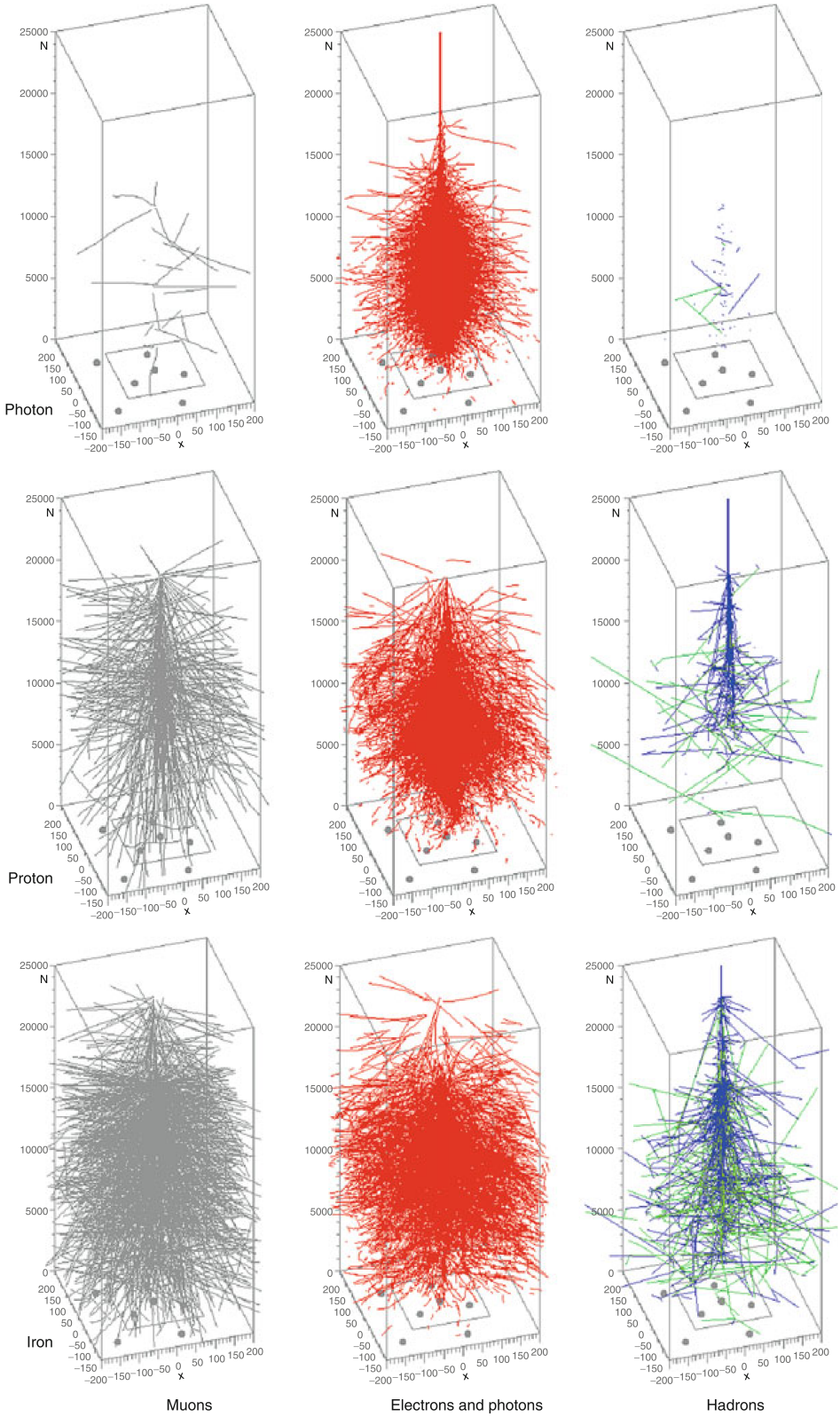


Figure 4.5.: Tracks of the secondary particles from the muonic, EM and hadronic shower components for simulated air showers, initiated by a photon, a proton and an iron nucleus. Simulations were done with CORSIKA and the height of each graph corresponds to an altitude of 25 km, and a width of 400 m. Figure taken from [291].

5. The Pierre Auger Observatory

The Pierre Auger Observatory was designed to deliver high statistics and high quality data on an unprecedented level about CRs at the highest energies ($> 10^{19}$ eV). For measurements in this energy region a large detector area was required, and to achieve high quality it was designed as a hybrid detector, combining a Surface Detector (SD) array with a fluorescence detector (FD), thereby providing a valuable measurement redundancy [292]. The observatory is located in the Argentinian Pampa Amarilla, near Malargüe, Mendoza Province, at an average altitude of $\sim$ 1400 m above sea level and covers an area of about $\sim$ 3000 km^2. A map of the detector is shown in figure 5.1.

Still under construction, the experiment started taking data in 2004, and was completed in 2008 [294]. Over the years, complementary detection systems were deployed to enhance the capabilities of the Observatory: extensions to lower energies with a more dense SD array and overlooked by three specific FDs called HEAT (as shown in fig. 5.1), the Auger Engineering Radio Array (AERA) to exploit the coherent radiation in EASs and demonstrate at a significant scale the physics capabilities of the radio technique, the Auger Muon and Infill for the Ground Array (AMIGA) to directly measure muons with buried scintillators [295]. New detectors are being deployed as a part of the Auger upgrade called *AugerPrime* (sec. 5.4). After a brief introduction of the fluorescence detector, the SD array, central for the presented work, is described.

5.1. The fluorescence detector

The FD is used to overlook the atmosphere above the detector field from four sites at the borders of the observatory and measure the longitudinal shower development by fluorescence light produced by EASs while passing through the Earth's atmosphere. Nitrogen (Ni) molecules in the air get exited by the charged particle component of the EAS which then emit fluorescence light in the UV range. 27 telescopes (4 sites with 6 telescopes each, plus 3 telescopes for HEAT) overlooking the SD array measure this fluorescence light, each covering a field of view of $30^\circ \cdot 30^\circ$ with 440 ($22 \cdot 20$) photomultipliers, whichs output signals are digitized every 100 ns. While the 24 FD telescopes are detecting fluorescence light at elevation angles in the range [1°, 31°], HEAT telescopes are detecting showers in the [30°, 60°] range, in order to focus on

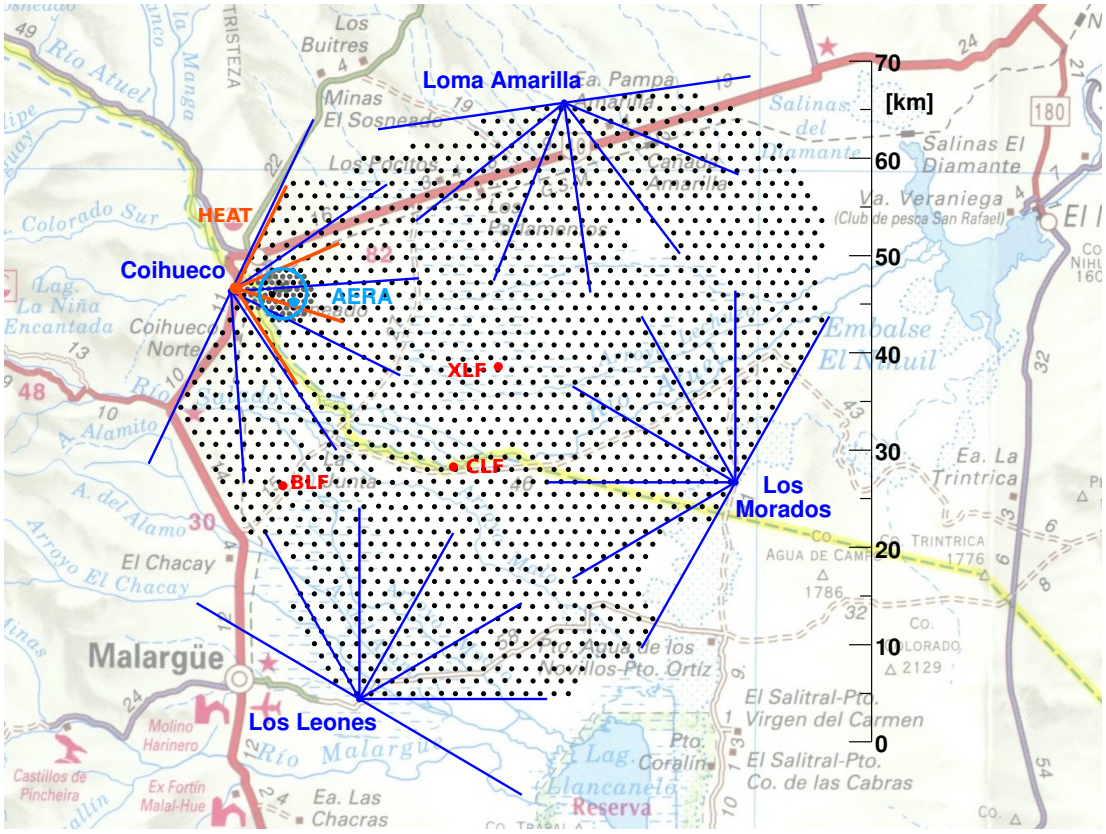


Figure 5.1.: A map of the Pierre Auger Observatory, where each black dot corresponds to one of the 1660 SD stations. The four FD sites together with the location of High Elevation Auger Telescopes (HEAT) are shown, and also the two laser facilities, CLF and XLF, near the Observatory center, together with the extensions Auger Engineering Radio Array (AERA) and Auger Muon and Infill for the Ground Array (AMIGA), and the balloon launching facility (BLF) [293].

lower energy events. Since the number of emitted fluorescence photons in an EAS is proportional to the EM energy loss in the atmosphere, it can be used to measure the total energy deposit of the charged shower particles, representing $\sim$ 90% of the primary particles energy, and therefore providing a nearly calorimetric measurement of the primary CR energy. The FD provides as well a very accurate determination of the shower maximum depth, and its data are crucial for investigation of the chemical composition of CRs. However, FD measurements are restricted to clear and moonless nights, resulting in duty cycle of $\sim$ 13%. To calibrate the FD, LASER shots are fired regularly during the FD operation, and additionally the atmospheric conditions above the Observatory are constantly monitored. More details can about the fluorescence detector of the Pierre Auger Observatory be found in [296].

5.2. Surface detector

A detailed description of the SD array is given in [297]. The SD array is built of 1660 WCDs, which are self-contained detector stations (WCD + local electronics, see figure 5.2), realized via a cylindrical, opaque polyethylene tank of 1.2 m height and 3.6 m diameter, filled with ultra-pure water. With a spacing of 1.5 km between the SD stations, the energy threshold for full trigger efficiency is $3 \cdot 10^{18}$ eV for zenith angles $\theta < 60°$, independent of the type of the primary particle initiating the EAS, see [298] and fig. 5.4. The SD station signals are transferred wireless, centralized and combined by the Central data acquisition system (CDAS) using a proprietary protocol.

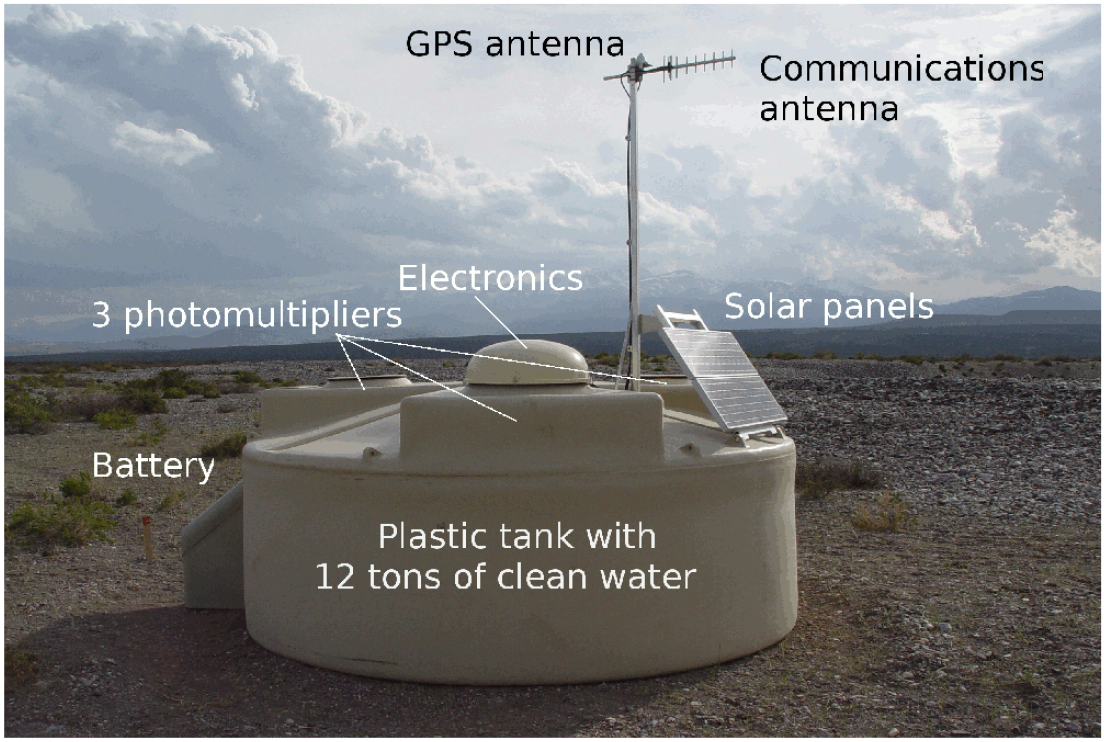


Figure 5.2.: A WCD station of the Pierre Auger Observatory with the various components of the system indicated. Figure taken from [299].

The WCDs sample the EASs at ground-level and are sensitive to both muonic and EM particles (including high energy photons converting into $e^{\pm}$ pairs in the water). The relativistic muons arrive with a mean energy of around 1 GeV, and electrons and photons with $\sim$ 10 MeV, producing a short Cherenkov light pulse in the water. In the WCD these light pulses (including multiple reflections) have a duration of about 100 ns. For stable signal measurements the walls of the tanks are covered with a Tyvek liner to diffuse the Cherenkov light, which, in combination with the ultra-pure water, increases and unifies the light collection of the detector.

Each WCD has three photomultiplier tubes (PMTs) of size 9 inch, which

are directed downwards into the water. The PMT signals are digitized by 10-bit flash analog–to–digital converters (FADCs) each 25 ns. A solar panel and battery allows each tank an autonomous operation. A GPS antenna provides timing and position information. Depending on the geometry and energy of a shower, the shower front triggers several tanks and produces a lateral distribution of measured signals across the SD array (c.f. figures 4.4 and 5.6), which in addition to the time structure of the signals contains physical information about the primary particle that induced the shower.

In addition, to assure a stable signal measurement, the physical conditions of each station and especially of the PMT data quality are monitored continuously in intervals of several minutes: Temperature, pressure, voltages, currents, battery status, water level, dynode-to-anode ratio (D/A), the vertical equivalent muon (VEM) peak, etc. [295].

To cope with large amounts of data and the bandwidth limitations of the wireless communication system, which must transmit to receivers at up to 40 km distance with a power consumption of less than 1 W [298], the recorded signals are transferred to the CDAS only if a shower trigger has been detected in three adjacent tanks simultaneously. Due to possible changes of the trigger thresholds with time, calibration quantities are continuously monitored for each station in the array. A summary about SD calibration and triggers is provided in the following.

5.2.1. Calibration

Since the recorded FADC traces also depend on the properties of a single SD station (i.e. the gain of the PMTs and the optical coupling of the PMTs to the water, the exact reflectivity of the inner liner, etc.), the signal measured for a single muon varies from station to station, the detectors are calibrated using low-energy background muons to correct for this bias.

For a common reference level between all WCDs, and to facilitate comparisons with detector simulations, each station regularly performs an automatic self-calibration. Therefore the counts of the FADCs are converted to units of the signal that would be deposited in a tank by a VEMs [295] in electronics units (i.e. integrated channels). The distribution of incident angles of background muons is dominated by vertical muons. A peak ($Q^{\text{peak}}_{\text{VEM}} = 1$ VEM in units of ADCT = ADC count $\cdot$ time bin [300]) in the distribution of the integrated FADC charge and the equivalent amplitude, $I^{\text{peak}}_{\text{VEM}}$ (in units of channels) [300], of a VEM-particle are determined to provide the reference pulse height needed for the triggering algorithms [295]. The atmospheric muons passing through each detector at a rate of $\sim$ 2500 Hz are used for a continuous online calibration each time a coincidence of signals with at least 5 FADC counts above the baseline level occurs (in all working PMTs), the signals are recorded and charge- and pulse height histograms are filled. Both of these calibration histograms contain two distinct peaks: one created by the distribution of low-

energy particles which accidentally satisfy the trigger requirements; the second due to muons traversing the detector - therefore used for the calibration. These background signals to build the histogram are collected for 60 s, then fitted (with a correction factor to account for the angles of the passing muons) and the peaks position is set as the VEM reference [295].

While $Q_{\mathrm{VEM}}^{\mathrm{peak}}$ can be directly converted to the charge of a VEM, the trigger levels of $I_{\mathrm{VEM}}^{\mathrm{peak}}$ rely on a peak estimate, which is defined by tuning its value until the calibration trigger rate peak is 70 Hz, otherwise this would require large dead times of the detector to remove the baseline. For this offline calibration of the signal the charge histograms for single PMTs, their sum and the pulse shape are saved, allowing a precision of the VEM unit calibration of $\sim 3\%$ [295].

Before the signal traces are averaged on the three PMTs, the VEM trace bin values (between a start-time and a stop-time) are summed up to a total signal of a station, scanned for signal fragments of consecutive bins above 3 FADC counts and merged into a single trace (with the start-time and stop-time assigned accordingly) [295].

5.2.2. The triggers

Triggers ([298]) can be seen as a filtering mechanisms, aiming to reduce the event rate from single station events and to retain as much usable air shower information as possible. The constraints of the wireless communication system for an event rate reduction on station level are realized by the first two levels of the five-level hierarchical SD trigger system: T1 and T2.

1. T1 and T2 triggers:
   A T1 level trigger can be produced by four algorithms applied to the raw SD data: Threshold (Th), Time over Threshold (ToT), Time over Threshold deconvoluted (ToTd), and Multiplicity of Positive Steps (MoPS). Only for the Th trigger there is a difference between the T1 and T2 level, while for the other algorithms the requirements of T1 and T2 are the same.

   - Th requires that all three PMTs of a station record simultaneously 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, where $I_{\mathrm{VEM}}^{\mathrm{peak}}$ is the pulse height value of a VEM. It selects stronger signals that are not necessarily spread in time, and in case of only one or two PMTs working, these values of the thresholds are increased, according to reduce the effect of random coincidences.
   - ToT requires at least two out of three PMTs to record a signal above $0.2 I_{\mathrm{VEM}}^{\mathrm{peak}}$ for 13 bins within a 3 μs window, selecting series of low signals spread in time.

In 2014, after a decade of SD observations, two new station-level triggers, called ToTd and MoPS, were added to the existing triggers. They were designed to be mostly sensitive to the low-energy EM component of EAS far from the shower axis and insensitive to the muons (which make up the dominant background), by identifying the low-amplitude, long-duration waveforms resulting from EM particles entering the water and rejecting the short spikes caused by single muons.

- ToTd is a refinement of ToT, implemented using the predictable time delay originating from the light reflecting on the liner surface, which prevents ToT conditions from being met by a small number of particles with large signals and thus long tails, via applying a deconvolution to the FADC traces to remove the tails and leaving only the peaks. This deconvolution reduces the signal in each FADC bin by a constant fraction of the signal in the preceding bin. Then applies a normalization term, followed by applying the triggering conditions of ToT to the deconvolved trace (with the same thresholds values as in the ToT description).
- In contrast to the other trigger algorithms, the MoPS trigger is completely independent of $I_{\mathrm{VEM}}^{\mathrm{peak}}$, but similar to the ToT and ToTd algorithms. MoPS selects signals produced by a series of low-energy particles, via selecting FADC traces that contain a certain number of positive steps, M, above a threshold within a moving time window, assuming that each of these steps corresponds to the arrival of a new particle in the detector. A positive step is defined therein as a sequence of bins in which the FADC trace increases with a minimum step size, y_{min}, to avoid statistical fluctuations and a maximum step size, y_{max}, to avoid muon-like signals. In addition a veto to prevent the counting of additional steps due to fluctuations in the tails of large peaks is included. A MoPS trigger requires at least two of the three PMTs of a station to satisfy the condition.

2. T3 trigger:
First, all T2 triggers, including their time stamps, are sent to the CDAS each second, where the global trigger condition T3 is constructed. This array trigger condition corresponds to the shower candidates and is based on spatial and temporal conditions of the T2s, where all clusters of T2 signals within a time window of ± 25 μs are examined for spatial coincidences. T3 has two modes (see fig. 5.3 a) and b)): (1) when at least three closely clustered stations trigger with at least two ToT triggers; (2) when at least four closely clustered stations with any T2 level trigger are found. With a T3 rate of ~ 0.1 Hz CDAS records all the traces from the T2 and T1 triggered stations within six crowns of the central station and applies two additional offline triggers [298].

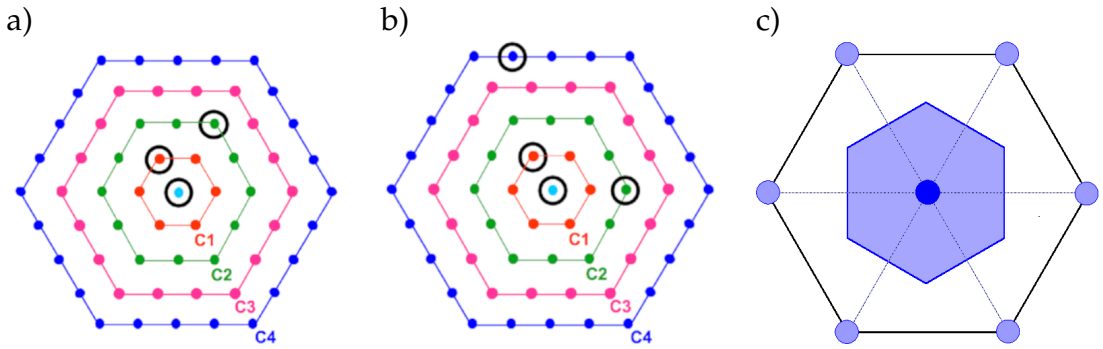


Figure 5.3.: Example of T3 configurations: **a)**: the 3-fold T3 mode ToT-2C1&-3C2, **b)**: the 4-fold mode 2C1&-3C2&-4C4, where C1, C2, C3, C4 indicate the first, second, third and fourth sets of neighbours, respectively, at 1.5, 3, 4.5 and 6 km from a given detector. **c)** Scheme of an hexagon of detectors: the elemental hexagon cell where a cell a_{cell} is the shaded area around the central detector. Figures taken from [298].

3. T4 trigger:
   It is a physics trigger, used to test if the signal timings can be fit to a plane shower front moving with the speed of light and makes sure that 99% of the stations containing a physical signal from the shower are kept; while keeping the number of random coincidences to less than one per day over the full detector array [298].

4. T5 trigger:
   Finally, the T5 trigger checks for 6 active stations surrounding the station with the highest signal to assure a good CR event detection and reconstruction accuracy; it is therefore called 6T5. Some studies have less requirements, as e.g. only four or five working adjacent stations, for which the corresponding triggers are called 4T5 or 5T5 respectively. For the cases when the shower arrived close to the border of the array, the fiducial trigger T5 was introduced [295, 298].

Further types of local triggers are the *scaler* triggers, which have a very low signal threshold and provide information on space weather via studying counting rates of low energy shower particles induced by primary CRs in the GeV energies [301], and the *calibration* triggers, which have a low signal threshold and are used to control the level of CR intensity at the observatory as well as to record the stations calibration histograms [301].

For air showers with zenith angle below 60° and energies above 3 EeV the trigger efficiency of an EAS event reaches almost 100% for the most conservative case of a primary proton, as can be seen in figure 5.4.

In comparison, photons reach full efficiency at about 10 EeV by reason of the delayed development of photon-induced showers due to LPM effect. The probability to trigger a single station reaches 100% above 10 VEM and the

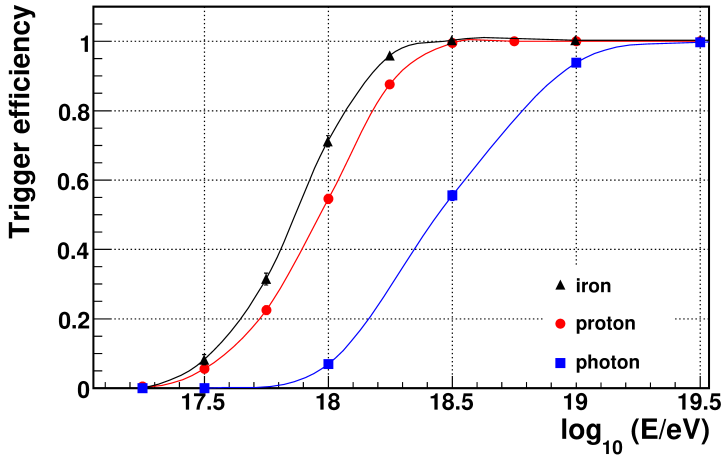


Figure 5.4.: SD trigger efficiency as a function of the simulated (true) energy, integrated for showers with $\theta < 60°$, for different primaries. Figure taken from [298].

lateral dependence for a fixed primary changes with energy and inclination angle [298].

5.2.3. Aperture and Exposure of the Surface Detector

The total SD aperture depends on the number of hexagons of active detector stations, where each of these hexagons consists of one central and six surrounding detector stations, and has an elemental cell with a size of $A_{\text{cell}} = 1.95\ \text{km}^2$. The aperture per hexagon, a_{cell}, is obtained by integrating A_{cell} for air showers with a zenith angle of θ over the used solid angle range:

$$a_{\text{cell}} = \int A_{\text{cell}} \cos(\theta)\, \mathrm{d}\Omega\,, \tag{5.1}$$

together with the assumption of an isotropic CR flux resulting in $a_{\text{cell}}(0° - 60°) = 4.59\ \text{km}^2\ \text{sr}$ and $a_{\text{cell}}(30° - 60°) = 2/3\, a_{\text{cell}}(0° - 60°) = 3.06\ \text{km}^2\ \text{sr}$. The accuracy of the aperture determination is found to be better than 1% [298].

The exposure A of the SD can then be calculated geometrically from the number of elementary cells N_{cells} (see figure 5.3 c)) with a 6T5 trigger and for events with full trigger efficiency (E > 3 EeV and $\theta < 60°$) for a period Δt, whereat each elementary cell contributes with a_{cell} to the exposure [298]:

$$A = N_{\text{cells}} \cdot a_{\text{cell}} \cdot \Delta t\,. \tag{5.2}$$

Furthermore, are time periods in which the SD array was not working properly labelled as *bad periods* and are removed from the aperture integration

[298].

5.2.4. Data reconstruction

The SD array samples the lateral distribution of EASs at discrete points, with the shower particle composition as a function of distance to the shower axis consisting of different components as described in sec. 4.1.

The reconstruction of an event from SD data can be broken down into three stages: (1) determination of the geometry of the EAS from the timing information and the positions of the involved single SD stations; (2) calculation of the lateral profile of the shower from the signals recorded by each station; (3) obtaining the energy of the primary particle initiating the EAS from the lateral profile. An example of an event falling within the SD array in top-down view is given in figure 5.5 a), with the projection of the shower axis on the ground ending at the impact-point of the shower-core. SD stations are coloured according to their trigger time. To illustrate the concept of the hybrid detection of the Pierre Auger Observatory, a CR event detected by all four FD telescopes and the SD array is shown in figure 5.5 b).

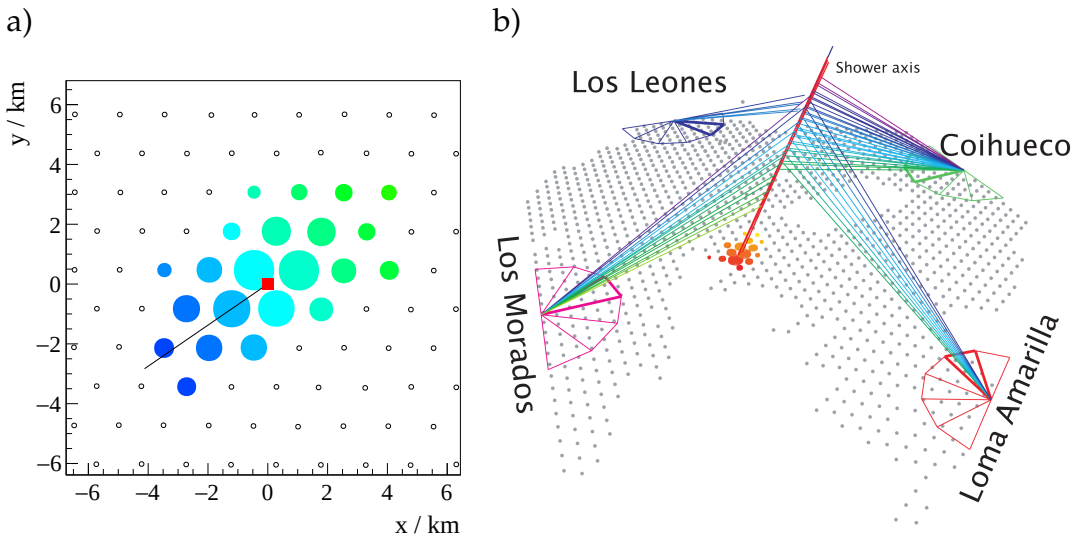


Figure 5.5.: a) A top-down view, zoomed-in on an event falling within the array, with the projection of the shower axis on the ground represented by the black line, ending at the impact-point of the shower-core (red square). The SD stations are coloured according to their trigger time (blue is early, green is late) and their area proportional to the logarithm of the signal amplitude. Figure taken from [286]. **b)** Illustration of a hybrid event, with the pixels in the camera of the FD tracing the shower profile – specifically, the energy loss of the shower as a function of its penetration in the atmosphere, together with the SD array detecting the particles from the same shower. The red line shows the trajectory of the shower, the shower axis.

Since each experiment and detector system has its own characteristics, observables specific for the Pierre Auger Observatory, which naturally arise from data reconstruction methods and used in the analysis presented in this thesis, are introduced further in this section. After a physical motivation of these observables with respect to their photon-hadron discrimination power, a more technical description on calculation and implementation is given.

5.2.4.1. Shower geometry

The geometry of an EAS is derived from a fit of the SD station signal times to a shower front moving with the speed of light. A schematic view of the shower geometry within a coordinate system showing "early" vs. "late" shower regions (connected to the different amount of atmosphere traversed by the particles reaching the SD stations in each region) is given in figure 5.6.

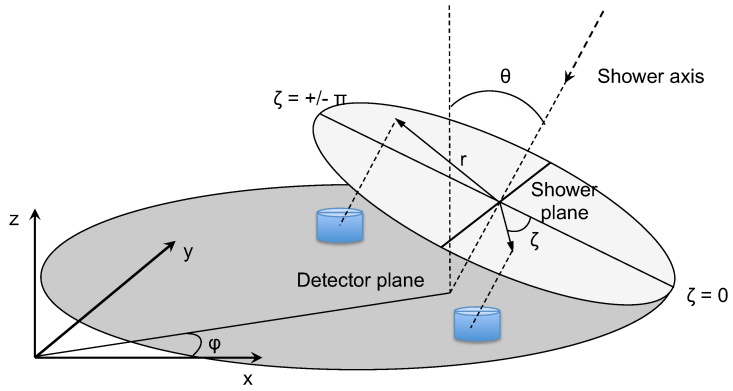


Figure 5.6.: Schematic representation of the shower geometry for incoming direction of the primary particle in two regions: "early" for $|\zeta| < \pi/2$, and "late" region for $|\zeta| > \pi/2$. Figure taken from [302].

If the number of triggered detector stations is sufficient, a spherical shower front is used to estimate the arrival time of the shower front particles at a given position, instead of a simple planar front (compare fig. 4.4) [295]. Fitted are the shower axis, the time at which the core hits the ground and the radius of curvature. The center of this spherical shower front is interpreted as a virtual point of shower origin and used together with the shower impact point on the ground to derive the arrival direction of the air shower. For events with more than three stations the angular resolution is better than 1.6°, and better than 0.9° for events with more than six stations [303].

- Radius of curvature
  Since photon induced EAS develop deeper in the atmosphere compared to nuclei primaries, they will have larger time delays Δt, and hence a smaller radius of curvature at fixed distance r to the shower axis compared to showers originating from larger heights (heavy particles) as

shown in figure fig. 4.4. The lack of muons in photon induced showers even amplifies this effect, because shower muons can reach the ground from still higher altitudes, further reducing the time delay.

The radius of curvature R as an event parameter is obtained via fitting the trigger times, t_i to a spherical model by minimizing

$$\chi^2 = \sum_{i=1}^{N} \frac{[c \cdot (t_i - t_0) - (R\vec{a} - \vec{x}_i)]^2}{c^2 \sigma_t^2} , \tag{5.3}$$

where c is the speed of light, t_0 is the arrival time of the shower to the center of curvature, $\vec{x}_i$ are the locations of the detector stations relative to the shower core, $\vec{a}$ is the unit vector along the shower axis, and σ_t is the uncertainty of the shower arrival time [304].

The Pierre Auger collaboration uses two reconstruction frameworks: the *Observer* and the *Herald* [286]. For this work the Observer reconstruction has been used.

The Observer reconstruction approximates the shower development as starting at time t_0 from a virtual point of origin $\vec{x}_0$ and propagating towards the ground in the shape of a sphere, concentrically inflating with the speed of light (fig. 4.4 right) with the advantage that this spherical fit can be performed without any prior knowledge of the impact point $\vec{x}_c$ or the shower axis [286]. The development of the shower front is described via four free parameters. For low-energy events having a station multiplicity of only three or four, there are not enough degrees of freedom (DoF) to solve for the shower-front curvature, and therefore a curved model with an R_0 fixed to a parametrization optimized using events with a larger number of stations is used for events with less than five triggered stations [286].

- Rise-time of the signal
  Also the spread in time of the signal can be used to discriminate between photon and hadron induced showers, where a larger spread of the signal intensity is expected in case of deep developing photon primaries, for a given distance to the shower axis. The rise time for a surface detector at distance r_i from the shower axis is defined as the time $t_{1/2}(r_i)$ it takes the total signal (i.e. the integrated WCD trace) to increase from 10% to 50%. It was already introduced in 1973 by Lapikens et al. [305]. When concerning the particles produced at a fixed path length ΔH and arriving at distance r from the shower axis at two different production depths $X_1 < X_2$ the rise time of the signal is increased for photon primaries developing deeper in the atmosphere.

  Since secondary EAS particles can be created at any time from the top of the atmosphere to the shower maximum, a particle that is created later

will arrive with a time delay as compared to a particle created closer to the initial interaction. A sketch of how such geometrical effects affect the temporal spread of the muons at a detector can be seen in figure 5.7.

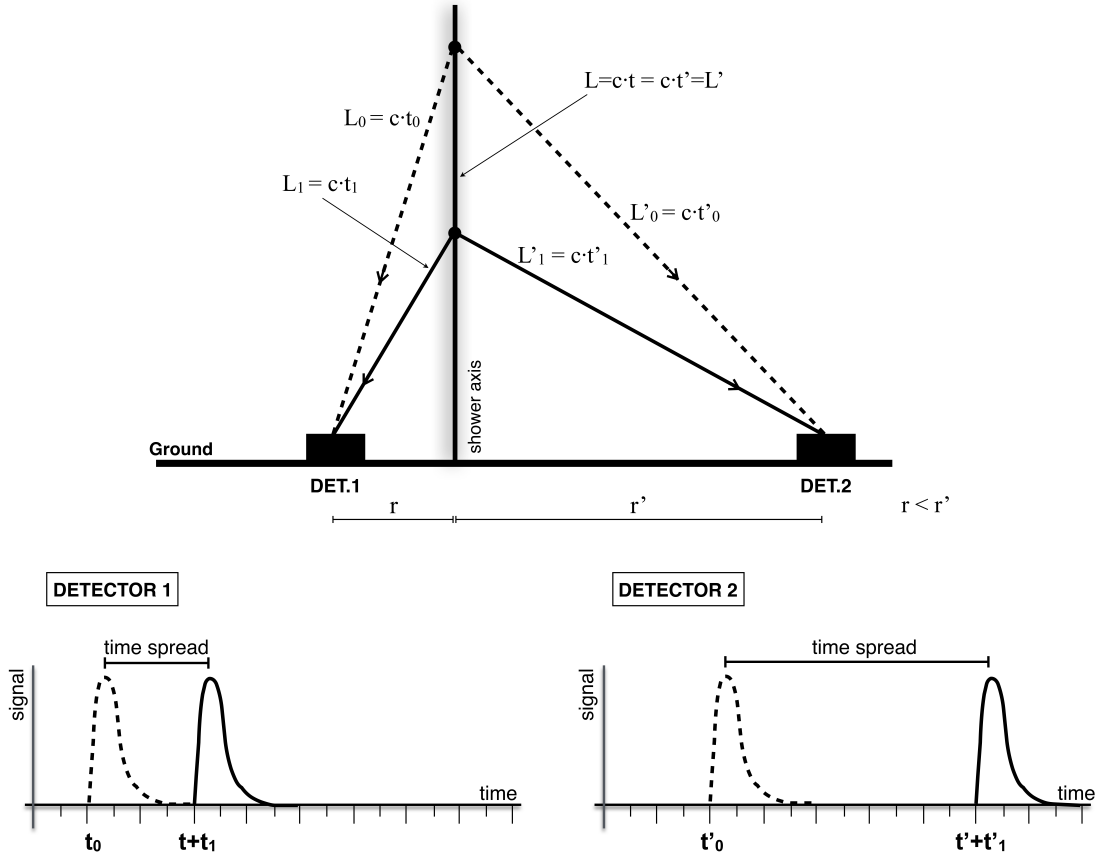


Figure 5.7.: Influence of geometrical effects on detector signals. Figure taken from [306].

Old showers reach their maximum earlier than young showers, and thus have a smaller time spread in the signal of SD stations. This in turn leads to showers from heavier primaries producing a smaller spread of arrival times and possessing a higher muonic content at SDs compared to lighter primaries. Muons, which are mostly produced early in the atmosphere, deposit a larger amount of energy in the detectors, and therefore appear in the SD signal as sharp peaks. Secondaries from EM induced showers on the other hand typically travel a shorter distance and represent the bulk of the SD signal. Furthermore, since heavier primaries have a larger muon content, the signal will have sharper peaks and thus a shorter value of the rise-time. This makes the rise-time parameter sensitive to both muons and EM particles detected by SD stations, and thereby very useful for discriminating between primary species. Also the geometrical asymmetry (fig. 5.6), arising from stations with different azimuth angles ζ around the shower core and increasing with the

shower inclination (early and late region, see fig. 5.6), has to be taken into account for the rise-time [302].

The rise-times of SD stations are further combined to an event rise-time parameter via several methods:

- ◇ The rise-time at 1000 m, $t_{1/2}(1000)$ or "RT1000", without the use of a benchmark function, is evaluated using a second degree polynomial, which is (χ^2-)fitted to the station rise-times as a function of the distance to the core and then evaluated at 1000 m from the shower core:

$$t_{1/2}(r) \;=\; 40\,\text{ns} \;+\; ar \;+\; br^2 \tag{5.4}$$

  Nevertheless, an approach like this is only practical at high energies, since high multiplicities are needed to estimate the rise-time at 1000 m by extrapolation [306].

- ◇ Another possibility to characterize an event by a single rise-time, via obtaining a large sample of data over a wide range of energies, can be done via description of the rise-times as a function of distance in a narrow range of energy, called "benchmark" functions. For this purpose, first a correction for the asymmetry effect is applied, then rise-times at particular SD stations are compared with the relevant times from the benchmark, $t_{1/2}^{\text{bench}}$, in units of the accuracy with which they are determined. The benchmarks are zenith-angle dependent. The term "Delta method" has been chosen for this approach. In short, for each measurement of $t_{1/2}$ at a single detector, i, an estimate of

$$\Delta_i \;=\; \frac{t_{1/2} \;-\; t_{1/2}^{\text{bench}}}{\sigma_{1/2}} \tag{5.5}$$

  is made. Then each shower is characterised by the average of these estimates for the N selected event stations Δ_s:

$$\Delta_s \;=\; \frac{1}{N} \sum_{i=1}^{N} \Delta_i \,. \tag{5.6}$$

  Further details on the Δ_s method, and why the choices for the benchmarks are most effective in dealing with signals from both high-gain and low-gain FADC traces of SD station PMTs for an extended dynamic range, are given in [306].

- ◇ A further way to obtain information from SD station rise-times, which was tested in this work, is done via using another rise-time relative observable, based on the concept of the asymmetry of the rise-time, $(\sec\theta)_{\text{max}}$ [302]. The relative rise-time called Δ_R addresses the issues of having low-end tail structures in the sig-

nal as well as the dependence on the zenith angle, with a similar treatment to the Delta method Δ_s, following similar steps in the calculation, but excluding their calculation of uncertainties. The dependence on the zenith angle is accounted for via a binning of the zenith angle, and the fitting with benchmark functions corrects for the dependence of the rise-time on the distance from the shower axis. A fit for a zenith angle bin applied to the Observer *v13r0* data production, the so-called SD burn sample, as used in this work, is shown in figure A.1. Similar to eq. (5.6), Δ_R is simply the arithmetic mean over all stations relative rise-times involved in a single event $\Delta_R = \frac{1}{N} \sum_{i=1}^{N} \Delta_i$. Further details on this method can be found in [302].

5.2.4.2. Lateral Distribution Function

After the geometry is known, the lateral distribution of the signals measured in the SD stations can be calculated. Describing the measured signals S_i, in the individual SD stations as a function of their distance r from the shower axis, the Lateral Distribution Function (LDF) uses a functional form approximation to the lateral air shower profile.

The Observer reconstruction uses a slightly modified Nishimura-Kamata-Greisen (NKG) function [307–309]

$$S(r) = S(r_{\text{opt}}) \left(\frac{r}{r_{\text{opt}}} \right)^{\beta} \left(\frac{r + r_1}{r_{\text{opt}} + r_1} \right)^{\beta+\gamma} , \tag{5.7}$$

where the shape parameters β and γ define the steepness of the LDF [286]. The parameter $S(r_{\text{opt}})$ is an estimator of the shower size, 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 [310]. $r_1 = 700$ m and $r_{\text{opt}} = 1000$ m are fixed for SD and the parameters β and γ are obtained from a parametrization depending on $S(1000)$ and θ. Therefore $S(1000)$ and the impact point of the shower on the ground are the only remaining free parameters which can be obtained from the fit, with an example of a LDF shown in fig 5.8 a), and an example of $S(1000)$ comparing photon, proton and Auger SD data distributions to each other can be seen in the appendix A.1.3. An example for the distributions of the LDF β parameter for photon, proton, and the Auger SD burn sample can be seen in the appendix A.1.2.

In the next step the fit to the LDF form is done via maximizing the log-likelihood $\ln \mathcal{L}$, which is composed as the product of probabilities P over the coordinates of the shower impact point $\vec{x}_c$ and the size $S(1000)$ [286].

Although in reality the showers are asymmetric due to a combination of the longitudinal evolution and geometrical effects related to the angles of incidence of the particles on the stations, the LDF model assumes that the

deposited signals in the stations are rotationally symmetric around the shower axis. This results in a corresponding shift of the impact-point and is also responsible for a very small, yet systematic difference in arrival directions [286].

- Number of Candidate Stations
  Another SD-related observable that can be used to search for photon candidate events is the **number of candidate stations**. The steeper LDF and the lower number of muons in photon induced EAS leads to a smaller event footprint, or in other words to a lower number of triggered SD stations than for hadronic induced EASs of the same energy and geometry. This simple parameter can therefore be used as a photon-sensitive separation parameter in data analyses. Candidate stations are SD stations that were not rejected during the reconstruction process by e.g. the absence of trigger or GPS data, bad calibration or belonging to the stations that are originally not triggered by the current EAS.

5.2.4.3. SD Energy Reconstruction

Geometry effects and the attenuation of the secondary particles from the EAS in the atmosphere decrease the value of $S(1000)$ with increasing zenith angle and since the more inclined showers are sampled at a later shower age, the more-attenuated EM cascade leads to a smaller observed shower size. This bias is minimized via conversion to a zenith-independent quantity S_{38} [286], the equivalent shower size if it had arrived at the median zenith angle of 38°, based on the Constant Intensity Cut (CIC) method [275, 311]:

$$S_{38} = \frac{S(1000)}{f_{CIC}(\theta)} , \tag{5.8}$$

with the third degree polynomial

$$\begin{aligned} f_{CIC}(\theta) &= 1 + ax + bx^2 + cx^3 \quad , \text{ with} \\ x &= (\cos\theta)^2 - (\cos 38^\circ)^2 , \end{aligned} \tag{5.9}$$

with $a = 0.980 \pm 0.004$, $b = -1.68 \pm 0.01$ and $c = -1.30 \pm 0.45$.

Via well measured hybrid events that allow for both a full SD reconstruction and a precise measurement of the calorimetric energy with the FD, the energy of the primary particle E_{FD} is calculated from S_{38} by a simple power law relation:

$$E_{FD} = A \cdot (S_{38}[\text{VEM}])^B , \tag{5.10}$$

with $A = (1.90 \pm 0.05) \cdot 10^{17}$ eV and $B = 1.025 \pm 0.007$, as shown in figure 5.8 b), where eq. (5.10) is fitted to data events with $E_{FD} > 3$ EeV using a maximum likelihood fit [73].

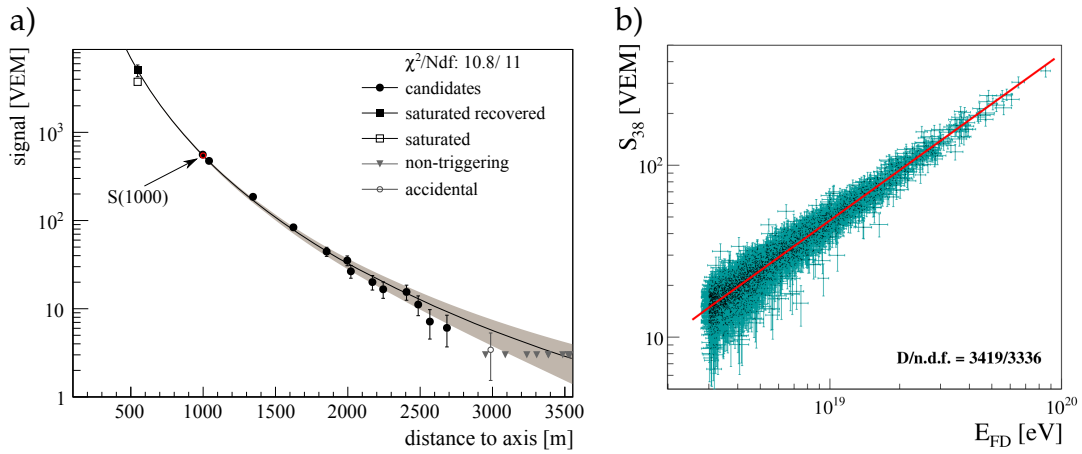


Figure 5.8.: a) Example of a LDF of a SD event and its NKG type parametrization (eq. (5.7)). Figure taken from [295]. **b)** Correlation between the SD (S_{38}) and FD energies. Figure taken from [73].

From the hybrid measurements the shower energy can be estimated from S_{38} with a statistical uncertainty of less than about 16% and a systematic uncertainty (dominated by the uncertainty in the absolute calibration of the FD) of about 14% [312].

Due to the unique properties of photon induced EAS, this energy reconstruction method introduces an underestimation of the energy when applied to photon primaries, and needs to be adapted to account for their different shower development [313], as developed (briefly discussed later) over the years [73].

Distributions of the reconstructed SD energy for photon, proton and Auger data sets used in this work is shown in the appendix A.1.4, and a comparison of $S(1000)$ vs. the reconstructed energy in a scatter plot is given in A.1.5.

5.3. The $\overline{\text{Off}}\underline{\text{line}}$ framework

The $\overline{\text{Off}}\underline{\text{line}}$ [314, 315] Auger analysis framework, of which the Observer event reconstruction is an essential part, was originally devised when the only existing systems were the SD and FD. Since then, it has been extended to handle the various extensions without requiring dramatic framework changes. The data reconstruction and simulation procedures used in analyses of the Pierre Auger Collaboration make use of this internally developed software framework, which has been designed to accommodate contributions from a large number of physicists developing C++ applications over the long lifetime of the experiment, incorporating an essential feature - a "plug-in" mechanism. It is composed of the main components of an event-based data structure, a time-dependent detector description and a collection of *modules* containing physics-related algorithms [315]. Furthermore does $\overline{\text{Off}}\underline{\text{line}}$ have external de-

pendencies which are not distributed along with it, but have to be installed separately or via the Auger Package Environment (APE) tool.

Simulations as well as data reconstruction of EASs are executed by running a sequence of modules, which include the reading of input data, simulation and reconstruction algorithms. Finally the reconstructed data is written out. For most types of analyses (depending on the detector components used and air shower types) a set of standard module sequences and configurations is provided, whereat each module, and the module sequence itself, can be configured using XML files.

The output of Offline is stored in Advanced Data Summary Tree (ADST) files, which is a standalone package based on the ROOT [316] toolkit. It provides several features, like a graphical display for reconstructed event properties, the *EventBrowser*, and offers high-level analyses on ADST-files directly without the need to modify any code and perform a full reconstruction of the raw data.

Simulations of the response of the WCDs of the SD array are performed using the Geometry and Tracking (Geant4) framework [317, 318]. As input air shower simulations, the Offline framework can use e.g. those created with CORSIKA [289, 290].

5.4. Upgrade of the observatory *AugerPrime*

The ongoing major upgrade of the Pierre Auger Observatory is called *AugerPrime* [319] and is intended to extend the physics capabilities of the Observatory with the aim to achieve further insights into the open questions about UHECRs (as described in ch. 1, 2 and 3).

The upgrade comprises installation of a surface scintillation detector (SSD) and a radio detector (RD) on top of each of the 1660 WCDs. Additionally a fourth, small PMT (sPMT) will extend the dynamic range of the WCDs, together with and upgrade of the SD station electronics which will enable a higher sampling rate of the FADCs of 120 Mhz (compared to the current one of 40 MHz), leading to a time resolution of 8.3 ns [319]. Moreover will AMIGA be finalized with underground scintillation detectors next to 61 SD stations of the SD-750 m array. Also, via reducing the supplied high voltage of the camera PMTs the duty cycle of the FDs will be extended by $\sim$ 50% compared to the current one of 19% [319].

In combination with the detector upgrade, also the EAS reconstruction will become more precise, and new composition-sensitive observables will be added. The main increase in performances of the upgrade in terms of mass composition studies relies on a more exact measurement of the muonic and EM components on a shower-by-shower basis, since the muonic component of the shower is a key observable for primary mass discrimination via interrelation between the shower development stage and the muon production in EAS.

The expected enhanced discrimination power between the muonic and EM signals of the upgraded SD stations is also particularly promising from the experimental point of view for photon studies. Indeed, AugerPrime will allow for event-by-event measurement of the muonic shower component, which will be very beneficial for searching for muon-poor photon-induced showers.

6. Applying machine learning for UHE photon search

The mass composition analyses of extreme energy CRs rely on simulations of CR-induced EASs, which implies some assumptions on the hadronic interactions of CRs at the energies inaccessible at man-made accelerators. When searching for UHE photon events in CR data, in the absence of an UHE photon beam, the simulations of air showers from UHE photons can be confronted to the real CR data events. Though a priori presenting less uncertainties from the side of MC simulations – as photon interactions are simpler than the hadronic ones – the photon analyses face another challenge. Namely, due to shower to shower fluctuations of CRs – especially of protons – CR events can mimic UHE photons. In such situation, advanced analysis methods, such as ML, can be of great help for mass-composition identification of individual CR events.

This chapter starts with the description of simulations: of air showers (section 6.1), and of the Pierre Auger SD response (section 6.2) to EASs. The reconstruction of data events and the data sample used in the analyses of UHE photon content are outlined in section 6.3. The review of the applied ML techniques for UHECR event classification is provided in section 6.5. The details of the analyses are presented in section 6.6.

6.1. Air shower simulations

The properties of EASs in the atmosphere have been evaluated in this work via detailed MC simulations with the already introduced (sec. 5.3) CORSIKA [289, 290] tool.

CORSIKA tracks the primary particle through the atmosphere, calculates its interactions with the air nuclei or its decay and repeats this process for each created secondary particle (except those discarded by the *thinning* [290] - see below). Secondary particles that are created in an EAS are recorded, including type, energy, location, direction and arrival times. The program is built of several blocks, in particular:

- A general program handles in- and output, decays of unstable particles, the deflection in Earth's MF and ionization energy losses.
- Hadronic interactions, with the possibility to chose the interaction model
  - ◇ at higher energies

- ◇ at lower energies

- Electromagnetic processes, which takes into account interaction and transport of $e^{\pm}$ and photons.

The "thinning" procedure [320] copes with the huge number of particles in EASs induced by UHECRs, since it would not be possible to generate, follow and store them within a reasonable time. In this case a particle is only followed until its energy reaches a predefined threshold, below which only a small predefined fraction of particles is followed. The thinning procedure provides a subsample of particles with statistical weights, allowing to obtain unbiased estimators of a full population of shower particles.

To obtain a fair description of the particles entering a ground detector, a backward "unthinning" or "resampling" procedure, that rescales the weight of the "thinned" shower particles in order to reproduce their local flux, along with all relevant distributions in the parameter space (energy, direction, time), is performed [321].

6.1.1. Utilized CORSIKA air shower libraries

As an input to detector simulations with Offline, libraries of simulated HE EAS induced by primary photons, protons, and iron nuclei, created within the Pierre Auger collaboration with the CORSIKA program were used. The following air shower libraries have been used, indicating the CORSIKA version used for their production:

- Photon_18.5_20.5_1e-6_CORSIKAv75600 (In total about 30k events)
- Iron_18.5_20.5_1e-6_CORSIKAv75600 (In total about 8.4k events)
- Proton_18.5_20.5_1e-6_CORSIKAv75600 (In total about 1k events)
- Photon_17.5_19.5_CORSIKAv75600 (In total about 10k events)

Highlighted in blue is the primary particle type. Further properties of the simulated shower input files, for the first three libraries are:

- Energy range: $18.5 < \log(E/eV) < 20.5$
- Spectral index: -1
- Zenith angle: $0° < \theta < 70°$
- Thinning: 1e-6
- CORSIKA version: v75600
- High-energy hadronic interactions: EPOS-LHC

- Low-energy hadronic interactions: Fluka (v. 2011.2c)

- Including preshower option in the case of primary photons

Photon simulations need a special treatment, since the hadronic energy reconstruction is not applicable to them due to their different shower development as described in ch. 3 and sec. 4.2.

6.2. Detector response

The detector response of the Pierre Auger Observatory's SD array has been simulated using the $\overline{\text{Off}}\underline{\text{line}}$ framework, introduced in sec. 5.3. $\overline{\text{Off}}\underline{\text{line}}$ starts from reading secondary particles of a CORSIKA shower at observation level, and simulates the time-dependent SD response, using Geant4. To enlarge the statistics, each CORSIKA shower has been resampled several times with random core locations. After a successful detector simulation the events are reconstructed. The total sequence of detector simulation and SD event reconstruction steps is done by the $\overline{\text{Off}}\underline{\text{line}}$ module *SdSimulationReconstruction*. The $\overline{\text{Off}}\underline{\text{line}}$ module sequences in relation to the two versions used in this work are given in the appendix B.1.1 and B.1.2. The key simulated and reconstructed parameters, including high level physics variables needed for physics analysis including additional low level data to facilitate the development of data selection cuts, are stored into the already introduced (sec. 5.3) ROOT based file format called ADST.

6.3. Data reconstruction

Real, raw data are similarly reconstructed with $\overline{\text{Off}}\underline{\text{line}}$ (cf. sec. 5.3), using a module sequence for reconstructing only the SD response:

```
<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>
```

The purpose of the most important modules is briefly summarized: the *EventFileReader* module reads in the raw files containing the shower particles at detector level;
EventCheckerOG checks the integrity of the file containing the event data to be reconstructed;
TriggerTimeCorection assigns proper time delays between different versions of the programmable logic device (PLD) trigger code, whereat the actual shift by the time offset is performed in the SdCalibrator;
SdCalibratorOG performs the calibration of the traces and start time computation;
SdPMTSignalShapeQualityChecker ensures high data quality of FADC signal shapes;
SdStationPositionCorrection attempts to correct for time offsets which are a result of a wrong position set in the radio station GPS unit when switched to the position fixed mode of operation;
SdEventSelectorOG carries out different actions, all of them related to the candidate station selection and the T4 and T5 trigger level, like inter alia lightning event removal, discarding of particular Surface Detector stations such as Engineering Array stations, discarding of Doublet and infill stations;
SdEventPosteriorSelector is based on a certain official event selection;

SdCompositionParameters calculates composition sensitive variables for an event; *DLECorrectionWG* corrects Direct Light Effects in the FADC traces;

6.4. Experimental data - SD burn sample

An Auger collaboration policy prevents possible negative impact on future data analyses by requiring the use of one common sub-sample of the data, which is defined for all photon analyses, in order to avoid i) the biases from previous results, as well as ii) the application of penalty factors, which would very fast decrease the significance of the results. The application of the analysis for testing and proving functionality on data has been performed on such a selection of $\sim$ 5% of all the SD data collected in a time period from 2004 to 2018, called *burn sample*.

For this work the SD data production used for the 36$^{\text{th}}$ International Cosmic Ray Conference (ICRC) in 2019 has been used. The SD burn sample is composed of specific events from the total SD dataset, composed of two parts:

- (A) from 01.01.2004 to 14.05.2013 – random extraction from a pre-selected set, with criteria: T4 & 6T5, $\theta_{\text{rec}} < 60^{\circ}$, Reconstructed LDF, $E_{\text{rec}} > 3$ EeV[1]
- (B) from 15.05.2013 to 30.06.2018 – with the event selection SdId % 50 == 0 applied, where SdId denotes an event number for events detected solely by Auger SD detector.

The result are around 100,000 SD events in total in the SD burn sample, approximately 55% in part A, 45% in part B.

[1] It was originally created with the CDAS version v1r0, which contains the PMT monitoring information needed to reject faulty PMTs, reconstructed with Offline v2r9p3 - revision 23943 corresponding to the version used for the ICRC2013

6.5. Machine learning techniques in data analysis

A few observables for the Pierre Auger Observatories SD array with discriminating power between photons and hadrons have been already introduced in section 5.2.4. Nevertheless, no single SD array observable has been found yet to satisfy the statistical and quality requirements for a discriminating statement about the primary particle being a photon. An analysis combining several observables was therefore chosen to provide a classification of primary particles being photons.

In this thesis several ML techniques have been tested in a MVA in order to classify events, measured with the SD array of the Pierre Auger Observatory, as photon candidate. These techniques were implemented and utilized via the Toolkit for Multivariate Analysis (TMVA) 4.2.0 [322] which provides a ROOT-integrated environment for ML. The best performing ML method was chosen to be realized in the "real-time" event stream of photon-candidate events to AMON, functionality has been proven (see sec. 6.6), and this analysis has been proposed for implementation to, and accepted by, the Pierre Auger Collaboration.

MVA is based on the principles of multivariate statistics, a subdivision of statistics. Many different MVA models exist, each with its own type of analysis, and one cannot know a priori, which method will be the best performing for a given problem. The several ML methods tested in the MVA in this work were compared to each other, and are briefly introduced in the following.

ML and statistics are closely related fields in terms of methods, but distinct in their principal goal, since statistics draws population inferences from a sample, while ML finds generalizable predictive patterns. ML is nowadays largely perceived as one of the main innovative technologies, as much as computers have been in the 1980s and 1990s.

The term "machine learning" was popularized in 1959 by Arthur Samuel [323] and is seen as a part of the academic discipline of artificial intelligence (AI), of which it grew out when some researchers were interested in having machines learn from data. Therefore, in short, ML algorithms build a model based on sample data, known as "training data", in order to make predictions or decisions without being explicitly programmed to do so. Used for a vast array of data processing tasks, which has included most scientific disciplines in recent years, ML encompasses a broad range of algorithms and modelling tools.

6.5.1. Approaches and techniques in machine learning

Although historically often considered to be built up on the two broad main categories of *supervised learning* and *unsupervised learning*, the tendency nowadays drifts to divide ML in three or more main subfields with additionally *semi-supervised learning*, *reinforcement learning*, or other approaches like *topic*

modelling, *meta learning*, and *deep learning*. However, one way of a graphical illustration of the topology of ML (by no means complete or immaculate) is shown in figure 6.1.

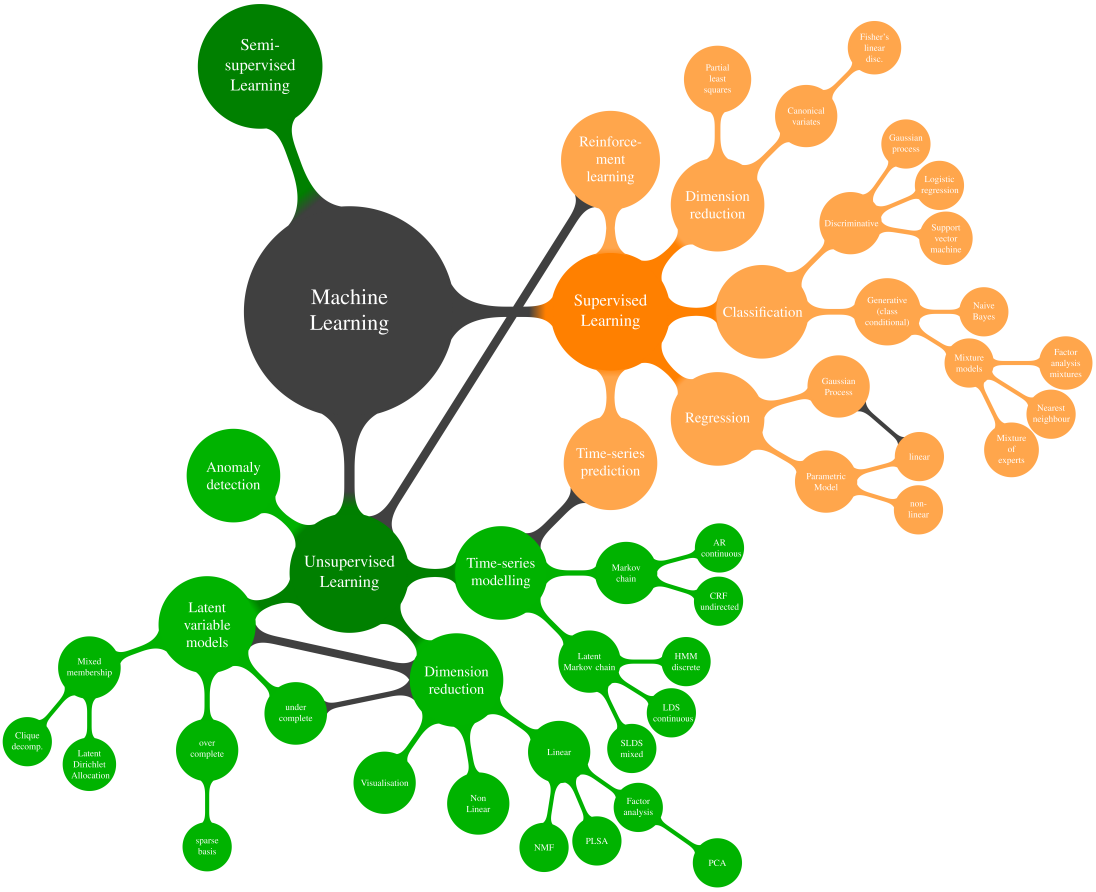


Figure 6.1.: An illustration of the topology in ML, denoting loose associations between subjects, taken from [324].

In unsupervised learning algorithms the aim is to find compact descriptions of the data with no labels given, leaving it on its own to find structure in its input, e.g. discovering hidden patterns in data, or feature learning. The goal in supervised learning can be seen as learning a general rule, which maps inputs to outputs, with the focus on accurate prediction. A commonality is the interest in methods that generalise well to previously unseen data, thereby distinguishing between data that is used to train a model and data that is used to test the performance of the trained model.

In this work the following methods were tested:

- Linear discriminant analysis (LDA): Fisher discriminants (Fisher), Fisher discriminant with Gauss-transformed input variables (FisherG), and boosted Fisher discriminants (BoostedFisher), which uses generalised MVA method boosting;

- neural networks (NNs): feedforward multilayer perceptron (MLP) Broyden-Fletcher-Goldfarb-Shanno (BFGS) = MLPBFGS, and NN with BFGS training method and Bayesian regulator (MLPBNN);
- decision trees (DTs): boosted decision trees (BDTs) using Gradient Boost (BDTG);
- Support Vector Machines (SVMs).

The basics of these methods will be described in the following section, in the appendix B.3, and comparisons of the applied methods are shown in figures 6.15, 6.21, 6.27 and 6.35), whereof the best results maximizing signal over the overall fluctuations were obtained using the BDTG method as shown in sec. 6.6. More details of these models utilized in this work can be found in [322].

6.5.1.1. Principal component analysis

In addition to the methods tested in this work, a method on which one of the earlier presented UHE photon content analysis results, figure 3.1, is based on, is the principal component analysis (PCA). This method is used in exploratory data analysis, common for dimensionality reduction of data while preserving as much of the data's variation as possible.

Projection methods such as PCA are in general aiming to find a mapping from the inputs in the original d-dimensional space to a new ($k < d$)-dimensional space, with minimum loss of information. In the sense of not using the output information, PCA can be classified an unsupervised method which maximizes the variance. The *principal components* are eigenvectors of the data's *covariance matrix* (with the largest eigenvalue), thus PCA is the simplest of the true eigenvector-based MVAs and is closely related to factor analysis. A PCA defines a new orthogonal coordinate system that optimally describes variance in a single dataset, and is sensitive to outliers, which would have a large effect on the variances and thus the eigenvectors [325]. In a simplified way a PCA can be understood as a rotation of the axis along the highest correlation between the variables, with the output being the projected value of an event on this principal component axis.

In [204], after applying several selection criteria on the datasets (in order to ensure the quality of the reconstruction), two observables, the Δ rise-time via benchmarks, and the modified photon LDF, L_{LDF} are transformed via a PCA, so that the mean is equal to zero and the standard deviation equal to one, resulting in gL_{LDF} and $g\Delta$. For the data the above described SD burn sample was used, and the scatter plot of the two observables is shown together with the new axis of the PCA in figure 6.2.

After the projection onto the new principal component axis, a cut, above which an event is considered as a photon candidate, is defined as the median of the MC photon events (on the new axis). From the real CR events 4 passed the photon cut above $E_\gamma > 10^{19}$ eV, 2 events above $2 \cdot 10^{19}$ eV and none above

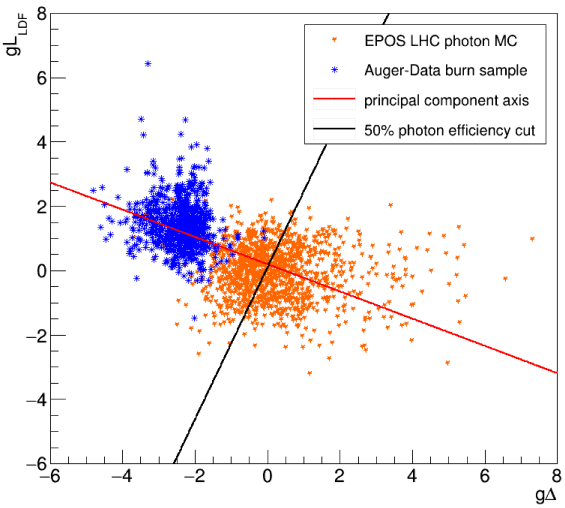


Figure 6.2.: The axis of the PCA in red is sown together with the data in blue and the MC prediction for photons in orange for the normalized observables L_{LDF} and $g\Delta$ [204].

$4 \cdot 10^{19}$ eV, which is consistent with the background expectations. Thereof the upper-limits on the photon flux were conservatively derived, for each range of energy $E_\gamma > E_0$ at a 95% CL, assuming an E^{-2} spectrum.

6.5.1.2. (Boosted) Decision Trees

Since in this work the method of BDTG was found to be the best performing and therefore chosen to be implemented for the planned data stream to AMON, it is described in more detail than the other methods.

As a supervised learning method, DTs is an efficient non-parametric method, which can be understood as a hierarchical data structure implementing the *divide-and-conquer* strategy and can be used for both classification and regression. In general the local region is identified in a sequence of recursive splits in a smaller number of steps, and a decision tree is composed of internal decision nodes and terminal leaves. The structure of the DT is not fixed a priori, but the tree grows as branches and leaves are added during learning, which depends on the complexity of the problem inherent in the data.

Simplified for explanation in the case of two input attributes (observables) x_1 and x_2, each event can be represented

$$\vec{x} = \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \tag{6.1}$$

with its label r denoting its type

$$r = \begin{cases} 1, & \text{or "yes", or signal ,} \\ 0, & \text{or "no", or background ,} \end{cases} \tag{6.2}$$

and defines a *two-class* problem. Each event is represented by such an ordered pair $(\mathbf{x}, r)$, with the training set containing N such examples

$$X = \{\mathbf{x}^t, r^t\}_{t=1}^N , \tag{6.3}$$

with t indexing different examples in the set (not representing time or any such order).

Since several measurements are made on each event an observation vector is generated, and the sample may be viewed as a data matrix

$$\mathbf{X} = \begin{pmatrix} X_{1,1} & X_{1,2} & \cdots & X_{1,d} \\ X_{2,1} & X_{2,2} & \cdots & X_{2,d} \\ \vdots & \vdots & \ddots & \vdots \\ X_{N,1} & X_{N,2} & \cdots & X_{N,d} \end{pmatrix} \tag{6.4}$$

where the d columns (also called inputs, features, or attributes) correspond to d variables denoting the result of measurements made on an event, and the N rows correspond to independent and identically distributed observations, examples, or instances on N events.

A test function $f_m(\mathbf{x})$ with discrete outcomes labelling the branches is implemented in each decision node m. Thereby at each node, a test is applied and one of the branches is taken depending on the outcome, starting at the root and repeated recursively until a leaf node is hit, at which point the value written in the leaf constitutes the output [325]. Each decision thereby divides the phase space into two distinct regions.

In the d-dimensional input space each node $f_m(\mathbf{x})$ defines a discriminant, which divides the input space into smaller regions that are further subdivided as a path is taken from the root down.

The method can be differentiated between *univariate* trees and *multivariate* trees, whereat in the case of a univariate tree, only one input dimension is used at a split, and in a multivariate tree, at a decision node, all input dimensions can be used – what makes it more general and the node becomes more flexible [325].

As an example, a binary *linear* multivariate node can be defined as

$$f_m(\mathbf{x}) \; : \; \mathbf{w}_m^T \mathbf{x} + w_{m0} > 0 , \tag{6.5}$$

defining a hyperplane with arbitrary orientation (where the w_i are the parameters to learn from data). A graphical illustration of the idea behind linear multivariate decision trees can be seen in figure 6.3.

Another and even more flexible node can be achieved by using a *nonlinear* multivariate node, as for example, with a quadratic

$$f_m(\mathbf{x}) \; : \; \mathbf{x}^T \mathbf{W}_m \mathbf{x} + \mathbf{w}_m^T \mathbf{x} + w_{m0} > 0 , \tag{6.6}$$

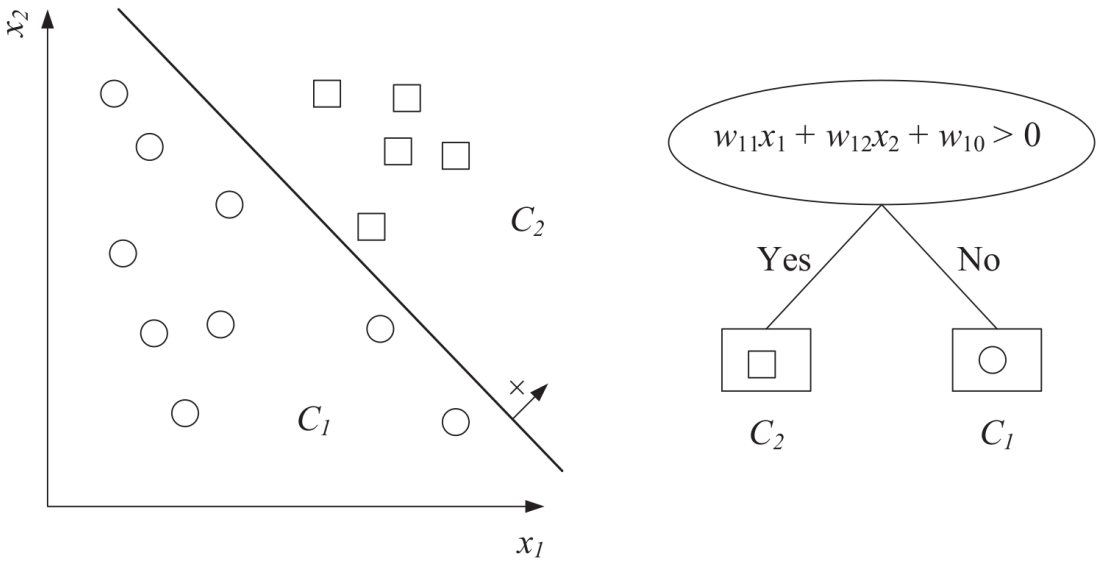


Figure 6.3.: An example of a linear multivariate decision tree, where the node can place an arbitrary hyperplane. Figure taken from [325].

Another way of having nonlinear decision nodes can be done via the use of a MLP, which is a linear sum of nonlinear basis functions (see sec. 6.5.1.5).

The successive nodes on a path from the root to a leaf further divide the trees, with leaf nodes defining polyhedra in the input space. The univariate numeric node defines a linear discriminant, with d possible orientations ($\mathbf{w}_m$) and $N_m - 1$ possible thresholds, enabling an exhaustive search. On the other hand, in a multivariate node there are $2^d \binom{N_m}{d}$ possible hyperplanes, what makes an exhaustive search less practical [325].

Many algorithms for learning multivariate DTs for classification have been proposed and are not explained here. It should just be noted, that the earliest was the multivariate version of the classification and regression tree (CART) algorithm [326].

DTs learn and respond quickly, and are accurate in many domains, and when written down as a set of IF-THEN rules, the tree can be understood and the rules can be validated by humans.

One can differentiate between hard and soft decision nodes. Compared to the above described hard decision nodes, in a soft decision tree all the branches are taken, and followed in parallel along all the paths and reaching all the leaves, but with different probabilities, where the output gets the weighted average of all the outputs in all the leaves with the weights corresponding to the probabilities accumulated over the paths.

Without caring about the distributions, a DT directly codes the discriminants which separate the class instances and therefore is *discriminant-based*, whereas the statistical methods are *likelihood-based*. An advantage of these discriminant-based methods is the direct estimation of the discriminants, by-

passing the estimation of class densities.

Since one single DT may be unstable in general, the concept is extended by *boosting* [327, 328] a DT to get several re-weighted trees which together form a *forest*, or training trees on a random subset leading to a random forest (RF). If small changes in the training set cause a large difference in the generated learner (= the learning algorithm has high variance), then a learning algorithm is an *unstable* algorithm.

A *weak learner* has an error probability less than 1/2, which makes it better than random guessing (on a two-class problem), while a *strong learner* has arbitrarily small error probability.

- **Boosting** is based on weak learners (high bias, low variance), which are in terms of DTs shallow trees, sometimes even as small as decision stumps (trees with two leaves). Boosting thereby reduces the error mainly by reducing bias (and also to some extent variance, by aggregating the output from many models). The concept of the combination of trees forming a forest is extended by boosting to get several re-weighted trees, resulting in the final response output called BDT. In boosting the next learner is trained on the mistakes of the previous learners, what generates complementary base-learners. It is especially susceptible to noise and outliers. Furthermore, a disadvantage of the original boosting method is that it requires a very large training sample.

- Another very popular model combining DTs to an ensemble (decision forest), where each DT is trained on a random subset of training set or a random subset of the input features, and their predictions are combined, the overall accuracy can be significantly increased and it is called the *RF*.

  In comparison to boosting the RF method uses fully grown decision trees (low bias, high variance), and tackles the error reduction task in the opposite way via reducing variance. The trees are made uncorrelated to maximize the decrease in variance, but the algorithm cannot reduce bias (which is slightly higher than the bias of an individual tree in the forest). Hence the need for large, unpruned trees, so that the bias is initially as low as possible.

It shall be noted that while boosting is sequential, a RF grows trees in parallel.

The shortcoming of DTs in their instability with respect to statistical fluctuations in the training sample can be overcome by the boosting technique. Thereby via re-weighting (boosting) versions of the training data and finally taking a weighted majority vote, the classification of typically weak multivariate methods can be enhanced with several methods of which three are the following.

- **Bagging** (short for *bootstrap aggregating*) is a voting method whereby base-learners are made different by training them over slightly different

training sets, and thereby a way of smearing over statistical representations of the training data. Generating complementary base-learners which is left to chance and to the unstability of the learning method is suited to stabilize the response of a classifier and often shows a performance increase. Recurrently the training sample is resampled in a way that the same event is allowed to be picked several times from the parent sample, thereby making the training sample a representation of the probability density distribution of the parent sample. In this way the classifiers are trained with different parent distributions and combined into a collection, which improves stability with respect to statistical fluctuations in the training sample.

- A statistical classification meta-algorithm called **AdaBoost** (short for Adaptive Boosting) was proposed in 1996 by Freund and Schapire [329] who therefore won the *Gödel Prize* in 2003.

  In AdaBoost the same training set is used over and over, thus it does not need to be large, but the classifiers should be simple so that they do not *overfit*. Many variants of AdaBoost have been proposed, where the main principle of the original idea is to modify the probabilities of drawing the instances as a function of the error. During the training misclassified events of a DT are given a higher event weight before training the following tree, which results in a better separation of misclassified events in the subsequent DT. This makes AdaBoost a voting method, once training is done. The success of AdaBoost can be explained to be due to its property of increasing the margin, which results in better separation of the training instances and an error is less likely. Thereby AdaBoost is similar to SVMs (see sec. 6.5.1.4).

  Based on an exponential loss function, this method has therefore the disadvantage that it lacks robustness in presence of outliers or mislabelled data points and consequently the performance degrades in noisy settings.

- **Gradient Boost**
  With the aim to improve the drawback of the AdaBoost method (originating from the exponential loss function), the gradient boosting algorithm [330, 331] uses a binomial log-likelihood loss function, which is a more robust loss function without giving up on the good out-of-the-box performance of AdaBoost. The idea of gradient boosting is to apply a steepest descent approach in an iterative procedure to minimize this loss function (functional gradient descent). The algorithm is thus called gradient boosted trees, and it usually outperforms random forest.

  Disadvantages arise since gradient boosting is in general less susceptible to over-training, and while it can increase the accuracy of a base learner, (e.g. DT or linear regression), it sacrifices intelligibility and in-

terpretability. Moreover the implementation may be more difficult due to the higher computational demand.

A further problem, mentioned already, is the BDT's tendency for *over-fitting*, which can be overcome by cross-validating with a separate data set in so called *pruning* methods.

6.5.1.3. Linear discriminant analysis (LDA)

Fisher discriminants [332] belong to the more general concept of LDA, a supervised method for dimensionality reduction for classification problems, which is used to find a linear combination of features that characterizes or separates two or more classes of objects or events. This is done on the basis of determining an axis in the hyperspace of input features and performing a projection of class outputs (signal and background) onto this axis. Beside being used as a linear classifier, the resulting combination is often used for dimensionality reduction prior to classification [325, 333].

These methods are closely related to PCA (and factor analysis), since both search for linear combinations of variables which best explain the data. What makes LDA a supervised technique in comparison to a PCA is that LDA chooses new axes such that the separability between classes is optimized, while PCA chooses new axes for dimensions preserving the variance (and hence the "shape") of the data.

Usually LDA is applied when groups are known a priori (unlike in cluster analysis). It creates one or more linear combinations of predictors, and a new latent variable for each function, which are called discriminant functions. The first of these created functions maximizes the differences between groups on that function, while in the following the second function maximizes differences on that function, and in addition must not be correlated with the previous function; the whole process is continued iteratively under these requirements. This maximizes the distance between the class distributions in a way that also the dispersion of each class is minimized, leading to a hyperspace axis, of which an illustration of two examples is shown in figure 6.4 – once for an unfavourable hyperplane selection, with overlap of classes and a large dispersion on the left, and on the right an optimal separation.

Best performance of LDA is achieved when the measurements made on independent variables for each observation result in continuous quantities, with advantages of simplicity of the classifier, good separation power, transparency, hard to over-train and fast training speed, but have worse performance for non-linear correlations.

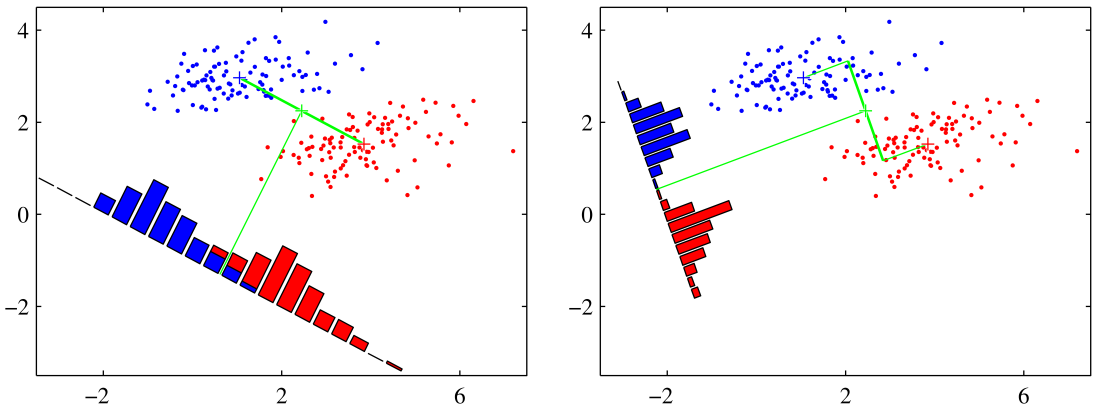


Figure 6.4.: Principle of selecting a hyperplane in two selection cases, where on the right the corresponding projection is based on the Fisher linear discriminant, resulting in a greatly improved class separation, more in the text. Figure taken from [333].

6.5.1.4. Support vector machines

The method of SVMs is based on potential functions, linear classifiers, and neighbour-based methods and the idea of generalizing linear models by mapping the data to a new space through non-linear basis functions, with the novelty of integrating this into a learning algorithm whose parameters are defined in terms of a subset of data instances (the so-called dual representation). Introduced by Vapnik et al. [334], SVMs, which can be generalized under the name kernel machine, are supervised learning models with associated learning algorithms. They are based on statistical learning frameworks (or Vapnik–Chervonenkis (VC) theory [335]), and one of the most robust prediction methods. This method maps training examples to points in space to maximize the width of the gap between the two categories (see fig. 6.5), where the SVM training algorithm builds a model that assigns new examples to one category or the other, thus SVMs as a binary linear classifier, are not fitting properly within a probabilistic framework. A binary classifier can be imagined as a function which decides whether or not an input (represented by a vector), belongs to some specific class.

Instead of using a projection like linear discrimination models, a maximum margin hyperplane (or decision boundary) is determined to maximize the distance between points of classes, minimizes simultaneously misclassification, and as an estimate for the expected error rate the number of support vectors can be used as an upper-bound (instead of the input dimensionality) [324].

Instances that are not "support vectors" carry no information, and therefore removing any subset of them would still get the same solution. This links SVMs to the (condensed) nearest neighbour algorithm (which stores only the instances neighbouring and hence constraining the class discriminant). This principle of considering only instances close to the boundary and discard-

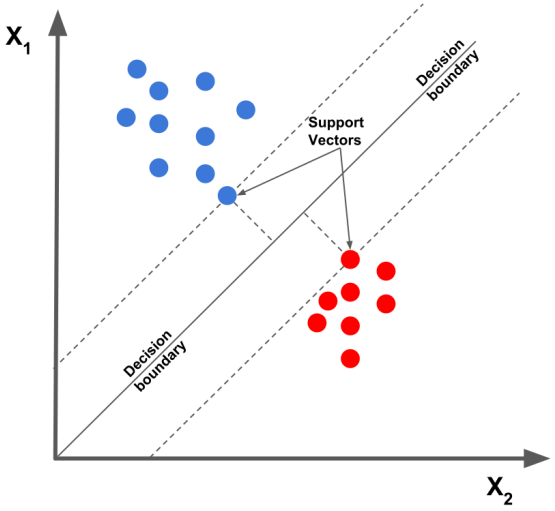


Figure 6.5.: Example of a SVM with the two class data sets separated by a hyperplane (decision boundary) maximizing the perpendicular distance to the closest point of both classes. Figure taken from [336].

ing those that lie in the interior makes SVMs a discriminant-based method, and enables to use a simpler classifier before, to filter out a large portion of such instances, which decreases the complexity of the optimization step of the SVM.

A further reason for the popularity of SVMs arises from its ability to efficiently perform a non-linear classification (via the "kernel trick"), implicitly mapping their inputs into high-dimensional feature spaces.

SVMs are potentially misled by outliers, since mislabelled points (or outliers) have a significant impact on the location of the decision boundary, but can be dealt with as a robust technique using the zero-one loss in which a mislabelled point contributes to only a relatively small loss [324].

The SVM's trade-off between margin maximization and error minimization can be understood in a way, that a too large data set results in a high penalty for non-separable points, subsequently storing too many support vectors and over-fitting, while a too small data set results in too simple solutions that under-fit.

6.5.1.5. Multilayer perceptrons / Neural Networks

The perceptron, introduced by Rosenblatt [337], can be compared in the context of NNs as an artificial neuron, and is categorized as supervised ML algorithm of binary classifiers. Since it is based on a linear function combining a set of weights with the feature vector, it is another example of a linear discriminant model. With one layer (thus also called a *single-layer perceptron*) between the input and output layer (as shown in figure 6.6) it is thereby the simplest *feedforward* NN [324, 325, 333].

The perceptron's successor, the feedforward NN with two or more hidden layers between the input and output layers (therefore called a MLP) is schematically shown in figure 6.7 and has (much) greater processing power than simple perceptrons, where already only a second layer of perceptrons is sufficient to solve a lot of otherwise non-separable problems.

Feedforward describes an architecture without closed directed cycles, to ensure that the outputs are deterministic functions of the inputs, and thereby differs from its descendant, the *recurrent* NNs.

Often called a misnomer in literature (arguing that the model comprises multiple layers of logistic regression models with continuous non-linearities rather than multiple perceptrons with discontinuous non-linearities), MLP is a class of feedforward artificial neural network (ANN) [324, 325, 333].

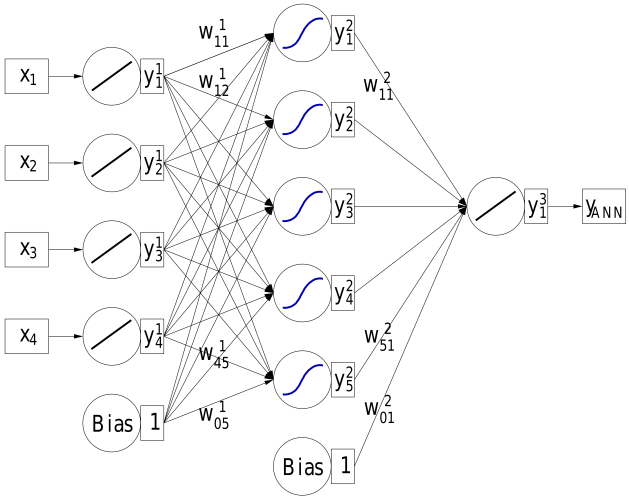


Figure 6.6.: Schematic illustration of a simple perceptron with one (hidden) layer between the input and output layer. Figure taken from [322].

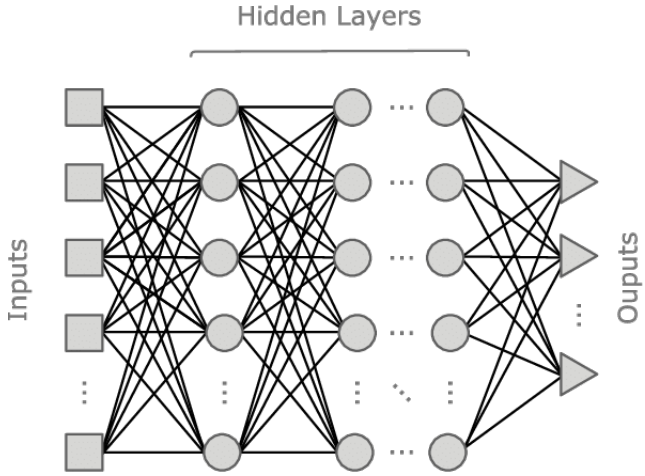


Figure 6.7.: Schematic illustration of a feedforward ANN (or MLP) with more hidden layers between the input and output layer.

In addition to the multiple layers to distinguish MLPs from the simple linear perceptron, each node (except for the input nodes) represents a neuron that uses a nonlinear activation function (AF), what enables it to classify data that is not linearly separable [324, 325, 333].

The AFs are the non linear transformation between the layers, and can be understood as decision of whether the neuron would fire or not. Each basis function itself is a non-linear function of a linear combination of the inputs, with the coefficients in the linear combination being adaptive parameters describing a NN as a series of functional transformations. Some examples of AFs are shown in figure 6.8, with the nowadays in deep NN very popular Rectified Linear Unit (ReLU) AF. Of course many other ways to construct parametric non-linear basis functions exist.

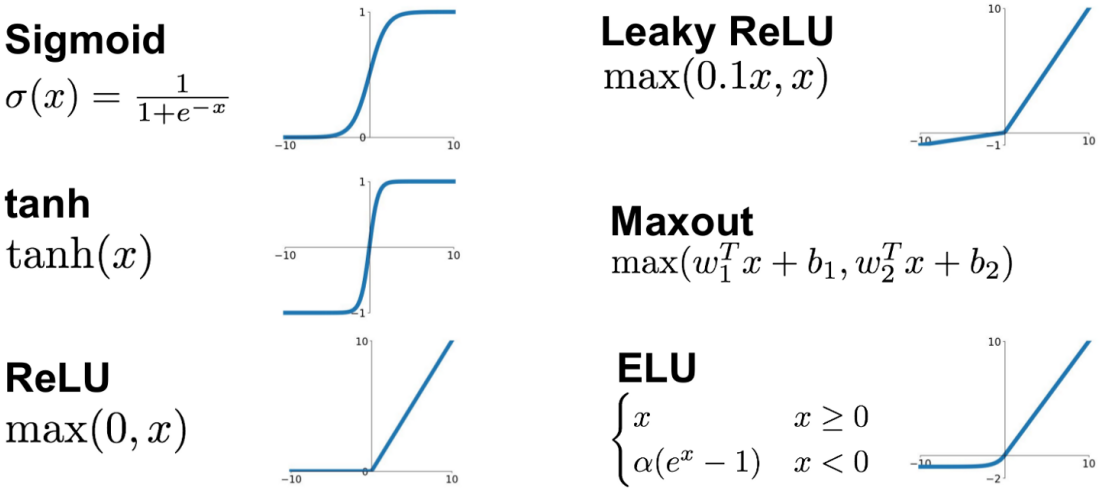


Figure 6.8.: Examples of some prominent AFs [338].

In the perceptron model of Rosenblatt the input vector $\mathbf{x}$ is first transformed using a fixed non-linear transformation, giving a feature vector $\vec{\phi}(\mathbf{x})$, which is then used to construct a generalized linear model of the form

$$y(\mathbf{x}) = f\left(\mathbf{w}^T \vec{\phi}(\mathbf{x})\right) , \tag{6.7}$$

where the non-linear AF $f(\cdot)$ is a step function.

The basis of feedforward NNs are linear combinations of fixed non-linear basis functions $\phi_j(\mathbf{x})$ in the form

$$y(\mathbf{x}, \mathbf{w}) = f\left(\sum_{j=1}^{M} w_j \phi_j(\mathbf{x})\right) , \tag{6.8}$$

where $f(\cdot)$ in the case of classification is a non-linear AF [333].

At the end of the feedforward propagation, when weights define the mapping of input features to output variables, a supervised learning technique

called *backpropagation* for training is utilized, which minimizes the error function (by working backwards) via recursively adjusting the weights in the network.

Another numerical optimization algorithm is the BFGS [339–342] algorithm, which is an iterative method for solving unconstrained non-linear optimization problems and usually performs a smaller number of iterations during training and is thus faster.

In a Bayesian view of learning, when training NNs, the parameters (connection weights w_i) are considered as random variables drawn from a prior distribution $p(w_i)$, and the posterior probability is computed given the data [324].

NNs are receptive for over-training and rather poorly transparent.

6.6. Analysis results

6.6.1. TMVA workflow

The ROOT TMVA utilized in this work provides a full workflow for MVA and is operated in two phases, namely the *training* and the *application* (or *classification*) phase [322]. Before data (simulations and/or real data) classification the training of the MVA needs to be performed with a training sample, which usually is a sample of known composition, having all necessary information for discrimination, the chosen observables "booked" for training. Prior to the data treatment, basic selection cuts are applied on events as e.g. a requirement of a successful reconstruction. A big advantage of the performed ML-based analysis, namely for the search for sub-threshold events, is that rather loose cuts can be applied, while stringent quality cuts would have cut down the available statistics significantly. The complete training sample, composed of all simulations of photon, proton and iron events, has been split into several sub-samples for learning (or training) and testing, so that no event is used twice. The various ML methods in the MVA have been applied to the training sample to "learn" how to discriminate between signal (photon) and background (proton) events. After the learning phase the MVA gets applied on an independent test sample to test the performance and possible effects of over-training.

To avoid a bias introduced when repeating the learning and testing phase, a common practice used in data analysis was applied, leaving the final sample to be classified, the SD burn data set, untouched till the optimal configuration of the MVA was found. This set of real data, of which the composition is unknown, was then classified in the final phase, the application phase, where it is up to user to decide where to place the cut on the response variable for signal selection.

6.6.2. Application of ML techniques for photon candidate event selection

As an overview of what is presented in the following sections, the main points of the work of application of ML techniques for selection of photon candidates among the Auger SD events can be broken down to:

- The increase of variables of the MVA (from 2 to 3 to 4) improved the performance, which was qualitatively determined with an increase of the integrated area under the receiver operating characteristic (ROC) curve, see sections 6.6.3 → 6.6.4 → 6.6.5, 6.6.6 and 6.6.6.1, and appendix A.5.

- Various MVA methods were tested and compared to each other, as described in section 6.5.

- The ML techniques were trained on simulations and after first application to classify pseudo-data sets (which are data sets composed of a combination of simulated SD events), see sections 6.6.3, 6.6.4, 6.6.5, real data – the SD burn data set (sec. 6.6.6 and 6.6.6.1) – were classified to test functionality.

- A systematic cross check was performed, in which the background was composed of 50% proton and 50% iron, as described in 6.6.6.1.

For the first implementation of the MVA with simulated events, the $\overline{\text{Off}}\underline{\text{line}}$ analysis framework version v3r3p4, with the utilized modules given in appendix B.1.1, and the following discrimination variables were used:

1. Rise-time ($t_{1/2}(1000)$, eq. (5.4), standard $\overline{\text{Off}}\underline{\text{line}}$ simulation output)
   a) "Recalculated" rise-time (see fig. 6.9 a))
   b) Δ_R-method (see fig. 6.12)
2. Curvature of the shower front (see fig. 6.9 b))
3. Number of Stations (NSt) (see fig. 6.16)

The observable termed "recalculated" rise-time was tested in the beginning, since the lower limit of the total SD station signal is set in $\overline{\text{Off}}\underline{\text{line}}$ to be 10 VEM, together with the requirement of at least three triggered stations, which removes many low energy events (with low signal-to-noise ratio (S-N-R)) during the calculation of $t_{1/2}(1000)$. As in [306], for the recalculated rise-time the limit was reduced to 5 VEM and SD station rise-times $t_{1/2}$ were recalculated via the approach from the $\overline{\text{Off}}\underline{\text{line}}$. It was probed if the analysis can benefit from using either the recalculated or the Δ_R rise-time (described in more detail in sec. 5.2.4.1), instead of the standard rise-time at 1000 m from the shower core $t_{1/2}(1000)$. It was found that the differences in the results of the MVA between these two methods are minimal and the improvement of using them instead of the $t_{1/2}(1000)$ is negligible, and that the real improvement comes with the combination of more variables (with separating power), what is presented in later sections.

The variables were used in the following four combinations:

A) 1.a) & 2.: Recalculated rise-time & Curvature, results in fig. 6.10 a) and b) and 6.11;

B) 1.b) & 2.: Δ_R rise-time & Curvature, results in fig. 6.13 a) and b) and 6.14;

C) 1.a) & 2. & 3.: Recalculated rise-time & Curvature & NSt, results in fig. 6.17 a) and b) and 6.18;

D) 1.b) & 2. & 3.: Δ_R rise-time & Curvature & NSt, results in fig. 6.19 a) and b) and 6.20;

For all four combinations the MVA distribution, the background discrimination vs. signal efficiency (ROC) curves, and the correlation matrices are given in the following sections. The correlation matrices show the linear correlation coefficients in % of a linear relationship between two variables, what is a measure of the strength and direction. Thereby the correlation coefficients between two variables x and y, r_{xy} are given by

$$r_{xy} = \frac{\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})}{(N-1)\sigma_x \sigma_y}, \tag{6.9}$$

with σ_x and σ_y the standard deviations of the samples and $\bar{x}$ and $\bar{y}$ the sample's mean values. Figures 6.11, 6.14, 6.18, 6.20, 6.24, 6.26, 6.36 show these correlations. Although only the linear correlations are presented in these mentioned figures (and can be described via eq. (6.9)), some ML techniques like BDTs or ANNs can also handle higher dimensional correlations. Differences in the correlation matrices of signal and background may already hint towards beneficial features for the MVA.

Further and more detailed checks of the classifiers like over-training, MVA response functions, etc., are presented for the main part of the analysis in sec. 6.6.6.

The test and comparison of several machine learning techniques against each other can be best seen (on a qualitative basis, but sufficient for the purpose of this work) in ROC curves, whereof the best performing (meaning largest area under the curve) method will be chosen for the final analysis of course. The signal (background) efficiency [343–345] $\epsilon_{S(B)}$[2] of a ML method is defined for a certain cut i via:

$$\epsilon_{S(B)}(i) = \frac{N_{S(B)}(i)}{N^{\text{tot}}_{S(B)}}, \tag{6.10}$$

with $N_{S(B)}(i)$ the number of events above the cut i, and $N^{\text{tot}}_{S(B)}$ the total number of signal (background) events[3]. Thereof the background (signal) rejection $r_{B(S)}$ is gained via $1 - \epsilon_{B(S)}$[4].

An example of the parameter combinations of some tested ML methods is given in the appendix B.3 in table B.1.

In the plots that show the cut efficiencies (figures 6.10 b), 6.13 b), 6.17 b), 6.19 b), 6.23 b), 6.25 b), 6.34 b) 6.38 b)), different cut options offered from the TMVA are given as well as the surviving number of data events and the

[2] In other scientific domains ϵ_S is often called "recall", "sensitivity" or in terms of the confusion matrix true positive rate (TPR), while ϵ_B "fall-out", type I error, or false positive rate (FPR).

[3] Counts are denoted with upper-case letters and rates are denoted with lower-case.

[4] r_B is often called "specificity", or true negative rate (TNR), and r_S "miss rate", type II error, or false negative rate (FNR).

resulting signal *purity* (in other scientific domains often called "precision") defined as

$$\mathcal{P} = \frac{S}{S+B}\,, \tag{6.11}$$

with S and B the number of signal and background events in the data sample, respectively. More details can be found in [322]. Together with the purity in these plots, superimposed are the signal efficiency ϵ_S, background efficiency ϵ_B, the signal efficiency times the purity $\epsilon_S \cdot \mathcal{P}$ and the significance of a cut at a certain response value. Via maximizing the significance often the best working point is achieved, with the definition of significance varying according to the actual problem. Most relevant for this work, and in general where $S \ll B$, choosing a significance is $S/\sqrt{B}$ may be best. Resulting values of events after cuts are given for the two most relevant cases discussed below in sec. 6.6.6 and 6.6.6.1 and a comparison of those values is presented in the appendix in table A.1.

Through all the plots in this section photons were used as signal in the MVA and are presented in blue colour, background in red and **data** and accordingly **pseudo-data** (or whatever was used to be classified) in **black**.

6.6.3. MVA with two observables

6.6.3.1. A: Recalculated rise-time & Curvature

The input variables for the MVA are shown in figures 6.9 a) and b), with their linear correlations shown in fig. 6.11 and the resulting MVA variable distributions in fig. 6.10 a) and cut efficiencies in b) for the method of BDTG, and further tested methods can be seen in the ROC curves hereof in fig. 6.15 a).

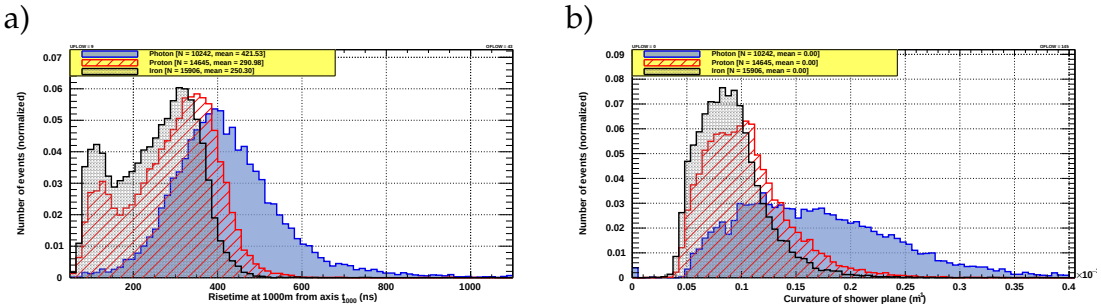


Figure 6.9.: Distributions of the **a)**: (recalculated) rise-time; **b)** curvature of the shower front [km^{-1}]

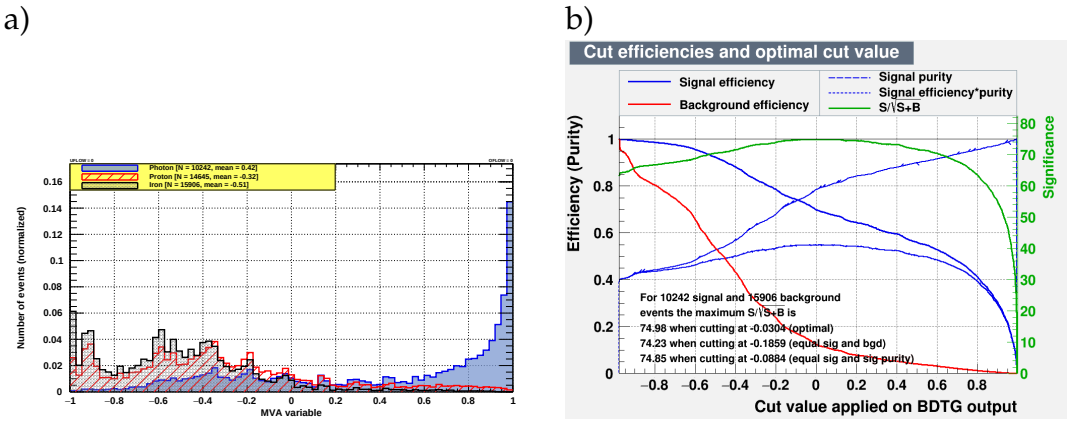


Figure 6.10.: Case A: rise-time recalculated and curvature with BDTG **a)** MVA distribution; **b)** cut efficiencies;

6.6.3.2. B: Δ_R rise-time & Curvature

In case B the Δ_R rise-time, shown in fig. 6.12 was combined as input variable with the curvature of fig. 6.9 b) in a MVA. Their linear correlations are shown in figures 6.14 and the resulting MVA variable distributions in fig. 6.13 a) and cut efficiencies in b) for the method of BDTG, and further tested methods can be seen in the ROC curves hereof in fig. 6.15 b).

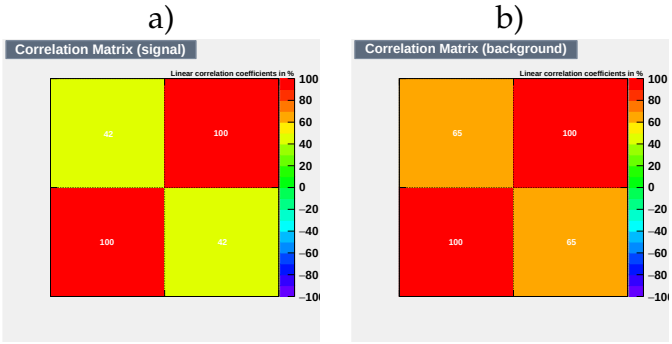


Figure 6.11.: Linear correlation matrices for case A: rise-time recalculated and curvature with BDTG **a)** signal, **b)** background, where a positive value indicates a positive correlation, with 100% indicating a full positive correlation, zero no correlation and -100% a full anti-correlation.

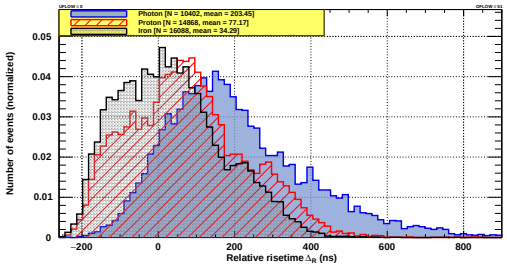


Figure 6.12.: Distributions of the rise-time Δ_R.

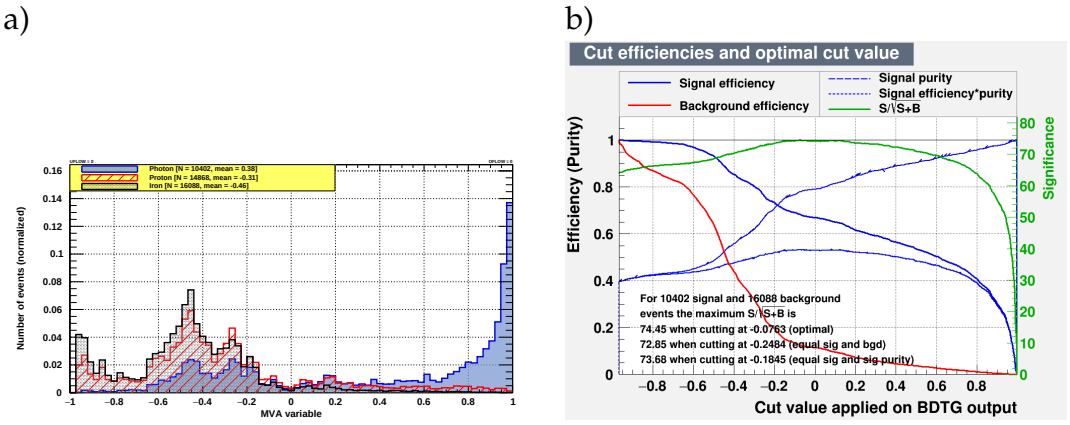


Figure 6.13.: Case B: rise-time Δ_R and curvature with BDTG **a)** MVA distribution; **b)** cut efficiencies;

6.6.3.3. ROC curves: case A vs. B

In figure 6.15 the ROC curves of the two cases A: Recalculated rise-time & Curvature and B: Δ rise-time & Curvature are given next to each other to be compared.

In comparing the ROC curves in fig. 6.15, case A seems slightly better than

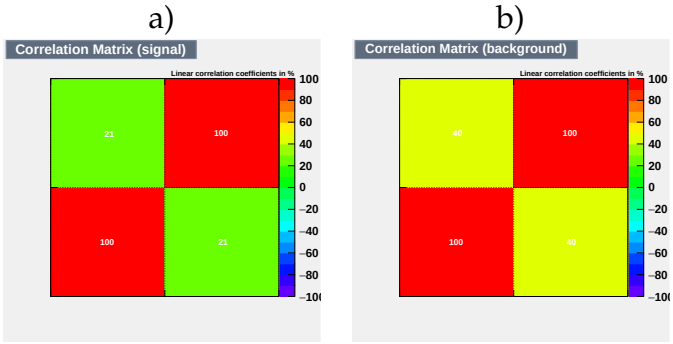


Figure 6.14.: Linear correlation matrices for case B: rise-time Δ_R and curvature with BDTG **a)** signal, **b)** background, where a positive value indicates a positive correlation, with 100% indicating a full positive correlation, zero no correlation and -100% a full anti-correlation.

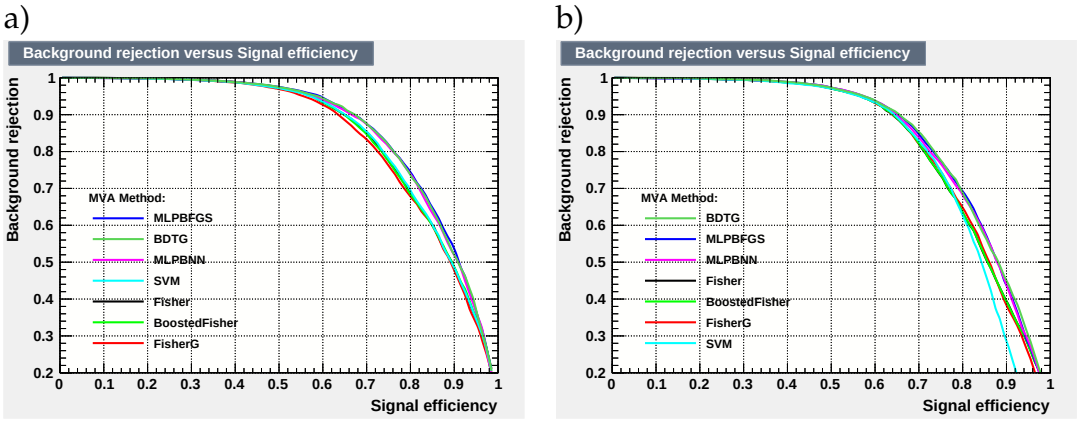


Figure 6.15.: ROC curves: **a)** case A: rise-time recalculated and curvature vs. **b)** case B: Δ_R and curvature; The various ML methods are ranked by the integral of the ROC diagram with the best method on top.

B, what originates from the two peaks left of the main peak of the rise-time in case A (and therefore better separation when comparing figures 6.10 a) with 6.13 a)).

A further comparison of the results of these two methods against each other and the other MVA results is given in the appendix in table A.1.

The MVA distributions and cut efficiencies given are for the method of BDTG, as this one delivered the best results. The MVA distributions and cut efficiencies for the other tested methods are not presented for the sake of the overview, only for the last section 6.6.6 they can be found in the appendix A.2.

6.6.4. MVA with three observables

6.6.4.1. C: recalculated rise-time & Curvature & NSt

In comparison to case A, in case C a third observable, the NSt (fig. 6.16), was added to the MVA. The linear correlations of the input variables are shown in fig. 6.18 and the resulting MVA variable distributions in fig. 6.17 a) and cut efficiencies in b) for the method of BDTG, and further tested methods can be seen in the ROC curves hereof in fig. 6.21 a).

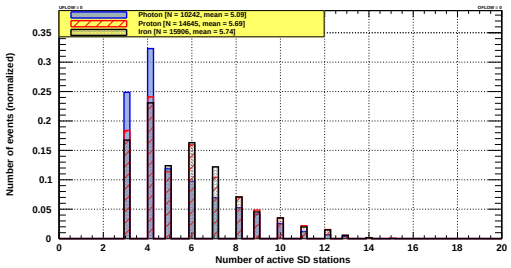
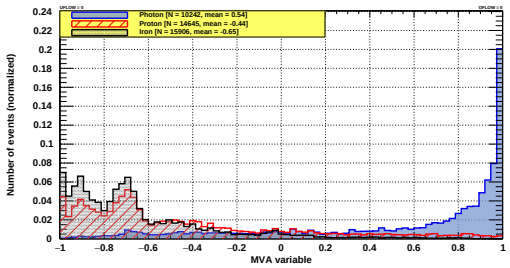


Figure 6.16.: Distribution of the NSt.

a)



b)

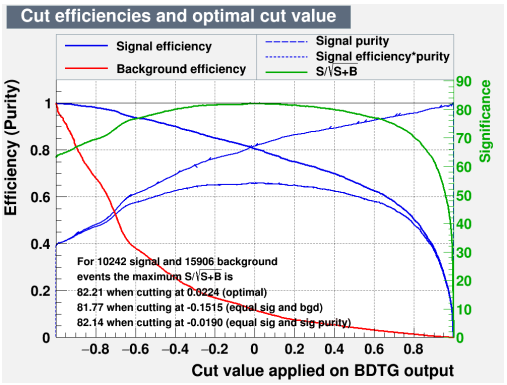


Figure 6.17.: Case C: rise-time, curvature and NSt with BDTG **a)** MVA distribution; **b)** cut efficiencies;

6.6.4.2. D: Δ_R rise-time & Curvature & NSt

In comparison to case B, in case D a third observable, the NSt (fig. 6.16), was added to the MVA. The linear correlations of the input variables are shown in fig. 6.20 and the resulting MVA variable distributions in fig. 6.19 a) and cut efficiencies in b) for the method of BDTG, and further tested methods can be seen in the ROC curves hereof in fig. 6.21 b).

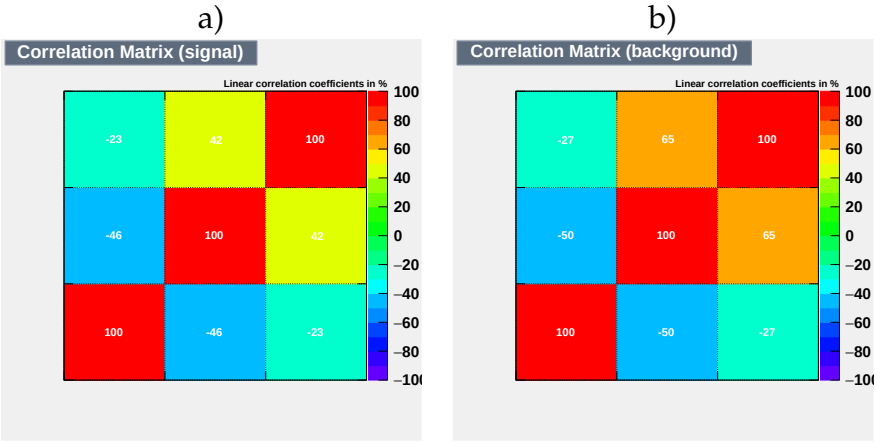


Figure 6.18.: Linear correlation matrices for case C: rise-time, curvature and NSt with BDTG **a)** signal, **b)** background, where a positive value indicates a positive correlation, with 100% indicating a full positive correlation, zero no correlation and -100% a full anti-correlation.

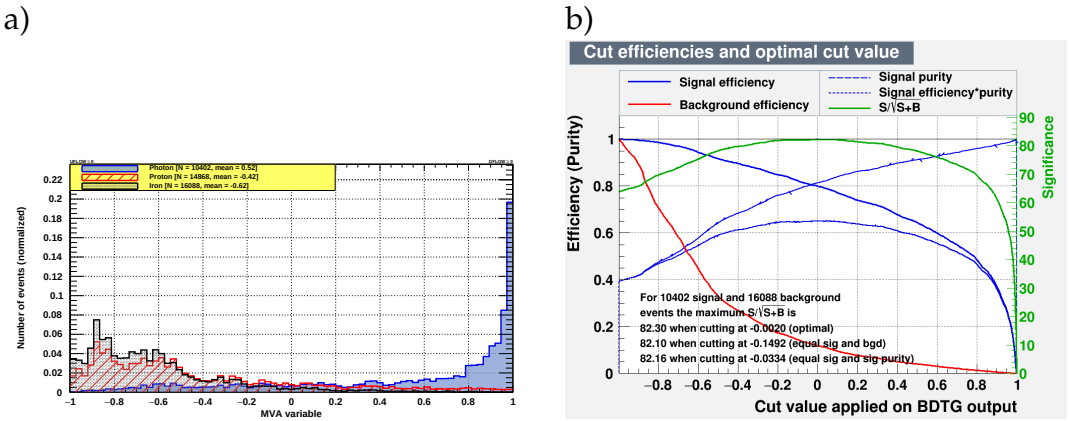


Figure 6.19.: Case D: Δ_R, curvature and NSt with BDTG **a)** MVA distribution; **b)** cut efficiencies;

6.6.4.3. ROC curves: case C vs. D

Here in figure 6.21 the ROC curves of the two cases C: Recalculated rise-time & Curvature & NSt and D: Δ_R rise-time & Curvature & NSt are given next to each other for comparison.

This combination of three observables (one additional variable) leads to better separation, same for case C and D, as compared to results presented above for the two observables case in sec. 6.6.3.1 and 6.6.3.2, and by comparing the ROC curves in figures 6.15 with 6.21.

A further comparison of the results of these two methods against each other, showing quantitatively the improvement resulting from the extension from two to three variables, is given in the appendix in table A.1.

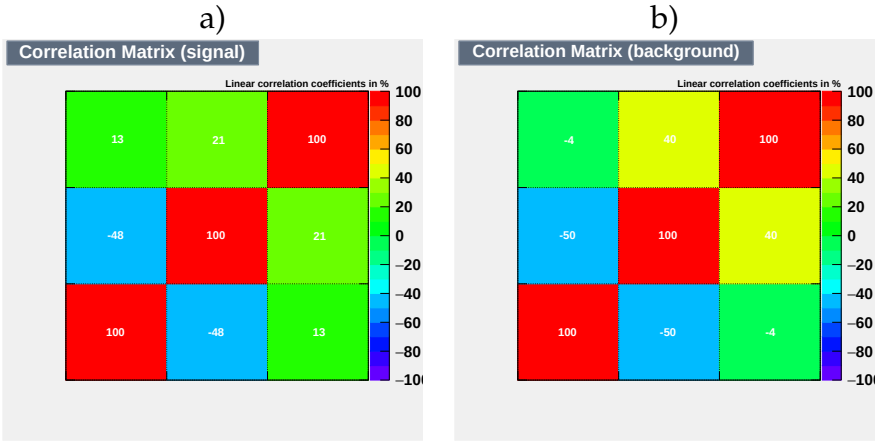


Figure 6.20.: Linear correlation matrices for case D: Δ_R, curvature and NSt with BDTG **a)** signal, **b)** background, where a positive value indicates a positive correlation, with 100% indicating a full positive correlation, zero no correlation and -100% a full anti-correlation.

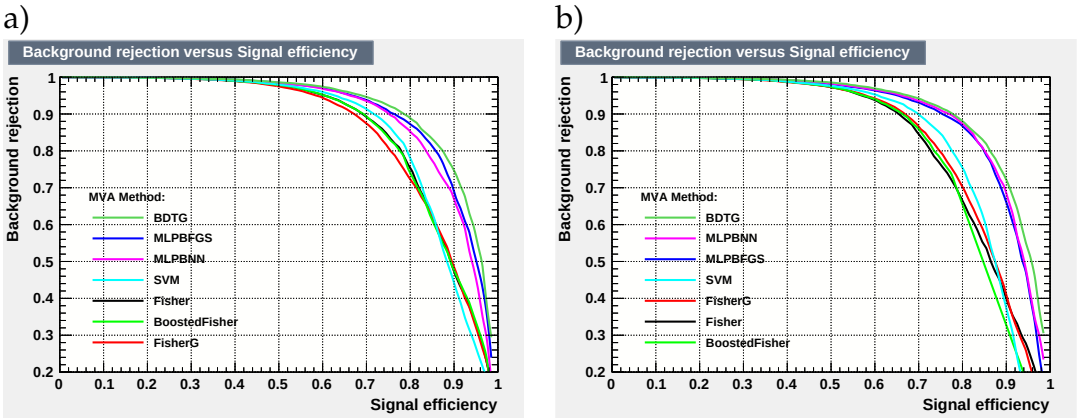


Figure 6.21.: ROC curves: **a)** case C: rise-time recalculated, curvature and NSt vs. **b)** case D: Δ_R, curvature and NSt; The various ML methods are ranked by the integral of the ROC diagram with the best method on top.

The MVA distributions and cut efficiencies given are for the method of BDTG, which delivered the best results. The MVA distributions and cut efficiencies for the other tested methods are not presented for the sake of the overview, only for the last section 6.6.6 they can be found in the appendix A.2.

6.6.5. MVA with four observables

Due to improvements of the Auger Offline analysis framework algorithms dealing with photomultiplier tube signals, a better single station treatment in simulations, etc., it was decided to run simulations with an updated version of the modules. The utilized modules in the Offline trunk version 32493 are given in appendix B.1.2.

The results and plots of this section present in comparison to the previous sections a much more developed analysis, containing beside a updated module sequence for simulations with Offline additional variables and an increased statistics.

Observables in this case were:

1. Rise-time ($t_{1/2}(1000)$, shown for comparison in fig. 6.22 a))
   a) Recalculated (see figure 6.22 c))
   b) Δ_R (see figure 6.22 d))
2. Curvature of shower front (see figure 6.22 e))
3. NSt (see figure 6.22 f))
4. Zenith angle (see figure 6.22 b))

with the tested combinations of variables:

A) 1.a) & 2. & 3 & 4

B) 1.b) & 2. & 3 & 4

The bumps in figure 6.22 a) for the two background components are present already in the rise-time coming from the Offline output. The inclined events ($\theta > 60°$) are reconstructed differently [302, 306], the calculation of rise-time might need some correction at large zenith angles. Low values of rise-time arise from a high muonic content, resulting in sharp peaks in PMT traces, and appear for events with large zenith angles (c.f. sec. 5.2.4.1). Indeed, it was found that the reason of those peaks was a mixture of events from a simulated data set containing events with large zenith angles (> 60°).

In the recalculated rise-time (fig. 6.22 c)) the mentioned bumps are still present, but of less amplitude, as well as in Δ_R (fig. 6.22 d)).

The break/dip for the NSt (figure 6.22 f)) occurs because there is a different distance cut as the energy increases [302, 306] (c.f. sec. 5.2.4.1). For events with $E < 10^{19.6}$ eV, the limits to the SD station distance from the shower axis are [300 m, 1400 m]. Above $10^{19.6}$ eV the limits are increased to [300 m, 2000 m]. This means that more candidate stations are considered as active (contributing to the event), if the energy is higher. If one produces an MVA analysis by limiting energies to below $10^{19.6}$ eV, this discontinuity vanishes. The reason for the distance cuts is the following: the lower range of distance,

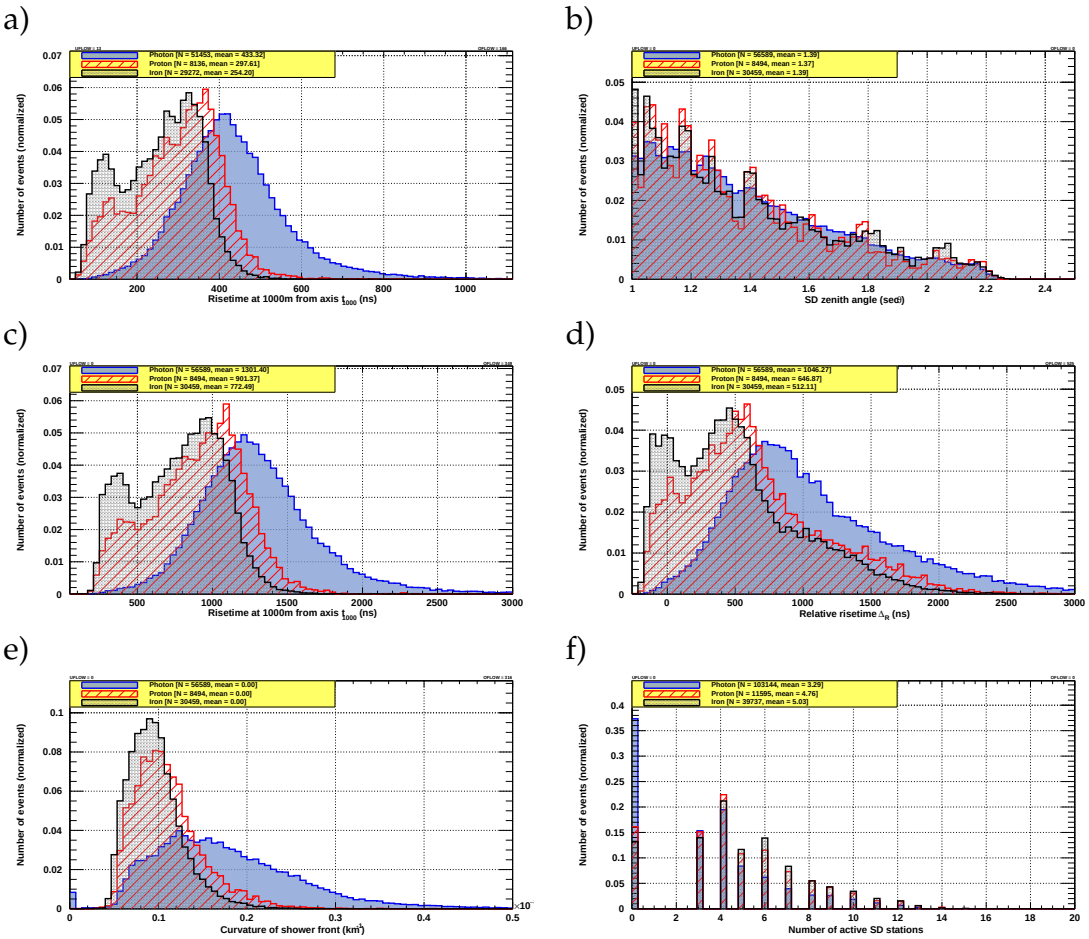


Figure 6.22.: Distributions for the observables **a)** $t_{1/2}(1000)$ **b)** Zenith angle, expressed as $\sec(\theta)$ **c)** recalculated rise-time **d)** Δ_R **e)** curvature of the shower front [1/km] **f)** NSt.

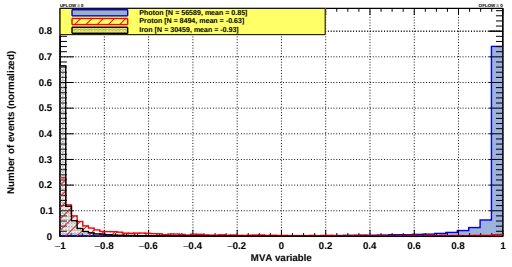
300 m, is selected to avoid the problems set by the inability of the data acquisition system to record fast pulses, while the upper range 1400 m is chosen to span what is consistent with unbiased selection. For the highest energies, this has been extended to 2000 m as the signal sizes in such events are sufficiently large to give accurate measurements. In the recalculation of the rise-time the reduction of the threshold to 5 VEM will most likely not cause a dip in the distribution, since it should increase the number of active stations for all energies.

It should be noted that for the goal of this work these features are of minor importance.

6.6.5.1. Case A: rise-time recalculated + curvature + NSt + θ

The linear correlations of the input variables are shown in fig. 6.24 and the resulting MVA variable distributions in fig. 6.23 a) and cut efficiencies in b) for the method of BDTG, and further tested methods can be seen in the ROC curves hereof in fig. 6.27 a).

a)



b)

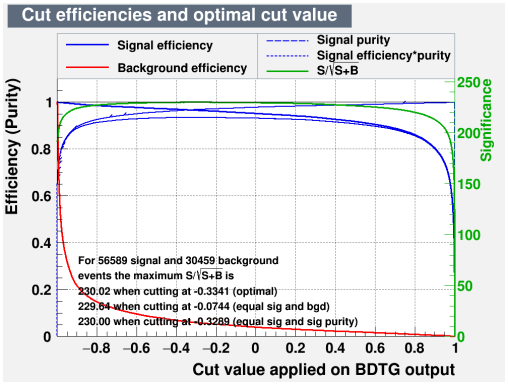


Figure 6.23.: Case A: recalculated rise-time, curvature, NSt and θ with BDTG **a)** MVA distribution; **b)** cut efficiencies;

6.6.5.2. Case B: Δ_R + curvature + NSt + θ

The linear correlations of the input variables are shown in fig. 6.26 and the resulting MVA variable distributions in fig. 6.25 a) and cut efficiencies in b) for the method of BDTG, and further tested methods can be seen in the ROC curves hereof in fig. 6.27 b).

6.6.5.3. ROC curves: case A vs. B

Results in case A are better than in case B, because the variable rise-time in A (fig. 6.22 c)) has a stronger pronounced signal/photon peaks below the

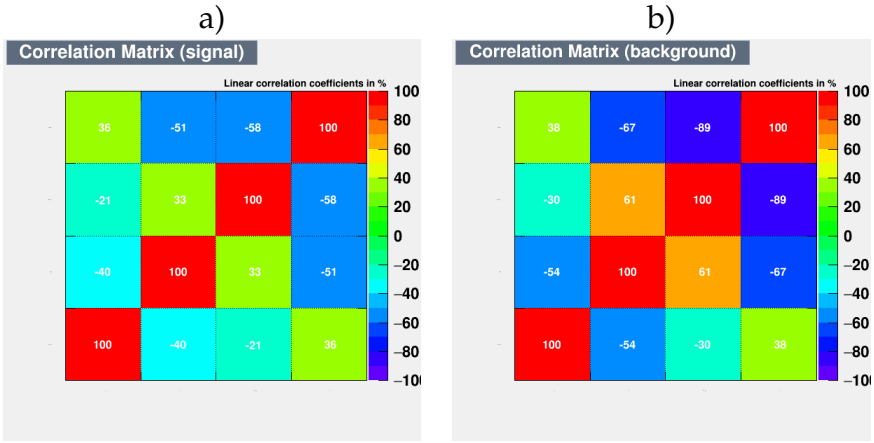


Figure 6.24.: Linear correlation matrices for case A: recalculated rise-time, curvature, NSt and θ with BDTG **a)** signal, **b)** background, where a positive value indicates a positive correlation, with 100% indicating a full positive correlation, zero no correlation and -100% a full anti-correlation.

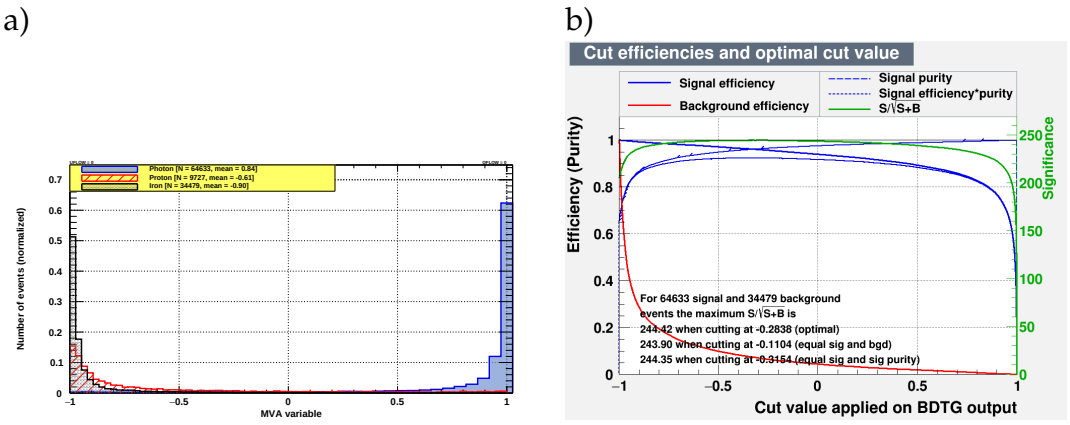


Figure 6.25.: Case B: Δ_R, curvature, NSt and θ with BDTG **a)** MVA distribution; **b)** cut efficiencies;

main peak in comparison to the background component and therefore better separation. Adding the zenith angle as observable to the MVA considerably improves the discrimination performances of the method (for both cases A & B), as compared to results presented above in sec. 6.6.3.1, 6.6.3.2, 6.6.4.1 and 6.6.4.2, when comparing the ROC curves of figures 6.27 with the ones from two variables in fig. 6.15 and three variables in fig. 6.21. The resulting improvement is further quantified in the appendix in table A.1.

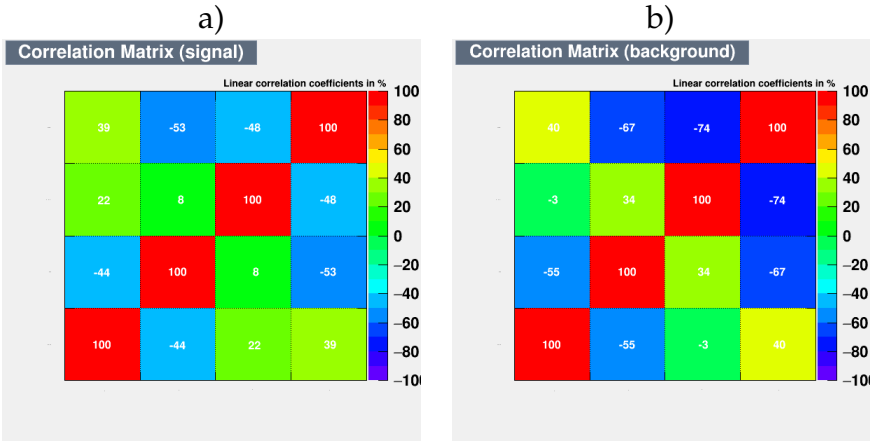


Figure 6.26.: Linear correlation matrices for case B: Δ_R, curvature, NSt and θ with BDTG **a)** signal, **b)** background, where a positive value indicates a positive correlation, with 100% indicating a full positive correlation, zero no correlation and -100% a full anti-correlation.

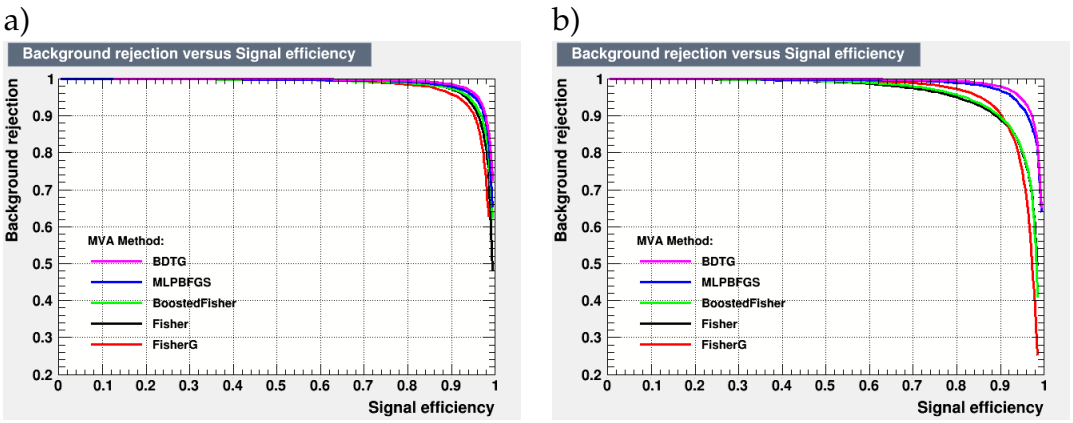


Figure 6.27.: ROC curves: **a)** case A: vs. **b)** case B: Δ_R, curvature and NSt; The various ML methods are ranked by the integral of the ROC diagram with the best method on top.

6.6.6. Application to the SD burn data

The **data** are shown in the following histograms in **black** colour.

As shown previously, the quality of the discrimination/identification, or "classification" of photon events increased with the use of more variables within the MVA, whereas also no significant time loss occurred with this inclusion.

Since protons have the smallest mass of all nuclei and are thus the most similar to photons in terms of air shower development in comparison to heavier nuclei, a pure proton background in the MVA represents the "worst case scenario" and thereby can be understood as a lower bound/limit for the sep-

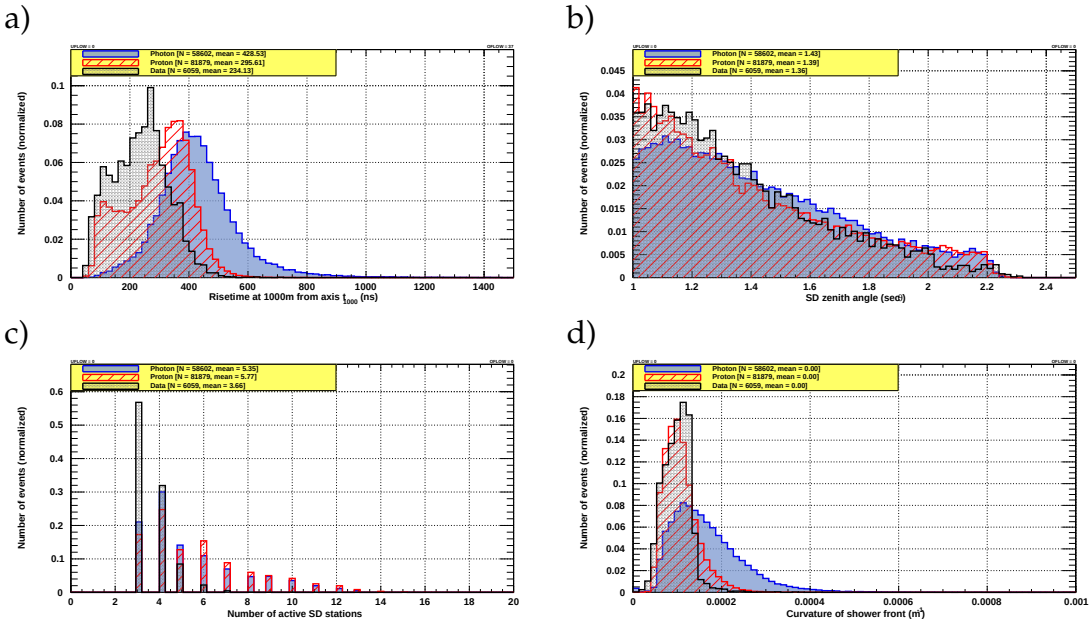


Figure 6.28.: Distributions for the observables **a)** $t_{1/2}(1000)$ **b)** Zenith angle, expressed as $\sec(\theta)$ **c)** NSt **d)** curvature of the shower front [1/km].

aration power of the MVA.

In this section the results and plots from simulations with the updated module sequence (see section 6.6.5) are shown (produced with Offline trunk version 32493). For this analysis I used the following variables in my MVA:

1. Rise-time $t_{1/2}(1000)$ (see figure 6.28 a))
2. Zenith angle of an incoming shower (see figure 6.28 b))
3. Number of stations retained in an event (see figure 6.28 c))
4. Curvature of shower front (see figure 6.28 d))

At this advanced stage of the analysis, additional checks have been performed, as shown for the various classifiers the TMVA response distribution in fig. 6.29, TMVA over-training checks in fig. 6.30, the convergence test for the MLPBNN in fig. 6.31, for the method of Fisher for signal and background in fig. 6.32 a) the probability density functions (PDFs), in b) the TMVA probability and in c) the TMVA rarity. In addition fig. A.9 shows the PCA transformations of the four input variables.

In short, the effect of over-training appears for ML techniques which have too few DoF, with the over-training's sensitivity depending on the method. This effect results in a seeming increase in the classification performance over the objectively achievable one, if measured on the training sample, and to an effective performance decrease when measured with an independent test sample. Fisher and LDAs can hardly be over-trained, and BDTs usually suffer

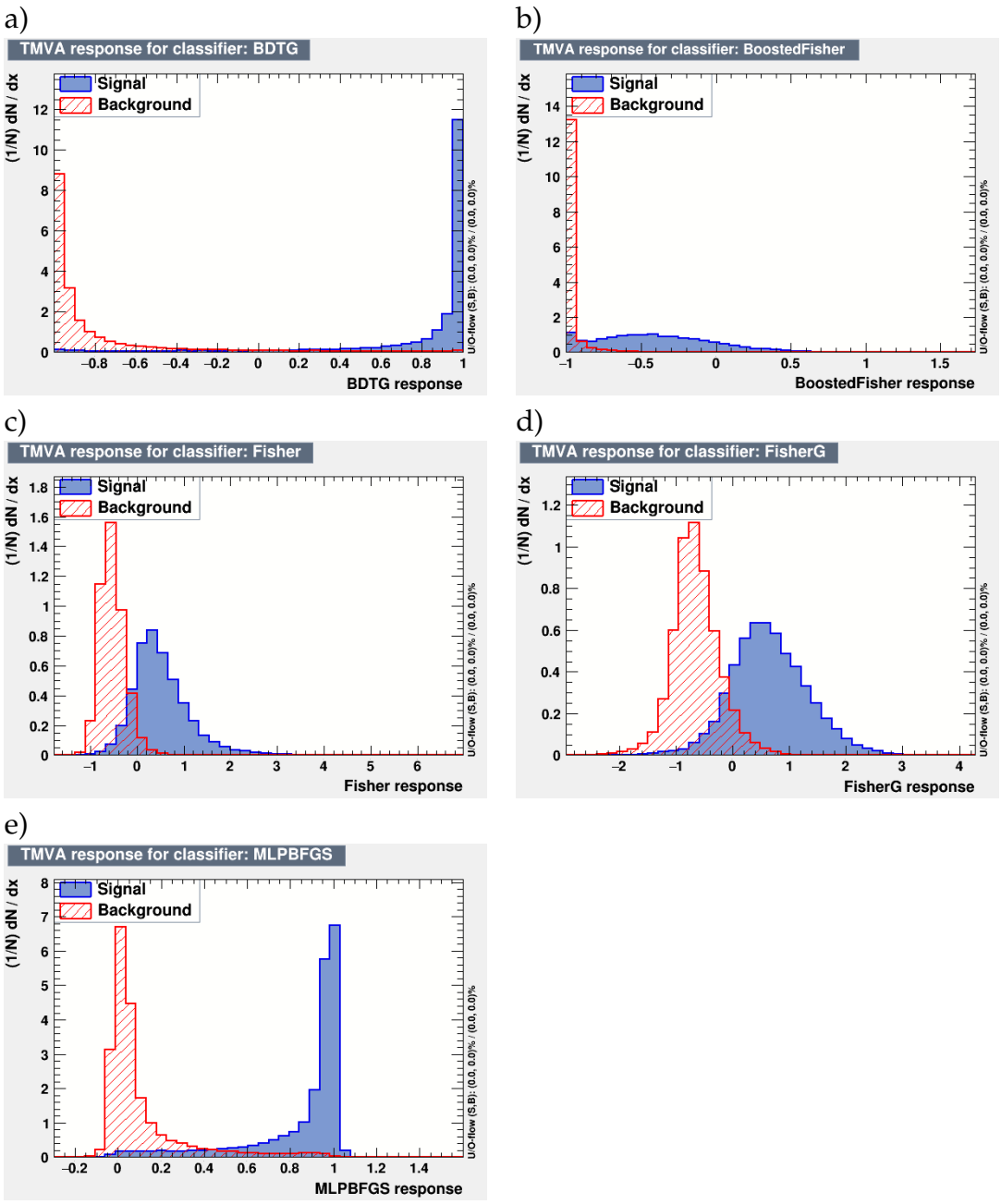


Figure 6.29.: TMVA response distributions **a)** BDTG **b)** BoostedFisher **c)** Fisher **d)** FisherG and e) MLPBFGS.

from at least partial over-training (without the appropriate counter measures) arising from their large number of nodes. The TMVA offers a convenient way to detect over-training and to measure its impact via comparing the performance results between training and test samples.

Figure 6.31 shows the MLP error-function convergence versus the training epoch for training and test events, called convergence test, which is performed at the same time as the over-training tests.

The TMVA also provides the classifier's signal and background PDFs (fig. 6.32 a)) $\hat{y}_{S(B)}$, which can be used to derive classification probabilities (fig. 6.32 b), as explained in [322]) for individual events, or to compute any kind of transformation of which the *Rarity* $\mathcal{R}(y)$ (of a classifier y) (fig. 6.32 c)) transformation is implemented, defined as:

$$\mathcal{R}(y) = \int_{-\infty}^{y} \hat{y}_B(y')\mathrm{d}y' \,, \tag{6.12}$$

defining $\mathcal{R}(y_B)$ for background events to be uniformly distributed between 0 and 1, while signal events cluster towards 1. Therefore the signal distributions can be directly compared among the various classifiers, with the stronger the peak towards 1, the better the discrimination. Furthermore, via exhibition of non-uniformity the Rarity can directly visualise deviations of a test background (physics data) from the training sample.

An interesting visualisation of the "inner kitchen" of the reported analysis obtained using the TMVA's *Shrink* method for BDTG is shown in figure 6.33.

The linear correlations of the input variables are shown in fig. 6.36 and the resulting MVA variable distributions in fig. 6.34 a) and cut efficiencies in b) for the method of BDTG. In addition a histogram of the SD energy distribution for the trained signal (= photons), background (= protons) and SD burn data is given in the appendix A.1.4.

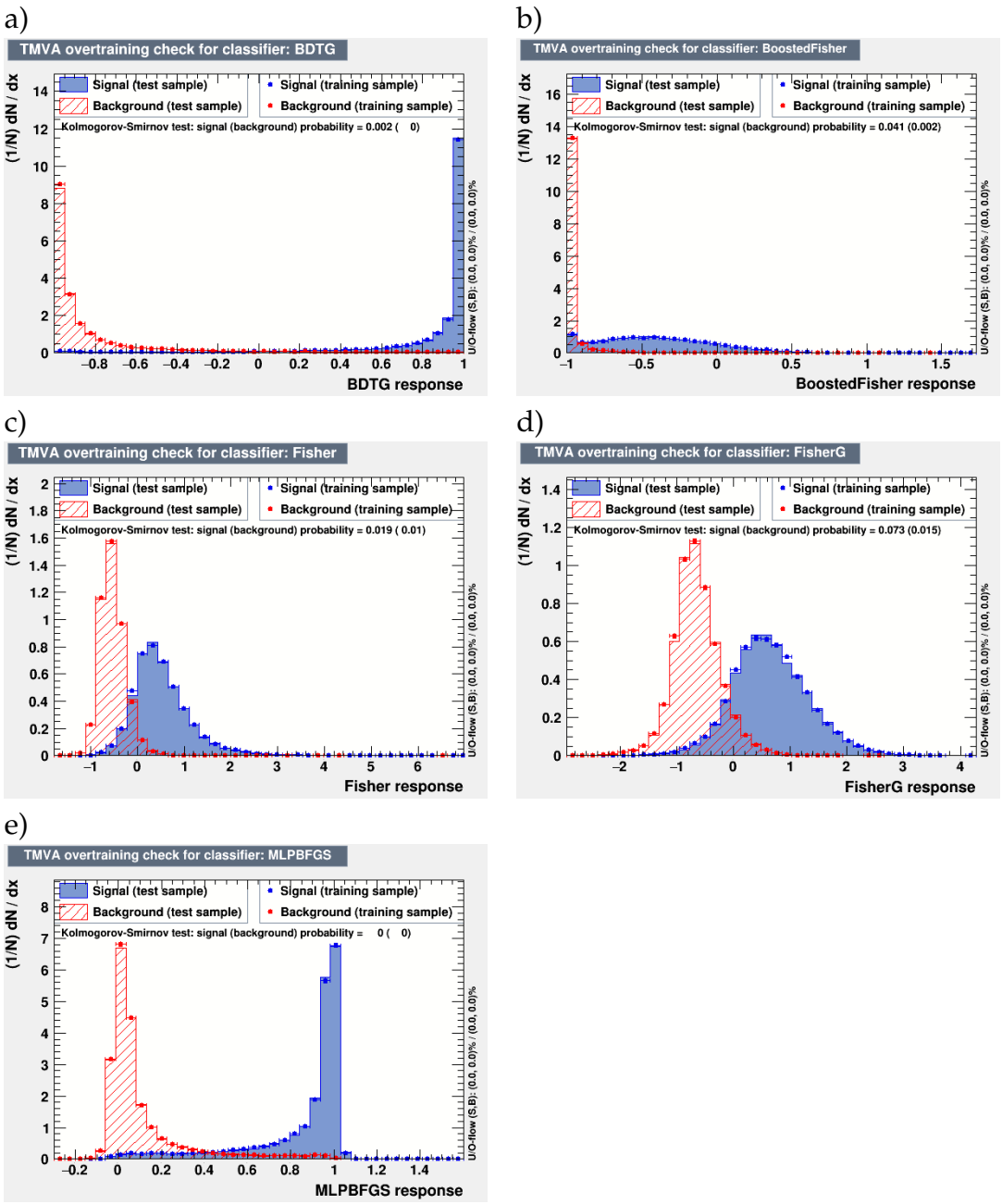


Figure 6.30.: TMVA over-training check for **a)** BDTG **b)** BoostedFisher **c)** Fisher **d)** FisherG and e) MLPBFGS. The good agreement between training and test distribution indicates that over-training is under control, where the shaded area represents the test sample while dots represent the training sample.

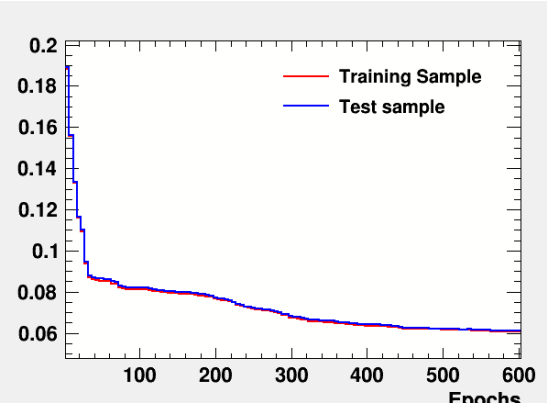


Figure 6.31.: TMVA convergence test for the MLPBNN.

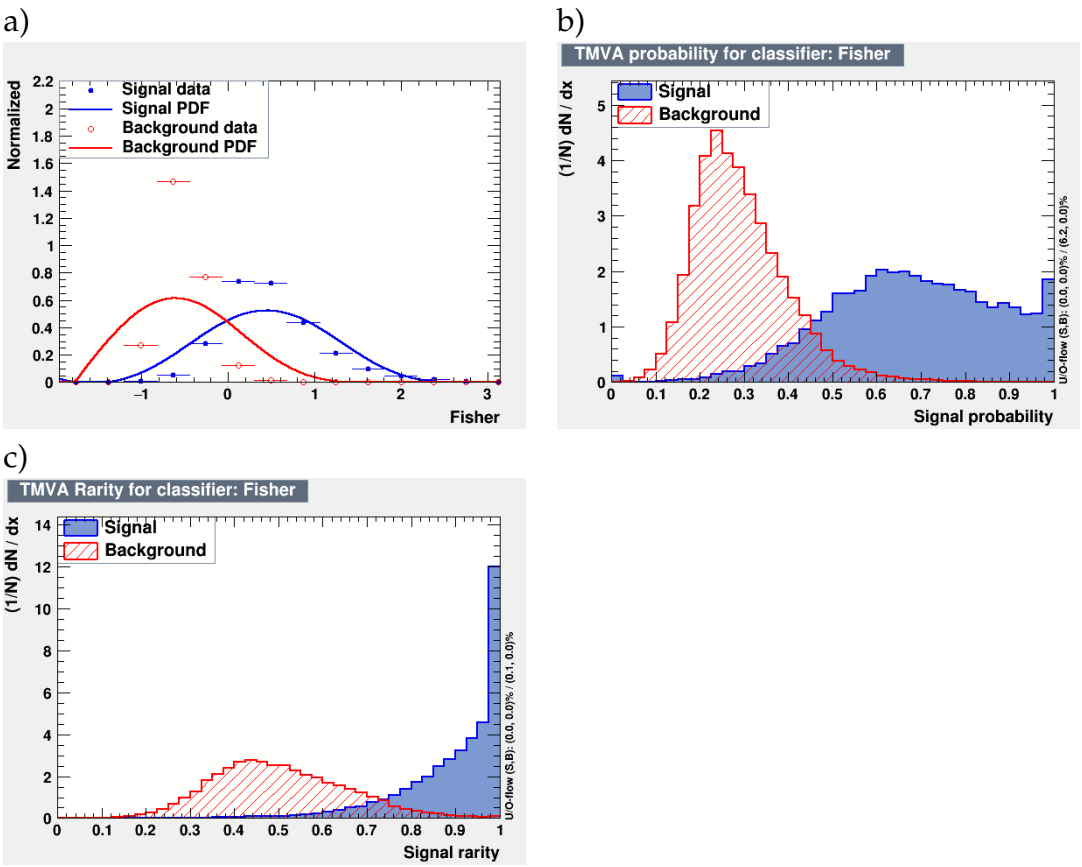


Figure 6.32.: For the method of Fisher for signal and background **a)** the PDFs, **b)** the TMVA probability and **c)** the TMVA rarity.

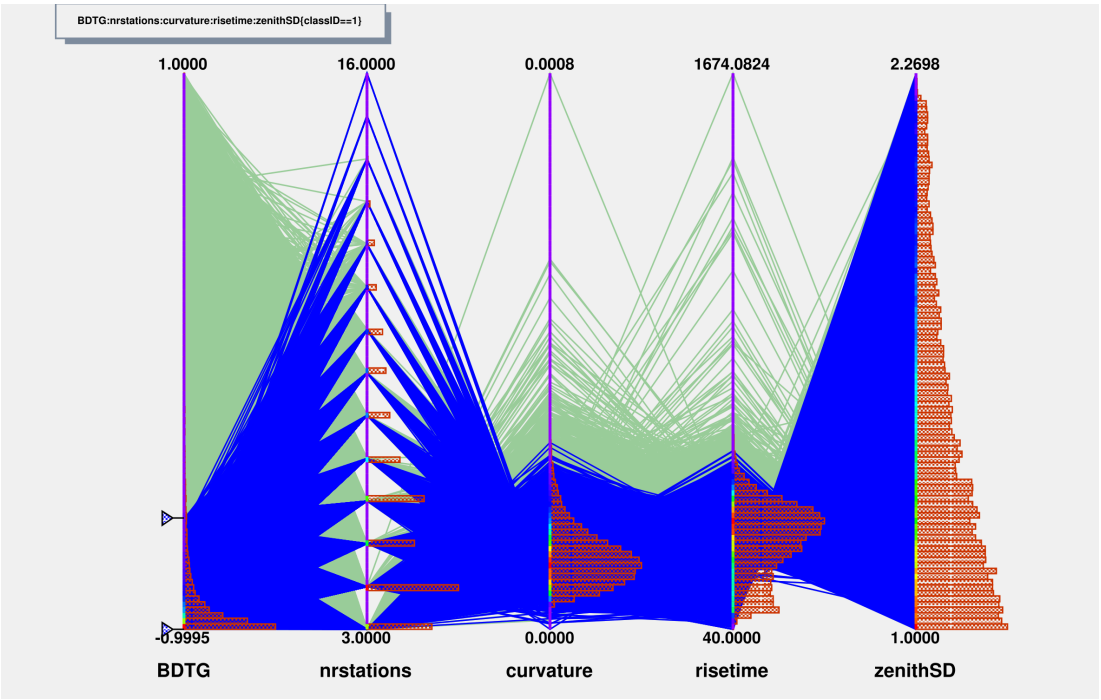


Figure 6.33.: Parallel coordinates representation for BDTG for signal (photons). Parallel coordinates of the Shrink method for photon (left) and proton (right) primaries. Green polylines are drawn for each event and connect parallel vertical axis, which represent the input observables and the classifier response. The individual value of an event is referred to via the position of the line on the vertical axis, and the frequency scale of each observable is represented by a red histogram superimposed to the vertical axis, and blue polylines mark events which are misclassified by the classifier response.

a) b)

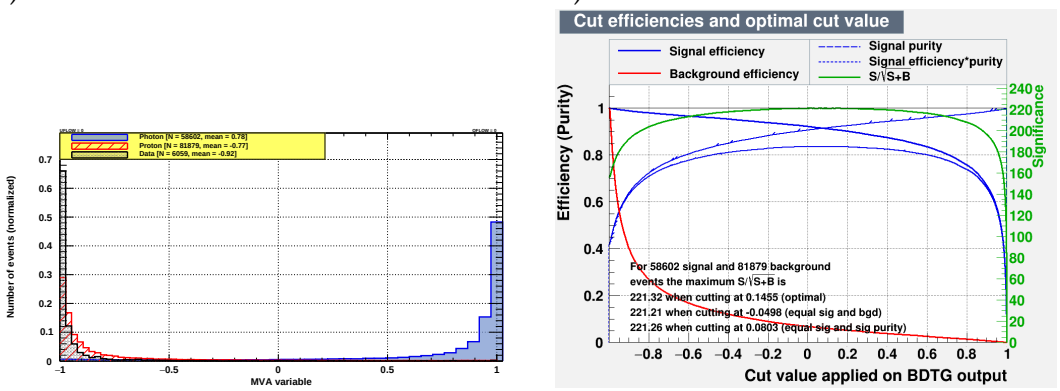


Figure 6.34.: Results of the four variables from fig. 6.28 combined in a MVA with BDTG **a)** MVA variable distribution showing excellent separation between signal (photons) and background (protons); **b)** cut efficiencies;

In figure 6.34 a) the MVA variable for the ML method of BDTG and its excellent separation power is presented, where the data distribution is as expected on the left side of the distribution for protons, since the mass composition measured with the Pierre Auger Observatory is heavier [71]. The MVA distribution and cut efficiency plots for the other tested methods that have been compared to each other, (shown in figures 6.35), can be found in the appendix A.2. Since ReLU (see sec. 6.5.1.5) is due to its success a very popular AF for deep NN, it was tested successfully in this work but without optimization, thereby still being less efficient than BDTG, it is shown in the appendix A.3.

In fig. 6.34 b) the cut efficiencies for the method of BDTG are shown. The resulting performances for the ML method of BDTG for the optimal cut value (of equal signal and signal purity, called "mean cut") of 0.0803 are:

- 53584 out of 58602 photons were correctly classified as photons ($\epsilon_S = 91.4\%$), the remaining 5018 were classified as protons ($r_S = 8.6\%$)
- 76918 out of 81879 protons were correctly classified as protons ($\epsilon_B = 93.9\%$), the remaining 4961 were classified as photons ($r_B = 6.1\%$)
- **95** real data events were classified as photons ($\mathcal{P} = 0.0157$) and 5964 were classified as protons (98.4%)

These resulting values are presented for further comparison with other MVA results in the appendix in table A.1.

This way of classifying SD events as photon candidates was proposed to be applied in the "real-time" photon candidate data stream to AMON. From the results above, these **95** as photons classified real data events, would be sent via

this new data stream to AMON. A **detailed technical proposal** (see appendix B.2) was submitted to the Pierre Auger Collaboration, describing the exact plans of how this stream will be implemented, together with the workflow of data handling and supplied by the details of the analysis presented here.

The good performance of the MVA as well as a direct comparison of the tested ML methods are shown in fig. 6.35 a) and b) via the inverse of the background efficiencies vs. the signal efficiency, and the ROC curve respectively.

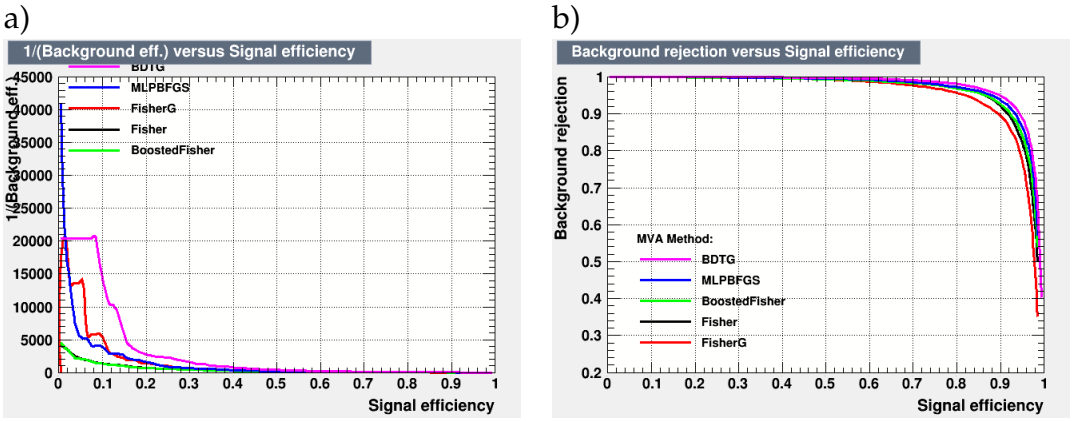


Figure 6.35.: a) inverse ROC curves, **b)** ROC curves; The various ML methods are ranked by the integral of the ROC diagram with the best method on top.

The correlation matrices for the BDTG method are presented in fig. 6.36.

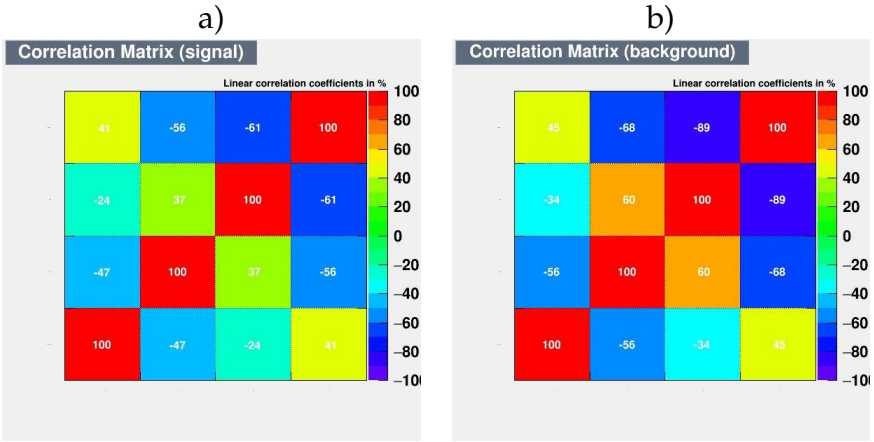


Figure 6.36.: Linear correlation matrices for with BDTG **a)** signal, **b)** background.

Since the latency is an important aspect for this work, also the time of raw data reconstruction has to be considered. To save some time it would be possible not to create complete ADST files, but via reconfiguring the module *SdCompositionParameters* one can extract only the needed parameters. However,

this is of insignificant magnitude for the overall achieved latency, because first of all the raw files are small and contain not a huge number of events, what refers to the second and main contributing factor in the latency - these raw SD files ("sd_online.root.nobackup", which are in the CDAS format) are obtained only every 15 min in the Observer production, and therefore stems an even-wise analysis impossible.

Due to this unfortunately big dead time till new raw files are received, further ideas to reduce analysis time, like e.g. leave out tests (*SdPMTSignalShapeQualityChecker*, *SdBadStationRejectorKG*) and apply them later on, have not been pursued.

6.6.6.1. Comparison against mixed proton-iron background

As a part of systematic cross checks, instead of the "worst case scenario" background for UHE photons, composed purely of protons presented above, a background composition of **50% proton** & **50% iron** was chosen, since for photons every primary CR nucleus heavier than protons is qualitatively considered to be iron. From this of course one expects a better separation power, i.e. more data events are expected to be classified by the MVA as photons. The same input variables (but different events from simulations, plots shown in 6.37 a-d)), the same data sample (SD burn sample), the same quality criteria were used as in the previous section.

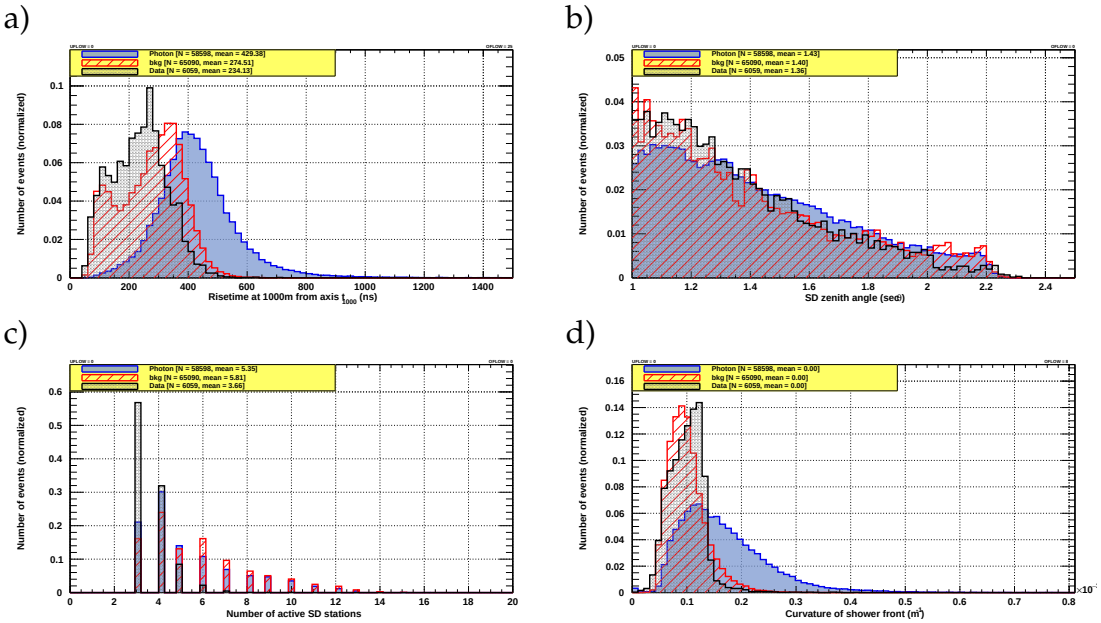


Figure 6.37.: Distributions for the observables with the background being composed of 50% proton and 50% iron **a)** $t_{1/2}(1000)$ **b)** Zenith angle, expressed as $\sec(\theta)$ **c)** NSt **d)** curvature of the shower front [1/km].

For this cross check only the best method of BDTG was applied. The resulting MVA variable distributions can be seen in fig. 6.38 a) and cut efficiencies in b) for the method of BDTG.

In comparison to the results in the previous section (sec. 6.6.6) the results of classification with choosing the cut the same as before to be the mean cut (equal signal and signal purity), with now a value of -0.0018, resulted in:

a) b)

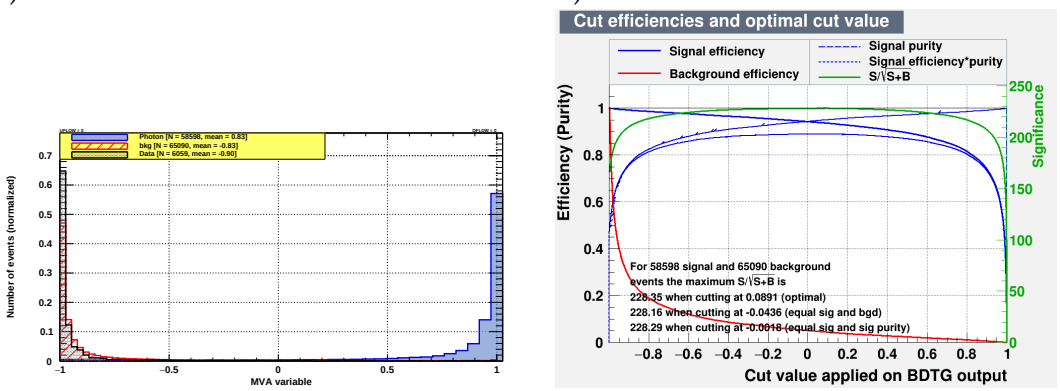


Figure 6.38.: Results of the four variables from fig. 6.28 combined in a MVA with BDTG **a)** MVA variable distribution showing excellent separation between signal (photons) and background (50% protons + 50% iron); **b)** cut efficiencies;

- Photons: out of 58598 events, 55240 classified as photons ($\epsilon_S = 94.27\%$), vs. 3358 classified as background ($r_S = 5.73\%$)
- Background: out of 65090 events, 3382 classified as photons ($r_B = 5.2\%$), vs. 61708 classified as background ($\epsilon_B = 94.8\%$)
- Data: out of 6059 events, **149** classified as photons ($\mathcal{P} = 0.0246$), vs. 5910 classified as background (97.54%)

This shows a significantly better separation than the case of training with a background composed of pure proton,
better signal efficiency: 94.27% (sec. 6.6.6.1) vs. 91.4% (sec. 6.6.6),
better purity: 2.46% vs. 1.57%, 149 vs. 95 events,
and smaller miss rate (type II error): 5.73% vs. 8.6%, as expected after training the MVA on a more heavy background sample (for comparison of the values, see also table A.1 in the appendix).

7. Conclusions

This work has its focus on the earliest known, and main messengers utilized to investigate the Universe: photons. While cosmic EM radiation has been observed up to now in an extremely wide energy range from below 10^{-8} eV to few 10^{15} eV, theoretical models of the UHECR production predict a certain photon fraction at EeV energies with a variety of prospects. The observation of photons in the highest energy regime would open a new observational window to the Universe and potentially uncloak new physics beyond the standard model.

The combination of UHE photons together with one or more of the other three messengers of natures fundamental forces, the most energetic particles known to date – CRs, neutrinos and GWs, as well as with lower energy EM radiation, can reach far beyond the capabilities of what one single messenger or a single experiment could ever achieve. As timing and directional coincidences of the data collected by different observatories may help to reveal new astrophysical objects and to enrich our understanding of underlying physical processes, the progress in all subtopics in this field, and especially international collaborations in MM studies, remains important.

To further support global endeavours of MM studies, this work, performed within the world's largest CR experiment – the Pierre Auger Observatory – presents the new advanced analysis methods to search for UHE photon candidates in the Auger data. The obtained results allow a straightforward implementation of UHE photon candidate event stream, to one of the global networks of observatories – the AMON.

An event of an EAS initiated by a primary particle interacting in the Earth's atmosphere, detected with the Pierre Auger Observatory's SD array, results in observables constituting the measured data. A challenge of data analysis is to distinguish between primary particle species of measured events. Several observables are known to incorporate discriminating power between potential photons and the predominant background of hadrons. Stand-alone, none of these observables can give satisfying discriminating statements about the nature of the primary particles.

To reach a compromise between a moderate level of analysis, a small latency, and still achieve sophisticated classification statements about sub-threshold photon candidate events, advantage was taken of the MVA approach. This was achieved in this work through the combination of several observables via the well developed ROOT integrated TMVA framework together with its straightforward usability, and was shown to satisfy the data analysis needs.

Various powerful ML techniques were probed and compared to each other, on MC simulations and real data, to find out which method performs best for this specific task. The best performing method was found to be the BDTG, the advanced classifier, which takes into account high-dimensional correlations using the gradient boost algorithm in combination with bagging-like resampling and a small shrinkage parameter (thus reducing the learning rate and hence making the method less susceptible for over-training). The presented work is relatively robust regarding its purpose.

The analysis methods developed in this work can be easily adapted to include potential future extensions. The ongoing upgrade of the Pierre Auger Observatory, *AugerPrime*, with additional RDs and the SSDs on top of each SD station, is in particular promising from an experimental point of view for future UHE photon searches, as it will immensely increase the power to discriminate photons from other primary species. Further mass-sensitive observables from the SD array data, which have been developed or are under construction, as well as observables from *AugerPrime* can be joined in the presented approach. In addition a reduction of the photon energy threshold can be achieved via the low-energy extensions of the SD 750 m array, including the underground muon detector array of the finalized AMIGA detector. Further possible extensions of this work arise from the great potential and constant improvement of the TMVA framework. Due to category extensions, dependencies of observables on the energy and zenith angle could be reduced, the performance of classification increased, and not just binary but even multi-class separation of primary particle types could be established in multiple primary candidate streams to global MM networks to fully exploit the Pierre Auger Observatory's detection capabilities.

Bibliography

[1] V. F. Hess. "*Über Beobachtungen der durchdringenden Strahlung bei sieben Freiballonfahrten*". In: *Physikalische Zeitschrift* **13** (1912), pp. 1084–1091.

[2] Z. Cao, F. A. Aharonian, Q. An, et al. "*Ultrahigh-energy photons up to 1.4 petaelectronvolts from 12 γ-ray Galactic sources*". In: *Nature* (2021). URL: https://doi.org/10.1038/s41586-021-03498-z.

[3] The Nobel Foundation". "*Nobel Lectures, Physics 1922-1941*". In: *Elsevier Publishing Company, Amsterdam* (1965).

[4] M. Curie. "*Rayons e'mis par les composés de l'uranium et du thorium*". In: *Comptes rendus, Paris* **126 (15)** (1898), pp. 1101–1103.

[5] C. T. R. Wilson. "*On the ionisation of atmospheric air*". In: *Proceedings of the Royal Society of London* **68 (444)** (1901), pp. 151–161.

[6] J. Elster and H. Geitel. "*Weitere Versuche über die Elektrizitätszerstreuung in abgeschlossenen Luftmengen*". In: *Phys. Ztschr.* **2 (38)** (1900), pp. 560–563.

[7] T. Wulf. "*Ein neues Elektrometer für statische Ladungen*". In: *Physikalische Zeitschrift* **8** (1907), pp. 246–248.

[8] T. Wulf. "*Beobachtungen über die Strahlung hoher Durchdringungsfähigkeit auf dem Eiffelturm*". In: *Phys. Ztschr.* **11 (18)** (1910), pp. 811–813.

[9] F. Linke. "*Luftelektrische Messungen bei zwölf Ballonfahrten*". In: *Abhandlungen der Gesellschaft der Wissenschaften in Göttingen, mathematisch-physikalische Klasse* **3** (1904), pp. 1–90.

[10] A. Gockel and T. Wulf. "*Beobachtungen über die Radioaktivität der Atmosphäre im Hochgebirge*". In: *Physikalische Zeitschrift* **9** (1908), pp. 907–911.

[11] D. Pacini. "*La radiation pénétrante sur la mer*". In: *Le Radium* **8** (1911), pp. 307–312.

[12] V. F. Hess. "*Über die Absorption der γ-Strahlen in der Atmosphäre*". In: *Physikalische Zeitschrift* **12** (1911), pp. 998–1001.

[13] V. F. Hess. "*Über Neuerungen und Erfahrungen an den Radiummessungen nach der γ-Strahlenmethode*". In: *Physikalische Zeitschrift* **14** (1913), pp. 1135–1141.

[14] W. Kolhörster. "*Messungen der durchdringenden Strahlung bis in Höhen von 9300 m*". In: *Verhandlungen der Deutschen Physikalischen Gesellschaft* **16** (1914), pp. 719–721.

[15] R. A. Millikan and G. H. Cameron. "*High frequency rays of cosmic origin III. Measurements in snow-fed lakes at high altitudes*". In: *Physical Review* **28**.5 (1926), pp. 851–868. DOI: 10.1103/PhysRev.28.851.

[16] J. Clay. "*Penetrating Radiation*". In: *Proceedings of the Section of Sciences* **30**.9–10 (1927), pp. 1115–1127.

[17] W. Bothe and W. Kolhörster. "*Das Wesen der Höhenstrahlung*". In: *Zeitschrift für Physik* **56** (1929), pp. 751–777.

[18] B. Rossi. "*On the Magnetic Deflection of Cosmic Rays*". In: *Phys. Rev.* **36** (3 1930), pp. 606–606. DOI: 10.1103/PhysRev.36.606. URL: https://link.aps.org/doi/10.1103/PhysRev.36.606.

[19] T. H. Johnson. "*The Azimuthal Asymmetry of the Cosmic Radiation*". In: *Phys. Rev.* **43** (10 1933), pp. 834–835. DOI: 10.1103/PhysRev.43.834. URL: https://link.aps.org/doi/10.1103/PhysRev.43.834.

[20] L. Alvarez and A. H. Compton. "*A Positively Charged Component of Cosmic Rays*". In: *Phys. Rev.* **43** (10 1933), pp. 835–836. DOI: 10.1103/PhysRev.43.835. URL: https://link.aps.org/doi/10.1103/PhysRev.43.835.

[21] B. Rossi. "*Directional Measurements on the Cosmic Rays Near the Geomagnetic Equator*". In: *Phys. Rev.* **45** (3 1934), pp. 212–214. DOI: 10.1103/PhysRev.45.212. URL: https://link.aps.org/doi/10.1103/PhysRev.45.212.

[22] A. H. Compton. "*A Geographic Study of Cosmic Rays*". In: *The Physical Review* **43** (1933), pp. 387–403.

[23] P. A. M. Dirac. "*A theory of electrons and protons*". In: *Proc. R. Soc. Lond. A* **126** (1930), pp. 360–365. DOI: 10.1098/rspa.1930.0013.

[24] C. D. Anderson. "*The positive electron*". In: *Physical Review* **43**.6 (1933), pp. 491–494. DOI: 10.1103/PhysRev.43.491.

[25] Wellcome Collection gallery (2018-03-30). King of Sweden presents Nobel prize. URL: https://wellcomecollection.org/works/hjvbu2f7.

[26] P. Auger, R. Maze, P. Jr. Ehrenfest, and A. Fréon. "*Les grandes gerbes de rayons cosmiques*". In: *Journal de Physique et Le Radium* **10** (1939), pp. 39–48.

[27] E. Regener and G. Pfotzer. "*Vertical Intensity of Cosmic Rays by Threefold Coincidences in the Stratosphere*". In: *Nature* **136** (1935), pp. 718–719.

[28] B. Rossi. "*Misure sulla distribuzione angolare di intensita della radiazione penetrante all'Asmara*". In: *Ricerca Scientifica* **5** (1 1934), pp. 579–589.

[29] H. J. Bhabha and W. Heitler. "*The Passage of Fast Electrons and the Theory of Cosmic Showers*". In: *Proceedings of the Royal Society of London A* **159** (1937), pp. 432–458.

[30] S. H. Neddermeyer and C. D. Anderson. "*Note on the Nature of Cosmic-Ray Particles*". In: *The Physical Review* **51** (1937), pp. 884–886.

[31] G. D. Rochester and C. C. Butler. "*Evidence for the existence of new unstable elementary particles*". In: *Nature* **160** (1947), pp. 855–857.

[32] P. Freier, E. J. Lofgren, E. P. Ney, F. Oppenheimer, H. L. Bradt, and B. Peters. "*Evidence for Heavy Nuclei in the Primary Cosmic Radiation*". In: *Phys. Rev.* **74** (2 1948), pp. 213–217. DOI: 10.1103/PhysRev.74.213. URL: https://link.aps.org/doi/10.1103/PhysRev.74.213.

[33] H. L. Bradt and B. Peters. "*Investigation of the Primary Cosmic Radiation with Nuclear Photographic Emulsions*". In: *Phys. Rev.* **74** (12 1948), pp. 1828–1837. DOI: 10.1103/PhysRev.74.1828. URL: https://link.aps.org/doi/10.1103/PhysRev.74.1828.

[34] E. Fermi. "*On the Origin of the Cosmic Radiation*". In: *Phys. Rev.* **75** (8 1949), pp. 1169–1174. DOI: 10.1103/PhysRev.75.1169. URL: https://link.aps.org/doi/10.1103/PhysRev.75.1169.

[35] G. T. Zatsepin and V. A. Kuzmin. "*Upper limit of the spectrum of cosmic rays*". In: *JETP Lett.* **4** (1966), pp. 78–80.

[36] K. Greisen. "*End to the Cosmic-Ray Spectrum?*" In: *Phys. Rev. Lett.* **16** (17 1966), pp. 748–750. DOI: 10.1103/PhysRevLett.16.748. URL: https://link.aps.org/doi/10.1103/PhysRevLett.16.748.

[37] C. Evoli. The Cosmic-Ray Energy Spectrum. accessed June 2021. 2020. URL: http://doi.org/10.5281/zenodo.4396125.

[38] T. Sanuki et al. "*Precise Measurement of Cosmic-Ray Proton and Helium Spectra with the BESS Spectrometer*". In: *The Astrophysical Journal* **545**.2 (2000), pp. 1135–1142. DOI: 10.1086/317873. URL: https://doi.org/10.1086/317873.

[39] H. S. Ahn et al. "*Measurements of cosmic-ray secondary nuclei at high energies with the first flight of the CREAM balloon-borne experiment*". In: *Astroparticle Physics* **30**.3 (2008), pp. 133–141. ISSN: 0927-6505. DOI: https://doi.org/10.1016/j.astropartphys.2008.07.010. URL: https://www.sciencedirect.com/science/article/pii/S092765050800114X.

[40] O. Adriani et al. "*Cosmic-Ray Electron Flux Measured by the PAMELA Experiment between 1 and 625 GeV*". In: *Phys. Rev. Lett.* **106** (20 2011), p. 201101. DOI: 10.1103/PhysRevLett.106.201101. URL: https://link.aps.org/doi/10.1103/PhysRevLett.106.201101.

[41] M. Aguilar et al. "*First Result from the Alpha Magnetic Spectrometer on the International Space Station: Precision Measurement of the Positron Fraction in Primary Cosmic Rays of 0.5–350 GeV*". In: *Phys. Rev. Lett.* **110** (14 2013), p. 141102. DOI: 10.1103/PhysRevLett.110.141102. URL: https://link.aps.org/doi/10.1103/PhysRevLett.110.141102.

[42] M. Tanabashi et al. "*Review of Particle Physics*". In: *Phys. Rev. D* **98**.3 (2018), p. 030001. DOI: 10.1103/PhysRevD.98.030001.

[43] C. Grupen. Astroparticle Physics. Berlin, Heidelberg: Springer-Verlag, 2005. ISBN: 3-540-25312-2.

[44] J.-L. Tassoul and M. Tassoul. A Concise History of Solar and Stellar Physics. Princeton University Press, 2014. ISBN: 9780691165929. DOI: doi:10.1515/9781400865390. URL: https://doi.org/10.1515/9781400865390.

[45] J. Linsley. "*Evidence for a Primary Cosmic-Ray Particle with Energy* 10^{20} *eV*". In: *Phys. Rev. Lett.* **10** (4 1963), pp. 146–148. DOI: 10.1103/PhysRevLett.10.146. URL: https://link.aps.org/doi/10.1103/PhysRevLett.10.146.

[46] M. A. Lawrence, R. J. O. Reid, and A. A. Watson. "*The cosmic ray energy spectrum above* $4 * 10^{17}$ *eV as measured by the Haverah Park array*". In: *Journal of Physics G: Nuclear and Particle Physics* **17**.5 (1991), pp. 733–757. DOI: 10.1088/0954-3899/17/5/019. URL: https://doi.org/10.1088/0954-3899/17/5/019.

[47] S. Yoshida et al. "*The cosmic ray energy spectrum above* 3×10^{18} *eV measured by the Akeno Giant Air Shower Array*". In: *Astroparticle Physics* **3**.2 (1995), pp. 105–123. ISSN: 0927-6505. DOI: https://doi.org/10.1016/0927-6505(94)00036-3. URL: https://www.sciencedirect.com/science/article/pii/0927650594000363.

[48] D. Bird et al. "*The cosmic-ray energy spectrum observed by the Fly's Eye*". In: *The Astrophysical Journal* **424** (1994), pp. 491–502. DOI: 10.1086/173906.

[49] A. Aab et al. "*Features of the Energy Spectrum of Cosmic Rays above* 2.5×10^{18} eV *Using the Pierre Auger Observatory*". In: *Phys. Rev. Lett.* **125** (12 2020), p. 121106. DOI: 10.1103/PhysRevLett.125.121106. URL: https://link.aps.org/doi/10.1103/PhysRevLett.125.121106.

[50] W. D. Apel et al. "*KASCADE-Grande measurements of energy spectra for elemental groups of cosmic rays*". In: *Astroparticle Physics* **47** (2013), pp. 54–66. ISSN: 0927-6505. DOI: https://doi.org/10.1016/j.astropartphys.2013.06.004. URL: https://www.sciencedirect.com/science/article/pii/S0927650513000996.

[51] S. Y. Lee. Accelerator Physics (second Edition). 2004. DOI: 10.1142/5761.

[52] B. Peters. "*Primary cosmic radiation and extensive air showers*". In: *Nuovo Cim* **22** (1961), pp. 800–819. DOI: https://doi.org/10.1007/BF02783106.

[53] V. Ptuskin, S. Rogovaya, V. Zirakashvili, L. Chuvilgin, G. Khristiansen, E. Klepach, and G. Kulikov. "*Diffusion and drift of very high energy cosmic rays in galactic magnetic fields*". In: *Astronomy and Astrophysics* **268** (Jan. 1993), pp. 726–735. ISSN: 0004-6361.

[54] J. R. Hörandel. "*Models of the knee in the energy spectrum of cosmic rays*". In: *Astroparticle Physics* **21**.3 (2004), pp. 241–265. ISSN: 0927-6505. DOI: https://doi.org/10.1016/j.astropartphys.2004.01.004. URL: https://www.sciencedirect.com/science/article/pii/S0927650504000209.

[55] J. Blümer, R. Engel, and J. R. Hörandel. "*Cosmic rays from the knee to the highest energies*". In: *Progress in Particle and Nuclear Physics* **63**.2 (2009), pp. 293–338. ISSN: 0146-6410. DOI: https://doi.org/10.1016/j.ppnp.2009.05.002. URL: https://www.sciencedirect.com/science/article/pii/S0146641009000362.

[56] W. D. Apel et al. "*Ankle-like feature in the energy spectrum of light elements of cosmic rays observed with KASCADE-Grande*". In: *Phys. Rev. D* **87** (8 2013), p. 081101. DOI: 10.1103/PhysRevD.87.081101. URL: https://link.aps.org/doi/10.1103/PhysRevD.87.081101.

[57] W. D. Apel et al. "*Kneelike Structure in the Spectrum of the Heavy Component of Cosmic Rays Observed with KASCADE-Grande*". In: *Phys. Rev. Lett.* **107** (17 2011), p. 171104. DOI: 10.1103/PhysRevLett.107.171104. URL: https://link.aps.org/doi/10.1103/PhysRevLett.107.171104.

[58] J. R. Hörandel. "*On the knee in the energy spectrum of cosmic rays*". In: *Astroparticle Physics* **19**.2 (2003), pp. 193–220. ISSN: 0927-6505. URL: http://dx.doi.org/10.1016/S0927-6505(02)00198-6.

[59] R. Aloisio, V. Berezinsky, P. Blasi, A. Gazizov, S. Grigorieva, and B. Hnatyk. "*dip in the UHECR spectrum and the transition from galactic to extragalactic cosmic rays*". In: *Astroparticle Physics* **27**.1 (2007), pp. 76–91. ISSN: 0927-6505. DOI: 10.1016/j.astropartphys.2006.09.004. URL: http://dx.doi.org/10.1016/j.astropartphys.2006.09.004.

[60] R. Aloisio, V. Berezinsky, and A. Gazizov. "*Transition from galactic to extragalactic cosmic rays*". In: *Astroparticle Physics* **39–40** (2012). Cosmic Rays Topical Issue, pp. 129–143. ISSN: 0927-6505. DOI: https://doi.org/10.1016/j.astropartphys.2012.09.007. URL: https://www.sciencedirect.com/science/article/pii/S092765051200182X.

[61] A. Castellina. "*Highlights from the Pierre Auger Observatory (ICRC2019)*". In: *PoS* **ICRC2019** (2019), p. 004. arXiv: 1909.10791 [astro-ph.HE].

[62] E. Roulet. "*Large-scale anisotropies above 0.03 EeV measured by the Pierre Auger Observatory*". In: *PoS* **ICRC2019** (2019), p. 408. DOI: 10.22323/1.358.0408.

[63] A. H. Compton and I. A. Getting. "*An Apparent Effect of Galactic Rotation on the Intensity of Cosmic Rays*". In: *Phys. Rev.* **47** (11 1935), pp. 817–821. DOI: 10.1103/PhysRev.47.817. URL: https://link.aps.org/doi/10.1103/PhysRev.47.817.

[64] L. J. Gleeson and W. I. Axford. "*The compton-getting effect*". In: *Astrophys Space Sci* **2** (1968), pp. 431–437. DOI: 10.1007/BF02175919. URL: https://doi.org/10.1007/BF02175919.

[65] Y. Zhang. "*Large-scale Cosmic Ray Anisotropy with Tibet air shower array*". In: *PoS* **ICRC2019** (2019), p. 488. DOI: 10.22323/1.358.0488.

[66] R. Aloisio, P. Blasi, and V. Berezinsky. "*Ultra high energy cosmic rays: implications of Auger data for source spectra and chemical composition*". In: *Journal of Cosmology and Astroparticle Physics* **2014**.10 (2014). ISSN: 1475-7516. DOI: 10.1088/1475-7516/2014/10/020. URL: https://www.osti.gov/biblio/22375834.

[67] A. Aab et al. "*Large-scale Cosmic-Ray Anisotropies above 4 EeV Measured by the Pierre Auger Observatory*". In: *The Astrophysical Journal* **868**.1 (2018), p. 4. DOI: 10.3847/1538-4357/aae689. URL: https://doi.org/10.3847/1538-4357/aae689.

[68] E. G. Berezhko. "*Maximum energy of cosmic rays accelerated by supernova shocks*". In: *Astroparticle Physics* **5**.3 (1996), pp. 367–378. ISSN: 0927-6505. DOI: https://doi.org/10.1016/0927-6505(96)00037-0. URL: https://www.sciencedirect.com/science/article/pii/0927650596000370.

[69] V. Berezinsky. "*On transiton from galactic to extragalactic cosmic rays*". In: *Journal of Physics: Conference Series* **47** (2006), pp. 142–153. DOI: 10.1088/1742-6596/47/1/018. URL: https://doi.org/10.1088/1742-6596/47/1/018.

[70] A. Aab et al. "*Combined fit of spectrum and composition data as measured by the Pierre Auger Observatory*". In: *JCAP* **2017**.4, 038 (Apr. 2017). Erratum: JCAP**2018**03,no.03,E02(2018), doi:10.1088/1475-7516/2018/03/E02, p. 038. DOI: 10.1088/1475-7516/2017/04/038. arXiv: 1612.07155 [astro-ph.HE].

[71] A. Yushkov. "*Mass Composition of Cosmic Rays with Energies above $10^{17.2}$ eV from the Hybrid Data of the Pierre Auger Observatory*". In: *PoS* **ICRC2019** (2019), p. 482. DOI: 10.22323/1.358.0482.

[72] G. R. Farrar, M. Unger, and L. A. Anchordoqui. Origin of the ankle in the ultra-high energy cosmic ray spectrum and of the extragalactic protons below it. 2015. arXiv: 1512.00484 [astro-ph.HE].

[73] V. Verzi. "*Measurement of the energy spectrum of ultra-high energy cosmic rays using the Pierre Auger Observatory*". In: *PoS* **ICRC2019** (2019), p. 450. DOI: 10.22323/1.358.0450.

[74] R. Aloisio, V. Berezinsky, and A. Gazizov. "*Ultra high energy cosmic rays: The disappointing model*". In: *Astroparticle Physics* **34**.8 (2011), pp. 620–626. ISSN: 0927-6505. DOI: https://doi.org/10.1016/j.astropartphys.2010.12.008. URL: https://www.sciencedirect.com/science/article/pii/S0927650510002434.

[75] A. Aab et al. "*Measurement of the cosmic-ray energy spectrum above* 2.5×10^{18} eV *using the Pierre Auger Observatory*". In: *Phys. Rev. D* **102** (6 2020), p. 062005. DOI: 10.1103/PhysRevD.102.062005. URL: https://link.aps.org/doi/10.1103/PhysRevD.102.062005.

[76] O. Deligny. "*The energy spectrum of ultra-high energy cosmic rays measured at the Pierre Auger Observatory and at the Telescope Array*". In: *PoS* **ICRC2019** (2019), p. 234. DOI: 10.22323/1.358.0234.

[77] M. Lemoine and G. Sigl. Physics and Astrophysics of Ultra High Energy Cosmic Rays. Vol. **576**. Lecture Notes in Physics. Berlin Heidelberg: Springer-Verlag, 2001. ISBN: 978-3-540-45615-5. DOI: 10.1007/3-540-45615-5.

[78] A. Venkatesan, M. Coleman Miller, and A. V. Olinto. "*Constraints on the Production of Ultra–High-Energy Cosmic Rays by Isolated Neutron Stars*". In: *The Astrophysical Journal* **484**.1 (1997), pp. 323–328. DOI: 10.1086/304304. URL: https://doi.org/10.1086/304304.

[79] D. Caprioli. "*An Original Mechanism for the Acceleration of Ultra-High-Energy Cosmic Rays*". In: *Nuclear and Particle Physics Proceedings* **297-299** (2018). Cosmic Ray Origin - Beyond the Standard Models, pp. 226–233. ISSN: 2405-6014. DOI: https://doi.org/10.1016/j.nuclphysbps.2018.07.032. URL: https://www.sciencedirect.com/science/article/pii/S2405601418301214.

[80] D. Caprioli. ""*ESPRESSO*" *ACCELERATION OF ULTRA-HIGH-ENERGY COSMIC RAYS*". In: *ApJL* **811**.2 (2015), p. L38. DOI: 10.1088/2041-8205/811/2/L38.

[81] A. M. Hillas. "*The Origin of Ultra-High-Energy Cosmic Rays*". In: *Annual Review of Astronomy and Astrophysics* **22**.1 (1984), pp. 425–444. DOI: 10.1146/annurev.aa.22.090184.002233. eprint: https://doi.org/10.1146/annurev.aa.22.090184.002233. URL: https://doi.org/10.1146/annurev.aa.22.090184.002233.

[82] R. Alves Batista et al. "*Open Questions in Cosmic-Ray Research at Ultrahigh Energies*". In: *Frontiers in Astronomy and Space Sciences* **6** (2019), p. 23. ISSN: 2296-987X. DOI: 10.3389/fspas.2019.00023. URL: https://www.frontiersin.org/article/10.3389/fspas.2019.00023.

[83] D. Giannios. "*UHECRs from magnetic reconnection in relativistic jets*". In: *Monthly Notices of the Royal Astronomical Society: Letters* **408**.1 (Oct. 2010), pp. L46–L50. ISSN: 1745-3925. DOI: 10.1111/j.1745-3933.2010.00925.x. eprint: https://academic.oup.com/mnrasl/article-pdf/408/1/L46/3266695/408-1-L46.pdf. URL: https://doi.org/10.1111/j.1745-3933.2010.00925.x.

[84] E. M. de Gouveia Dal Pino, R. A. Batista, G. Kowal, T. Medina-Torrejón, and J. C. Ramirez-Rodriguez. Magnetic Reconnection, Cosmic Ray Acceleration, and Gamma-Ray emission around Black Holes and Relativistic Jets. 2019. arXiv: 1903.08982 [astro-ph.HE].

[85] M. Lyutikov and R. Ouyed. "*Inductive acceleration of UHECRs in sheared relativistic jets*". In: *Astroparticle Physics* **27**.6 (2007), pp. 473–489. ISSN: 0927-6505. DOI: 10.1016/j.astropartphys.2007.01.008. URL: http://dx.doi.org/10.1016/j.astropartphys.2007.01.008.

[86] C. Guépin, B. Cerutti, and K. Kotera. "*Proton acceleration in pulsar magnetospheres*". In: *Astronomy and Astrophysics* **635** (Mar. 2020), A138. ISSN: 1432-0746. DOI: 10.1051/0004-6361/201936816. URL: http://dx.doi.org/10.1051/0004-6361/201936816.

[87] A. M. Bykov and I. N. Toptygin. "*Effect of shocks on interstellar turbulence and cosmic-ray dynamics*". In: *Astrophys Space Sci* **138** (1987), pp. 341–354. DOI: 10.1007/BF00637855. URL: https://doi.org/10.1007/BF00637855.

[88] W. I. Axford. "*The Acceleration of Galactic Cosmic Rays*". In: *Symposium - International Astronomical Union* **94** (1981), pp. 339–358. DOI: 10.1017/S0074180900074866.

[89] R. D. Blandford and J. P. Ostriker. "*Particle acceleration by astrophysical shocks.*" In: *ApJL* **221** (Apr. 1978), pp. L29–L32. DOI: 10.1086/182658.

[90] J. A. Peacock. "*Fermi acceleration by relativistic shock waves*". In: *MNRAS* **196** (July 1981), pp. 135–152. DOI: 10.1093/mnras/196.2.135.

[91] M. Nagano and A. A. Watson. "*Observations and implications of the ultrahigh-energy cosmic rays*". In: *Rev. Mod. Phys.* **72** (3 July 2000), pp. 689–732. DOI: 10.1103/RevModPhys.72.689. URL: https://link.aps.org/doi/10.1103/RevModPhys.72.689.

[92] A. Vilenkin and E. P. S. Shellard. Cosmic Strings and Other Topological Defects. Cambridge University Press, 1994. ISBN: 978-0-521-65476-0.

[93] C. T. Hill. "*Monopolonium*". In: *Nuclear Physics B* **224**.3 (1983), pp. 469–490. ISSN: 0550-3213. DOI: https://doi.org/10.1016/0550-3213(83)90386-3. URL: https://www.sciencedirect.com/science/article/pii/0550321383903863.

[94] T. Weiler. "*Resonant Absorption of Cosmic-Ray Neutrinos by the Relic-Neutrino Background*". In: *Phys. Rev. Lett.* **49** (3 July 1982), pp. 234–237. DOI: 10.1103/PhysRevLett.49.234. URL: https://link.aps.org/doi/10.1103/PhysRevLett.49.234.

[95] M. V. Medvedev. "*Constraint on electromagnetic acceleration of highest energy cosmic rays*". In: *Physical Review E* **67**.4 (2003). ISSN: 1095-3787. DOI: 10.1103/physreve.67.045401. URL: http://dx.doi.org/10.1103/PhysRevE.67.045401.

[96] P. Gorham, C. Hebert, K. Liewer, C. Naudet, D. Saltzberg, and D. Williams. "*Experimental Limit on the Cosmic Diffuse Ultrahigh Energy Neutrino Flux*". In: *Physical review letters* **93** (July 2004), p. 041101. DOI: 10.1103/PhysRevLett.93.041101.

[97] G. B. Gelmini, O. E. Kalashev, and D. V. Semikoz. "*GZK photons as ultra-high-energy cosmic rays*". In: *Journal of Experimental and Theoretical Physics* **106** (2008), p. 1061. DOI: 10.1134/S106377610806006X. URL: https://doi.org/10.1134/S106377610806006X.

[98] J. Ellis, V. E. Mayes, and D. V. Nanopoulos. "*Ultrahigh-energy cosmic rays particle spectra from crypton decays*". In: *Phys. Rev. D* **74** (11 Dec. 2006), p. 115003. DOI: 10.1103/PhysRevD.74.115003. URL: https://link.aps.org/doi/10.1103/PhysRevD.74.115003.

[99] V. Berezinsky, M. Kachelrieß, and A. Vilenkin. "*Ultrahigh Energy Cosmic Rays without Greisen-Zatsepin-Kuzmin Cutoff*". In: *Phys. Rev. Lett.* **79** (22 Dec. 1997), pp. 4302–4305. DOI: 10.1103/PhysRevLett.79.4302. URL: https://link.aps.org/doi/10.1103/PhysRevLett.79.4302.

[100] V. A. Kuzmin and V. A. Rubakov. "*Ultrahigh-energy cosmic rays: A Window to postinflationary reheating epoch of the universe?*" In: *Phys. Atom. Nucl.* **61** (1998), p. 1028. arXiv: astro-ph/9709187.

[101] P. Blasi, R. Dick, and E. W. Kolb. "*Ultra-high energy cosmic rays from annihilation of superheavy dark matter*". In: *Astroparticle Physics* **18**.1 (2002), pp. 57–66. ISSN: 0927-6505. DOI: https://doi.org/10.1016/S0927-6505(02)00113-5. URL: https://www.sciencedirect.com/science/article/pii/S0927650502001135.

[102] M. Birkel and S. Sarkar. "*Extremely high energy cosmic rays from relic particle decays*". In: *Astroparticle Physics* **9**.4 (1998), pp. 297–309. ISSN: 0927-6505. DOI: https://doi.org/10.1016/S0927-6505(98)00028-0. URL: https://www.sciencedirect.com/science/article/pii/S0927650598000280.

[103] J. Abraham et al. "*Anisotropy studies around the galactic centre at EeV energies with the Auger Observatory*". In: *Astroparticle Physics* **27**.4 (2007), pp. 244–253. ISSN: 0927-6505. DOI: https://doi.org/10.1016/j.astropartphys.2006.11.002. URL: https://www.sciencedirect.com/science/article/pii/S0927650506001733.

[104] E. Waxman. "*Cosmological Gamma-Ray Bursts and the Highest Energy Cosmic Rays*". In: *Phys. Rev. Lett.* **75** (3 1995), pp. 386–389. DOI: 10.1103/PhysRevLett.75.386. URL: https://link.aps.org/doi/10.1103/PhysRevLett.75.386.

[105] D. F. Torres and L. A. Anchordoqui. "*Astrophysical origins of ultrahigh energy cosmic rays*". In: *Reports on Progress in Physics* **67**.9 (July 2004), pp. 1663–1730. DOI: 10.1088/0034-4885/67/9/r03. URL: https://doi.org/10.1088/0034-4885/67/9/r03.

[106] W. Baade and F. Zwicky. "*On Super-Novae*". In: *Proceedings of the National Academy of Sciences* **20**.5 (1934), pp. 254–259. ISSN: 0027-8424. DOI: 10.1073/pnas.20.5.254. eprint: https://www.pnas.org/content/20/5/254.full.pdf. URL: https://www.pnas.org/content/20/5/254.

[107] V. Ptuskin, V. Zirakashvili, and E.-S. Seo. "*SPECTRUM OF GALACTIC COSMIC RAYS ACCELERATED IN SUPERNOVA REMNANTS*". In: *The Astrophysical Journal* **718**.1 (June 2010), pp. 31–36. DOI: 10.1088/0004-637x/718/1/31. URL: https://doi.org/10.1088/0004-637x/718/1/31.

[108] J. W. Hewitt, F. Acero, T. J. Brandt, J. Cohen, F. de Palma, and F. Giordano. The 1st Fermi LAT SNR Catalog: Probing the Gamma-ray Population. 2013. arXiv: 1307.6570 [astro-ph.HE].

[109] M. Ackermann et al. "*Detection of the Characteristic Pion-Decay Signature in Supernova Remnants*". In: *Science* **339**.6121 (2013), pp. 807–811. ISSN: 0036-8075. DOI: 10.1126/science.1231160. eprint: https://science.sciencemag.org/content/339/6121/807.full.pdf. URL: https://science.sciencemag.org/content/339/6121/807.

[110] E. Striani, M. Tavani, et al. "*THE CRAB NEBULA SUPER-FLARE IN 2011 APRIL: EXTREMELY FAST PARTICLE ACCELERATION AND GAMMA-RAY EMISSION*". In: *The Astrophysical Journal* **741**.1 (Oct. 2011), p. L5. DOI: 10.1088/2041-8205/741/1/l5. URL: https://doi.org/10.1088/2041-8205/741/1/l5.

[111] M. Cardillo, E. Amato, and P. Blasi. "*Supernova remnant W44: a case of cosmic-ray reacceleration*". In: *AA* **595** (2016), A58. DOI: {10.1051/0004-6361/201628669}. URL: %7Bhttps://doi.org/10.1051/0004-6361/201628669%7D.

[112] M. Tavani, A. Bulgarelli, et al. "*Discovery of Powerful Gamma-Ray Flares from the Crab Nebula*". In: *Science* **331**.6018 (2011), pp. 736–739. ISSN: 0036-8075. DOI: 10.1126/science.1200083. eprint: https://science.sciencemag.org/content/331/6018/736.full.pdf. URL: https://science.sciencemag.org/content/331/6018/736.

[113] A. Giuliani, M. Tavani, et al. "*AGILE detection of GeV γ-ray emission from the SNR W28*". In: *AA* **516** (2010), p. L11. DOI: {10.1051/0004-6361/201014256}. URL: %7Bhttps://doi.org/10.1051/0004-6361/201014256%7D.

[114] accessed May 2021. URL: https://scitechdaily.com/stepping-inside-supernova-remnant-cassiopeia-a/.

[115] J. Abraham et al. "*Correlation of the highest energy cosmic rays with nearby extragalactic objects*". In: *Science* **318** (2007), pp. 938–943. DOI: 10.1126/science.1151124. arXiv: 0711.2256 [astro-ph].

[116] S. Mollerach and E. Roulet. "*Progress in high-energy cosmic ray physics*". In: *Progress in Particle and Nuclear Physics* **98** (Jan. 2018), pp. 85–118. ISSN: 0146-6410. DOI: 10.1016/j.ppnp.2017.10.002. URL: http://dx.doi.org/10.1016/j.ppnp.2017.10.002.

[117] P. Abreu et al. "*The ultra-high-energy cosmic-ray sky above 32 EeV viewed from the Pierre Auger Observatory*". In: *PoS* **ICRC2021** (2021), p. 307. DOI: 10.22323/1.395.0307.

[118] A. Aab et al. "*An Indication of Anisotropy in Arrival Directions of Ultra-high-energy Cosmic Rays through Comparison to the Flux Pattern of Extragalactic Gamma-Ray Sources*". In: *The Astrophysical Journal* **853**.2 (Feb. 2018), p. L29. ISSN: 2041-8213. DOI: 10.3847/2041-8213/aaa66d. URL: http://dx.doi.org/10.3847/2041-8213/aaa66d.

[119] C. Guépin, K. Kotera, E. Barausse, K. Fang, and K. Murase. "*Ultra-high-energy cosmic rays and neutrinos from tidal disruptions by massive black holes*". In: *A&A* **616** (2018), A179. DOI: 10.1051/0004-6361/201732392. URL: https://doi.org/10.1051/0004-6361/201732392.

[120] S. Inoue. "*Origin of ultra-high energy cosmic rays in the era of Auger and Telescope Array*". In: *Journal of Physics: Conference Series* **120**.6 (July 2008), p. 062001. DOI: 10.1088/1742-6596/120/6/062001. URL: https://doi.org/10.1088/1742-6596/120/6/062001.

[121] J. P. Rachen and P. L. Biermann. "*Extragalactic ultrahigh-energy cosmic rays. 1. Contribution from hot spots in FR-II radio galaxies*". In: *Astron. Astrophys.* **272** (1993), pp. 161–175. arXiv: astro-ph/9301010.

[122] G. Ghisellini. "*The jet/disk connection in blazars*". In: *AIP Conference Proceedings* **1242**.1 (2010), pp. 43–54. DOI: 10.1063/1.3460151. eprint: https://aip.scitation.org/doi/pdf/10.1063/1.3460151. URL: https://aip.scitation.org/doi/abs/10.1063/1.3460151.

[123] F. M. Rieger, V. Bosch-Ramon, and P. Duffy. "*Fermi acceleration in astrophysical jets*". In: *Astrophysics and Space Science* 309 (2007), pp. 119–125. DOI: 10.1007/s10509-007-9466-z. URL: https://doi.org/10.1007/s10509-007-9466-z.

[124] B. Fanaroff and J. Riley. "*The Morphology of Extragalactic Radio Sources of High and Low Luminosity*". In: *Monthly Notices of the Royal Astronomical Society* **167** (Apr. 1974), 31P–36P. DOI: 10.1093/mnras/167.1.31P.

[125] P. L. Biermann and P. A. Strittmatter. "*Synchrotron Emission from Shock Waves in Active Galactic Nuclei*". In: *ApJ* **322** (Nov. 1987), pp. 643–649. DOI: 10.1086/165759.

[126] A. Aab et al. "*SEARCHES FOR ANISOTROPIES IN THE ARRIVAL DIRECTIONS OF THE HIGHEST ENERGY COSMIC RAYS DETECTED BY THE PIERRE AUGER OBSERVATORY*". In: *The Astrophysical Journal* **804**.1 (Apr. 2015), p. 15. DOI: 10.1088/0004-637x/804/1/15. URL: https://doi.org/10.1088/0004-637x/804/1/15.

[127] L. Merten et al. "*Scrutinizing FR 0 radio galaxies as ultra-high-energy cosmic ray source candidates*". In: *Astropart. Phys.* **128** (2021), p. 102564. DOI: 10.1016/j.astropartphys.2021.102564. arXiv: 2102.01087 [astro-ph.HE].

[128] S. Guillot et al. "*NICER X-Ray Observations of Seven Nearby Rotation-powered Millisecond Pulsars*". In: *The Astrophysical Journal* **887**.1 (Dec. 2019), p. L27. DOI: 10.3847/2041-8213/ab511b. URL: https://doi.org/10.3847/2041-8213/ab511b.

[129] The Nobel Prize in Physics. URL: https://www.nobelprize.org/prizes/physics/1974/summary.

[130] W. Fong, T. Laskar, et al. "*The Broadband Counterpart of the Short GRB 200522A at z = 0.5536: A Luminous Kilonova or a Collimated Outflow with a Reverse Shock?*" In: *The Astrophysical Journal* **906**.2 (Jan. 2021), p. 127. DOI: 10.3847/1538-4357/abc74a. URL: https://doi.org/10.3847/1538-4357/abc74a.

[131] A. L. Piro and J. A. Kollmeier. "*ULTRAHIGH-ENERGY COSMIC RAYS FROM THE "EN CAUL" BIRTH OF MAGNETARS*". In: *The Astrophysical Journal* **826**.1 (July 2016), p. 97. DOI: 10.3847/0004-637x/826/1/97. URL: https://doi.org/10.3847/0004-637x/826/1/97.

[132] V. M. Kaspi and A. M. Beloborodov. "*Magnetars*". In: *Annual Review of Astronomy and Astrophysics* **55**.1 (2017), pp. 261–301. DOI: 10.1146/annurev-astro-081915-023329. URL: https://doi.org/10.1146/annurev-astro-081915-023329.

[133] K. Kotera and J. Silk. "*ULTRAHIGH-ENERGY COSMIC RAYS AND BLACK HOLE MERGERS*". In: *The Astrophysical Journal* **823**.2 (May 2016), p. L29. ISSN: 2041-8213. DOI: 10.3847/2041-8205/823/2/l29. URL: http://dx.doi.org/10.3847/2041-8205/823/2/L29.

[134] I. Bartos and M. Kowalski. Multimessenger Astronomy. 2399-2891. IOP Publishing, 2017. ISBN: 978-0-7503-1369-8. DOI: 10.1088/978-0-7503-1369-8. URL: http://dx.doi.org/10.1088/978-0-7503-1369-8.

[135] The Nobel Prize in Physics. URL: https://www.nobelprize.org/prizes/physics/1993/summary.

[136] M. Burgay, N. D'Amico, A. Possenti, et al. "*An increased estimate of the merger rate of double neutron stars from observations of a highly relativistic system*". In: *Nature* 426 (2003), pp. 531–533. URL: https://doi.org/10.1038/nature02124.

[137] B. P. Abbott et al. "*Multi-messenger Observations of a Binary Neutron Star Merger*". In: *The Astrophysical Journal* **848**.2 (Oct. 2017), p. L12. DOI: 10.3847/2041-8213/aa91c9. URL: https://doi.org/10.3847/2041-8213/aa91c9.

[138] M. Amenomori et al. "*First Detection of Photons with Energy beyond 100 TeV from an Astrophysical Source*". In: *Phys. Rev. Lett.* **123** (5 2019), p. 051101. DOI: 10.1103/PhysRevLett.123.051101. URL: https://link.aps.org/doi/10.1103/PhysRevLett.123.051101.

[139] B. Gendre et al. "*THE ULTRA-LONG GAMMA-RAY BURST 111209A: THE COLLAPSE OF A BLUE SUPERGIANT?*" In: *The Astrophysical Journal* **766**.1 (Mar. 2013), p. 30. DOI: 10.1088/0004-637x/766/1/30. URL: https://doi.org/10.1088/0004-637x/766/1/30.

[140] M. Vietri. "*On the acceleration of ultrahigh-energy cosmic rays in gamma-ray bursts*". In: *Astrophys. J.* **453** (1995), pp. 883–889. DOI: 10.1086/176448. arXiv: astro-ph/9506081.

[141] M. Palla, F. Matteucci, F. Calura, and F. Longo. The nature of GRB host galaxies from chemical abundances. 2019. arXiv: 1903.01353 [astro-ph.GA].

[142] Ph. Podsiadlowski, P. A. Mazzali, K. Nomoto, D. Lazzati, and E. Cappellaro. "*The Rates of Hypernovae and Gamma-Ray Bursts: Implications for Their Progenitors*". In: *The Astrophysical Journal* **607**.1 (Apr. 2004), pp. L17–L20. DOI: 10.1086/421347. URL: https://doi.org/10.1086/421347.

[143] E. Waxman. "*Gamma-ray bursts: Potential sources of ultra high energy cosmic-rays*". In: *Nuclear Physics B - Proceedings Supplements* **151**.1 (Jan. 2006), pp. 46–53. ISSN: 0920-5632. DOI: 10.1016/j.nuclphysbps.2005.07.008. URL: http://dx.doi.org/10.1016/j.nuclphysbps.2005.07.008.

[144] B. Paczynski. "*Gamma-ray bursters at cosmological distances*". In: *ApJL* **308** (Sept. 1986), pp. L43–L46. DOI: 10.1086/184740.

[145] M. J. Rees and P. Mészáros. "*Relativistic fireballs: energy conversion and time-scales*". In: *Monthly Notices of the Royal Astronomical Society* **258**.1 (Sept. 1992), 41P–43P. ISSN: 0035-8711. DOI: 10.1093/mnras/258.1.41P. eprint: https://academic.oup.com/mnras/article-pdf/258/1/41P/2897945/mnras258-041P.pdf. URL: https://doi.org/10.1093/mnras/258.1.41P.

[146] P. Mészáros. "*Gamma-ray bursts*". In: *Reports on Progress in Physics* **69**.8 (July 2006), pp. 2259–2321. ISSN: 1361-6633. DOI: 10.1088/0034-4885/69/8/r01. URL: http://dx.doi.org/10.1088/0034-4885/69/8/R01.

[147] D. Ryu, H. Kang, E. Hallman, and T. W. Jones. "*Cosmological Shock Waves and Their Role in the Large-Scale Structure of the Universe*". In: *The Astrophysical Journal* **593**.2 (Aug. 2003), pp. 599–610. ISSN: 1538-4357. DOI: 10.1086/376723. URL: http://dx.doi.org/10.1086/376723.

[148] C. A. Norman, D. B. Melrose, and A. Achterberg. "*The Origin of Cosmic Rays above 10 18.5 eV*". In: *ApJ* **454** (Nov. 1995), p. 60. DOI: 10.1086/176465.

[149] H. Kang, D. Ryu, and T. W. Jones. "*Cluster Accretion Shocks as Possible Acceleration Sites for Ultra–High-Energy Protons below the Greisen Cutoff*". In: *The Astrophysical Journal* **456** (Jan. 1996), p. 422. ISSN: 1538-4357. DOI: 10.1086/176666. URL: http://dx.doi.org/10.1086/176666.

[150] S. Inoue, G. Sigl, F. Miniati, and E. Armengaud. Ultrahigh energy cosmic rays as heavy nuclei from cluster accretion shocks. 2007. arXiv: 0711.1027 [astro-ph].

[151] S. Inoue, G. Sigl, F. Miniati, and E. Armengaud. "*Ultrahigh energy cosmic rays as heavy nuclei from cluster accretion shocks*". In: *arXiv e-prints*, astro-ph/0701167 (Jan. 2007), astro–ph/0701167. arXiv: astro-ph/0701167 [astro-ph].

[152] M. Lemoine, G. Sigl, A. V. Olinto, and D. N. Schramm. "*Ultra-high-Energy Cosmic-Ray Sources and Large-Scale Magnetic Fields*". In: *ApJ* **486**.2 (1997), p. L115. DOI: 10.1086/310847.

[153] P. Blasi and A. V. Olinto. "*Magnetized local supercluster and the origin of the highest energy cosmic rays*". In: *Phys. Rev. D* **59** (2 Dec. 1998), p. 023001. DOI: 10.1103/PhysRevD.59.023001. URL: https://link.aps.org/doi/10.1103/PhysRevD.59.023001.

[154] G. R. Farrar and M. S. Sutherland. "*Deflections of UHECRs in the Galactic magnetic field*". In: *JCAP* **2019**.5, 004 (May 2019), p. 004. DOI: 10.1088/1475-7516/2019/05/004. arXiv: 1711.02730 [astro-ph.HE].

[155] L. M. Widrow. "*Origin of galactic and extragalactic magnetic fields*". In: *Rev. Mod. Phys.* **74** (2002), pp. 775–823. DOI: 10.1103/RevModPhys.74.775. arXiv: astro-ph/0207240.

[156] A. A. Ruzmaikin, D. D. Sokoloff, and A. M. Shukurov. Magnetic fields of galaxies. Astrophysics and Space Science Library. Springer, 1988, p. 159. ISBN: 9027724504.

[157] P. P. Kronberg. "Extragalactic magnetic fields". In: *Rep. Prog. Phys.* 57.4 (1994), p. 325. DOI: 10.1088/0034-4885/57/4/001.

[158] N. E. Yanasak et al. "*Measurement of the Secondary Radionuclides ^{10}Be, ^{26}Al, ^{36}Cl, ^{54}Mn, and ^{14}C and Implications for the Galactic Cosmic-Ray Age*". In: *ApJ* **563**.2 (2001), pp. 768–792. DOI: 10.1086/323842.

[159] M. Simon and A. Molnar. "*A New Thought on the Energy Dependence of the $^{10}Be/^{9}Be$ Ratio*". In: *Proceedings, 28th International Cosmic Ray Conference (ICRC 2003): Tsukuba, Japan, July 31-August 7, 2003*. Frontiers Science Series **4** (2003), p. 1937.

[160] T. Vernstrom, G. Heald, F. Vazza, T. Galvin, J. West, N. Locatelli, N. Fornengo, and E. Pinetti. Discovery of Magnetic Fields Along Stacked Cosmic Filaments as Revealed by Radio and X-Ray Emission. 2021. arXiv: 2101.09331 [astro-ph.CO].

[161] R. Durrer and A. Neronov. "*Cosmological magnetic fields: their generation, evolution and observation*". In: *AAPR* **21**, 62 (June 2013), p. 62. DOI: 10.1007/s00159-013-0062-7. arXiv: 1303.7121 [astro-ph.CO].

[162] W. A. Hiltner. "*Polarization of Radiation from Distant Stars by the Interstellar Medium*". In: *Nature* **163**.4138 (Feb. 1949), p. 283. DOI: 10.1038/163283a0.

[163] W. A. Hiltner. "*Polarization of Light from Distant Stars by Interstellar Medium*". In: *Science* **109**.2825 (Feb. 1949), p. 165. DOI: 10.1126/science.109.2825.165.

[164] J. S. Hall. "*Observations of the Polarized Light from Stars*". In: *Science* **109**.2825 (Feb. 1949), pp. 166–167. DOI: 10.1126/science.109.2825.166.

[165] H. Alfvén. "*On the Solar Origin of Cosmic Radiation*". In: *Physical Review* **75**.11 (June 1949), pp. 1732–1735. DOI: 10.1103/PhysRev.75.1732.

[166] E. Fermi. "*On the Origin of the Cosmic Radiation*". In: *Physical Review* **75**.8 (Apr. 1949), pp. 1169–1174. DOI: 10.1103/PhysRev.75.1169.

[167] L. Jr. Davis and J. L. Greenstein. "*The Polarization of Starlight by Aligned Dust Grains.*" In: *ApJ* **114** (Sept. 1951), p. 206. DOI: 10.1086/145464.

[168] T. R. Jaffe. "*Practical Modeling of Large-Scale Galactic Magnetic Fields: Status and Prospects*". In: *Galaxies* **7**.2 (2019). ISSN: 2075-4434. DOI: 10.3390/galaxies7020052. URL: https://www.mdpi.com/2075-4434/7/2/52.

[169] P. P. Kronberg. "*Extragalactic magnetic fields*". In: *Rept. Prog. Phys.* **57** (1994), pp. 325–382. DOI: {10.1088/0034-4885/57/4/001}.

[170] T. Stanev. "*Ultra–High-Energy Cosmic Rays and the Large-Scale Structure of the Galactic Magnetic Field*". In: *The Astrophysical Journal* **479**.1 (Apr. 1997), pp. 290–295. ISSN: 1538-4357. DOI: 10.1086/303866. URL: http://dx.doi.org/10.1086/303866.

[171] G. E. Hale. "*On the Probable Existence of a Magnetic Field in Sun-Spots*". In: *ApJ* **28** (Nov. 1908), p. 315. DOI: 10.1086/141602.

[172] H. Bethe and W. Heitler. "*On the stopping of fast particles and on the creation of positive electrons*". In: *Proc. R. Soc. Lond. A* **146** (1934), pp. 83–112. DOI: 10.1098/rspa.1934.0140. URL: http://doi.org/10.1098/rspa.1934.0140.

[173] C. Heiter, D. Kuempel, D. Walz, and M. Erdmann. "*Production and propagation of ultra-high energy photons using CRPropa 3*". In: *Astroparticle Physics* **102** (2018), pp. 39–50. ISSN: 0927-6505. DOI: https://doi.org/10.1016/j.astropartphys.2018.05.003. URL: https://www.sciencedirect.com/science/article/pii/S0927650518300069.

[174] K. A. Olive. "*Review of Particle Physics*". In: *Chinese Physics C* **38**.9 (2014), p. 090001. DOI: 10.1088/1674-1137/38/9/090001. URL: https://doi.org/10.1088%2F1674-1137%2F38%2F9%2F090001.

[175] C. Patrignani et al. "*Review of Particle Physics*". In: *Chin. Phys. C* **40**.10 (2016), p. 100001. DOI: 10.1088/1674-1137/40/10/100001.

[176] D. Harari, S. Mollerach, and E. Roulet. "*On the ultrahigh energy cosmic ray horizon*". In: *JCAP* **2006**.11, 012 (Nov. 2006), p. 012. DOI: 10.1088/1475-7516/2006/11/012. arXiv: astro-ph/0609294 [astro-ph].

[177] R. Aloisio. "*Acceleration and propagation of ultra high energy cosmic rays*". In: *PTEP* **2017**.12 (2017), 12A102. DOI: 10.1093/ptep/ptx115. arXiv: 1707.08471 [astro-ph.HE].

[178] P. Tinyakov. "*Latest results from the telescope array*". In: *Nuclear Instruments and Methods in Physics Research A* **742** (Apr. 2014), pp. 29–34. DOI: 10.1016/j.nima.2013.10.067.

[179] B. R. Martin and G. Shaw. Particle Physics. Third edition. John Wiley and Sons, Ltd, 2008. ISBN: 978-0-470-03293-0.

[180] D. Allard. "*Extragalactic propagation of ultrahigh energy cosmic-rays*". In: *Astroparticle Physics* **39-40** (2012). Cosmic Rays Topical Issue, pp. 33–43. ISSN: 0927-6505. DOI: https://doi.org/10.1016/j.astropartphys.2011.10.011. URL: https://www.sciencedirect.com/science/article/pii/S092765051100199X.

[181] J. L. Puget, F. W. Stecker, and J. H. Bredekamp. "*Photonuclear interactions of ultrahigh energy cosmic rays and their astrophysical consequences.*" In: *ApJ* **205** (Apr. 1976), pp. 638–654. DOI: 10.1086/154321.

[182] L. Cazon. "*Working Group Report on the Combined Analysis of Muon Density Measurements from Eight Air Shower Experiments*". In: *PoS* **ICRC2019** (2020), p. 214. DOI: 10.22323/1.358.0214. arXiv: 2001.07508 (astro-ph.HE).

[183] S. Ostapchenko. "*Monte Carlo treatment of hadronic interactions in enhanced Pomeron scheme: QGSJET-II model*". In: *Physical Review D* **83**.1 (Jan. 2011). ISSN: 1550-2368. DOI: 10.1103/physrevd.83.014018. URL: http://dx.doi.org/10.1103/PhysRevD.83.014018.

[184] T. Pierog, I. Karpenko, J. M. Katzy, E. Yatsenko, and K. Werner. "*EPOS LHC: Test of collective hadronization with data measured at the CERN Large Hadron Collider*". In: *Physical Review C* **92**.3 (Sept. 2015). ISSN: 1089-490X. DOI: 10.1103/physrevc.92.034906. URL: http://dx.doi.org/10.1103/PhysRevC.92.034906.

[185] F. Riehn, R. Engel, A. Fedynitch, T. K. Gaisser, and T. Stanev. "*A new version of the event generator Sibyll*". In: *PoS* **ICRC2015** (2016), p. 558. DOI: 10.22323/1.236.0558. arXiv: 1510.00568 [hep-ph].

[186] T. Pierog. "*Air Shower Simulation with a New Generation of post-LHC Hadronic Interaction Models in CORSIKA*". In: *PoS* **ICRC2017** (2017), p. 1100. DOI: 10.22323/1.301.1100.

[187] M. Amenomori et al. "*Anisotropy and Corotation of Galactic Cosmic Rays*". In: *Science* **314**.5798 (2006), 439–443. ISSN: 0036-8075. DOI: {10.1126/science.1131702}. eprint: {https://science.sciencemag.org/content/314/5798/439.full.pdf}. URL: %7Bhttps://science.sciencemag.org/content/314/5798/439%7D.

[188] M. Amenomori et al. "*New estimation of the spectral index of high-energy cosmic rays as determined by the Compton-Getting anisotropy*". In: *Astrophys. J. Lett.* **672** (2008), p. L53. DOI: 10.1086/525526. arXiv: 0711.2002 (astro-ph).

[189] R. Abbasi, IceCube Collaboration, et al. "*Observation of Anisotropy in the Galactic Cosmic-Ray Arrival Directions at 400 TeV with IceCube*". In: *ApJ* **746**.1, 33 (Feb. 2012), p. 33. DOI: 10.1088/0004-637X/746/1/33. arXiv: 1109.1017 [hep-ex].

[190] M. G. Aartsen, IceCube Collaboration, et al. "*Anisotropy in Cosmic-Ray Arrival Directions in the Southern Hemisphere Based on Six Years of Data from the IceCube Detector*". In: *ApJ* **826**.2, 220 (Aug. 2016), p. 220. DOI: 10.3847/0004-637X/826/2/220. arXiv: 1603.01227 [astro-ph.HE].

[191] A. Chiavassa et al. "*KASCADE-Grande experiment measurements of the cosmic ray spectrum and large scale anisotropy*". In: *Nuclear and Particle Physics Proceedings* **279-281** (2016). Proceedings of the 9th Cosmic Ray International Seminar, pp. 56–62. ISSN: 2405-6014. DOI: https://doi.org/10.1016/j.nuclphysbps.2016.10.009. URL: https://www.sciencedirect.com/science/article/pii/S2405601416301900.

[192] V. S. Ptuskin, I. V. Moskalenko, F. C. Jones, A. W. Strong, and V. N. Zirakashvili. "*Dissipation of Magnetohydrodynamic Waves on Energetic Particles: Impact on Interstellar Turbulence and Cosmic-Ray Transport*". In: *ApJ* **642**.2 (May 2006), pp. 902–916. DOI: 10.1086/501117. arXiv: astro-ph/0510335 [astro-ph].

[193] R. Aloisio and V. Berezinsky. "*Diffusive propagation of UHECR and the propagation theorem*". In: *Astrophys. J.* **612** (2004), 900–913. DOI: 10.1086/421869. arXiv: astro-ph/0403095.

[194] O. Deligny. "*Measurements and implications of cosmic ray anisotropies from TeV to trans-EeV energies*". In: *Astroparticle Physics* **104** (Jan. 2019), pp. 13–41. ISSN: 0927-6505. DOI: 10.1016/j.astropartphys.2018.08.005. URL: http://dx.doi.org/10.1016/j.astropartphys.2018.08.005.

[195] E. P. George. "*Cosmic rays measure overburden of tunnel*". In: *Commonwealth Engineer* **455** (1955).

[196] C. L. Morris, C. C. Alexander, et al. "*Tomographic Imaging with Cosmic Ray Muons*". In: *Science & Global Security* **16**.1-2 (2008), pp. 37–53. DOI: 10.1080/08929880802335758.

[197] F. Riggi, M. Bandieramonte, U. Becciani, et al. "*Multiparametric approach to the assessment of muon tomographic results for the inspection of a full-scale container*". In: *Eur. Phys. J. Plus* **136** (2021), p. 139. DOI: 10.1140/epjp/s13360-020-00970-z.

[198] K. Jourde, D. Gibert, J. Marteau, J. de Bremond d'Ars, and J.-C. Komorowski. "*Muon dynamic radiography of density changes induced by hydrothermal activity at the La Soufrière of Guadeloupe volcano*". In: (2016). arXiv: 1606.03003 [physics.geo-ph].

[199] S. Kedar, H. K. M. Tanaka, C. J. Naudet, C. E. Jones, J. P. Plaut, and F. H. Webb. "*Muon radiography for exploration of Mars geology*". In: *Geoscientific Instrumentation, Methods and Data Systems* **2**.1 (2013), pp. 157–164. DOI: 10.5194/gi-2-157-2013. URL: https://gi.copernicus.org/articles/2/157/2013/.

[200] R. Stone. "*Fukushima Cleanup Will Be Drawn Out and Costly*". In: *Science* **331**.6024 (2011), pp. 1507–1507. ISSN: 0036-8075. DOI: 10.1126/science.331.6024.1507. eprint: https://science.sciencemag.org/content/331/6024/1507.full.pdf. URL: https://science.sciencemag.org/content/331/6024/1507.

[201] P. C. Burns, R. C. Ewing, and A. Navrotsky. "*Nuclear Fuel in a Reactor Accident*". In: *Science* **335**.6073 (2012), pp. 1184–1188. ISSN: 0036-8075. DOI: 10.1126/science.1211285. eprint: https://science.sciencemag.org/content/335/6073/1184.full.pdf. URL: https://science.sciencemag.org/content/335/6073/1184.

[202] R. Delunel et al. "*Late-Pleistocene catchment-wide denudation patterns across the European Alps*". In: *Earth-Science Reviews* **211** (2020), p. 103407. ISSN: 0012-8252. DOI: https://doi.org/10.1016/j.earscirev.2020.103407. URL: http://www.sciencedirect.com/science/article/pii/S0012825220304530.

[203] K. Morishima, M. Kuno, A. Nishio, et al. "*Discovery of a big void in Khufu's Pyramid by observation of cosmic-ray muons*". In: *Nature* **552** (2017), pp. 386–390. DOI: 10.1038/nature24647.

[204] J. Rautenberg. "*Limits on ultra-high energy photons with the Pierre Auger Observatory*". In: *PoS* **ICRC2019** (2019), p. 398. DOI: 10.22323/1.358.0398.

[205] Aab, A. and others. "*A search for point sources of EeV photons*". In: *Astrophys. J.* **789**.2 (2014), p. 160. DOI: 10.1088/0004-637X/789/2/160. arXiv: 1406.2912 [astro-ph.HE].

[206] A. Aab, P. Abreu, M. Aglietta, I. Al Samarai, I. F. M. Albuquerque, I. Allekotte, A. Almela, J. Castillo, J. Alvarez-Muñiz, G. A. Anastasi, et al. "*A Targeted Search for Point Sources of EeV Photons with the Pierre Auger Observatory*". In: *The Astrophysical Journal* **837**.2 (Mar. 2017), p. L25. ISSN: 2041-8213. DOI: 10.3847/2041-8213/aa61a5. URL: http://dx.doi.org/10.3847/2041-8213/aa61a5.

[207] M. Risse and P. Homola. "*Search for ultra-high energy photons using air showers*". In: *Modern Physics Letters A* **22**.11 (Apr. 2007), pp. 749–766. ISSN: 1793-6632. URL: http://dx.doi.org/10.1142/S0217732307022864.

[208] F. W. Stecker and M. H. Salamon. "*Photodisintegration of ultrahigh-energy cosmic rays: A New determination*". In: *Astrophys. J.* **512** (1999), pp. 521–526. DOI: 10.1086/306816. arXiv: astro-ph/9808110.

[209] R. Aloisio, V. Berezinsky, and M. Kachelriess. "*Fragmentation functions in supersymmetric QCD and ultrahigh energy cosmic ray spectra produced in top-down models*". In: *Phys. Rev. D* **69** (9 May 2004), p. 094023. DOI: 10.1103/PhysRevD.69.094023. URL: https://link.aps.org/doi/10.1103/PhysRevD.69.094023.

[210] P. Bhattacharjee and G. Sigl. "*Origin and propagation of extremely high-energy cosmic rays*". In: *Phys. Rept.* **327** (2000), pp. 109–247. DOI: 10.1016/S0370-1573(99)00101-5. arXiv: astro-ph/9811011.

[211] T. Stanev. High Energy Cosmic Rays. second edition. Springer-Verlag Berlin Heidelberg, 2010. ISBN: 978-3-540-85148-6. DOI: 10.1007/978-3-540-85148-6. URL: https://doi.org/10.1007/978-3-540-85148-6.

[212] V. A. Kostelecký. "*Gravity, Lorentz violation, and the standard model*". In: *Physical Review D* **69**.10 (May 2004). ISSN: 1550-2368. DOI: 10.1103/physrevd.69.105009. URL: http://dx.doi.org/10.1103/PhysRevD.69.105009.

[213] D. Mattingly. "*Modern tests of Lorentz invariance*". In: *Living Rev. Rel.* **8** (2005), p. 5. DOI: 10.12942/lrr-2005-5. arXiv: gr-qc/0502097.

[214] S. Liberati and L. Maccione. "*Lorentz Violation: Motivation and New Constraints*". In: *Annual Review of Nuclear and Particle Science* **59**.1 (Nov. 2009), pp. 245–267. ISSN: 1545-4134. DOI: 10.1146/annurev.nucl.010909.083640. URL: http://dx.doi.org/10.1146/annurev.nucl.010909.083640.

[215] S. R. Coleman and S. L. Glashow. "*High-energy tests of Lorentz invariance*". In: *Phys. Rev. D* **59** (1999), p. 116008. DOI: 10.1103/PhysRevD.59.116008. arXiv: hep-ph/9812418.

[216] L. Maccione and S. Liberati. "*GZK photon constraints on Planck scale Lorentz violation in QED*". In: *JCAP* **08** (2008), p. 027. DOI: 10.1088/1475-7516/2008/08/027. arXiv: 0805.2548 [astro-ph].

[217] S. R. Coleman and S. L. Glashow. "*Evading the GZK cosmic ray cutoff*". In: (Aug. 1998). arXiv: hep-ph/9808446.

[218] S. Chadha and H. B. Nielsen. "*Lorentz invariance as a low energy phenomenon*". In: *Nuclear Physics B* **217**.1 (1983), pp. 125–144. ISSN: 0550-3213. DOI: https://doi.org/10.1016/0550-3213(83)90081-0. URL: https://www.sciencedirect.com/science/article/pii/0550321383900810.

[219] F. R. Klinkhamer, M. Niechciol, and M. Risse. "*Improved bound on isotropic Lorentz violation in the photon sector from extensive air showers*". In: *Physical Review D* **96**.11 (Dec. 2017). ISSN: 2470-0029. DOI: 10.1103/physrevd.96.116011. URL: http://dx.doi.org/10.1103/PhysRevD.96.116011.

[220] J. S. Díaz, F. R. Klinkhamer, and M. Risse. "*Changes in extensive air showers from isotropic Lorentz violation in the photon sector*". In: *Physical Review D* **94**.8 (Oct. 2016). ISSN: 2470-0029. DOI: 10.1103/physrevd.94.085025. URL: http://dx.doi.org/10.1103/PhysRevD.94.085025.

[221] T. C. Weekes, M. F. Cawley, D. J. Fegan, et al. "*Observation of TeV Gamma Rays from the Crab Nebula Using the Atmospheric Cerenkov Imaging Technique*". In: *ApJ* **342** (jul 1989), p. 379. DOI: 10.1086/167599.

[222] URL: https://fermi.gsfc.nasa.gov.

[223] F. Longo. "*On overview on Gamma Ray Astrophysics in the Fermi era*". In: *Frascati Phys. Ser.* **66** (2018). Ed. by R. Fusco Femiano, G. Mannocchi, A. Morselli, and G. C. Trinchero, pp. 117–131. URL: https://www.lnf.infn.it/sis/frascatiseries/Volume66/Volume66.pdf.

[224] URL: https://www.mpi-hd.mpg.de/hfm/HESS/HESS.shtml.

[225] URL: https://wwwmagic.mpp.mpg.de/.

[226] URL: https://veritas.sao.arizona.edu/.

[227] URL: https://www.hawc-observatory.org.

[228] K. Hirata et al. "*Observation of a neutrino burst from the supernova SN1987A*". In: *Phys. Rev. Lett.* **58** (14 Apr. 1987), pp. 1490–1493. DOI: 10.1103/PhysRevLett.58.1490. URL: https://link.aps.org/doi/10.1103/PhysRevLett.58.1490.

[229] R. M. Bionta et al. "*Observation of a neutrino burst in coincidence with supernova 1987A in the Large Magellanic Cloud*". In: *Phys. Rev. Lett.* **58** (14 Apr. 1987), pp. 1494–1496. DOI: 10.1103/PhysRevLett.58.1494. URL: https://link.aps.org/doi/10.1103/PhysRevLett.58.1494.

[230] URL: https://icecube.wisc.edu.

[231] M. G. Aartsen et al. "*Evidence for High-Energy Extraterrestrial Neutrinos at the IceCube Detector*". In: *Science* **342** (2013), p. 1242856. DOI: 10.1126/science.1242856. arXiv: 1311.5238 [astro-ph.HE].

[232] The IceCube Collaboration et al. "*Multimessenger observations of a flaring blazar coincident with high-energy neutrino IceCube-170922A*". In: *Science* **361**.6398 (2018). ISSN: 0036-8075. DOI: 10.1126/science.aat1378. eprint: https://science.sciencemag.org/content/361/6398/eaat1378.full.pdf. URL: https://science.sciencemag.org/content/361/6398/eaat1378.

[233] IceCube Collaboration et al. "*Neutrino emission from the direction of the blazar TXS 0506+056 prior to the IceCube-170922A alert*". In: *Science* **361**.6398 (2018), pp. 147–151. ISSN: 0036-8075. DOI: 10.1126/science.aat2890. eprint: https://science.sciencemag.org/content/361/6398/147.full.pdf. URL: https://science.sciencemag.org/content/361/6398/147.

[234] H. Poincaré. "Sur la dynamique de l'électron". In: *C.R.T.* **140** (1905), pp. 1504–1508. URL: https://www.academie-sciences.fr/pdf/dossiers/Poincare/Poincare_pdf/Poincare_CR1905.pdf.

[235] A. Einstein. "*Näherungsweise Integration der Feldgleichungen der Gravitation*". In: *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin* (Jan. 1916), pp. 688–696. DOI: 10.1002/3527608958.ch7.

[236] A. Einstein. "*Über Gravitationswellen*". In: *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin* (Jan. 1918), pp. 154–167. DOI: 10.1002/3527608958.ch12.

[237] B. P. Abbott et al. "*Observation of Gravitational Waves from a Binary Black Hole Merger*". In: *Phys. Rev. Lett.* **116** (6 Feb. 2016), p. 061102. DOI: 10.1103/PhysRevLett.116.061102. URL: https://link.aps.org/doi/10.1103/PhysRevLett.116.061102.

[238] A. Aab et al. "*Limits on point-like sources of ultra-high-energy neutrinos with the Pierre Auger Observatory*". In: *Journal of Cosmology and Astroparticle Physics* **2019**.11 (Nov. 2019), pp. 004–004. DOI: 10.1088/1475-7516/2019/11/004. URL: https://doi.org/10.1088/1475-7516/2019/11/004.

[239] A. Aab et al. "*Probing the origin of ultra-high-energy cosmic rays with neutrinos in the EeV energy range using the Pierre Auger Observatory*". In: *Journal of Cosmology and Astroparticle Physics* **2019**.10 (Oct. 2019), pp. 022–022. DOI: 10.1088/1475-7516/2019/10/022. URL: https://doi.org/10.1088/1475-7516/2019/10/022.

[240] T. K. Gaisser, F. Halzen, and T. Stanev. "*Particle astrophysics with high energy neutrinos*". In: *Physics Reports* **258**.3 (1995), pp. 173–236. ISSN: 0370-1573. DOI: https://doi.org/10.1016/0370-1573(95)00003-Y. URL: https://www.sciencedirect.com/science/article/pii/037015739500003Y.

[241] M. G. Aartsen et al. "*The IceCube Neutrino Observatory: Instrumentation and Online Systems*". In: *JINST* **12**.03 (2017), P03012. DOI: 10.1088/1748-0221/12/03/P03012. arXiv: 1612.05093 [astro-ph.IM].

[242] M. Ageron, J. A. Aguilar, I. Al Samarai, A. Albert, F. Ameli, M. André, M. Anghinolfi, G. Anton, S. Anvar, M. Ardid, and et al. "*ANTARES: The first undersea neutrino telescope*". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **656**.1 (Nov. 2011), pp. 11–38. ISSN: 0168-9002. DOI: 10.1016/j.nima.2011.06.103. URL: http://dx.doi.org/10.1016/j.nima.2011.06.103.

[243] S. Aiello, S. E. Akrame, F. Ameli, E. G. Anassontzis, M. Andre, G. Androulakis, M. Anghinolfi, G. Anton, M. Ardid, J. Aublin, et al. "*Sensitivity of the KM3NeT/ARCA neutrino telescope to point-like neutrino sources*". In: *Astroparticle Physics* **111** (Sept. 2019), pp. 100–110. ISSN: 0927-6505. DOI: 10.1016/j.astropartphys.2019.04.002. URL: http://dx.doi.org/10.1016/j.astropartphys.2019.04.002.

[244] A. Aab et al. "*Improved limit to the diffuse flux of ultrahigh energy neutrinos from the Pierre Auger Observatory*". In: *Phys. Rev. D* **91** (9 May 2015), p. 092008. DOI: 10.1103/PhysRevD.91.092008. URL: https://link.aps.org/doi/10.1103/PhysRevD.91.092008.

[245] P. Abreu et al. "*Search for ultrahigh energy neutrinos in highly inclined events at the Pierre Auger Observatory*". In: *Phys. Rev. D* **84** (12 Dec. 2011), p. 122005. DOI: 10.1103/PhysRevD.84.122005. URL: https://link.aps.org/doi/10.1103/PhysRevD.84.122005.

[246] K.-H. Kampert, M. Alejandro Mostafa, E. Zas, The Pierre Auger Collaboration, et al. "*Multi-Messenger Physics With the Pierre Auger Observatory*". In: *Frontiers in Astronomy and Space Sciences* **6** (2019), p. 24. ISSN: 2296-987X. DOI: 10.3389/fspas.2019.00024. URL: https://www.frontiersin.org/article/10.3389/fspas.2019.00024.

[247] L. A. Anchordoqui, H. Goldberg, and T. J. Weiler. "*Auger Test of the Cen A Model of Highest Energy Cosmic Rays*". In: *Phys. Rev. Lett.* **87** (8 July 2001), p. 081101. DOI: 10.1103/PhysRevLett.87.081101. URL: https://link.aps.org/doi/10.1103/PhysRevLett.87.081101.

[248] A. M. Taylor, J. A. Hinton, P. Blasi, and M. Ave. "*Identifying Nearby Accelerators of Ultrahigh Energy Cosmic Rays Using Ultrahigh Energy (and Very High Energy) Photons*". In: *Phys. Rev. Lett.* **103** (5 July 2009), p. 051102. DOI: 10.1103/PhysRevLett.103.051102. URL: https://link.aps.org/doi/10.1103/PhysRevLett.103.051102.

[249] P. Abreu, M. Aglietta, M. Ahlers, E. J. Ahn, I. F. M. Albuquerque, D. Allard, I. Allekotte, J. Allen, P. Allison, et al. "*A SEARCH FOR POINT SOURCES OF EeV NEUTRONS*". In: *The Astrophysical Journal* **760**.2 (Nov. 2012), p. 148. ISSN: 1538-4357. DOI: 10.1088/0004-637x/760/2/148. URL: http://dx.doi.org/10.1088/0004-637X/760/2/148.

[250] Pierre Auger Collaboration, P. Abreu, et al. "*A Search for Point Sources of EeV Neutrons*". In: *ApJ* **760**.2, 148 (Dec. 2012), p. 148. DOI: 10.1088/0004-637X/760/2/148. arXiv: 1211.4901 [astro-ph.HE].

[251] O. Adriani et al. "*A statistical procedure for the identification of positrons in the PAMELA experiment*". In: *Astroparticle Physics* **34**.1 (Aug. 2010), pp. 1–11. DOI: 10.1016/j.astropartphys.2010.04.007. arXiv: 1001.3522 [astro-ph.HE].

[252] M. Ackermann, M. Ajello, A. Allafort, W. B. Atwood, L. Baldini, G. Barbiellini, D. Bastieri, K. Bechtol, R. Bellazzini, B. Berenji, et al. "*Measurement of Separate Cosmic-Ray Electron and Positron Spectra with the Fermi Large Area Telescope*". In: *Physical Review Letters* **108**.1 (Jan. 2012). ISSN: 1079-7114. DOI: 10.1103/physrevlett.108.011103. URL: http://dx.doi.org/10.1103/PhysRevLett.108.011103.

[253] K. Blum, B. Katz, and E. Waxman. "*AMS-02 Results Support the Secondary Origin of Cosmic Ray Positrons*". In: *Phys. Rev. Lett.* **111** (21 Nov. 2013), p. 211101. DOI: 10.1103/PhysRevLett.111.211101. URL: https://link.aps.org/doi/10.1103/PhysRevLett.111.211101.

[254] I. Andreoni and J. Cooke. "*The Deeper Wider Faster Programme: Chasing the Fastest Bursts in the Universe*". In: *Proceedings of the International Astronomical Union* **14**.S339 (Nov. 2017), pp. 135–138. ISSN: 1743-9221. DOI: 10.1017/s1743921318002399. URL: http://dx.doi.org/10.1017/S1743921318002399.

[255] URL: http://www.amon.psu.edu/.

[256] M. W. E. Smith, D. B. Fox, D. F. Cowen, P. Mészáros, G. Tešić, J. Fixelle, I. Bartos, P. Sommers, A. Ashtekar, G. Jogesh Babu, et al. "*The Astrophysical Multimessenger Observatory Network (AMON)*". In: *Astroparticle Physics* **45** (May 2013), pp. 56–70. ISSN: 0927-6505. DOI: 10.1016/j.astropartphys.2013.03.003. URL: http://dx.doi.org/10.1016/j.astropartphys.2013.03.003.

[257] H. A. Ayala Solares. "*AMON Multimessenger Alerts: Past and Future*". In: *Galaxies* **7**.1 (2019). ISSN: 2075-4434. DOI: 10.3390/galaxies7010019. URL: https://www.mdpi.com/2075-4434/7/1/19.

[258] H. A. Ayala Solares, S. Coutu, et al. "*The Astrophysical Multimessenger Observatory Network (AMON): Performance and science program*". In: *Astroparticle Physics* **114** (2020), pp. 68–76. ISSN: 0927-6505. DOI: https://doi.org/10.1016/j.astropartphys.2019.06.007. URL: https://www.sciencedirect.com/science/article/pii/S0927650519301227.

[259] S. D. Barthelmy, P. Butterworth, T. L. Cline, N. Gehrels, G. J. Fishman, C. Kouveliotou, and C. A. Meegan. "*BACODINE, the Real-Time BATSE Gamma-Ray Burst Coordinates Distribution Network*". In: *APSS* **231**.1–2 (Sept. 1995), pp. 235–238. DOI: 10.1007/BF00658623.

[260] S. D. Barthelmy, P. Butterworth, T. L. Cline, N. Gehrels, F. Marshall, T. Takeshima, V. Connaughton, R. M. Kippen, C. Kouveliotou, and C. R. Robinson. "*The GRB coordinates network (GCN): A status report*". In: *AIP Conference Proceedings* **428**.1 (1998), pp. 99–103. DOI: 10.1063/1.55426. eprint: https://aip.scitation.org/doi/pdf/10.1063/1.55426. URL: https://aip.scitation.org/doi/abs/10.1063/1.55426.

[261] A. Keivani, H. Ayala, and J. DeLaunay. "*Astrophysical Multimessenger Observatory Network (AMON): Science, Infrastructure, and Status*". In: *PoS* **ICRC2017** (2017), p. 629. DOI: 10.22323/1.301.0629.

[262] P. W. Gorham. "*On the possibility of radar echo detection of ultra-high energy cosmic ray- and neutrino-induced extensive air showers*". In: *Astroparticle Physics* **15**.2 (Apr. 2001), pp. 177–202. ISSN: 0927-6505. DOI: 10.1016/s0927-6505(00)00143-2. URL: http://dx.doi.org/10.1016/S0927-6505(00)00143-2.

[263] P. W. Gorham, N. G. Lehtinen, G. S. Varner, J. J. Beatty, A. Connolly, P. Chen, M. E. Conde, W. Gai, C. Hast, C. L. Hebert, et al. "*Observations of microwave continuum emission from air shower plasmas*". In: *Physical Review D* **78**.3 (Aug. 2008). ISSN: 1550-2368. DOI: 10.1103/physrevd.78.032007. URL: http://dx.doi.org/10.1103/PhysRevD.78.032007.

[264] G. A. Askaryan. "Hydrodynamic radiation from the tracks of ionizing particles in stable liquids". In: *The Soviet Journal of Atomic Energy* 3 (2 1957), pp. 921–923. DOI: 10.1007/BF01480076. URL: https://doi.org/10.1007/BF01480076.

[265] G. A. Askariyan, B.A. Dolgoshein, A.N. Kalinovsky, and N.V. Mokhov. "*Acoustic detection of high energy particle showers in water*". In: *Nuclear Instruments and Methods* **164**.2 (1979), pp. 267–278. ISSN: 0029-554X. DOI: https://doi.org/10.1016/0029-554X(79)90244-1. URL: https://www.sciencedirect.com/science/article/pii/0029554X79902441.

[266] F. D. Kahn and I. Lerche. "Radiation from Cosmic Ray Air Showers". In: *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences* **289**.1417 (1966), pp. 206–213. ISSN: 00804630. DOI: 10.2307/2415075. URL: http://www.jstor.org/stable/2415075.

[267] A. Haungs, D. Kang, S. Schoo, D. Wochele, J. Wochele, W. D. Apel, J. C. Arteaga-Velázquez, K. Bekk, M. Bertaina, J. Blümer, et al. "*The KASCADE Cosmic-ray Data Centre KCDC: granting open access to astroparticle physics research data*". In: *The European Physical Journal C* **78**.9 (Sept. 2018). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-018-6221-2. URL: http://dx.doi.org/10.1140/epjc/s10052-018-6221-2.

[268] W. Heitler. The quantum theory of radiation. Vol. **138**. International Series of Monographs on Physics. Oxford: Oxford University Press, 1936, pp. 483–484. DOI: 10.1038/138483a0.

[269] J. Matthews. "*A Heitler model of extensive air showers*". In: *Astroparticle Physics* **22**.5 (2005), pp. 387–397. ISSN: 0927-6505. DOI: https://doi.org/10.1016/j.astropartphys.2004.09.003. URL: https://www.sciencedirect.com/science/article/pii/S0927650504001598.

[270] K.-H. Kampert and M. Unger. "*Measurements of the cosmic ray composition with air shower experiments*". In: *Astroparticle Physics* **35**.10 (2012), pp. 660–678. ISSN: 0927-6505. DOI: https://doi.org/10.1016/j.astropartphys.2012.02.004. URL: https://www.sciencedirect.com/science/article/pii/S0927650512000382.

[271] J. R. Horandel. "*On total inelastic cross-sections and the average depth of the maximum of extensive air showers*". In: *J. Phys. G* **29** (2003), pp. 2439–2464. DOI: 10.1088/0954-3899/29/11/001. arXiv: astro-ph/0309010.

[272] J. R. Horandel. "*Cosmic rays from the knee to the second knee: 10**4 to 10**18-eV*". In: *Mod. Phys. Lett. A* **22** (2007), pp. 1533–1552. DOI: 10.1142/S0217732307024139. arXiv: astro-ph/0611387.

[273] R. Ulrich, R. Engel, and M. Unger. "*Hadronic multiparticle production at ultrahigh energies and extensive air showers*". In: *Phys. Rev. D* **83** (5 Mar. 2011), p. 054026. DOI: 10.1103/PhysRevD.83.054026. URL: https://link.aps.org/doi/10.1103/PhysRevD.83.054026.

[274] R. Engel, D. Heck, and T. Pierog. "*Extensive Air Showers and Hadronic Interactions at High Energy*". In: *Annual Review of Nuclear and Particle Science* **61**.1 (Nov. 2011), pp. 467–489. DOI: 10.1146/annurev.nucl.012809.104544.

[275] T. K. Gaisser and A. M. Hillas. "*Reliability of the Method of Constant Intensity Cuts for Reconstructing the Average Development of Vertical Showers*". In: 15th International Cosmic Ray Conference **8** (Jan. 1977), p. 353.

[276] L. D. 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*". In: *Dokl. Akad. Nauk SSSR* **92** (1953), p. 535. DOI: 10.1016/b978-0-08-010586-4.50080-8. URL: https://cds.cern.ch/record/436540.

[277] L. D. Landau and I. I. Pomeranchuk. "*Electron-cascade processes at ultra-high energies*". In: *Dokl. Akad. Nauk SSSR* **92** (1953), p. 735. DOI: 10.1016/b978-0-08-010586-4.50081-x. URL: https://cds.cern.ch/record/436781.

[278] A. B. Migdal. "*Bremsstrahlung and Pair Production in Condensed Media at High Energies*". In: *Physical Review* **103**.6 (Sept. 1956), pp. 1811–1820. DOI: 10.1103/PhysRev.103.1811.

[279] P. Arnold and S. Iqbal. "*The LPM effect in sequential bremsstrahlung*". In: *JHEP* **04** (2015). [Erratum: JHEP 09, 072 (2016)], p. 070. DOI: 10.1007/JHEP09(2016)072. arXiv: 1501.04964 [hep-ph].

[280] T. Stanev, C. Vankov, R. E. Streitmatter, R. W. Ellsworth, and T. Bowen. "*Development of ultrahigh-energy electromagnetic cascades in water and lead including the Landau-Pomeranchuk-Migdal effect*". In: *Phys. Rev. D* **25** (1982), pp. 1291–1304. DOI: 10.1103/PhysRevD.25.1291.

[281] J. Alvarez-Muñiz. "*EeV Hadronic Showers in Ice: The LPM effect*". In: *26th International Cosmic Ray Conference (ICRC26), Volume 1*. Vol. **1**. International Cosmic Ray Conference. Jan. 1999, p. 506. arXiv: astro-ph/9906347 [astro-ph].

[282] T. ERBER. "*High-Energy Electromagnetic Conversion Processes in Intense Magnetic Fields*". In: *Rev. Mod. Phys.* **38** (4 Oct. 1966), pp. 626–659. DOI: 10.1103/RevModPhys.38.626. URL: https://link.aps.org/doi/10.1103/RevModPhys.38.626.

[283] B. McBreen and C. J. Lambert. "*Interactions of high-energy ($E > 5 \times \times 10^{19}$ eV) photons in the Earth's magnetic field*". In: *Phys. Rev. D* **24** (9 Nov. 1981), pp. 2536–2538. DOI: 10.1103/PhysRevD.24.2536. URL: https://link.aps.org/doi/10.1103/PhysRevD.24.2536.

[284] P. Homola, M. Risse, R. Engel, D. Góra, J. Pe˛kala, B. Wilczyńska, and H. Wilczyński. "*Characteristics of geomagnetic cascading of ultra-high energy photons at the southern and northern sites of the Pierre Auger Observatory*". In: *Astroparticle Physics* **27**.2 (2007), pp. 174–184. ISSN: 0927-6505. DOI: https://doi.org/10.1016/j.astropartphys.2006.10.005. URL: https://www.sciencedirect.com/science/article/pii/S0927650506001551.

[285] M. Settimo. "*Search for ultra-High Energy Photons with the Pierre Auger Observatory*". In: *PoS* **Photon 2013** (2014), p. 062. DOI: 10.22323/1.192.0062.

[286] A. Aab, P. Abreu, M. Aglietta, J. M. Albury, I. Allekotte, A. Almela, J. Castillo, J. Alvarez-Muñiz, R. Alves Batista, G. A. Anastasi, et al. "*Reconstruction of events recorded with the surface detector of the Pierre Auger Observatory*". In: *Journal of Instrumentation* **15**.10 (Oct. 2020), P10021–P10021. ISSN: 1748-0221. DOI: 10.1088/1748-0221/15/10/p10021. URL: http://dx.doi.org/10.1088/1748-0221/15/10/P10021.

[287] X. Bertou, P. Billoir, and S. Dagoret-Campagne. "*LPM effect and pair production in the geomagnetic field: a signature of ultra-high energy photons in the Pierre Auger Observatory*". In: *Astroparticle Physics* **14**.2 (2000), pp. 121–130. ISSN: 0927-6505. DOI: https://doi.org/10.1016/S0927-6505(00)00112-2. URL: https://www.sciencedirect.com/science/article/pii/S0927650500001122.

[288] S. R. Klein. "*e+ e- pair production from 10-GeV to 10-ZeV*". In: *Radiat. Phys. Chem.* **75** (2006), pp. 696–711. DOI: 10.1016/j.radphyschem.2005.09.005. arXiv: hep-ex/0402028.

[289] J. N. Capdevielle, P. Gabriel, H. J. Gils, P. Grieder, D. Heck, J. Knapp, H. J. Mayer, J. Oehlschläger, H. Rebel, G. Schatz, and T. Thouw. "*Extensive air shower simulations with the CORSIKA program*". In: ed. by Lawrence Jones. Vol. **276**. American Institute of Physics Conference Series. June 1993, pp. 545–553. DOI: 10.1063/1.43851.

[290] D. Heck, J. Knapp, J. N. Capdevielle, G. Schatz, and T. Thouw. "*CORSIKA: a Monte Carlo code to simulate extensive air showers.*" In: (1998). Technical Report FZKA 6019.

[291] R. Engel. "*Indirect Detection of Cosmic Rays*". In: Irène Buvat (eds.) Claus Grupen. Handbook of particle detection and imaging. Berlin: Springer-Verlag, 2012. ISBN: 3642132715.

[292] "*The Pierre Auger Project Design Report*". In: (Jan. 1996). URL: https://inspirehep.net/files/6359807356b173c82673ecdcab43558d.

[293] D. Verberic. Maps of the Pierre Auger Observatory. 2020. URL: https://web.ikp.kit.edu/darko/auger/auger-array/auger_array-pdf/.

[294] The Pierre Auger Collaboration, J. Abraham, et al. Operations of and Future Plans for the Pierre Auger Observatory. 2009. arXiv: 0906.2354 [astro-ph.IM].

[295] "*The Pierre Auger Cosmic Ray Observatory*". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **798** (Oct. 2015), pp. 172–213. ISSN: 0168-9002. DOI: 10.1016/j.nima.2015.06.058. URL: http://dx.doi.org/10.1016/j.nima.2015.06.058.

[296] J. Abraham et al. "*The fluorescence detector of the Pierre Auger Observatory*". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **620**.2–3 (Aug. 2010), pp. 227–251. ISSN: 0168-9002. DOI: 10.1016/j.nima.2010.04.023. URL: http://dx.doi.org/10.1016/j.nima.2010.04.023.

[297] I. Allekotte et al. "*The surface detector system of the Pierre Auger Observatory*". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **586**.3 (Mar. 2008), pp. 409–420. ISSN: 0168-9002. DOI: 10.1016/j.nima.2007.12.016. URL: http://dx.doi.org/10.1016/j.nima.2007.12.016.

[298] J. Abraham et al. "*Trigger and aperture of the surface detector array of the Pierre Auger Observatory*". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **613**.1 (Jan 2010), pp. 29–39. ISSN: 0168-9002. DOI: 10.1016/j.nima.2009.11.018. URL: http://dx.doi.org/10.1016/j.nima.2009.11.018.

[299] R. Smida. "*Observation of Ultra-high Energy Cosmic Rays*". In: *29th International Conference on Physics in Collision*. 1 2010. arXiv: {1001.1660} (astro-ph.HE).

[300] X. Bertou, P. S. Allison, C. Bonifazi, P. Bauleo, C. M. Grunfeld, M. Aglietta, F. Arneodo, D. Barnhill, J. J. Beatty, N. G. Busca, et al. "*Calibration of the surface array of the Pierre Auger Observatory*". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **568**.2 (Dec. 2006), pp. 839–846. ISSN: 0168-9002. DOI: 10.1016/j.nima.2006.07.066. URL: http://dx.doi.org/10.1016/j.nima.2006.07.066.

[301] The Pierre Auger collaboration. "*The Pierre Auger Observatory scaler mode for the study of solar activity modulation of galactic cosmic rays*". In: *Journal of Instrumentation* **6**.01 (Jan. 2011). DOI: 10.1088/1748-0221/6/01/p01003. URL: https://doi.org/10.1088/1748-0221/6/01/p01003.

[302] A. Aab et al. "*Azimuthal Asymmetry in the Risetime of the Surface Detector Signals of the Pierre Auger Observatory*". In: *Phys. Rev. D* **93**.7 (2016), p. 072006. DOI: 10.1103/PhysRevD.93.072006. arXiv: {1604.00978} (astro-ph.HE).

[303] C. Bonifazi. "*The angular resolution of the Pierre Auger Observatory*". In: *Nuclear Physics B - Proceedings Supplements* **190** (May 2009), pp. 20–25. ISSN: 0920-5632. DOI: 10.1016/j.nuclphysbps.2009.03.063. URL: http://dx.doi.org/10.1016/j.nuclphysbps.2009.03.063.

[304] 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*". In: *Astropart. Phys.* **29** (2008), 243–256. DOI: 10.1016/j.astropartphys.2008.01.003. arXiv: 0712.1147 [astro-ph].

[305] J. Lapikens, A. A. Watson, P. Wild, and J. G. Wilson. "*The relation of longitudinal development to shower front structure*". In: vol. **4**. 1973, pp. 2582–2587.

[306] A. Aab et al. "*Inferences on mass composition and tests of hadronic interactions from 0.3 to 100 EeV using the water-Cherenkov detectors of the Pierre Auger Observatory*". In: *Phys. Rev. D* **96**.12 (2017), p. 122003. DOI: 10.1103/PhysRevD.96.122003. arXiv: {1710.07249} (astro-ph.HE).

[307] K. Kamata and J. Nishimura. "*The Lateral and the Angular Structure Functions of Electron Showers*". In: *Progress of Theoretical Physics Supplement* **6** (Feb. 1958), pp. 93–155. ISSN: 0375-9687. DOI: 10.1143/PTPS.6.93. eprint: https://academic.oup.com/ptps/article-pdf/doi/10.1143/PTPS.6.93/5270594/6-93.pdf. URL: https://doi.org/10.1143/PTPS.6.93.

[308] K. Greisen. "*The extensive air showers*". In: *Prog. Cosmic Ray Phys.* **3** (1956), p. 1.

[309] K. Greisen. "*Cosmic Ray Showers*". In: *Annual Review of Nuclear Science* **10**.1 (1960), pp. 63–108. DOI: 10.1146/annurev.ns.10.120160.000431. eprint: https://doi.org/10.1146/annurev.ns.10.120160.000431. URL: https://doi.org/10.1146/annurev.ns.10.120160.000431.

[310] D. NEWTON, J. KNAPP, and A. WATSON. "*The optimum distance at which to determine the size of a giant air shower*". In: *Astroparticle Physics* **26**.6 (Jan. 2007), pp. 414–419. ISSN: 0927-6505. DOI: 10.1016/j.astropartphys.2006.08.003. URL: http://dx.doi.org/10.1016/j.astropartphys.2006.08.003.

[311] J. Hersil, I. Escobar, D. Scott, G. Clark, and S. Olbert. "*Observations of Extensive Air Showers near the Maximum of Their Longitudinal Development*". In: *Phys. Rev. Lett.* **6** (1 Jan. 1961), pp. 22–23. DOI: 10.1103/PhysRevLett.6.22. URL: https://link.aps.org/doi/10.1103/PhysRevLett.6.22.

[312] A. Schulz. "*The measurement of the energy spectrum of cosmic rays above 3x10^{17} eV with the Pierre Auger Observatory*". In: **ICRC2013** (2013), p. 0769. URL: https://inspirehep.net/files/1548368d85c92c23c61a9e2ade55ee95.

[313] P. Billoir, C. Roucelle, and J.-C. Hamilton. "*Evaluation of the Primary Energy of UHE Photon-induced Atmospheric Showers from Ground Array Measurements*". In: (Jan. 2007). arXiv: astro-ph/0701583.

[314] S. Argirò, S. L. C. Barroso, J. Gonzalez, L. Nellen, T. Paul, T. A. Porter, L. Prado Jr., M. Roth, R. Ulrich, and D. Veberic. "*The offline software framework of the Pierre Auger Observatory*". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **580**.3 (Oct. 2007), pp. 1485–1496. ISSN: 0168-9002. DOI: 10.1016/j.nima.2007.07.010. URL: http://dx.doi.org/10.1016/j.nima.2007.07.010.

[315] J. G. Gonzalez. "*The Offline Software of the Pierre Auger Observatory: Lessons Learned*". In: (Aug. 2012). arXiv: 1208.2154 [astro-ph.IM].

[316] R. Brun and F. Rademakers. "*ROOT — An object oriented data analysis framework*". In: *Nuclear Instruments and Methods in Physics Research A* **389**.1 (Feb. 1997), pp. 81–86. DOI: 10.1016/S0168-9002(97)00048-X.

[317] G. Simone. "*GEANT4: Simulation for the Next Generation of HEP Experiments*". In: *Computing in High Energy Physics: CHEP '95 - Proceedings of the International Conference. Edited by SHELLARD RONALD ET AL. Published by World Scientific Publishing Co. Pte. Ltd*. Jan. 1996, pp. 147–152. DOI: 10.1142/9789814447188_0024.

[318] S. Agostinelli et al. "*GEANT4–a simulation toolkit*". In: *Nucl. Instrum. Meth. A* **506** (2003), 250–303. DOI: 10.1016/S0168-9002(03)01368-8.

[319] The Pierre Auger Collaboration, A. Aab, et al. The Pierre Auger Observatory Upgrade - Preliminary Design Report. 2016. arXiv: 1604.03637 [astro-ph.IM].

[320] A. M. Hillas. "*Shower simulation: lessons from MOCCA*". In: *Nuclear Physics B - Proceedings Supplements* **52**.3 (1997), pp. 29–42. ISSN: 0920-5632. DOI: https://doi.org/10.1016/S0920-5632(96)00847-X. URL: https://www.sciencedirect.com/science/article/pii/S092056329600847X.

[321] P. Billoir. "*A sampling procedure to regenerate particles in a ground detector from a "thinned" air shower simulation output*". In: *Astroparticle Physics* **30**.5 (2008), pp. 270–285. ISSN: 0927-6505. DOI: https://doi.org/10.1016/j.astropartphys.2008.10.002. URL: https://www.sciencedirect.com/science/article/pii/S0927650508001394.

[322] A. Hoecker et al. "*TMVA - Toolkit for Multivariate Data Analysis*". In: *arXiv e-prints*, physics/0703039 (Mar. 2007), physics/0703039. arXiv: physics/0703039 [physics.data-an].

[323] A. L. Samuel. "*Some Studies in Machine Learning Using the Game of Checkers*". In: *IBM Journal of Research and Development* **3**.3 (1959), pp. 210–229. DOI: 10.1147/rd.33.0210.

[324] D. Barber. *Bayesian reasoning and machine learning*. Cambridge University Press, 2011. ISBN: 0521518148.

[325] E. Alpaydin. *Introduction to Machine Learning*. 3rd Edition. The MIT Press, 2014. ISBN: 0262028182.

[326] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone. Classification and Regression Trees. 1st ed. Belmont, CA: Wadsworth International Group, 1984. ISBN: 9781315139470. URL: https://doi.org/10.1201/9781315139470.

[327] R. E. Schapire. "*The strength of weak learnability*". In: *Machine Learning* **5** (2 1990), pp. 197–227. DOI: 10.1007/BF00116037. URL: https://doi.org/10.1007/BF00116037.

[328] Y. Freund. "*Boosting a Weak Learning Algorithm by Majority*". In: *Information and Computation* **121**.2 (1995), pp. 256–285. ISSN: 0890-5401. DOI: https://doi.org/10.1006/inco.1995.1136. URL: https://www.sciencedirect.com/science/article/pii/S0890540185711364.

[329] Y. Freund and R. E. Schapire. "*Experiments with a New Boosting Algorithm*". In: ICML'96 (1996), pp. 148–156. DOI: 10.5555/3091696.3091715.

[330] L. Breiman. Arcing The Edge. Technical Report 486. Ann. Prob. 26, 1683-1702 (1998). Statistics Department, University of California, Berkeley, 1997. URL: https://statistics.berkeley.edu/tech-reports/486.

[331] J. H. Friedman. "*Stochastic gradient boosting*". In: *Computational Statistics & Data Analysis* **38**.4 (2002). Nonlinear Methods and Data Mining, pp. 367–378. ISSN: 0167-9473. DOI: https://doi.org/10.1016/S0167-9473(01)00065-2. URL: https://www.sciencedirect.com/science/article/pii/S0167947301000652.

[332] R. A. Fisher. "*The use of multiple measurements in taxonomic problems*". In: *Annals of Eugenics* **7**.2 (1936), pp. 179–188. DOI: https://doi.org/10.1111/j.1469-1809.1936.tb02137.x. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1469-1809.1936.tb02137.x. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1469-1809.1936.tb02137.x.

[333] C. M. Bishop. Pattern Recognition and Machine Learning. Springer New York, 2011. ISBN: 0387310738.

[334] V. Vapnik and C. Cortes. "*Support-vector networks*". In: *Machine Learning* **20** (3 1995), pp. 273–297. DOI: 10.1007/BF00994018. URL: https://doi.org/10.1007/BF00994018.

[335] V. N. Vapnik. The Nature of Statistical Learning Theory. 2nd ed. Statistics for Engineering and Information Science. Name Springer, New York, 2000. ISBN: 978-0-387-98780-4. URL: https://doi.org/10.1007/978-1-4757-3264-1.

[336] URL: https://www.learnopencv.com/wp-content/uploads/2018/07/support-vectors-and-maximum-margin.png.

[337] F. Rosenblatt. "*The perceptron: A probabilistic model for information storage and organization in the brain*". In: *Psychological Review* **65**.6 (1958), pp. 386–408. DOI: doi:10.1037/h0042519.

[338] URL: https://miro.medium.com/max/2400/0*44z992IXd9rqyIWk.png.

[339] C. G. BROYDEN. "*The Convergence of a Class of Double-rank Minimization Algorithms 1. General Considerations*". In: *IMA Journal of Applied Mathematics* **6**.1 (Mar. 1970), pp. 76–90. ISSN: 0272-4960. DOI: 10.1093/imamat/6.1.76. eprint: https://academic.oup.com/imamat/article-pdf/6/1/76/2233756/6-1-76.pdf. URL: https://doi.org/10.1093/imamat/6.1.76.

[340] R. Fletcher. "*A new approach to variable metric algorithms*". In: *The Computer Journal* **13**.3 (Jan. 1970), pp. 317–322. ISSN: 0010-4620. DOI: 10.1093/comjnl/13.3.317. URL: https://doi.org/10.1093/comjnl/13.3.317.

[341] D. Goldfarb. "*A family of variable-metric methods derived by variational means*". In: *Math. Comp.* **24** (1970), pp. 23–26. URL: https://doi.org/10.1090/S0025-5718-1970-0258249-6.

[342] D. F. Shanno. "*Conditioning of quasi-Newton methods for function minimization*". In: *Math. Comp.* **24** (1970), pp. 647–656. URL: https://doi.org/10.1090/S0025-5718-1970-0274029-X.

[343] T. Fawcett. "*An introduction to ROC analysis*". In: *Pattern Recognit. Lett.* **27** (2006), pp. 861–874. DOI: 10.1016/j.patrec.2005.10.010.

[344] M. Sokolova and G. Lapalme. "*A systematic analysis of performance measures for classification tasks*". In: *Inf. Process. Manag.* **45** (2009), pp. 427–437. DOI: 10.1016/j.ipm.2009.03.002.

[345] A. Valassi. "*Binary classifier metrics for optimizing HEP event selection*". In: *EPJ Web Conf.* **214** (2019), p. 06004. DOI: 10.1051/epjconf/201921406004. URL: https://doi.org/10.1051/epjconf/201921406004.

List of Figures

2.1. Nobel prize ceremony 1936. Front row from left to right: Prof. Otto Loewi, Sir Henry Dale (both laureates in medicine), Prof. Peter Debye (laureate in chemistry), Dr. C.D. Anderson, Prof. V.F. Hess. Figure taken from [25]. 8

2.2. Compilation of the full range CR energy spectrum measured by several experiments (after 2000). Figure taken from [37]. . . 10

2.3. Differential energy spectrum of charged cosmic rays, measured from 10^{13} eV to few 10^{20} eV by cosmic ray observatories. The spectrum has been multiplied by $E^{2.6}$ in order to put more emphasis on the features of the steep spectrum which are otherwise difficult to discern. These features are: the "Knee" at around $4 \cdot 10^{15}$ eV, the "2^{nd} Knee" at 10^{17} eV and the "Ankle" at $4 \cdot 10^{18}$ eV. Figure taken from [42]. 12

2.4. Amplitude (left) and phase (right) of the equatorial dipole amplitude. Plot taken from [62]. 14

2.5. The flux of CRs above 8 EeV in equatorial coordinates, where the dashed line represents the galactic plane and the star indicates the GC. Plot taken from [62]. 14

2.6. The energy spectrum measured at the Pierre Auger Observatory, scaled with E^3 and fitted with a sequence of four power laws represented by the red line. The spectral index γ_i of the power law corresponds to the energy intervals identified with the numbers ($i = 1, 2, 3, 4$), the shaded band represents the statistical uncertainty of the fit, and the upper limits are at the 90% CL. Plot taken from [49]. 16

2.7. Comparison between the spectrum derived with the Pierre Auger Observatory in blue and the one derived at the Telescope Array, scaled with E^3. The blue graph referred to by the words "this work" in the legend refers to the article [75] from which the plot is taken. 17

2.8. Hillas diagram, representing source classes shown as function of their characteristic size and magnetic field strength (quoted values are in the comoving frame of the source). The diagonal lines indicate the limit beyond which confinement of protons (red) and iron (blue) nuclei with energy 10^{20} eV is possible. Two different shock velocities β (in units of c) are indicated as solid and dashed lines. Figure taken from [82]. 20

2.9. This is a mosaic image of Cas A, composed of X-ray data from NASA's Chandra X-ray Observatory mission, infrared data from the Spitzer Space Telescope, and optical data from other telescopes [114]. 25

2.10. This image of Centaurus A is composed of images obtained with three instruments, operating at very different wavelengths and thereby revealing the lobes and jets emanating from the active galaxy's central black hole: in orange the 870-micron submillimetre data, from LABOCA on APEX, in blue X-ray data from the Chandra X-ray Observatory, and visible light data from the WFI on the MPG/ESO 2.2 m telescope show the background stars and the Galaxy's characteristic dust lane in "true colour". 27

2.11. Schematic depiction of compact BS mergers. Figure taken from [134]. 28

2.12. The pulsar in the center of the Crab Nebula. Optical data from the Hubble Space Telescope (in red) are combined with X-ray images from Chandra X-ray Observatory (in blue). 29

2.13. Schematic depiction in Feynman diagrams of the PP process in lowest order. Figure taken from [179]. 37

2.14. (In the nucleus rest frame) **a)** Contributions of different energy loss processes to the energy evolution of the energy loss length χ_{loss} of protons. **b)** Contributions of PP and PD to the total attenuation length of iron nuclei with evolution of the logarithm of the Lorentz factor on the abscissa. Both figures taken from [180]. 38

2.15. Energy loss lengths χ_{loss} of different species. Figure taken from [180]. 38

2.16. The relative abundances of elements in the solar system and in low energy galactic CRs (up to Ni), normalized to a value of 10^6 for Si. Figure taken from [116]. 39

2.17. The CR spectra for individual primary particle types as measured by various experiments for energies up to and above the knee, including the uncertainties of the hadronic interaction models that were used to perform the simulations on which the analysis is based. Figure taken from [116]. 40

2.18. The Pierre Auger Observatory's measurements of $\langle X_{\mathrm{max}} \rangle$ (**left**) and $\sigma(X_{\mathrm{max}})$ (**right**) compared to the predictions for proton and iron nuclei of three different hadronic interaction models. The hadronic interaction models QGSJETII-04 [183], EPOS-LHC (V3400) [184], and Sibyll 2.3c [185] are commonly used to simulate air showers and are all based on the simple parton model associated with the Gribov-Regge multiple scattering approach, further explained in their respective references as well as in [186]. Figure taken from [71]. 41

3.1. Upper limits on the integrated photon flux at 90% CL by KASCADE-Grande and EAS-MSU and 95% CL by TA and the Pierre Auger Observatory compared to model predictions [204]. 48

3.2. Photon flux, scaled with E^2 **a)** from protons and **b)** from iron. Black dots show the total photon flux. Both figures taken from [173]. 49

3.3. Interaction lengths for cosmic-ray photons interacting with cosmic photon backgrounds CBR. The relevant processes are PP (PP, black solid line) and double PP (DPP, black dashed line). Figure taken from [173]. 51

3.4. **a)** Upper limit on diffuse fluxes, compared to the differential limits obtained by other experiments and astrophysical and cosmogenic neutrino models. In red are shown the integral and differential upper limit for the normalisation constant assuming a E^{-2} energy flux for a single flavour for all the three neutrino channels, and the dashed red line is the ES channel alone, which contributes the strongest to the limit [239]. **b)** Auger 90% CL upper limit on the fluxes of point-like sources as a function of equatorial declination obtained from the non-observation of neutrinos. Note the different energy ranges where the limits of each observatory apply. See [238] and references therein. 57

3.5. Rates of selected streams received by AMON. The plot shows the number of events per day received during the month of June 2018. Figure taken from [257]. 60

4.1. Schematic view of an EAS, with the hadron, muon (and neutrino), and EM components. Figure taken from [267]. 63

4.2. Schematic depiction of the dominant Feynman diagrams for the bremsstrahlung process for an electron in the EM field of a nucleus. Figure taken from [179]. 64

4.3. Average (left) lateral and (right) longitudinal shower profiles for vertical, proton-induced showers at 10^{19} eV, with energy thresholds in the simulation of 0.25 MeV for γ and $e^{\pm}$, and 0.1 GeV for muons and hadrons. Figure taken from [274]. . . . 68

4.4. Schematic view of (left) a plane shower front approximation, (right) a spherical shower front, where the front reaches the station i at position $\vec{x}_i$ at time t_i. Figure taken from [286]. 71
4.5. Tracks of the secondary particles from the muonic, EM and hadronic shower components for simulated air showers, initiated by a photon, a proton and an iron nucleus. Simulations were done with CORSIKA and the height of each graph corresponds to an altitude of 25 km, and a width of 400 m. Figure taken from [291]. 72

5.1. A map of the Pierre Auger Observatory, where each black dot corresponds to one of the 1660 SD stations. The four FD sites together with the location of HEAT are shown, and also the two laser facilities, CLF and XLF, near the Observatory center, together with the extensions AERA and AMIGA, and the BLF [293]. 74
5.2. A WCD station of the Pierre Auger Observatory with the various components of the system indicated. Figure taken from [299]. 75
5.3. Example of T3 configurations: **a)**: the 3-fold T3 mode ToT-2C1&-3C2, **b)**: the 4-fold mode 2C1&-3C2&-4C4, where C1, C2, C3, C4 indicate the first, second, third and fourth sets of neighbours, respectively, at 1.5, 3, 4.5 and 6 km from a given detector. **c)** Scheme of an hexagon of detectors: the elemental hexagon cell where a cell a_{cell} is the shaded area around the central detector. Figures taken from [298]. 79
5.4. SD trigger efficiency as a function of the simulated (true) energy, integrated for showers with $\theta < 60^{\circ}$, for different primaries. Figure taken from [298]. 80
5.5. **a)** A top-down view, zoomed-in on an event falling within the array, with the projection of the shower axis on the ground represented by the black line, ending at the impact-point of the shower-core (red square). The SD stations are coloured according to their trigger time (blue is early, green is late) and their area proportional to the logarithm of the signal amplitude. Figure taken from [286]. **b)** Illustration of a hybrid event, with the pixels in the camera of the FD tracing the shower profile – specifically, the energy loss of the shower as a function of its penetration in the atmosphere, together with the SD array detecting the particles from the same shower. The red line shows the trajectory of the shower, the shower axis. 81
5.6. Schematic representation of the shower geometry for incoming direction of the primary particle in two regions: "early" for $|\zeta| < \pi/2$, and "late" region for $|\zeta| > \pi/2$. Figure taken from [302]. 82

5.7. Influence of geometrical effects on detector signals. Figure taken from [306]. 84
5.8. **a)** Example of a LDF of a SD event and its NKG type parametrization (eq. (5.7)). Figure taken from [295]. **b)** Correlation between the SD (S_{38}) and FD energies. Figure taken from [73]. 88

6.1. An illustration of the topology in ML, denoting loose associations between subjects, taken from [324]. 97
6.2. The axis of the PCA in red is sown together with the data in blue and the MC prediction for photons in orange for the normalized observables L_{LDF} and $g\Delta$ [204]. 99
6.3. An example of a linear multivariate decision tree, where the node can place an arbitrary hyperplane. Figure taken from [325]. 101
6.4. Principle of selecting a hyperplane in two selection cases, where on the right the corresponding projection is based on the Fisher linear discriminant, resulting in a greatly improved class separation, more in the text. Figure taken from [333]. 105
6.5. Example of a SVM with the two class data sets separated by a hyperplane (decision boundary) maximizing the perpendicular distance to the closest point of both classes. Figure taken from [336]. 106
6.6. Schematic illustration of a simple perceptron with one (hidden) layer between the input and output layer. Figure taken from [322]. 107
6.7. Schematic illustration of a feedforward ANN (or MLP) with more hidden layers between the input and output layer. 107
6.8. Examples of some prominent AFs [338]. 108
6.9. Distributions of the **a)**: (recalculated) rise-time; **b)** curvature of the shower front [km^{-1}] 114
6.10. Case A: rise-time recalculated and curvature with BDTG **a)** MVA distribution; **b)** cut efficiencies; 114
6.11. Linear correlation matrices for case A: rise-time recalculated and curvature with BDTG **a)** signal, **b)** background, where a positive value indicates a positive correlation, with 100% indicating a full positive correlation, zero no correlation and -100% a full anti-correlation. 115
6.12. Distributions of the rise-time Δ_R. 115
6.13. Case B: rise-time Δ_R and curvature with BDTG **a)** MVA distribution; **b)** cut efficiencies; 115
6.14. Linear correlation matrices for case B: rise-time Δ_R and curvature with BDTG **a)** signal, **b)** background, where a positive value indicates a positive correlation, with 100% indicating a full positive correlation, zero no correlation and -100% a full anti-correlation. 116

6.15. ROC curves: **a)** case A: rise-time recalculated and curvature vs. **b)** case B: Δ_R and curvature; The various ML methods are ranked by the integral of the ROC diagram with the best method on top. 116
6.16. Distribution of the NSt. 117
6.17. Case C: rise-time, curvature and NSt with BDTG **a)** MVA distribution; **b)** cut efficiencies; 117
6.18. Linear correlation matrices for case C: rise-time, curvature and NSt with BDTG **a)** signal, **b)** background, where a positive value indicates a positive correlation, with 100% indicating a full positive correlation, zero no correlation and -100% a full anti-correlation. 118
6.19. Case D: Δ_R, curvature and NSt with BDTG **a)** MVA distribution; **b)** cut efficiencies; 118
6.20. Linear correlation matrices for case D: Δ_R, curvature and NSt with BDTG **a)** signal, **b)** background, where a positive value indicates a positive correlation, with 100% indicating a full positive correlation, zero no correlation and -100% a full anti-correlation. 119
6.21. ROC curves: **a)** case C: rise-time recalculated, curvature and NSt vs. **b)** case D: Δ_R, curvature and NSt; The various ML methods are ranked by the integral of the ROC diagram with the best method on top. 119
6.22. Distributions for the observables **a)** $t_{1/2}(1000)$ **b)** Zenith angle, expressed as $\sec(\theta)$ **c)** recalculated rise-time **d)** Δ_R **e)** curvature of the shower front [1/km] **f)** NSt. 121
6.23. Case A: recalculated rise-time, curvature, NSt and θ with BDTG **a)** MVA distribution; **b)** cut efficiencies; 122
6.24. Linear correlation matrices for case A: recalculated rise-time, curvature, NSt and θ with BDTG **a)** signal, **b)** background, where a positive value indicates a positive correlation, with 100% indicating a full positive correlation, zero no correlation and -100% a full anti-correlation. 123
6.25. Case B: Δ_R, curvature, NSt and θ with BDTG **a)** MVA distribution; **b)** cut efficiencies; 123
6.26. Linear correlation matrices for case B: Δ_R, curvature, NSt and θ with BDTG **a)** signal, **b)** background, where a positive value indicates a positive correlation, with 100% indicating a full positive correlation, zero no correlation and -100% a full anti-correlation. 124
6.27. ROC curves: **a)** case A: vs. **b)** case B: Δ_R, curvature and NSt; The various ML methods are ranked by the integral of the ROC diagram with the best method on top. 124
6.28. Distributions for the observables **a)** $t_{1/2}(1000)$ **b)** Zenith angle, expressed as $\sec(\theta)$ **c)** NSt **d)** curvature of the shower front [1/km]. 125

6.29. TMVA response distributions **a)** BDTG **b)** BoostedFisher **c)** Fisher **d)** FisherG and e) MLPBFGS. 126
6.30. TMVA over-training check for **a)** BDTG **b)** BoostedFisher **c)** Fisher **d)** FisherG and e) MLPBFGS. The good agreement between training and test distribution indicates that over-training is under control, where the shaded area represents the test sample while dots represent the training sample. 128
6.31. TMVA convergence test for the MLPBNN. 129
6.32. For the method of Fisher for signal and background **a)** the PDFs, **b)** the TMVA probability and **c)** the TMVA rarity. 129
6.33. Parallel coordinates representation for BDTG for signal (photons). Parallel coordinates of the Shrink method for photon (left) and proton (right) primaries. Green polylines are drawn for each event and connect parallel vertical axis, which represent the input observables and the classifier response. The individual value of an event is referred to via the position of the line on the vertical axis, and the frequency scale of each observable is represented by a red histogram superimposed to the vertical axis, and blue polylines mark events which are misclassified by the classifier response. 130
6.34. Results of the four variables from fig. 6.28 combined in a MVA with BDTG **a)** MVA variable distribution showing excellent separation between signal (photons) and background (protons); **b)** cut efficiencies; 131
6.35. **a)** inverse ROC curves, **b)** ROC curves; The various ML methods are ranked by the integral of the ROC diagram with the best method on top. 132
6.36. Linear correlation matrices for with BDTG **a)** signal, **b)** background. 132
6.37. Distributions for the observables with the background being composed of 50% proton and 50% iron **a)** $t_{1/2}(1000)$ **b)** Zenith angle, expressed as $\sec(\theta)$ **c)** NSt **d)** curvature of the shower front [1/km]. 134
6.38. Results of the four variables from fig. 6.28 combined in a MVA with BDTG **a)** MVA variable distribution showing excellent separation between signal (photons) and background (50% protons + 50% iron); **b)** cut efficiencies; 135

A.1. Fits of benchmark function (red) to non-saturated data (black points) and fits of high-gain saturated benchmark function (blue) to high-gain saturated data (gray points). 182
A.2. Comparison of the LDF β parameter for photon, proton and the Auger SD burn sample distributions. 183
A.3. Comparison of $S(1000)$ of photons, protons and the Auger SD burn sample. 183

A.4. Comparison of the reconstructed SD energy of photons, protons and the Auger SD burn sample. 184
A.5. Comparison of the photon, proton and the Auger SD burn sample distributions. 184
A.6. Results of the following methods: a) MVA for Fisher, b) Fisher cut efficiencies, c) MVA for FisherG, d) FisherG cut efficiencies, and continuation with further methods in the next figure A.7. . 185
A.7. Continuation from fig.: A.6: e) MVA for BoostedFisher, f) BoostedFisher cut efficiencies, g) MVA for MLPBFGS, h) MLPBFGS cut efficiencies, i) MVA for MLPBNN, j) MLPBNN cut efficiencies. 186
A.8. MLP BFGS (with ReLU as AF) MVA variable distribution showing second best separation between signal (photons) and background (hadrons). 187
A.9. PCA transformation of the four input variables. 188

Appendix

A.

A.1. Further plots

A.1.1. Benchmark function fits

A benchmark function fit, applied to the v13r0 production of Pierre Auger data, of which the description can be found in section 5.2.4.1 is shown here. The fits are split into high-gain saturated benchmark functions and non-saturated benchmark functions [302, 306], and applied to SD station rise-times $t_{1/2}$, obtained from high-gain saturated or non-saturated PMT traces, respectively. The fits were produced for ten zenith angle bins between 0° and 60°, and for a reference energy bin between $10^{18.9}$ eV and $10^{19.1}$ eV. The example in figure A.1 shows the bin $\sec\theta$ [1.1, 1.2].

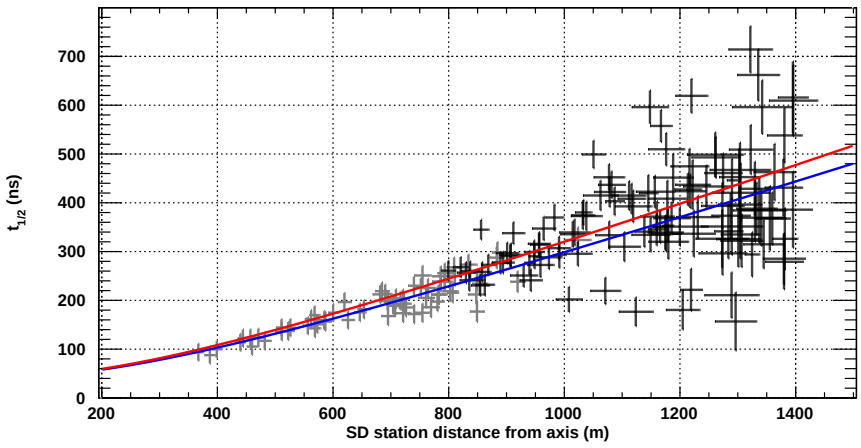


Figure A.1.: Fits of benchmark function (red) to non-saturated data (black points) and fits of high-gain saturated benchmark function (blue) to high-gain saturated data (gray points).

A.1.2. LDF β

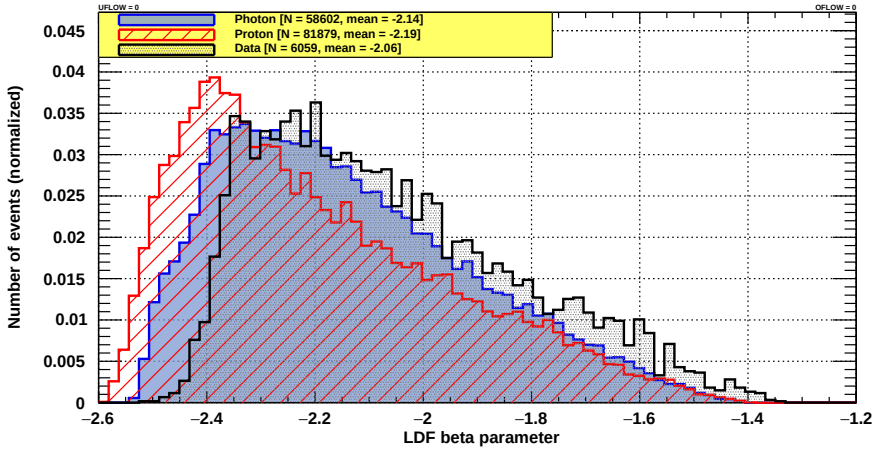


Figure A.2.: Comparison of the LDF β parameter for photon, proton and the Auger SD burn sample distributions.

A.1.3. $S(1000)$

Comparison of $S(1000)$ SD energy estimator for the configuration used in sec. 6.6.6.

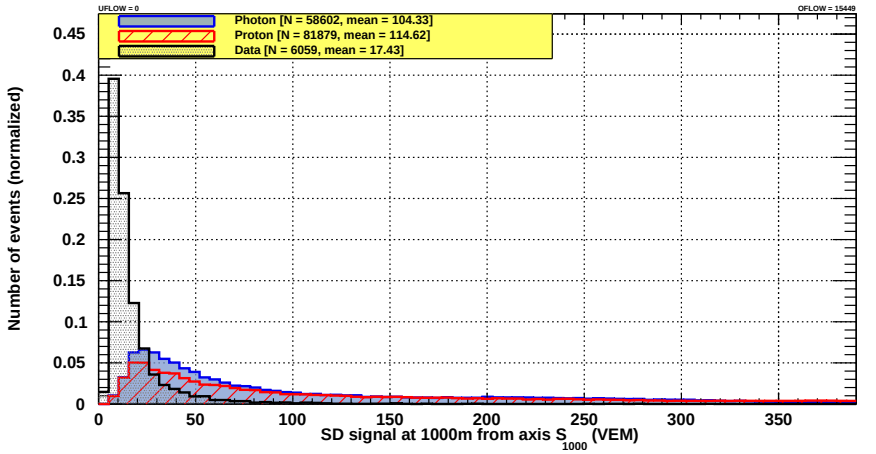


Figure A.3.: Comparison of $S(1000)$ of photons, protons and the Auger SD burn sample.

A.1.4. S_{38}

Comparison of the S_{38} SD energy estimator for the configuration used in sec. 6.6.6.

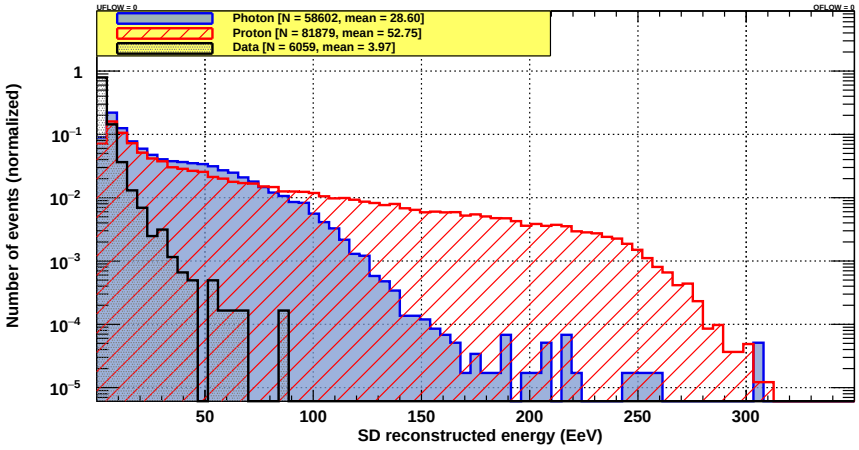


Figure A.4.: Comparison of the reconstructed SD energy of photons, protons and the Auger SD burn sample.

A.1.5. $S(1000)$ vs. S_{38}

$S(1000)$ from fig. A.3 plotted in a scatter plot vs. S_{38} from fig. A.4 for the configuration used in sec. 6.6.6.

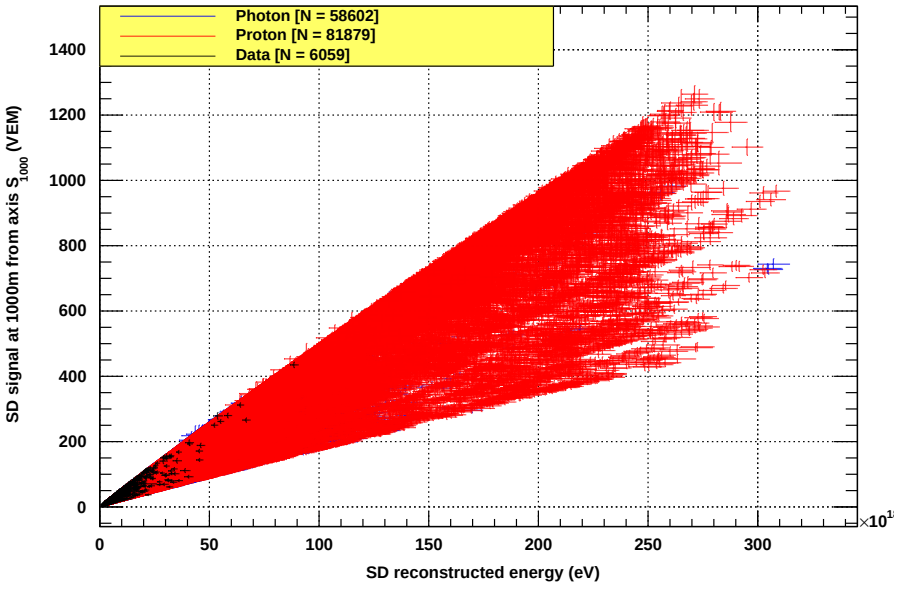


Figure A.5.: Comparison of the photon, proton and the Auger SD burn sample distributions.

A.2. MVA distributions and cut efficiencies for tested methods on SD burn sample

The MVA variable distributions and the cut efficiencies for other tested methods can be seen in figures A.6 and A.7.

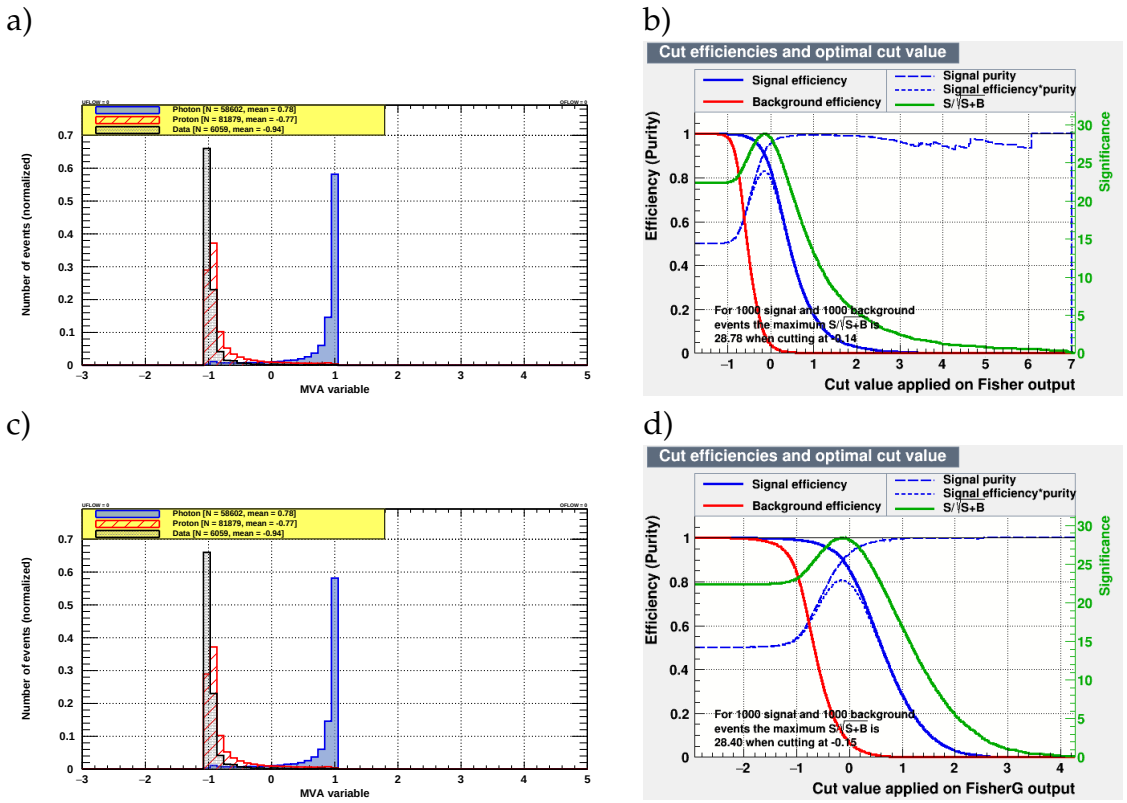


Figure A.6.: Results of the following methods: a) MVA for Fisher, b) Fisher cut efficiencies, c) MVA for FisherG, d) FisherG cut efficiencies, and continuation with further methods in the next figure A.7.

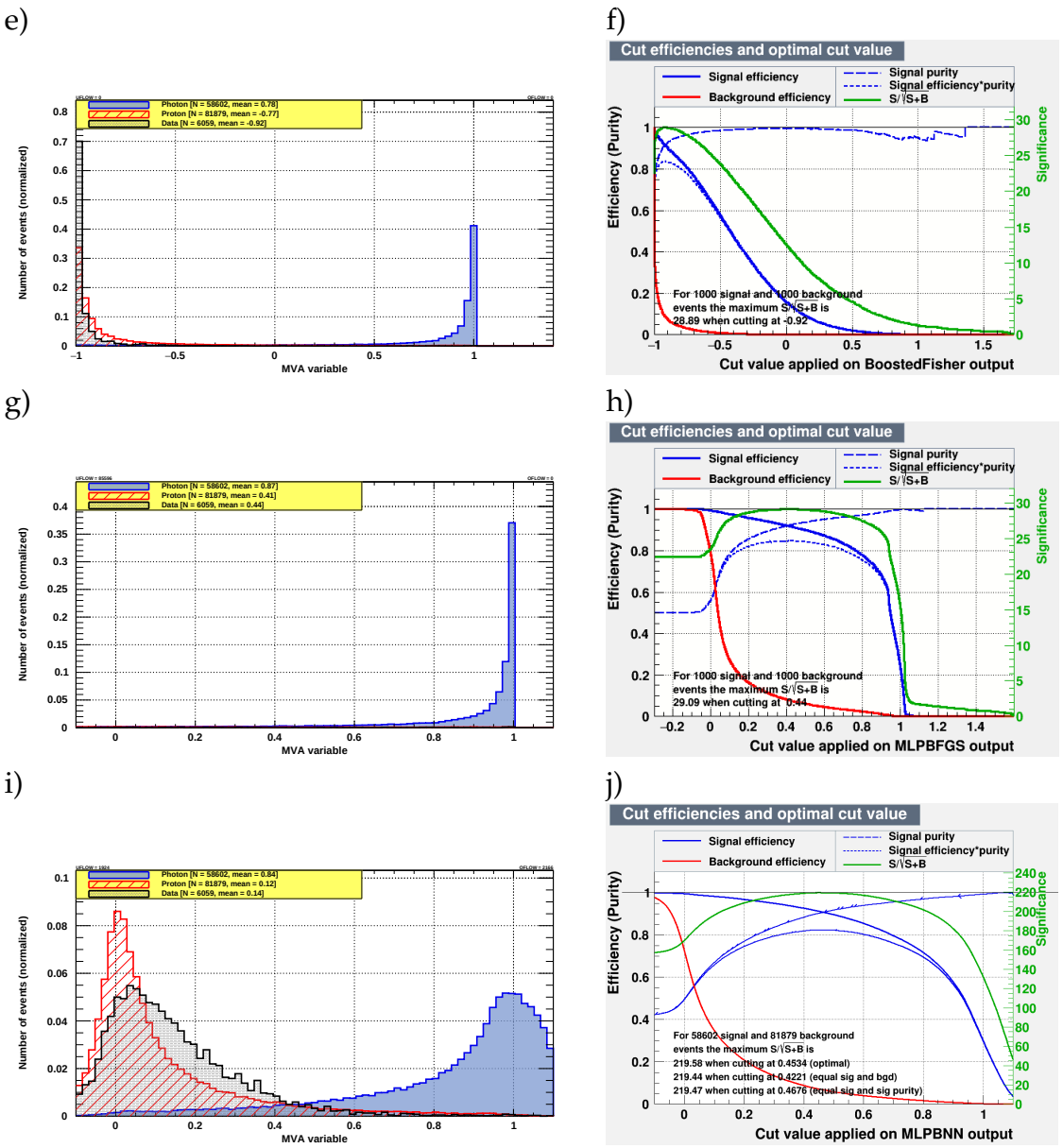


Figure A.7.: Continuation from fig.: A.6: e) MVA for BoostedFisher, f) BoostedFisher cut efficiencies, g) MVA for MLPBFGS, h) MLPBFGS cut efficiencies, i) MVA for MLPBNN, j) MLPBNN cut efficiencies.

A.3. ReLU activating function in MLPBFGS for MVA

To test a potential optimization, the AF of the NN was changed from the hyperbolic tangent to the ReLU function (see sec. 6.5.1.5). As a consequence, the neural network results became very good, the second best of the tested methods, but still could not reach the performance of the BDTG, though being a fast and good alternative to it. The BFGS feedforward MLP MVA variable distribution, with ReLU is shown in figure A.8.

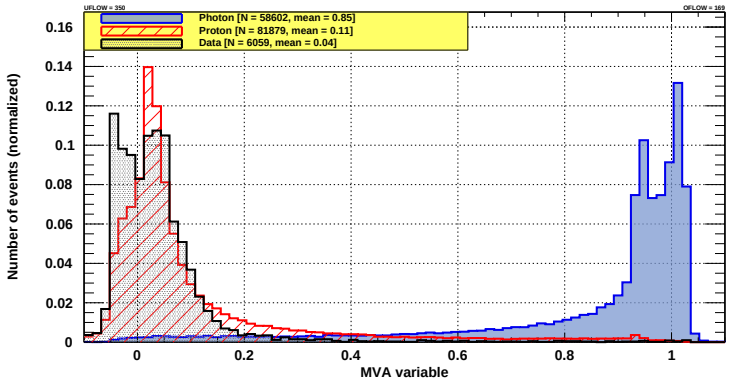


Figure A.8.: MLP BFGS (with ReLU as AF) MVA variable distribution showing second best separation between signal (photons) and background (hadrons).

A.4. PCA transformed MVA input variables

The PCA transformed input variables, computed by the ROOT class TPrincipal, can be seen in figure A.9.

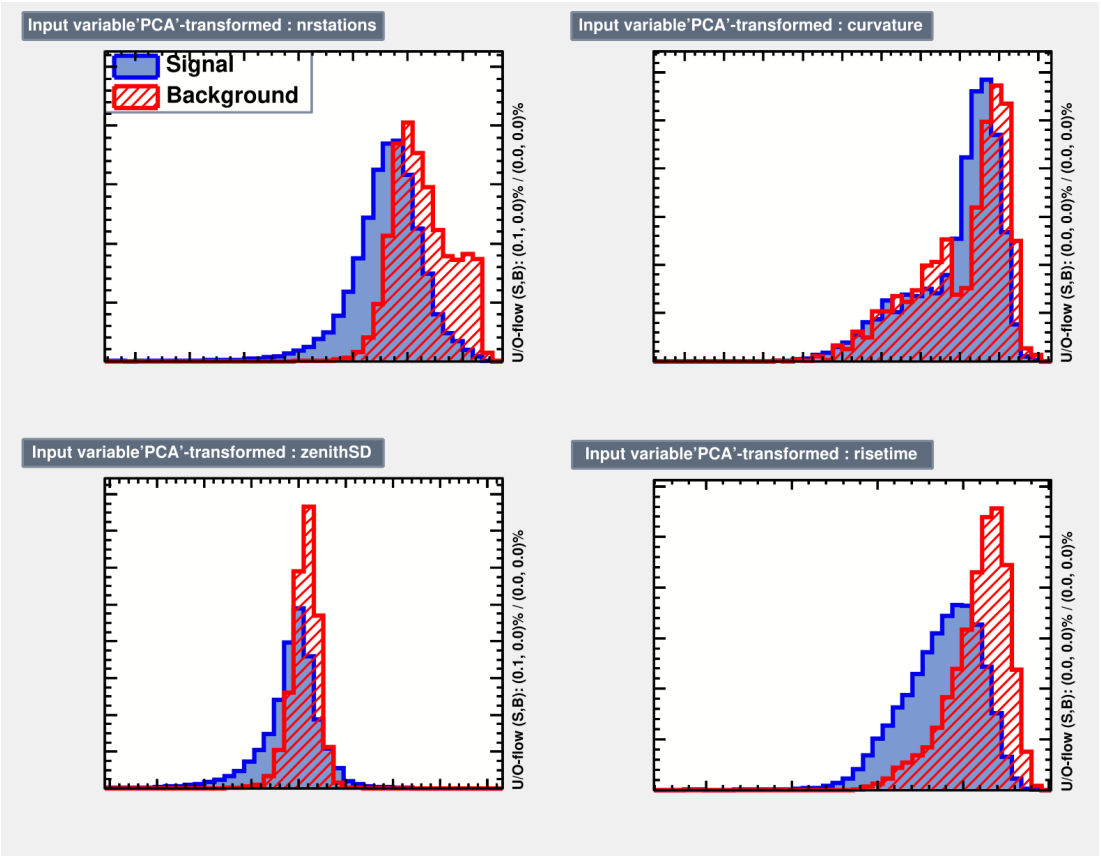


Figure A.9.: PCA transformation of the four input variables.

A.5. Comparison of methods

Values in table A.1 are given only for the ML method of BDTG, and #Obs is the number of observables in the MVA. The column "Section" redirects to sections with details on the observables used. Presented for comparison are the values of the signal (background) efficiency $\epsilon_{S(B)}$ (eq. (6.10)), the background (signal) rejection $r_{B(S)}$, the purity $\mathcal{P}$ (eq. (6.11)) and the signal-to-noise ratio (S-N-R).

#Obs	Section	ϵ_S	$\mathbf{r_S}$	ϵ_B	$\mathbf{r_B}$	$\mathcal{P}$	S-N-R
2	6.6.3.1	74.35%	25.65%	75.58%	24.42%	0.0966	74.85
	6.6.3.2	72.45%	27.55%	74.19%	25.81%	0.1124	73.68
3	6.6.4.1	81.51%	18.49%	80.61%	19.39%	0.0589	82.14
	6.6.4.2	80.72%	19.28%	80.37%	19.63%	0.0592	82.16
4	6.6.5.1	96.34%	3.66%	83.64%	16.36%	0.0183	230.00
	6.6.5.2	96.45%	3.55%	76.38%	23.62%	0.0635	245.16
	6.6.6	91.4%	8.6%	93.9%	6.1%	0.0157	221.26
	6.6.6.1	94.27%	5.73%	94.8%	5.2%	0.0246	228.29

Table A.1.: Comparison of values of the presented MVAs for the ML method of BDTG.

B.

B.1. Offline module sequences

B.1.1. Simulations with Offline version: ape-0.99-v3r3p4

```
<loop numTimes="unbounded" pushEventToStack="no">
  <module> CachedShowerRegeneratorOG </module>
  <module> G4TankSimulatorOG         </module>
</loop>

<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> SdCalibratorOG                  </module>
<module> SdSignalRecoveryKLT             </module>
<module> SdMonteCarloEventSelectorOG     </module>
<module> SdEventSelectorOG               </module>
<module> SdPlaneFitOG                    </module>

<module> LDFFinderKG                     </module>
<module> SdEventPosteriorSelectorOG      </module>
<module> Risetime1000LLL                 </module>

<module> EventFileExporterOG             </module>
<module> RecDataWriterNG                 </module>
```

B.1.2. Simulations with Offline trunk version 32493

```
<module> EventGeneratorOG           </module>
<module> SdAccidentalInjectorKG     </module>
<module> TabulatedTankSimulatorKG   </module>

&SdSimulation;

<module> ClearParticleLists              </module>
<module> CentralTriggerSimulatorXb       </module>
<module> CentralTriggerEventBuilderOG    </module>
<module> EventBuilderOG                  </module>
<module> EventCheckerOG                  </module>
<module> SdCalibratorOG                  </module>

<module> SdPMTSignalShapeQualityChecker  </module>
<module> SdSignalRecoveryKLT             </module>
<module> SdMonteCarloEventSelectorOG     </module>
<module> SdEventSelectorOG               </module>

<module> SdPlaneFitOG                    </module>
<module> LDFFinderKG                     </module>
<module> DLECorrectionWG                 </module>
<module> SdCompositionParameters         </module>

<try>
  <module> ScintillatorLDFFinderKG </module>
</try>

<module> SdEventPosteriorSelectorOG      </module>
<module> Risetime1000LLL                 </module>

<module> EventFileExporterOG             </module>
<module> RecDataWriterNG                 </module>
```

B.2. Proposal for the technical implementation

The plans for the technical implementation of the developed analysis for the "real-time" photon candidate event stream was provided to the collaboration in a detailed internal document, of which the workflow of data handling and manipulation and some details are given here in a brief overview.

- The analysis is thought to run in Malargüe on the computer where also other similar analyses (such as the neutrino analyses) run.
- Needed software: standard unix shell with ssh, g++ compiler, to be compatible with the ROOT version used, ROOT 5.XX, e.g. 5.34.38 (not ROOT 6), and Offline.
- The current SD data file "sd_online.root.nobackup", which is updated every 15 min, will be copied via a shell script to the working directory, to not cause any damage on this original file. The file "sd_online.root.nobackup" is in the CDAS format.
- A shell script will set the environmental variables and then start the analysis described below.
- The mentioned file with the latest SD events will then be reconstructed using Offline and the output file in ADST format (containing the reconstructed event parameters) will be created. A sequence of modules for an SD only reconstruction with Offline (the file ModuleSequence.xcd) is provided in the sec. 6.3.
- The file in the ADST format, containing parameters of the reconstructed SD events (from a previous step), will be rewritten into a ROOT file with a much simpler structure, as needed for an input file for the ROOT TMVA toolkit which will perform the MVA analysis of the event parameters.

  Event selection will be applied at this step. The corresponding quality and MVA classification cuts on events can be changed easily in the future if needed. However, since the goal of the proposed photon candidate data stream to AMON is to select *sub-threshold* events, rather loose cuts will be applied.

  First, events with bad reconstruction will be discarded. Further quality criteria depend on the trigger level, e.g. 4T5, 5T5, 6T5, where it is to be discussed how stringent this should be. (Presented results so far were created with a simple T5 condition.)

  Other cuts arise from limitations due to used variables, as for example the curvature of the shower front is energy dependent, a cut will be applied, so that only events with energies $> 10^{18}$ eV will be used.

- MVA:
  The presented MVA will be applied to the reconstructed event parameters obtained at the previous step. As described in sec. 6.6, the workflow of the MVA will consist of two steps:
  - Training:
    In an optimal outcome of the analysis training is thought to be done once, for example at a computer of the University of Nova Gorica, otherwise on-site (what would cost very little additional time). Simulations of photons will be used as signal and trained against a set of proton simulations, as presented in sec. 6.6.6. A training test will be done on the computer in Malargüe to check for feasibility of the completely autonomous on-site running option.
  - Classification:
    The classification of the SD events will happen on the computer in Malargüe. The rewritten data file, after the cuts, will be combined with the MVA file trained at the University of Nova Gorica. Therefore simulations (the training result) and real data files must be in exactly the same format and combined to one file as input for the MVA. Since the data file containing the events is updated every 15 min, and cuts are applied, it will not consist of a high amount of events, and therefore the step of classification will not take long.
- The MVA creates a new file, containing all the results. Events that have been classified in the MVA as photons (MVA variable > MVA-cut-value) can then with a script be selected and written into a photon candidate event list. The format of this output event list (ASCII text, root, FITS, etc.) as well as the information of selected events to be sent to AMON can be changed easily in the future if needed.

B.3. Settings of tested ML methods

Method	Options
Fisher	VarTransform=None CreateMVAPdfs PDFInterpolMVAPdf=Spline2 NbinsMVAPdf=50 NsmoothMVAPdf=10
FisherG	VarTransform=Gauss
BoostedFisher	Boost_Num=20 Boost_Transform=log Boost_Type=AdaBoost Boost_AdaBoostBeta=0.2 !Boost_DetailedMonitoring
MLPBFGS	NeuronType=tanh VarTransform=N NCycles=600 HiddenLayers=N+5 TestRate=5 TrainingMethod=BFGS !UseRegulator
MLPBNN	UseRegulator *
BDTG	NTrees=1000 MinNodeSize=2.5% BoostType=Grad Shrinkage=0.10 UseBaggedBoost BaggedSampleFraction=0.5 nCuts=20 MaxDepth=2
SVM	Gamma=0.25 Tol=0.001 VarTransform=Norm

Table B.1.: Overview of applied settings for the ML methods used in this work, together with a brief description (further details can be found in [322]). *otherwise same as MLPBFGS.

The description of the options in table B.1 is mostly directly taken from [322]:

"V" indicates the "Verbose flag" (default = true) for the output representation, and since "!V" was used for all methods for a better overview it is

mentioned here.

The booking option "H" indicates help messages, which assist with the performance tuning and which are printed on standard output [322], and was not used only for BDTG.

"Boost_Num" is the number of times the classifier is boosted, where the algorithm is iterated until an error rate of 0.5 is reached or until "Boost_Num" iterations occurred.

"Boost_Type" defines the boosting type for the classifiers.

"Boost_AdaBoostBeta" Defines the AdAboost parameter that sets the effect of every boost step on the events' weights.

"UseBaggedGrad" uses only a random subsample of all events for growing the trees in each iteration and is only valid for "GradBoost".

"Shrinkage" defines the learning rate for the GradBoost algorithm.

"nCuts" is the number of steps during node cut optimisation.

"MaxDepth" is the maximal depth of the decision tree which is allowed.

"NCycles" defines the number of training cycles.

"HiddenLayers" specifies the hidden layer architecture, where "N" stands for number of variables; any integers may also be used.

"NeuronType" defines the neuron AF type, with predefined values "sigmoid", "linear", "tanh", "radial". In this work for all the plots "tanh" was used, only for MLPBFGS in the additional try (sec. A.3) the ReLU function was chosen.

"TrainingMethod" allows values of BP, GA or BFGS, which mean train with Back-Propagation (BP), BFGS algorithm, or Genetic Algorithm (GA - slower and worse).

"TestRate" defines that a test for over-training is performed at each #th epochs.

"UseRegulator" indicates weather or not to use a Bayesian regulator.

"Boosting", which was originally introduced in the TMVA for simple classifiers as the DTs, for the ML method of Fisher cannot enhance the performance in a linear way. This is due to the fact that the Fisher method is equivalent to the linear discriminant (LDA), since the output of a Fisher discriminant represents a linear combination of the input variables, a linear combination of different Fisher discriminants is simply again a Fisher discriminant. Therefore Fisher (and linear discriminants in general) is efficiently boosted not by applying the linear combination on the discriminant's output, but via a fully "non-linear" transformation directly applied to the Fisher classifier response value with a *linear* (= no transformation), *step* (= the output is -1 below the step and +1 above; default), or, as used in this work the *log*, which is a logarithmic transformation of the output [322].

For SVMs "Tol", which is the Tolerance parameter and "Gamma", which is the RBF kernel parameter Gamma, were used.

Missing explanations of options are either assumed to be self-explanatory or can be found in [322].

Acronyms

ADST Advanced Data Summary Tree 89, 93, 132, 193

AERA Auger Engineering Radio Array 73, 74, 175

AF activation function 108, 131, 176, 179, 187, 196

AGN active galactic nucleus 18, 19, 24–26

AI artificial intelligence 96

AMIGA Auger Muon and Infill for the Ground Array 73, 74, 89, 138, 175

AMON Astrophysical Multimessenger Observatory Network 2, 3, 58–60, 96, 99, 131, 132, 137, 193, 194

ANN artificial neural network 107, 112, 176

APE Auger Package Environment 89

Ar Argon 61

B boron 39

BDT boosted decision tree 98, 102, 104, 112, 125, 197

BDTG BDTs using Gradient Boost 98, 99, 114–119, 122–124, 126–128, 130–132, 134, 135, 138, 176–178, 187, 189, 195, 196

Be beryllium 39

BFGS Broyden-Fletcher-Goldfarb-Shanno 98, 109, 126, 128, 178, 179, 187, 195, 196, 200

BH black hole 19, 25, 28, 55

BLF balloon launching facility 74, 175

BNS binary neutron star 28, 29

BoostedFisher boosted Fisher discriminants 97, 126, 128, 178, 195

BS binary system 28, 55, 173

C carbon 40

CART classification and regression tree 101

Cas A Cassiopeia A 24, 25, 173

CBR cosmic background radiation 15, 34, 35, 44, 48, 50–53, 174

CDAS Central data acquisition system 75, 76, 78, 95, 133, 193

CDM cold dark matter 22, 30

CF Cosmic Filament 32

CIC Constant Intensity Cut 87

CL confidence level 16, 48, 57, 58, 99, 172, 174

CMB cosmic microwave background 9, 15, 34–38, 48, 51

CORSIKA Cosmic Ray Simulations for Kascade 71, 72, 89, 91–93, 175

CR cosmic ray 1–3, 5, 7, 9–15, 17, 18, 20, 21, 23, 24, 27, 28, 30, 31, 33–37, 39–47, 50, 54, 55, 57, 58, 60–62, 68, 73, 74, 79–81, 91, 98, 134, 137, 172, 173

DGH down-going high 56

DGL down-going low 56

DM Dark Matter 30

DoF degrees of freedom 83, 125

DPP double pair production 50

DSA diffusive shock acceleration 20

DT decision tree 98, 99, 101–103, 196

DWF Deeper, Wider, Faster 58

EAS extensive air shower 1, 2, 8, 53–57, 61–64, 66–68, 70, 71, 73–75, 78, 79, 81–83, 87–89, 91, 92, 137, 174

EFT Effective Field Theory 53

EGMF extragalactic MF 32, 33, 43, 44

EM electromagnetic 10, 12, 24, 27, 33, 35, 51, 54, 57, 62–72, 74, 75, 78, 84, 87, 89, 90, 137, 174, 175

ES Earth-skimming 56, 57, 174

FADC flash analog–to–digital converter 76–78, 85, 89, 94

FD fluorescence detector 73, 74, 81, 87–89, 175, 176

Fe iron 39, 40

Fisher Fisher discriminants 97, 125, 126, 128, 129, 178, 195, 196

FisherG Fisher discriminant with Gauss-transformed input variables 97, 126, 128, 178, 195

FNR false negative rate 112

FOV field of view 56, 59

FPR false positive rate 112

FR Fanaroff-Riley 26

FRM Faraday rotation measure 33, 34

GC Galactic Center 13, 14, 34, 172

GCN/TAN GRB coordinates network/transient astronomy network 59

Geant4 Geometry and Tracking 89, 93

GMF galactic MF 31, 33, 44

GR General Relativity 52, 55

GRB gamma-ray burst 19, 24, 30, 59, 199, 200

GUT Grand Unified Theory 21, 22

GW gravitational wave 54, 55, 58, 137

GZK Greisen–Zatsepin–Kuzmin 9, 11, 15, 22, 35, 36, 43, 47, 48, 52, 53

H hydrogen 34, 39, 40

HE high energy 36, 37, 47, 50, 51, 53, 55, 56, 58, 70, 92

He helium 39, 40

HEAT High Elevation Auger Telescopes 73, 74, 175

IACT imaging atmospheric Cherenkov telescope 54, 61

ICRC International Cosmic Ray Conference 95

ICS inverse Compton scattering 34, 37, 50, 51

IR infrared 34, 47, 51

ISM Interstellar Medium 26

ISS International Space Station 9

LDA linear discriminant analysis 97, 104, 125, 196

LDF Lateral Distribution Function 86–88, 98, 176, 178, 183

LGRB long GRB 30

LHAASO Large High Altitude Air Shower Observatory 47

Li lithium 39

LIV Lorentz invariance violation 52–54

LPM Landau-Pomeranchuk-Migdal 69, 70, 79

LSC local supercluster 31, 44

MBR molecular bremsstrahlung radiation 62

MC Monte Carlo 65, 91, 98, 99, 138, 176

MDR modified dispersion relations 53

MF magnetic field 2, 17, 26, 28, 31–34, 42–45, 47, 55, 61, 91, 198, 199

MHD magneto-hydro-dynamic 17

ML machine learning 2, 3, 91, 96, 97, 106, 110–112, 116, 119, 124, 125, 131, 132, 138, 176–178, 189, 195, 196

MLP multilayer perceptron 98, 101, 107, 108, 126–128, 176, 178, 179, 187, 195, 196

MLPBNN NN with BFGS training method and Bayesian regulator 98, 125, 129, 178, 195

MM multi–messenger 2, 54, 55, 58–60, 137, 138

MoPS Multiplicity of Positive Steps 77, 78

MoU Memorandum of Understanding 58, 60

MVA multivariate analysis 2, 3, 96–98, 110–120, 122–125, 127, 131, 132, 134, 135, 137, 176–179, 187, 189, 193, 194

N nitrogen 40, 61

NGC New General Catalogue of Nebulae and Clusters of Stars 26

Ni nickel 39, 73, 173

NKG Nishimura-Kamata-Greisen 86, 88, 176

NN neural network 98, 106–109, 131, 187, 200

NSt Number of Stations 111, 117–125, 134, 177, 178

O oxygen 40, 61

PCA principal component analysis 98, 99, 104, 125, 176, 179, 188

PD Photodisintegration 35, 37, 38, 49, 173

PDF probability density function 125, 127, 129, 178

PLD programmable logic device 94

PMT photomultiplier tube 75–78, 85, 89, 95, 120, 182, 202

PP pair production 35–38, 49–51, 53, 70, 71, 173, 174

QCD Quantum Chromodynamics 47, 50

QED quantum electrodynamics 7, 8, 47, 52–54

QFT quantum field theory 52

QG Quantum Gravity 52–54

RD radio detector 89, 138

ReLU Rectified Linear Unit 108, 131, 179, 187, 196

RF random forest 102

RG radio galaxy 26

rms root mean square 44

ROC receiver operating characteristic 110, 112, 114–119, 122–124, 132, 177, 178

S-N-R signal-to-noise ratio 111, 189

SD surface detector 2, 56, 57, 73–78, 80–82, 84–91, 93, 95, 96, 98, 110, 111, 120, 127, 131, 133, 134, 137, 138, 175, 176, 178, 179, 182–184, 193, 194

SHDM Super-Heavy Dark Matter 21, 22

Si silicon 39, 173

SM Standard Model of particle physics 21, 52

SMBH supermassive black hole 18, 24

SNR Supernova Remnant 24, 59

sPMT small PMT 89

SSD surface scintillation detector 89, 138

SVM Support Vector Machine 98, 103, 105, 106, 176, 195, 196

TA Telescope Array 16, 36, 47, 48, 174

TD Topological Defect 21, 22

TDE tidal disruption event 25

Th Threshold 77

TMVA Toolkit for Multivariate Analysis 96, 110, 112, 125–129, 137, 138, 178, 193, 196

TNR true negative rate 112

ToT Time over Threshold 77–79, 175

ToTd Time over Threshold deconvoluted 77, 78

TPR true positive rate 112

UHE ultra–high energy 1–3, 9, 15, 16, 18, 21–23, 25–28, 31, 36, 37, 43, 47–58, 60, 68–70, 91, 98, 134, 137, 138

UHECR ultra–high energy cosmic ray 1, 2, 5, 9, 11, 14–19, 21–26, 29–32, 34–36, 41–45, 47, 48, 50–53, 55, 56, 58, 69, 89, 91, 92, 137

URB universal radio background 51

UV ultra–violet 34, 45, 47, 73

VC Vapnik–Chervonenkis 105

VEM vertical equivalent muon 76, 77, 79, 111, 122

VHE very–high energy 54, 55

WCD water Cherenkov detector 1, 75, 76, 83, 89, 175