# Search for gamma-rays in the EeV sky at the Pierre Auger Observatory using universality

Pierpaolo Savina

Thèse de doctorat

NNT: 2021UPASP041

# Search for gamma-rays in the EeV sky at the Pierre Auger Observatory using Universality

*Recherche de photons de très haute énergie à l'Observatoire Pierre Auger en utilisant l'universalité*

**Thèse de doctorat de l' Università del Salento et de l'université Paris-Saclay**

École doctorale n° 576, Particules, Hadrons, Énergie et Noyau: Instrumentation, Image, Cosmos et Simulation (PHENIICS)
Spécialité de doctorat: astroparticules et cosmologie
Unité de recherche: Université Paris-Saclay, CNRS, IJCLab, 91405, Orsay, France.
Référent: : Faculté des sciences d'Orsay

**Thèse présentée et soutenue à Paris-Saclay, le 21/05/2021, par**

Pierpaolo SAVINA

**Composition du jury**

| | |
|---|---|
| **Reza ANSARI**<br>Professeur, Université Paris Saclay | Président |
| **Yvonne BECHERINI**<br>Professeure, Linnaeus University | Rapporteur & examinatrice |
| **Karl-Heinz KAMPERT**<br>Professeur, Bergische Universität | Rapporteur & examinateur |
| **Stefania SPAGNOLO**<br>Professeure associé, Universitá del Salento | Examinatrice |
| **Eric NUSS**<br>Professeur, Université Paris Saclay | Examinateur |

**Direction de la thése**

| | |
|---|---|
| **Piera Luisa GHIA**<br>Directrice de recherche, IN2P3 CNRS Paris | Directrice de thèse |
| **Lorenzo PERRONE**<br>Professeure associé , Universitá del Salento | Co-directeur de thèse |

université PARIS-SACLAY
ÉCOLE DOCTORALE
Physique et ingénierie : électrons, photons, sciences du vivant (EOBE)

**Titre:** Recherche de photons de très haute énergie à l'Observatoire Pierre Auger en utilisant l'universalité

**Mots clés:** Universalité, Observatoire Pierre Auger, Photons, Rayons Cosmiques

**Résumé:** Les rayons cosmiques d'ultra-haute énergie (énergie supérieure à $10^{18}$ eV) peuvent produire des gammas ultra énergétiques via leurs interactions avec les particules présentes dans les environnements de leurs sources, ou avec les photons de basse énergie constituant des fonds diffus dans l'Univers lors de leur propagation vers la Terre. La détection de tels photons cosmogéniques permettrait de sonder davantage les rayons cosmiques de très haute énergie car leur flux dépend des caractéristiques des sources ainsi que de la nature des noyaux parents. D'autre part, les photons d'ultra haute énergie pourraient également sonder de la nouvelle physique, car leur détection pourrait montrer la présence de matière noire composée de particules super lourdes se désintégrant en photons. Aux énergies les plus hautes, les rayons cosmiques et les photons sont mesurés grâce aux gerbes atmosphériques produites lors de leur interaction dans l'atmosphère terrestre. Les signatures clés permettant de distinguer les gerbes de photons du fond dominant dû aux hadrons sont une plus grande profondeur atmosphérique du maximum de la gerbe ($X_{\max}$) et un nombre plus faible de muons. Ces deux observables peuvent être mesurées à l'Observatoire Pierre Auger, qui combine, dans un instrument hybride, un détecteur de fluorescence (FD) et un réseau au sol de détecteurs de particules (SD). Dans les événements hybrides, le FD mesure $X_{\max}$, tandis que le SD permet d'estimer le contenu en muons. Dans cette thèse, développée au sein de la Collaboration Auger, nous avons conçu une nouvelle variable sensible à la masse, $F_\mu$, liée au contenu muonique de la gerbe, en exploitant la propriété d'universalité des gerbes atmosphériques, en combinaison avec la reconstruction des événements hybrides. Cette nouvelle variable est ensuite combinée avec $X_{\max}$, l'observable sensible à la masse par excellence, dans une méthode d'analyse qui augmente le pouvoir de séparation entre photons et hadrons primaires afin de pouvoir identifier des photons d'énergies au-delà de $10^{18}$ eV.

Dans le chapitre 1, nous discutons de la phénoménologie des rayons gamma d'ultra-haute énergie, en termes de leur production et de leur propagation, ainsi que du principe de leur détection par la discrimination entre les gerbes atmosphériques generées par des photons et celles produites par des hadrons. Le chapitre 2 se concentre sur l'Observatoire Pierre Auger, décrivant ses caractéristiques techniques et une sélection des principaux résultats obtenus en plus de 15 ans de fonctionnement. Dans le chapitre 3, nous expliquons le concept d'universalité des gerbes atmosphériques et comment il est appliqué pour construire un modèle, basé sur cette universalité, des signaux dans les détecteurs du réseau au sol. Le chapitre 4 introduit la première partie de ce travail de thèse: la construction des simulations et de l'ensemble des données qui sont utilisées dans les chapitres suivants. Une attention particulière est accordée à la procédure de simulation et à la sélection des événements. Dans le chapitre 5, nous utilisons le concept d'universalité, en combinaison avec la reconstruction hybride, pour mettre en place une technique permettant de dériver $F_\mu$ à partir du signal enregistré dans les stations SD individuelles. Le potentiel de $F_\mu$ comme discriminateur photon-hadron est également évalué. Dès lors, le chapitre 6 décrit la combinaison de $X_{\max}$ et de $F_\mu$, en utilisant des techniques d'analyse multivariée. Une approche basée sur les données pour l'estimation du fond attendu est exploitée pour dériver la sélection des photons. Enfin, dans le chapitre 7, la sélection des photons est appliquée aux données hybrides. Comme aucun photon d'ultra haute énergie n'est observé, des limites supérieures de leur flux sont obtenues et discutées.

**Title:** Search for gamma-rays in the EeV sky at the Pierre Auger Observatory using Universality

**Keywords:** Universality, Photons, Cosmic rays, Pierre Auger Observatory

**Abstract:** The search for photons of ultra-high energy (UHE), above $10^{18}$ eV, lies in the highest energy range of gamma-ray astronomy, a prominent discipline in multimessenger astronomy, as at high energies it is pursued primarily in the context of the astrophysics of cosmic rays, the progenitors of the gamma rays. The highest energy cosmic rays are expected to produce UHE gamma-rays, either in interactions with the source ambient, or with the soft universal background photons in their propagation to Earth. These *cosmogenic photons* probe ultra-high energy cosmic rays as their flux depends on the characteristics of the sources, as well as on the nature of the parent nuclei. UHE-photons could also probe new physics, as their detection could be a smoking gun for dark matter composed of super-heavy particles decaying in photons. At UHE, cosmic rays and photons are measured through the extensive air showers produced when entering the Earth's atmosphere. The key signatures to distinguish photons showers from the overwhelming background due to hadrons are a larger atmospheric depth at the shower maximum ($X_{\rm max}$) and a lower number of muons. These two observables can be measured at the Pierre Auger Observatory, which combines, into a *hybrid* instrument, a fluorescence detector (FD) with a ground array of particle detectors (SD). In hybrid events, FD measures $X_{\rm max}$, while from the SD the muon content can be estimated. In this thesis, developed in the Auger Collaboration, we have conceived a new mass-sensitive variable, $F_\mu$, related to the muonic content of the shower, by exploiting the air-shower universality property, in combination with the reconstruction of hybrid events. This new variable is then combined with $X_{\rm max}$, the mass-sensitive observable by excellence, into an analysis method that increases the photon/hadron separation power for the search of photons with energies above $10^{18}$ eV.

In chapter 1, we discuss the phenomenology of UHE gamma rays, in terms of their production and propagation, as well as the principle of their detection through the discrimination of photons showers from hadron ones. Chapter 2 will focus on the Pierre Auger Observatory, describing its technical features and some of the key results obtained in more than 15 years of operation. In chapter 3 we explain the universality concept and how it is applied to build a universality-based model of the signals in the detectors of the ground array. Chapter 4 introduces the first part of this thesis work: the building of the simulations and of the data sets that are used in the following chapters. Particular attention is devoted to the simulation procedure, and to the event selection. In chapter 5, we use the universality concept, in combination with the hybrid reconstruction, to set up a technique to derive $F_\mu$ from the signal recorded in individual SD stations. The potential of $F_\mu$ as photon-hadron discriminator is also assessed. Chapter 6 describes the combination of $X_{\rm max}$ and $F_\mu$, by using multivariate analysis techniques. A data-driven approach for the estimation of the background expected is exploited to derive the photon selection cut. Finally, in chapter 7 the photon selection cut is applied to the hybrid data. As no UHE photons are observed, upper limit to their flux are set, and the physics implications are discussed.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France

# CONTENTS

# 1 GAMMA RAYS AS MESSENGERS OF COSMIC RAYS

## Contents



The quest for the origin of cosmic rays intrinsically implies a multi-messenger approach. Due to magnetic fields that permeate the universe, cosmic rays, which are mostly charged ions, do not point back to the sources. Direct information about their acceleration sites can however be obtained by searching for the neutral particles, $\gamma$-rays and neutrinos, generated by the interactions of cosmic rays at the acceleration sites, or during their propagation.

The study of cosmic gamma rays, which is the objective of $\gamma$-ray astronomy, is a prominent discipline in the context of *multimessenger astronomy*, because at high energies, above 100 MeV, it is pursued primarily in the context of the astrophysics of cosmic ray particles, the progenitors of the $\gamma$-rays. This thesis work lies in the highest energy range of this discipline, as it concerns the search for photons of energies above $10^{18}$ eV, in order to probe the origin and nature of the highest energy cosmic rays.

An overview of $\gamma$-ray astronomy is presented in section 1.1, where the detection techniques and the main results are outlined in the whole energy range, from 100 keV up to highest energies.

The next two sections are in turn specialised on ultra-high energy photons. First, their phenomenology is discussed in section 1.2, from production mechanisms to propagation to Earth. Then, section 1.3 illustrates the phenomenon of extensive air showers, that the extremely rare UHE photons produce interacting with the atmosphere, and explains how photon-showers can be distinguished from the overwhelming background produced by charged cosmic rays.

## 1.1 GAMMA-RAY ASTRONOMY: AN OVERVIEW

*Multimessenger astronomy* connects different kinds of observations of the same astrophysical event or system. It is a relatively recent experimental opportunity, connecting traditional astrophysics observations with the new observational windows opened by $\gamma$-ray and neutrino detectors, as well as gravitational wave observatories.

To solve the outstanding problem of the identification of the sources of the cosmic rays, multi-messenger astronomy may not just be an advantage, but rather a necessity, because the directions of cosmic rays, mostly charged particles, are scrambled by the magnetic field of our Galaxy and by the inter-galactic ones. Gamma-rays and/or neutrinos, produced by the interactions of cosmic rays, instead point back to the sources. Either space-borne or ground-based instruments are used to measure these three messengers, $\gamma$-rays, cosmic rays and neutrinos, depending on the their energy. This is illustrated with the position of the cartoons on the top of figure 1.1a with respect to the energy scale at the bottom of the same figure. The middle panel shows examples of sky-maps of gamma-rays, neutrinos and cosmic rays observed at different, increasing energies.

As this thesis work focuses on the (search for) highest energy $\gamma$-rays, this section is devoted to an overview of $\gamma$-ray astronomy. This is a discipline that studies photons with energies ranging from 100 keV up to the highest energies. Table 1.1 shows the classical subdivision of the huge energy range into bands, a classification that depends on the different detection techniques, which in turn are determined by the interactions that photons have with matter. The Earth's atmosphere effectively blocks all gamma-ray radiation: a first broad separation is related to the detecting environment. Up to 100 GeV or so, space-borne techniques are viable: first rockets and balloons, and later satellites, have been used to do $\gamma$-astronomy in this energy range. At larger energies, instead, given the rapidly decreasing fluxes, $\gamma$-rays are studied through ground-based experiments, by exploiting the phenomenon of extensive air showers, which are formations of secondary particles produced from the interactions of $\gamma$-rays with the atmosphere.

| **Band** | **Energy range** | **Detecting environment** |
|---|---|---|
| Low Energy (LE) | 0.1 MeV – 10 MeV | space |
| Medium Energy (ME) | 10 MeV – 30 MeV | space |
| High Energy (HE) | 30 MeV – 100 GeV | space |
| Very High Energy (VHE) | 100 GeV – 100 TeV | ground |
| Ultra High Energy (UHE) | > 100 TeV | ground |

**Table 1.1:** *Conventional classification of gamma rays in energy bands.*

The detection techniques employed in each energy band are related to the interactions that photons have with matter: figure 1.1b shows all the contributions to the photon total cross-section as a function of the energy in a light element (carbon, top panel) and in a heavy element (lead, bottom panel).

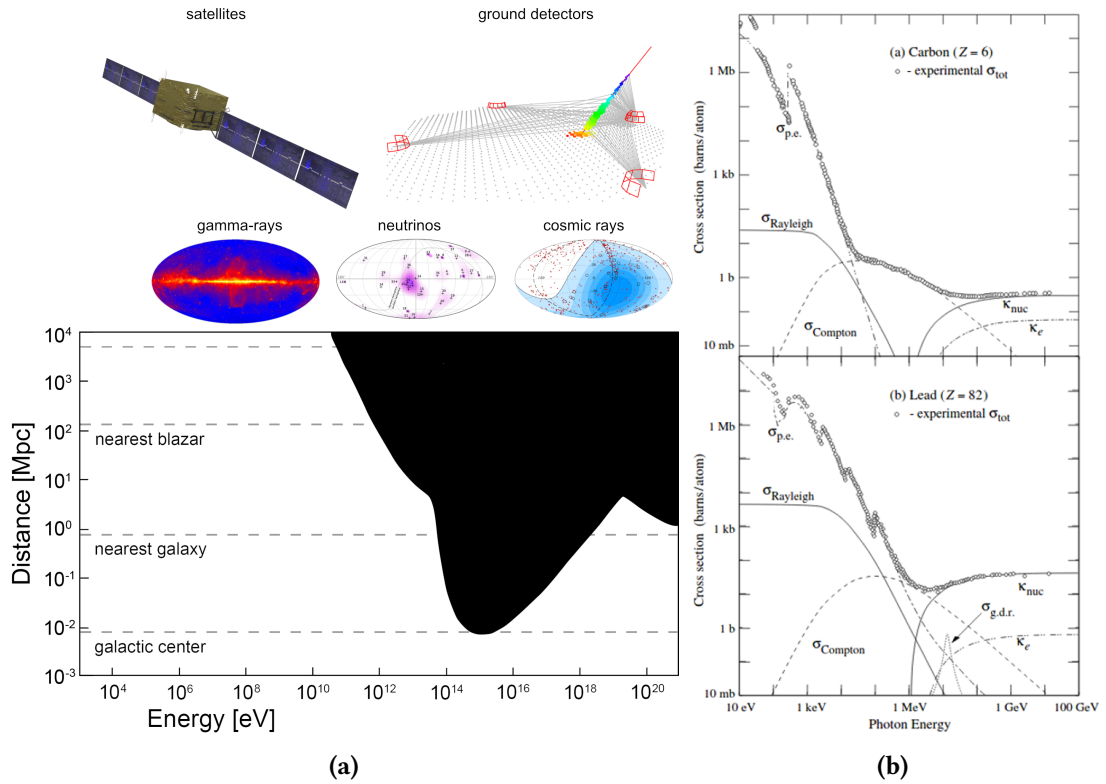



**Figure 1.1:** ***(a)*** *Top: Cartoons of different type of instruments used to probe high-energy astrophysical messengers. Below* $10^{14}$ eV *space experiments are viable, while at higher energies ground and underground experiments are exploited. Middle: sky maps obtained with Fermi-LAT gamma-rays [1], Icecube neutrinos [2] and Auger cosmic rays [3]. Bottom: Schematic view of the mean free path of gamma-rays as a function of energy [4].* ***(b)*** *Photon total cross-sections as a function of energy in carbon (top) and in lead (bottom). The contribution of different processes [5] are also shown: photoelectric effect ($\sigma_{p.e.}$); coherent elastic scattering with atoms ($\sigma_{Rayleigh}$); incoherent scattering ($\sigma_{Compton}$); pair production in the nuclear electromagnetic field ($k_{nuc}$); pair production in the electron electromagnetic field ($k_e$); photonuclear interactions ($\sigma_{g.d.r.}$).*

Below 30 MeV, i.e., in the *low-energy* and *medium-energy* band, the dominant processes are the photoelectric and the Compton effects. These energy regions are mostly explored by Compton telescopes, whose main goal is the measurement of the nuclear lines.

Above 30 MeV, instead, the pair production becomes the dominant process. *High-energy* telescopes, operating in the energy range between 30 MeV and 100 GeV, identify the electron pair in balloon- or satellite-borne detectors.

Above 100 GeV, $\gamma$-rays are studied at ground, using indirect methods, by measuring either the light produced by air shower or the particles forming the showers. *Very-high and ultra-high energy* $\gamma$-rays interact with atmosphere producing an electromagnetic shower of secondary particles, as described in the section 1.3. The secondary charged particles in the shower produce Cherenkov light if their relativistic velocity $\beta > 1/n$, where $n$ it is the refractive index of the air. The

Cherenkov light is emitted at an angle $\theta_C$, with $\cos(\theta_C) = 1/\beta n$. As the refraction index of the atmosphere change with the atmospheric depth, the Cherenkov angle increases, resulting in a enlightened ring-like ground region of radius $\simeq 120\,\mathrm{m}$. The most notable, and successful, $\gamma$-ray technique is the so-called imaging air-Cherenkov which exploits the difference in the shape of the air showers, and hence of the Cherenkov image, between primary photons and hadrons, thereby highly suppressing the background. However, since the $\gamma$-ray spectrum quickly decreases with energy, such technique is effective only up to 100 TeV or so. At higher energies, arrays of particles detectors, that can cover very large areas, are thus exploited. Several methods to discriminate photon- from hadron-showers have been developed with these instruments, at different energies: a brief review will be given in Chapter 5, and the development of one of them is one of the goals of this thesis.

The different $\gamma$-ray bands differ not only in detection techniques but also in the inferences, in terms of the cosmic distances that can be probed. The bottom panel of figure 1.1a shows the mean free path for photons as a function of the energy. At the highest energies, as it will be discussed in section 1.2.2, $\gamma$-rays during their propagation suffer a non-negligible absorption from the emission point to the Earth, due to their interaction with the photons of the cosmic microwave background (CMB) and of the extra-galactic background light (EBL). The figure shows schematically the distances, and hence the objects, that can be probed at different energies. For example, for energies $\sim 1$ TeV, the mean free path is $\sim 100\,\mathrm{Mpc}$, therefore the nearest blazars can be studied. The mean free path then reaches a minimum, of the order of the distance from the Galactic center, around 1 PeV, and then increases again up to few Mpc above 1 EeV.

The understanding of the $\gamma$-ray sky has improved over the last decades. In the following, a selection of experimental results is presented, with reference to figure 1.2 where the $\gamma$-ray sky observed with different instruments, at different energies, is shown.

Following the Hubble space telescope, the Compton Gamma-Ray Observatory (CGRO) was the second of NASA's great observatories to cover the widest interval of the electromagnetic radiation [12]. It was launched using the Space Shuttle Atlantis in 1991 and operated successfully until it was de-orbited on 2000. The CGRO carried four instruments, covering a range from 15 keV to more than 30 GeV. The four instruments were: the Burst And Transient Source Experiment (BATSE), designed to measure low-energy $\gamma$-rays; the Oriented Scintillation Spectrometer Experiment (OSSE) [13] that mapped the 0.5 MeV line from positron annihilation and provided measurements of soft $\gamma$-ray sources; the Compton Telescope (COMPTEL) detected medium-energy $\gamma$-rays. Among its results, COMPTEL mapped the distribution of radioactive Aluminum-26 in the Galaxy, showing the locations of newly formed material [14]; the Energetic Gamma Ray Experiment Telescope (EGRET) was the high-energy instrument on CGRO, covering the energy range from 20 MeV to 30 GeV. EGRET reported the first picture of the entire high-energy $\gamma$-ray sky [7], shown in figure 1.2a. In this map, the Milky way runs horizontally across the figure with the Galactic center lying in the middle. The most prominent feature is the presence of a large background of diffuse photons, in particular along the Galactic plane. Over the diffuse background, some per-

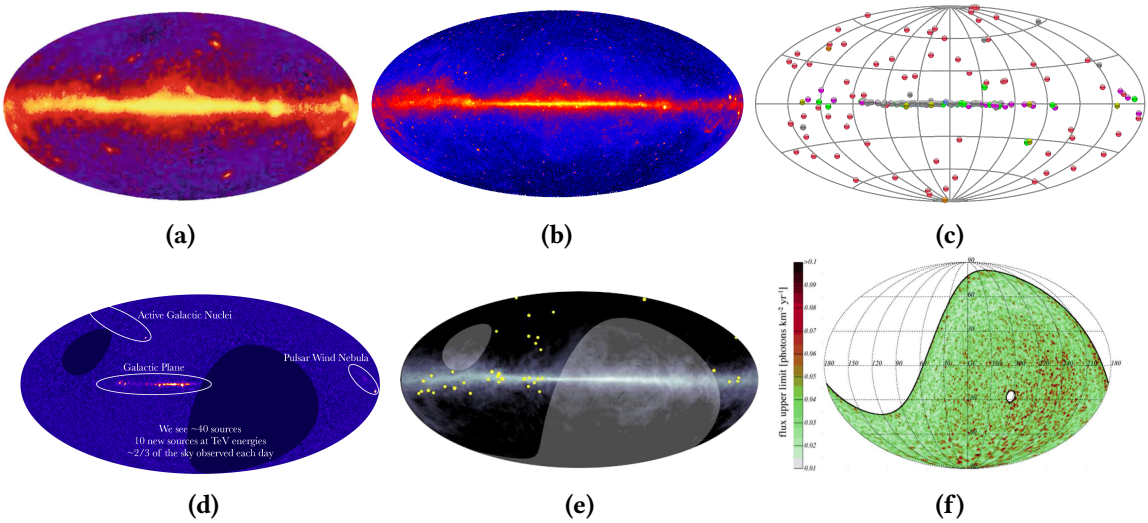



**Figure 1.2:** ***(a)*** *The $\gamma$-ray sky above* 100 MeV *seen by EGRET, shown in Galactic coordinates. From [7]* ***(b)*** *The $\gamma$-ray sky above* 1 GeV *seen by the Fermi-LAT telescope, shown in Galactic coordinates. From [1].* ***(c)*** *Map of the TeV sources, from [8].* ***(d)*** *HAWC sky after 3 years of data acquisition. From [9].* ***(e)****. Tibet AS-$\gamma$ sky, from [10]. Ultrahigh-energy diffuse gamma rays (yellow points) are distributed along the Milky Way galaxy. The background color contour shows the atomic hydrogen distribution in the galactic coordinates. The gray shaded area shows the regions outside of the field of view.* ***(f)*** *Photon flux upper limits (red), measured by the Pierre Auger Observatory. The white region is outside the field of view of the observatory. From [11].*

sistent sources are evident. In the Galactic plane the brightest sources were identified with pulsars. Many of the bright sources away from the Galactic plane are blazars.

After EGRET, a milestone in high-energy $\gamma$-ray astronomy was the launch of the Fermi satellite, instrumented with the Large Area Telescope (LAT) [15]. LAT is a detector designed to measure $\gamma$-rays in the energy range from 20 MeV to more than 300 GeV. A large effective area and a better particle tracking, resulting in an improved background rejection. The sky seen by Fermi-LAT above 1 GeV is reported in figure 1.2b. Like in the EGRET map, the sky is composed of a broad band of diffuse emission along the Galactic plane, stemming from interactions between Galactic cosmic rays and interstellar gas and dust, as well as of various individual sources [1]. Fermi-LAT has extended the range of observations of the diffuse emission of the Galactic disk to several hundred GeV.

Thanks to the advent of the Imaging Air Cherenkov Telescopes technique and to high-altitude (low energy threshold) air shower particle arrays, the very high-energy $\gamma$-ray astronomy rapidly evolved. Imaging Cherenkov detectors are essentially wide-field optical telescopes consisting of a large reflector of about 10 m radius, reflecting the light into a high-speed multi-PMT camera. There are currently three major imaging atmospheric Cherenkov telescope system in operation: the High-Energy Stereoscopic System (HESS) Observatory [16], located in Namibia, with four telescopes arranged in the form of a square of 120 m and a huge central telescope with a 600 $m^2$ area; the Very Energetic Radiation Imaging Telescope Array System (VERITAS) [17], located in Ari-

zona, instrumented with four telescopes; the Major Atmospheric Gamma-ray Imaging Cherenkov (MAGIC) [18], in the Canary islands, consisting of two large telescopes. The sky map, in Galactic coordinates, resulting from the combination of the measurements of these instruments, is in figure 1.2c, where the sources detected above 1 TeV are shown [8]. Different $\gamma$-ray sources, Galactic and extra-galactic, have been observed in the TeV/multi-TeV energy range, namely pulsar wind nebulas (purple), blazars (red), starburst galaxies (orange). Some of them (grey) are still non-identified.

Figure 1.2d shows instead the sky map resulting from the observations of the HAWC (High-Altitude Water Cherenkov Observatory) experiment [9]. Located in Sierra Negra, Mexico, it consists of 300 water-Cherenkov tanks that sample the particles of the air showers produced by the gamma rays in atmosphere [19]. Arrays of particles detectors operate at a higher energy threshold than IACTs, given that they require the shower particles to reach the ground. For that reason, these detectors are typically located at very-high altitudes: the site of HAWC is at an altitude of 4100 m, which allows it to operate in the range between 500 GeV and 10 TeV. The sky map, in Galactic coordinates, is derived from the all-sky search for point sources with index -2.7, by using the data collected between 2014 and 2017. The inner Galactic plane is clearly visible, as well as the Crab and Geminga in the outer Galactic plane. Outside the Galactic plane, Markarian 421 and Markarian 501 stand out [9].

The highest energy gamma-rays have been observed by another high-altitude shower array, the Tibet AS-$\gamma$ experiment, located at 4300 m of altitude in Tibet. The collaboration has very recently reported [10] the detection of a diffuse emission of gamma rays with energies between 0.1 and 1 PeV. Figure 1.2e shows the corresponding sky map, where the blue dots indicate the directions of the observed gamma-rays: most of them are clustered in the vicinity of the Galactic plane (yellow shaded area). This is the first observation of diffuse emission at such energies, at which gamma rays are likely generated by PeV cosmic rays. All $\gamma$-rays above 400 TeV are observed apart from known TeV $\gamma$-ray sources and are compatible with expectations from the hadronic emission scenario in which $\gamma$-rays originate from the decay of neutral pions produced in the interaction of protons with the interstellar medium in the Galaxy.

Gamma-rays of higher energies can also be detected (or at least searched for) by giant instruments like the Pierre Auger Observatory or the Telescope Array [20]. Auger and the use of its data to search for gamma rays at the highest energy are at the core of this thesis and will be discussed extensively in the next chapters. The Auger sensitivity as $\gamma$-ray observatory is illustrated in figure 1.2f, where a sky map of photon-flux directional upper limits (in red) is shown.

## 1.2 ULTRA-HIGH ENERGY GAMMA-RAYS

In this section we focus on the highest energy photons, namely those with energies above $10^{18}$ eV. Like at lower energies, they play a role in the understanding of the highest energy cosmic rays, providing information on their nature, sources and propagation, complementary to those that can be attained by the study of charged cosmic rays. The main mechanisms of ultra-high energy

(UHE) photon production, either in the cosmic-ray sources, or in their travel to Earth, are outlined in section 1.2.1. At such high energies, $\gamma$-rays can interact with the soft background photons permeating the Universe: this fact is discussed in section 1.2.2.

### 1.2.1 UHE-gammas production

UHE photons are expected to contribute to the flux of cosmic rays, due mainly to the decay of neutral pions produced by a *primary process* [21], i.e.,

$$\text{primary process} \rightarrow \pi^0(+\pi^{\pm}) + ... \rightarrow \gamma_{UHE}(+\nu_{UHE}) + ... \tag{1.1}$$

The nature of the primary process depends, on the one hand, on the theoretical model adopted to describe the origin of ultra-high energy cosmic rays (UHECRs), to which UHE photons are connected. On the other hand, the primary process can be related to the propagation of the UHECRs from the source to the Earth.

The different theories and models that explain how cosmic rays gain their energies follow two main approaches: the *bottom-up* approach, where low-energy particles are accelerated to ultra-high-energies, and the *top-down* approach, where super-massive particles are postulated and decay (or annihilate) directly into the UHECRs that are observed at Earth.

In the **bottom-up models** [22–24], the production of the $\pi^0$ is related to the interactions of UHECRs with the source ambient. If the interactions take place in the proximity of astrophysical sources, the observation of UHE $\gamma$-rays would provide a direct signature of the presence in astrophysical environments of nuclei accelerated up to the highest energies. The neutral pions, needed for the photon production, are originated in proton-proton collisions via

$$p + p \rightarrow \pi^0, \pi^{\pm}, K^0, k^{\pm}, p, n, ... \tag{1.2}$$

Because of the similarity with the process of production of secondary hadrons in a fixed-target accelerator experiment, a process like 1.2 is usually referred as an *astrophysical beam-dump mechanism* [25, 26].

A second process that produces $\pi^0$, and that can occur in the proximity of astrophysical sources, is the so-called *photo-pion production*, which is due to high-energy protons interacting with low-energy photons in the surroundings of sources. Around astrophysical sources there is usually a high density of radio, infrared, visible, and ultraviolet photons (*ambient photons*), with which UHECRs can interact. The photo-pion production occurs through the $\Delta^+$ resonance:

$$p + \gamma \rightarrow \Delta^+ \rightarrow \pi^0 + p \tag{1.3a}$$

$$\rightarrow \pi^+ + n \tag{1.3b}$$

The photo-pion production cross-section is two orders of magnitude smaller than the cross section of the beam-dump process, it is therefore important in astrophysical environments where the target photon density is much higher than the matter density.

In turn, in the theoretical models that follow the **top-down approach** [27], the primary process is related to the decay or the annihilation of primordial relics such as topological defects (TD) [28,29], Super-Heavy Dark Matter (SHDM) [31,32] or relic neutrinos [33]. The first ones, the cosmological topological defects, may be produced in symmetry-breaking transitions in the early universe formation. Like more familiar transitions in condensed-matter systems, these may have led to the formation of defects of different type: cosmic strings or vortices, domain walls, monopoles, or combinations of these. In many cases, such defects are stable. If they exist, they would constitute a uniquely direct connection to the highly energetic events of the early Universe [28]. SHDM is described, in some models, as a non-thermal relic[1], with a lifetime much greater than the age of the universe [32]. In such models, the relic particles are clustered as cold dark matter in our Galaxy, and UHE photons, as their decay products, would be observed at Earth with little processing. Finally, in the Z-burst scenario, photons are generated via the resonant production of Z bosons by UHE neutrinos annihilating on the relic neutrino background [33]. In topological-defects and Z-burst models models, UHE photons would be more difficult to detect because they would be injected at larger distance from the Earth. Overall, in all top-down models, the decay products are leptons and quarks. Quarks then hadronize mostly into pions, with the neutral ones leading to predicted fluxes of UHE photons [35] that are, for some of the models, two orders of magnitude larger than that due to the the so-called cosmogenic photons, discussed below. The flux, in fact, depends on a variety of assumptions made, among which the density and lifetime of the postulated particles.

Finally, the other relevant primary processes that generates UHE photons are related to the **propagation of UHECRs** from the sources to Earth. The main interaction mechanism is the so-called Greisen, Kuzmin, Zatsepin (GZK) process [36,37], in which protons interact with the cosmic microwave background photons (CMB). This process is a photo-pion production that occurs if the proton energy is large enough to reach in the centre-of-mass system the resonant production of the $\Delta^+$ hadron, in which neutral and charged pions are generated. This process becomes energetically possible above a threshold energy of about $E_p \simeq 5 \times 10^{19}\,\mathrm{eV}$. For nuclei, the energy threshold for the GZK process is higher, but for these particles, also the *photo-disintegration* processes with soft photons can take place:

$$A + \gamma_{\mathrm{CMB}} \rightarrow (A - nN) + nN \tag{1.4}$$

with $A$ denoting a nucleus with mass number $A$ and $N$ denoting a nucleon. The dominant process is one nucleon emission ($n = 1$). Note that while the interaction with the soft photons of infrared extragalactic background light can be neglected in the case of protons, for heavy nuclei it cannot. Less energy is required for photo-disintegration in the center of mass, as compared to photo-pion production, so that both the CMB and infrared photons are relevant targets for nuclei. As the dominating interaction for heavy nuclei is photo-disintegration, which does not produce photons, and as the produced nucleons are often below the energy threshold for photo-pion production,

[1] A thermal relic is a particle in local thermal equilibrium in the early universe. The thermal relic assumption, with the observed value [30] of the dark matter abundance, constrains the mass of a thermal relic as $m < 100\,\mathrm{TeV}$.

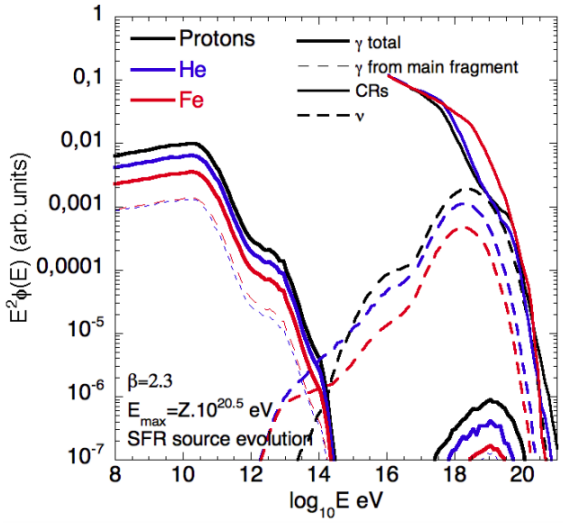



**Figure 1.3:** *Cosmic ray, neutrino (summed over all flavors) and photon spectra assuming three compositions: pure proton (black), pure helium (violet) and pure iron (red) at the source, a source spectral index $\beta = 2.3$ and a maximum energy at the source $E_{max}(Z) = Z \times 10^{20.5}$ eV. The same cosmic ray luminosity between $10^{16}$ eV and $E_{max}(Z)$ is assumed. From [39].*

fewer UHE photons are produced, so that the photon flux depends on the nature of the travelling cosmic ray.

As an example, we show in figure 1.3 the expected fluxes of cosmic rays, neutrinos and photons, when assuming three different compositions of the cosmic rays, namely proton (black), helium (violet) and iron nuclei (red), for a source luminosity distribution following the star formation rate. The same luminosity is injected between $10^{16}$ eV and $E_{\rm max}(Z) = Z \times 10^{20.5}$ eV for the three composition models. As one can see, at ultra-high energies, the expected photon spectra are a factor between about $10^{-6}$ and $10^{-5}$ smaller than that from cosmic rays, depending on their species. The production of secondaries from pure proton and nuclei is different, as the energy transferred to photons (solid lines) and neutrinos (dashed lines) becomes lower as the mass of the parent nucleus increases. It has to be noted that, in this case, a source spectral index $\beta = 2.3$ has been assumed. In fact, the flux prediction for cosmogenic photons is very complex, as it much depends on a variety of factors. If the spectrum is harder, for example, then the predictions would be more similar for proton and nuclei, as the secondary nucleons would be more energetic and would be able to still produce photons via photo-pion production. Overall, generally speaking, and as evinced in other studies (e.g., [38]), a heavy- or intermediate-mass composition leads to approximately an order of magnitude suppression of the photon fraction with respect to a light composition.

### 1.2.2 UHE gamma propagation

The flux of ultra-high energy gamma-rays that might be observed on Earth depends not only on the models for their production, but also on their propagation from the sources to the Earth. Differently from charged cosmic rays, UHE-photons, which are neutral particles, are not subject to any deflections in magnetic fields. However, like charged particles, UHE photons can interact

with photons from the CMB and thus lose energy during propagation, producing electron-positron pairs [40]:

$$\gamma_{UHE} + \gamma_{background} \rightarrow e^+ + e^- + ... \rightarrow \gamma_{\text{GeV-TeV}} + ... \tag{1.5}$$

The threshold energy, $E_{\text{thr}}$, for pair production in the interaction with a background photon of energy $\epsilon$ is:

$$E_{\text{thr}} = \frac{M_e^2}{\epsilon} \simeq \frac{2.6 \times 10^{11}\,\text{eV}^2}{\epsilon[\text{eV}]} \tag{1.6}$$

where $M_e$ denotes the electron mass [41]. Hence, for photons with energy $\approx 10^{19}$ eV propagating through the Universe, background photons with an energy $\epsilon \lesssim 10^{-6}$ eV, corresponding to a radio-frequency of 100 MHz, play an important role. However, the Universal Radio Background (URB) is not well known, mainly because, for measurements of the URB, it is very difficult to disentangle the Galactic component from the extragalactic one [41]. For smaller energies of the primary photon, also interactions with the CMB and the Infrared (IR) background become increasingly important. The energy loss lengths of photons due to interactions with the background radiation fields are shown in figure 1.4. Typical energy loss lengths for UHE photons range between 7 Mpc – 15 Mpc at $10^{19}$ eV [21]. For comparison, also the energy loss lengths of protons are reported.

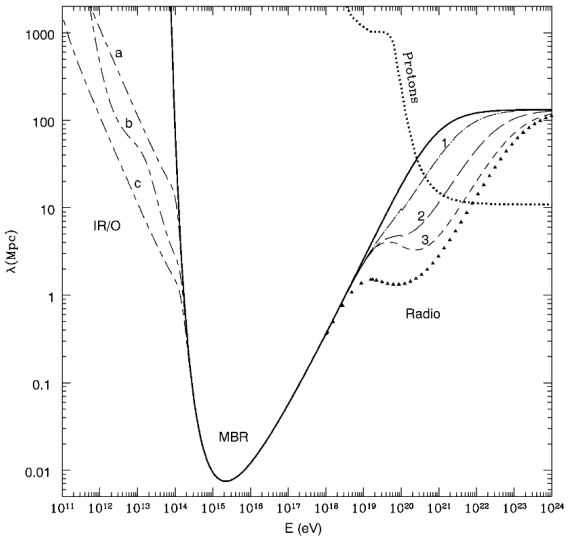



**Figure 1.4:** *Pair production mean free path $\lambda$, for UHE photons as a function of the energy [43]. Below $10^{14}$ eV, $\gamma$-rays interact primarily with Infrared or Optical (IR/O) photons, between $10^{14}$ eV and $10^{19}$ eV they interact with the Cosmic Microwave Background Radiation, above $10^{19}$ eV they interact with radio photons. Curves a, b and c show $\lambda$ for different IR/O background models. Curves 1, 2 and 3 show $\lambda$ for different radio background models. The protons mean free path (dotted line) is reported for comparison.*

The energy distribution between the electron and the positron produced in the pair production process according to equation 1.5 is not symmetric due to the very high center-of-mass energy. One of the particles carries away most of the energy of the primary UHE photon [41]. This leading particle can then undergo inverse Compton scattering processes with background photons:

$$e^{\pm} + \gamma_{\text{background}} \rightarrow e^{\pm} + \gamma_{\text{UHE}} \tag{1.7}$$

In this process, most of the energy of the electron, or positron, is transferred to the upscattered background photon, which can be then considered an UHE photon [41]. Through repeated cycles

of pair production and inverse Compton scattering, an electromagnetic cascade develops. This cascade stops when the photon energies reach the TeV to GeV range, where the Universe becomes increasingly transparent to photons, as shown in figure 1.4 [21]. In addition, adiabatic energy losses due to the expansion of the Universe have to be taken into account. The energy loss length for this mechanism can be estimated to be about 4000 Mpc for the Einstein-de Sitter model of a flat, matter-dominated, Universe, assuming a Hubble constant $H_0 = 75\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ [42].

## 1.3 EXTENSIVE AIR SHOWERS: GAMMAS VERSUS HADRONS

As we have seen in section 1.1, to study gamma rays at energies above 100 TeV, ground-based, large-area detectors need to be used, that exploits the phenomenon of extensive air showers, which is the subject of this section.

When entering the Earth's atmosphere, a high-energy cosmic ray, or a gamma ray, interacts with atmospheric nuclei, producing a cascade of secondary particles that propagates longitudinally at the speed of light along the initial direction of the primary particle. This cascade is what is called an *extensive air shower*. The shower development, as well as its particle content, depends on the nature of the primary particle. In the case of a primary photon, an almost pure electromagnetic shower is generated. In turn, if the primary is a hadron, the collision processes are hadron-dominated and form a hadronic cascade. Muons and neutrinos result mostly from the decay of charged pions and kaons of such cascade, forming a superimposed shower of muons and neutrinos (the latter remain essentially undetected with standard shower detection instrumentation). Neutral pions and, to a lesser extent muons, decay into electromagnetic channels, creating an electromagnetic shower. The two air-shower types are illustrated in figure 1.5, where the secondary products generated by a photon (top) and a proton (bottom) with the same energy, $10^{13}$ eV, are shown. For each primary, the three shower-components are displayed: muonic (left), electromagnetic (middle) and hadronic component (right). Hadron-initiated showers and photons showers, as well as the differences between them, are discussed below.

**Note**. The development of an extensive-air-shower in the atmosphere is described not as a function of the altitude, $h$, but rather in terms of the atmospheric depth $X$. $X$ is a measure of the matter traversed by the particle, and it is calculated as:

$$X = \int_l^\infty \rho(l)\mathrm{d}l \tag{1.8}$$

where the integration is carried out over the atmospheric density profile $\rho(h)$ along the path length $l$ considered.

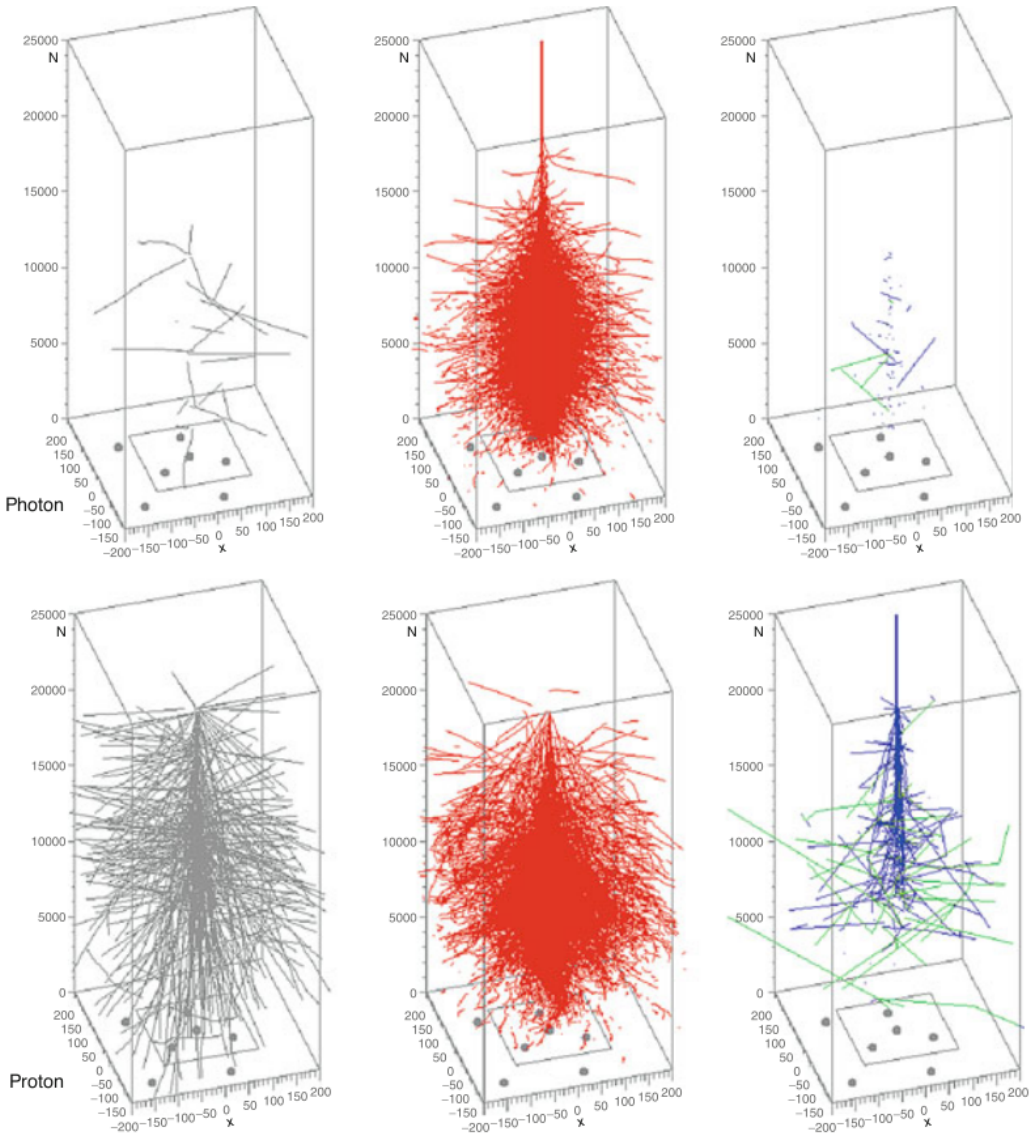



**Figure 1.5:** *Tracks of the secondary particles in vertical air showers initiated by a photon (top row) and a proton (bottom row). For both primaries, the three images show the muonic, electromagnetic and hadronic components, from left to right, respectively [44]. The height of each graph corresponds to an altitude of* 25 km, *while the width corresponds to* 400 m.

### 1.3.1 Photon-induced air showers

Electromagnetic showers generated by high energy photons (either primaries, or originating from the decay of neutral pions) can be mathematically described using differential *transport equations*, which include particle energy losses and production. One of the most notable description was de-

veloped in the 1950s by Rossi and Greisen [45]. Here, the simplified description of the development of an e.m. shower developed by Heitler [46] is used. Two alternating processes that occur in the magnetic field of a nucleus are considered: the pair production, $\gamma \rightarrow e^+e^-$, and the bremsstrahlung emission, $e^\pm \rightarrow e^\pm\gamma$, from an electron or a positron. The latter loses energy also by ionisation. The $e^\pm$ energy-loss can be described as:

$$-\frac{dE}{dX} \simeq \alpha(E) + \frac{E}{X_0} \tag{1.9}$$

where $\alpha(E)$ is the term related to the ionization/excitation energy loss, which depends on the energy, and $X_0 = 37\,\mathrm{g\,cm^{-2}}$ is the mean free path of an electron in air. In turn, the mean free path for pair production by a photon, $\lambda_R$, is given by $\lambda_R = (9/7)X_0$, where the factor $9/7$ is the ratio between the cross-section for bremsstrahlung and for pair production. Thus, each secondary particle produced in a photon-initiated cascade approximately interacts after a distance $d = \ln(2)X_0$ [47]. At each step, the amount of particles in the shower continues to double, until the energy of the secondary particles remains above the critical energy, $E_c = 85\,\mathrm{MeV}$, which is the energy at which the electrons lose energy by ionisation and by bremsstrahlung at the same rate. At this stage, where the energy per particle is $E_c$, the number of particles is maximum, $N_{\max} = E_0/E_c$, where $E_0$ is the energy of the primary photon. The atmospheric depth where $N_{\max}$ is reached can be expressed as

$$X_{\max} = \frac{\lambda_R}{\ln(2)} \ln\left(\frac{E_0}{E_c}\right) \tag{1.10}$$

Therefore, this simplified model predicts that the maximum number of particles is proportional to the primary energy while the depth of the maximum is proportional to its logarithm.

The electromagnetic component of the shower also suffers multiple scattering and its spread can be described in terms of the Molière radius, $R_\mathrm{M}$, which is the characteristic unit of length in the scattering theory of Molière. It amounts to $9.5\,\mathrm{g\,cm^{-2}}$, and it can be expressed in meters, accounting for a direct dependence on air temperature and an inverse dependence on air pressure. It thus increases with altitude and at a depth of $875\,\mathrm{g\,cm^{-2}}$, for example, it is about 80 m. This quantity is important because it characterises the lateral distribution of the electromagnetic particles in the air shower (see section 5.1).

Despite photon-initiated cascades are almost purely electromagnetic, muon pairs can be produced during the shower development. The cross-section for the muon pairs production is however suppressed by a factor $(m_e/m_\mu)^2$, where $m_e$ and $m_\mu$ are, respectively, the electron and the muon masses. The cross-section for photonuclear interactions, which mainly transfer energy to secondary hadrons (and these subsequently to muons), is expected to be more than two orders of magnitude below the pair production cross-section. As a consequence, a muon and a hadronic components are expected in a photon-shower, even if much reduced compared to the electromagnetic component, as it can be seen in the example shown in the top panel of figure 1.5.

At the highest energies, additional processes, which are unique to photon-initiated air showers, can have an effect on the shower development, and, consequently, on the average $X_{\max}$.

Above $10^{18}$ eV, the shower development is influenced by a process identified by Landau, Pomeranchuk and Migdal [48], the *LPM effect*, that takes place at high energies, or at high matter densities. This consists in a reduction of the pair-production and bremsstrahlung cross-sections due to destructive interference from several scattering centers. Thus, the LPM effect delays the development of the shower, leading to a larger $X_{\max}$, because the processes that degrade the energy carried by the shower particles are suppressed. In the case of pair production, the reduced cross section can be approximated by:

$$\sigma_{\mathrm{LPM}} = \sigma_{\mathrm{BH}} \sqrt{\frac{E_\gamma E_{\mathrm{LPM}}}{E_e(E_\gamma - E_e)}} \tag{1.11}$$

where $\sigma_{\mathrm{BH}}$ denotes the cross-section calculated from the Bethe-Heitler formula, while $E_\gamma$ and $E_e$ are, respectively, the energies of the incident photon and of the electron created in the pair production process. The parameter $E_{\mathrm{LPM}}$ can be calculated as

$$E_{\mathrm{LPM}} \sim 7.7\,\mathrm{TeV\,cm^{-1}}\,\frac{\lambda_{\mathrm{R}}}{\rho} \tag{1.12}$$

where $\lambda_{\mathrm{R}}$ and $\rho$ are the radiation length and the density of air [21]. From equation 1.11 it follows that the cross-section is reduced for the creation of an electron-positron pair with the energy equally distributed between the two particles. Therefore, an asymmetric energy distribution is favored in the pair production process. The cross-section for the bremsstrahlung process is suppressed in a similar way [21].

Above $10^{19}$ eV, while nuclear primaries propagate through the geomagnetic field nearly with no interactions, UHE photons may convert in the geomagnetic field into an electron-positron pair, which then emits synchrotron radiation, leading to an electromagnetic cascade above the atmosphere, the so-called *preshower* [49, 50]. Consequently, the electromagnetic particles generated reach the top atmosphere and then, in turn, initiate electromagnetic cascades. Since each electromagnetic particle that reaches the top of the atmosphere carries a fraction of the energy of the original photon primary, the individual showers develop higher in the atmosphere, leading to a smaller $X_{\max}$. The separation of the preshower particles entering the atmosphere is well below current detector resolutions, both in transverse distance and time, so that the subsequent showers are observed as a single air shower event [21]. The local differential conversion probability for a photon of energy $E$ depends on the parameter:

$$\chi = \frac{E_0}{m_e c^2}\frac{B_\perp}{B_c} \qquad B_c \approx 4.414 \times 10^{13}\,\mathrm{G} \tag{1.13}$$

where $E_0$ is the energy of the parent particle, $m_e$ the electron mass, $B_c$ a costant and $B_\perp$ is the local magnetic field component transverse to the direction of the particles motion [21]. Given this dependence on the transverse component, preshower characteristics depend on the arrival direction of the particle.

### 1.3.2 Hadron-induced air showers

As the development of showers initiated by protons or heavier nuclei is characterized by the hadronic interactions, hadron-initiated showers have a more complex description. The Heitler model, however, can be extended to describe hadronic-initiated cascades [51]. The first phase of the cascade production is dominated by the production of baryons and mesons; the number of charged hadrons, $n_{ch}$, produced in a single interaction can be parametrized, as found in the $p\bar{p}$ and $pp$ data collisions [52], as

$$n_{ch} \propto E_0^{0.2} \tag{1.14}$$

where $E_0$ is the energy of the interacting particle in the frame of an external observer. The total number of hadrons, $n_h$, including the neutral ones, is $n_h = (3/2)n_{ch}$. The produced particles are mostly pions. A significant fraction of the total energy is carried away by a single *leading particle*, which uses its energy for a successive interaction after traversing on average one interaction length, $\lambda_I$. This process continues until the hadron energy falls below a threshold, $E_c^\pi$, that typically is between 20 GeV and 30 GeV. Below the threshold energy, the decay length of the pions becomes smaller than the interaction length. The neutral pions thus start to decay in two photons, initiating an almost pure electromagnetic cascade, as described above, while the $\pi^\pm$ decay as

$$\pi^\pm \rightarrow \mu^\pm \overset{(-)}{\nu}_\mu \tag{1.15}$$

Hence, the basic properties of a cascade induced by a primary hadron with energy $E_0$ can be understood by studying the decay channels for neutral and charged pions. A simplified cascade model thus consists of two interrelated processes: the development of a hadronic shower, whose observables are mostly the daughter muons, and an electromagnetic shower. The number of muons can be therefore estimated from the average number of charged particles produced during the shower development, until the shower reaches the critical energy $E_c^\pi$. Assuming that all pions considered decay into muons, the total number of muons in a proton-initiated shower can be calculated as

$$N_\mu^p = \left(\frac{E_0}{E_c^\pi}\right)^\beta \qquad \text{with} \qquad \beta = \frac{\ln(n_{ch})}{\ln(\frac{3}{2}n_{ch})} \approx 0.85 \tag{1.16}$$

Therefore, the number of muons increases almost linearly with the primary energy [51].

Figure 1.6 shows the average longitudinal development of the single shower-components as a function of the atmospheric depth and of the altitude, for vertical protons with a primary energy of $10^{19}$ eV. The electromagnetic component quickly becomes the dominant shower component. Electrons and positrons, in fact, account for about 90% of all the particles in a well-developed hadron-initiated air shower in the lower half of the atmosphere, with photons being even more abundant, as one can see from the figure.

Therefore, also in the case of a hadronic primary, the longitudinal development of the shower is in fact mostly characterized by that of the electromagnetic component. To estimate the atmospheric depth at which a proton-induced shower reaches its maximum, $X_{\max}^p$, we consider the

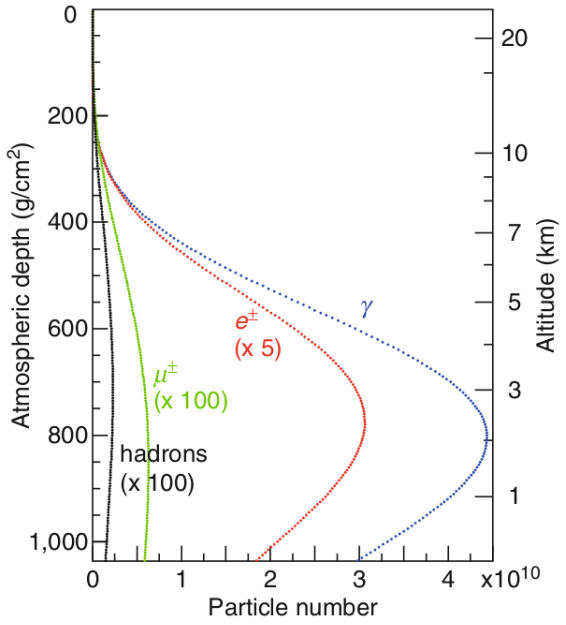



**Figure 1.6:** *Average number of particles as a function of atmospheric depth (left y-axis) and altitude (right y-axis) in simulated showers induced by vertical protons with energies of* $10^{19}$ eV*: hadrons (black, scaled by 100), muons (green, scaled by 100),* $e^{\pm}$ *(red, scaled by 5) and photons (blue) [44].*

cascade as the superposition of many individual showers. The $n_0 = 1/2N_{ch}$ neutral pions, produced in the first interaction, generate $2n_0$ $\gamma$-rays starting the electromagnetic cascade at the same position in the atmosphere. A simple estimate of $X^p_{\max}$ can be derived as

$$X^p_{\max} \simeq \lambda_I + X_0 \ln\left(\frac{E_0}{3n_{ch}E_c}\right) \tag{1.17}$$

To extend this simplified approach from primary protons to nuclei, the *superposition model* is used. This assumes that a nucleus with atomic mass number $A$ and energy $E_0$ is equivalent to $A$ individual nucleons, each having an energy $E_0/A$, and acting independently. The resulting shower is treated as the sum of $A$ individual proton-induced showers, all starting at the same point. According to the simplified Heitler model, the average atmospheric depth of the superposition of all showers is then given by

$$X^A_{\max} \simeq \lambda^A_I + X_0 \ln\left(\frac{E_0}{3An_{ch}E_c}\right) \tag{1.18}$$

The dependence on $A$ implies that, on average, showers initiated by nuclei develop higher in the atmosphere, i.e., they have their maximum at a smaller $X_{\max}$, than showers initiated by protons with the same primary energy. Also the number of muons depends on the mass of the primary particle, as

$$N_\mu = A^{1-\beta} N^p_\mu \tag{1.19}$$

### 1.3.3 Difference between photon- and hadron-induced air showers

The bulk of all air showers that reach ground are produced by primary hadrons. These showers produce a very high level of background events that mask the extremely rare UHE-gamma-induced air showers that are predicted to be, as we have seen above in the case of cosmogenic photons, much less than 1 per mil of all showers, making it very difficult to filter them out. To discriminate gamma-rays from hadrons, it is thus necessary to adopt selection methods that rely on the distinct properties of the showers generated by the two kind of primaries, illustrated above.

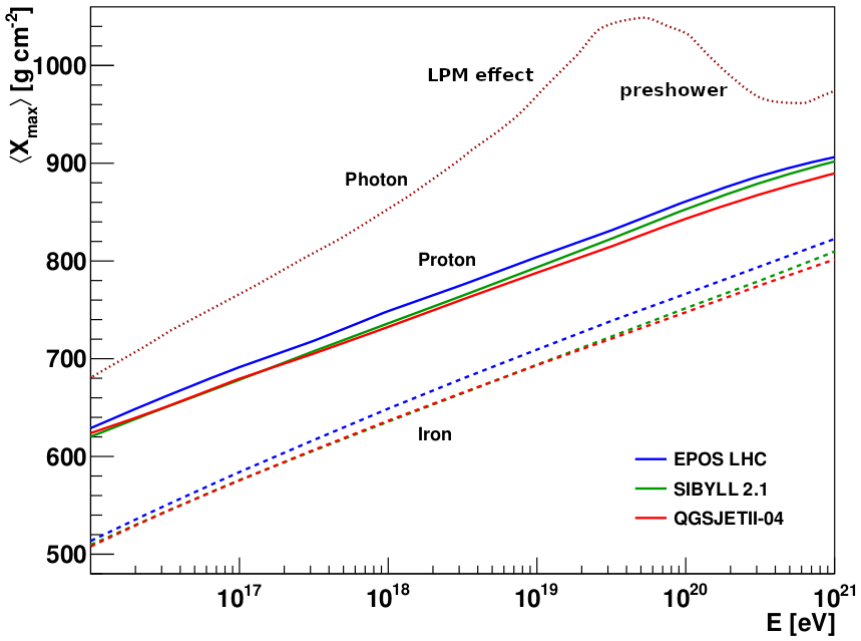



**Figure 1.7:** *Average atmospheric depth of the shower maximum,$\langle X_{max} \rangle$, as a function of the primary energy for extensive air showers initiated by photons, protons and iron nuclei as primary particles (adapted from [53]). For protons and iron nuclei, three different hadronic interaction models have been used: EPOS LHC [54], SYBILL 2.1 [55] and QGSJETII-04 [56]. At the highest energies, the LPM and the preshower effects have been taken into account for primary photons. Due to the preshower effect, $X_{max}$ for energies above about $10^{19}$ eV are also dependent on the location and on the incoming direction of the primary photon. For this plot, the location of the Pierre Auger Observatory in Malargüe, Argentina (see chapter 2), is used.*

The main signatures of a photon-induced shower are due to the fact that it lacks completely the high-energy hadronic core, and thus has a negligible hadronic content. These facts impact the spatial structure of the shower, its particle content and its development. The structure of gamma ray showers is more compact than that of hadron showers, because the electromagnetic cascade products do not acquire on average the large transverse momenta that are typical for hadrons emerging from production reactions. Also, in gamma ray showers the particles are strongly bunched in a thin disk, as they lack the tail of late, trailing hadrons and muons. The two most relevant differences between photon and hadrons showers are, in fact, those on the particle content, namely

the muon content, and on the longitudinal development. On the one hand, since the radiation length is more than two orders of magnitude smaller than the mean free path for photo-nuclear interaction, in photon showers the transfer of energy to the hadron/muon channel is reduced with respect to the bulk of hadron-induced air showers, resulting in a lower number of muons. On the other hand, as the development of photon showers is delayed by the typically small multiplicity of electromagnetic interactions, they reach the maximum development of the shower $X_{\mathrm{max}}$ deeper in atmosphere than shower initiated by hadrons.

The average $X_{\mathrm{max}}$ obtained from Monte Carlo simulations for extensive air showers induced by photon, proton and iron primaries as a function of the primary energy is shown in figure 1.7. Three different hadronic interaction models are used for the simulations of showers generated by protons and iron nuclei. The differences between the predictions of the individual models can be as large as $20\,\mathrm{g\,cm^{-2}}$. Instead, the choice of a specific hadronic interaction model does not influence the $X_{\mathrm{max}}$ for photon primaries, since these air showers are almost purely electromagnetic: the differences in the average $X_{\mathrm{max}}$ predicted by the three models are less than $5\,\mathrm{g\,cm^{-2}}$ for such showers [21]. In turn, the LPM and the preshower effects, as indicated in the figure, are accounted for in the simulations of photon-initiated air showers, because they have an influence on the average $X_{\mathrm{max}}$. The figure shows that there is a large separation between the average $X_{\mathrm{max}}$ in case of photon and proton primaries, for any of the hadronic models considered: the difference is about $60\,\mathrm{g\,cm^{-2}}$ at $10^{16}\,\mathrm{eV}$, and it increases with energy. The separation is, as expected, much larger when the curves relative to the photons are compared to that due to iron nuclei.

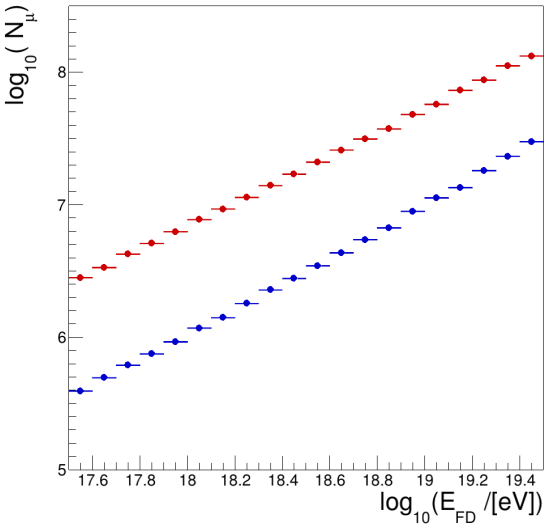



**Figure 1.8:** *Average number of muons that reach the ground as a function of the energy in simulated proton- (red) and photon-initiated (blue) air showers, using EPOS-LHC as hadronic model [54].*

The average number of muons as a function of the energy is, in turn, shown in figure 1.8 for simulated proton and photon showers. The hadronic model used is EPOS-LHC. Also in this case, there is a large separation between the average $N_\mu$ predicted for photon and proton primaries: the difference is of about one order of magnitude in number of muons, and it decreases with energy.

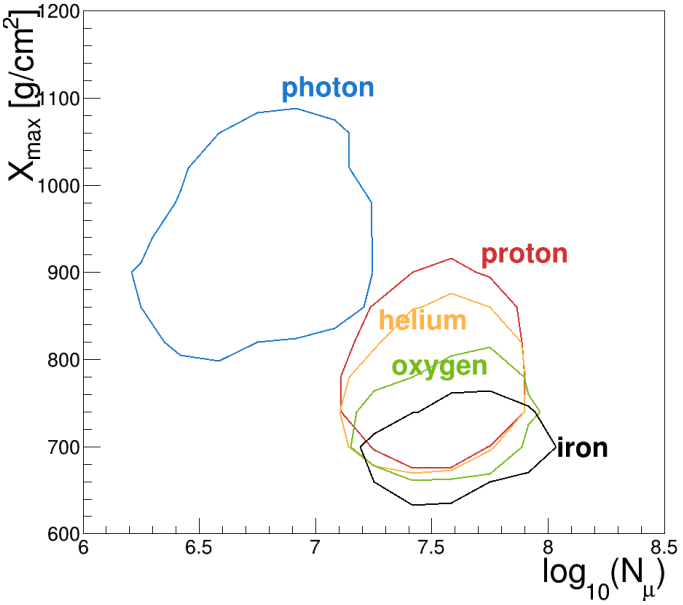



**Figure 1.9:** *$X_{max}$ and $N_\mu$ distributions for air-shower initiated by different primaries, with an energy ranging between $10^{18.5}$ eV and $10^{19.0}$ eV. Contour lines enclose the 90% of the distribution for each primary type. From [57].*

Finally, $X_{\max}$ and $N_\mu$ are shown together in figure 1.9, where the contour plots are obtained using simulated air-showers induced by photons (blue), protons (red), helium (yellow), oxygen (green) and iron (black) nuclei, with energies between $10^{18.0}$ eV – $10^{18.5}$ eV. Each contour includes 90% of the distributions for each primary particle. One can see that photon-initiated showers are well separated from those initiated by hadrons, even accounting for the shower-to-shower fluctuations.

It has to be noted that all figures shown in this section include Monte Carlo quantities, i.e., no detector or reconstruction effects are accounted for. The intent here is, in fact, that of visualising the large photon/hadron separation power of the two variables independently, and the even larger one when combining the two of them. It is also that of preparing the ground for the thesis work, presented in next chapters, the goal of which is the search for gamma rays with energies above $10^{18}$ eV in data from the Pierre Auger Observatory (described in chapter 2) by exploiting the differences in $X_{\max}$ and $N_\mu$ between showers initiated by photons and hadrons.

The search will be performed with the so-called *hybrid* events, i.e, those that have been detected simultaneously with the two main instruments of the observatory, the surface array and the fluorescence detector. A new mass-sensitive parameter, related to the shower muon content, $F_\mu$, will be conceived, which can be derived in hybrid events by exploiting the principle of shower *universality*, illustrated in chapter 3. It will be shown, in chapter 5, how $F_\mu$ can be calculated from the signal recorded by individual surface detectors, profiting of universality-based models that predict the muon content starting from the geometry, the energy and $X_{\max}$ of the shower, parameters that are reconstructed in hybrid events with high resolution. The combination of $F_\mu$ with $X_{\max}$, which will be shown to yield an improved photon/hadron separation with respect to earlier-used methods, will then be used to search for UHE photons with higher sensitivity than in the past.

# 2 THE PIERRE AUGER OBSERVATORY

## Contents



The Pierre Auger Observatory was built to study the most energetic extensive air-showers (EAS) generated by the interaction of the cosmic rays (CRs) in the atmosphere. The main aim of the observatory is the study of the energy spectrum, of the arrival directions and of the chemical composition of the cosmic rays above $10^{17}$ eV [58]. Taking data since 2004, it is located in the province of Mendoza, Argentina, between latitudes 35.0° and 35.3° S and between longitudes 69.0° and 69.4° W. Its mean altitude is 1400 m, corresponding to an atmospheric overburden of 875 g cm$^{-2}$. A map of the observatory is shown in figure 2.1.

The density of the secondary particles that reach the ground is sampled by the surface detector (SD), that is a triangular grid of 1660 water-Cherenkov detectors (WCDs, black dots in the map), operating 24 h per day. 1600 stations are separated by the nearest neighbours by 1500 m, covering an area of $\sim$ 3000 km$^2$. The standard SD array is complemented by a smaller array with the stations separated by 750 m (*infill array*). A description of the surface detector, as well as of the reconstruction of SD data, are presented in section 2.1.

The SD is overlooked by the Fluorescence Detector (FD), which measures the fluorescence light produced in the interaction of the cascades in the atmosphere. It is composed by 27 telescopes, positioned in 4 sites (blue dots in the map) on the boundaries of the array. The FD measures the longitudinal profile of the showers. It is operational only during moonless nights with optimal

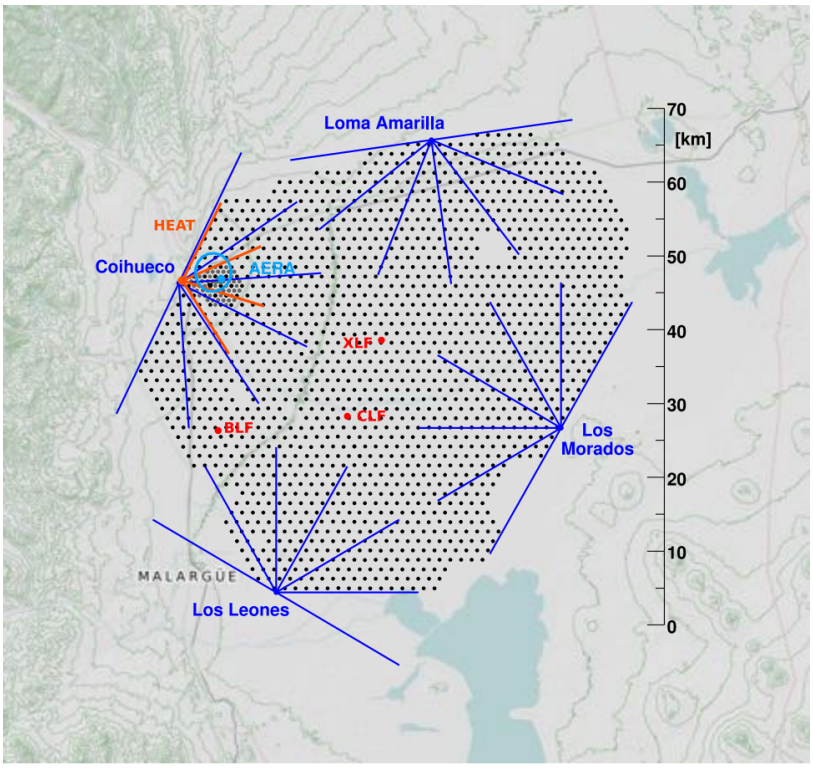



**Figure 2.1:** *Map of the Pierre Auger Observatory. Each black point corresponds to one of the* 1660 *WCD of the surface detector. The four FD sites are represented by blue dots, with the blue lines corresponding to the field of view of their 6 telescopes. The Coihueco site holds also the 3 High Elevation Atmosphere Telescopes (HEAT). The lasers (CLF and XLF) and the Balloon Lauching (BLF) Facilities are shown with red dots. Near Coihueco there is also the AERA site equipped with radio stations. From [59].*

atmospheric conditions, resulting in a duty cycle of about 15 %. The SD and FD are used to observe air-showers in complementary ways, providing important cross-checks and measurement redundancy. In particular, the *hybrid* events, i.e. those measured simultaneously by SD and FD, are reconstructed with an accuracy that is better than that achieved by the two instruments independently. The FD detector as well as the hybrid reconstruction are detailed in section 2.2.

In addition to the shower detectors of the observatory, a comprehensive set of instruments for monitoring the atmospheric conditions above the array has been developed and installed, as varying atmospheric conditions influence the development and the detection of extensive air showers. The instruments dedicated to the atmospheric monitoring are presented in section 2.3.

The observatory is now in a phase of upgrade. Its aim is to extend the measurement of the mass composition at the highest energies, where the FD, the data of which are at the basis of current mass studies, loses statistical power due to its limited duty cycle. The upgrade, shortly illustrated in section 2.4, consists mainly in the installation of scintillators and radio antennas on top of the SD stations.

Finally, a selection of Auger results is presented in section 2.5,

## 2.1 THE SURFACE DETECTOR

The surface detector is an array of water Cherenkov detectors (WCD). A photo of one of them is shown in the left panel of figure 2.2. As one can see from the right panel of the figure, a WCD consists of a 3.6 m diameter water tank containing a sealed liner with a reflective inner surface. The liner contains 12 000 l of ultra-pure water. Three $\sim$ 20 cm diameter photomultiplier tubes (PMTs) are symmetrically distributed on the surface of the liner at a distance of 1.20 m from the tank center-axis and oriented toward the bottom of the tank. Each WCD is autonomous. A solar power system provides the power supply for the PMTs and electronics package consisting of a processor, GPS receiver, radio transceiver and power controller.

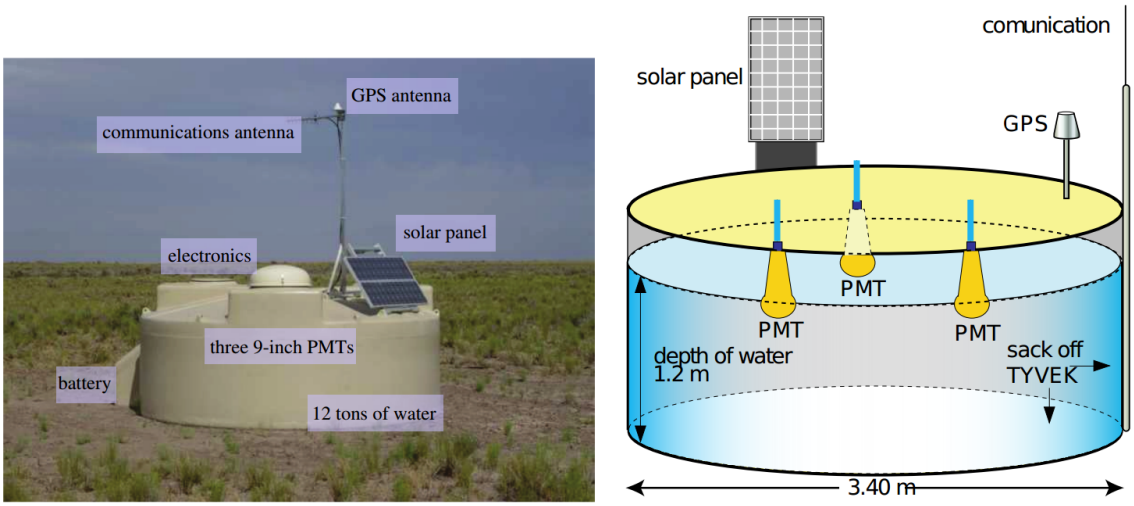



**Figure 2.2:** *Left: photo of a WCD in the field. Right: schematic view of a WCD, showing its main components. From [60].*

The PMTs record the Cherenkov light produced by the passage of relativistic charged particles through the water. Note that the tank height of 1.2 m makes it also sensitive to high energy photons, which convert to electron–positron pairs in the water volume. Each PMT has two outputs: an AC coupled anode signal is provided (*low-gain*) together with the signal at the last dynode (*high-gain*), that is amplified and inverted by the PMT base electronics, thus giving a signal of 32 times with respect to the charge of the anode. Six identical channels of electronics are provided to digitize the anode amplified dynode signals from each of the PMTs. The analog signals are then fed to 10 bit 40 MHz flash Analog to Digital Converter (FADC).

The calibration of the traces collected by the 3 PMTs is explained in section 2.1.1. This is preliminary to the processing of the signals, which is described in section 2.1.2. Air-shower induced

events can be identified with a trigger/selection system described in section 2.1.3. Finally, the reconstruction of the selected events is outlined in section 2.1.4.

### 2.1.1 SD Calibration

The remoteness and large number of the SD detectors require a robust, automatic self-calibration procedure. To achieve it, the atmospheric muons that pass through each WCD at a rate of approximately 2500 Hz are used. The average charge, and amplitude of the signal, measured for a vertical and through-going muon, named the *vertical-equivalent muon* (VEM) is the primary quantity used to calibrate the SD. This, in its normal configuration, has no way to select only vertical and through-going muons; however, the distribution of the light from atmospheric muons produces a peak in the charge and amplitude distributions, which corresponds to those. Examples of charge and pulse amplitude histograms produced in a WCD (black lines) are shown in figure 2.3. The second peak is that due to vertical through-going atmospheric muons, as one can see by comparing its position to that of the red dashed histogram, produced in an external muon telescope selecting only vertical and central muons. The small shift between the two is due to the convolution of the photo-electrons statistics with an asymmetric peak in the distributions of the track-lengths and to light collection effects.

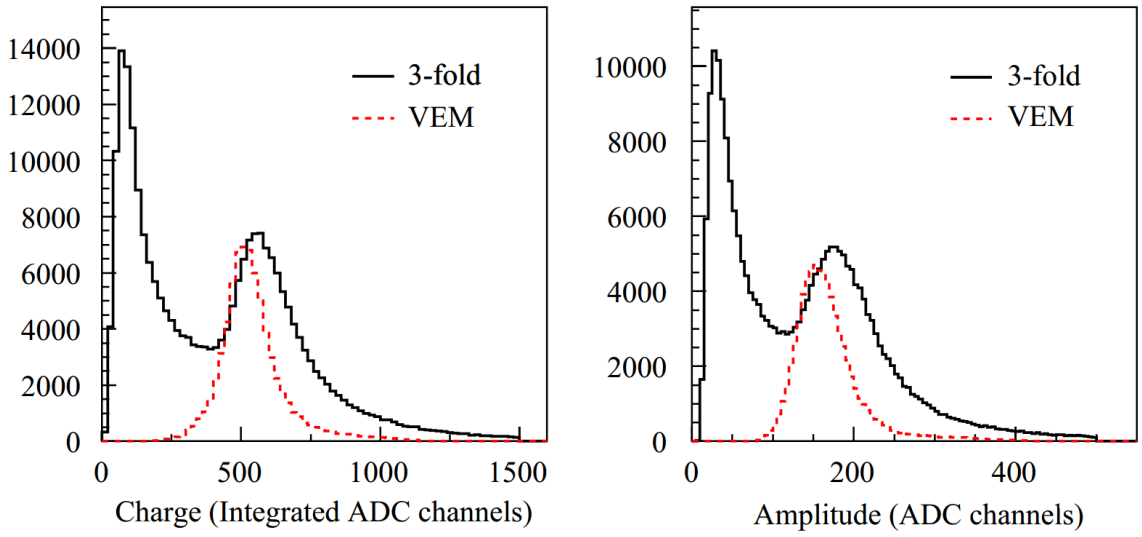



**Figure 2.3:** *Charge distribution (left) and signal amplitude distribution by a WCD (black line). The dashed histograms are produced by an external muon telescope providing the trigger to select only vertical and central muons. The first peak in the black histograms is caused by the convolution of the trigger with a steeply falling distribution from low-energy particles. From [61]*

The peak, $I_{\mathrm{VEM}}$, in the pulse-amplitude histograms, which are continuously built locally in the WCD every minute, provides the common reference unit for the triggers thresholds (see 2.1.3), thus

guaranteeing the trigger uniformity over the whole array. In turn, the peak, $Q_{\mathrm{VEM}}$, in the charge histogram is the quantity used to convert the integrated FADC channels into a physics quantity, common to all WCDs. When a shower trigger takes place, not only the FADC traces of each PMT in each triggered WCD are acquired, but also the charge histograms built in the minute prior to the event. From a fit of such histogram to a parabola with a negative curvature, the value of $Q_{\mathrm{VEM}}$ is derived, for each PMT in each station, and used to calibrate the FADC traces in VEM units, from which, as shown in the next section, the signal size is derived.

Note that, in addition to the conversion from integrated channels to VEM units, the calibration procedure deals with the conversion from the raw FADC traces into integrated channels. The parameters needed for this are the baselines of all six FADC inputs, and the gain ratio between the dynode and anode, $D/A$. The former are computed by recording FADC traces in both anode and dynode channels over a 1-minute interval, and building baseline histograms, which are also acquired with each event. The $D/A$, included too in the data stream, is determined by averaging large pulses, so that both the dynode and anode channels output a signal.

### 2.1.2 SD signal

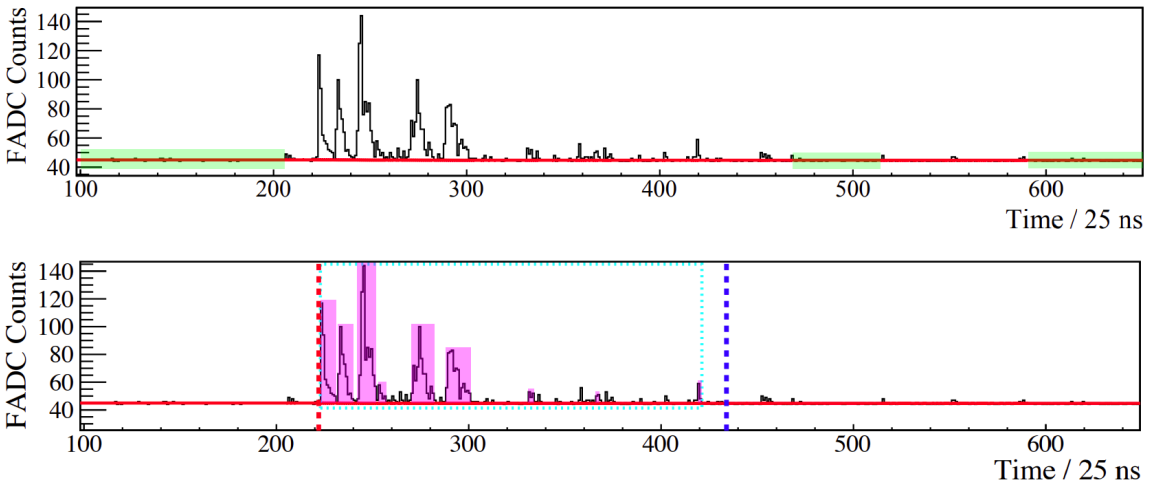



**Figure 2.4:** *Top: illustration of the baseline estimation (red), constructed from the individual segments (highlighted in green) and interpolation, on a high-gain FADC trace of a PMT. Bottom: identification of signal fragments (magenta filled boxes) and the final merged signal window (large dotted cyan box). Based on the number of peaks, this trace probably contains ∼ 5 muons with some additional electromagnetic component. Trace start and stop times are denoted with the vertical dashed red and blue lines, respectively. From [62].*

An example of a high-gain FADC trace recorded from one PMT in one WCD is shown in figure 2.4. The first step in its processing concerns the baselines. Although information on those are acquired on-line from the WCD, on an event-by-event basis, the baselines are in fact determined again off-line, for all six channels, prior to the processing of the signals. This is because the baseline

of a trace can be affected by the signals themselves. The off-line algorithm allows for a changing baseline along the trace. Constant baseline segments are identified as a sequence of bins varying in amplitude by less than that of the electronic noise, i.e., 2 ADC counts. These *baseline segments*, illustrated by the green boxes in the top panel of figure 2.4, are then combined to form the total baseline.

The next step is the extraction of the relevant signal from the traces of individual PMTs. This is done by identifying the start- and stop-time of the signal, performed on the high-gain channel, due to its superior resolution. As for the baselines, the FADC traces are first scanned to identify candidate signal fragments, which consist of consecutive bins with amplitudes of at least 3 ADC counts above the baseline. Examples of identified signal fragments are indicated in figure 2.4 on the bottom panel, with magenta boxes.

After the subtraction of the baseline in each PMT, traces are calibrated in VEM units using the value of $Q_{\mathrm{VEM}}$ extracted from the charge histograms explained in the previous section. The signal size of a station is then obtained by integrating the VEM-calibrated trace, which consists of the bin-by-bin average of the high-gain (or low-gain, if the high gain is saturated) traces of the working PMTs between the previously determined start and stop times. The uncertainty in the signal size is derived according to studies using *twin* stations, i.e., pairs of stations placed at $\sim 11$ m apart and thus sampling the same part of a shower. The uncertainty in the signal $S$ is then modeled with a Poisson-like parametrization, as a function of the zenith angle $\theta$ [63]:

$$\sigma_S = f_S(\theta)\sqrt{S} \qquad f_S(\theta) = 0.34 + \frac{0.46}{\cos\theta} \tag{2.1}$$

### 2.1.3 SD trigger

The SD trigger has a hierarchical structure, which is detailed in figure 2.5. Locally in the WCD, there are two trigger levels, T1 and T2 [64]. The T1 has 2 modes. The first one is a threshold trigger (TH) requiring the coincidence of the three PMT signals above 1.75 $I_{\mathrm{VEM}}$. It is used to select large signals that are not necessarily spread in time, being particularly effective for the detection of very inclined showers that have penetrated through a large atmospheric depth and are consequently dominantly muonic. The threshold has been adjusted to reduce the rate of atmospheric muon triggers from about 2.5 kHz to 100 Hz. The second T1 mode is a time-over-threshold trigger (ToT) requiring that at least 13 bins within a 3 µs window exceed a threshold of 0.2 $I_{\mathrm{VEM}}$ in coincidence for two out of the three PMTs. The ToT trigger selects sequences of small signals spread in time, and is thus efficient for the detection of vertical events, and more specifically for stations near the core of low-energy showers, or stations far from the core of high-energy showers. The rate of the ToT trigger averages to 1.2 Hz. The second trigger level, called T2, is applied to decrease the global rate of the T1 trigger down to about 20 Hz, which is the maximum rate allowed by the bandwidth of the communication system from WCDs to CDAS. While all T1-ToT triggers are promoted to

T2-ToT, only T1-TH triggers passing a single threshold of 3.2 $I_{\mathrm{VEM}}$ in coincidence for the three PMTs pass this second level and become T2-TH.

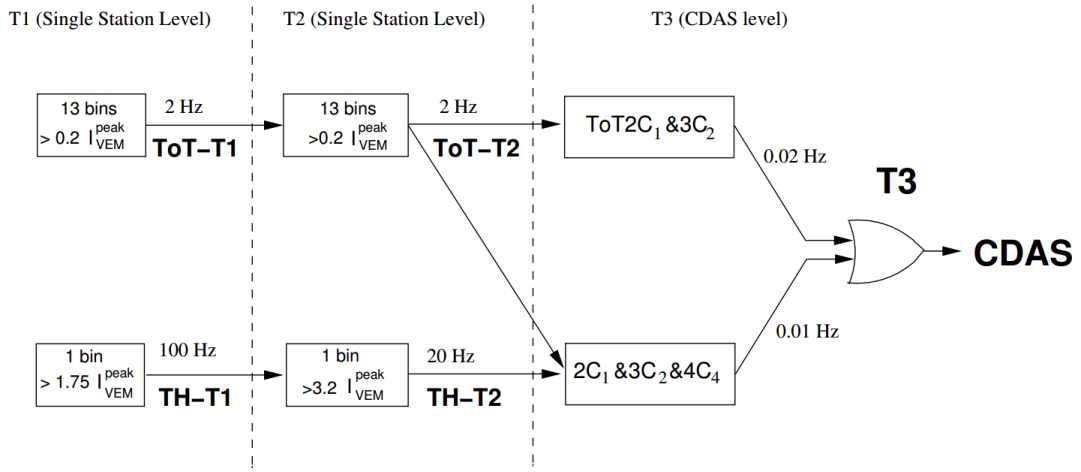



**Figure 2.5:** *Scheme of the hierarchical structure of the SD trigger. The first two steps are local triggers performed in each station. The aim of this chain is to reject the largest fraction of atmospheric events. The T3 trigger is activated when there is a correlation in space and time in the signal of a group of stations. From [64]*

Note that, starting in 2013, two additional T1 triggers [65] were set in operation in the SD array, the so-called deconvolved ToT (ToTd) and the Multiplicity of Positive Steps (MoPS). The first one is basically the same as the ToT but a de-convolution is applied on the trace to account for the exponential decay of the signals. The second one is an algorithm that counts how many positive steps above a certain threshold are present in the trace. They are aimed to further reduce the influence of muons in the trigger, so to make a further reduction of its threshold possible, thus extending the capacities of the SD to measure lower energy showers. As signals from these recent T2s are not used in this work thesis, they are not considered in the following.

All stations passing a T2 trigger send their timestamp to the Central Data Acquisition System (CDAS). The CDAS combines local trigger information to determine an array trigger (T3): if spatial and temporal coincidences are identified, the data acquisition is started. As quite loose temporal constraints are applied at T3 level, a large number of acquired events is still due to chance coincidences. Two higher levels of *triggers* (event selection) are thus applied to the T3 events. A *physics trigger* (T4) selects shower events by requiring that the start times of the signals in the stations are compatible with a plane shower front moving at the speed of light. Finally, a further *fiducial trigger* requires that all the stations around the one with the highest signal are in acquisition. This condition not only ensures an accurate reconstruction of the impact point on the ground, but it also allows for a purely geometrical calculation of the aperture/exposure [66].

### 2.1.4 SD Event Reconstruction

The timing and the size of the signal measured in each station, as well as accurate knowledge of the positions of the stations, are the key inputs for the reconstruction of the arrival directions and the sizes of the showers selected with the T4 criterion. To reconstruct these quantities, a simplified model of air showers is adopted that allows to separate the process into two parts. First, from the timing information of the stations, the geometry of the shower is determined, namely the direction of the shower axis and the position of the impact point of the shower core on the ground. Using this geometry, the second step consists of fitting the signal sizes as a function of distances from the shower axis to an empirically-derived functional form describing the average lateral distribution of particles.

The mean of the station positions weighted with the signal gives a first approximation of the core position. Using the start time of the signals, and considering a plane shower front, the arrival direction of the shower are then calculated. For events with enough triggered stations, these times are described by a more detailed concentric-spherical model, which approximates the evolution of the shower front with a speed-of-light inflating sphere.

The impact points of the air showers on the ground are then obtained from fits of the signals in SD stations. This fit of the lateral distribution function (LDF) is based on a maximum likelihood method which also takes into account the probabilities for the stations that did not trigger and the stations close to the shower axis with saturated signal traces. The saturation is caused by the overflow of the FADC read-out electronics with finite dynamic range and a modification of the signal due to the transition of the PMTs from a linear to a non-linear behavior.

An example of a SD event produced by a cosmic ray with energy $(104 \pm 11)$ EeV and zenith angle of $(25.1 \pm 0.1)^\circ$ is shown in figure 2.6a. The lateral distribution of the signals is shown in figure 2.6b. The function employed to describe the lateral distribution of the signals on the ground is a modified Nishimura-Kamata-Greisen (NKG) function [67, 68]

$$S(r) = S_{1000} \cdot f_{\mathrm{LDF}}(r) \qquad f_{\mathrm{LDF}}(r) = \left(\frac{r}{1000}\right)^{\beta} \left(\frac{r+700}{1700}\right)^{\beta+\gamma}$$

where the signal $S$ as a function of the core distance $r$ is parametrized as the product of the signal expected at 1000 m and the shape of the LDF, $f_{\mathrm{LDF}}$; $\beta$ and $\gamma$ are parameters that give the slope of the LDF [69]. $S_{1000}$ is the parameter used as energy estimator. 1000 m is used as a reference distance because the fluctuations in the expected signal due to a lack of knowledge of the lateral distribution function are minimized [70].

The angular resolution of the reconstructed arrival direction is of the order of $1^\circ$ and approaches $0.5^\circ$ for the largest shower sizes. The resolution of the impact point varies from about 100 m at the lowest energies to about 50 m at the highest energies. The resolution in the reconstructed shower size, $S_{1000}$, improves from about 15%, for the smallest shower sizes down to about 6% for the largest ones.

Due to the atmospheric attenuation, $S_{1000}$ depends on the zenith angle. Assuming an isotropic flux of primary cosmic rays at the top of the atmosphere, the shape of the attenuation curve is

extracted from the data using the Constant Intensity Cut (CIC) method [71]. The zenith angle dependency is accounted using the $S_{1000}$ expected at $\theta = 38$, labeled as $S_{38}$:

$$S_{38} = \frac{S_{1000}}{1 + ax + bx^2 + cx^3}$$

where $x = \cos^2(\theta) - \cos^2(38°)$, $a = 0.980 \pm 0.004$, $b = -1.68 \pm 0.01$, $c = -1.30 \pm 0.45$ [71]. This approach has recently been improved by considering an energy dependent parametrization of the CIC [72].

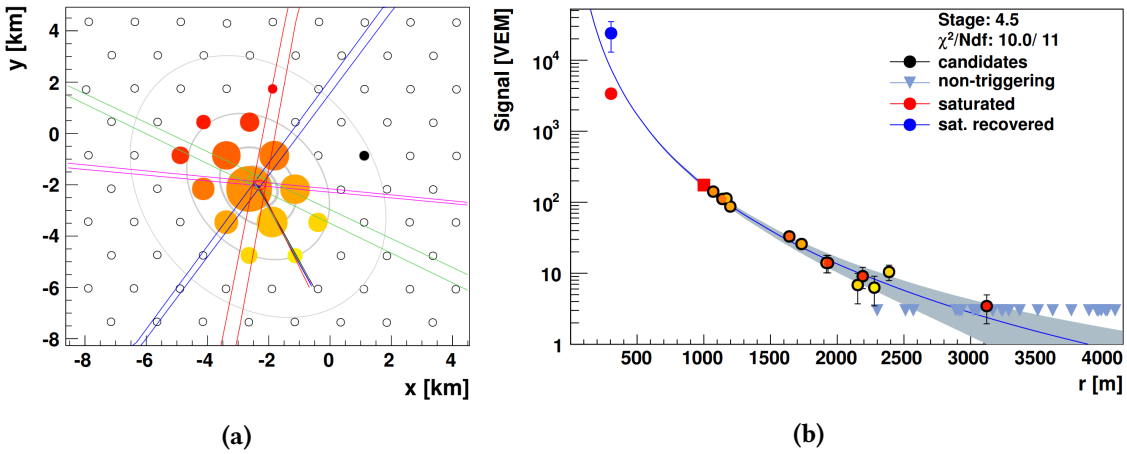



**Figure 2.6:** ***(a)****: SD event display; the circle size is proportional to the logarithm of the signal amplitude, colors (from yellow to red) the arrival directions.* ***(b)****: fit to the Lateral Distribution Function.*

With this procedure, the minimally biased, zenith-independent energy estimator $S_{38}$ is obtained. This can be directly calibrated by the nearly-calorimetric energy measurement of the FD using hybrid events (see section 2.2.3). The power-law calibration curve

$$E[\text{EeV}] = A \cdot S_{38}^B$$

where $A = 0.190 \pm 0.005$ and $B = 1.025 \pm 0.007$ [73] is then used to assign the energy to all SD events.

## 2.2 THE FLUORESCENCE DETECTOR

The FD measures the fluorescence light (wavelength ranging from 300 nm to 430 nm) due to the interaction of the secondary particles, produced during the shower development, with the nitrogen molecules in the atmosphere [74]. The FD is composed of 4 units, located in 4 sites, Los Leones, Los Morados, Loma Amarilla and Coihueco, that overlook the SD array (figure 2.1). Each site (as shown in figure 2.7) contains 6 independent telescopes. At the Coihueco site there are the

3 additional High Elevation Auger Telescopes (HEAT) dedicated to the measuring of low energy showers.

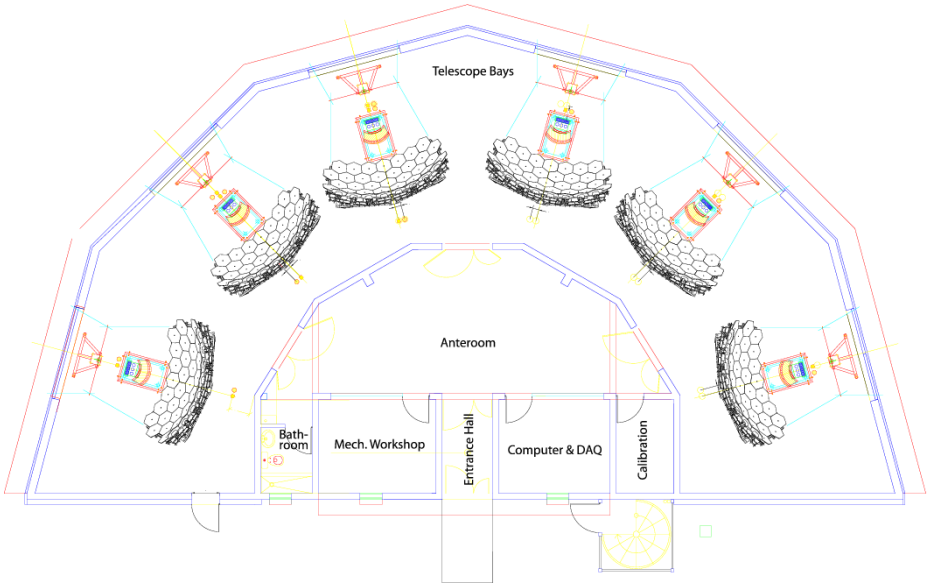



**Figure 2.7:** *Schematic view of an "eye" of the fluorescence detector. From [74].*

A single telescope has a field of view of $30° \times 30°$ in azimuth and elevation, designed to measure the longitudinal development of the shower. The telescopes face towards the interior of the array so that the combination of the six telescopes provides $180°$ coverage in azimuth.

The details of the fluorescence detector telescope are shown in figure 2.8a. The telescope design is based on Schmidt optics because it reduces the coma aberration of large optical systems. Nitrogen fluorescence light, emitted isotropically by an air shower, enters through a circular diaphragm of 1.1 m radius covered with a filter glass window. The filter reduces the background light flux and thus improves the ratio between the signal and noise of the measured air shower signal. It also serves as a window over the aperture which keeps the space containing the telescopes and electronics climate controlled. A corrector ring, divided in 24 sectors, with internal and external radius of respectively 0.85 m and 1.1 m is introduced to increase the field of view of the detector. The shutters are closed during daylight and also close automatically at night when the wind becomes too high or rain is detected.

The light is focused by a segmented spherical mirror of 3.4 m radius of curvature onto a spherical focal surface with radius of curvature 1.7 m. The average reflectivity of cleaned mirror segments at 370 nm is more than 90 %. Alignment of individual mirror segments was cross-checked with a laser on site. Moreover, additional methods using data measured by telescopes were used, such as star tracking, Central Laser Facility (CLF) and eXtreme Laser Facility (XLF) shots (section 2.3), or a comparison of FD and SD geometry reconstruction.

The camera body hosts 440 hexagonal PMTs (also called *pixels*, with a 40 mm diameter and arranged in a $22 \times 20$ matrix. Each pixel has a field of view that corresponds to an angular size of 1.5°. The head electronics for each PMT is connected to a distribution board located just behind the camera body. Each board serves 44 PMTs, providing high and low voltage and receiving the output signals. The signal is then shaped and digitized in the front-end electronics (FE) unit, where threshold and geometry triggers are also generated. Analog boards in the FE unit are designed to handle the large dynamic range required for air fluorescence measurements; this means a range of 15 bits and 100 ns timing. To maximize the light collection, each PMT has 6 light collectors, known as *mercedes*, shown in figure 2.8b. Each mercedes is a plastic structure covered by an high reflectivity material. It has three arms half pixel long, with a triangular section. The mercedes increase the efficiency of the pixels from 50% to 90% (figure 2.8c).

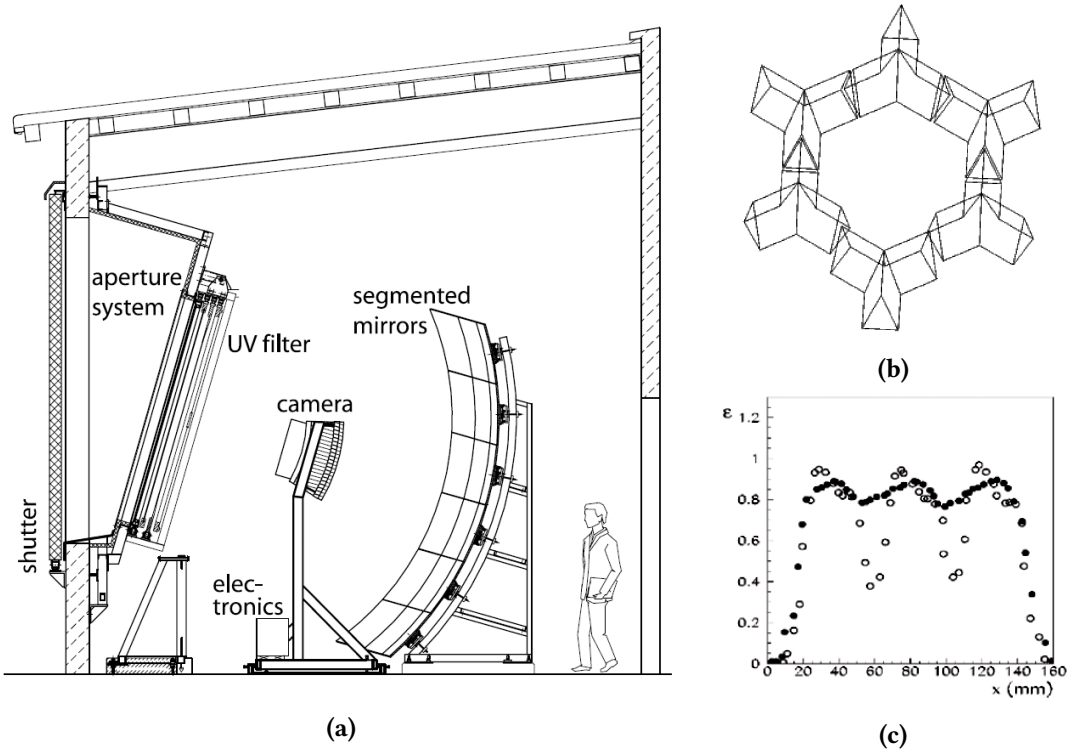



**Figure 2.8:** ***(a)*** *Schematic view of a FD telescope. From left to the right: the aperture system controlled using shutters opened during the data acquisition; the UV filter and the corrector ring; the camera with a matrix of 440 PMT and a segmented mirror. From [74].* ***(b)****: Mercedes scheme, plastic structures that increase the efficiency of the light collection. From [75].* ***(c)****: Light collection efficiency using the mercedes (full dots) compared to the efficiency without mercedes (empty dots). From [75].*

The calibration and the trigger of the FD are presented, respectively, in sections 2.2.1 and 2.2.2. In section 2.2.3, the information from FD and SD are combined to perform the hybrid reconstruc-

tion. This allows for the determination of the cosmic-ray energy with a nearly calorimetric measurement, thus providing the energy scale of the observatory, which is illustrated in section 2.2.4.

### 2.2.1 FD Calibration

The reconstruction of the shower longitudinal profile requires the conversion of the FADC counts to a light flux for each pixel that receives a portion of the signal from a shower. A calibrated large-diameter, drum-shaped light source provides each pixel of the fluorescence telescopes with an absolute, end-to-end calibration . The illumination of each pixel with the same flux of photons allows for the valuation of their response, including the effects of diaphragm area projection, optical filter transmittance, mirror reflectivity, pixel light collection efficiency and area, cathode quantum efficiency, PMT gain, preamp and amplifier gains, and digital conversion.

The drum light source consists of a pulsed UV LED, emitting in a narrow band around 365 nm, mounted in a cylindrical shell illuminating the interior of the 2.5 m diameter cylindrical drum, 1.4 m deep. The front face of the drum is made of a thin sheet of Teflon, which transmits light diffusively. The drum is positioned at the entrance of the telescope under calibration, filling the aperture, providing uniform illumination to each pixel over the full acceptance of the telescope. The drum light source intensity is calibrated [76] using a calibrated photodiode as a reference. Absolute calibration constants are obtained from the ratio of the known pulsed flux of photons emitted by the drum and the corresponding ADC pulse integrals of the camera pixels. From the end-to-end calibration, the appropriate constants are found to be approximately 4.5 photons/ADC count for each pixel. To derive a flux of photons for the observed physics events, the integrated ADC number is multiplied by this constant and divided by the area of the aperture. The flux in photons per $m^2$ perpendicular to the arrival direction is thus obtained. A novel technique for the absolute calibration is currently being developed. The main idea is use a smaller light source and take many measurements at different positions on the aperture. The sum of all measurements is expected to resemble the measurements taken with the drum calibration. The realization of this new method is done with an LED inside of an integrating sphere, which is mounted on a scanner. The scanner consists of two independent linear stages which can move the sphere on the aperture of the telescope.

A relative optical calibration system [74] is in turn used to monitor the long-term time variations in the calibration of the telescopes. The relative calibration is performed at the beginning and at the end of each night of data taking. In each building, three light sources, coupled to optical fibers, distribute light signals to three destinations on each telescope: (i) a pulsed LED light source send signals at the center of the mirror with the light directed towards the camera; (ii) a second light source sends signals from the center of two sides of the camera, toward the mirror; (iii) a third source send signals to ports on the sides of the entrance aperture where the light is directed toward reflective targets mounted on the telescope doors, from which it is reflected back into the telescopes. Drifts of the temporal performance of pixels, mirror and aperture components can be identified by comparing measurements from the three light sources. As an example of application

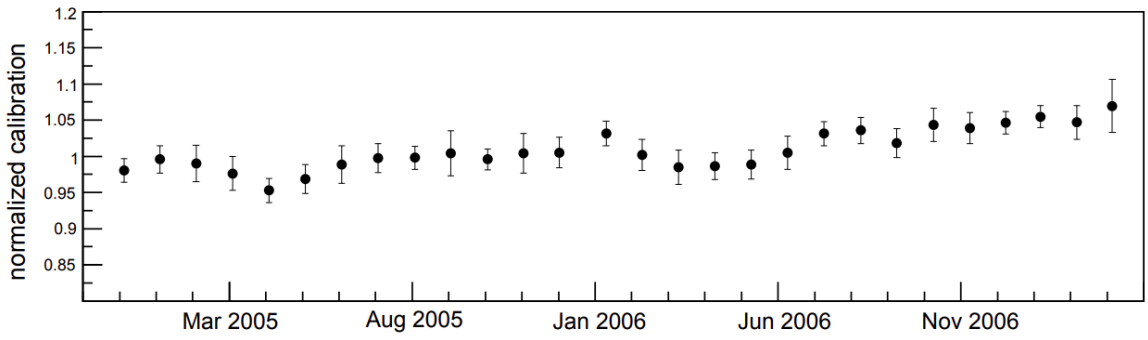



**Figure 2.9:** *Relative calibration factors for camera no. 4 in Los Leones. Each point represents the average over one data-taking shift, which lasts for about 14 night). The error bars represent the 1 s.d. night-by-night fluctuations. From [77].*

of the relative calibration is shown in figure 2.9, where the calibration variations over about two years are shown for one camera. To reduce systematic uncertainties, the relative calibration is used to make corrections to the absolute calibrations.

### 2.2.2 FD Trigger

The fluorescence detector has a multi-leveled trigger. The first 2 levels are implemented in the hardware: the First Level Trigger (FLT) checks the pixel signal while the Second Level Trigger (SLT) checks the geometrical configuration of the triggered pixels. The Third Level Trigger (TLT) instead is a software algorithm aimed to reduce noise events.

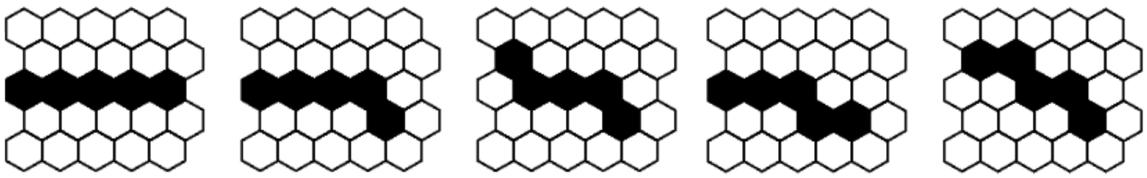


**Figure 2.10:** *Fundamental patterns that define the SLT configurations. From [74].*

The FLT operates on a dynamic threshold that depends on the background light conditions: its aim is that of keeping the rate at the level of about 100 Hz. The background light levels seen by each PMT are monitored through the analysis of the variance of the ADC counts.

The SLT searches in a time window of 20 µs after the activation of the FLT, for track segments of at least 5 pixels in length within a camera. The 5 different fundamental patterns composed by 5 pixel are shown in figure 2.10. Considering the possible rotations and reflections, 108 differ-

ent configurations are probed. The SLT requires 4 FLT pixels out of the 5 included in a specific configuration, and reduces the trigger rate between 0.1 Hz to 10 Hz.

The TLT is a software algorithm designed to clean the air shower data stream of noise events that survive the low-level hardware triggers. It is optimized for the fast rejection of triggers caused by lightning or by muon impacts on the camera, or by randomly activated pixels. The rate of the events selected by the third level trigger is $\sim 0.03$ Hz. TLTs are combined with the SD T3 triggers (section 2.1.3) by the Central Data Acquisition System (CDAS) to form hybrid events.

### 2.2.3 Hybrid Reconstruction

The hybrid reconstruction is based on fluorescence detector data with additional timing information from the surface detector.

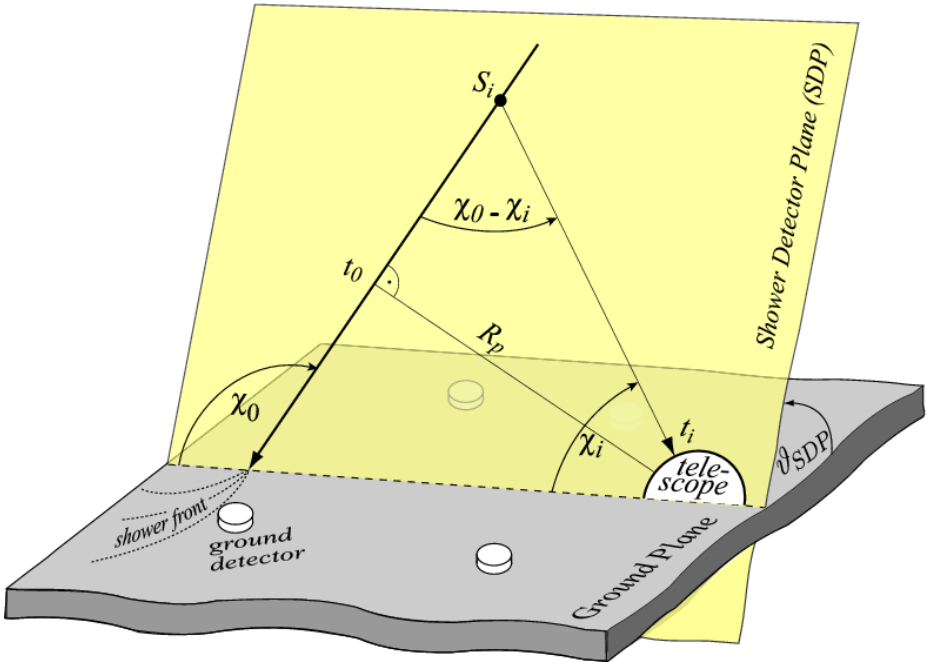



**Figure 2.11:** *Schematic view of the geometric reconstruction of the shower. The Shower-Detector Plane (SDP) is determined using the camera position and the track formed by the triggered pixels. $R_p$ is the perpendicular distance from the camera to the track, $\chi_0$ is the angle between the shower axis with the ground in the SDP, $t_0$ is the time when the shower front on the axis passes the point of closest approach $R_p$. The $i$-th pixel is identified by a trigger time $t_i$, a pointing direction in the SDP with respect to the ground $\chi_i$ and a projected length $S_i$ along the shower track. From [74].*

Firstly, the shower-detector plane (SDP), is calculated (see figure 2.11). The SDP is the plane that includes the location of the FD site and the line of the shower axis. Within this plane a three-dimensional reconstruction of the shower-arrival direction is achieved by determining the geometry from the arrival times of the shower light as a function of viewing angle and from the

time of arrival of the shower front at ground level as measured by the surface-detector station closest to the shower axis. With reference to figure 2.11, for each time bin $t_i$, a vector pointing from the telescope to the shower is defined, and the signals of all photomultipliers (PMTs) pointing to the same direction within a given opening angle, $\chi_i$, are convoluted to reconstruct the overall signal. This angle is determined event-by-event by maximizing the ratio of the signal to the accumulated noise from the night sky background. The angular resolution of the detector operating in hybrid mode is typically 0.6°.

Once the geometry is known, the second step is the conversion of the PMT signals at each time bin into the energy deposited by the shower as a function of slant depth. Every time bin is projected to a path of length $S_i$ along the shower track. The slant depth, $X_i$, is inferred by integrating the atmospheric density through $S_i$. During its path from the shower axis to the FD, light is attenuated due to scattering on air and aerosols. The light emitted on the shower track at time bin $t_i$ can be calculated from the measured light at the aperture corrected by this attenuation factor. The detected photons correspond to different light emission mechanisms and can reach the telescope directly or by scattering in the atmosphere. The light from the shower is composed of fluorescence and Cherenkov photons. The production yield of the former is proportional to the energy deposited by the shower particles within the volume under study, and the latter depends on the number of charged particles above the energy threshold for Cherenkov emission. Fluorescence light is emitted isotropically along the shower track. While high-energy charged particles emit Cherenkov light in a forward-concentrated beam. Even if the shower does not point directly to the detector, a fraction of this beam will be scattered into the field of view. This fraction is calculated taking into account the characteristics of both molecular and aerosol scattering in the atmosphere.

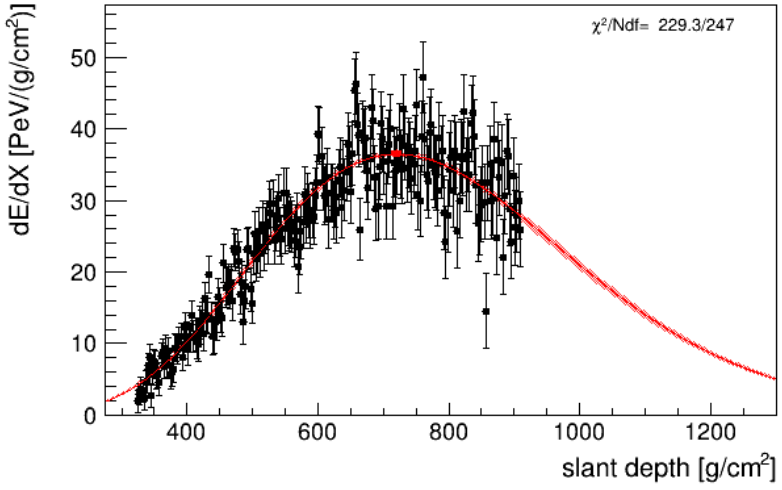



**Figure 2.12:** *Reconstruction of the longitudinal energy deposit $dE/dX$ as a function of the slant depth $X$. The shape is fitted using a Gaisser-Hillas function. From the fit, the depth at the shower maximum $X_{max}$ and the calorimetric energy $E_{cal}$ are obtained. From [74].*

The Cherenkov and fluorescence light produced by an air shower are connected to the energy deposit by a set of equations [78]. By using an estimation of the maximum energy deposit of the longitudinal profile $(dE/dX)_{\max}$, the normalized energy deposit profile, defined as

$$(dE/dX)' \equiv (dE/dX)/(dE/dX)_{\max}. \tag{2.2}$$

can be described by a Gaisser-Hillas function [79], written as a function of parameters $R$ and $L$ [80],

$$(dE/dX)' = \left(1 + R\frac{X'}{L}\right)^{R^{-2}} e^{-\frac{X'}{RL}} \tag{2.3}$$

where $R = \sqrt{\lambda/|X_0'|}, L = \sqrt{|X_0'|\lambda}, X_0' \equiv X_0 - X_{\max}$ and $\lambda$ is a shape parameter. In this notation, the Gaisser-Hillas function is a Gaussian with standard deviation $L$, multiplied by a term that distorts it, with the asymmetry governed by $R$. Thus, the profile of the the energy deposit as a function of slant depth is finally obtained by fitting the number of photoelectrons detected in the PMTs using the equation 2.3. The proportionality between the number of fluorescence photons and the energy deposit is given by the fluorescence yield [81], which depends on the molecular properties of the atmosphere, while the statistical uncertainty is calculated from the Poisson uncertainty of photoelectrons detected by the PMTs. The integral of the Gaisser-Hillas function gives the calorimetric energy $E_{cal}$ of the shower, that must be corrected for the *invisible* energy carried away by neutrinos and muons [82]. The invisible energy, $E_{\text{inv}}$, is calculated by using a data-driven approach, described in [83].

The invisible energy estimation is based on two alternative data-driven methods, the first using "vertical" events ($\theta < 60°$) and the second one using inclined events ($60° < \theta < 80°$) [83]. The method developed on inclined showers uses measurements of the muon number at ground level. In these showers, the electromagnetic component is largely absorbed by the atmosphere and the signal in the SD detectors is dominated by muons. In vertical showers the invisible energy, $E_{\text{inv}}$, is calculated through its correlation with $S_{1000}$, Finally for both methods, a parametrization of the invisible energy as a function of the calorimetric energy is provided. The two methods are compared and give consistent results within uncertainties. The advantages of using a data-driven estimation of $E_{\text{inv}}$ are that simulations are known to be deficient in their estimation of the muon content of air showers [84] (critical to the invisible energy), and that the estimation naturally takes into account the evolving mass composition of the cosmic rays with energy.

### 2.2.4 Energy Scale

The energy scale of the Pierre Auger Observatory is based on FD observations, as they provide a near-calorimetric measurement of the primary energy. This is transferred to the surface detector through the hybrid measurement of air showers (see section 2.1.4).

The main sources of the 14% systematic uncertainty associated with the FD energy [85,86] are reviewed here, and are shown in table 2.1:

| | |
|---|---|
| Fluorescence yield | 3.6 % |
| Atmosphere | 3.4 % - 6.2 % |
| FD calibration | 9.9 % |
| FD profile reconstruction | 6.5 % - 5.6 % |
| Invisible energy | 3.0 % - 1.5 % |
| Statistical error of the SD calibration fit | 0.7 % - 1.8 % |
| Stability of the energy scale | 5.0 % |
| Total | 14.0 % |

**Table 2.1:** *Contributions to the systematic uncertainties in the energy scale. See text for more details. From [85].*

- The parameters characterizing the fluorescence yield, as measured in the AIRFLY experiment [87]. They include the absolute normalization of the wavelength spectrum, the relative intensities in different spectral bands, and their dependencies on pressure, temperature and humidity. The total energy systematic associated with the fluorescence yield is 3.6%.

- The atmosphere. Energy systematics from the atmosphere include those associated with scattering of light by molecules and aerosols, and those connected to the quenching of fluorescence light at the source. The total systematic ranges between 3.4% and 6.2% (smaller uncertainty at lower energies), dominated by the uncertainty in aerosol optical depth.

- The FD calibration. This contributes 9.9% to the total systematic. It includes uncertainties in the absolute calibration using the drum method and the nightly relative calibration done with fixed telescope-based light sources. Uncertainty in the wavelength-dependence of the telescope efficiency (including filter, lens, mirror and camera) is also included.

- The profile reconstruction. Several uncertainties contribute to an energy systematic of 6.5% – 5.6%, slightly larger at lower energies. The main contributor is an uncertainty in light collection, given that the image spot is a convolution of the optical point spread function and the finite width of the shower image. Also included are a small systematic associated with the model for multiple-scattered light, and a contribution to account for systematics from the constraints placed on parameters of the Gaisser-Hillas function, mainly important for close-by showers where a smaller range of atmospheric depths is viewed by the FD.

- The invisible energy. Analysis of the systematic uncertainties on the invisible energy [72] shows an uncertainty in the total energy which decreases with energy from 3 % to 1.5 %.

The stability of the energy scale is monitored through the ratio between the reconstructed FD energy $E_{FD}$ and the energy estimator $S_{38}$ (see section 2.1.4) as a function of time. Its long term behaviour is illustrated in figure 2.13. A clear seasonal modulation is visible, that peaks during the Southern hemisphere's winter months. Also, a downward long term drift is observed. The seasonal modulation shown in figure 2.13 is related to the variation of the temperature inside the

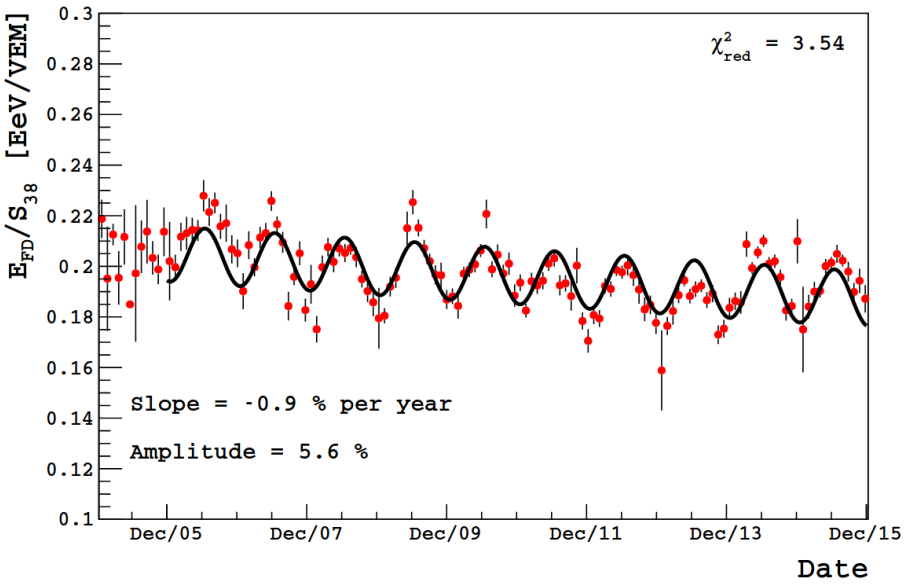



**Figure 2.13:** *Average ratio between the FD energy and $S_{38}$ as a function of time, for hybrid data independently reconstructed with the SD. The black line is a fit to the sum of a linear and a sinusoidal function. From [88].*

FD buildings (the plot is instead corrected for the seasonal modulation associated to $S_{38}$), while the long term drift is due to a combination of aging effects from both the SD and the FD. A related systematic contribution of 5 % is part of of the systematic budget to account for stability of the energy scale over time.

## 2.3 ATMOSPHERE MONITORING

For the indirect detection of cosmic rays through extensive air showers, the atmosphere is in practice used as a giant calorimeter. Its variations have an impact both on the shower development and on the response of the detectors: an extensive program to monitor the atmosphere is implemented at the Auger site.

The knowledge of atmospheric state variables is a fundamental ingredient for the cosmic ray event reconstruction. The state variables influence the development of extensive air showers and the amount of light emission, but they are also needed for the analyses of aerosol and cloud measurements. A simple and robust measurement of the state variables is done with ground-based weather stations. Several **weather stations** are operated, at each FD site and at the Central Laser Facility site: they record temperature, pressure, humidity, and wind speed every 5 min.

Aerosols and clouds represent the most dynamic monitoring challenges for the observatory. Installations to determine the optical scattering and absorption behavior of the atmosphere, as

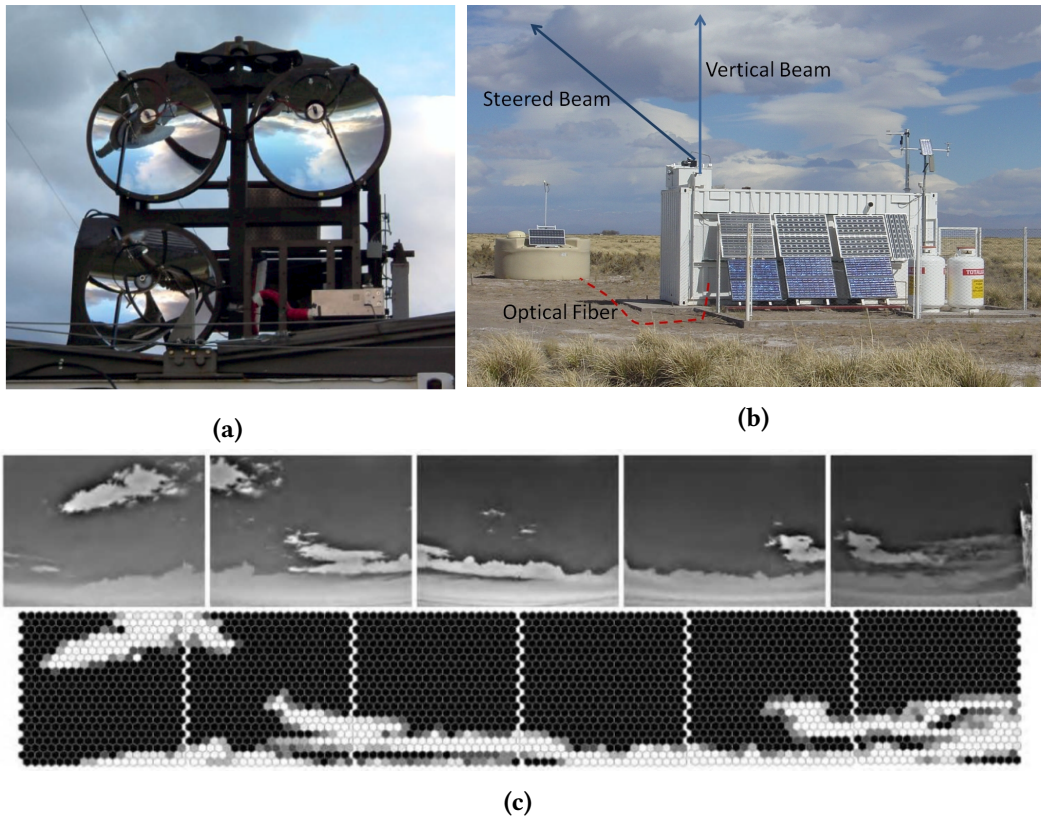



**Figure 2.14:** ***(a)****: Lidar station at Los Leones.* ***(b)****: Central Laser Facility.* ***(c)****: Clouds observed by the IR cloud camera. Up: raw image. Bottom: pixel mask for each telescope of the FD-site.*

well as to identify clouds in the field of view of the FD, are operated. The facilities are shown in the overview of the observatory in figure 2.1.

At each FD site, besides the weather stations, there are a lidar station and an infrared camera for cloud monitoring [89]. Moreover, two Laser Facilities, called CLF (shown in figure 2.14b) and XLF, are located at the center of the array: they are used to measure the aerosol contamination along the line of sight of each FD eye.

The **lidar** (shown in figure 2.14a) is the principal detector used for the study of the atmospheric conditions. It measures aerosols and clouds in the atmosphere using the back-scattering light fired with pulsed UV lasers into different directions. Each lidar is instrumented with 3 parabolic mirrors that focus the back-scattering light in 3 PMTs. The **CLF and XLF** emit UV laser pulses at 355 nm every 15 min during the data-taking at several angles and energies and can be observed from the FD. This laser pulses are characterized by a particular timestamp.

An **infrared camera** sensitive to the temperature differences between clouds and clear sky is mounted on the roof of each FD site (figure 2.14c). The cameras scan the field of view of the FDs every 5 minutes, and also generate a full sky scan every 15 minutes. However, the cloud

cameras are not able to detect the absolute distance of clouds. Information from the CLF, lidars and cloud-cameras are stored in several databases for crossing all the available information in the event selection procedure.

## 2.4 UPGRADE

The huge collecting area and the hybrid detection strategy of the Auger observatory have yielded important advances in the measurements of UHECRs, as it will be shown in the next section. The data of the observatory have also largely affected the traditional intuition on UHECRs, in particular suggesting that the most energetic cosmic rays are mostly heavy nuclei [90]. However, the FD, on whose data the mass measurements is based, loses statistical power above 40 EeV, due to its limited duty cycle. To extend the study of mass composition at the highest energies with a large statistics, an upgrade of the SD is ongoing, aimed at improving the measurement of the shower muonic-component, which is the principal mass-sensitive observable for EAS arrays.

The key element of the upgrade is the installation of plastic scintillator detectors (SSD), each seen by a 1.5-inches PMT, on top of the SD stations, as shown in figure 2.15. Since the SSD and the WCD have different sensitivity to muons and electromagnetic particles that reach the ground, the combination of the two independent measures will allow for the reconstruction of the two shower components.

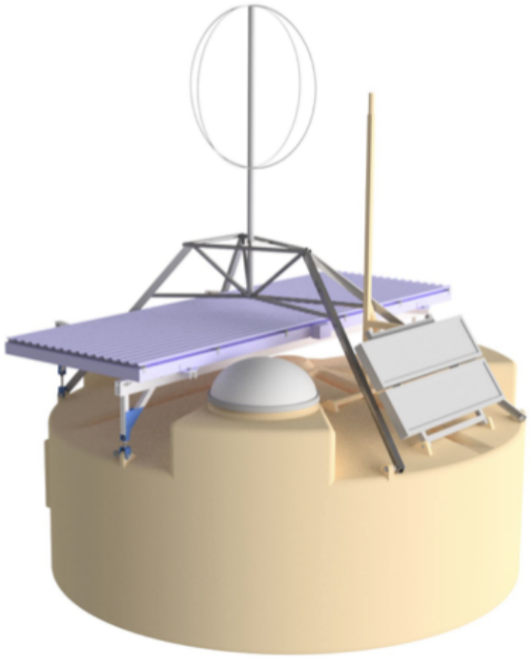


**Figure 2.15:** *3D view of the upgraded SD station. From [91].*

The upgrade of the SD is also the occasion to modernize its electronics. To increase the dynamic range of the SD stations, a fourth small PMT is foreseen in the WCD. With an active area of about 1/80 with respect to the large WCD PMTs, it potentially allows for an equivalent dynamic range extension, which is particularly advantageous for the measurement of very large showers, as those produced by the highest energy cosmic rays, as well as the sampling of the signal near the shower core. The upgrade of the SD electronics is not only adapted to read-out the two additional PMTs, but also to process faster the signals, namely at 120 MHz. The local-station trigger- and processing-capabilities are also increased, by using a more powerful local-station processor and FPGA (field-programmable gate array), allowing for the implementation of more complex trigger algorithms.

The upgrade of the observatory includes also the deployment of an additional instrument, dubbed Auger Muon detector for the Infill Ground Array (AMIGA), designed for the direct measurement of the muonic component of the shower in the infilled area of the SD. Besides addressing the study of mass composition at low energies, around 0.1 EeV, these muon measurements will also serve as a verification of the methods envisioned to extract the muonic signals from the combination of SSDs and WCDs.

Finally, an additional aspect of the upgrade is the installation, on the top surface of each SD station, of a circular loop radio antenna (RD) – also shown in figure 2.15 – operating in the frequency range from 30 to 80 MHz. As the radio signal emitted by air showers travels unimpeded through the atmosphere, the antennas can measure the electromagnetic component of horizontal showers. The combination RD-WCD thus well complements the SD-SSD one, which is in turn used to analyse the particle contents of the vertical showers.

## 2.5 RESULTS OF THE PIERRE AUGER OBSERVATORY: A SELECTION

This section gives an overview about some selected results obtained from the data of the Pierre Auger Observatory. The focus is on results in connection with this thesis work.

As discussed in chapter 1, the search for UHE photons has a potential in the identification of the astrophysical sources of UHECRs. The origin of UHECRs is also investigated by studying the distribution of their arrival directions at the highest energies. This challenge is considerable because the flux falls rapidly with increasing energy, and because UHECRs experience substantial magnetic deflections due to their charge. The resistance of UHECRs to deflections from their trajectory, due to the Galactic or extra-galactic magnetic fields, depends on the energy $E$ and charge $eZ$ of the particle, and it is expressed in terms of the *rigidity*, defined as $R = E/eZ$.

At low rigidity/energy (around a few EeV) the astrophysical sources cannot be searched for directly, due to too large cosmic-ray deflections. In this regime, in turn, the **search for large-scale anisotropy** is of interest, as this can be reflective of either a collective motion of cosmic rays (e.g., of their propagation) or of the global distribution of their sources, or of both. Such a search has been performed at the observatory, using the technique of the harmonic analysis of the

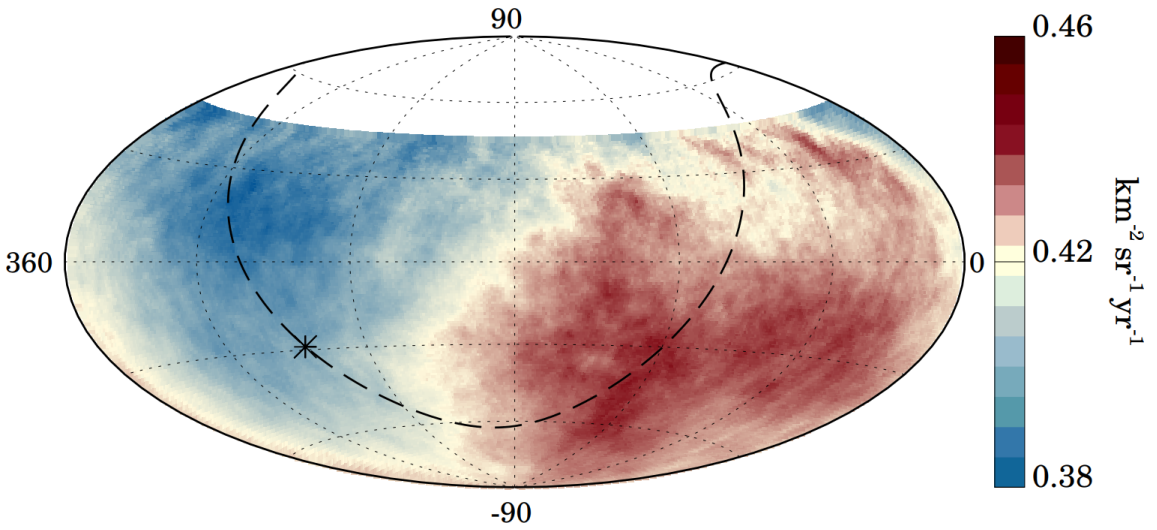



**Figure 2.16:** *Sky map in equatorial coordinates. It shows the cosmic rays flux with energies above* $8 \times 10^{18}$ eV. *The Galactic plane is represented with a dashed line, and the Galactic center is indicated with a star. The dipolar component has a direction lying* $\sim 125°$ *from the Galactic Center direction, indicating an extra-galactic origin for this flux. From [92].*

counting rate, both in right ascension and in azimuth angle, which is sensitive to non-uniformity in declination. For energies above 8 EeV, the amplitude of the first harmonic in right ascension is significant at a level of more than $5.2\sigma$, demonstrating the existence of an anisotropy in arrival directions. By combining the analysis in right ascension with that in azimuth, the anisotropy can be characterised as a dipole with an amplitude of $\approx 6.5\%$ toward right ascension $\approx 100°$ and declination $\approx -24°$. For illustration, the distribution of events in equatorial coordinates, smoothed with a 45° radius top-hat function to better display the large-scale features, is shown in figure 2.16. The dipole direction is $\approx 125°$ away from the Galactic center, thus indicating an extra-galactic origin for these UHE particles.

At higher rigidity/energy (tens of EeV), instead, the distribution of the UHECR arrival directions might show anisotropy at smaller angular scales, mirroring the inhomogeneous distribution of the nearby extra-galactic matter. For protons, the typical angular deflections caused by magnetic fields would be of the order of a few degrees, being Z times larger in the case of nuclei with atomic number Z. The harmonic analysis has thus been complemented in Auger by the **search for anisotropy at small and intermediate**[1] **angular scales** [93]. Two prominent classes of extragalactic sources detected by Fermi-LAT have been considered in the analysis: active galactic nuclei (AGNs) and starburst galaxies (SBGs) since these populations are well-motivated physically.

[1] With "intermediate", an angular scale is denoted, which is larger than the experimental resolution of the SD detector ($\sim 1°$) and smaller than large-scale patterns ($\gtrsim 45°$).

AGNs are considered because their jets and radio lobes satisfy the Hillas criterion for shock acceleration [94]. SBGs are regions of intense star formation, and potentially have increased rates of extreme events associated with the deaths of short-lived, massive stars, such as gamma-ray bursts, hypernovae, and magnetars [95, 96].

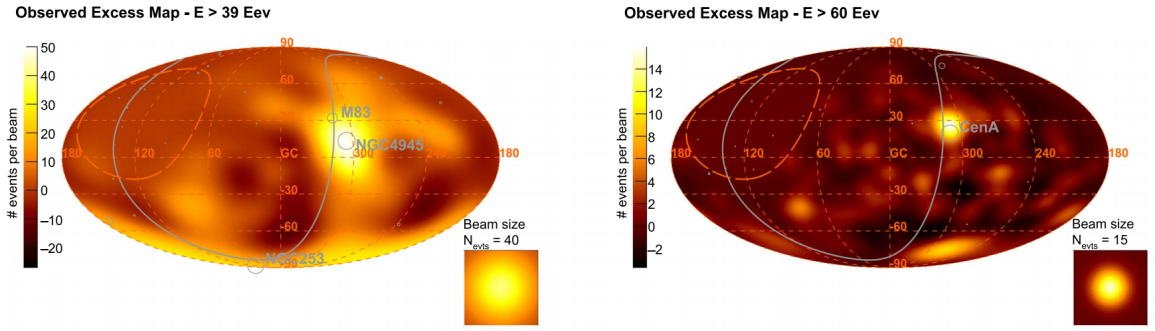



**Figure 2.17:** *Sky maps in equatorial coordinates. Observed excess map obtained with SBGs above* 39 EeV *(left) and $\gamma$ AGN above* 60 EeV *(right). These maps are obtained subtracting the estimated isotropic component. The Galactic plane is shown as a solid gray line. The orange dashed line delimits the field of view of the Auger SD. From [97].*

Two sky models based on these two populations of extra-galactic gamma-ray emitters have been built, with two free parameters, the fraction of events correlating with the astrophysical objects, and the angular scale characterizing the clustering of cosmic rays around those objects. A maximum-likelihood ratio test has been used to evaluate the best values of these parameters and to quantify the strength of each model by contrast with isotropy. The different degrees of anisotropy obtained from the two catalogs can be understood from figure 2.17, where the observed excess-maps obtained with SGBs (left) above 39 EeV and with AGNs (right) above 60 EeV are shown. 39 EeV and 60 EeV are the energy thresholds that yield the maximum test-statistic value for the SBGs and AGNs models, corresponding to 4 s.d and 2.7 s.d., respectively. The maximum deviation from isotropy for the SBGs (AGNs) model is found at an angular scale of about 13 (7) degrees. A hotspot in the direction of the Centaurus A/M83/NGC 4945 group is visible when comparing data to the SBGs model. The AGN model on the right is also dominated by Centaurus A, which is 7 and 13 degrees away from NGC 4945 and M83, respectively. The larger significance associated to the SBGs model is because it additionally captures the excess close to the Galactic South Pole, interpreted as contributions from NGC 1068 and NGC 253.

The dipole detected for cosmic rays above 8 EeV is in a direction about 55° away from that of the dipole of the 2MASS redshift survey [92], which traces the distribution of extragalactic matter. The agreement between the directions of the dipoles is improved if one assumes a cosmic-ray mass-composition heavier than protons. Similarly, the indication of anisotropy at intermediate angular scales (around 10 degrees) when comparing the arrival directions of UHECRs with the SBGs and AGNs models, suggests that the Galactic and/or extragalactic magnetic fields have a non-negligible effect on the cosmic-ray trajectories. Overall, these two facts are compatible with scenarios in which the extra-galactic cosmic rays composition is not of pure protons, but it is mixed,

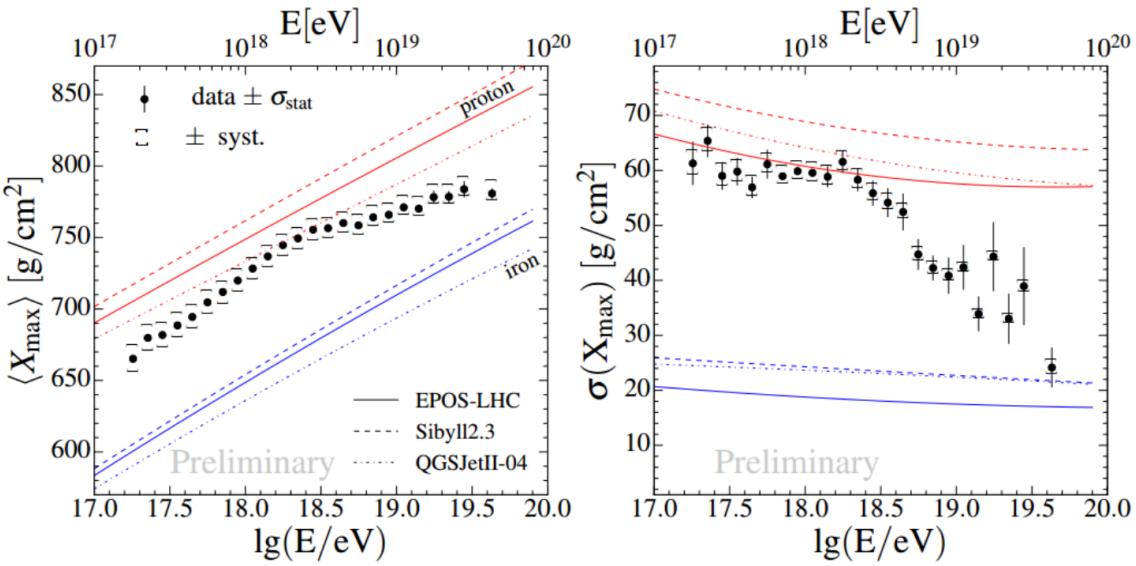
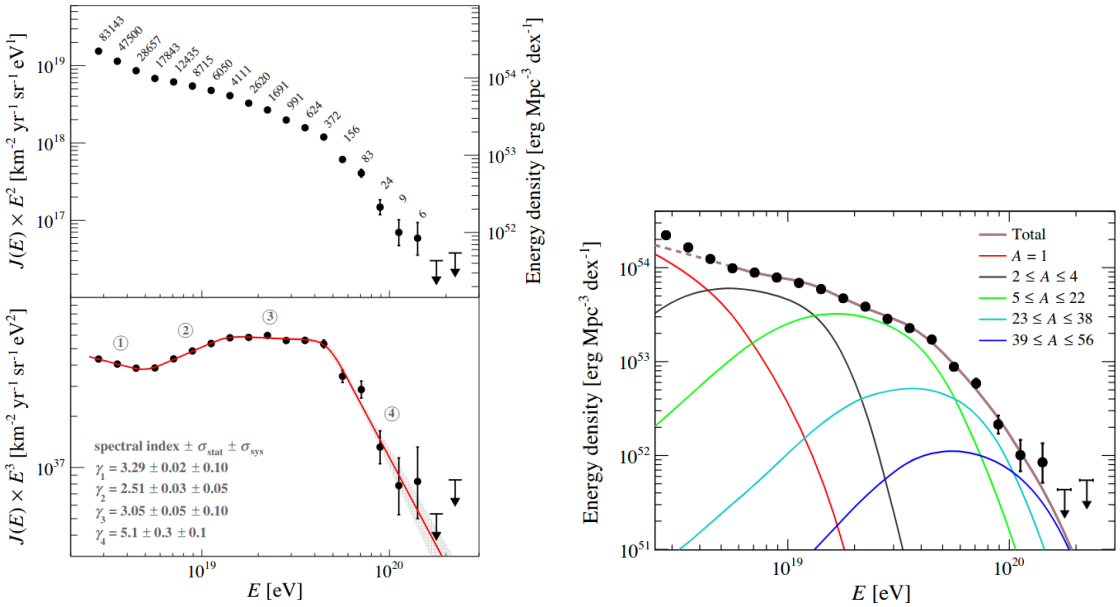



**Figure 2.18:** *Energy evolution of the first central moments of the $X_{max}$ distribution (average on the left panel and standard deviation on the right panel) compared to air shower simulations proton (red) and iron (blue) primaries. EPOS-LHC (solid line), Sibyll2.1 (dashed line) and QGSJetII-04 (dotted line) are considered. From [90].*

up to the highest energies. This is actually consistent with the outcome of **mass composition studies** performed with Auger data [90]. These are based on the measurement of the depth at the shower maximum, $X_{\text{max}}$, with the FD. $X_{\text{max}}$ is a mass-dependent observable as it is proportional to the logarithm of the mass $A$ of the primary particle. Information on the mass composition can then be inferred by comparing the measured $X_{\text{max}}$ distributions with those predicted from simulations of different species of primary cosmic rays, using different models for the hadronic interactions in the atmosphere. The two panels in figure 2.18 show the energy evolution of the two first moments of the $X_{\text{max}}$ distributions obtained from data (black points) compared to those expected from simulations of primary protons (red lines) and iron nuclei (blue lines). Although the hadronic models slightly differ on the predicted values of the moments, the data above $10^{18}$ eV trend clearly point to a composition of UHECRs evolving toward heavier nuclei as the energy increases [90].

The measurement of the flux of cosmic rays as a function of energy, i.e., of the **UHECR energy spectrum**, is also important to probe the sources of cosmic rays, as its structures may indicate either changes in their origin, or nature, or propagation. The spectrum measured with the high-statistics data from the SD is shown in figure 2.19a. Described as a sequence of four power laws with smooth transitions [99] (red line in the bottom panel), it presents different features. A hardening of the spectrum from $\gamma_1 = 3.29 \pm 0.02 \pm 0.10$ to $\gamma_2 = 2.51 \pm 0.03 \pm 0.05$, denoted as *ankle*, takes place at the energy $E_{12} = (5.0 \pm 0.1 \pm 0.8) \times 10^{18}$ eV. The softening at $E_{23} = (13 \pm 1 \pm 2) \times 10^{18}$ eV, where the spectral index changes from from $\gamma_2$ to $\gamma_3 = 3.05 \pm 0.05 \pm 0.10$, has been observed by Auger for the first time, thanks to the statistical power of its data accumulated




**(a)** **(b)**

**Figure 2.19:** ***(a top)**: Energy spectrum scaled by $E^2$. The number of detected events is indicated for each energy bin. In this representation the data provide an estimation of the differential energy density per decade. **(a bottom)**: Energy spectrum scaled by $E^3$ fitted with a sequence of four power laws (red line). The numbers ($i = 1, \ldots, 4$) in the circles identify the energy intervals where the spectrum is described by a power law with spectral index $\gamma_i$. The shaded band indicates the statistical uncertainty of the fit. Upper limits are at the 90 % c.l. **(b)**: Energy density obtained considering the $X_{max}$ evolution as a function of the energy. The dashed curve shows the energy range that is not used in the fit and where an additional component is needed for describing the spectrum. From [98].*

out of an exposure of $\approx 60000$ km$^2$ sr y. The spectrum then softens further above a suppression energy of $E_{34} = (46 \pm 3 \pm 6) \times 10^{18}$ eV with $\gamma_4 = 5.1 \pm 0.3 \pm 0.1$, confirming with higher precision previous reports of the strong attenuation of the flux at the highest energies [100–102]. The newly found feature at $E_{23}$ indicates that, in fact, the suppression happens in two steps.

The features observed in the energy spectrum can be better explained by relating them with the above-discussed data on mass composition and on arrival directions. As experimentally demonstrated with the discovery of significant directional anisotropies, cosmic rays at the highest energy, above about 10 EeV, are of extragalactic origin. Consequently, to reach Earth they must cross the background photon fields permeating the extragalactic space. In particular, the cosmic microwave background (CMB) photons induce pion-production with protons colliding at around $5 \times 10^{19}$ eV and photo-disintegration of heavier nuclei at a roughly similar threshold, leading to the expectation of a spectral steepening (the Greisen-Zatsepin-Kuz'min (GZK) effect, see section 1.2.1). The observed suppression could then be a propagation effect. On the other hand, if UHECRs are accelerated in astrophysical sources to a maximum energy proportional to their charge (i.e., to the

same maximum-rigidity), the observed suppression of the flux might also be a consequence of the maximum acceleration energy reached by the sources.

We show in Fig. 2.19b the best reproduction of the Auger data by simultaneously fitting the energy spectrum above $5{\times}10^{18}$ eV and the distribution of $X_{\mathrm{max}}$ (figure 2.18, EPOS LHC as model of hadronic interactions). The fitted model considers several nuclear components injected at the sources with a power-law spectrum and with the maximal energy of the sources exponentially cutting off. As one can see, the abundance of nuclear elements at the sources is dominated by intermediate-mass nuclei accelerated to $\approx 5\ Z{\times}10^{18}$ eV and escaping from the source environments with a very hard spectrum. In this scenario, the steepening observed above $\approx 5{\times}10^{19}$ eV results from the combination of the maximum energy of acceleration of the heaviest nuclei at the sources and the GZK effect. The newly observed steepening at $\approx 10^{19}$ eV reflects the interplay between the flux contributions of the helium and carbon-nitrogen-oxygen components injected at the source with their distinct cut-off energies, shaped by photodisintegration during the propagation.

Complementary information on the origin of the flux suppression are provided by the search for UHE photons and neutrinos. Interactions between UHECRs and photons of the CMB lead to emission of cosmogenic neutrinos and photons, whose flux is dependent on the mass of UHECRs, being more suppressed for heavier primaries. The search for UHE photons and neutrinos is also relevant with respect to some models of dark matter, such as super-heavy relic particles from the early universe, the decay of which might result in photons and neutrinos dominating the final state.

The search for photons in Auger has yielded the most stringent upper limits to the photon flux at energies above $E > 10^{18}\,\mathrm{eV}$, as shown in figure 2.20. The blue arrows represent limits obtained with hybrid data. The black arrows correspond to limits obtained with SD data. The limits are compared to different predictions for the GZK flux, depending on the UHECR mass [35, 103], as well as to expectations for a variety of top-down models of UHECR production(TD, Z-Burst, SHDM I [104] and SHDM II [105]). On the one hand, the attained sensitivity allows for testing photon fractions of about 0.1 % thus exploring the region of photon fluxes predicted in some optimistic astrophysical scenarios (e.g., GZK proton-I) [35]. On the other hand, topological defects and Z-burst models result to be excluded, while stringent limits on some SHDM models are set. The case of a SHDM particle with mass $M_\chi = 4.5\times10^{22}\,\mathrm{eV}$ and life-time $\tau_\chi = 2.2\times10^{22}$ yr is only marginally compatible with the limits which will be obtained in this work and severely constrained by the limits from the surface detector data [106], in agreement with the interpretation of the Planck results in [107]. Also, constraints on the lifetime-and-mass parameter space of SHDM particles can be imposed [108].

UHE neutrinos, with energies above $10^{17}$ eV, have been searched for with SD data. The identification is efficiently performed for neutrinos of all flavors interacting in the atmosphere at large zenith angles, as well as for Earth-skimming $\tau$ neutrinos with nearly tangential trajectories relative to the Earth. The search strategy consists in selecting showers that exhibit a broad time structure in the signals induced in the SD stations. Such signals are indicative of an early stage

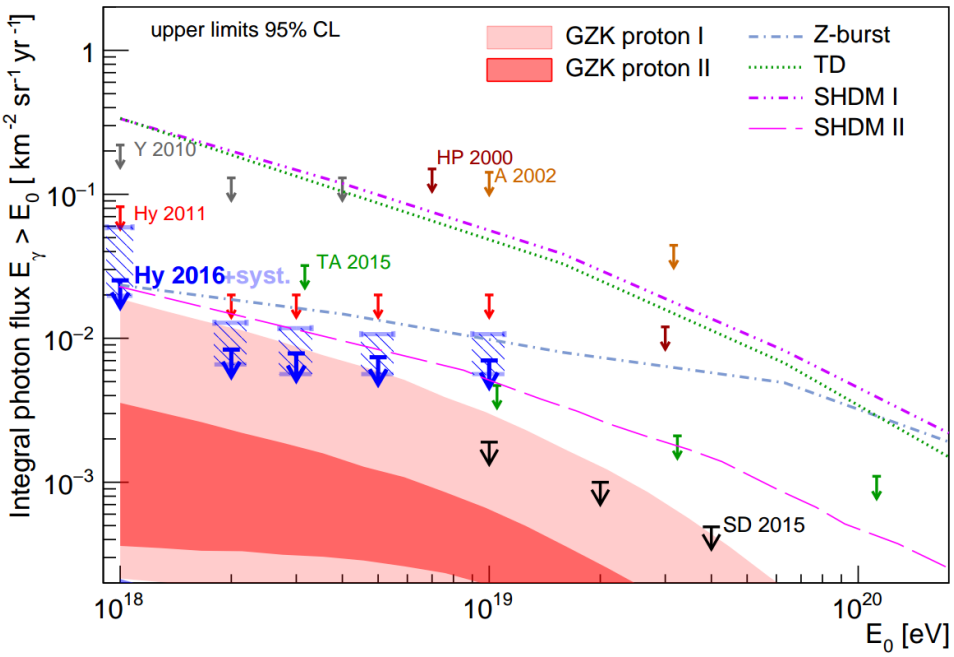



**Figure 2.20:** *Upper limits on the integral photon flux derived from 9 years of hybrid data (blue arrows, Hy 2016) for a photon flux $E^2$ and no background subtraction. The limits obtained when the detector systematic uncertainties are taken into account are shown as horizontal segments (light blue) delimiting a dashed-filled box at each energy threshold. Previous limits from Auger: (SD [106] and Hybrid 2011 [109]), for Telescope Array (TA) [110], AGASA (A) [111], Yakutsk (Y) [112] and Haverah Park (HP) [113] are shown for comparison. None of them includes systematic uncertainties. From [114].*

of development of the shower, a signature of the shower developing close to the ground, like it would happen for neutrino showers, which can initiate very deep in the atmosphere. Figure 2.21 show the attained upper bounds to cosmogenic neutrinos. These start to constrain astrophysical models that aim at describing the UHECR flux suppression above $4 \times 10^{19}$ eV by energy losses of protons in the CMB [123].

As already discussed the study of the arrival directions of cosmic rays has revealed deviations from isotropy that, combined with photon and neutrino searches, could be of relevance for multi-messenger astronomy. Many TeV $\gamma$-sources are observed at energy fluxes of the order of $1\,\mathrm{eV\,cm^{-2}\,s^{-1}}$. Such sources would be visible to the Auger Observatory as strong photon and Galactic neutron sources if their energy spectrum would continue with a Fermi-like energy distribution up to about $10^{17}$ eV. Their absence suggests that their maximum source energy does not reach out to the threshold energy of the Observatory and/or that their spectrum is significantly softer than a $-2$ spectrum. Point source searches includes also mergers of compact binaries alerted by gravitational wave interferometers. The most spectacular event so far was the neutron star merger GW170817 at a distance of about 40 Mpc. Within the predefined $\pm 500$ s search window, the Auger Observatory reached a neutrino flux sensitivity above 100 PeV that was over an

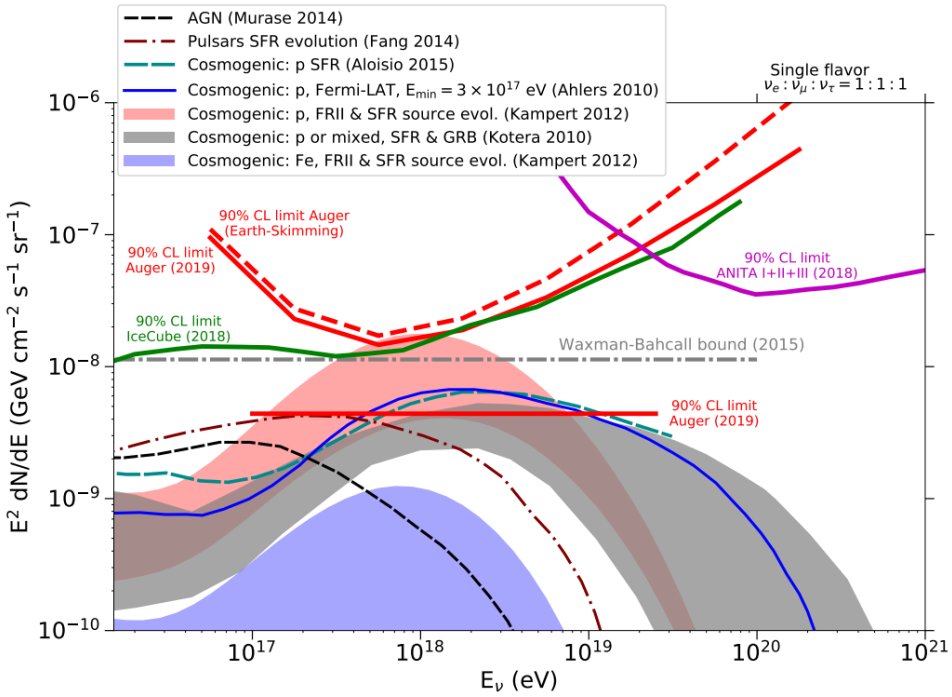



**Figure 2.21:** *Pierre Auger Observatory integral and differential upper limits (90 % C.L.) to the the diffuse flux of UHE neutrinos (solid red lines). The differential limits obtained by IceCube [115] (solid green) and ANITA I+II+III [116] (solid dark magenta) are also shown. The expected neutrino fluxes for several cosmogenic [118–120] and astrophysical models of neutrino production, as well as the Waxman-Bahcall bound [121, 122]are also plotted. All limits and fluxes are converted to single flavor. From [123].*

order of magnitude higher than of any other neutrino observatory presently operated. Again, the absence of neutrinos at Auger, IceCube and ANTARES allowed constraining the jet properties of the neutron star merger [124].

In this thesis work a new method is studied so to improve the search for UHE photons with hybrid data. The concept of universality of air-showers, which will be introduced in the next chapter, is at the basis of the new technique, the development of which will be explained in the following chapters.

# 3 AIR-SHOWER UNIVERSALITY: THE AUGER MODEL

## Contents

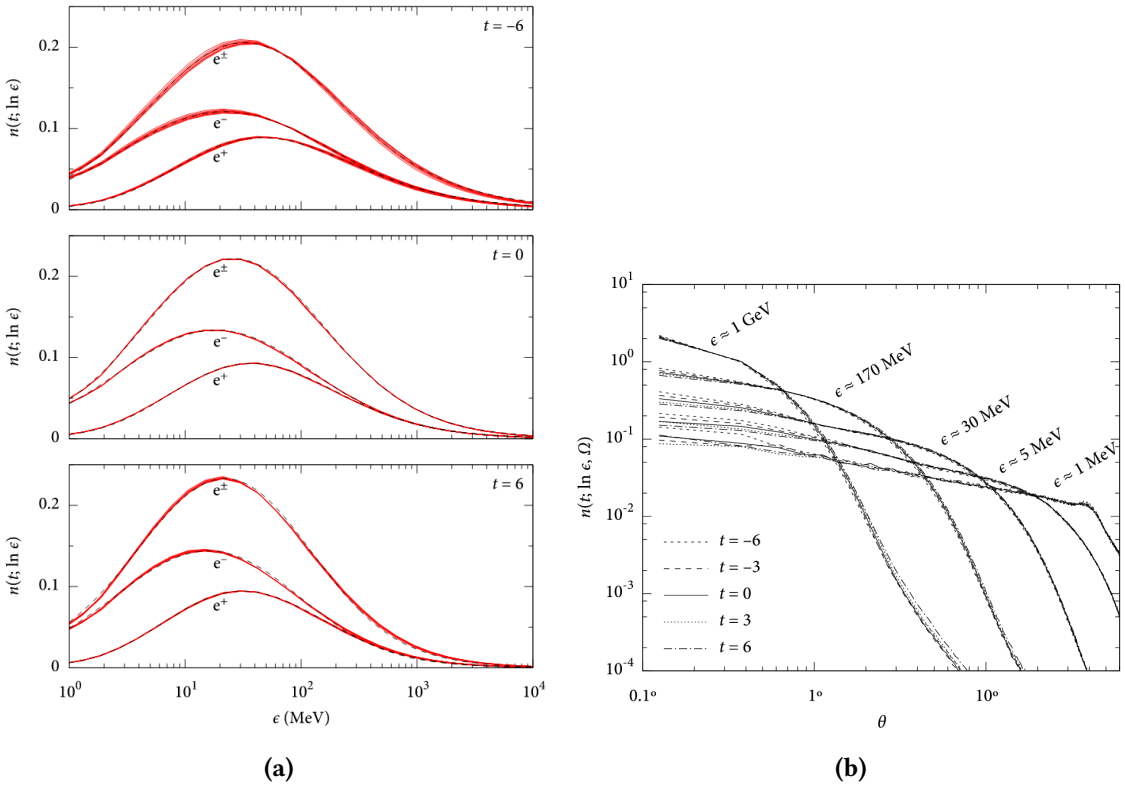


A high-energy cosmic ray that hits the atmosphere produces a large amount of secondary particles (for example, in a proton shower with energy of $10^{19}$ eV, about $10^{10}$ particles reach the ground at the altitude of the Pierre Auger Observatory). This bunch of particles can be considered, in effect, as a "thermodynamic" system, which can be described by few parameters. This property is known as air-shower *universality*. This concept is at the basis of one of the observables that will be used in this thesis to discriminate photon-induced showers from those initiated by nuclei. This chapter is thus meant to provide an introduction to the universality concept and to its application to the analysis of Auger data.

The general idea behind universality [125] as illustrated in section 3.1, is that, for electromagnetic cascades, the lateral and longitudinal development, along with the energy and the angular distribution of the secondary particles, depends only on two parameters, namely the energy of the primary cosmic ray and on the stage of development of the induced shower. As explained in section 3.2, within the Auger Collaboration this concept has been extended to hadronic cascades [126], so to include the dependency on the mass of the primary cosmic ray. This is done by introducing an additional quantity related to the mass, i.e., the shower muon content. A university-based description of a shower has then been built, based on the superposition of four different components, which have a universal behavior depending only on the energy and stage of development. As a consequence, a model [127, 128] has been developed for the prediction of the signals measured at ground with the Auger water-Cherenkov detectors (WCDs). The sections 3.3 and 3.4, describe the parameterizations used to describe the signal size and its temporal shape, respectively. Finally, the application of such model to Auger data allows for the reconstruction of the parameters that characterise the primary cosmic ray, as shortly outlined in the last section, 3.5.

## 3.1 THE CONCEPT OF AIR-SHOWER UNIVERSALITY

The discovery of the air-shower universality property dates back more than 50 years [131]. The general idea behind universality is that the energy spectrum of the secondary particles produced during the shower development as well as their angular and lateral distributions depend only on the energy of the primary and the stage of shower development [125, 132–136]. The universal behavior is a consequence of the huge amount of particles produced in an extensive air-shower at ultra-high energies: the very large number of interactions minimize the relevance of single fluctuations, allowing for a smoothing of the properties of the showers, which makes universal models viable.




**(a)** **(b)**

**Figure 3.1:** *Universality of the e.m. component of showers.* ***(a)****: average energy distribution (red bands) for electrons ($e^+$), positrons ($e^-$) and their sum ($e^\pm$) at different stages of the shower development ($t = -6,0,6$, where $t$ is defined in equation 3.1) for different primaries (p, Fe, $\gamma$). The dashed lines are the parametrizations of the energy spectra given in [136]. Figure from [136];* ***(b)****: angular distribution of electrons at different energies as a function of momentum angle to the shower axis, for proton-initiated showers at* $10^{18}$ eV*, and for different shower stages. From [136].*

Detailed Monte Carlo Simulations have been exploited in [136] for a complete investigation of the universality behaviour of the electron-positron component in extensive air-showers as a function of energy, mass, zenith angle of the primary particle and of the evolution stage of the

shower. In [136] the stage of the shower development is expressed through the *relative evolution stage*, $t$, defined as

$$t = \frac{X - X_{\max}}{X_0} \tag{3.1}$$

where $X_0 \simeq 36.7\,\mathrm{g\,cm^{-2}}$ is the radiation length for electrons in the air. With this definition of age, $t = 0$ corresponds to the shower maximum, $X_{\max}$. Positive values of $t$ represent a development stage after the shower maximum (*old* shower) while negative values indicate that the maximum in development is not reached yet (*young* shower).

A universal behaviour is evident in the normalized energy spectra of the secondary $e^{\pm}$ at different $t$ values, as shown in the three panels of figure 3.1a. The red bands represent the distributions in showers initiated by different primary species (p, Fe, and $\gamma$) and energies ($10^{17}$, $10^{18}$ and $10^{19}$ eV). The parametrizations developed in [136] are represented with dashed lines and show a very high accuracy, with deviations lower than 10 %.

Another evidence of universality is visible in figure 3.1b, which shows the distribution of the secondary electrons as a function of momentum angle $\theta$ between the secondary electron momentum and the shower axis for $10^{18}$ eV primary protons at different stages. Only a small range of angular distances is shown, because the majority of electrons and positrons stay close to the shower axis during the development. The dispersion near the axis is small, with an increase only for the lowest energy particles (small $\epsilon$ values). As reported in [136], the effects of primary mass and shower stage are smaller than the differences between individual showers; moreover the angular spectra are independent of the primary zenith angle or energy. This implies that the parameterizations of the angular spectra can be considered as functions only of the particles angle and energy.

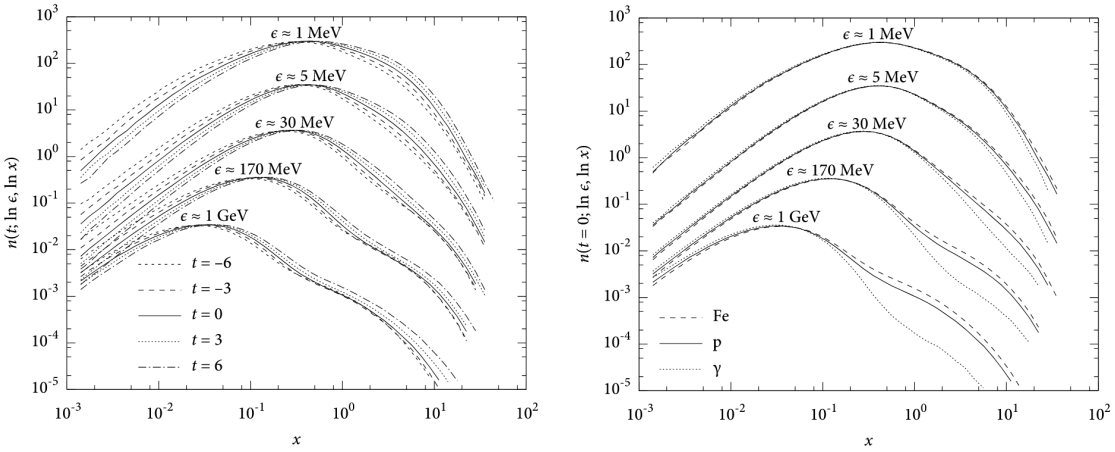



**Figure 3.2:** ***(a)****: Lateral spread distribution, $x = r/r_M$, of the e.m. particles in proton-initiated shower with energy $E = 10^{18}$ eV for different shower stages;* ***(b)****: Lateral spread distribution, $x$, at the shower maximum ($t = 0$) for different primaries (p, Fe, $\gamma$) at $E = 10^{18}$ eV. From [136].*

Another aspect relevant to universality is the lateral spread of the e.m. particles, which is shown in figure 3.2a for proton-initiated showers at $10^{18}$ eV at different shower stages. A universal description can be found after expressing the lateral spread in terms of the Moliére radius, $r_M$, as $x = r/r_M$, where $r$ is the distance of the secondary $e^{\pm}$ from the shower axis. A dependence on the shower evolution is evident. However, when including in the parametrisation the parameter $t$, no statistically significant dependencies of the lateral spread are found on either the primary zenith angle or energy. A violation of universality can in turn be observed in the figure 3.2b, where the lateral spread is shown for different primaries species. A dependence on the primary mass is visible at large values of $x$: this effect is related to the separate origin of the particles contributing to the lateral spread distribution. At small $x$ values, $e^{\pm}$ are mostly produced through bremsstrahlung and pair creation processes, while those at large $x$ values are mostly due to pions [136]. The pions, produced in the first stages of the shower development in hadronic showers, generate a muonic component, which depends on the mass of the primary particle.

Such universality violation thus called for deeper studies [137, 138], aimed at extending the universality concept to the muonic component. In [137], in particular, it was shown that the distributions of muons present an approximately universal behavior when the development of this component is described in terms of $X' = X - X^{\mu}_{\max}$ and when the number of muons, $N_{\mu}$, is expressed in terms of $N'_{\mu} = N_{\mu}/N^{\max}_{\mu}$, where $X$ is the slant depth traversed in atmosphere, $X^{\mu}_{\max}$ is the depth of the maximum of the muonic component, and $N^{\max}_{\mu}$ is the number of muons at $X^{\mu}_{\max}$.

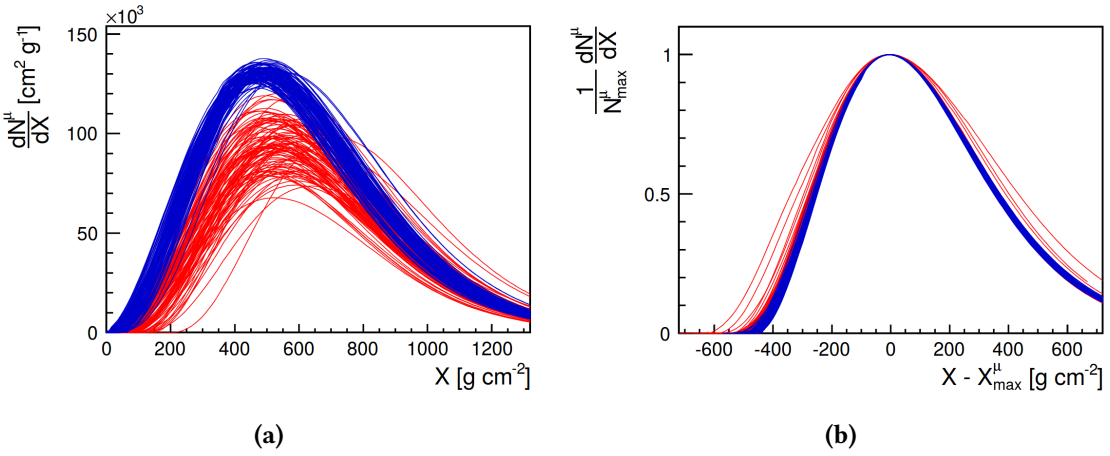



**Figure 3.3:** ***(a)****: Muon production profiles as a function of the slant depth (X), for proton-initiated (red) and iron-initiated (blue) showers at* $10^{19}$ *eV simulated with QGSJet-II.03.* ***(b)****: The same muon production profiles, shown in* $(X', N')$ *coordinates. From [137].*

The universal behavior of the muonic component is illustrated in figure 3.3. In figure 3.3a, the profiles of the muonic component are shown as a function of the slant depth $X$ for proton-initiated (red) and iron-initiated (blue) simulated showers. In turn, in figure 3.3b the same profiles are shown in the coordinates $X'$ and $N'$, from which a rather universal shape is obtained.

As it will be shown in the next section, the universality behavior, common to the electromagnetic and muonic shower-components, the former linked to the energy of the primary cosmic ray, the latter to its mass, can be exploited to build a model to reconstruct extensive air showers measured in ultra-high energy cosmic-ray experiments.

## 3.2 UNIVERSALITY-BASED MODEL OF GROUND DETECTOR SIGNALS

The universality property of electromagnetic cascades implies that the shower development can be described in terms of the energy of the primary cosmic ray and of the shower development stage only. Therefore, a description of the electromagnetic signal in any detector can be modeled, by including, in the description, the geometrical configuration of the shower with respect to the detector.

This approach has been extended to hadronic showers (see e.g., [139–141]) by introducing in the signal modelisation an additional quantity related to the shower muon content, which is in turn sensitive to the mass of the primary cosmic ray.

In a first formulation of the universality-based model, the signal in a generic detector was described by three components: the muonic component ($S_\mu$), the electromagnetic component ($S_{e\gamma}$), deriving from the decays of the highest energy $\pi^0$ and the electromagnetic one deriving from the muon decays ($S_{e\gamma(\mu)}$). These two additional components are also dependent only on energy and the shower stage development, like the pure electromagnetic one. The muonic component is actually introduced in the model as a scale factor, $F_\mu$, namely

$$F_\mu = \frac{S_\mu}{S_\mu^{\mathrm{ref}}} \tag{3.2}$$

where $S_\mu^{\mathrm{ref}}$ is the reference muon signal, taken as the average one produced by a proton of energy $10^{19}$ eV, simulated with QGSJetII-03 [127] as a reference hadronic model hadronic. Since the signal $S_\mu$ is proportional to the density of muons $\rho_\mu$, namely

$$F_\mu = \frac{S_\mu}{S_\mu^{\mathrm{ref}}} \propto \frac{\rho_\mu}{\rho_\mu^{\mathrm{ref}}} \tag{3.3}$$

$F_\mu$ represents the deviation of the muon density from the expectation of a reference primary (proton) and a reference hadronic model (QGSJetII-03).

However, it was found that in the three-components model the electromagnetic component deviates from the modeled signal as a function of the distance to the ground of the shower maximum, as shown in figure 3.4. While the e.m. signal, according to universality, should depend only on energy and development stage, one can see instead a dependence on the primary mass and on the hadronic model. The signal from iron-showers simulated with the same model, QGSJet-03, deviates by 10% from the prediction of the reference primary and model (grey continuous line), the deviation growing up to 40% when the iron-shower is simulated with EPOS-LHC [144].

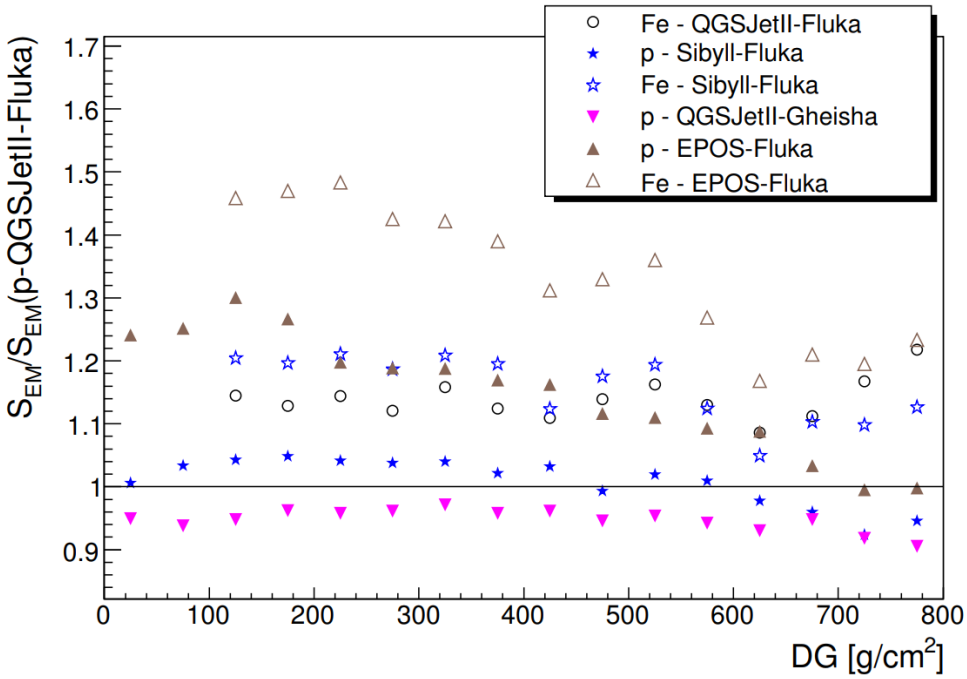



**Figure 3.4:** *Electromagnetic signals for showers generated by proton and iron primaries as a function of the distance to the ground (DG) of the shower maximum. Showers have been simulated using different models for the hadronic interactions (QGSJetII-03 [56], Sybill [55], EPOS [54]). The signals are normalised to those generated in proton-showers simulated with QGSJetII.*

Such violation of the universality has been investigated with simulations, from which it was found that a significant fraction of electromagnetic secondary particles derives in fact from low-energy hadronic interactions. This component, so-called *jet component*, is produced at a late stage of the shower development, from jets with high transverse momentum. Those jets have a large angle with respect to the shower axis, thus impacting the signals in stations far from the core. The electromagnetic particles created by those jets can be distinguished from the rest of the electromagnetic cascade by using the projected impact point of their *mother particles*, as sketched in figure 3.5a. The direction of the mother particle at the last interaction point is extrapolated to ground level where the distance to the shower core, $r_{\text{proj}}$, of the impact point is calculated. Figure 3.5b shows the distribution of $r_{\text{proj}}$ for electromagnetic particles falling in detectors at different radial distances to the core, $r$. For each $r$, the signal distribution shows a peak, which becomes more important at larger values of $r$.

A more robust and universal description of the shower signals thus requires the introduction of this fourth component. As demonstrated in [127], the signal size in an air shower can be described as the superposition of four different components: muons ($\mu$); electromagnetic particles produced in the high energy $\pi^0$ decays ($e\gamma$); electromagnetic particles produced in the decays or the interactions of muons ($e\gamma(\mu)$); electromagnetic particles produced by the jet component

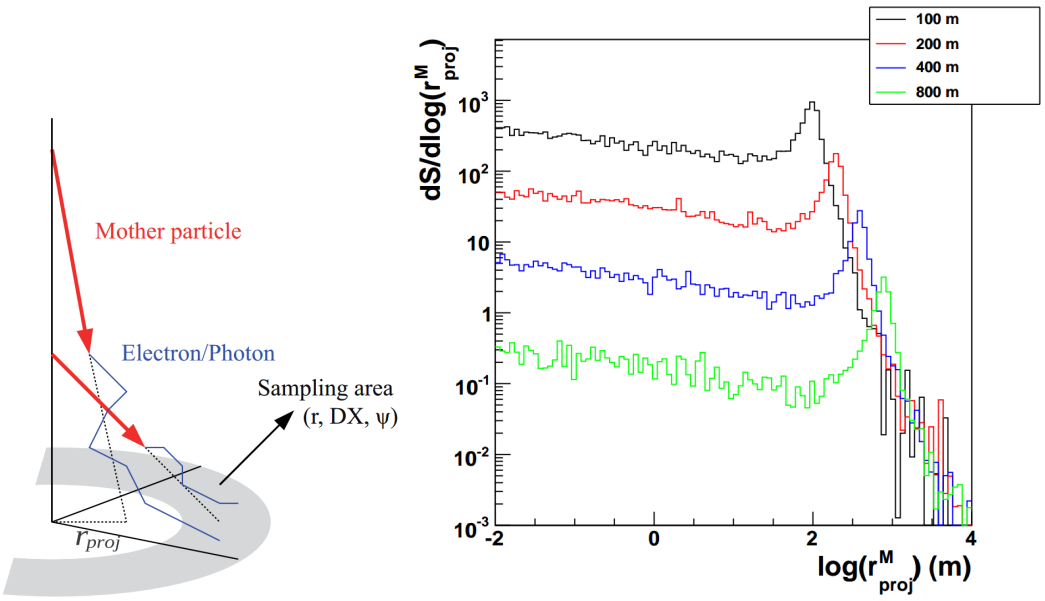
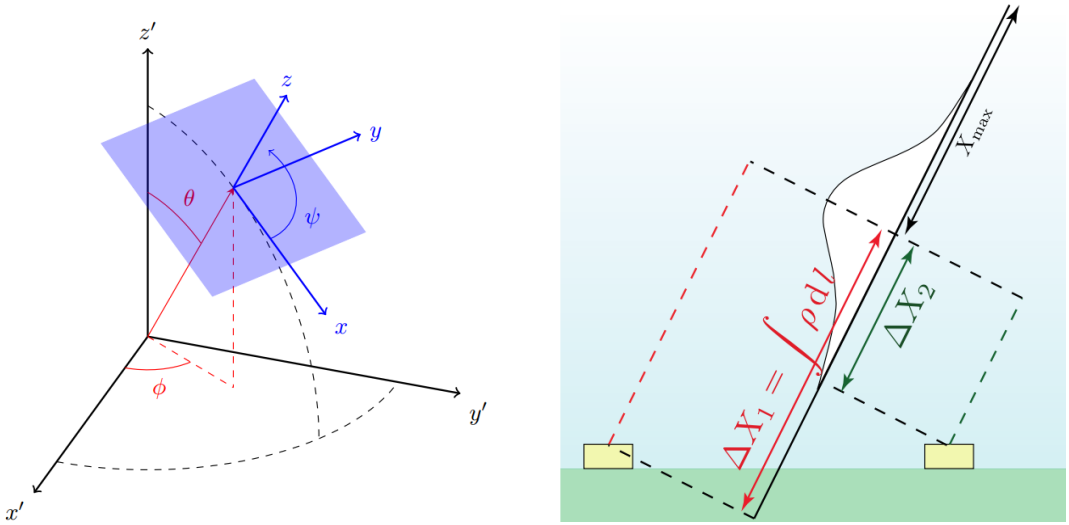



**Figure 3.5:** ***(a)****: Schematic description of the jet component. The electromagnetic component from hadron jets is distinguished from the rest of the cascade by the projected impact point of the mother particle. From [145].* ***(b)****: Distribution of the projected radius, $r_{proj}$, for the electromagnetic particles in detectors at different radial distances to the shower core. Figure from [143].*

($e\gamma$(had)). Each component has a universal behaviour, depending only on the primary energy and stage of development of the shower, while the mass of the primary is accounted for by $F_\mu$.

## 3.3 UNIVERSALITY-BASED PARAMETERIZATION OF THE SIGNAL SIZE

The universality-based model of ground detectors signal, introduced in the previous section, has been adopted for the Auger water-Cherenkov detectors of the Pierre Auger Observatory in [127]. A parameterization of each of the four universal components has been derived by using proton simulations based on the hadronic interaction model QGSJetII-03. The simulated energies range between $10^{18.5}$ eV and $10^{20}$ eV, and the zenith angles between 0° and 60°.




**(a)** **(b)**

**Figure 3.6:** ***(a)****: View of the shower coordinate system centered into the core. The core distance $r$ is calculated as $r = \sqrt{x^2 + y^2}$. The stations are projected on the plane $z = 0$. The position of a station on the plane $z$ is identified by $r$ and the angle $\psi$.* ***(b)****: $\Delta X$ distance from $X_{max}$. This distance is defined as the integral of the atmospheric depth along the shower axis. $\Delta X$ can be different for stations at the same radial distance from the core but at a different azimuth. From [142].*

For a given energy, $E$, depth at the shower maximum, $X_{\text{max}}$, relative muon content, $F_\mu$, zenith angle, $\theta$, and density of the air at the ground, $\rho^{\text{air}}_{\text{ground}}$, the universality ansatz for the signal description in a detector at $(r, \psi)$ in the shower coordinate system (see figure 3.6a) is

$$S\left(r,\psi,E,X_{\text{max}},F_\mu,\theta,\rho^{\text{air}}_{\text{ground}}\right) = \sum_{i=1}^{4} S^i_0(r,\Delta X,E)\cdot f^i_{\text{mod}}(r,\psi,\theta)\cdot f^i_{\text{atm}}(r,\rho^{\text{air}}_{\text{ground}}) \cdot f^i_{\text{conv}}(r,\psi,\theta)\cdot f^i_{F_\mu\text{fluct}}(r,F_\mu) \tag{3.4}$$

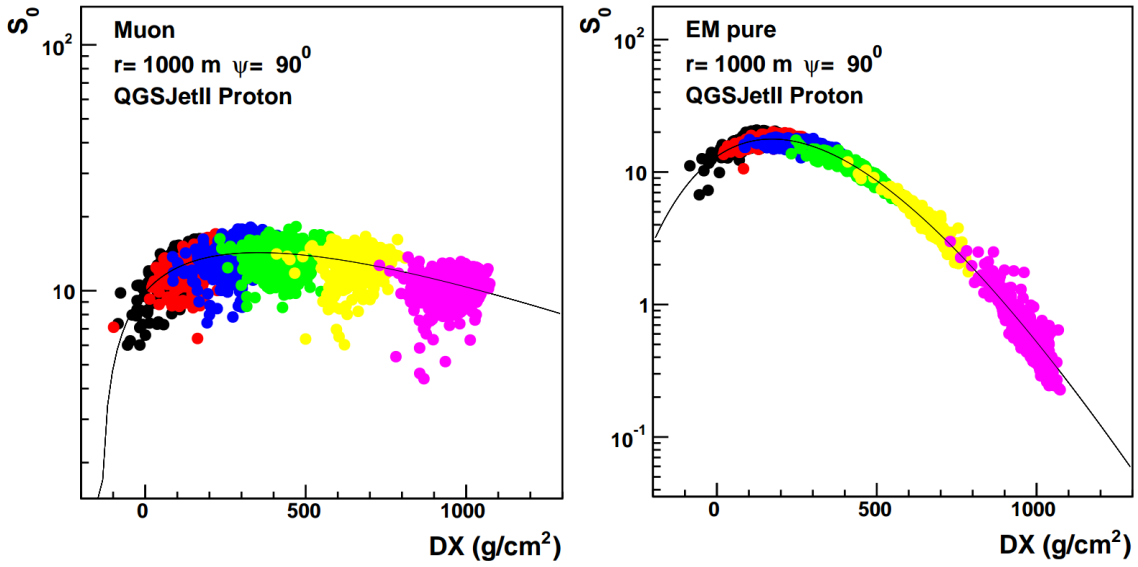
where $i$ runs over the 4 shower components, $\Delta X$ is the stage of development of the shower accounted as the difference between $X_{\text{max}}$ and the slant depth $X$ (see figure 3.6b). The description of the signal given in equation 3.4 is calculated by factorizing different terms. For each component, $S_0$ is the *ideal* signal, $f_{\text{conv}}$ is the conversion factor to a realistic detector, $f_{\text{atm}}$ and $f_{\text{mod}}$ account for the atmospheric effects and $f_{F_\mu\text{fluct}}$ takes into account the correlations with $F_\mu$. The different terms are detailed in the following paragraphs.

### *The ideal signal,* $S_0$

The term $S_0$ is the signal produced by each component in a ideal spherical detector of $10\,\mathrm{m}^2$ area and without ground. The detector response thus does not depend on the incoming direction of the shower particles. For each component, $S_0$ is fitted to a modified Gaisser-Hillas function:

$$S_0^i(\Delta X, E) = S_{\max}\left(\frac{E}{10^{19}\,\mathrm{eV}}\right)^{\gamma}\left(\frac{\Delta X - \Delta X_0}{\Delta X_{\max} - \Delta X_0}\right)^{\frac{\Delta X_{\max}-\Delta X}{\lambda(E)}} e^{\frac{\Delta X_{\max}-\Delta X}{\lambda(E)}} \tag{3.5}$$

where $\lambda(E) = \lambda_0 + f_\lambda \log_{10}(E/10^{19}\,\mathrm{eV})$. $f_\lambda = 0$ for the muonic component and for the muon decay products. $S_{\max}$, $\Delta X_{\max}$, $\gamma$, $\lambda_0$ are parametrized for each component as a function of the core distance. The fluctuations of the signal $S_0$ are obtained by the distributions of the signals within $\Delta X$ windows of $\pm 20\,\mathrm{g\,cm}^{-2}$ and they are almost independent of the energy.




**(a)** $\mu$ **(b)** $e\gamma$

**Figure 3.7:** *Longitudinal evolution of the signal in an ideal detector for the muonic (a) and the pure electromagnetic (b) components, in the case of* $10^{19}$ *eV proton-initiated showers at different angles* $\theta$*:* 12° *(black dots);* 25° *(red);* 36° *(blue);* 45° *(green);* 53° *(yellow);* 60° *(magenta). From [146].*

As examples, the longitudinal evolution of $S_0$ for the muonic and the pure electromagnetic components are shown in figure 3.7 for proton primaries of $10^{19}$ eV, at a core distance of 1000 m and for different zenith angles, represented with different colors. The black line is the result of the fit for each component. The attenuation of the signal due to the muonic component is weaker with respect to the electromagnetic one. $e\gamma(\mu)$ and $e\gamma(\mathrm{had})$ behave like the muonic and the pure electromagnetic components, respectively.

### *Conversion to a realistic signal with $f_{conv}$*

The factor $f_{\rm conv}$ takes into account the changes in the detector response and in the projected area with the incoming direction of the shower particle and the truncation of the distribution of incoming directions of the secondary particles due to the presence of the ground. In [127] it is found that the signal in a real detector is given by

$$f_{\rm conv} = \int_{p_z^{\rm cut}}^{1} \frac{dS_0}{dp_z}(p_z, r, \Delta X) \cdot A_{\rm mod}(\theta_p) \cdot T_{\rm mod}(\theta_p, r, \Delta X, p_z) dp_z \tag{3.6}$$

where $p_z$ is the cosine of the angle between the particle direction and the shower axis, $\theta_p$ is the zenith angle of the shower particle, $p_z^{\rm cut}$ is the value of $p_z$ at $\theta_p^{\rm cut} = 90^\circ$ (no upward-going particles due to the presence of the ground), $A_{\rm mod}$ is the projected area of the detector in the shower particle direction and $T_{\rm mod}$ is the convolution of the energy spectrum with the detector response.

### *The atmospheric effects factors $f_{mod}$ and $f_{atm}$*

Atmospheric effects affect the signal size. The electromagnetic signals are modified by daily and seasonal modulations of the air density a couple of radiation lengths above ground, which in turn change the Moliere radius and thus the lateral spread. The instantaneous density variation with respect to the average one is accounted for in $f_{\rm atm}$. Its normalization, $f_{\rm type\ I}$, has been derived for the three e.m. components in [127] and it is shown in the left panel of figure 3.8 as a function of core distance.

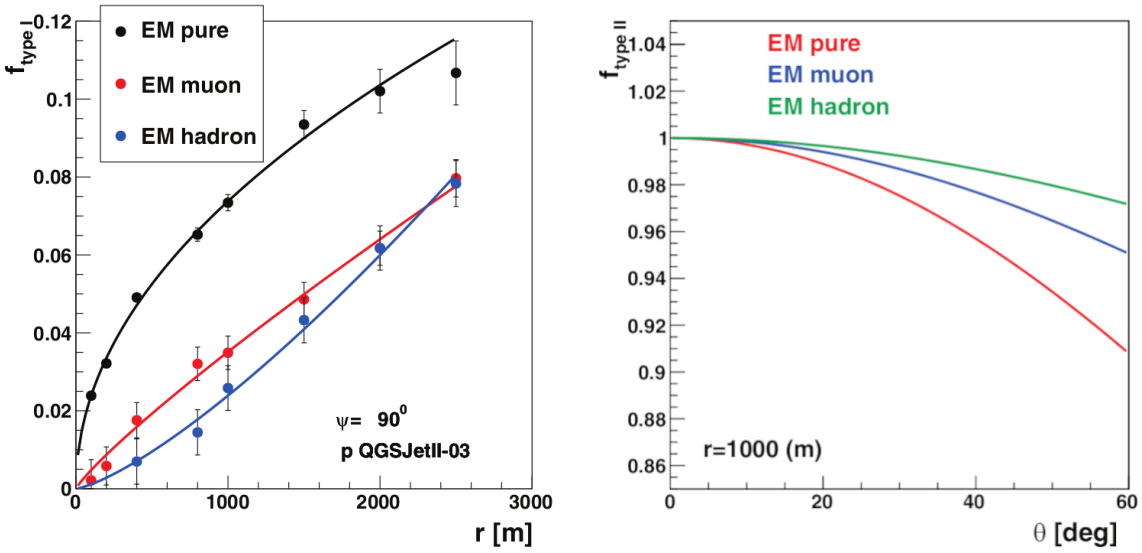



**Figure 3.8:** *Correction factors for the atmospheric effects, see text. From [127].*

The changes in air density depend also on the zenith angle of the shower, thus affecting differently the signal at different $\theta$: this dependence is accounted for by an additional term, $f_{\rm type\ II}$,

shown in the right panel of figure 3.8 as a function of the zenith angle. This description is however pertinent only for the pure e.m. component, which has diffusive characteristics. This is not appropriate for the e.m. component originating from muon decays, as the decay probability scales with traversed geometrical distance instead of integrated density. Also, the e.m. component due to low-energy hadron component has a jet-like behavior, as it has been explained above. These effects too have been studied in [127] and are accounted for in the term $f_{\mathrm{mod}}$.

### *Correlation with the relative muonic composition* $f_{F_\mu \mathrm{fluct}}$

Figure 3.9 shows the correlation between $S_{0,i}/S^{\mathrm{ref}}_{0,i}$ and $S_{0,\mu}/S^{\mathrm{ref}}_{0,\mu}$ where $S^{\mathrm{ref}}_{0,i}$ is the parameterized value of $S_{0,i}$, with QGSJetII-03 proton simulations as reference.

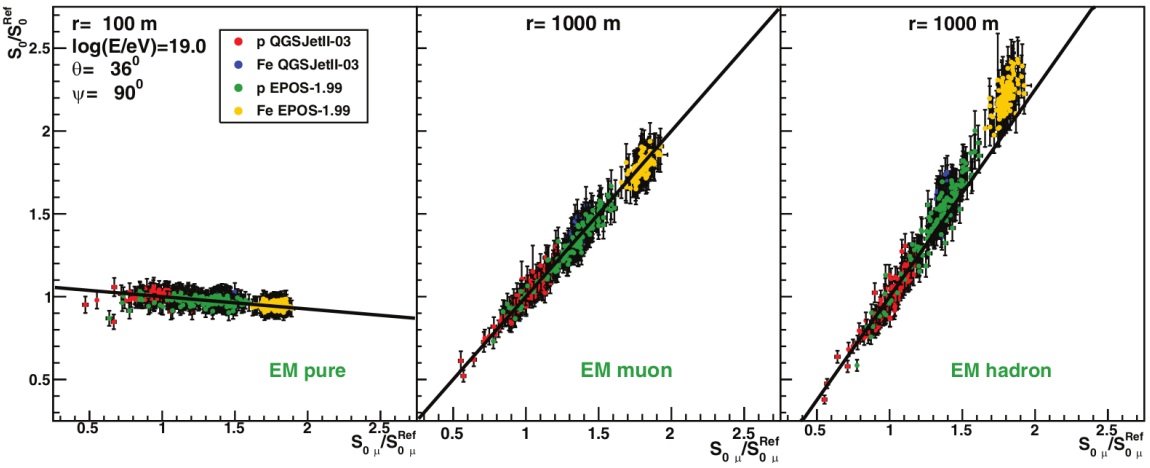



**Figure 3.9:** *Correlation between $S_{0,i}/S^{ref}_{0,i}$ and $S_{0,\mu}/S^{ref}_{0,\mu}$. Left: pure electromagnetic component, signal at a core distance $r = 100\,\mathrm{m}$. Middle: electromagnetic component from muons, signal at $r = 1000\,\mathrm{m}$. Right: electromagnetic component from low-energy hadrons, signal at $r = 1000\,\mathrm{m}$. The zenith of the simulated showers is $\theta = 36^\circ$ and the energy is $E = 10^{19}\,\mathrm{eV}$. Different colors represent different primaries and hadronic models. From [127].*

The solid line is a fit to the reference model and primary and corresponds to

$$\frac{S_{0,i}}{S^{\mathrm{ref}}_{0,i}} = 1 + \alpha_i \left( \frac{S_{0,\mu}}{S^{\mathrm{ref}}_{0,\mu}} - 1 \right) \tag{3.7}$$

For the pure electromagnetic and the $e\gamma(\mu)$ components, $\alpha$ does not exhibit a dependence on the core distance: the derived values of $\alpha$ are respectively $-0.075$ and 1. In the case of the electromagnetic component from low energy hadrons instead, $\alpha$ is parameterized as

$$\alpha_{e\gamma(\mathrm{had})} = 1.25 - 0.13 e^{-6 \frac{r}{1000\,\mathrm{m}}} \tag{3.8}$$

Finally, the term by $f_{F_\mu \mathrm{fluct}}$ is given, for each e.m. component, by

$$f_{F_\mu \mathrm{fluct}} = 1 + \alpha (F_\mu - 1) \tag{3.9}$$

## 3.4 UNIVERSALITY-BASED PARAMETERIZATION OF THE SIGNAL SHAPE

The same scheme used for the parametrisation of the signal size, illustrated in the previous section, can be used also to describe the arrival time distribution of the different shower components as measured in the WCD. This is described in the following, based on [128].

Muons and electromagnetic particles have different interactions in the atmosphere, and thus they propagate differently. Because the muons are relatively energetic, the effects of velocity difference, of deflections in the geomagnetic field and of Coulomb scattering, are small. By contrast, the electrons and photons of an air shower have mean energies of about 10 MeV so that the arrival of the electromagnetic component of the shower is delayed with respect to the muons because of the multiple scattering of the electrons. The delay of the electromagnetic component with respect to the muons also increases with distance. The timing distribution of the different components depends not only on the core distance, but also on $X_{\max}$. Like for the model of the signal size, the parametrization of the signal timing is obtained by Monte Carlo simulations.

The FADC traces of the WCD are simulated in bins of 25 ns, based on the properties of the electronics of the surface detector. The trace observed in a WCD depends on the timing distribution of the secondary particles and on the response of the detector to each component. In particular, it depends on the type, the energy, the momentum and the arrival direction of the particle. A vertical centered through-going muon with energy of 1 GeV generates on average $\sim 80$ photoelectrons (PEs) in a PMT. For a muon with an arbitrary direction, the number of PEs depends mostly on the track length inside the detector, i.e., on the angle and the position of incidence. The number of measured PEs fluctuates with a Poisson distribution due to the statistics of the photon Cherenkov production along the track inside the detector. The photons produced inside the detector are either reflected or absorbed before being detected by the PMT[1]. The typical signal of a particle in the WCD has a fast rise in the first $\sim 15$ ns followed by a slower quasi-exponential decrease, with a time constant of about $\sim 70$ ns, determined by the absorption coefficient of the light in the water and on the diffusive walls of the tank. Thus, the analytic description of the expected traces from a muon is a convolution of several approximated distributions. A semi-analytical model of the detector response to a single vertical centered through-going muon is reported in [147], although the fluctuations are difficult to be reproduced.

As ansatz for the parameterization of the shape of the signal $S(t)$ a log-normal (3.10) or a generalized gamma (3.11) distributions are assumed:

$$f(t;m,s) = \begin{cases} \frac{1}{t \cdot s\sqrt{2\pi}} \exp\left(-\frac{(\ln(t)-m)^2}{2s^2}\right) & t > 0 \\ 0 & t < 0 \end{cases} \tag{3.10}$$

[1] The PMT efficiency is $\sim 30\%$

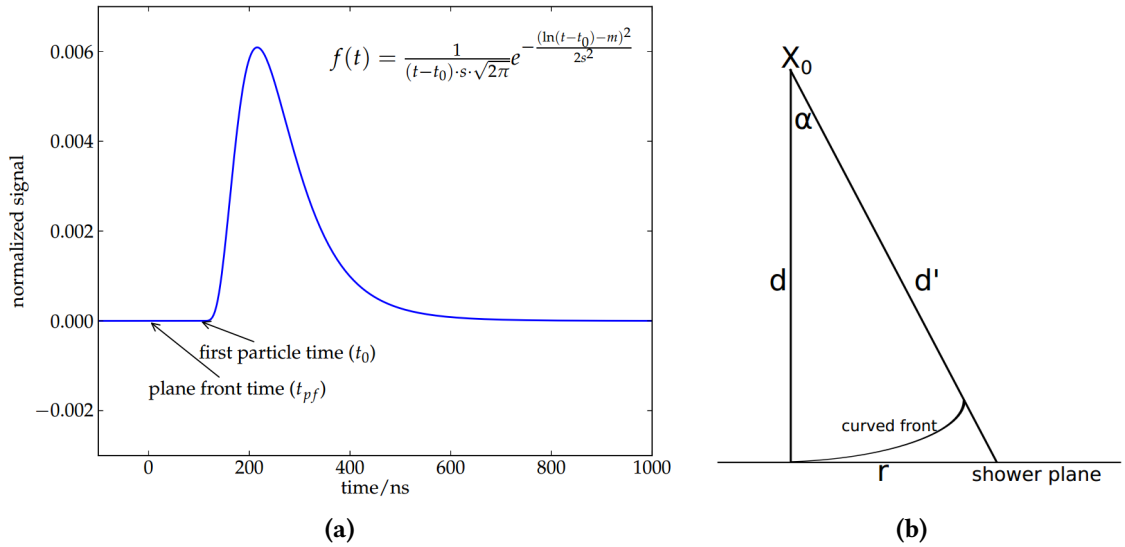



**Figure 3.10:** ***(a)****: Functional form of the signal parameterized with a log-normal function (eq. 3.10).* ***(b)****: Geometrical scheme used to derive the arrival time of the particles at a WCD. From [146].*

$$f(t;m,s,l) = \begin{cases} \frac{|l|}{t\cdot s}\frac{1}{\Gamma\left(\frac{1}{l^2}\right)}\exp\left(\frac{l\frac{\ln(t)-m}{s}+\ln\left(\frac{1}{l^2}\right)-\exp\left(l\frac{\ln(t)}{s}\right)}{l^2}\right) & l\neq 0 \\ \frac{1}{t\cdot s\sqrt{2\pi}}\exp\left(-\frac{(\ln(t)-m)^2}{2s^2}\right) & l=0 \end{cases} \tag{3.11}$$

While the log-normal has two shape parameters $m$ and $s$ (proportional to the mean and the standard deviation), the generalized gamma distribution has an additional parameter $l$. For $l = 0$ the 3.11 distribution reduces to a log-normal with equivalent parameters $m$ and $s$. Depending on the quality of the description of the traces desired, either of the two functions is chosen (log-normal example in figure 3.10a).

All times are expressed relative to the time when the shower core hits the ground. For the parameterization of the traces, the start times are calculated relative to the arrival time of the first particle that reach the station. The shower front can be parameterized as a parabola with different curvatures depending on the shower component considered (see figure 3.10b). The arrival time of the first particle, $t_0$, can be calculated from the distance of the WCD to the shower first interaction point $X_0$ ($d'$ in the figure), that in the simulation is known. Referring again to the figure 3.10b, the time delay $c\delta t$ between the plane and the curved shower front can be obtained as

$$c\delta t = d' - d \simeq \frac{r^2}{2d} \Rightarrow t_0 = \frac{1}{2c}\frac{r^2}{|\vec{P}_{\text{station}} - \vec{P}_{\text{first}}|} \tag{3.12}$$

where $\vec{P}_{\text{station}}$ is the position of the station and $\vec{P}_{\text{first}}$ is the position of first interaction point.

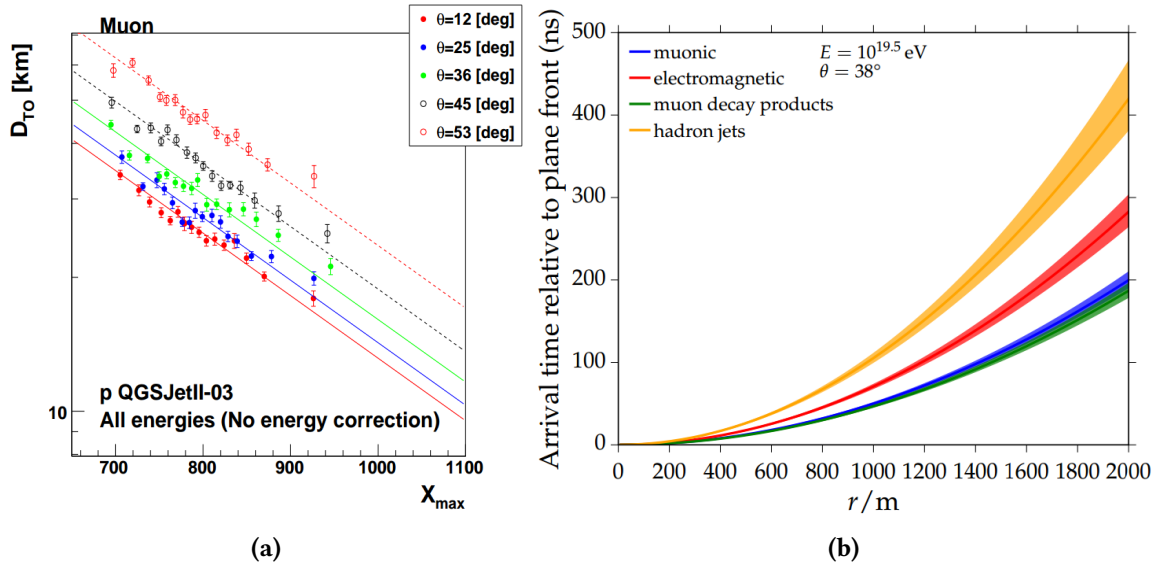



**Figure 3.11:** ***(a)**: Distance in km between the first interaction point and $X_{max}$ as a function of the shower maximum for the muonic component in QGSJetII-03 proton-showers with different zenith angles. **(b)**: Start time model as a function of the core distance. Times are given relative to the arrival time of a plane front traveling at the speed of light c. The model is plotted for a fixed energy and zenith angle as denoted in the plot. Colors indicate the different shower components. The shaded bands around each model line represent the variation due to different azimuthal angles. From [146].*

The difference in height between the first interaction point and the depth of the shower maximum, $X_{\mathrm{max}}$, is a relevant parameter, that depends on the particle component and the properties of the primary particle. As an example, figure 3.11a shows the height difference between $X_0$ and the shower maximum as a function of $X_{\mathrm{max}}$ for the muonic component in proton showers at different zenith angles. One can note the logarithmic decrease with $X_{\mathrm{max}}$, that is connected to the energy spectrum of the muons that reach the detector.

The actual start times from the curvature model are plotted in figure 3.11b for the different shower components. As muons propagate through the atmosphere almost without interaction, they arrive (blue band) before the more frequently scattered electromagnetic particles, except for the muon-decay products (green band) that, when the muons decay close to ground, arrive almost at the same time (if they appear to arrive even earlier in the figure, this is due to small inaccuracies of the model).

Finally, examples of fits to the time-shape of signals in WCDs from showers initiated by different primaries (proton, carbon, iron), are illustrated in figure 3.12, where each panel shows a different component. In this case, for primary energies of $10^{19}$ eV and an average distance of 700 m from the core, the log-normal model works well, and the fits with a generalized gamma distribution yield no improvement.

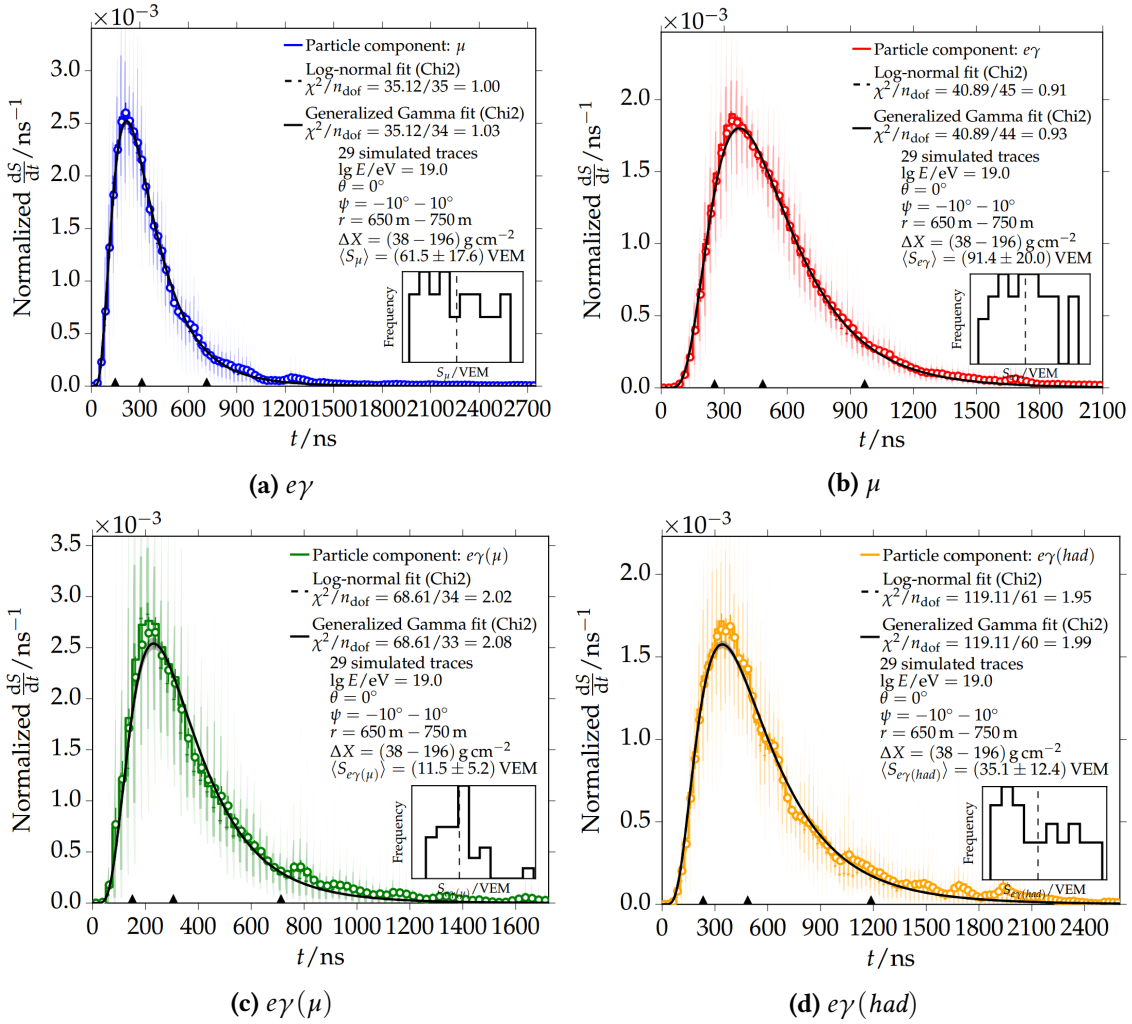



**Figure 3.12:** *Average time distributions of ground signals components in air showers initiated by primary particles with energy of* $10^{19}$ eV *and zenith angle* $\theta = 0°$. *The simulation sample includes proton, carbon and iron as primary species. From [146].*

## 3.5 UNIVERSALITY-BASED RECONSTRUCTION OF THE SURFACE DETECTOR DATA

In this final section, the application of the universality-based model to the reconstruction of the showers detected with the Auger surface detector is shortly outlined.

The model, as explained above, allows for the description of the size and the time-shape of the signals, generated in the WCDs by the secondary particles in an air-shower, as a function of physical variables related to the properties of the primary particle. These parameters are the

primary energy $E$, the depth at the shower maximum $X_{\max}$, the relative muon content $F_\mu$ and the shower geometry. Hence, by a fitting procedure of the model on the surface detector signals, the physical parameters which best reproduce the data can be retrieved [143, 146, 148].

The standard reconstruction of the SD data (summarised in section 2.1.4) provides the initial values for two of the parameters, namely the geometry and the energy. Then, in the universality-based reconstruction, a likelihood fit with multiple parameters is performed, namely the core position, the arrival time of the core, the zenith and azimuth angles, the depth of the shower maximum and of the first interaction point, the number of muons and the primary energy. The parameters can either be let free, or partially fixed. Specifically, two modes are implemented. In the so-called *iterative-reconstruction* mode, there are different steps; in each of them some selected physical quantities are fitted, while others are fixed or constrained to the SD reconstructed parameters or to the results of a previous step. In the so-called *global reconstruction* mode instead, only the energy is fixed, namely to the value provided by the SD reconstruction, while all the other parameters are fitted simultaneously.

In both the iterative and global reconstructions, the total likelihood of the fit is calculated from two contributions. One derives from the the *LDF fit*, where the signal sizes of the SD stations are fitted with a lateral distribution function resulting from the sum of the LDFs of each shower component. The second one derives from the *time-shape fit* where the FADC traces are fitted, bin by bin, to the sum of 4 parameterized log-normal distributions, due to each component.

The universality-based reconstruction has been originally oriented to analyze the SD highest energy events ($E > 10^{19}$ eV) because of their large station multiplicity. With time, the method has been extended to SD events of lower energies, down to $10^{18.5}$ eV. The quality and the stability of the reconstruction however decreases with the number of triggered SD stations decreasing. In chapter 5, an alternative reconstruction technique will be introduced, that will allow for the use of the university-based the model at even lower energies (down to $10^{18}$ eV) by inferring accurate information from even one station only. As it will be shown, this is possible by exploiting the universality concept in combination with the the cosmic-ray parameters reconstructed by the fluorescence detector, i.e., the energy and the depth of the shower maximum, in order to determine the relative muon content $F_\mu$.

# 4 HYBRID EVENTS: DATA AND SIMULATIONS

## Contents



This chapter details the construction of the Auger data set to be used in this thesis work for the search for UHE primary photons. For such an analysis, a crucial element is the detailed study of the development of air-showers induced by different primary particles and of the properties of such showers as observed in the Auger detectors. No UHE photons have been so far unambiguously identified in data. Therefore, Monte Carlo (MC) simulations are required, so to develop improved analysis techniques - such as that presented in the next chapters - with the aim of discriminating photon-induced showers from the most numerous hadron-induced ones.

In this thesis work *hybrid* data are used, i.e., those events that have been detected with the FD in coincidence with at least one SD station. Section 4.1 describes the initial Auger data set, from which the final sample will be extracted. Section 4.2, in turn, details the procedure followed in order to obtain a simulation sample representative of the Auger data. Both sets, data and simulations, contain however all events observed by the FD and the SD, regardless of the quality of the reconstruction, if any. The objective of section 4.3 is that of defining and explaining the quality cuts needed to guarantee the quality of the hybrid reconstruction. Finally, the quality cuts are applied to data and simulations: the final data sets, used in the rest of this thesis work, are described in section 4.4.

## 4.1 THE INITIAL HYBRID DATA SET

The data taking of the Pierre Auger observatory is an automated process, combining different systems. On the one hand, the Central Data Acquisition System (CDAS), which runs continuously in the Computer Center in Malargue, manages and assembles the triggers from the SD detectors. On the other hand, local acquisition systems, installed at each FD site, record FD data, which are then transferred to the Computer Center. Although the FD data acquisition is independent from the CDAS, hybrid coincidences are identified on line within the SD data stream. The merging of the FD and SD raw data into hybrid raw data is made offline [149], by combining events characterized by a TLT FD trigger (see section 2.2.2) with at least one T1 SD trigger (see section 2.1.3). An automated event reconstruction is then applied to the hybrid raw data, based on the Auger Observer framework [150].

Data taking started in 2004 with 154 SD detectors and one FD site, Los Leones, in operation. The rest of the SD stations and FD sites were installed gradually: the observatory was completed in June 2008 and running has been on-going since that date. The analysis presented in this thesis work is based on hybrid data collected from 1 January 2005 to 31 December 2017. The end date is defined by the fact that the atmospheric databases relative to year 2018 and following, necessary for the reconstruction of FD events, were not yet made available to the Collaboration at the time of this thesis work. The initial data set consists of almost 3 millions shower-candidates events. Given that for the automated reconstruction it is sufficient that only very loose criteria are fulfilled, a further selection of these events needs to be applied, which is explained in details in section 4.3.

| **Initial Data Set** | **Number of events** |
|---|---|
| Full sample | 2990303 |
| Burn sample | 149428 |
| Search Sample | 2840875 |

**Table 4.1:** *Number of events in the full, burn and search hybrid data samples.*

It is important to note that such a selection results in a final data set (described in section 4.4) with quality requirements generally suitable for Auger analyses. In turn, a specific selection and analysis for the search of UHE photons is adopted in this thesis work (see chapter 6). The selected search strategy is, in particular, that of a *blind analysis* [151]. This is an optimal way to reduce or eliminate biases, due to experimenters expectations, in the results in the search for rare events. Information which may influence the results, such as the data, are thus withheld until the analysis is concluded. In this thesis work, in particular, a sub-sample of the data, corresponding to 5 % of the total and called *burn sample* (BS), will be used to study the analysis technique and the background. To form the BS, the event identifier (ID) is used: the selected events have an ID that is multiple of 20. The BS is only used in the study of the performances of the photon search analysis and is excluded from the photon search. Table 4.1 reports the number of the hybrid events in the three initial data sets, i.e., full, burn and search samples.

## 4.2 THE INITIAL HYBRID SIMULATED-DATA SETS

The simulation of events in the Pierre Auger Observatory is a two-step process: first the propagation of extensive air showers within the atmosphere is simulated, then the generated showers are used as input for the simulation of the detector response. Showers are simulated with the Monte Carlo simulation software CORSIKA as described in section 4.2.1, while the simulation of the detector response is based on the Auger Offline Software framework as documented in section 4.2.2.

### 4.2.1 Shower simulation

The simulated showers (*shower library*) used in the following have been produced using CORSIKA 7.6400 [152], a software that handles simulation of particle cascades in the atmosphere. The secondary particles created in an extensive air shower are tracked explicitly until they interact with other particles or decay. These processes are treated according to the current state of knowledge. At low energies, i.e., below $\sim 80\,\mathrm{GeV}$, hadronic interactions are handled by the MC generator Fluka2011.2b.6 [153]. At higher energies, different models can be chosen to describe the hadronic interactions, such as Sybill [55], QGSJet [56] or EPOS-LHC [54].

In this thesis work, simulations are produced with EPOS-LHC (acronym for Energy conserving quantum mechanical multi-scattering approach, based on Partons, Off-shell remnants and Splitting parton ladders) that is a model tuned to describe the largest body of accelerator data. EPOS is based on a microscopic pomeron model in which the pomeron-parton coupling, including momentum sharing, is explicitly calculated [154]. To obtain a better description of the data, the authors introduced parameterizations to modify the baseline predictions of the model where needed [154]. Hadron production in heavy-ion collisions measured at RHIC is very important for tuning this model. Also, the implementation of collective flow effects in interactions with high parton density is unique to EPOS. The reason for the choice of using this model of hadronic interaction is that the average number of muons simulated in an air-showers is closer to the number observed in the data with respect to the other models [155].

The number of secondary particles produced in an air cascade is proportional to the primary energy, $E_0$, therefore the simulation of an high-energy shower can be very time-consuming. To reduce the computing time required, a thinning procedure is introduced. All the particles emerging from an interaction below an adjustable fraction of the primary energy, $\epsilon_{th} = E/E_0$, where $E$ is the energy of the parent particle, are subjected to the thinning algorithm. Only one of these particles is followed, with an associated weight factor $w_i$, considered to ensure the energy conservation.

A CORSIKA simulation is steered via an input file, where the parameters for the simulation, such as the primary energy, the zenith angle or the thinning fraction, are set. An example of an input file used for the simulations that are discussed in this thesis is reported in appendix A. Two simulation samples have been generated, using photons and protons as primary particles. Only protons have been considered, because, being the lightest nuclei, hence the most penetrating ones

in the atmosphere, they constitute the main source of background in ultra-high-energy photon searches.

| Energy range | Protons | Photons |
|---|---|---|
| $10^{17.5}\,\text{eV} < E < 10^{18.0}\,\text{eV}$ | 10000 | 10000 |
| $10^{18.0}\,\text{eV} < E < 10^{18.5}\,\text{eV}$ | 5000 | 10000 |
| $10^{18.5}\,\text{eV} < E < 10^{19.0}\,\text{eV}$ | 5000 | 10000 |
| $10^{19.0}\,\text{eV} < E < 10^{19.5}\,\text{eV}$ | 5000 | 10000 |

**Table 4.2:** *Number of showers generated for each energy bin and primary particle.*

The characteristics of the generated showers are the following:

- The energy ranges from $10^{17.5}$ eV to $10^{19.5}$ eV. This range is because above $10^{18.0}$ eV the hybrid trigger efficiency for the standard hybrid detector is full for hadron primaries, while above $10^{19.5}$ eV the sensitivity of the SD only detector is higher.

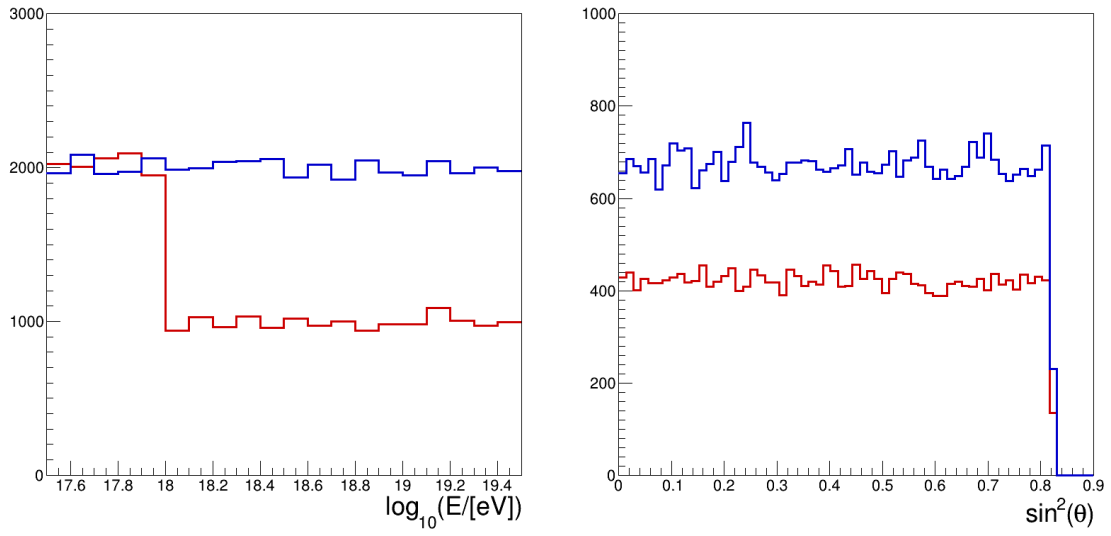
- The energy distribution follows a power law spectrum $E^{-\Gamma}$ with $\Gamma = 1.0$ (see figure 4.1a). A flat spectrum in $\log_{10}(E)$ is a compromise between the much steeper real spectrum and the computing time, thus allowing to obtain a large enough simulation sample also at the highest energies. Simulations are actually performed in 4 energy sub-ranges with a constant width of 0.5 in terms of $\log_{10}(E[\text{eV}])$. The number of showers generated in each energy bin for each primary particle is reported in table 4.2. The larger number of simulations in the first energy sub-range is due to tha fact that proton-initiated and photon-initiated showers at these energies trigger the SD detector less than proton showers above $10^{18}$ eV, due to the lower number of secondary particles generated during the shower development and to the smaller footprint on the ground.

- The zenith angle is distributed according to a $\sin\theta\cos\theta$ distribution from 0° to 65°, corresponding to an isotropic flux projected on a flat experiment (figure 4.1b). More inclined showers are not included because of the absorption of the electromagnetic component in the atmosphere and of the resultant small trigger efficiency for photons at the lowest energies.

- Four different atmospheric profiles are used for the simulation, representing the four season of the year at the Malargüe site, namely January for the summer, March for the autumn, August for the winter and September for the spring. This serves to account the seasonal effects on the shower development, which is affected by the strong seasonal variations of air density and temperature.

- The used thinning fraction is $\epsilon_{th} = 10^{-6}$, which is the most common choice [156].

Every CORSIKA simulation outputs two files. The first one, containing information about the longitudinal development of the shower, is used to simulate the FD detector response. The sec-




**(a)** **(b)**

**Figure 4.1:** *Distributions of the energies (4.1a) and zenith angles (4.1b) of the generated showers for primary photons (blue histograms) and protons (red histograms).*

ond one, containing information related to the position, energy and momentum of the secondary particles that reach the ground, is used for the simulation of the SD detector response.

### 4.2.2 Simulation of the detector response

The Auger Offline software framework ( $\overline{\text{Off}}\underline{\text{line}}$) [157] is used to simulate the responses of the FD and SD detectors to the simulated showers generated with CORSIKA. Like for the data, it is also used to reconstruct the hybrid simulated data.

The $\overline{\text{Off}}\underline{\text{line}}$ comprises three principal parts: 1. a collection of processing modules which can be assembled and sequenced through instructions provided in an XML file; 2. an event data model designed to handle both raw data and simulations, which, on the one hand, mimics the structure of the detector systems, and, on the other hand, collects all the simulation/reconstruction information; 3. a detector description which provides a gateway to data describing the configuration and performance of the observatory, as well as atmospheric conditions, as a function of time, Each processing module is independent and has a specific task. Different simulation/reconstruction chains can be implemented by combining the modules in the so-called “Module Sequences”. The simulated/reconstructed events at the end of the chain are exported in Advanced Data Summary Tree (ADST) files [158]. The ADST file format is based on the ROOT framework [159], and contains all the events information, from low-level to high-level quantities. The simulation in this work

is implemented with the module sequence reported in appendix B, and, in particular, the actual status and the time evolution of the detector are included.

The *FD simulation* [160] reproduces all the physical processes involved in the fluorescence technique. It includes the generation of fluorescence and Cherenkov photons in the atmosphere, their propagation through the air to the telescope aperture, the ray-tracing of photons in the Schmidt optics of the telescopes, and the simulation of the response of the electronics and of the multi-level trigger. The FD simulation has been designed to reproduce the actual sequencing of the detector status with a resolution of 10 min. First a time is randomly generated within the sidereal time interval considered. Then the status information about each telescope is retrieved from a file where the information, of the electronics, DAQ and communication systems are tabulated. Moreover, the data from the atmospheric monitoring system is used to set the hourly aerosol density profile and the monthly mean molecular atmosphere. Based on the actual status of the detector, only a sub-sample of the events is subjected to the detector simulation. In the case of FD time-dependent simulations, the values of the PMTs variance, baseline and trigger threshold averaged over 10 min are considered.

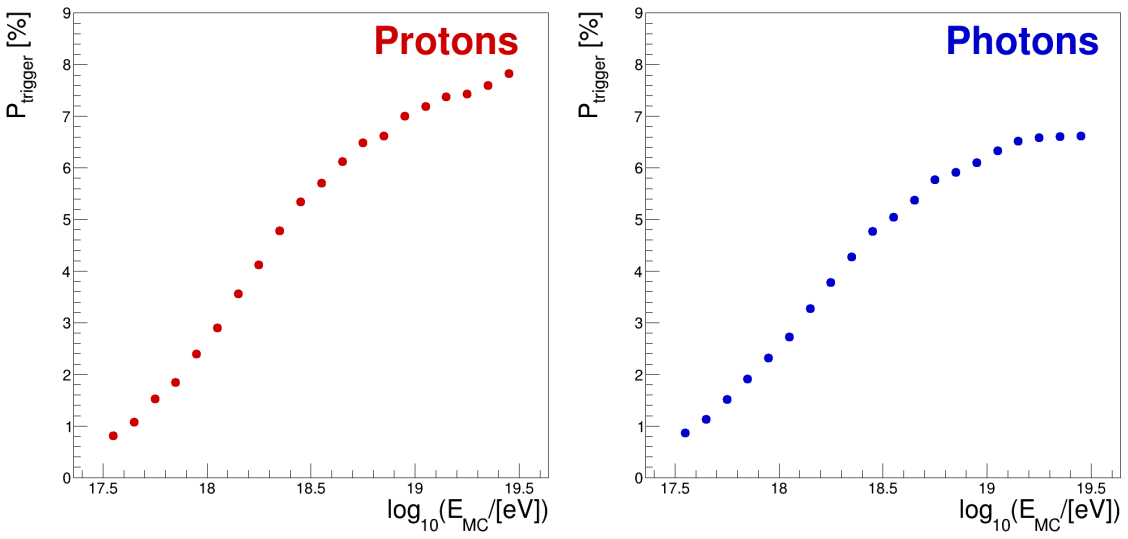



**Figure 4.2:** *Hybrid-trigger probability as a function of energy for simulated showers initiated by protons (4.2a) and photons (4.2b).*

The *SD simulation* is based on Geant4 [161]. A dedicated processing module produces Cherenkov photons along the path of the injected particle and tracks them through the water in the SD tank until they are absorbed or reach the active photo-cathode area of a PMT. Thus, the generated photo-electrons as a function of time are processed by a different module simulating the PMTs and electronics response. To account for the growth of the array with time and for the down-

time of stations, either due to malfunctioning or maintenance, during data taking, only active SD detectors are considered during simulation.

To increase the number of detector simulations, the detector response is simulated using every CORSIKA shower multiple times, by assigning to it a random core in a $80\,\mathrm{km} \times 80\,\mathrm{km}$ square centered on the surface array. Such surface is larger than the actual one, so to properly sample edge effects. A random time between 01 January 2005 00:00:00 and 31 December 2017 23:59:59 is assigned to each of these simulations. To avoid a waste of computation power, the status of the involved FD is checked at the randomly generated time. The simulation is performed only if the FD is in operation, otherwise the event is discarded, being accounted only as "generated event" in the final sample[1].

| **Energy range** $[\log_{10}(E/[\mathrm{eV}])]$ | **Shower usage** |
|---|---|
| 17.5 – 17.6 | 633 |
| 17.6 – 17.7 | 466 |
| 17.7 – 17.8 | 346 |
| 17.8 – 17.9 | 273 |
| 17.9 – 18.0 | 211 |
| 18.0 – 18.1 | 171 |
| 18.1 – 18.2 | 141 |
| 18.2 – 18.3 | 120 |
| 18.3 – 18.4 | 104 |
| 18.4 – 18.5 | 94 |
| 18.5 – 18.6 | 88 |
| 18.6 – 18.7 | 81 |
| 18.7 – 18.8 | 77 |
| 18.8 – 18.9 | 76 |
| 18.9 – 19.0 | 72 |
| 19.0 – 19.1 | 69 |
| 19.1 – 19.2 | 68 |
| 19.2 – 19.3 | 67 |
| 19.3 – 19.4 | 66 |
| 19.4 – 19.5 | 64 |

**Table 4.3:** *Number of times that the simulation of the detector response is performed using the same CORSIKA shower (shower usage) as a function of the energy.*

The number of times a single shower is injected in the detector simulation, (*shower usage*) is shown in table 4.3 for 20 energy bins. The shower usage is calculated starting from the probability $\mathcal{P}$ that a shower, with a certain energy and angle, triggers the hybrid detector in the configuration

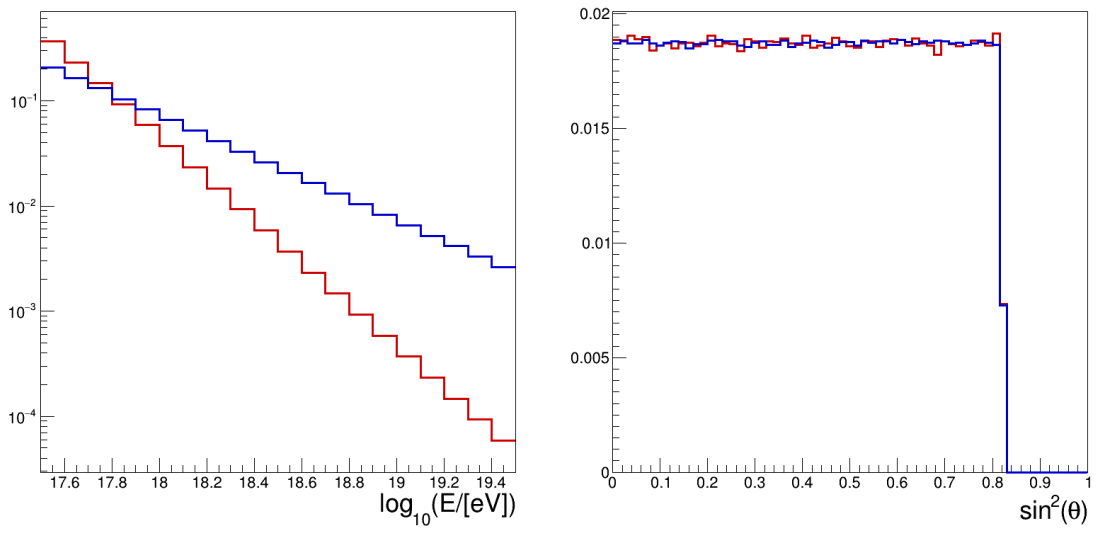
[1] The event must necessarily accounted as "generated" in view of the calculation of the analysis exposure described in section 7.2.

corresponding to a random time between 01 January 2005 00:00:00 and 31 December 2017 23:59:59, and assigning to it a random impact position in an 80 km × 80 km square centered on the SD array. $\mathcal{P}$ is derived from the ratio between the number of triggered events and that that are injected in the detector simulation: it is shown as a function of energy in the two panels a and b of figure 4.2, for primary protons and photons, respectively.

A weight is finally associated to each triggered event, so to obtain a realistic description of the energy spectra. Namely, each weight $w$ depends on the probability, $\mathcal{P}$, and on the true MC energy, $E$, as:

$$w = \mathcal{P} \cdot \frac{E[\mathrm{GeV}]^{-\gamma}}{E[\mathrm{GeV}]^{-1}} \tag{4.1}$$

where $\gamma$ is the spectral index. The spectral index of the cosmic ray spectrum shows small variations with respect to the average value of 2.7 (see section 2.5). Hence, an average spectral index $\gamma = 2.7$ is used to describe the proton energy distribution. The photon spectrum, on the other hand, is characterized by $\gamma = 2.0$, accordingly to the expected spectral index at the accelerating sources.



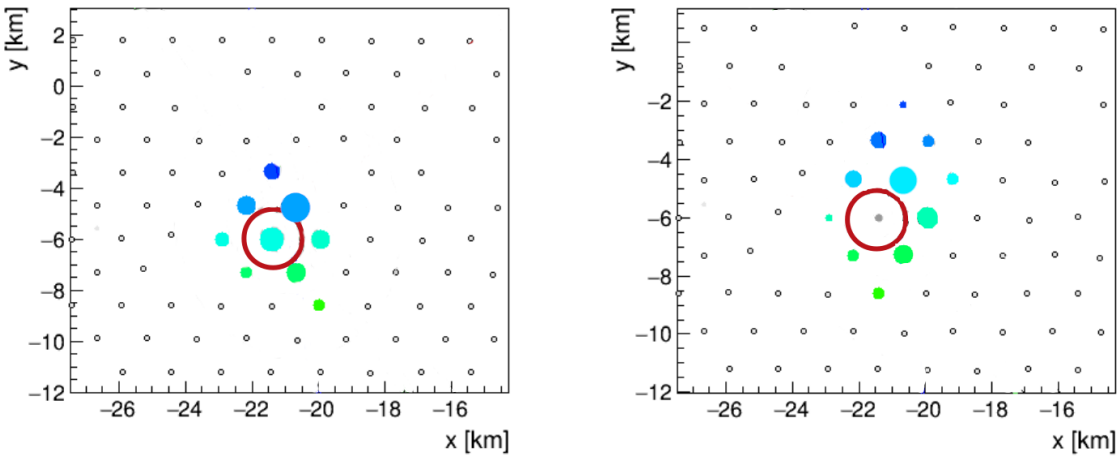

**(a)** **(b)**

**Figure 4.3:** *Distributions of the energies (4.3a) and zenith angles (4.3b) of the simulated showers including the detector response, for primary protons (blue histograms) and photons (red histograms). The number of events is weighted according to the equation 4.1 with $\gamma = 2$ for photons and $\gamma = 3$ for protons.*

Figures 4.3a and 4.3b show, respectively, the energy and zenith distributions of the generated Auger events after the weight correction.

### *Correction to the simulation of the SD detector status*

A by-product of the thesis work on the simulations production is the identification of a bug in the treatment of the status of the SD array, namely, it was found that all SD detectors resulted to be




(a) (b)

**Figure 4.4:** *Illustrations of the bug found in the simulation of the status of the SD stations. In figure 4.4a all the stations are active, while in figure 4.4b, after the bug correction, the station indicated with the red circle inactive stations is in fact non active (grey dot)*

always active. This bug has been corrected through the implementation of an additional $\overline{\text{Off}}\underline{\text{line}}$ module, which now accounts correctly for the status of the SD stations. In figure 4.4a and 4.4b, one can see the effect of the correction. The colored dots correspond to triggered stations: the color represents the trigger time (green to blue) while the size is proportional to the signal size. The station highlighted with the red circle, which resulted to be triggered in the figure 4.4a, before the bug correction, is in fact non operational (gray dot in figure 4.4b) once the actual status of the SD array is accounted for.

## 4.3 DATA SELECTION

The reconstruction of the events that form the hybrid data and the simulation sets, described in the two previous sections, fulfill only very loose criteria. This section details the further selection that is applied to both of them so to provide the quality needed in view of the analysis presented in this thesis work.

The selection cuts explained in the following are derived from other Auger analyses based on hybrid events [90, 114]. They are mostly meant to ensure a good resolution on reconstructed shower parameters, such as energy, $X_{\text{max}}$ and geometry, although few of them are specialized for photons searches. Part of this thesis work has been the optimisation of the cuts with the aim of increasing as much as possible the number of events while preserving an accurate enough reconstruction.

The event selection is divided into four levels: the pre-selection level, the geometry level, the profile level and the atmospheric level. At the pre-selection level, events are rejected if the reconstruction process failed or if they have been recorded during time periods with known detection system problems (e.g., problems with the communication system or with unstable PMTs). At the geometry and profile levels, several cuts are imposed to remove events without a reliable reconstruction of shower geometry and longitudinal profile, respectively. Finally, at the atmospheric level, events are removed if recorded when the atmospheric conditions might have an influence on the measurement (i.e., when clouds obscure part of the longitudinal profile), or when too many aerosols are present in the lower layers of the atmosphere.

### 4.3.1 Pre-selection Level

| Cut | Value |
|---|---|
| eyeCut | 1111 |
| minRecLevel | 1 |
| badFDPeriodRejection | true |
| badSDPeriodRejection | true |
| good10MHzCorrection | true |
| isCLF | false |
| isXLF | false |
| skipSaturated | true |
| hybridTankTrigger | 2 |
| hasMieDatabase | true |

**Table 4.4:** *Selection criteria applied on the simulation and data samples at the pre-selection level*

The selection criteria applied at the pre-selection level are listed in table 4.4.

Events recorded by the HEAT telescopes are removed by *eyeCut*, while *minRecLevel* remove all the events that does not include at least one triggered SD station.

Then, events are discarded if acquired during periods without an absolute calibration for the respective FD telescope or with known detector problems, such as unstable PMTs or erroneous GPS systems, by applying the *badFDPeriodRejection* and the *badSDPeriodRejection* cuts.

Certain problems in the detector electronics of the FD can be corrected offline, for example an erroneous clock on the FLT board. Events for which this correction is not possible are discarded (*good10MHzCorrection*).

Hybrid events require at least one triggered SD station, used for the geometry reconstruction. The

*hybridTankTrigger* cut rejects all the events where the so-called hybrid station was triggered by a MoPS or a TOTd trigger. The signal in such stations is always below 4 VEM, which is too small for the photon-search analysis, as it will be shown in the chapter 5 [162] .

Events with at least one saturated pixel are removed by *skipSaturated*.

Lasers from CLF and XLF are discarded by the *isCLF* and *isXLF* cuts, respectively.

Finally, *hasMieDatabase* excludes all the events where no atmospheric monitoring of the aerosol is available, for the time of the recorded event, since the atmospheric transmission is an important input for the FD reconstruction.

### 4.3.2 Geometry Level

| Cut | Value |
|---|---|
| maxCoreTankDist | 1500 m |
| maxZenithFD | 60° |
| angleTrackLength | 15° |
| HDSpectrumDistance2012 | true |

**Table 4.5:** *Selection criteria applied on the simulation and data samples at the geometry level.*

At the geometry level, events are removed if the reconstructed geometry, which is the basis for the reconstruction of the shower properties, is not accurate enough. The selection criteria applied at the geometry level are listed in 4.5.

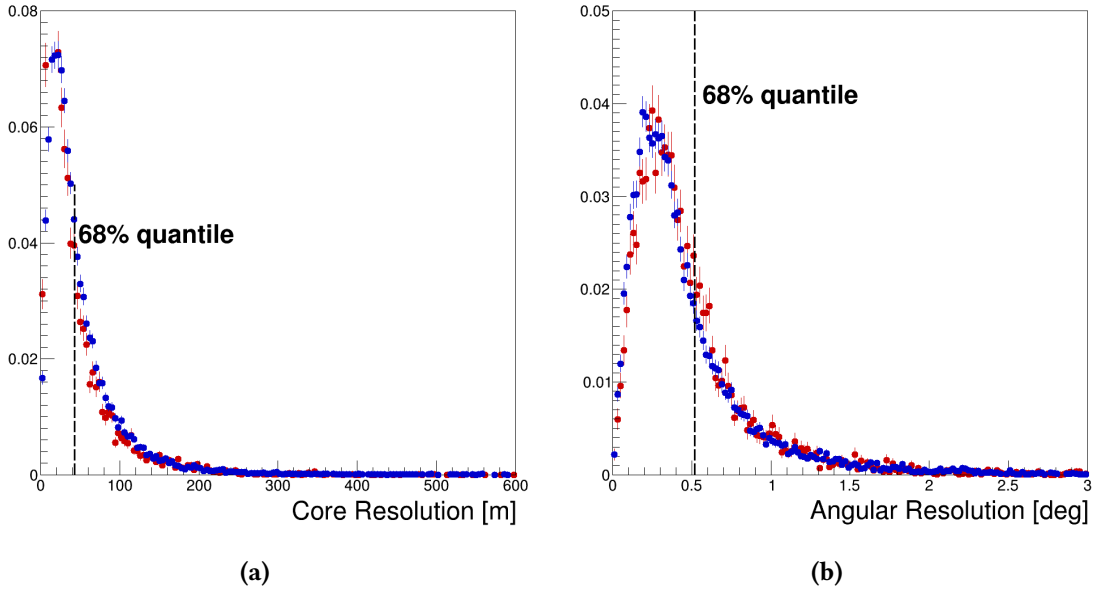
To avoid reconstructing events too far from the boundaries of the SD array, the station selected in the hybrid reconstruction is required to be within 1500 m of the shower axis by using the *maxCoreTankDist* cut.

Only events with zenith angle up to 60° are considered (*maxZenithFD*). More inclined events are not included because of the absorption of the electromagnetic components of the EAS in the atmosphere and the resultant small trigger efficiency for photons in particular at the lowest energies considered in the photon-search analysis.

The angular track length, defined as the angular separation between the highest and lowest FD pixels in the track, is required to be larger than 15° (*angleTrackLength*).

Events are selected, by the *HDSpectrumDistance2012* cut, if they land within a fiducial distance from the telescope for which the FD trigger efficiency is flat within 5 % [163] when shifting the energy scale by its systematic uncertainty, i.e. ±14 % (see section 2.2.4). This distance, parameter-

ized in different energy intervals, is based on simulations and is mostly independent of the mass composition and hadronic models.




(a) (b)

**Figure 4.5:** *Protons (red) and photons (blue) resolution in the core position (4.5a) and the arrival direction (4.5b) reconstructions.*

A resolution better than 40 m in the core position and of 0.5° (figure 4.5a) in the arrival direction (figure 4.5b) are obtained with these cuts for events with energy above $10^{18}$ eV. The resolution in the arrival direction is calculated using the angle between the reconstructed and the true Monte Carlo shower axis, while the core resolution is calculated as the distance between the true Monte Carlo and the reconstructed core position.

### 4.3.3 Profile Level

| Cut | Value |
|---|---|
| profileChi2Sigma | 3.0 −1.1 |
| minViewAngle | 20° |
| xMaxInFOV | 0 |
| XmaxErrorLessThenXmax | true |
| maxDepthHole | 20 % |
| calEnergyError | 20 % |

**Table 4.6:** *Selection criteria applied on the simulation and data samples at the profile level.*

The accuracy in the measurement of the longitudinal profile of a shower affects the resolutions on the reconstructed energy and depth at the shower maximum, $X_{\max}$. The selection criteria applied at the profile level are listed in 4.6.

The goodness of the profile fit (see section 2.2.3) is tested through the *profileChi2Sigma* cut by requiring

$$\frac{\chi^2 - N_{\text{dof}}}{\sqrt{2N_{\text{dof}}}} - \left\langle \chi^2 \right\rangle < \sigma_{\chi^2} \tag{4.2}$$

where $N_{\text{dof}}$ is the number of degree of freedom in the fit. $\left\langle \chi^2 \right\rangle$ and $\sigma_{\chi^2}$ are respectively $-1.1$ and 3.0 [149].

The *minViewAngle* cut requires a viewing angle between the shower axis and the telescope larger than 20°, and rejects events pointing toward the FD and having a large Cherenkov light contamination.

Biases in the reconstruction of the longitudinal profile are limited by requiring an $X_{\max}$ observed in the telescope field of view (*xMaxInFOV*) and gaps in the profile (*maxDepthHole*) shorter than 20 % of the total observed length.

*XmaxErrorLessThenXmax* instead, rejects all the events with a non-converging Gaisser-Hillas fit, while events are selected if the relative uncertainty on the calorimetric energy is smaller than 20 % by applying the cut *calEnergyError*.

These criteria ensure a resolution of the calorimetric energy (figure 4.6a) at level of 5 %, and a bias below 2 %. The $X_{\max}$ resolution (figure 4.6b), instead, is below 14 g cm$^{-2}$ with a bias of about 5 g cm$^{-2}$. Both the calorimetric energy and the $X_{\max}$ resolutions are derived by using the Monte Carlo simulations.

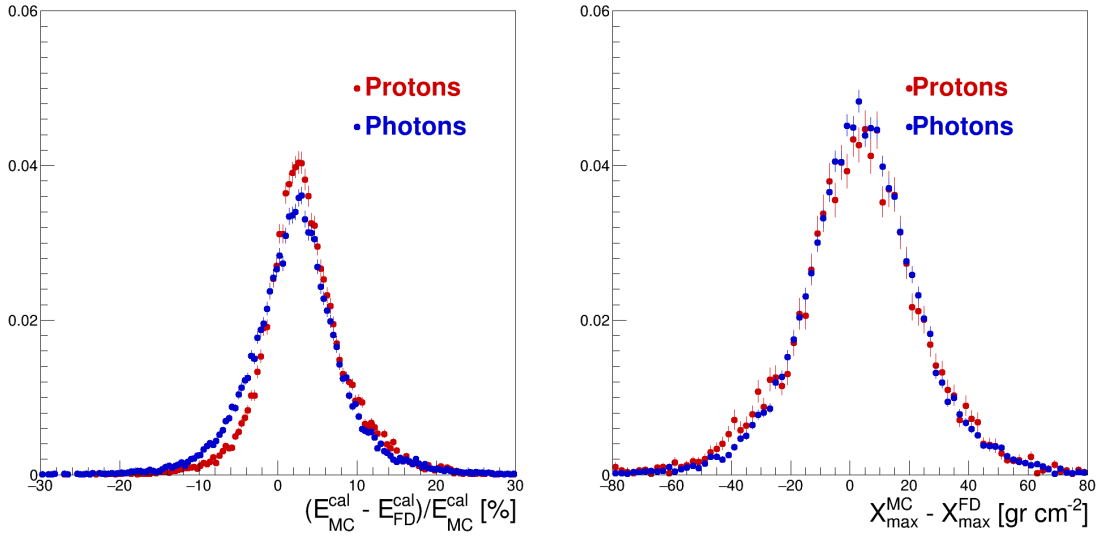
### 4.3.4 Atmospheric level

| Cut | Value |
|---|---|
| cloudCutXmaxPRD14 | 0 |
| maxVAOD | 0.1 |

**Table 4.7:** *Selection criteria applied on the simulation and data samples at the atmospheric level.*

The atmospheric level cuts, listed in table 4.7, are applied to exclude a possible reflection or shadowing of the light from the shower due to clouds, by combining information from the two laser facilities, the lidars and the cloud monitoring devices described in section 2.3.

By using the *cloudCutXmaxPRD14* [164], events are accepted if no cloud is detected along the




**(a)** **(b)**

**Figure 4.6:** *Resolutions in energy (4.6a) and $X_{max}$ (4.6b) as a function of the energy, for showers initiated by protons (red points) and photons (blue points).*

direction to the shower in either the telescope projection (cloud camera) or ground-level projection (GOES). When none of these requirements are met, events are rejected if either the cloud camera or GOES indicates the presence of clouds in their respective projections. When no data from these monitors are available, the event is accepted only if during the data taking the average cloud fraction as reported by lidars is below 25 %.

Time periods with poor viewing conditions are excluded by the *maxVAOD* cut, requiring that the measured vertical aerosol optical depth (VAOD), integrated from the ground to 3 km is smaller than 0.1.

## 4.4 DATA AND SIMULATIONS: THE ANALYSIS SETS

This section finally presents the data and simulations sets that will be used in the following. They are obtained by applying the event selection criteria, described in the previous section, to the samples illustrated in sections 4.1 and 4.2.2.

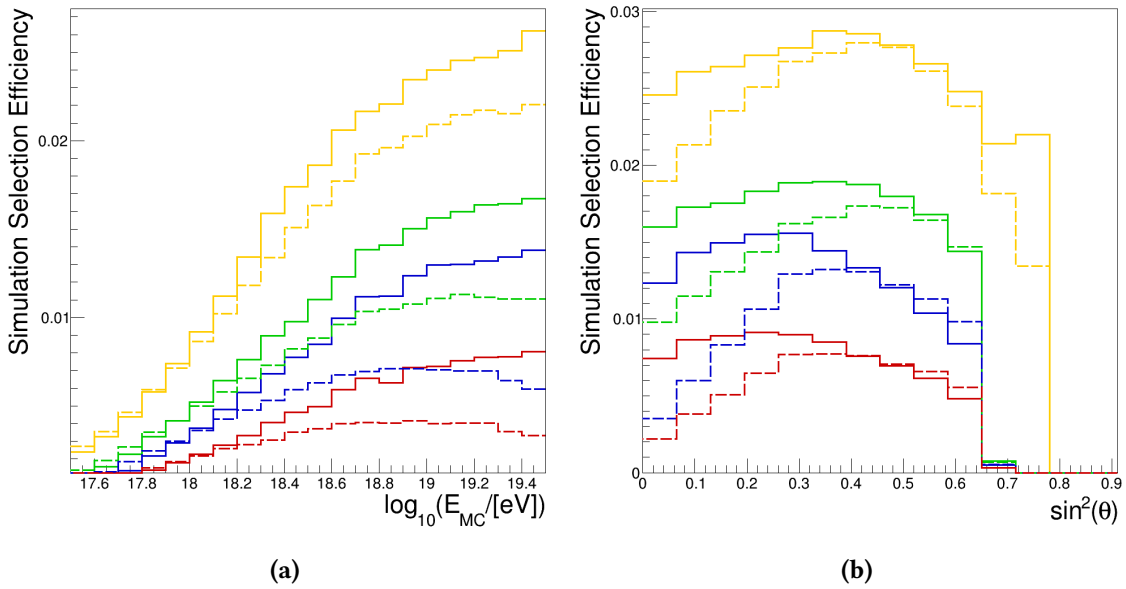
Table 4.8 shows the effect of each on the three sets introduced before, namely the full and burn hybrid data, as well as the proton- and photon-initiated simulated data. Overall, out of the 2990303 events in the full data set, 133741 events are selected, which corresponds to about 4.5 %. In the simulation samples, $\sim$ 0.5 % of the data are selected, resulting in 20593 proton-initiated events and 29187 photon-initiated ones.

| | Data | | | | Simulations | | | |
|---|---|---|---|---|---|---|---|---|
| | Burn Sample | | Full Data Sample | | Protons | | Photons | |
| | $N$ | $\epsilon$ [%] | $N$ | $\epsilon$ [%] | $N$ | $\epsilon$ [%] | $N$ | $\epsilon$ [%] |
| **Raw Sample** | 149428 | | 2990303 | | 2708365 | | 5388615 | |
| minRecLevel | 134836 | 90.2 | 2697253 | 90.2 | 335845 | 12.4 | 644734 | 12.0 |
| badSDPeriodRejection | 131147 | 97.3 | 2624095 | 97.3 | 320186 | 95.3 | 614744 | 95.2 |
| eyeCut | 117329 | 89.4 | 2347273 | 89.4 | 118214 | 36.9 | 203457 | 33.1 |
| badFDPeriodRejection | 111733 | 95.2 | 2236951 | 95.3 | 112651 | 95.3 | 193835 | 95.2 |
| good10MhzCorrection | 110784 | 99.2 | 2217195 | 99.1 | 112651 | 100.0 | 193835 | 100.0 |
| noBadPixelInPulse | 110781 | 100.0 | 2217126 | 100.0 | 112651 | 100.0 | 193835 | 100.0 |
| isCLF | 101021 | 91.2 | 2022980 | 91.2 | 112651 | 100.0 | 193835 | 100.0 |
| isXLF | 97911 | 96.9 | 1958792 | 96.8 | 112651 | 100.0 | 193835 | 100.0 |
| skipSaturated | 97617 | 99.7 | 1952737 | 99.7 | 111349 | 98.8 | 191382 | 98.7 |
| hybridTankTrigger | 61499 | 63.0 | 1229240 | 62.9 | 103617 | 93.1 | 173654 | 90.7 |
| hasMieDatabase | 55183 | 89.7 | 1103316 | 89.8 | 89121 | 86.0 | 149824 | 86.3 |
| **Preselection** | 55183 | 36.9 | 1103316 | 36.9 | 89121 | 3.2 | 149824 | 2.8 |
| maxCoreTankDist | 55023 | 99.7 | 1100180 | 99.7 | 87554 | 98.2 | 148967 | 99.4 |
| maxZenithFD | 52870 | 96.1 | 1057246 | 96.1 | 80071 | 91.5 | 136296 | 91.5 |
| angleTrackLength | 37923 | 71.7 | 759832 | 71.9 | 52849 | 66.0 | 82053 | 60.2 |
| HDSpectrumDistance2012 | 19585 | 51.6 | 393651 | 51.8 | 46003 | 87.0 | 72352 | 88.2 |
| **Geometry** | 19585 | 35.5 | 393651 | 35.7 | 46003 | 51.6 | 72352 | 48.3 |
| profileChi2Sigma | 18331 | 93.6 | 368736 | 93.7 | 44672 | 97.1 | 70020 | 96.8 |
| minViewAngle | 17263 | 94.2 | 347097 | 94.1 | 42349 | 94.8 | 66770 | 95.4 |
| xMaxInFoV | 10097 | 58.5 | 202250 | 58.3 | 35840 | 84.6 | 49963 | 74.8 |
| XmaxErrorLessThenXmax | 10096 | 100.0 | 202214 | 100.0 | 35838 | 100.0 | 49961 | 100.0 |
| maxDepthHole | 9938 | 98.4 | 198939 | 98.4 | 35367 | 98.7 | 49391 | 98.9 |
| calEnergyError | 9938 | 100.0 | 198933 | 100.0 | 35367 | 100.0 | 49391 | 100.0 |
| **Profile** | 9938 | 50.7 | 198933 | 50.5 | 35367 | 76.9 | 49391 | 68.3 |
| cloudCutXmaxPRD14 | 7064 | 71.1 | 142039 | 71.4 | 22399 | 63.3 | 31751 | 64.3 |
| maxVAOD | 6675 | 94.5 | 133741 | 94.4 | 20593 | 91.9 | 29187 | 91.9 |
| **Atmosphere** | 6675 | 67.1 | 133741 | 67.2 | 20593 | 58.2 | 29187 | 59.1 |
| **Selected Sample** | 6675 | 4.5 | 133741 | 4.5 | 20593 | 0.8 | 29187 | 0.5 |

**Table 4.8:** *Hybrid data and simulations: event selection criteria, number of events after each cut and selection efficiency with respect to the previous cut.*

As one can see, events are mostly removed at the pre-selection level, in particular by the *hybridTankTrigger* cut that excludes stations triggered by a TOTd or a MOPS trigger (selection efficiencies of 62.9 %). Due to the nature of these two triggers, which select very low signals in the SD sta-

tions, the removed events are mostly those with energy below $10^{18}$ eV. The differences found in the *minRecLevel* and *eyeCut* selection efficiencies in data and simulations are due to the presence of all the events generated in the raw simulations samples. Let us remind that the simulation is performed on a surface much larger than that of the SD and during an observation time (day and night) much larger than that of the FD (operating only in moonless clear nights). Therefore a large fraction of events do not actually trigger either an SD station or one of the FD telescopes. At the geometry level, events are mostly removed by the *HDSpectrumDistance* cut. This fiducial cut is used to avoid trigger-threshold effects and to achieve an exposure which is independent of the energy scale uncertainties. At the profile level, most events are removed by xMaxInFoV. This ensures that $X_{\max}$ is observed in the geometrical field of view of the telescopes and hence is intended to prevent biases in the profile fit. Finally, at the atmospheric level, for the events that survive all the previous levels, it is the cloud-coverage cut that causes the largest loss of events.

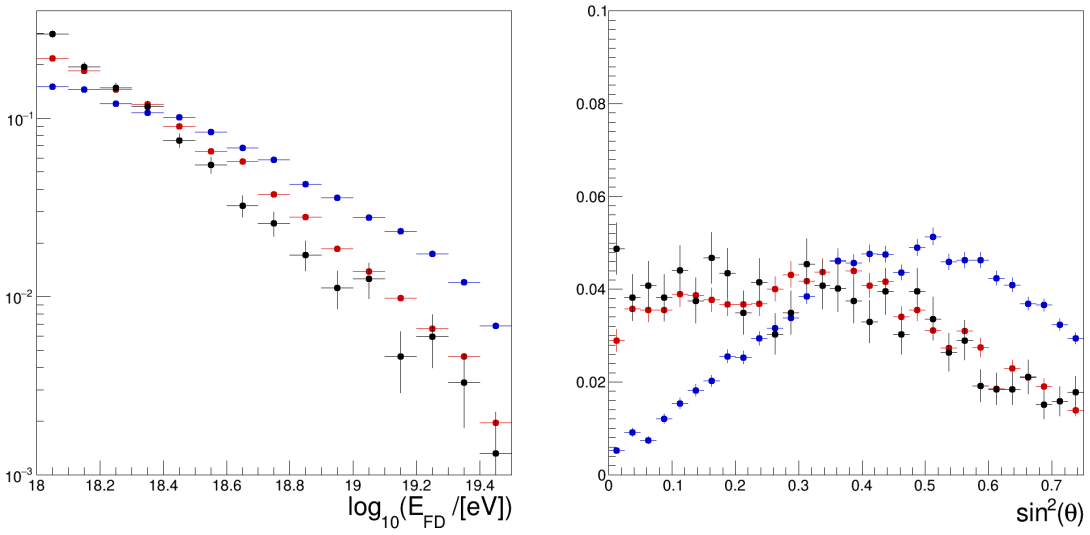



**(a)** **(b)**

**Figure 4.7:** *Protons (solid line) and photons (dashed line) simulation selection efficiency as a function of the energy (4.7a) and of the zenith angle (4.7b), at each cut level: preselection level (yellow), geometry level (green), profile level (blue) and atmospheric level (red).*

The selection efficiency for each selection level is shown as a function of the true Monte Carlo energy and of the zenith angle in figure 4.7a and 4.7b for simulated showers, proton-initiated (solid line) and photon-initiated (dashed line). At the geometric level, the selection efficiency for photon showers has a different shape than for those initiated by protons. This is mostly due to the XmaxInFoV cut which rejects ultra-high-energy photons, as these often reach $X_{\max}$ below the FD field of view.

Finally, figures 4.8a and 4.8b show the energy and zenith distributions of the selected events for the burn sample (black), proton-showers (red) and photon-showers (blue). The distributions are




**(a)** **(b)**

**Figure 4.8:** *Energy (4.8a) and zenith distribution (4.8b) of the selected samples, for the burnt sample (black), protons (red) and photons (blue).*

normalized to 1. As already observed in figure 4.7a, the vertical photon-showers are suppressed, because they land before reaching their maximum.

These samples, data and simulations, will be used in the next chapters for the study of an analysis technique devoted to the search for UHE photons. In particular, simulations will be used in chapter 5 for the study of a mass-sensitive variable, $F_\mu$, related to the muonic content of an air-shower. Simulations will be also used in chapter 6, together with the burn sample to define the photon selection technique. Finally, the results and physical implications of this analysis will be inferred from the analysis data sample in chapter 7.

# 5 STUDY OF THE RELATIVE MUON CONTENT, $F_\mu$, FOR PHOTON-HADRON SEPARATION

## Contents



The signatures of a photon-induced air shower are a deeper atmospheric depth of the shower maximum, $X_{\max}$, and a lower number of muons compared to showers of the same primary energy initiated by the much more numerous hadrons. $X_{\max}$ is measured with the FD, while the muonic content can be estimated through mass-sensitive SD observables. Profiting from the hybrid nature of the Auger Observatory, this thesis work, which exploits hybrid events, aims at combining FD- and SD-based observables to improve the photon-hadron separation power. This chapter presents the study of a new mass-sensitive SD parameter to be used together with $X_{\max}$.

First, a short review of the most relevant parameters measured with the SD is presented in section 5.1. They generally have a complex dependence on the shower geometry, on the energy $E$, and on $X_{\max}$. One of the goals of this thesis is to identify a new SD parameter, $F_\mu$, that, related to the muonic content of a shower, is almost independent of $X_{\max}$, $E$ and shower geometry, by using the universality paradigm in combination with the hybrid reconstruction. Section 5.2 illustrates how, and with which accuracy, $F_\mu$ can be estimated from the SD signal in hybrid events, even in the case of only one triggered SD station. The potential of $F_\mu$ as photon-hadron discriminator is assessed in section 5.3 by using Monte Carlo simulations, while in section 5.4 the method is applied to hybrid data and consistency checks are performed.

## 5.1 SD OBSERVABLES FOR PHOTON-HADRON SEPARATION

Measuring electron and muon numbers in showers is the most common technique applied to particles arrays to infer the cosmic ray composition and to discriminate primary photons from hadrons. Although the SD WCDs do not allow for a direct measurement of the two kind of particles, mass-sensitive variables can be built with WCD data by exploiting the different behavior of the muonic

and electromagnetic components. Such difference has an impact both on the timing structure of the signals in the WCDs, and on their lateral distribution.

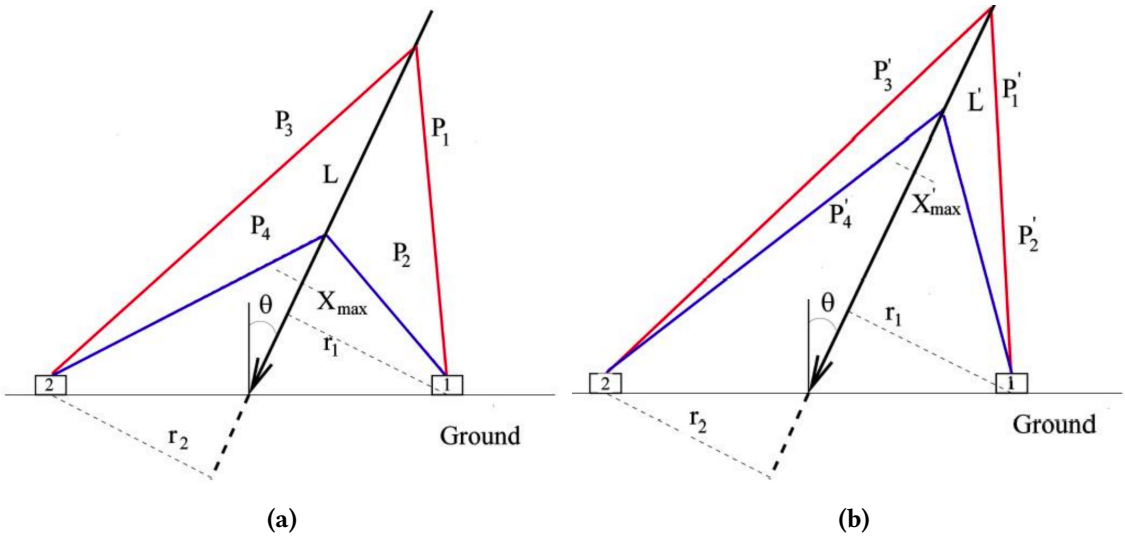
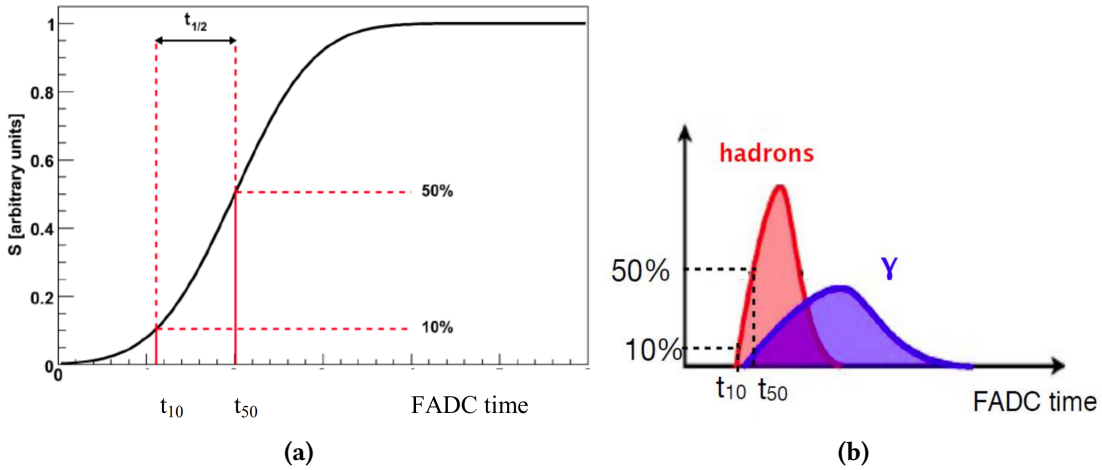



**Figure 5.1:** *Schematic diagrams showing the relationship between the shower geometry and the delay of particles arriving on the ground. The shower on the left penetrates deeper in the atmosphere than the shower on the right.*

The spread of the arrival times of particles at the ground carries information about the longitudinal development of a shower. This can be explained if we consider the geometry of the particles in the shower, illustrated in figure 5.1, where the average paths of particles produced in a deep shower (figure 5.1a) and in a shallow one (figure 5.1b) are shown schematically. The two showers develop in the direction of the shower axis, towards ground. One can see that the particles arriving in the detectors from lower down in the shower (paths L+P2 or L+P4) arrive later at a detector than those that are produced higher up (path P1 or P3). Therefore, a spread in the arrival times of particles is observed at the surface detectors. From simple geometry it is clear that the delay, i.e., the spread in the signal, in the left case is larger than the right one, i.e., it is larger for showers with a deeper $X_{\text{max}}$. The information related to the arrival times of particles are registered in the FADC traces of the SD stations. However, the full time structure of the signal is not used, because the late part carries little information about the development of the shower since it is dominated by electromagnetic particles, which arrive later than muons due to their multiple scattering. The **risetime** of the FADC trace, $t_{1/2}$, is thus used as a mass-sensitive observable, defined as the difference of the 50 % and 10 % time quantiles of the signal trace, as illustrated in figure 5.2a. As shown in the cartoon in figure 5.2b, the time spread of the trace, hence its risetime, is smaller for a hadron shower (red) than for a photon shower (blue) because the latter is characterized by a deeper $X_{\text{max}}$ and a lower muonic content. The risetime is a function of distance, zenith angle, and energy.

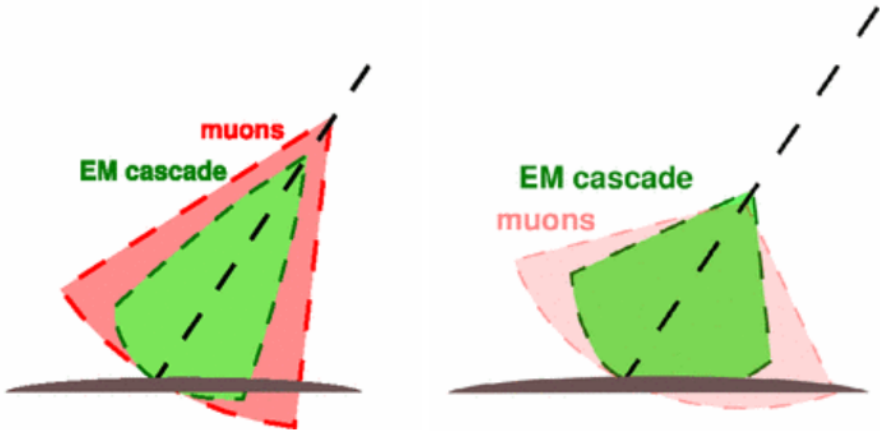
Another parameter used for photon-hadron separation is the **radius of curvature** of the shower front. The shape of the shower front (formed by the particles arriving first) carries in-




(a) (b)

**Figure 5.2:** *Definition of $t_{1/2}$ as the difference between the 10 % and 50 % time quantiles of the signal trace (5.2a). The differences between a signal induced by a hadron-initiated cascade (red) and a photon-initiated cascade are highlighted in (5.2b). From [165].*

formation of the geometry of the shower and the mass composition of the primary. The particles that arrive earliest at the ground in hadron showers are muons mostly produced in the first interactions. Photon showers develop deeper in the atmosphere and with fewer muons, so that the first particles arriving are mostly electromagnetic component, created at altitudes much lower than the first interactions and scattered before reaching the ground. Consequently, photon showers have more curved shower front with smaller radius of curvature than hadrons as schematically shown in figure 5.3. The radius of curvature depends on primary energy and zenith angle.




(a) (b)

**Figure 5.3:** *Illustrative view of the front of a hadron- (5.3a) and photon-induced (5.3b) shower. From [166].*

The lateral distribution function (LDF) describes the relationship between the signal size and the distance of the detector away from the shower axis (core). Besides being a tool to reconstruct the shower size (as discussed in chapter 2), the shape of the LDF also contains information about the mass of the primary particle. The *fall-off* of the signals with the distance is, in fact, related to the physics processes involved during the shower development, and to the transverse momentum of the secondary particles. In particular, the deflection due to the hadronic interaction is larger than that due to the electromagnetic scattering and is responsible for the lateral distribution of muons. As a consequence, the signal falls faster from the shower axis for the electromagnetic component than for the muon component, so that the **steepness of the LDF** is also a parameter used for photon-hadron separation. As illustration to this concept, figure 5.4 shows the average LDF for vertical proton-induced showers (red) and photon-induced ones (blue) with energy $E = 10^{18.5}$ eV: the LDF is flatter for proton showers as they contain more muons.

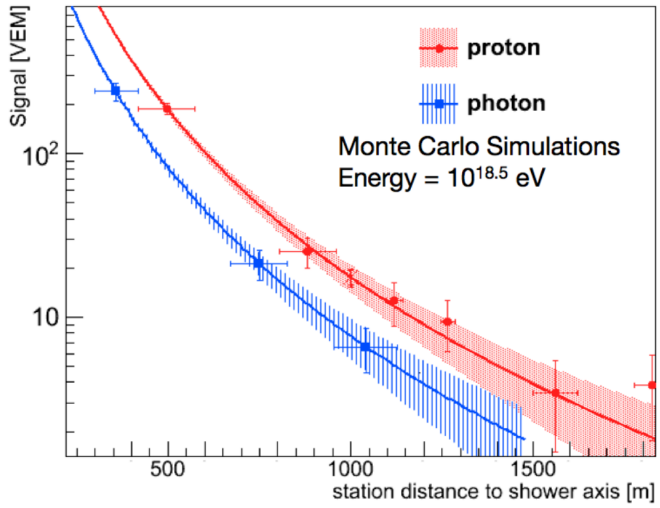



**Figure 5.4:** *Lateral distribution function (LDF) for proton (red) and photon (blue) vertical simulated showers with the same Monte Carlo energy $E = 10^{18.5}$ eV. From [109].*

Although the photon/hadron separation can benefit from the parameters presented so far, these can only be reliably estimated when a minimum of SD detectors have triggered. For example, at least five detectors are needed to determine the curvature of the shower front with acceptable accuracy. This effectively places the energy threshold typically at about 10 EeV when UHE photons are searched with analyses using such variables. To fully exploit the hybrid approach, that naturally allows to reach lower energies down to 1 EeV, SD-based variables have been used in the past, that require fewer SD stations [114]. The most successful one is the so-called **$S_b$ parameter** [167], which is sensitive to different lateral distribution functions, due to the presence/absence of the flatter muon component. As a result of both, the smaller signal in the stations, on average, and the steeper lateral distribution function, smaller values of $S_b$ are expected for photon primaries. This parameter, which does not require a minimum number of stations, is defined as:

$S_b = \sum_i S_i \, (r_i/1000\,\mathrm{m})^b$, where the sum extends over all the triggered stations, $S_i$ is the signal measured at the $i$-th station, $r_i$ is the distance to the shower axis and $b$ is a variable exponent. It has been found that the optimal separation between data and photons is obtained by setting $b = 4$ [114]. Figure 5.5 shows the $S_b$ distributions with $b = 4$ for primary photons (red) and data (black).

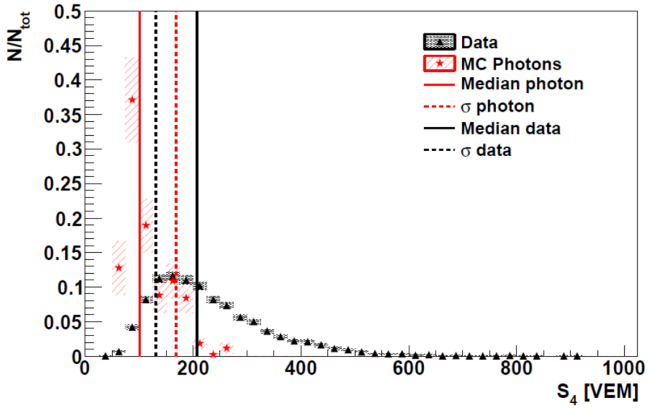



**Figure 5.5:** *$S_b$ distribution with $b = 4$ for photon simulations (red) and data (black) with energies $E > 10^{19}$ eV. Since both distributions are asymmetric and non-Gaussian, one standard deviation is defined as the position such that the area enclosed between the medians (solid lines) and $1\sigma$ (dashed lines) corresponds to 34 % of the total area. From [168].*

As one can see from the figure, the separation power of $S_b$ is quite limited. In the following, we present a new SD parameter, $F_\mu$, that can also be reconstructed from the signal of a single station and, as it will be proven has a better separation power compared to $S_b$.

## 5.2 $F_\mu$ ESTIMATION IN HYBRID EVENTS

In this section, using the concept of air-shower universality, we define a parameter related to the muonic content of a shower, $F_\mu$, to be used in combination with $X_{\mathrm{max}}$, obtained with the hybrid reconstruction, to search for UHE photons.

The study of $F_\mu$ is performed by using the sample of simulated hybrid events that have been described in the previous chapter. One of such events, shown in figure 5.6a, is used as an example to illustrate in the following the procedure of reconstruction of $F_\mu$. The primary particle in this case is a proton with Monte Carlo energy $E_{\mathrm{MC}} = 9.55 \times 10^{18}$ eV. The zenith angle of its true arrival direction is $\theta_{\mathrm{MC}} = 40.3^\circ$ and the azimuth is $\phi_{\mathrm{MC}} = 140.2^\circ$. The Monte Carlo longitudinal profile is shown as blue line in the left panel of the figure: the shower maximum is at a depth $X^{\mathrm{MC}}_{\mathrm{max}} = 822.4\,\mathrm{g\,cm^{-2}}$. Ten SD stations have triggered the SD, as illustrated in the right panel. The impact point of the shower is shown as a red point, being at $(-8.96, 24.25)$ km with respect to the center of the array, the distance of each station from the core being shown in the third column

of table 5.1. The dimension of the circles representing the stations is proportional to the size of signals, their values being listed in the first column of table table 5.1. The simulated event is processed using the hybrid reconstruction explained in section 2.2.3. The geometric reconstruction yields parameters very similar to the true ones: $\theta_{\rm REC} = (40.4 \pm 0.5)^\circ$, $\phi_{\rm REC} = (140.5 \pm 0.3)^\circ$ and $(-8.93 \pm 0.04, 24.25 \pm 0.05)$ km for the core position. Similarly, the reconstruction of the longitudinal profile (shown as a red line in the left panel of the figure) provides very consistent values for the energy and the depth of the shower maximum, namely $E_{\rm REC} = (9.59 \pm 0.71) \times 10^{18}$ eV and $X^{\rm REC}_{\rm max} = (813 \pm 13)\,{\rm g\,cm^{-2}}$.

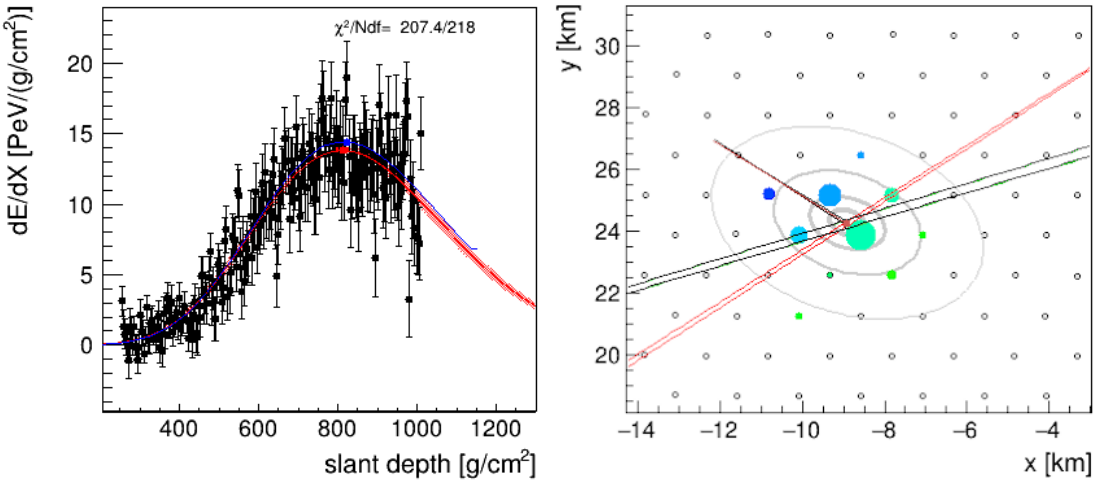



**Figure 5.6:** *Example of a hybrid simulated event generated by a primary proton with $E_{MC} = 9.55 \times 10^{18}$ eV and zenith angle $\theta = 40.3^\circ$. (a) Longitudinal profile reconstruction. The blue line represents the Monte Carlo profile, while the reconstructed one is shown by the red line. The red shaded area represents the uncertainties in the profile reconstruction. The blue and red points represent the Monte Carlo and reconstructed $X_{max}$, respectively. (b) Event footprint on the SD. Each circle represents one of the 10 triggered stations. The colors (from blue to green) represent the trigger times while the dimension is proportional to the signal size.*

Below, in section 5.2.1 we explain the technique to obtain a station-wise estimation of $F_\mu$ from the signal of each SD station. Then, in section 5.2.2 the station-wise $F_\mu$ are combined to estimate the event-wise $F_\mu$.

### 5.2.1 Station-wise estimation of $F_\mu$

#### $F_\mu$ reconstruction

The universality-based model introduced in chapter 3 predicts that the shower signal in any detector can be described as the superposition of four components: muons ($S_\mu$); $e^\pm$ and $\gamma$ from high energy pions ($S_{e\gamma}$); $e^\pm$ and $\gamma$ from muon decays ($S_{e\gamma(\mu)}$); $e^\pm$ and $\gamma$ due to low energy hadrons

| Signal Size (VEM) | Predicted Signal (VEM) | Distance (m) | $F_\mu^{\text{station}}$ | MC $F_\mu^{\text{station}}$ |
|---:|---:|---:|---:|---:|
| 752 | 847 | 381 | 1.34 | 1.64 |
| 87 | 103 | 827 | 1.37 | 1.28 |
| 25 | 33 | 1135 | 1.24 | 1.14 |
| 15 | 13 | 1420 | 2.22 | 2.30 |
| 8 | 8 | 1640 | 1.64 | 1.57 |
| 6 | 7 | 1560 | 1.34 | 1.38 |
| 4 | 2 | 2130 | 3.15 | 3.18 |
| 4 | 1 | 3130 | 7.11 | 7.17 |
| 3 | 8 | 1529 | 0.61 | 0.66 |
| 3 | 7 | 2000 | 0.58 | 0.56 |

**Table 5.1:** *Characteristics of the SD stations in the event in figure 5.6: observed and predicted signal, core distance, reconstructed and true* $F_\mu^{station}$.

($S_{e\gamma(\text{had})}$). Each i-th signal component, $S^i_{\text{comp}}$, has a universal behavior depending only on $E$, $X_{\text{max}}$, and geometry. The relative contributions of each of the four components, $f^i_{F_\mu\,\text{fluct}}$, instead, depend on the mass of the primary particle, through a parameter representing the number of muons in the shower, $F_\mu$. The predicted signal, $S_{\text{pred}}$, can then be expressed as:

$$S_{\text{pred}} = \sum_{i=1}^{4} f^i_{F_\mu\,\text{fluct}} \cdot S^i_{\text{comp}} = \sum_{i=1}^{4} \Big(1 + (F_\mu - 1)\alpha_i\Big) S^i_{\text{comp}} \tag{5.1}$$

where $i$ runs over the four components and $\alpha_i$ takes into account the correlation of the i-th component with $F_\mu$. $S^i_{\text{comp}}$, in turn, has been parametrised using QGSJetII-03 proton simulations (see again chapter 3) as $S^i_{\text{comp}} = S^i_0 \cdot f^i_{\text{mod}} \cdot f^i_{\text{atm}} \cdot f^i_{\text{conv}}$ where $S_0$ is the signal in an ideal detector, converted to a realistic one by $f_{\text{conv}}$, and corrected for atmospheric effects by $f_{\text{mod}}$ and $f_{\text{atm}}$.

As the reconstruction of hybrid events provide $E$, $X_{\text{max}}$, and the shower geometry, $S^i_{\text{comp}}$, which depends exclusively on those, can be directly calculated. It is thus of interest to look at the behavior of the different components as a function of these parameters, shown in figure 5.7. In all panels the red, green, blue, yellow and black lines represent $S_{e\gamma}$, $S_\mu$, $S_{e\gamma(\mu)}$, $S_{e\gamma(\text{had})}$ and $S_{\text{pred}}$, respectively. The top-left panel (a) shows the evolution of the predicted signals at 1000 m from the core for a shower with zenith angle of 45° whose maximum is at $X_{\text{max}} = 750\,\text{g}\,\text{cm}^{-2}$, as a function of energy: the signal size increases with energy, as expected. Similarly expected is the decrease of the average signal with $X_{\text{max}}$, as illustrated in the top-right panel (b), where proton showers with zenith $\theta = 45°$ and $E = 10^{19.0}\,\text{eV}$ are considered. Larges values of $X_{\text{max}}$ correspond to "younger" showers, i.e., less developed. Note that for showers that reach their maximum very deep in atmosphere, the contribution to the signal due to the jet component increases, as expected in very "young" showers. The evolution of the signals as a function of zenith is shown in the bottom-left panel (c), for deep ($X_{\text{max}} = 750\,\text{g}\,\text{cm}^{-2}$, solid lines) and shallow ($X_{\text{max}} = 1000\,\text{g}\,\text{cm}^{-2}$,

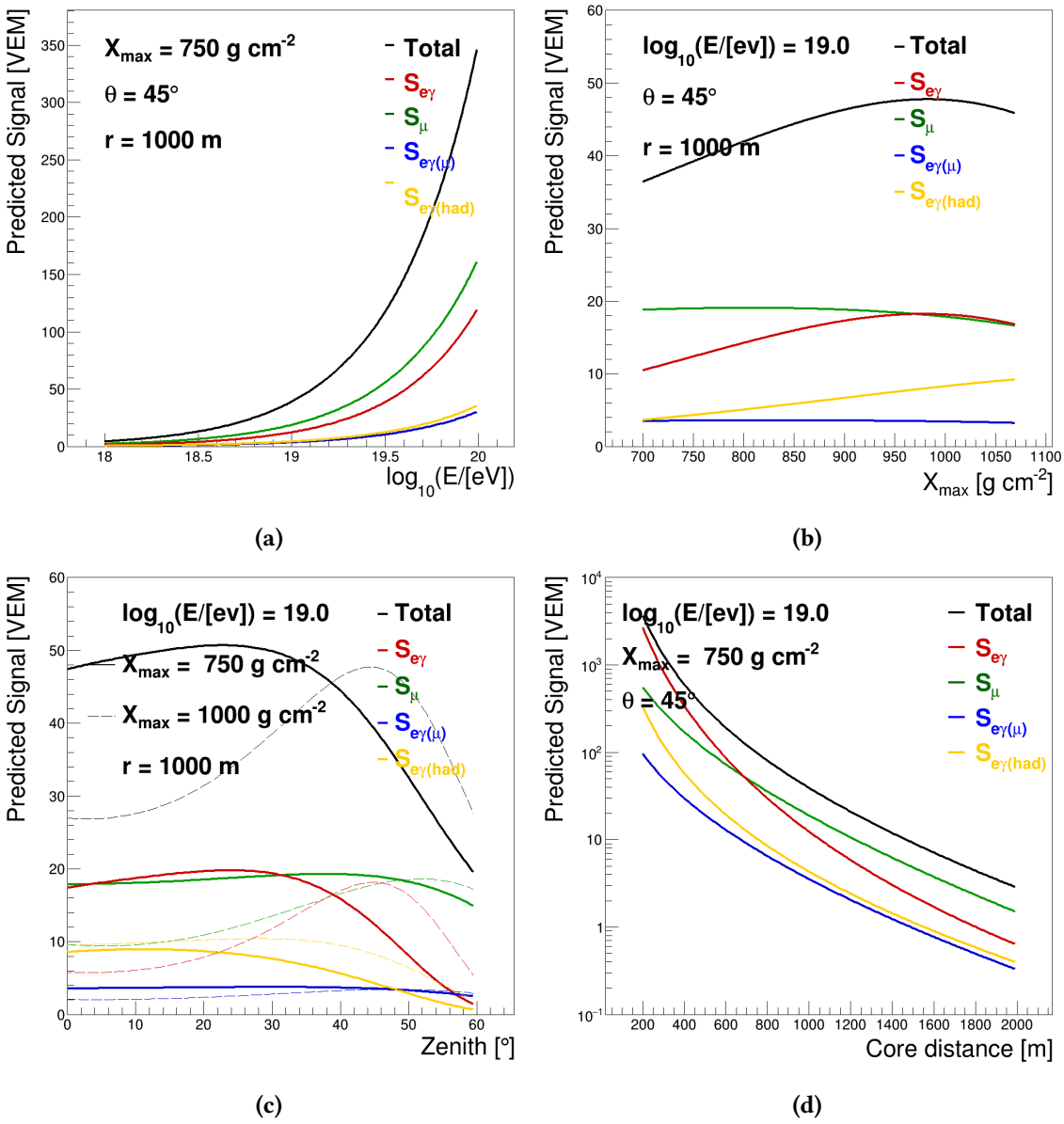



**Figure 5.7:** *Evolution of signal components as a function of energy (5.7a), $X_{max}$ (5.7b), zenith (5.7c) and core distance (5.7d). The predicted total signal (black line) is shown together with each parametrized signal component: $e^{\pm}\gamma$ (red), $\mu$ (green), $e^{\pm}\gamma(\mu)$ (blue) and $e^{\pm}\gamma(had)$ (yellow).*

dashed lines) showers. Both for deep and shallow horizontal showers, the pure e.m. component is absorbed by the atmosphere, while the muon component is dominant. Note that at higher zenith angles, the e.m. component in shallow showers (dashed lines) shows a peak due to the fact that they are more developed, because the portion of atmosphere traversed by the shower is larger.

Finally, the fall-off of the signal as a function of the distance to the core is shown in the bottom-right panel (d): the e.m. contribution is more important near the shower core, while, far from the core, the muonic component is dominant.

| | $S^i_{\text{comp}}$ (VEM) | $\alpha_i$ |
|---|---|---|
| $e\gamma$ | 448 | -0.075 |
| $\mu$ | 151 | 1.000 |
| $e\gamma(\mu)$ | 25 | 1.000 |
| $e\gamma$(had) | 55 | 1.237 |
| core distance | 381 m | |
| signal size | 752 VEM | |
| $\mathbf{F_\mu^{station}}$ | **1.34** | |

**Table 5.2:** *$S^i_{comp}$ and $F_\mu^{station}$ calculated for the station with the largest signal size in the event shown in figure 5.6b*

$S^i_{\text{comp}}$ can be calculated for each component and for each station involved in a hybrid event, by using the hybrid-reconstruction parameters. As an example, the values of $S^i_{\text{comp}}$ for the four components, calculated from the signal of the station with the largest signal in the event shown in figure 5.6a, are given in table 5.2, where the values for the coefficients $\alpha_i$, as provided in chapter 3, are also tabulated. For each hybrid event, and for each station, a station-wise estimation of the relative muon content, $F_\mu^{\text{station}}$, can be obtained by replacing the total predicted signal, $S_{\text{pred}}$, with the reconstructed signal, $S_{\text{rec}}$, and by inverting equation 5.1, thus obtaining:

$$F_\mu^{\text{station}} = \frac{S_{\text{rec}} - \sum_{i=1}^{4}(1-\alpha_i)S^i_{\text{comp}}(r,\theta,\phi,E,X_{\max})}{\sum_{i=1}^{4}\alpha_i S^i_{\text{comp}}(r,\theta,\phi,E,X_{\max})} \tag{5.2}$$

The value of $F_\mu^{\text{station}}$ calculated for each station in the event shown in figure 5.6b is reported in the fourth column of table 5.1. For the sake of comparison, the Monte Carlo expected value of $F_\mu^{\text{station}}$ is also shown, in the fifth column, which is well consistent with the calculated one. The largest difference is for the station closest to the core: this will be discussed in section 5.2.2.

The resulting distribution of $F_\mu^{\text{station}}$, calculated for each station in the sample of simulated hybrid events, is finally shown in figure 5.8 for primary protons (red) and photons (blue). The medians of the distributions (dashed vertical lines) are 1.3 and 0.15 for protons and photons, respectively.

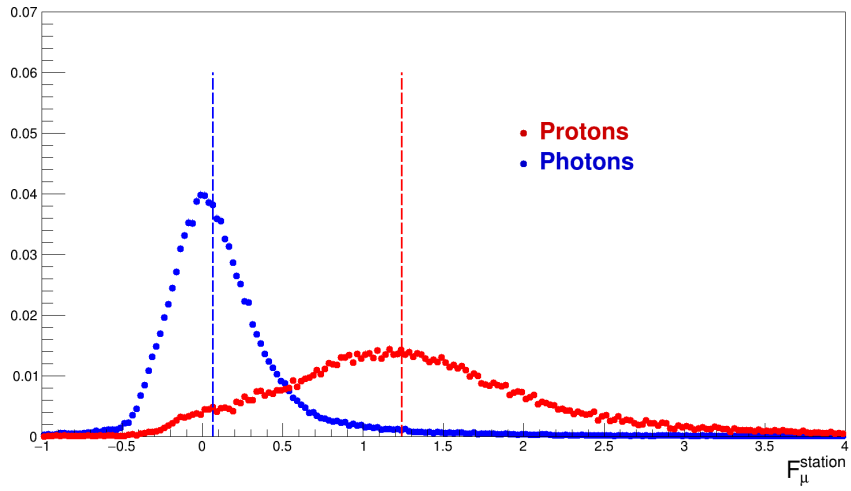



**Figure 5.8:** *Distributions of $F_\mu^{station}$ for protons (red) and photons (blue). The dashed vertical lines show the median of the distributions.*

### *Uncertainties in $F_\mu$ reconstruction*

The uncertainty in the estimation of $F_\mu^{\text{station}}$, $\sigma_{F_\mu}$, derives, on the one hand, from that of the reconstruction of $S_{\text{rec}}$ and, on the other hand, from that of the reconstructed hybrid parameters. For the sake of clarity, in the following we denote the uncertainties of each parameters with $\sigma$, and their contribution to $\sigma_{F_\mu}$ with $\delta$.

The contribution to $\sigma_{F_\mu}$ due to the reconstructed signal, $\delta_{\text{S rec}}$, is obtained by propagating analytically the signal uncertainty, $\sigma_S$ (see equation 2.1), into 5.2:

$$\delta_{\text{S rec}} = \frac{(0.34 + 0.46 \sec\theta)}{\sum_i \alpha_i S^i_{\text{comp}}} \sqrt{S_{\text{rec}}} \tag{5.3}$$

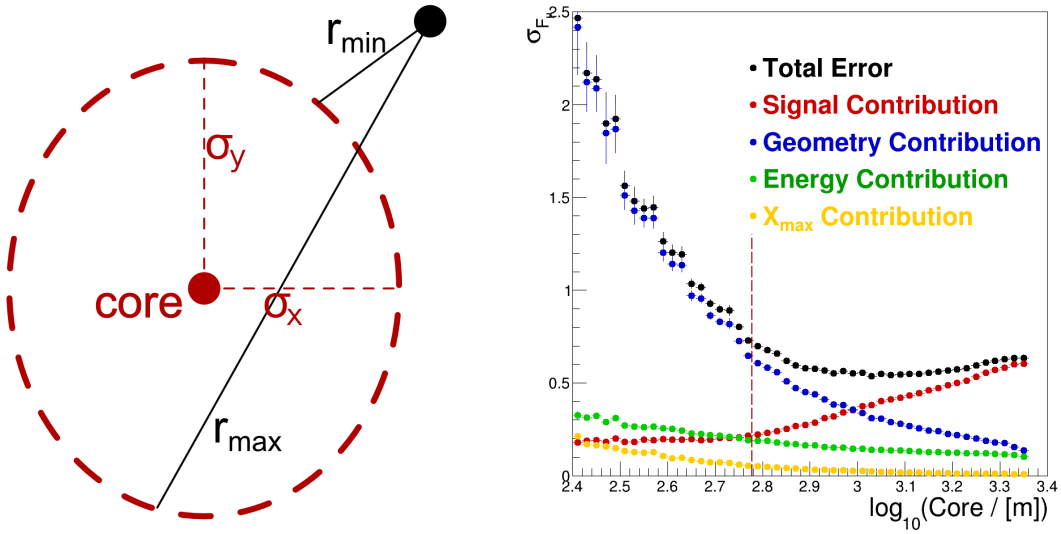
In turn, the contribution to $\sigma_{F_\mu}$ due to the hybrid reconstruction, $\delta_{\text{hyb rec}}$, is derived as the sum of different sub-terms, related to each reconstructed parameter, i.e.,

$$\delta_{\text{hyb rec}} = \sqrt{\delta_E^2 + \delta_{X_{\text{max}}}^2 + \delta_\theta^2 + \delta_\phi^2 + \delta_r^2}$$

where $\delta_E$, $\delta_{X_{\text{max}}}$, $\delta_\theta$, $\delta_\phi$ and $\delta_r$ account for the uncertainties $\sigma_E$, $\sigma_{X_{\text{max}}}$, $\sigma_\theta$, $\sigma_\phi$ and $\sigma_r$, respectively (see chapter 4).

$\delta_\theta$, $\delta_\phi$, $\delta_E$ and $\delta_{X_{\text{max}}}$ are derived by changing the zenith, or the azimuth, or the energy or the $X_{\text{max}}$ of the event by $\pm\sigma$ and then calculating a new estimation of $F_\mu^{\text{station}}$.

$\delta_r$ is calculated similarly, by using an estimation of the maximum and minimum allowed core distances, $r_{\text{max}}$ and $r_{\text{min}}$, in the calculation of $F_\mu^{\text{station}}$. These two distances depend on the uncertainties on the core coordinates, $(x_{\text{core}}, y_{\text{core}})$: they are estimated following the procedure illustrated in figure 5.9a, i.e., by finding the maximum and minimum distances of a station located in




(a) (b)

**Figure 5.9:** ***(a):*** *Scheme of the geometry used for the calculation the the minimum and maximum distances of an SD station to the shower core (see text for details);* ***(b):*** $\sigma_{F_\mu}$ *(black points) as a function of the distance to the shower core. The colored points show the different contributions: signal (red), geometry (blue), energy (green) and* $X_{max}$ *(yellow).*

$(x_{\text{station}}, y_{\text{station}})$ from an ellipse centered in $(x_{\text{core}}, y_{\text{core}})$, the uncertainty of which, $\sigma_x$ and $\sigma_y$, are the semi-axis. This is achieved by finding the stationary points of the function

$$\sqrt{(x - x_{\text{station}})^2 + (y - y_{\text{station}})^2}$$

constrained by the ellipse, by using the Lagrange multipliers method [169]:

The contributions of the different reconstructed parameters to $\sigma_{F_\mu}$ are shown in figure 5.9b as a function of the distance to the shower core. The black points correspond to total uncertainty of $F_\mu$, calculated as $\sigma_{F_\mu} = \sqrt{\delta^2_{\text{S rec}} + \delta^2_{\text{hyb rec}}}$. The uncertainty of the geometry results to be the most relevant contribution near the shower core, due to the steepness of the lateral distribution function. In turn, the contribution due to the uncertainty of the signal reconstruction is the most important one at large distances. The vertical dashed line will be discussed in section 5.2.2.

### 5.2.2 Event-wise estimation of $F_\mu$

About 50% of the hybrid events used in this work include more than one triggered station. In this section it is shown how an event-wise $F_\mu$ is estimated starting from $F_\mu^{\text{station}}$. To select the stations to be used in the event-wise estimation of $F_\mu$, criteria on signal size and core distance are defined.

Simulated proton events are used for studying the performance of the $F_\mu^{\text{station}}$ reconstruction as a function of the predicted signal, $S_{\text{pred}}$, and of the radial distance from the shower axis, $r_{\text{core}}$.

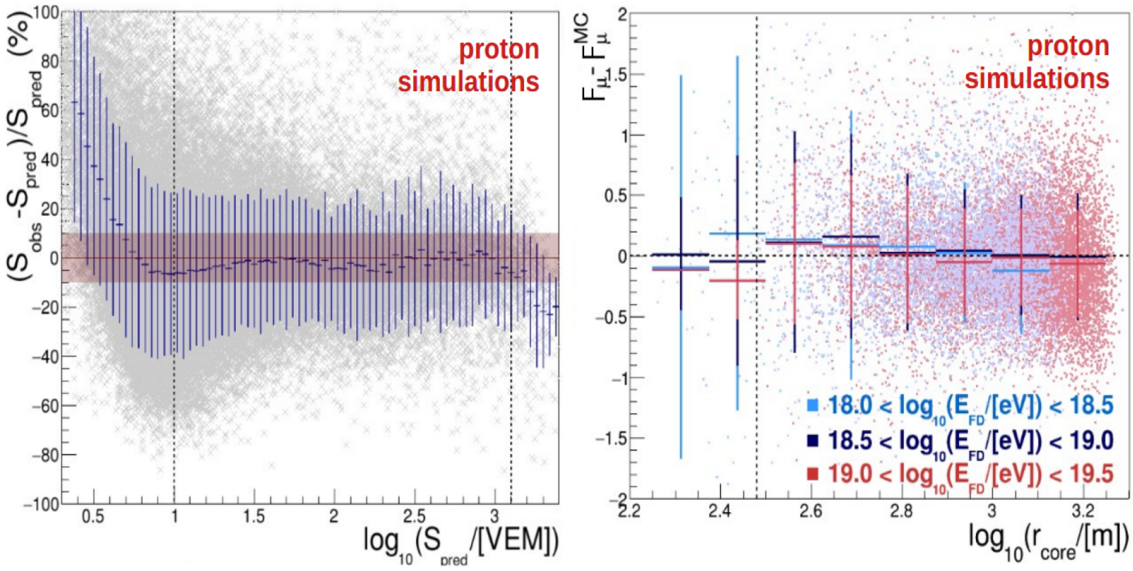
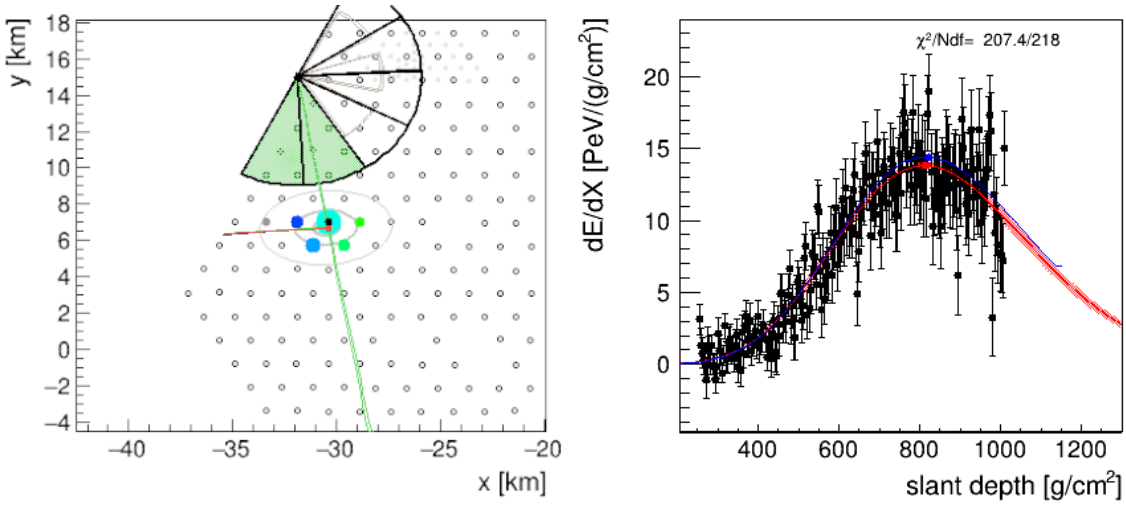



**Figure 5.10:** *Check for biases due to detector effects on the universal description of the total signal in the case of EPOS-LHC protons in the* $10^{18.0}$ eV – $10^{19.5}$ eV *energy range.* ***(5.10a)*** *Relative difference between the predicted and observed signal as a function of the predicted signal. Between* 6 VEM *and* 800 VEM*: the accuracy of the parametrization is better than* 10%*. Trigger (left) or saturation (right) effects are visible outside the region enclosed by the dash lines.* ***(5.10b)****: Difference between the calculated and the true* $F_\mu$ *as a function of the core distance. Vertical bars represent the standard deviation of the distribution in each bin for both plots.*

Figure 5.10a shows the relative difference between the reconstructed and the predicted signal as a function of the predicted signal in individual stations, calculated using equation 5.1 with the true Monte Carlo $F_\mu^{\text{MC}}$. The vertical bars represent the standard deviation of the distribution in each bin. For signals between 6 VEM and 800 VEM, the accuracy of the model prediction is better than 10 % (red band). For smaller (larger) signals, instead, trigger (saturation) effects truncate the distribution on one side producing a visible bias.

In figure 5.10b the difference between the reconstructed and the true value of $F_\mu$ is shown as a function of the distance from the shower axis. $F_\mu^{\text{station}}$ does not present any significant bias. However, due to the steepness of the lateral distribution close to the axis, for stations at small distances from the core the signal prediction can be strongly affected by the resolution on the core reconstruction. An example of a simulated event with a SD station associated with a distorted $F_\mu^{\text{station}}$ due to the core resolution is shown in figure 5.11. This event has an energy $E = 2.87 \times 10^{18}$ eV and $X_{\text{max}} = 842\,\text{g}\,\text{cm}^{-2}$. The hottest station is at 439 m from the core and has a reconstructed signal of 207 VEM, while the predicted one is 344 VEM. The calculated $F_\mu^{\text{station}}$ for this station is 0.28, while the Monte Carlo expected value is 1.06. Note that the difference between calculated




(a) (b)

**Figure 5.11:** *Example of hybrid event in which the signal estimation for the station closest to the core (cyan circle in the left panel) is biased (see text for details).*

and Monte Carlo $F_\mu^{\text{station}}$ remarked in the hottest station in the exemplary event of figure 5.6b is also due to its closeness to the core. Thus, to avoid such a bias in the $F_\mu^{\text{station}}$ reconstruction, a cut in distance is introduced, namely stations with $r_{\text{core}} < 600$ m (red vertical line in figure 5.9b) are excluded from the analysis.

The stations satisfying the criteria on signal, $6\,\text{VEM} < S_{\text{rec}} < 800\,\text{VEM}$, and on distance, $r > 600\,\text{m}$, are used for the event-wise estimation of $F_\mu$, which is obtained as

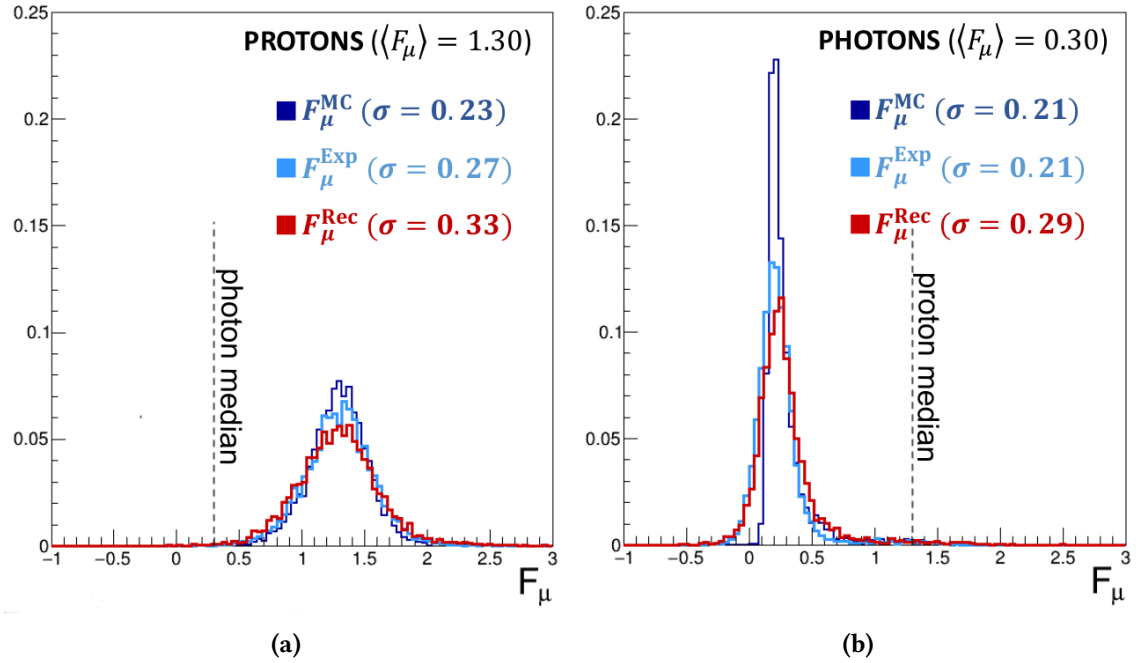
$$F_\mu = \frac{\sum_i F_\mu^{\text{station,i}} \cdot \left(\sigma_{F_\mu}^i\right)^{-2}}{\sum_i \left(\sigma_{F_\mu}^i\right)^{-2}} \tag{5.4}$$

where $i$ runs over the selected stations. By using the station selection criteria introduced above and the equation 5.4, the $F_\mu$ associated to the event reported in figure 5.6 results to be $F_\mu = 1.20$. $F_\mu^{\text{station}}$ is in turn used as estimator of $F_\mu$, if only one station passes the selection criteria.

## 5.3 $F_\mu$: A PARAMETER FOR PHOTON-HADRON SEPARATION

In this section we discuss the parameter $F_\mu$ in terms of photon/hadron separation power. This is illustrated in figure 5.12, where the $F_\mu$ distribution is presented for primary protons (5.12a) and photons (5.12b). The $F_\mu$ distribution is shown at different levels: $F_\mu^{\text{MC}}$ (dark blue) is the Monte Carlo value, $F_\mu^{\text{Exp}}$ (light blue) is the value reconstructed using as input for the model the true values of

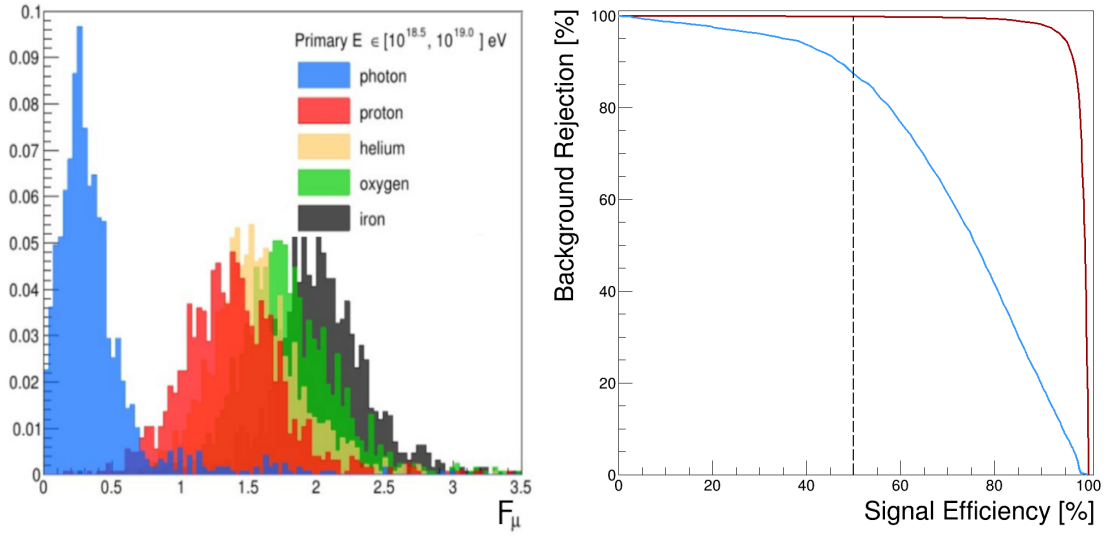
energy, $X_{\max}$ and geometry, while for $F_\mu^{\text{rec}}$ (red) the hybrid reconstructed values are used. For a given primary, all distributions have the same mean values thus proving that the reconstruction method is unbiased. The spread, $\sigma$, of $F_\mu^{\text{MC}}$ is mainly due to shower to shower fluctuations. For $F_\mu^{\text{Exp}}$ also sampling fluctuations can be observed. In the $F_\mu$ distribution the effects of the resolution of the hybrid reconstruction are observed additionally. Photon and proton distributions can be easily compared through the dashed black lines, that show the mean value of the distribution of the other primary type considered and are in both cases at more than $3\sigma$ from the mean.




**(a)** **(b)**

**Figure 5.12:** *$F_\mu$ distributions for protons (5.12a) and photons (5.12b) primaries. In dark blue the distribution of $F_\mu^{MC}$. In light blue the distribution of $F_\mu$ reconstructed with the true Monte Carlo values of energy, $X_{max}$ and geometry as input for the calculation of $S_{pred}$. In red the distribution of $F_\mu$ when the hybrid reconstructed values are used (realistic case). Dashed black lines represent the mean value of $F_\mu$ for the other primary type.*

The separation power of $F_\mu^{\text{rec}}$ is also highlighted in figure 5.13a where the well-separated distributions of $F_\mu^{\text{rec}}$ for simulated showers initiated by photon (blue) and proton (red) are shown. Here we display also the $F_\mu^{\text{rec}}$ distributions for simulated showers initiated by other, heavier, nuclei: helium (orange), oxygen (green) and iron (black). These showers are simulated with CORSIKA [152] using EPOS-LHC [54] as high-energy hadronic model. For each pirmary nucleus 5000 showers are generated with energy ranging from $10^{18.0}$ eV to $10^{18.5}$ eV, and zenith angle between 0° and 65°. Each shower is then used once for the simulation of the detector response. An ideal detector is considered, i.e., the actual status of the FD and SD is not accounted for. Although the separation among different nuclei is not as large as between photons and protons, one can clearly see that the mean value of $F_\mu^{\text{rec}}$ increases with mass, as expected, as it is a proxy for the number of muons.




**(a)** **(b)**

**Figure 5.13:** ***(a):*** *$F_\mu$ distributions for showers generated by different primaries: photon (blue), proton (red), helium (yellow), oxygen (green) and irons (black). The primary energy ranges from $10^{18.0}$ eV to $10^{18.5}$ eV.* ***(b):*** *Background rejection as a function of signal efficiency for $F_\mu$ (red) and the $S_b$ parameter (light blue).*

To quantify the separation between proton- and photon-induced air shower events, we use the merit factor, $\eta$, as a measure for the separation power of an observable. The merit factor is defined as

$$\eta = \frac{\left|\left\langle F_\mu^\gamma \right\rangle - \left\langle F_\mu^p \right\rangle\right|}{\sqrt{\sigma_\gamma^2 + \sigma_p^2}} \tag{5.5}$$

where $\left\langle F_\mu^\gamma \right\rangle$ and $\left\langle F_\mu^p \right\rangle$ are the median of the $F_\mu$ distributions for photon- and and proton-showers, respectively, while $\sigma_\gamma$ and $\sigma_p$ are the corresponding standard deviations. The merit factor is a measure of the overlap between two distributions. Two identical distributions have a corresponding merit factor equal to 0, increasing as the separation increases. The merit factor for the two distributions is $\eta = 2.5$. For comparison, the merit factor of the corresponding $S_b$ parameter for the same data sets has been calculated, which results to be $\eta = 1.5$.

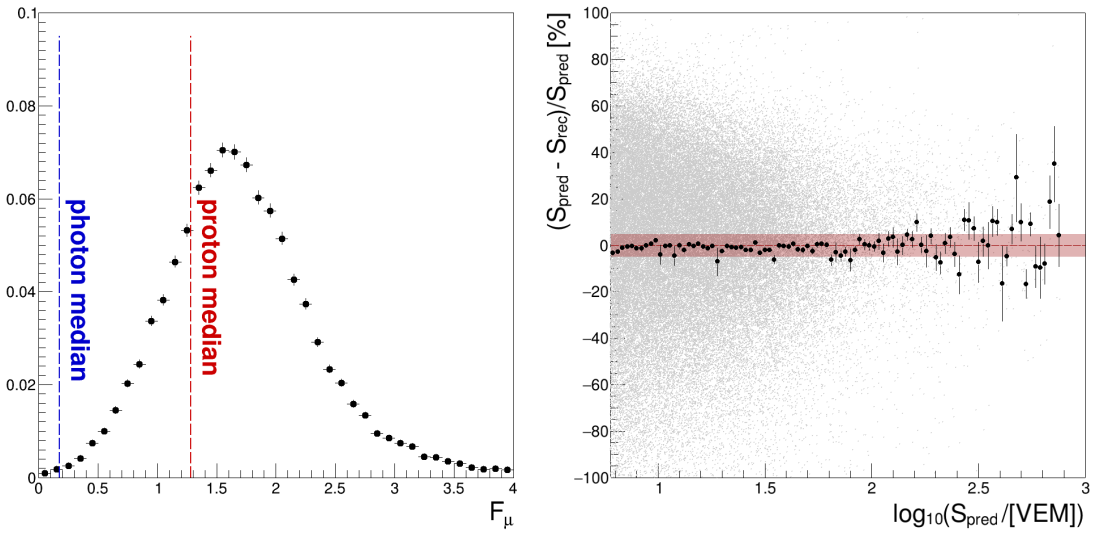
The merit factor, however, accounts only for the mean and width of the distributions, and not for their shape. A second, supplementary, measure of the separation power is the background rejection, i.e., the fraction of events in the proton-shower distribution rejected by a given cut value on the observable, as a function of the signal efficiency, i.e., the fraction of events in the photon-shower distribution that pass the given cut. This is reported as a red line in figure 5.13b. As reference value for the separation power, the background rejection at a signal efficiency of 50% (i.e. the cut value corresponds to the median of the photon distribution) is usually taken. A values of

99.5 % is obtained. As before, the corresponding background rejection for the observable $S_b$ has also been determined for comparison and it is shown in the same figure as a light blue line. For the same efficiency of 50%, the background rejection results to be 87.0 %.

It thus can be concluded that, for the same conditions, the separation power of $F_\mu$, in terms of both $\eta$ and background rejection, is larger than that of $S_b$, a parameter used earlier for photon searches with hybrid events over the same energy range considered here.

## 5.4 $F_\mu$: APPLICATION TO DATA

As a conclusion of this chapter, in this section the technique for the reconstruction of $F_\mu$ is applied to the hybrid data, presented in section 4.1. Such application allows, on the one hand, to verify the consistency of the method and, on the other hand, to check possible long-term effects, given the 13-years span of the data set.




(a) (b)

**Figure 5.14:** *(5.14a): Distribution of $F_\mu$ estimated from hybrid data. The vertical dashed lines are the medians of the proton (red) and the photon (blue) distributions. (5.14b): Relative difference between the predicted and the measured signals as a function of the predicted signal. The red band corresponds to 5% difference.*

First, we show in figure 5.14a the distribution of $F_\mu$ calculated from data. The median of the distribution (dashed vertical line) is 1.9. It is interesting to note that this is about the mean value of $F_\mu$ for showers initiated by oxygen nuclei (see figure 5.13a). However, in the energy range of the data, between $10^{18}$ eV and $10^{19}$ eV, the mass composition, as inferred from $X_{max}$ data [172], is mixed, but still dominated by protons. The (too) large value of the median is coherent with the

excess in the number of muons, measured with several methods in Auger data, with respect to simulations, independently from the hadronic-interaction model [155].

The check of the method consistency is carried out by comparing the signal predicted by universality with that reconstructed in the SD stations satisfying the criteria defined in section 5.2.2. The difference between the two signals as a function of the predicted one is shown in figure 5.14b. To calculate the predicted signal from data, the median of the $F_\mu$ distribution, 1.90, is used. As one can see from the figure, the average relative difference between the reconstructed and predicted signal is better than 5 % (highlighted by the red band). This is a remarkable achievement, as the universality model has been in fact designed to work at energies above $10^{18.5}$ eV. It implies that the number of secondary particles, produced in showers in the energy range considered in this analysis, is sufficient to satisfy the physics requirements of universality.

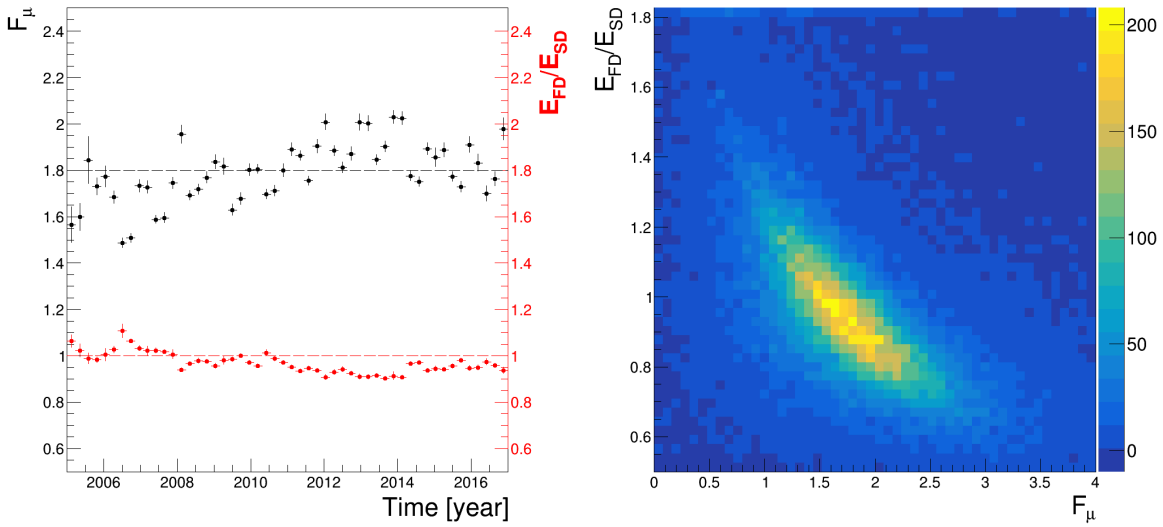



**Figure 5.15:** *(a). Evolution of the monthly average of $F_\mu$ (black points) and of the ratio between SD and FD energies (red points) over years 2014-2017. (b) Correlation between the monthly average of $F_\mu$ and the ratio between SD and FD energies.*

Potential long-term effects on the $F_\mu$ estimation, due either to the aging of the SD detectors or of the FD telescopes, are investigated by studying the variation of the average value of $F_\mu$ as a function of time. Figure 5.15a shows the behavior of the monthly average of $F_\mu$ (black points) from 2004 to 2017. An increasing trend is apparent, that one can compare with the evolution of the average ratio of the SD and FD energy (red) over time, in the same period. A seasonal effects is visible, as well as a decreasing drift of about $\sim 1$ % per year. The correlation between $F_\mu$ and the energy ratio is highlighted in figure 5.15b, where one can see that the two effects are, in fact, totally anti-correlated. Given that the drift in the energy scale is accounted for in the budget of

the systematic uncertainty of the energy, and given that the observed trend in time of $F_\mu$ is due to such drift, no systematic uncertainty due to this correlation will be associated to $F_\mu$.

In the next chapter, the technique for the search of UHE photons based on the combination of two mass-sensitive observables, the newly defined $F_\mu$ parameter and the traditional $X_{\max}$, will be described.

# 6 COMBINATION OF $F_\mu$ WITH $X_{MAX}$ TO SEARCH FOR UHE PHOTONS

## Contents



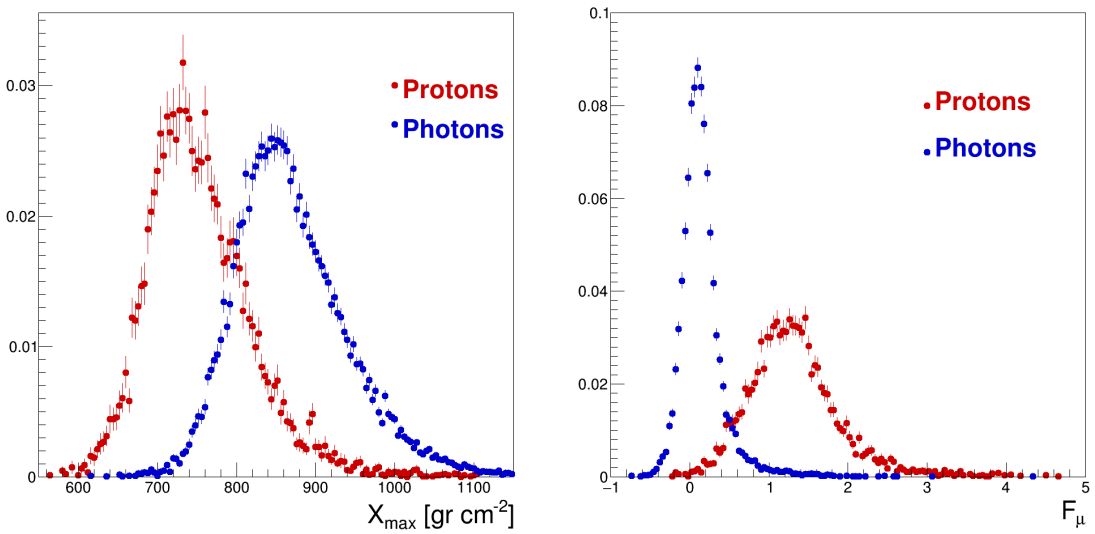
The new SD-based mass-sensitive observable, $F_\mu$, discussed in the previous chapter, has been shown to provide a very good photon-hadron separation, even when derived from the signal of one SD station only, in hybrid events. To fully exploit the hybrid approach, in this chapter $F_\mu$ is combined with another mass-sensitive variable measured with the FD, $X_{max}$, the depth of the shower maximum, so to further improve the photon-hadron separation power.

The combination of the two observables is discussed in section 6.1, where simulated hybrid events generated by protons and photons are used to test two multivariate classification methods, namely the boosted decision tree and the Fisher discriminant analysis. The reasons to choose the Fisher analysis, as well as its separation power, are explained. In the following section 6.2, we illustrate the approach used to determine the Fisher-discriminant threshold to select photon-candidates. Not only simulations are used, but also hybrid data, namely the burnt sample, presented in Chapter 4. From the Fisher distribution of simulated protons, we determine the functional form that best describes the background. Then, by fitting the function to the burnt sample, and extrapolating it to the full sample, a data-driven parametrisation of the background is obtained. Such a parametrization is then used to obtain the threshold value of the Fisher discriminant for the photon-candidate selection, as well as to calculate the number of the false-positive background events expected in the full data sample. Finally, in section 6.3, the selection is applied to the burnt sample as a verification of the analysis, before unblinding the full data set in the next chapter.

## 6.1 THE COMBINATION OF $F_\mu$ WITH $X_{MAX}$

In this section we define the method to combine $F_\mu$ with $X_{\max}$ for the search for photon primaries with hybrid data. For each event, $X_{\max}$ is directly provided by the hybrid reconstruction, while $F_\mu$ is calculated as described in the previous chapter. The performance of the two combined observables in terms of photon-hadron discrimination is expected to be better than for each of them separately. The separation power of the two variables is studied by using simulations of showers initiated by photons (signal) and by protons (background) in the energy range between $10^{18.0}$ eV and $10^{19.5}$ eV. After describing the two variables and their relation in section 6.1.1, two multivariate classification methods are considered to combine them, both performed by using the Toolkit for Multivariate Data Analysis in ROOT (TMVA) [173]. Following the approach used in previous photon searches with hybrid data [114], we first consider, in section 6.1.2, the Boosted Decision Tree, trained on different sets of variables. Then, in 6.1.3, we study the performance of the linear Fisher Discriminant Analysis, as this is the most appropriate method when the input parameters are uncorrelated, like $F_\mu$ and $X_{\max}$ are.

### 6.1.1 The $F_\mu$-$X_{\max}$ approach

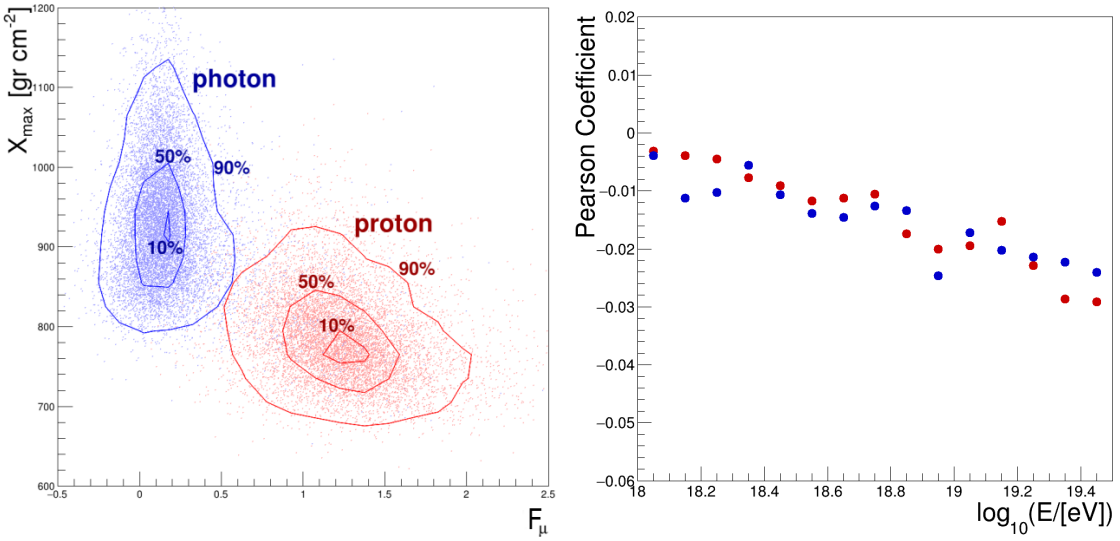



(a) (b)

**Figure 6.1:** *$X_{max}$ (6.1a) and $F_\mu$ (6.1b) normalized distributions for simulated photons (blue) and protons (red).*

To discuss the potential performance of the combination of $F_\mu$ and $X_{\max}$, we first show together, in figure 6.1, the distributions of the two reconstructed parameters in simulated hybrid events

initiated by protons (red) and photons (blue). The simulations are re-weighted to realistic energy spectra $E^{-\gamma}$, i.e., $\gamma = 2.7$ for protons and $\gamma = 2.0$ for photons.



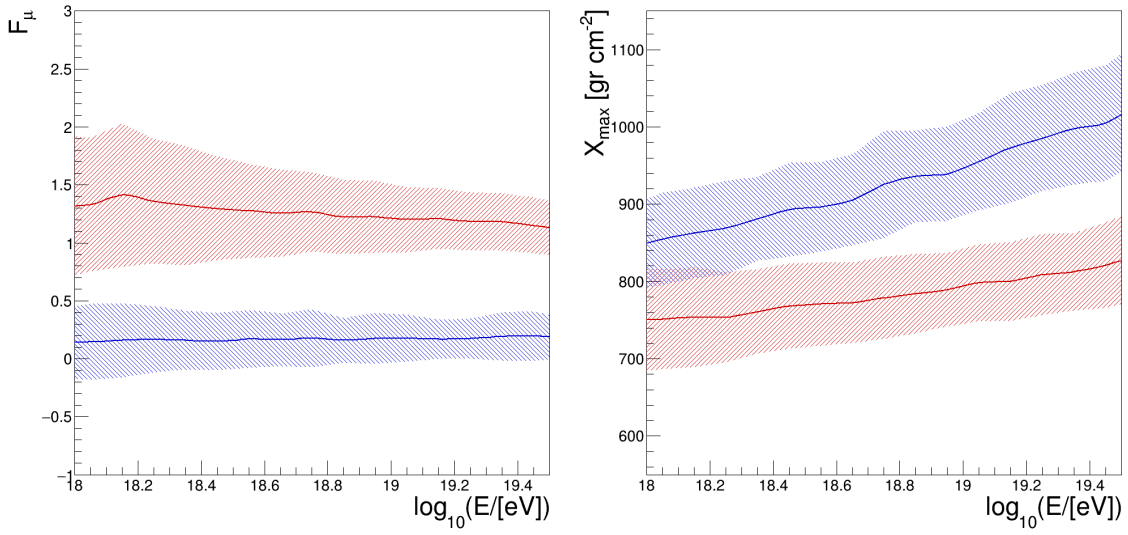

**(a)** **(b)**

**Figure 6.2:** ***(a):*** *$X_{max}$-$F_\mu$ distributions for photons (blue) and protons (red). Contour lines enclose the 90 %, 50 % and 10 % of the distributions of the events, re-weighted to a realistic power law spectrum $E^{-\gamma}$ ($\gamma = 2.7$ for protons and $\gamma = 2.0$ for photons).* ***(b):*** *Pearson correlation coefficient as a function of energy for photons (blue) and protons (red).*

The proton and photon distributions are already well-separated for both variables independently. Such separation becomes more striking when combining the two parameters, as shown in figure 6.2a. The blue (red) contour lines enclose the 90 %, 50 % and 10 % of the photon (proton) distributions. One can thus see that they have two clearly separated peaks, with minimal overlapping tails, a fact that has been already observed in figure 1.9, where the combination of the number of muons, $N_\mu$, with $X_{\max}$ has been shown.

The figure shows also that $F_\mu$ and $X_{\max}$ do not show any significant degree of correlation. To quantify this, we calculate the Pearson correlation coefficient, $r$, defined as the covariance of the two variables divided by the product of their standard deviations [174]. By construction, $r$ ranges between $-1$ and 1. A full correlation (anticorrelation) is characterized by $r = 1$ ($r = -1$), while for $r = 0$ the two variables are not correlated. Figure 6.2b shows the $r$ coefficient as a function of the energy for photons (blue) and protons (red). The observed $|r|$ is $< 0.03$ over the entire energy range, confirming that the correlation between $X_{\max}$ and $F_\mu$ is negligible.

Finally, for the sake of the combination of the two variables in a multivariate analysis, the behavior of the two variables as a function of energy is studied, as any energy dependence would reflect in the separation power. Figure 6.3 shows the evolution of $F_\mu$ (left panel) and of $X_{\max}$ (right panel) over energy, for protons (red) and photons (blue). While $F_\mu$ is almost independent from the




**(a)** **(b)**

**Figure 6.3:** *Evolution of the average $F_\mu$ (left) and $X_{max}$ (right) over energy for protons (red) and photon (blue). The shaded areas enclose one standard deviation of the distribution.*

primary energy for both primaries, $X_{\max}$ is linearly increasing with the logarithm of the energy, faster for photons than for protons. This implies that, given that the separation power of $X_{\max}$ increases with energy, the latter is to be used as additional input parameters in the MVA.

Note that, in the following, we do not use $E_{\mathrm{FD}}$, the energy reconstructed with FD, as estimator for the energy of each event. This is because $E_{\mathrm{FD}}$ includes an invisible energy correction calculated from data, which mostly consist of air showers initiated by protons and nuclei. If that is used for photon primaries, it would lead to an overestimate of the primary energy. An estimator of the primary photon energy, $E_\gamma$, is thus defined as $E_\gamma = (1 + 1\,\%)E_{\mathrm{cal}}$ , where the calorimetric energy, $E_{\mathrm{cal}}$, is corrected by a 1 % term that corresponds to the invisible energy contribution in electromagnetic showers [175]. In the following, unless differently specified, $E_\gamma$ is used as default for simulations and data, independently of the nature of the primary particle.

### 6.1.2 Boosted Decision Tree

A decision tree is a binary structured classifier, trained using signal/background simulations, denoted as *training sample*. Repeated binary (yes/no) decisions are taken on one single variable at a time until a stop criterion, which represents the classification target, is fulfilled. Thus, the phase-space is split into many regions that are eventually classified as signal or background, depending on the majority of training events that end up in the final *leaf* node. The output of a decision tree is then associated to the probability that an event belongs to the signal class.

The *boosting* is a technique used to improve the classification performance as well as to increase the stability with respect to the fluctuations of the training sample. The boosting operation is performed by applying the selection algorithm to re-weighted versions of the training sample. The output of the boosted decision tree (BDT) is then given by the weighted average of the single decision trees.

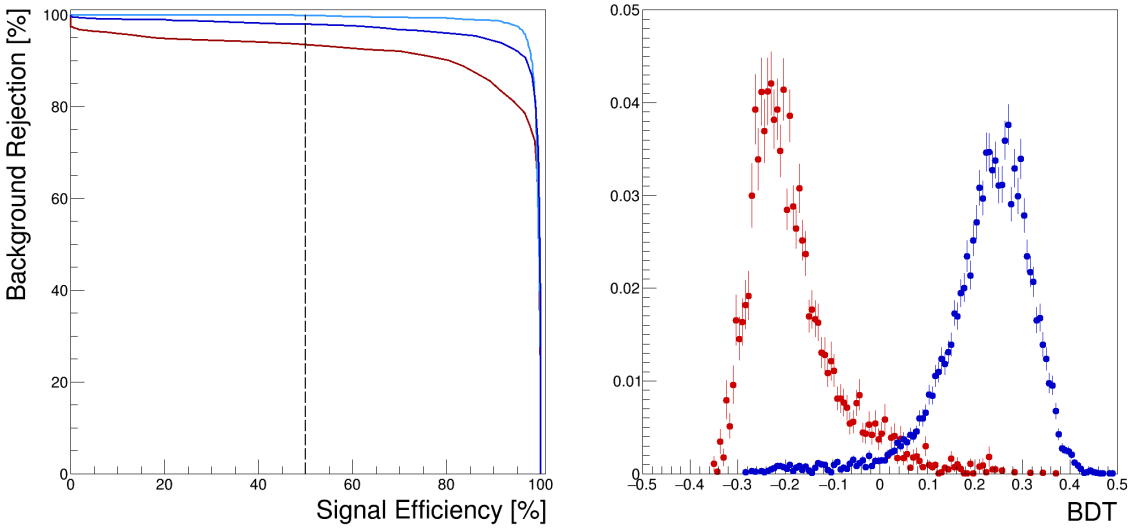



**Figure 6.4:** ***(6.5a):*** *Background rejection as a function of signal efficiency obtained with the BDT applied to the three sets of parameters (a), (b), (c) (see text), represented by the solid red, light blue and dark blue, respectively. The background rejection at a signal efficiency of 50% is indicated by the dashed black line.* ***(6.5b):*** *BDT distribution obtained from the (b) set of parameters, for protons (red) and photons (blue).*

In the literature, decision trees are sometimes referred to as the best "out of the box" classifiers. This is because little tuning is required in order to obtain reasonably good results due to the simplicity of the method. Due to these reasons, and also because in previous Auger analyses using hybrid events to identify primary photons, a study of the BDT performance has been carried on this thesis. The BDT technique has been applied on three sets of parameters:

(a) The first set includes the three parameters discussed above, $F_\mu, X_{\max}, \log_{10}(E_\gamma)$.

(b) The second set is that used in [114], for the sake of comparison. There, the classification has been performed using $\log_{10}(S_b), X_{\max}, \log_{10}(E_\gamma), \theta, n_{\text{station}}$, where $\theta$ is the zenith angle, $n_{\text{station}}$ is the number of stations triggered in an event, and $S_b$ is the SD-based parameter, explained in section 5.1), that, being sensitive to the shape of the lateral distribution function, is sensitive to the primary mass.

(c) The third set is composed by $F_\mu, X_{\max}, \log_{10}(E_\gamma), \theta, \log_{10}(S_b)$. It serves as a test of a possible increase in the separation power of the analysis when combining $F_\mu$ with $S_b$.

The performances of the BDT in the three cases are shown in figure 6.4a in terms of background (proton) rejection as a function of signal (photon) efficiency. The red, light blue and dark blue lines correspond to parameter sets (a), (b) and (c), respectively. The vertical black line indicates the background rejection at a signal efficiency of 50%, as a reference. The best classification performances are obtained for the (b) parameter set (light blue line). This is due to the fact that BDTs are in general very performing when there are correlations between the input parameters, which is the case for this set, given that $S_b$ is strongly correlated with the number of stations, the zenith angle and the energy. For the sake of completeness, figure 6.4b shows the BDT distribution obtained by using the (b) parameter set, for protons (red) and photons (blue). The BDT performances are in turn the worst ones, as expected, for the (a) set (red line) because, as we have seen, $X_{\max}$ and $F_\mu$ are essentially non-correlated. Since the BDT trained on the (b) set of parameters is found to best perform, it will be used for the comparison with the Fisher discriminant analysis in the next section.

### 6.1.3 Fisher Discriminant Analysis

In the Fisher Discriminant Analysis (FDA) [176], the parameters provided as input for the analysis are combined linearly to obtain the Fisher discriminant, $f$. The event classification is then performed in the transformed $f$ space. The linear discriminant analysis identifies an axis in the hyperspace of the input variables such that, when projecting the output classes (signal and background) upon this axis, the separation between the two classes is maximized, while the dispersion of the simulated events within each class is minimized. Fisher discriminants have the best separation performances in the case of uncorrelated, or linearly correlated input observables following a gaussian distribution, while no discrimination at all is achieved when a variable has the same sample mean for signal and background, even if the shapes of the distributions are very different.

The advantages of the Fisher discriminant, $f$, are that it provides a robust event classification for uncorrelated input observables, which is the case for $F_\mu$ and $X_{\max}$, and that it can be calculated analytically for each event. Namely, $f$ is calculated as

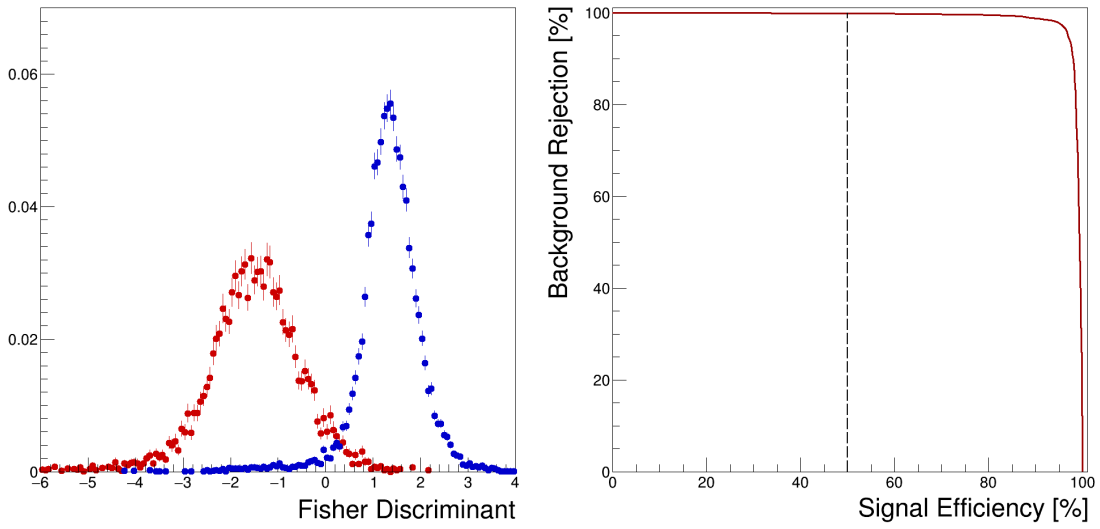
$$f = c_0 + \sum_{i=1}^{N} c_i x_i \tag{6.1}$$

where $x_i$ is the i-th of $N$ variables used as input for the analysis, and $c_i$ are the correspondent Fisher coefficients. The offset $c_0$ centers the mean of the overall signal/background $f$ distribution at zero. Here, the FDA is performed using $F_\mu$, $X_{\max}$ and $\log_{10}(E_\gamma)$ as three input parameters. Table 6.1 includes the coefficients obtained from the TMVA. For each observable, the ranking is also calculated and indicated in the table. This is related to the discrimination power of each parameter: in our case, the most discriminating one results to be $F_\mu$, as expected, thanks to its lack of energy dependence.

| Parameter | Fisher coefficient | Ranking |
|---|---|---|
| $F_\mu$ | -1.669 | 0.513 |
| $X_{\max}$ | +0.007 | 0.349 |
| $\log_{10}(E_\gamma)$ | -0.453 | 0.019 |
| bias | +3.399 | |

**Table 6.1:** *Fisher coefficients obtained from the TMVA.*

In figure 6.5a the resulting distributions of the Fisher discriminant are shown for photons (blue) and protons (red). The corresponding behavior of background rejection as a function of the signal efficiency is in turn shown in figure 6.5b.




**(a)** **(b)**

**Figure 6.5:** ***(6.5a):*** *Performance of the combination of $F_\mu$, $X_{max}$ and $E$ with the Fisher analysis: distribution of the discriminant for photons (signal, blue) and protons (background, red).* ***(6.5b):*** *Background rejection as a function of signal efficiency obtained with the Fisher analysis.*

The background rejection vs signal efficiency curve obtained with the FDA is compared in figure 6.6 with that obtained with the best performing BDT, reported in section 6.1.2. The inset highlights the region where the background rejection is larger than 98%: the two curves (red for FDA, light blue for BDT) are very similar, the background rejection being found to be, for both analyses, around 99.90 % for a signal efficiency of 50 % and 98.80 % for a signal efficiency of 90 %.

In conclusion, the combination of $X_{\max}$ and $F_\mu$ in a Fisher analysis that accounts also for the energy is as performing as a BDT including $S_b$ and other variables, as it was done in previous

searches for UHE photons based on hybrid events. Overall, the advantage of the approach adopted in this thesis work is twofold. On the one hand, the SD-based variable used here, $F_\mu$, depends only weakly on shower geometry or its energy, in contrast with $S_b$. On the other hand, such lack of dependencies and the lack of correlation with $X_{\max}$ allows us to use a method, the FDA, which is simpler, more robust and more transparent than the BDT [173].

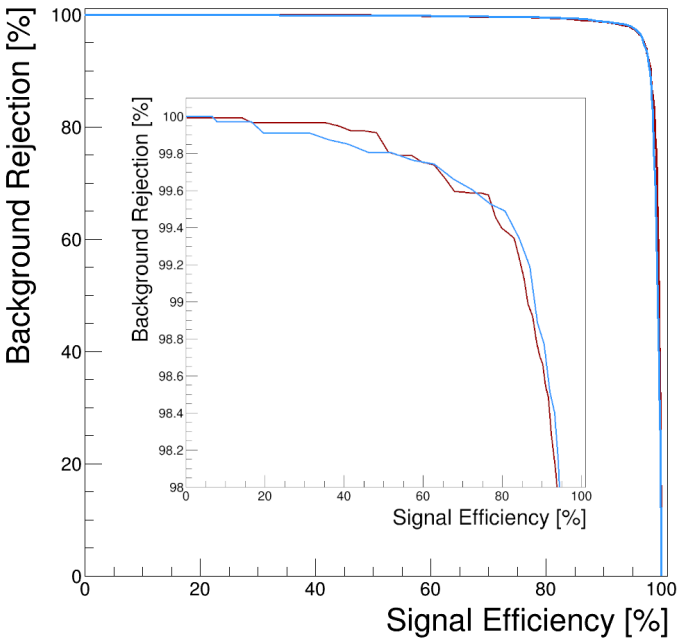



**Figure 6.6:** *Background rejection as a function of signal efficiency obtained with the Fisher analysis (red) and the BDT (light blue). The inset shows the background-rejection region above* 98 %.

## 6.2 THE PHOTON-SELECTION CUT

In this section, we describe the method adopted for determining the photon- in the Fisher discriminant. To this aim, not only simulated hybrid events are used, but also data. As the selected strategy for the search for UHE photons is that of a blind analysis, only the burnt sample is exploited here, corresponding to 5 % of the total data set, i.e, 6675 events. The Fisher distributions for the simulations (protons in red, photons in blue) and the burnt sample are shown in figure 6.7.

First, in section 6.2.1 a model for the functional form of the background is developed by using the proton simulations. This model is then fitted to the events of the burnt sample so to obtain a data-driven description of the background, not depending either on assumptions on cosmic ray composition, or on hadronic interaction models. The developed model is then used in section 6.2.2, to determine the threshold value of the Fisher discriminant most efficient for the photon

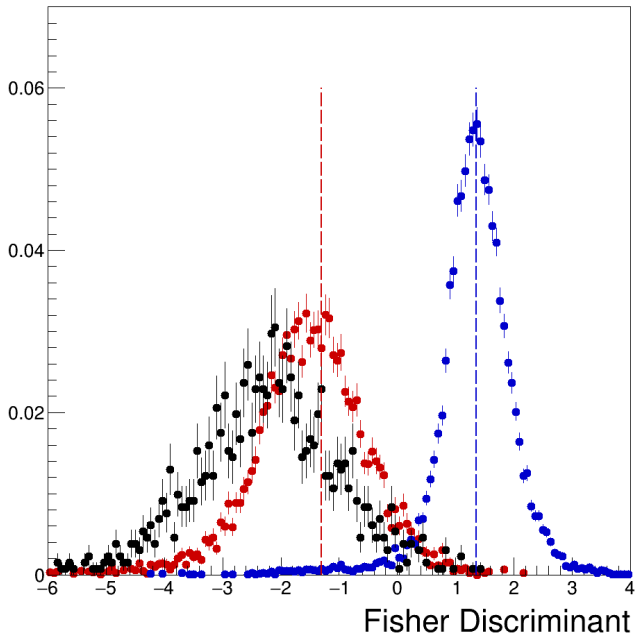



**Figure 6.7:** *Distribution of the Fisher discriminant for simulated photons (signal, blue) and protons (background, red), and for the burnt sample (black). The vertical red line marks the tail of the proton distribution, the blue one indicates the median of the photon distribution.*

selection, as well as to evaluate the number of false positive background events expected out of the unblinding of the total hybrid data sample.

### 6.2.1 Study of the background

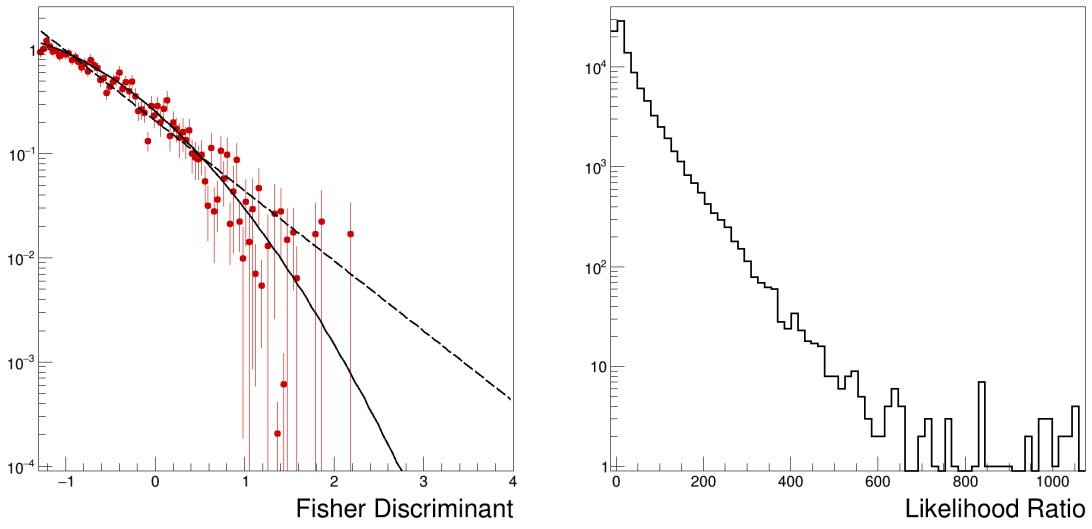
The goal of this section is to describe the distribution of the Fisher discriminant for the background. This is achieved in two steps.

In the first step, we study its shape by profiting of the statistics offered by the proton simulations. Only the rightmost tail of the Fisher distribution is considered, specifically only the events with a Fisher discriminant $f > -1.3$, indicated by the red vertical line in figure 6.7. This value of the Fisher discriminant is used because below $f_0 = -1.3$ the photon selection efficiency is almost 100 %.

The tail of the proton distribution is highlighted in figure 6.8a, where two exponential functions, $m_1$ and $m_2$, tested for its description, are superimposed (black dashed and solid lines, respectively):

$$m_1(f|B) = N_1(B)e^{-Bf} \tag{6.2a}$$

$$m_2(f|A,B) = N_2(A,B)e^{-(Af^2+Bf)} \tag{6.2b}$$




(a) (b)

**Figure 6.8:** ***(6.8a):*** *Tail of the Fisher distribution for protons. The two black lines represent the fits of the two functions, $m_1$ (dashed) and $m_2$ (solid), discussed in the text.* ***(6.8b):*** *Likelihood-ratio distribution from the fit of $m_1$ and $m_2$ to $10^6$ simulated samples (see text for details).*

where $A$ and $B$ are shape parameters, and $N_1$ and $N_2$ are the normalizations of $m_1$ and $m_2$, respectively. $N_1$ and $N_2$ are calculated as a function of the parameters $A$ and $B$, by requiring that the integral of $m_1$ and $m_2$ is equal to the number of events $N$ that have a value of the Fisher discriminant above $f_0$, thus obtaining

$$N_1(A) = \frac{N_0 B}{e^{-B f_0}} \tag{6.3a}$$

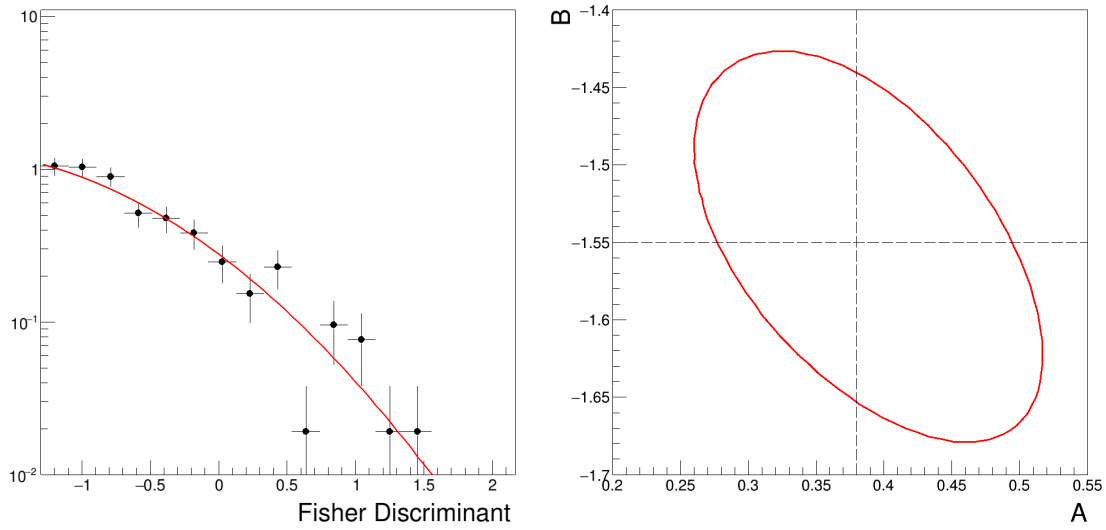
$$N_2(A,B) = \frac{N\sqrt{A}}{e^{B^2/4A}\,\mathrm{erfc}\left(\frac{B}{2\sqrt{A}}\left(\frac{2A}{B}f_0 - 1\right)\right)} \tag{6.3b}$$

where erfc is the complementary error function. The parameters obtained from an unbinned likelihood fit of $m_1$ and $m_2$ to the tail of the Fisher distribution are reported in table 6.2.

| | **A** | **B** |
|---|---|---|
| $m_1$ | | 1.55 |
| $m_2$ | 0.42 | -1.73 |

**Table 6.2:** *Values of the parameters A and B obtained from an unbinned likelihood fit to the tail of the Fisher distribution of protons, i.e., to events with $f > -1.3$.*

The best-fit model is determined by using a likelihood-ratio test [177], in which two hypotheses on the shape of the tail distribution are compared: the null-hypothesis, $H_0$, according to which it is described by $m_1$, i.e., $m_1(f|B) = m_2(f|A = 0, B)$; the alternative hypothesis, $H_1$, according to which it is described by $m_2(f|A \neq 0, B)$. The likelihood ratio $\mathcal{L}_{\text{ratio}}(\text{BS})$ results to be $\approx 4000$. The p-value, $p_{\text{value}}(\text{BS})$, associated to $\mathcal{L}_{\text{ratio}}(\text{BS})$ is derived by applying the likelihood-ratio test on simulated samples of Fisher values, generated according to the $m_1$ model and then fitted with both models. Each sample consists of 30000 events (*realizations*). The resulting distribution of the likelihood-ratios, based on 1 000 000 realizations, is shown in figure 6.8b. As the maximum value attained in $10^6$ trials is about 1000, i.e., $p_{\text{value}}(1000) \approx 10^{-6}$, then $p_{\text{value}}(\text{BS}) < 10^{-6}$, i.e., the $m_1$ model is discarded in favor of $m_2$.

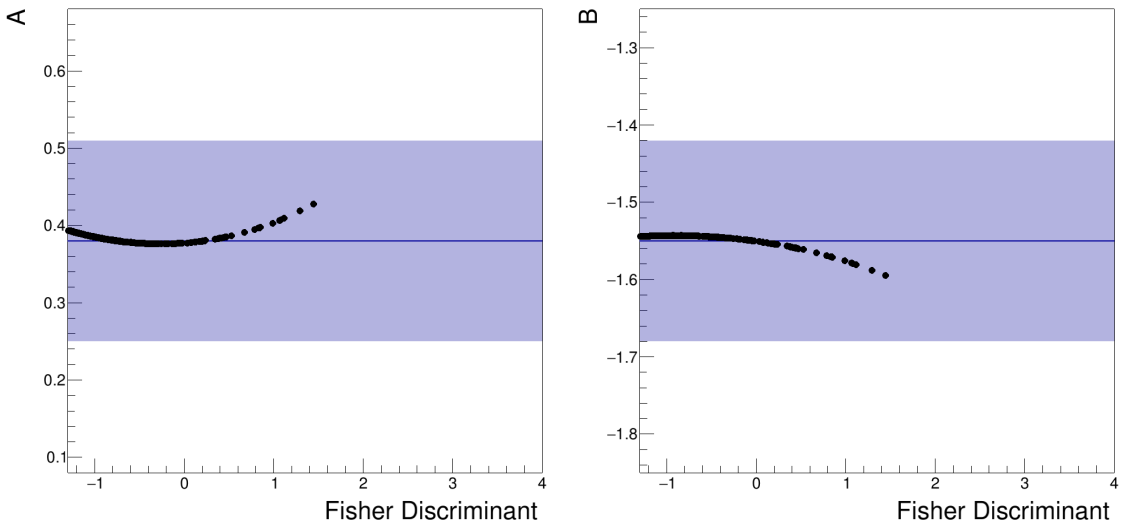



**(a)** **(b)**

**Figure 6.9:** ***(6.9a):*** *Tail of the Fisher distribution for the burnt sample (black points). The red line represents the fit of the $m_2$ function to the data.* ***(6.9b):*** *1-sigma contour plot of the errors of the parameters $A$ and $B$. The dashed lines indicate the values obtained from the fit of $m_2$ on the burnt sample.*

After having derived the shape of the background from proton simulations, the second step in the characterisation of the background involves the burnt sample. To finalize the estimation of the background, we fit the $m_2$ model on the burnt sample distribution, as it is shown in figure 6.9a. The best-fit values of $A$ and $B$ are 0.38 and −1.55, respectively: they are represented by the grey dashed lines in figure 6.9b, together with the red ellipses that marks the 1-sigma contour of the statistical errors.

A possible photon contamination in the burnt sample cannot however be excluded: a possibly related systematic effect has thus been studied by using a jackknifing technique [178]. This is a re-sampling technique, which involves a leave-one-out strategy for the estimation of the parameters (in this case, $A$ and $B$) in a data set of $N$ observations. The values of $A$ and $B$ calculated as a




**(a)** **(b)**

**Figure 6.10:** *Values of the parameters A (6.10a) and B (6.10b) as calculated with the jackknifing technique (see text), as a function of the Fisher discriminant. The blue lines show the values obtained from the fit on the burnt sample, while the blue shaded areas show the statistical uncertainties.*

function of the Fisher discriminant $f$ with this technique are shown in figures 6.10a and 6.10b, respectively. The blue shaded area represents the statistical uncertainties obtained from the fit. As one can see, the systematic deviations from the central values (marked by the blue lines) are negligible with respect to the statistical ones. A systematic uncertainty of 0.01 is then derived from the width of the distribution of the deviations when projected on the y-axis.

The parameters of the burnt sample obtained from the fit of the model $m_2$ to the burnt sample distribution are thus:

$$A = \quad 0.38 \pm (\text{stat})\ 0.13 \pm (\text{syst})\ 0.01 \tag{6.4a}$$

$$B = -1.55 \pm (\text{stat})\ 0.13 \pm (\text{syst})\ 0.01 \tag{6.4b}$$

Finally, to extrapolate the parametrisation of the Fisher distribution of the background to the full hybrid data set, the normalization of the function $m_2$ is scaled to the number of total events by setting $N = N_{\text{data}}$ in equation 6.3b, where $N_{\text{data}} = 1328$. The distribution of the Fisher discriminant for the extrapolated background is shown as a blue line in figure 6.11. The uncertainties, $\sigma_f$, in the extrapolation, represented by the blue band, is calculated as

$$\sigma_f = \sum_{i,j=A,B} \frac{\partial m_2}{\partial i} k_{ij} \frac{\partial m_2}{\partial j} \tag{6.5}$$

where $i$ and $j$ runs over the parameters $A$ and $B$, and $k_{ij}$ are the elements of the covariance matrix:

$$K = \begin{pmatrix} 0.0165 & -0.0086 \\ -0.0086 & 0.0158 \end{pmatrix} \tag{6.6}$$

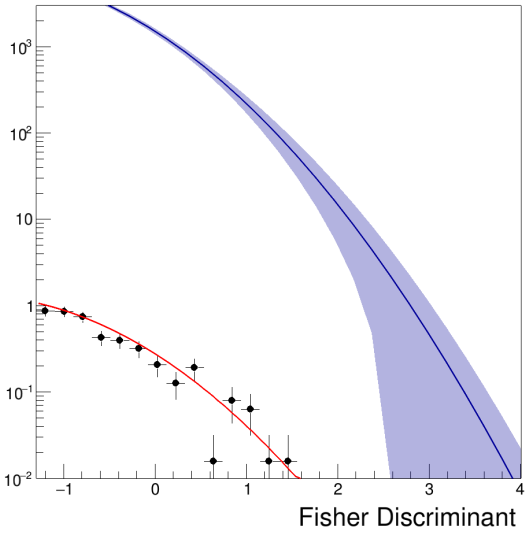



**Figure 6.11:** *Distribution of the Fisher discriminant for the background expected in the full hybrid data sample (blue line). The blue shaded area shows the statistical uncertainties on the expected background. For comparison, the corresponding distribution for the burnt sample is also shown (black points), together with the fit to the $m_2$ function (red line).*

### 6.2.2 Determination of the photon-selection cut

The background extrapolation obtained in the previous section 6.2.1 is used in this section to determine the Fisher discriminant cut, $f_\gamma$, to select photon candidates.

First, we search for a region of the phase space where there is no background contamination by examining, on the one hand, the photon selection efficiency $\varepsilon_\gamma$ above $f$ (figure 6.12a), and, on the other hand, the number of expected background events, $N_b$, above $f$, calculated from the integral of the extrapolated background (red line in figure 6.12b).

Given that the expectation of $N_b = 0$ from the background extrapolation realizes only when $f$ approaches infinity, the non-contaminated region is defined by requiring $N_b = 1$, corresponding to $f = 2.36$. As the photon selection efficiency for $f > 2.36$ results to be less than 0.04 %, we choose an alternative way to determine $f_\gamma$, namely, by minimizing the background to signal ratio.

The determination of $f_\gamma$ is achieved by computing the number of events corresponding to upper bound of the Rolke 95 % confidence interval [179] with respect to a given number of expected background events. The number of events at such upper bound is shown as a function of

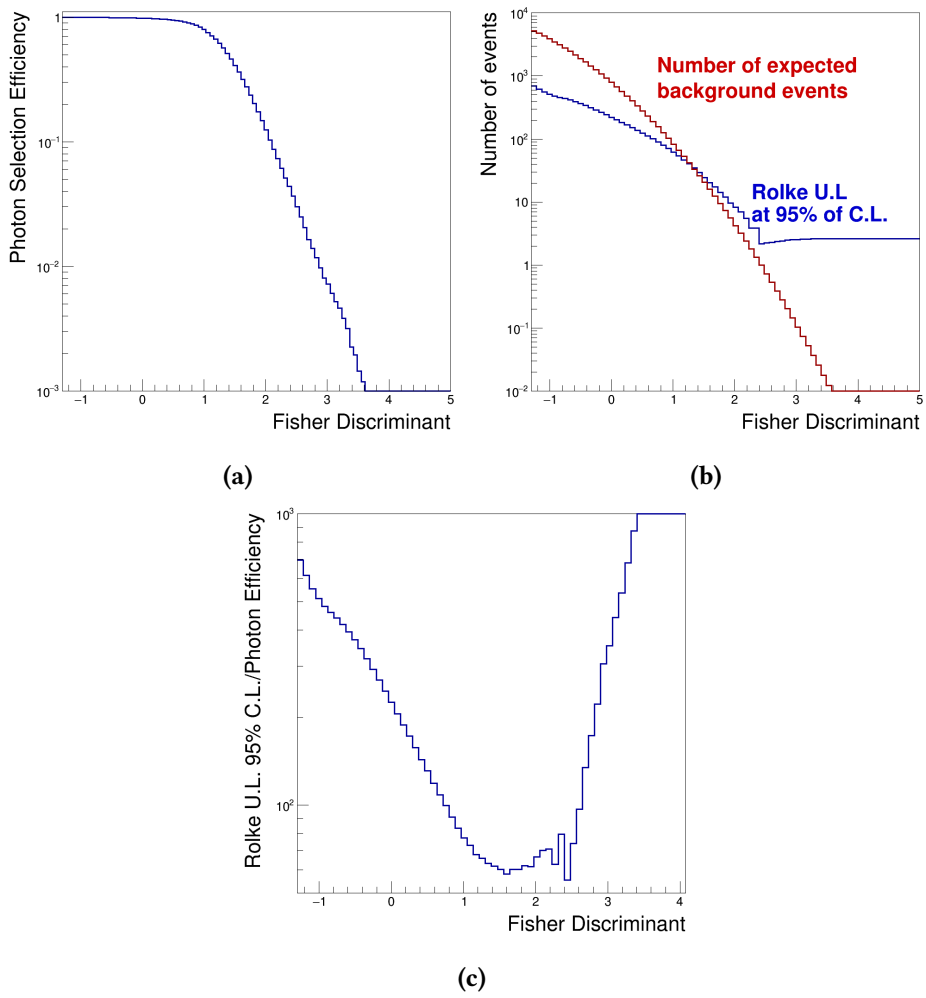



**Figure 6.12:** ***(6.12a):*** *Photon selection efficiency as a function of the Fisher-discriminant threshold.* ***(6.12b):*** *Number of expected background events (red) and number of events at the upper bound of the Rolke 95 % confidence interval (blue) as a function of the Fisher-discriminant threshold.* ***(6.12c):*** *Ratio between the number of expected background events and the photon selection efficiency as a function of the function of the Fisher-discriminant threshold.*

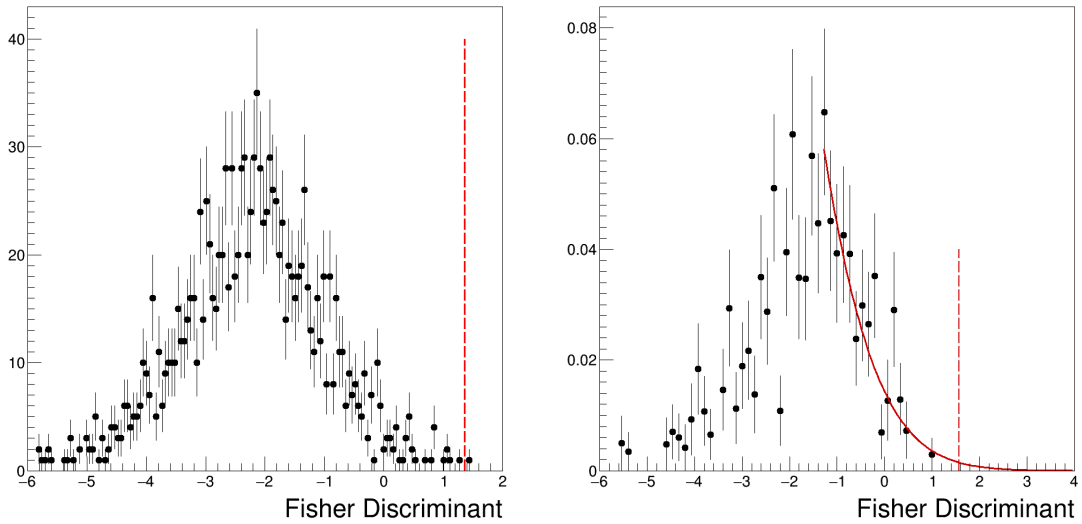
the Fisher discriminant as a blue line in figure 6.12b. The calculation has been performed assuming $N_b$ background events, with the uncertainty computed from equation 6.5, and zero signal events.

The ratio between the number of events at the upper bound of the Rolke confidence interval and the photon selection efficiency is a proxy for the background to signal ratio. This ratio, reported in figure 6.12 as a function of the Fisher discriminant threshold, presents a minimum at $f \simeq 1.6$. Since this minimum approximately corresponds to the median of the Fisher discriminant distribution for photons, $f \simeq 1.36$ (blue vertical line in figure 6.7), the value of such median is selected as $f_\gamma$. This approach has two benefits. The first is, as already discussed, that the back-

ground to signal ratio is minimized. Secondly, the upper limits are optimized in the case of a non-identification of a photon signal. The number of expected false-positive events in the full hybrid data set can be calculated by integrating the function describing the extrapolated background above $f_\gamma$, and it is found to be $30 \pm 15$.

## 6.3 THE PHOTON-SELECTION CUT APPLIED TO THE BURNT SAMPLE

To verify the photon-selection method developed in this chapter, in this last section we apply it on the burnt sample. The distribution of the Fisher discriminant in the burnt sample is shown in figure 6.13a with black dots. The vertical dashed red line corresponds to $f_\gamma$, i.e., the median of the photon distribution. As it can be seen, one event in the sample has a Fisher discriminant above $f_\gamma$, i.e., it is selected as photon-candidate. This is, in fact, consistent with the expectation of 1 false-positive candidate, as obtained by considering the background parametrization, described in section 6.2.1, and taking into account the number of events in the burnt sample.




(a) (b)

**Figure 6.13:** ***(6.13a):*** *Fisher-discriminant distribution in the burnt sample. The red dashed line corresponds to $f_\gamma$. One event is selected as a photon-candidate.* ***(6.13b):*** *Fisher-discriminant distribution in proton simulations characterized by the same geometry and energy of the selected candidate. The superimposed solid line is the result of fit to the background parametrization. The vertical dashed red line corresponds to the value of the Fisher discriminant of the selected candidate.*

Although compatible with a fluctuation of the background, the candidate event has been subjected to additional checks.

On the one hand, the reconstruction of the selected candidate has been inspected: the energy is $E = (1.87 \pm 0.21)$ EeV and the depth of the maximum is $X_{max} = (996 \pm 30)\,g\,cm^{-2}$. The event includes 3 triggered SD stations: table 6.3 shows the distance to the shower axis, the signal size and the reconstructed $F_\mu^{station}$ for each of them. For the calculation of the value of $F_\mu$ of the event, $F_\mu = 0.43$, only one station is used: the hottest station is rejected because too close to the core, while the lowest-signal station does not pass the cut on signal size.

| Signal Size (VEM) | Core Distance (m) | $F_\mu^{station}$ |
|---:|---:|---:|
| 766.9 | 343 | 6.07 |
| **3.8** | **1041** | **0.43** |
| 3.4 | 1821 | 7.81 |

**Table 6.3:** *Reconstructed parameters from the triggered SD stations in the candidate event in the burnt sample.*

On the other hand, it is of interest to quantify the probability that the selected candidate is actually a background event. To this aim, 1000 proton showers, characterized by the same geometrical configuration and energy of the candidate, have been simulated. The simulations are generated with CORSIKA by using EPOS-LHC as high energy hadronic model and following the same approach described in section 4.2. The resulting Fisher-discriminant distribution is shown in figure 6.13b, with black points. The dashed red vertical line represents the value of the Fisher discriminant of the candidate event. The solid red line superimposed to the distribution is the fit to the background parametrization, from which we derive the p-value associated to the Fisher value of the candidate, $p_{value} = 1.58 \times 10^{-2}$, which confirms that the candidate is a background fluctuation.

After the verification of the method on the burnt sample, the photon-selection procedure will be at last applied to the full hybrid data sample in the next and final chapter 7, where the results of the selection and the related physical implications will be illustrated.

# 7 SEARCH FOR UHE PHOTONS IN THE UNBLINDED HYBRID DATA SET

## Contents



In this chapter, the photon-search analysis discussed in chapter 6 is finally applied after unblinding the full hybrid data sample, which consists of 38 430 events with energy above $10^{18}$ eV collected between 1 January 2005 to 31 December 2017. This is presented in section 7.1, where the distributions of the data parameters relevant to the search for photons are shown, including that of the Fisher discriminant. After applying to the latter the photon-cut discussed in the previous chapter, 22 events are selected as *photon candidates*. Their Fisher distribution results well-consistent with the expectations from fluctuations of the hadronic background. The candidate events are however inspected, with particular attention to the most significant one.

Since no significant excess of a photon signal with respect to the background has been identified, in section 7.2 upper limits to the differential flux of UHE photons are calculated. Systematic effects which may impact the derived limits are also studied. Finally, we conclude in section 7.3, by discussing the physics implications of the upper limits obtained.

## 7.1 UNBLINDING THE DATA

The unblinding of the data, after application of the selection criteria described in section 5.2.1, results in a sample of 31877 hybrid events, which are visualised in figure 7.1, where the distributions of the main reconstructed parameters are shown, namely of energy, zenith angle, $X_{\max}$ and $F_{\mu}$. In the two latter, which are the parameters entering the Fisher analysis used to search for photons, the values of the median of the distributions for protons (dashed red line) and photons (blue dashed line) are superimposed for reference.

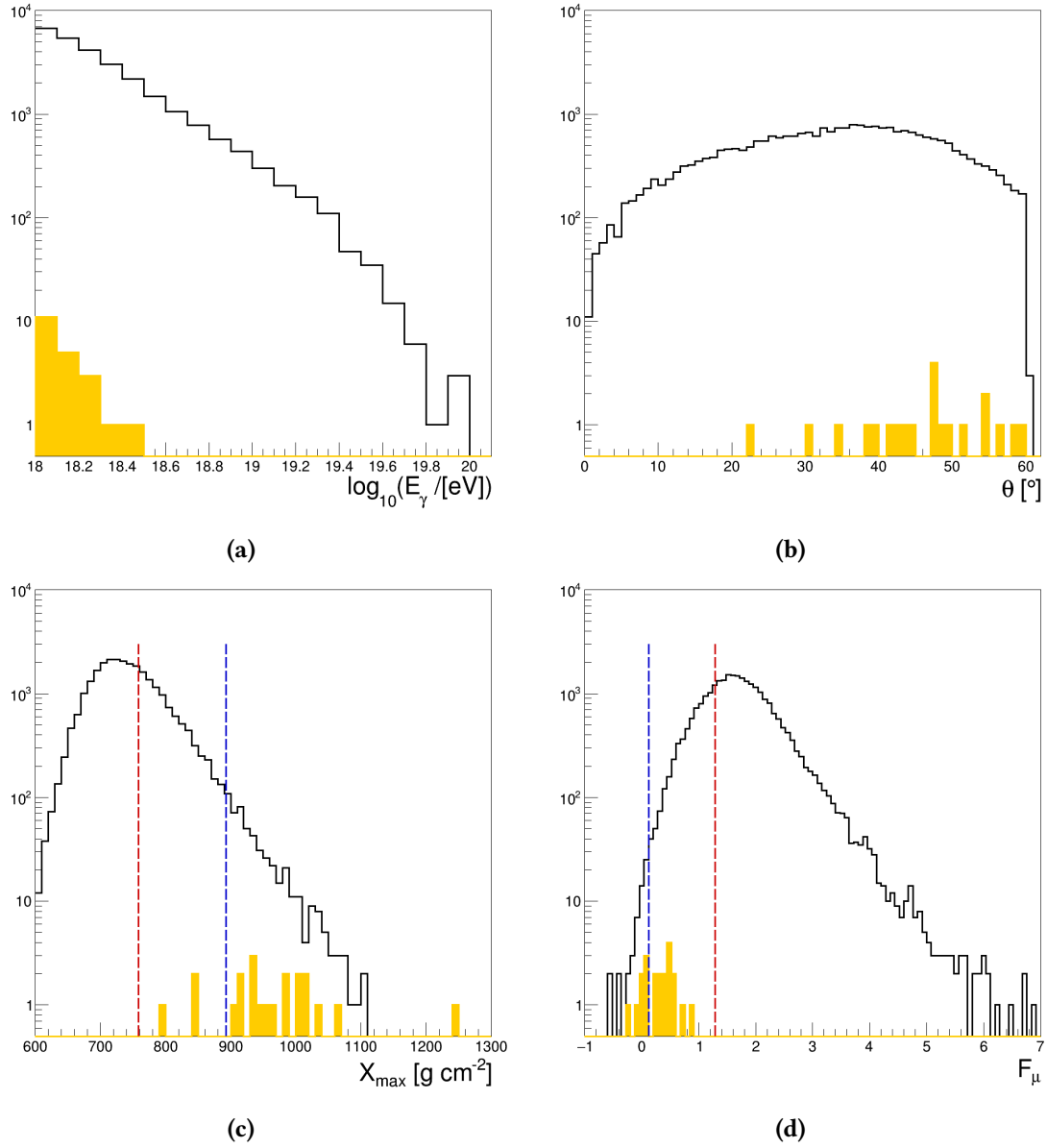
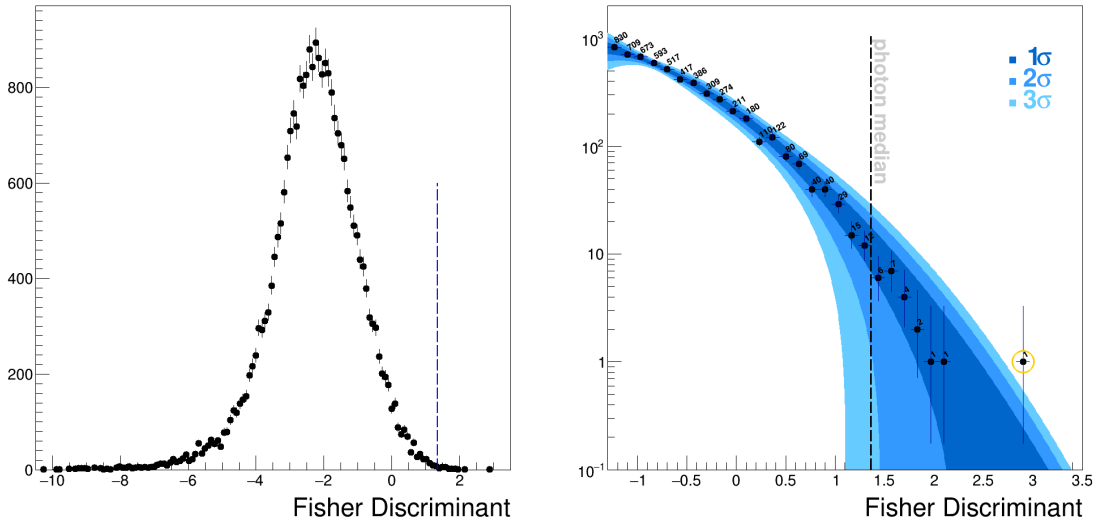



**Figure 7.1:** *Energy (7.1a), zenith (7.1b), $X_{max}$ (7.1c) and $F_\mu$ (7.1d) distributions of the hybrid data sample. The values of the reconstructed parameters of the 22 selected events are shown in yellow. For comparison, the proton (red) and photon (blue) medians are reported with a dashed vertical line in the $X_{max}$ and $F_\mu$ distributions.*

The Fisher-discriminant distribution obtained by combining $X_{\max}$ and $F_\mu$ in the data set is displayed in figure 7.2a, where the vertical dashed-blue line represents the photon selection cut, which, as discussed in the previous chapter, is taken as the median of the Fisher distribution of simulated photons. Figure 7.2b is a zoom on the tail of the Fisher distribution, that includes $\sim 5600$

events: the black points are the value obtained from data, while the shaded blue bands represent those expected from the background, including the uncertainty in its estimation at different sigma levels.




(a) (b)

**Figure 7.2:** *(7.2a): Fisher discriminant distribution of the selected hybrid data sample; (7.2b): Tail of the Fisher discriminant distribution of the hybrid data sample (black dots). The vertical dashed line represented the photon-median cut. The shaded blue areas show the 1, 2, 3 $\sigma$ uncertainties in the expected background.*

As one can see, the data distribution is coherent with that from the background expectation. The median selection cut yields 22 candidates, that is well-consistent with the expectation of $30 \pm 15$ false-positive candidates, as calculated in the previous chapter by considering the background parametrization.

The values of the reconstructed parameters of the 22 selected events are shown in yellow in the four panels of figure 7.1: such characteristics are more extensively discussed below, in section 7.1.1. Also, the features of the most peculiar event, whose Fisher discriminant is the one that deviates most from the background expectations (indicated with a yellow circle) in figure 7.2b are inspected in section 7.1.2.

### 7.1.1 The 22 candidates

For the sake of completeness, we inspect here the distributions of key parameters of the candidate events. The distributions of energy, zenith angle, $X_{\max}$ and $F_\mu$ for the 22 candidates are shown in figure 7.3: in all panels the values of the most significant candidate are highlighted in yellow.

Figure 7.3a shows that all the candidates selected are characterized by small values of $E_\gamma$. As one can see from figure 7.3b, the selected candidates are all characterized by a deep $X_{\max}$, as

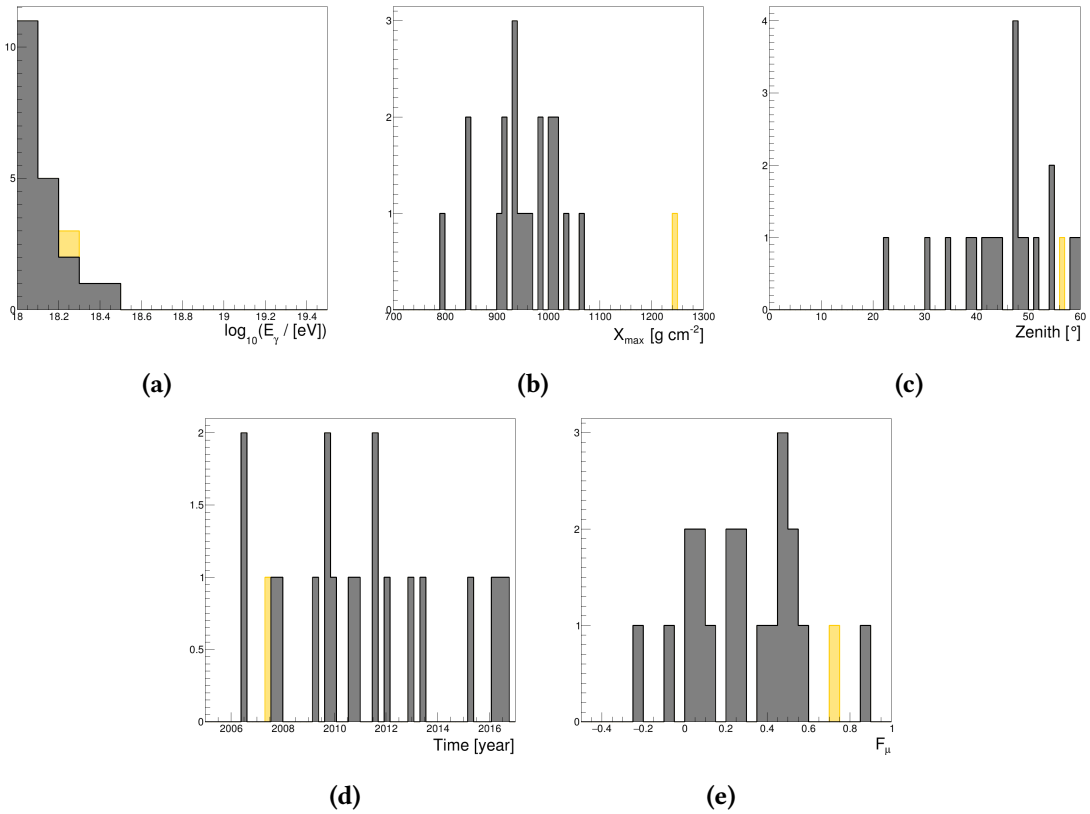



**Figure 7.3:** *$X_{max}$ (7.3b) and zenith (7.3c) distributions of the selected candidates. The most significant candidate, discussed in section 7.1.2, is highlighted in yellow. Observation time (7.3d) and $F_\mu$ (7.3e) distributions of the selected candidates. The values of the most significant candidate, discussed in section 7.1.2, are highlighted in yellow.*

expected for photon candidates. This is reflected in the zenith distribution (figure 7.3c), which shows that the selected candidates appear preferably at larger zenith angles. This is also due to the fact, discussed in section 4.3, that, to guarantee a reliable $X_{\mathrm{max}}$ reconstruction, it is required that the depth of the shower maximum is observed in the FD field of view. Therefore, vertical events with a deep $X_{\mathrm{max}}$ cannot be observed.

As for the $F_\mu$ distribution shown in figure 7.3e, one can note that the candidate events exhibit a small $F_\mu$ compared to the median of the protons, which is also expected for photon candidates. However, it has also to be noted that all values of $F_\mu$ sit in the tail of the proton distribution, i.e., overall the $F_\mu$ distribution of the selected events is consistent with that of background, reported in figure 6.1b.

The arrival times of the candidates have also been examined: figure 7.3d shows the distribution of the observation times. It appears flat over time, i.e., no particularly significant clustering can be identified.

The general characteristics of the candidate events are summarised in table 7.1, where the $E_\gamma$, $X_{\max}$, $F_\mu$, the GPS time and the Fisher discriminant value are reported for each candidate. The reconstructed longitudinal profile and the footprint on the array for each of them are presented in the appendix C.

| $\log_{10}(E_\gamma/[\text{eV}])$ | $X_{\max}$ [g cm$^{-2}$] | $F_\mu$ | $\theta$ [°] | GPS | Fisher |
|---|---|---|---|---|---|
| 18.31 | 987.7 | 0.42 | 38.7 | 834996450 | 1.57 |
| 18.01 | 1039.9 | 0.39 | 47.6 | 835412500 | 2.12 |
| 18.24 | 1245.2 | 0.75 | 56.7 | 863837908 | 2.87 |
| 18.02 | 907.6 | 0.22 | 43.6 | 870750320 | 1.46 |
| 18.00 | 913.4 | 0.29 | 47.8 | 881735354 | 1.40 |
| 18.10 | 938.9 | 0.11 | 39.0 | 922084511 | 1.84 |
| 18.29 | 1008.7 | 0.52 | 47.8 | 939970475 | 1.57 |
| 18.01 | 1010.4 | 0.59 | 59.3 | 940132288 | 1.58 |
| 18.21 | 796.3 | -0.23 | 22.7 | 947908557 | 1.36 |
| 18.01 | 1019.9 | 0.52 | 49.6 | 970117635 | 1.75 |
| 18.14 | 984.7 | 0.45 | 47.3 | 971249641 | 1.57 |
| 18.17 | 935.6 | 0.07 | 30.8 | 993100676 | 1.86 |
| 18.02 | 1109.3 | 1.01 | 57.2 | 993881848 | 1.57 |
| 18.20 | 944.3 | 0.20 | 54.6 | 996371961 | 1.68 |
| 18.08 | 932.7 | 0.02 | 44.2 | 1008567108 | 1.96 |
| 18.04 | 967.5 | 0.48 | 35.0 | 1036824689 | 1.45 |
| 18.04 | 1061.8 | 0.86 | 41.7 | 1056592884 | 1.47 |
| 18.48 | 1001.9 | 0.45 | 51.8 | 1110436364 | 1.55 |
| 18.04 | 954.3 | 0.29 | 54.5 | 1141435435 | 1.67 |
| 18.12 | 917.0 | 0.07 | 48.1 | 1151733711 | 1.74 |
| 18.07 | 847.4 | 0.01 | 58.5 | 1154937152 | 1.38 |
| 18.05 | 849.9 | -0.07 | 42.4 | 1181870094 | 1.54 |

**Table 7.1:** *List of the generalities of the selected candidates.*

Finally, it is of interest to have an overview of the arrival directions of the candidate events. Using the known position and orientation of the Earth at the time of arrival of the photon candidates and the reconstructed event geometry, the arrival direction of the events have been determined in Galactic coordinates, which are oriented such that the primary direction is aligned with the direction of the center of the Milky Way and the fundamental plane is in the Galactic plane.

The resulting sky map is shown in figure 7.4: the black dots indicate the directions, in Galactic coordinates, of the candidate events (the most significant one is in yellow), while the graded blue area represents the directional exposure to photons of the hybrid system (discussed in details in section 7.2.1). The apparent alignment of 8 events within 15° of the Galactic plane, has stimulated a further study of the distribution of the arrival directions. This study has been carried out by

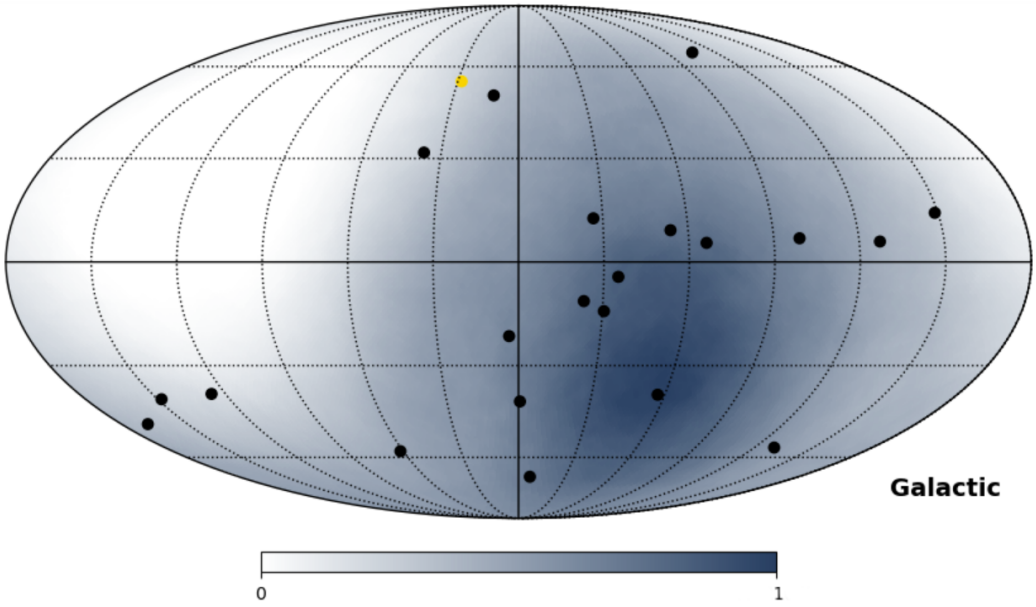



**Figure 7.4:** *Sky map in Galactic coordinates of the selected events (black dots). The blue shaded area shows the hybrid directional exposure to photons. The most significant event is highlighted in yellow.*

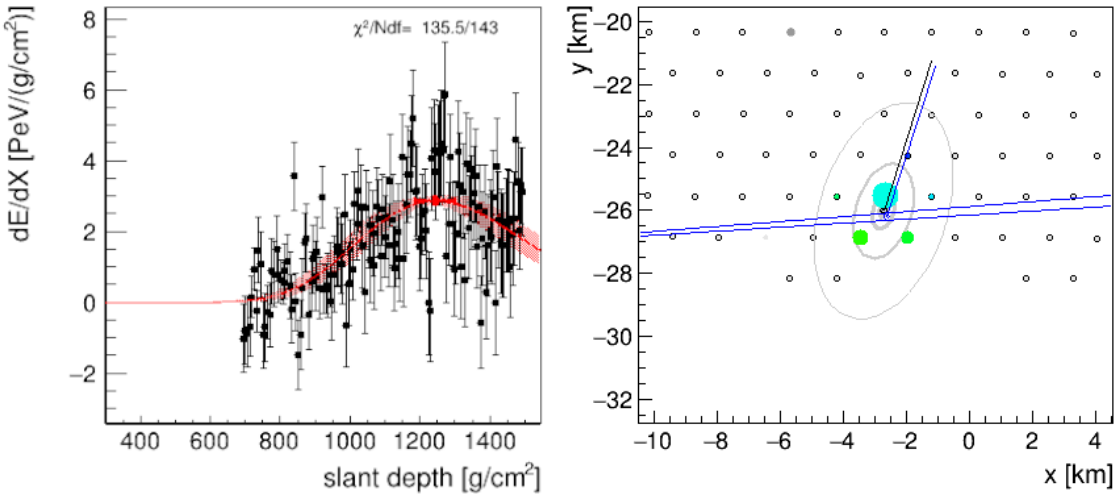
simulating 10000 realisations of the candidate distribution using the hybrid photon exposure as a model, so to assess the probability to observe such clustering. The probability obtained is $\sim 3\,\%$, from which it follows that the observed distribution is not indicative of a clustering around the Galactic plane. The same exercise has been performed at different angular distances from the Galactic plane, finding similar results[1].

### 7.1.2 The most significant candidate

As one can see from figure 7.3b, the candidate event with the Fisher value most deviating from the expectations has the peculiarity of having a very deep $X_{\mathrm{max}}$. This fact motivated a further inspection, although the significance of the event in itself is moderate (1 event observed against 0.1 expected).

The event, labeled with an ID $=$ 3478968 was detected on May, 22nd 2007 at 02:58:14 UTC. Figure 7.5a shows its reconstructed profile: the hybrid reconstruction yields an energy $E_\gamma = (1.73 \pm 0.16) \times 10^{18}\,\mathrm{eV}$, a depth of the shower maximum $X_{\mathrm{max}} = (1245 \pm 57)\,\mathrm{g\,cm^{-2}}$ and a zenith angle $\theta = (56.7 \pm 1.0)^\circ$. The footprint of the event on the SD array is in turn shown in figure 7.5b, where one can see that the shower triggered 6 stations. Out of them, only 3 stations pass the selection criteria, described in section 5.2.2, for the $F_\mu$ calculation. The station with the

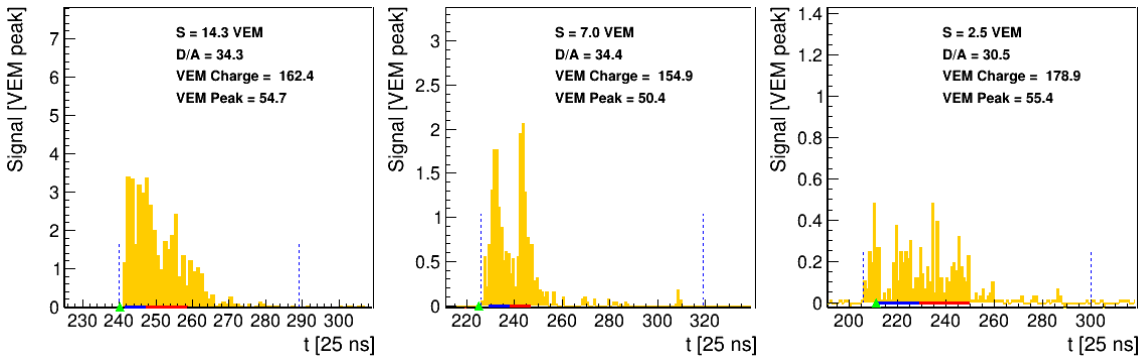
[1] The exhaustive search for directional excesses of photons, associated or not with known astrophysical objects and structures, goes beyond the scope of this thesis work. It will be part of further future studies.




(a) (b)

**Figure 7.5:** *Longitudinal profile (7.5a) of the most significant candidate. The red line indicates the reconstructed profile, while the red shaded area represents the uncertainty. The red point shows the reconstructed $X_{max}$. (7.5b): Event footprint on the SD array. Each colored circle indicate a triggered station. The colors (from blue to green) represent the trigger times while the dimensions are proportional to the signal size.*

largest signal size is, in fact, rejected because too close to the core, while other two stations are rejected because of the small signal size. The signals in the stations considered are shown in figure 7.6. The $F_\mu$ associated to this event is $0.75 \pm 0.41$. By combining it with the value of $X_{\max}$, the resulting value of the Fisher discriminant is $f \simeq 2.87$.




(a) (b) (c)

**Figure 7.6:** *Reconstructed signals in the stations used in the most significant candidate for the $F_\mu$ calculation. The vertical dashed lines show the start and the stop time of the signals, while the blue and the red horizontal segments show, respectively the risetime and the falltime of the traces.*

As already noted, this event is characterized by a very deep $X_{\max}$. Actually, the reconstructed depth of the shower maximum is deep even when compared to the expected $X_{\max}$ for photons-initiated showers, as one can see from figures 7.7a and 7.7b, where the $X_{\max}$ of the candidate (dashed vertical line) is compared with the simulated $X_{\max}$ distributions for protons and photons, respectively. The latter have been obtained by simulating $\sim 1\,000\,000$ showers initiated by protons and photons with the same energy and geometry of the candidate, using the CONEX [181] software.

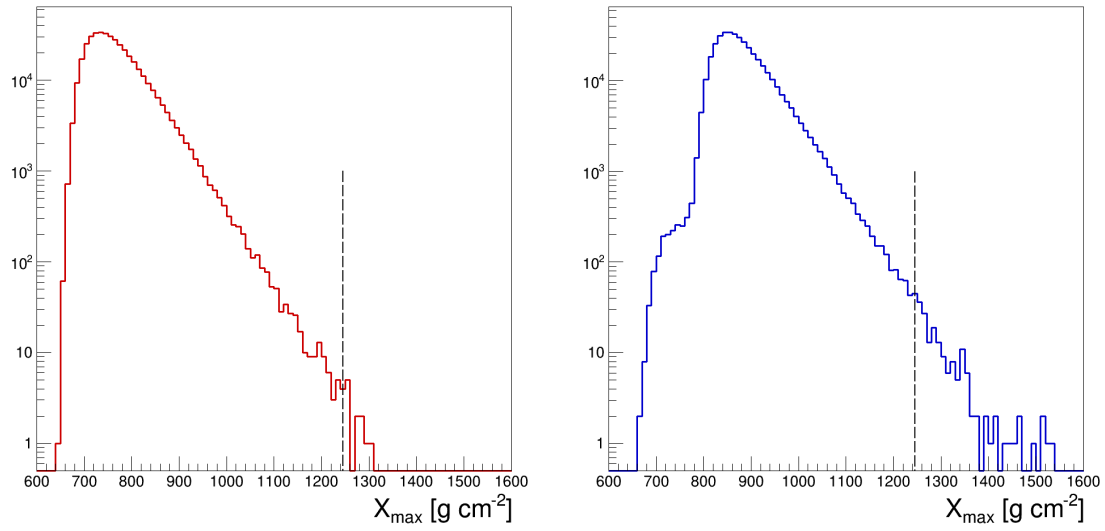



**Figure 7.7:** *$X_{max}$ distributions for protons (7.7a) and photons (7.7b). The vertical dashed line shows the $X_{max}$ reconstructed in the peculiar candidate. The events are simulated using CONEX [181], for primaries with the same energy and geometry of the candidate.*

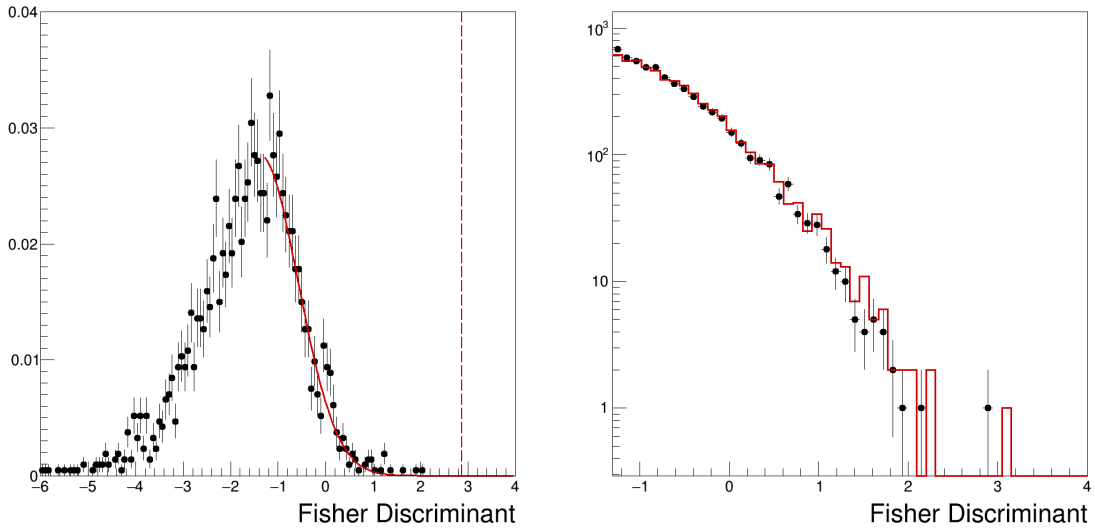
Hence, several cross-checks have been carried out, to study this peculiar event in details. First, we found out that this event, acquired in 2007, was not selected as a candidate in previous Auger photon searches with hybrid data [114, 180]. The main motivation for this resides in the fact that since the time of publication, the profile reconstruction has improved, as discussed in section 2.2.3, and the analysis strategy has evolved.

As a second cross-check, the profile was reconstructed with alternative reconstruction methods, such as the Gaisser-Hillas fit or the monocular profile constrained geometry fit [183][2]. With this test we verified that all reconstruction methods give consistent values of $X_{\max}$.

Thirdly, the atmospheric conditions at the time of the event were checked and found to be optimal, with a measured VAOD of 0.02 and no cloud coverage.

[2] The PCGF uses an additional condition on the shape of reconstructed longitudinal profile to be in the form of Gaisser-Hillas function. A scan through possible geometries inside a shower–detector plane is performed and the most likely geometry, based on the pixel timing information as well as on the additional condition on profile description, is selected.

Finally, the event was also cross-checked by using SD-based information. Namely, the risetime of the signals in the triggered stations (see section 5.1) has been analysed, resulting to be consistent with an event developing late in the atmosphere.




**(a)** **(b)**

**Figure 7.8:** *(7.8a): Fisher discriminant distribution in proton simulations characterized by the same geometry and energy of the most significant candidate. The vertical dashed red line corresponds to the value of the Fisher discriminant of the candidate. (7.8a): Example of one realization (red) of the data sample, out of the* 100 000*, simulated to account for the look-elsewhere effect. The realizations are obtained using the background extrapolation as model. For comparison the data distribution is reported in black.*

To quantify the probability that the selected candidate is actually a background event the same strategy as described in section 6.3 is adopted. 2000 proton showers, characterized by the same geometrical configuration and energy of the candidate, are simulated. The simulations are generated with CORSIKA by using EPOS-LHC as high energy hadronic model and following the same approach described in section 4.2.2. The resulting Fisher discriminant distribution is shown with black points in figure 7.8a. The dashed red vertical line represents the value of the Fisher discriminant of the candidate event. The solid red curve superimposed to the distribution is the fit to the background parametrization, from which we derive a *local* significance associated to the Fisher value of the candidate, above $3\sigma$.

Since the photon signal is identified as an excess with respect to the expected background, in a position of the region considered not known a priori, the significance must account for the fact that spurious signals with seemingly high significance can be found in other positions. This is known as *look-elsewhere* effect [184]. Therefore, we derived the probability to have an event with a Fisher discriminant $f > 2.87$, by simulating 100 000 realizations of the data sample using the extrapolated background as a model: $\sim 25\,\%$ of the realizations met our criteria, thus a claim for a

photon observation is not possible. In figure 7.8b it is shown an example of a particular simulated realization (red) with a Fisher above 2.87, compared to the data distribution (black).

## 7.2 UPPER LIMITS ON THE DIFFUSE PHOTON FLUX

Since, as discussed above, no significant excess of photons has been observed with respect to the background, upper limits on the diffuse UHE-photon flux, $\Phi_\gamma^{\text{C.L.}}$, are derived in this section as:

$$\Phi_\gamma^{\text{C.L.}}(E_\gamma > E_\gamma^0) = \frac{N_\gamma^{\text{C.L.}}(E_\gamma > E_\gamma^0)}{\mathscr{E}_\gamma(E_\gamma > E_\gamma^0)} \tag{7.1}$$

where $N_\gamma^{\text{C.L.}}$ is the upper limit on the number of photons at a certain confidence level above a energy threshold $E_\gamma^0$, and $\mathscr{E}_\gamma$ is the hybrid exposure to photons above the same $E_\gamma^0$.

The calculation of the exposure, that for the hybrid detector depends on the energy, is described in section 7.2.1. This will allow for the determination of the upper limits, derived in section 7.2.2.

### 7.2.1 The hybrid photon exposure

The exposure is defined as the time-integrated acceptance of the experiment. For the hybrid detector, the photon exposure is calculated as [182]:

$$\mathscr{E}_\gamma = \int_T \int_\Omega \int_S \varepsilon_\gamma(E_\gamma, t, \theta, \phi, x, y) \cos(\theta) dS d\Omega dt \tag{7.2}$$

where $\varepsilon_\gamma$ is the overall photon efficiency, including detection, reconstruction, and selection of events. $\varepsilon_\gamma$ is a function of the zenith, $\theta$, and azimuth, $\phi$, of the impact position, $x, y$, of the time, $t$, and of the energy, $E_\gamma$.

The configurations of the FD and the SD have been continuously changing over the period of data used in this thesis work. In the first period, until 2008, the SD and the FD were under construction, so that both the number of SD stations and of the FD telescopes gradually increased with the time. Furthermore, even in the final configuration, some SD stations may be temporarily out of service at any time. Similarly, during nightly operations, individual telescopes may be sometimes deactivated because of sky or weather conditions, or of hardware failures. Therefore, to properly take into account the actual varying detector configurations, the exposure of the hybrid detector is calculated using the full Monte Carlo simulations, described in section 4.2.2, that reproduce the exact conditions of the observatory, i.e., the actual sequence of the configurations of SD and FD.

Given a set of $N$ simulated events in an energy bin $E_\gamma^j$ and in a zenith bin $\theta_i$, and generated on an area $S_{\text{gen}}$ within the time interval $T$, the *raw exposure* is calculated numerically as

$$\mathscr{E}_\gamma(E_\gamma^j) = 2\pi S_{\text{gen}} T \sum_{i,j} \frac{n(E_\gamma^j, \cos\theta_i)}{N(E_\gamma^j, \cos\theta_i)} \cos\theta_i \Delta\cos\theta_i \tag{7.3}$$

where $n$ denotes the number of events that fulfill the selection criteria presented in section 4.3, the station selection cuts described in section 5.2.2, and the photon selection discussed in section 6.2.2.

As the upper limits on the flux are calculated above fixed energy thresholds, to calculate the total exposure above such thresholds the raw exposure $\mathscr{E}_\gamma$ is weighted with a power law spectrum assuming a spectral index $\Gamma = 2$, namely:

$$\mathscr{E}_\gamma^{\text{weighted}}(E_\gamma > E_\gamma^0) = \frac{\int_{E_\gamma^0}^{+\infty} E_\gamma^{-\Gamma}\, \mathscr{E}_\gamma(E'_\gamma) dE_\gamma}{\int_{E_\gamma^0}^{+\infty} E_\gamma^{-\Gamma} dE_\gamma} \tag{7.4}$$

Since the calculated raw exposure is not a continuous function, a cubic spline is used to interpolate between the data points, the integration is performed by using a Gauss-Legendre quadrature. The resulting behavior of $\mathscr{E}_\gamma^{\text{weighted}}$ as a function of the energy threshold is shown as a solid line in figure 7.9. The grey shaded band represents the systematic uncertainty of ±6.4 % [182] due to the on-time and trigger efficiency. The systematic uncertainty associated to the choice of $\Gamma = 2$ is discussed in the next section.

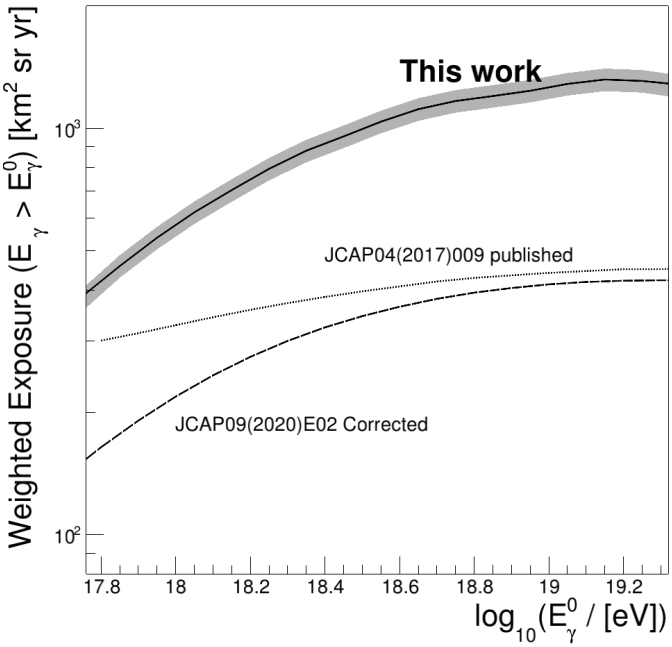



**Figure 7.9:** *Weighted (7.9) hybrid exposure for primary photons (solid line) in the time interval 1 January 2005 - 31 December 2017, assuming a power-law spectrum with $\Gamma = 2$. In figure 7.9 the systematic uncertainty due to the on-time and the trigger efficiency are shown as a gray band. The dotted (dashed) line represents the published [114] (corrected [185]) weighted exposure. See text for details.*

The obtained exposure is compared with the exposure relative to the data set used in a previous Auger search for UHE photons [114] (dashed line): the current exposure results to be $\sim 2.5$ larger, mostly due to the use of 4 additional years of data.

The figure shows also a third dotted line, which corresponds to the exposure actually published in [114]. A by-product of the calculation performed in this thesis work has in fact been the identification of an overestimation of the published weighted exposure. In particular, as one can see, the *corrected* weighted exposure decreases much more quickly than the published one for smaller threshold energies. At a threshold energy of $10^{18}$ eV, the published exposure is overestimated by a factor of 1.5. Towards higher threshold energies, the difference between the corrected and published exposure decreases, being only about 7 % larger at a threshold energy of $10^{19}$ eV. The corrected exposure, and hence the corrected upper limits, have been the subject of an *erratum corrige* publication [185].

### 7.2.2 Upper limits calculation

The calculation of upper limits is carried out through equation 7.1 for 5 different energy thresholds, $E_\gamma^0$, the same as in [114]. They are listed in the first column of table 7.2.

| $E_\gamma^0$ [EeV] | $N_b(E_\gamma > E_\gamma^0)$ | $N_\gamma(E_\gamma > E_\gamma^0)$ | $N_\gamma^{95\,\%}(E_\gamma > E_\gamma^0)$ | $\mathscr{E}_\gamma^{\text{weighted}}(E_\gamma > E_\gamma^0)$ [km$^2$ sr yr] | $\Phi_\gamma^{95\,\%}(E_\gamma > E_\gamma^0)$ [km$^{-2}$ sr$^{-1}$ yr$^{-1}$] |
|---|---|---|---|---|---|
| 1.0 | $30 \pm 15$ | 22 | 23.38 | 579 | 0.0403 |
| 2.0 | $6 \pm 6$ | 2 | 9.53 | 840 | 0.0113 |
| 3.0 | $0.7 \pm 1.9$ | 0 | 3.42 | 976 | 0.0035 |
| 5.0 | $0.06 \pm 0.25$ | 0 | 2.59 | 1141 | 0.0023 |
| 10.0 | $0.02 \pm 0.06$ | 0 | 2.62 | 1263 | 0.0021 |

**Table 7.2:** *Upper limits on the diffuse flux of UHE photons (last column). The different energy thresholds are listed in the first column. The expected number of background events, of photon candidates and of its 95 % upper limit are shown in columns 2 to 4. The exposure is given in the 5th column.*

$N_\gamma^{95\,\%}$ (shown in the fourth column of the table) is computed at each energy threshold as the Rolke upper limit [179] at 95 % C.L. of the observed number of photon candidates (listed in the third column), taking into account the expected number of background events. Rolke's method is chosen because it allows to include the uncertainties on the estimation on $N_b$, which in our case are not negligible. The expected number of background events, reported in the second column, is calculated from the parameterisation of the background, given in section 6.2.1, after normalizing it to the number of hybrid events above each $E_\gamma^0$. The fifth column of table 7.2 shows the exposure.

Various sources of systematic uncertainties in the calculated upper limits have been investigated. The main source of uncertainties are the systematic uncertainties of the reconstructed shower parameters, namely energy $X_{\text{max}}$ and $F_\mu$, that enter the Fisher discriminant analysis. The impact of the systematic uncertainty of the energy scale, 14 % (see section 2.2.4), is evaluated by

shifting all energy values upward or downward. The number of candidates changes by $^{+5}_{-8}$ in the first bin, and by $^{+3}_{-1}$ in the second bin. Consequently the value of the correspondent upper limits change, respectively, by the 25 % and the 10 %. For $X_{\mathrm{max}}$, the systematic uncertainty is less than $10\,\mathrm{g\,cm^{-2}}$ at all energies. To evaluate the impact on the upper limits, the reconstructed $X_{\mathrm{max}}$ values of the data sample are changed by $\pm\Delta X_{\mathrm{max}} = 10\,\mathrm{g\,cm^{-2}}$: the number of photon candidates is found to change by $^{+6}_{-2}$ in the first energy interval, and leaves unaffected the limits at larger energy thresholds. Consequently, the upper limits change by about the 15 %. As for $F_\mu$, the most notable systematic effect is the observed drift over time, shown in section 5.4. As discussed there, this is totally correlated to the drift in the energy scale, which is due to the ageing of the two detectors. As such systematic uncertainty is already accounted for in the systematics of the energy scale, no additional systematic uncertainty associated to $F_\mu$ is included in that of the upper limits.

Another source of uncertainty is the unknown photon spectral index, which reflects into a change in the exposure. Differences of 15 % and 20 % are found in the first two energy intervals when changing the spectral index from 2 to 1.5 and 2.5, respectively. The lack of knowledge of $\Gamma$ may also have an impact on the analysis, because a different spectral index changes the shape of the distributions used as input variables for the MVA method. However, in the case of the Fisher Discriminant Analysis, the impact of the shape change has been verified to be negligible compared to the exposure effect.

Finally, the photon/proton discrimination may be impacted by the choice of the hadronic model, as different models have different predictions for $X_{\mathrm{max}}$ and the number of muons in showers generated by hadronic primaries. The uncertainties in the modelling of proton- and nucleon-induced air showers, i.e., the background events for the analysis presented in this thesis, may have an impact on the training of the Fisher discriminant analysis. However, the simulations used in this work are produced with EPOS-LHC as hadronic interaction model. The air showers predicted from hadronic primaries with this model are characterized by the largest muon component and the deepest $X_{\mathrm{max}}$ with respect to other models. This implies, as also shown in [114], that the upper limits produced with this model are more conservative. A new shower-library with a second hadronic interaction model will be built to verify this fact, prior to the publication of this thesis work.

## 7.3 PHYSICS IMPLICATIONS

In this final section, the derived upper limits are discussed in the context of other experimental results and of different astrophysical scenarios.

They are drawn, with black symbols, in figure 7.10, where, for comparison, the limits obtained with other data sets of Auger (blue and dark green), and with Telescope Array data (light green), are also shown. As one can see, the limits obtained in this thesis work lower by a factor of about 3.0 those attained with the previous hybrid photon search [114], shown with the blue symbols. It is of interest to note that even at the highest energies covered with the hybrid data, i.e., at $10^{19}$ eV,

the upper limits obtained from the hybrid search are more stringent then those obtained with the surface detectors, both of Telescope Array (light green symbols) and of Auger (dark green symbols). This is particularly remarkable in comparison with the latter, obtained with an exposure which is 40 times larger than that of the hybrid system. This is due to the intrinsically better power of separation between photon and hadron showers of fluorescence detectors, that can observe the full development of the cascade, differently from surface arrays.

The upper limits derived with this thesis work, and, more in general, with Auger data, show that the Auger observatory is actually the most sensitive gamma-rays instrument operating in the energy range above $10^{18}$ eV. Before discussing the limits in the context of specific models of gamma-ray production at such energies, one can get an intuition of the Auger sensitivity by comparing them with the (optimistic) extrapolation of fluxes more familiar in gamma-ray astronomy. On the one hand, the dashed grey line in the figure represents the extrapolation of the "guaranteed" diffuse flux from the Galactic plane, due to the interaction of cosmic rays with the gas in the disc. This is expected to be at the level of about $6.4 \times 10^{-5}$ of the cosmic rays flux [186]. Note that the highest energy observation of such flux is that from the Tibet AS-gamma experiment, in the energy range between 100 TeV and 1 PeV. On the other hand, the yellow band indicates the extrapolation of the flux observed by the HAWC experiment from the source J1825-134 [187], found to be coincident with a giant molecular cloud. Although the source is not visible at the latitude of Auger, it is taken as an exemplary one as its energy spectrum extends well beyond 200 TeV without a break or a cutoff.

Moving to specific models for the production of ultra-high energy gamma-rays, figure 7.10 shows the flux predictions due to different production mechanisms. As discussed in chapter 1, conventional production processes are those that take place either at the level of the acceleration of ultra-high energy cosmic rays, or during their propagation. Less conventional processes are those that predict fluxes of photons in non-acceleration mechanisms, such as the decay or annihilation of exotic particles.

During their propagation, ultra-high energy cosmic rays interact with low-energy universal photons (e.g., those forming the cosmic microwave background or the infrared extragalactic background), and produce a flux of secondary UHE photons. Although such *cosmogenic* photons interact themselves with the soft photons of the universal backgrounds, they can propagate for several Mpc without being absorbed. The expected photon fluxes depend on the cosmic-ray composition, on their maximal acceleration energy at the sources, on the emissivity distribution and on the cosmological evolution of the production sites. Two expectations for cosmogenic photons [190] are shown in the figure with the two shaded red and blue areas, for a pure proton and a pure iron composition, respectively. In the two models a strong evolution of the sources is assumed, which is only partially constrained by the limits on the neutrino flux above 10 PeV [190]. The energy spectra at the sources are modeled as power-laws with a spectral index $\Gamma = 2$ and an exponential cutoff at the maximal energy $Z \cdot E_{\mathrm{max}}$, with $E_{\mathrm{max}} = 10^{21}$ eV. As one can see, the two predictions differ by approximately one order of magnitude, with the largest expected being that from pure protons. The uncertainty bands of these models are due to the assumption on the spatial distri-

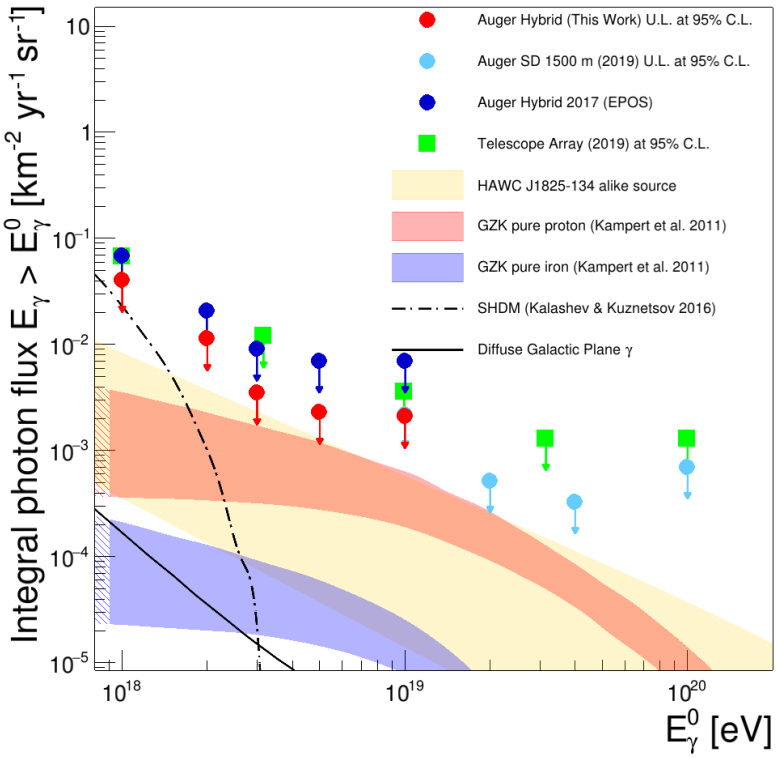



**Figure 7.10:** *Photon flux limits at* 95 % *C.L. obtained with this thesis work (black symbols). Red and dark green symbols represent the limits attained by other Auger analysis [114,188]. The light green symbols show the limits derived from Telescope Array data [189]. Predictions of UHE photon fluxes [190] are indicated as colored bands, for comparison. See text for details.*

bution of sources, on the photon background fields in the infrared and radio regime and on the intergalactic magnetic fields. The derived upper limits are approaching the region of the photon fluxes predicted in the pure proton composition scenario. A much larger increase of the exposure is required to probe the pure iron scenario.

Other predictions of UHE photons arise from non-acceleration models. The detection of a flux of UHE photons could thus be a smoking gun for dark matter (DM) composed of super-heavy particles. Models of super-heavy DM particles, first put forward in the 90s [191–195], were recently revived as an alternative to the weakly-interactive massive particles. On the theoretical aspect, the various motives for DM particles to be related to the Planck scale or the GUT scale can can be detected at the Observatory. Of particular interest would be the detection of UHE photons from regions of denser DM density such as the center of our Galaxy. Although SHDM particles do not decay in a standard way because they are protected in the perturbative domain by the conservation of quantum numbers, they can disintegrate through non perturbative effects [196]. Photons can then be produced through QCD decay channels or leptonic channels either via direct

gamma production or by means of interactions of products (i.e. electrons) with cosmic microwave background (CMB) and galactic media. From the absence of photons with energies above 1 EeV, constraints on the properties of DM particles in terms of mass $M_X$ and lifetime $\tau_X$ can be inferred, as illustrated in figure 7.11 where the 95 % C.L. allowed regions of the mass and lifetime of the particles are shown (the grey area is excluded). The strongest constrain over the whole mass range is $\tau_X > 3 \times 10^{22}$ yr at $M_X \approx 10^{20}$ eV. From these constrains, couples of $M_X$ and $\tau_X$ values can be selected and the corresponding flux of photons can be obtained. Such a flux is illustrated as the black line in figure 7.11. As one can see, it is only marginally compatible with the limits derived in this work and severely constrained by the limits from the surface detector data.

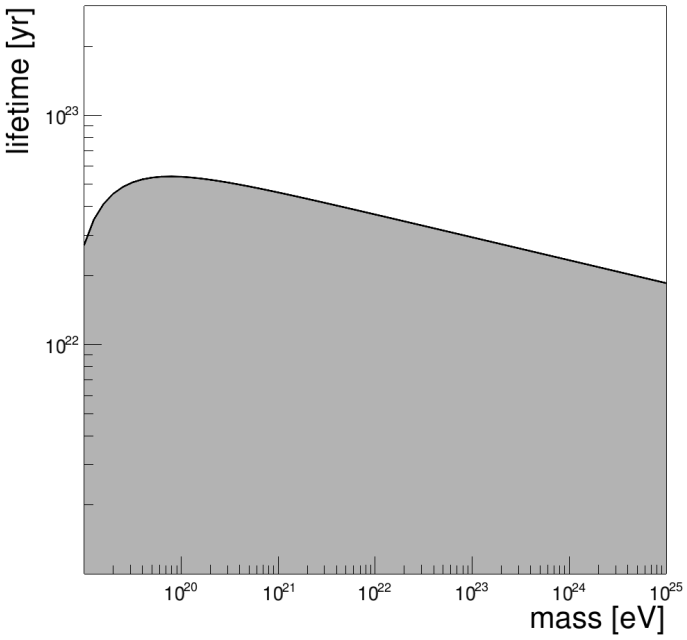



**Figure 7.11:** *Constraints on mass $M_X$ and lifetime $\tau_X$ of super heavy dark matter. Grey area is excluded. See text for details*

# 8 SUMMARY AND OUTLOOK

This thesis work was aimed at the search for ultra-high-energy (UHE) cosmic photons, in an energy range, above $10^{18}$ eV, which is well beyond that attainable with current human-made accelerators.

Although at such energies photons propagating in the cosmic vacuum may interact with those of the cosmic microwave background, thus limiting the explored horizon to a few Mpc, the implications of such a search remain relevant both for cosmic-ray physics and for fundamental one. Due to magnetic fields that permeate the universe, cosmic rays, which are mostly charged ions, do not point back to the sources, even at the highest energies. Therefore, the quest for the origin of UHE cosmic rays intrinsically implies a multi-messenger approach, since, on the one hand, direct information about their acceleration sites can be obtained by searching for the neutral particles, $\gamma$-rays and neutrinos, generated by the interactions of cosmic rays at the acceleration sites, via the so-called *astrophysical beam dump* process. On the other hand, UHE photons are expected to be produced by ultra-high energy cosmic rays in interaction with the soft photons of the universal backgrounds in their propagation to Earth, via the Greisen-Zatsepin-Kuzmin effect. These *cosmogenic photons* also probe ultra-high energy cosmic rays (CRs) as their flux depends on the characteristics of the sources, as well as on the nature of the parent nuclei. In addition, UHE photons might probe new physics, as their detection would be a smoking gun for dark matter composed of super-heavy particles decaying in photons.

Due to the steepness of both the cosmic ray and cosmic photon fluxes, this search can only be done through large ground-based detectors that exploits the phenomenon of extensive air-showers. The identification of photon primaries, when detectors of this kind are used, relies on the ability to distinguish the showers generated by photons from those initiated by the overwhelming background of nuclei. Since the radiation length in atmosphere is more than two orders of magnitude smaller than the mean free path for photo-nuclear interaction, in photon showers the transfer of energy to the hadron/muon channel is reduced with respect to the bulk of hadron-induced air showers, resulting in a lower number of secondary muons. Additionally, as the development of photon showers is delayed by the typically small multiplicity of electromagnetic interactions, they reach the maximum development of the shower, $X_{\mathrm{max}}$, deeper in the atmosphere with respect to showers initiated by hadrons.

These two observables can be measured at the Pierre Auger Observatory, that is the world's largest cosmic rays detector, which employs a *hybrid* detecting technique for the observation of extensive air showers, by combining a fluorescence detector (FD) with a ground array of particle detectors (SD). The FD provides a direct observation of the longitudinal shower profile, which allows for the measurement of the energy, $E$, and of the $X_{\mathrm{max}}$ of a shower, while the SD samples

the secondary particles at ground level. Although the SD observes showers at a fixed depth, the longitudinal development is embedded in the signals detected.

In this thesis work I used the so-called hybrid events, i.e., those detected simultaneously by the FD and the SD, and I conceived a new variable, $F_\mu$, measured with the SD and related to the muonic content of the shower, by exploiting the so-called air-shower universality property in combination with the reconstruction of hybrid events. To fully exploit the hybrid approach, I then combined $F_\mu$ with $X_{\max}$, yielding a photon/hadron separation power better than that of the two individual observables. Finally, I searched for photons by applying this analysis method to the hybrid events with energies above $10^{18}$ eV collected at the Pierre Auger Observatory between January 2005 and December 2017.

## 8.1 MUONIC CONTENT ESTIMATION FROM UNIVERSALITY

The general idea behind air-shower universality is that the energy spectrum of the secondary particles produced during the shower development, as well as their angular and lateral distributions, depends only on the energy of the primary and on the stage of the shower development. As a consequence, the distribution of secondary particles produced in the cascade can be described at different stages of the shower development, so that the distribution of secondary particles at the ground can be predicted.

By exploiting this property, a universality-based model has been developed within the Pierre Auger Collaboration, that allows to predict the signals induced by the secondary particles in the SD stations. This model describes the signal size as the superposition of four components: muons ($S_\mu$); $e^\pm$ and $\gamma$ from high energy pions ($S_{e\gamma}$); $e^\pm$ and $\gamma$ from muon decays ($S_{e\gamma(\mu)}$); $e^\pm$ and $\gamma$ due to low energy hadrons ($S_{e\gamma(\mathrm{had})}$). Each i-th signal component, $S^i_{\mathrm{comp}}$, has a universal behavior depending only on the primary energy $E$, on $X_{\max}$, and on the geometrical configuration of the shower. The relative contributions of each of the four components, $f^i_{F_\mu\,\mathrm{fluct}}$, instead, depend on the mass of the primary particle, through a parameter representing the number of muons in the shower, $F_\mu$. The predicted signal, $S_{\mathrm{pred}}$, can then be expressed as:

$$S_{\mathrm{pred}} = \sum_{i=1}^{4} f^i_{F_\mu\,\mathrm{fluct}}(F_\mu) \cdot S^i_{\mathrm{comp}} \tag{8.1}$$

where $i$ runs over the four components, while $S^i_{\mathrm{comp}}$, in turn, is the contribution of each component and has been parametrised using QGSJetII-03 proton simulations.

In this thesis work I exploited this model in the case of hybrid events. As their reconstruction provides $E$, $X_{\max}$, and the shower geometry, $S^i_{\mathrm{comp}}$ can be directly calculated for each station involved in a hybrid event. Given the reconstructed signal in a station of the SD, $S^i_{\mathrm{rec}}$, $F_\mu$ can be calculated for each station in each event by fixing $S_{\mathrm{rec}} = S_{\mathrm{pred}}$ in equation 9.1. By using simulations, I studied the performance of this procedure, as well as the photon/hadron separation power of $F_\mu$.

To these aims, I produced a set of simulated hybrid events following the two-steps process used in the Pierre Auger Collaboration. First, I simulated extensive air showers developping within the atmosphere and then I input the generated showers into the simulation of the detectors response. The simulated showers were produced using CORSIKA, a software that handles simulation of particle cascades in the atmosphere, adopting EPOS-LHC as model for the description of the hadronic interactions. The showers were generated in the energy and zenith-angle ranges of interest for hybrid events, i.e., between $10^{17.5}$ eV and $10^{19.5}$ eV, with angles $0^\circ$ and $65^\circ$. I then used the Auger Offline software framework to simulate the responses of the FD and SD detectors to the CORSIKA showers, as well as their reconstruction. The detectors simulation were performed accounting for the real configuration of the observatory, that is, considering the actual status and the time evolution of the detectors. I then applied to the events the selection criteria in use in the Collaboration to ensure an accurate reconstruction of the geometry and of the longitudinal profile. Only showers with reconstructed zenith $\theta < 60^\circ$ and with energies above $10^{18}$ eV were considered.

Using these simulations, I first determined the selection criteria for the signals in the SD stations. By studying the difference between the reconstructed and predicted $F_\mu$, I found that biases appeared in two ranges of signals, namely below 6 VEM[1], due to trigger effects, and above 800 VEM because of saturation effects. Stations could thus only be used if the signal size were in the range between 6 VEM and 800 VEM. Moreover, I also studied the performance of the $F_\mu$ calculation as a function of the distance of the SD stations from the shower axis, finding that the estimation in stations at a distance less that 600 m was largely affected by the core reconstruction resolution, due to the steepness of the lateral distribution of the particles close to the shower axis. I used the simulations also to evaluate the uncertainty in the estimation of $F_\mu$, which is due, on the one hand, to that of the reconstruction of $S_{\text{rec}}$ and, on the other hand, to that of the reconstructed hybrid parameters. The derived uncertainty served to calculate $F_\mu$ when more then one station was selected in an event, i.e., with an average weighted by the uncertainty.

Finally, the simulations allowed for assessing the photon-hadron separation power of $F_\mu$. After finding that the photon and proton $F_\mu$ distributions were well separated, I quantified the separation power by studying the background rejection, i.e., the fraction of events in the proton-shower distribution rejected by a given cut value on $F_\mu$, as a function of the signal efficiency, i.e., the fraction of events in the photon-shower distribution that pass the same cut. As reference value for the separation power, I took the value of the background rejection at a signal efficiency of 50 % (i.e., the cut value corresponding to the median of the photon distribution): the obtained background rejection at the photon median was found to be 99.5 %.

[1] The signals reconstructed in the SD stations are measured in units of the signal produced by a vertical muon traversing the detector (VEM).

## 8.2 A NEW ANALYSIS TECHNIQUE FOR THE UHE-PHOTON SEARCH

Having demonstrated that $F_\mu$, in hybrid events, provides a very good photon-hadron separation even when derived from the signal of one SD station only, to fully exploit the hybrid approach I combined it, through a multi-variate analysis (MVA) technique, with $X_{\max}$ so to further improve the photon-hadron separation power. I first studied the two observables independently, finding that $F_\mu$ is almost independent from the primary energy for both proton and photons primaries. In turn, $X_{\max}$ resulted to be linearly increasing with the logarithm of the energy, faster for photons than for protons, implying that its separation power increases with energy. I also found that the two observables were uncorrelated. I thus chose to combine $F_\mu$, $X_{\max}$ and energy linearly into a Fisher discriminant analysis, that yields the Fisher discriminant, $f$. Fisher discriminants have indeed the best separation performances in the case of uncorrelated, or linearly correlated, input observables following a gaussian distribution, while no discrimination at all is achieved when a variable has the same sample mean for signal and background, even if the shapes of the distributions are very different. The advantages of the Fisher discriminant, $f$, are that it provides a robust event classification for uncorrelated input observables, which is the case for $F_\mu$ and $X_{\max}$, and that it can be calculated analytically for each event.

To verify the performances of the combination, I used again the simulations. The Fisher discriminant distributions obtained for the proton and photon simulations turned out to be well separated, resulting into a background rejection of $\simeq 99.90\,\%$ for a signal efficiency of 50 %. Overall, and as expected, the performance of the two combined observables in terms of photon-hadron discrimination resulted to be better than for each of them separately.

In turn, to study the expected background I chose a data-driven approach, that is, I used hybrid data. As the ultimate search for photons would have been developed following the prescription of a blind analysis, I exploited only a sub-sample of the data, so-called *burnt sample*, corresponding to 5% of the full data set. I first compared the Fisher distribution of the burnt sample with that of simulated photons, and found them well separated, so that the events contained in the burnt sample could be safely considered as background events. Due to the limited number of events in the burnt sample, in a preliminary step I used proton simulations to derive the functional form of the background. By considering only the rightmost tail of the Fisher distribution, specifically only the events with a Fisher discriminant $f_0 > -1.3$, I found that the background distribution could be described by

$$m(f|A,B) = N(A,B)e^{-(Af^2+Bf)} \tag{8.2}$$

with A and B as shape parameters, and $N$ as the normalization of the background calculated as a function of the parameters $A$ and $B$, namely:

$$N(A,B) = \frac{N\sqrt{A}}{e^{B^2/4A}\mathrm{erfc}\left(\frac{B}{2\sqrt{A}}\left(\frac{2A}{B}f_0 - 1\right)\right)} \tag{8.3}$$

I then fitted this model to the events of the burnt sample, so to obtain a description of the background free from either assumptions on cosmic ray composition, or on hadronic interaction mod-

els. This background estimation was used to derive the photon selection cut, $f_\gamma$. The determination of $f_\gamma$ was achieved by studying a proxy for the signal-to-background ratio. The signal was determined by the selection efficiency of photons at different thresholds of the Fisher discriminant, while the background was calculated from the extrapolation of the estimation to the whole data set. The signal to background ratio was found to show a maximum at $f \simeq 1.6$. Since this value approximately corresponded to the median of the Fisher discriminant distribution for photons, $f \simeq 1.36$, the value of such median was finally selected as $f_\gamma$.

## 8.3 RESULTS AND POSSIBLE FUTURE PERSPECTIVES

I finally applied the analysis developed to the whole hybrid data set, that consisted of about 32000 events. Five different energy thresholds of 1, 2, 3, 5, and 10 EeV were considered. The number of air-shower events with photon-like characteristics found at each energy threshold were, respectively, 22, 2, 0, 0, 0. These numbers were found to be in agreement with the expected background events from a pure hadronic background determined for each energy threshold, respectively: $30 \pm 15$, $6 \pm 6$, $0.7 \pm 1.9$, $0.06 \pm 0.25$ and $0.02 \pm 0.06$. Since no significant excess of a photon signal with respect to the background was identified, I calculated upper limits to the differential flux of UHE photons, that resulted to be, respectively, 0.0403, 0.0113, 0.0035, 0.0049, 0.0021 in units of $\mathrm{km}^{-2}$ $\mathrm{sr}^{-1}$ $\mathrm{yr}^{-1}$.

The derived upper limits were first discussed in the context of other experimental results. Compared to previous analyses, they resulted to be the most stringent ones, in particular improving those attained in previous searches by the Auger collaboration by a factor of $\sim 3$. The improvement between 1 and 2 EeV was mainly due to the fact that the background extrapolation obtained from the burnt sample has been accounted in the upper limits calculation. Above 3 EeV, instead, since the background expectation is compatible with 0, the improvement of the calculated upper limits was mainly related to an increase of the exposure.

The attained upper limits were then discussed in the context of expectations from different production processes for UHE photons. During their propagation, ultra-high energy cosmic rays interact with low-energy universal soft photons (i.e., those forming the cosmic microwave background or the infrared extragalactic background), and produce a flux of secondary UHE photons. Although such cosmogenic photons interact themselves with the soft photons of the universal backgrounds, they can propagate for several Mpc without being absorbed. The expected photon fluxes depend on the cosmic-ray composition. The derived upper limits were found to be approaching the region of the photon fluxes predicted in a pure proton composition scenario. A much larger increase of the exposure would be required instead to probe the pure iron scenario.

A comparison of the achieved upper limits was also performed with expectations from processes that predict fluxes of photons in non-acceleration mechanisms, such as the decay or annihilation of exotic particles, most notably the dark matter (DM) composed of super-heavy particles. The latter were recently revived as an alternative to weakly-interactive massive particles.

Although they do not decay in a standard way, because they are protected in the perturbative domain by the conservation of quantum numbers, they can disintegrate through non perturbative effects, yielding a photon flux that is in the reach of the Auger observatory. From the absence of photons with energies above 1 EeV, I could set constraints on the properties of DM particles in terms of mass $M_X$ and lifetime $\tau_X$.

The results obtained in this thesis work will lead to a publication by the Pierre Auger Collaboration. Before publication, it requires a more in-depth study of the systematic uncertainties related to the hadronic interaction model used. The uncertainties in the modelling of proton- and nucleon-induced air showers, i.e., the background events for the analysis presented in this thesis, may have in fact an impact on the training of the Fisher discriminant analysis. The simulations used in this work were produced with EPOS-LHC as hadronic interaction model. The air showers predicted from hadronic primaries with this model are characterized by the largest muon component and the deepest $X_{\mathrm{max}}$ with respect to other models. This implies that the upper limits produced with this model are expected to be the most conservative ones, but a cross-check by repeating the analysis chain with another model will certainly be of interest.

Future directional, all-sky and targeted, searches for EeV photon are also anticipated by using the analysis technique developed in this thesis work. Sources producing particle fluxes according to an $E^{-2}$ energy spectrum inject equal energy into each decade. A measured energy flux of $1\,\mathrm{eV\,cm^{-2}\,s^{-1}}$ in the TeV decade would result in the same energy flux in the EeV decade if one assumes that the spectrum continues to such high energies and that sources are also present in the nearby universe, so that energy losses en route to Earth are negligible. Directional searches are also motivated by the study of the dark matter: of particular interest would be the detection of UHE photons from regions of denser DM density such as the center of our Galaxy.

The developed analysis has also the potential to be extended down to $10^{17.5}$ eV, profiting from the low energy enhancements of the Pierre Auger Observatory, namely the High Elevation Altitude Telescopes, i.e., three additional fluorescence telescopes with an elevated field of view, overlooking a denser SD array, in which the stations are separated by 750 m . The combination of these two instruments allow for the measurement of showers in the energy range from below the so-called second knee of the cosmic-ray spectrum up to the ankle. As a preliminary step for the application of the technique developed in this work at lower energies, a validation of the universality-based model at low energies is mandatory.

Finally, the technique developed in this work can profit vastly from the planned upgrades of the detector systems. The key element of the upgrade is the installation of plastic scintillator detectors (SSD) on top of the SD stations, which have different sensitivity to muons and electromagnetic particles that reach the ground. Therefore, a better separation between the shower components can intrinsically be achieved with the combined measurement, setting an ideal ground for the use of the universality concept and thus leading to a significant boost in the resolution of $F_\mu$.

# 9 RÉSUMÉ ET PERSPECTIVES

Ce travail de thèse avait pour but de rechercher des photons cosmiques d'ultra-haute énergie (UHE), dans une gamme d'énergie supérieure à $10^{18}$ eV qui se situe bien au-delà de ce qui peut être atteint avec les accélérateurs construits par l'homme actuellement.

Bien qu'à de telles énergies les photons se propageant dans le vide cosmique puissent interagir avec ceux du fond diffus cosmologique, limitant ainsi leur horizon exploré à quelques Mpc, les implications d'une telle recherche restent pertinentes tant pour la physique des rayons cosmiques que pour la physique fondamentale. En raison des champs magnétiques qui imprègnent l'univers, les rayons cosmiques, qui sont pour la plupart des ions chargés, ne pointent pas vers leurs sources, même aux plus hautes énergies. Par conséquent, la recherche de l'origine des rayons cosmiques d'ultra-haute énergie implique intrinsèquement une approche multi-messagers, puisque, d'une part, des informations directes sur les sites d'accélération peuvent être obtenues en recherchant les particules neutres, les rayons gamma et les neutrinos, générés par les interactions des rayons cosmiques via le processus dit *astrophysical beam dump*. D'autre part, on s'attend à ce que les photons UHE soient produits par les rayons cosmiques de très haute énergie interagissant avec les photons des fonds universels lors de leur propagation vers la Terre, via l'effet Greisen-Zatsepin-Kuzmin. Ces photons cosmogéniques sondent également les rayons cosmiques d'ultra-haute énergie, car leur flux dépend des caractéristiques des sources ainsi que de la nature des noyaux parents. En outre, les photons UHE pourraient permettre de sonder une nouvelle physique, car leur détection en provenance du centre galactique serait une observation décisive en faveur de l'existence d'une matière noire composée de particules super lourdes se désintégrant en photons.

En raison de la petitesse des flux de rayons cosmiques et de photons cosmiques, cette recherche ne peut se faire qu'à l'aide de grands détecteurs au sol qui exploitent le phénomène des grandes gerbes atmosphériques. L'identification des photons primaires, lorsque des détecteurs de ce type sont utilisés, repose sur la capacité à distinguer les gerbes générées par les photons de celles initiées par le fond écrasant des noyaux. Étant donné que la longueur de rayonnement dans l'atmosphère est inférieure de plus de deux ordres de grandeur au libre parcours moyen pour l'interaction photonucléaire, dans les gerbes à photons, le transfert d'énergie vers le canal hadron/muon est réduit par rapport à l'ensemble de celles induites par les hadrons, ce qui entraîne un nombre plus faible de muons secondaires. De plus, étant donné que le développement des gerbes de photons est retardé par la petite multiplicité d'interactions électromagnétiques, celles-ci atteignent le développement maximal, $X_{\mathrm{max}}$, plus profondément dans l'atmosphère par rapport à celles initiées par des hadrons.

Ces deux observables peuvent être mesurées à l'observatoire Pierre Auger, qui est le plus grand détecteur de rayons cosmiques au monde et qui utilise une technique *hybride* pour l'observation des gerbes atmosphériques, en combinant un détecteur de fluorescence (FD) avec un réseau au

sol de détecteurs de particules (SD). Le FD fournit une observation directe du profil longitudinal de la gerbe, ce qui permet d'en mesurer l'énergie, $E$, et $X_{\max}$, tandis que le SD échantillonne les particules secondaires au niveau du sol. Bien que le SD observe les gerbes à une profondeur fixe, le développement longitudinal est intégré dans les signaux détectés.

Dans ce travail de thèse, j'ai utilisé les événements dits hybrides, c'est-à-dire ceux qui sont détectés simultanément par le FD et le SD, et j'ai conçu une nouvelle variable, $F_\mu$, mesurée avec le SD et liée au contenu muonique de la gerbe, en exploitant la propriété dite d'universalité des gerbes, en combinaison avec la reconstruction d'événements hybrides. Pour exploiter pleinement l'approche hybride, j'ai ensuite combiné $F_\mu$ avec $X_{\max}$, ce qui donne un pouvoir de séparation photon/hadron meilleur que celui des deux observables individuelles. Enfin, j'ai recherché des photons en appliquant cette méthode d'analyse aux événements hybrides d'énergie supérieure à $10^{18}$ eV collectés à l'observatoire Pierre Auger entre janvier 2005 et décembre 2017.

## 9.1 ESTIMATION DU CONTENU MUONIQUE À PARTIR DE L'UNIVERSALITÉ

L'idée générale qui sous-tend l'universalité des gerbes atmosphériques est que le spectre d'énergie des particules secondaires, ainsi que leurs distributions angulaires et latérales, ne dépendent que de l'énergie du rayon cosmique primaire et du stade de développement de la gerbe. En conséquence, la distribution des particules secondaires produites dans la cascade peut être décrite à différents stades du développement, de sorte que la distribution des particules secondaires au sol puisse être prédite.

En exploitant cette propriété, un modèle basé sur l'universalité a été développé au sein de la Collaboration Pierre Auger, qui permet de prédire les signaux induits par les particules secondaires dans les stations du SD. Ce modèle décrit la taille du signal comme la superposition de quatre composantes : muons ($S_\mu$) ; $e^\pm$ et $\gamma$ provenant de pions de haute énergie ($S_{e\gamma}$) ; $e^\pm$ et $\gamma$, provenant de la désintégration des muons ($S_{e\gamma(\mu)}$) ; $e^\pm$ et $\gamma$ dus aux hadrons de basse énergie ($S_{e\gamma(\mathrm{had})}$). Chaque $i$-ème composante du signal, $S^i_{\mathrm{comp}}$, a un comportement universel qui ne dépend que de l'énergie primaire $E$, de $X_{\max}$, et de la configuration géométrique de la gerbe. Les contributions relatives de chacune des quatre composantes, $f^i_{F_\mu\,\mathrm{fluct}}$, au contraire, dépendent de la masse de la particule primaire, à travers un paramètre représentant le nombre de muons dans la gerbe, $F_\mu$. Le signal prédit, $S_{\mathrm{pred}}$, peut alors être exprimé comme :

$$S_{\mathrm{pred}} = \sum_{i=1}^{4} f^i_{F_\mu\,\mathrm{fluct}}(F_\mu) \cdot S^i_{\mathrm{comp}}, \tag{9.1}$$

où $i$ s'étend sur les quatre composantes, tandis que $S^i_{\mathrm{comp}}$ est la contribution de chaque composante et a été paramétrée à l'aide de simulations de protons basées sur le modèle d'interactions QGSJetII-03.

Dans ce travail de thèse, j'ai exploité ce modèle dans le cas d'événements hybrides. Comme leur reconstruction fournit $E$, $X_{\max}$, et la géométrie de la gerbe, $S^i_{\mathrm{comp}}$ peut être directement calculé

pour chaque station impliquée dans un événement hybride. Étant donné le signal reconstruit dans une station du SD, $S^i_{\text{rec}}$, $F_\mu$ peut être calculé pour chaque station dans chaque événement en fixant $S_{\text{rec}} = S_{\text{pred}}$ dans l'équation 9.1. En utilisant des simulations, j'ai étudié les performances de cette procédure, ainsi que le pouvoir de séparation photon/hadron de $F_\mu$. Pour atteindre ces objectifs, j'ai produit un ensemble d'événements hybrides simulés en suivant le processus en deux étapes utilisé par la collaboration Pierre Auger. Tout d'abord, j'ai simulé le développement dans l'atmosphère de gerbes engendrées par des photons et des protons, puis je les ai introduites dans la simulation de la réponse des détecteurs. Les gerbes simulées ont été produites à l'aide de CORSIKA, un logiciel qui gère la simulation de cascades de particules dans l'atmosphère, en adoptant EPOS-LHC comme modèle pour la description des interactions hadroniques. Les gerbes ont été générées dans les intervalles d'énergie et d'angle zénithal d'intérêt pour les événements hybrides, c'est-à-dire entre $10^{17.5}$ eV et $10^{19.5}$ eV, avec les angles entre 0° et 65°. J'ai ensuite simulé les réponses des FD et SD aux gerbes CORSIKA, ainsi que leur reconstruction, avec le logiciel développé par la Collaboration dans ce but. La simulation des détecteurs a été réalisée en tenant compte de la configuration réelle de l'observatoire, c'est-à-dire en considérant l'état réel et l'évolution temporelle des détecteurs. J'ai ensuite appliqué aux événements les critères de sélection utilisés par la Collaboration pour assurer une reconstruction précise de la géométrie et du profil longitudinal. Seules les gerbes avec un zénith reconstruit $\theta < 60°$ et avec des énergies supérieures à $10^{18}$ eV ont été considérées.

En utilisant ces simulations, j'ai d'abord déterminé les critères de sélection des signaux dans les stations du SD. En étudiant la différence entre le $F_\mu$ reconstruit et celui prédit, j'ai constaté que des biais apparaissaient dans deux gammes de signaux, à savoir en deça de 6 VEM[1], en raison des effets de déclenchement, et au-delà de 800 VEM en raison des effets de saturation. Les stations ne pouvaient donc être utilisées que si la taille du signal se situait entre 6 VEM et 800 VEM. J'ai également étudié la performance du calcul de $F_\mu$ en fonction de la distance des stations du SD par rapport à l'axe de la gerbe, en constatant que l'estimation dans les stations à une distance inférieure à 600 m était largement affectée par la résolution du coeur de la gerbe reconstruit, en raison de la très pentue distribution latérale des particules près de l'axe. J'ai également utilisé les simulations pour évaluer l'incertitude de l'estimation de $F_\mu$, qui est due, d'une part, à celle de la reconstruction de $S_{\text{rec}}$ et, d'autre part, à celle des paramètres hybrides reconstruits. L'incertitude dérivée a servi à calculer $F_\mu$ lorsque plus d'une station était sélectionnée dans un événement, en utilisant une moyenne pondérée par l'incertitude.

Enfin, les simulations ont permis d'évaluer le pouvoir de séparation photon-hadron de $F_\mu$. Après avoir constaté que les distributions de $F_\mu$ des photons et des protons étaient bien séparées, j'ai quantifié le pouvoir de séparation en étudiant le rejet du bruit de fond, c'est-à-dire la fraction d'événements dans la distribution des protons rejetés par une certaine valeur de coupure sur $F_\mu$, en fonction de l'efficacité du signal, c'est-à-dire, la fraction d'événements dans la distribution de la gerbe de photons qui passent la même coupure. Comme valeur de référence pour le pouvoir de séparation, j'ai adopté pour valeur du rejet du bruit de fond une efficacité du signal de 50 %

[1] Les signaux reconstruits dans les stations du SD sont mesurés en unité du signal produit par un muon vertical traversant le détecteur (VEM).

(c'est-à-dire la valeur de coupure correspondant à la médiane de la distribution des photons) : le rejet du bruit de fond obtenu pour la médiane des photons s'est révélé être de 99.5 pourcent.

## 9.2 UNE NOUVELLE TECHNIQUE D'ANALYSE POUR LA RECHERCHE DE PHOTONS UHE

Ayant démontré que $F_\mu$, dans les événements hybrides, fournit une très bonne séparation photon-hadron, même lorsqu'il est dérivé du signal d'une seule station du SD, pour exploiter pleinement l'approche hybride, je l'ai combinée, par le biais d'une technique d'analyse multivariée (MVA), avec $X_{\max}$ afin d'améliorer ultérieurement l'analyse photon-hadron.

J'ai d'abord étudié les deux observables indépendamment, et j'ai constaté que $F_\mu$ était presque indépendant de l'énergie primaire pour les protons et les photons. D'un autre côté, $X_{\max}$ s'est avéré augmenter linéairement avec le logarithme de l'énergie, de façon plus rapide pour les photons que pour les protons, ce qui implique que son pouvoir de séparation augmente avec l'énergie. J'ai également constaté que les deux observables n'étaient pas corrélées. J'ai donc choisi de combiner $F_\mu$, $X_{\max}$ et l'énergie de manière linéaire dans une analyse discriminante de Fisher, qui donne le discriminant de Fisher, $f$. Les discriminants de Fisher ont en effet les meilleures performances de séparation dans le cas d'observables non corrélées, ou corrélées linéairement, suivant une distribution gaussienne. Les avantages du discriminant de Fisher, $f$, sont une classification robuste des événements pour des observables non corrélées, ce qui est le cas pour $F_\mu$ et $X_{\max}$, et un calcul analytique pour chaque événement.

Pour vérifier les performances de la combinaison, j'ai utilisé à nouveau les simulations. Les distributions discriminantes de Fisher obtenues avec les simulations de protons et de photons se sont avérées bien séparées, ce qui a entraîné un rejet du bruit de fond de $\simeq 99.90\,\%$ pour une efficacité du signal de 50 %. Dans l'ensemble, et comme prévu, les performances des deux observables combinées se sont révélées excellentes en terme d'efficacité photon-hadron, meilleures que celles des deux observables séparément.

En revanche, pour étudier le fond attendu, j'ai choisi une approche axée sur les données, en utilisant les données hybrides. Comme la recherche ultime de photons a été développée suivant la prescription d'une analyse en aveugle, je n'ai exploité qu'un sous-échantillon de données, appelé *échantillon brûlé*, correspondant à 5% de l'ensemble des données. J'ai d'abord comparé les distributions de Fisher de l'échantillon brûlé et des photons simulés, et je les ai trouvées bien séparées, de sorte que les événements contenus dans l'échantillon brûlé pouvaient être considérés sans risque comme des événements de fond. En raison du nombre limité d'événements dans l'échantillon brûlé, dans une étape préliminaire j'ai utilisé des simulations de protons pour dériver la forme fonctionnelle du fond. En ne considérant que l'extrême queue de la distribution de Fisher, c'est-

à-dire uniquement les évènements avec un discriminant de Fisher $f_0 > -1.3$, j'ai trouvé que la distribution du fond pouvait être décrite par

$$m(f|A,B) = N(A,B)e^{-(Af^2+Bf)}, \tag{9.2}$$

avec $A$ et $B$ comme paramètres de forme, et $N$ comme normalisation du fond calculé en fonction des paramètres $A$ et $B$, à savoir :

$$N(A,B) = \frac{N\sqrt{A}}{e^{B^2/4A}\mathrm{erfc}\left(\frac{B}{2\sqrt{A}}\left(\frac{2A}{B}f_0 - 1\right)\right)}. \tag{9.3}$$

J'ai ensuite ajusté ce modèle aux événements de l'échantillon brûlé afin d'obtenir une description du fond libre de toute hypothèse sur les rayons cosmiques ou sur les modèles d'interaction hadronique. Cette estimation du bruit de fond a été utilisée pour dériver le seuil de sélection des photons, $f_\gamma$. La détermination de $f_\gamma$ a été réalisée en étudiant un *proxy* pour le rapport signal/bruit de fond. Le signal a été déterminé par l'efficacité de la sélection des photons pour différents seuils du discriminant de Fisher, tandis que le fond a été calculé à partir de l'extrapolation de l'estimation à l'ensemble des données. Le rapport signal/bruit de fond a montré un maximum à $f \simeq 1,6$. Puisque cette valeur correspondait approximativement à la médiane du discriminant de Fisher pour les photons, $f \simeq 1.36$, la valeur de cette médiane a finalement été choisie comme $f_\gamma$.

## 9.3 RÉSULTATS ET PERSPECTIVES POSSIBLES

J'ai finalement appliqué l'analyse développée à l'ensemble des données hybrides, qui comprenaient environ 32 000 événements. Cinq différents seuils d'énergie de 1, 2, 3, 5, et 10 EeV ont été considérés. Le nombre d'événements avec des caractéristiques de type photon trouvés pour chaque seuil d'énergie était, respectivement, 22, 2, 0, 0, 0. Ces nombres se sont avérés en accord avec les événements de fond attendus d'un fond hadronique pur déterminé pour chaque seuil d'énergie, respectivement : $30\pm15$, $6 \pm 6$, $0{,}7pm1{,}9$, $0.06pm0.25$ et $0.02\pm0.06$. Puisqu'aucun excès significatif d'un signal de photons par rapport au fond n'a été identifié, j'ai calculé des limites supérieures au flux différentiel de photons UHE, qui ont résulté être, respectivement, 0,0403, 0,0113, 0,0035, 0,0049, 0,0021 en unité $\mathrm{km}^{-2}\ \mathrm{sr}^{-1}\ \mathrm{yr}^{-1}$.

Les limites supérieures dérivées ont d'abord été discutées dans le contexte des résultats d'autres expériences. Par rapport aux analyses précédentes, elles se sont avérées être les plus strictes, améliorant en particulier celles obtenues lors des recherches précédentes de la collaboration Auger par un facteur de $\sim 3$. L'amélioration entre 1 et 2 EeV est principalement due au fait que l'extrapolation du bruit de fond obtenue à partir de l'échantillon brûlé a été prise en compte dans le calcul des limites supérieures. Au-delà de 3 EeV, au contraire, puisque l'extrapolation du bruit de fond est compatible avec 0, l'amélioration des limites supérieures calculées est principalement liée à une augmentation de l'exposition.

Les limites supérieures atteintes ont ensuite été discutées dans le contexte des prédictions de production de photons par de différents processus. Au cours de leur propagation, les rayons cosmiques à ultra-haute énergie interagissent avec des photons universels de faible énergie (c'est-à-dire ceux qui forment le fond cosmologique micro-ondes ou le fond extragalactique infrarouge), et produisent un flux de photons UHE secondaires. Bien que ces photons cosmogéniques interagissent eux-mêmes avec les photons des fonds universels, ils peuvent se propager sur plusieurs Mpc sans être absorbés. Les flux de photons attendus dépendent de la composition des rayons cosmiques. Les limites supérieures dérivées s'approchent de la région du flux attendu dans un scénario de composition purement protonique. Une augmentation beaucoup plus importante de l'exposition serait, au contraire, nécessaire pour sonder le scénario du fer pur.

Une comparaison des limites supérieures obtenues a également été effectuée avec les prévisions de flux de photons par des mécanismes de non-accélération, comme la désintégration ou l'annihilation de particules exotiques, notamment de matière noire (DM) composée de particules super lourdes. Ces dernières ont récemment été ravivées comme une alternative aux particules massives faiblement interactives. Bien qu'elles ne se désintègrent pas de manière standard, parce qu'elles sont protégées dans le domaine perturbatif par la conservation de nombre(s) quantique(s), elles peuvent se désintégrer via des effets non perturbatifs, donnant un flux de photons qui est éventuellement à la portée de l'observatoire Auger. A partir de l'absence de photons d'énergie supérieure à 1 EeV, j'ai pu déterminer des contraintes sur les propriétés des particules DM en termes de masse $M_X$ et de durée de vie $\tau_X$.

Les résultats obtenus dans ce travail de thèse donneront lieu à une publication par la Collaboration Pierre Auger. Avant cette publication, une étude plus approfondie des incertitudes systématiques liées au modèle d'interaction hadronique utilisé est nécessaire. Les incertitudes dans la modélisation des gerbes atmosphériques induites par des protons et des nucléons, c'est-à-dire, les événements de fond pour l'analyse présentée dans cette thèse, peuvent avoir un impact sur l'entraînement de l'analyse discriminante de Fisher. Les simulations utilisées dans ce travail ont été produites avec EPOS-LHC comme modèle d'interaction hadronique. Les gerbes prédites à partir des primaires hadroniques avec ce modèle sont caractérisées par la plus grande composante muonique et un $X_{\mathrm{max}}$ plus profond par rapport aux autres modèles. Cela implique que les limites supérieures produites avec ce modèle devraient être les plus conservatrices, mais une vérification croisée en répétant la chaîne d'analyse avec un autre modèle sera certainement intéressante.

La technique d'analyse développée dans ce travail de thèse permet également d'anticiper les futures recherches directionnelles, dans tout le ciel et ciblées, de photons EeV. Les sources produisant des flux de particules selon un spectre d'énergie $E^{-2}$ injectent une énergie égale dans chaque décade. Un flux d'énergie mesuré de $1\,\mathrm{eV cm^2/s}$ dans la décade TeV se traduirait par le même flux d'énergie dans la décade EeV, si l'on suppose que le spectre se poursuit jusqu'à des énergies aussi élevées et que des sources sont également présentes dans l'univers proche, de sorte que les pertes d'énergie en route vers la Terre sont négligeables. Les recherches directionnelles sont également motivées par l'étude de la matière noire : une recherche de photons UHE provenant du

centre de notre galaxie, où la densité de DM est attendue être la plus élevée, serait particulièrement intéressante.

L'analyse développée a également le potentiel d'être étendue jusqu'à $10^{17.5}$ eV, en profitant des améliorations à basse énergie de l'Observatoire Pierre Auger, à savoir les trois télescopes à fluorescence supplémentaires avec un champ de vue élevé, surplombant un réseau SD plus dense, dans lequel les stations sont séparées de 750 m . La combinaison de ces deux instruments permet de mesurer les gerbes dans la gamme d'énergie allant de ce qu'on appelle le deuxième genou du spectre des rayons cosmiques jusqu'à la cheville. Comme étape préliminaire à l'application de la technique développée dans ce travail à des énergies plus basses, une validation du modèle basé sur l'universalité à basse énergie est obligatoire.

Enfin, la technique développée dans ce travail peut profiter largement des mises à niveau prévues des systèmes de détection. L'élément clé est l'installation de scintillateurs plastique (SSD) au-dessus des stations SD, qui ont une sensibilité différente aux muons et aux particules électromagnétiques qui atteignent le sol. Par conséquent, une meilleure séparation entre les composants de la gerbe pourra intrinsèquement être obtenue avec la mesure combinée, créant un terrain idéal pour l'utilisation du concept d'universalité et conduisant ainsi à une augmentation significative de la résolution de $F_\mu$.

# A EXAMPLE OF A CORSIKA STEERING CARD

In this section, an example of a CORSIKA input file used for the air shower simulations that are discussed in this thesis is given. The main steering parameters are discussed in section 4. For a more detailed explanation of all steering parameters, see [197].

```
RUNNR 34984
NSHOW 1
EVTNR 1
PRMPAR 14
ESLOPE -1.0
ERANGE 2.115e+09 2.115e+09
SEED 16545944 0 0
SEED 16546257 0 0
SEED 16546645 0 0
THIN 1.000E-06  2.115e+04 5.000E+3
THINH 1.000E+00 1.000E+02
THETAP 57.578 57.578
PHIP 166.129 166.129
EPOS T 0
EPOSIG T
FIXCHI 0.0
ATMOD 22
OBSLEV 1.452E+05
MAGNET 1.94E+01 -1.41E+01
ECUTS 5.00E-02 1.00E-02 2.50E-04 2.50E-04
MUADDI T
MUMULT T
HADFLG 0 0 0 0 0 2
ELMFLG F T
STEPFC 1.0
RADNKG 5.0E+05
EPOPAR input ../epos/epos.param
EPOPAR fname inics ../epos/epos.inics
EPOPAR fname iniev ../epos/epos.iniev
EPOPAR fname initl ../epos/epos.initl
EPOPAR fname inirj ../epos/epos.inirj
EPOPAR fname inihy ../epos/epos.ini1b
EPOPAR fname check none
EPOPAR fname histo none
EPOPAR fname data  none
EPOPAR fname copy  none
LONGI T 1.0 T T
ECTMAP 2.5E+05
```

```
MAXPRT 1
DATBAS T
PAROUT T T
DIRECT /storage/gpfs_data/auger/psavina/RealMC/jobs/candidates/3478968/out/
USER psavina
HOST condor
DEBUG F 6 F 100000
EXIT
```

**Listing A.1:** *CORSIKA steering card for a proton shower*

# B SIMULATION MODULE SEQUENCE

In this section, the XML file used for steering the simulation and reconstruction of air shower events with Offline is listed. In the steering file, the sequence of modules which are used during the run time is specified. For a more detailed description of the modules and their implementation, see [198].

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

<!DOCTYPE sequenceFile [
  <!ENTITY % fd SYSTEM "/opt/exp_software/auger/ICRC2019/v3r99p1-install/share/auger-
offline/config/standardFdSequences.dtd">
  %fd;

  <!ENTITY % sd SYSTEM "/opt/exp_software/auger/ICRC2019/v3r99p1-install/share/auger-
offline/config/standardSdSequences.dtd">
  %sd;
] >

<sequenceFile
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:noNamespaceSchemaLocation='/opt/exp_software/auger/ICRC2019/v3r99p1-install/sha
re/auger-offline/config/ModuleSequence.xsd'>

  <enableTiming/>

  <moduleControl>

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

      <module> EventFileReaderOG </module>
      <module> MCShowerCheckerOG </module>

      <!-- increase numTimes if you want to throw the shower
           into the array more than once -->
      <loop numTimes="171" pushEventToStack="yes">

        <module> EventGeneratorOG </module>
        <try>

          &FdSimulation;

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

```
    <module> TabulatedTankSimulatorKG </module>

    &SdSimulation;

    <module> ClearParticleLists              </module>

    <module> T2LifeStationRemoverLG </module>

    <module> CentralTriggerSimulatorXb       </module>
    <module> CentralTriggerEventBuilderOG    </module>
    <module> EventBuilderOG                  </module>

    <module> EventCheckerOG                  </module>
    <module> SdCalibratorOG                  </module>
    <module> FdCalibratorOG                  </module>

    <try>
      <module> FdEyeMergerKG           </module>
      <module> FdPulseFinderOG         </module>
      <module> PixelSelectorOG         </module>
      <module> FdSDPFinderOG           </module>
      <module> FdAxisFinderOG          </module>
      <module> HybridGeometryFinderOG  </module>
      <module> HybridGeometryFinderWG  </module>
      <module> FdApertureLightKG       </module>
      <module> FdEnergyDepositFinderKG </module>
    </try>

    <try>
      <module> SdSignalRecoveryKLT             </module>
      <module> SdMonteCarloEventSelectorOG     </module>
      <module> SdEventSelectorOG               </module>

      <module> SdPlaneFitOG                    </module>
      <module> LDFFinderKG                     </module>

      <!--
      <try>
        <module> ScintillatorLDFFinderKG </module>
      </try>
      -->
      <module> SdEventPosteriorSelectorOG      </module>
      <module> Risetime1000LLL                 </module>
      <module> UniversalityFitter              </module>
      <!-- export in Offline format -->
    </try>
    <module> EventFileExporterOG             </module>
  </try> <!-- catch trigger Continues -->

  <!-- export the ADST -->
```

```
        <module> RecDataWriterNG                </module>

      </loop>
    </loop>
  </moduleControl>
</sequenceFile>
```

**Listing B.1:** *Module sequence used for the simulation of the detector response*

# C GENERALITIES OF THE SELECTED EVENTS

The generalities of the selected photon candidates, obtained in chapter 7, are listed event by event in the following.

**EVENT 2377631** This event, labeled with an ID = 2377631 was detected on June 22nd, 2006 at the 07:27:16 UTC. Figure C.1a shows its reconstructed profile: the hybrid reconstruction yields an energy $E_\gamma = 2.03 \times 10^{18}$ eV, a depth of the shower maximum $X_{\max} = 988\,\mathrm{g\,cm^{-2}}$ and a zenith angle $\theta = 39.0°$. The footprint of the event on the SD array is in turn shown in figure C.1b.

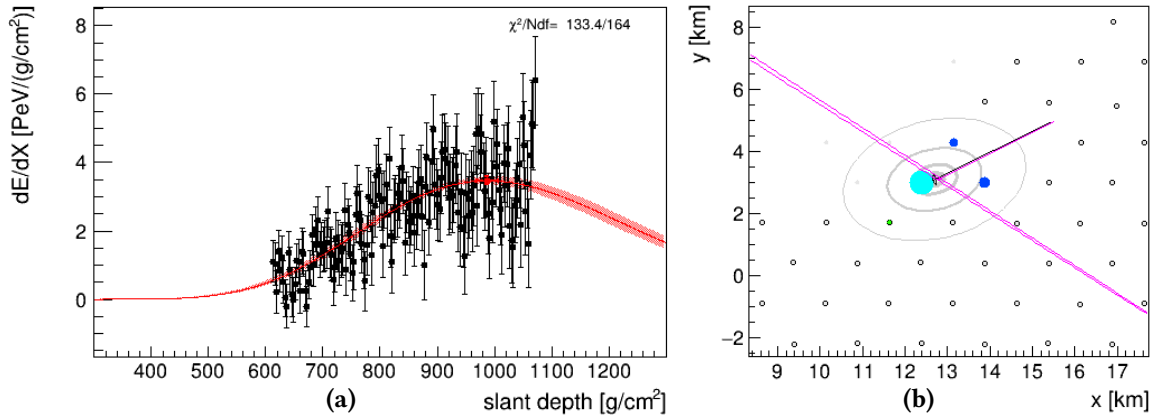



**Figure C.1:** *Longitudinal profile (C.1a) of the candidate 2377631. The red line indicates the reconstructed profile, while the red shaded area represents the uncertainty. The red point shows the reconstructed $X_{max}$. (C.1b) Event footprint on the SD array. Each colored circle indicate a triggered station. The colors (from blue to green) represent the trigger times while the dimensions are proportional to the signal size.*

**EVENT 2399211** This event, labeled with an ID = 2399211 was detected on June 27th, 2006 at the 03:01:26 UTC. Figure C.2a shows its reconstructed profile: the hybrid reconstruction yields an energy $E_\gamma = 1.02 \times 10^{18}$ eV, a depth of the shower maximum $X_{\max} = 1040\,\mathrm{g\,cm^{-2}}$ and a zenith angle $\theta = 47.6°$. The footprint of the event on the SD array is in turn shown in figure C.2b

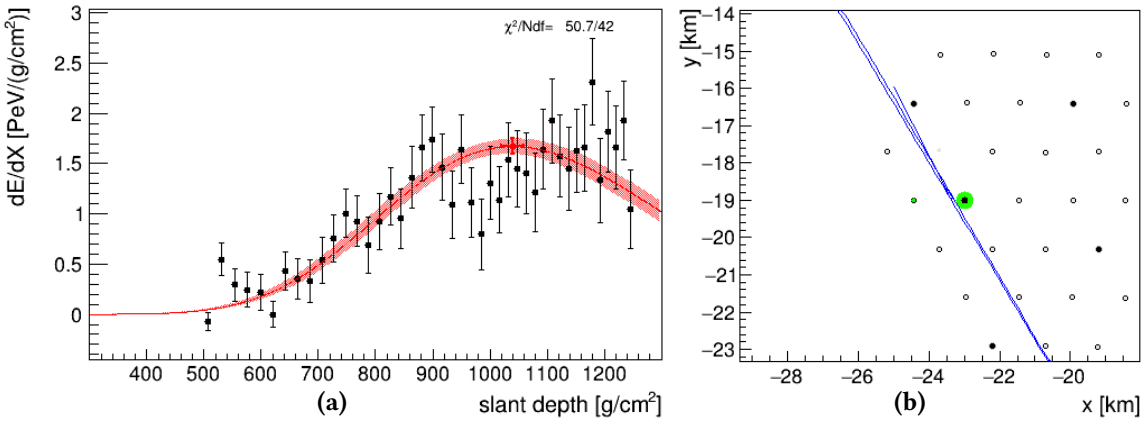



**Figure C.2:** *Longitudinal profile (C.2a) of the candidate 2399211. The red line indicates the reconstructed profile, while the red shaded area represents the uncertainty. The red point shows the reconstructed $X_{max}$. (C.2b). Event footprint on the SD array. Each colored circle indicate a triggered station. The colors (from blue to green) represent the trigger times while the dimensions are proportional to the signal size.*

**EVENT 3478968** This event, labeled with an ID = 3478968 was detected on 10th August, 2007 at the 03:05:06 UTC. Figure C.3a shows its reconstructed profile: the hybrid reconstruction yields an energy $E_\gamma = 1.70 \times 10^{18}\,\mathrm{eV}$, a depth of the shower maximum $X_{\mathrm{max}} = 1245\,\mathrm{g\,cm^{-2}}$ and a zenith angle $\theta = 56.7°$. The footprint of the event on the SD array is in turn shown in figure C.3b

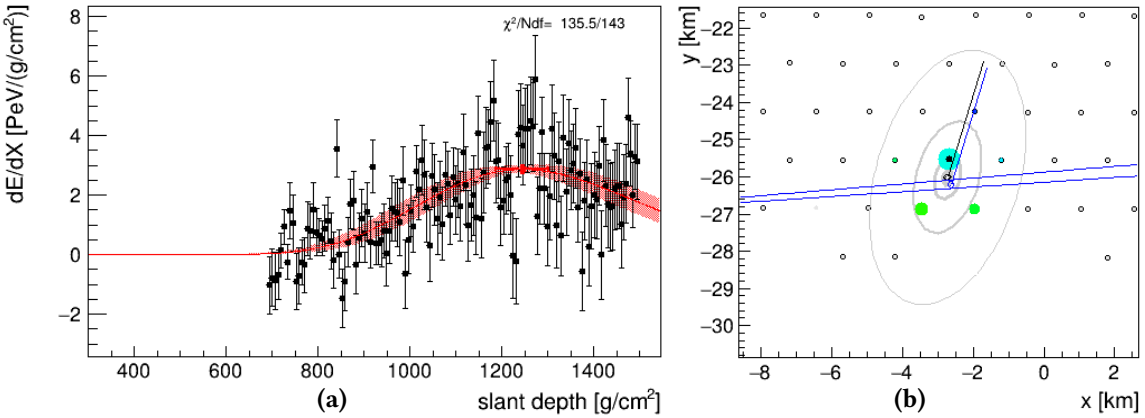



**Figure C.3:** *Longitudinal profile (C.3a) of the candidate 3478968. The red line indicates the reconstructed profile, while the red shaded area represents the uncertainty. The red point shows the reconstructed $X_{max}$. (C.3b). Event footprint on the SD array. Each colored circle indicate a triggered station. The colors (from blue to green) represent the trigger times while the dimensions are proportional to the signal size.*

**EVENT 3805261** This event, labeled with an ID = 3805261 was detected on December 15th, 2007 at the 06:29:00 UTC. Figure C.4a shows its reconstructed profile: the hybrid reconstruction

yields an energy $E_\gamma = 1.04 \times 10^{18}\,\text{eV}$, a depth of the shower maximum $X_{\text{max}} = 908\,\text{g}\,\text{cm}^{-2}$ and a zenith angle $\theta = 43.6°$. The footprint of the event on the SD array is in turn shown in figure C.4b

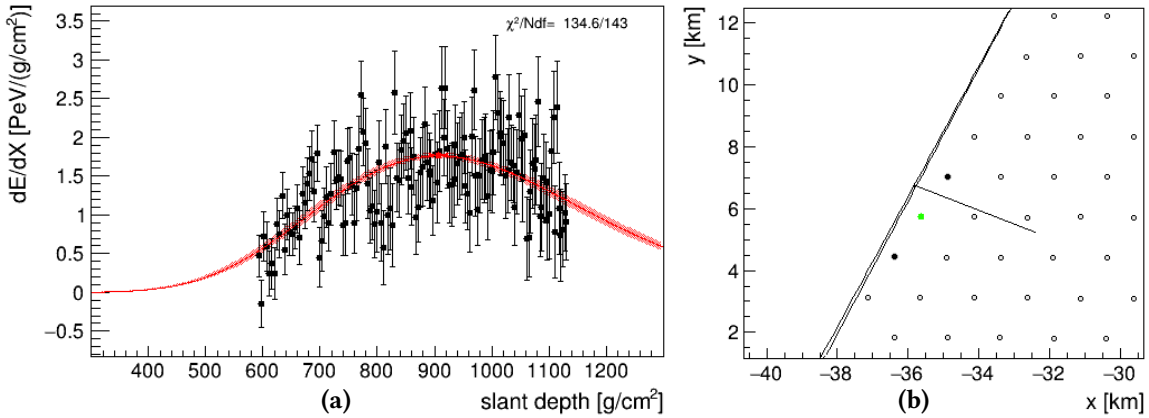



**Figure C.4:** *Longitudinal profile (C.4a) of the candidate 3805261. The red line indicates the reconstructed profile, while the red shaded area represents the uncertainty. The red point shows the reconstructed $X_{max}$. (C.4b). Event footprint on the SD array. Each colored circle indicate a triggered station. The colors (from blue to green) represent the trigger times while the dimensions are proportional to the signal size.*

**EVENT 4329935** This event, labeled with an ID = 4329935 was detected on March 26th, 2009 at the 06:34:56. Figure C.5a shows its reconstructed profile: the hybrid reconstruction yields an energy $E_\gamma = 0.99 \times 10^{18}\,\text{eV}$, a depth of the shower maximum $X_{\text{max}} = 913\,\text{g}\,\text{cm}^{-2}$ and a zenith angle $\theta = 47.8°$. The footprint of the event on the SD array is in turn shown in figure C.5b

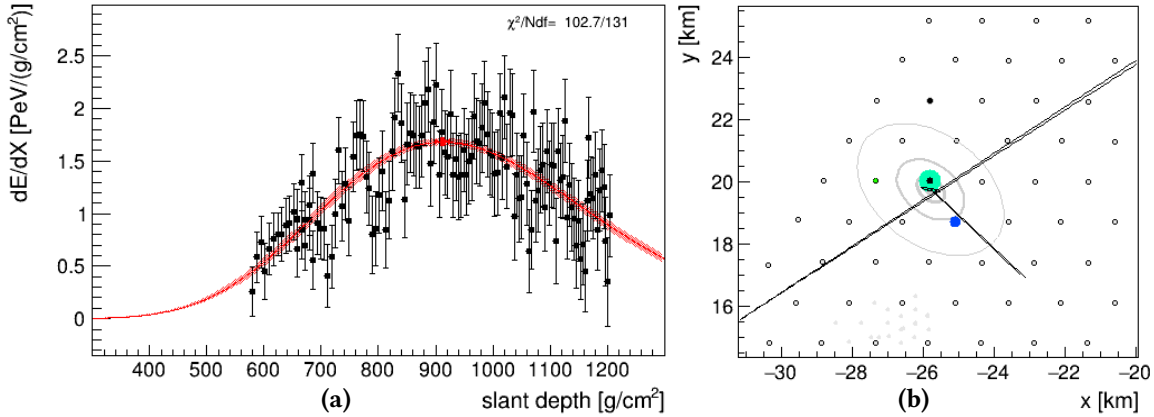



**Figure C.5:** *Longitudinal profile (C.5a) of the candidate 4329935. The red line indicates the reconstructed profile, while the red shaded area represents the uncertainty. The red point shows the reconstructed $X_{max}$. (C.5b). Event footprint on the SD array. Each colored circle indicate a triggered station. The colors (from blue to green) represent the trigger times while the dimensions are proportional to the signal size.*

**EVENT 7467206** This event, labeled with an ID = 7467206 was detected on the March 26th 2009 at the 06:34:56 UTC. Figure C.6a shows its reconstructed profile: the hybrid reconstruction yields an energy $E_\gamma = 1.25 \times 10^{18}$ eV, a depth of the shower maximum $X_{\max} = 939\,\mathrm{g\,cm^{-2}}$ and a zenith angle $\theta = 39.0°$. The footprint of the event on the SD array is in turn shown in figure C.6b

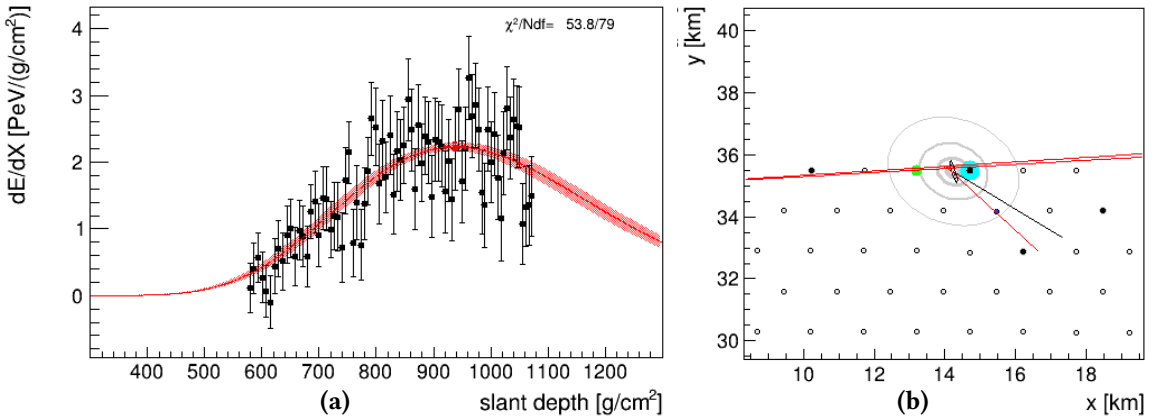



**Figure C.6:** *Longitudinal profile (C.6a) of the candidate 7467206. The red line indicates the reconstructed profile, while the red shaded area represents the uncertainty. The red point shows the reconstructed $X_{max}$. (C.6b). Event footprint on the SD array. Each colored circle indicate a triggered station. The colors (from blue to green) represent the trigger times while the dimensions are proportional to the signal size.*

**EVENT 8542419** This event, labeled with an ID = 8542419 was detected on October 19th, 2009 at the 06:54:20. Figure C.7a shows its reconstructed profile: the hybrid reconstruction yields an energy $E_\gamma = 1.93 \times 10^{18}$ eV, a depth of the shower maximum $X_{\max} = 1009\,\mathrm{g\,cm^{-2}}$ and a zenith angle $\theta = 47.8°$. The footprint of the event on the SD array is in turn shown in figure C.7b

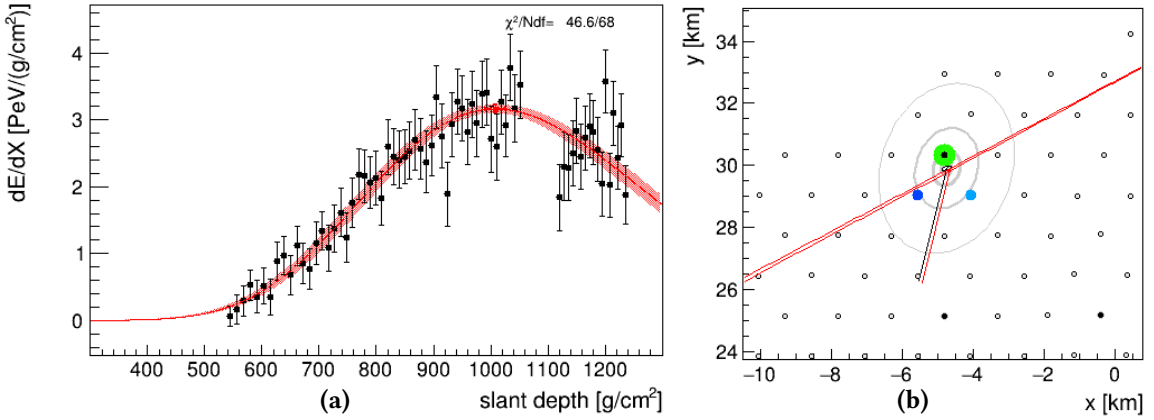



**Figure C.7:** *Longitudinal profile (C.7a) of the candidate 8542419. The red line indicates the reconstructed profile, while the red shaded area represents the uncertainty. The red point shows the reconstructed $X_{max}$. (C.7b). Event footprint on the SD array. Each colored circle indicate a triggered station. The colors (from blue to green) represent the trigger times while the dimensions are proportional to the signal size.*

**EVENT 8556435** This event, labeled with an ID = 8556435 was detected on October 21st, 2009 at the 03:51:13. Figure C.8a shows its reconstructed profile: the hybrid reconstruction yields an energy $E_\gamma = 1.00 \times 10^{18}$ eV, a depth of the shower maximum $X_{\max} = 1010\,\mathrm{g\,cm^{-2}}$ and a zenith angle $\theta = 59.3°$. The footprint of the event on the SD array is in turn shown in figure C.8b

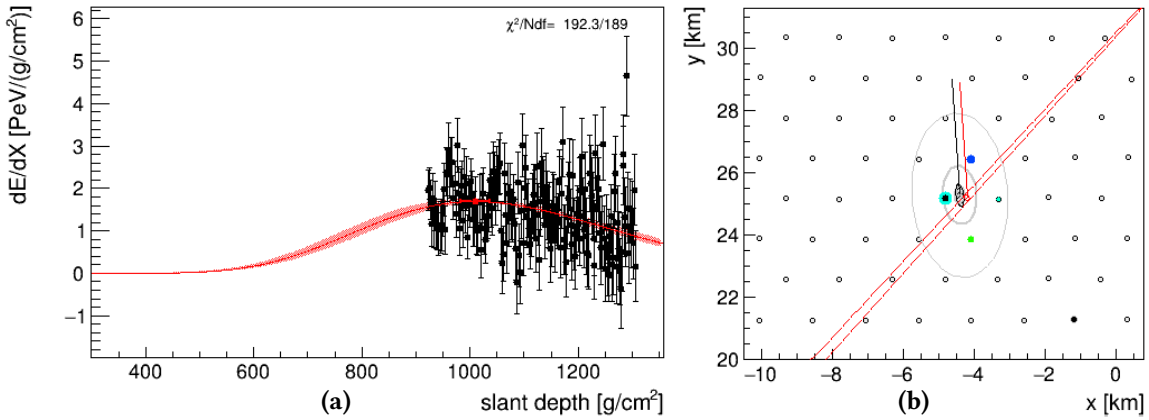



**Figure C.8:** *Longitudinal profile (C.8a) of the candidate 8556435. The red line indicates the reconstructed profile, while the red shaded area represents the uncertainty. The red point shows the reconstructed $X_{max}$. (C.8b). Event footprint on the SD array. Each colored circle indicate a triggered station. The colors (from blue to green) represent the trigger times while the dimensions are proportional to the signal size.*

**EVENT 9037274** This event, labeled with an ID = 9037274 was detected on January 19, 2010 at the 03:55:42 UTC. Figure C.9a shows its reconstructed profile: the hybrid reconstruction yields an energy $E_\gamma = 1.63 \times 10^{18}$ eV, a depth of the shower maximum $X_{\max} = 796\,\mathrm{g\,cm^{-2}}$ and a zenith angle $\theta = 22.7°$. The footprint of the event on the SD array is in turn shown in figure C.9b

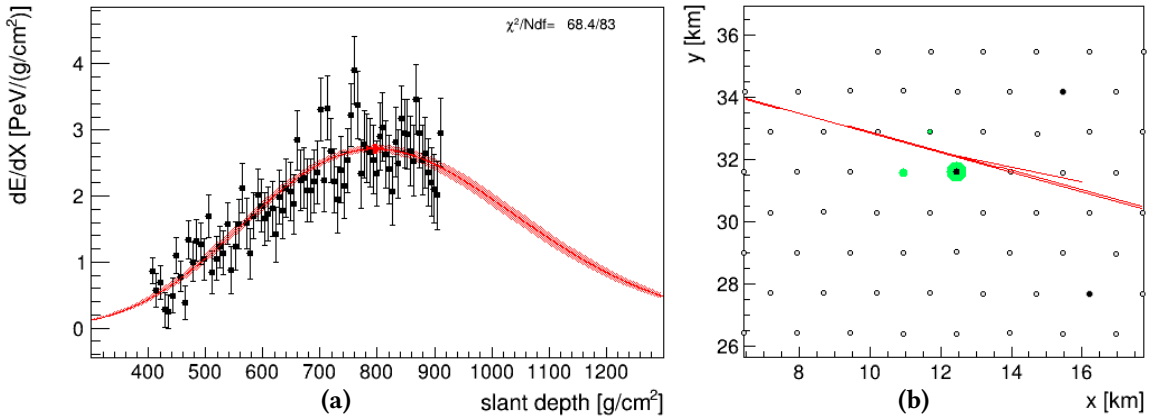



**Figure C.9:** *Longitudinal profile (C.9a) of the candidate 9037274. The red line indicates the reconstructed profile, while the red shaded area represents the uncertainty. The red point shows the reconstructed $X_{max}$. (C.9b). Event footprint on the SD array. Each colored circle indicate a triggered station. The colors (from blue to green) represent the trigger times while the dimensions are proportional to the signal size.*

**EVENT 10341437** This event, labeled with an ID = 10341437 was detected on October 3rd, 2010 at the 05:07:00 UTC. Figure C.10a shows its reconstructed profile: the hybrid reconstruction yields an energy $E_\gamma = 1.02 \times 10^{18}$ eV, a depth of the shower maximum $X_{\max} = 1020\,\mathrm{g\,cm^{-2}}$ and a zenith angle $\theta = 49.6°$. The footprint of the event on the SD array is in turn shown in figure C.10b

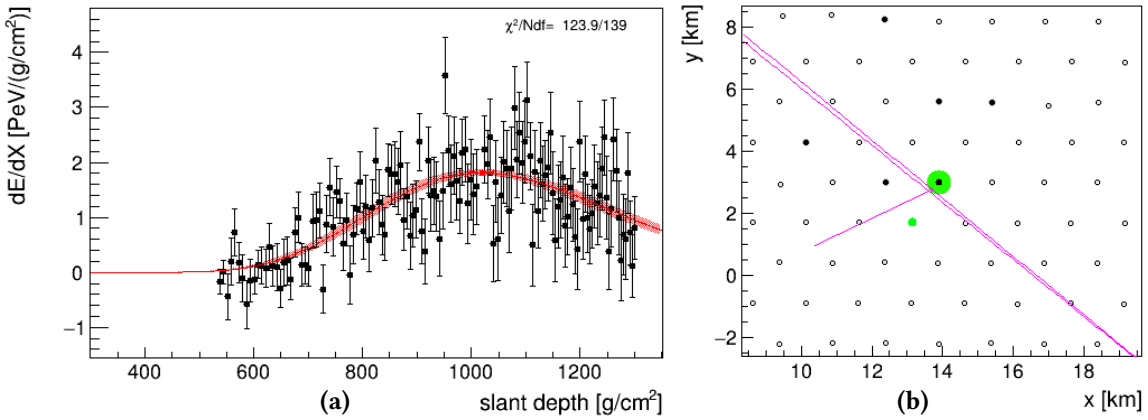



**Figure C.10:** *Longitudinal profile (C.10a) of the candidate 10341437. The red line indicates the reconstructed profile, while the red shaded area represents the uncertainty. The red point shows the reconstructed $X_{max}$. (C.10b). Event footprint on the SD array. Each colored circle indicate a triggered station. The colors (from blue to green) represent the trigger times while the dimensions are proportional to the signal size.*

**EVENT 10436573** This event, labeled with an ID = 10436573 was detected on October 16th, 2010 at the 07:33:46 UTC. Figure C.11a shows its reconstructed profile: the hybrid reconstruction yields an energy $E_\gamma = 1.38 \times 10^{18}$ eV, a depth of the shower maximum $X_{\max} = 985\,\mathrm{g\,cm^{-2}}$ and a zenith angle $\theta = 47.4°$. The footprint of the event on the SD array is in turn shown in figure C.11b.

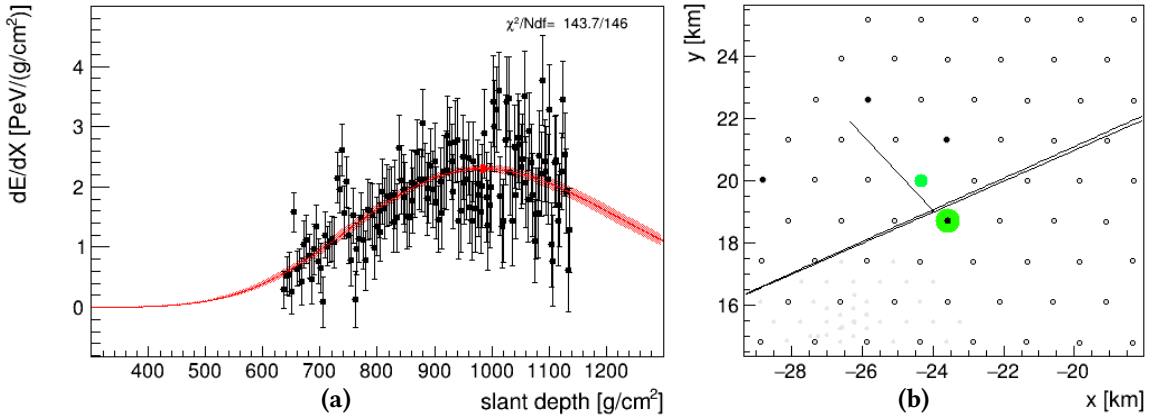



**Figure C.11:** *Longitudinal profile (C.11a) of the candidate 10436573. The red line indicates the reconstructed profile, while the red shaded area represents the uncertainty. The red point shows the reconstructed $X_{max}$. (C.11b). Event footprint on the SD array. Each colored circle indicate a triggered station. The colors (from blue to green) represent the trigger times while the dimensions are proportional to the signal size.*

**EVENT 12008234** This event, labeled with an ID = 12008234 was detected on June 26th, 2011 at the 05:17:41 UTC. Figure C.12a shows its reconstructed profile: the hybrid reconstruction yields an energy $E_\gamma = 1.45 \times 10^{18}\,\text{eV}$, a depth of the shower maximum $X_{\text{max}} = 935\,\text{g}\,\text{cm}^{-2}$ and a zenith angle $\theta = 30.9°$. The footprint of the event on the SD array is in turn shown in figure C.12b

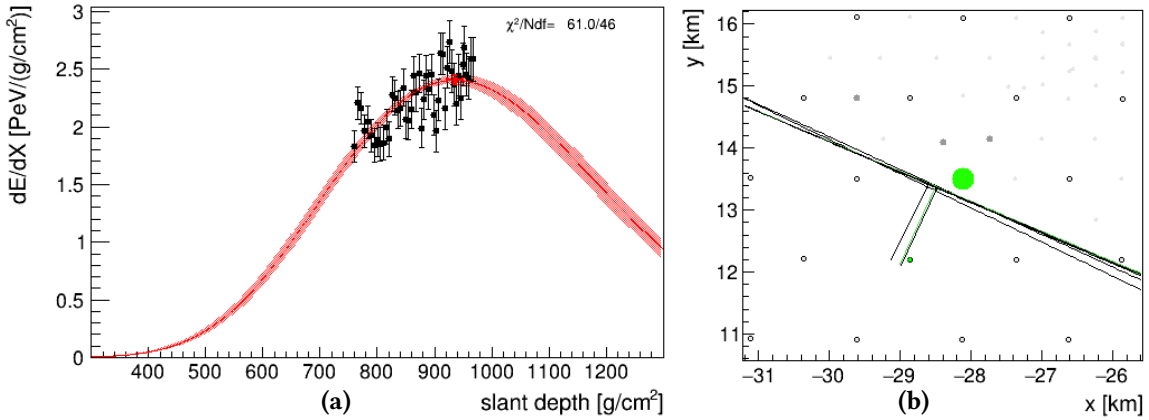



**Figure C.12:** *Longitudinal profile (C.12a) of the candidate 12008234. The red line indicates the reconstructed profile, while the red shaded area represents the uncertainty. The red point shows the reconstructed $X_{max}$. (C.12b). Event footprint on the SD array. Each colored circle indicate a triggered station. The colors (from blue to green) represent the trigger times while the dimensions are proportional to the signal size.*

**EVENT 12094673** This event, labeled with an ID = 12094673 was detected on 5th July, 2011 06:17:13 UTC. Figure C.13a shows its reconstructed profile: the hybrid reconstruction yields an

energy $E_\gamma = 1.04 \times 10^{18}$ eV, a depth of the shower maximum $X_{\max} = 1109\,\mathrm{g\,cm^{-2}}$ and a zenith angle $\theta = 57.2°$. The footprint of the event on the SD array is in turn shown in figure C.13b

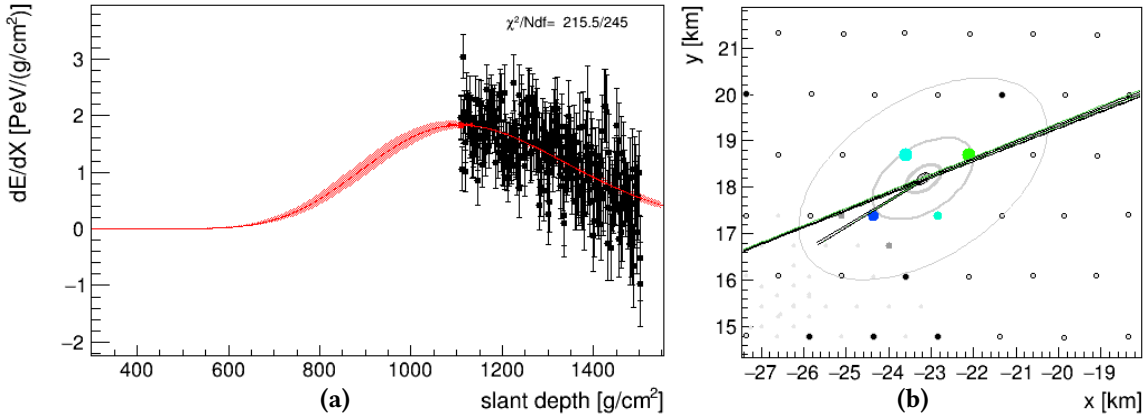



**Figure C.13:** *Longitudinal profile (C.13a) of the candidate 12094673. The red line indicates the reconstructed profile, while the red shaded area represents the uncertainty. The red point shows the reconstructed $X_{max}$. (C.13b). Event footprint on the SD array. Each colored circle indicate a triggered station. The colors (from blue to green) represent the trigger times while the dimensions are proportional to the signal size.*

**EVENT 12324657** This event, labeled with an ID = 12324657 was detected on August 3rd, 2011 at the 01:59:06 UTC. Figure C.14a shows its reconstructed profile: the hybrid reconstruction yields an energy $E_\gamma = 1.57 \times 10^{18}$ eV, a depth of the shower maximum $X_{\max} = 944\,\mathrm{g\,cm^{-2}}$ and a zenith angle $\theta = 54.6°$. The footprint of the event on the SD array is in turn shown in figure C.14b

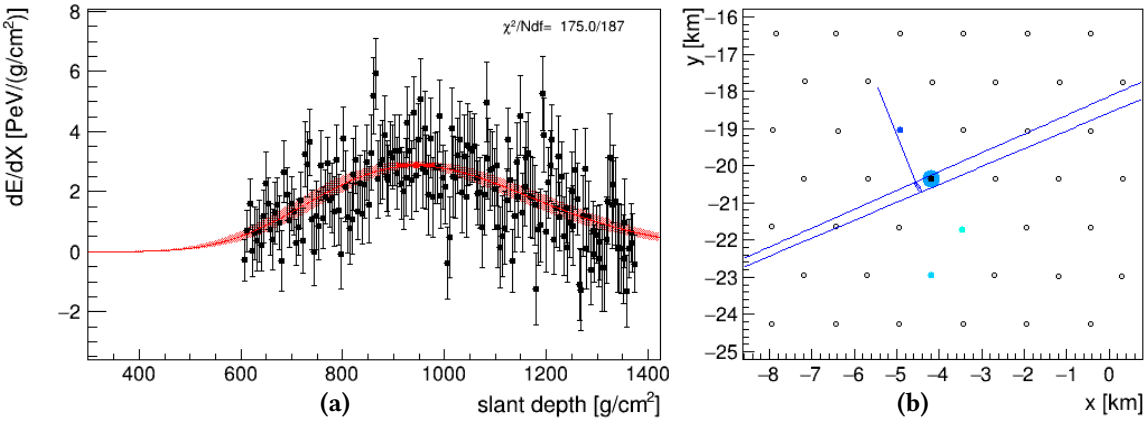



**Figure C.14:** *Longitudinal profile (C.14a) of the candidate 12324657. The red line indicates the reconstructed profile, while the red shaded area represents the uncertainty. The red point shows the reconstructed $X_{max}$. (C.14b). Event footprint on the SD array. Each colored circle indicate a triggered station. The colors (from blue to green) represent the trigger times while the dimensions are proportional to the signal size.*

**EVENT 13439833** This event, labeled with an ID = 13439833 was detected on December 22nd, 2011 at the 05:31:33. Figure C.15a shows its reconstructed profile: the hybrid reconstruction yields an energy $E_\gamma = 1.19 \times 10^{18}$ eV, a depth of the shower maximum $X_{\max} = 933\,\mathrm{g\,cm^{-2}}$ and a zenith angle $\theta = 44.2°$. The footprint of the event on the SD array is in turn shown in figure C.15b

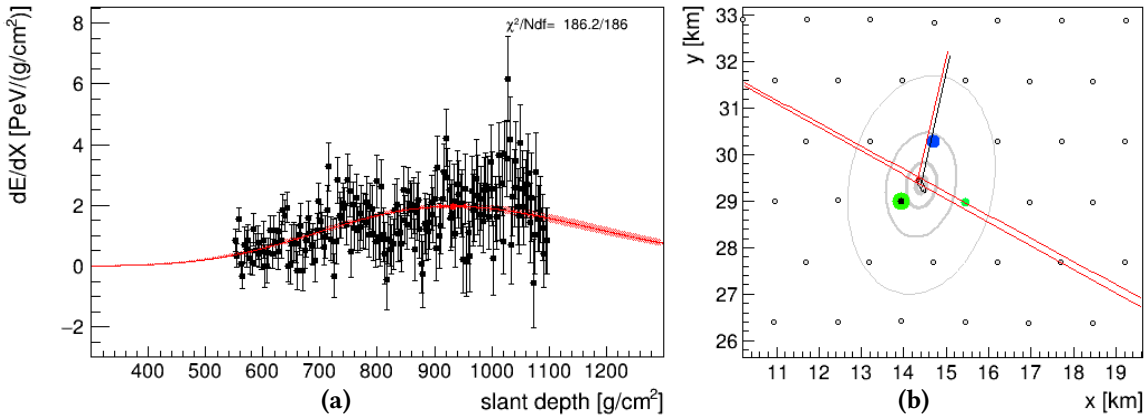



**Figure C.15:** *Longitudinal profile (C.15a) of the candidate 13439833. The red line indicates the reconstructed profile, while the red shaded area represents the uncertainty. The red point shows the reconstructed $X_{max}$. (C.15b). Event footprint on the SD array. Each colored circle indicate a triggered station. The colors (from blue to green) represent the trigger times while the dimensions are proportional to the signal size.*

**EVENT 20509091** This event, labeled with an ID = 20509091 was detected on November 13th, 2012 at the 06:51:13 UTC. Figure C.16a shows its reconstructed profile: the hybrid reconstruction yields an energy $E_\gamma = 1.09 \times 10^{18}$ eV, a depth of the shower maximum $X_{\max} = 967\,\mathrm{g\,cm^{-2}}$ and a zenith angle $\theta = 34.9°$. The footprint of the event on the SD array is in turn shown in figure C.16b

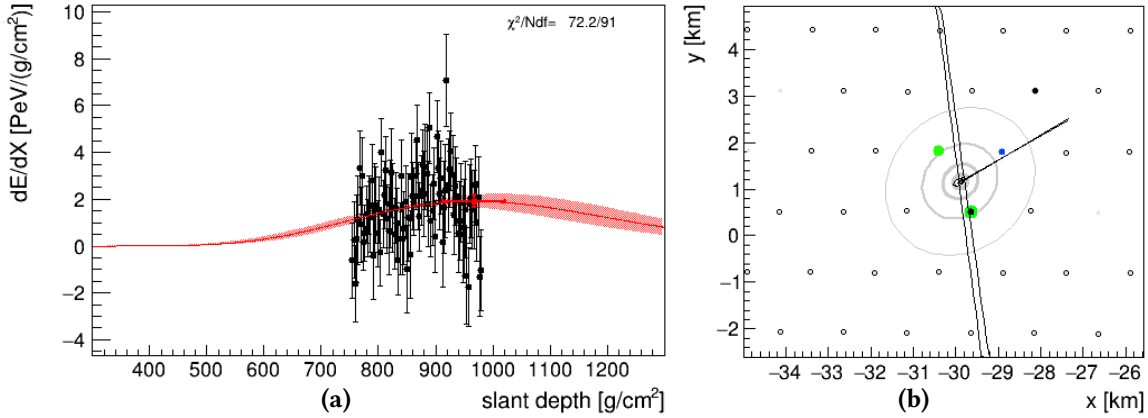



**Figure C.16:** *Longitudinal profile (C.16a) of the candidate 20509091. The red line indicates the reconstructed profile, while the red shaded area represents the uncertainty. The red point shows the reconstructed $X_{max}$. (C.16b). Event footprint on the SD array. Each colored circle indicate a triggered station. The colors (from blue to green) represent the trigger times while the dimensions are proportional to the signal size.*

**EVENT 22130573** This event, labeled with an ID = 22130573 was detected on June 30th, 2013 at the 02:01:08 UTC. Figure C.17a shows its reconstructed profile: the hybrid reconstruction yields an energy $E_\gamma = 1.10 \times 10^{18}$ eV, a depth of the shower maximum $X_{\max} = 1062\,\mathrm{g\,cm^{-2}}$ and a zenith angle $\theta = 41.7°$. The footprint of the event on the SD array is in turn shown in figure C.17b

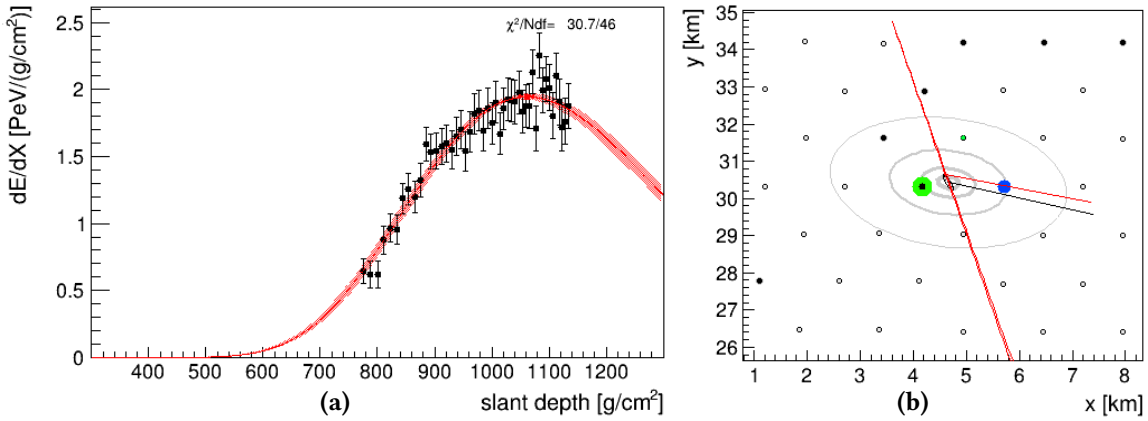



**Figure C.17:** *Longitudinal profile (C.17a) of the candidate 22130573. The red line indicates the reconstructed profile, while the red shaded area represents the uncertainty. The red point shows the reconstructed $X_{max}$. (C.17b). Event footprint on the SD array. Each colored circle indicate a triggered station. The colors (from blue to green) represent the trigger times while the dimensions are proportional to the signal size.*

**EVENT 31926471** This event, labeled with an ID = 31926471 was detected on March 15th, 2015. Figure C.18a shows its reconstructed profile: the hybrid reconstruction yields an energy $E_\gamma = 3.02 \times 10^{18}$ eV, a depth of the shower maximum $X_{\max} = 1002\,\mathrm{g\,cm^{-2}}$ and a zenith angle $\theta = 51.9°$. The footprint of the event on the SD array is in turn shown in figure C.18b

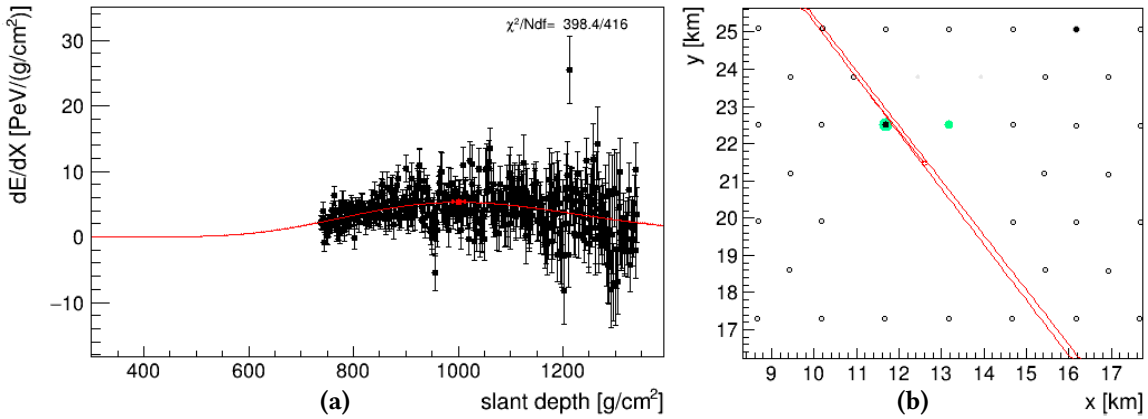



**Figure C.18:** *Longitudinal profile (C.18a) of the candidate 31926471. The red line indicates the reconstructed profile, while the red shaded area represents the uncertainty. The red point shows the reconstructed $X_{max}$. (C.18b). Event footprint on the SD array. Each colored circle indicate a triggered station. The colors (from blue to green) represent the trigger times while the dimensions are proportional to the signal size.*

**EVENT 37073026** This event, labeled with an ID = 37073026 was detected on March 8th, 2016 at the 01:23:38 UTC. Figure C.19a shows its reconstructed profile: the hybrid reconstruction yields an energy $E_\gamma = 1.09 \times 10^{18}$ eV, a depth of the shower maximum $X_{\max} = 954\,\mathrm{g\,cm^{-2}}$ and a zenith angle $\theta = 54.5°$. The footprint of the event on the SD array is in turn shown in figure C.19b

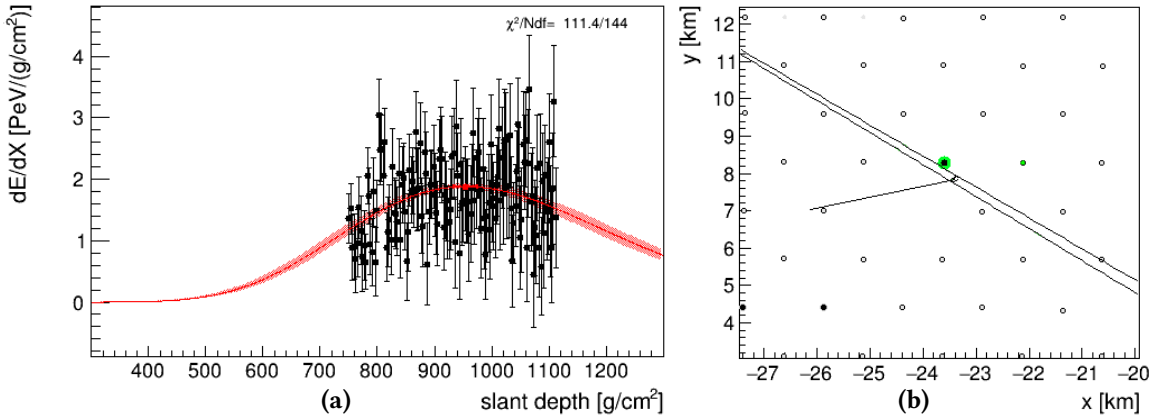



**Figure C.19:** *Longitudinal profile (C.19a) of the candidate 37073026. The red line indicates the reconstructed profile, while the red shaded area represents the uncertainty. The red point shows the reconstructed $X_{max}$. (C.19b). Event footprint on the SD array. Each colored circle indicate a triggered station. The colors (from blue to green) represent the trigger times while the dimensions are proportional to the signal size.*

**EVENT 38530128** This event, labeled with an ID = 38530128 was detected on July 5th, 2016 at the 06:01:34 UTC. Figure C.20a shows its reconstructed profile: the hybrid reconstruction yields an energy $E_\gamma = 1.32 \times 10^{18}$ eV, a depth of the shower maximum $X_{\max} = 917\,\mathrm{g\,cm^{-2}}$ and a zenith angle $\theta = 48.1°$. The footprint of the event on the SD array is in turn shown in figure C.20b

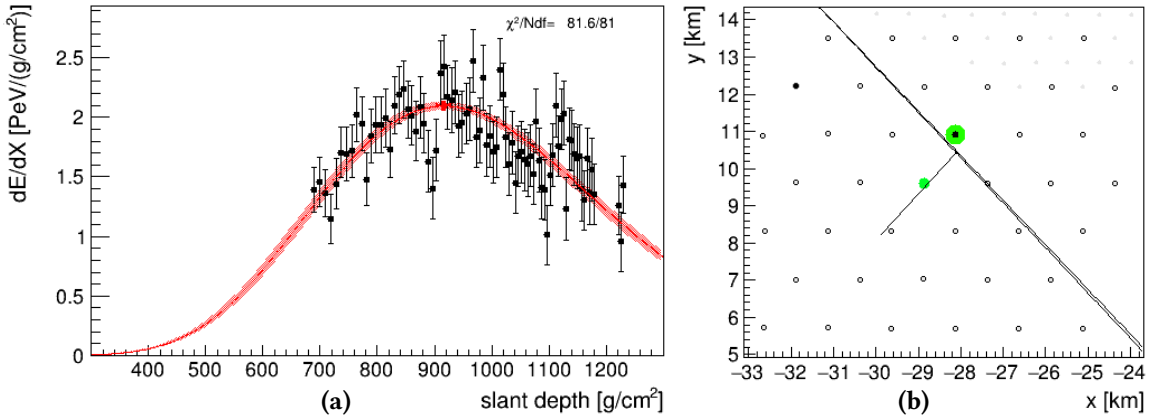



**Figure C.20:** *Longitudinal profile (C.20a) of the candidate 38530128. The red line indicates the reconstructed profile, while the red shaded area represents the uncertainty. The red point shows the reconstructed $X_{max}$. (C.20b). Event footprint on the SD array. Each colored circle indicate a triggered station. The colors (from blue to green) represent the trigger times while the dimensions are proportional to the signal size.*

**EVENT 39077494** This event, labeled with an ID = 39077494 was detected on August 11th, 2016 at the 07:52:15 UTC. Figure C.21a shows its reconstructed profile: the hybrid reconstruction yields an energy $E_\gamma = 1.17 \times 10^{18}\,\mathrm{eV}$, a depth of the shower maximum $X_{\max} = 847\,\mathrm{g\,cm^{-2}}$ and a zenith angle $\theta = 58.5°$. The footprint of the event on the SD array is in turn shown in figure C.21b

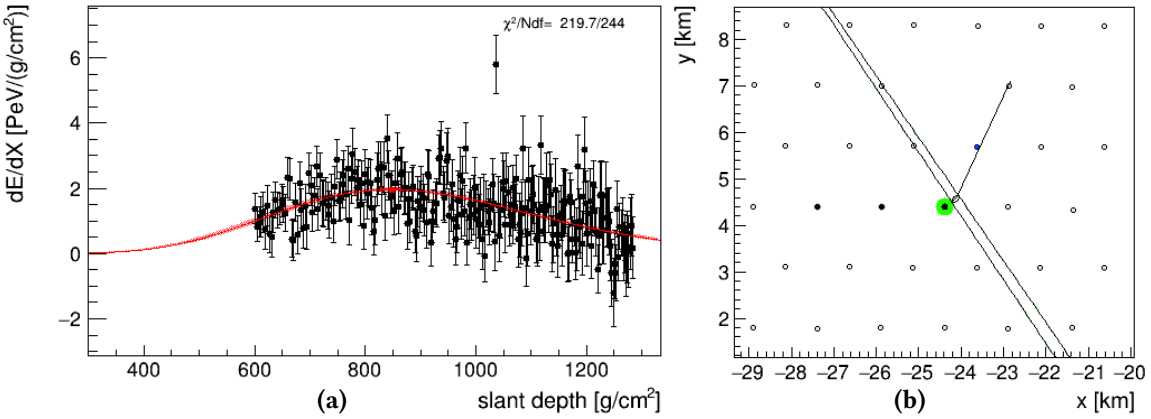



**Figure C.21:** *Longitudinal profile (C.21a) of the candidate 39077494. The red line indicates the reconstructed profile, while the red shaded area represents the uncertainty. The red point shows the reconstructed $X_{max}$. (C.21b). Event footprint on the SD array. Each colored circle indicate a triggered station. The colors (from blue to green) represent the trigger times while the dimensions are proportional to the signal size.*

**EVENT 43283582** This event, labeled with an ID = 43283582 was detected on June 19th, 2017 at the 01:14:36 UTC. Figure C.22a shows its reconstructed profile: the hybrid reconstruction yields an energy $E_\gamma = 1.11 \times 10^{18}\,\mathrm{eV}$, a depth of the shower maximum $X_{\max} = 850\,\mathrm{g\,cm^{-2}}$ and a zenith angle $\theta = 42.4°$. The footprint of the event on the SD array is in turn shown in figure C.22b

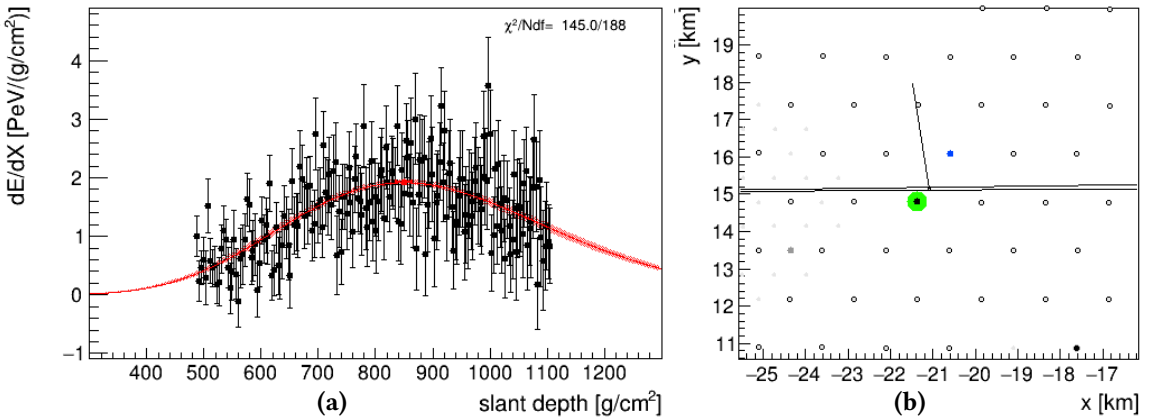



**Figure C.22:** *Longitudinal profile (C.22a) of the candidate 43283582. The red line indicates the reconstructed profile, while the red shaded area represents the uncertainty. The red point shows the reconstructed $X_{max}$. (C.22b). Event footprint on the SD array. Each colored circle indicate a triggered station. The colors (from blue to green) represent the trigger times while the dimensions are proportional to the signal size.*

# BIBLIOGRAPHY


[1] L. Timbaldo for the Fermi Collaboration, "A tale of cosmic rays narrated in gamma rays by Fermi", Highlight talk on the 33rd ICRC, Rio de Janeiro (2013);

[2] IceCube Collaboration, "Evidence for High-Energy Extraterrestrial Neutrinos at the IceCube Detector", Science (342), 2013, 1242856;

[3] The Pierre Auger Collaboration, "Large-scale cosmic-ray anisotropies above 4 EeV measured by the Pierre Auger Observatory", ApJ (868), 2018, 4;

[4] IceCube collaboration, "Energy and wavelength spectra", https://icecube.wisc.edu/science/highlights/neutrino_astronomy;

[5] M. Tanabashi *et al.* (Particle Data Group), "Review of Particle Physics", Phys. Rev. D. (98), 2018, 454, and references therein;

[6] T. C. Weekes, "Very High Energy Gamma Ray Astronomy", Publ. I.O.P. (U.K.) 2003;

[7] D. J. Thompson, "Gamma ray astrophysiscs: the EGRET results", Rep. Prog. Phys. (71), 2008, 116901;

[8] S. P. Wakely, and D. Horan, "TevCat: An online catalog for Very High Energy Gamma-Ray Astronomy", International Cosmic Ray Conference (3), 2008, 1341-1344, http://tevcat.uchicago.edu;

[9] The HAWC Collaboration, "The 2HWC HAWC Observatory Gamma-Ray Catalog", ApJ (843), 2017, 1;

[10] M. Amenomori et al. (Tibet AS$\gamma$ Collaboration), "First detection of sub-PeV diffuse gamma rays from the Galactic disk: Evidence for ubiquitous galactic cosmic rays beyond PeV energies", Phys. Rev. Lett. (126), 2021, 141101;

[11] The Pierre Auger Collaboration, "A search for point sources of EeV photons", ApJ (789), 2014, 160;

[12] N. Gehrels et al., "The Compton Gamma Ray Observatory", ApJS (92), 1994, 351 – 362;

[13] W. N. Johnson et al., "The Oriented Scintillation Spectrometer Experiment - Instrument description", ApJS (86), 1993, 693 – 712;

[14] R. Diehi et al., "COMPTEL Observations of the 1.809 MeV Gamma-Ray Line from Galactic 26Al", ApJS (92), 1994, 429;

[15] Fermi LAT Collaboration, "The Large Area Telescope on the Fermi Gamma-Ray Space Telescope Mission", Astrop. J. (697), 2009, 1071 – 1102;

[16] J.A. Hinton, "The status of the H.E.S.S. project", New Astron.Rev. (48), 2004, 331-337;

[17] Krennrich, F. et al., "VERITAS: The Very Energetic Radiation Imaging Telescope Array System", New Astronomy Reviews (48), 1999, 637 – 641;

[18] MAGIC collaboration, "Technical status of the MAGIC telescopes", Proc. International Cosmic Rays Conference, 2009;

[19] G. Sinnis, "The HAWC TeV Gamma-ray Observatory", Proceedings of SciNeGHE, September 2010, Trieste, Italy;

[20] H. Kawai et al., "Telescope Array Experiment", Nuclear Physics B: Proceedings Supplements (175), 2008, 220 – 226;

[21] M. Risse, P. Homola, "Search for ultra-high energy photons using air showers", Modern Physics Letter A (22), 2007, 749;

[22] W. F. G. Swann, "A Mechanism of Acquirement of Cosmic-Ray Energies by Electrons", Phys. Rev. (43), 1933, 217 – 220;

[23] K. S. Cheng et al., "Energetic radiation from rapidly spinning pulsars. II. Vela and Crab", Astrop. J. (300), 1986, 522 – 539;

[24] A. Venkatesan, "Constraints on the production of ultra-high-energy cosmic rays by isolated neutron stars", Astrop. J. (484), 1997, 323 – 328;

[25] M. Spurio, "Particles and Astrophysics A Multi-Messenger Approach", Springer, 2014, 243-321, ISBN 978-3-319-08051-2;

[26] F. Halzen, "High-energy neutrino astrophysics", Nature Phys, (13), 2017, 232 – 238;

[27] P. Bhattacharjee, "Origin and Propagation of Extremely High Energy Cosmic Rays", Phys. Rept. (327), 2000, 109-247;

[28] C. T. Hill, "Monopolonim", Nucl. Phys. B (224), 1983, 469-490;

[29] M.B. Hindmarsh, T. W. B. Kibble, "Cosmic strings", Rept. Prog. Phys. (58), 1995, 477-562;

[30] J. Dunkley et al. "Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Likelihoods and Parameters from the WMAP data" Astrophys. J. Suppl. (180), 2009, 306-329;

[31] M. Birkel, S. Sarkar, "Extremely high energy cosmic rays from relic particle decays", Astropart. Phys. (9), 1998, 297;

[32] P. Blasi, R. Dick, E. W. Kolb, "Ultra-high energy cosmic rays from annihilation of superheavy dark matter" Astropart. Phys. (18), 2002, 57;

[33] D.Fargion, B. Mele, A. Salis, "Ultrahigh energy neutrino scattering onto relic light neutrinos in galactic halo as a possible source of highest energy extragalactic cosmic rays", Astrophys. J. (517), 1999, 725-733;

[34] B. Fuchs, "DARK98 Proceedings the Second International Conference on Dark Matter in Astro and Particle Physics", 1999;

[35] G. Gelmini et al., "GZK Photons as Ultra-High-Energy Cosmic Rays", J. Exp. Theor. Phys. (106), 2008, 1061 – 1082;

[36] K. Greisen, "End to the Cosmic Ray Spectrum?", Phys. Rev. Lett. (16), 1966, 748;

[37] G.T. Zatsepin, V. A. Kuzmin, "Upper limit of the spectrum of cosmic rays", JETP Lett. (4), 1966, 78-80;

[38] D. Hooper et al., "Cosmogenic photons as a test of ultra-high energy cosmic ray composition", Astrop. phys. (34), 2011, 340 – 343;

[39] G. Decerprit et al., "Constraints on the origin of ultra-high-energy cosmic rays from cosmogenic neutrinos and photons", Astronomy & Astrophysics (535), 2011, A66;

[40] R. Ruffini, G.V. Vereshchagin, S. S. Xue "Cosmic absorption of ultra high energy particles" Astrophys. Space Sci. (361), 2016, 82;

[41] P. Bhattacharjee et al., "Origin and propagation of extremely high-energy cosmic ray", Physics Reports (327), 2000, 109 – 247,

[42] T. Stanev, "High Energy Cosmic Rays", second edition, Springer-Verlag, 2010.

[43] P. S. Coppi, F. A. AHARONIAN, "Constraints on the very high energy emissivity of the universe from the diffuse GeV Gamma-ray baclground", Astrophys. J. (487), 1997, L9-L12;

[44] R. Engel, "Indirect Detection of Cosmic Rays", Handbook of Particle Detection and Imaging, Springer-Verlag, 2012.

[45] B. Rossi, "High Energy Particles", Prentice-Hall, Englewood Cliffs, NJ, 1952;

[46] W. Heitler, "Quantum theory of radiation.", Oxford University press, 1944;

[47] S. Braibant, G. Giacomelli, M. Spurio, "Particles and Fundamental Interactions", Springer, 2012, 43, ISBN 978-94-007-2464-8;

[48] A. B. Migdal, "Bremsstrahlung and Pair Production in Condensed Media at High Energies", Phys. Rev. (103), 1956, 1811 – 1820;

[49] J. K. Daughterty, A. K. Harding, "Pair production in superstrong magnetic fields", Astrophys. J. (273), 1983, 761;

[50] The Pierre Auger Collaboration, "Characteristics of geomagnetic cascading of ultra-high energy photons at the southern and northern sites of the Pierre Auger Observatory", Astropart. phys. (27), 2007, 174;

[51] J. Matthews, "A Heitler model of extensive air showers", Astropart. phys. (22), 2005, 387 – 397;

[52] S. Eidelman et al. (Particle data group), "Review of particle physics", Phys. Rew. Lett. B (592.1), 2004;

[53] M. Settimo for the Pierre Auger Collaboration, "Search for ultra-High Energy Photons with the Pierre Auger Observatory", Proceedings of the International Conference on the Structure and the Interactions of the Photon (Photon 2013)

[54] T. Pierog, K. Werner, "EPOS Model and Ultra High Energy Cosmic Rays", Nucl. Phys. Proc. Suppl. (196), 2009, 102-105;

[55] E. Ahn, R. Engel, T. K. Gaisser, P. Lipari, T. Stanev, "Cosmic ray event generator Sibyll 2.1", Phys. Rev. D (80), 2009, 094003;

[56] S. Ostapchenko, "Status of QGSJET", AIPConf. Proc. (928), 2007, 118-125;

[57] P. Savina et al., "Searching for UHE photons in the EeV range: a two-variable approach exploiting air-shower universality.", PoS(ICRC2019)414, 2019;

[58] J. Abraham, "Properties and performances of the prototype instrument for the Pierre Auger Observatory", NIM A (586), 2007 409-420;

[59] The Pierre Auger Collaboration, "The Pierre Auger Observatory Upgrade - Preliminary Design Report", arXiv:1604.03637v1, 2016;

[60] I. Allekotte et al., "The Surface Detector System of the Pierre Auger Observatory", NIM A (586), 2007, 409 - 420;

[61] X. Bertou et al., "Calibration of the surface array of the Pierre Auger Observatory", NIM A (A), 2006, 839 - 846;

[62] The Pierre Auger Collaboration, "Reconstruction of the events recorded with the surface detector of the Pierre Auger Observatory", JINST (15), 2020, P10021;

[63] R. Hiller, M. Roth, "An update on the signal accuracy using the infill array", Auger Internal Note - GAP2012-012, 2012;

[64] The Pierre Auger Collaboration, "Trigger and aperture of the surface detector array of the Pierre Auger Observatory", NIM A (613), 2010, 29 - 39;

[65] The Pierre Auger Collaboration, "New proposal to improve the local trigger of the Surface Detector array of the Pierre Auger Observatory", Auger Internal Note - GAP2011-089, 2011;

[66] J. Abraham, et al., "Trigger and aperture of the surface detector array of the Pierre Auger Observatory", Nuclear Instruments and Methods in Physics Research Section A (613), 2010, 29.

[67] J. Nishimura, K. Kamata, "The lateral and the angular structure functions of electron showers", Prog. Theor. Phys. Supp. (6), 1958, 93 - 155;

[68] K. Greisen, "Progress in Elementary Particle and Cosmic Ray Physics", North - Holland Publishing Company (3), 1956.

[69] M. Roth, for the Pierre Auger Collaboration, "The lateral distribution function of shower signals in the surface detector of the Pierre Auger Observatory", HE1.1, Proc. 28th Int. Cosmic Ray Conf. (333), 2003;

[70] D. Newton et al., "The optimum distance at which to determine the size of a giant air shower", Astroparticle Physics (26), 2007, 414;

[71] R. Pesce for the Pierre Auger Collaboration, "Energy calibration of data recorded with the surface detectors of the Pierre Auger Observatory: an update", Proceeding of the 32nd ICRC, 2011, 13, arXiv:1107.4809;

[72] The Pierre Auger Collaboration, "Measurement of the cosmic ray energy spectrum above $2.5 \times 10^{18}$ eV using the Pierre Auger Observatory", Physical Review D (102), 2020, 062005;

[73] A. Schulz, "Measurement of the cosmic ray spectrum above $3 \times 10^{17}$ eV with the Pierre Auger Observatory", Proceeding og the 33rd ICRC, 2013, arXiv:1307.5059;

[74] The Pierre Auger Collaboration, "The fluorescence detector of the Pierre Auger Observatory", NIM A (620), 2010, 227 - 251;

[75] The Pierre Auger Collaboration, "Technical Design Report", 2004;

[76] J. T. Brack et al., "Absolute calibration of a large-diameter light source", Journal of Instrumentation (8), 2013, P05014;

[77] The Pierre Auger Collaboration, "The Absolute, Relative and Multi-Wavelength Calibration of the Pierre Auger Observatory Fluorescence Detectors", Proceedings of the 30th ICRC (4), 2008, 343 - 346;

[78] M. Unger et al., "Reconstruction of Longitudinal Profiles of Ultra-High Energy Cosmic Ray Showers from Fluorescence and Cherenkov Light Measurements", Nucl. Instrum. Meth. A (588), 2008, 433;

[79] T. Gaisser and A. Hillas, "Reliability of the method of constant intensity cuts for reconstructing the average development of vertical showers", Proceeding of 15th ICRC (8), 1977, p 353;

[80] The Pierre Auger Collaboration, "Measurement of the average shape of longitudinal profiles of cosmic-ray air showers at the Pierre Auger Observatory", JCAP (03), 2019, 018;

[81] M. Ave et al., "Precise measurement of the absolute fluorescence yield of the 337 nm band in atmospheric gases", Astropart. Phys. (42), 2013, 90;

[82] M. Tueros, "Estimate of the non-calorimetric energy of showers observed with the fluorescence and surface detectors of the Pierre Auger Observatory", 2013, p. 11

[83] The Pierre Auger Collaboration, "Data-driven estimation of the invisible energy of cosmic ray showers with the Pierre Auger Observatory", Phys. Rew. D (100), 2019, 082003;

[84] The Pierre Auger Collaboration, "Muons in air showers at the Pierre Auger Observatory: Mean number in highly inclined events" Phys. Rev. D (91), 2015, 032003; Erratum Phys. Rev. D (91), 2015, 059901;

[85] V. Verzi for the Pierre Auger Observatory, "The Energy Scale of the Pierre Auger Observatory", ICRC 2013, 0928;

[86] B. R. Dawson for The Pierre Auger Collaboration, "The Energy Scale of the Pierre Auger Observatory" PoS(ICRC2019)231;

[87] F. Arciprete et al., "AIRFLY: Measurement of the fluorescence yield in atmospheric gases", Nuclear Physics B (150), 2006, 186 – 189;

[88] P. Nguyen, "Energy Systematics and Long Term Performance of the Pierre Auger Observatory's Fluorescence Telescopes", PhD Thesis, University of Adelaide;

[89] The Pierre Auger Collaboration, "The Rapid Atmosphere Monitoring System of the Pierre Auger Observatory", JINST (7), 2012, P09001, arXiv:1208.1675;

[90] J. Bellido for the Pierre Auger Collaboration, "Depth of maximum of air-shower profiles at the Pierre Auger Observatory: Measurements above $10^{17.2}$ eV and Composition Implications", PoS(ICRC2017)506;

[91] The Pierre Auger Collaboration, "The Pierre Auger Observatory Upgrade: Auger-Prime. Preliminary Design Report", arXiv:1604.03637, 2016;

[92] The Pierre Auger Collaboration, "Searches for anisotropies in the arrival directions of the highest energy cosmic rays detected by the Pierre Auger Observatory", ApJ (804), 2015, 15;

[93] The Pierre Auger Collaboration, "An Indication of Anisotropy in Arrival Directions of Ultra-high-energy Cosmic Rays through Comparison to the Flux Pattern of Extragalactic Gamma-Ray Sources", Apjl (853), 2018, L29;

[94] A. M. Hillas, "The Origin of ultra-high-energy cosmic rays", Ann. Rev. Astron. Astrophys. (22), 1984, 425 – 444;

[95] P. L. Biermann et al., "The Nature and Origin of Ultra-High Energy Cosmic Ray Particles", Frascati Phys.Ser. 64 (2017) 103-121;

[96] D. A. Perley et al., "Long-Duration Gamma-Ray Burst Host Galaxies in Emission and Absorption", Space Science Reviews (202), 2016, 111 – 142;

[97] The Pierre Auger Collaboration, "An Indication of Anisotropy in Arrival Directions of Ultra-high-energy Cosmic Rays through Comparison to the Flux Pattern of Extragalactic Gamma-Ray Sources", Astrophys. J. Lett. (853), 2018, L29;

[98] The Pierre Auger, "Features of the energy spectrum of cosmic rays above

[99]

[100] R. U. Abbasi et al., "First Observation of the Greisen-Zatsepin-Kuzmin Suppression", Phys. Rev. Lett. (100), 2008, 101101;

[101] The Pierre Auger Collaboration, "Observation of the Suppression of the Flux of Cosmic Rays above $4 \times 10^{19}$ eV", Phys. Rev. Lett. (101), 061101;

[102] Telescope Array Collaboration, "The cosmic-ray energy spectrum observed with the surface detector of the telescope array experiment", ApJL (768), 2013, L1;

[103] B. Sarkar et al., "Ultra-High Energy Photon and Neutrino Fluxes in Realistic Astrophysical Scenarios" Proceedings of the 32nd Int. Cosmic Ray Conf. (2), 2011, 198;

[104] J. Ellis et al., "Ultrahigh-energy cosmic rays particle spectra from crypton decays", Phys. Rev. D (74), 2006, ) 115003;

[105] R. Aloisio, S. Matarrese, A.V. Olinto, "Super Heavy Dark Matter in light of BICEP2, Planck and Ultra High Energy Cosmic Rays Observations", J. Cosmol. Astropart. P. (08), 2015, 24;

[106] C. Bleve for the Pierre Auger Collaboration, "Updates on the neutrino and photon limits from the Pierre Auger Observatory", PoS(ICRC2015) 1103 (2015);

[107] The Planck Collaboration, "Planck 2015 results. XX. Constraints on inflation", A&A (594), 2016, A20;

[108] O. E. Kalashev and M. Yu. Kuznetsov, "Constraining heavy decaying dark matter with the high energy gamma-ray limits", Phys. Rev. D (94), 2016, 063535;

[109] M. Settimo for the Pierre Auger Collaboration, "An update on a search for ultra-high energy photons using the Pierre Auger Observatory", Proceedings of the 32nd Int. Cosmic Ray Conf. (2), 2011, 55;

[110] Rubtsov. G.I. et al. for the Telescope Array Collaboration "Telescope Array search for photons and neutrinos with the surface detector data" PoS(ICRC2015) 331;

[111] K. Shinozaki et al., "Upper Limit on Gamma-Ray Flux above 1019 eV Estimated by the Akeno Giant Air Shower Array Experiment", Astrophys. J. (571), 2002, L117;

[112] A. Glushkov et al., "Constraints on the flux of primary cosmic-ray photons at energies E>$10^{18}$ eV from Yakutsk muon data", Phys. Rev. D (82), 2010, 041101: 1;

[113] M. Ave et al., "New Constraints from Haverah Park Data on the Photon and Iron Fluxes of Ultrahigh-Energy Cosmic Rays", Phys. Rev. Lett. (85), 2000, 2244;

[114] The Pierre Auger Collaboration, "Search for photons with energies above $10^{18}$ eV using the hybrid detector of the Pierre Auger Observatory", JCAP04(2017)009;

[115] K.S. Capelle, J.W. Cronin, G. Parente, E. Zas, "On the detection of ultra high energy neutrinos with the Auger observatory", Astropart. Phys. (8), 1998, 321;

[116] I. Allekotte et al., "The Surface Detector System of the Pierre Auger Observatory,", Nucl. Instrum. Meth. A (586), 2008, 409;

[117] IceCube Collaboration, "Constraints on Ultrahigh-Energy Cosmic-Ray Sources from a Search for Neutrinos above 10 PeV with IceCube", Phys. Rev. Lett. (117), 2016, 241101; Erratum: Phys. Rev. Lett. (119), 2017, 259902;

[118] The Pierre Auger Collaboration, "Evidence for a mixed mass composition at the ankle in the cosmic-ray spectrum", Phys. Lett. B (762), 2016, 288;

[119] The Pierre Auger Collaboration, "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", Phys. Rev. D (96), 2017, 122003;

[120] K. Murase, Y. Inoue, and C. D. Dermer, "Diffuse neutrino intensity from the inner jets of active galactic nuclei: Impacts of external photon fields and the blazar sequence", Phys. Rev. D (90), 2014, 023007;

[121] K. Fang, K. Kotera, K. Murase, and A. V. Olinto, "Testing the Newborn Pulsar Origin of Ultrahigh Energy Cosmic Rays with EeV Neutrinos", Phys. Rev. D 90, 103005 (2014); Erratum: Phys. Rev. D 92 (2014) 129901;

[122] L. Lu for the IceCube Collaboration, "Recent IceCube results evidences of neutrino emission from the blazar TXS 0506+056 and searches for Glashow resonance", talk given at the Ultra High Energy Cosmic Rays 2018, Paris, France, 8 - 12 October 2018;

[123] The Pierre Auger Collaboration, "Probing the origin of ultra-high-energy cosmic rays with neutrinos in the EeV energy range using the Pierre Auger Observatory", JCAP (10), 2019, 022, arXiv:1906.07422;

[124] K. H. Kampert, M. A. Mostafa, E. Zas, and The Pierre Auger Collaboration, "Multi-Messenger Physics With the Pierre Auger Observatory", Front. Astron. Space Sci. (6), 2019, 24;

[125] P. Lipari, "Concept of "age" and "universality" in cosmic rays showers", PRD (79), 2009, 063001;

[126] The Pierre Auger Collaboration, "Applying EAS Universality to ground detector data", Proc. 30th ICRC, 2007, arXiv:07061990;

[127] M. Ave et al., "A generalized description of the signal size in extensive air shower detectors and its applications", Astroparticle Physics (87), 2017, p. 23 - 39;

[128] M. Ave et al., "A generalized description of the time dependent signals in extensive air shower detectors and its applications", Astroparticle Physics (88), 2017, p. 46 - 59;

[129] R. Aloisio et al., "SimProp: a simulation code for Ultra-High Energy Cosmic Ray Propagation", JCAP (1210), 2012, 7, arXiv:1204.2970.

[130] K. Karl-Heinz et al., "CRPropa 2.0 – a Public Framework for Propagating High Energy Nuclei, Secondary Gamma Rays and Neutrinos", Astropart. Phys. (42), 2013, 41 - 51, arXiv:1206.3132;

[131] B. Rossi, K. Greisen, "Cosmic-Ray Theory", Rev. Mod. Physics (13), 1941, 240;

[132] A. M. Hillas, "Angular and energy distributions of charged particles in electron photon cascades in air", J. Phys. (G8), 1982, 1461 - 1473;

[133] M. Giller, A. Kacperczyk, J. Malinowski, W. Tkaczyk, G. Wieczorek, "Similarity of extensive air showers with respect to the shower age", J- Phys. (G31), 2005, 947 - 958;

[134] F. Nerling, J. Blümer, R. Engel, M. Risse, "Universality of electron distributions in high-energy air showers: Description of cherenkov light production", Astropart. Phys. (24), 2006, 421-437;

[135] P. Lipari, "Universality of cosmic ray shower development", Nucl. Phys. Proc. Suppl. (196), 2009, 309 - 318;

[136] S. Lafebre, R. Engel, H. Falcke, J. Hörandel, T. Huege, J. Kuijpers, R. Ulrich, "Universality of electron-positron distributions in extensive air showers", Astropart. Phys. (31), 2009, 243 - 254;

[137] S. Andringa et al., “The muonic longitudinal shower profiles at production”, Astropart. Phys. (35), 2012, 821 - 827;

[138] L. Cazon et al., “A model for the transport of muons in extensive air showers”, Astropart. Phys. (36), 2012, 211 - 223;

[139] F. Schmidt for the Pierre Auger Collaboration, “Applying extensive air shower universality to ground detector data”, Proc. 30th Int. Cosmic Ray Conf., 2007;

[140] M. Ave, R. Engel, J. Gonzalez, D. Heck, T. Pierog, M. Roth, “Extensive air shower universality of ground particle distributions”, Proc. of 31st Int. Cosmic Ray Conf. 2011;

[141] D. Maurel, M. Roth, J. Gonzalez, “Universality of the time structure of ground particle distributions and its application to the reconstruction of extensive air showers”, ICRC Proceedings 2013;

[142] M. Ave, M. Roth, A. Schulz, “A Universal description of temporal and lateral distributions of ground particles in extensive air showers”, Proceedings of 34th ICRC, 2015;

[143] D. Maurel, “Mass composition of ultra-high energy cosmic rays based on air shower universality”, Auger Internal Note GAP-2013-105, 2013;

[144] M. Ave et al., “Can EPOS reproduce the Auger SD and Hybrid Data?”, Auger Internal Note GAP-2007-098;

[145] M. Ave et al., “Prediction of the tank response $S(r, DX, E|\theta, \psi)$ from shower universality”, Auger Internal Note GAP-2011-087;

[146] A. Schulz, “Measurements of the Energy Spectrum and Mass Composition of Ultra-high Energy Cosmic Rays”, Auger Internal Note GAP-2016-021, 2016;

[147] Rémi Bardenet et al., “Single muon response: the signal model”, Auger Internal Note GAP-2010-110, 2010;

[148] G. A. Anastasi, “Reconstruction of events from the surface detector of the Pierre Auger Observatory using air shower Universality”, Auger Internal Note GAP-2019-035;

[149] The Pierre Auger Observatory, “The Pierre Auger Cosmic Ray Observatory”, Nucl. Inst. Meth. A (768), 2015, 172 – 213;

[150] H. Dembinski et al., “The Auger Observer”, 2014, http://augerobserver.fzk.de/.;

[151] A. Roodman, “Blind Analysis in Particle Physics”, ECONF C030908:TUIT001,2003;

[152] D.Heck, J. Knapp, J. Capdevielle, G. Schatz, T. Thouw, “CORSIKA: A Monte Carlo code to simulate extensive air showers”, Wissenschaftliche Berichte, Forschungszentrum Karlsruhe FZKA 6019 (1998);

[153] A. Fassó, et al. “The physics model of Fluka: status and recent development”, Computing in High Energy and Nuclear Physics 2003 Conference (CHEP2003), La Jolla, CA, USA, 2003, arXiv:hep-ph/0306267;

[154] M. Hladik, H. J. Drescher, S. Ostapchenko, T. Pierog, K. Werner, “Self-Consistency Requirement in High-Energy Nuclear Scattering”, Phys. Rev. Lett. (86), 2001, 3506;

[155] The Pierre Auger Observatory, “Measurement of the fluctuations in the number of muons in extensive air showers with the Pierre Auger Observatory”, Accepted for publication in PRL;

[156] The Pierre Auger Collaboration, “A thinning method using weight limitation for air-shower simulations”, Astropart.Phys. (15), 2001, 259-273;

[157] S. Argirò et al., “The offline software framework of the Pierre Auger Collaboration, Nucl. Instrum. Meth. A (580), 2007, 1485, arXiv:0707.1652;

[158] I. C. Maris, F. Schussler, R. Ulrich, M. Unger, “ADST and EventBrowser Reference Manual: Data Summary Trees and Shower Visualization for Reconstructed and Simulated Auger Events”, Auger Internal Note, 2010, GAP-2006-081;

[159] R. Brun, F. Rademakers, “ROOT - n object oriented data analysis framework”, Nucl. Instrum. Meth. A (389), 1997, 81 – 86;

[160] L. Prado et al., “Simulation of the fluorescence detector of the Pierre Auger Observatory”, Nucl. Instrum. Meth. A (545), 2005, 632;

[161] S. Agostinelli et al., “GEANT4 – a simulation toolkit”, Nucl. Instrum. Meth. A (506), 2003, 250;

[162] Alan Coleman, “Measurement of the Cosmic Ray flux above 100 PeV at the Pierre Auger Observatory”, PhD thesis, Pennsylvania State University;

[163] Fabian Schüssler, “Measurement of the Energy Spectrum of Ultra-High Energy Cosmic Rays using Hybrid Data of the Pierre Auger Observatory”, PhD thesis, Karlshruhe Institute of Technology;

[164] The Pierre Auger Collaboration, “Depth of Maximum of Air-Shower Profiles at the Auger Observatory: Measurements at Energies above $10^{17.8}$ eV”, Phys. Rev. D (90), 2014, 122005;

[165] C.Wileman, PhD thesis, University of Leeds (2008);

[166] The Telescope Array Collaboration, “Upper limit on the flux of photons with energies above $10^{19}$ eV using the Telescope Array surface detector” Phys. Rev. D (88), 2013, 112005;

[167] G. Ros et al., “A new composition-sensitive parameter for Ultra-High Energy Cosmic Rays”, Astopart. Phys. (35), 2011, 140;

[168] N. Krohm and C. Bleve et al., Auger Internal Note, GAP 2012-059 (2012);

[169] Dimitri P. Bertsekas, "Constrained Optimization and Lagrange Multiplier Methods", Academic Press, 1982, 95 – 157;

[170] R. U. Abbasi et al. [Telescope Array Collaboration], Astropart. Phys. (110), 2019, 8, arXiv:1811.03920;

[171] F. Nerling, J. Blumer, R. Engel, M. Risse, "Universality of electron distributions in high-energy air showers - description of Cherenkov light production", Astropart. Phys. (24), 2006, 421 - 437;

[172] The Pierre Auger Collaboration, "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", Phys. Rev. D (96), 2017, 122003;

[173] A. Hoecker et al., "TMVA User guide", CERN, Geneva, 2013.

[174] D. Freedman, R. Pisani, R. Purves, "Statistics (international student edition)", Pisani, R. Purves, 4th edition (2007), 125-130;

[175] T. Pierog, R. Engel, D. Heck, "Impact of Uncertainties in Hadron Production on Air-Shower Predictions", Czech.J.Phys. 56 (2006) A161-A172;

[176] Fisher, R. A., "The use of Multiple Measurements in Taxonomic Problems", Annals of Eugenics. 7 (2), 179 – 188;

[177] L. Lista, "Statistical Methods for Data Analysis in Particle Physics", Springer, 2016, ISBN 978-3-319-20176-4;

[178] B. Efron, C. Stein, "The Jackknife Estimate of Variance", Ann. Statist. 9(3), 1981, 586 – 596;

[179] W. A. Rolke et al., "Limits and Confidence Intervals in the Presence of Nuisance Parameters", Nucl.Instrum.Meth. A(551), 2005, 493 – 503;

[180] The Pierre Auger Collaboration, "Upper limit on the cosmic-ray photon fraction at EeV energies from the Pierre Auger Observatory", Astropart. Phys. (31), 2009, 399 – 406;

[181] T. Bergmann et al., "One-dimensional hybrid approach to extensive air shower simulation", Astropart. Phys. (26), 2007, 420 – 432.

[182] The Pierre Auger Collaboration, " The exposure of the hybrid detector of the Pierre Auger Observatory", Astropart. Phys. (34), 2011, 368.

[183] Vladimír Novotný for the Pierre Auger Collaboration, "Measurement of the spectrum of cosmic rays above $10^{16.5}$ eV with Cherenkov–dominated events at the Pierre Auger Observatory", PoS(ICRC2019)374;

[184] L. Lyons, "Open statistical issues in particle physics", Annals of Applied Statistics (2), 2008, 887 – 915;

[185] The Pierre Auger Collaboration, "Search for photons with energies above $10^{18}$ eV using the hybrid detector of the Pierre Auger Observatory", JCAP09(2020)E02;

[186] V. Berezinsky, et al., "Diffuse radiation from cosmic ray interactions in the galaxy", Astropart.Phys. (1), 1993, 281 – 288;

[187] A. Albert et al., "Evidence of 200 TeV photons from HAWC J1825-134", Astrophys. J. Lett. (907), 2021, L30;

[188] Julian Rautenberg for the Pierre Auger Collaboration, "Limits on ultra-high energy photons with the Pierre Auger Observatory", PoS(ICRC2019)398;

[189] Telescope Array Collaboration, "Constraints on the diffuse photon flux with energies above $10^{18}$ eV using the surface detector of the Telescope Array experiment", Astropart. Phys. (110), 2019, 8;

[190] K.-H. Kampert et al., "Ultra-High Energy Photon and Neutrino Fluxes in Realistic Astrophysical Scenarios", Proceedings ICRC (2), 2011, 198;

[191] J. R. Ellis, et al., "Confinement of fractional charges yields integer charged relics in string models", Phys. Lett. B (247), 1990, 257;

[192] J. R. Ellis, et al., "Astrophysical constraints on massive unstable neutral relic particles", Nucl. Phys. B 373 (1992) 399;

[193] V. Berezinsky, M. Kachelriess and A. Vilenkin, "Ultrahigh-energy cosmic rays without GZK cutoff", Phys. Rev. Lett. (79), 1997, 4302;

[194] D. J. Chung, E. W. Kolb and A. Riotto, "Superheavy dark matter", Phys. Rev. D (59), 1998, 023501;

[195] V. Kuzmin and V. Rubakov, "Ultrahigh-energy cosmic rays: A Window to postinflationary reheating epoch of the universe?", Phys. Atom. Nucl. 61 (1998) 1028 [astro-ph/9709187];

[196] M. Garny, M. Sandora and M. S. Sloth, "Planckian Interacting Massive Particles as Dark Matter", Phys. Rev. Lett. (116), 2016, 101302;

[197] D. Heck and T. Pierog, "Extensive Air Shower Simulation with CORSIKA: A User's Guide", Institut für Kernphysik, Karlsruher Institut für Technologie (KIT), 2013.

[198] The Pierre Auger Collaboration, "Documentation of the Auger $\overline{\text{Off}}\underline{\text{line}}$ Software Framework", 2014, https://www.auger.unam.mx/AugerWiki/OfflineSoftware.