



MEASUREMENT OF THE MUON ATMOSPHERIC PRODUCTION DEPTH WITH THE WATER CHERENKOV DETECTORS OF THE PIERRE AUGER OBSERVATORY

Laura Molina Bueno

Universidad de Granada

1st September 2015

Advisor:
Prof. Antonio Bueno Villar

Co-advisor:
Dr. Sergio Pastor Carpi

Departamento de Física Teórica y del Cosmos & CAFPE

Editorial: Universidad de Granada. Tesis Doctorales
Autora: Laura Molina Bueno
ISBN: 978-84-9125-244-3
URI: <http://hdl.handle.net/10481/40899>

La doctoranda Laura Molina Bueno y los directores de la tesis Antonio Bueno Villar y Sergio Pastor Carpi garantizamos, al firmar esta tesis doctoral, que el trabajo ha sido realizado por la doctoranda bajo la dirección de los directores de la tesis y hasta donde nuestro conocimiento alcanza, en la realización del trabajo, se han respetado los derechos de otros autores al ser citados, cuando se han utilizado sus resultados o publicaciones.

Granada, 1 de septiembre de 2015.

Directores de la tesis

Doctoranda

Fdo: Antonio Bueno Villar, Sergio Pastor Carpi

Laura Molina Bueno

D. Antonio Bueno Villar, Catedrático de Universidad,

CERTIFICA: que la presente tesis doctoral, MEASUREMENT OF THE MUON ATMOSPHERIC PRODUCTION DEPTH WITH THE WATER CHERENKOV DETECTORS OF THE PIERRE AUGER OBSERVATORY, ha sido realizada por D^a Laura Molina Bueno bajo su dirección en el Dpto. de Física Teórica y del Cosmos de la Universidad de Granada, así como que ésta ha realizado una estancia en el extranjero por un periodo superior a tres meses en la *Universidad Pierre y Marie Curie-LPNHE*.

Granada, 1 de septiembre de 2015

Fdo: Antonio Bueno Villar

Contents

1	Ultra High Energy Cosmic Rays	1
1.1	Features of the cosmic ray spectrum	1
1.1.1	The Knee	2
1.1.2	The Ankle	3
1.1.3	The end of the cosmic ray energy spectrum	3
1.2	Origin of cosmic rays	4
1.3	UHE photon and neutrino limits	5
1.4	Arrival direction distribution	6
1.5	Extensive air showers	8
1.5.1	Heitler model of electromagnetic showers	8
1.5.2	Extension of the Heitler model to hadronic showers	10
1.6	Mass composition of UHECR	11
2	The Pierre Auger Observatory	15
2.1	Surface Detector	16
2.1.1	Calibration	17
2.1.2	Trigger	18
2.2	Fluorescence Detector	21
2.2.1	Trigger	22
2.2.2	Calibration	23
2.2.3	Fluorescence yield	23
2.2.4	Atmospheric monitoring	23
2.2.5	FD reconstruction	25
3	SD event reconstruction	31
3.1	Event selection	31
3.2	Geometry reconstruction	32
3.3	Reconstruction of the lateral signal distribution	32
3.3.1	Signal variance	33
3.3.2	Angular resolution	40
3.3.3	LDF maximum likelihood fit	42
3.4	Energy estimation	43

4	Revision of the Muon Arrival Time Model	49
4.1	MPD algorithm	49
4.2	Classic (previous) kinematic delay parameterization	53
4.3	New approach	58
5	Muon Production Depth in an event by event basis	65
5.1	MPD features	65
5.1.1	Motivation of X_{\max}^{μ} as mass composition observable	66
5.2	MPD reconstruction for individual events	68
5.2.1	Breakdown of the different contributions to X_{\max}^{μ} bias	71
5.2.2	Final bias	84
5.3	Contributions to X_{\max}^{μ} resolution	85
5.4	Application to SD Data	86
6	Average MPD analysis	91
6.1	Reconstruction of the average MPD distribution	91
6.2	MC studies	92
6.2.1	Bias contributions to the average X_{\max}^{μ}	92
6.2.2	Deconvolution	95
6.2.3	Final bias and shape after all the reconstruction chain	99
6.3	Application to SD data	104
6.3.1	Systematic uncertainties	104
6.3.2	Estimation of cosmic ray mass composition through the mean logarithmic mass, $\ln A$	108
7	New detector configurations and potential improvements of MPD studies	115
7.1	Implementation of new configurations in Offline software	115
7.2	Resolution of the maximum of the MPD distributions	121
	Conclusions and future prospects	123
	Conclusiones y perspectivas futuras	127
A	Example of outliers	131
A.1	1500 m Array and different angular ranges	132
	List of figures	135
	List of tables	142
	Bibliography	154

Introduction

Ultra-high-energy cosmic rays (UHECR) are particles of uncertain origin and composition, with energies above 1 EeV (10^{18} eV or 0.16 J). The measured flux of UHECR is a steeply decreasing function of energy. Above 1 J, we expect to collect one event per year per km^2 per steradian. This low flux makes it impossible to detect them above the atmosphere. This kind of extremely energetic primary particle thus gives rise to huge shower containing billions of daughter particles. We have no first-hand access to the identity of the primary particle and are therefore forced to build huge arrays of particle detectors at the ground if we want to study the nature of this non-thermal sort of radiation that continuously bombards the Earth's atmosphere. The fact that we can only record and study those secondary particles, customarily known as Extensive Air Showers (EAS), add extra difficulties in our quest to understand what are UHECR, where they are produced and what mechanisms are at work to deliver such extraordinary energies, which are far from being matched by any man-made particle accelerators.

The field of UHECR is therefore not short in supply of unanswered questions. Not surprisingly it was and continues to be a very active field, where large international collaborations assemble to understand the physics behind this extreme manifestation of the non-thermal Universe. The largest and most sensitive apparatus built to date to record and study EAS is the Pierre Auger Observatory. Covering 3000 km^2 it was devised to reveal the nature of charged cosmic rays thanks to the simultaneous use of two detection techniques: the detection of fluorescence light and the sampling of the particles that reach the ground. It is thus a hybrid detector with improved capabilities since its calibration is data-driven and for this purpose it does not rely on cumbersome simulations affected by large uncertainties.

The Pierre Auger Observatory has produced the largest and finest amount of data ever collected for UHECR. A broad physics program is being carried out covering all relevant topics of the field. Among them, one of the most interesting is the problem related to the estimation of the mass composition of cosmic rays in this energy range. Currently the best measurements of mass are those obtained by studying the longitudinal development of the electromagnetic part of the EAS with the Fluorescence Detector. However, the collected statistics is small, specially at energies above several tens of EeV. Although less precise, the volume of data gathered with the Surface Detector is nearly a factor ten larger than the fluorescence data. So new ways to study composition with data collected at the ground are under investigation.

The subject of this thesis follows one of those new lines of research. Using prefer-

entially the time information associated with the muons that reach the ground, we try to build observables related to the composition of the primaries that initiated the EAS. A simple phenomenological model relates the arrival times with the depths in the atmosphere where muons are produced. The experimental confirmation that the distributions of muon production depths (MPD) correlate with the mass of the primary particle was done in [1]. This opened the way to a variety of studies of which this thesis is a continuation of the original work with the aim of enlarging and improving its range of applicability.

This document is organized as follows: chapter 1 contains introductory text to the most important milestones reached in cosmic ray physics. In chapter 2 the Pierre Auger Observatory and the main features of this hybrid detector are described. Since this thesis is based on the analysis of the data registered by the Surface Detector, we discuss in depth in chapter 3 how the properties of the primary cosmic rays are reconstructed using the information provided by the Water-Cherenkov Detectors (WCD) [2, 3]. In chapter 4 we revisit the phenomenological model which is at the root of the analysis and discuss a new way to improve some aspects of the model [4]. In chapter 5 we carried out a thorough revision of the original analysis with the aim to understand the different contributions to the total bias and resolution when building MPDs on an event-by-event basis [5, 6]. Chapter 6 is focused on an alternative way to build MPDs: we consider average MPDs for ensembles of air-showers with the aim of enlarging the range of applicability of this kind of analyses. Finally, in chapter 7 we analyze how different improvements in the WCD electronics and its internal configuration affect the resolution of the MPD [7]. We conclude summarizing the main results and discussing potential ways to improve MPD-based mass composition studies.

Introducción

Los rayos cósmicos de ultra alta energía (UHECR de sus siglas en inglés) son partículas de origen y composición incierta cuyas energías se encuentran por encima de 1 EeV (10^{18} eV o 0.16 J). Su flujo es una función fuertemente decreciente con la energía. Por encima de 1 J, esperamos medir una de dichas partículas por año, km^2 y esterreo radián. Esto hace impensable detectar de manera directa en las capas altas de la atmósfera estas partículas primarias de energías extremas, antes de que interaccionen y den lugar a cascadas de billones de partículas secundarias. En definitiva no tenemos información directa sobre la partícula primaria y por eso estamos obligados a construir enormes sistemas de detectores en el suelo si pretendemos estudiar la naturaleza de este tipo de radiación no térmica que continuamente bombardea la atmósfera terrestre. El hecho de que sólo podamos estudiar los productos de la interacción, conocidos habitualmente como Cascada Extensa de Partículas (EAS), añade dificultades extra a nuestra investigación para conocer qué son los UHECR, dónde se producen y qué mecanismos le confieren tan extraordinarias energías, las cuales somos incapaces de alcanzar con la tecnología actual de aceleradores de partículas.

El campo de los UHECR no está exento de poseer un buen número de preguntas sin respuesta. Por tanto no es sorprendente que dicho campo fuera y siga siendo muy activo, en el cual grandes colaboraciones internacionales trabajan unidas para entender los misterios de estas manifestaciones extremas del Universo no térmico. El detector más grande y sensible hasta ahora construido para el estudio de los EAS es el Observatorio Pierre Auger. Cubre una superficie de 3000 km^2 y fue diseñado para revelar los secretos de los rayos cósmicos cargados mediante el uso de dos técnicas de detección: la medida de la luz de fluorescencia producida en la atmósfera y la detección de una parte de las partículas que llegan al suelo. Es por tanto un detector híbrido capaz de realizar calibraciones a partir de los datos experimentales recogidos. Esto reduce los sistemáticos asociados pues no dependen de complicadas simulaciones plagadas de grandes incertidumbres.

Pierre Auger ha sido capaz de recoger el conjunto de datos más grande y de mejor calidad en la historia de los UHECR. Gracias a ello se está llevando a cabo un amplio programa de física que cubre los asuntos más relevantes del campo. Entre las líneas de investigación más interesantes se encuentra el estudio de la composición en masas de los rayos cósmicos. Actualmente las mejores inferencias en lo que a masas se refiere son las que se obtienen a través de las medidas hechas con el Detector de Fluorescencia. Este estudia el desarrollo longitudinal de la parte electromagnética de la cascada. Sin embargo el conjunto de datos recogidos no es muy grande, especialmente a las más altas energías

que es la zona de mayor interés. Aunque de menor precisión, el conjunto de los datos del Detector de Superficie es casi un factor diez más grande que el conjunto de datos de fluorescencia. Por tanto nuevas formas de inferir las masas de los primarios se están desarrollando basándose en la información dada por los detectores de superficie.

Esta tesis sigue una de esas nuevas líneas de investigación. Usando principalmente la información temporal de los muones detectados en el suelo, tratamos de construir observables físicos relacionados con la composición del primario que inició la cascada. Un modelo fenomenológico simple relaciona los tiempos de llegada con las profundidades atmosféricas a las que se producen los muones. La confirmación experimental de que las distribuciones de producción de muones (MPD) están correlacionadas con la masa de la partícula primaria fue hecha por primera vez en [1]. Este trabajo abrió nuevas líneas de estudio, siendo esta tesis una continuación de ese trabajo original con el objetivo de ver si podemos ampliar y mejorar el rango de aplicación de esta técnica.

El presente documento se organiza como sigue: el capítulo 1 contiene una somera descripción de los hitos más importantes alcanzados en la historia de la Física de Rayos Cósmicos. El capítulo 2 explica qué es y de qué partes está compuesto el Observatorio Pierre Auger. Ya que esta tesis está basada en los datos registrados por el Detector de Superficie, en el capítulo 3 discutimos en profundidad cómo se reconstruyen los sucesos a partir de la información que proporcionan los detectores Cherenkov de agua (WCD) [2, 3]. En el capítulo 4 revisamos el modelo fenomenológico que constituye la base de este trabajo e introducimos formas de mejorar algunos aspectos del mismo [4]. Una revisión completa del análisis original aplicado a sucesos individuales se lleva a cabo en el capítulo 5 con el fin de evaluar cuáles son las diferentes fuentes que contribuyen al sesgo y la resolución en la medida de los máximos de las MPD reconstruidas [5, 6]. El capítulo 6 se centra en una forma alternativa de usar MPD: construimos MPD promedio para conjuntos de sucesos de energías similares con el fin de aumentar el rango de aplicabilidad de este tipo de análisis. Finalmente en el capítulo 7 presentamos posibles mejoras al análisis de las MPD mediante la mejora de la electrónica y/o la estructura interna de los WCD [7]. Este documento se cierra con las conclusiones más importantes de los estudios realizados y unas indicaciones sobre posibles líneas de trabajo para el futuro.

Ultra High Energy Cosmic Rays

1

Astroparticle Physics started more than 100 years ago with the seminal work of V. Hess [8]. Nowadays one of the biggest mysteries of this field is associated to cosmic rays of ultra-high energy. Their energy is several orders of magnitude larger than the limits reached in particle accelerators (a cosmic ray of 10^{17} eV produces an energy in center of mass equivalent to the 14 TeV reached at LHC with proton-proton collisions). As a consequence, when drawing conclusions from observations the main subject of uncertainty comes from the theoretical models used which are extrapolations of much lower-energy data. We are still unable to identify unequivocally the source of production, mechanisms of acceleration or chemical composition of UHECR [9, 10, 11, 12, 13].

The exploration of cosmic rays began as a mixture of physics and environmental studies more than a century ago. After the discovery of natural radioactivity in 1896 by Becquerel [14], it was noticed that between 10 and 20 ions were generated per cubic centimeter of air every second. The main question that physicists wanted to address was if this ionization comes from the Earth. However, Victor Hess in 1912 [8], by flying ionization chambers in a series of balloon flights, found that the ionization of air at 5000 m was more than twice that at sea-level. Kolhörster measured an even higher ionization rate in another flight. These measurements demonstrated that this radiation comes from outside the Earth. The term *cosmic rays* was coined by R.A. Millikan in 1926 given that gamma rays were the most penetrating radiation known at that moment [15].

In 1930s Rossi, Schmeiser, Bothe, Kolhörster and Auger realized that these particles reaching the ground were correlated in time, leading to the discovery of EAS. They were given this name because when the cosmic ray interacts in the atmosphere generates a subsequent shower of particles which can be detected at the ground covering several square kilometers. The first claim for an EAS detection above 10^{20} was done by Linsley in 1963 at Volcano Ranch [16]. The details of how EAS evolve in the atmosphere will be discussed in a forthcoming section.

1.1 Features of the cosmic ray spectrum

The energy spectrum of cosmic rays (primary particle flux as a function of energy, J) can be described as a rather featureless power law function extending from energies around

10^9 eV up to 10^{20} eV. The exponent of this power law (the *spectral index*), is close to 3 ($J \propto \frac{dN}{dE} \propto E^{-3}$). The flux falls by 25 orders of magnitude over 11 decades of energy. This behaviour is expected in the case of stochastic acceleration of charged particles at astrophysical shocks [17].

Figure 1.1 shows the cosmic ray flux (number of particles per unit of area, energy, solid angle and time) as measured by different experiments. Three regions of the spectrum exhibit a particularly interesting deviation from the average behavior, the “Knee” (at $\sim 10^{15}$ eV), the “Ankle” (at $\sim 10^{18}$ eV) and the highest region of the spectrum. The properties of these regions will be explored in the following sections. In order to see clearly the different features the flux is multiplied by $E^{2.65}$. A change in the acceleration mechanism, or in the the propagation processes through the medium or in the hadronic interaction cross section with increasing energy can explain the different spectral features.

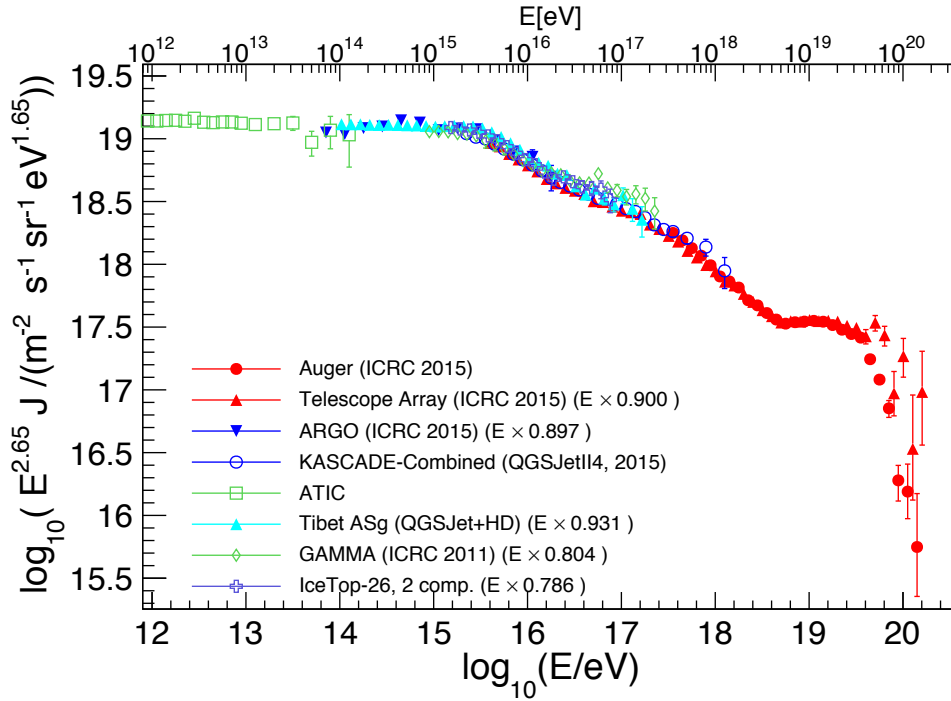


Figure 1.1: Spectrum of cosmic rays measured by different experiments [18].

1.1.1 The Knee

The knee is the point where the spectral index changes from ≈ 2.7 to ≈ 3.1 (at $\sim 10^{15}$ eV) (see [19] for more details). There, the flux of particles is about one per square meter and per year. Several experiments have confirmed the existence of this change in the spectrum: Yakutsk [20] and Akeno [21], KASCADE [22] and its extension KASCADE-Grande [23]

In some models this feature is related to acceleration processes. It is assumed that cosmic rays are accelerated in shock fronts from supernova explosions and the drop in the spectrum occurs when it reaches the maximum energy. In other models the knee is caused

by the leakage of cosmic rays from our Galaxy or to interactions with background particles in the Galaxy.

Another scenario consider that the knee is the result of a change of the hadronic interactions at these energies [24]. In this case this feature is not a characteristic of the spectrum itself, it is the result of its observation at Earth.

1.1.2 The Ankle

The ankle corresponds to a flattening of the spectrum to a spectral index which is again close to $\gamma \approx 2.7$ at an energy $\sim 10^{18}$ eV. This change in the spectrum was clearly observed by the Fly's Eye experiment [25], and different explanations are assumed to be the responsible of this feature.

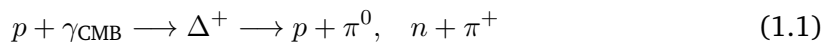
Some models associate this change to the transition from galactic to extragalactic origin of the cosmic rays [26, 27]. In the *ankle model* the extragalactic component is thought to have a pure proton composition. The ankle is the point where the two components contribute equally to the total flux. However, in the *mixed composition model* [28] is favoured the possibility of a mixed composition above 10^{19} eV.

On the other hand, the *dip model* [29, 30, 31], predicts that the extragalactic component, composed mainly by protons, dominates at much lower energies. Thus, in the ankle region the galactic component already disappeared. The change in the spectral index is just a propagation effect: the interaction of protons with the cosmic microwave background (CMB) produces an electron-positron pair and the energy of the primary is reduced. This causes a suppression of the flux at high energies, increasing it at low ones, and therefore causing the appearance of the ankle.

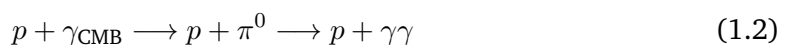
1.1.3 The end of the cosmic ray energy spectrum

The spectrum measured by the Pierre Auger Observatory has established the suppression of the cosmic ray flux at 4×10^{19} eV unambiguously [32]. However, why it occurs is still a mystery.

Greisen [33], and independently Zatsepin and Kuz'min [34] predicted in 1966 after the discovery of the CMB by Penzias and Wilson [35] that the energy spectrum should steepen near 4×10^{19} eV if the sources are distributed uniformly throughout the Universe. This effect is commonly known as the GZK suppression. Cosmic rays lose energy on their path to Earth interacting with the CMB. The reaction is:



The proton loses about 20% of its energy in this process and it limits the distance from which a high-energy particle have traveled before detection to ~ 100 Mpc (see figure 1.2). The pions resulting from the GZK reaction are producing photons and neutrinos of very high energy:



resulting in two cosmogenic photons, and the second one

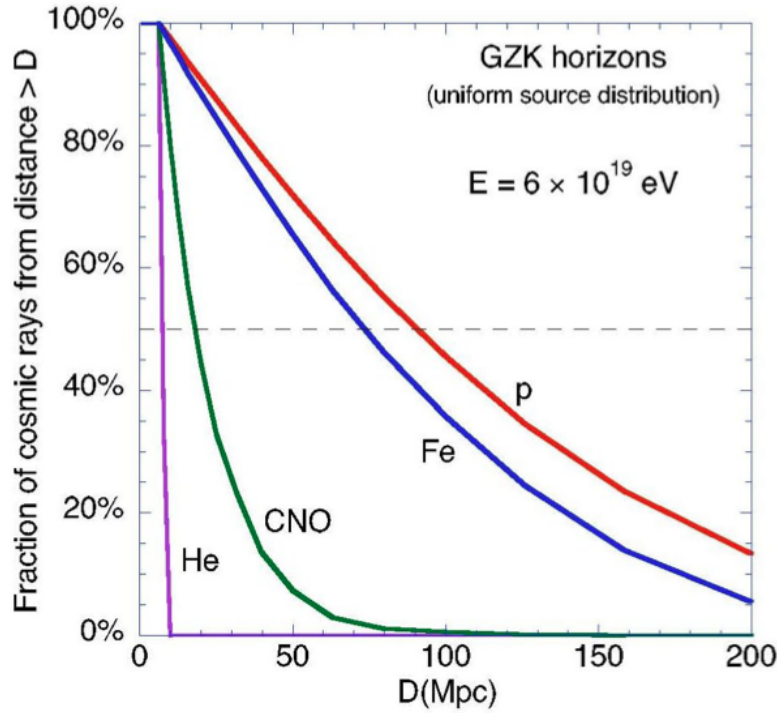
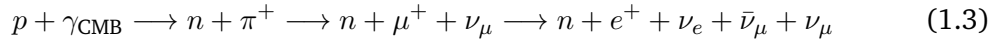


Figure 1.2: Illustration of the survival of nuclei $> 6 \times 10^{19}$ eV as a function of distance. Effectively only proton and iron nuclei can reach Earth from sources more than ~ 50 Mpc away [12].



which produces three neutrinos in the final state.

Nevertheless, the suppression does not imply that the GZK is present. The *disappointing model* [36] attributes this suppression to the maximum energy to which sources can accelerate particles. Therefore the suppression is still an open question awaiting for a definitive answer.

1.2 Origin of cosmic rays

Most of the low energy cosmic rays, up to some GeV, come from the Sun. However, at higher energies they must have another origin as there is not any process involving such a huge energy taking place in the Sun. In a synchrotron the maximum energy reached depends on the magnetic field and the radius, $E_{max} = ZeBR\beta c$ (Z is the charge of the nucleus, e is the charge of the electron, B is the magnetic field in a region of radius R and βc is the velocity of the particle). In a single-shot acceleration process, such as might occur near a neutron star, a high voltage could be generated between the pole and the equator: the maximum energy can be described with the same equation. In diffusive shock

acceleration essentially the same relationship can be applied: here $E_{max} = kZeBR\beta c$ with the constant $k < 1$ and β denoting the speed of the shock. According to this condition, the so called *Hillas plot* indicates the sites where particles can be accelerated (see figure 1.3) [37]. Very few sites can generate particles with energies around or above 100 EeV: either this occurs on highly condensed objects with huge B or enormously extended objects.

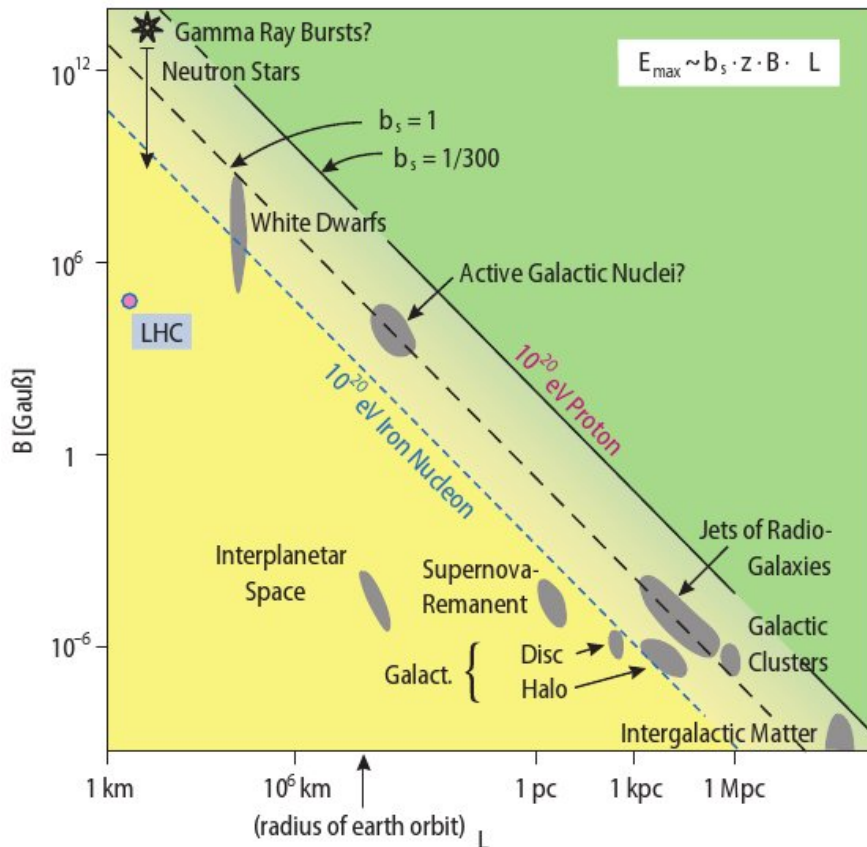


Figure 1.3: The Hillas plot represents the strength of the magnetic field versus the size of possible candidates for ultra-high energy cosmic ray acceleration. Objects below the diagonal can be ruled out.

1.3 UHE photon and neutrino limits

As we pointed out in section 1.1.3, in the GZK reactions photons and neutrinos of very high energy can be produced. The discovery of them would favour a light cosmic ray composition at ultra-high energies. This is due to the fact that photo-pion production for heavy primaries has a much higher energy threshold.

In figure 1.4 are shown the latest limits on the fluxes of photons [38, 39, 40, 41] and neutrinos [42, 43, 44, 45] including observations from the Pierre Auger Observatory [46], Anita-II [47] and IceCube [48]. Model scenarios for sources of UHECR like the

top-down, in which the observed particles are produced by the decay of extremely heavy particles are strongly disfavoured by the actual limits. Moreover, the flux limits begin to reach the predicted secondary fluxes for models in which the suppression of cosmic ray flux is originated by the GZK energy loss process for a proton dominated flux [49, 50, 51, 11].

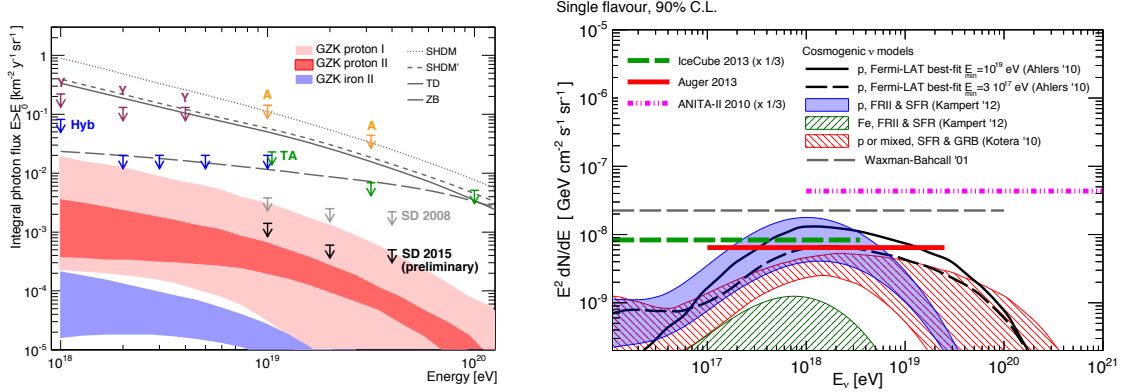


Figure 1.4: Limits on the flux of photons (left) [38, 39, 40, 41] and neutrinos (right) [52] obtained from the Pierre Auger Observatory. The data are shown together with the current limits from other experiments [53, 54, 55, 56, 57, 58, 59] and some examples of predicted fluxes.

1.4 Arrival direction distribution

The arrival direction distribution is one of the key observables to search for the transition from galactic to extragalactic cosmic rays, and for sources or source regions of UHECR.

The Pierre Auger collaboration has observed a dipole structure on a large scale above an energy of $E > 8 \times 10^{18}$ eV [60]. The amplitude of the first harmonic in right ascension was determined to be $r_{1\alpha} = (4.4 \pm 1.0) \times 10^{-2}$, having a chance probability of $P(\geq r_{1\alpha}) = 6.4 \times 10^{-5}$. Under the assumption that the only significant contribution to the anisotropy is from the dipolar component, this observation would correspond to a dipole of amplitude $d = 0.073 \pm 0.015$ pointing to $(\alpha, \delta) = (95^\circ \pm 13^\circ, -39^\circ \pm 13^\circ)$. The origin of this anisotropy is subject to ongoing discussions. It could arise, for example, from an inhomogeneity of the distribution of nearby sources, see [61].

Applying the Rayleigh [65] and the differential East-West [66] methods to the arrival directions of cosmic rays observed with the Pierre Auger Observatory [63] the measured amplitude of the dipole (as well as the corresponding upper limit) and its phase angle have been determined. They are shown in Fig. 1.5 as a function of energy.

The phase angle exhibits a smooth change with energy. It points near the Galactic center below 10^{18} eV, suggesting component with a galactic origin. The phase angle points in the opposite direction at higher energy, possibly manifesting a signature of the inhomogeneous distribution of nearby extragalactic matter. Given that the phase angle is statistically more sensitive than the dipole amplitude [65], a prescribed test to determine the statistical significance of the observed transition in the phase is being performed, and will run

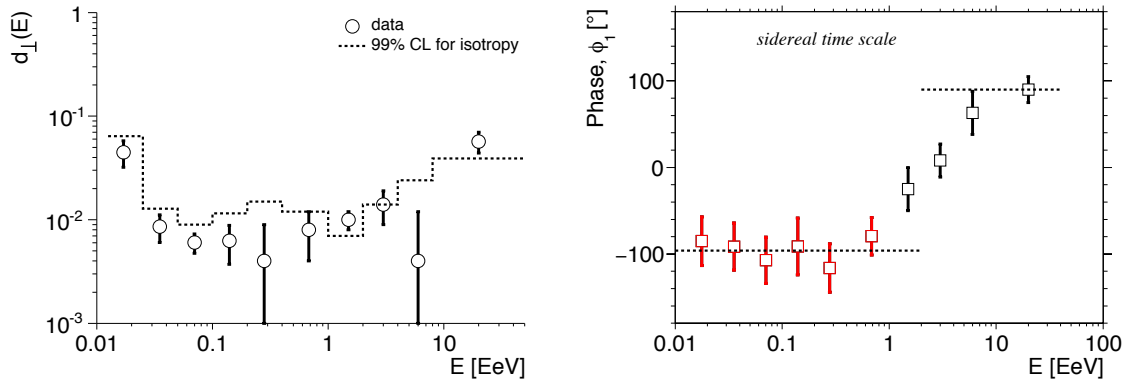


Figure 1.5: Large scale anisotropy search. Left: 99% limits on the dipole anisotropy in the equatorial plane for the collected statistics until end of 2014 (dashed line) and values of the dipole amplitude d_{\perp} . Right: estimated phase angles ϕ_1 . The red points of the equatorial phase are from the analysis of the infill array. The data shown [62] is an update of the analyses [63, 64].

until the end of 2015.

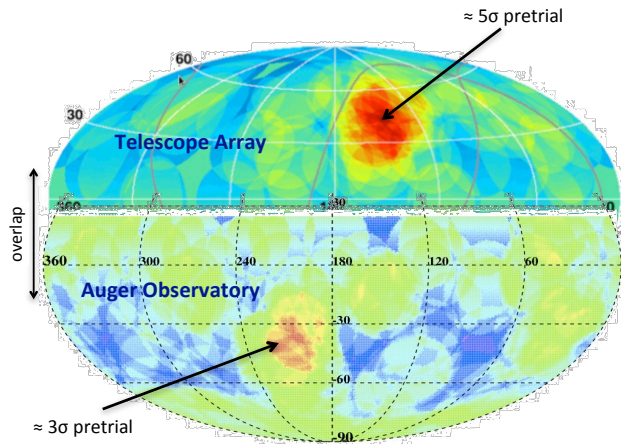


Figure 1.6: Regions of over-density observed after $\sim 20^\circ$ -smearing of the arrival directions of particles with $E > 5.5 \times 10^{19}$ eV. The results from the northern hemisphere are from the TA Collaboration [67].

Up to now it has not been possible to establish small-angle correlations of the arrival direction distribution of Auger data with possible sources or source regions beyond any doubt [68], even though there are some intriguing indications [69, 70, 71]. Here we only want to mention the 20° region of over-density observed around the direction of Centaurus A [68]. Although not being a statistically significant excess beyond 3σ , it is interesting to note that the TA Collaboration has recently reported a “hot spot” of similar intermediate angular scale [67], see figure 1.6.

1.5 Extensive air showers

When a cosmic ray interacts with an atom of the atmosphere a cascade of secondary particles is produced (i.e. an EAS is created). The generated number of particles is a vast number: about 10^{10} particles for events having 10^{19} eV. Extensive air showers can be described as the superposition of different components (see figure 1.7). The most important ones are the hadronic, muonic and electromagnetic cascades. Photons and electrons/positrons represent 99% of these particles and they transport 85% of the total energy. They initiate what is called the *electromagnetic cascade*. The remaining particles are muons, pions and in smaller proportion neutrinos and baryons.

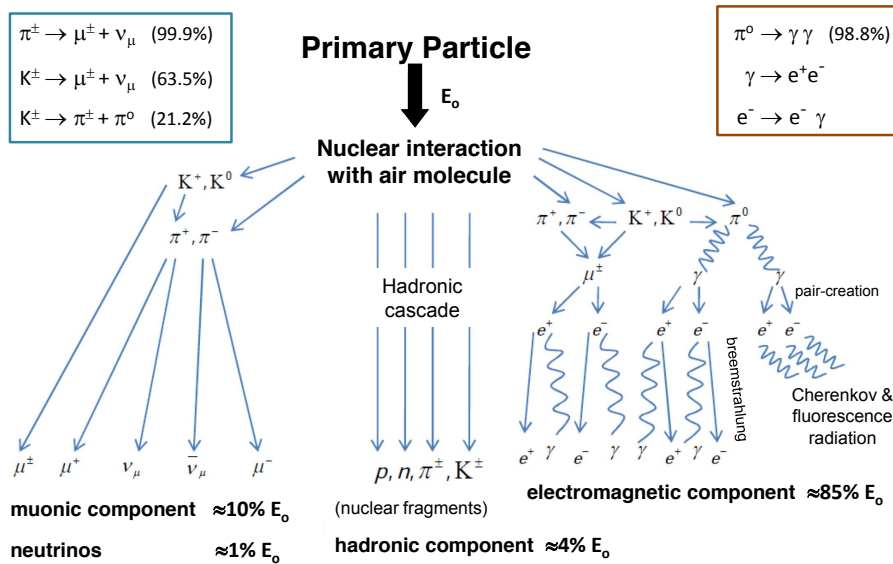


Figure 1.7: Main components of extensive air showers.

The basic properties of the development of a cascade can be described with a model developed by Heitler [72]. This model was extended by Matthews [73] to include the description of hadronic showers.

1.5.1 Heitler model of electromagnetic showers

In this simplified model at each step all particles interact producing two secondary particles of equal energy. Electrons, positrons and photons interact after traveling an interaction length $d = \lambda_r \ln 2$, where λ_r is the radiation length of the medium ($\lambda_r = 37 \text{ g/cm}^2$ in air). After each step electrons divide their energies in half via *bremsstrahlung* emission of a single photon while photons produce a pair e^+e^- of equal energy. After n steps the particle number is $N_n = 2^n$ and their individual energy is E_0/N_n . The process ends when the individual energy drops below the critical value (energy at which the rate of energy

loss via *bremstrahlung* is equal to the rate by ionization), $E_c^\gamma \sim 80$ MeV in air. At this point the electromagnetic cascade reaches its maximum. Figure 1.8 shows an illustration of the whole process.

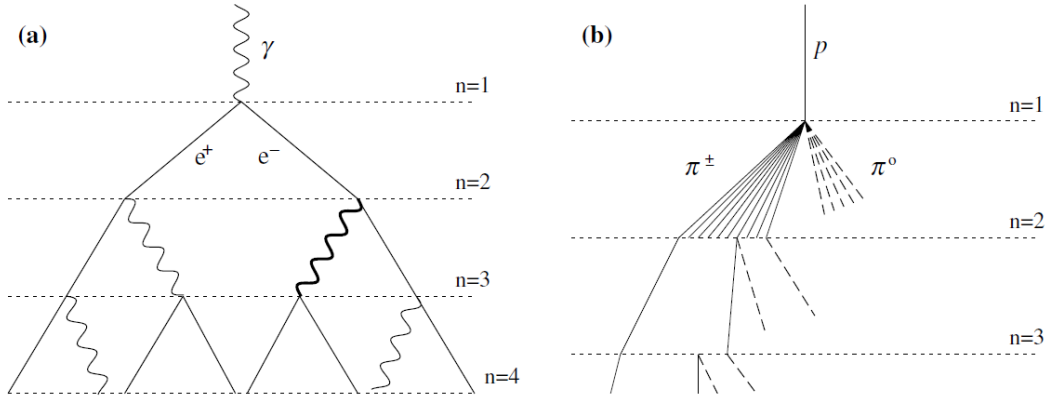


Figure 1.8: Schematic representation of the Heitler model for electromagnetic (a) and hadronic (b) cascade development.

Although the model is conceptually very simple, it correctly describes some of the main features of electromagnetic cascades:

- The number of particles at the shower maximum is proportional to the energy of the primary particle.

$$N_{\max} = E_0/E_c^\gamma \quad (1.4)$$

- The depth of maximum of the shower evolves logarithmically with energy. The depth of maximum shower development, X_{\max} , is determined by the number of interactions needed to reach the critical energy:

$$n_{\max} = \ln(N_{\max})/\ln 2 = \ln(E_0/E_c^\gamma)/\ln 2 \quad (1.5)$$

$$X_{\max} = X_0 + n_{\max}d = X_0 + \lambda_r \ln(E_0/E_c^\gamma) \quad (1.6)$$

where X_0 is the depth of the first interaction.

- The rate of evolution of X_{\max} with energy is called the *elongation rate* and is defined as:

$$D_{10} = \frac{dX_{\max}}{d \log E_0} = \lambda_r \ln 10 \approx 85 \text{ g/cm}^2 \quad (1.7)$$

The results given by the model were confirmed by simulations. However, it presents some discrepancies such as an overestimation of the number of particles by a factor 2-3 and the ratio of electrons and positrons to photons by a factor 10-12. This is due to the fact that absorption of particles below the critical energy is not considered in the model [73]. Also, more than one photon can be emitted during *bremstrahlung* and electrons lose their energy much faster than photons.

1.5.2 Extension of the Heitler model to hadronic showers

In analogy with the electromagnetic cascade, the hadronic component can be described assuming that after each interaction the main products are pions. This was done by Matthews [73] and is thoroughly described in [9]. In this extension, the relevant parameter is the hadronic interaction length λ_I . After each step of thickness $\lambda_I \ln 2$, $2N_\pi$ charged pions are produced and N_π neutral ones. The π_0 will decay and go into the electromagnetic cascade and charged pions interact further producing the hadronic cascade (see figure 1.8b). In this case, the end of the cascade is determined by the energy at which pions start to decay into muons.

The number of muons in the shower is obtained assuming that all pions decay into muons when they reach the critical energy, E_c^π . Thus, $N_\mu = (2N_\pi)^{n_c}$, where $n_c = \frac{\ln(E_0/E_c^\pi)}{\ln(3N_\pi)}$ is the number of steps for the pions to reach E_c^π . Introducing $\beta = \ln 2N_\pi / \ln 3N_\pi$ the number of muons can be written as: $N_\mu = \left(\frac{E_0}{E_c^\pi}\right)^\beta$.

The determination of the position of the shower maximum is more complex in the case of hadronic shower than in the electromagnetic one. The larger cross-section and the larger multiplicity at each step will reduce the value of the maximum and the energy evolution of those quantities will modify the rate of change of X_{\max} with energy. In addition the inelasticity of the interaction will also modify both the position of the maximum and the elongation rate. The energy transfer from the hadronic component to the electromagnetic one at each step and a correct superposition of each electromagnetic sub-shower to compute X_{\max} is needed. A good approximation is the assumption made in [73]. There, the effect of the hadronic cascade is considered only in the first interaction. Therefore, for proton showers

$$D_{10}^p = D_{10}^\gamma + \frac{dX_0}{d \log E_0} \quad (1.8)$$

where D_{10}^γ is the elongation rate for electromagnetic showers and $X_0 = \lambda_I \ln 2$ the depth of the first interaction. Introducing a realistic parameterization of the dependence of λ_I as a function of the energy, such as the one given in [74], the elongation rate is $D_{10}^p \approx 62 \text{ g/cm}^2$. Moreover, since hadronic interaction models predict an approximately logarithmic decrease of λ_I with energy, D_{10}^p is approximately constant.

An important consequence of equation (1.8) was noted by Linsley in 1977: the *Elongation Rate Theorem* [75]. This theorem stipulates that regardless of the particular parameterization of λ_I that is chosen, (1.8) will always decrease with increasing energy, and thus the second term in the equation is always negative. Therefore, the elongation rate for electromagnetic showers is always bigger than the one for hadronic shower.

The extension of this description to air showers initiated by different nuclear primaries can be done with the theoretical framework called *superposition model*. In this model, a primary nucleus of mass A and energy E is described as the superposition of A nucleons of energy $E' = E/A$ (for more details see e.g. [11]). Showers from heavy nuclei will develop higher, faster and with less shower to shower fluctuations than showers initiated by lighter nuclei. From these simple assumptions some of the most important phenomena that are correctly described are:

- Nuclei initiated showers will be on average less penetrating than those generated by protons of the same energy

$$X_{\max}^A(E_0) = X_{\max}^p(E_0/A) = X_{\max}^p(E_0) - \lambda_r A \quad (1.9)$$

- The number of muons is larger for heavier primaries than for light primaries of the same energy

$$N_{\mu}^A(E_0) = \sum_i^A N_{\mu}^p(E_0/A) = N_{\mu}^p(E_0) A^{1-\beta} \quad (1.10)$$

- The elongation rate is the same regardless of the mass of the primary. It shows up as parallel lines in an X_{\max} vs energy plot.

$$D_{10}^A = \frac{dX_{\max}^A}{d \log E_0} = \frac{d(X_{\max}^p - \lambda_r A)}{d \log E_0} = \frac{dX_{\max}^p}{d \log E_0} = D_{10}^p \quad (1.11)$$

- The shower to shower fluctuations of X_{\max} is smaller for heavy nuclei than for light ones.

The superposition model is a simplification and cannot fully describe hadronic EAS, as it does not account for nuclear effects such as re-interaction in the target nucleus or nuclear fragmentation. In order to consider all of these processes and others, more realistic transport codes are used, such as CORSIKA [76], AIRES [77] or COSMOS [78], together with hadronic interaction models like EPOS [79], QGSJET-II [80] or SIBYLL [81] (see e.g. [82] for a comprehensive review of air shower simulations).

1.6 Mass composition of UHECR

The composition of cosmic rays of energies below the knee can be measured directly by space-based experiments. However, at higher energies, the only possibility is to characterise the properties of the EAS generated by different primaries. Nevertheless, due to fluctuations of the properties of the first few hadronic interactions in the cascade, the primary mass can not be measured on an event-by-event basis. It must be inferred statistically from the distribution of shower maxima of a set of air showers. The longitudinal profile of the energy deposit of an air shower as a function of the atmospheric slant depth can be directly measured with fluorescence telescopes. The depth at which the energy deposit reaches the maximum is called X_{\max} . The global X_{\max} distribution results from the superposition of the distributions produced by different nuclei of mass A_i :

$$f(X_{\max}) = \sum_i p_i f_i(X_{\max}) \quad (1.12)$$

where p_i represents the fraction of primary particle of type i . Showers generated by protons have an average value of X_{\max} about 100 g/cm^2 larger than showers produced by iron nuclei of the same energy, so it is possible to do inferences about mass composition with this observable. This was achieved for the first time by the Fly's Eye collaboration

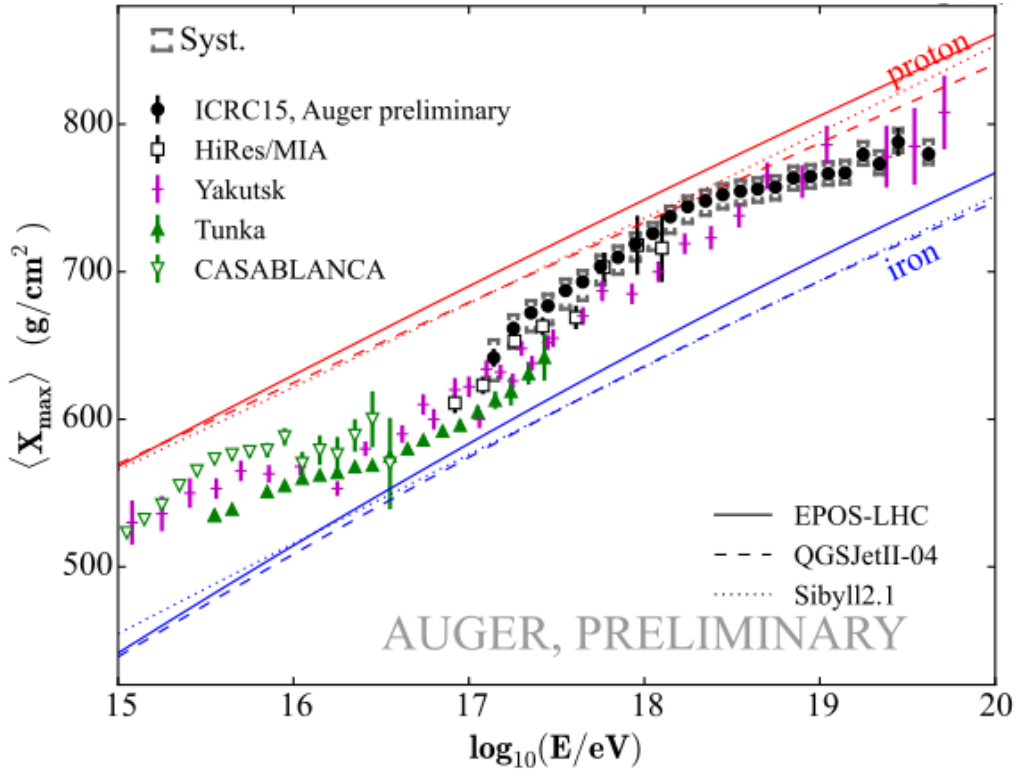


Figure 1.9: Mean of the measured X_{\max} distributions as a function of energy compared to air-shower simulations for proton and iron primaries for different experiments.[85].

[83] and followed by different experiments since then [84]. In figure 1.9 is illustrated the rate of change of the average shower maximum per decade of energy for different experiments. If the composition is pure (i.e. only one primary species) the slope should be constant with energy. However, in the elongation rate one can observe changes which are not compatible with a pure composition. Nevertheless, the interpretation of data values depends on model predictions. At energies above 10^{17} eV models rely on extrapolations from measurements done at LHC energies and therefore are subject to large systematic uncertainties.

Since each experiment has different selection cuts and systematic uncertainties, to allow a straightforward comparison of the results the first two moments of the distribution of X_{\max} (the mean and the shower-to-shower fluctuations) can be related to the first two moments of the distribution of $\ln A$. The transformation of X_{\max} measurements into an average value of $\ln A$ is done using the following equation:

$$\langle \ln A \rangle = \frac{\langle X_{\max} \rangle_p - \langle X_{\max} \rangle_{\text{data}}}{\langle X_{\max} \rangle_p - \langle X_{\max} \rangle_{\text{Fe}}} \ln 56 \quad (1.13)$$

With such a simple transformation, we can compare (in a model dependent way) whether the measurements of different experiments yield compatible estimations for the

mass of UHECR. As a matter of fact the situation concerning mass composition is nowadays unclear and more precise measurements are required. The main subject of this thesis is an attempt to contribute to the ongoing discussion about composition. In what follows we will focus on making inferences about the masses of UHECR using new observables extracted from the data collected by the Surface Detector of the Pierre Auger Observatory.

The Pierre Auger Observatory

2

The Pierre Auger Observatory has been conceived to measure the main properties (flux, arrival direction and mass composition) of cosmic rays from 10^{18} eV up to the largest energies with high statistical significance [86].



Figure 2.1: Map of the Pierre Auger Observatory. Red dots indicate SD detectors and green lines show the field of view of the FD telescopes.

The Observatory is located at the "pampa amarilla" near Malargüe, in the Province of Mendoza (Argentina), at a mean altitude of 1400 m (depth = 879 g/cm^2). It was completed in 2008 and is taking data since 2004. The site is relatively flat and near the base of the Andes mountains. It covers a ground area of 3000 km^2 and contains 1600 water-Cherenkov detectors, the Surface Detector (SD), arranged on a triangular grid, with a 1.5 km separation among detectors, overlooked from 4 sites by optical stations (figure 2.1), the Fluorescence Detector (FD). These stations contain 6 telescopes each, designed to detect air-fluorescence light emitted by atmospheric nitrogen when it is excited by charged particles. The WCD detect particles at the ground (mainly muons, electrons, positrons and

photons) via the Cherenkov light emitted while they cross the ultra-pure water inside the detectors. The precise calorimetric information provided by the FD allows a calibration of the SD array in a data-driven way, thus avoiding the systematics associated with the use of simulated EAS.

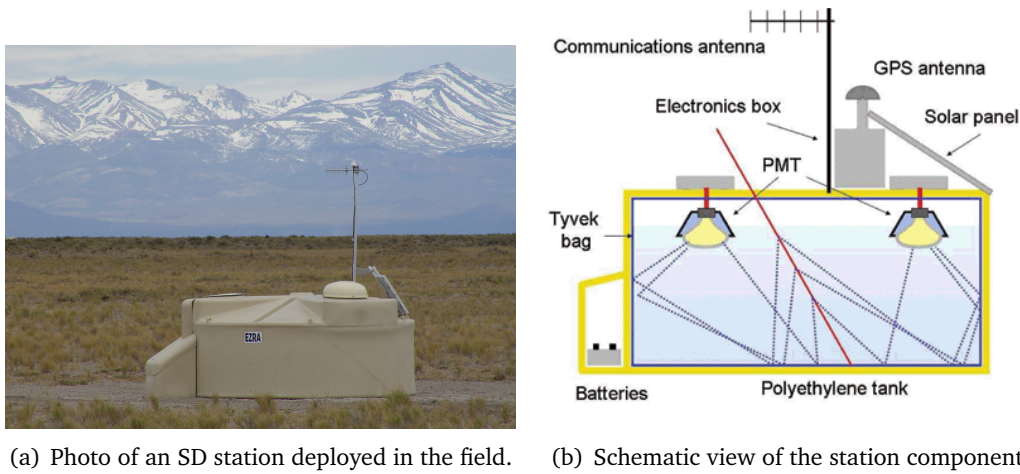
Several enhancements have been built recently at the Observatory. First of all, to study the energy spectrum down to energies of 0.1 EeV, a small portion of the SD array ($\simeq 24 \text{ km}^2$, known as the *Infilled array*) has been equipped with 61 WCD laid out in a denser grid of 750 m spacing. The *Infilled array* is overlooked from one observation site by three high-elevation telescopes (HEAT) [87] whose field of view cover elevations from 30° to 60° , allowing the study of lower energy showers. With the aim of measuring with better precision the muon content of the recorded air showers, the AMIGA project is deploying scintillation muon counters (each with an area of 30 m^2 and buried 2.3 m underground) near the WCD of the *Infilled array* [88]. Other programs are focused on the detection of the radio emission that takes place while the shower of particles evolves in the atmosphere. The Auger Engineering Radio Array (AERA) operates in the 30 to 80 MHz frequency range to detect the radio emission. It consists of 153 self-triggered antennas covering approximately 17 km^2 . We are also exploring the possibility of detecting microwave emission (GHz range) with an array of 61 antennas (EASIER) covering 100 km^2 , and with imaging parabolic dish detectors AMBER and MIDAS [86].

2.1 Surface Detector

To guarantee a high rate of events at the highest energies a large area is required. The Surface Detector of the Pierre Auger Observatory covers an area of more than 3000 km^2 . The spacing between WCD, 1.5 km, is the result of a compromise between cost considerations and the energy threshold. The surface is covered by WCD because of their low cost and robustness. Moreover, it is a well known detector for air-showers: the same principle was used successfully in other experiments like Haverah Park [89].

Each of the SD stations (see figure 2.2) is a cylindrical polyethylene tank of 10 m^2 base and 1.2 m height, with three 9 inch (22.9 cm Photonis XP1805) photomultipliers (PMTs) and filled with 12000l of purified water (resistivity: 5 - 15 $\text{M}\Omega \text{ cm}$). The water is enclosed in a Tyvek liner which reflects Cherenkov light produced by traversing particles, with high reflectivity (about 98% depending on wavelength). The signal registered by the PMTs is extracted both from the anode and the last dynode, the latter being amplified by a nominal factor of 32. This allows a large dynamic range extending from a few up to 10^5 photoelectrons. All channels are connected to a 5 pole anti-aliasing filter with a cutoff at 20 MHz digitized using 10 bit Flash Analog Digital Converters (FADCs), with a sampling rate of 40 MHz.

The station is equipped with two solar panels installed on top of the tank which charge 12V batteries. They provide a power supply of 10W. Global Positioning System (GPS) units establish the timing of the FADC signals providing one pulse per second output and software corrections. This signal is used to synchronize a 100 MHz clock that serves to time-tag the trigger. A radio antenna transmits the signals to the closest FD site, where a communication tower sends the data to the Central Data Acquisition System (CDAS).



(a) Photo of an SD station deployed in the field. (b) Schematic view of the station components.

Figure 2.2: View and scheme of an SD station of the Pierre Auger Observatory.

2.1.1 Calibration

The SD obtains a measurement of the Cherenkov light produced by shower particles passing through the detector. The unit used to measure the Cherenkov light is the signal produced by a vertical and central through-going muon (VCT), named vertical equivalent muon (VEM). The goal of the calibration procedure is to measure the value of 1 VEM in FADC units.

Atmospheric muons passing through the WCD give an excellent method to estimate the value of a VEM since they produce a peak in a charge histogram. In addition, the atmospheric muons calibration assures the uniformity of the signal size over the entire array. The calibration of each detector is performed locally and automatically so each detector is independent of the others. The most important quantities to calibrate an SD station are the average charge, Q_{VEM} , and the amplitude of the signal, $I_{\text{VEM}}^{\text{peak}}$, produced in a PMT by a VEM [90].

In practice, the SD can not select only VCT muons. However, the charge distribution produced by atmospheric muons produces a peak on both the charge distribution, $Q_{\text{VEM}}^{\text{peak}}$, and the pulse height, $I_{\text{VEM}}^{\text{peak}}$, (see figure 2.3) which are proportional to the expected values for a vertical muon. This peak $Q_{\text{VEM}}^{\text{peak}}$ is at approximately 1.09 VEM for the sum of the three PMTs, and at 1.03 VEM for each PMT. This shift is caused by the asymmetry in the muon track length distribution and by local light collection effects. The calibration to VEM units can be described in three steps:

- Set up the gains of the PMTs to have the $I_{\text{VEM}}^{\text{peak}}$ at 50 channels (ch). It translates in a mean gain of $\sim 3.4 \times 10^5$ for a mean $n_{pe}/\text{VEM} \sim 94$ photoelectrons (pe).
- Calibrate the electronics so that $I_{\text{VEM}}^{\text{peak}}$ is well determined. The average value of $I_{\text{VEM}}^{\text{peak}}$ for the PMTs of the SD is currently 46 ± 4 ch.

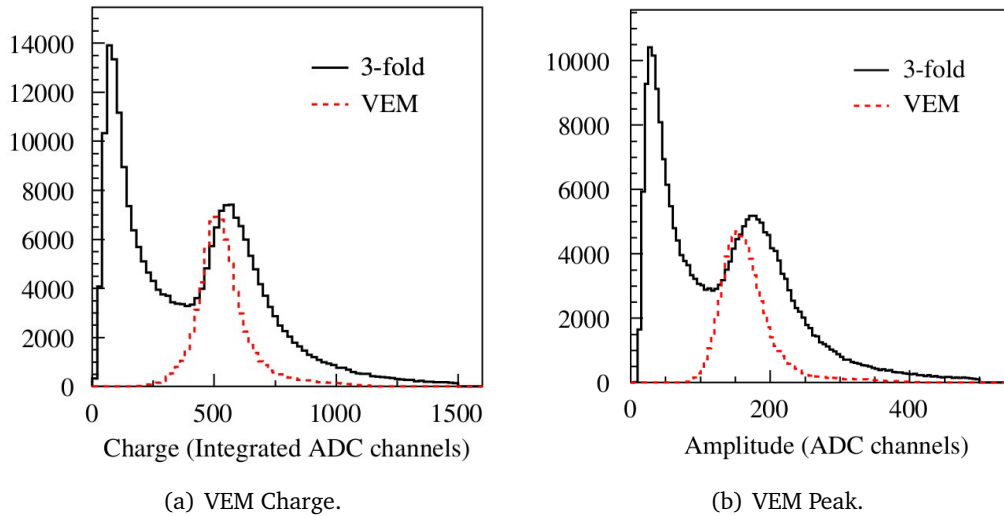


Figure 2.3: Charge and pulse height histograms for an SD station with a 3-fold trigger (signal in all 3 PMTs). The signal is the sum of the three PMTs. In the solid histogram the second peak is produced by vertical through-going atmospheric muons, while the first peak is a trigger effect. The dashed histogram is produced by vertical and central muons (VEMs) selected with an external muon telescope.

- Determine the value of $Q_{\text{VEM}}^{\text{peak}}$ from the charge distribution, and use it to establish the conversion to VEM units.

The calibration must also be able to convert the raw FADC traces into integrated channels. For this, the baseline of all six FADC inputs are needed, and the gain ratio between the dynode and anode (D/A). Averaging large pulses and performing a linear time-shifted fit we obtain both D/A and the phase delay between the dynode and anode. This method determines the D/A with a 2% resolution.

The calibration parameters, $Q_{\text{VEM}}^{\text{peak}}$ and $I_{\text{VEM}}^{\text{peak}}$, are determined every 60 seconds and sent to the CDAS together with every triggering event.

2.1.2 Trigger

The SD trigger is configured to detect the cosmic rays in a wide range of primary energies, for vertical and very inclined showers with an efficiency larger than 95% for cosmic rays above $10^{18.5}$ eV. The wireless communication system represents the main constraint to the rate of recordable events. To satisfy both physical and technical requirements the trigger system has been designed in a hierarchical form [91].

Station triggers

Two trigger levels, **T1** and **T2** are performed locally. At T1 level, different trigger modes are implemented to detect, in a complementary way, the electromagnetic and muonic

components of an EAS. The threshold trigger, **Thr1**, requires a coincidence between the 3 PMTs each above 1.75 VEM. This trigger is used to select large signals that are not necessarily spread in time corresponding to the muonic component. It also reduces the rate due to atmospheric muons from ~ 3 kHz to ~ 100 Hz. The Time over Threshold (**ToT**) trigger asks for a coincidence of at least two PMTs with more than 12 FADC bins above 0.2 VEM above the baseline within a window of 120 time bins. It is optimised for selecting small but spread signals, typical from high energy distant EAS or from low energy showers dominated by the electromagnetic component. Two additional triggers were implemented and operated since June 2013 in the local station software to increase the sensitivity of the SD to low energy air-showers and enlarge the distance on the ground up to which we can still observe the particles that are reaching the ground. These triggers are the Time over Threshold deconvoluted, **ToTd**, and the Multiplicity of positive steps **MoPS** both designed to increase the sensitivity of the SD to signals dominated by the electromagnetic components [92, 93, 94].

The second level, T2, is coded in the station software to reduce to about 20 Hz the rate of signals per detector. All TOT-T1 triggers are promoted to the T2 level. However, the Thr1 should pass a higher threshold of $3.2 I_{\text{VEM}}^{\text{peak}}$ in coincidence in the 3 PMTs. Only T2 triggers are used for the definition of a first level event trigger, T3.

Event trigger

The lowest event trigger (**T3**) looks for spatial and temporal coincidences in the T2 signals, and tries to associate them to an air shower. All data satisfying the T3 trigger are stored. Two kind of patterns are taken into account, 3-fold and 4-fold:

- 3-fold: One of the detectors must have one of its closest neighbours and one of its second closest neighbour triggered. The allowed time window considers the distance among stations in the following way: two stations have to be in the first two crowns around the first one considered. If this criteria is passed, the pattern is tagged as T3-3ToT. This trigger is extremely relevant since 90% of the selected events are showers and is mostly efficient for vertical showers.
- 4-fold: Four stations with T2 (Thr2 or ToT) have to be in coincidence with a moderate compactness requirement. In this case, the fourth station is being accepted if it is within four crowns around the reference station. This condition is only relevant for showers of large zenith angle with wide-spread topological patterns. The efficiency of this trigger is only 2%.

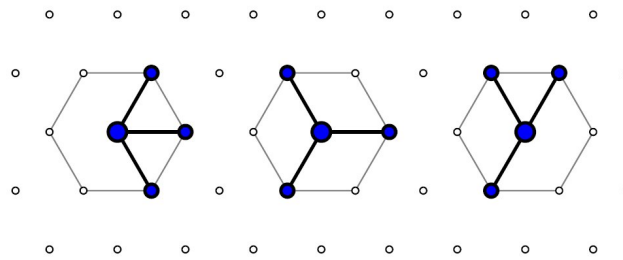
Physics trigger

The T3 trigger does not necessarily guarantee a relevant air shower. Thus, the physics trigger, **T4** is necessary to select real showers from the stored T3 data. Two criteria are defined:

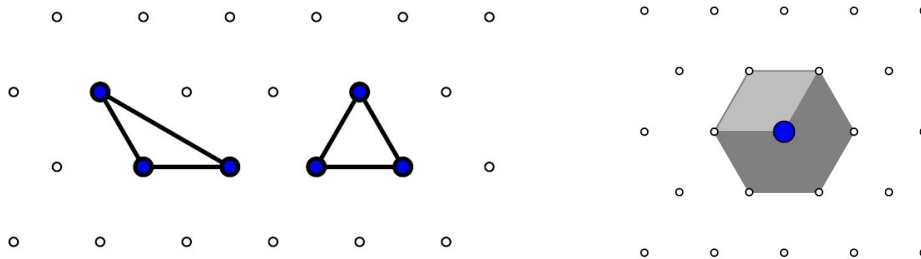
- 3ToT (see figure 2.4(b)): it requires at least 3 stations with a ToT trigger and it is a stricter version of T3-3ToT in a non-aligned compact configuration. About 99% of the vertical events are selected with this trigger condition [95, 96].

- 4C1 (see figure 2.4(a)): events with four stations with any T2 trigger and a configuration of one station with 3 close neighbours. This condition is only important for nearly-horizontal air showers that do not fulfill the 3ToT condition.

Compatibility in time between stations forming the trigger is required in every T4 event. The difference in their start time has to be smaller than the distance between them divided by the speed of light in vacuum, allowing for a marginal limit of 200 ns.



(a) Examples of T4-4C1 configurations.



(b) Examples of T4-3ToT configurations.

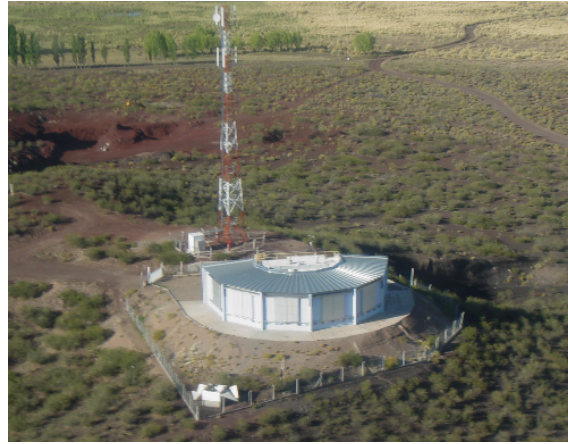
(c) 6T5 and 5T5 quality configurations.

Figure 2.4: T4 and T5 configurations. 2.4(a): The three minimal compact configurations for the T4-4C1 trigger. 2.4(b): The two minimal compact configurations for the T4-3ToT configuration. 2.4(c): Example of the 6T5 hexagon (shadow) and the 5T5 hexagon (dark shadow).

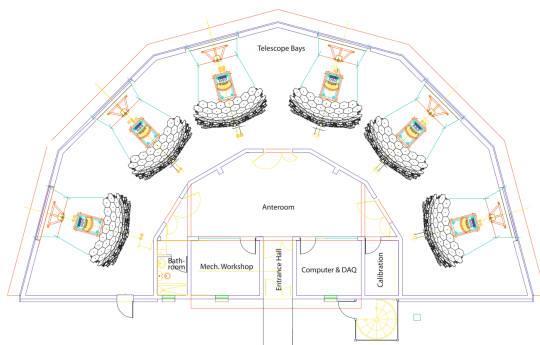
Quality trigger

The quality trigger **T5** is the highest level of trigger in the Pierre Auger Observatory. It mainly excludes events falling close to the border or in any hole of the array, where due to a possible missing signal, the reconstruction of the air shower variables has a worse resolution. To fulfill this trigger the highest signal station should be surrounded at the time of triggering by six working stations (not necessarily triggered). This is the so-called **6T5** trigger. A less restrictive criteria requires only five stations working in the first crown, **5T5**. Due to the large number of stations, around 1% of the detectors may not work at any moment, even with constant maintenance.

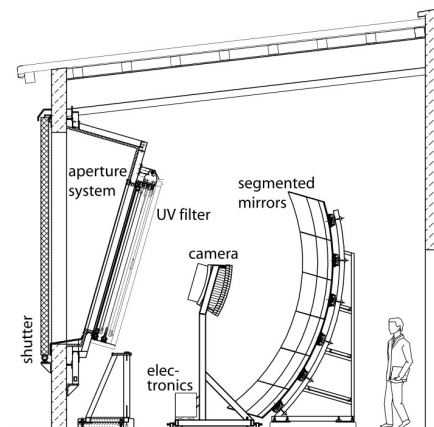
2.2 Fluorescence Detector



(a) Aerial photo of the FD site at Los Leones.



(b) Schematic top view of an FD site.



(c) Schematic lateral view of an FD telescope.

Figure 2.5: View and schemes of an FD site of the Pierre Auger Observatory.

Charged particles generated from a cosmic ray shower traversing the atmosphere excite atmospheric nitrogen molecules. These molecules emit fluorescence light with discrete wavelengths between 300 and 430 nm. The FD measures the longitudinal development profile $dE/dX(X)$ of the air shower as a function of the atmospheric slant depth. It also gives a calorimetric measurement of the energy of the primary cosmic ray.

There are 24 Schmidt telescopes located on four observation sites: Los Leones, Los Morados, Loma Amarilla and Coihueco, all of them situated on hills between 40 and 150 metres high surrounding the SD array. Each telescope has a field of view of $30^\circ \times 30^\circ$. The optical system consists of a 1.7 m diameter diaphragm. The collected light is reflected by a 12 m^2 spherical mirror with a radius of curvature of 3.4 m and then focused onto a camera comprised of 440 hexagonal PMTs in a 22×20 matrix (see figure 2.5). Each PMT

has a diameter of 45 mm and a quantum efficiency of about 25%. The light collected by the PMTs is converted into an electrical signal, and finally digitised by analogue-to-digital converters (ADCs) [97].

2.2.1 Trigger

The fluorescence detector trigger consists of four levels:

- **First level trigger (FLT):** The functions of the FLT are implemented in 4 FPGA (Field Programmable Gate Array), each controlling 6 channels, whose main task is to generate the pixel trigger using a threshold cut on the integrated ADC signal. The threshold is dynamically adjusted to maintain a pixel trigger rate of 100 MHz. The FADC values are integrated over 10 bins improving the signal to noise ratio by a factor of $\sqrt{10}$.
- **Second level trigger (SLT):** Their functions are also implemented in FPGA logic. It consists of identifying geometry patterns of pixels that look similar to straight lines. The algorithm searches for track segments of at least five pixels length in any of the patterns shown in figure 2.6, overlapping in a time window of 1 to 32 μ s. Patterns created by rotation or mirror reflection of these are also accepted. The rate is 0.1 Hz

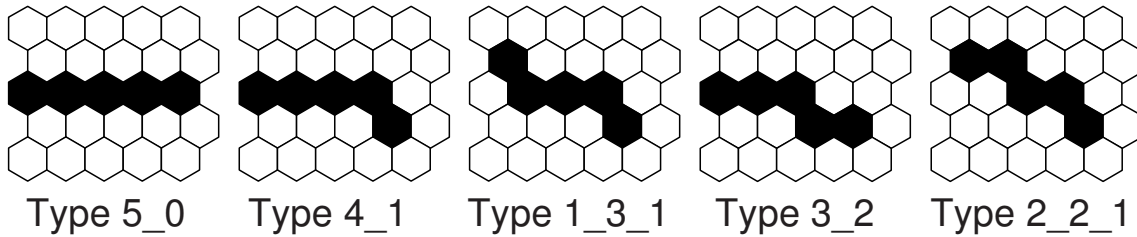


Figure 2.6: Basic patterns of triggered pixels considered by the Second Level Trigger in the FD.

- **Third level trigger (TLT):** It is a software algorithm designed to clean the air shower data stream of noise events that survive the low-level hardware triggers. It checks for the time structure of an event. In optimal data-taking conditions, only one or two events per minute and per telescope will survive SLT. The average trigger rate is 0.02 Hz per mirror.
- **Hybrid trigger (T3):** Events passing the TLT in each telescope are sent to the EyePC through the DAQ subnet. The software merges coincident events from adjacent telescopes and sends a hybrid trigger, called a T3, to the central data acquisition system (CDAS). When this trigger occurs the SD array is also read. Its purpose is to record hybrid events below 3×10^{18} eV, where the SD is not fully efficient. At these low energies, only one or two SD stations survive the trigger conditions, but this information together with the FD measurement is enough to ensure a high-quality hybrid reconstruction.

2.2.2 Calibration

The reconstruction of an air shower longitudinal profile depends on the ability to convert ADC counts into light flux at the telescope aperture for each channel that receives a portion of the signal from a shower. A calibration procedure is necessary to evaluate the response of each pixel to a given flux of incident photons.

The calibration of the FD is performed following two complementary strategies: an absolute calibration called “drum calibration” and a step by step sequence of dedicated calibration runs which are performed at the beginning of every data-taking night. The drum calibration consists in beaming a 2.5 diameter calibrated light source onto each telescope. It provides an uniform response for each pixel of 5 photons per ADC. This sort of calibration is performed only once every two or three years.

2.2.3 Fluorescence yield

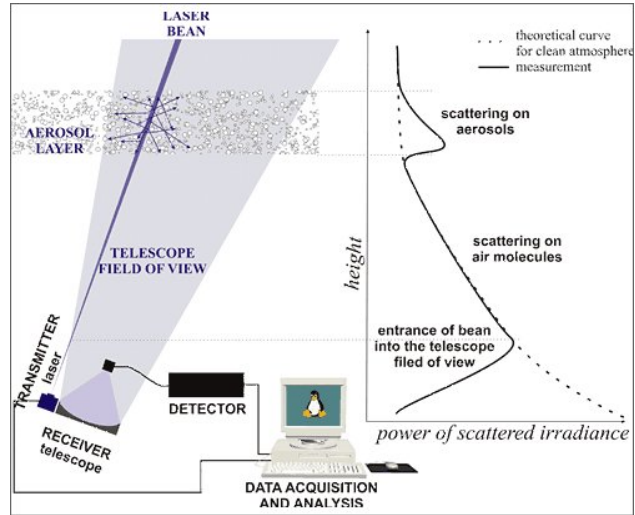
As we have described previously, the photons measured by the FD are emitted isotropically and in the wavelength range between 300 and 430 nm. The fluorescence yield is a key ingredient for the FD reconstruction because it is the proportionality factor between the number of photons emitted and the energy deposited in the atmosphere. The parameters that characterize the fluorescence yield include an absolute normalization of the wavelength spectrum, the relative intensities in different spectral bands, and their dependencies on pressure, temperature and humidity. The measurement of the absolute yield taken in the reconstruction is the one made by the Airfly collaboration. Nowadays the most accurate measurement in the 337 nm band, which has an uncertainty of 4% [98].

2.2.4 Atmospheric monitoring

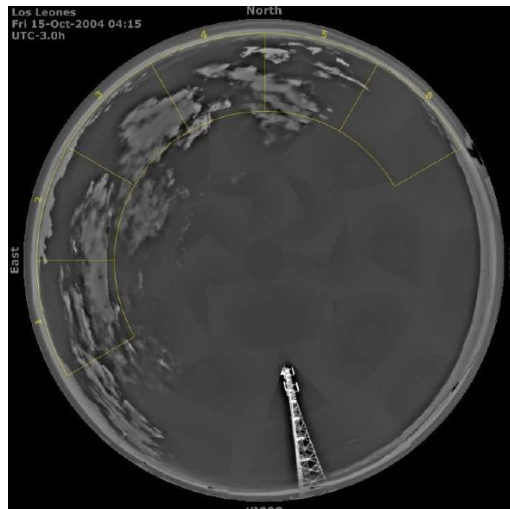
In addition to the absolute calibration of the telescope, a detailed analysis of the atmosphere above the SD array is a key to obtain the absolute number of photons emitted at the shower axis.

The rate of development of an air shower depends strongly on the atmospheric density and temperature as a function of altitude, and both evolve significantly over time, both on a short, daily and on a large, yearly scale. Moreover, there are additional effects affecting the production of light by secondary particles. The most relevant are Mie and Rayleigh scattering, both depending on the amount of aerosols in the air, and also light absorption.

The atmospheric monitoring in the Pierre Auger Observatory [99] is performed by four light detection and ranging (LIDAR) stations adjacent to each FD building, and also by the central laser facility (CLF), located near the centre of the SD array. In the four LIDARs, a beam of ultra-violet (UV) laser light is directed into the atmosphere at periodic intervals (see figure 2.7(a)). Each LIDAR has a PMT that detects the backscattered light from the UV laser pulses. The intensity and direction of the returning light collected by the LIDAR mirrors is used to measure the optical transmission conditions near the FD telescopes [100]. This allows the estimate of the aerosol content of the atmosphere. Infra-red cloud cameras and meteorological weather stations are also used to measure cloud coverage, humidity and other parameters in the vicinity of each FD site (see picture 2.7(b)).



(a) Scheme of the LIDAR operation.



(b) Example of an infra-red cloud camera shot in Los Leones.

Figure 2.7: The LIDAR and cloud cameras are part of the atmospheric monitoring used in the Pierre Auger Observatory.

Similarly, the CLF shoots a laser light into the atmosphere following a predetermined sequence of directions and zenith elevations every hour. The reconstructed energy and direction is compared to the true values, with a typical discrepancy of about 15%, due to the atmospheric effects previously mentioned. The CLF laser shots are used as well to calibrate the GPS timing of both the FD and the SD.

2.2.5 FD reconstruction

In the FD reconstruction, the first step consists in the reconstruction of the geometry of the event:

Geometry reconstruction

A hybrid detector achieves the best geometrical accuracy using timing informations of both components, FD pixels and SD stations. In addition, the hybrid capability extends the sensitivity of the detector at energies below 3 EeV, an energy range not accesible for the SD alone. Each element records a pulse of light from which can be determined not only the amount of light, but also the arrival time of the signal and its uncertainty. This temporal information is utilised to reconstruct the shower axis through a χ^2 minimisation procedure [97].

The first step is the determination of the shower detector plane (SDP). It is defined as the plane between the location of the eye and the shower axis (see fig 2.8). Experimentally, it is the plane passing through the detector that contains most of the pointing directions of the FD pixels centred on the shower axis.

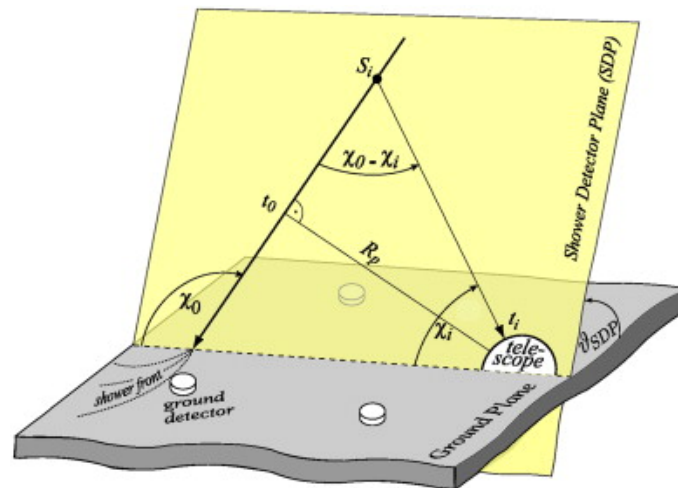


Figure 2.8: Shower-detector plane parameters.

The next step is to reconstruct the shower axis from the pixel timing information. It must lay within the SDP, and is defined by two parameters: R_p , the perpendicular distance from the camera to the track, and χ_0 , the angle the track forms with the horizontal line in the SDP. To minimize the uncertainty in the arrival direction and core position the FD timing information is used simultaneously with one of the SD stations, in what is called

the *hybrid reconstruction* (see figure 2.9). If the reconstruction is performed using only an FD telescope, is called *monocular reconstruction*. The uncertainty in the arrival direction and core position is larger (see figure 2.10).

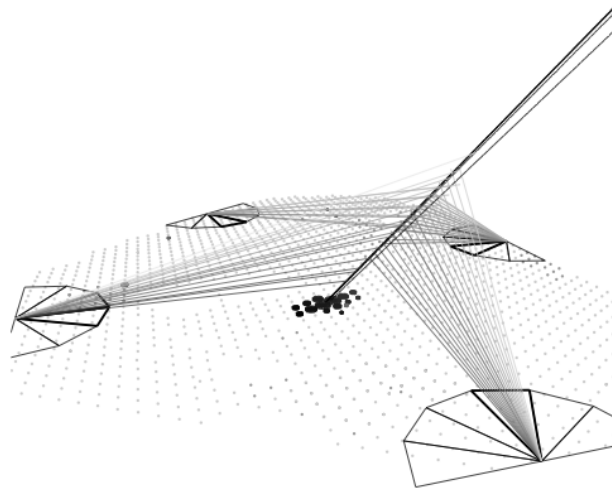


Figure 2.9: Geometry reconstruction of an event observed by four telescopes and the Surface Detector [86].

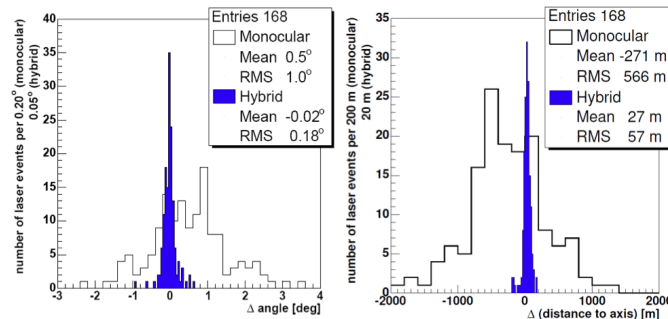


Figure 2.10: Angular resolution (left) and core location accuracy (right) for monocular and hybrid reconstruction.

As the energy of the primary cosmic ray raises, there is an increasing probability for an event to be observed simultaneously by two or more FD sites (stereo, triple and four-fold events). These events are particularly interesting for the evaluation of resolution, as they offer the chance to observe an incoming shower from different directions, performing an independent hybrid reconstruction for each site. Alternatively, it is also possible to reconstruct the geometry of the shower in a multiple-eye mode by intersecting the different shower detector planes and without information coming from the SD. In the stereo mode, the resolution on the geometry depends mainly on the number of pixels used to find the SDP. A minimum of 6 pixels in each FD station is required to obtain a good reconstruction. The angular resolution of the FD is below 0.6° for events with energy above 3×10^{18} eV

(see figure 2.11) [101].

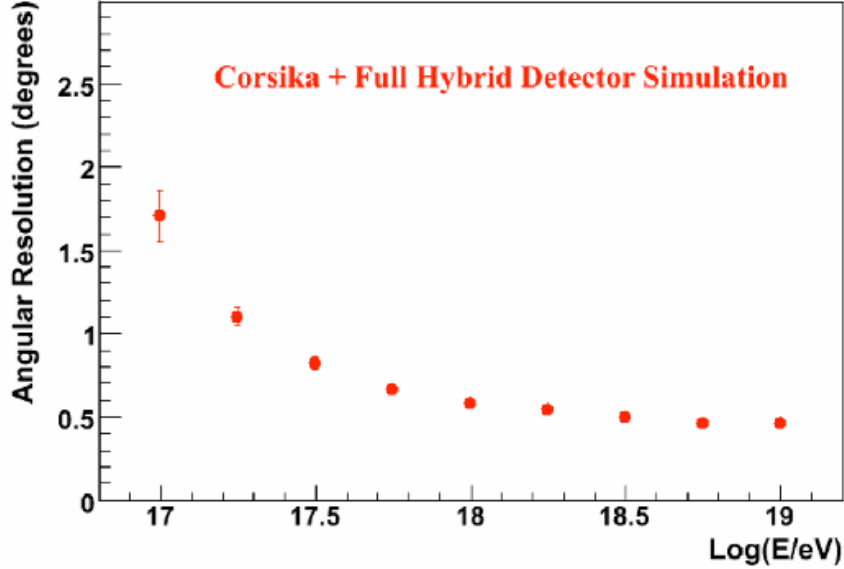


Figure 2.11: Angular resolution of hybrid events [101].

Profile and energy reconstruction

Once the geometry of the event has been reconstructed, the amount of light collected by the pixels can be converted into deposited energy as a function of the slant depth. The atmospheric depth is the integral of the atmospheric density as a function of the altitude. The slant depth, X , is defined as the atmospheric depth measured along the shower axis.

$$X = \int_z^{\infty} \rho(z') dz' \quad (2.1)$$

where ρ is the density of the atmosphere at distance z' from ground measured along the shower axis.

To obtain the amount of deposited energy by the shower, the light attenuation from the shower to the telescope needs to be estimated. Moreover, the different contributions to the produced light must be considered: fluorescence, multiple-scattering and Cherenkov (both direct and scattered) light. The energy deposited at slant depth X_i is then expressed as

$$\frac{dE}{dX_i} = N_e(X_i) \int_0^{\infty} f_e(E, X_i) \frac{dE}{dX_e}(E, X_i) dE \quad (2.2)$$

where $f_e(E, X_i)$ is the normalised electron energy distribution, dE/dX_e the energy loss of a single electron with energy E and $N_e(X_i)$ represents the number of electrons and positrons above a certain constant energy cut-off [102].

The calorimetric energy of the shower is estimated by fitting a Gaisser-Hillas function [103] to the reconstructed profile and integrating it (see figure 2.12):

$$f_{GH}(X) = \frac{dE}{dX_{\max}} \left(\frac{X - X_0}{X_{\max} - X_0} \right)^{(X_{\max} - X_0)/\lambda} \exp^{(X_{\max} - X)/\lambda} \quad (2.3)$$

where X_{\max} represents the depth where the number of particles is maximum and dE/dX_{\max} the maximum energy deposit.

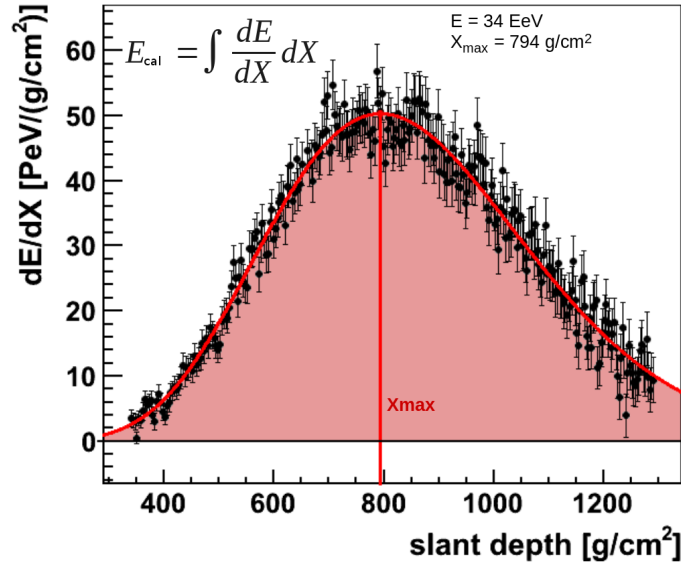


Figure 2.12: Energy reconstruction in the FD. The calorimetric energy is obtained by the integration of the fit to the profile.

Finally, the total energy is obtained by correcting the calorimetric energy by a correction factor f_{inv} which accounts for the energy not included in the electromagnetic component of the shower. It is called the *invisible energy* [104], E_{inv} . It includes two facts: the deposition of energy in the atmosphere for muons is negligible, and then they release most of their energy at the ground. Also, neutrinos carry an amount of energy that escapes undetected. This correction is obtained from data themselves reducing the dependence on the hadronic interaction models. It is estimated for each shower using the FD measurement of the longitudinal profile and $S(1000)$, the SD signal at 1000 m from the core position of the shower at the ground. E_{inv} can only be reliably estimated above 3×10^{18} eV, the threshold for the SD array full efficiency, because for lower energies it is biased by upward fluctuations of the shower signals. The FD detects showers at lower energies than the SD. To be able to estimate E_{inv} at these energies, it is parameterized with an analytical function above 3×10^{18} eV and extrapolated down to 10^{17} eV. The *golden hybrid events* (the same set of events used to calibrate the SD energy) are used to obtain the analytical function:

$$E_{\text{inv}} = a_0(E_{\text{cal}}[\text{EeV}])^{a_1} \quad (2.4)$$

where $a_0 = (0.174 \pm 0.001) \times 10^{18}$ eV and $a_1 = (0.914 \pm 0.008)$ [105].

The resolution in the measurement of the energy achieved by the FD depends on the uncertainties associated to variations in the atmosphere, ranging from 4.5% at 3×10^{18} eV to 6.9% at 10^{20} eV, the invisible energy (1.5%) and the geometry reconstruction, which ranges from 5.2% to 3.3% for the same energy interval. The resulting overall energy resolution is almost constant with energy in the range [3, 100] EeV, and lies between 7% and 8% [106].

SD event reconstruction

3

The measurement of the energy and the arrival direction of the cosmic rays producing air showers that have triggered the Surface Detector array is based on the sizes and times of signals registered from individual SD stations. By sampling both the arrival times and the deposited signal in the detector array, the geometry, the arrival direction and the shower size of the incident cosmic ray can be determined.

The reconstruction technique used depends upon the zenith angle (θ) of the direction of the incident cosmic ray which defines the amount of atmosphere traversed by the shower, and therefore the level of attenuation of the shower components. We distinguish between *vertical*, showers with $\theta < 60^\circ$, and *inclined* reconstruction, with $60^\circ < \theta < 80^\circ$.

To build the relevant physics quantities we use a dedicated software called Offline¹ [107]. It is a framework divided in modules that contains all the tools required to simulate and reconstruct the events. In the following sections we will describe the main details of the SD reconstruction performed in Offline [108].

3.1 Event selection

All the stations passing the trigger criteria are not part of the event. To determine which ones belong to the actual air shower a selection at the station and PMT level is necessary. *Candidate* stations in the reconstruction procedure are the ones compatible in space and time with the propagation of a plane shower front.

The reasons to remove a station from the reconstruction are bad calibration and/or accidental timing information. Atmospheric muons can also trigger a station. This accidental triggered stations are removed using a compactness criteria. In addition to the regular stations, some are placed in special configurations. For signal and timing accuracy studies one or two are located at 11 m off the standard ones (they formed the *doublets* or *triplets*). These stations are not used in the reconstruction.

¹Another SD reconstruction software, CDAS, is available in the Pierre Auger Collaboration, but it will not be discussed here.

3.2 Geometry reconstruction

A shower track (see figure 3.1) can be described as a point \vec{x}° moving at the speed of light c along the shower axis \hat{a} and arriving at the origin at time t_0 :

$$-\hat{a} \left(\vec{x}^\circ(t) - \vec{b}^\circ \right) = c(t - t_0) \quad (3.1)$$

where \vec{b}° is the signal-weighted barycentre, taken as the first approximation at the origin from where all the distances are measured. Similarly, the weighted bary-time is set as time origin. These quantities are replaced by more accurate estimations after every iteration of the reconstruction chain.

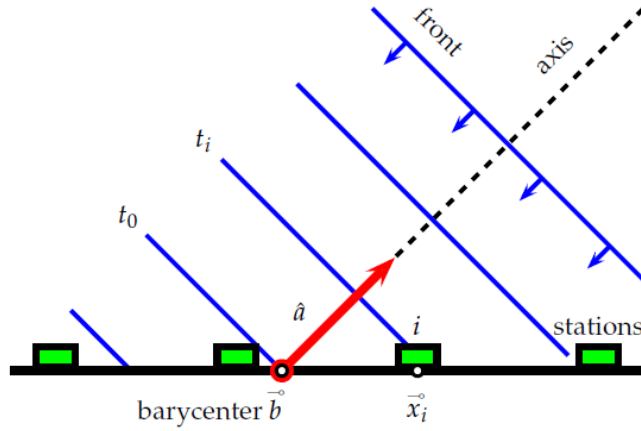


Figure 3.1: Schematic view of a plane shower front.

The arrival direction is approximated by fitting a planar shower front to the start times of the signals recorded by WCD. The *shower plane* is a plane perpendicular to the shower axis, moving along with the same speed and containing the shower front.

A more realistic description of the shower front is based on a curved front fit, as depicted in figure 3.2. The idea is to extend the plane fit method with a parabolic term that describes the curvature of the shower front near the impact point.

3.3 Reconstruction of the lateral signal distribution

The Lateral Distribution Function (LDF) characterises the dependence with the radial distance of the measured signal. In the Pierre Auger Observatory, this dependence is modelled with a modified Nishimura-Kamata-Greisen function [109]:

$$S(r) = S(r_{opt}) \left(\frac{r}{r_{opt}} \right)^\beta \left(\frac{r + r_1}{r_{opt} + r_1} \right)^{\beta+\gamma} \quad (3.2)$$

where r_{opt} is the optimum distance, $r_1 = 700$ m and $S(r_{opt})$ is an estimator of the shower size used for energy assignment. For the SD array with station spacing of 1.5 km the optimum distance [110] is $r_{opt} = 1000$ m and the shower size is thus $S(1000)$. For the

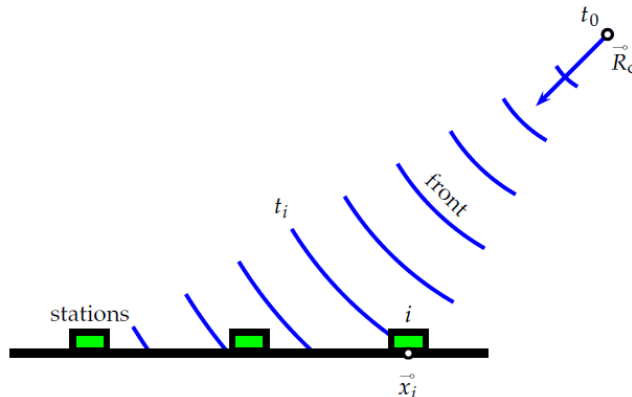


Figure 3.2: Schematic view of a spherical shower front.

Infilled array, whose spacing between detectors is 750 m the optimum distance is 450 m, and the shower size $S(450)$. The parameter β depends on the zenith angle and the shower size. *Vertical events* are observed at an earlier shower age than the inclined ones. The LDF of vertical events is steep due to the different contributions from the muonic and the electromagnetic component at the ground. For events with only three stations, the reconstruction of the air showers can be obtained by fixing the two parameters, β and γ to a parameterization obtained using events with a number of stations larger than four.

The model to determine the uncertainty in the signal has been updated including the new triggers ToTd and MoPs. A detailed description is given in the next section.

3.3.1 Signal variance

In this section we will describe a new way to estimate the uncertainty of the signal, σ_i , [2] measured by the WCD and which is a key ingredient in the fit of the LDF. In early works [111, 112, 113, 114, 115, 116] the signal variance was deduced for signals larger than 10 VEM and it was shown that the variance of the signal is proportional to the signal with a linear dependence on $\sec(\theta)$. The dependency on the zenith angle is caused by the change in the nature of the particles reaching the ground (a muonic signal has a larger variance than an electromagnetic one), and by the larger track-length of inclined muons within the water.

A recent study [117] has disentangled the contributions from the electromagnetic and the muonic components in the signal variance, showing that we expect dependencies of the signal variance on the distance to the shower axis, the azimuth of the station in the shower plane, and other variables sensitive to the electromagnetic to muonic ratio. The implementation of new triggers in the local station software requires to validate the signal variance for low signals and to understand if the sensitivity of the new triggers to mostly electromagnetic particles (as in case of MoPs) involves a new parameterization of the signal variance.

In order to obtain the signal variance we perform an un-binned maximum likelihood minimization in the signal range of full trigger efficiency. Taking into account the station

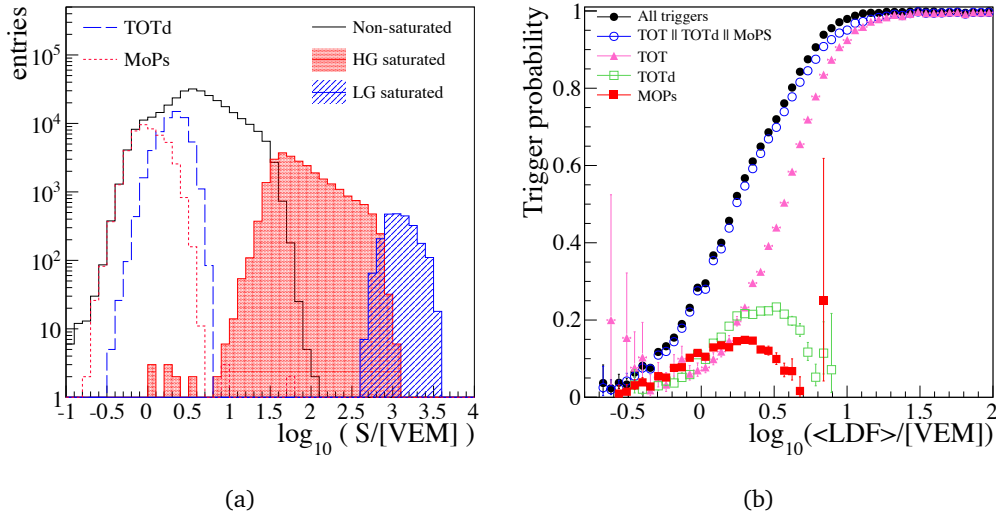


Figure 3.3: (a) Measured signal spectrum (b) Trigger probability as a function of the mean expected signal.

trigger probability we investigate the low signal region via a toy MC. For this study we are using 6 months of the *Infilled* data, starting from the implementation of the new triggers, end of July 2013 until March 2014 and the entire data period for the old triggers, from 2008. Data were reconstructed with the *Offline* framework and saved in Advanced Data Summary Tree (ADST) format [118, 119]. We revise the current variance model for signals larger than 14 VEM, described in the first section. A posteriori we investigate with a toy MC if the signal variance holds over the entire signal range.

Signal variance in the region of full trigger efficiency

The dynamic range for the observed signals extends from about 0.3 VEM up to 3000 VEM (see figure 3.3(a)). There are several transitions in the signal spectrum: from the new triggers, TOTd and MOPs, to the standard triggers TOT/Thr2 at about 3 VEM, then after about 40 VEM the high gain saturated signals start to dominate while at 400 VEM we observe signals saturated also in the anode (low gain) channel. The station will always trigger for a signal larger than 12 VEM independently of the zenith angle [3, 120] (see figure 3.3(b)). The different contributions to the signal spectrum define our selection of the signals to obtain the signal variance model.

The signal variance is obtained from the doublet stations, that are separated 11 m on the ground. In this study we use the doublets which are located in the *Infilled* array. The two stations, being located so close to each other, even if they do not measure the same particles, observe the same spot of the air-shower and have the same relative contributions from the electromagnetic and the muonic parts. Close to the shower axis the signal decreases rapidly with increasing distance, and a difference in the location of the stations of 11 m becomes important producing large differences between the measured signals. Nevertheless, one can correct for this bias by using the expected signals [116]. Moreover, to

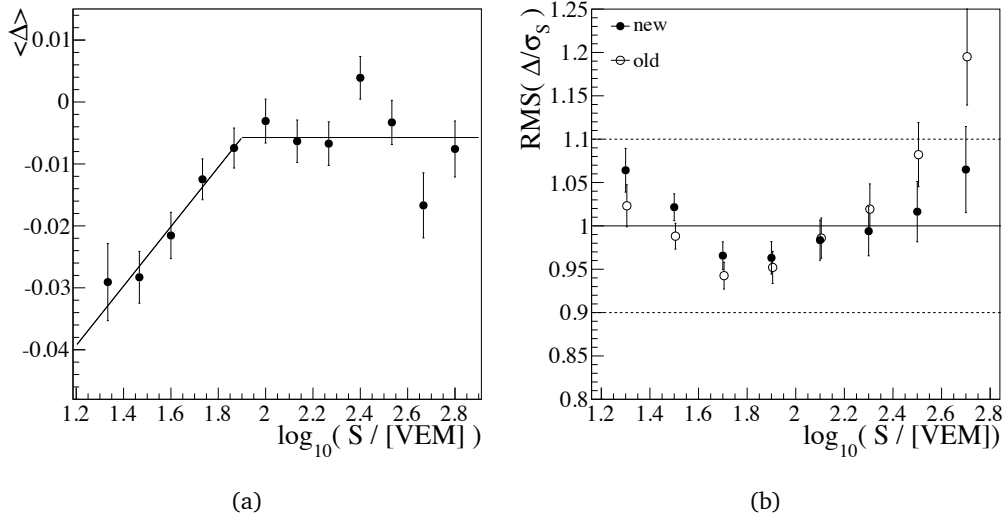


Figure 3.4: (a) Mean of Δ as a function of the estimated signal. (b) RMS of the distribution Δ divided by the signal uncertainty for the new (solid circles) and the previous parameterization (open circles).

avoid biasing the variance distributions, we are restricting our analysis to doublets located at more than 150 m from the shower axis. The signals are shifted to the mean distance, $\langle r \rangle$, using the lateral distribution function values, $f(r)$,

$$S = S_{\text{meas}} + [f(r) - f(\langle r \rangle)] \quad (3.3)$$

where r is the distance to the shower axis.

Assuming that the two stations measure the same physical quantity, we can define the relative signal resolution as

$$\frac{\sigma}{S} = \text{RMS}(\Delta) \quad \text{with} \quad \Delta = \frac{1}{\sqrt{2}} \frac{S_1 - S_2}{\langle S \rangle}, \quad (3.4)$$

where $S_{1,2}$ correspond to the signals in the pair of detectors.

The signal variance besides the expected uncertainty given by the number of particles that enter the WCD should have a contribution from the detector resolution:

$$\sigma_S^2 = \sigma_{\text{sh}}^2 + \sigma_{\text{det}}^2 \quad (3.5)$$

The last term has been already proposed in [111, 112] and it was suggested that it is related to the calibration uncertainty, which is 2% [121]. The first term related to the air-shower physics, to the Poisson fluctuation of the particles entering the detector, should be proportional to the expected signal S_{exp} ,

$$\sigma_S^2 = f^2 \cdot S_{\text{exp}} + c^2 S^2 \quad (3.6)$$

In [117] it has been shown that the spectral factor, f , depends on the ratio between the electromagnetic and the muonic ratio, and thus it shows dependencies on the early-late

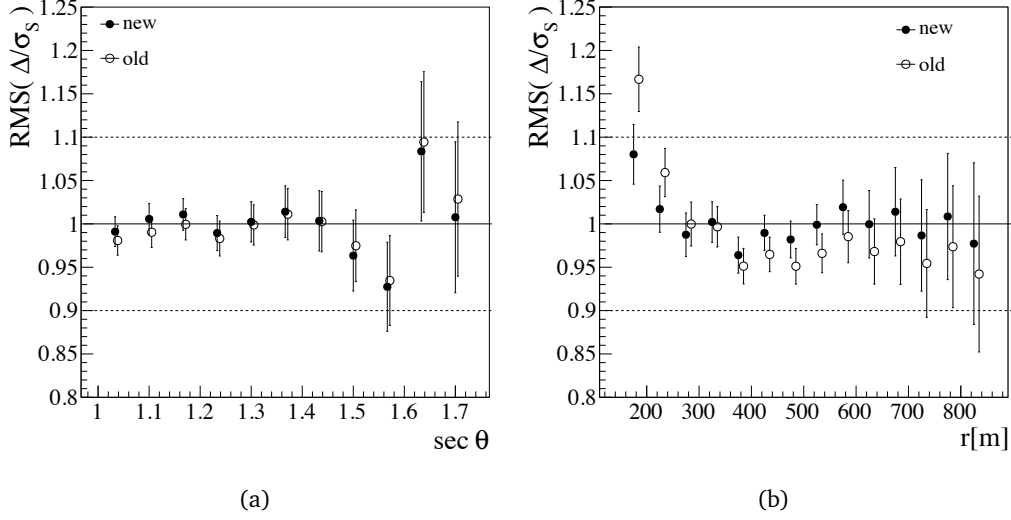


Figure 3.5: The RMS of the distribution Δ divided by the signal uncertainty for the new (solid circles) and the previous parameterization (open circles) as a function of $\sec\theta$, (a), and of distance to axis, (b).

asymmetry of the air-shower, on the distance to the shower core and on the zenith angle. In this study we neglect all these dependencies and we apply the classical approach of a simple linear relation on the secant of the zenith angle:

$$f = a \cdot (1 + b(\sec\theta - \sec 35^\circ)) \quad (3.7)$$

where a is given as the uncertainty factor at 35° to reduce the correlation between the parameters a and b . To obtain the parameterization we maximize the likelihood:

$$\mathcal{L} = \prod_{i=1}^N \mathcal{L}_i = \prod_{i=1}^N \frac{1}{\sqrt{2\pi} \sigma} \exp\left(-\frac{(\Delta_i - \langle\Delta\rangle)^2}{2\sigma^2}\right) \quad (3.8)$$

where N is the number of doublets, σ the relative uncertainty of the signal and Δ is given by equation 3.4. $\langle\Delta\rangle$ represents the mean (or the bias) of Δ and it is illustrated in figure 3.4(a). The observed bias is present in all pairs and we are using always for S_1 the doublet with the lowest ID² which were deployed before their twin companion. It might be caused by the ageing of the detectors. The bias is less than 3.5% and has been described with a linear function up to $\log_{10}(S_{\text{exp}}/\text{VEM}) = 1.9$ and then a constant of 0.008:

$$\langle\Delta\rangle = -(0.10 \pm 0.01) + (0.05 \pm 0.01) \log_{10}(S_{\text{exp}}) \quad (3.9)$$

We use data from the *Infilled array* that fulfill the 5T5 trigger condition, have an energy larger than 0.3 EeV and a zenith angle smaller than 55° . We reject outliers for which the relative signal difference between the doublets is larger than 1.5. An example of this

²ID is a numerical flag identifying every single WCD. Represented by a cardinal number, they were assigned chronologically in terms of deployment date.

type of outliers is given in appendix A. We also reject the doublets for which the time difference between them is larger than 200 ns. We obtain the variance model in the region of full trigger efficiency: in order to avoid a cut on the signal which might influence the distributions in the cut region we select the pairs which are at a distance where the trigger efficiency is more than 98% as given by the lateral trigger probabilities (LTP) described in [3].

The minimization leads to the parameters:

$$a = 0.865 \pm 0.007, \quad b = 0.593 \pm 0.006, \quad c = 0.023 \pm 0.002 \quad (3.10)$$

The RMS of the pull distribution, Δ/σ , is illustrated in figure 3.4(b) as a function of the signal for the new parameterization compared to the one from a previous parameterization [116]. As can be observed both parameterizations present a small bias evolving with signal. Nevertheless the new variance, being smaller between 40 and 100 VEM, leads to a value closer to the expected one. The high signal region is also better described by the new parameterization. In figure 3.5(a) the same variable shows no dependency on the zenith angle. There is a dependency on the distance to the shower axis (see figure 3.5(b)). It might indicate that the signal variance should be larger than the one obtained, but this can be an effect of the lateral distribution correction. Let us note that in principle by going closer to the shower axis the signal variance factor, f , should decrease due to the larger contribution of the electromagnetic component.

We have obtained the signal variance with the *Infilled array* for energies spanning the 0.4-1 EeV range. As in principle the signal variance should have an energy dependency [117], we have checked at energies above 3 EeV if the same parameterization holds. In appendix A.1 the *Infilled* data are compared to the larger array. There is a very good agreement between the two data samples. As a remark the statistics of the doublets in the regular array is a factor 10 smaller using data starting from 2004.

Signal variance for low signals

The signal variance in the whole range is shown in figure 3.6(a). In the previous section we obtained the relative signal uncertainty as a function of the expected signal above 14 VEM. We can distinguish three different regions. Above 400 VEM the trend is almost flat because the detector resolution contribution becomes significant. The region between 14 and 1000 VEM is the physical region where the Surface Detector is fully efficient. Below 12 VEM we are dominated by trigger effects.

The distribution of Δ in each bin can be described by a Gaussian with a mean of zero. Going to lower signals, the distribution starts to be truncated due to trigger effects, therefore the RMS decreases as a consequence of not measuring all the signals. Furthermore, the bias in Δ distribution starts to be significant. The combination of these two facts results in the change of trend that we observe. For the case of standard triggers, TOT and T2, the probability of having a triggered station for signals smaller than 3 VEM is almost zero (see figure 3.3(b)). However, in the case of the new triggers, the non-zero probability extends up to 0.5 VEM. Even so, as the contribution of new triggers is not very large, still 12 VEM are needed to reach full trigger efficiency.

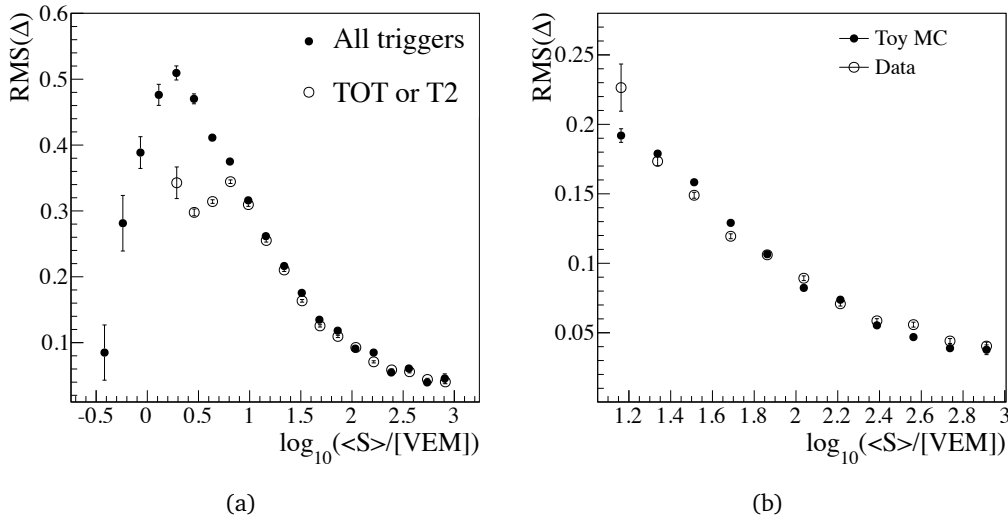


Figure 3.6: (a) $\text{RMS}(\Delta)$ versus mean signal for old triggers (open circles), TOT and T2, and including the new ones, TOTd and MoPS. (b) $\text{RMS}(\Delta)$ versus mean signal for the toy MC (solid circles) and for data (open circles), in the region where the trigger efficiency is larger than 98%.

A toy MC is built in order to understand the signal variance in the low signal region. The first ingredient to calculate the relative signal accuracy is the incident signal distribution. We draw randomly the signal from the measured distribution weighted by the trigger probability. To obtain the two *measured* signals, we smear this value with a normal distribution with a sigma given by the variance parameterization. To include the trigger probability we use the parameterizations described in [3].

In figure 3.6(b) the signal uncertainty as a function of the mean signal, considering only the signals with a trigger probability larger than 98%, is illustrated. The variance estimation was obtained in the radial distance where the LTPs are larger than 98%. As can be seen the agreement is very good over the whole signal range. In addition, the cut in LTP is equivalent to the cut in the trigger signal probability, as in both cases signals above 14 VEM are selected.

Considering all the signals we build the $\text{RMS}(\Delta)$ as a function of the signal for triggers TOT and T2 (figure 3.7(a)), and include new triggers, TOTd and MoPs as well (figure 3.7(c)). In both cases the agreement between the toy MC and data for signals larger than 0.6 VEM is very good. The ratio of the relative signal accuracy and the parameterization as a function of the mean signal is illustrated in figure 3.7(b) for standard triggers and in 3.7(d) for all of them. In both cases we reach a plateau for signals larger than 12 VEM where the array is fully efficient, as it is expected. These plots have been done in the zenith range between 0 and 55 degrees. As the trigger probability parameterization depends on zenith angle we have included in appendix A.1 the plots in 5 different ranges: the results are compatible in all cases. From these results we can conclude that the signal variance obtained in the signal region between 14 to 1000 VEM holds in the entire signal range.

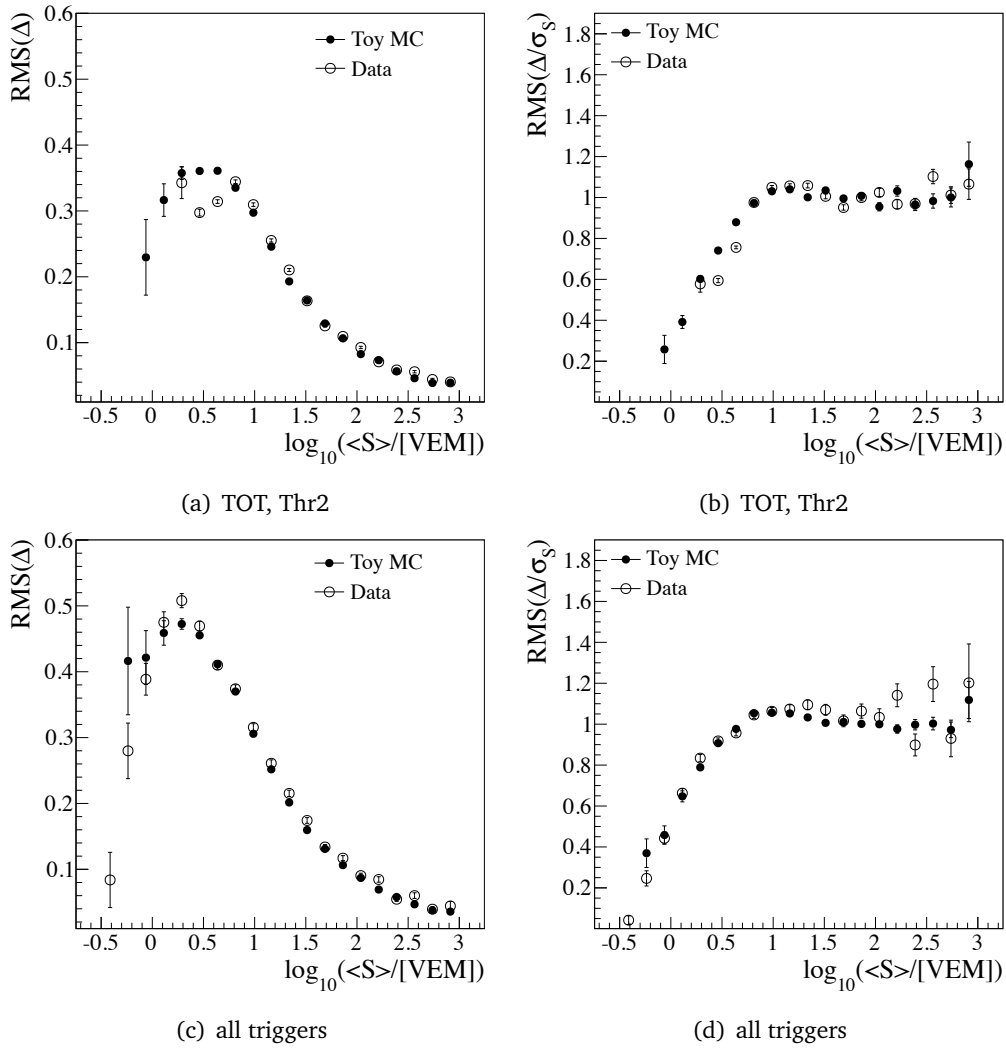


Figure 3.7: (a,c): The $\text{RMS}(\Delta)$ versus mean signal compared to the toy MC. (b,d): the RMS of Δ divided by the parameterization. A good agreement is observed between data and the toy MC.

Impact on the *Infilled reconstruction* The signal variance model obtained in this work does not differ too much from the one that is used in the official analysis and thus the impact on the reconstruction is not expected to be large. We have reconstructed the *Infilled* data with both signal variance models. The mean change on the energy is less than 0.5% with a spread of less than 1%. The core is shifted less than 20 m with a negligible impact on the angle. The only effect is, as expected, on the energy uncertainty. Having a signal variance that is smaller for parts of the signal spectrum, the χ^2 has larger values for the new parameterization, which leads to a reduction of the energy uncertainty by about 10% at 10^{16} eV, 3% at $10^{17.5}$ eV and less than 0.5% above 10^{18} eV. Nevertheless very few events (0.02%) have a change in the location of the impact point on the ground of more than

20 m.

3.3.2 Angular resolution

To estimate the angular resolution of *vertical events*³ a single station time variance model (TVM) is used [122] to take into account the size of the total signal and the time evolution of the signal trace.

The arrival direction of cosmic rays detected with the SD is determined from the arrival time of the shower front (T_s) in each SD station. The accuracy in the determination of T_s is limited by the physics of the shower (particle density, shower front thickness and shower-to-shower fluctuations) and by detector resolution. In the TVM the variance $V[T_s]$ of the arrival time of the shower front can be described by:

$$V[T_s] = a^2 \left(\frac{2T_{50}}{n} \right)^2 \frac{n-1}{n+1} + b^2 \quad (3.11)$$

where n is the estimated number of particles entering in the station; T_{50} the time interval to reach 50% of the total signal. The term b^2 is related to the time accuracy of the detector, the GPS uncertainty and the resolution of the FADC. The two parameters (a and b) have been obtained from a fit to the data collected by the *doublet stations* (as in the case of the signal uncertainty) because we can assume that they are sampling the same region of the shower. These parameters are expressed as a function of $\cos \theta$:

$$a^2 = -0.49 \cos^2 \theta + 0.22 \cos \theta + 0.65 \quad (3.12)$$

$$b^2 = 412 \cos^2 \theta - 209 \cos \theta + 141 \text{ [ns}^2\text{]} \quad (3.13)$$

The angular resolution will be determined, on an event by event basis, from the zenith (θ) and the azimuth (ϕ) uncertainties obtained from the geometrical reconstruction using the relation:

$$F(\eta) = \frac{1}{2}(V[\theta] + \sin^2(\theta)V[\phi]) \quad (3.14)$$

where η is the space-angle and $V[\theta]$ and $V[\phi]$, the zenith and azimuth variances. If θ and $\phi/\sin(\theta)$ have Gaussian distributions with variance σ^2 , then $F(\eta) = \sigma^2$ and η has a distribution proportional to $e^{-\eta^2/2\sigma^2} d(\cos \eta) d\phi$. The angular resolution (AR) is defined as the angular radius that contains 68% of the showers coming from a given point source [123]. Then:

$$AR = 1.5\sqrt{F(\eta)} \quad (3.15)$$

It depends strongly on the detector timing resolution and weakly on the shower front model and the core position uncertainty.

In figure 3.8 it is illustrated the angular resolution as a function of the zenith angle for different station multiplicities. The resolution is better than 1.6° , and better than 0.9° for events with more than six stations. In figure 3.9 we show the angular resolution as a function of the energy of the air-shower for different angle data sets.

³For *inclined events* a different approach is used for the reconstruction.

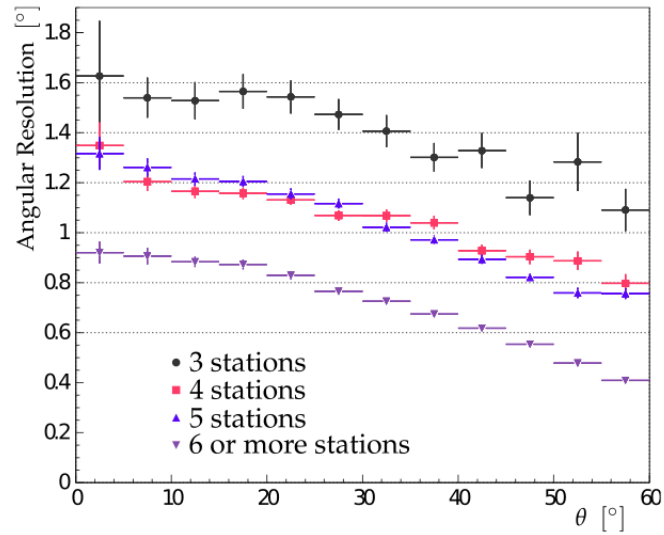


Figure 3.8: Angular resolution as a function of the zenith angle θ for events with an energy above 3 EeV, and for various station multiplicities [86].

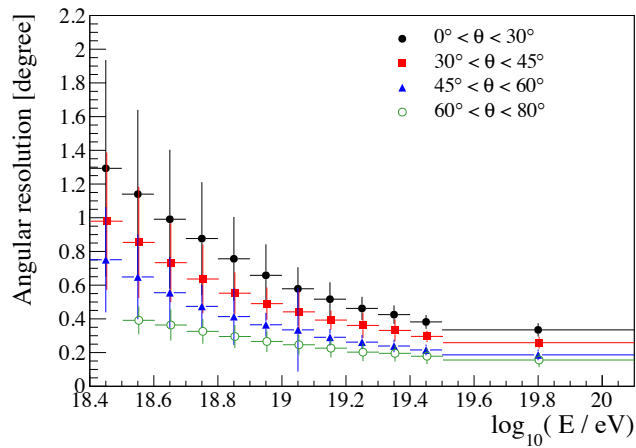


Figure 3.9: Angular resolution as a function of the primary energy for different angle intervals.

3.3.3 LDF maximum likelihood fit

The LDF fit is based on a log-likelihood maximization. It allows us to include: non-triggering stations with signal below trigger threshold, small or large signal, and the signal of saturated stations. Therefore, it is important to correctly estimate the uncertainty in the signal for each station as discussed previously.

The energy deposit in the WCD, or equivalently the number of registered photo-electrons depends on the particle type, injection point and incident angle. Although the signal is calibrated in VEM there is no simple conversion between photo-electrons and particles.

The main contributions to the total signal measured in a detector are the electromagnetic and the muonic part:

$$S = S_\mu + S_{e/\gamma} \quad (3.16)$$

The signal $S_{e/\gamma}$ is much smaller on average than S_μ and the mean conversion factor for electrons and muons is smaller than 1 VEM. The total number of particles that have produced the signal is estimated as:

$$n = pS \quad (3.17)$$

where n is the number of particles, S the collected signal and p the *Poisson factor*, approximated in this minimal model by:

$$p = p(\theta) = \begin{cases} \frac{1}{\sqrt{\sigma_S(\theta)}} & , \text{ if } \sigma_S(\theta) \geq 1 \\ 1 & , \text{ otherwise} \end{cases} \quad (3.18)$$

independent of primary energy, mass and distance to the core.

The likelihood function to maximize, gathering the sampled information of detector i at distance r_i is:

$$L = \prod_i f_P(n_i, \mu_i) \prod_i f_G(n_i, \mu_i) \prod_i F_{\text{sat}}(n_i, \mu_i) \prod_i F_{\text{zero}}(n_i, \mu_i) \quad (3.19)$$

and thus the log-likelihood function is

$$\ell = \sum_i \ln f_P(n_i, \mu_i) + \sum_i \ln f_G(n_i, \mu_i) + \sum_i \ln F_{\text{sat}}(n_i, \mu_i) + \sum_i \ln F_{\text{zero}}(n_i, \mu_i) \quad (3.20)$$

with n_i the effective number of particles detected in the detector as obtained using equation (3.17), and μ_i the corresponding LDF expectation. The four members of this equation represent each of the four possibilities for a station signal:

- **Small signals:** ($n < 30$ particles) A Poissonian distribution is assumed

$$\ln f_P(n_i, \mu_i) = n_i \ln \mu_i - \mu_i - \sum_{j=1}^{n_i} \ln j \quad (3.21)$$

- **Large signals:** For signals larger than 15 VEM ($n > 30$ particles), a Gaussian behaviour is assumed

$$\ln f_G(n_i, \mu_i) = -\frac{(n_i - \mu_i)^2}{2\sigma_i^2} - \ln \sigma_i - \frac{1}{2} \ln 2\pi \quad (3.22)$$

- **Saturated signals:** The saturated signal, n_i , represents a lower limit to the actual signal. The saturation of the station is caused by the overflow of the FADC read-out electronics with a finite dynamic range and a modification of the signal due to the transition of the PMTs from a linear to a non-linear behaviour. We integrate f_G over all possible values above the lower limit, to get an estimate of the probability of detecting a larger signal

$$F_{\text{sat}}(n_i, \mu_i) = \int_{n_i}^{\infty} f_G(n, \mu_i) dn = \frac{1}{2} \operatorname{erfc} \left(\frac{n_i - \mu_i}{\sqrt{2}\sigma_i} \right) \quad (3.23)$$

where $\operatorname{erfc}(x) = 1 - \operatorname{erf}(x)$ is the complementary error function. If the missing signal is recovered using *saturation recovery* [124], it is used as a lower limit if the rise of the LDF turns out to be too large.

- **Zero-signal stations:** We assume a threshold of $n_i \geq n_{\text{th}} \equiv 3$ to trigger a detector. Therefore, we have to sum over all Poissonian probabilities with a predicted number of particles μ_i and actual number of particles below n_{th}

$$\ln F_{\text{zero}}(n_{\text{th}}, \mu_i) = -\mu_i + \ln \sum_{n=0}^{n_{\text{th}}} \frac{\mu_i^n}{n!} \quad (3.24)$$

3.4 Energy estimation

The aim of the LDF estimation is to determine the core location and $S(r_{\text{opt}})$ (see equation 3.2). As we discussed in section 3.3, r_{opt} depends on the spacing between WCD, being 1000 m for 1.5 km ($S(450)$ for the *Infilled array*). This distance is chosen because there the interpolated signal has less dependence on the LDF function assumption and on the primary.

At different zenith angles air showers are observed at the ground at different ages. Therefore $S(1000)$ depends on the zenith angle, θ , of the shower (see figure 3.10). For the same primary particle energy $S(1000)$ for small zenith angles is larger than for bigger ones. The signals $S(1000)$ and $S(450)$ are corrected for their zenith angle dependence with a third degree polynomial in $x = (\cos^2(\theta) - \cos^2(38))$ [125]:

$$f(\theta) = 1 + ax + bx^2 + cx^3 \quad (3.25)$$

where

$$a = 0.980 \pm 0.004$$

$$b = -1.68 \pm 0.01$$

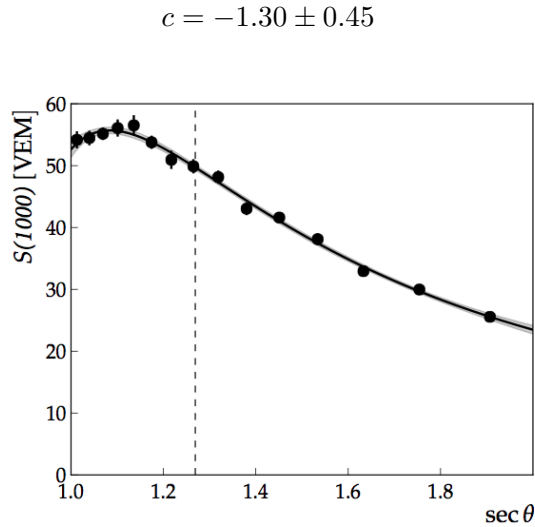


Figure 3.10: Attenuation curve described by a third degree polynomial function in $x = (\cos^2(\theta) - \cos^2(38))$. In this example the polynomial coefficients are deduced from $S(1000)$ dependence at $S_{38} \sim 50$ VEM which corresponds to an energy of about 10.5 EeV [86].

The equivalent signal at median zenith angle of 38° (35°), S_{38} (S_{35}), will be the energy estimator used for *vertical events*. Note that for the 750 m array, only events with zenith angle below 55° are accepted.

Inclined air-showers are characterised by the dominance of secondary muons at the ground, as the electromagnetic component is largely absorbed in the large atmospheric depth traversed by the shower. The reconstruction of *inclined events* is performed with a different method. It is based on the estimate of the relative muon content, N_{19} , with respect to a simulated proton shower with energy 10^{19} eV [126].

A subset of high-quality events is used, the so-called *Golden hybrids events*, to calibrate the SD energy estimators with the calorimetric measurement obtained with the FD, E_{FD} . They are high-quality hybrid T5 events which have an energy above the SD full efficiency trigger threshold [91] registered between January 2004 and December 2013. The correlation between the two variables is obtained from a maximum likelihood method which takes into account the evolution of uncertainties with energy, as well as event migrations due to the finite energy resolution of the SD. The relation between the different SD energy estimators and E_{FD} is well described by a single power-law function,

$$E_{FD} = A(S)^B \quad (3.26)$$

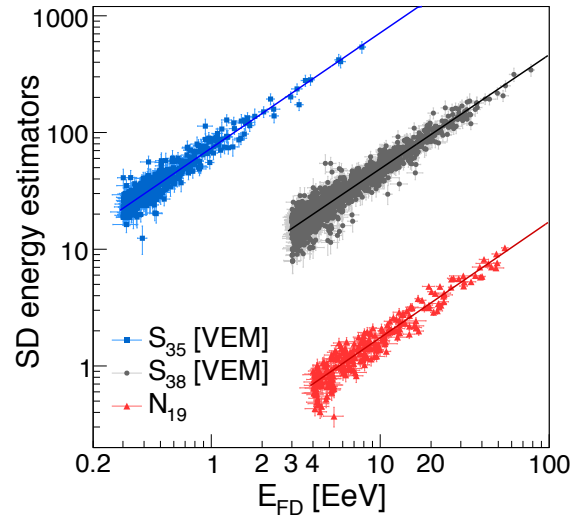
where the resulting parameters for the different observables are given in table 3.1. The correlation between the different energy estimators and E_{FD} is shown in figure 3.11(a). In the table is also illustrated the energy threshold for the different data sets. In the case of the SD data: *vertical*, *inclined* and *Infilled* this value represents the energy at which the

trigger efficiency is more than 98%. In the case of hybrid measurement at this energy the dependency of the flux on the primary mass is less than 10%. The exposures for the different data sets are also shown in figure 3.11(b).

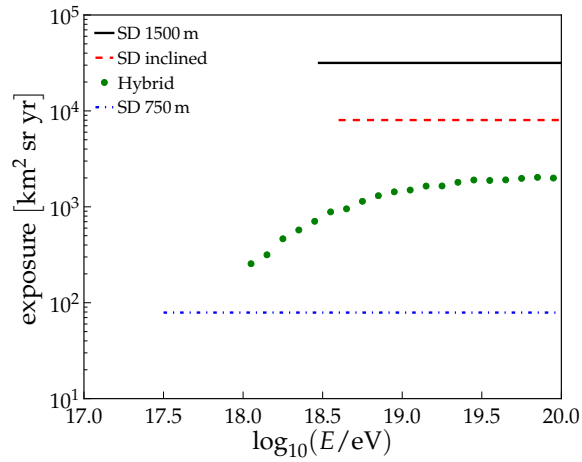
The resolution of the final SD energy estimator can be inferred from the distribution of the ratio E_{SD}/E_{FD} . Using the FD energy resolution of 7.6%, the resulting SD energy resolution with its statistical uncertainty is also shown in table 3.1. The systematic uncertainty on the SD energy due to the calibration is better than 2% over the whole energy range. The main contribution to the systematic uncertainty in the SD energy comes from E_{FD} which is 14% (see figure 3.12).

	Auger SD			Auger Hybrid
	1500m vertical	1500m inclined	750m vertical	
Data taking period	01/2004 - 12/2014	01/2004 - 12/2013	08/2008 - 12/2014	11/2005 - 12/2013
Exposure [km ² sr yr]	42500±1300	10900±300	150±5	1500±20 at 10 ¹⁹ eV
Zenith angles [°]	0 - 60	60 - 80	0 - 55	0 - 60
Threshold energy E_{eff} [eV]	3×10^{18}	4×10^{18}	3×10^{17}	10^{18}
No. of events ($E > E_{\text{eff}}$)	82318	11074	29585	11155
No. of hybrid events	1731	255	469	-
Energy scale (A) [EeV]	0.1871±0.0004	5.71±0.09	12.87±0.63 [PeV]	-
Energy scale (B)	1.023±0.006	1.01±0.02	1.013±0.013	-
Energy resolution [%]	15.3±0.4	19±1	13±1	-

Table 3.1: Summary of the experimental parameters regarding energy measurement in the Pierre Auger Observatory [32].



(a)



(b)

Figure 3.11: *Top*: Energy calibration in the Pierre Auger Observatory. Correlation between S_{38} , S_{35} and N_{19} with the energy measured by the FD [32]. *Bottom*: Integrated exposures for the different data sets [106].

Systematic uncertainties on the energy scale	
Absolute fluorescence yield	3.4%
Fluor. spectrum and quenching param.	1.1%
Sub total (Fluorescence yield - sec. 2)	3.6%
Aerosol optical depth	3% ÷ 6%
Aerosol phase function	1%
Wavelength depend. of aerosol scatt.	0.5%
Atmospheric density profile	1%
Sub total (Atmosphere - sec. 3)	3.4% ÷ 6.2%
Absolute FD calibration	9%
Nightly relative calibration	2%
Optical efficiency	3.5%
Sub total (FD calibration - sec. 4)	9.9%
Folding with point spread function	5%
Multiple scattering model	1%
Simulation bias	2%
Constraints in the Gaisser-Hillas fit	3.5% ÷ 1%
Sub total (FD profile rec. - sec. 5)	6.5% ÷ 5.6%
Invisible energy (sec. 6)	3% ÷ 1.5%
Stat. error of the SD calib. fit (sec. 7)	0.7% ÷ 1.8%
Stability of the energy scale (sec. 7)	5%
Total	14%

Figure 3.12: Systematic uncertainties on the energy scale [105].

Revision of the Muon Arrival Time Model

4

When building muon production depth (MPD) distributions, the main source of distortion comes from effects like the under-sampling of the detector, the tank response to muons and the geometry resolution [1, 127, 128, 129]. To a lesser extent, the algorithm used to convert the times recorded by the SD detectors into muon production distances also deforms the average profile¹. In this chapter, we quantify among other things, the distortion introduced by the latter effect using events with energies and zenith angles similar to those used in the standard analysis (i.e., $E > 20$ EeV and $\theta \in [55^\circ, 65^\circ]$) and two different distance cuts: $r > 1000$ m or $r > 1700$ m. In particular, we follow a new approach to parameterize the kinematic delay that results in a sensible reduction of the bias associated to the reconstructed X_{\max}^μ .

4.1 MPD algorithm

With the SD detectors we can measure the positions and the times of muons reaching the ground. Starting from this information the challenge is to reconstruct the longitudinal development of the muonic cascade (i.e. the MPD distribution).

Muons mainly come from the decay of pions (99% branching ratio) and kaons (64% branching ratio). Also kaons will decay to charged pions with a 28.5% probability. Hence, we can assume that muons mainly come from pions. The muon parent particle travels at a certain angle from the shower axis. The muon time measured at the ground therefore contains the path traveled before decaying by the parent particle:

$$t_{SD} = t_\pi + t_\mu \quad (4.1)$$

In our calculation it is mandatory to set a reference start time. This time is the arrival time of the shower front at the position of each muon at the ground (t_{front}). Thus, we will be interested in relative quantities (i.e. time delays). They are defined as:

$$\tau = (t_\pi + t_\mu) - t_{\text{front}} \quad (4.2)$$

¹i.e., the sum of all the MPD distributions for a particular energy bin

The breakdown of the different contributions to the total delay is summarized in what follows [130]:

- **Geometric delay, τ_g :** The minimum muon time of flight assuming they are massless particles (their velocity equals speed of light) that suffer no interactions and therefore follow trajectories that are straight lines. This delay can be deduced by means of elementary geometry arguments (see figure 4.1):

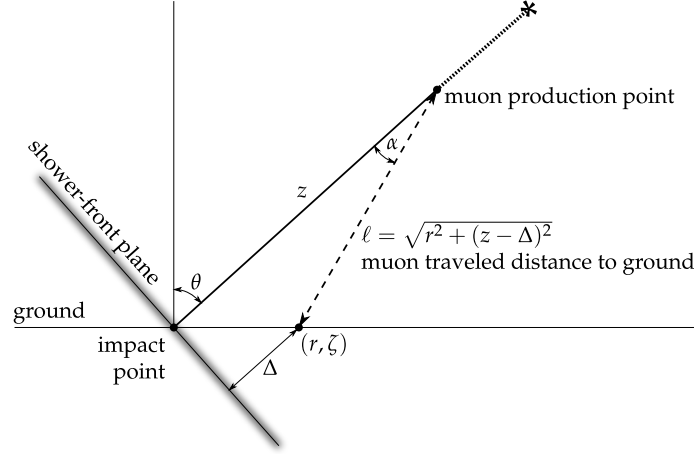


Figure 4.1: Geometry used to obtain the muon traveled distance.

$$l = \sqrt{r^2 + (z - \Delta)^2} = ct_{min} \quad (4.3)$$

where t_{min} would be the minimum arrival time of the muon at the ground, r the distance to the shower core and z the production point of the muon along the shower axis.

The difference of this time with respect to the arrival time of the shower front is the geometric delay:

$$c\tau_g = (ct_{min} + ct_\pi) - ct_{front} = (ct_{min} + ct_\pi) - (z_{1st} - \Delta) = \sqrt{r^2 + (z - \Delta)^2} - (z - \Delta) \quad (4.4)$$

where Δ represents the distance from the point at the ground to the shower front, defined as a plane traveling at the speed of light perpendicularly to the shower axis. For a horizontal ground plane:

$$\Delta(r, \zeta) = r \cos \zeta \tan \theta \quad (4.5)$$

θ being the zenith angle of the air-shower, ζ the azimuth and r the distance to the shower axis. ct_π is the path traveled by the parent particle. Given the position of the first interaction point on the shower axis, z_{1st} , then

$$ct_\pi = z_{1st} - z. \quad (4.6)$$

- **Kinematic delay, τ_{kin} :** Muons are massive particles so they do not travel at the speed of light and their velocities depend on energy. They lose energy mainly because of inelastic collisions with atomic electrons in the air. The difference with respect to a particle traveling at c is referred to as τ_{kin} .
- **Multiple scattering, τ_{MS} :** Muons do not travel following straight lines. Whilst traversing the air they can be deflected mainly due to Coulomb scattering from nuclei.
- **Geomagnetic field, τ_B :** Muons are charged particles so the magnetic field of the Earth affects their trajectories, changing the impact point on the ground and delaying the arrival time. The longer the path of the muon, the larger this effect. Hence it is especially important for very inclined events.

The total delay of the muon at the ground is:

$$\tau = \tau_g + \tau_{kin} + \tau_B + \tau_{Rem} \quad (4.7)$$

where τ_{Rem} accounts for multiple scattering and any remanent delay due to model approximations.

The goal of the reconstruction is to obtain the production distance from the measured delay which contains all the contributions. The simplest way to do it will be by inverting equation (4.4), the mapping between production distances and geometric delays:

$$z = \frac{1}{2} \left(\frac{r^2}{c\tau_g} - c\tau_g \right) + \Delta \quad (4.8)$$

where τ_g can be obtained from equation (4.7):

$$\tau_g = \tau - \tau_{kin} - \tau_B - \tau_{Rem} \quad (4.9)$$

Nevertheless, to use this expression we need to evaluate the different corrections to the total delay. Figure 4.2 shows the relative contributions for each of those corrections.

The effect of multiple scattering is negligible for muons in air and the deflection suffered by them caused by the geomagnetic field is also small at this zenith range. Hence the contribution of the last two terms, $\tau_B + \tau_{Rem}$ is of the order of a few percent respect to the other ones (see figure 4.2).

Therefore, the main correction to the geometric delay is the kinematic one. Assuming constant energy losses along their flying path, $E_f = E_i - \rho a l$, it can be expressed as:

$$\tau_{kin} = \frac{1}{c\rho a} \left[\sqrt{E_i^2 - [mc^2]} - \sqrt{E_f^2 - [mc^2]} \right] - \frac{l}{c} \quad (4.10)$$

where E_i is the muon energy at production, E_f the muon energy at ground, a the stopping power (considered constant here) and ρ the atmospheric density. At larger distances (more than 1000 m) and for zenith angles $\theta = 60^\circ$, the kinematic delay typically amounts to less than 30% of the total delay [130], becoming negligible as we go farther from the core. It

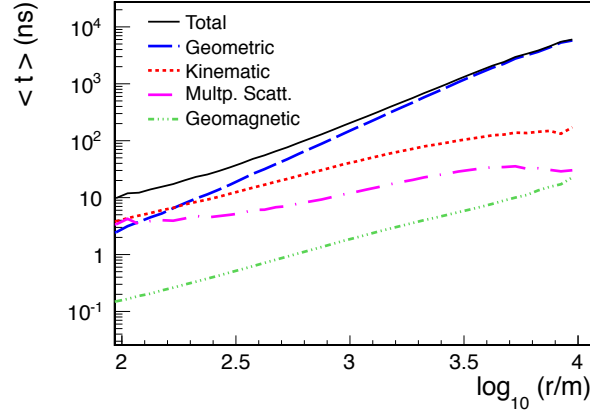


Figure 4.2: Average time delay of muons with a breakdown of the different contributions. The curves correspond to a proton-initiated shower with a zenith angle of 60° and primary energy of $E = 10$ EeV [130].

dominates close to the core, where obviously the geometric delay tends to zero and the mean muon delay is dominated by the huge amount of low energy muons (see figure 4.3).

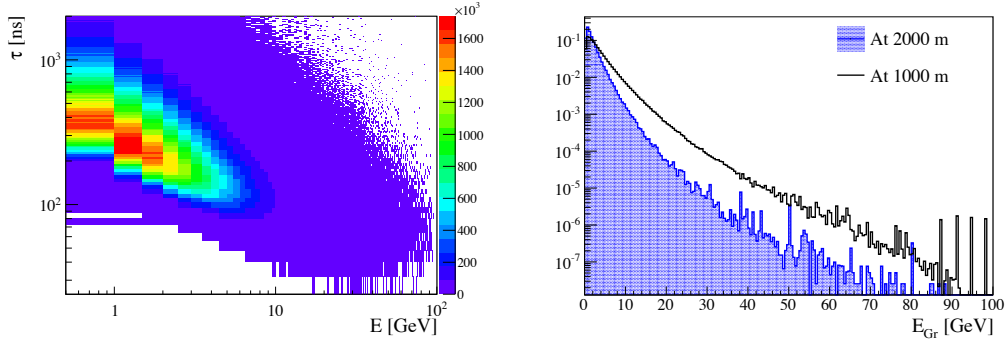


Figure 4.3: *Left*: Delay vs energy at the ground for distances to the shower core larger than 1000 m. *Right*: Ground energy distribution at 1000 m (empty histogram) and at 2000 m (filled histogram).

Therefore the estimation of the production distance will require an accurate determination of the kinematic delay:

$$\tau_g \approx \tau - \tau_{kin} \quad (4.11)$$

where τ_{kin} can be obtained using equation (4.10). Now, equation (4.8) can be rewritten as:

$$z \approx \frac{1}{2} \left[\frac{r^2}{c\tau - c\tau_{kin}} - (c\tau - c\tau_{kin}) \right] + \Delta \quad (4.12)$$

In figure 4.4 we illustrate the goodness of this approach to deduce the true MPD distribution. Dashed line represents the distribution without applying the kinematic correction, assuming muons traveling at c . The shape is distorted and widened compare to the true distribution (continuous line). A dramatic overestimation of X_{\max}^{μ} is observed. However, if we apply the kinematic correction of equation (4.10) (filled histogram) we reproduce almost perfectly the true distribution at the two selected distances. The difference with respect to the true one comes from the approximations done.

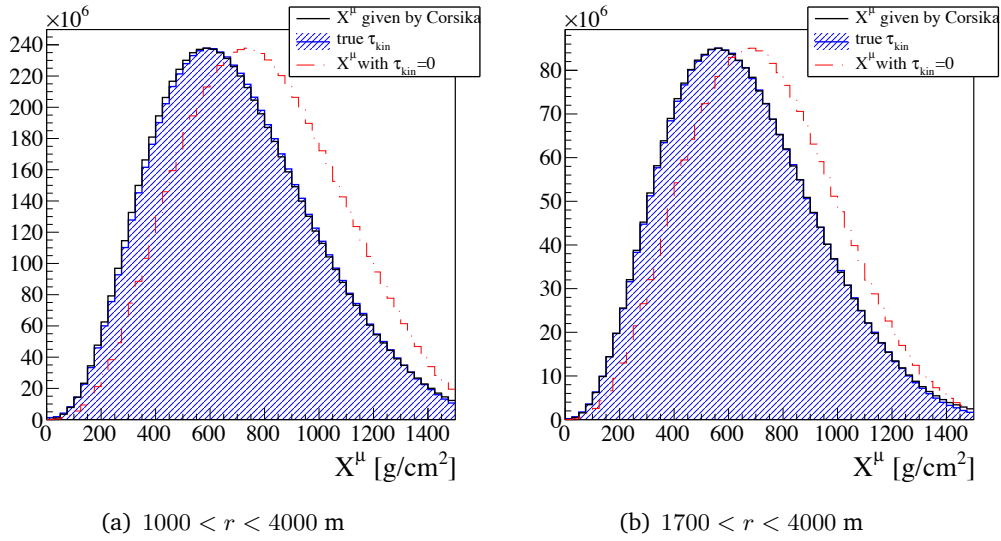


Figure 4.4: MPD distributions obtained for proton QGSJET-II-04 at $10^{19.75}$ eV. Continuous line represents the distribution obtained directly from the simulations (the true distribution). The dashed line is the one obtained using only the geometric delay, and the filled histogram is the result after correcting by the kinematic delay, using equation (4.10).

In what follows we further investigate the role of the kinematic delay in the reconstruction of the muon production distances.

4.2 Classic (previous) kinematic delay parameterization

In order to compute the kinematic delay, an estimation of the muon energy is required. Unfortunately, with the Auger SD detectors we can not get a direct measurement of this quantity. Thus, to account for it in the reconstruction algorithm, an approximation is performed. In the standard analysis, the average kinematic delay $\langle t_{kin} \rangle$ is parameterized using an approximation of the muon energy spectrum [131] (in this reference they referred to the kinematic delay as $\langle t_{\varepsilon} \rangle$). In [132, 131] it is parameterized as a function of the distance to the shower core and the production distance arriving to the following expression:

$$\langle t_{\varepsilon} \rangle = \frac{1}{2c} \frac{r^2}{l} \epsilon(r, z - \Delta) \quad (4.13)$$

where the function $\epsilon(r, z - \Delta)$ is:

$$\epsilon(r, z - \Delta) = p_0(z) \left(\frac{r}{[\text{m}]} \right)^{p_1} \quad (4.14)$$

with

$$\log_{10}(p_0(z)) = -0.6085 + 1.955 \log_{10}(z/[\text{m}]) - 0.3299 \log_{10}^2(z/[\text{m}]) + 0.0186 \log_{10}^3(z/[\text{m}]) \quad (4.15)$$

$$\log_{10}(p_1) = -1.176 \quad (4.16)$$

For the case where $z - \Delta \gg r$, it gives the ratio of the average kinematic delay to the geometric delay at a given position [131]. Figure 4.5 shows the evolution of ϵ as a function of r and $z - \Delta$. As expected, the value of ϵ is larger near the shower core, where the kinematic delay contribution is more significant.

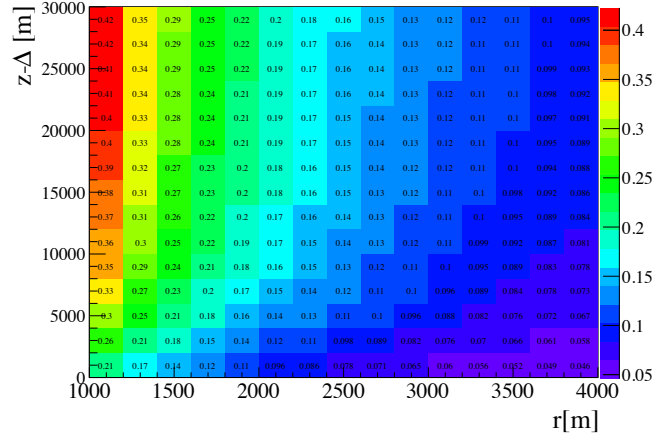


Figure 4.5: ϵ as a function of r and as a function of $z - \Delta$ for proton QGSJET-II-04 at $10^{19.75}$ eV and around 60° .

Therefore, we rewrite equation (4.8) as:

$$z \approx \frac{1}{2} \left[\frac{r^2}{c\tau - c\langle t_\epsilon \rangle} - (c\tau - c\langle t_\epsilon \rangle) \right] + \Delta \quad (4.17)$$

This equation does not have an analytical solution because $\langle t_\epsilon \rangle$ depends on z . An iterative process can be applied to solve it:

- As a first approximation, z is obtained taking zero kinematical delay.
- Then the average kinematic delay is computed for this value of z and r by using equation (4.13).
- A first estimation of the geometric delay is obtained: $\tau_g \approx \tau - \langle t_\epsilon \rangle$.
- Finally, the obtained value is substituted in equation (4.17)

This process converges quickly and a couple of iterations are enough to obtain an accurate estimate of z .

The effect of this correction over the average MPD distribution at the two selected distances is presented in figure 4.6 for proton QGSJET-II-04 air-showers with energies equal to $10^{19.75}$ eV. The continuous line corresponds to the average distribution obtained after correcting with the true kinematic delay, while the broken line is the result obtained after applying equation (4.13). The correction is overestimating the kinematic delay which implies that the depth is underestimated, especially close to the shower core, where τ_{kin} is more important. Moreover, if we compute the contributions to the mean and the RMS of X_{max}^μ , at 1000 m the difference with the true value is -70 g/cm² and the width is distorted from 283 g/cm² to 302 g/cm². For the standard analysis distance, $r > 1700$ m, the contributions are -30 g/cm² in the mean value and the RMS changes from 269 g/cm² to 276 g/cm².

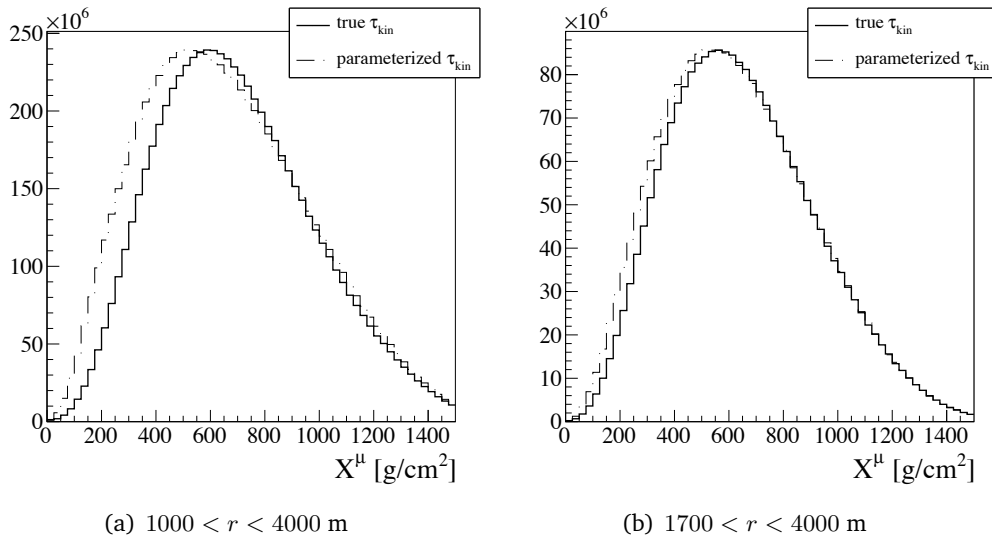


Figure 4.6: Effect of the kinematic delay parameterization over the average MPD distribution in the standard analysis, at distances larger than 1700 m (right plot), and from 1000 m (left plot). Black line is the distribution obtained using the true kinematic delay and the dot-dashed line is the one obtained using the parameterization.

To understand this distortion the average profile is studied in different energy and production distance ranges. In figure 4.7, we plot the MPD distribution obtained with the true kinematic delay and the one from the parameterized delay at different production distance regions. The correction works better for small z (larger depths). At far distances from the core, we are dominated by low energy muons which implies smaller z . This effect combined with the narrower energy distributions obtained at farther distances, makes the distortion on the profile smaller for the standard analysis distance, 1700 m (see figure 4.6). In fact, if the same plot is done for different energy regions, see figure 4.8, we see that low energy muons are better reproduced. Furthermore, one can observe that low energy muons are the main contribution to the full MPD distribution. Muons below 2 GeV represent 55% of the total at 1000 m and 67% at 1700 m.

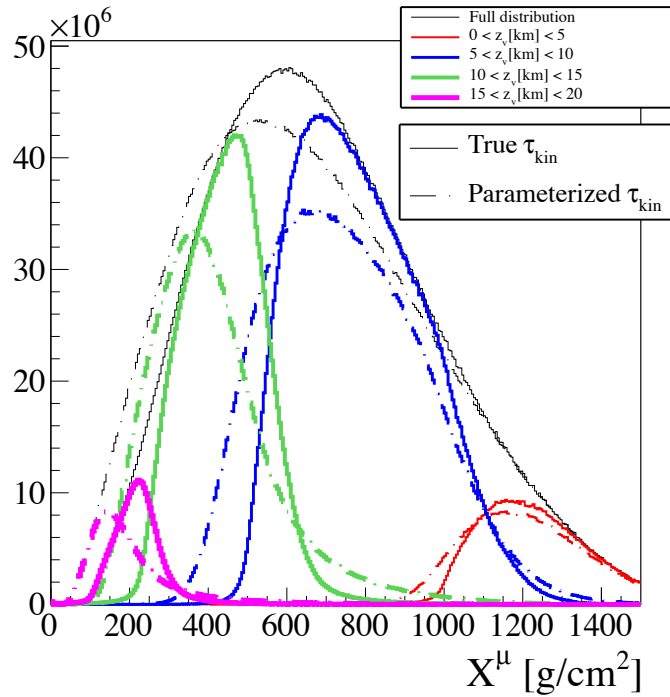
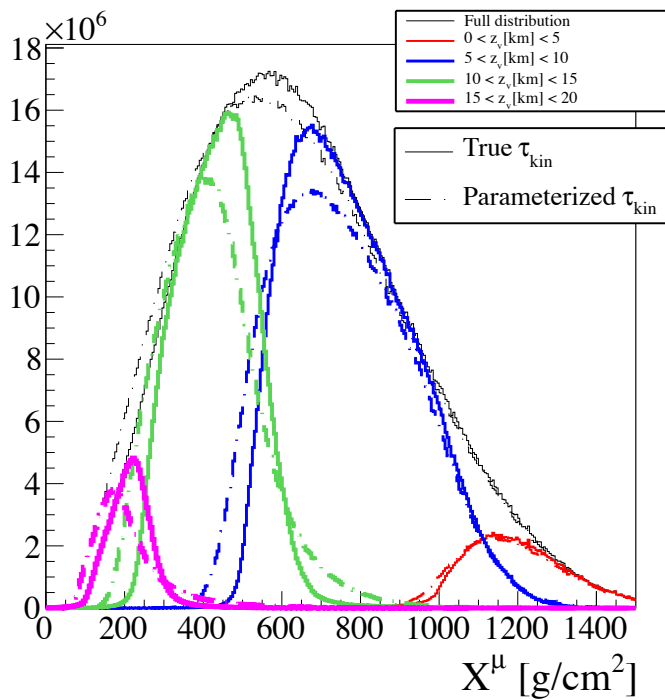
(a) $1000 < r < 4000$ m(b) $1700 < r < 4000$ m

Figure 4.7: (Similar to figure 4.6) Kinematic delay parameterization effect over the average MPD distribution for different production distances ranges.

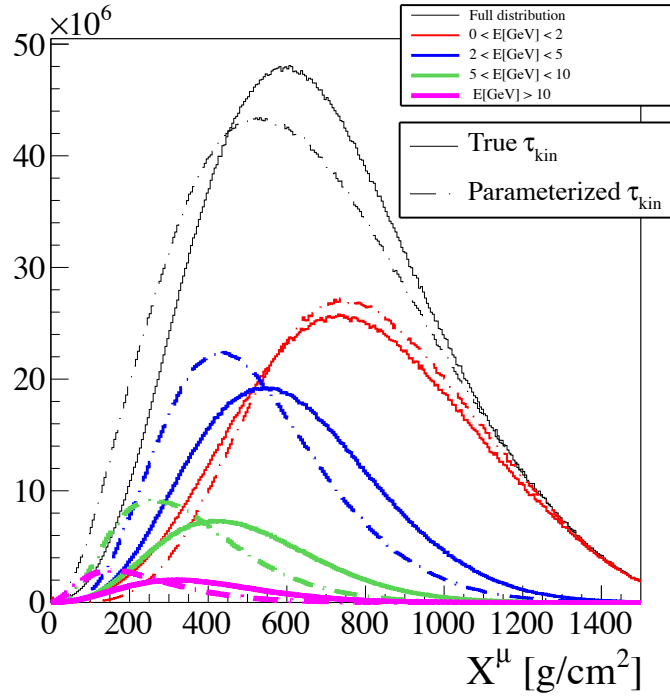
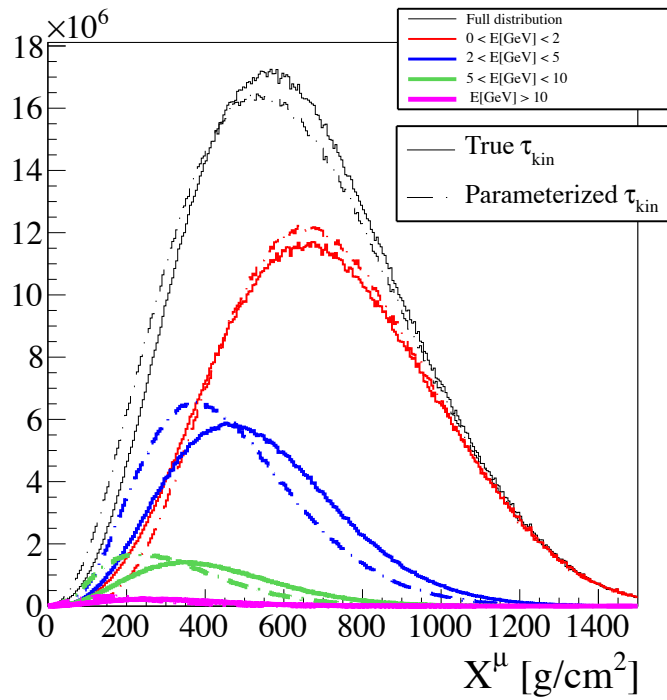
(a) $1000 < r < 4000$ m(b) $1700 < r < 4000$ m

Figure 4.8: (Similar to figure 4.6) Kinematic delay parameterization effect over the average MPD distribution effect for different ground energy ranges.

4.3 New approach

As we want to reproduce as accurately as possible the shape of the average MPD distribution we study a new approach to perform the correction of the kinematic effects in the distribution. In addition, we want to minimize the contribution of these effects in the average X_{\max}^{μ} . Equation 4.8 relates times with production distances. Thus, to determine the true production distance, z , the key point is to be able to have an accurate estimation of the geometric delay. As already discussed, for this zenith range the main correction to τ_g comes from the kinematic delay, thus it can be approximated as: $\tau_g \approx \tau - \tau_{kin}$. In the WCD we measure the delay of the muon with respect to the arrival time of the shower front plane, τ_m . In addition, from CORSIKA simulations we can estimate the true kinematic delay of the muon. Thus, we can tabulate the kinematic delay as a function of the measured delay and as a function of the distance to the shower core. Sixty seven proton QGSJET-II-04 showers at 10 EeV and around 60° have been used to tabulate the average kinematic delay (see figure 4.9). The table is built for distances between 1000 m and 4000 m in intervals of 100 m. The measured delay is tabulated between 0 and 5000 ns in intervals of 100 ns. That region was selected to minimize the average bias and to avoid discarding any of the generated muons.

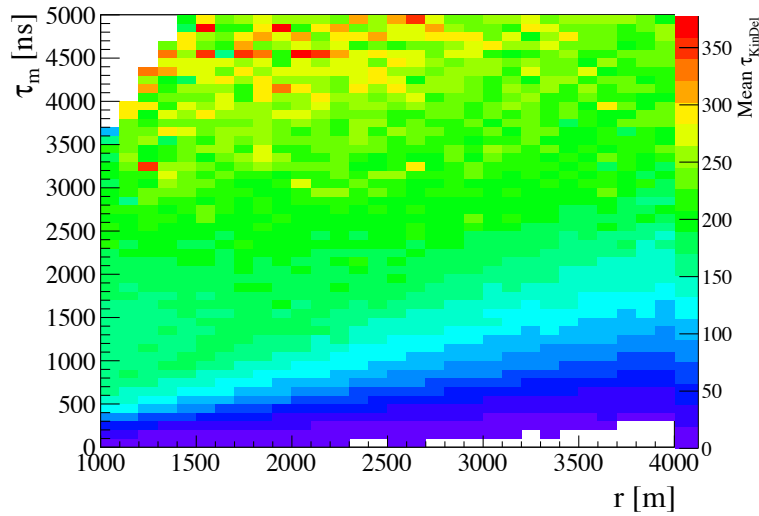


Figure 4.9: Average kinematic delay as a function to the measured delay and the distance to the shower core.

In this new approach we do not need an iterative process as was needed with the standard parameterization, we just correct the measured delay obtained from the reconstruction with the average kinematic delay obtained from the table:

- We obtain τ_m as the delay of the muon with respect to the arrival time of the shower front plane, assuming zero kinematical delay.
- The average kinematic delay is computed for that value of τ_m and r using the values in figure 4.9.

- The geometric delay is obtained as: $\tau_g \approx \tau_m - \langle \tau_{kinDel} \rangle$.
- Finally, the computed τ_g value is substituted in equation (4.17)

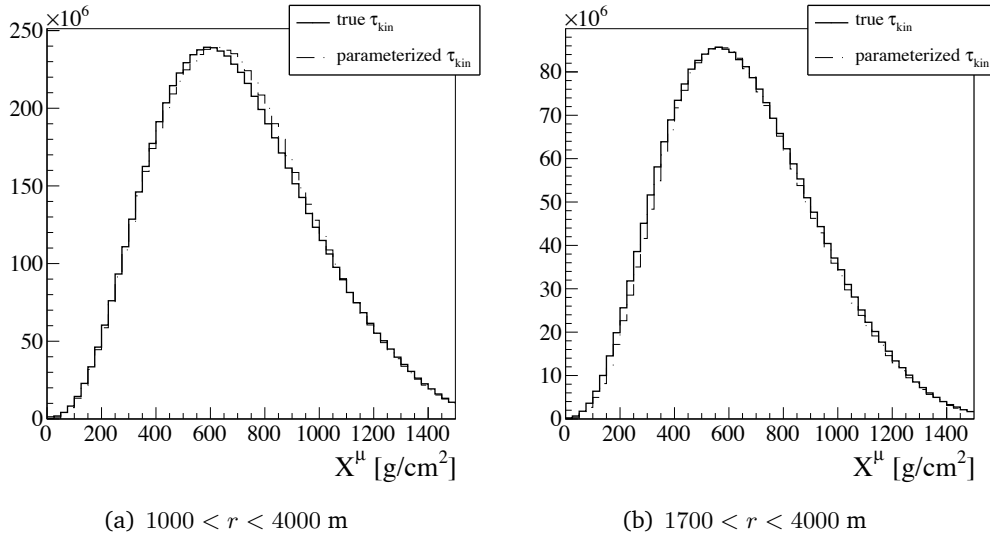


Figure 4.10: Effect of the new approximation used to compute the kinematic delay for two distance ranges. Black line is the distribution obtained using the true kinematic delay from the simulations and the dot-dashed line is the one obtained using the new approximation.

Figure 4.10 illustrates the effect of our kinematic delay correction for the two selected distances to the core for proton QGSJET-II-04 at $10^{19.75}$ eV. Now, the distortion of the profile has been sensibly diminished reproducing quite accurately the shape of the original distribution. Applying the new parameterization, the bias to the mean X_{max}^μ and its RMS are significantly reduced. At 1000 m, the bias amounts to 7 g/cm² and for the standard analysis distance cut it is 5 g/cm². Regarding the RMS, it changes from 283 g/cm² to 280 at 1000 m, and from 269 g/cm² to 264 g/cm² at 1700 m. In addition, a breakdown of the different production distance and energy contributions is represented in figures 4.11 and 4.12. The improvement when compared to the results in figures 4.7 and 4.8 is remarkable.

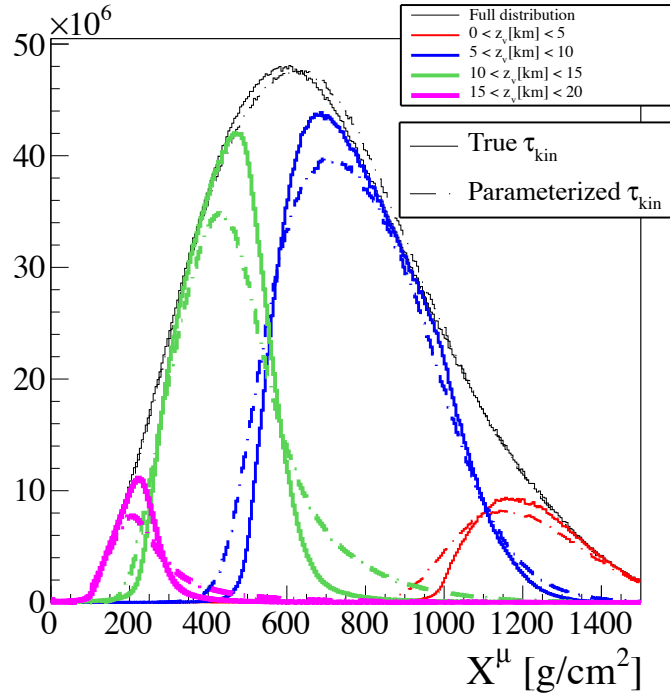
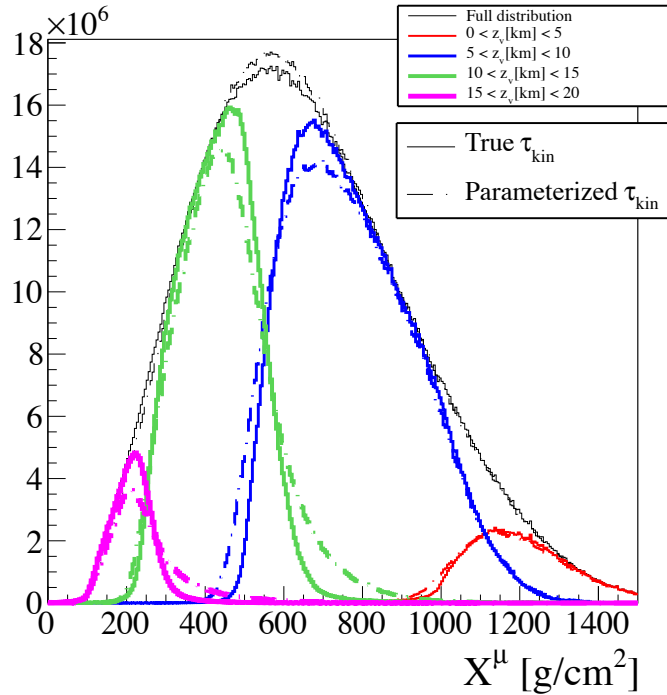
(a) $1000 < r < 4000$ m(b) $1700 < r < 4000$ m

Figure 4.11: (Similar to figure 4.10) Effect of the new kinematic delay approximation for different production distances ranges. Curves correspond to proton showers with an energy of $10^{19.75}$ eV.

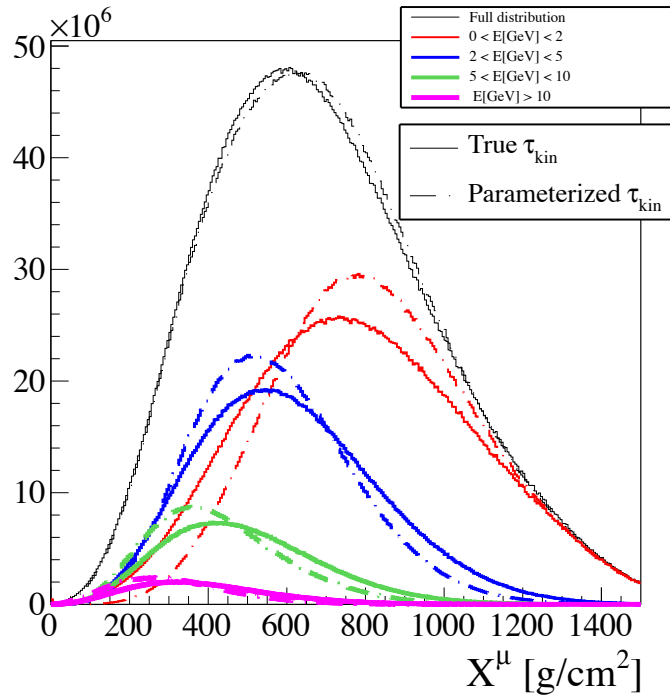
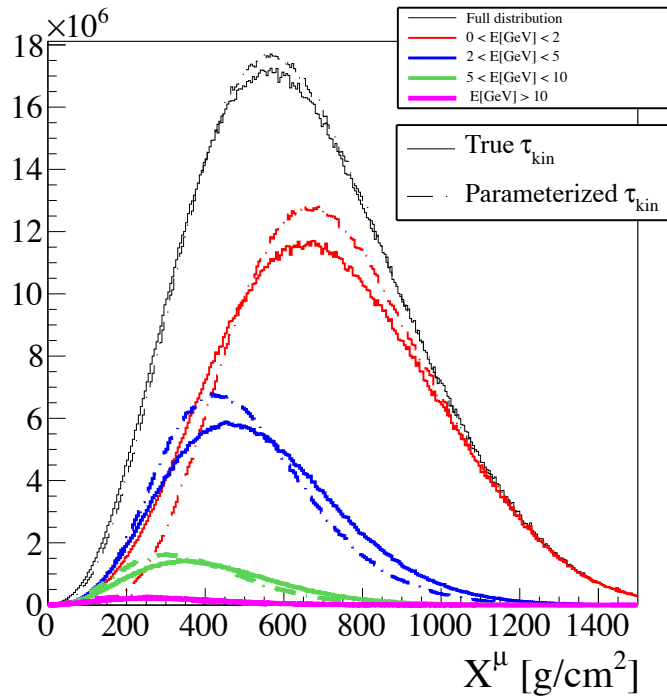
(a) $1000 < r < 4000$ m(b) $1700 < r < 4000$ m

Figure 4.12: (Similar to figure 4.10) Effect of the new kinematic delay approximation for different ground energy ranges. Curves correspond to proton showers with an energy of $10^{19.75}$ eV.

We have studied the systematic effects associated to our parameterization. It was done using proton QGSJET-II-04 air-showers around 10^{19} eV and 60° . Below, we list and quantify the most relevant sources contributing to the overall systematic uncertainty:

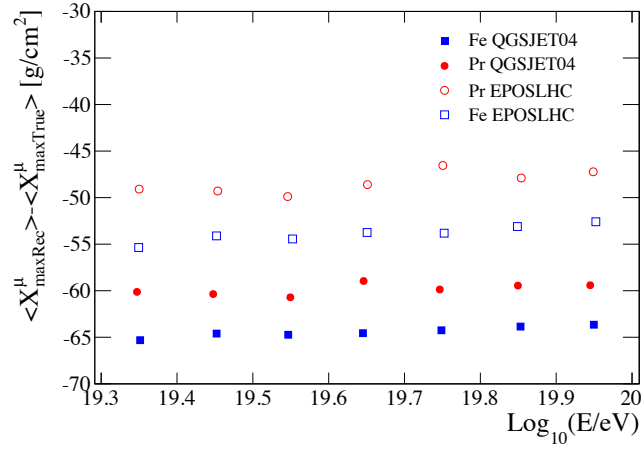
- **Energy:** We have applied this parameterization in our energy range (above 20 EeV). We have found a maximum difference between the reconstructed and the true X_{\max}^μ of 4 g/cm^2 compared to the bias obtained at the reference energy for the parameterization (i.e., 10 EeV).
- **Primary particle:** When we apply our parameterization to reconstruct iron QGSJET-II-04 air showers, the difference in X_{\max}^μ compared to the proton value is 10 g/cm^2 . In an attempt to reduce this number a parameterization using a mixed composition (50% proton and 50% iron) was performed obtaining no improvement with respect to the parameterization done with proton.
- **Hadronic model:** We have reconstructed EPOS LHC air-showers using our parameterization. We find a deviation in the value of X_{\max}^μ of 1 g/cm^2 .

Table 4.1 summarizes all the sources contributing to the systematic uncertainty in the bias of X_{\max}^μ . The overall value amounts to 10 g/cm^2 .

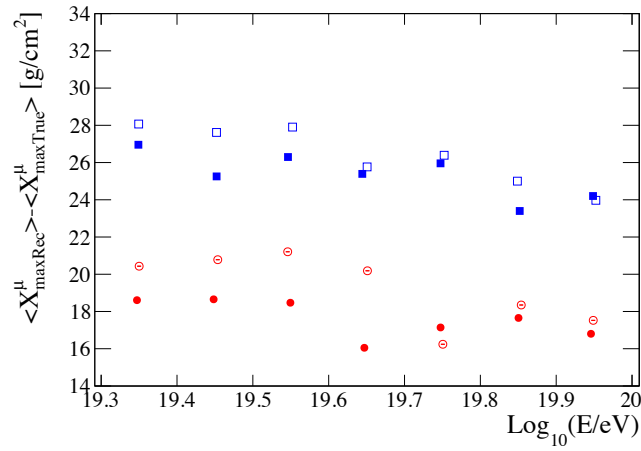
Source	Sys. uncertainty [g/cm^2]
Energy	4
Primary	9
Hadronic model	1
Total	10

Table 4.1: Contributions to the systematic uncertainty in the new estimation of τ_{kin} for the reconstruction of X_{\max}^μ .

Throughout the chapter we were taking as reference the MPD obtained using the time model and the true kinematic delay to isolate the effects of the kinematic corrections. However, when we perform the MPD reconstruction our reference lines are the maximum of the distributions obtained directly from CORSIKA. In figure 4.13 we compute the difference between our reference value, the true X_{\max}^μ , and the one obtained after correcting the kinematic effects with each of the discussed parameterizations, the old and the new one proposed in this work. As shown in figure 4.13 the bias is reduced significantly when the new approach is used. The spread between primaries and models is similar, around 10 g/cm^2 .



(a) Old parameterization



(b) New parameterization

Figure 4.13: Difference between the X_{\max}^μ obtained from the reference distribution, $X_{\max\text{True}}^\mu$, and the one using each one of the kinematic delay parameterizations (the top plot corresponds to the old one and the bottom to the new one) for all species.

Muon Production Depth in an event by event basis

5

In the previous chapter we have described the theoretical framework supporting the MPD analysis. In this chapter we focus our interest in the reconstruction of the distributions at detector level. The features of the MPD distributions and the different steps followed in the event by event reconstruction are discussed in [1] and in [133]. Here, we assess the effect that each one of the intermediate steps have in the total bias and resolution we get by going from the true distribution (the one obtained directly from CORSIKA) to the final one after the whole reconstruction chain has been applied.

5.1 MPD features

The MPD is reconstructed using the FADC signals provided by the water-Cherenkov detectors. The finite area of the detectors induces fluctuations due to the different muon samples being collected. In addition, the shape of the MPD distribution observed from different positions at the ground varies because of differences in the probability of in-flight decay and because muons are not produced isotropically around the shower axis. It is an integration over distance, r , which enables the estimation of the dN_μ/dX distribution or MPD distribution (where N_μ refers to the number of produced muons). However, for discrete detector arrays, measurements at just a handful of r values are available and are limited to a small number of muons due to the finite collection surface (for Auger detectors it amounts to 10 m^2 cross-section for vertical incidence).

Another important feature of the MPD distribution observed at the ground is its dependence on the zenith angle. There are two reasons for this. The first one is due to the fact that inclined events mostly evolve in a less dense atmosphere than more vertical ones. This makes pions reach their critical energies (ϵ_c^π) at higher altitude resulting in a shallower development of the shower. While the shapes of the muon distribution at production are almost unaffected by this difference in reaching ϵ_c^π [130], their production depths are shifted. For proton showers at 10^{19} eV, the difference between the distribution of maxima for vertical and 60° events is approximately 20 g/cm^2 . The second and main reason for the zenith angle dependence of the MPD distribution is a consequence of the muon-decay probability. This effect influences not only the location of the maximum but also the shape of the observed MPD distribution. Figure 5.1 demonstrates this dependence

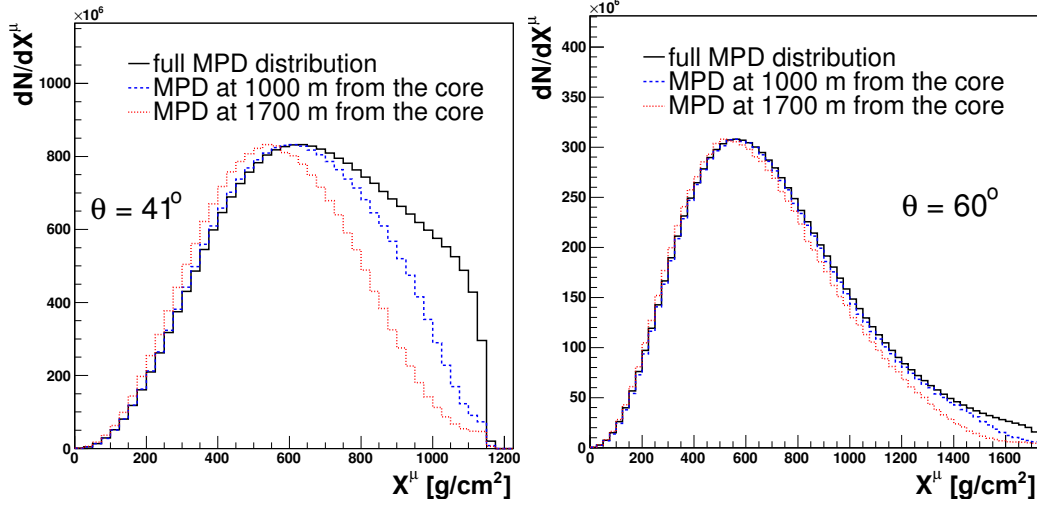


Figure 5.1: MPD distributions produced by an iron shower with energy $10^{19.5}$ eV impinging with two different zenith angles: 41° (left) and 60° (right). We use EPOS LHC[79] to model high-energy hadronic interactions. The shape of the MPD distribution shows a dependence with the distance to the shower core. This dependence is more pronounced for events with smaller inclinations. The histograms are normalized to have the same maxima [1].

for MPDs extracted from simulations at different zenith angles and at different distances r from the core. For zenith angles about 40° and lower, the shape of the MPD and the position of its maximum are a function of r . However, at zenith angles around 60° and above, the differences between the MPD distributions reconstructed at different distances to the core are small. At such angles, large z values dominate, diminishing the dependence of the traveled distance l on r .

5.1.1 Motivation of X_{\max}^μ as mass composition observable

With the aim of obtaining useful physics information from the MPD distribution, for each event we make a fit of the muon longitudinal development profile with a Gaisser-Hillas function [103]:

$$\frac{dN}{dX} = \frac{dN_{max}}{dX} \left(\frac{X - X_0}{X_{max}^\mu - X_0} \right)^{\frac{X_{max}^\mu - X_0}{\lambda}} e^{-\frac{X_{max}^\mu - X}{\lambda}} \quad (5.1)$$

Of the four parameters, X_{\max}^μ accounts for the point along the shower axis where the production of muons reaches its maximum as the shower develops through the atmosphere. This parameter will be the main physics observable for composition studies. In figure 5.2 we show the distribution of X_{\max}^μ for simulated proton and iron showers with 30 EeV of energy between 55° and 65° . The best set of parameters that describes a longitudinal muon profile (either at CORSIKA or at Offline level) is obtained through a log-likelihood maximization of the Gaisser-Hillas function. At CORSIKA level we fit all

the Gaisser-Hillas parameters. However, at detector level, the number of muons is not very large as we have explained previously. This under-sampling does not yield reliable estimates when all four parameters of the function are fitted. The solution adopted in the standard analysis to overcome this problem is to fix X_0 during the fitting procedure. According to simulations the preferred X_0 value depends on the nature of the primary particle. Trying to avoid a composition-dependent bias, this parameter is fixed to the average value obtained for proton and iron nuclei, $X_0 = -45 \text{ g/cm}^2$. This choice also constrains the λ values due to the correlation between X_0 and λ . Nevertheless, it does not represent a large systematic uncertainty source given the weak correlation between X_{max}^μ and X_0 . A shift of 10 g/cm^2 in X_0 corresponds to a variation of 1.5 g/cm^2 in the value of X_{max}^μ . Thus, as the mean difference in the X_0 values for proton and iron primaries is about 30 g/cm^2 , a maximum bias of 3 g/cm^2 is expected. The MPD distribution fit is performed in an interval of depths ranging from 0 to 1200 g/cm^2 , and it contains the entire range of possible values of X_{max}^μ .

The distribution of muon production depths varies as a function of the mass of the particle that initiates the atmospheric cascade. For heavier particles, the average value of X_{max}^μ is smaller and the distribution is narrower compared with that of lighter particles. The same behaviour is observed when considering different energies for the primary particle. According to simulations, X_{max}^μ allows us to study the mass composition of UHECR data collected by a surface array of particle detectors.

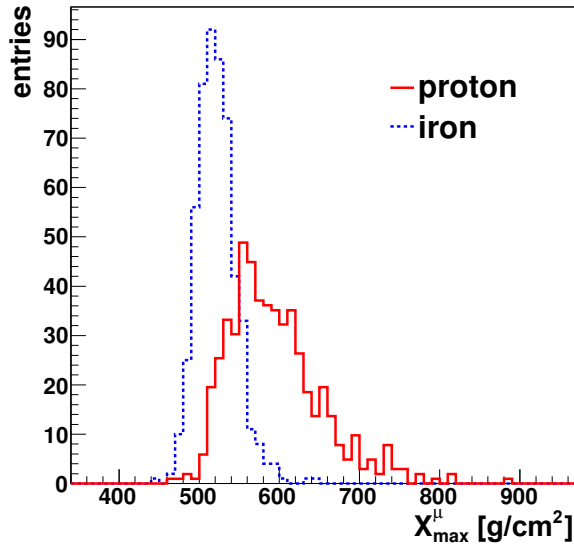


Figure 5.2: X_{max}^μ distribution for proton and iron showers simulated at 30 EeV with EPOS LHC at zenith angles between 55° and 65° . The mean value and the RMS of the distributions show a clear dependence on the mass of the primary cosmic ray. For the construction of the MPDs, only muons reaching the ground at distances greater than 1700 m were considered [1].

The MPD of an EAS accounts for its hadronic development in a similar way as the longitudinal development measured by the FD accounts for the electromagnetic (EM) one.

The EM component of a cascade is generated by photons from π_0 decays while muons mainly come from the decay of charged pions (99% of them have a muon in the final state) and kaons (64% of them have a final state muon). As already pointed out, also kaons will decay to charged pions with a 28.5% probability. Hence, we can assume that muons mainly come from charged pions. Analyzing the evolution of the electromagnetic part of the cascade we observe that the position of X_{max} depends mainly on the primary cross-section and the critical energy of pions ϵ_c^π (linked with the multiplicity), while the elongation rate will depend on the evolution of those quantities with energy. The point where the number of produced muons reaches the maximum must be regulated by the same effects, so we expect a similar evolution with energy, as it is observed in figure 5.3.

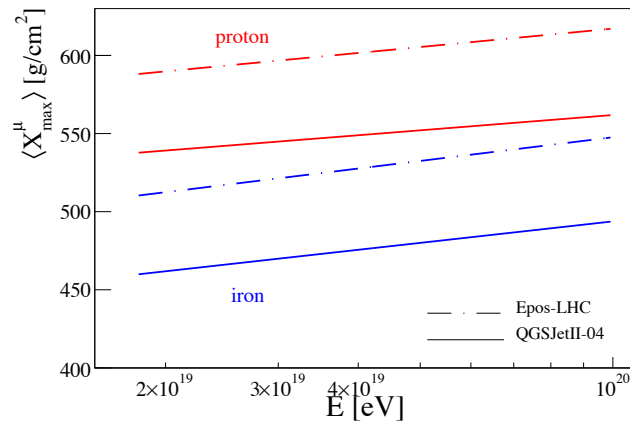


Figure 5.3: Evolution of X_{max}^μ as a function of the energy for proton and iron and two hadronic models.

5.2 MPD reconstruction for individual events

The signals recorded by the SD result from a mixture of muons and EM particles. The reconstruction of the MPD distribution for a given event requires the selection of the signal coming from muons. The EM signal is thus a background that must be eliminated. In inclined showers, around or above 60° , the EM component is heavily absorbed by the atmosphere. In addition, the dependence of the MPD shape with the distance to the shower axis r decreases drastically as θ increases. For these reasons, the standard MPD analysis [1] was performed in the interval of zenith angles $[55^\circ, 65^\circ]$.

In general, the EM signals are broader in time and with smaller amplitudes. A cut on signal threshold that rejects all time bins with signals below a certain value ($S_{threshold}$) helps to diminish the residual EM contamination. This $S_{threshold}$ was set to 15% of the maximum (peak) of the recorded signal. This cut, apart from minimizing potential baseline fluctuations, guarantees muon fractions above 85%, regardless of the energy and mass of the primary particle.

For each entry in the MPD distribution, the uncertainty introduced in X^μ (δX^μ) is a

function of the time resolution (δt) and the accuracy of the reconstruction of the shower angle and core location. The uncertainty in time gives rise to an uncertainty in the reconstruction of X^μ that decreases quadratically with r , and increases with X^μ as:

$$\delta X^\mu = \frac{2X^\mu h_0}{r^2 \cos \theta} \ln^2 \left(\frac{X^\mu \cos \theta}{h_0 \rho_0} \right) c \delta t \quad (5.2)$$

where an exponential atmospheric density $\rho(z) = \rho_0 \exp(-z \cos \theta / h_0)$ has been assumed. Near the shower axis, muons arrive closer in time, hence the impact of time resolution on the estimation of the production depth is larger here than far from the core. The contribution of the geometric reconstruction to δX^μ also increases as we get closer to the core (with a linear dependence on r)

Thus, to keep the distortions of the reconstructed MPD small, only detectors far from the core are considered. This cut diminishes the efficiency in the reconstruction and also affects the resolution as it reduces the number of accepted muons. However, the reconstruction efficiency improves with energy, since the number of muons becomes larger as energy increases. Since the number of muons at the ground level is a function of the mass of the primary particle this cut has to be carefully chosen (see figure 5.4). The optimal value which minimizes the reconstruction bias in X_{\max}^μ and is independent of primary and energy is $r_{cut} = 1700$ m. As we do not have triggered stations beyond 4000 m we build the MPD distributions, for both true and reconstructed level, by counting muons in the distance range $1700 \text{ m} < r < 4000 \text{ m}$.

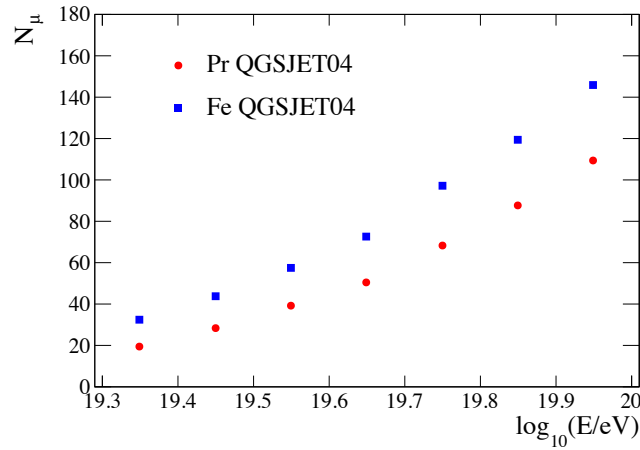


Figure 5.4: Average number of muons contributing to the MPD for proton and iron QGSJET-II-04 showers between 1700 and 4000 m.

To build the MPD distribution, every time bin of the FADC traces is converted into an MPD entry. Muon arrival times are smeared by light propagation inside the detector and the electronics response. To compensate for this *detector effect*, we subtract an offset t_{shift} to each time bin. Since in the SD the rise time of the muon signal is much shorter than the decay length because of the small path traversed by the muon in the detector, the t_{shift}

value depends on the $S_{threshold}$. Simulating vertical muons that pass through an Auger surface detector for $S_{threshold} = 15\%$, we obtain $t_{shift} = 73$ ns.

The MPD distribution for a single detector is obtained as the average of the MPD distributions yielded by each working PMT.

Finally, a simple set of selection criteria is applied in the set of events with a reconstructed MPD distribution to select those with a reliable measurement of the longitudinal profiles:

- **Energy cut:** Since the number of muons increases with the energy of the primary, events with energies below 20 EeV have very few entries per MPD resulting in a very poor determination of X_{max}^{μ} .
- X_{max}^{μ} **uncertainty:** Events whose relative uncertainty $\delta X_{max}^{\mu}/X_{max}^{\mu}$ is larger than a certain value ϵ_{max} (see table 5.1) are rejected.

In tables 5.2 and 5.3 are summarized the selection efficiencies after all the reconstruction cuts for both primaries and hadronic models.

$\log_{10}(E/eV)$	ϵ_{max} [%]
[19.3, 19.4]	15
[19.4, 19.6]	11
]19.6, 19.7]	10
]19.7, 19.8]	8
> 19.8	7

Table 5.1: Maximum relative uncertainties allowed in the estimation of X_{max}^{μ} . The value chosen for ϵ_{max} ensures no selection bias between the different primary species.

$\log_{10}(E/eV)$	Pr efficiency ϵ_{Pr} [%]	Fe efficiency ϵ_{Fe} [%]	ϵ_{Pr-Fe} [%]
19.35	88	93	5
19.45	84	93	9
19.55	90	97	7
19.65	94	99	5
19.75	95	99	4
19.85	96	99	3
19.95	98	99	1

Table 5.2: Efficiencies for proton and iron QGSJET04 showers after all the reconstruction cuts.

Figure 5.5 shows the absolute value of the mean bias after reconstruction. It is smaller than 10 g/cm², regardless of hadronic model, energy and atomic mass of the simulated primary particle. The resolution ranges from 100 (80) g/cm² for proton (iron) at the lower

$\log_{10}(E/eV)$	Pr efficiency ϵ_{Pr} [%]	Fe efficiency ϵ_{Fe} [%]	$\epsilon_{Pr}-\epsilon_{Fe}$ [%]
19.35	88	95	7
19.45	86	95	9
19.55	92	98	6
19.65	94	99	5
19.75	93	99	6
19.85	96	99	3
19.95	98	99	1

Table 5.3: Efficiencies for proton and iron EPOS LHC showers after all the reconstruction cuts.

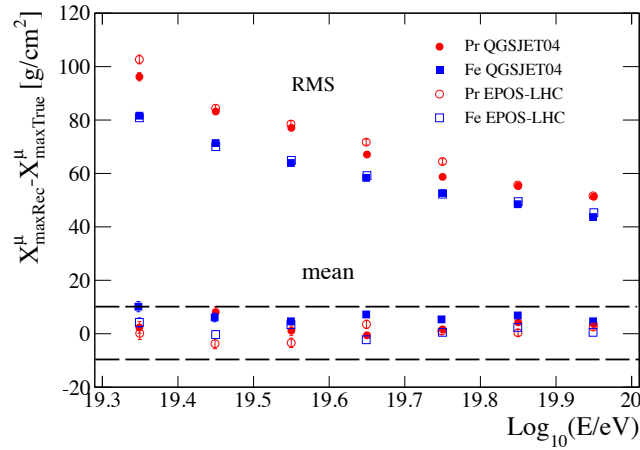


Figure 5.5: Standard Analysis bias and RMS. In this analysis MPDs are reconstructed individually for each event.

energies to about 50 g/cm^2 at the highest energies. The improvement of the resolution with energy is a direct consequence of the increase in the number of recorded muons.

Several sources contribute to the total resolution and to the bias of X_{\max}^{μ} . Based on simulations, we have decoupled and estimated the importance of each of them for this particular analysis. This is the main subject of study of the present chapter.

5.2.1 Breakdown of the different contributions to X_{\max}^{μ} bias

Figure 5.6 shows the bias for X_{\max}^{μ} as a function of energy after the full event reconstruction. We have decoupled each one of the sources contributing to this bias for a particular case, proton QGSJET-II-04 showers. In addition, we have studied the possible dependences of the sources with energy and the selected hadronic model.

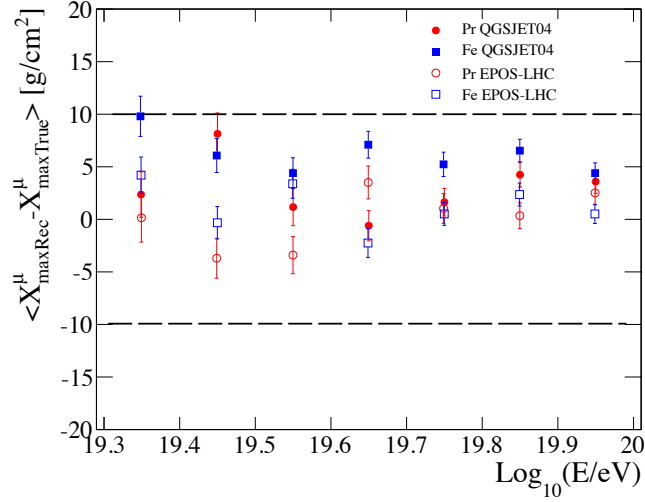


Figure 5.6: Standard analysis bias.

Accuracy of the method, Δ_{Method}

CORSIKA provides the exact production point of each muon, X^μ . We fit the distribution built with those values with a GH (equation 5.1). The maximum of this distribution will be our reference value hereafter. From now on, we will refer to it as $X_{\max}^{\mu \text{True}}$.

To be able to reconstruct the production point of a muon from its measured time and position in the WCD, we use the phenomenological model described in the previous chapter. The influence of those approximations on the value of the reconstructed X_{\max}^{μ} is considered to be the accuracy of the method.

We will start applying the model at CORSIKA level, where all the information of the muon (energy, position, time...) is available. To obtain the contribution of the model approximations to the bias of X_{\max}^{μ} we evaluate the difference between the maximum of the MPD calculated using the algorithm with the true muon information and our reference value (see figure 5.7). The total bias is less than 6 g/cm² independent of energy, as shown in figure 5.7.

Influence of the kinematic delay parameterization, $\Delta_{\text{kin.Del.Par}}$

The kinematic processes suffered by muons while traversing the atmosphere are a function of their energy (E_μ). To correct for their effects we need to know E_μ . The WCD do not allow to measure this quantity. For this reason, in the standard analysis a parameterization of the mean kinematic delay as a function of the distance to the shower core and the production distance is used [132, 131]. We can quantify the effect of this parameterization in the bias of the reconstructed X_{\max}^{μ} . In this case, instead of using the true kinematic delay, we will use its parameterization to obtain X_{\max}^{μ} . The difference between this value and our reference is illustrated in figure 5.8. The parameterization underestimates the reconstructed values by 30 g/cm².

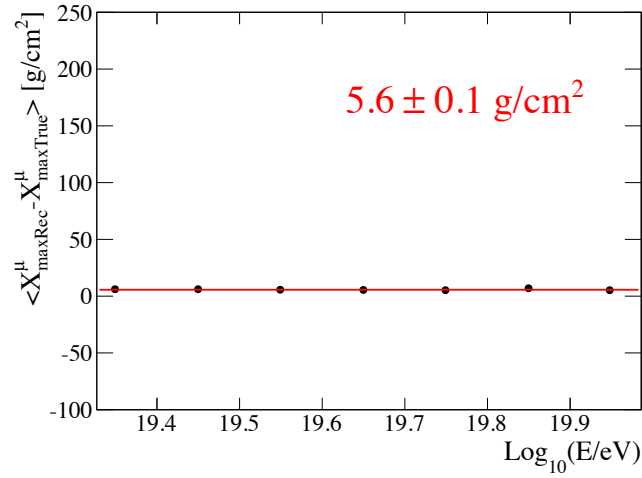


Figure 5.7: Contribution to the bias of X_{\max}^μ reconstructed from the approximations done in the muon arrival time model.

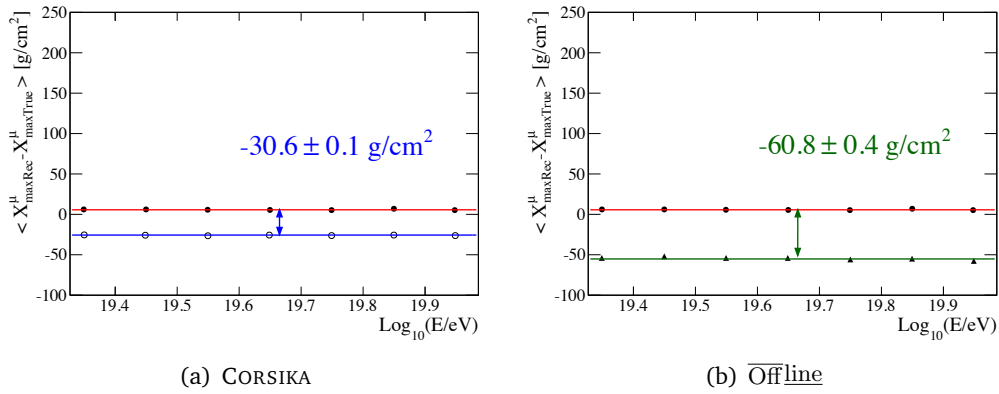


Figure 5.8: Effect of the kinematic delay parameterization in the mean value of $X_{\max\text{Rec}}^\mu$ at CORSIKA (left plot) and at Offline level (right plot).

In Offline simulations we have access to the true properties of the muon at the ground (times, geometry, ...). Thus, the impact of the kinematic delay parameterization can likewise be assessed at Offline level¹. At detector level, the number of muons contributing to each MPD is very small in comparison with CORSIKA. For proton showers we have 20 muons on average at 20 EeV. This comes from the spacing between the WCD and our distance cut. To reduce possible distortions due to the lack of statistics, particularly at low energies, we simulate the showers in a dense array² where we enlarge the number of

¹The true kinematic delay can not be evaluated at Offline level because we can not access the muon energy at production.

²WCD separated 750 m instead of 1500 m.

muons by a factor 3 (see figure 5.9). The observed underestimation of X_{\max}^{μ} due to the kinematic delay parameterization is two times bigger when we evaluate it using Offline simulations (see figure 5.8 right).

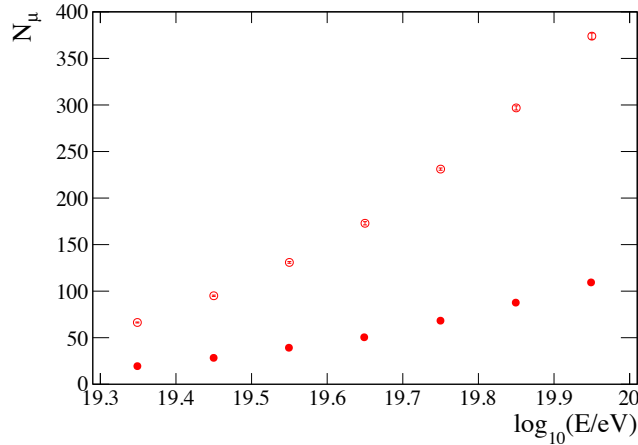


Figure 5.9: Number of muons as a function of the energy in the standard array (filled dots) and in a dense array (empty dots) for protons simulated using QGSJET04.

This discrepancy must come from subtle differences between the information obtained from CORSIKA and the true information obtained from Offline simulations. In principle, apart from the number of muons contributing to each MPD, the conditions in CORSIKA and in Offline (using the true information) are the same. However, to produce the CORSIKA files we use the *thinning* algorithm because the CPU time required for the simulation of full (non-thinned) EAS on a large scale is unmanageable. In this approximation, among the vast number of particles generated in the simulation, only a subsample of them are followed up to the end. Statistical weights are assigned to the surviving particles in representation of those ones that, in similar regions of phase space, have been removed along the procedure. In particular, we use simulated showers with a thinning level of 10^{-6} . The choice of this thinning level can affect the MPD distribution. We evaluate this effect comparing showers at CORSIKA level with a thinning of 10^{-5} , 10^{-6} and 10^{-7} . The differences obtained are smaller than 2 g/cm^2 , see figure 5.10.

The surface of each WCD is small compared to their spacing. In these conditions the simulation output contains very few particles falling exactly onto each detector. To deal with this problem it is implemented the un-thinning or resampling procedure. Its purpose is to use a wider sampling in the ground particle file to obtain a fair simulation of the particles entering in the detector. A reliable simulation of any detector response depends on the quality of the estimation of the local density of particles at the position of the detector. Within this context, these fluxes are obtained as averages over *more* or *less* extended areas around the location of interest. The module doing this in the Offline is the CachedShowerRegenerator. For the injection of a realistic flux of particles in a specific surface detector, the resampling method defines a sampling-area (see figure 5.11) which is defined by two parameters: a radial distance (δ) and an angular aperture (α) from the

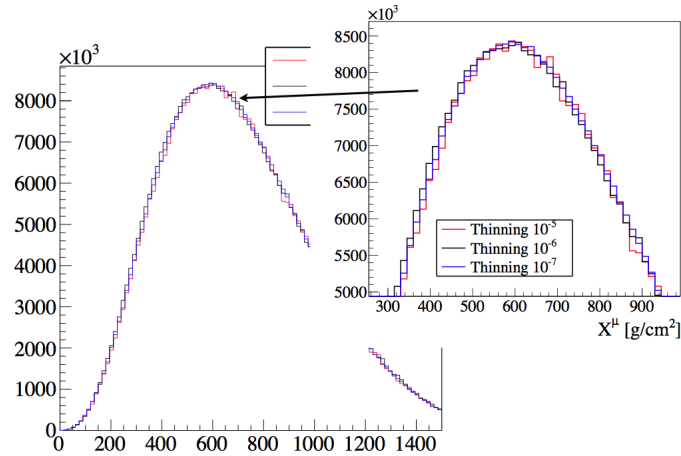


Figure 5.10: Differences in the average MPD distribution for 25 proton QGSJET-II-04 showers at $10^{19.55}$ at 60° generated with a thinning level of 10^{-5} (lightest line), 10^{-6} (dark line) and 10^{-7} (lighter line).

detector position (in the shower reference frame). The election of the values of δ and α is delicate: they must be small enough to guarantee the validity of the local averaging (times, energies, flux, etc.) and big enough to collect sufficient particles and avoid large artificial fluctuations (caused primarily by large remaining weights). Once this region is defined, all the particles within it are cloned according to their weights, keeping their times and momenta. Thereby, the mean number of particles hitting one single detector is given by the sum of all the particle weights inside the sampling region scaled by the ratio of the area of the detector and the area of the sampling region.

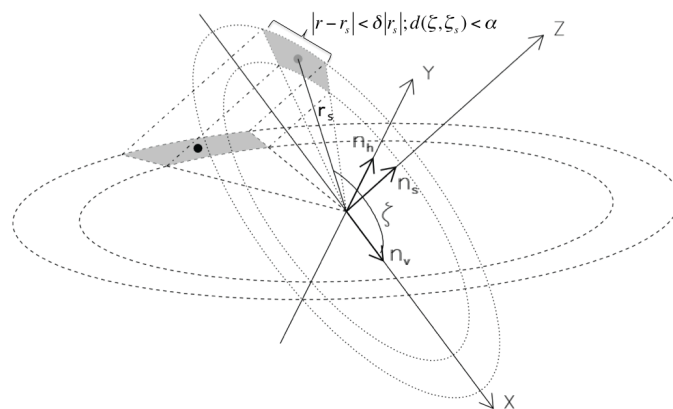


Figure 5.11: Definition of the sampling region (shadow) at a given position (marked with a point) [134].

We studied the temporal structure of the simulated air-showers before (at CORSIKA level) and after (at Offline level) the application of the resampling (or un-thinning) method summarized above. Even when we choose a sampling region, where it is adequate to assume a small variation of the curvature of the shower-front within it, a time correction must be applied to preserve the intrinsic structure of the shower:

$$t' = t_p + \vec{n}_s \cdot (\vec{r}_p - \vec{r}_s)/c \quad (5.3)$$

where t_p is the time of the particle, \vec{n}_s is the unitary vector describing the *plane front*, \vec{r}_p is the position of the particle and \vec{r}_s is the position of the detector. This correction preserves the delay $\tau = t_p - t_{\text{front}}$ of each resampled particle w.r.t. the front plane. A smearing procedure is also applied over the arrival times of the different clones to avoid possible artificial piling of particles in the simulation. For a precise description of the resampling procedure and their possible consequences we refer the reader to [134].

The resampling technique is the art of describing the physical properties of a point p_s by the averaged physical properties of an extended area A_s around it. In Offline, p_s is located in the exact geometric center of A_s . When an EAS develops in the atmosphere, the azimuthal dependence of the density of particles is small, so we do not expect any significant influence of the parameter α in the resampled results (a value of $\alpha < 20^\circ$ describes well the asymmetries [134]). However, the steep lateral distribution function of the particles $\rho(r)$ (there we will focus on the effect in muons), together with the increasing time delay as a function of r , introduce a non-trivial effect in the resampled arrival time distributions: the larger density close to the core favours a narrower distribution, while the summing over slices around p_s , with different time distributions, increases artificially the spread of the times. Figure 5.12 (left) shows the origin of this distortion.

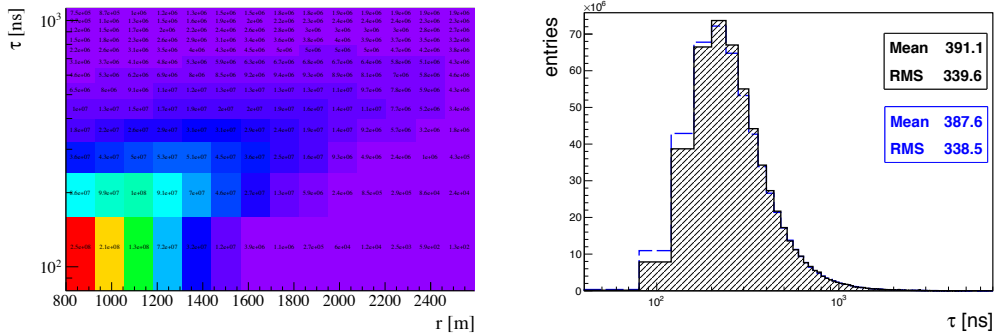


Figure 5.12: *Left*: Typical arrival time structure of muons in an EAS. *Right*: Distribution of the delay of the muons integrated in a sampling region centered at $r_s = 1450$ m using a value of $\delta = 1\%$ (dashed) and $\delta = 10\%$ (empty). The underestimation in the later is visible. Both figures have been obtained using QGSJET-II-04 with proton primaries at $10^{19.5}$ eV and $\theta = 60^\circ$.

In Offline simulations, the values of the parameters used to define the sampling region are $\delta = 10\%$ and $\alpha = 8.6^\circ$. As we have said, we will only concentrate on the effect of δ , so, in the following, we will define our sampling zones as full crowns (i.e. $\alpha = \pi$) removing

any dependence on α , while we increase the number of particles. Using CORSIKA, we “mimic” the $\overline{\text{Offline}}$ conditions at a distance r_s from the core, by integrating the number of particles (muons) in the crown defined by $|r_s - r| < 0.1|r_s|$. We define the *true* values when $\delta = 1\%$. Under that condition, we have verified that the weighted/integrated value of r in the crown correlates perfectly with r_s at each position (correlation factor ≈ 1), so we assume no bias.

Figure 5.12 (right) shows the “*true*” and the “ $\overline{\text{Offline}}$ ” distributions of muon delays in a sampling region centered at $r_s = 1450$ m, and for a sample of proton showers simulated using QGSJET-II-04 with $E = 10^{19.5}$ eV and zenith angle $\theta = 60^\circ$. As we see in the figure 5.12 (left), the number of muons decreases rapidly with r while the delay τ increases. The resampling procedure favors a bias towards smaller values of τ for large values of δ , as we see in the right-panel of that figure.

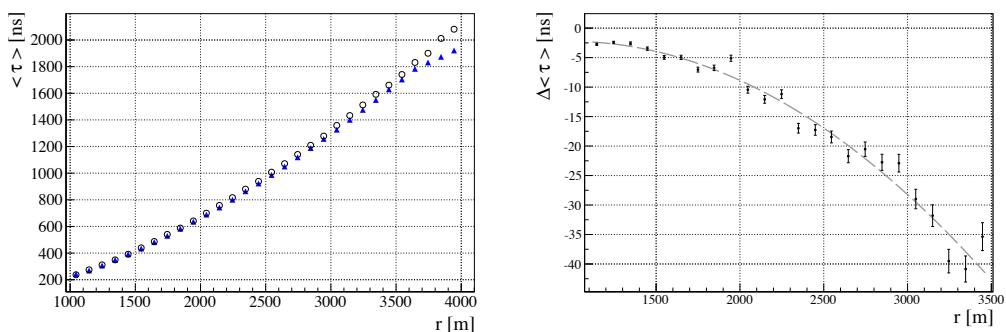


Figure 5.13: *Left*: Averaged time front curvatures obtained for two different definitions of the sampling region: $\delta_{\text{true}} = 1\%$ (circles) and $\delta_{\text{off}} = 10\%$ (triangles). *Right*: Averaged values of the expected bias in the muon delay in $\overline{\text{Offline}}$ by the resampling procedure. The fit to a quadratic function is also shown just to guide the eyes. This figure has been obtained using proton showers simulated with CORSIKA with QGSJET-II-04, $E = 10^{19.5}$ eV and $\theta = 60^\circ$.

The averaged delay of the muons (for the same sample of simulations) as a function of the distance to the shower axis is shown in figure 5.13. In the left panel, we show the results for two values of δ : 1% (circles) and 10% (triangles). To quantify the magnitude of this bias and its dependence with the position, in the panel on the right we show the residuals of $\langle \tau \rangle$ ($\overline{\text{Offline}} - \text{true}$) as a function of r . It is important to notice that close to the core (< 2000 m) the underestimation of $\langle \tau \rangle$ is small (< 10 ns) for the $\overline{\text{Offline}}$ value of δ . The value of the bias increases with r (approximately with r^2) becoming very large at large distances from the core, but its relative importance decreases given the attenuation on the number of muons with r .

As expected, the dependence of the bias with the particular choice of δ is large, as shown in figure 5.14 (left), where the results obtained for half (5%) and twice (20%) the value used in $\overline{\text{Offline}}$ are shown. In [134] some corrections are suggested, based on general assumptions that preserve the shower-to-shower fluctuations, to reduce the different biases. In figure 5.14 (right) we show the result of applying two of those corrections: (i)

to correct for the number of particles, each weight of the particles at ground is re-scaled by the factor $(r_p/r_s)^{\eta-1}$, with $\eta = 3.5$ (empty circles) (ii) a correction directly to the delay can be also applied by $\tau' = (\tau - \tau_f(r_p))(r_s/r_p) + \tau_f(r_s)$, where $\tau_f(r) = r^2/(2Rc)$ and R is the radius of curvature (empty triangles). In the evaluation of this correction we have used $R = 11.5$ km [135]. Figure 5.14 (right) shows that both corrections work essentially at the same level in absolute value. Any of them remedies most of the bias introduced by the resampling algorithm with $\delta = 10\%$. If we compare with the figure on the left, the effect of applying these corrections is similar to reducing the value of δ to a value of 5% (i.e. bias smaller than 10 ns for distances closer than 3000 m). In the Offline no correction to the resampling is applied.

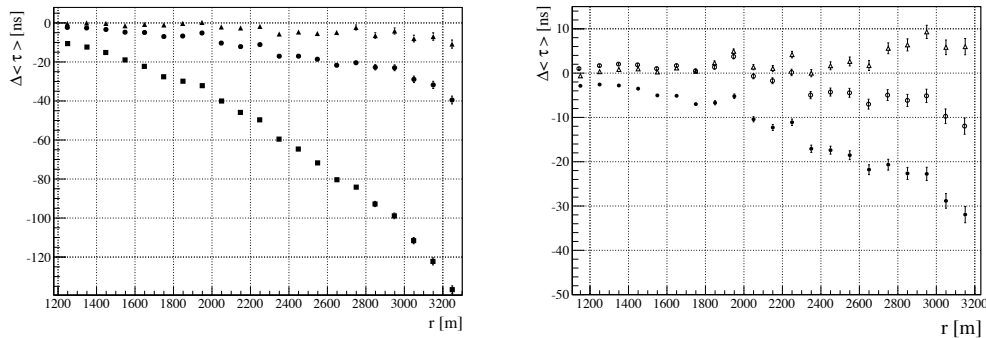


Figure 5.14: *Left:* $\Delta \langle \tau \rangle$ vs r for three different values of δ : 5% (triangles), 10% (circles) and 20% (squares) respectively. *Right:* Bias in the arrival time of the muons with (empty symbols) and without (full circles) applying the corrections suggested in [134]. The improved estimation of the arrival times when we use a correction is remarkable.

We have studied the dependency of $\langle \tau \rangle(r)$ (figure 5.13) as a function of the primary (iron vs proton), the energy of the primary (20 to 100 EeV) and the hadronic model (QGSJET-II-04 vs EPOS LHC). We have not found any significant dependency: no bias and a width RMS < 4 ns for all the considered cases (see figure 5.15).

For completeness, we have also studied the effect of choosing particular values of the azimuth ζ rather than integrating in a full crown. With the aim of covering both *early* and *late* regions, we have compared our results, integrating over all ζ values, with four different sampled regions centered at $\zeta = 0, \pi/4, \pi/2$ and $3\pi/4$ and with a value of $\alpha = 20^\circ$ for all of them. The results for this comparison are shown in figure 5.16. Only at large distances from the core ($r > 3000$ m) an extra bias ($\Delta \langle \tau \rangle_{\zeta_i - 2\pi} > 10$ ns) is induced for different azimuthal locations.

The simulations used in this work were done with Offline- icrc13 where the value of δ was set to 10%. We fit the bias in the delay as a function of the distance (figure 5.13 right) between 1000 m and 4000 m and we correct the delays by this quantity. The difference obtained by using the true information and the kinematic delay parameterization in Offline with respect to CORSIKA can be explained by this effect of the un-thinning procedure. If we plot figure 5.8 right after applying this correction the result is shown in figure 5.17. Now, the observed discrepancy when using the true information in both CORSIKA

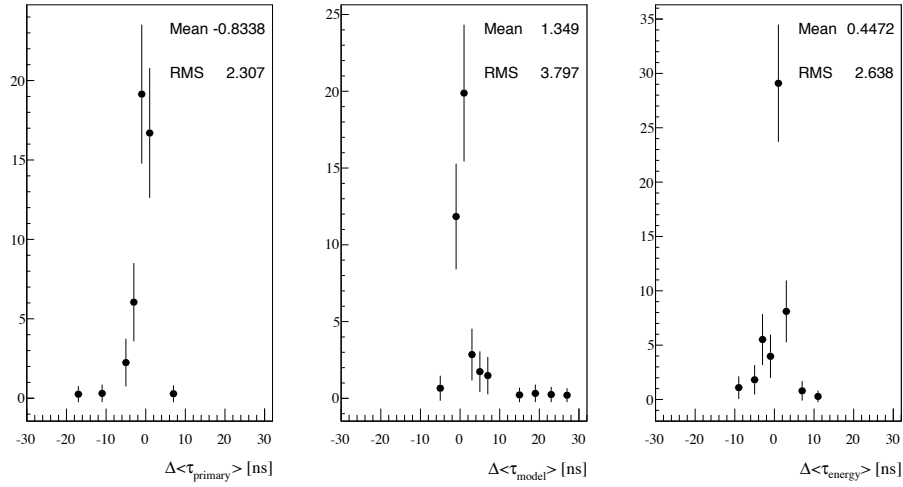


Figure 5.15: Comparison of the values of the muon delays for proton and iron primaries (*Left*), QGSJET-II-04 and EPOS LHC hadronic models (*Center*) and different energies (*Right*). A negligible bias is found for all the considered cases.

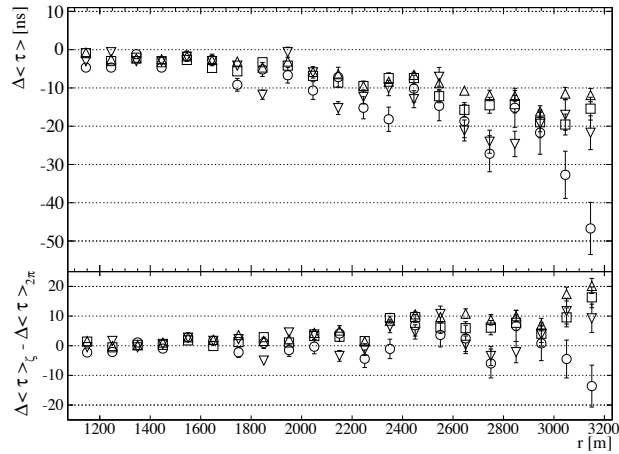


Figure 5.16: Dependency of the bias in τ with the azimuthal position. In the bottom panel we compare the results when we integrate over a full crown ($\alpha = \pi$) with the results at four sampling regions: $\zeta = 0$ (empty circles), $\pi/4$ (empty inverse-triangles), $\pi/2$ (empty squares) and $3\pi/4$ (empty triangles) for a value of $\alpha = 20^\circ$ and $\delta = 10\%$.

and $\overline{\text{Offline}}$ is at the level of 5 g/cm^2 , making comparable both approaches.

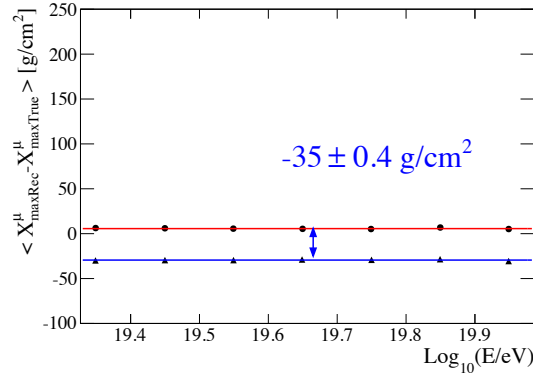


Figure 5.17: Effect of the kinematic delay parameterization in the bias of $X_{\max\text{Rec}}^\mu$ at Offline level after resampling correction.

Detector effects

The next step is to decouple the contributions to the bias in X_{\max}^μ coming from detector effects: the accuracy in the geometry reconstruction, Δ_{geometry} and the response of the WCD to muons, Δ_{signal} .

The uncertainty in the determination of the core position and the angular reconstruction affects the value of $X_{\max\text{Rec}}^\mu$. Both quantities are correlated in the SD reconstruction so it is very difficult to disentangle the influence of any of them separately in the bias of X_{\max}^μ (see figure 5.18). Because of that we will study the influence of both quantities together in the estimation of X_{\max}^μ .

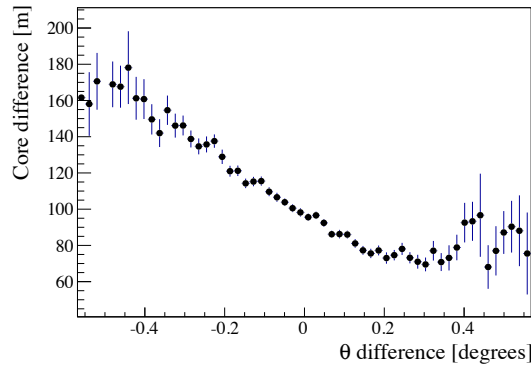


Figure 5.18: Correlation between the uncertainty in core and zenith determination.

In this case, to reconstruct our observable we use the true time of muons and the reconstructed geometry. The result are the empty triangles of figure 5.19. The difference between the empty and the filled ones (the result obtained in the previous step) is the contribution coming from the addition of the geometry reconstruction. It amounts to

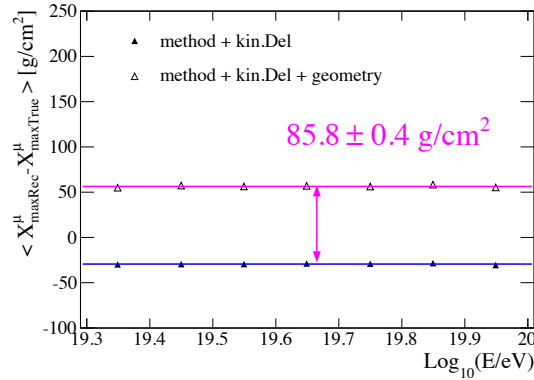
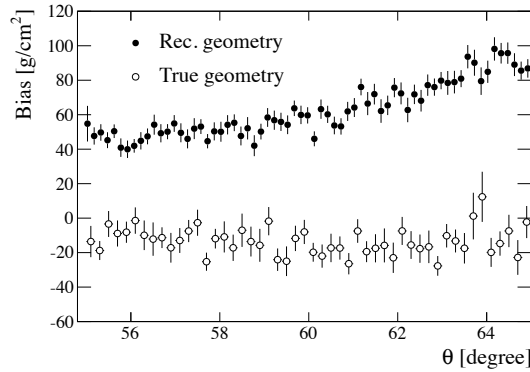


Figure 5.19: Contribution to the bias from the geometry reconstruction.

86 g/cm².

Figure 5.20: Bias vs θ comparing true and reconstructed geometry.

Furthermore, plotting the bias in X_{\max}^{μ} as a function of the zenith angle for the cases where the true and the reconstructed geometry are used, we find a trend starting around 61° when the reconstructed geometry is considered (filled dots of figure 5.20)³.

The fact that the response of the WCD to muons shows a spread on time affects our estimation of X_{\max}^{μ} . If we compare the reconstructed X_{\max}^{μ} after all the reconstruction chain (i.e., adding the signal to all the other contributions) with its true value, the result are the squares that we show in figure 5.21. The difference between the squares and the empty triangles is the bias added by this last contribution. It amounts to 126 g/cm² and it is by far the largest contribution to the bias.

In table 5.4 and figure 6.5 we summarize and quantify all the identified sources contributing to the bias of X_{\max}^{μ} .

³Investigations using the inclined reconstruction as implemented in [Offline icrc13](#) have shown that the same trend is observed.

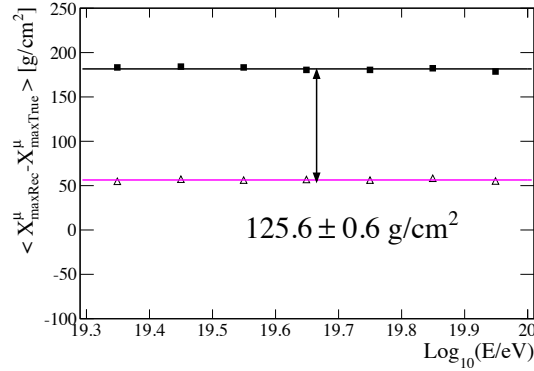
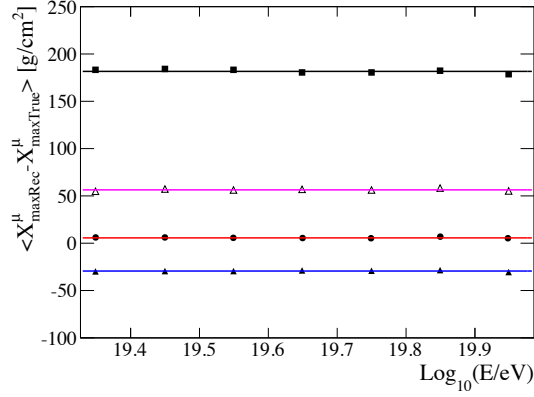
Figure 5.21: Signal contribution to X_{\max}^{μ} bias.

Figure 5.22: Summary of all the bias contributions. Filled dots correspond to the method accuracy. The filled triangles include the kinematic delay parameterization plus the method. The empty triangles contain the method, the kinematic delay parameterization and the geometry reconstruction. Finally, the squares represent the bias after the whole chain is applied (i.e. including the method, the kinematic delay parameterization, the geometry reconstruction and the effect of the signal spread).

Source	Bias [g/cm ²]	Source	Bias [g/cm ²]
Δ_{Method}	6	Δ_{Method}	6
+ $\Delta_{\text{kin.Del.Par}}$	-29	+ $\Delta_{\text{kin.Del.Par}}$	-29
+ Δ_{geometry}	56	+ Δ_{Signal}	114
+ Δ_{Signal}	182	+ Δ_{geometry}	182

Table 5.4: Cumulative contributions to X_{\max}^{μ} bias.

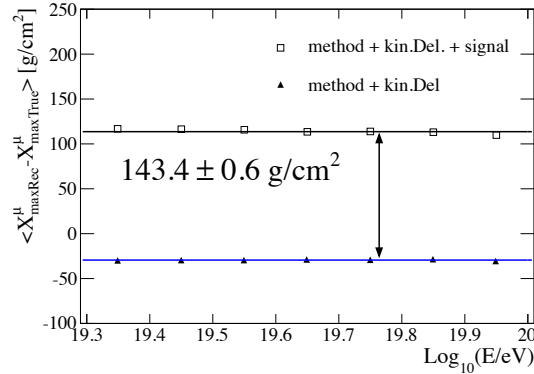


Figure 5.23: Signal contribution to X_{\max}^{μ} bias where true geometry is considered.

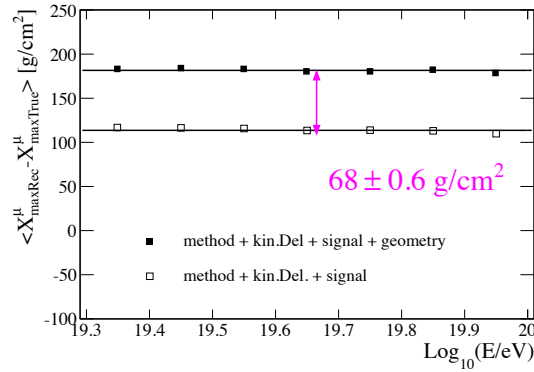


Figure 5.24: Contribution to the bias due to the geometry reconstruction.

To quantify in first place the contribution from the detector response we have inverted the order, i.e. we first evaluate this effect using the true geometry. In this case the WCD response makes a difference in the X_{\max}^{μ} bias of 143 g/cm^2 (see figure 5.23), while the remaining geometry bias is 68 g/cm^2 , as illustrated in figure 5.24. The difference in the values coming just by switching the order is caused by the correlation between the geometry reconstruction and the detector response.

The response of the muon in the WCD is an asymmetric function. When we use the true time of the muons we have delta functions. Thus, it is not straightforward to compare these effects following the proposed approaches. In fact, we can study the dependency of our bias with the asymmetry of the signal. For example, if we cut at a certain signal threshold we will remove part of the tails. We have evaluated the geometry reconstruction bias for different signal thresholds and we have observed a dependence on that. When we cut at 50% of the peak, the difference between the true and the reconstructed geometry increases by 10 g/cm^2 , as we expect from the result obtained when using the true time of the muon. In table 5.5 we quote the evolution of this bias as a function of the signal

threshold used for a bin of proton QGSJET-II-04 at $\text{Log}(E/\text{eV}) = 19.55$.

Signal Threshold [%]	Bias [g/cm^2]
0	67.6
15	72.3
30	74.5
50	75.2

Table 5.5: Bias due to geometry reconstruction depending on the signal threshold applied.

All in all, if we sum up all the bias contributions we obtain a value of $182 \text{ g}/\text{cm}^2$.

5.2.2 Final bias

We show in figure 5.25 the bias after all the reconstruction chain and once the resampling correction has been applied to the standard analysis (cutting the signal at 15% of the peak and applying the time shift). The results for proton and iron (for both EPOS LHC and QGSJET-II-04) show that, after corrections and cuts are applied, we are left with a residual bias of about $25 \text{ g}/\text{cm}^2$.

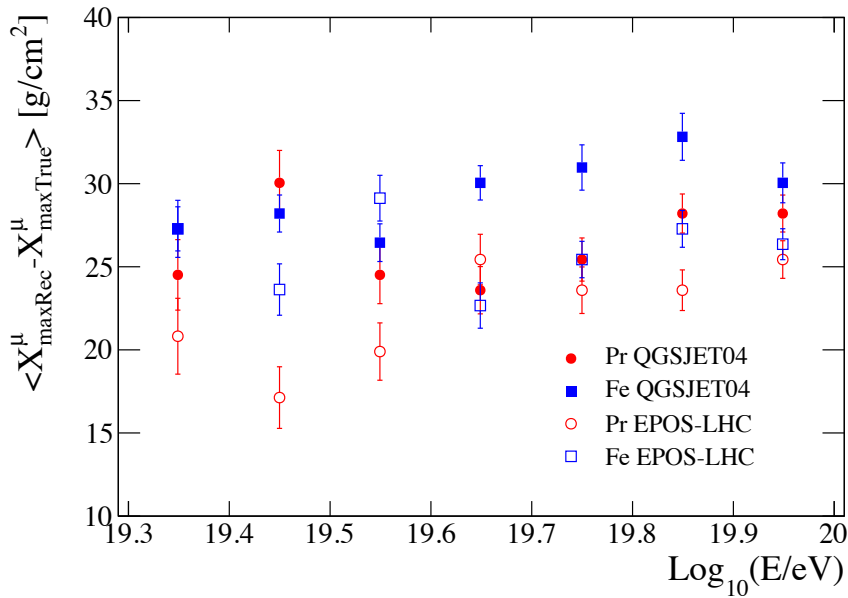


Figure 5.25: Bias in the standard analysis conditions after all the reconstruction chain for proton and iron nuclei for both hadronic models.

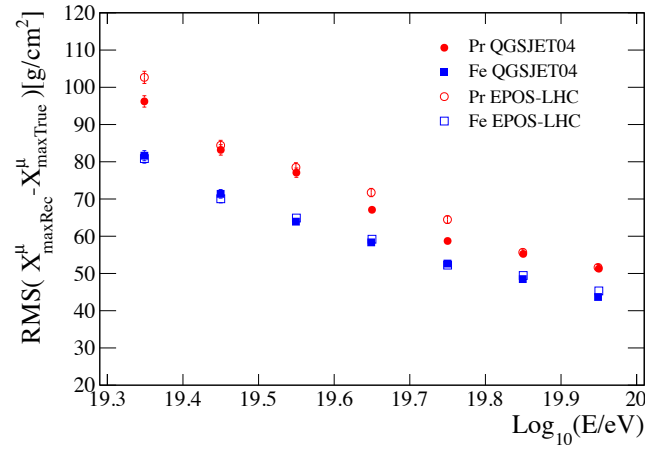
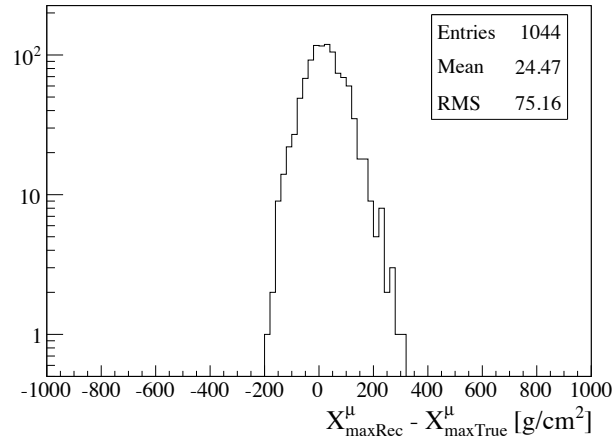


Figure 5.26: Standard analysis RMS.

5.3 Contributions to X_{\max}^{μ} resolution

Several sources contribute to the total resolution in the measurement of X_{\max}^{μ} . In figure 5.26 is illustrated the RMS for all the species. Based on simulations we have decoupled each of the sources as we did for the case of the bias in X_{\max}^{μ} . For this study we have taken a bin of proton QGSJET-II-04 at $10^{19.55}$ eV. The histogram for this energy bin of the difference between the reconstructed and the generated value is shown in figure 5.27.

Figure 5.27: Difference between X_{\max}^{μ} reconstructed and the true one after all the reconstruction chain for proton QGSJET-II-04 at $10^{19.55}$ eV.

The total RMS of X_{\max}^{μ} has four contributions:

$$RMS_{\text{Total}} = \sqrt{RMS_{\text{Method+Kin.Del.Par}}^2 + RMS_{\text{Time}}^2 + RMS_{\text{geometry}}^2 + RMS_{\text{Statistics}}^2} = 75.2 \text{ g/cm}^2 \quad (5.4)$$

The influence of the method accuracy plus the kinematic delay parameterization in the RMS, $RMS_{\text{Method+Kin.Del.Par}}$, is negligible. We have built the histogram of the difference between X_{max}^μ reconstructed using the model and the parameterized kinematic delay, and the reference one. The RMS of this histogram is 6.5 g/cm^2 .

The largest contribution comes from the number of selected muons, $RMS_{\text{Statistics}}$. To estimate this quantity we have compared the contribution of the method to the RMS at CORSIKA level (built with a large number of muons) with the one obtained in Offline simulations selecting the true geometry and the true times, RMS_{Offline} :

$$RMS_{\text{Statistics}} = \sqrt{RMS_{\text{Offline}}^2 - RMS_{\text{Method+Kin.Del.Par}}^2} = \sqrt{56.04^2 - 6.5^2} = 55.7 \text{ g/cm}^2 \quad (5.5)$$

The time uncertainty of the detector also affects the accuracy in the reconstruction of X_{max}^μ . To evaluate this number we have compared the RMS of the distribution obtained using the true time of the muon plus the reconstructed geometry (at Offline level) with the one obtained after all the reconstruction chain:

$$RMS_{\text{Time}} = \sqrt{RMS_{\text{Total}}^2 - RMS_{\text{True Time}}^2} = \sqrt{75.16^2 - 65.67^2} = 36.6 \text{ g/cm}^2 \quad (5.6)$$

Finally, the accuracy of the reconstruction of the shower geometry influences the determination of X_{max}^μ too. We can estimate this effect comparing the RMS of the distribution obtained using the true time and the true geometry (at Offline level) with the one obtained using only the true time:

$$RMS_{\text{geometry}} = \sqrt{RMS_{\text{True Time}}^2 - RMS_{\text{True Time + True Geometry}}^2} = \sqrt{65.67^2 - 56.04^2} = 34.2 \text{ g/cm}^2 \quad (5.7)$$

The contributions of the geometry and the time uncertainty have been minimized due to the optimization of the distance cut. Both effects are larger when we go closer to the shower core.

Finally, Table 5.6 summarizes the contributions to the RMS. The main source is the limited number of muons contributing to each MPD distribution.

5.4 Application to SD Data

In [1] the bias due to the resampling procedure was assumed to be negligible. The residuals between the bias obtained with and without the resampling correction are shown in figure 5.28 for both primaries and hadronic models. A linear fit to the biases for proton (iron) QGSJET-II-04 as a function of the energy leads to an estimation of $\Delta(X_{\text{Rec}}^\mu - X_{\text{True}}^\mu) = 23.9 \pm 0.8$ ($\Delta(X_{\text{Rec}}^\mu - X_{\text{True}}^\mu) = 23.7 \pm 0.7$) g/cm^2 . Those values do not depend

Source	RMS [g/cm ²]
Method + Kin.delay parameterization	6.5
Statistics	55.7
Detector response	36.6
Geometry reconstruction	34.2
Total	75.2

Table 5.6: Summary of all the identified RMS contributions for proton QGSJET-II-04 at $\text{Log}(E/\text{eV}) = 19.55$.

on the hadronic interaction model, since for the case of EPOS LHC we obtain $\Delta(X_{\text{Rec}}^{\mu} - X_{\text{True}}^{\mu}) = 23.9 \pm 0.8 \text{ g/cm}^2$ for proton and $\Delta(X_{\text{Rec}}^{\mu} - X_{\text{True}}^{\mu}) = 25.1 \pm 0.7 \text{ g/cm}^2$ for iron. This number is above the systematic uncertainties, which are 17 g/cm^2 . The evolution of the measured X_{max}^{μ} as a function of the energy after correcting this bias is shown in the bottom panel of figure 5.29. Figure 5.30 illustrates the conversion to $\langle \ln A \rangle$. The tension between data and predictions is evident for the case of EPOS LHC. The mean $\ln A$ values extracted from the measurements of $\langle X_{\text{max}} \rangle$ and $\langle X_{\text{max}}^{\mu} \rangle$ are incompatible with each other. EPOS LHC in combination with Fluka 2011.2b.4 as a low-energy interaction model does not offer a consistent description of the electromagnetic and muonic components of the EAS. With QGSJET-II-04/Fluka, we obtain compatible values for $\ln A$ within 1.5σ , but it should be noted that, in contrast to EPOS LHC, this model has problems to describe in a consistent way the first two moments of the $\ln A$ distribution obtained from the X_{max} measurements done with the FD [136]. From the comparisons shown in [136] and here we see that none of the interaction models recently tuned to LHC data provide a consistent description of the Auger data on EM and MPD profiles.

Therefore, X_{max}^{μ} is a valuable observable to test the validity of the assumptions used to model hadronic interactions at ultra-high energies.

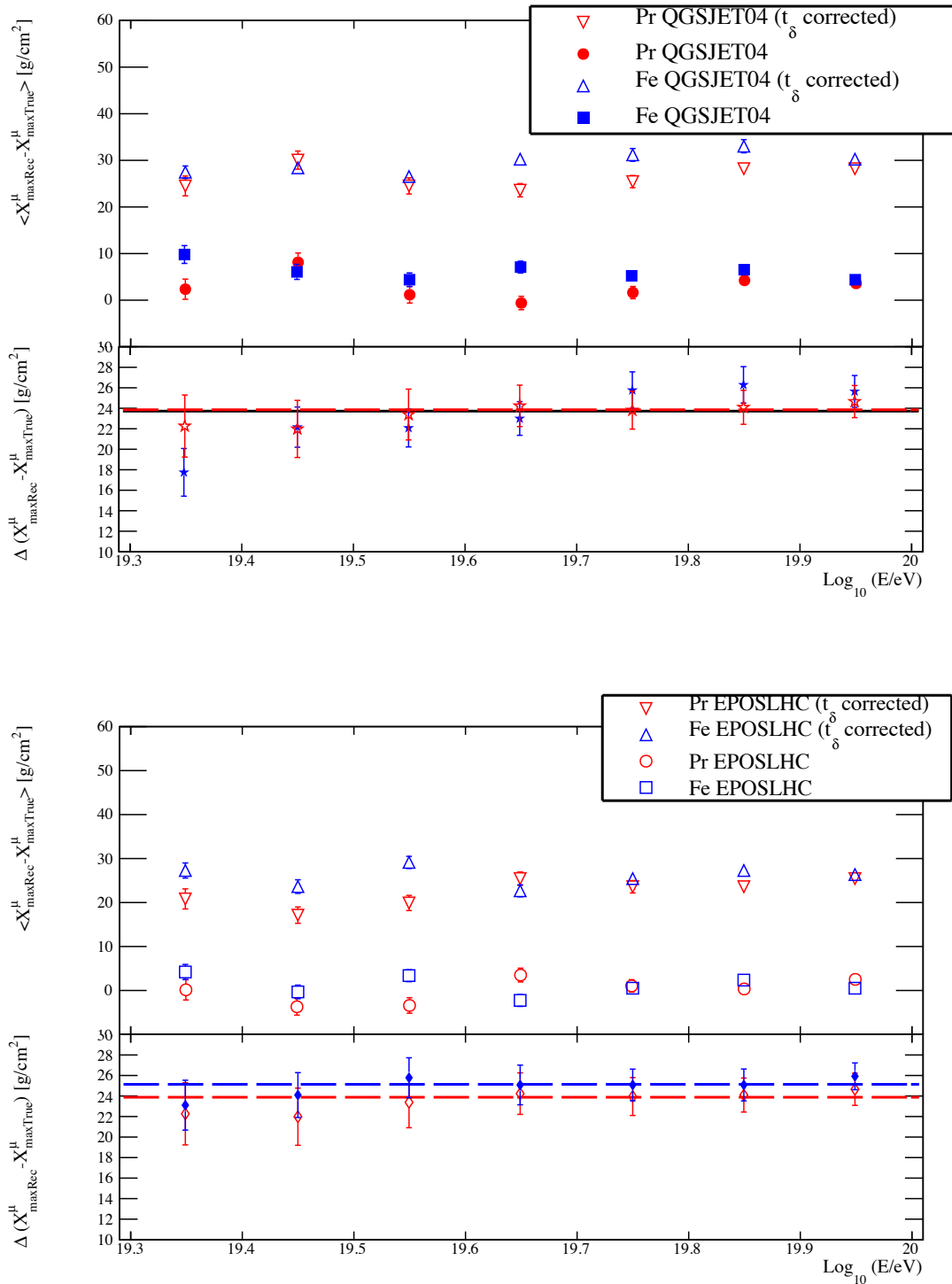


Figure 5.28: Bias in the standard analysis conditions after all the reconstruction chain for proton and iron nuclei QGSJET-II-04 (top) and EPOS LHC (bottom) after applying the resampling correction and before. The bottom pad shows the residual of both biases for both primaries and models.

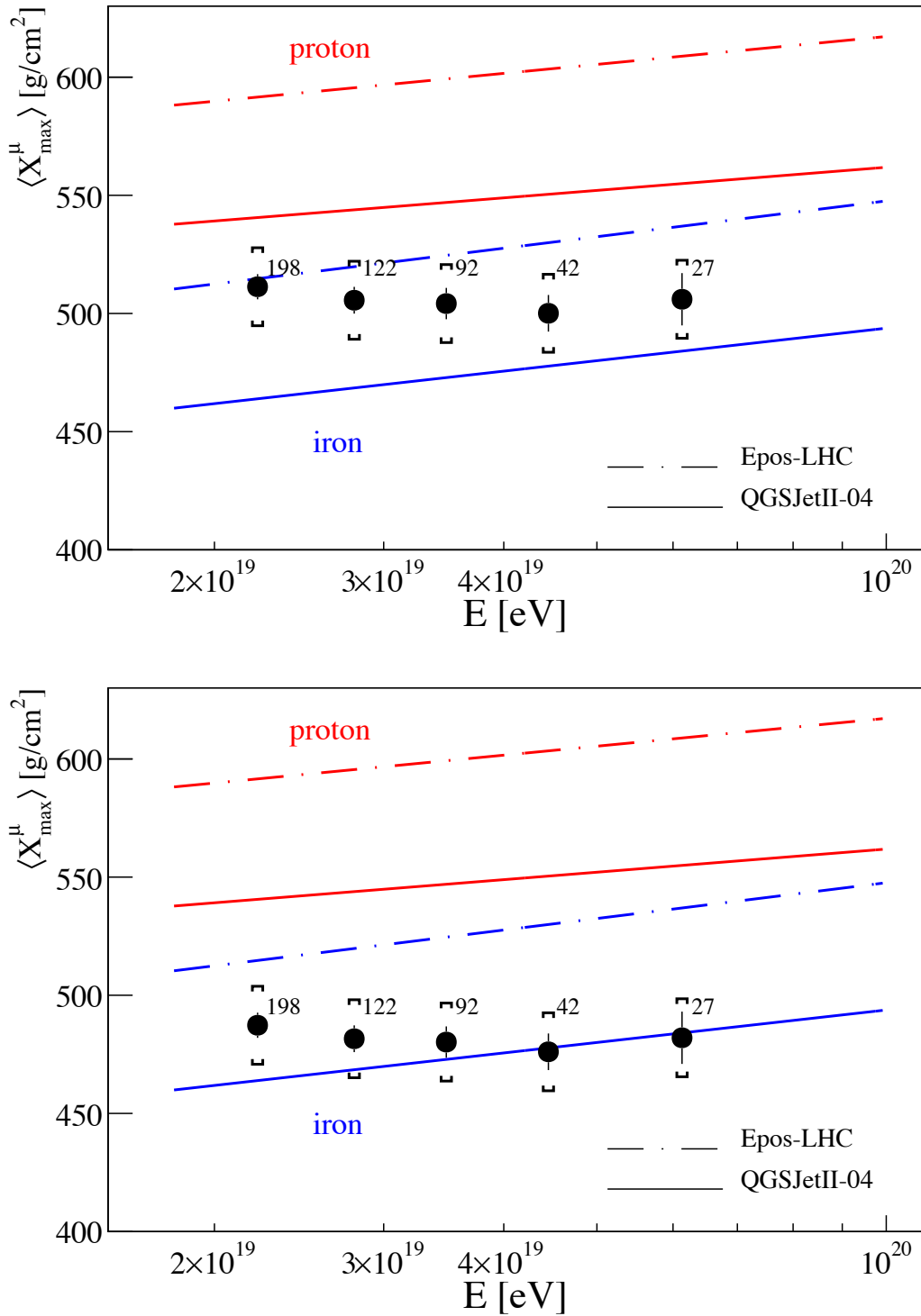


Figure 5.29: $\langle X_{\max}^{\mu} \rangle$ as a function of energy before (upper plot [1]) and after (bottom plot) the resampling correction. The prediction of different hadronic models for proton and iron are shown. Numbers indicate the statistics in each energy bin and brackets represent the systematic uncertainty.

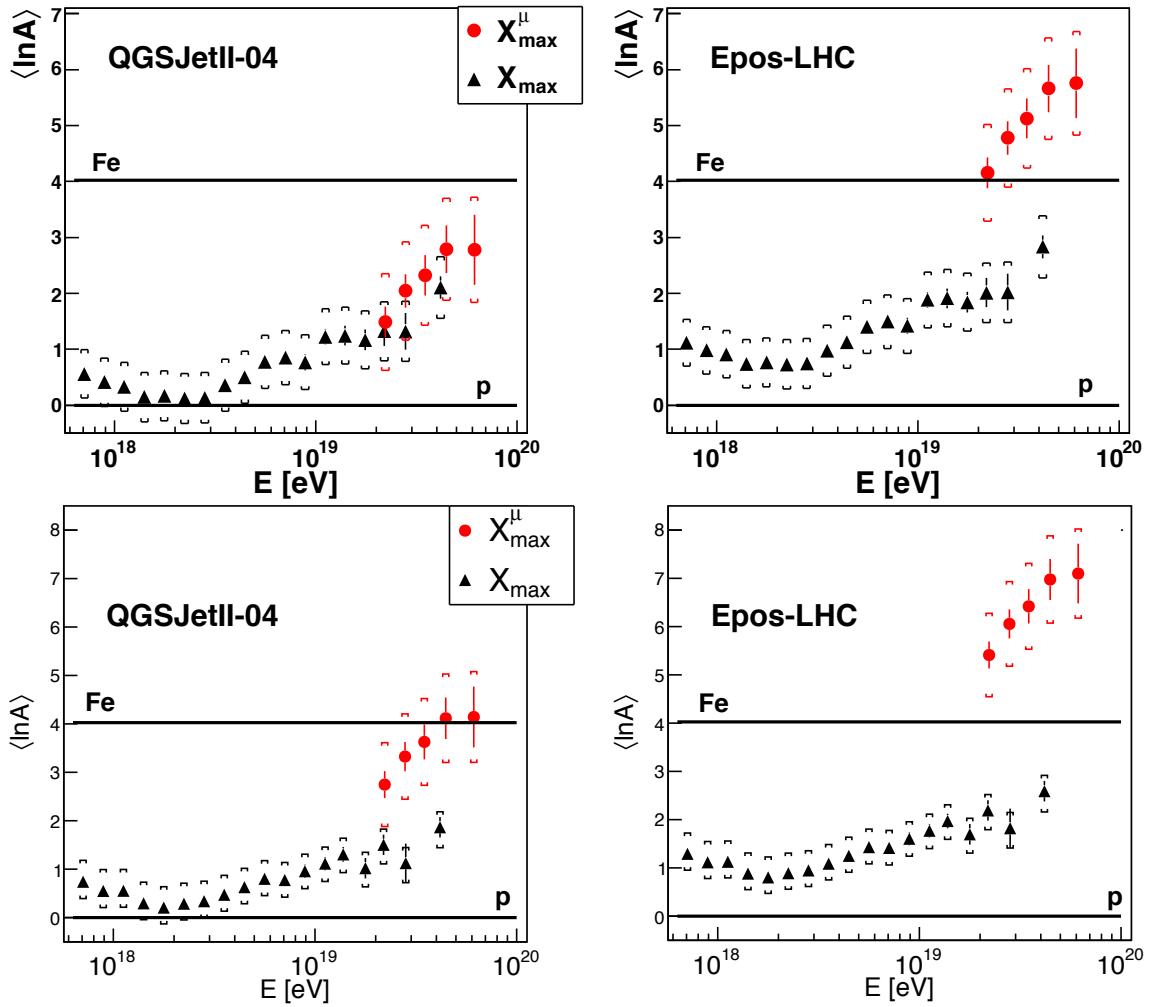


Figure 5.30: Conversion of $\langle X_{\max}^{\mu} \rangle$ (circles) and $\langle X_{\max} \rangle$ [137] to $\langle \ln A \rangle$, as a function of energy before (top) and after (bottom [1]) bias correction. On the left (right) plot we use QGSJET-II-04 (EPOS LHC) as the reference hadronic model. Brackets correspond to the systematic uncertainties.

Average MPD analysis

6

In chapter 5 we have decoupled the main sources contributing to the resolution found in the event-by-event measurement of X_{\max}^{μ} . The main contribution is due to the lack of muons entering in each MPD distribution. The energy threshold in the MPD analysis is also a consequence of the scarce number of available muons (for protons we have on average around 20 muons at 20 EeV) [1]. Therefore, to be able to extend the energy range of applicability of MPD studies we need to enlarge the number of muons. Previous works address this issue by going closer to the shower core [127, 128, 129]. We use here a different approach to bypass this problem. Instead of using an event-by-event observable, we study the average MPD distribution of all the events that belong to a particular energy bin. We investigate the feasibility of reducing the distance cut by using all the stations above 1000 m. We are going to explore how low in energy we can go with this new approach. The analysis is focused on events with the same zenith range, around 60° , as those used for the published MPD analysis. No attempt is made in this analysis to enlarge the angular window. Since our interest now focuses on average MPD distributions, this sort of study introduces an additional complication: not only the mean value of X_{\max}^{μ} matters, now we must reproduce as accurately as possible the shape of the MPD distribution, as well.

6.1 Reconstruction of the average MPD distribution

In chapter 5 we have explained the main features of the MPD distributions and their reconstruction for an event-by-event analysis. The maximum of the distribution is an SD observable that can be used as a discriminant for mass composition. In the analysis described in this chapter, we add up all the distributions belonging to a given energy bin and we build an average profile (see figure 6.1). From now on, *the average* X_{\max}^{μ} , $\langle X_{\max}^{\mu} \rangle$, will be our physics observable. As in the standard analysis, the best set of parameters that describes the average longitudinal muon profile (at CORSIKA and at Offline level) are obtained through a log-likelihood maximization of the GH function. We fit all the parameters in a chosen interval of depths ranging from 0 to 1200 g/cm^2 . With this choice we guarantee that we contain the entire range of possible $\langle X_{\max}^{\mu} \rangle$ values.

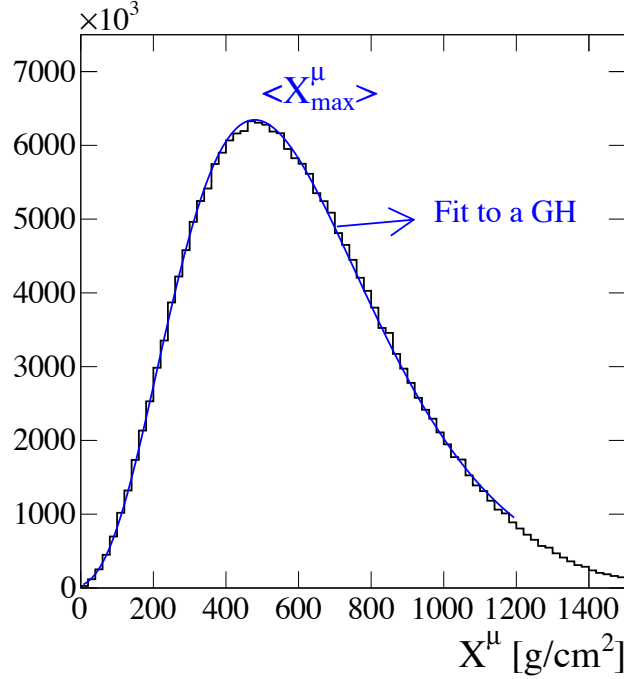


Figure 6.1: Sum of all the distributions for iron QGSJET-II-04 at $10^{19.55}$ eV obtained from CORSIKA. The maximum obtained from a GH fit (solid line) will be our SD observable, $\langle X_{\max}^\mu \rangle$.

6.2 MC studies

6.2.1 Bias contributions to the average X_{\max}^μ

We start this study by decoupling each one of the contributions that enter on the average X_{\max}^μ bias. We build MPD profiles for a distance range spanning from 1000 to 4000 m and carefully assess the contribution each step of the reconstruction adds to the total bias. This time care has to be taken in identifying if the chosen reconstruction procedure induces any unwanted modification in the shape of the distribution. We take proton QGSJET-II-04 showers at $10^{19.55}$ eV as reference.

Influence of the accuracy of the method and the new kinematic delay parameterization, $\Delta_{\text{Method}} + \Delta_{\text{kin.Del.Par}}$

In chapters 4 and 5 we quantified the effect of the phenomenological model used to relate arrival times to muon production distances. In the first place we need a reference to compare to. The reference average profiles per energy bin are obtained from the information provided by CORSIKA, black line of figure 6.2. We fit the distribution with a GH and its maximum will be our reference value from now on, $\langle X_{\max\text{True}}^\mu \rangle$.

To obtain the contribution of the model approximations to the average bias of X_{\max}^μ ,

we evaluate the difference between the maximum of the average MPD calculated using the time model with the true muon information (filled blue histogram of figure 6.2) and our reference value. The total bias is around 11 g/cm^2 .

To correct the kinematic effects suffered by muons while traversing the atmosphere we need to parameterize them. For the current analysis we use the new kinematic delay parameterization described in chapter 4. We demonstrated that this new approach does not distort the profile and reduces significantly the bias. The result is the red filled histogram of figure 6.2. The difference between the maximum of the CORSIKA distribution and the red one will be the cumulative bias coming from the method and the new kinematic delay parameterization, which in this distance range amounts to 20 g/cm^2 .

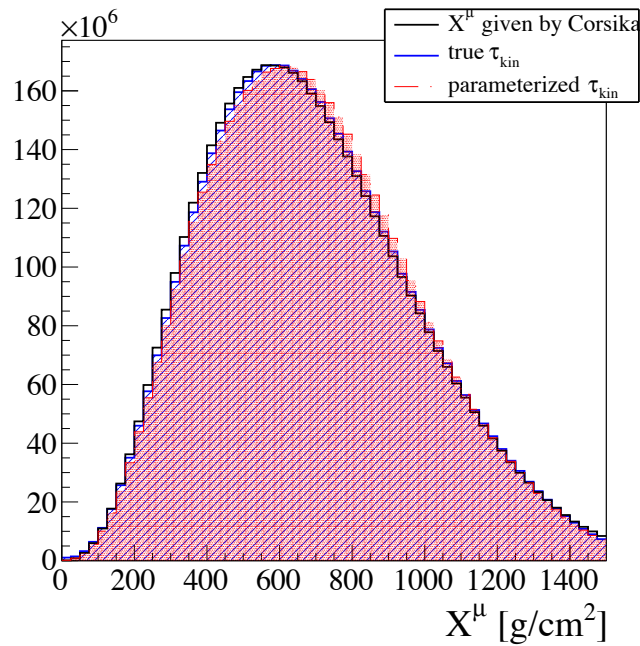


Figure 6.2: Average MPD distribution for proton QGSJET-II-04 at $10^{19.55}$ eV. Black line is the distribution directly obtained from CORSIKA, the blue filled histogram is the result correcting the delay with the true kinematic delay, and the red filled one is the result after applying the new parameterization described in chapter 4.

Detector effects

The next step is to decouple the contributions to the bias in $\langle X_{\max}^\mu \rangle$ coming from detector effects: the accuracy in the geometry reconstruction, Δ_{geometry} and the response of the WCD to muons, Δ_{signal} . The simulations used in this analysis were performed using Offline-icrc13, thus we take into account the resampling correction described in the previous chapter [5].

Following the same approach adopted for the event-by-event analysis, we obtain the average distribution using the true time of the muons and the reconstructed geometry. The difference between this maximum and our reference will represent the bias coming from

the geometry reconstruction, Δ_{geometry} . This bias has a dependency with the distance to the shower core as it is illustrated in figure 6.3.

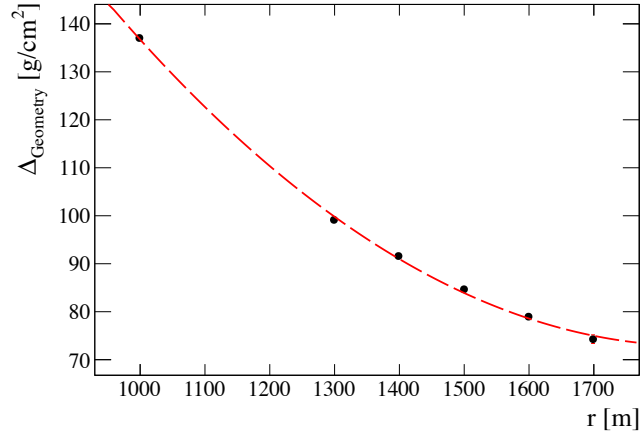


Figure 6.3: Bias due to the geometry reconstruction as a function of the distance to the shower core.

As already discussed the shape of the distribution is another source of concern. We know that the geometry reconstruction distorts the profile and this effect is larger the closer we get to the shower core. In figure 6.4 we compare the average profile at two distances from the shower core using the true geometry and the reconstructed one. Starting at 1500 m we do not match the true distribution but the distortion is certainly less significant than when the lower limit in distance starts at 1000 m.

We fix our distance cut as a trade-off between two requirements: i) the distortion in the profile and the average bias has to be as small as possible; ii) the number of contributing muons should be the largest possible. According to those requirements we choose as the minimum distance to the core 1500 m. Hereafter our integration range will be $1500 < r[\text{m}] < 4000$. In this distance range the contribution to the bias coming from the geometry reconstruction is 85 g/cm^2

The fact that the response of the WCD to muons shows a spread on time is the largest contribution to the average X_{max}^μ , similar to what happens for the case of the event-by-event analysis. If we compare the reconstructed average X_{max}^μ after all the reconstruction chain (i.e., adding the signal to all the other contributions) with its true value, the bias obtained is 124 g/cm^2 .

In table 6.1 and figure 6.5 we summarize and quantify all the identified sources contributing to the average X_{max}^μ bias in this zenith and distance range. The right plot of figure 6.5 is the average profile resulting after all the reconstruction steps.

To quantify first the effect of the detector response and immediately after the effect of the geometry reconstruction we have inverted the order in which those effects are evaluated. Doing this, Δ_{signal} amounts to 140 g/cm^2 . The difference obtained in the results (see table 6.1) comes from the correlation between the geometry reconstruction and the detector response.

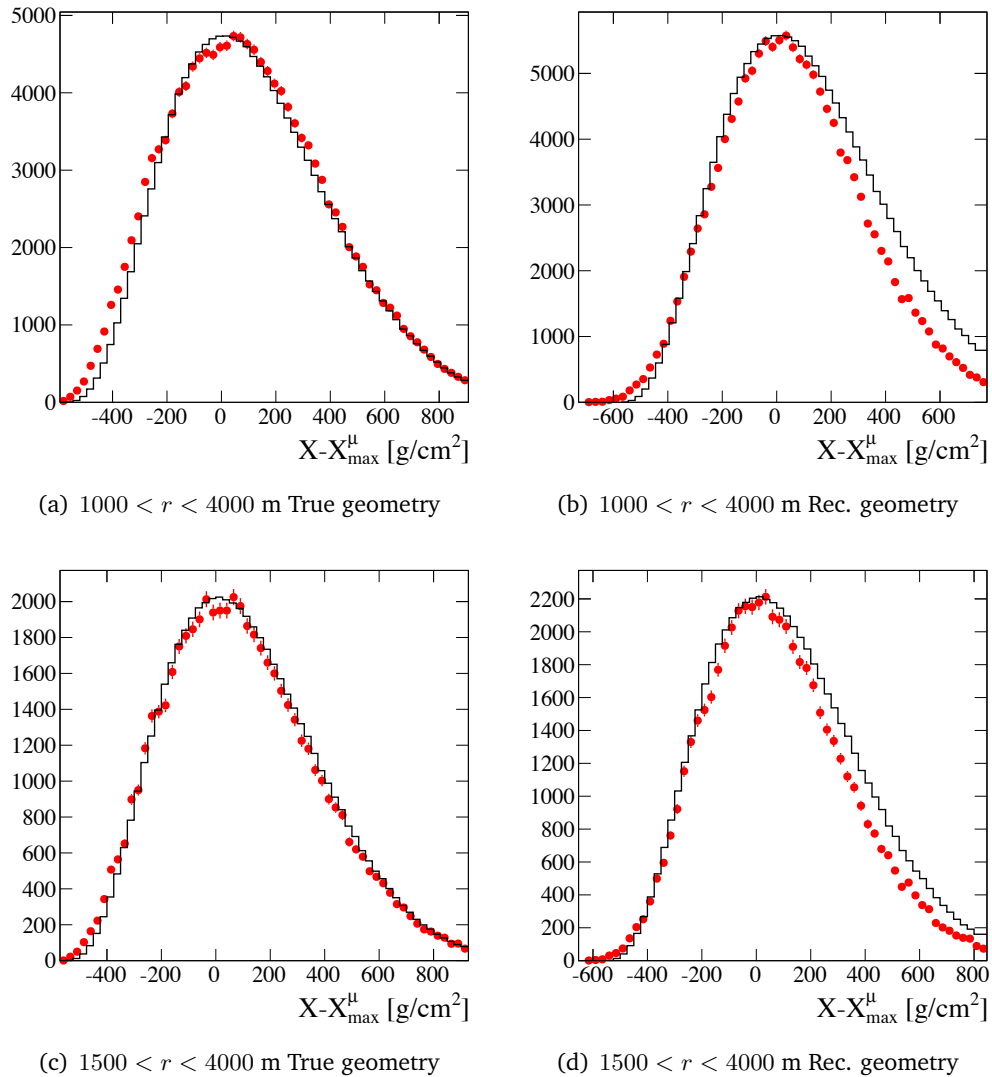


Figure 6.4: *Left*: Average distribution using the model, the new kinematic delay parameterization, the true time and the true geometry. Dots represent the reconstructed distribution and the line is the histogram directly obtained from CORSIKA. *Right*:. The same but using the reconstructed geometry instead of the true one.

6.2.2 Deconvolution

The main contribution to the average X_{\max}^{μ} bias comes from the WCD response. The time response of the detector to vertical muons is illustrated in figure 6.6 and can be described with the following equation:

$$F[t] = A * (e^{-t/65} - e^{-t/4}) \quad (6.1)$$

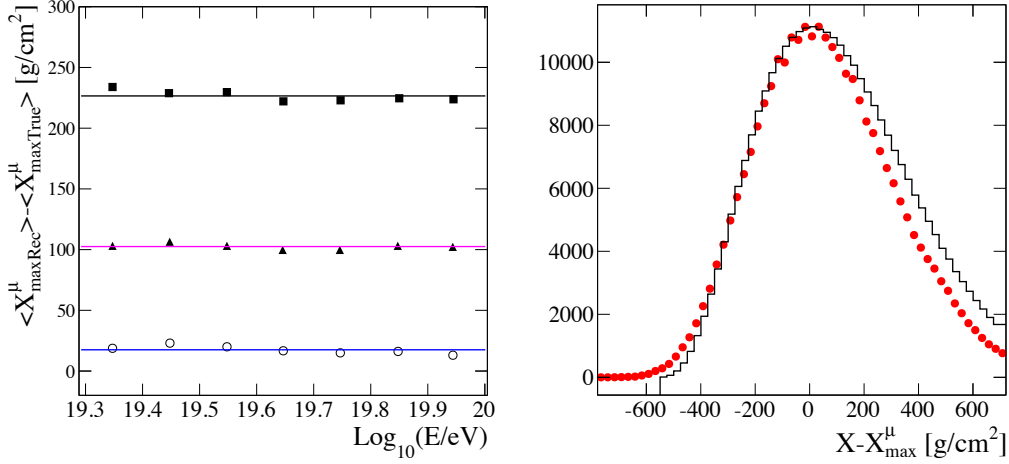


Figure 6.5: *Left* Cumulative bias contributions to the average X_{\max}^{μ} bias. Empty dots correspond to the method accuracy plus the new kinematic delay parameterization. The filled triangles include the method, the kinematic delay parameterization and the geometry reconstruction. Finally, the squares represent the bias after all effects are included. *Right* Average distribution at $10^{19.55}$ eV after all the reconstruction chain.

Source	Bias [g/cm ²]	Source	Bias [g/cm ²]
$\Delta_{\text{Method}} + \Delta_{\text{kin.Del.Par}}$	18	$\Delta_{\text{Method}} + \Delta_{\text{kin.Del.Par}}$	18
+ Δ_{geometry}	103	+ Δ_{signal}	158
+ Δ_{signal}	227	+ Δ_{geometry}	227

Table 6.1: Cumulative contributions to the average X_{\max}^{μ} bias between 1500 and 4000 m.

where A is the normalization and t the time both measured in ns. The time distribution of the signal at the WCD is convolved with this response. As it was shown in [129] we can use the *gold deconvolution* algorithm implemented in the ROOT package [138] (for more details see [139]) to deconvolute the trace and to obtain a good approximation to what the muon signal ought to be. In figure 6.7 we give an example of the effect of the deconvolution. In the right plot we apply the deconvolution to a muonic signal, red histogram, registered in one of the PMTs of a station at 1510 m from the core and with a total signal of 7 VEM. In the left plot we apply the deconvolution to the total trace. There the distinction between peaks produced by muons or by photons is not straightforward.

To minimize the effect of small peaks resulting from the residual electromagnetic contamination or baseline fluctuations we cut the deconvoluted signal at a certain threshold. We cut at $5\% \langle S_{\text{peak}} \rangle$, see figure 6.8. We chose this number since it guarantees that the value of the difference between the integrated signal from the peaks and the muonic signal is less than 5%. Likewise with this choice the difference between proton and iron biases is minimized. Note that if we cut very hard we will increase the difference between

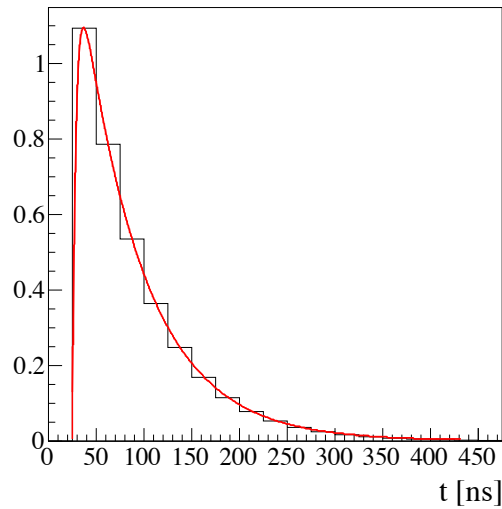


Figure 6.6: Histogram of the response to 10000 muons in the WCD. The line is the fit to a function $A * (e^{-t/65} - e^{-t/4})$.

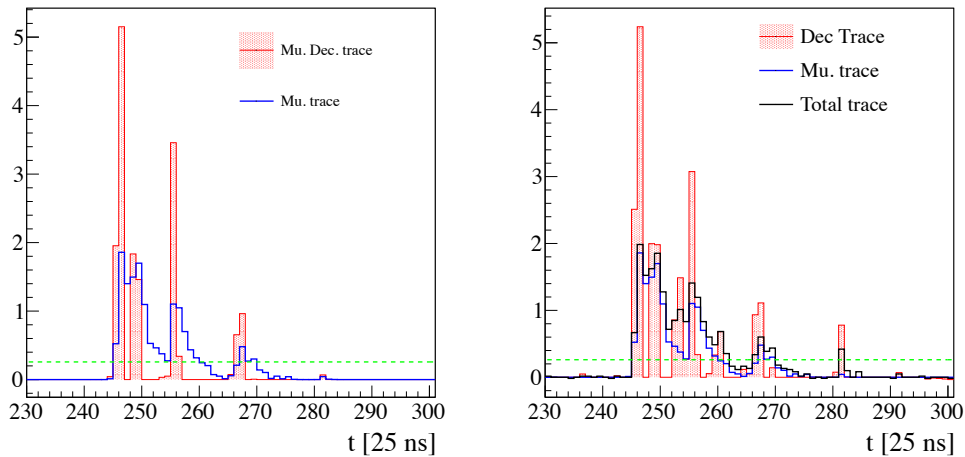


Figure 6.7: *Left*: Muon (blue line) and Deconvoluted muonic trace (filled red histogram) in one PMT of a station at 1510 m to the shower core and with a signal of 7 VEM. *Right*: The same PMT but including the total trace (black line) and applying the deconvolution to the total trace.

primaries.

In figure 6.9 we compare the bias obtained using the true geometry and the raw signal with the one obtained using the true geometry and the deconvoluted signal with the threshold cut. As it is illustrated in the plot the deconvolution reduces the bias by a factor 2. However, we still have a bias because we apply the deconvolution in the traces of the

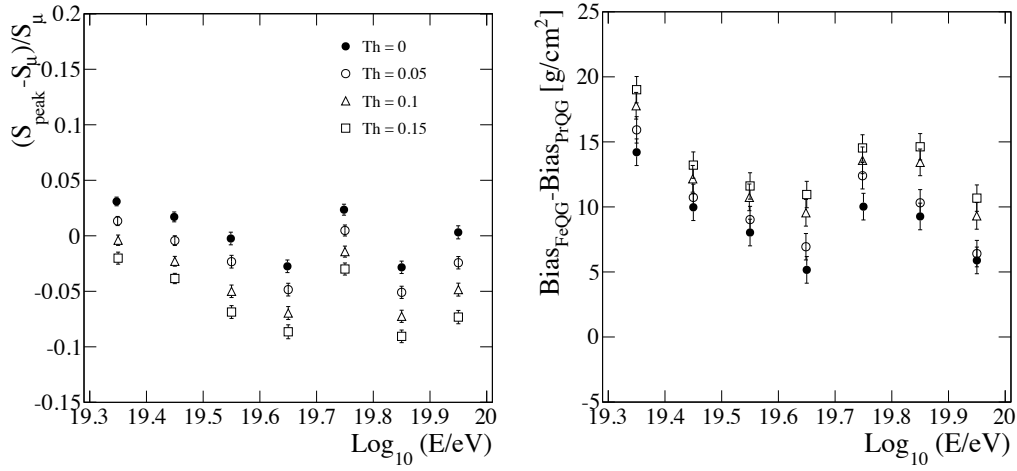


Figure 6.8: *Left*: Difference between the signal obtained from the integration of the peaks and the muonic signal for different threshold cuts. *Right*: Difference between proton and iron biases for different threshold cuts.

WCD. Those traces include two effects: a filter response and a binning of 25 ns (FADC of 40 MHz). This translates into a persistent spread in time. Therefore muons do not appear as delta-functions in our analysis. The actual situation is illustrated in figure 6.10. The empty histogram shows the response to crossing muons. The blue filled histogram is the result of the deconvolution in the trace of 10000 vertical muons after cutting at 5% $\langle S_{\text{peak}} \rangle$. The mean is at 37.5 ns. We will use this number as a *time shift*, t_{shift} , equivalent to the one used in the standard analysis.

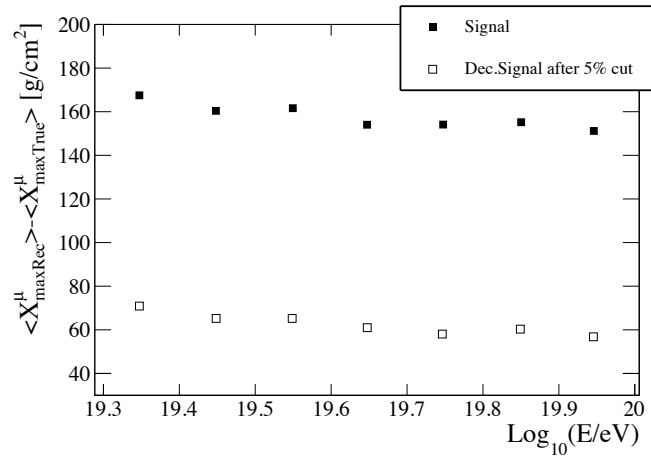


Figure 6.9: Bias in the average X_{max}^{μ} coming from the method plus the new kinematic delay parameterization and the signal (filled squares) compared to the one obtained in the same conditions but after applying the deconvolution and the threshold cut.

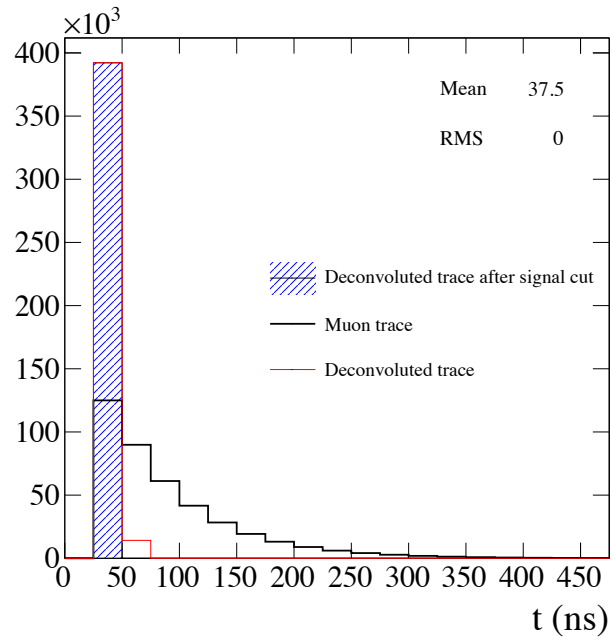


Figure 6.10: Histogram of the response of muons in the WCD after deconvoluting the trace (empty histogram) and after cutting the deconvoluted trace at 5% of the peak (filled histogram).

Finally, in figure 6.11 we compare the average distribution using always the true geometry. The raw signal is used to produce the left plot; the right plot is the result of applying the deconvolution, the threshold cut and the time shift. Our approach clearly removes the distortion and generates a faithful reproduction of the shape of the distribution.

6.2.3 Final bias and shape after all the reconstruction chain

We discussed before that candidate stations are those whose distance to the reconstructed core position is larger than 1500 m. To avoid unwanted fluctuations due to threshold effects, we disregard those detectors whose signal is smaller than 3 VEM. These cuts also have the effect of fixing our energy threshold, simply because not all the events have a station at this distance range; particularly when it comes to low energies. In figure 6.12 we show the efficiency for both primaries and models and for data as a function of the energy. To guarantee a selection efficiency larger than 95% we only take events above 10 EeV. No biases among species is introduced and data follows reasonable well the trend observed in simulations.

Figure 6.13 illustrates the final bias for all the considered species once the whole reconstruction and analysis chain is applied for the chosen distance and energy ranges. We show in the right plot the average bias for all the considered cases. We observe that the bias shows a soft dependence with the logarithm of the energy of the primary particle. The behaviour of this bias has been parameterized with the following linear function (the result of the fit is the red line shown in figure 6.13 right):

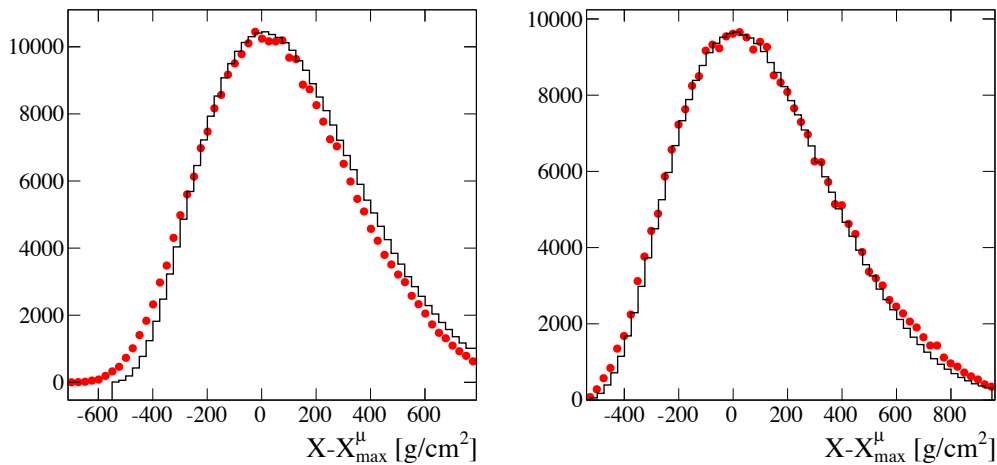


Figure 6.11: *Left*: Average distribution using the model, the new kinematic delay parameterization, the signal and the true geometry. Dots represent the reconstructed distribution and the line is the histogram directly obtained from CORSIKA. *Right*:. The same but applying the deconvolution and the cut at 5% and the time shift.

$$Bias[\log_{10}(E/eV)] = (296 \pm 10) - (11 \pm 0.5) \times \log_{10}(E/eV) \text{ [g/cm}^2\text{]} \quad (6.2)$$

Figure 6.14 shows the final distribution obtained after all the reconstruction chain has been applied. The distortion of the profile comes from the approach used to reconstruct the geometry of the event. To extract valuable physics information from the shape of the distribution we must improve the geometry reconstruction. Particularly in this zenith range where the official reconstruction for horizontal and vertical events do not show an optimal performance. Improving the geometry reconstruction is something that goes beyond the scope of this study. Therefore we will not discuss this issue any further. We have to focus the goal of the analysis on the determination of the average X_{\max}^{μ} .

The resulting biases obtained for all the primaries and hadronic models have been corrected using equation 6.2. The spread for all the primaries and models is $\pm 15 \text{ g/cm}^2$ (see top panel of figure 6.15). In the bottom plot we have compared the reconstructed maximum obtained from simulations (dots) with the expectation values obtained from CORSIKA (lines) and both are in good agreement within our uncertainty of $\pm 15 \text{ g/cm}^2$. The approach is therefore robust and can be applied to experimental data in order to extract valuable information about the mass composition of UHECR. Note that the method can be applied to SD data with energies well beyond those reached by the composition studies done with the FD.

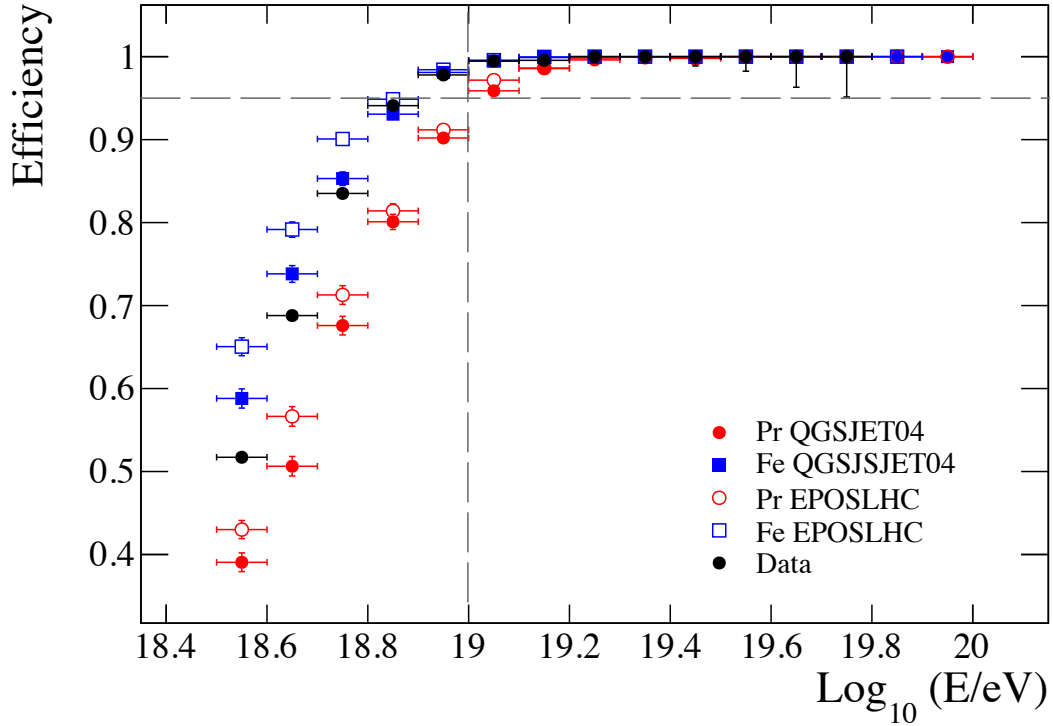


Figure 6.12: Efficiency for proton and iron QGSJET-II-04 and EPOS LHC and for data as a function of the energy.

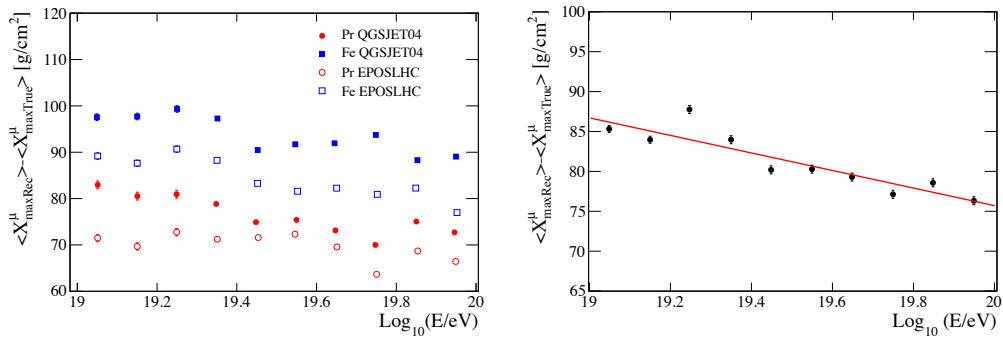


Figure 6.13: *Left* Bias after corrections for all primaries and models. *Right* Average bias for all the cases. The line represents the fit to a linear function.

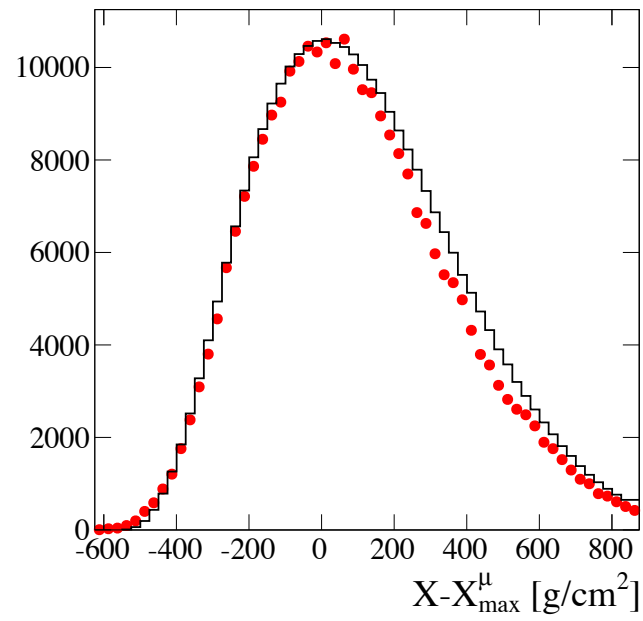


Figure 6.14: Average distribution after all the reconstruction chain.

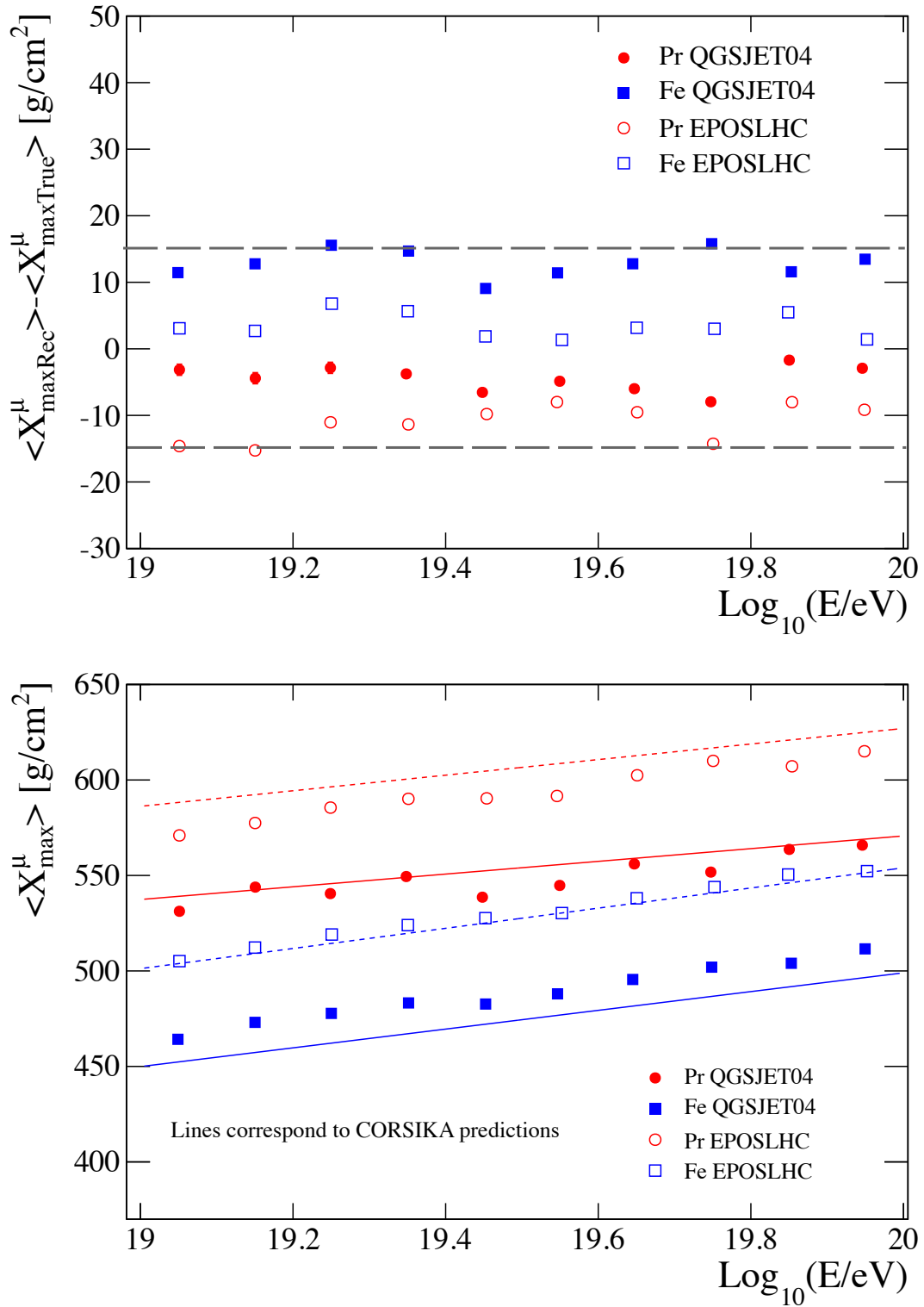


Figure 6.15: *Top:* Final bias for all the species after correcting the bias with equation 6.2. *Bottom:* The reconstructed $\langle X_{\max}^\mu \rangle$ after applying the bias correction for both primaries and models obtained from simulations compared to CORSIKA predictions.

6.3 Application to SD data

The analysis chain described in the previous sections has been applied to the data recorded with the Surface Detector of the Pierre Auger Observatory. The data set comprises events recorded between 1 January 2004 and 31 October 2014. To guarantee a notable quality of our working sample, lightning events and bad periods have been excluded with minimum effect on the resulting statistics. We require that all the events pass the 6T5 criteria to avoid the influence of the edges of the array. We require also that at least the event has one station at $r > 1500$ m with $S > 3\text{VEM}$. The total number of events and its selection efficiency is summarized in table 6.2:

Cut	Events	Efficiency [%]
$E > 10 \text{ EeV}$ and $\theta \in [55^\circ, 65^\circ]$	2668	100
6T5 trigger	2351	88.1
$r > 1500$ m and $S > 3\text{VEM}$	2345	87.9

Table 6.2: Cuts applied to data and its relative selection efficiency.

The distributions for each energy bin are shown in figures 6.16 and 6.17. The quoted errors correspond to the statistical uncertainties associated to the total number of muons that give rise to the average distributions. The conversion factor between the MPD distribution and the number of muons comes from a very rough estimation obtained from simulations, where the true information of the recorded particles is available. The set of average maxima as a function of the energy is shown in figure 6.18. The values come from a fit with a GH function and after applying the bias correction given in equation 6.2. Data shows a very mild dependence with energy, exception made of the data point corresponding to the most energetic events (there are 32 for energies above $10^{19.7}$ eV). Their average maximum suggests they are composed of a majority of lighter primaries that penetrate more in the atmosphere. However to make solid inferences about mass composition we have to compare this data with the predictions offered by the state-of-the-art models of ultra-high energy hadronic interactions.

6.3.1 Systematic uncertainties

We now turn our attention to identify and evaluate the contributions of the different sources of systematics that influence the measurement of the average X_{max}^μ . In what follows, we only discuss the most relevant contributions:

- **Reconstruction, hadronic model and primary mass.** As it is illustrated in figure 6.15 the difference between the generated and the reconstructed $\langle X_{\text{max}}^\mu \rangle$ is bracketed by $\pm 15 \text{ g/cm}^2$ for proton and iron primaries of both hadronic models (QGSJET-II-04 and EPOS LHC) with energies above 10 EeV and with zenith angles between 55 and 65°. We take this value as an estimate of the overall systematic uncertainty due to the reconstruction effects, differences in the hadronic interaction models, and differences due to the unknown nature of the primary particle.

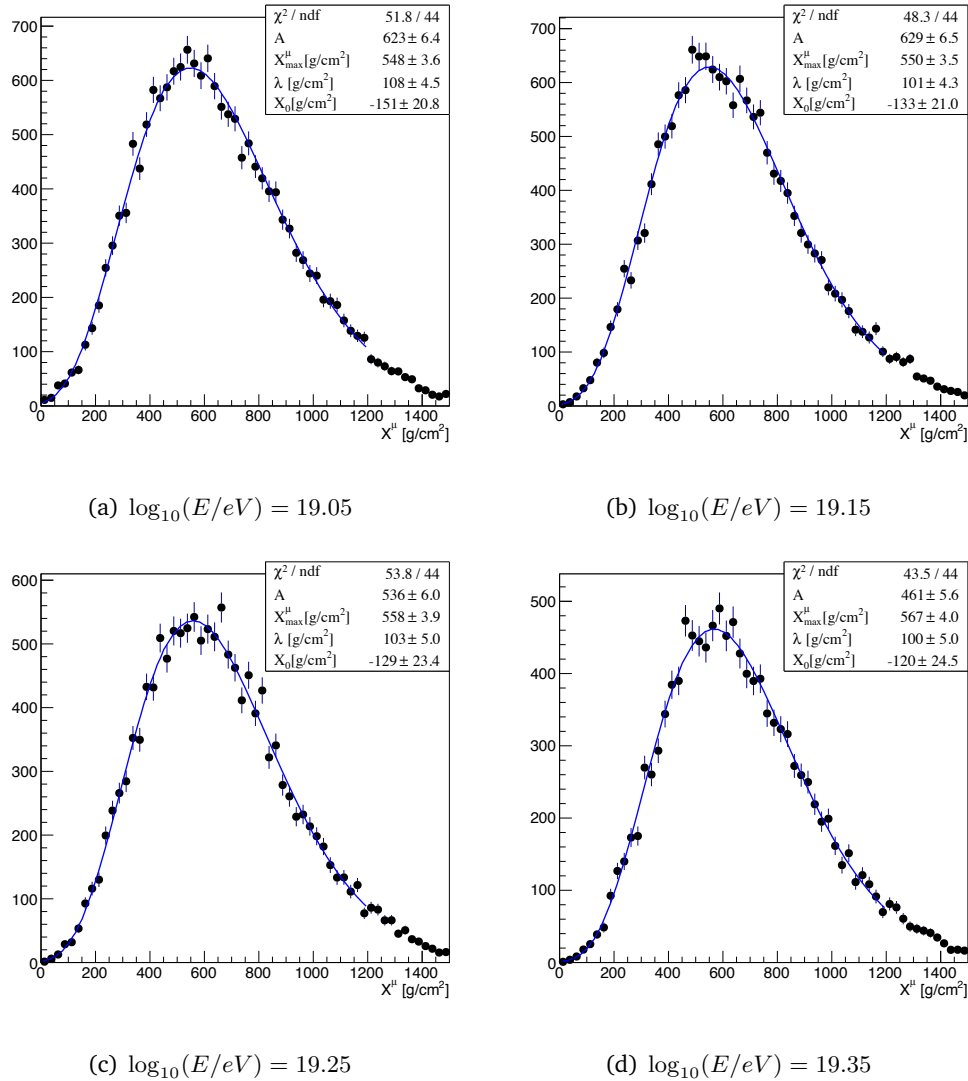


Figure 6.16: Application to experimental data: Average distributions.

- **Seasonal effect.** The data show a dependence of unknown origin of the measured average X_{max}^{μ} value with seasons. This effect is shown in figure 6.19 where the measured $\langle X_{\text{max}}^{\mu} \rangle$ is deeper in summer. Monte Carlo events generated for different seasons do not show such a modulation (either at generation or after detector effects are considered). Several tests probed unsuccessfully to identify the source of this discrepancy, and therefore we include the amplitude of this effect as a systematic uncertainty.
- **Time variance model.** The uncertainty on the arrival time of the EAS front has been modeled from data themselves. It influences the reconstruction of the curvature and of the impact point on the ground, and it has a direct effect on the reconstruction of the average maximum of the MPD. To obtain the associated systematic uncertainty

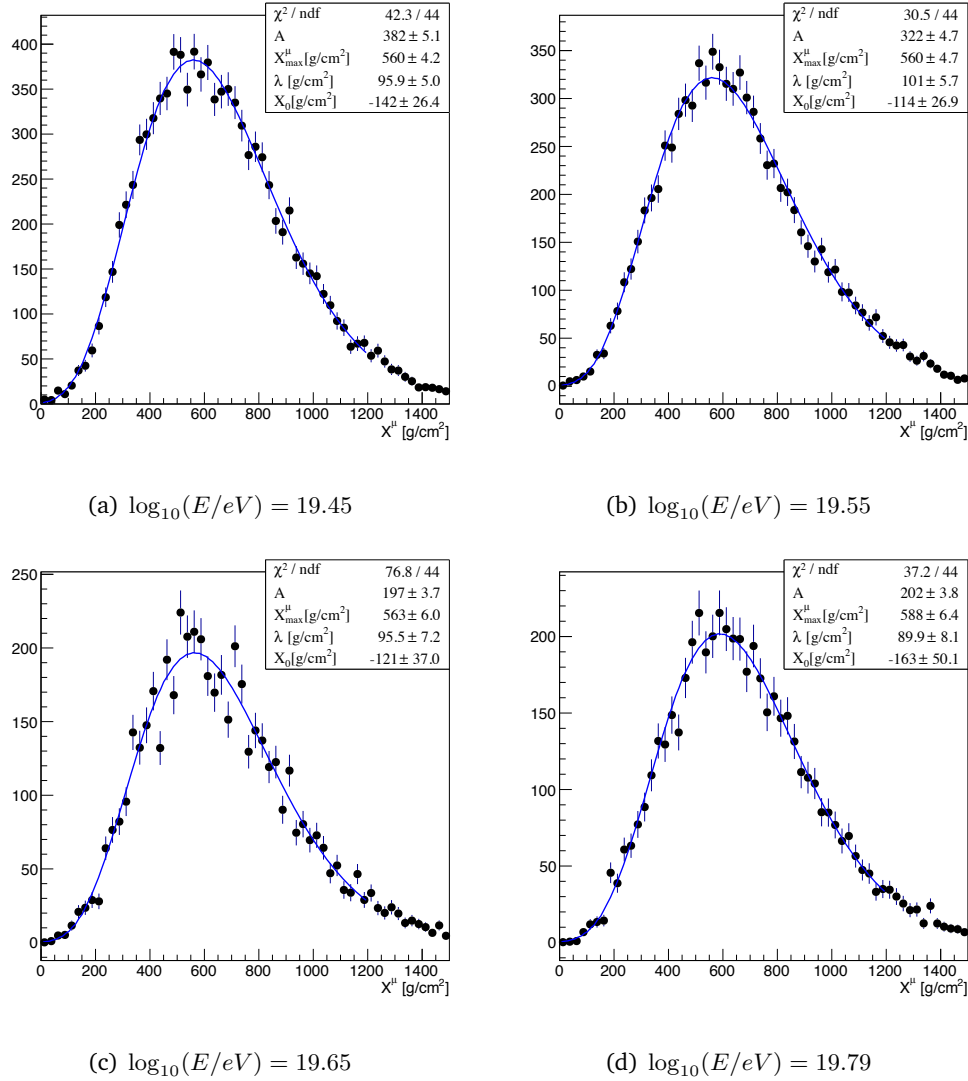


Figure 6.17: Application to experimental data: Average distributions.

introduced by this model we compared two different parameterizations of the time variance model [122]. They both have a common contribution from the resolution on the absolute time given by the GPS and from the 40 MHz sampling of the FADCs, but they differ in the modeling of the fluctuations of the arrival time of the first particle. We include as a systematic uncertainty the difference between the two models on the determination of the $\langle X_{\text{max}}^{\mu} \rangle$. It amounts to 5 g/cm^2 .

- **Accidental signals.** In real events, a background of random accidental signals might appear. The most frequent source of random noise is created by single particles (generally isolated atmospheric muons) and, more rarely, by a bunch of particles arriving at the same time from a low-energy shower close to a SD station. In general, it is very difficult to identify and take into account all possible sources of accidental sig-

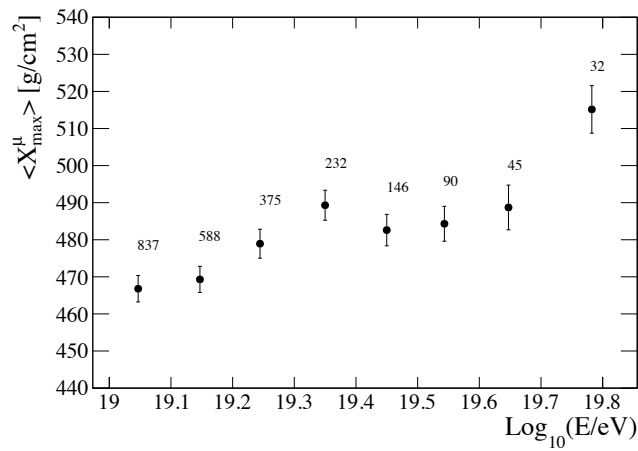


Figure 6.18: Evolution of the average X_{max}^{μ} as a function of the primary energy for the data of the Pierre Auger Observatory between 1 January of 2004 and 31 October 2014. The number of events in each energy bin is indicated.

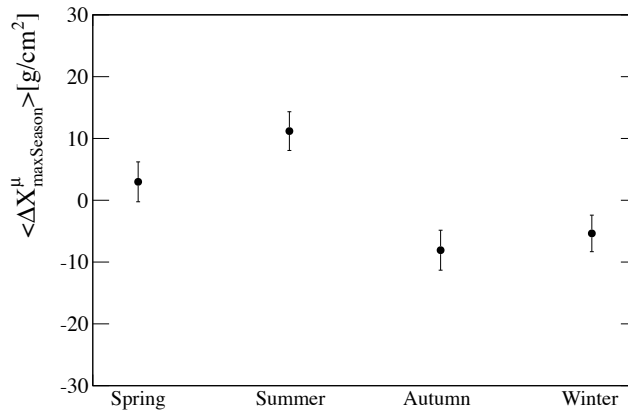


Figure 6.19: Seasonal effect.

nals. They can appear at any time and at any location in the SD array, completely uncorrelated with the genuine primary shower signal. Random accidental signals can have a damaging effect on the data quality, since they can trigger some stations of the array, distorting the reconstruction of the showers. In the analysis, the main impact comes from a possible underestimation of the start time of the traces due to an accidental signal prior to the true one. Using an unbiased sample of random accidental signals extracted from data events collected in the SD stations, we have studied the influence of accidental signals in the Monte Carlo reconstructions. Regardless of the energy and primary mass, we have found a systematic underestimation of \sim

4.5 g/cm² in the determination of the average X_{\max}^{μ} . We have corrected for this bias in data.

- **Energy scale.** The systematic uncertainty on the energy scale is 14%. We have studied the effect of this uncertainty in the determination of the average X_{\max}^{μ} . In figure 6.20 is illustrated the elongation rate using the reconstructed energies, and the reconstructed energies plus (minus) 14%. The RMS of the points in the bottom panel of figure 6.20 is taken as a systematic and it amounts to ~ 7 g/cm².
- **Atmospheric profile.** For the reconstruction of the MPD profiles, the atmospheric conditions at the Auger site, mainly height-dependent atmospheric profiles, have to be well known. To quantify the influence of the uncertainty in the reconstructed atmospheric profiles on the value of $\langle X_{\max}^{\mu} \rangle$, a direct comparison of GDAS data¹ with local atmospheric measurements² has been performed obtaining a small shift of 2 g/cm².
- **Ageing effect.** As shown in figure 6.21 no significant effect due to the ageing of the detector is observed. A fit to a linear function has been performed being the slope compatible with 0. We observed a deviation for 2004. A constant fit has been applied including and removing this year, and in both cases the effect is below 1 g/cm². In any case we decided to remove this year as a safety prescription. The effect of removing this year has negligible effects on the evolution of the average X_{\max}^{μ} with energy (see figure 6.22). The data sample is reduced from 2,345 to 2,319 events (approximately 1%).

Table 6.3 summarizes all the sources to the systematic uncertainty affecting our measurement. The overall value amounts to 20.5 g/cm².

Source	Sys. uncertainty [g/cm ²]
Reconstruction, hadronic model and primary mass	15
Seasonal effect	11
Energy scale	7
Time variance model	5
Total	20.5

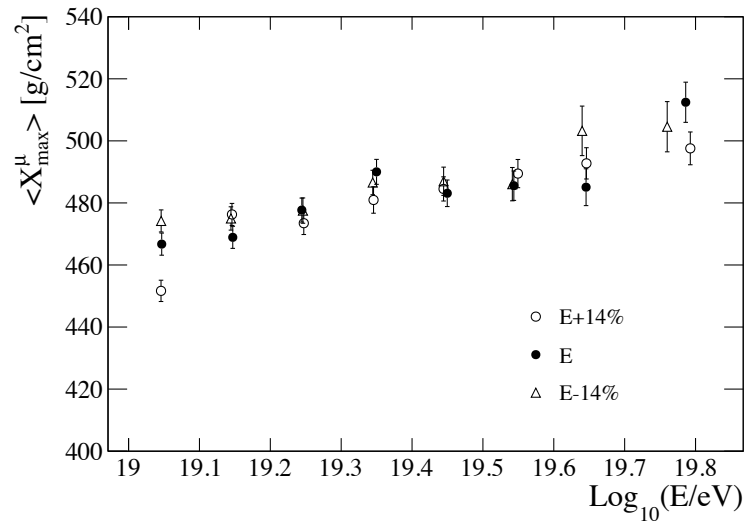
Table 6.3: Contributions to the systematic uncertainty in the reconstruction of the average X_{\max}^{μ} .

6.3.2 Estimation of cosmic ray mass composition through the mean logarithmic mass, $\ln A$

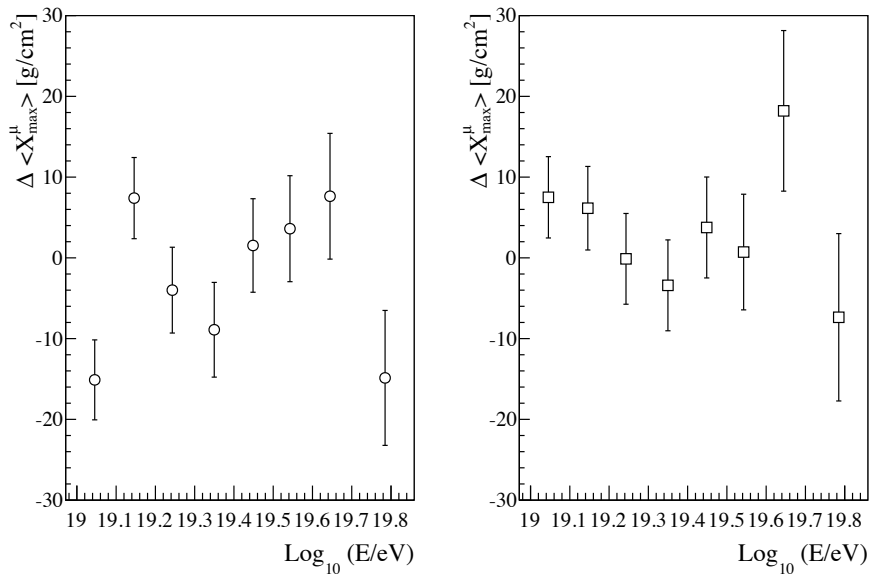
It is not possible to infer the mass of our selected data sample without the help of air shower simulations. Our measurements are compared to the predictions given by the

¹GDAS is a publicly available data set containing all main state variables dependent on altitude with a validity of 3 hours for each data set [140, 141].

²Intermittent meteorological radio soundings with permanent ground-based weather stations.



(a) $\langle X_{\max}^{\mu} \rangle$ as a function of the energy and as a function of the energy $\pm 14\%$.



(b) Differences in $\langle X_{\max}^{\mu} \rangle$ for E+14% (left) and E-14% (right).

Figure 6.20: Energy scale systematic uncertainty.

latest versions of the packages QGSJET-II-04 and EPOS LHC, which have been tuned to reproduce the most recent results published by the experiments collecting data at the LHC. Figure 6.23 shows the result of such comparison. Black dots correspond to data. The bars represent the systematic uncertainty. Events are grouped according to their $\log E$ value in bins of width 0.1. The rightmost bin contains all the events whose energy is larger

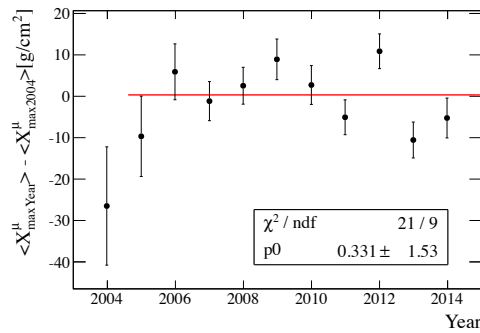
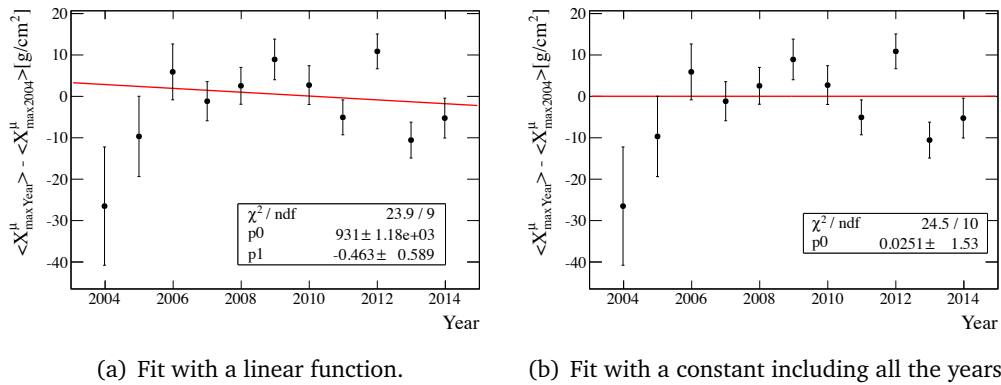
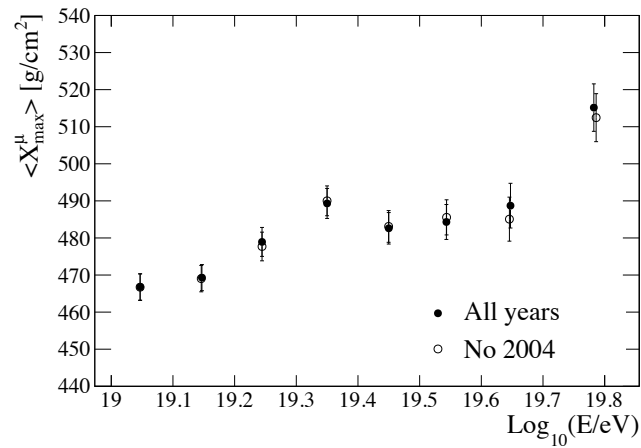


Figure 6.21: Aging effect.

Figure 6.22: Average X_{\max}^{μ} as a function of the primary energy for all the data set (filled dots) and excluding 2004 (empty dots).

than 50 EeV. Solid lines represent the predictions of QGSJET-II-04. Data is well bracketed by the values corresponding to proton and iron primaries. According to this model the data is compatible with a composition largely dominated by heavy elements. Interestingly enough is the change of trend manifested in case the most energetic events are considered. Here it seems that a lighter component starts to emerge as energy increases and thus a deeper average X_{\max}^{μ} value is obtained.

For the case of EPOS LHC predictions and data are at odds. Data is heavier than the expectations corresponding to iron primaries (the most-massive stable element known). Therefore, as we saw already for the case of the event-by-event analysis, the average MPD can also be used to test and constrain the assumptions contained in the packages that model hadronic interactions at ultra-high energies.

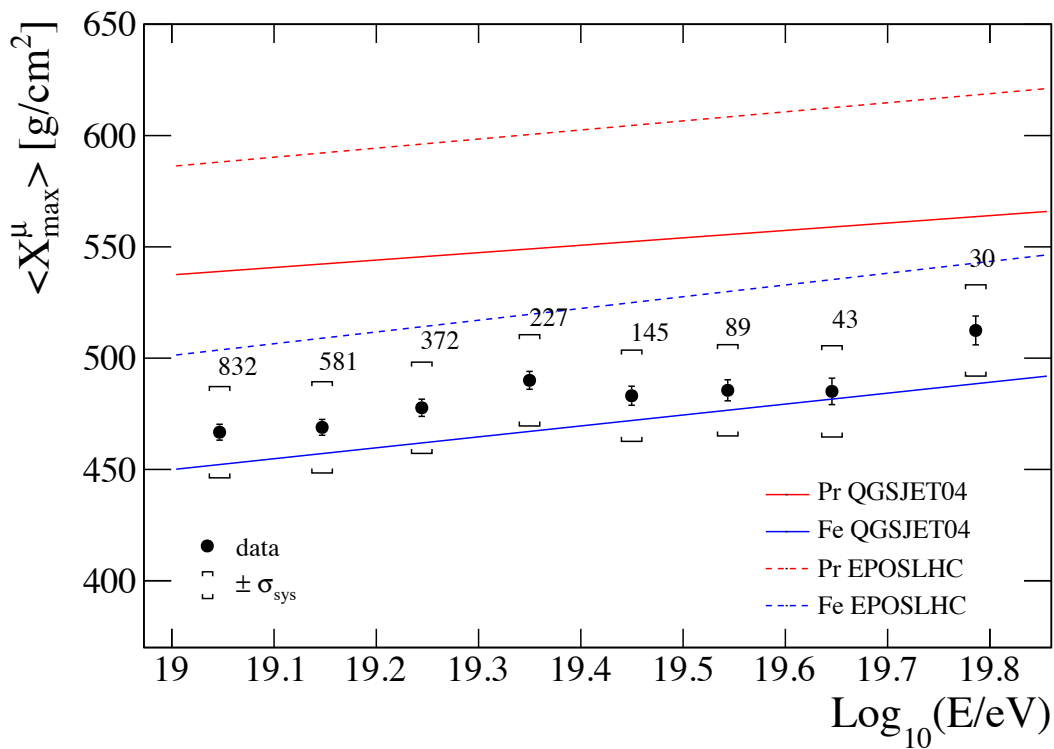


Figure 6.23: Energy dependence of the maximum obtained from average MPD distributions. Data is compared to expectations from simulations.

It is interesting to compare the shapes of the eight average MPD distributions obtained from our selected data sample to simulations. We look for features that might evidence possible shortcomings in the description that simulations do of muon profiles when compared to data. Figures 6.24 and 6.25 show that the general features of the average muons profiles obtained for data (black dots) follow similar trends to the ones corresponding to iron and proton (histograms). The maxima of the distribution are clearly shifted towards the values favoured by heavy elements for the whole range of considered energies.

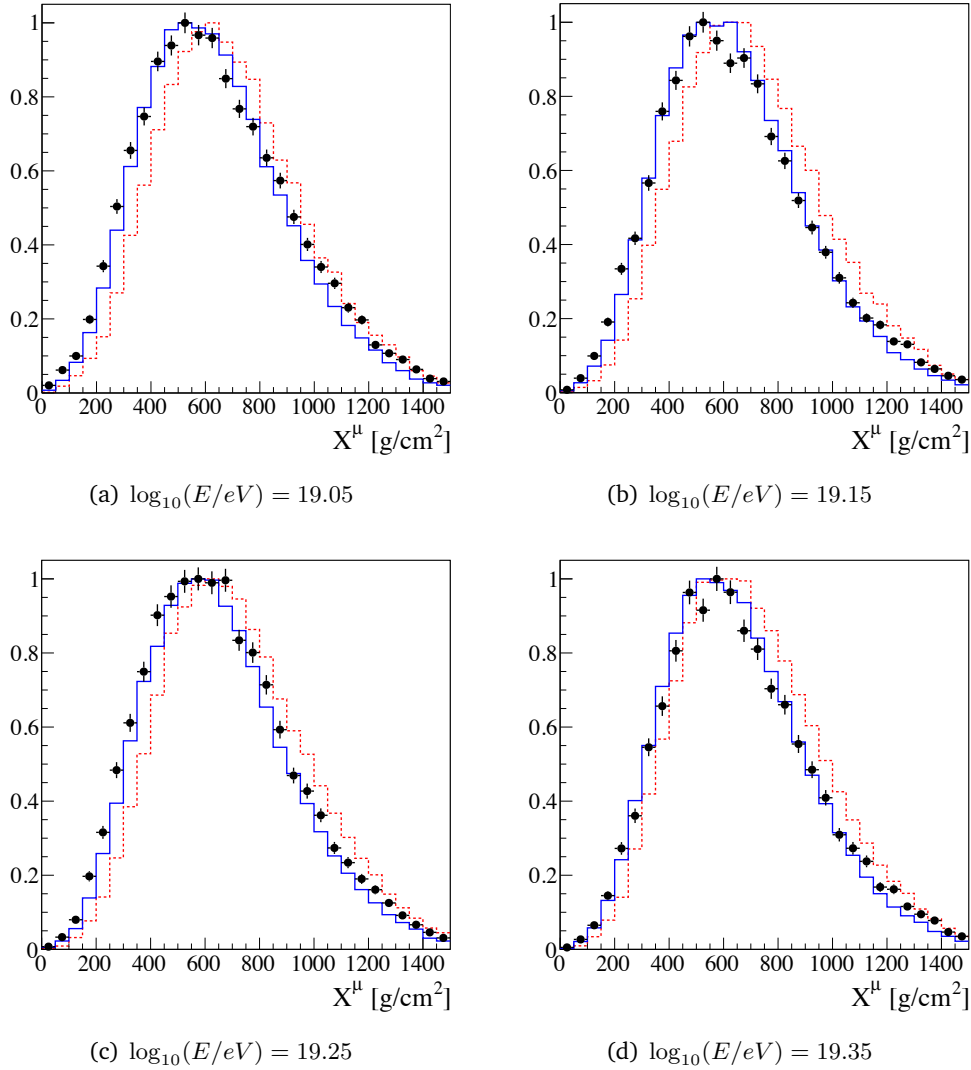


Figure 6.24: Average distributions for data (dots) and QGSJET-II-04 predictions (the continuous line corresponds to iron and the dotted line to proton).

Since the average X_{\max}^μ is strongly correlated with the mass of the primary particle, we can convert the values of this observable into average values of the logarithm of the atomic mass number A , following the prescriptions explained in previous chapters. Figure 6.26 shows the result of such conversion for the two models considered throughout this thesis. Compared to the event-by-event MPD analysis, the study of the average MPD allows to go down up to energies of 10 EeV. The mass estimations of this analysis agree well with the results of the event-by-event study. The values at the highest energies are compatible within the quoted systematic errors. It is interesting to note that for the range of energies where the FD analysis and the present study overlap, the evolution of $\langle \ln A \rangle$ as a function of the energy follows a similar trend. When looking at the predictions for

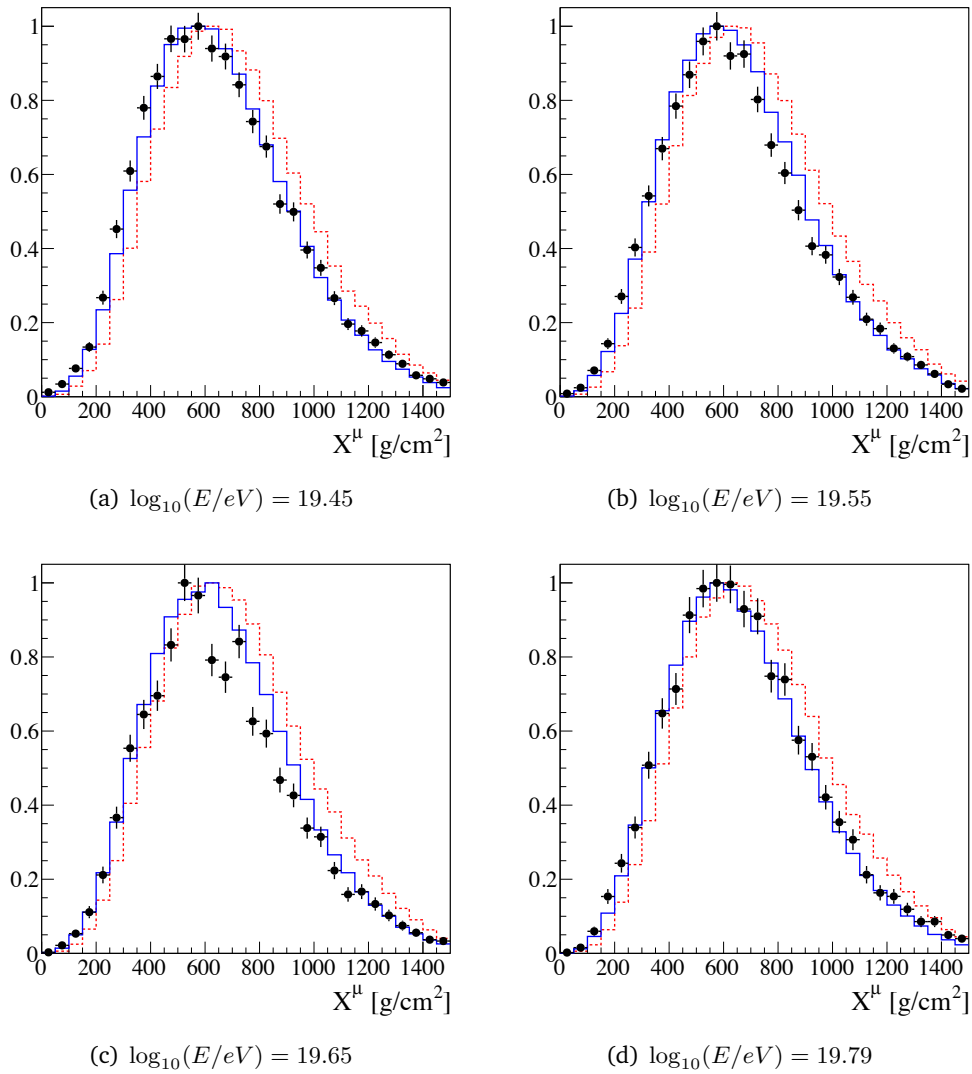


Figure 6.25: Average distributions for data (dots) and QGSJET-II-04 predictions (the continuous line corresponds to iron and the dotted line to proton).

each particular model, our average MPD measurement favours a heavy composition for the case of QGSJET-II-04. The disagreement with the FD values is at the level of 1σ . For the case of EPOS LHC, the average mass is above the value corresponding to iron (thick horizontal line) and the disagreement with FD values varies around 3σ .

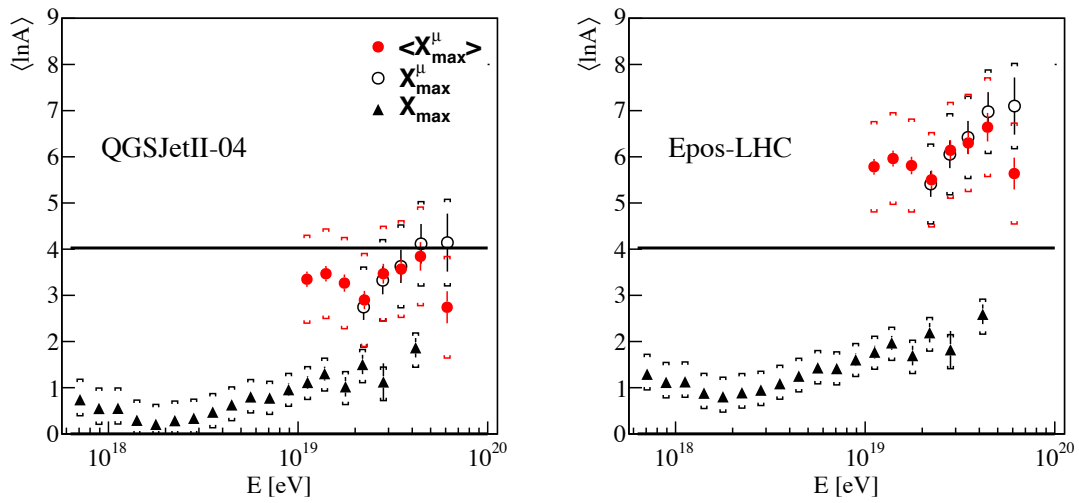


Figure 6.26: Conversion of $\langle X_{\max}^{\mu} \rangle$ (filled circles), X_{\max}^{μ} [1] and $\langle X_{\max} \rangle$ [136] to $\langle \ln A \rangle$, as a function of energy. On the left (right) plot we use QGSJET-II-04 (EPOS LHC) as the reference hadronic model. Brackets correspond to the systematic uncertainties.

New detector configurations and potential improvements of MPD studies

7

During the last two years the Pierre Auger collaboration has been actively devising proposals for the upgrade of the Surface Detector. The final decision was to use scintillators on top of the WCD and to update the current electronics [142]. Those configurations had to be simulated to evaluate the expected improvements with respect to the current detector.

For reasons that will become more evident later in the chapter, one of the proposals that could influence more MPD-based analyses was to change the internal properties of the liner in order to reduce the time response of the WCD [143]. Motivated by this idea we proceed to implement the required changes in the `Offline` [107] software to be able to perform Geant4 simulations. All the changes have been implemented in a new branch, `v2r8-blacktank-experimental` [144]. The configuration of the electronics has been updated to match the features of the proposed upgrade.

In this chapter we document the changes done in the branch and discuss the cross-checks we performed to understand the impact the proposed changes will have in the physics. We also describe how the resolution on the reconstruction of the maximum of the muon production distance is affected by those changes. Finally we redo this analysis considering the case of a *dense array* of WCD.

7.1 Implementation of new configurations in Offline software

As introduced in chapter 3, the Auger offline software, `Offline`, is a framework which provides the mechanisms for encapsulating in modules the steps involved in the simulation and reconstruction of EAS. It also provides the means to configuring and sequencing these modules by external files. It has three principal components: a Framework, a detector description and an event description. We applied the required changes in each one of the components.

Framework changes The detector properties are coded inside the Framework. Because the physical properties of the Surface Detector were not expected to be changed, some of them were hard-coded. The main changes that had to be tracked all along the different modules were the length of the trace, the sampling frequency and the dynamic range. Another WCD property that has been changed at Framework level is the number of PMTs.

With this we are able to accommodate different detectors configurations like a Layered Surface Detector [145]. These configurations can be set in the SModels.xml card:

```
<nrOfPMTs> 3 </nrOfPMTs>
<frequency unit="MHz"> 120. </frequency> <!-- traceLength in bins -->
<traceLength> 2304 </traceLength> <!-- number of bits in the FADC -->
<dynamicRange> 12 </dynamicRange>
<topPMTsIds> 1 2 3 </topPMTsIds>
<bottomPMTsIds> </bottomPMTsIds>
```

All these changes have been propagated throughout the simulation and reconstruction modules. One further change has been performed by [146] to be able to use a different size for the PMT, however it is not discussed here since it has no interest for what follows. The variables are all now accessible via functions implemented in the SDetector.

Changes in the simulation modules The standard SD simulation chain implemented in the Offline software comprises the following modules:

```
<loop numTimes="unbounded" pushEventToStack="no">
<module> CachedShowerRegeneratorOG </module>
<module> G4TankSimulatorOG </module>
</loop>

<module> SdSimulationCalibrationFillerOG </module>
<module> SdPMTSimulatorOG </module>
<module> SdFilterFADCSimulatorMTU </module>
<module> SdBaselineSimulatorOG </module>
<module> TankTriggerSimulatorOG </module>
<module> TankGPSSimulatorOG </module>
<module> CentralTriggerSimulatorXb </module>
```

The ones in bold characters have been modified according to the new detector properties, while keeping the old configuration unaffected.¹

G4TankSimulatorOG. In the Geant4 module that deals with the particle and light propagation and interaction inside the WCD, we have implemented the option to support different number of PMTs and to define complete photon absorption for the bottom, top and lateral sides of the liner. These can be configured via a xml card.

The effect of the *black top* liner is that the light propagated to the top of the WCD is getting absorbed and therefore the number of photoelectrons arriving at the PMTs is reduced. Injecting vertical central through-going muons with an energy of 1 GeV we obtained for the case of a black top WCD 21 photoelectrons to be compared with the 80 photoelectrons obtained for a standard WCD. One very important property of the black top WCD is that it reduces the time response of the detector to a muon as illustrated in figure 7.1. The time response for the standard WCD is 70 ns and in case of the black top WCD it is 13 ns.

SdSimulationCalibrationFillerOG/SdFilterFADCSimulatorMTU. The strategy adopted for the SD simulations in Auger is that the calibration constants are simulated just once and then the same calibration numbers are used for all the detectors. This is done so

¹This has been cross-checked from station level up to the high-level reconstruction.

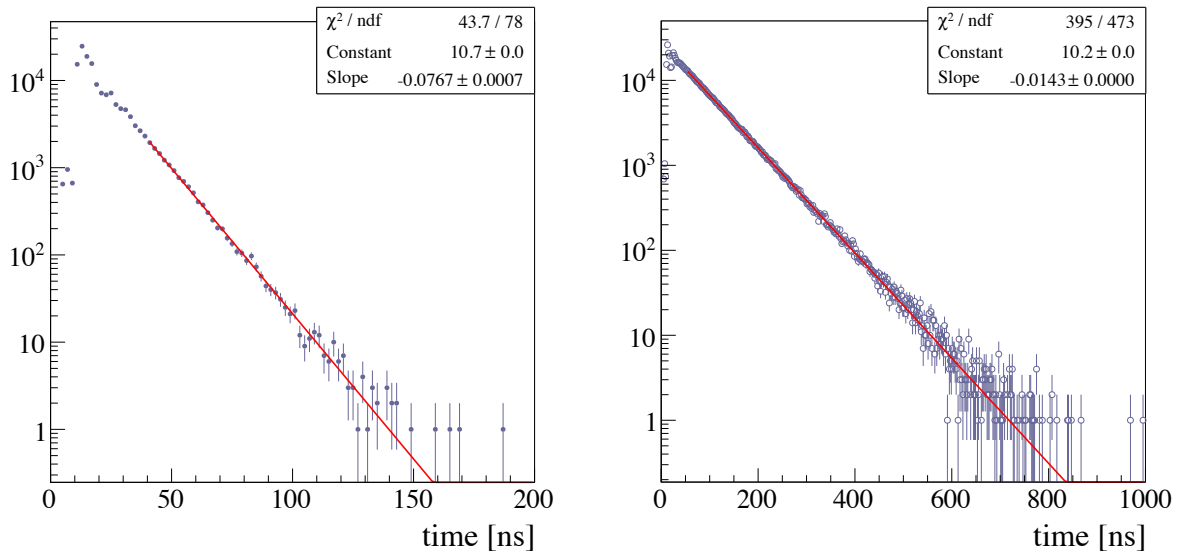


Figure 7.1: The distribution of photoelectrons in time for a black top liner (left) and for the standard liner (right).

	St. 40 MHz	St. 120 MHz	BT 40 MHz	BT 120 MHz
Charge	187	472	49	163
Peak	53	68	32	56
Nr bins	13	25	5	8

Table 7.1: Calibration constants for current liner (St.) and for a black top liner (BT) for two sampling frequencies. The last line represents the number of bins above threshold requested by the trigger.

because the differences between the response of the detectors are ignored in the simulations. The calibration constants, mainly the VEM peak and charge, are obtained by injecting 5000 vertical central muons.² The simulation of the Bessel filter present in the detector electronics has been changed to take into account automatically the FADCs frequency via a simple time scaling. Some of the most relevant values are given in table 7.1. The charge ratio between the standard WCD and the black top WCD is consistent with the ratio of photoelectrons.

TankTriggerSimulatorOG. The station trigger condition in Auger is based on obtaining a trigger rate of about 1 Hz. The TOT trigger requires a number of bins (13 currently) within a window of $3 \mu\text{s}$ with a threshold of 0.2 VEM in time coincidence between at least 2 PMTs. We have investigated a new trigger condition, TOT2, which is similar to the TOT trigger condition, but does not require the time coincidence between the PMTs.

²A new module has been implemented that can inject the spectrum of particles, but it is not described here.

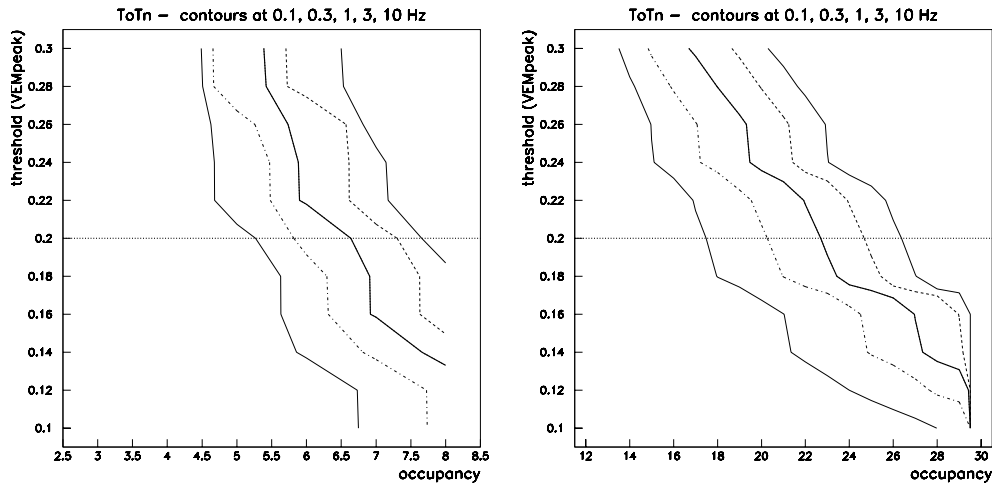


Figure 7.2: Example of trigger studies at 100 MHz for a black top liner (left) and for the standard liner (right). The occupancy represents the number of bins above threshold requested by the trigger. The contour lines are ordered as follows from left to right: 10 for the smallest occupancy up to 0.1 Hz for the largest.

In figure 7.2 the trigger rate as a function of the occupancy and the VEM threshold is illustrated as an example at 100 MHz. The number of bins for different frequencies and configurations are given in the last line table 7.1.

Reconstruction Modules The standard reconstruction module sequence looks like:

```

<module> SdCalibratorOG </module>
<module> SdEventSelectorOG </module>
<module> SdMonteCarloEventSelectorOG </module>
<module> SdPlaneFitOG </module>
<module> LDFFinderKG </module>
<module> SdEventPosteriorSelectorOG </module>
<module> RecDataWriterNG </module>

```

The main change has been made in SdCalibratorOG. The signal start time has been previously determined from the station which has been used for obtaining the calibrated total signal. We have changed the requisite that the start time is always obtained from the high-gain trace. This is necessary as there is a time delay between the dynode and the anode trace, which is not usually taken into account. The module has also been changed to account automatically for the different dynamic ranges and for different frequencies. The baseline subtraction algorithm depends heavily on parameters specific to different WCD responses. This algorithm has not been tuned for the new configuration, but a constant baseline subtraction is applied. The total signal and start time differences between 40 MHz and 120 MHz as a function of signal are illustrated in figure 7.3. The agreement

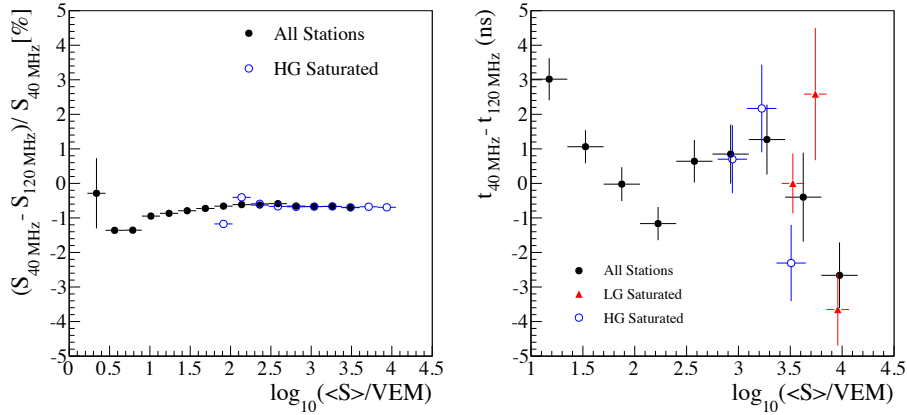


Figure 7.3: Station level variables. Total signal differences between stations versus total signal using 120 MHz and 40 MHz with 12 bits and white top.

	Mean	RMS
$\Delta S(1000)/S(1000)$ [%]	-0.81 ± 0.04	0.59 ± 0.03
$\Delta E/E$ [%]	-0.83 ± 0.05	0.72 ± 0.03
Angle [deg.]	0.152 ± 0.007	0.114 ± 0.005
Core [m]	1.49 ± 0.09 m	1.36 ± 0.06 m

Table 7.2: Reconstruction differences between 40 and 120 MHz for the standard WCD.

between the signals is very good, with the mean differences being smaller than 1% and 5 ns respectively.

The SdPlaneFitOG and LDFFinderKG are indirectly affected by the changes in the trunk. One of the modifications that affects marginally the angular reconstruction is related to the inclusion in the time variance model of the changes in frequency and GPS resolution. The reconstruction of the lateral distribution function is also slightly affected when one changes the dynamic range of the signal reading. Modifications have also been made in RecDataWriterNG to include new variables like a changing number of PMTs and frequencies. The EventBrowser can display now automatically different detector configurations.

All the changes have been checked up to the finest details. We have simulated 200 air-showers, half with protons as primary particle and half with iron. We used CORSIKA (Fluka/QGSJet II) with an energy of 30 EeV and an isotropic angular distributed between 0 and 60 degrees. The particles on the ground were fed to the detector simulations using permutations between different configurations: 40 MHz, 120 MHz, black top liner (BT), 10 bit FADCs and 12 bit FADCs .

The results for the comparisons of the main variables, i.e. $S(1000)$, energy, angle and core location, are given in table 7.2 for two frequencies, 40 MHz and 120 MHz. The differences are less than 1% in $S(1000)$ and energy, with an angular difference of less

	Mean	RMS
$\Delta S(1000)/S(1000)$ [%]	-0.8 ± 0.3	3.88 ± 0.18
$\Delta E/E$ [%]	-0.8 ± 0.3	4.09 ± 0.19
Angle [deg]	0.248 ± 0.014	0.216 ± 0.010
Core [m]	17.5 ± 1.0 m	15.4 ± 0.7 m

Table 7.3: Reconstruction differences between standard and black top WCD at 120 MHz.

Frequency	Mean		RMS	
	40 MHz	120 MHz	40 MHz	120 MHz
Candidates	16.5 ± 0.3	17.8 ± 0.4	5.4 ± 0.2	5.5 ± 0.3
Candidates (BT)	17.8 ± 0.4	19.3 ± 0.1	5.5 ± 0.3	6.1 ± 0.1
Core	66 ± 3 m	66 ± 3 m	45 ± 2 m	45 ± 2 m
Core (BT)	66 ± 3 m	62 ± 3 m	45 ± 2 m	45 ± 2 m
Energy [%]	-32.9 ± 0.8	-32.1 ± 0.8	12.4 ± 0.6	12.4 ± 0.6
Energy (BT) [%]	-32.1 ± 0.8	-31.3 ± 0.7	12.4 ± 0.6	11.1 ± 0.5
Angle [deg.]	0.33 ± 0.01	0.31 ± 0.01	0.21 ± 0.01	0.19 ± 0.01
Angle (BT) [deg.]	0.31 ± 0.01	0.31 ± 0.01	0.19 ± 0.01	0.19 ± 0.01

Table 7.4: Resolutions for 120 MHz white and black top (BT). The candidates represent the mean number of stations used in the reconstruction, and the rest represent the differences to the generated values.

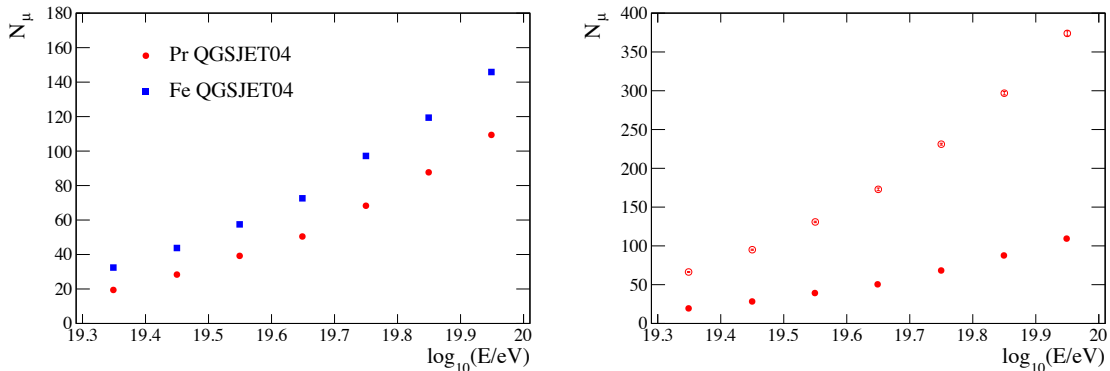


Figure 7.4: (left) Number of muons entering in the MPD distributions as a function of energy for proton and iron QGSJET-II-04 and the standard array. (right) The same for proton QGSJET-II-04 for the standard array (filled dots) and the dense one (empty dots).

than 0.2 degrees and a distance between the reconstructed core locations of less than 2 m. Similar results are obtained when comparing the standard liner configuration with a black top liner (see tables 7.3 and 7.4).

7.2 Resolution of the maximum of the MPD distributions

We showed in chapter 5 that the resolution on X_{\max}^{μ} has different contributions:

$$\sigma^2(X_{\max}^{\mu}) = \sigma_{\text{Method}}^2 + \sigma_{\text{Time}}^2 + \sigma_{\text{Geometry}}^2 + \sigma_{\text{Statistics}}^2. \quad (7.1)$$

Of all those terms the dominant contribution comes from the lack of muons contributing to each MPD distribution. In the standard event-by-event MPD analysis only stations far from the shower axis are employed. This distance cut imposes also the energy threshold: events at energies larger than 20 EeV do have enough muons (20 on average in the case of proton QGSJET-II-04, see figure 7.4). The contributions of the geometry and the time uncertainty have been minimized by optimizing the cut on the distance to the shower axis and the signal threshold: effects are larger when we use stations closer to the shower core.

We want to understand how new detector configurations modify the resolution of an analysis based on the MPD method. The results are illustrated in figure 7.5. The $\sigma(X_{\max}^{\mu})$ has been calculated using the new electronics configuration. In the standard analysis, for the 40 MHz, only stations that are at a distance larger than 1700 m are entering the final MPD distributions. This cut is mainly due to the effect of the geometry resolution on transforming the time of the muons at the ground to atmospheric depth and time response of the detector. For the case of the 120 MHz electronics, the only changes relevant to the MPD analysis are the accuracy of the GPS time tagging and a faster response of the Besel filter. These allow to perform the analysis with stations starting from 1500 m from the air-shower axis which leads to a slight improvement of the resolution of X_{\max}^{μ} . It varies

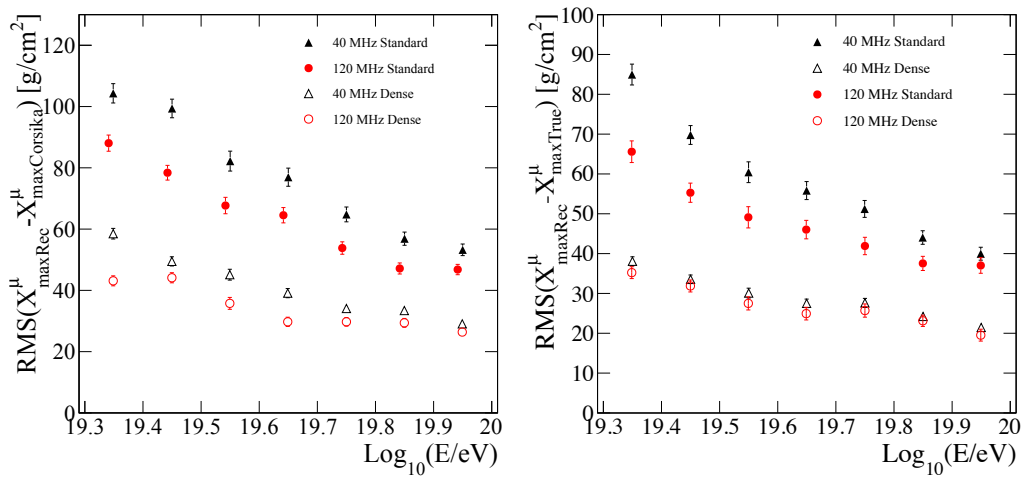


Figure 7.5: Resolution for the maximum of the muon production distance for proton (left) and iron (right) EPOS 1.99.

from about 60 g/cm² to 50 g/cm² at the highest energies. Our resolution is still far from the values of the fluorescence detector measurement of X_{\max} (around 20 g/cm²).

One way to improve the $\sigma(X_{\max}^{\mu})$ is to change, for individual muons, the properties of the time distribution of the photons that reach the PMTs. A black top liner does exactly this: reducing the time response of the detectors as has been previously explained. We have tested the reconstruction at an energy of 30 EeV using a black top liner configuration and 120 MHz electronics. The $\sigma(X_{\max}^{\mu})$ improves for this energy from 70 g/cm² to 50 g/cm². As this study has been performed during the decision taking process for the prototypes of the upgrade and the black top liner was not followed by the collaboration, we did not analyze the evolution of the resolution for other energies but we are persuaded that significant improvements can be obtained for the entire energy interval. Therefore blackening the top of the inner liner is an option that helps improving the expected resolution of MPD analyses.

Of course one of the best ways to increase the resolution is to collect more muons at the ground. This can be done by deploying a *dense array* like the proposal in [147]. For that array configuration, it features 750 m separation between the detectors, we have calculated the $\sigma(X_{\max}^{\mu})$. Having a factor of about three increase in the number of muons (see figure 7.4) the $\sigma(X_{\max}^{\mu})$ decreases to less than 45 g/cm² for the entire energy range. At higher energies the expected resolutions are similar to those obtained in the measurement of X_{\max} with the Fluorescence Detector (see figure 7.5).

Conclusions and future prospects

This thesis is a contribution to the ongoing questions related to the mass composition of the UHECR and how accurately hadronic interaction models describe data. This work builds on the information recorded by the SD of the Pierre Auger Observatory, that is why our starting point was a revision of the SD event reconstruction.

■ Improvements in the SD event reconstruction

- The impact point on the ground of the air-shower and its energy are determined through a fit of the lateral distribution of the signals. One of the key ingredients for a reliable estimation of those quantities is the uncertainty of the WCD signals.
- Using the *Infill* data we revised the signal variance model and updated it to include the new triggers implemented in the array, the ToTd and MOPs. The new model was tested in the interval of full trigger efficiency for the SD. A new component of 2.3% was added to the signal uncertainty that might be related to the uncertainty of the calibration. The new signal variance model does not change significantly the reconstructed variables and it has only a mild impact on the uncertainties.
- With the help of a toy Monte Carlo which includes the effects of the trigger on the signal variance distributions, we showed that the signal variance form is valid in the entire signal range. It can be extrapolated to the very low signals (i.e. down to 0.5 VEM) observed with the new triggers ToTd and MOPs.

The next part of the thesis discusses in detail how the recorded arrival times of muons provide a variable which is highly sensitive to the nature of UHECRs.

■ Update of the Muon Arrival Time Model: a description of a new kinematic delay parameterization

- We revised the muon arrival time distribution model with special focus on how it affects the bias of X_{\max}^{μ} and the shape of the MPD distribution. In particular, the kinematic delay approximation was carefully studied. We found that the classic approach overestimates this contribution, and therefore the reconstructed z are in general underestimated.

- To reduce the distortions coming from this effect, we explored new ways to derive the kinematic delays. The average kinematic delay obtained from simulations was tabulated as a function of the measured delay in the detectors and the distance to the shower core (for $r > 1000$ m and $\theta \sim 60^\circ$). As a result, the absolute value of X_{\max}^μ bias was reduced by a factor of 3, and the distortion in the shape of the MPD distribution was successfully removed.

■ Revision of the MPD analysis on an event-by-event basis: Decoupling the main contributions to X_{\max}^μ bias and RMS.

- This work was the first one able to fully understand the different contributions to the X_{\max}^μ bias. Among all the contributing sources, it was found that the response of the muon in the WCD is the main source of bias.
- We studied the impact the un-thinning or resampling procedure, performed in the official Auger simulation software, on the muon arrival times in the WCD. An underestimation of these times, increasing almost quadratically with the distance to the shower core, r , was found. It ranges from a few nanoseconds, for small values of r , up to several tens of nanoseconds for larger distances. We proved that for a fixed zenith angle, around 60° , this underestimation is universal (i.e. independent of the nature of the primary particle, the energy and the hadronic model used in the simulations), being an effect of the algorithm itself.
- The resampling procedure done in the Auger software induces a bias in the determination of X_{\max}^μ . The estimated final bias after the whole reconstruction chain is 24 g/cm^2 which has to be taken into account when analyzing the data. After correcting for this effect in the data, the tension with hadronic models is aggravated.
- The different contributions to X_{\max}^μ resolution were fully understood. The main contribution is due to the lack of muons contributing to each MPD distribution.

■ Average MPD analysis

- With the aim of enlarging the energy range of applicability of MPD studies and to alleviate the problem of lack of statistics to build MPD distributions, we proposed to study the behaviour with energy of average MPD distributions (i.e. we build single distributions with all the air showers belonging to a particular energy bin).
- As our interest was to study the shape of the MPD, we decoupled each contribution to the bias of the average X_{\max}^μ and identified all the distortion sources. The main contribution to the bias is the response of the WCD to muons, as in the event-by-event case. However, we found that the geometry reconstruction also modifies the shape. To minimize the effect of the reconstruction and its influence on the bias of the average X_{\max}^μ we restricted the analysis to distances larger than 1500 m.
- The distance cut imposes the energy threshold. Requesting a selection efficiency larger than 95% implies that the average X_{\max}^μ can be extended down to only 10 EeV which is a factor two better than the baseline event-by-event analysis.

-
- We fine-tune and apply the deconvolution to remove part of the MPD distortion induced by the WCD response. Nevertheless, to be able to study the shape of the distribution at very high level of detail an improvement of the geometry reconstruction in this zenith range region is mandatory. This will also help the event-by-event analysis.
 - We analyzed the data collected by the Pierre Auger Observatory between January 2004 and October 2014. After the selection cuts we are left with 2319 events above 10 EeV in a zenith angle range around 60° .
 - The analysis was performed in the same zenith angle range as in [1] but in the future it can be extended to other regions.
 - The main sources of systematic uncertainties yield an overall value of $\sim 20.5 \text{ g/cm}^2$ being dominated by the contribution of the hadronic models.
 - The $\ln A$ obtained from the reconstructed data is compatible with the result published in [1], confirming the tension between hadronic interaction models and the data which persists also between 10 and 22 EeV. Therefore, it is difficult to draw solid conclusions about the mass composition of UHECR.

■ New detector configurations and its influence on MPD analysis

- The MPD analysis was applied to different detector configurations which we implemented in the official Auger software. We have shown that an intrinsically faster time response of the WCD to particles (for example when using a black-top liner) might improve the resolution of X_{max}^μ . Increasing only the sampling frequency does not provide a significant improvement. However, we conclude that only a much denser array could provide a resolution compatible with the FD measurements thanks to the sizable increase in the number of detected muons at the ground.

Conclusiones y perspectivas futuras

Esta tesis es un intento de contribuir al debate acerca de la composición de los rayos cósmicos de ultra alta energía y cómo de precisos son los modelos hadrónicos a la hora de describir los datos. Este trabajo se ha hecho con la información registrada en el detector de superficie del Observatorio Pierre Auger por lo que el punto de partida ha sido una revisión de la reconstrucción de los sucesos registrados por dicho detector.

■ Mejoras en la reconstrucción de los sucesos del detector de superficie

- El punto de impacto en el suelo y la energía de las cascadas se determinan a partir del ajuste a una función de distribución lateral de las señales registradas. Una de las claves en el proceso de ajuste para estimar de forma precisa estas cantidades es la determinación de los errores de las señales de los detectores
- Usando los datos registrados por el detector *Infilled*, se ha revisado y actualizado el modelo que describe las varianzas de las señales para así incluir los nuevos triggers que se han implementado en el detector de superficie, ToTd y MOPs. El modelo se ha revisado en la región de máxima eficiencia de trigger para el detector de superficie. Una nueva componente que podría estar relacionada con incertidumbres en la calibración del 2.3 % se ha añadido al error de la señal. El nuevo modelo para describir los errores en la señal no cambia mucho las variables reconstruidas y tiene un leve impacto en sus incertidumbres.
- Con la ayuda de un *toy Monte Carlo* que contiene los efectos de trigger en las distribuciones de la varianza de las señales, hemos demostrado que el modelo funciona en todo el rango de señales. Además puede extrapolarse hasta señales muy bajas donde los nuevos triggers ToTd y MOPs son sensibles.

La siguiente parte de la tesis describe en detalle cómo los tiempos de llegada de los muones registrados en los detectores dan lugar a una variable muy sensible a la composición química de los rayos cósmicos de ultra alta energía.

■ Actualización del modelo de los tiempos de llegada de los muones: descripción de un nuevo retraso cinemático.

- El modelo se ha revisado haciendo hincapié en cómo afecta al sesgo de X_{\max}^{μ} y a la forma de la distribución de la MPD. En particular, se ha estudiado la aproxima-

ción que se usa para estimar el retraso cinemático. Hemos demostrado que la actual descripción sobrestimaba esta contribución subestimando los valores de las z reconstruidas.

- Una nueva forma de calcular el retraso cinemático se ha desarrollado para reducir las distorsiones que este provocaba en el perfil. Se ha tabulado el retraso cinemático promedio con simulaciones en función del retraso medido en los detectores y la distancia al eje de la cascada (para $r > 1000$ m, $\theta \sim 60^\circ$). Hemos reducido el valor absoluto del sesgo en un factor 3, y eliminado la distorsión en el perfil de la distribución de la MPD.

■ Revisión del análisis de la MPD suceso a suceso: Separación de las principales contribuciones al sesgo y a la resolución de X_{\max}^μ .

- Este trabajo ha sido el primero en desacoplar las diferentes contribuciones al sesgo de X_{\max}^μ . Entre todas ellas, hemos visto que la principal es debida a la respuesta de los detectores a los muones.
- Hemos estudiado el impacto del proceso de *un-thinning* o *resampling* que se lleva a cabo en la versión `Offline-icrc2013` en el tiempo de llegada de los muones a los detectores. Encontramos una subestimación de estos tiempos que aumentaba cuadráticamente con la distancia al eje de la cascada, variando de unos pocos nanosegundos cerca del eje a decenas de nanosegundos a distancias lejanas. Probamos que esta subestimación es universal (independiente de la naturaleza de la partícula primaria, la energía y el modelo hadrónico usado en las simulaciones) siendo un efecto del algoritmo en sí.
- El proceso de *resampling* realizado en `Offline-icrc2013` induce un sesgo en la determinación de X_{\max}^μ . El sesgo final después de toda la cadena de reconstrucción es de 24 g/cm^2 y tiene que tomarse en cuenta al analizar los datos. Tras tener en cuenta este efecto en los datos, se agrava la tensión con los modelos hadrónicos.
- Se han estudiado las diferentes contribuciones a la resolución de X_{\max}^μ , siendo el principal factor la falta de muones que contribuyen a cada distribución de la MPD.

■ Análisis de la distribución promedio de la MPD

- Con el propósito de extender el rango de aplicabilidad de los estudios de MPD y para aliviar la falta de estadística a la hora de construir cada distribución de MPD, hemos propuesto estudiar el comportamiento con la energía de la distribución promedio de la MPD (es decir, construimos la distribución resultante de todos los sucesos que pertenecen a un intervalo de energía, en particular).
- Como nuestro objetivo era estudiar el perfil de la distribución, desacoplamos cada una de las contribuciones al sesgo del máximo de la distribución promedio e identificamos todas las fuentes de distorsión. La principal contribución al sesgo es la respuesta de los detectores a los muones, al igual que en el análisis suceso a suceso. Sin embargo, encontramos que la reconstrucción de la geometría también puede

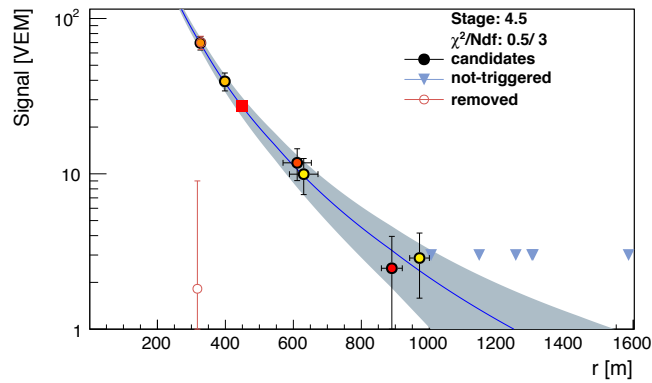
modificar el perfil. Para minimizar estos efectos de la reconstrucción y su influencia en el sesgo de X_{\max}^{μ} hacemos el análisis a distancias mayores de 1500 m.

- El corte en distancia impone un corte en energía. Requerimos que la eficiencia de selección sea mayor del 95 %, lo que implica seleccionar sucesos con energía mayor de 10 EeV.
- Aplicamos la deconvolución para eliminar la distorsión del perfil provocada por la señal pero para ser capaces de estudiar el perfil de la distribución es necesaria una mejora de la reconstrucción geométrica en esta región de ángulo cenital. Esto ayudaría también al análisis suceso a suceso.
- Hemos analizado los datos registrados por el Observatorio Pierre Auger entre enero de 2004 y octubre de 2014. Después de los cortes de selección tenemos 2319 sucesos por encima de 10 EeV.
- El análisis se ha desarrollado en la misma región de ángulo cenital que [1] pero en un futuro podría extenderse a otras regiones.
- Las principales fuentes de sistemáticos resultan en un valor total de $\sim 20,5 \text{ g/cm}^2$ dominado por las contribuciones de los modelos hadrónicos.
- El $\ln A$ obtenido de la reconstrucción de los datos es compatible con el resultado publicado en [1], confirmando así la tensión entre los modelos hadrónicos y los datos. Por tanto, es difícil extraer conclusiones sólidas acerca de la masa de los rayos cósmicos de ultra alta energía.

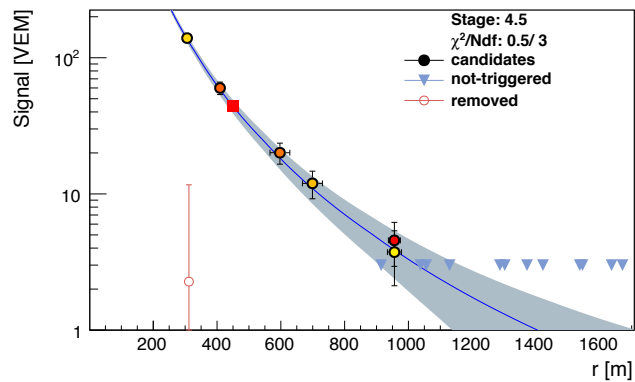
■ Nuevas configuraciones de detector y su influencia en el análisis de la MPD

- Se ha elaborado el análisis de la MPD para diferentes configuraciones de detector. En cuanto a la mejora en la resolución de X_{\max}^{μ} , una respuesta temporal más rápida de los detectores a las partículas, por ejemplo un liner con la superficie opaca, ayudaría a mejorarla. Aumentar sólo la frecuencia de muestreo no produce mejoras apreciables en la resolución. Sin embargo, tener muchos más detectores, un array denso, nos permitiría tener muchos más muones y así obtener una resolución compatible con la de las medidas del FD.

Example of outliers



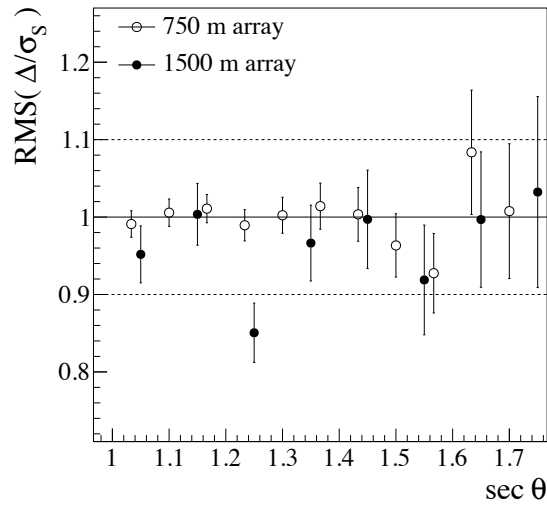
(a)



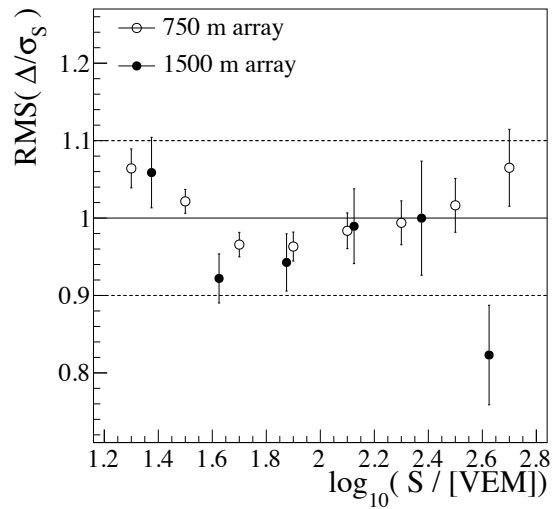
(b)

Figure A.1: Lateral distribution function for events where the difference between signals is larger than 1.5. (a) Event 6466282. The signal candidate station of the doublet, station 734, observed 69.6 VEM, while the partner, station 93, has a signal of 1.8 VEM. (b) Event 14995513. The candidate station, 669, has a signal of 139.4 VEM, while the partner, station 80, measured 2.3 VEM.

A.1 1500 m Array and different angular ranges



(a)



(b)

Figure A.2: The RMS of the distribution Δ divided by the signal uncertainty for the 750 m array (open circles) and for the 1500 m array (solid circles) as a function of secant θ (a) and of signal (b).

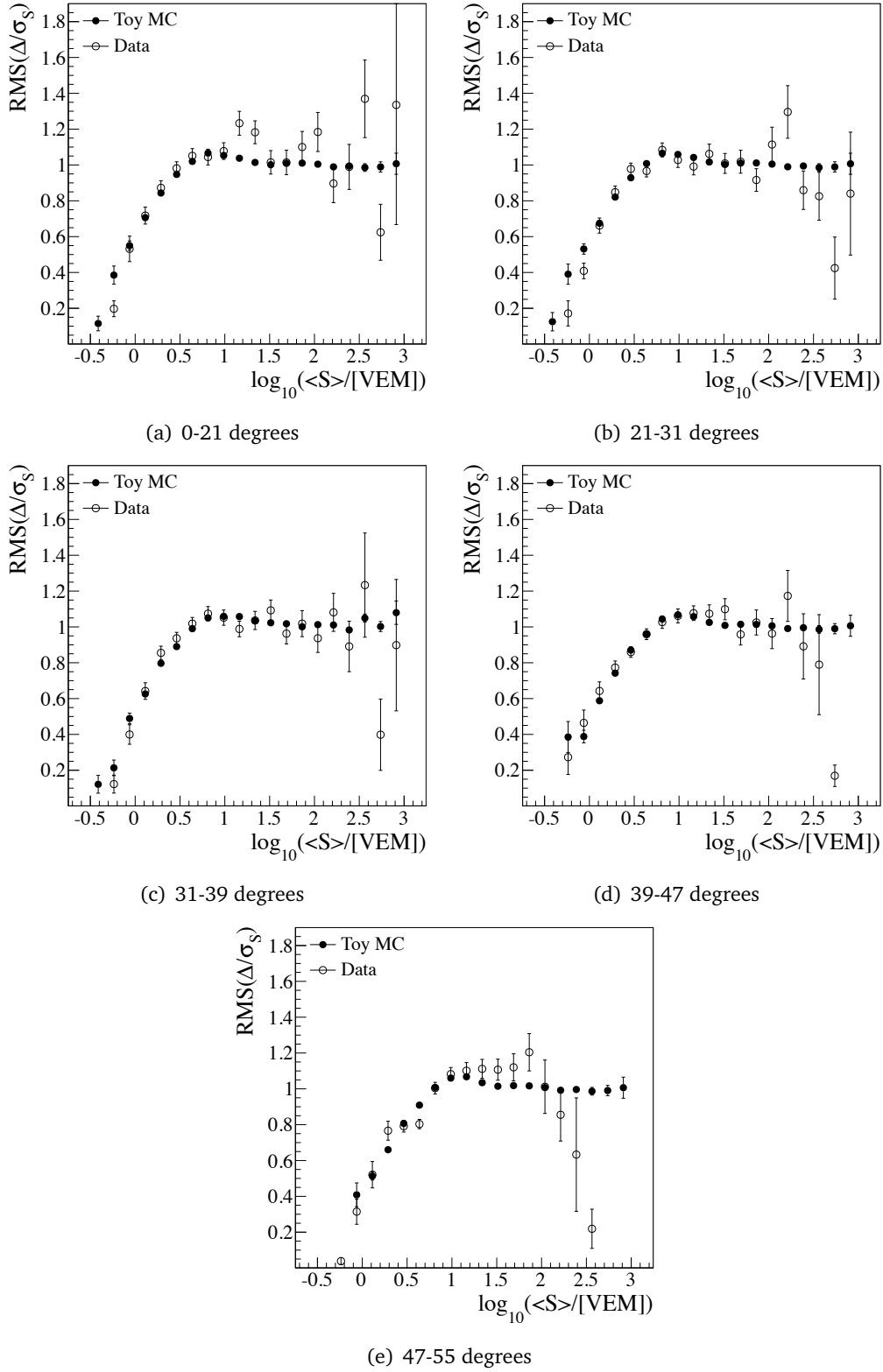


Figure A.3: Same as figure 3.6(a) for different ranges of zenith angle for all the triggers. As we require that a station is not saturated in the anode channel, at the highest signals we are selecting just the lower fluctuations and thus the $RMS(\Delta)$ is underestimated.

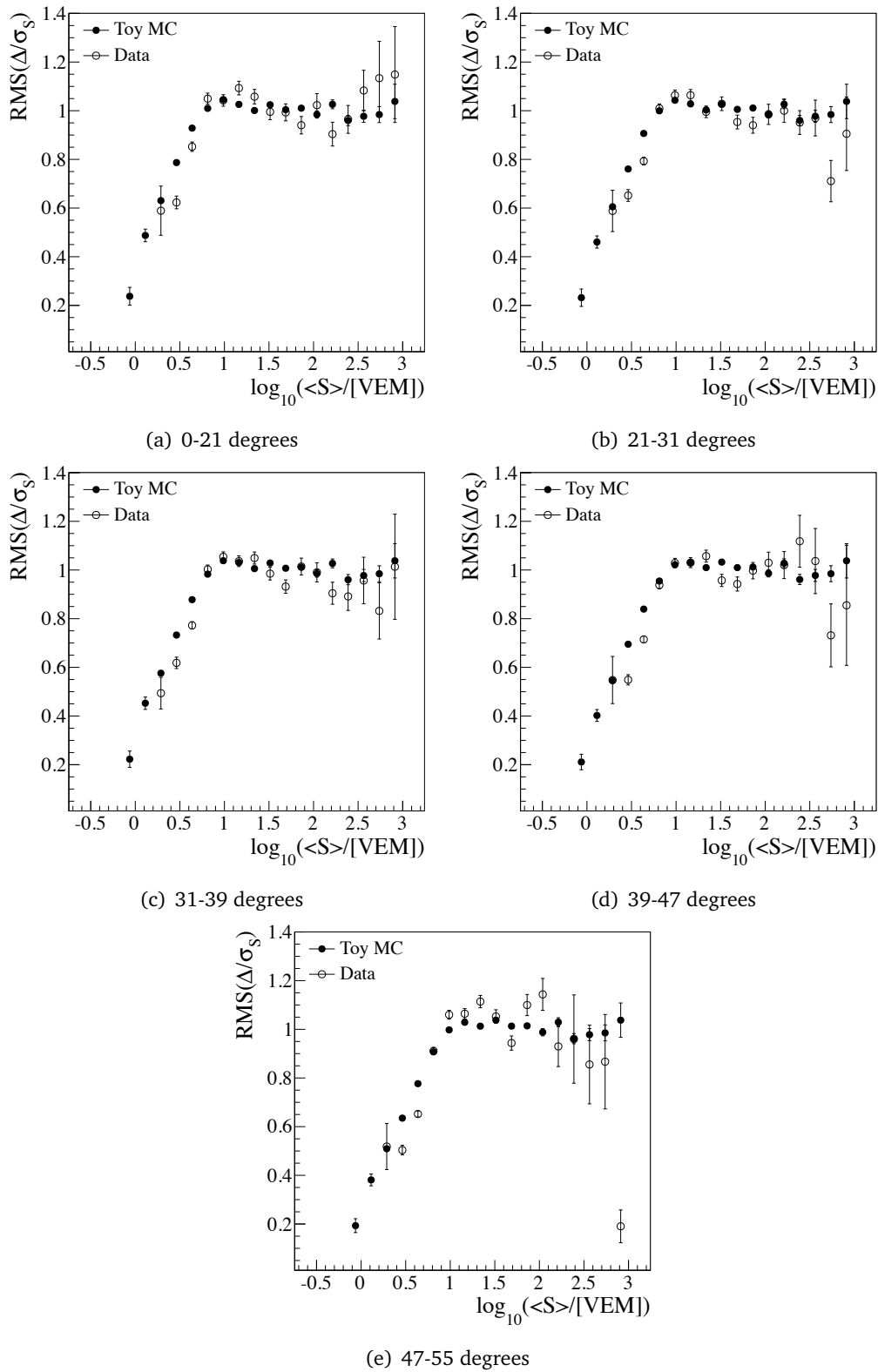


Figure A.4: Same as Fig. A.3 for old triggers.

List of Figures

1.1	Spectrum of cosmic rays measured by different experiments [18].	2
1.2	Illustration of the survival of nuclei $> 6 \times 10^{19}$ eV as a function of distance. Effectively only proton and iron nuclei can reach Earth from sources more than ~ 50 Mpc away [12].	4
1.3	The Hillas plot represents the strength of the magnetic field versus the size of possible candidates for ultra-high energy cosmic ray acceleration. Objects below the diagonal can be ruled out.	5
1.4	Limits on the flux of photons (left) [38, 39, 40, 41] and neutrinos (right) [52] obtained from the Pierre Auger Observatory. The data are shown together with the current limits from other experiments [53, 54, 55, 56, 57, 58, 59] and some examples of predicted fluxes.	6
1.5	Large scale anisotropy search. Left: 99% limits on the dipole anisotropy in the equatorial plane for the collected statistics until end of 2014 (dashed line) and values of the dipole amplitude d_{\perp} . Right: estimated phase angles ϕ_1 . The red points of the equatorial phase are from the analysis of the infill array. The data shown [62] is an update of the analyses [63, 64].	7
1.6	Regions of over-density observed after $\sim 20^{\circ}$ -smearing of the arrival directions of particles with $E > 5,5 \times 10^{19}$ eV. The results from the northern hemisphere are from the TA Collaboration [67].	7
1.7	Main components of extensive air showers.	8
1.8	Schematic representation of the Heitler model for electromagnetic (a) and hadronic (b) cascade development.	9
1.9	Mean of the measured X_{\max} distributions as a function of energy compared to air-shower simulations for proton and iron primaries for different experiments.[85].	12
2.1	Map of the Pierre Auger Observatory. Red dots indicate SD detectors and green lines show the field of view of the FD telescopes.	15
2.2	View and scheme of an SD station of the Pierre Auger Observatory.	17

2.3	Charge and pulse height histograms for an SD station with a 3-fold trigger (signal in all 3 PMTs). The signal is the sum of the three PMTs. In the solid histogram the second peak is produced by vertical through-going atmospheric muons, while the first peak is a trigger effect. The dashed histogram is produced by vertical and central muons (VEMs) selected with an external muon telescope.	18
2.4	T4 and T5 configurations. 2.4(a): The three minimal compact configurations for the T4-4C1 trigger. 2.4(b): The two minimal compact configurations for the T4-3ToT configuration. 2.4(c): Example of the 6T5 hexagon (shadow) and the 5T5 hexagon (dark shadow).	20
2.5	View and schemes of an FD site of the Pierre Auger Observatory.	21
2.6	Basic patterns of triggered pixels considered by the Second Level Trigger in the FD.	22
2.7	The LIDAR and cloud cameras are part of the atmospheric monitoring used in the Pierre Auger Observatory.	24
2.8	Shower-detector plane parameters.	25
2.9	Geometry reconstruction of an event observed by four telescopes and the Surface Detector [86].	26
2.10	Angular resolution (left) and core location accuracy (right) for monocular and hybrid reconstruction.	26
2.11	Angular resolution of hybrid events [101].	27
2.12	Energy reconstruction in the FD. The calorimetric energy is obtained by the integration of the fit to the profile.	28
3.1	Schematic view of a plane shower front.	32
3.2	Schematic view of a spherical shower front.	33
3.3	(a) Measured signal spectrum (b) Trigger probability as a function of the mean expected signal.	34
3.4	(a) Mean of Δ as a function of the estimated signal. (b) RMS of the distribution Δ divided by the signal uncertainty for the new (solid circles) and the previous parameterization (open circles).	35
3.5	The RMS of the distribution Δ divided by the signal uncertainty for the new (solid circles) and the previous parameterization (open circles) as a function of $\sec\theta$, (a), and of distance to axis, (b).	36
3.6	(a) RMS(Δ) versus mean signal for old triggers (open circles), TOT and T2, and including the new ones, TOTd and MoPS. (b) RMS(Δ) versus mean signal for the toy MC (solid circles) and for data (open circles), in the region where the trigger efficiency is larger than 98%.	38
3.7	(a,c): The RMS(Δ) versus mean signal compared to the toy MC. (b,d): the RMS of Δ divided by the parameterization. A good agreement is observed between data and the toy MC.	39
3.8	Angular resolution as a function of the zenith angle θ for events with an energy above 3 EeV, and for various station multiplicities [86].	41
3.9	Angular resolution as a function of the primary energy for different angle intervals.	41

3.10	Attenuation curve described by a third degree polynomial function in $x = (\cos^2(\theta) - \cos^2(38))$. In this example the polynomial coefficients are deduced from $S(1000)$ dependence at $S_{38} \sim 50$ VEM which corresponds to an energy of about 10.5 EeV [86].	44
3.11	<i>Top</i> : Energy calibration in the Pierre Auger Observatory. Correlation between S_{38} , S_{35} and N_{19} with the energy measured by the FD [32]. <i>Bottom</i> : Integrated exposures for the different data sets [106].	46
3.12	Systematic uncertainties on the energy scale [105].	47
4.1	Geometry used to obtain the muon traveled distance.	50
4.2	Average time delay of muons with a breakdown of the different contributions. The curves correspond to a proton-initiated shower with a zenith angle of 60° and primary energy of $E = 10$ EeV [130].	52
4.3	<i>Left</i> : Delay vs energy at the ground for distances to the shower core larger than 1000 m. <i>Right</i> : Ground energy distribution at 1000 m (empty histogram) and at 2000 m (filled histogram).	52
4.4	MPD distributions obtained for proton QGSJET-II-04 at $10^{19.75}$ eV. Continuous line represents the distribution obtained directly from the simulations (the true distribution). The dashed line is the one obtained using only the geometric delay, and the filled histogram is the result after correcting by the kinematic delay, using equation (4.10).	53
4.5	ϵ as a function of r and as a function of $z - \Delta$ for proton QGSJET-II-04 at $10^{19.75}$ eV and around 60°	54
4.6	Effect of the kinematic delay parameterization over the average MPD distribution in the standard analysis, at distances larger than 1700 m (right plot), and from 1000 m (left plot). Black line is the distribution obtained using the true kinematic delay and the dot-dashed line is the one obtained using the parameterization.	55
4.7	(Similar to figure 4.6) Kinematic delay parameterization effect over the average MPD distribution for different production distances ranges.	56
4.8	(Similar to figure 4.6) Kinematic delay parameterization effect over the average MPD distribution effect for different ground energy ranges.	57
4.9	Average kinematic delay as a function to the measured delay and the distance to the shower core.	58
4.10	Effect of the new approximation used to compute the kinematic delay for two distance ranges. Black line is the distribution obtained using the true kinematic delay from the simulations and the dot-dashed line is the one obtained using the new approximation.	59
4.11	(Similar to figure 4.10) Effect of the new kinematic delay approximation for different production distances ranges. Curves correspond to proton showers with an energy of $10^{19.75}$ eV.	60
4.12	(Similar to figure 4.10) Effect of the new kinematic delay approximation for different ground energy ranges. Curves correspond to proton showers with an energy of $10^{19.75}$ eV.	61

4.13	Difference between the X_{\max}^{μ} obtained from the reference distribution, $X_{\max}^{\mu True}$, and the one using each one of the kinematic delay parameterizations (the top plot corresponds to the old one and the bottom to the new one) for all species.	63
5.1	MPD distributions produced by an iron shower with energy $10^{19.5}$ eV impinging with two different zenith angles: 41° (left) and 60° (right). We use EPOS LHC[79] to model high-energy hadronic interactions. The shape of the MPD distribution shows a dependence with the distance to the shower core. This dependence is more pronounced for events with smaller inclinations. The histograms are normalized to have the same maxima [1].	66
5.2	X_{\max}^{μ} distribution for proton and iron showers simulated at 30 EeV with EPOS LHC at zenith angles between 55° and 65° . The mean value and the RMS of the distributions show a clear dependence on the mass of the primary cosmic ray. For the construction of the MPDs, only muons reaching the ground at distances greater than 1700 m were considered [1].	67
5.3	Evolution of X_{\max}^{μ} as a function of the energy for proton and iron and two hadronic models.	68
5.4	Average number of muons contributing to the MPD for proton and iron QGSJET-II-04 showers between 1700 and 4000 m.	69
5.5	Standard Analysis bias and RMS. In this analysis MPDs are reconstructed individually for each event.	71
5.6	Standard analysis bias.	72
5.7	Contribution to the bias of X_{\max}^{μ} reconstructed from the approximations done in the muon arrival time model.	73
	level (right plot).73figure.5.8	
5.9	Number of muons as a function of the energy in the standard array (filled dots) and in a dense array (empty dots) for protons simulated using QGSJET04.	74
5.10	Differences in the average MPD distribution for 25 proton QGSJET-II-04 showers at $10^{19.55}$ at 60° generated with a thinning level of 10^{-5} (lightest line), 10^{-6} (dark line) and 10^{-7} (lighter line).	75
5.11	Definition of the sampling region (shadow) at a given position (marked with a point) [134].	75
5.12	<i>Left:</i> Typical arrival time structure of muons in an EAS. <i>Right:</i> Distribution of the delay of the muons integrated in a sampling region centered at $r_s = 1450$ m using a value of $\delta = 1\%$ (dashed) and $\delta = 10\%$ (empty). The underestimation in the later is visible. Both figures have been obtained using QGSJET-II-04 with proton primaries at $10^{19.5}$ eV and $\theta = 60^{\circ}$	76
	by the resampling procedure. The fit to a quadratic function is also shown just to guide the eyes. This figure has been obtained using proton showers simulated with CORSIKA with QGSJET-II-04, $E = 10^{19.5}$ eV and $\theta = 60^{\circ}$.77figure.5.13	

- 5.14 *Left*: $\Delta\langle\tau\rangle$ vs r for three different values of δ : 5% (triangles), 10% (circles) and 20% (squares) respectively. *Right*: Bias in the arrival time of the muons with (empty symbols) and without (full circles) applying the corrections suggested in [134]. The improved estimation of the arrival times when we use a correction is remarkable. 78
- 5.15 Comparison of the values of the muon delays for proton and iron primaries (*Left*), QGSJET-II-04 and EPOS LHC hadronic models (*Center*) and different energies (*Right*). A negligible bias is found for all the considered cases. . . 79
- 5.16 Dependency of the bias in τ with the azimuthal position. In the bottom panel we compare the results when we integrate over a full crown ($\alpha = \pi$) with the results at four sampling regions: $\zeta = 0$ (empty circles), $\pi/4$ (empty inverse-triangles), $\pi/2$ (empty squares) and $3\pi/4$ (empty triangles) for a value of $\alpha = 20^\circ$ and $\delta = 10\%$ 79
- 5.17 level after resampling correction.80figure.5.17
- 5.18 Correlation between the uncertainty in core and zenith determination. . . . 80
- 5.19 Contribution to the bias from the geometry reconstruction. 81
- 5.20 Bias vs θ comparing true and reconstructed geometry. 81
- 5.21 Signal contribution to X_{\max}^μ bias. 82
- 5.22 Summary of all the bias contributions. Filled dots correspond to the method accuracy. The filled triangles include the kinematic delay parameterization plus the method. The empty triangles contain the method, the kinematic delay parameterization and the geometry reconstruction. Finally, the squares represent the bias after the whole chain is applied (i.e. including the method, the kinematic delay parameterization, the geometry reconstruction and the effect of the signal spread). 82
- 5.23 Signal contribution to X_{\max}^μ bias where true geometry is considered. 83
- 5.24 Contribution to the bias due to the geometry reconstruction. 83
- 5.25 Bias in the standard analysis conditions after all the reconstruction chain for proton and iron nuclei for both hadronic models. 84
- 5.26 Standard analysis RMS. 85
- 5.27 Difference between X_{\max}^μ reconstructed and the true one after all the reconstruction chain for proton QGSJET-II-04 at $10^{19.55}$ eV. 85
- 5.28 Bias in the standard analysis conditions after all the reconstruction chain for proton and iron nuclei QGSJET-II-04 (top) and EPOS LHC (bottom) after applying the resampling correction and before. The bottom pad shows the residual of both biases for both primaries and models. 88
- 5.29 $\langle X_{\max}^\mu \rangle$ as a function of energy before (upper plot [1]) and after (bottom plot) the resampling correction. The prediction of different hadronic models for proton and iron are shown. Numbers indicate the statistics in each energy bin and brackets represent the systematic uncertainty. 89
- 5.30 Conversion of $\langle X_{\max}^\mu \rangle$ (circles) and $\langle X_{\max} \rangle$ [137] to $\langle \ln A \rangle$, as a function of energy before (top) and after (bottom [1]) bias correction. On the left (right) plot we use QGSJET-II-04 (EPOS LHC) as the reference hadronic model. Brackets correspond to the systematic uncertainties. 90

6.1	Sum of all the distributions for iron QGSJET-II-04 at $10^{19.55}$ eV obtained from CORSIKA. The maximum obtained from a GH fit (solid line) will be our SD observable, $\langle X_{\max}^{\mu} \rangle$	92
6.2	Average MPD distribution for proton QGSJET-II-04 at $10^{19.55}$ eV. Black line is the distribution directly obtained from CORSIKA, the blue filled histogram is the result correcting the delay with the true kinematic delay, and the red filled one is the result after applying the new parameterization described in chapter 4.	93
6.3	Bias due to the geometry reconstruction as a function of the distance to the shower core.	94
6.4	<i>Left:</i> Average distribution using the model, the new kinematic delay parameterization, the true time and the true geometry. Dots represent the reconstructed distribution and the line is the histogram directly obtained from CORSIKA. <i>Right:</i> . The same but using the reconstructed geometry instead of the true one.	95
6.5	<i>Left</i> Cumulative bias contributions to the average X_{\max}^{μ} bias. Empty dots correspond to the method accuracy plus the new kinematic delay parameterization. The filled triangles include the method, the kinematic delay parameterization and the geometry reconstruction. Finally, the squares represent the bias after all effects are included. <i>Right</i> Average distribution at $10^{19.55}$ eV after all the reconstruction chain.	96
6.6	Histogram of the response to 10000 muons in the WCD. The line is the fit to a function $A * (e^{-t/65} - e^{-t/4})$	97
6.7	<i>Left:</i> Muon (blue line) and Deconvoluted muonic trace (filled red histogram) in one PMT of a station at 1510 m to the shower core and with a signal of 7 VEM. <i>Right:</i> The same PMT but including the total trace (black line) and applying the deconvolution to the total trace.	97
6.8	<i>Left:</i> Difference between the signal obtained from the integration of the peaks and the muonic signal for different threshold cuts. <i>Right:</i> Difference between proton an iron biases for different threshold cuts.	98
6.9	Bias in the average X_{\max}^{μ} coming from the method plus the new kinematic delay parameterization and the signal (filled squares) compared to the one obtained in the same conditions but after applying the deconvolution and the threshold cut.	98
6.10	Histogram of the response of muons in the WCD after deconvoluting the trace (empty histogram) and after cutting the deconvoluted trace at 5% of the peak (filled histogram).	99
6.11	<i>Left:</i> Average distribution using the model, the new kinematic delay parameterization, the signal and the true geometry. Dots represent the reconstructed distribution and the line is the histogram directly obtained from CORSIKA. <i>Right:</i> . The same but applying the deconvolution and the cut at 5% and the time shift.	100
6.12	Efficiency for proton and iron QGSJET-II-04 and EPOS LHC and for data as a function of the energy.	101

6.13	<i>Left</i> Bias after corrections for all primaries and models. <i>Right</i> Average bias for all the cases. The line represents the fit to a linear function.	101
6.14	Average distribution after all the reconstruction chain.	102
6.15	<i>Top</i> : Final bias for all the species after correcting the bias with equation 6.2. <i>Bottom</i> : The reconstructed $\langle X_{\max}^{\mu} \rangle$ after applying the bias correction for both primaries and models obtained from simulations compared to CORSIKA predictions.	103
6.16	Application to experimental data: Average distributions.	105
6.17	Application to experimental data: Average distributions.	106
6.18	Evolution of the average X_{\max}^{μ} as a function of the primary energy for the data of the Pierre Auger Observatory between 1 January of 2004 and 31 October 2014. The number of events in each energy bin is indicated.	107
6.19	Seasonal effect.	107
6.20	Energy scale systematic uncertainty.	109
6.21	Aging effect.	110
6.22	Average X_{\max}^{μ} as a function of the primary energy for all the data set (filled dots) and excluding 2004 (empty dots).	110
6.23	Energy dependence of the maximum obtained from average MPD distributions. Data is compared to expectations from simulations.	111
6.24	Average distributions for data (dots) and QGSJET-II-04 predictions (the continuous line corresponds to iron and the dotted line to proton).	112
6.25	Average distributions for data (dots) and QGSJET-II-04 predictions (the continuous line corresponds to iron and the dotted line to proton).	113
6.26	Conversion of $\langle X_{\max}^{\mu} \rangle$ (filled circles), X_{\max}^{μ} [1] and $\langle X_{\max} \rangle$ [136] to $\langle \ln A \rangle$, as a function of energy. On the left (right) plot we use QGSJET-II-04 (EPOS LHC) as the reference hadronic model. Brackets correspond to the systematic uncertainties.	114
7.1	The distribution of photoelectrons in time for a black top liner (left) and for the standard liner (right).	117
7.2	Example of trigger studies at 100 MHz for a black top liner (left) and for the standard liner (right). The occupancy represents the number of bins above threshold requested by the trigger. The contour lines are ordered as follows from left to right: 10 for the smallest occupancy up to 0.1 Hz for the largest.	118
7.3	Station level variables. Total signal differences between stations versus total signal using 120 MHz and 40 MHz with 12 bits and white top.	119
7.4	(left) Number of muons entering in the MPD distributions as a function of energy for proton and iron QGSJET-II-04 and the standard array. (right) The same for proton QGSJET-II-04 for the standard array (filled dots) and the dense one (empty dots).	121
7.5	Resolution for the maximum of the muon production distance for proton (left) and iron (right) EPOS 1.99.	122

-
- A.1 Lateral distribution function for events where the difference between signals is larger than 1.5. (a) Event 6466282. The signal candidate station of the doublet, station 734, observed 69.6 VEM, while the partner, station 93, has a signal of 1.8 VEM. (b) Event 14995513. The candidate station, 669, has a signal of 139.4 VEM, while the partner, station 80, measured 2.3 VEM. 131
- A.2 The RMS of the distribution Δ divided by the signal uncertainty for the 750 m array (open circles) and for the 1500 m array (solid circles) as a function of secant θ (a) and of signal(b). 132
- A.3 Same as figure 3.6(a) for different ranges of zenith angle for all the triggers. As we require that a stations is not saturated in the anode channel, at the highest signals we are selecting just the lower fluctuations and thus the $RMS(\Delta)$ is underestimated. 133
- A.4 Same as Fig. A.3 for old triggers. 134

List of Tables

3.1	Summary of the experimental parameters regarding energy measurement in the Pierre Auger Observatory [32].	45
4.1	Contributions to the systematic uncertainty in the new estimation of τ_{kin} for the reconstruction of X_{max}^{μ}	62
5.1	Maximum relative uncertainties allowed in the estimation of X_{max}^{μ} . The value chosen for ϵ_{max} ensures no selection bias between the different primary species.	70
5.2	Efficiencies for proton and iron QGSJET04 showers after all the reconstruction cuts.	70
5.3	Efficiencies for proton and iron EPOS LHC showers after all the reconstruction cuts.	71
5.4	Cumulative contributions to X_{max}^{μ} bias.	82
5.5	Bias due to geometry reconstruction depending on the signal threshold applied.	84
5.6	Summary of all the identified RMS contributions for proton QGSJET-II-04 at $\text{Log}(E/\text{eV}) = 19.55$	87
6.1	Cumulative contributions to the average X_{max}^{μ} bias between 1500 and 4000 m.	96
6.2	Cuts applied to data and its relative selection efficiency.	104
6.3	Contributions to the systematic uncertainty in the reconstruction of the average X_{max}^{μ}	108
7.1	Calibration constants for current liner (St.) and for a black top liner (BT) for two sampling frequencies. The last line represents the number of bins above threshold requested by the trigger.	117
7.2	Reconstruction differences between 40 and 120 MHz for the standard WCD.	119
7.3	Reconstruction differences between standard and black top WCD at 120 MHz.	120
7.4	Resolutions for 120 MHz white and black top (BT). The candidates represent the mean number of stations used in the reconstruction, and the rest represent the differences to the generated values.	120

Bibliography

- [1] Pierre Auger Collaboration, A. Aab *et al.*, *Muons in air showers at the Pierre Auger Observatory: Measurement of atmospheric production depth*, Phys. Rev. **D90**, 012012 (2014), arXiv:1407.5919.
- [2] L. Molina Bueno, P. Billoir, and I. C. Mariş, GAP Report No. 035, 2014 (unpublished).
- [3] M. Settimo, P. Billoir, I. Maris, and L. Molina Bueno, GAP Report No. 014, 2013 (unpublished).
- [4] A. Bueno, D. Garcia-Gamez, and L. Molina Bueno, GAP Report No. 020, 2015 (unpublished).
- [5] D. Garcia-Gamez and L. Molina Bueno, GAP Report No. 006, 2015 (unpublished).
- [6] A. Bueno, D. Garcia-Gamez, and L. Molina Bueno, GAP Report No. 034, 2015 (unpublished).
- [7] L. Molina Bueno, P. Billoir, A. Bueno, D. Garcia-Gamez, and I. Maris, GAP Report No. 066, 2015 (unpublished).
- [8] V. F. Hess, *Über Beobachtungen der durchdringenden Strahlung bei sieben Freiballonfahrten*, Phys. Z. **13**, 1084 (1912).
- [9] T. Stanev, *High Energy Cosmic Rays* (Springer Berlin Heidelberg, 2010).
- [10] A. Letessier-Selvon and T. Stanev, *Ultrahigh energy cosmic rays*, Rev. Mod. Phys. **83**, 907 (2011).
- [11] K. H. Kampert and M. Unger, *Measurements of the cosmic ray composition with air shower experiments*, Astropart. Phys. **35**, 660 (2012).
- [12] A. Watson, *High-Energy Cosmic Rays and the Greisen-Zatsepin-Kuzmin Effect*, (2013), arXiv:1310.0325.
- [13] K.-H. Kampert and A. Watson, *Extensive Air Showers and Ultra High-Energy Cosmic Rays: A Historical Review*, (2012), arXiv:1207.4827.
- [14] H. Becquerel, *Sur les radiations émises par phosphorescence*, Comptes rendus de l'Acad. des Sciences **420** (1896).

- [15] R. A. Millikan and G. H. Cameron, *High Frequency Rays of Cosmic Origin III. Measurements in Snow-Fed Lakes at High Altitudes*, Phys. Rev. **28**, 851 (1926).
- [16] J. Linsley, *Evidence for a Primary Cosmic-Ray Particle with Energy 10^{20} eV*, Phys. Rev. Lett. **10**, 146 (1963).
- [17] E. Fermi, *On the origin of the Cosmic Radiation*, Phys. Rev. **75**, 1169 (1949).
- [18] Courtesy of Ioana Mariş.
- [19] J. R. Hörandel, *Models of the knee in the energy spectrum of cosmic rays*, Astropart. Phys. **21**, 241 (2004).
- [20] A. A. Ivanov, S. P. Knurenko, and I. Y. Sleptsov, *Measuring extensive air showers with Cherenkov light detectors of the Yakutsk array: the energy spectrum of cosmic rays*, New J. Phys. **11**, 065008 (2009).
- [21] M. Nagano *et al.*, *Energy spectrum of primary cosmic rays above $10^{17.0}$ eV determined from extensive air shower experiments at Akeno*, J. Phys. G Nucl. Partic. **18**, 423 (1992).
- [22] T. Antoni *et al.*, *KASCADE measurements of energy spectra for elemental groups of cosmic rays: Results and open problems*, Astropart. Phys. **24**, 1 (2005).
- [23] The KASCADE-Grande Collaboration, W. D. Apel *et al.*, *KASCADE-Grande measurements of energy spectra for elemental groups of cosmic rays*, Astropart. Phys. **47**, 54 (2013), arXiv:1306.6283.
- [24] A. D. Erlykin and A. W. Wolfendale, *Properties of cosmic ray interactions at PeV energies*, Astropart. Phys. **18**, 151 (2002), arXiv:astro-ph/0112553.
- [25] D. J. Bird *et al.*, *Evidence for correlated changes in the spectrum and composition of cosmic rays at extremely high energies*, Phys. Rev. Lett. **71**, 3401 (1993).
- [26] A. M. Hillas, *Cosmic Rays: Recent Progress and some Current Questions*, (2006), arXiv:astro-ph/0607109.
- [27] D. De Marco and T. Stanev, *On the shape of the ultrahigh energy cosmic ray spectrum*, Phys. Rev. D **72**, 081301 (2005).
- [28] D. Allard, E. Parizot, and A. Olinto, *On the transition from galactic to extragalactic cosmic-rays: Spectral and composition features from two opposite scenarios*, Astropart. Phys. **27**, 61 (2007).
- [29] V. Berezhinsky, *Transition from galactic to extragalactic cosmic rays*, (2007), arXiv:0710.2750.
- [30] V. Berezhinsky, S. Grigorieva, and B. Hnatyk, *Extragalactic UHE proton spectrum and prediction for iron-nuclei flux at $10^8 - 10^9$ GeV*, Astropart. Phys. **21**, 617 (2004).

- [31] V. Berezhinsky, *Extragalactic UHE proton spectrum and prediction of flux of iron-nuclei at 108–109 GeV*, Nucl. Phys. B Proc. Sup. **151**, 497 (2006).
- [32] I. Valiño for the Pierre Auger Collaboration, *The flux of ultra-high energy cosmic rays after ten years of operation of the Pierre Auger Observatory*, Proceedings of the 34th International Cosmic Ray Conference, The Hague (2015).
- [33] K. Greisen, *End to the cosmic ray spectrum?*, Phys. Rev. Lett. **16**, 748 (1966).
- [34] G. Zatsepin and V. Kuzmin, *Upper limit of the spectrum of cosmic rays*, JETP Lett. **4**, 78 (1966).
- [35] A. A. Penzias and R. W. Wilson, *A Measurement of excess antenna temperature at 4080-Mc/s*, Astrophys. J. **142**, 419 (1965).
- [36] R. Aloisio, V. Berezhinsky, and A. Gazizov, *Ultra high energy cosmic rays: The disappointing model*, Astropart. Phys. **34**, 620 (2011).
- [37] A. Hillas, *The origin of ultra-high-energy cosmic rays*, Annual review of astronomy and astrophysics. **22**, 425 (1984).
- [38] Pierre Auger Collaboration, J. Abraham *et al.*, *An upper limit to the photon fraction in cosmic rays above 10^{19} eV from the Pierre Auger observatory*, Astropart. Phys. **27**, 155 (2007), astro-ph/0606619.
- [39] Pierre Auger Collaboration, J. Abraham *et al.*, *Upper limit on the cosmic-ray photon flux above 10^{19} eV using the surface detector of the Pierre Auger Observatory*, Astropart. Phys. **29**, 243 (2008), arXiv:0712.1147.
- [40] Pierre Auger Collaboration, J. Abraham *et al.*, *Upper limit on the cosmic-ray photon fraction at EeV energies from the Pierre Auger Observatory*, Astropart. Phys. **31**, 399 (2009), arXiv:0903.1127.
- [41] Pierre Auger Collaboration, M. Settimo *et al.*, *An update on a search for ultra-high energy photons using the Pierre Auger Observatory*, Proc of 32nd Int. Cosmic Ray Conf., Beijing, China **2**, 55 (2011), arXiv:1107.4805.
- [42] Pierre Auger Collaboration, J. Abraham *et al.*, *Upper limit on the diffuse flux of UHE tau neutrinos from the Pierre Auger Observatory*, Phys. Rev. Lett. **100**, 211101 (2008), arXiv:0712.1909.
- [43] Pierre Auger Collaboration, P. Abreu *et al.*, *A Search for Ultra-High Energy Neutrinos in Highly Inclined Events at the Pierre Auger Observatory*, Phys. Rev. **D84**, 122005 (2011), arXiv:1202.1493.
- [44] Pierre Auger Collaboration, P. Abreu *et al.*, *Search for point-like sources of ultra-high energy neutrinos at the Pierre Auger Observatory and improved limit on the diffuse flux of tau neutrinos*, Astrophys. J. **755**, L4 (2012), arXiv:1210.3143.
- [45] Pierre Auger Collaboration, P. Abreu *et al.*, *Ultra-high Energy Neutrinos at the Pierre Auger Observatory*, Adv. High Energy Phys. **2013**, 708680 (2013), arXiv:1304.1630.

- [46] Pierre Auger Collaboration, P. Abreu *et al.*, *A Search for Ultra-High Energy Neutrinos in Highly Inclined Events at the Pierre Auger Observatory*, Phys. Rev. D **84**, 122005 (2011), arXiv:1202.1493.
- [47] ANITA Collaboration, P. W. Gorham *et al.*, *New Limits on the Ultra-high Energy Cosmic Neutrino Flux from the ANITA Experiment*, Phys. Rev. Lett. **103**, 051103 (2009), arXiv:0812.2715.
- [48] IceCube Collaboration, A. Achterberg *et al.*, *First year performance of the IceCube neutrino telescope*, Astropart. Phys. **26**, 155 (2006), arXiv:astro-ph/0604450.
- [49] G. Gelmini, O. Kalashev, and D. V. Semikoz, *GZK Photons as Ultra High Energy Cosmic Rays*, J. Exp. Theor. Phys. **106**, 1061 (2008), arXiv:astro-ph/0506128.
- [50] M. Ahlers, L. Anchordoqui, M. Gonzalez-Garcia, F. Halzen, and S. Sarkar, *GZK Neutrinos after the Fermi-LAT Diffuse Photon Flux Measurement*, Astropart. Phys. **34**, 106 (2010), arXiv:1005.2620.
- [51] K. Kotera, D. Allard, and A. Olinto, *Cosmogenic Neutrinos: parameter space and detectability from PeV to ZeV*, JCAP **1010**, 013 (2010), arXiv:1009.1382.
- [52] Pierre Auger Collaboration, P. Pieroni, *Ultra-high energy neutrinos at the Pierre Auger Observatory*, Proceedings of the 33rd International Cosmic Ray Conference, Rio de Janeiro, Brazil (2013), arXiv:1307.5059.
- [53] IceCube Collaboration, M. Aartsen *et al.*, *Probing the origin of cosmic-rays with extremely high energy neutrinos using the IceCube Observatory*, Phys. Rev. **D88**, 112008 (2013), arXiv:1310.5477.
- [54] ANITA Collaboration, P. Gorham *et al.*, *New Limits on the Ultra-high Energy Cosmic Neutrino Flux from the ANITA Experiment*, Phys. Rev. Lett. **103**, 051103 (2009), arXiv:0812.2715.
- [55] ANITA Collaboration, P. Gorham *et al.*, *The Antarctic Impulsive Transient Antenna Ultra-high Energy Neutrino Detector Design, Performance, and Sensitivity for 2006-2007 Balloon Flight*, Astropart. Phys. **32**, 10 (2009), arXiv:0812.1920.
- [56] ANITA Collaboration, P. Gorham *et al.*, *Erratum: Observational Constraints on the Ultra-high Energy Cosmic Neutrino Flux from the Second Flight of the ANITA Experiment*, Phys. Rev. **D85**, 049901 (2012), arXiv:1011.5004.
- [57] RICE Collaboration, I. Kravchenko *et al.*, *Updated Results from the RICE Experiment and Future Prospects for Ultra-High Energy Neutrino Detection at the South Pole*, Phys. Rev. **D85**, 062004 (2012), arXiv:1106.1164.
- [58] HiRes Collaboration, R. U. Abbasi *et al.*, *An upper limit on the electron-neutrino flux from the HiRes detector*, Astrophys. J. **684**, 790 (2008), arXiv:0803.0554 [astro-ph].
- [59] Telescope Array Collaboration, T. Abu-Zayyad *et al.*, *Upper limit on the flux of photons with energies above 10^{19} eV using the Telescope Array surface detector*, Phys. Rev. **D88**, 112005 (2013), arXiv:1304.5614.

- [60] Pierre Auger Collaboration, A. Aab *et al.*, *Large scale distribution of ultra high energy cosmic rays detected at the Pierre Auger Observatory with zenith angles up to 80°*, *Astrophys. J.* **802**, 111 (2015), arXiv:1411.6953.
- [61] D. Harari, S. Mollerach, and E. Roulet, *Anisotropies of ultrahigh energy cosmic rays diffusing from extragalactic sources*, *Phys. Rev.* **D89**, 123001 (2014), arXiv:1312.1366.
- [62] I. Al Samarai for the Pierre Auger Collaboration, *Large-scale cosmic rays anisotropies searches with Auger data*, Proceedings of the 34th International Cosmic Ray Conference, The Hague (2015).
- [63] Pierre Auger Collaboration, P. Abreu *et al.*, *Search for First Harmonic Modulation in the Right Ascension Distribution of Cosmic Rays Detected at the Pierre Auger Observatory*, *Astropart. Phys.* **34**, 627 (2011), arXiv:1103.2721.
- [64] Pierre Auger Collaboration, I. Sidelnik, *Measurement of the first harmonic modulation in the right ascension distribution of cosmic rays detected at the Pierre Auger Observatory: towards the detection of dipolar anisotropies over a wide energy range*, Proceedings of the 33rd International Cosmic Ray Conference, Rio de Janeiro, Brazil (2013), arXiv:1307.5059.
- [65] J. Linsley, *Fluctuation effects on directional data*, *Phys. Rev. Lett.* **34**, 1530 (1975).
- [66] R. Bonino *et al.*, *The East-West method: an exposure-independent method to search for large scale anisotropies of cosmic rays*, *Astrophys. J.* **738**, 67 (2011), arXiv:1106.2651.
- [67] Telescope Array Collaboration, R. Abbasi *et al.*, *Indications of Intermediate-Scale Anisotropy of Cosmic Rays with Energy Greater Than 57 EeV in the Northern Sky Measured with the Surface Detector of the Telescope Array Experiment*, *Astrophys. J.* **790**, L21 (2014), arXiv:1404.5890.
- [68] Pierre Auger Collaboration, A. Aab *et al.*, *Searches for Anisotropies in the Arrival Directions of the Highest Energy Cosmic Rays Detected by the Pierre Auger Observatory*, *Astrophys. J.* **804**, 15 (2015), arXiv:1411.6111.
- [69] Pierre Auger Collaboration, J. Abraham *et al.*, *Correlation of the highest energy cosmic rays with nearby extragalactic objects*, *Science* **318**, 938 (2007), arXiv:0711.2256.
- [70] Pierre Auger Collaboration, J. Abraham *et al.*, *Correlation of the highest-energy cosmic rays with the positions of nearby active galactic nuclei*, *Astropart. Phys.* **29**, 188 (2008), arXiv:0712.2843.
- [71] Pierre Auger Collaboration, P. Abreu *et al.*, *Update on the correlation of the highest energy cosmic rays with nearby extragalactic matter*, *Astropart. Phys.* **34**, 314 (2010), arXiv:1009.1855.

- [72] W. Heitler, *The Quantum Theory of Radiation*. Dover Books on Physics and Chemistry (Dover Publications, 1954).
- [73] J. Matthews, *A Heitler model of extensive air showers*, *Astropart. Phys.* **22**, 387 (2005).
- [74] J. Alvarez-Muniz, R. Engel, T. Gaisser, J. A. Ortiz, and T. Stanev, *Hybrid simulations of extensive air showers*, *Phys. Rev. D* **66**, 033011 (2002).
- [75] J. Linsley, *Structure of large air showers at depth 834 g/cm². III - Applications*, *Int. Cosm. Ray Conf.* **12**, 89 (1977).
- [76] D. Heck, G. Schatz, T. Thouw, J. Knapp, and J. Capdevielle, *CORSIKA: A Monte Carlo code to simulate extensive air showers*, FZKA-6019 (1998).
- [77] S. Sciutto, *AIRES: A System for air shower simulations. User's guide and reference manual. Version 2.2.0.*, (1999), arXiv:astro-ph/9911331.
- [78] S. Roh *et al.*, *Comparison of CORSIKA and COSMOS simulations*, *American Institute of Physics Conference Series* **1367**, 177 (2011), arXiv:1104.1005.
- [79] T. Pierog and K. Werner, *Muon Production in Extended Air Shower Simulations*, *Phys. Rev. Lett.* **101**, 171101 (2008).
- [80] S. Ostapchenko, *Monte Carlo treatment of hadronic interactions in enhanced Pomeron scheme: QGSJET-II model*, *Phys. Rev. D* **83**, 014018 (2011).
- [81] E. Ahn, R. Engel, T. K. Gaisser, P. Lipari, and T. Stanev, *Cosmic ray interaction event generator SIBYLL 2.1.*, *Phys. Rev. D* **80**, 094003 (2009).
- [82] J. Knapp, D. Heck, S. Sciutto, M. Dova, and M. Risse, *Extensive air shower simulations at the highest energies*, *Astropart. Phys.* **19**, 77 (2003).
- [83] T. K. Gaisser *et al.*, *Cosmic-ray composition around 10¹⁸ eV*, *Phys. Rev. D* **47**, 1919 (1993).
- [84] M. Nagano and A. A. Watson, *Observations and implications of the ultra high-energy cosmic rays*, *Rev. Mod. Phys.* **72**, 689 (2000).
- [85] A. Porcelli for the Pierre Auger Collaboration, *Measurements of X_{max} above 10¹⁷ eV with the fluorescence detector of the Pierre Auger Observatory*, Talk at the 34th International Cosmic Ray Conference, The Hague (2015).
- [86] Pierre Auger Collaboration, A. Aab *et al.*, *The Pierre Auger Cosmic Ray Observatory*, Accepted for publication in *Nucl. Instrum. Meth.* (2015), arXiv:1502.01323.
- [87] T. Mathes for the Pierre Auger Collaboration, *The HEAT Telescopes of the Pierre Auger Observatory. Status and First Data*, Proceedings of the 32nd International Cosmic Ray Conference Proceedings, Beijing (2011).

- [88] F. Sanchez for the Pierre Auger Collaboration, *The AMIGA Detector of the Pierre Auger Observatory: An Overview*, Proceedings of the 32nd International Cosmic Ray Conference, Beijing (2011).
- [89] M. A. Lawrence, R. J.O. Reid, and A. A. Watson, *The cosmic ray energy spectrum above 4×10^{17} eV as measured by the Haverah Park array*, *J.Phys.G* **17**, 733 (1991).
- [90] X. Bertou *et al.*, *Calibration of the surface array of the Pierre Auger Observatory*, *Nucl. Instrum. Meth. A* **568**, 839 (2006).
- [91] Pierre Auger Collaboration, J. Abraham *et al.*, *Trigger and aperture of the surface detector array of the Pierre Auger Observatory*, *Nucl. Instrum. Meth.* **A613**, 29 (2010), arXiv:1111.6764.
- [92] P. Billoir, GAP Report No. 179, 2009 (unpublished).
- [93] P. Billoir, P. Ghia, D. Nitz, and R. Sato, GAP Report No. 032, 2011 (unpublished).
- [94] P. Billoir, GAP Report No. 089, 2011 (unpublished).
- [95] P. Ghia, GAP Report No. 018, 2004 (unpublished).
- [96] P. Ghia and G. Navarra, GAP Report No. 007, 2003 (unpublished).
- [97] Pierre Auger Collaboration, J. Abraham *et al.*, *The fluorescence detector of the Pierre Auger Observatory*, *Nucl. Instrum. Meth. A* **620**, 227 (2010), arXiv:0907.4282.
- [98] AIRFLY Collaboration, M. Ave *et al.*, *Precise measurement of the absolute fluorescence yield of the 337 nm band in atmospheric gases*, *Astropart. Phys.* **42**, 90 (2013), arXiv:1210.6734.
- [99] Pierre Auger Collaboration, P. Abreu *et al.*, *The rapid atmospheric monitoring system of the Pierre Auger Observatory*, *JINST* **7**, P09001 (2012), arXiv:1208.1675.
- [100] L. Wiencke and the Pierre Auger Collaboration, *Atmospheric calorimetry above 10^{19} eV: Shooting lasers at the Pierre Auger Cosmic-Ray Observatory*, *J. Phys. Conf. Ser.* **160**, 012037 (2009), arXiv:0807.2884.
- [101] C. Bonifazi for the Pierre Auger Collaboration, *The angular resolution of the Pierre Auger Observatory*, *Nucl. Phys. B Proc. Sup.* **190**, 20 (2009), arXiv:0901.3138, Proceedings of the 31st International Cosmic Ray Conference.
- [102] M. Unger, GAP Report No. 010, 2006 (unpublished).
- [103] T. Gaisser and A. Hillas, *Reliability of the method of constant intensity cuts for reconstructing the average development of vertical showers*, *Int. Cosm. Ray Conf.* **8**, 353 (1977).
- [104] M. Tueros for the Pierre Auger Collaboration, *Estimate of the non-calorimetric energy of showers observed with the fluorescence and surface detectors*, Proceedings of the 33rd International Cosmic Ray Conference, Rio de Janeiro (2013).

- [105] V. Verzi for the Pierre Auger Collaboration, *The energy scale of the Pierre Auger Observatory*, Proceedings of the 33rd International Cosmic Ray Conference, Rio de Janeiro (2013).
- [106] Pierre Auger Collaboration, A. Aab *et al.*, *The Pierre Auger Observatory: Contributions to the 33rd International Cosmic Ray Conference (ICRC 2013)*, (2013), arXiv:1307.5059.
- [107] S. Argirò *et al.*, *The offline software framework of the Pierre Auger Observatory*, Nucl. Instrum. Meth. A **580**, 1485 (2007), arXiv:0707.1652.
- [108] D. Veberic and M. Roth, GAP Report No. 035, 2005 (unpublished).
- [109] Pierre Auger Collaboration, M. Roth, *The Lateral distribution function of shower signals in the surface detector of the Pierre Auger Observatory*, p. 333 (2003), arXiv:astro-ph/0308392.
- [110] D. Newton, J. Knapp, and A. A. Watson, *The Optimum Distance at which to Determine the Size of a Giant Air Shower*, Astropart. Phys. **26**, 414 (2007), arXiv:astro-ph/0608118.
- [111] M. Ave, P. Bauleo, and T. Yamamoto, GAP Report No. 030, 2003 (unpublished).
- [112] A. Castellina, GAP Report No. 031, 2003 (unpublished).
- [113] P. Bauleo, A. Castellina, R. Knapik, G. Navarra, and J. Harton, GAP Report No. 047, 2004 (unpublished).
- [114] Pierre Auger Collaboration, J. Abraham *et al.*, *The accuracy of signal measurement with the water Cherenkov detectors of the Pierre Auger Observatory*, Nucl. Instrum. Meth. **A578**, 180 (2007).
- [115] M. Rammes, *Investigating signal fluctuations of the surface detector array of the Pierre Auger Observatory using pair tanks*, PhD thesis, University of Siegen, (2008), Master Thesis.
- [116] R. Hiller and M. Roth, GAP Report No. 012, 2012 (unpublished).
- [117] P. Billoir, GAP Report No. 005, 2001 (unpublished).
- [118] I.C. Mariş, F. Schüessler, R. Ulrich and M. Unger, GAP Report No. 081, 2006 (unpublished).
- [119] I.C. Mariş, S. Müller, F. Schüessler, R. Ulrich and M. Unger, GAP Report No. 012, 2009 (unpublished).
- [120] I. Lhenry-Yvon, GAP Report No. 021, 2014 (unpublished).
- [121] C. Jillings *et al.*, GAP Report No. 037, 2002 (unpublished).
- [122] C. Bonifazi, A. Letessier-Selvon, and E. Santos, *A model for the time uncertainty measurements in the Auger surface detector array*, Astropart. Phys. **28**, 523 (2008).

- [123] C. Bonifazi and A. Letessier-Selvon, GAP Report No. 016, 2006 (unpublished).
- [124] D. Veberic for the Pierre Auger Collaboration, *Estimation of Signal in Saturated Stations of Pierre Auger Surface Detector*, Proceedings of the 33rd International Cosmic Ray Conference, Rio de Janeiro (2013).
- [125] Pierre Auger Collaboration, P. Abreu *et al.*, *The Pierre Auger Observatory I: The Cosmic Ray Energy Spectrum and Related Measurements*, (2011), arXiv:1107.4809.
- [126] Pierre Auger Collaboration, A. Aab *et al.*, *Reconstruction of inclined air showers detected with the Pierre Auger Observatory*, JCAP **1408**, 019 (2014), arXiv:1407.3214.
- [127] E. dos Santos, GAP Report No. 014, 2014 (unpublished).
- [128] B. Zamorano, GAP Report No. 039, 2014 (unpublished).
- [129] L. Collica, GAP Report No. 103, 2014 (unpublished).
- [130] L. Cazon, R. Conceição, M. Pimenta, and E. Santos, *A model for the transport of muons in extensive air showers*, Astropart. Phys. **36**, 211 (2012).
- [131] L. Cazón, R. Vázquez, and E. Zas, *Depth development of extensive air showers from muon time distributions*, Astropart. Phys. **23**, 393 (2005).
- [132] L. Cazón, R. Vázquez, A. Watson, and E. Zas, *Time structure of muonic showers*, Astropart. Phys. **21**, 71 (2004).
- [133] D. Garcia Gamez, *Muon Arrival Time distributions and its relationship to the mass composition of Ultra High Energy Cosmic Rays: An application to the Pierre Auger Observatory*, PhD thesis, Universidad de Granada, (2010), PhD Thesis.
- [134] P. Billoir, *A sampling procedure to regenerate particles in a ground detector from a 'thinned' air shower simulation output*, Astropart. Phys. **30**, 270 (2008).
- [135] Ioana Mariş, private communication.
- [136] Pierre Auger Collaboration, A. Aab *et al.*, *Depth of maximum of air-shower profiles at the Pierre Auger Observatory. I. Measurements at energies above $10^{17.8}$ eV*, Phys. Rev. **D90**, 122005 (2014), arXiv:1409.4809.
- [137] L. Collica for the Pierre Auger Collaboration, *Measurement of the muon content in air showers at the Pierre Auger Observatory*, Proceedings of the 34th International Cosmic Ray Conference, The Hague (2015).
- [138] ROOT Data analysis Framework,
<https://root.cern.ch/drupal/>.
- [139] M. Morhac *et al.*, *Efficient one- and two-dimensional Gold deconvolution and its application to gamma-ray spectra decomposition*, Nucl. Instrum. Meth. A **401**, 385 (1997).

-
- [140] NOAA Air Resources Laboratory (ARL), Global Data Assimilation System (GDAS1) Archive Information,
<http://ready.arl.noaa.gov/gdas1.php>.
- [141] Pierre Auger Collaboration, P. Abreu *et al.*, The Pierre Auger Observatory IV: Operation and Monitoring, in *Proceedings, 32nd International Cosmic Ray Conference (ICRC 2011)*, 2011, arXiv:1107.4806.
- [142] P. Auger, Upgrade design report,
https://www.auger.unam.mx/AugerWiki/UTF_Home.
- [143] P. Billoir, GAP Report No. 106, 2014 (unpublished).
- [144] v2r8-blacktank-experimental,
<https://devel-ik.fzk.de/svn/auger/Offline/branches/v2r8-blacktank-experimental>.
- [145] A. Letessier-Selvon, P. Billoir, M. Blanco, I. C. Mariş, and M. Settimo, *Layered water Cherenkov detector for the study of ultra high energy cosmic rays*, Nucl. Instrum. Meth. **A767**, 41 (2014), arXiv:1405.5699.
- [146] Scherini, V. and Perrone, L. and Maldera, S., GAP Report No. 089, 2014 (unpublished).
- [147] O. Deligny, P. Ghia, Antoine, Letessier-Selvon, and A. Watson, GAP Report No. 040, 2013 (unpublished).