Universidad de Alcalá

DEPARTAMENTO DE FÍSICA

TESIS DOCTORAL

Composition and energy determination in cosmic ray surface arrays: an application to the Pierre Auger Observatory

Dirigida por los doctores:
L. del Peral Gochicoa, G. A. Medina Tanco,
M. D. Rodríguez Frías y F. Arqueros Martínez.

Germán Ros Magán

Alcalá de Henares. Septiembre de 2009.

Universidad de Alcalá

DEPARTAMENTO DE FÍSICA

Composition and energy determination in cosmic ray surface arrays: an application to the Pierre Auger Observatory

Memoria presentada para optar
al Grado de Doctor en Física por:
Germán Ros Magán
Alcalá de Henares. Septiembre de 2009.

Fdo: Germán Ros Magán

Agradecimientos

El camino recorrido durante estos años, hasta llegar aquí, ha sido muy gratificante, tanto por el desafío personal que ha supuesto en todo momento como por la gente que he tenido la oportunidad de conocer. Al comienzo de esta aventura, uno no sabe dónde se mete ni hacia dónde le conducirá, pero para mí, a pesar de las dificultades y de algunos momentos duros, ha sido la mejor experiencia de mi vida y ha estado repleta de regalos.

Primero y ante todo, he de dar las gracias a mi familia: a mis padres que me han permitido seguir el camino que he querido, dándome todo su apoyo y que se han preocupado tanto o más que yo en algunos momentos; y, por supuesto, a mi hermano y a mis abuelos, a los que tanto necesito.

A Luis y a Loly, mis directores de tesis en el día a día en la Universidad de Alcalá, agradecerles la confianza que han depositado en mí, por guiarme, ayudarme en todo lo que está en su mano y por haber conseguido siempre lo mejor para mí. Me gustaría agradecer también a Fernando, por quien me inicié en el mundo de la investigación y llegué a esto de los rayos cósmicos.

A Gustavo, he de agradecerle muchas cosas, entre ellas el haber sabido sacar lo mejor de mí y haber creído en mí más que yo mismo, indicándome el camino correcto, ofreciéndome su apoyo. Además, me ha dado su amistad y me ha abierto la puerta al grupo de gente que he conocido en la UNAM y en México, que tanto significan para mí. Aprovecho para dar las gracias a todos ellos: Marat, Mané, Chamín, Daniel, Andrea, Fede, Gabi, Cinzia, Natalia, Alondra, Camila ... Podía haber sido muy difícil pasar 14 meses allí, y sin embargo, han sido unos meses maravillosos gracias a todos vosotros. Quiero destacar un poquito a Marat por acogerme desde el primer día, a Chamín por enseñarme que México es mucho más que el DF, a Daniel porque conocerle y trabajar al lado de una persona como él ha sido un verdadero un placer y a Natalia por haberme dedicado gran parte de su tiempo y ayudado a conocer un país y una gente tan maravillosos como México.

A mis amigos de siempre, a los de la Complu, muchos de los cuales están en este mismo camino, y a los nuevos que he descubierto en Ciencias durante la tesis, gracias por haber estado ahí. Sé que puedo contar con todos vosotros. Gracias también Eduardo por haberme ayudado en la correción de la tesis y a Izas, por hacer que el inglés suene algo mejor.

He de mencionar a las instituciones que me han apoyado y financiado en estos años. A la Comunidad de Madrid por la concesión de una F.P.I. y varias ayudas de viaje, al Departamento de Física y a la Universidad de Alcalá, donde he pasado la mayor parte de este tiempo siempre muy a gusto y que también ha permitido que pueda asistir a varios congresos y estancias, al programa HELEN que me ha permitido realizar diversas estancias en el extranjero que han sido muy necesarias para este resultado final y, por último, al Instituto de Ciencias Nucleares de la Universidad Nacional Autónoma de México, al que pertenezco como estudiante asociado de postgrado, que me ha acogido durante más de un año de gran provecho profesional y personal.

Pero esta tesis además me tenía guardada una gran sorpresa, el mayor regalo, un día María Ángeles entró en mi despacho y sé que no se irá nunca. Eres lo más bonito del mundo.

Contents

1 Introduction **1**

2 Ultra-high Energy Cosmic Rays **5**

2.1 History and cosmic ray discoveries 5

2.2 Ultra-high energy cosmic rays physics 7

2.2.1 Candidate sources and acceleration mechanisms 7

2.2.2 Is cosmic ray astronomy possible?. Propagation, magnetic fields and the GZK effect 14

2.2.3 Energy and composition 20

2.3 Extensive air showers 29

2.3.1 Phenomenology 29

2.3.2 Longitudinal development 33

2.3.3 Lateral development 35

2.3.4 Detection Techniques 38

3 Pierre Auger Observatory **43**

3.1 Background and advantages of a hybrid detector 43

3.2 Surface detectors 46

3.3 Fluorescence detectors 48

3.4 Composition Observables 51

3.4.1 Parameters from the fluorescence technique 51

3.4.2 Parameters from the surface detectors 53

3.4.3 Why the issue of composition is so difficult? 61
3.5 Pierre Auger South Observatory: Results 64
3.5.1 UHECRs: spectrum 64
3.5.2 UHECRs: composition 66
3.5.3 Upper limit on the diffuse flux of ultra-high energy neutrinos 72
3.5.4 Search for UHECR sources and anisotropies 74

4 Optimum distance in surface arrays 79
4.1 Motivation 80
4.2 Algorithm to determine the optimum distance 83
4.3 r_{opt} dependencies on energy, zenith angle and array size 88
4.4 Influence of r_{opt} on the reconstructed energy 94
4.4.1 Energy error distribution functions 95
4.4.2 Bias in the reconstructed energies 95
4.4.3 Reconstruction of a rapidly changing spectrum 100
4.5 Summary and discussion 105

5 A new parameter for composition discrimination 107
5.1 Motivation 108
5.2 Analytical study 110
5.2.1 Optimization assuming Auger tanks 112
5.2.2 Modifying the slope of the LDF 114
5.2.3 Modifying the muon content of the simulated showers 116
5.3 Numerical analysis 118
5.3.1 Optimization and comparison with the analytical result 119
5.3.2 Influence of the detectors far from the shower axis 120
5.3.3 Energy and zenith angle dependence 122
5.4 Application 124
5.5 Summary 130

6 On-going work and perspectives 133

7 Conclusions and Outlook 139

A Optimum distance at Auger North array 147

A.1 Introduction . . . 147
A.2 Algorithm: r_{opt} and energy determination . . . 148
A.3 r_{opt} dependence on energy and zenith angle . . . 153
A.4 Energy error distributions . . . 154
  A.4.1 Shape of the energy error distributions . . . 157
  A.4.2 Bias in the reconstructed energies . . . 157
A.5 Summary and discussion . . . 162

B Reconstruction of surface events at Auger 165

B.1 Station and event selection . . . 166
B.2 Plane fit to the shower front . . . 169
B.3 The lateral distribution function . . . 171
B.4 Maximum Likelihood . . . 172
B.5 Curvature shower front . . . 175
B.6 Fit stages . . . 176
B.7 Energy estimation . . . 177

C Resumen 179

C.1 Antecedentes . . . 179
C.2 Distancia óptima en los experimentos de superficie . . . 182
C.3 Un nuevo parámetro para estudios de composición . . . 186
C.4 Conclusiones . . . 192

Bibliography 201

List of Figures

2.1 V. Hess in the balloon flights which led to the discovery of cosmic rays . . 7
2.2 First and second order Fermi acceleration mechanisms 11
2.3 Hillas Plot: possible sources of UHECRs 12
2.4 Deflection of UHECRs by magnetic fields 15
2.5 Reduction of the primary proton energy due to GZK effect 17
2.6 Several interactions of proton and iron nucleus with the CMB 19
2.7 Energy spectrum of cosmic rays 21
2.8 Energy spectrum of cosmic rays measured by most of the experiments in the whole energy range 22
2.9 Energy spectrum of cosmic rays at the highest energies multiplied by $E^{2.7}$. 22
2.10 The ankle and the dip models for the transition region of the spectrum . . 25
2.11 The mixed composition model for the transition region of the spectrum . . 26
2.12 The trans-GZK recovery of the spectrum 28
2.13 Sketch of an Extensive Air Shower and energy flow between different shower components 31
2.14 Longitudinal development of an EAS 35
2.15 Lateral development of an EAS 37

3.1 Picture of hybrid detection of an EAS 45
3.2 Layout of the Auger South Observatory 46
3.3 An Auger surface water Cherenkov detector 47
3.4 Sketch of the Fluorescence Telescope 49

3.5 Camera, mirror and diaphragm of the fluorescence telescopes 50

3.6 Elongation Rate from different experiments 52

3.7 Geometrical correlation between the longitudinal development of a shower and the time structure 55

3.8 Rise time benchmark to determine the parameter $< \Delta >$ 56

3.9 Radius of curvature as a mass sensitive parameter 58

3.10 XAsymMax parameter for composition studies 60

3.11 Composition measurements by several experiments 63

3.12 Auger Spectrum of UHECRs 65

3.13 Vertical SD, horizontal SD and hybrid Auger spectra 66

3.14 X_{max} and RMS of the X_{max} distribution as a function of energy determined by Auger 67

3.15 Composition from surface parameters by Auger 70

3.16 Upper limits on the fraction of photons as primaries of UCHERs of energies above 10^{18} eV reported by Auger 71

3.17 Upper limit on the diffuse flux of ultra-high energy neutrinos published by Auger 73

3.18 Auger search for excess of cosmic rays in the galactic center region 75

3.19 Correlation between the arrival direction of the highest energy cosmic rays detected by Auger and the position of Active Galactic Nuclei 77

4.1 Examples of the fitting procedure to find r_{opt} 87

4.2 r_{opt} vs. energy for different array spacings and zenith angles 89

4.3 r_{opt} vs. energy. Events with and without saturated detectors are shown separately. 90

4.4 Distance of the triggered stations to the shower axis for events without saturation 90

4.5 r_{opt} vs. zenith angle for different array spacings and energies 91

4.6 r_{opt} vs. zenith angle for 1500 m separation array and for saturated and non-saturated events separately 92
4.7 Distribution functions of the energy errors 96
4.8 Bias in the inferred energy by using both reconstruction methods 97
4.9 Scatter plots of reconstructed energy vs. real energy 98
4.10 Optimum distance as a function of primary energy in a 1 km array 99
4.11 Bias for a different value of r_0 100
4.12 Fraction of events with saturated detectors as a function of energy 101
4.13 Fits of the energy error distributions with the AGG function 102
4.14 Reconstruction of a realistic spectrum 103

5.1 Lateral distribution functions of the muon, electromagnetic and total signal in the Cherenkov detectors for proton and iron primaries 111
5.2 Distance of the stations to the shower axis 114
5.3 Merit factor for S_b as a function of b 115
5.4 Merit factor of S_b as a function of b when modifying the slope of the Iron and proton LDFs 116
5.5 Mean value of S_3 for protons and iron nuclei when the number of muons is modified 117
5.6 Merit factor of S_b as a function of b when the number of muons is modified 118
5.7 Merit factor of S_b as a function of b obtained from simulated data 120
5.8 Distance of the furthest triggered station to the shower axis 121
5.9 Merit factor of S_3 as a function of energy for several cuts 122
5.10 S_3 vs. $sec(\theta)$ 123
5.11 $\log(S_3/VEM)$ vs. $\log(E/eV)$ 124
5.12 $\log(S_3/VEM)$ vs. $\log(S_{38}/VEM)$ 124
5.13 $t_{1/2}(1000)$ as function of $sec(\theta)$ 126
5.14 Inferred vs. true proton fraction using the rise time, X_{max} and S_3. The hadronic model used is QGSJet-II 128

5.15 Inferred vs. true proton fraction using the rise time, X_{max} and S_3. The hadronic model used is Sibyll 2.1 129
5.16 Fits of the parameter distributions using an AGG function 130
5.17 Error in the inferred proton abundance determined by using S_3, X_{max} and the rise time for 1 and 5 years of Auger exposure 131

6.1 Examples of the LDF fit performed by Auger and the method developed here to find r_{opt} 136

A.1 Comparison between both LDFs considered 151
A.2 Number of triggered stations vs. energy for both LDFs 152
A.3 Fraction of saturated events as a function of energy for both LDFs 152
A.4 r_{opt} dependence on the primary energy in a square grid 154
A.5 r_{opt} dependence on primary energy in a square grid for saturated and non-saturated events 155
A.6 r_{opt} dependence on zenith angle in a square grid 156
A.7 Energy error distributions using the Old LDF 158
A.8 Energy error distributions using the New LDF 159
A.9 Bias introduced by both reconstruction methods 160
A.10 Scatter plot of r_{opt} as a function of energy 161
A.11 Inferred energy vs. real energy 161

B.1 T4 Trigger: 3TOT compact congurations 168
B.2 T4 Trigger: 4C1 congurations 169
B.3 Sketch of the plane front arrival 170
B.4 Sketch of the spherical shower front development. 176

List of Tables

2.1 Experiments based on an array of surface detectors 39
2.2 Experiments based on the fluorescence technique 41

Chapter 1

Introduction

Indirect detection of radiation from outside the Earth was first discovered by Victor Hess in 1912 [1]. These high energy particles were called *Cosmic Rays* (CR) and still represent a big challenge in physics. Since then, a huge progress has been made from both the theoretical and the experimental point of view but fundamental questions still remain open: Where do cosmic rays come from? How are they accelerated to such high energies? What is the composition of the most energetic cosmic rays? How do we interpret the features observed in the energy spectrum? In addition, these questions are intrinsically correlated making the problem even more complicated to solve.

The energy spectrum of cosmic rays extends over twelve orders of magnitude from 10^9 to more than 10^{20} eV. They hit the Earth's atmosphere at the rate of about 1000 per square meter per second. Cosmic rays at energies up to 10^{15} eV could be detected by direct measurements with balloons launched at high altitudes in the atmosphere or with satellites. However, the spectrum decreases as $\sim E^{-2.7}$, where E is the energy of the particle, so that at high energies the flux is so low that direct measurements are not feasible. On the other hand, once high energy cosmic rays hit the upper atmosphere, the sequence of interactions and cascades of particles create the so-called *extensive air shower* (EAS). Thus, the properties of the primary cosmic ray at high energies could be indirectly determined by studying the subsequently produced air shower.

Two different techniques are traditionally used to study the extensive air showers.

First, telescopes could collect the fluorescence light emitted by atmospheric Nitrogen molecules after they have been excited by the cascade particles. That is the method used by Fly's Eye [2] and HiRes [3] experiments. Second, one can use an array of detectors located at ground level, such as scintillators (e.g. AGASA [4]) or water Cherenkov tanks (e.g. Haverah Park [5]). In the case of surface detectors, the shower front is sampled at a discrete set of points at a single observation level, where cascade particles deposit energy at the detectors. The Pierre Auger Observatory [6] represents a step forward in the study of CRs because it combines both techniques, the fluorescence telescopes and the array of water Cherenkov tanks. Therefore, Auger can detect a sizable fraction of events simultaneously with both techniques (*hybrid events*), significantly improving the measurements of the cascade properties. Additionally, the surface array at Auger is the largest ever made, providing the statistics needed to study the ultra-high energy cosmic rays (UHECRs), whose energies are higher than 10^{18} eV. In the near future, observations of UHECRs may be possible from space by observing the air fluorescence and the reflected Cherenkov light produced by the cascade. In this direction the JEM-EUSO experiment, which will be located at the International Space Station, is in phase B [7].

In the present work, we concentrate on the surface array technique, dealing with two of the problems related to UHECRs: the determination of the energy spectrum and the chemical composition.

Regarding the energy spectrum determination, a new method to improve the inference of the primary particle energy is suggested. Experiments based on surface array of detectors, are able to measure the lateral distribution of particles (i.e. the measured signal or particle density as a function of the distance to the shower axis) and to use the inferred signal at a characteristic distance as energy estimator. This characteristic distance is considered as a fixed parameter for all the showers independently on their energy or direction. On the other hand, we propose to calculate a specific point in the lateral distribution of particles for each individual shower and demonstrate that the interpolated signal at this distance is a better energy estimator. First, we focus on pure surface array experiments and follow the procedure developed by AGASA. Later, this study is applied to the future

Pierre Auger North Observatory so that the implications of its different array geometry and different energy calibration (obtained from hybrid data instead from Monte Carlo simulations as pure surface arrays do) are also analyzed.

The problem of composition is also tackled. A new family of parameters, which make exclusive use of surface data, are proposed and applied to the Pierre Auger South Observatory. We perform analytical and numerical studies of the composition estimators in order to assess their reliability, stability and possible optimization. The effects of experimental uncertainties, intrinsic fluctuations and reconstruction errors are taken into account. In particular, special attention is paid to the effect of a possible underestimation of the size of the muon component in the simulated showers, as it is suggested by experimental evidence. The potential discrimination power of an optimized realization of these parameters is compared on a simplified, albeit quantitative way, with that expected from other surface and fluorescence estimators obtained in similar experimental conditions.

This PhD. thesis is organized as follows. Chapter 2 starts with the history of cosmic rays and the significant discoveries that were carried out since the beginning of their study. A review on the physics related to UHECRs such as their energy spectrum, origin, composition and propagation are given. Finally, a brief description of the phenomenology of the EAS and of the different techniques to detect them are explained. In Chapter 3 the Pierre Auger South Observatory is reviewed. First, the fluorescence telescopes and the water Cherenkov tanks are described. Second, the main composition observables from both techniques are discussed and, finally, the main results published by the Auger collaboration are presented. Chapter 4 is devoted to the question of energy spectrum determination from surface arrays assuming an AGASA-like experiment. The new parameter proposed for composition studies in surface array experiments, is presented in Chapter 5. Extensive analytical and numerical studies are shown. Perspectives, on-going work and the conclusions are presented in Chapters 6 and 7 respectively. In addition, the study shown in Chapter 4 is applied to the Pierre Auger North Observatory in Appendix A, and the standard reconstruction procedure of surface events at Auger is discussed in Appendix B.

Chapter 2

Ultra-high Energy Cosmic Rays

2.1 History and cosmic ray discoveries

The study of cosmic rays started approximately in 1900 as a result of the observation of ionization in gases contained in closed vessels. First hypothesis to explain this phenomena were that it was the consequence of radioactive radiation coming from the surface of the Earth, from the walls of the vessel or from radioactive emanations in the gas. In order to rule out these hypothesis, balloon flights were undertaken. They led to the definite discovery of the cosmic rays by Victor Hess in 1912 [1] (Fig. 2.1), who observed that the ionization rate at altitude around 5 km was several times that the observed at sea level, and therefore, the radiation must come from outside the Earth. The term *Cosmic Rays* to this radiation was coined by Robert Millikan.

The interaction of the Earth's magnetic field on charged particles propagation trough the atmosphere was discovered in 1927. It was demonstrated that it affects the cosmic rays that come from the East differently than those from the West, so that it was proved that cosmic rays are mainly charged particles.

The discovery of cosmic rays was an invaluable tool for early particle physicists because they are the most energetic particles of the Universe and, when hitting the atmosphere, provide the circumstances for the creation of previously undiscovered particles. In 1931, Anderson [8] gave the proof of the existence of a positively charged particle with an identical mass as the electron using cosmic rays. This particle had been previously proposed by Dirac [9] and it was correctly interpreted later as an anti-electron, called positron. Anderson and Hess shared the Nobel prize in 1936 for their work.

In 1937 Anderson and Neddermeyer, and at the same time Street and Stevenson [10], discovered a particle with the same mass as the one that Yukawa had proposed associated with the strong nuclear force [11]. It was in 1947 when it was discovered that they are two different particles with similar masses that abound in cosmic ray air showers, called muon (μ) and pion (π). The latter was the one proposed by Yukawa. In 1947, a new type of particle was discovered that was different from the previously known ones. It was a new particle with the mass of at least twice that of pions later called the kaon (K^0) [12]. It is formed by strange quarks, and it was the first of these kind of particles that were discovered using cosmic rays.

A crucial advance by Pierre Auger and collaborators took place in 1938 [13]. They observed an unexpectedly high rate of coincidences among counters located at the same altitude and separated by large distances using electronics with microsecond timing. They correctly interpreted this result proving the existence of Extensive Air Showers (EAS) generated by a single particle, the cosmic ray, entering in the atmosphere. The interaction of a cosmic ray of high enough energy with an atmospheric nucleus cause a cascade of particles falling to the Earth's surface at the same time. On the basis of their measurements, and using a simple model of shower development and the distance between counters, they were able to estimate that the energy of this primary particle should be around 10^{15} eV.

Since the second half of the 20^{th} century, the search for the high energy cosmic rays began. Large array of surface detectors were first used encouraged by Bassi et al. at MIT in 1953 [14], who were able to reconstruct the original direction of the cosmic ray from the timing information in their array of scintillation detectors. In 1963, Linsley, using the Volcano Ranch array, detected for the first time a cosmic ray with an energy of 10^{20} eV [15]. Five years later, Tanahashi detected an air shower from an incident cosmic ray of 10^{19} eV using a different technique: fluorescence in the atmosphere [16]. That method

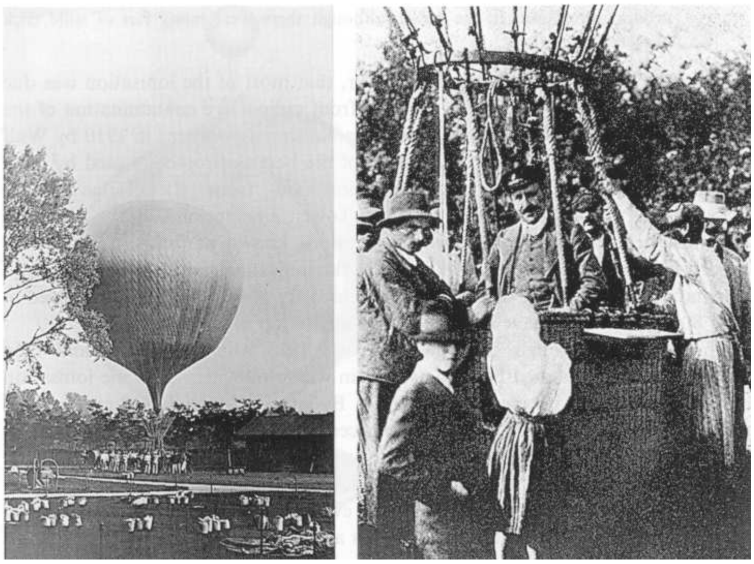

Figure 2.1: V. Hess in the balloon flights which led to the discovery of cosmic rays.

was inspired in the work of Suga and Chudakov who first proposed that the atmosphere could be used as a large scintillator for air shower detection. In a giant step forward, Volcano Ranch recorded a fluorescence event in coincidence with the ground array [17]. That was the first *hybrid* event: an event recorded by two different techniques of detection at the same time and at the same location. The Pierre Auger Observatory [6] is going to study the final region of the cosmic rays energy spectrum, those with energies above 10^{18} eV, which are called ultra-high energy cosmic rays (UHECRs). It started taking data in 2004 and uses the hybrid technique, that together with its huge array, provides the best chance to go further in cosmic ray discoveries.

2.2 Ultra-high energy cosmic rays physics

2.2.1 Candidate sources and acceleration mechanisms

After almost a century since the discovery of cosmic rays, only the Sun has been identified as a source of charged cosmic rays. However, cosmic rays of energies above 10^9 eV cannot be solar in origin since the flux does not exhibit day-night variations. Only some candidates have been found for high energy cosmic rays using mainly theoretical arguments. On the other hand, a large number of gamma ray sources have been identified in past decades by dedicated experiments as Whipple [18], Hess [19] or Magic [20]. In this Section the difficulties to establish the possible sources of UHECRs are analyzed.

UHECRs are extragalactic

Charged cosmic rays are deflected by magnetic fields changing their trajectory. At energies above 10^{18} eV the Larmor radius of a proton in a magnetic field of 1 μG (the typical value of the Galaxy) is around 1 kpc, comparable to the size of the Galaxy (more details are in Section 2.2.2 where the galactic magnetic field is explained). Therefore, the bulk of the cosmic rays of energies lower than 10^{18} eV are considered of galactic origin, probably produced at supernovae (SN). It is still not clear what the maximum acceleration energy achievable by SN is. Recently, it has been argued that SN cannot accelerate nuclei to energies above a few $Z \times 10^{15}$ eV, where Z is the atomic number [21]. More optimistic calculations predict a maximum energy around $Z \times 10^{17}$ eV [22]. Consequently, the majority of the ultra-high energy cosmic rays must be of extragalactic origin.

Astrophysical vs. exotic models

The models devoted to explain the acceleration of cosmic rays to ultra high energies could be divided in two groups. First, the *bottom-up* models, where these particles are accelerated in an astrophysical object. They are studied later in detail along this Section. Second, the so-called *top-down* models, which proposed a more speculative scenarios. One is the production of UHECRs from the decay and annihilation of Super-Heavy Dark Matter particles, which are remnants of the early Universe [23]. Others, called Topological Defect models [24], suggest that unknown X particles are emitted by topological defects formed in the early stages of the Universe, such as magnetic monopoles, cosmic strings and necklaces (a closed loop of cosmic string). The X particles decay and, as by-products, energetic photons, neutrinos and charged leptons together with a small fraction of nucleons are produced with energies up to the X mass without any acceleration mechanism.

Other is the Z-burst model [25]. According to this model, ultra-high energy neutrinos are generated from remote sources somewhere in the Universe. These neutrinos annihilate with the relic neutrinos, which are remnants of the Big Bang, generating Z^0 bosons. The Z^0 boson decays and generates a flux of nucleons, pions, photons and neutrinos. The problem in this model is that no astrophysical source is yet known to meet the requirements for the Z-burst hypothesis.

Most top-down models and the Z-burst model were formulated to avoid the energy loss of cosmic rays due to the interaction with the microwave background radiation, the so-called GZK effect (Section 2.2.2), motivated by the AGASA experiment that did not detect this effect which would have caused a sharp suppression in the spectrum at the highest energies (Section 2.2.3). Even more exotic models were proposed to that end. For example, some theories predict a Lorentz invariance violation that suppresses the cross section for inelastic collision between nucleons and microwave background photons [26].

All these models, except for bottom-up ones, involve that a large fraction of the flux of UHECRs must be gamma-rays. For example, top-down models predict around 10% of gammas at 10 EeV and 50% at 100 EeV [27]. However, this is not confirmed by recent results published by the Pierre Auger Observatory, where the upper limits on the fraction of photons as primaries have been estimated at 1% below 10 EeV, 4% below 20 EeV and 21% below 40 EeV [28], although they are dependent on the choice of the hadronic model used in the analysis (Section 3.5.2). In addition, the GZK effect has been recently confirmed by Auger and HiRes experiments [29, 30] (Section 3.5.1). Therefore, these models are disfavored at the energies around $10^{19} - 10^{19.5}$ eV whereas they could not be rejected definitely. They could be also important at even higher energies. A thorough review on them could be found in [31]. Therefore, we focus on *bottom-up* models hereafter.

Fermi acceleration mechanism in astrophysical objects

If *bottom-up* models are assumed, What will be the possible acceleration mechanisms? The most plausible acceleration mechanism of cosmic rays in astrophysical objects is the one introduced by Fermi in 1949 [32]. Collisions with magnetic clouds accelerate particles by wave-particle resonances in the source plasma. During these resonant encounters, particles can either gain or lose energy. Since the acceleration efficiency goes as the square of the magnetic cloud velocity v ($\Delta E/E \propto \beta^2$, where $\beta = v/c$), the process is known as Fermi-II or second-order Fermi acceleration. The average energy gain is positive in every collision, but slow and small since it is of the second order in β (and $\beta << 1$). In addition, energy losses are significant and mainly caused by ionization and the radiation generated when particle trajectories bend.

This mechanism is modified by the Fermi shock wave acceleration which is much more efficient. It is referred to as first order Fermi acceleration because it is linear with the speed of the shock wave ($\Delta E/E \propto \beta$), resulting in faster acceleration. A shock wave passes through a medium of gas or dust and creates a density gradient at the shock front. The shock wave creates kinetic energy in the medium and there is a resulting net motion as it passes. Particles diffuse and randomly travel in the medium. They have a probability to hit the shock front being accelerated, and then scatter back downstream passing the shock front again gaining more energy. The acceleration continues until energy losses match energy gains, which depends on ambient conditions.

Both processes are schematically shown in Fig. 2.2. The mechanisms are similar but in a different scenario (magnetized clouds or shocks) which essentially modifies the distribution in the number of encounters and the energy gain in each one. The second order Fermi acceleration is often unduly neglected, whereas it cannot be ruled out from the viewpoint of efficiency. Its main defect is that the resulting energy spectral index depends on cloud properties while the first order process gives a universal index as it is observed experimentally. Both processes are more efficient when the flow speed is close to the velocity of light. But in the relativistic regime, the expansion of the first order and second order Fermi processes is not obvious and the theory must be reconsidered.

Other options to accelerate particles to more than EeV energies are direct and fast acceleration achieved by a strong electromagnetic field as it could happen in Gamma Ray Bursts, and the existence of a strong rotating magnetic field (for example in pulsars) which results in a large electromotive force. However, both have several problems as it

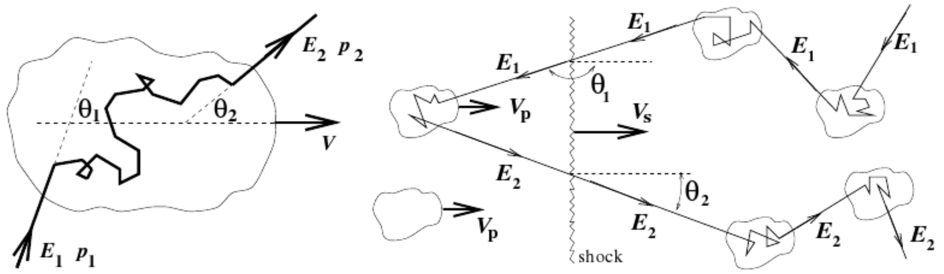

Figure 2.2: Left: sketch of the second order Fermi acceleration mechanism occurring in a moving magnetized cloud. Right: first order Fermi acceleration occurring in strong plane shocks. From [33].

will be commented next, where possible sources are analyzed.

Source candidates

What are the possible astrophysical objects for the origin of UHECRs? Even though the actual acceleration mechanisms are unknown one can rely on very basic arguments to characterize possible source scenarios. Hillas [34] proposed that in order to be able to accelerate charged particles they have to be at least partially confined into some acceleration region and that the maximum achievable energy is given by

$$E_{max}(EeV) \simeq \beta \, Z \, B(\mu G) \, L(kpc) \tag{2.1}$$

where β is the characteristic velocity of particles or fields driving the acceleration in a shock front, Z is the charge of the accelerated particle and B the magnetic field needed to keep the particles inside the acceleration region of size L. This relation is the basis for the so-called Hillas plot shown in Fig. 2.3. It shows that to achieve a given maximum energy, one must have acceleration sites that have either a large magnetic field or a large size of the acceleration region. Only a few astrophysical sources such as active galaxies, hot spots of radio-galaxies, gamma ray bursts and compact objects like neutron stars, seem to satisfy the conditions necessary for acceleration of protons up to 10^{20} eV (diagonal line). Some remarks about them are given in the following:

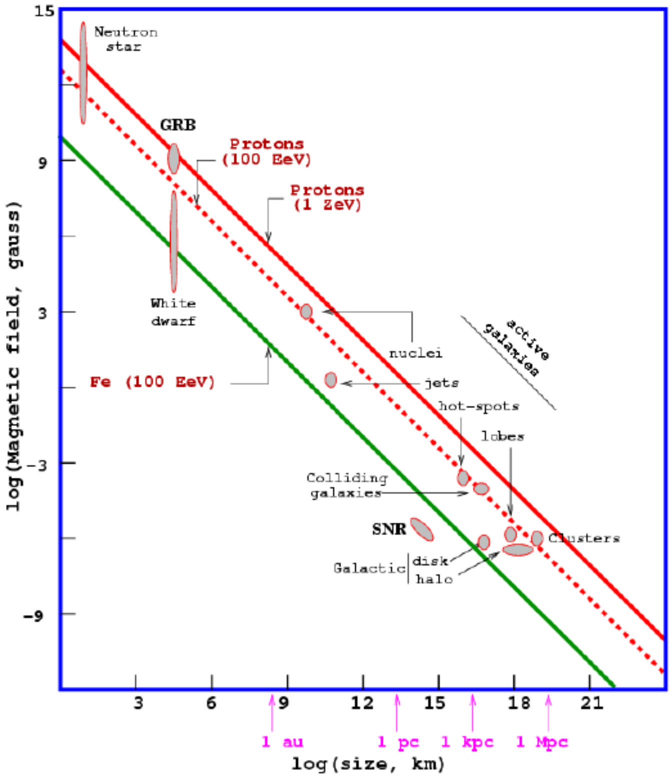

Figure 2.3: Adapted Hillas plot of the magnetic field strength required to accelerate protons and iron to a given energy as a function of the confinement region size. Objects must lie above the given lines in order to be able to accelerate particles to the given energies. From [35].

- Pulsars (B $\sim 10^{13}$ Gauss, L $\sim$ 10 km): they have a strong rotating magnetic field which results in a large electromotive force. This can trap the particle while accelerating it to high energies. However, there are some problems with this model. For example, the power law spectrum observed in cosmic rays is not immediately obvious in this scenario and, the acceleration occurs in a dense region of space where chances for energy loss are high due to meson photo-production, photo-nuclear fission and pair creation. These affect the energy spectrum and the composition of the resulting cosmic rays which are not in agreement with experimental data.

- Gamma Ray Burst (GRB, B $\sim 10^{9}$ Gauss, L $\sim 10^{4} - 10^{5}$ km): The origin of the detected gamma ray bursts can be explained by the collapse of massive stars or mergers of black holes or neutron stars. A relativistic shock is caused by a relativistic fireball in a pre-existing gas, such as a stellar wind, producing or accelerating electrons/positrons to very high energies. The observed gamma-rays are emitted by relativistic electrons via synchrotron radiation and inverse Compton scattering. The detected GRBs release energy up to 10^{51} erg/s which would account for the luminosity required for cosmic rays above 10^{19} eV if the GRBs are uniformly distributed (independently of redshift). However, recent studies indicate that their redshift distribution seems to follow the average star formation rate of the Universe and that GRBs are more numerous at high redshifts. In addition, no correlation between Auger data and GRBs has been reported recently [36].

- Active Galactic Nuclei (AGN, B $\sim 10^3$ Gauss, L $\sim 10^{10}$ km): AGNs are one of the most favored sources for cosmic rays at the highest energies [37]. AGNs are powered by the accretion of matter onto a super massive black hole of $10^6 - 10^8$ solar masses. Typical values of the central engine are $L \sim 10^{-2}$ pc and $B \sim 5$ G, which make possible the confinement of protons up to 10^{20} eV . The main problem here is the large energy loss in a region of high field density, which would limit the maximum energy achievable for protons and forbid the escape for heavy nuclei. Another solution is that the acceleration occurs in AGN jets, where particles are injected with Lorentz factors larger than 10 and energy losses are less significant.
- Cluster of Galaxies (B $\sim 10^{-6}$ Gauss, L $\sim$ 0.1 - 1 Mpc): Galaxy clusters are reasonable sites for ultra-high energy cosmic rays acceleration since particles with energy up to 10^{20} eV can be contained by cluster fields ($\sim 5\mu$G) in a region of size up to 500 kpc. Acceleration in clusters of galaxies could be originated by the large scale motions and the related shock waves resulting from structure formation in the Universe. However, losses due to interactions with the microwave background during the propagation inside the clusters limit UHECRs in cluster shocks to reach 10 EeV.
- Radio Galaxies Hot spots (RGH, B $\sim$ 0.1 - 1 mGauss, L $\sim$ 1 kpc) and Radio Galaxies

Lobes (RGL, B $\sim$ 0.1 μGauss, L $\sim$ 100 kpc). In Fanaroff-Riley II galaxies there are regions of intense synchrotron emission observed within their lobes, known as *hot spots*, and they are produced when the jet ejected by a central super massive black hole interacts with the intergalactic medium generating turbulent fields. The result is a strong shock responsible for particle re-acceleration and magnetic field amplification. The acceleration of particles up to ultra relativistic energies in the hot spots is achieved by repeated scattering through the shock front, similar to the Fermi acceleration mechanism. For typical hot-spot conditions, a maximum acceleration energy for protons is around $5 \cdot 10^{20}$ eV.

All these hypothetical sources are in the border to be able to accelerate particles to the measured energies of UHECRs. In addition, problems as particle injection and the dynamics of the acceleration are still unsolved, as well as propagation processes and the magnetic fields involved.

2.2.2 Is cosmic ray astronomy possible?. Propagation, magnetic fields and the GZK effect

Once candidate sources have been explained in previous Section, a new question appears. Is it possible to make cosmic ray astronomy?. In order to answer, the propagation of cosmic rays from their source to Earth must be studied, which involve to consider how magnetic fields could affect their trajectory and which interactions may suffer. This Section is devoted to these questions.

Magnetic fields

UHECR astronomy would be possible if the original particle direction during its travel from the source to the Earth were conserved. Unfortunately, charged cosmic rays are deflected by the galactic and extragalactic magnetic fields. The problem could be solved if all the quantities needed to determine the primary deflection were known, such as the

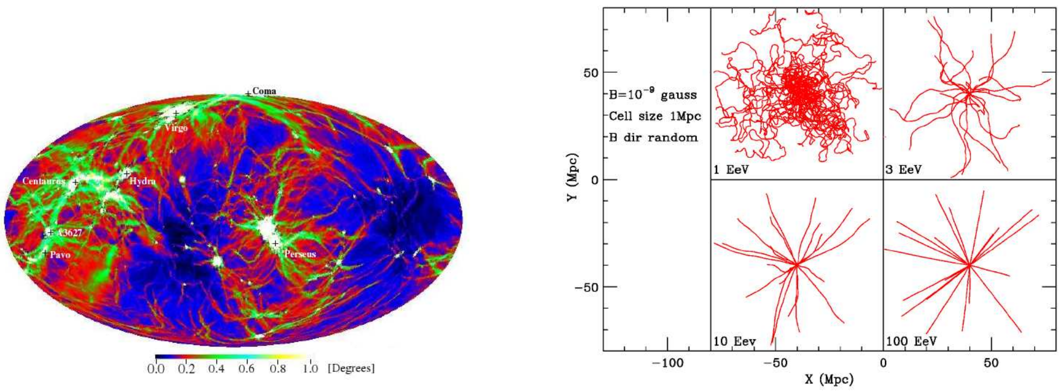

Figure 2.4: (a) Deflections of UHECRs in the local supercluster (from [38]). (b) Projected view of 20 trajectories of proton primaries emanating from a point source for several energies. Each proton is tracked until it reaches a physical distance from the source of 40 Mpc (from [39]).

strength and orientation of the magnetic field, the charge of the cosmic particle and the distance between the source and the Earth.

The magnetic field of the Galaxy can be described as the superposition of two components, one regular and one chaotic. The regular component has an intensity of some few μG and lies on the galactic plane. The chaotic component has an intensity of the same order of magnitude but it is produced from magnetic clouds generated from the motion of ionized gas. If only the regular component is considered, the characteristic deflection of a particle of energy E in the magnetic field B is given by the Larmor radius as $R_L(kpc) \simeq \frac{E(EeV)}{ZB(\mu G)}$. Given a nucleus of charge Z, as the energy increases, the gyroradius of the nucleus becomes comparable or larger than the transversal dimension of the confinement region and, consequently, the nucleus can escape from the Galaxy. Therefore, at energies above 10^{17} eV protons could escape while Iron nuclei are confined inside the Galaxy at least up to energies around 10^{19} eV.

At high energies, as well as the galactic cosmic rays are able to escape from the Galaxy, extragalactic particles are able to penetrate in the galactic confinement region. That is possible if extragalactic particles are able to reach our Galaxy. In fact, in the energy range between $5 \cdot 10^{17}$ and the $3 \cdot 10^{18}$ eV (the energies corresponding to the second knee and the ankle of the spectrum respectively, as will be explained later), all kind of nuclei, starting from protons up to Iron nuclei, are able to arrive from the local universe.

On the other hand, the extragalactic magnetic fields are almost unknown but an upper limit of around 1 nG is usually accepted.

An estimation of the deflection angle in a constant magnetic field which perpendicular component to the particle momentum is $B_{\perp}$ over a distance d is given by [40]:

$$\theta(E,d) = 0.52^{o} Z \left(\frac{E}{10^{20}eV}\right)^{-1} \left(\frac{B_{\perp}}{10^{-9}G}\right) \left(\frac{d}{1Mpc}\right) \tag{2.2}$$

In case of a proton of $\sim 10^{20}$ eV, the deviation is less than 1^{o} in two scenarios: in the Galaxy where magnetic field is typically of $\sim \mu G$ on a distance $\sim kpc$, or outside the Galaxy where the extragalactic magnetic field is the order of $\sim nG$ over a distance of the order of Mpc. The predicted cosmic ray deflections in the local supercluster are shown in Fig. 2.4(a), where values between 0 and 1° are found. As it can be seen in Eq. 2.2, higher the energy of the cosmic ray is, lower is the deflection (see Figure 2.4(b)), so that the door is opened to make cosmic ray astronomy at the highest energies. More detailed calculations could be found in [31, 41, 42].

Interactions of CRs during propagation and the GZK effect

Unfortunately, propagation through galactic and extragalactic magnetic fields is not the only problem. Cosmic rays may interact with background radiation fields like the cosmic microwave background (CMB), the infrared background (IB) and radio background (RB), losing energy. Other energy losses are due to the Hubble expansion of the Universe and due to the interaction with dust, but they are not significant at the energies of our interest. The most important interaction at the highest energies is the *GZK effect* proposed by Greisen, Zatsepin, and Kuz'min [43, 44] just a bit later than the CMB discovery by Penzias and Wilson in 1965 [45]. They independently pointed out that this radiation would make the Universe opaque to cosmic rays of sufficiently high energy. Protons with an energy

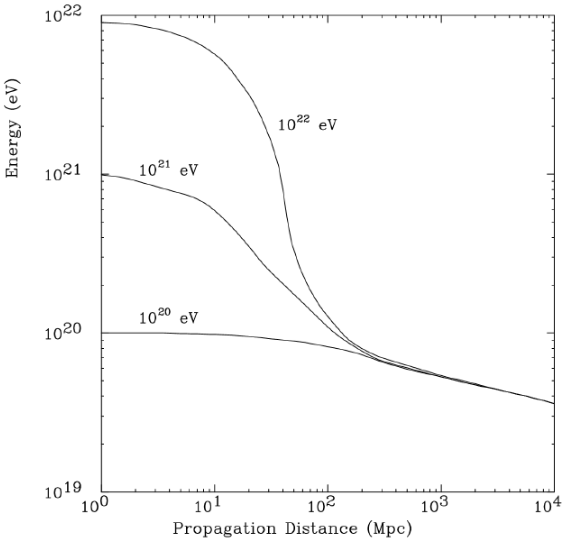

Figure 2.5: Reduction of primary proton energy due to GZK effect. At 10^{22} eV particle would be reduced in energy to 10^{20} eV after traveling $\sim$ 100 Mpc (from [46]).

exceeding $E \sim 5 \cdot 10^{19}$ eV (called *GZK threshold*) have a large probability to interact with the CMB photons, losing energy by pion photo-production:

$$\begin{aligned} p + \gamma_{CMB} &\rightarrow p + \Pi^0 \\ &\rightarrow n + \Pi^+ \end{aligned} \tag{2.3}$$

These interactions occur via the Δ^+ resonance whose cross section at that energy is very high ($\sim 10^{-28} cm^{-2}$). Assuming typical value for the CMB photon density ($400 cm^{-3}$), the mean free path[1] for a proton can be estimated as $\sim$ 8 Mpc. The energy loss per interaction for the proton is $\sim$ 20%, giving an attenuation length[2] of the order of some tenths of Mpc, beyond which the proton energy falls below the GZK threshold. Fig. 2.5 shows how the energy of a proton degrades due to successive interactions with the CMB. On the other hand, the neutron decay length ($n \rightarrow p + e^- + \overline{\nu_e}$) is about 1 Mpc at 10^{20} eV, so that it decays before interacting.

If cosmic rays are protons, another energy loss process will be important between 10^{18} eV and the GZK threshold. It is the photo-pair production when protons interact with

---

[1] The average distance covered by a particle between subsequent interactions

[2] The distance at which the probability that a particle has not been absorbed drops to $1/e$

photons of the CMB producing a electron-positron pair ($p + \gamma_{CMB} \rightarrow p + e^{+} + e^{-}$). The energy loss in each interaction is small. It may, however, contribute to the shape of the spectrum at these energies if the primaries are protons from distant sources. At lower energies the attenuation length tends to become constant and equal to the energy loss due to the expansion of the universe ($\sim$ 4 Gpc).

If cosmic rays are nuclei of mass A, they will undergo due to photo-disintegration and pair production, both with CMB and IR backgrounds:

$$\begin{aligned} A + \gamma_{CMB,IR} &\rightarrow (A-1) + N \\ &\rightarrow (A-2) + 2N \\ &\rightarrow A + e^{+} + e^{-} \end{aligned} \tag{2.4}$$

Since the energy is shared between nucleons, the threshold energy for these processes increases compared to that of protons. The inelasticity is lower by a factor $\sim 1/A$, while the cross section increases with Z^2 . This means that the loss length, in case of heavy nuclei, will be smaller ($\sim$ 1 Mpc) with respect to protons, but it occurs at a higher energies.

Finally, if cosmic rays are photons, the dominant interaction is pair production with the cosmic background photons ($\gamma + \gamma_{CMB,RB} \rightarrow e^{+} + e^{-}$). Pair creation with the CMB is important above $4 \cdot 10^{14}$ eV while attenuation from pair creation with the radio background dominates the energy loss above $2 \cdot 10^{19}$ eV. On the other hand, at energies higher than 10^{22} eV, the attenuation length grows to values of order of 100 Mpc, making possible the hypothesis of photons as primary of extremely-high energy cosmic rays. In fact, these photons could produce secondary photons at energies higher than the GZK-threshold. If this were the case, a secondary photon spectrum $\propto E^{-2}$ should be observed, independently on the source spectrum.

Is charged particle astronomy possible?

The interaction processes explained below are summarized in Fig. 2.6 and they have significant implications on the cosmic ray spectrum and on the possibility of making

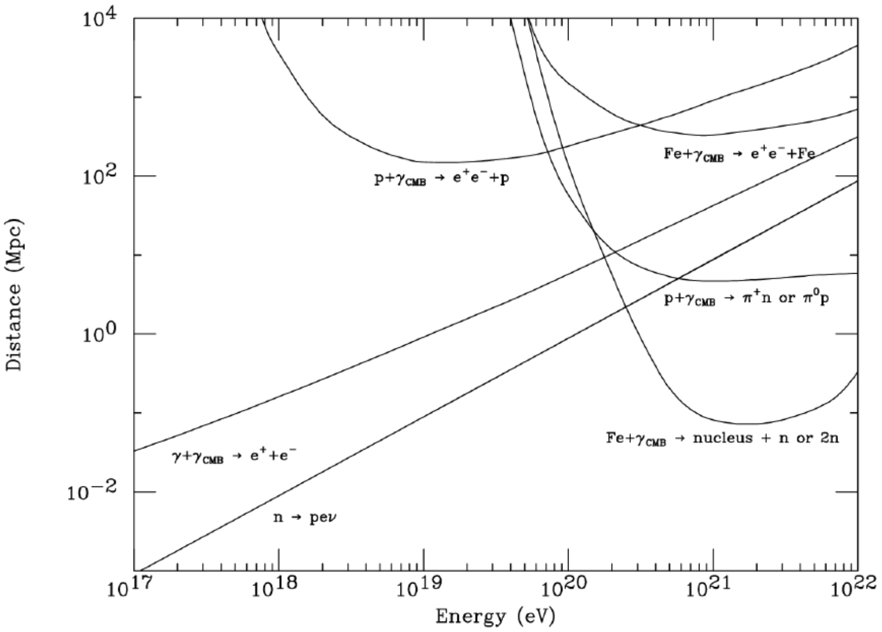

Figure 2.6: Several interactions with the CMB. The curves labeled $p+\gamma_{CMB} \rightarrow e^++e^-+p$ and $Fe+\gamma_{CMB} \rightarrow e^++e^-+Fe$ are the distances for which the proton and the iron nucleus lose $1/e$ of their energy due to pair production. $p+\gamma_{CMB} \rightarrow N+\pi$ is the mean free path for photo-pion production. $Fe+\gamma_{CMB} \rightarrow nucleus+n$ or $2n$ is the mean free path for spallation. $\gamma+\gamma_{CMB} \rightarrow e^++e^-$ is the mean free path for pair creation for photons with the CMB. $n \rightarrow p+e+\nu$ is the mean decay length for a neutron. Figure from [47].

cosmic ray astronomy. First, due to the GZK effect, the observed spectrum should not extend, except at greatly reduced flux, beyond about several times 10^{19} eV. This expected suppression in the energy flux is known as the *GZK suppression.* Nevertheless, it does not mean that no event could be detected above this energy. In fact, some of them have been detected and are known as *Super-GZK events.* Second, Super-GZK events must have a nearby origin, cosmologically speaking, closer than one hundred of Mpc, usually called the *GZK-sphere* or the *GZK-horizon.* Otherwise, their energy would have been reduced below the GZK threshold due to this effect.

Besides their interaction with cosmic photon backgrounds, charged particles are also affected by the presence of magnetic fields in the media they traverse. The intensity and topology of this fields is mostly unknown in the intergalactic medium but, if nG intensities are assumed, as a naive interpretation of Faraday rotation measurements would suggest, then protons at the highest energies could have gyroradii in excess of 100 Mpc. Therefore, charged cosmic rays originated at sources located at less that a few tens of Mpc, should keep enough directional information at Earth to produce observable anisotropy and make their astrophysical counterparts visible. The later opens the possibility of a charged particle astronomy in a similar sense as traditional photon-astronomy. At energies below few 10^{19} eV, the deflection due to the intergalactic and galactic magnetic fields combined is probably too large and only lower momenta of anisotropy can be expected even with very high statistics available. It is very likely that the same considerations apply, even at the highest energies, for nuclei heavier than protons. In any case, even for protons, only sources inside the GZK-sphere ($\lesssim$ 100 Mpc) could be explored due to the combined effects of intervening radiation backgrounds and magnetic fields.

It must be noted that some proposed exotic neutral particles, if they do exist, and neutrinos, if detected in enough quantities, may help probe a deeper portion of universe while keeping directional information.

2.2.3 Energy and composition

The energy spectrum of cosmic rays is almost featureless. Extending from 10^9 up to 10^{20} eV, the spectrum follows a power law $\frac{dN(E)}{dE} \propto E^{-\alpha}$, where the index α is almost constant and close to 3.0 in the whole energy range. The flux decreases 24 orders of magnitude along this energy range (Fig. 2.7). This behavior is expected in the case of stochastic acceleration of charged particles at astrophysical shocks as explained previously. The measured spectrum by the majority of the experiments is shown in Fig. 2.8.

However, there exist some deviations from this power law fall-off at high energies. They are more clear in Fig. 2.9 where the flux is multiplied by $E^{2.7}$. The first change is at $\sim 3 \cdot 10^{15}$ eV and it is called the *knee*. A second possible steeping, the *second knee*, occurs at energies around $\sim 5 \cdot 10^{17}$ eV. Another break where the spectrum turns up again,

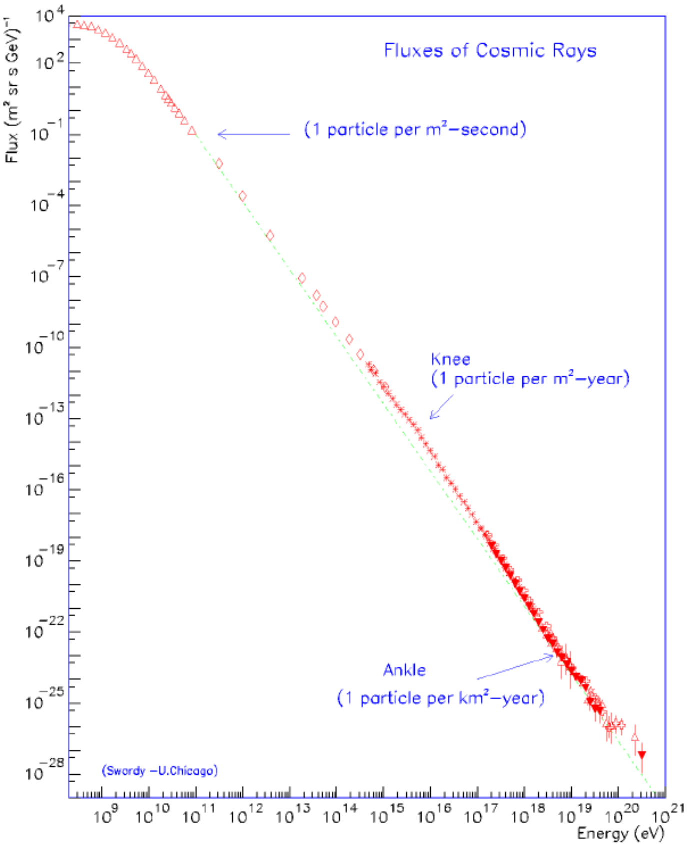

Figure 2.7: Energy spectrum of cosmic rays.

usually called the *ankle*, occurs at $\sim 3 \cdot 10^{18}$ eV. The other key point is the highest energy region around and beyond the GZK-threshold. The spectral features might be interpreted as a change in the acceleration mechanism at the sources, as a propagation effect or as a change in composition or in the hadronic interaction processes involved. As can be seen in Fig. 2.9 several disagreements exist at the highest energies between different experiments: the normalization of the flux, the position of the ankle and the existence or not of the GZK suppression. We discuss these energy regions in detail in the following.

Galactic cosmic rays: the knee

The knee, where the spectral index α increases significantly from 2.7 to about 3.1, is considered to be related with the limit of acceleration of lighter cosmic rays in the Galaxy.

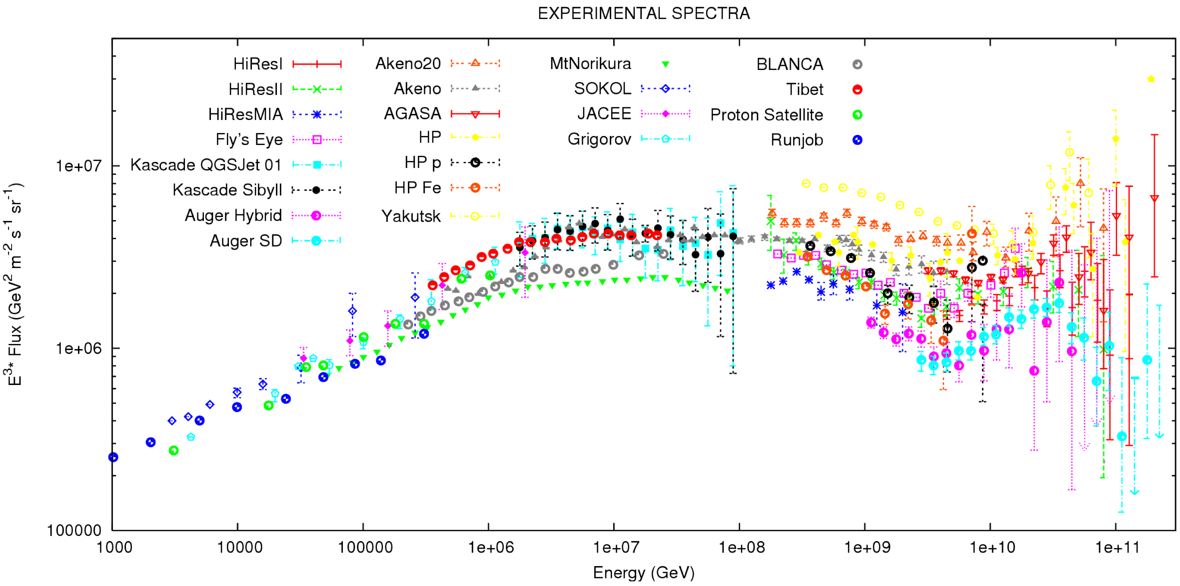

Figure 2.8: The flux of cosmic rays as determined by the majority of the experiments. Vertical axis is multiplied by E^3. Taken from [35].

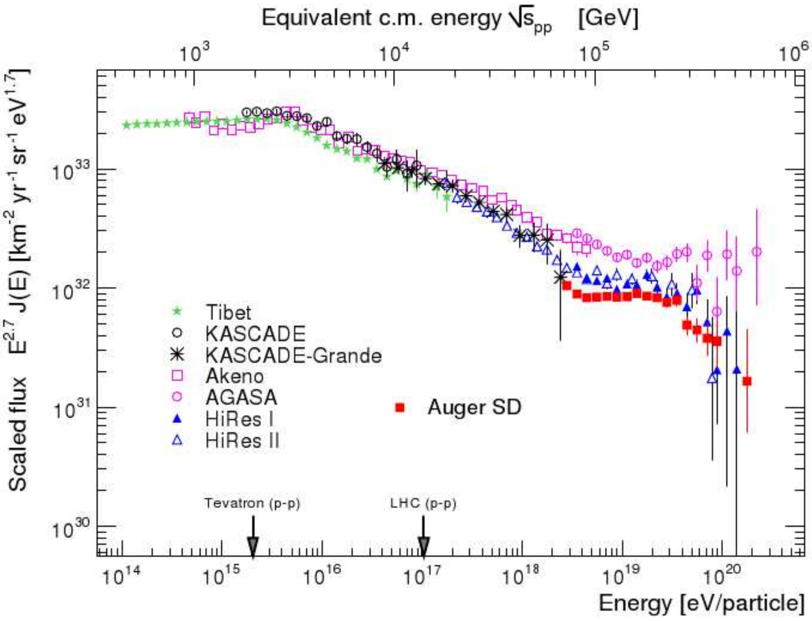

Figure 2.9: The flux of UHECRs as determined by the several experiments. Vertical axis is multiplied by $E^{2.7}$ in order to make more visible the different features.

As it was shown in Eq. 2.1 the maximum acceleration is proportional to the atomic number of the element. In these models the energy achievable by nuclei is rigidity[3] dependent. The knees of the spectrum of nuclei of charge Z are related to the proton knee energy $E^{Z}_{knee} = ZE^{p}_{knee}$, where $E^{p}_{knee} \simeq 3 \cdot 10^{15}$ eV. Beyond the highest energy knee $E^{Fe}_{knee} \sim \cdot 10^{17}$ eV, the total galactic flux, which is dominated by the Iron component, must be steeper.

Possible acceleration sites are supernova remnants (SNR) [48]. It could be also related to the limitation of the galactic magnetic fields to bind the nuclei into the Galaxy when they reach these energies. Particles produced in SNR will be confined until a certain energy at which their Larmor radius becomes comparable to the size of the Galaxy. At this point particles will be able to leak out of the Galaxy producing a break in the spectrum that has been identified in the knee structure (details in [49]). This is assume as the standard scenario.

However, other possibilities have been proposed. Another scenario assumes that the knee might be caused by a sudden change of the hadronic interactions at these energies [50]. In this case, the knee observed is not a characteristic of the spectrum itself, but of its observation at Earth. Finally, the knee can be interpreted also as a propagation effect due to a change in the regime of diffusion in the galactic magnetic field [51]. For a complete review about the observation and theoretical models for the interpretation of the knee see [52].

Transition region from galactic to extragalactic cosmic rays: from the knee to the ankle

Assuming the standard scenario, the transition region from the galactic to extragalactic origin of cosmic rays occurs between the knee and the ankle. Thus, a drop of the heavy components at an energy scaled with the charge is expected. As commented before, if the knee is caused by light elements, another knee-like feature would be observed for the heaviest elements at higher energy. This is a possible explanation for the second knee where the spectrum steepens to $\alpha \sim 3.3$. This second knee should be in the region between

---

[3] The rigidity is the momentum of the element over its charge.

10^{17} to 10^{18} eV. It has been observed by the Fly's Eye [53] and Akeno [54] experiments, but if it exists or not is still not clear. It is also being searched by KASCADE-Grande experiment [55].

The ankle, where the spectrum turns up with spectral index $\alpha \sim 2.7$, was first found by AGASA at 10^{19} eV [56], in agreement with Yakutsk [57]. However, it has been observed at around $3 \cdot 10^{18}$ eV by Fly's Eye [58], Haverah Park [59], Hires [30] and Auger [29].

Different interpretations of the transition region of the spectrum, between the knee and the ankle, have been proposed, with the consequent predictions on the cosmic rays composition. The standard interpretation for the ankle, called *the ankle model* (Fig. 2.10-left), is that the flat extragalactic component crosses the steep galactic spectrum, generating this feature at 10^{19} eV where the two components contribute equally to the total flux [60, 61]. The extragalactic component is thought to have a pure proton composition, so that the main problem in this model is how to explain a heavier composition up to 10^{19} eV.

An alternative explanation is the *dip model*, recently proposed [62, 63] (Fig. 2.10-right). It is build from the hypothesis that the extragalactic component, that is composed mainly of protons, starts to dominate at lower energies and the transition from the galactic to extragalactic CRs takes place at around $5 \cdot 10^{17}$ eV (second knee), so that in the ankle region the galactic component has already vanished. The spectral index change in the ankle is just a propagation effect: protons passing through the cosmic microwave background loose energy via e^-/e^+ production and this causes a flux suppression at higher energies and an accumulation at a slightly lower energy.

If composition above 10^{19} eV is not proton-like, another model called *mixed composition model* [64] is favored, because it assumes that extragalactic CRs have a mixed composition as the galactic component. As in the ankle model, the intersection of the galactic and extragalactic components gives origin to the the dip structure, but with the advantage of a lower transition energy at around $E \sim 3 \cdot 10^{18}$ eV (Fig. 2.11), which softens the requirement of additional acceleration mechanisms and is more compatible with recent results by HiRes and Auger. The predicted spectrum and mass composition depends

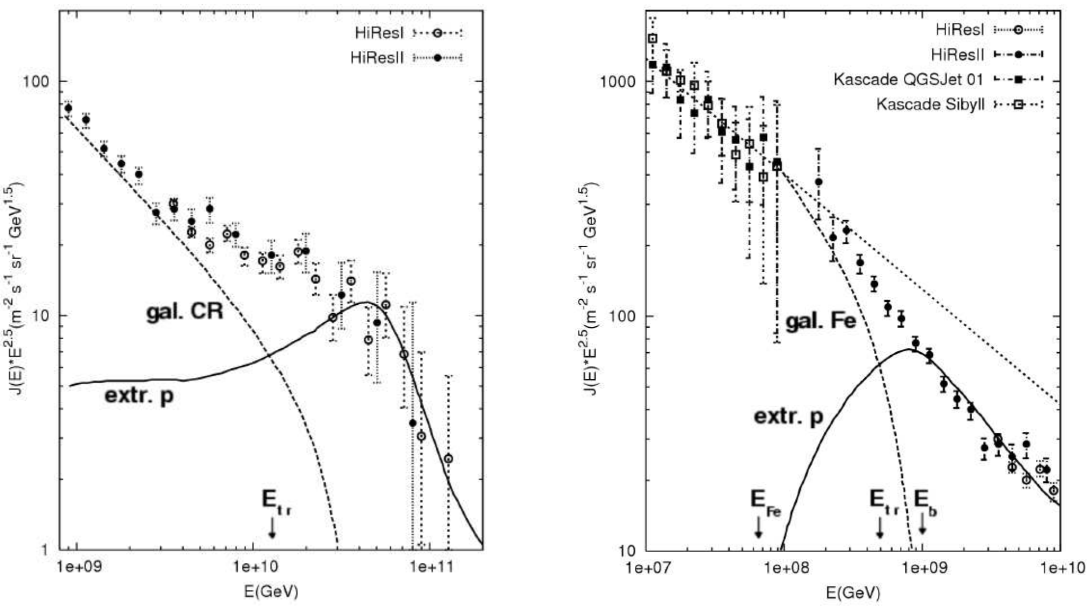

Figure 2.10: Transition models: ankle model (left) and dip model (right). In the left panel, the extragalactic proton spectrum and the galactic component are shown, and the transition energy E_{tr} is around 10^{19} eV . In the right panel, the extragalactic proton spectrum and the galactic component (dominated by Iron nuclei above E_{Fe}) are shown, as well as the transition energy E_{tr} that in this model is at the second knee. Data of KASCADE and HiResI, HiResII monocular spectra are shown. Taken from [35].

on several parameters (cosmological and describing the source composition), making the model very flexible and able to reproduce many composition profiles.

The three models could be experimentally distinguished through accurate measurements of the spectrum, although the most discriminant feature is the chemical composition. Since in the ankle model the transition takes place at around 10^{19} eV, the galactic heavy component dominates up to the ankle energy. At higher energies the extragalactic component begins to dominate and the composition becomes proton dominated. Consequently, the composition in the dip region is dominated by heavy nuclei. On the contrary, in the dip model, as the transition is completed at energy around 10^{18} eV, the composition in the ankle region is proton dominated. The composition in the ankle region (Iron/proton) is a strong discriminant between both models.

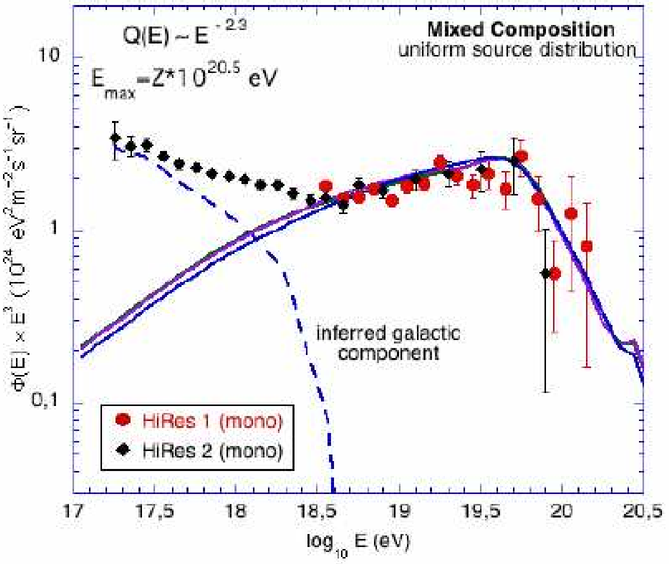

Figure 2.11: The mixed composition model for the transition region of the spectrum under certain conditions. At $E > 4 \cdot 10^{19}$ the spectrum is characterized by GZK suppression. At energy $3 \cdot 10^{18}$ eV the transition to pure extragalactic component is completed. HiResI, HiResII monocular spectra are shown. Taken from [35].

In the case of the mixed model, the transition between galactic and extragalactic (mixed composition) component occurs at $3 \cdot 10^{18}$ eV . Consequently in the dip region the chemical composition is mixed, while at lower energies the Galactic heavy component dominates. This model predicts a slower decrease of the Iron component and a slower increase of the proton fraction in the transition energy range. At higher energies, due to photo-disintegration of the nuclei, the composition get lighter and at $E > 3 \cdot 10^{19}$ eV becomes strongly proton-dominated.

A full discussion about the transition region could be seen at [65, 35].

The GZK suppression

At higher energies, there were a controversy about the existence or not of the GZK suppression. AGASA reported a continuation of the cosmic ray flux in form of a power law [56], while HiRes observed a suppression above $\sim 5 \cdot 10^{19}$ eV. Both use different techniques for detection and suffer a rapid increase with energy of the systematic uncertainties, mainly due to the lack of statistics. The dispute has been settled down by the recent results of the HiRes and Pierre Auger Observatory collaborations, where the flux suppression is determined with 5 and 6 standard deviations of significance respectively [30, 29].

Beyond the GZK

The trans-GZK complex is affected by acceleration mechanisms, chemical composition of particles, cosmological evolution of the accelerating objects, and the existence of new physics. First, the theoretical upper limit is set by the product of the size of the objects and strength of the magnetic field in it, as shown in the Hillas diagram (Fig. 2.3). Therefore, if extreme energy particles are accelerated by the bottom-up processes in these known astronomical objects, it is highly likely that acceleration limit should be around 10^{20} eV, and a deep cut-off should exist in the energy spectrum. However, if the existence of the measured suppression is due to the GZK effect instead of a limit in the acceleration processes in the sources, it would exist a recovery in the spectrum around $3 \cdot 10^{20}$ eV as a consequence of particles coming from sources closer to the GZK-horizon (see Fig. 2.12). In addition, in this scenario a bump would exist in the flux just at energies where the GZK begins, as a consequence of higher energetic particles that interact with the CMB photons and degrade their energy to that point (the cross section of this interaction is much lower below the GZK-threshold, see Fig. 2.6).

Unfortunately, the discovery of the recovery would not solve this issue, because the recovery would also exists if the acceleration limit is still higher than the GZK energy. If this is the case, the existence of new categories of unknown objects located in the blank region at the upper right corner of the Hillas diagram is strongly suggested, otherwise, the top-down scenario must hold.

Additionally, chemical composition of particles affects the shape of the trans-GZK complex, since it shifts in energy as the nucleus mass grows up. On the other hand, if we know the trans-GZK complex in detail, we could get some information on the chemical composition of the particles. If protons dominate and nucleus are negligible in

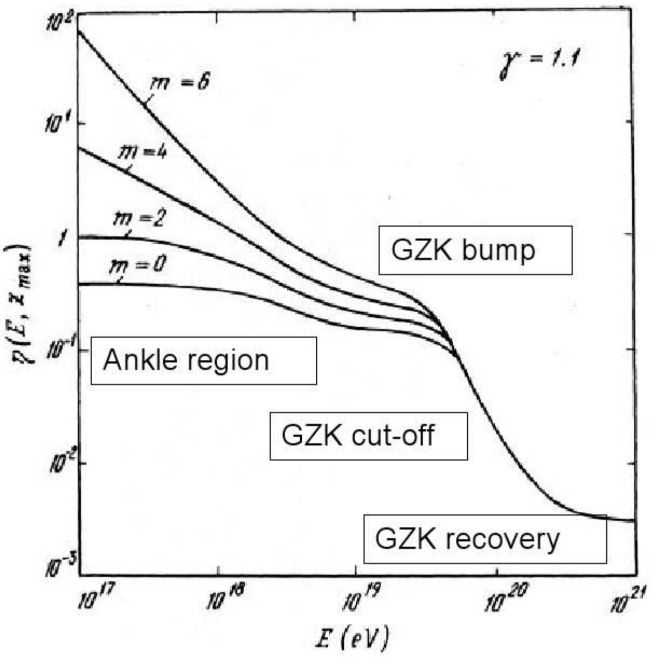

Figure 2.12: Theoretically predicted modification function of the spectrum of extreme energy particles due to the effect of propagation through the space. Here, m is a parameter that represents the degree of evolution of the sources. Taken from [66].

the chemical composition of the extreme energy particles, such a composition is difficult to be explained by the bottom-up scenario; it could be an evidence of the top-down scenario. If the nucleus component is comparable to solar abundance, it is the evidence of the acceleration in objects with standard chemical composition, such as galaxies including their nuclei. If nucleus components are more abundant compared to the solar abundance, it is the evidence of acceleration in metal-rich environment such as supernovae or gamma-ray bursts.

Ground based experiments, such as Auger and Telescope Array, have major problems to perform the analysis described above, mainly because the observation area of such experiments is too small to get high enough statistics. The trans-GZK profile will be measured by the JEM-EUSO experiment (still in *R&D* phase), which achieves by far (700 times of AGASA) the large exposure required by observing from space [7].

The field of cosmic rays still represent a big challenge for physicists, both from theoretical and experimental point of view. Many people is working on this area (around 1000 papers were submitted to the International Cosmic Ray Conference that took place is Mexico in 2007), and important advances have been performed in last decade (see Section 3.5 where Auger results are shown). However, as it has been shown along this Section, most of the questions are only partially answered and the puzzle is still uncompleted.

2.3 Extensive air showers

2.3.1 Phenomenology

The flux of cosmic rays at energies higher than 100 TeV is so low that direct measurements are not useful. Fortunately, when a cosmic ray come into Earth, it interacts with a nucleus of the atmosphere, mainly Nitrogen or Oxygen, producing an extensive air shower (EAS). The first interaction is hadronic and usually takes place in the upper atmosphere (20-30 km depending on the energy and the mass of the primary). The particles can undergo due to all kind of nuclear reactions leading to the production of nuclear fragments and secondary particles. The primary energy is shared among these secondaries and, due to the enormous amount of energy available, they have a large probability to interact with other nuclei in the atmosphere and produce new particles before decaying into (mainly) photons, muons, electrons and neutrinos. Therefore, the secondary products again interact with the molecules of the atmosphere, emitting further secondary particles. In this way the particle multiplicity is increasing dramatically, leading to several million or billion of secondary particles which are heading towards the Earth's surface with almost the speed of light. This cascade process is repeated until the energy of the secondary particles reaches the energy thresholds of the different processes involved. Then, the particles are absorbed, mainly by ionization, so the number of particles is reduced. To sum up, the number of particles increases, reaches a maximum at a certain depth in the atmosphere (called X_{max}), and decreases later.

The theory of electromagnetic interactions is assumed to be still valid at these energies, but the hadronic interactions are more problematic because they require far extrapolations of empirical models tuned on experimental data at lower energy. In fact, data from accelerators is used but the energetic range studied is limited; the center of mass energy in a nucleon-nucleon collision is given by $\sqrt{s} \sim \sqrt{2m_nE}$. For example, the energetic limit foreseen for LHC ($\sim$ 14 TeV) corresponds to a nucleon energy of 10^{17} eV . The needed extrapolations induce uncertainties on the first steps of the cascade, which cannot be directly observed.

Nevertheless, simple models describing the development of the showers exist. The toy model suggested by Heitler [67] provides the macroscopic characteristics of an electromagnetic showers, while the development of hadronic showers induced by protons could be described by a similar model [68]. Even if they cannot replace detailed simulations, these simple models predict the most important features of the cascades.

The nucleus-air interactions could be described by applying the superposition model. Assuming that the incoming projectile is a nucleon or a nucleus with atomic number A (in practice $A \leq 56$, because nuclei heavier than iron are not abundant), the primary interaction is hadronic and, as a first approximation, a nucleus A of energy E is equivalent to the superposition of A independent nucleons each with energy E/A (the binding energy of nucleons is $\sim$ 8 MeV, consequently at high energies they can be considered as free nucleons).

In Fig. 2.13-left a sketch of the different components of an extensive air shower is shown and the energy flow between different air shower components is in Fig. 2.13-right. In more detail, the different interactions involved are explained next.

The primary interaction produces a large number of secondaries, mainly pions (π^0, π^+, π^-) and kaons (K^+, K^-), which give rise to further hadronic interactions, and so on: this is the *hadronic cascade*. This component of the shower is also fed, in a low fraction, by the photons via two pion photo-production processes $(\gamma + \gamma \rightarrow \pi^+ + \pi^-)$ in the presence of a nucleus and the muon when collides with a nucleon. The lifetime of neutral pions is $\tau_0 = 8.4 \cdot 10^{-17}$ s, so that they immediately decay into photons before interacting.

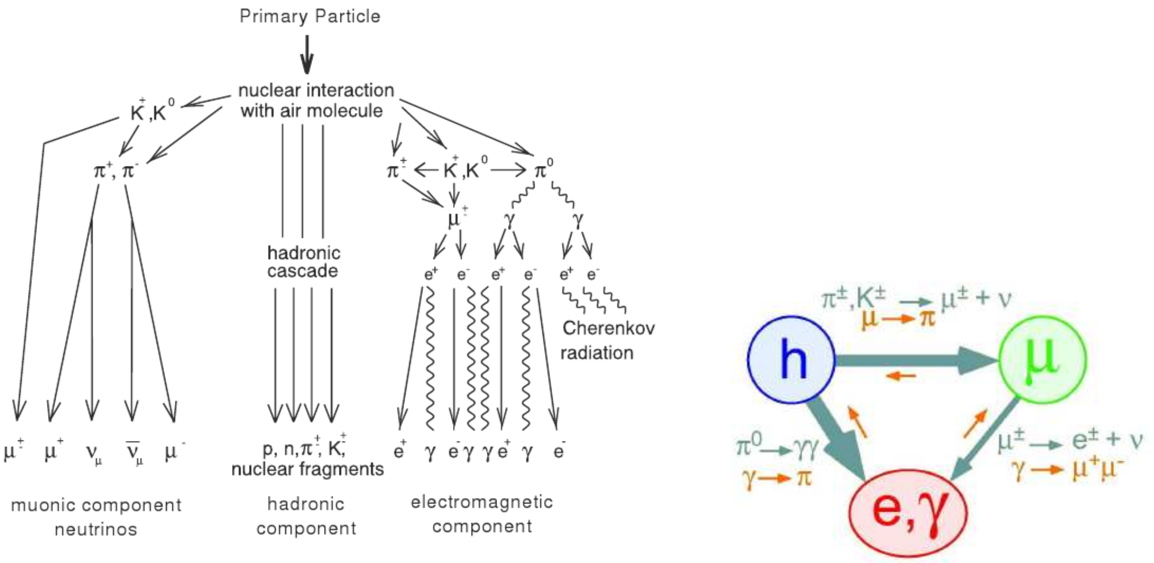

Figure 2.13: Left: Sketch of the different components in an extensive air shower. Right: Energy flow between different air shower components. The thickness of the arrows illustrates the amount of energy transferred in the given direction by the stated processes (from [69]).

Therefore, at each step of the hadronic cascade, about 1/3 of the energy is transferred to photons, giving raise to the *electromagnetic cascade*:

$$\pi^0 \rightarrow \gamma + \gamma \quad (\sim 98.8\%) \tag{2.5}$$

$$\pi^0 \rightarrow e^+ + e^- + \gamma \quad (\sim 1.2\%) \tag{2.6}$$

where the branching ratios of the two decay channels are given in the brackets. The hadronic cascade ends up with the decay of charged pions into muons (as will be shown later), at intermediate altitudes around 6 km with large spread. Therefore, only few nucleons, pions, kaons and nuclear fragments (the hadronic cascade) reach the ground and very colimated with shower axis.

Photons produce e^+/e^- pairs and Compton electrons, and electrons/positrons radiate through bremsstrahlung on atmospheric nuclei which leads to the emission of further photons, which afterwards again may produce additional e^+/e^- pairs, and so on:

$$\gamma + N \rightarrow N + e^+ + e^- \tag{2.7}$$

$$e^\pm + N \rightarrow N + e^\pm + \gamma \tag{2.8}$$

This chain reaction proceeds until the energy of the electrons and positrons drops below the critical energy[4], which is around 85 MeV in air. Then, the ionization energy loss starts to dominate the Bremsstrahlung process. Pure electromagnetic cascades can also be initiated directly by high energy photons or electrons. A small fraction of the electromagnetic component is re-injected in the hadronic cascade because of hadronic interactions of the photons as shown below. As consequence, a shower induced by a primary photon would also develop an hadronic cascade.

The decay of secondary pions and kaons generate neutrinos and the *muon component* via:

$$\pi^\pm \rightarrow \mu^\pm + \nu_\mu(\overline{\nu_\mu}) \quad (\sim 99.99\%) \quad (\tau_0 = 2.6 \cdot 10^{-8}\ s) \tag{2.9}$$

$$K^\pm \rightarrow \mu^\pm + \nu_\mu(\overline{\nu_\mu}) \quad (\sim 63.54\%) \quad (\tau_0 = 1.2 \cdot 10^{-8}\ s) \tag{2.10}$$

$$K^\pm \rightarrow \pi^\pm + \pi^0 \quad (\sim 20.68\%) \tag{2.11}$$

The muons are produced with typical energy of few GeV, increasing with the altitude of production. In addition, they inherit the transverse momentum of their parents (a few hundred of MeV), so their divergence[5] is relatively small and strongly anti-correlated to their energy (more energetic muons are close to shower axis). Few muons are also produced via the electromagnetic interactions of photons ($\gamma + \gamma \rightarrow \mu^+ + \mu^-$). A large fraction of muons reach the ground before decaying, with a non-negligible energy loss ($\simeq 2MeV/(gcm^{-2})$). Muons, however, may decay in flight when their energy drops below 10 GeV, producing a second source of neutrinos and e^+/e^-:

$$\mu^\pm \rightarrow e^\pm + \nu_e(\overline{\nu_e}) + \nu_\mu(\overline{\nu_\mu}) \quad (\sim 99\%) \tag{2.12}$$

---

[4] The critical energy E_c for the electrons is defined as the energy at which the loss by ionization ($\propto Z$) equals the loss by radiation ($\propto Z^2$). $E_c \propto 1/Z$

[5] the angle respect to the shower axis

As shown, in the decay of pions, kaons and muons, a great number of neutrinos/antineutrinos are produced. They are usually called *atmospheric neutrinos*.

2.3.2 Longitudinal development

Fluorescence telescopes detect the fluorescence light emitted by Nitrogen molecules excited by the passage of an EAS. Most of the fluorescence light is emitted in the near ultraviolet, between 300 and 400 nm. Oxygen molecules do not emit a significant amount of fluorescence light in this range. The amount of fluorescence light emitted is proportional to the energy lost by the electromagnetic shower particles. The proportionality constant is called the fluorescence yield, which is measured by different experiments in the wavelength bands of interest for EAS experiments [70, 71, 72]. The excited Nitrogen molecules rapidly de-excite such that the fluorescence light is emitted very close to the actual shower particles. The fluorescence light is emitted isotropically.

The longitudinal profile is the number of particles of a shower as a function of the amount of matter crossed, and it is observed with the fluorescence telescopes. In order to measure the total thickness of air crossed by the shower, the atmospheric slant depth is used. The atmospheric slant depth at a given height z measured from the ground level, depends on the shower geometry and on the atmosphere density $\rho(z)$:

$$X(z) = \int_{z}^{\infty} \rho(z') \frac{dz'}{cos\theta} \tag{2.13}$$

where θ is the zenith angle and the density dependence on the altitude is parametrized using atmospheric models.

The general shape of the longitudinal profile is shown in Fig. 2.14. It grows up while the energy of the particles is high enough to produce more particles, it reaches a maximum and later decreases because the energy of the particles is too low to generate more and they are absorbed mainly by ionization. The longitudinal profile may be adequately parametrized by the Gaisser-Hillas function [73]:

$$N(X) = N_{max} \left(\frac{X - X_0}{X_{max} - X_0} \right)^{\frac{X_{max} - X_0}{\lambda}} \exp \left(\frac{X_{max} - X_0}{\lambda} \right) \tag{2.14}$$

where N(X) is the number of the shower charged particles (mainly electrons/positrons) at an atmospheric depth X, X_{max} is the atmospheric depth at which the number of particles of the shower reaches its maximum N_{max}, X_0 is the depth corresponding to the first interaction of the primary particle in the atmosphere and λ is the interaction length. $X_{max} - X_0$ depends on the energy and the nature of the primary and it is an indicator of the stage of evolution usually called the *age* of the shower. X_{max} increases logarithmically with energy: $X_{max} \simeq X_i + 55 \cdot log_{10}(E)\ gcm^{-2}$, where the value of X_i depends on the nature of the primary (an average value is 700 gcm^{-2}) [74]. In practice, at a given energy, $X_{max}(Fe) - X_{max}(p) \simeq 100\ gcm^{-2}$. This is the essential feature allowing a discrimination between protons, light and heavy nuclei. In our energy range, X_{max} is of the order of 700 to 800 gcm^{-2}, that is less than the total vertical thickness of the atmosphere[6]. Then, at any zenith angle, the maximum of the shower is above the ground. The trend of X_{max} as a function of energy is called the *elongation rate*, as it is a very useful tool to study the composition of UHECRs (see Section 3.5.2).

Once the fit of the longitudinal development is performed using the Gaisser-Hillas function, the energy of the electromagnetic component is calculated by integrating the shower profile:

$$E_{em} = 2.2\,MeV/(gcm^{-2}) \int_0^{\infty} N(X)dX \tag{2.15}$$

where 2.2 $MeV/(gcm^{-2})$ is the average energy deposited by one charged particle of the cascade in the atmosphere in each depth interval of 1 gcm^{-2} [75]. This value is the critical energy of electrons divided by its radiation length in air ($\sim$ 37.1 gcm^{-2}). A crucial property to develop fluorescence detectors is the fact that they are only sensitive to the energy deposited in the atmosphere by the electromagnetic component and most of the energy is finally carried by this component. In [68] it is shown that, for example, for primaries around 10^{17} eV the energy carried by the electromagnetic component is around 90% of the primary energy. From this point of view, the atmosphere behaves as a giant calorimeter with a good linearity.

Nevertheless, not all the energy is going into the fluorescence light, since part is carried

---

[6] atmospheric depth at sea level is around 1030 gcm^{-2}

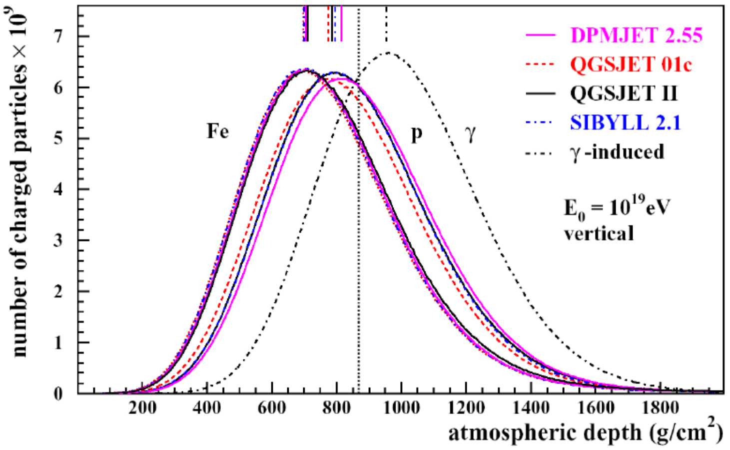

Figure 2.14: Simulated longitudinal profiles for proton, Iron and photon primaries with an initial energy of 10^{19} eV and arriving at a zenith angle $\theta = 0^o$. Figure taken from [78].

away by neutrinos, muons and hadrons. This contribution is known as the *missing energy* since it is invisible to the current generation of fluorescence telescopes. This fraction of missing energy has been parametrized as a function of the electromagnetic energy using shower simulations in [76] and, for example, it amounts to about 10% at 10^{19} eV. This correction is model and primary dependent and it could be included in the algorithms to determine the primary energy in fluorescence experiments by modifying the factor 2.2 by 2.65 $MeV/(gcm^{-2})$ in Eq. 2.15 [77].

2.3.3 Lateral development

Whatever the physical process during the cascade development, the products generated have, in general, a moderate transverse momentum regardless their energy. Then, most high energy particles are colimated along the initial axis, constituting the *core* of the shower. The lateral extension of the core depends on the mean free path, so that it is proportional to the inverse of the density. It may be expressed in terms of the Moliere radius r_M, such that 90% of the energy is contained within a distance from the shower axis $r < r_M$ (in atmospheric showers $r_M \simeq 60 - 100m$). However, low energy photons and electrons, as well as muons, extend far away from core, constituting the *halo* which has a detectable density up to a few kilometers from the axis depending on the primary energy.

The electromagnetic part of the halo increases with the depth, reaching its maximum at around $X_{max} + 100\ g/cm^2$ and then decreasing rapidly (it is completely extinguished at $X_{max} + 1000\ g/cm^2$). Most muons travel beyond the electromagnetic cascade, giving a *muonic tail*, with an increasing spread with core distance due to a simple straight-forward propagation. After a long range, lower energy muons (with larger divergence) decay. Independently of the electromagnetic cascade, the muons generate an *electromagnetic tail* through their decay (see Eq. 2.12), $\delta-$ray production and radiative processes (bremsstrahlung and pair production) which are important above few tens of GeV. $\delta-$ray production is characterized by very fast electrons produced by energetic charged particles as the muons, knocking orbiting electrons out of atoms. Collectively, these electrons are defined as delta radiation when they have sufficient energy to ionize further atoms through subsequent interactions on their own.

For high energetic events, the electromagnetic component dominates the region near the shower axis and the muonic component is more important in the far region. The density of particles for vertical showers with energies of 10 EeV and for different primaries is shown in Fig. 2.15.

The surface detector only samples the shower front of the air showers. The total number of particles reaching the ground has to be inferred by fitting the individual sampled densities to an assumed lateral distribution function. As the lateral distribution of particles is mainly determined by Coulomb scattering of the dominant electromagnetic component, it can be approximated by the Nishimura-Kamata-Greisen (NKG) function [79, 80]:

$$\rho(r) = k \left(\frac{r}{r_M}\right)^{-\beta} \left(1 + \frac{r}{r_M}\right)^{-(\beta-\gamma)} \tag{2.16}$$

where β and γ are parameters determined experimentally and k is proportional to the shower size. Many other functions have been used instead of the NKG one (see for example [81]). The lateral distribution is subject to fluctuations, from the detector response

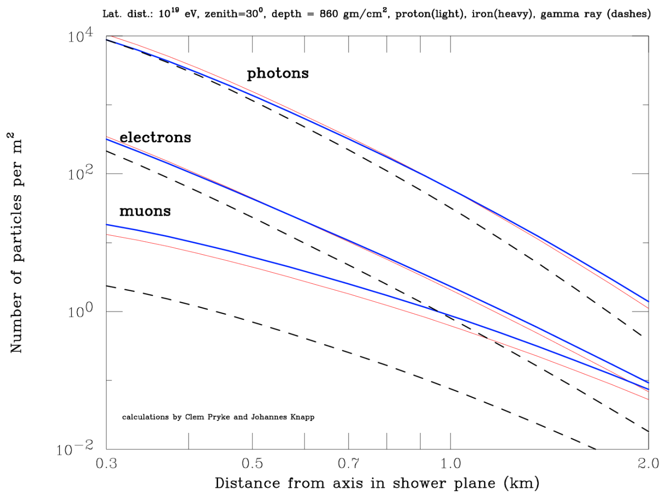

Figure 2.15: Simulation of the lateral distribution of shower particles. Curves are obtained averaging on many showers. The photon, electron and muon densities in the plane perpendicular to the shower axis are shown for shower induced by protons (red line), Iron nuclei (blue line) and gamma rays (dashed line) of energy 10^{19} eV.

itself and from shower-to-shower. In addition, uncertainties in core location would significantly modify the integral of this function and the signal measurements are poor at distances lower to the Moliere radius because the array size is usually larger and/or the detectors may saturate. Therefore, a procedure similar as the one used in the fluorescence technique may introduce huge uncertainties. Instead, Hillas proposed to use the density at a certain distance from the shower axis where these uncertainties minimized [82]. In surface experiments, a transformation from the signal at this distance into the energy is usually done by Monte-Carlo (MC) air shower simulations. Unfortunately, the lateral distribution of particles obtained from simulations does not describe perfectly the data [83, 84] and MC are based on hadronic interaction models that introduce important uncertainties. A complete study of the optimum distance to use as shower energy estimator is one of the goals of this thesis (Chapter 4). On the other hand, at Auger the conversion from the signal to the cosmic ray energy could be done independently of MC simulations, using coincident observations of the same air shower with the hybrid technique.

2.3.4 Detection Techniques

In this Section the different techniques for measuring the properties of the extensive air showers are discussed: an array of detectors located at ground level, the fluorescence technique and some comments about the radio and acoustic detection of UHECRs.

Surface arrays

Surface detectors (SD) are the most common way to measure UHECRs. The first SD was built by Pierre Auger in 1939 [85] using Geiger-Müller counters just after the discovery of the existence of the EAS. As this detection technique could not provide information about the direction of the shower, it has been replaced by scintillators or water Cherenkov tanks. A modern SD for the detection of UHECRs consists on an array of hundreds of radiation detector stations. The principle components of a SD station are a target material (generally either water or a plastic scintillator), a photomultiplier tube, and a data acquisition system. Radiation passing through a detector station causes a short burst of light in the target material which is amplified and converted to an electric signal by the photomultiplier tube. The signal level is a function of the number and type of particles which penetrate the detector. The passage of a shower front is detected when adjacent detector stations measure signals coincident in time.

Surface detectors may cover a large area. The area required depends on the rate of events which are expected to be detected, and for ultra-high energy cosmic rays, it must be several square kilometers. Larger the collecting area, higher is the energy range that could be explored. To increase the maximum energy that could be studied by 10 times, it is required a 1000 times increase in the collecting area (the flux falls as $\sim E^{-3}$). The separation of the detectors is chosen to match the scale of the footprint of the showers and it is usually of the order of several hundred meters. The low energy threshold of a SD is determined by the detector spacing because a lower energy shower has a smaller footprint on the ground and, thus, a smaller array would be required to resolve it. The size of the detectors is chosen appropriately for the shower component to be studied. It is generally

Table 2.1: Past, present, and planned UHECR surface detector experiments. Some information from [86].

Experiment	**Location**	**Area** (km^2)	**Number detectors**	**Detector type**	**Start**	**End**
Volcano Ranch	New Mexico	8	19	Scintillator	1959	1963
SUGAR	Australia	70	47	Scintillator	1968	1979
Haverah Park	United Kingdom	12	34	Water Cherenkov	1968	1987
Yakutsk	Siberia	10	68	Scintillator	1969	Running
AGASA	Japan	100	111	Scintillator	1990	2004
KASCADE	Germany	0.04	252	Several (see text)	1996	Running (2009)
Auger South	Argentina	3000	1600	Water Cherenkov	2004	Running
Telescope Array	Utah	760	576	Scintillator	2007	Running
Auger North	Colorado	$\sim$ 10000	$\sim$ 3900	Water Cherenkov	Planning Stage	

of the order of 10 m^2 for charged particles and ideally larger for muons. The vertical atmospheric depth of the SD site is selected according to the depth of shower maximum of the most energetic cosmic rays of interest. Otherwise, a higher energy EAS which has not yet reached its maximum will look very similar to a lower energy EAS which is past maximum.

The shower arrival direction is calculated from the station timing and the energy from the signal inferred at a certain distance as it was explained in previous Section. Several

SD arrays have been performed and the most significant are summarized in Table 2.1. Volcano Ranch [17] was the pioneer and used scintillators as detectors. A new scintillator array, the Telescope Array [87] is currently being commissioned. Haverah Park [5] was the first that used water Cherenkov tanks instead of scintillators. SUGAR [88] was the only array in the South Hemisphere before Auger. Several techniques like electron and muon counters together with a hadron calorimeter are used in the KASCADE [89] and KASCADE-Grande [55] experiments. Pierre Auger South Observatory is the largest ever built, covering around 3000 km^2 with 1600 water tanks. The most important advantage of SD arrays is that they could have nearly a 100% duty cycle and they are only mildly affected by atmospheric conditions such as temperature and pressure.

Fluorescence Telescopes

The fluorescence detector (FD) technique was first successfully used at the University of Utah in 1976 [90]. Fluorescence experiments are formed by an array of wide field ultraviolet telescopes. The optics typically consist on a fast large aperture primary mirror on a fixed mount. Sometimes a Schmidt correcting lenses are used. The camera is an array of phototubes.

An EAS is detected when there is a signal in several adjacent phototubes in the camera. The axis of an EAS is determined by the light track on the camera and the timing of the light pulses in the phototubes. As explained before, since electromagnetic particles dominate in an EAS, the energy of the primary cosmic ray is very nearly proportional to the total amount of fluorescence light produced. The largest systematic error in the FD energy determination comes from the uncertainty in the fluorescence yield parameter. The main advantages of FD are that the energy measurement does not rely on Monte-Carlo simulations that are model dependent, and that X_{max} is the most useful composition estimator.

A FD can only operate on clear dark nights. The moon is usually required to be below the horizon or less than half illuminated, and the clarity of the atmosphere including clouds must be closely monitored. Therefore, the duty cycle of FD is about 10% compared to

Table 2.2: Experiments based on the fluorescence technique . Some information from [86].

Experiment	Location	Start	End
Fly's Eye	Utah	1981	1992
HiRes	Utah	1998	2005
Auger South	Argentina	2004	Running
Telescope Array	Utah	2007	Running
Auger North	Colorado	Planning Stage	
JEM-EUSO	International Space Station	Planning Stage	

SD. Table 2.2 shows all UHECR experiments which have used or are planning to use the FD technique.

Radio and acoustic detection

An alternative technique is the detection of EAS making use of the radio frequency pulses generated by charged particles in the geomagnetic field [91]. Cosmic ray showers induce radio pulses through several mechanisms, but the dominant process is coherent synchrotron emission by the electron-positron pairs propagating in the Earth magnetic field. Radio pulses in the range 200 - 100 MHz induced by EAS were first measured in 1966 [92] but only recently this technique has been proposed as the next generation of detection.

The main advantages of radio detection compared to other techniques are that radio signals are not absorbed nor deflected on their path and the amplitude of the signal is proportional to the primary energy of the incoming particle. Moreover, the duty cycle is in principle 100% which guarantees the large data volume necessary for statistical analysis.

The radio technique has started a new generation of experiments as the Low-Frequency Array (LOFAR) [93]. LOFAR, with its 100 stations of 100 dipoles antennas distributed on an area of radius around 400 km, predicts to observe EAS up to 10^{20} eV at a rate of 1 event per year. Several other experiments are incorporating this new technique, as KASCADE, Icecube [94] and the Pierre Auger experiment. An array of about 20 km^2 to be deployed at the Pierre Auger site is currently in R&D phase [95, 96], and the main objectives are the investigation of the radio emission from an air shower at the highest energies, the exploration of the capability of the radio-detection technique, and composition and energy measurements between $10^{17.4}$ and $10^{18.7}$ eV, where the transition from galactic to extragalactic origin of cosmic rays is expected.

Another technique which is now being explored, especially in combination with large scale neutrino telescopes under the sea or ice, is based on acoustic sensors. They detect the signals produced by high-energy particle cascades which for short time heat the traversed volume. This effect, leading to a pressure pulse with amplitude dependent on primary energy, was suggested by Askaryan [97] in 1957. The absorption length for acoustic waves is one order of magnitude larger than that of Cerenkov radiation. Therefore, acoustic signals can be detected at larger distances ($\sim$ 1000 m) with respect to Cherenkov light. Current tests of detector prototypes along with studies of the background are underway at various sites, see [98] for a detailed review.

Chapter 3

Pierre Auger Observatory

3.1 Background and advantages of a hybrid detector

Previous experiments have brought some light in certain topics about UHECRs but they have added more intrigue too. For instance, the published spectra from the AGASA and HiRes collaborations show significant discrepancy between them: i) the spectra are different in a factor of 2 in flux or 30% in energy or a combination of both, with AGASA systematically higher, ii) the GZK steeping was detected by HiRes, while no suppression was found by AGASA, and iii) there were no coincidence in the measured position of the ankle (it is at 10^{19} eV for AGASA and at $10^{18.5}$ eV for HiRes). Each experiment uses a different detection technique which make the comparison more complex. Regarding cosmic rays origin, AGASA [99] and SUGAR [88] reported some excess in the region of the Galactic Center (GC), but these results were not completely accepted mainly due to the lack of statistics. Finally, the discrepancy is clear between all the experiments about the chemical composition of UHECRs (as will be shown later in the Fig. 3.11). Therefore, conclusions about the physics involved in cosmic ray production and propagation to Earth were also not reliable. A new advance from the experimental point of view were needed.

The Pierre Auger Observatory represents a big step towards revealing the mysteries of UHECRs. Previous experiments used the surface detectors (SD) or the fluorescence detectors (FD) in isolation. However, Auger is the first which combines both simultaneously. It is a *hybrid* detector. The Pierre Auger Observatory was originally designed as a pair of detector systems, one in the northern and one in the southern hemisphere giving full sky coverage. The northern site is still in the planning phase and it will be located in Southeast Colorado, USA [100]. The construction of the southern site was completed in 2008 and it is taking data since 2004 using a partial array. It is located on the Pampa Amarilla, near the town of Malargüe in Mendoza Province, Argentina, at an altitude of around 1400 m over the sea level. The site was selected because it is extremely flat (varying by less than 20 m in altitude from one end to the other), the altitude is suitable to detect UHECRs (875 gcm^{-2} of atmospheric depth, a bit larger than the maximum of the longitudinal development of the showers of interest) and the atmospheric conditions are appropriate (clear sky and low pollution).

As it was commented before, the main advantage of Auger is that it is a hybrid detector which involves several benefits. The fluorescence technique allows the measurement of the longitudinal development of the shower, so that the energy determination is almost model independent while, on the other hand, SD only can estimate the energy under assumptions on primary identity and on the hadronic interaction model for the shower development in the atmosphere. However, the FD has a small duty cycle of approximately 10% (the telescopes can only be used on clear moonless nights), while SD array is working almost 100% of the time providing huge statistics. Auger hybrid observatory can then be used to detect 10% of the showers with both techniques. Thus, the SD data with the energy scale set by cross calibration with the FD, could be used to measure anisotropy and the energy spectrum of CRs accurately with high statistics. In addition, a hybrid detector can measure lower energy showers than a SD or FD alone, because only one triggered surface detector is needed to perform a hybrid reconstruction properly (pure SD reconstruction would require at least 3 triggered detectors and pure FD reconstruction could suffer significant uncertainties in the geometrical fit). Furthermore, the determination of the fluorescence detector aperture is not straightforward, but the surface detector has a well-defined aperture independently of primary energy or weather conditions. Another benefit

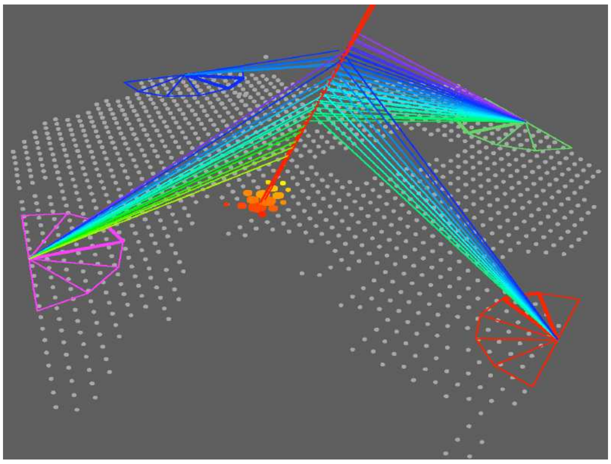

Figure 3.1: Picture of hybrid detection of an EAS in Auger by the SD array and the four telescopes eyes.

is that cross checks between FD and SD reconstructions are also useful to identify sources of systematic uncertainties.

The measurement of a cosmic ray shower with a single fluorescence telescope leads to large inaccuracies in the shower geometry, that translates into a poor measurement of the longitudinal profile and primary energy. The accuracy of the shower axis estimation for an FD event is greatly improved when SD stations are used in the reconstruction. The SD stations give the time and the place where the shower front passes through the ground level that is a strong constraint to the geometry. Consequently, geometric reconstruction using both detectors (hybrid reconstruction) is more accurate than using either detector by itself. Finally, with a hybrid detector the two best composition indicators could be measured, the shower muon content and the depth of shower maximum. FD measure directly and accurately the depth of shower maximum, while the SD measurements could do it only indirectly. On the other hand, the Auger SD is capable of measuring the shower muon content and many other composition sensitive parameters that could be defined as will be shown in Section 3.4. The muon content is not obvious to be determined from water Cherenkov tanks because, as will be commented later, they have a large sensitivity to the electromagnetic component of the shower, whose signal should be extracted from

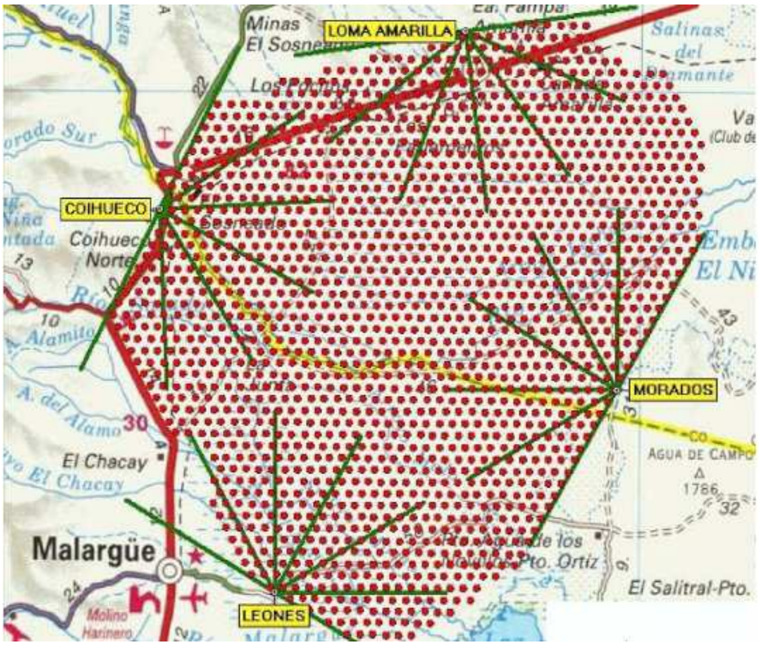

Figure 3.2: The layout of the Auger Observatory. Each red dot is the position of a SD station. The positions of the FD buildings are also shown. The green line segments show the field of view of the 24 FD telescopes.

the total one to get the muonic part. An extension of Auger called AMIGA [101], that consists on buried scintillators, is now being developed. It could determine accurately the muon content of the EAS.

The Auger layout is in Fig. 3.2 where each circle is an SD tank and the lines the field of view of the telescopes. Auger could give a boost to solve the problems related to UHECRs and, in fact, some discrepancies between previous results have been clarified. The Auger main results will be shown later in Section 3.5. Next, the surface and the fluorescence detectors are presented briefly.

3.2 Surface detectors

The Auger SD is a sparse array of 1600 water Cherenkov tanks or stations on a 3000 km^2 hexagonal grid. The spacing between each station is 1.5 km. The full layout of the SD

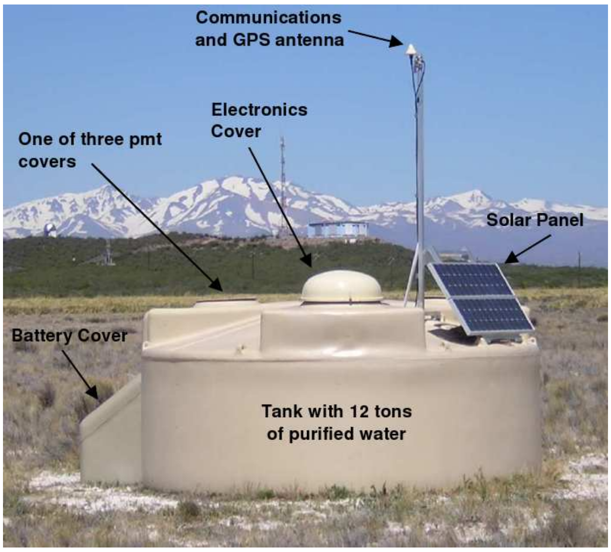

Figure 3.3: An Auger SD station. The Los Leones FD telescope building is in the background.

is shown in Fig. 3.2 and a single station in Fig. 3.3. Each water Cherenkov tank is 1.55 m high with a diameter of 3.6 m, giving a total surface area of $\sim 10\ m^2$. Twelve tons of high purity water are deposited in the tank filling it to a height of around 1.2 m.

The detectors measure the Cherenkov light emissions produced when charged particles pass through the pure water with a velocity greater than that of light in this medium. The particles at ground level in an EAS are mainly photons, electrons and muons. The electrons have mean energy of around 10 MeV and muons around 1 GeV. The incident photons interact via Compton scattering and pair production creating relativistic electrons which then produce Cherenkov light. The detectors are a few radiation lengths thick so that they absorb the majority of the electromagnetic component. However, muons pass straight through the tank, so that a signal proportional to the track length is recorded.

The light is detected by three photomultipliers. The inner part of the tank is covered with a Tyvek bag to increase the diffuse reflectivity of the walls for the Cherenkov UV-light. Due to the large distances, the detectors are controlled and read out by mobile phone technology. Furthermore, they have to be completely self-sustaining concerning the power supply which is achieved by solar panels and batteries. The time synchronization of the individual detectors needed to determine the incoming direction of the EAS is done by GPS signals.

Water Cherenkov tanks have an increased sensitivity to the electromagnetic portion of the EAS in comparison to scintillator based detectors, due to a greater response to photons. The detection of photons is important because, as it has been shown experimentally, at distances beyond ~ 100 m from the shower core, photons are 10 times more numerous than electrons (see Fig. 2.15). Another feature of water Cherenkov tanks is that they are more sensitive to inclined events than scintillator detectors, allowing approximately twice the sky coverage for a similar sized array.

The electronics of the surface detectors are explained in [102]. Detailed description about the trigger system and event reconstruction (including tank selection, geometrical reconstruction and the fit of the lateral distribution of particles) are given in Appendix B. Regarding the energy determination, the calibration and monitoring of surface detectors, details could be found in [103, 104, 105, 106].

3.3 Fluorescence detectors

The FD is composed of 24 telescopes disposed in groups of 6 in 4 buildings. Each building is located in one of the hills on the borders of the site, providing and extra elevation between 60 m and 200 m. Each telescope is placed on a separate bay and observes a different volume of air through a window on its bay. This window could be closed with a shutter to protect the camera from daylight. With the shutter closed, the number of background photons incident on the corresponding camera are about a factor of 5 smaller than the normal background during data taking. Shutters can be operated remotely from the main campus and closed automatically if an outside light sensor detects too much light. Since shutters close slowly, a set of fast dropping curtains were installed in front of the diaphragm of the telescopes that promptly protect the camera in case of too large illumination. The field of view of each telescope spans 30 degrees in azimuth and 28.6 degrees in elevation and overlooks the SD. Thus, the region of maximum development

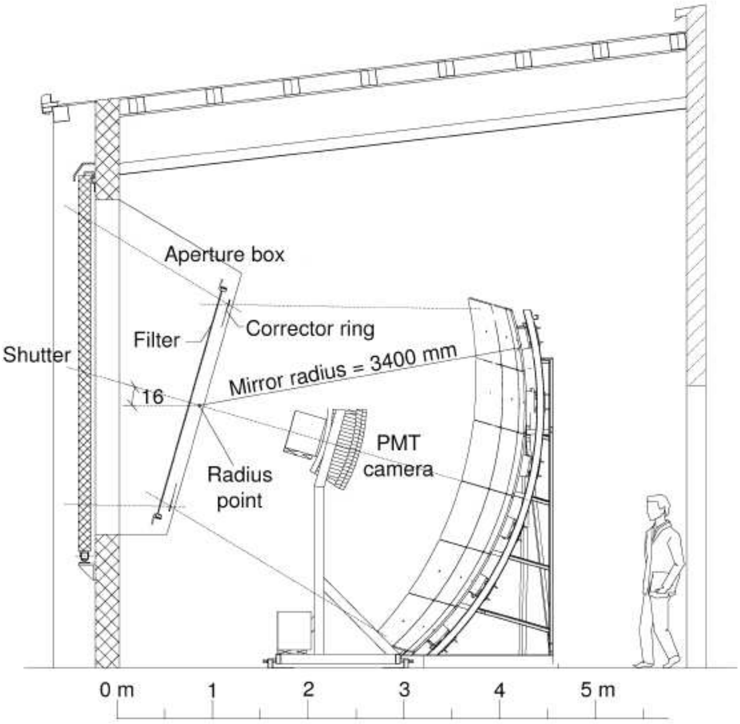

Figure 3.4: Sketch of the Fluorescence Telescope showing the spherical mirror, the photomultiplier camera and the UV-filter.

of the EAS from primaries of EeV energies is inside their field of view. The Auger Collaboration has decided to further expand its energy range down to 10^{17} eV after completion of the southern observatory, so that three additional fluorescence telescopes with an elevated field of view from 30^o to 60^o above the horizon have been installed. This enhancement is called High Elevation Auger Telescopes (HEAT) [107].

Fig. 3.4 illustrates the elements of the fluorescence telescopes. Each telescope is composed of a spherical mirror focusing light onto a camera, placed at the focal surface. Mirrors are about 3.5 m × 3.5 m in size. The camera is an array of 440 quasi-regular hexagonal pixels made of 22 rows and 20 columns. Each pixel has a size of 1.5^o and is instrumented with a photomultiplier tube (see Fig. 3.5). A circular diaphragm placed at the center of curvature of the mirror reduces the coma aberration and the addition of a set of lenses (called *corrector rings*) to control the spherical aberration allow for a

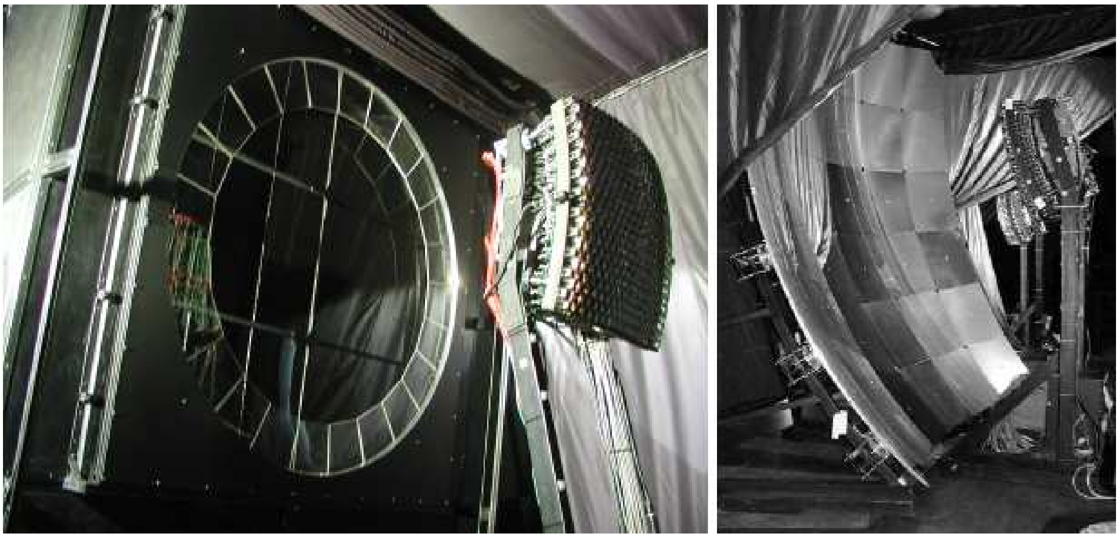

Figure 3.5: Camera, mirror and diaphragm with the corrector ring of the fluorescence telescopes.

total collecting area of 3 m^2, taking the shadow of the camera into account. This system results in a spot size of about 0.5 degrees. An UV filter is placed in front of the diaphragm. This filter transmits the UV fluorescence light while blocking most of the visible night-sky background. The filter also acts as a window allowing for climate control inside the buildings. Each telescope array building has a main computer to handle high level triggering and event data storage, a wireless communication link with the Central Data Acquisition System (CDAS) for event data transfer and SD triggering, and a GPS receiver to provide absolute timing information. Each camera has a dedicated data acquisition system.

Fluorescence telescopes are used to observe the longitudinal development of the shower via the fluorescence photons. As the EAS develops, atmospheric nitrogen (N_2 and N_2^+) is excited mainly by electrons and positrons below 1 GeV. The de-excitation results in the emission of photons with wavelengths of between 300-400 nm. The fluorescence photons can be gathered up from any point by appropriate telescopes. In addition, since most charged shower particles travel faster than the speed of light in air, Cherenkov light will be emitted during propagation as well. The Cherenkov contribution is proportional to the number of charged particles above a certain threshold and, since electromagnetic particles dominate the shower profile, this contribution is proportional to the number of shower electrons. In comparison with fluorescence isotropic light, Cherenkov emission in air is strongly forward peaked along the particle direction. A detailed description of the fluorescence and Cherenkov productions in EAS can be seen in [103].

The Auger FD do not operate if the moon is in the field of view of the telescopes or due to bad weather conditions (rain or clouds). Each operation period is called an FD shift ($\sim$ 18 days). The duty cycle of the Auger FD is approximately 13%.

The electronics of FD are explained in [108]. Details about the optics, trigger conditions, data acquisition, calibration and monitoring could be found in [103, 104, 106].

3.4 Composition Observables

This Section is devoted to explain the main composition observables used by the Pierre Auger Observatory. First, the parameters from the fluorescence technique are shown. Second, those obtained from the surface detectors information are explained. In Chapter 5 of this thesis a new surface parameter is presented.

3.4.1 Parameters from the fluorescence technique

From the fluorescence telescopes two are the parameters traditionally used: the position of the maximum in the longitudinal development, X_{max}, and the number of particles at this maximum, N_{max}. Both are obtained from the Gaisser-Hillas fit (see Eq. 2.14). However, N_{max} has not been used much as a composition estimator (an example is [109]), but it has been also proposed as an energy estimator [110]. The point of first interaction in the atmosphere will be probably the best composition indicator but, unfortunately, it is not measurable. Therefore, we focus on X_{max}.

The average value $< X_{max} >$ at a certain energy E is related to the mean logarithmic mass $< lnA >$ via:

$$< X_{max} >= D_p \left[ln(E/E_0) - lnA \right] + c_p \tag{3.1}$$

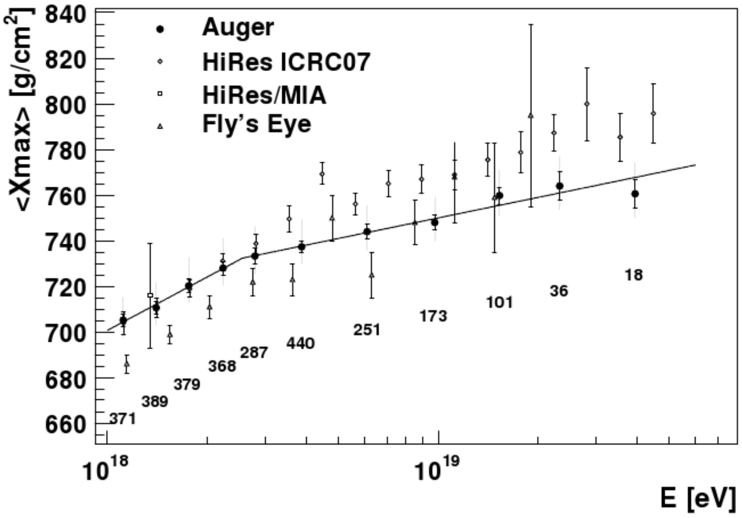

Figure 3.6: Elongation Rate from different experiments: Fly's Eye, HiRes-MIA, HiRes and Auger (ICRC 2009). The numbers below denote the events in each bin of the Auger data and the line is a fit of the elongation rate with two linear functions. The error bars denote the statistical uncertainty. Taken from [116].

where D_p denotes the *elongation rate* of a proton, and c_p is the average depth of a proton with reference energy E_0. The elongation rate is defined as the rate of change of X_{max} with the logarithm of the energy $D_p \equiv \frac{dX_{max}}{dlog_{10}(E)}$ and it was first proposed in [111]. Both, D_p and c_p, depend on the nature of the hadronic interactions. The elongation rate comparing data and simulations is the best technique nowadays to determine the composition trend of UHECRs. In order to ensure a good elongation rate, the X_{max} resolution must be around 20 gcm^{-2}, which is achieved by Auger [112], and the estimated uncertainties of the shower maximum and total energy must be smaller than 40 gcm^{-2} and 20% respectively. In Fig. 3.6 the elongation rate determined by Fly's Eye [113], HiRes-MIA [114], HiRes [115] and Auger [116] are shown.

The width of the X_{max} distribution is another composition sensitive parameter [116, 117] since heavy nuclei are expected to produce smaller shower-to-shower fluctuations than protons. The latest results about the elongation rate and the RMS of the X_{max} distribution by Auger will be shown and discussed later in Section 3.5.

3.4.2 Parameters from the surface detectors

Surface detectors make a discrete sampling of the shower front at ground level, so that the main SD observables are related to the temporal and spatial distribution of particles. The most important between them are summarized next.

Muon number N_μ:

It is the most promising parameter, essentially because the differences between iron and proton are predicted to be very significant. To evaluate quantitatively the relative difference in the muon production in heavy nucleus showers vs. protons showers, first note that the muon production in proton showers increases with energy as $E^{0.85}$ [118]. Considering a nucleus initiated shower of energy E as a superposition of A showers, each with energy E/A, the total number of muons is $N_\mu^A \propto A(E/A)^{0.85}$, or comparing to proton showers $N_\mu^A = A^{0.15} N_\mu^p$. Thus, it is expected that an iron shower (A=56) will produce about 80% more muons than a proton shower of the same energy.

The muon content is easily determined from scintillators buried at ground or covered with a shield, so that absorbs the electromagnetic component is absorbed. However, it is not straightforward but not from water Cherenkov tanks which, additionally, are very sensitive to the photon component of the shower. Therefore, for the latter is more common to use the muon to electromagnetic ratio [119] because both components are measured efficiently by these detectors. However, several methods are now being tested [120, 121] to determine directly the number of muons too. AMIGA is designed to measure accurately the muon content of the EAS [101, 122]. In [123] it is showed that most of the SD mass sensitive parameters depend strongly on X_{max} and N_μ.

Rise time:

The spread of the arrival time of the shower particles at a fixed core distance increases for smaller production heights, so the rise time is expected to be smaller for heavy primaries that develop higher in the atmosphere. This is the geometrical reason that relates the rise time with the stage of the shower development, and hence, with X_{max}. A sketch is shown in Fig. 3.7, where it is assumed that the particles travel un-scattered after production in the cascade which is almost true for the rise time since the initial part of the signal is dominated by muons which travel almost straightforward.

Here, the steps followed by Auger in [124] to obtain the rise time are presented. First, the rise time of a single station is defined as the time it takes to increase from 10% to 50% of the total signal deposited in the station. For non vertical showers a slight dependence of the rise time on the internal azimuth angle of the stations within the shower plane is expected. The internal azimuth angle is defined as the clockwise angle between the projection of the shower axis on the ground and the line connecting the shower core and the station. The reason is that the shower has to traverses more atmosphere to reach the stations on the exterior side (or late part) of the shower compared to those on the interior side (or early part) of it. Thus, the observed rise time of each station is corrected depending on its internal azimuth angle and the zenith angle of the shower (more details in [124]). The rise time at a fix distance from the shower axis (usually $r_0 = 1000m$), $t_{1/2}(r_0)$, is finally obtained by fitting the corrected rise time of each triggered station using the function $t_{1/2}(r) = (40 + ar + br^2)$ ns. Parameters a and b are free in the fit, and the function is anchored at 40 ns at $r = 0$ because this is the mean single particle response in the water Cherenkov detectors. Only stations in the range from 600 to 1500 m from the shower axis and signal greater than 10 VEM (Vertical Equivalent Muon, i. e. the signal deposited by one vertical muon in an Auger water tank) are included in the fit in order to avoid signals dominated by large fluctuations. At least three stations are needed in the fit.

In addition, great advance has been performed in the study of rise time in last years, mainly by the Leeds and Karlsrühe Auger groups. First, trace cleaning in FADC traces is applied. The water-Cherenkov tanks used in the surface array are susceptible to the unavoidable background noise produced by isolated low energy cosmic rays. These can cause peaks in the FADC trace which are not deposited energy associated with the EAS. Trace cleaning is the term applied to the process of identifying and removing these parts of the signal. Tanks far from the shower core, and therefore with smaller signal sizes,

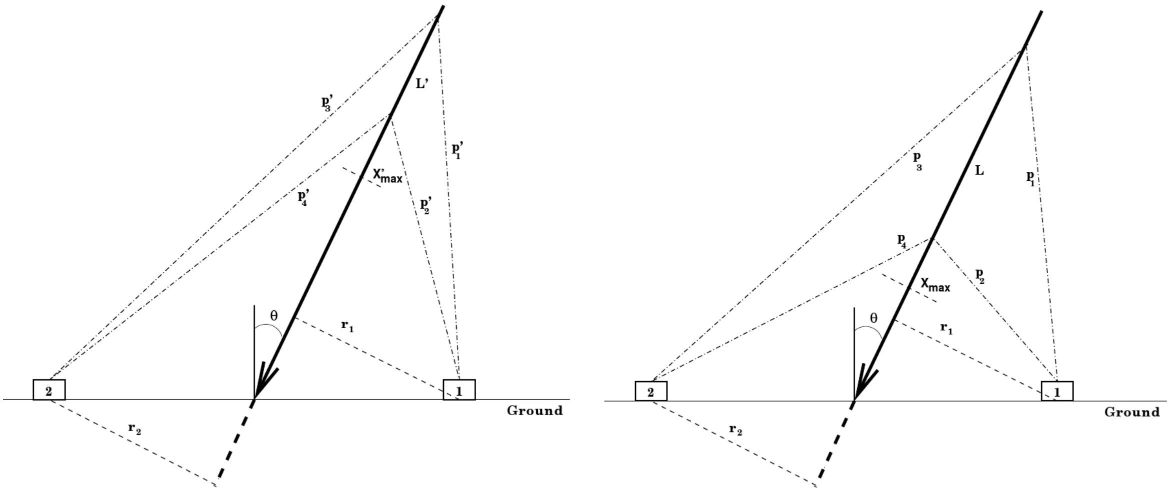

Figure 3.7: Geometrical illustration of the correlation between the longitudinal development of a shower and the time structure measured in surface detectors. Left: a shower developing at small atmospheric depths. Right: a cascade developing at larger column densities. The difference in path lengths p_1' and p_2' is much smaller than for p_1 and p_2. This is also true for the differences between p_3' and p_4' with p_3 and p_4. Consequently, the distribution in arrival times and thus the signals recorded from EAS at large values of X_{max} are expected to be longer than for smaller values. Taken from [125].

are most susceptible to these background effects. Second, a deconvolution algorithm is performed. The time structure of the signal recorded when particles are detected in surface tanks is artificially lengthened due to detector effects. This is a consequence of the physical size of the tanks and the bandwidth limitations of the electronics which is corrected by the deconvolution algorithm. Some improvements have been also done in the asymmetry correction explained before and in order to consider temperature and pressure effects. Details about these procedures could be found in [126, 127].

Currently, the tendency is try to measure X_{max} indirectly from water Cherenkov detectors. The most investigated method is via the rise time [128, 129, 130]. The method consists on determining the average value of the rise time as a function of the core distance (r) and the zenith angle (θ) for a given reference energy (10^{19} eV), the so-called benchmark. Then, for each selected detector in a given event, the deviation of the measured

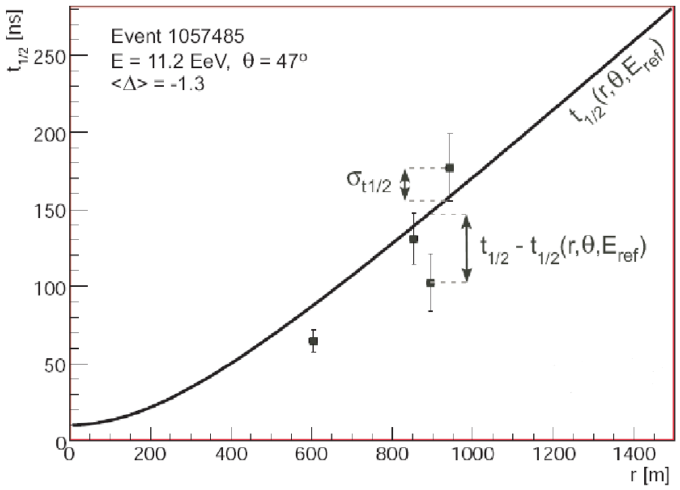

Figure 3.8: Rise time vs distance to the core. The curve is the benchmark rise time and the data points represent the measurements of rise time of each detector with uncertainties for this particular event. Figure from [131].

rise time from the benchmark function is calculated in units of measurement uncertainty and averaged for all detectors in the event as shown in Eq. 3.2 and Fig. 3.8, enabling a new observable, $< \Delta_i >$ to be introduced,

$$< \Delta_i >= \frac{1}{N} \sum_{i=1}^{N} \frac{t^i_{1/2} - t_{1/2}(r, \theta, E_{ref})}{\sigma^i_{1/2}(\theta, r, S)} \tag{3.2}$$

where $\sigma^i_{1/2}(\theta, r, S)$ stands for the uncertainty parametrized as a function of zenith angle, distance to the core and signal (S) of each detector. The $< \Delta_i >$ are expected to be larger for showers developing deeper in the atmosphere than the reference rise time. The results obtained with this parameter will be shown later in Section 3.5.

Other time parameters:

In addition to the rise time, many other time parameters could be defined. For example the *fall time*, defined as the time that the signal takes to increase from 50% to 90% of the total one deposited in the station, and many other combinations changing these percentages. t10-30 is the most useful one and it uses the first portion of the signal. Following the procedure previously shown for the rise time, another Δ parameter could be defined. The t10-30 is more muon dominated and then may show smaller fluctuations and it is less sensitive to the asymmetry effect. An analysis could be found in [132].

Radius of Curvature:

The curvature of the shower front is the distance from the first interaction point to the impact point of the shower. The shower can be approximated as a developing sphere from its starting point in time and space. Away from the shower core, the first particles to reach the ground are muons as they propagate unaffected by scattering. Because the majority of muons are created relatively early in the shower and at large distances from the ground, the shower front can be modeled as the surface of a sphere formed by the geometrical paths of the muons produced at a point source. After the muon production stage finishes and the shower propagates to the ground, this surface expands, so the radius of the sphere increases and becomes thinner due to the absorption of the electromagnetic particles.

Since the shower front curvature is related to the initial point of the shower, it is also to the depth of the shower maximum. For example, an iron primary has a shorter mean free path in air than proton or photon primaries. Then, the initial point of the shower is higher in the atmosphere, resulting in a shallower X_{max} and a longer radius of the shower front curvature. Photon showers develop closer to the ground (deeper X_{max}) and have a smaller radius of curvature. A sketch is shown inf Fig. 3.9. The width of the shower front has also been proposed as a mass composition parameter.

The muon content is also related to the shower front curvature through the arrival times at ground of the shower particles. Electromagnetic particles are scattered multiple times, so that they must travel farther through the atmosphere and take longer to reach the ground when compared with muons which travel in straight lines. Thus, farther from the core, if the shower is more electromagnetic (light primaries), the relative delay is larger than for muon rich showers (heavy primaries), resulting in a larger reconstructed shower front curvature (lower radius).

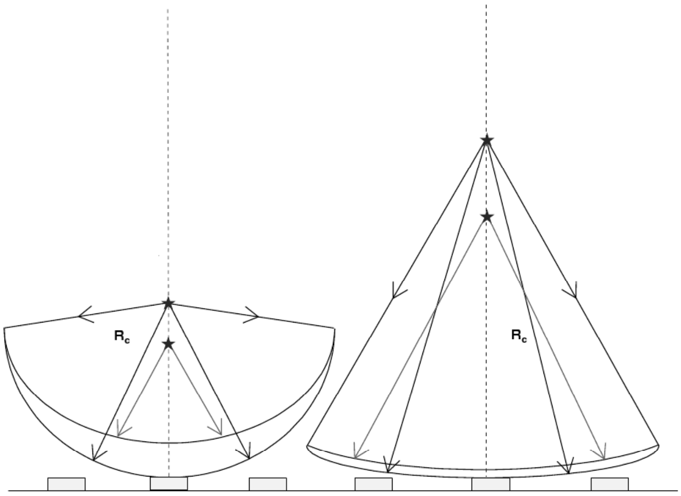

Figure 3.9: A schematic demonstration of the radius of curvature as a mass sensitive parameter. The diagram on the left shows a shower which develops late in the atmosphere as is expected for low mass primary particles. The radius of the sphere associated with particles from the first interaction is small. The diagram on the right is used to show a shower which develops much earlier in the atmosphere, typical for heavier primaries, and it has a much larger radius than previously. Therefore, one would expect larger radii of curvature to indicate increasing changes in the average mass composition of UHECRs. Taken from [126].

The slope of the lateral distribution of particles:

The lateral distribution of shower particles in a cascade arises mainly from the distribution in transverse momentum of the pions, which decay produce the electromagnetic component of the shower and the lateral spread of this component due to multiple scattering.

The density of particles in the shower falls off more rapidly with distance in late developing showers. In deep showers (light primaries) much of the electromagnetic component created close to the core arrives at detection level, but becomes more and more attenuated at larger distances, so that the density of particles decreases accordingly. When a heavy nuclei is the primary, the cascade develops earlier, so that the low energy electromagnetic component is more attenuated even at distances relatively close to the core compared to cascade initiated by light primaries. Consequently, for heavy primaries the gradient in the distribution of showers particles is flatter.

This parameter has been used by several experiments, such as Volcano Ranch [133] and Haverah Park [134], for composition determination of UHECRs. Other approach is for example in [135]. However, at Auger the slope of the lateral distribution function is fix in the fit using a parametrization obtained from data which depends on the signal at 1000 meters from shower core S(1000) and the zenith angle of the shower (see Appendix B), so it is not very useful for composition analysis.

Azimuthal asymmetry on the rise time:

Evidence of azimuthal asymmetries in the signal size were first observed at Haverah Park [136] and asymmetries in the time structure of the signals were found for the first time in the Pierre Auger Observatory [137]. In fact, the observed azimuthal asymmetry of the time distributions from the signal of inclined showers (incoming direction with zenith angle larger than 60^o) is an unique feature of the Pierre Auger Observatory which carries very valuable information related with the chemical composition of cosmic rays.

This method does not allow a measurement of the primary mass for individual showers. Instead, average values are obtained for each data sample. Next, the procedure is summarized. For every sample of events in a given energy and zenith angle bin, the mean value of the rise time divided by the core distance $< t_{1/2}/r >$, using the stations that fulfilled some required cuts, is fitted to a functional form $< t_{1/2}/r > (\zeta) = a + bcos\zeta$, where ζ is the azimuth angle in the shower plane. The evolution of b/a with zenith angle is an indicator of the shower development, and then, the $sec(\theta)$ value at which b/a reaches its maximum, the so-called XAsymMax (see Fig. 3.10), is expected to be correlated with X_{max}. XAsymMax has a low sensitivity to the uncertainty in the reconstructed energy and, in addition, the difference between iron and proton does not depend on the absolute number of muons at ground. It mainly depends on the evolution of the asymmetry of the electromagnetic component of the shower. Details about the procedure could be found in [138] and how to infer X_{max} from it in [139]. The results obtained with this parameter

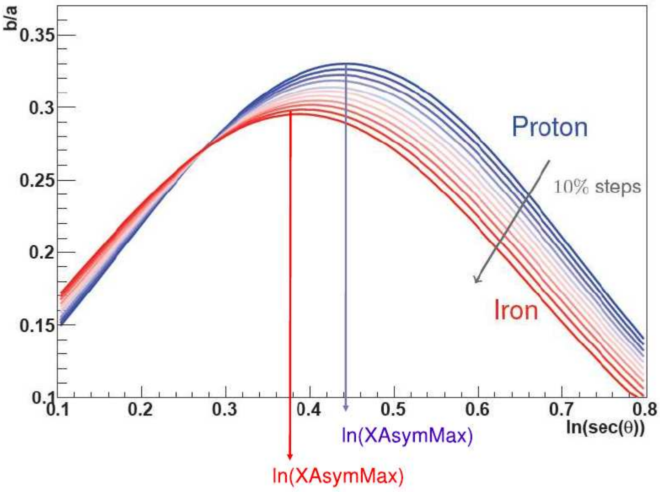

Figure 3.10: Asymmetry development for the different samples with mixed composition, going from pure proton to pure iron in steps of 10%. The positions of the maximum for the different primaries are marked. Figure from [131].

will be shown later in Section 3.5.

Muon Production Distance:

In [140], it was shown that the arrival time structure observed in muons is a transformation of the muon production distance (MPD) distribution. The basic idea comes from the fact that muons travel following straight lines which start in a narrow neighborhood of the shower axis. The more parallel the trajectories to the shower axis are, the less time it takes for the particles to reach ground. When the observation point at ground is fixed, the muons produced at small (large) MPD will deviate more (less) from shower axis and, therefore, they will travel more (less) distance and, hence, will have more (less) delay. In [141, 142] is shown how to determine the distance along the shower axis where muon was produced from the time stamp of muons in a detector array.

S_b:

This is a new family of surface parameters that is presented in this thesis in Chapter 5, where it is mainly applied to the Pierre Auger South Observatory.

3.4.3 Why the issue of composition is so difficult?

Composition is probably the most difficult issue from the experimental point from view. Two are the most important problems to identify the primary particle from the extensive air shower it produces when traverses the atmosphere whatever the parameter used: the large shower to shower fluctuations and the uncertainties about the hadronic interaction models at such high energies.

The experimental observables are limited by the intrinsic fluctuations. Regarding X_{max}, simulations show that proton primaries with energies near 10^{19} eV produce an X_{max} distribution with mean near 700 gcm^{-2}, while iron primaries of the same energy will have an X_{max} mean of 80 to 100 gcm^{-2} shallower. Because an iron nucleus produces an EAS which is basically a superposition of 56 lower energy showers, the fluctuations of X_{max} around the mean for iron are smaller than for protons, with Iron fluctuations having a standard deviation of near 30 gcm^{-2} and protons 70 gcm^{-2}. These fluctuations are of the same magnitude as the difference in the mean X_{max} of both primaries. Similar behavior is expected for any parameter considered, essentially, because it is a consequence of the intrinsic fluctuations in the shower development which translate into any observable.

The uncertainties in the hadronic interaction models (HIM) add more disadvantages. At least three main families of HIMs are used: QGSJet [143, 144], Sibyll [145] and EPOS [146] (discussion can be found in [147] and [148]). In UHECRs the relevant center of mass energy at 10^{20} eV is beyond the one at the LHC, so the systematic uncertainty introduced by the extrapolation of the hadronic interaction models tuned at much lower energies is extremely difficult to quantify. In addition, a significant discrepancy in the number of muons produced in the EAS development exists between models [149, 150].

Additionally, other problems appear related with shower simulations. First, artificial fluctuations are introduced by the thinning [151] and unthinning [152] algorithms that are needed to perform the simulations without huge CPU time and disk space. Second, additional dispersion in the sensitive composition estimators is caused when the showers are re-used, a common practice in composition studies due to the long CPU time required for shower simulations [153].

Finally, another difficulty arises due to the uncertainties in energy determination. It causes that showers of different energy were binned together, reducing the discrimination power of the composition sensitive parameters.

The difficulty is clear when comparing results published by different experiments. Different methods of measuring the primary mass give different answers and the interpretations are usually dependent on the hadronic model assumed. The Pierre Auger Observatory by measuring the elongation rate, concludes using QGSJET-II that their data suggest a moderate lightening of primary CRs at low energies and an almost constant composition at high energies, whereas the EPOS elongation rate is clearly larger than the measured one at high energies, which would indicate a transition from light to heavy elements [154] (see Fig. 3.14-top, more details in Section 3.5.2). HiRes using stereo measurements on X_{max} suggests light composition above 10^{18} eV and a change from heavy to light composition in the range $10^{17} - 10^{18}$ eV. Studying fluctuations in X_{max} above 10^{18} eV they obtained 80% protons for QGSJet and 60% for Sibyll [155]. AGASA used the number of muons at ground to determine that their data is marginally consistent with a transition between heavy and light composition between an energy range of $2 \cdot 10^{17}$ to $2 \cdot 10^{19}$ eV [156] and determined an upper limit for the iron fraction of 35% in the range $10^{19} - 10^{19.5}$ eV and 76% in the range $10^{19.5} - 10^{20}$ eV [157]. Even more disappointing is the fact that Volcano Ranch and Haverah Park, using both the slope of their lateral distribution function and with the same hadronic model used for the interpretation, are on disagreement. While Haverah Park has reported an iron fraction around 66% in the energy range 0.2-1 EeV [134], Volcano Ranch reported 79% iron fraction from $10^{17.7}$ to 10^{19} eV [133]. Some of these results are shown together in Fig. 3.11 where the inferred iron fraction from 10^{17} to 10^{20} eV is shown as determined by different experiments.

Just one point seems to be clear, photons represent a very low fraction of primaries at the highest energies, as it has been demonstrated by Auger [159, 124] (Section 3.5.2).

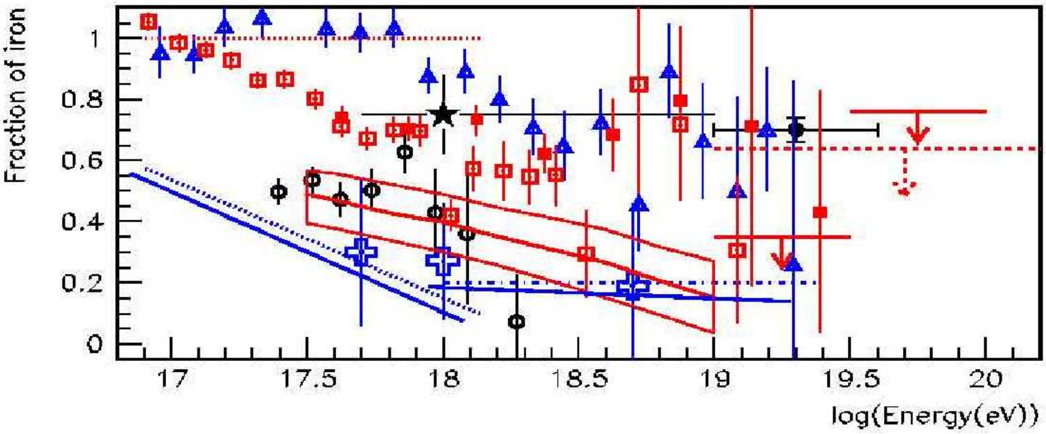

Figure 3.11: Iron fraction from various experiments. Muon densities (red): AGASA A1 (solid line with band) and upper limits above 10^{19} eV (solid lines) with QGSJET98; AGASA A1 (squares) and A100 (filled squares) deduced using Mocca-SIBYLL; Yakutsk upper limit (dashed line) using QGSJET01. MIA (dotted line) QGSJET98 iron fraction estimate. Geometrically-based methods (black): Volcano Ranch LDF (star), Haverah Park LDF (circles) and rise time (point) with QGSJET01. X_{max} (blue): HiRes (dot dash line) and Yakutsk (crosses) derived from fluctuation in the X_{max} distributions. HiRes (triangles) using Mocca-SIBYLL, HiRes-MIA (dotted line) using QGSJET98, HiRes (solid lines) using QGSJET01 deduced from X_{max} mean values. Figure taken from [158].

Another way to deal with this problematic issue has been to apply different techniques for data analysis. Most are statistical methods, such as the Principal Component Analysis, Linear Discriminant Analysis [160], Multiparametric Topological Analysis [109] or Likelihood methods [161]. Interesting efforts have been performed using Neural Networks as well [162].

3.5 Pierre Auger South Observatory: Results

In this Section the main results published by the Pierre Auger Observatory are presented. A brief comment and the corresponding references are given.

3.5.1 UHECRs: spectrum

Confirmation of the GZK suppression

In [163] the Pierre Auger Observatory reports the measurement of the energy spectrum of cosmic rays above $2.5 \cdot 10^{18}$ eV derived from 20.000 events recorded between 1 January 2004 and 31 August 2007. Hybrid calibration was performed to SD data [164]. Systematic uncertainties on the energy scale due to the calibration procedure are 7% at 10^{19} eV and 15% at 10^{20} eV, while a 22% systematic uncertainty in the absolute energy scale comes from the FD energy measurement. The spectral index γ of the particle flux, $J \propto E^{-\gamma}$, at energies between $4 \cdot 10^{18}$ eV and $3 \cdot 10^{19}$ eV is 2.69 ± 0.02(stat) $\pm$ 0.06(syst), steepening to 4.2 ± 0.4(stat) $\pm$ 0.06(syst) at higher energies. The hypothesis of a single power law is rejected with a significance greater than 6 standard deviations (6σ). The result is shown in Fig. 3.12.

Several astrophysical implications of the measured spectrum exist. If a mixed composition model is assumed at the sources with nuclear abundances similar to those of the low-energy galactic cosmic rays, good agreement is found down to energies close to the energy of the ankle, where another component (possibly an extra-galactic one) emerges. Another set of models (the ankle and dip models, Section 2.2.3) which assume only proton primaries is tested. To reproduce Auger spectrum by these models, a much stronger source evolution would be needed. The elongation rate observed by the FD also disfavors the pure proton assumption as it will be shown later. A hypothetical model of a pure iron composition is able to describe the measured spectrum above the ankle under certain conditions, below which an additional component is required. More details in [165, 166].

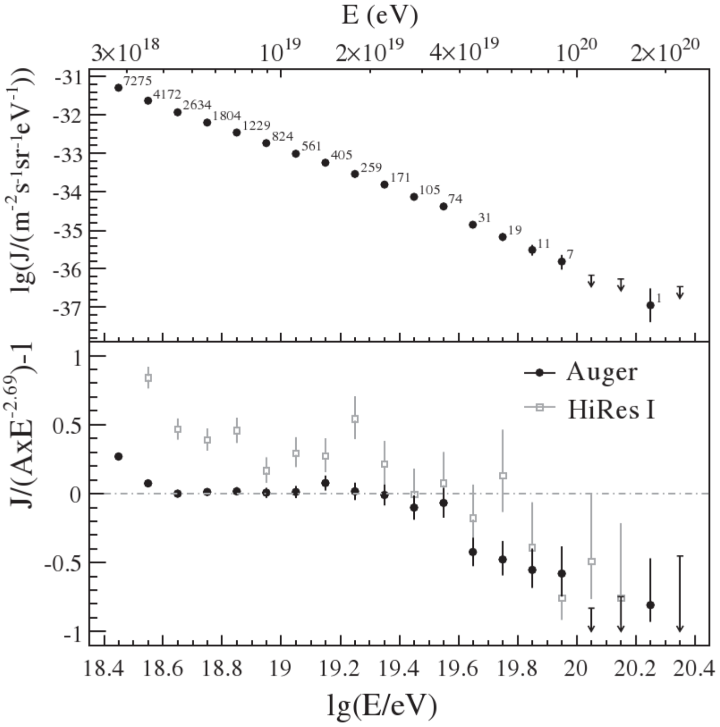

Figure 3.12: Upper panel: The differential flux J as a function of energy, with statistical uncertainties. Lower Panel: The fractional differences between Auger and HiRes I compared with a spectrum with an index of 2.69. Taken from [163].

Agreement between vertical SD, inclined SD and hybrid spectra

At the ICRC in Mexico (2007), Auger presented three independent measurements of the energy spectrum [165]. One is based on the high statistics provided by the surface detector data, using only vertical showers (below 60^o). The second uses almost horizontal showers detected by the SD and, the third, considers the hybrid data where the precision of the fluorescence measurements is enhanced by additional information from the surface array. The complementarity of the three approaches is emphasized and results are compared. Each spectrum implies different reconstruction methods. An agreement between them is shown in Fig. 3.13. The existence of the ankle at $\sim 10^{18.5}$ eV and the existence of the GZK suppression were confirmed. These results have been recently updated with higher statistics at the ICRC in Lodz (2009), where the spectrum obtained from surface data

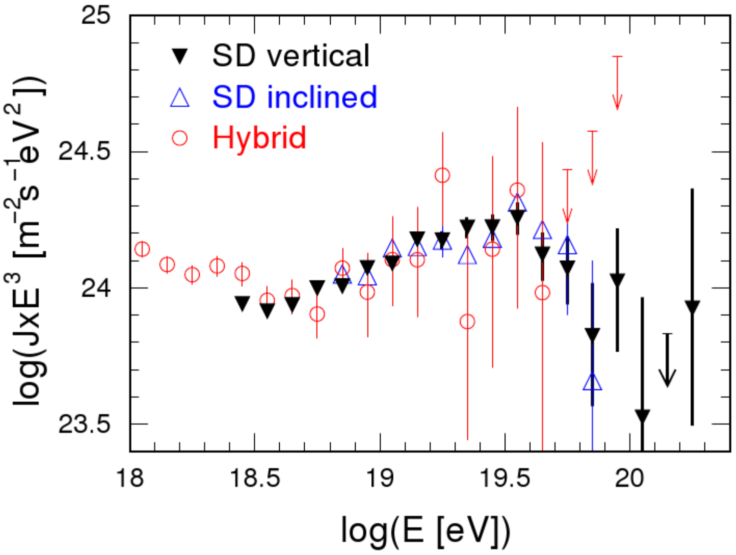

Figure 3.13: The energy spectrum multiplied by E^3 derived from SD using showers at zenith angles above (filled triangles) and below (opened triangles) 60^o, together with the spectrum derived from the hybrid data set (red circles). Arrows indicate 84% CL upper-limits. Taken from [165]. Presented by Auger at 30^{th} ICRC, 2007.

and hybrid data are shown separately and also both combined [166]. In addition, the spectrum obtained from inclined events were updated in this conference [167].

3.5.2 UHECRs: composition

X_{max} and X_{max} fluctuations

In [116] Auger presented a study of the cosmic ray composition using events recorded in the hybrid mode. They are observed by the fluorescence and the surface detectors simultaneously, so the depth of shower maximum, X_{max}, is measured directly by the FD while the SD stamp assures good reconstruction quality. The average X_{max} and the width of the X_{max} distribution at a given energy, are both correlated with the cosmic rays mass composition. Protons penetrate deeper in the atmosphere (larger values of X_{max}) and have wider X_{max} distributions than heavy nuclei. In [116] the elongation rate previously published in [154] were updated and the results with X_{max} fluctuations were

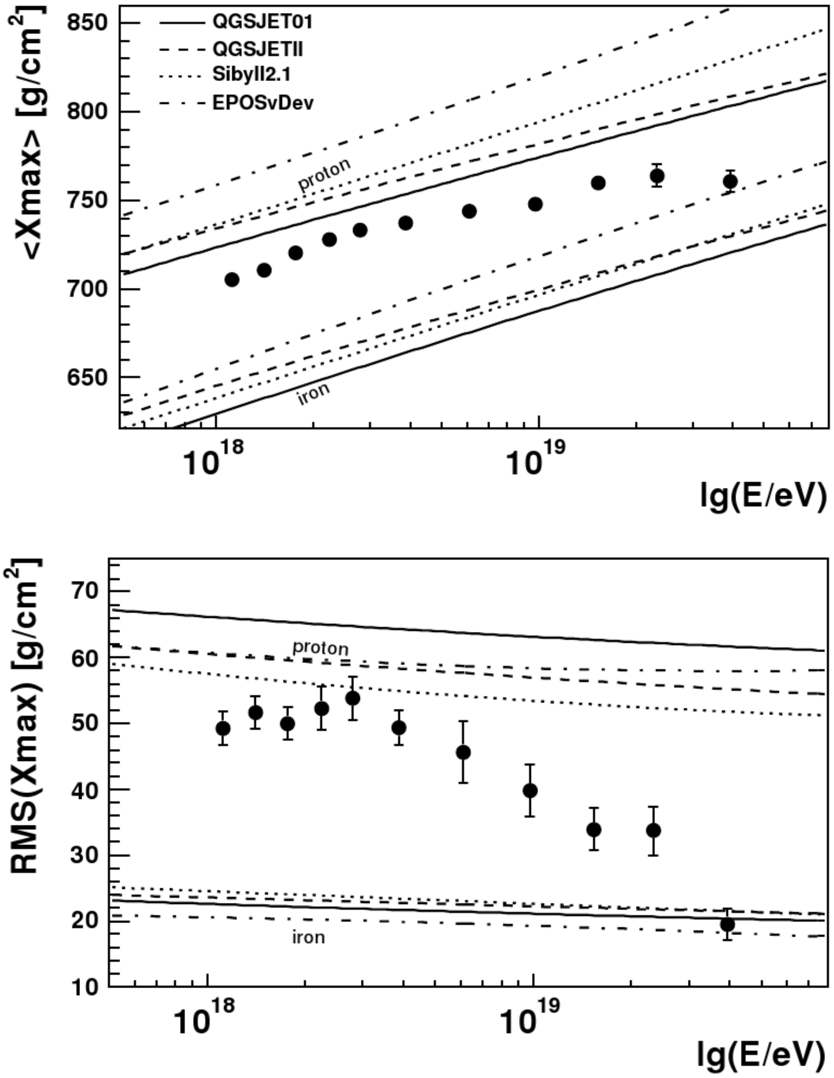

Figure 3.14: X_{max} (Top) and RMS of the X_{max} distribution (down) as a function of energy compared to predictions from different hadronic interaction models. Taken from [116].

first presented. Both are shown in Fig. 3.14. The cosmic ray composition is studied in different energy ranges by comparing the observed average $< X_{max} >$ with predictions from air shower simulations for different nuclei. The change of $< X_{max} >$ with energy is used to derive estimates of the change in primary composition. Several quality and anti-bias cuts were applied to data and systematic uncertainties were also studied. In Fig. 3.14 (upper panel) the updated mean X_{max} as a function of energy is shown. The

X_{max} measurement favors a mixed composition at all energies. A simple linear fit, $X_{max} = D_{10} \cdot log_{10}(E/\text{eV}) + c$, yields an elongation rate (variation of X_{max} per decade of energy) of 42 ± 3 (stat.) gcm^{-2}/decade, but it does not describe the data very well (χ^2 /Ndof = 20/9, P=1%). Allowing for a break in the elongation rate at an energy E_b leads to a satisfactory fit (χ^2 /Ndof = 3/7, P=85%). Then, $D_{10} = 78 \pm 13$ (stat.) gcm^{-2}/decade below $E_b = 10^{18.4} \pm 0.1$ eV, and $D_{10} = 30 \pm 4$ (stat.) gcm^{-2}/decade above this energy. E_b coincides with the measured position of the ankle (this fit was shown in Fig. 3.6).

Regarding the fluctuations of X_{max}, Fig. 3.14 (low panel) shows the root-mean-square (RMS) of the observed X_{max} distribution as a function of energy after taking into account the detector resolution. This resolution is slight energy dependence. The width of the X_{max} distributions at lower energies is about 52 gcm^{-2} and it remains constant up to energies around $10^{18.4}$ eV. Above this energy, the width starts to become gradually smaller. At the highest energy bin ($10^{19.3}$ - 10^{20} eV), the width is 22 gcm^{-2} remarkably consistent with pure iron composition. If the cosmic ray composition were made of only proton and iron, the RMS of the X_{max} distribution for the mixed composition will remain consistent with the proton one, unless the proton fraction becomes smaller than 50%.

A comparison with previous measurements as HiRes and Flys Eye was shown previously in Fig. 3.6. The results of the three experiments are compatible within their systematic uncertainties. It is worthwhile noting that although the Auger data presented in Fig. 3.6 have been collected during the construction of the observatory, their statistical precision already exceed that of preceding experiments.

In addition, in [168] several statistical methods to determine the mass composition making use of the mean and fluctuations of X_{max} are shown. They are a log-likelihood fit of the X_{max} distributions, multi-topological analysis of a selection of parameters describing the shower profile and another method using the momentums of the X_{max} distribution. These methods give primary consistent fractions that allow to reproduce the measured elongation rate reported by Auger at the ICRC 2007, independently from the hadronic model and from the applied set of analysis cuts. The comparison confirms the published Auger results independently of Monte Carlo techniques.

Composition using the surface parameters

To study the nuclear mass composition of UHECRs with the surface detector, two parameters have been selected in [131], Δ as defied in Eq. 3.2 and XAsymMax. Both have been defined previously in Section 3.4.2. The evolution of XAsymMax with energy comparing data and different hadronic models is shown in Fig. 3.15-top, which is in agreement with the elongation rate presented before. It represents a pure SD method to infer the primary composition trend. On the other hand, using hybrid events a calibration of Δ and XAsymMax with X_{max} could be obtained and, therefore, the elongation rate could be determined from both parameters as shown in Fig. 3.15-down. Both results are compatible with composition trends indicated from the direct measurements of X_{max} from the FD detectors.

An upper limit to the photon fraction in cosmic rays above 10^{18} eV

In 2007 Auger reported an upper limit to the photon fraction in cosmic rays above 10^{19} eV based on observations of the depth of the shower maximum performed with the hybrid detector [159]. An upper limit of 16% (at 95% c.l.) was derived. This was the first limit on photons obtained by observing the fluorescence light profile of air showers. This upper limit confirms and improves previous results from the Haverah Park and AGASA surface arrays (shown in Fig. 3.16)

Later, in 2008, Auger reported the upper limit obtained by using data from the surface detector, and selecting the rise time and the radius of curvature as parameters to discriminate between photons and hadrons [124]. Systematic uncertainties were considered and different hadronic models used. The limits were compared with previous ones determined by other experiments and with theoretical calculus from top-down models. These limits (labeled as Auger SD in Fig. 3.16) improve significantly upon bounds from previous experiments and put strong constraints on these models. While a minor contribution from top-down models to the observed UHE cosmic-ray flux might still be allowed within the limits derived, current top-down models do not appear to provide an adequate

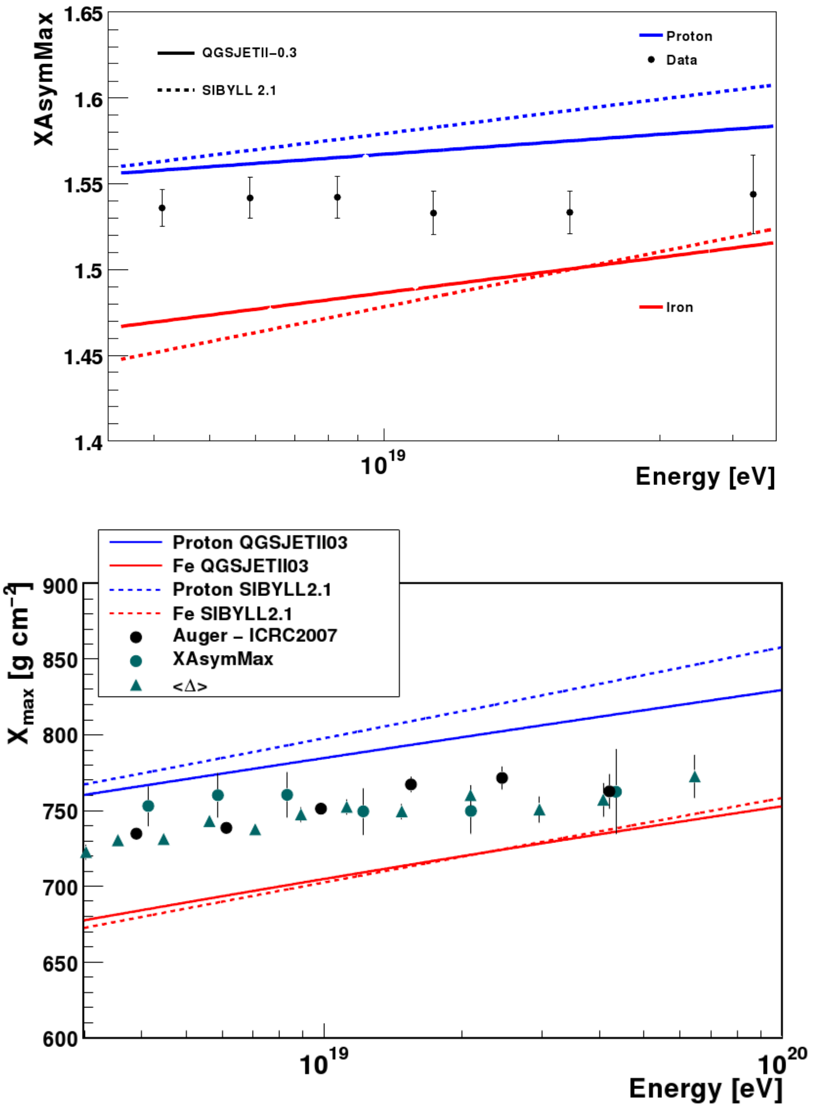

Figure 3.15: Top: Position of maximum asymmetry vs. primary energy for different models and primaries. Lines correspond to fitted distributions of MC samples for proton (blue) and iron (red) primaries. Down: X_{max} from Δ and XAsymMax vs. energy. Predictions for a pure iron and pure proton composition according to different models as well as results from direct measurement of X_{max} using the FD are shown for comparison. Uncertainties are only statistical. Taken from [131].

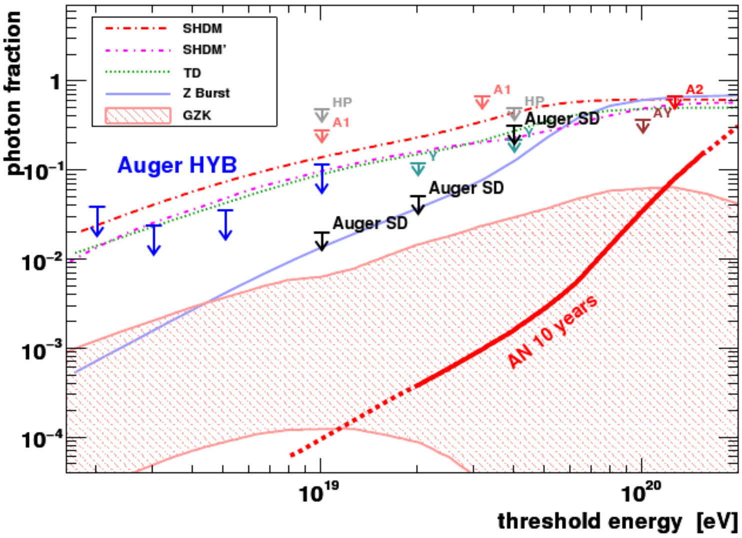

Figure 3.16: Upper limits on the photon fraction in the integral cosmic ray flux from different experiments. The limits from the Auger surface detector are labeled Auger SD and the limits from [169] as Auger HYB. The line labeled as AN indicates sensitivity of the northern Auger Observatory after 10 years of operation. The other lines indicate predictions from top-down models and the shaded region shows the expected GZK photon fraction. Figure is taken from [169].

explanation of the origin for the bulk of the highest-energy cosmic rays. This result join with [170], where the discovery of a correlation between the direction of most energetic cosmic rays and the positions of nearby AGNs is shown, suggest an astrophysical origin of UHE cosmic rays.

Finally, Auger extended the results to EeV energies using X_{max} in 2009 [169]. The resultant 95% c.l. upper limits on the photon fractions are 3.8%, 2.4%, 3.5% and 11.7% for the primary energies above 2, 3, 5 and 10 EeV respectively (labeled as Auger HYB in Fig. 3.16).

These photon limits also help to reduce certain systematic uncertainties in other analysis of air shower data: (i) regarding the energy spectrum, the Auger method of reconstructing the energy spectrum does not suffer from a large contamination from photons at EeV energies; (ii) about nuclear primary composition, the interpretation of observables sensitive to the primary particle (for instance the observed average X_{max}) in terms of a nuclear primary composition can only be marginally biased by contributions from photons; (iii) the possible contamination from photons was one of the dominant uncertainties for deriving the proton-air cross-section [171], and this uncertainty is now significantly reduced.

3.5.3 Upper limit on the diffuse flux of ultra-high energy neutrinos

In [172] is showed that the surface detector array of the Pierre Auger Observatory is sensitive to Earth-skimming tau neutrinos that interact in Earth's crust. Tau leptons from ν_τ charged-current interactions can emerge and decay in the atmosphere producing a nearly horizontal shower with a significant electromagnetic component. The data collected between 1 January 2004 and 31 August 2007 were used to place an upper limit on the diffuse flux of ν_τ at EeV energies.

The procedure devised to identify neutrino candidate events within the data set is based on an end-to-end simulation of the whole process, from the interaction of the ν_τ inside the Earth to the detection of the signals in the tanks. The first step is the calculation of the τ flux emerging from the Earth and modeling of the showers from τ decays in the atmosphere. Later, it is needed to evaluate the response of the SD to such events. A set of conditions has been designed and optimized to select showers induced by Earth-skimming ν_τ, rejecting those induced by UHECR. Over the period analyzed, no candidate events were found that fulfilled the selection criteria. Based on this, the Pierre Auger Observatory data were used to place a limit on the diffuse flux of UHE ν_τ. For this purpose the exposure of the detector was also evaluated.

Later, Auger updates this result. First, using data until April 2008 in [173]. Second, using data from 1 Jan 04 - 28 Feb 09 (corresponds to $\sim$ 2 years of full operation) in 2009 [174], where again no candidate was found. The updated limit based on Earth-skimming up-going neutrinos is $k < 4.7^{2.2}_{6.9} \times 10^{-8}$ GeVcm^{-2}s^{-1}sr^{-1}, where the upper/lower values

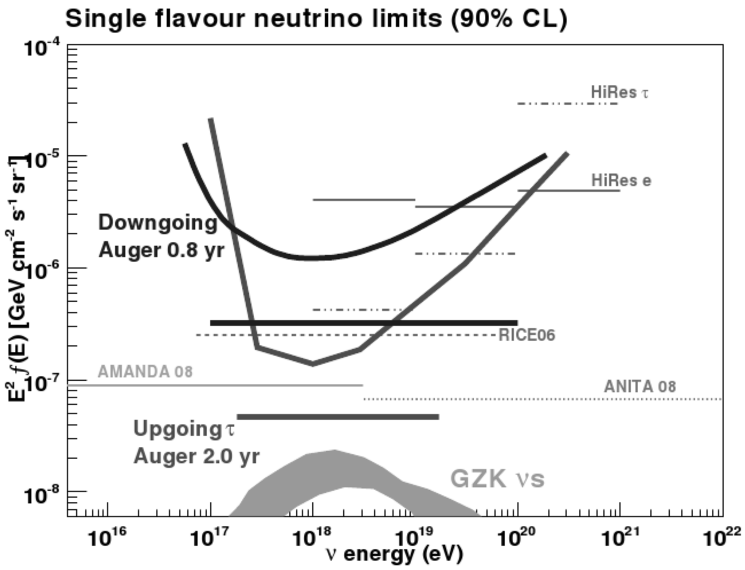

Figure 3.17: Differential and integrated upper limits (90% C.L.) from the Pierre Auger Observatory for a diffuse flux of down-going ν in the period 1 Nov 07 - 28 Feb 09 and up-going ν_τ in the period 1 Jan 04 - 28 Feb 09. Limits from other experiments are also plotted. A theoretical flux for GZK neutrinos is shown. Taken from [174].

correspond to best/worse scenario of systematics.

In addition, in [174] it is shown that the SD of the Pierre Auger Observatory is also sensitive to *down-going* neutrinos of all flavors interacting in the atmosphere, and inducing showers close to the ground. Down-going neutrinos of any flavor interacting through charged or neutral current, may induce showers in the atmosphere that can be detected using the SD. A set of conditions has been designed to select inclined showers initiated by down-going neutrinos, different from the one applied to search for up-going neutrinos (details in [175, 176]). To identify down-going neutrinos, the exposure of the SD array to UHE neutrinos is calculated. Assuming a $f(E_\nu) = k \cdot E_\nu^{-2}$ differential neutrino flux, Auger obtained a 90% C.L. limit in the all-flavor neutrino flux using down-going showers of $k < 3.2 \times 10^{-7}$ GeVcm^{-2}s^{-1}sr^{-1}.

In Fig. 3.17 both limits are shown and also in the differential format to show the range in energies at which the sensitivity of the Pierre Auger Observatory to down-going and

Earth-skimming ν peaks.

3.5.4 Search for UHECR sources and anisotropies

Anisotropy studies around the galactic center at EeV energies

The Galactic Center (GC) region constitutes an attractive target for cosmic ray anisotropy studies at EeV energies. These may be the highest energies for which the galactic component of the cosmic rays is still dominant. Moreover, since the GC harbors the very massive black hole associated with the radio source Sagittarius A*, as well as the expanding supernova remnant Sagittarius A East, it contains objects that might be candidates for powerful cosmic ray accelerators. The recent high significance observation by H.E.S.S. of a TeV γ ray source near the location of Sagittarius A* [177], further motivates the search for excesses in this direction.

AGASA experiment reported a 4.5σ excess of cosmic rays with energies in the range $10^{18} - 10^{18.4}$ eV in a 20^o radius region centered at right ascension and declination coordinates $(\alpha, \delta) \approx (280^o, -17^o)$, in which the number of observed and expected events are $n^{obs}/n^{exp} = 506/413.6 = 1.22 \pm 0.05$, where the error quoted is the one associated with Poisson background fluctuations [99]. Later searches near this region with a re-analysis of SUGAR data [88], though with smaller statistics, failed to confirm these findings, but reported a 2.9σ excess flux of cosmic rays with energies in the range $10^{17.9} - 10^{18.5}$ eV in a region of 5.5^o radius centered at $(\alpha, \delta) \approx (274^o, -22^o)$, for which they obtained $n^{obs}/n^{exp} = 21.8/11.8 = 1.85 \pm 0.29$.

In [178] the first 2.3 years of Auger data are analyzed to search for anisotropies at EeV energies near the direction of the GC, which is well within the field of view of the Observatory. The exposure of the surface array in this part of the sky is significantly larger in this period than that of the fore-runner experiments. The results do not support previous findings of localized excesses in the AGASA and SUGAR data (see Fig. 3.18). Also hybrid events, which have better pointing accuracy but are less numerous than those of the surface array alone, do not show any signicant localized excess from this direction.

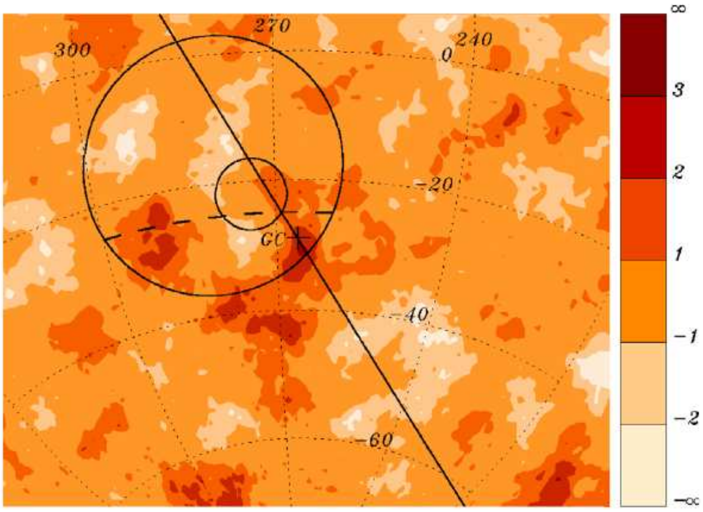

Figure 3.18: Map of cosmic ray overdensity significances near the Galactic Center (GC) region on top-hat windows of 5^o radius. The GC location is indicated with a cross, lying along the galactic plane (solid line). Also the regions where the AGASA experiment found their largest excess as well as the region of the SUGAR excess are indicated. No significant excess is found by Auger. Taken from [178].

Discriminating potential astrophysical sources of the highest energy cosmic rays

Several papers were presented regarding this topic in the ICRC 2009. In [179], the distribution of arrival directions of the highest energy cosmic rays detected by the Pierre Auger Observatory from 1 January 2004 to 31 March 2009, is compared to that of populations of potential astrophysical sources. For this purpose, two complementary statistical tests allowing to describe and quantify the degree of compatibility between data and a given sources catalog are used. These tests were applied to active galactic nuclei detected in X-rays by SWIFT-BAT experiment and to galaxies found in the HI Parkes and in the 2 Micron All-Sky Surveys. These tests show that Auger data are different from isotropic expectations and compatible with all models, indicating that the UHECRs may follow the distribution of local matter.

In [36] Auger used the cosmic ray data from the surface detector to search for cosmic rays that correlate with the time and position of GRBs. No such correlations were found. As a separate analysis, the data was used to look for increases in the average trigger rate of the surface detectors, which would indicate the occurrence of a GRB. No burst with characteristics similar to those expected for GRBs was observed.

Correlation of the highest energy cosmic rays with nearby extragalactic objects

Data collected by the Pierre Auger Observatory between 1 January 2004 and 31 August 2007, provide evidence for anisotropy in the arrival directions of the cosmic rays with the highest energies, which are correlated with the positions of relatively nearby active galactic nuclei (AGN) from the 12^{th} edition of the catalog of quasars and AGN by Véron-Cetty and Véron (VCV catalog) [170, 180]. The correlation has maximum significance for cosmic rays with energy greater than $\sim 6 \cdot 10^{19}$ eV and AGN at a distance less than ~ 75 Mpc. Anisotropy is confirmed at a confidence level of more than 99% through a test with parameters specified a priory, using an independent data set. The result is shown in Fig. 3.19. The observed correlation is compatible with the hypothesis that cosmic rays at the highest energies originate from extra-galactic sources close enough so that their flux is not significantly attenuated by interaction with the cosmic background radiation (the Greisen-Zatsepin-Kuz'min effect). The angular scale of the correlation observed is a few degrees, which suggests a predominantly light composition unless the magnetic fields are very weak outside the thin disk of our galaxy. This result do not identify AGN as the sources of cosmic rays unambiguously, because other candidate sources which are distributed as nearby AGN are not ruled out and the catalog used is not complete. In [180] it is also discuss the prospect of unequivocal identification of individual sources of the highest-energy cosmic rays within a few years of continued operation of the Pierre Auger Observatory.

This result has been supported recently by two contributions presented at the ICRC 2009. In [181] the previous study is updated using data collected between 1 January 2004 and 31 March 2009. The arrival directions were correlated with the positions of nearby objects from the same catalog as before. This catalog is not an unbiased statistical sample,

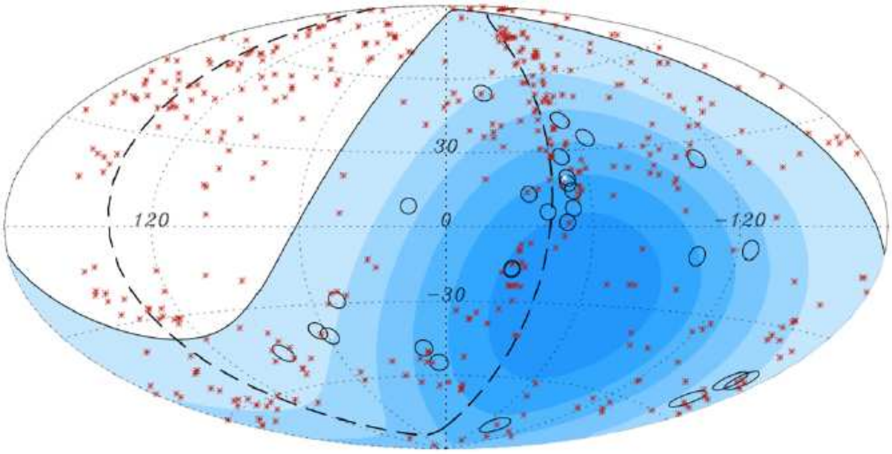

Figure 3.19: Aitoff projection of the celestial sphere in galactic coordinates with circles of 3.2^o centered at the arrival directions of 27 cosmic rays detected by the Pierre Auger Observatory with reconstructed energies $E > 57$ EeV. The positions of the 442 AGN (292 within the field of view of the Observatory) with redshift $z \leq 0.017$ $(D < 71Mpc)$ from the 12th edition of the catalog of quasars and active nuclei are indicated by asterisks. The solid line draws the border of the field of view for the southern site of the Observatory (with zenith angles smaller than 60^o). The dashed line is, for reference, the super-galactic plane. Darker color indicates larger relative exposure. Each colored band has equal integrated exposure. Centaurus A, one of the closest AGN, is marked in white. Taken from [180].

since it is neither homogeneous nor statistically complete. This is not an obstacle to demonstrate the existence of anisotropy if CRs arrive preferentially close to the positions of nearby objects in this sample. The nature of the catalog, however, limits the ability of the correlation method to identify the actual sources of cosmic rays. The observed correlation identifies neither individual sources nor a specific class of astrophysical sites of origin. It provides clues to the extragalactic origin of the CRs with the highest energies and suggests that the suppression of the flux is due to interaction with the cosmic background radiation. The correlation found has a less than 1% probability to occur by chance if the arrival directions are isotropically distributed. The evidence for anisotropy and for correlation with objects in the VCV catalog has not strengthened since the analysis reported in [170]. The largest excess over angular scales of order 20^o in the data set occurs close to the direction of the radio source Cen A, a region dense in potential sources. This excess is based on a posteriori data but suggests that this region of the sky warrants further study.

The intrinsic anisotropy of Auger data at the highest energies is also confirmed in [182] where several techniques are discussed. They have been developed for determining the intrinsic anisotropy of sparse ultra-high-energy cosmic ray data sets, including autocorrelation, an improved two point method (2pt+) and a three point method (3pt). Monte Carlo simulations indicate that we need more events in the highest energy bin to be able to measure the intrinsic anisotropy. The data provide an indication of intrinsic anisotropy at the highest energies, with a minimum found by both statistical estimators (2pt+ and 3pt) at about 52 EeV, which corresponds to the 70 events with highest energy.

Chapter 4

Optimum distance to determine the cosmic ray energy in surface arrays

In most high energy cosmic ray surface arrays, the primary energy is currently determined from the value of the lateral distribution function at a fixed distance from the shower core, r_0. The value of r_0 is mainly related to the geometry of the array and is, therefore, considered as fixed independently of the shower energy or direction. We argue, however, that the dependence of r_0 on energy and zenith angle is not negligible. Therefore, in the present work we propose a new optimum distance, which we call r_{opt}, specifically determined for each individual shower, with the objective of optimizing the energy reconstruction. This parameter may not only improve the energy determination, but also allow a more reliable reconstruction of the shape and position of rapidly varying spectral features. We show that the use of a specific r_{opt} determined on a shower-to-shower basis, instead of using a fixed characteristic value, is of particular benefit in dealing with the energy reconstruction of events with saturated detectors, which are in general a large fraction of all the events detected by an array as energy increases. Furthermore, the r_{opt} approach has the additional advantage of applying the same unified treatment for all detected events, regardless of whether they have saturated detectors or not.

In this Chapter surface arrays of triangular geometry (as Auger South Observatory or AGASA experiment) are considered. The application to a square array, as the future Pierre Auger North Observatory, is shown in the Appendix A.

4.1 Motivation

The procedure to determine the primary energy in surface arrays is a two step process. First, the lateral distribution function (LDF), i.e. the shower particle density or signal versus distance to the shower axis, is fitted assuming a known functional form. This fit suffers from uncertainties related to the statistical shower fluctuations, the uncertainties in the core location and the ignorance of the exact form of the LDF. The normalization constant of the LDF of an extensive air shower is a monotonous (almost linear) increasing function of the energy of the primary cosmic ray. Therefore, Hillas [82] proposed to use the interpolated signal at some fixed *characteristic distance* from the shower core, $S(r_0)$, at which fluctuations in the LDF are minimal. The uncertainty due to the lack of knowledge of the LDF is also minimized by this procedure [81]. The use of the signal interpolated at r_0, $S(r_0)$, is widely used as energy estimator by surface detector arrays. AGASA [183, 184], Yakutsk [185] and Haverah Park [186], for example, choose $r_0 = 600$ m, while Auger uses 1000 m due to its larger array spacing [187]. The characteristic distance r_0 is mainly, although not completely, determined by the geometry of the array. Thus, the same value of r_0 is used to estimate the energy for all the showers, regardless of primary energy or incoming direction. In the second step, there are at least two possible approaches to calibrate $S(r_0)$ as a function of primary energy: either via Monte Carlo simulations or, as in the case of Auger, by using the almost calorimetric measurement obtained from the fluorescence observation of high quality hybrid events [164].

As an alternative, but motivated by Hillas' original idea [82], in the present work we focus in the shower-to-shower determination of an *optimum distance* to the core, which we name hereafter r_{opt}, at which the interpolation of the signal is the best energy estimator for each individual shower, regardless of whether this point is actually the one that minimizes shower to shower LDF fluctuations.

We perform a detailed study of r_{opt} as a function of array spacing, primary energy and the zenith angle of the incoming cosmic ray and demonstrate that, although array geometry is an important underlying factor, the dependence of r_{opt} on the remaining parameters is not negligible. We study the bias associated with both techniques, r_0 and r_{opt}, and show that, if the dynamical range of the detector covers a wide interval of energies, it is much safer to estimate an r_{opt} for the energy reconstruction of each individual event than to fix a single r_0 for the whole data set. In fact, not only the bias as a function of energy can be kept negligible over at least 2.5 decades in energy, but also the error distribution functions are much better behaved, i.e. without appreciable kurtosis or skewness and very much Gaussian-like in the mentioned energy range. The latter has a potential impact in the reconstruction accuracy of the energy spectrum. We demonstrate this by applying a fixed r_0 as well as a shower-to-shower r_{opt}, to a simplified version of the actual energy spectrum between ~ 1 and ~ 100 EeV.

A further advantage of the r_{opt} approach is the straightforward treatment of events with saturated detectors. The problem of saturation is very common in all surface experiments, specially when dealing with high energy vertical showers. In fact, at the highest energies inside the designed dynamical range of any experiment, usually events with saturated detectors can account for a large, if not dominant, fraction of all the observed events. Different strategies have been used to deal with them. In some cases saturated detectors are directly discarded from the LDF fit, while in others the saturation value is used as a lower limit to the true signal during the fitting procedure. The Auger Collaboration is developing at present special, more sophisticated algorithms to estimate the signal of a saturated detector [106] in order to more properly account for them in the LDF fit. We show here that it is actually not possible to define a single characteristic r_0 distance for both kinds of events. In fact, even if well defined medians values of r_{opt} for events with and without saturated detectors do exist, the dispersions around the median at any energy are so large that both sets cannot be clearly differentiated as to use, for example, just two fixed distances instead of a single one. Nevertheless, using a shower-to-shower r_{opt} distance, the inferred energy is unbiased for events with and without saturated detectors. This reconstruction strategy allows for an homogeneous treatment of the data set regardless of the increasing number of events with saturated detectors when the energy increases.

In a recent work [187], Newton and co-workers also estimated an optimal shower-to-shower distance, but used a different algorithm and with a somewhat different scope. They were mainly concerned with demonstrating the existence of a single distance for any given shower at which fluctuations in the LDF are minimum. By assuming that such fluctuations can be well described by the fluctuations of just one parameter, the slope of the LDF, externally fed into their procedure, and using a combination of simulations and semi-analytical analysis, they claim that, regardless of the functional form of the LDF considered, there exist a *convergence* point of the LDFs, at a characteristic distance they call optimal, where shower-to-shower fluctuations are minimal. Their results, combined together for a mix of energies drawn from a flat spectrum, seem to support their claim and lead them to the conclusion that a single fixed distance, depending only on the geometry and spacing of a given array, would be a good choice for the energy determination in the whole energy range of the experiment. Furthermore, it is not clear from their study how to deal with the events with saturated detectors in the later scenario.

Alternatively, in the present work we do not constrain the parameters of the LDF, which are an output of the fit to the simulated data. We introduce instead reliable error estimations for the reconstruction of the core position, as calculated by [188] for arrays of varying spacing as a function of energy. Furthermore, our final scope is the determination of energy all along the dynamical range of an experiment, and not the study of the manifestation of signal fluctuations in the LDF. Therefore, we study in detail the dependence of r_{opt} and of its distribution function as a function of energy, zenith angle and array spacing. This study is performed for events with and without saturated detectors. We also give a comparative description of error and biases for the fix distance and the r_{opt} distance approaches in that parameter space. In the same line, we further extend our analysis to the reconstruction of a simulated energy spectrum of known shape, and show what the potential effects are of using each technique.

The Chapter is organized as follows. Section 4.2 describes our general algorithm. Two different detector arrays are considered, scintillators and water Cherenkov tanks. In Section 4.3, in order to study the r_{opt} dependencies with array spacing and the energy and incoming direction of primary cosmic ray, water Cherenkov (Auger-like) stations have been used. In Section 4.4 we deal with the issue of energy determination. In that analysis, (AGASA-like) scintillators are considered. A general discussion and conclusions are given in Section 4.5. While different detectors are used in Sections 4.3 and 4.4, the algorithm to find r_{opt} is the same for both and the results and conclusions are not affected by the array under consideration.

4.2 Algorithm to determine the optimum distance

The basic idea of our algorithm is to estimate the optimum distance to the core at which to determine the energy of a shower under the most realistic possible conditions.

We assume a certain analytical LDF as the intrinsic average lateral distribution of particles inside the shower front as a function of the distance to the core. For the chosen energy and geometry of the event (azimuth, zenith and true core position), this function is used to estimate the average LDF value at the actual position of each detector. Afterward, a signal is calculated using the previous average value as the mean of a Poissonian distribution. If the calculated signal falls between a minimum threshold and an upper limit corresponding to a saturation condition, it is assigned to the detector. In case of saturated detectors, the event is kept, but the saturated detectors are not used in the subsequent analysis, i.e. in fitting the LDF. An event flag indicates when saturation occurs. Once a set of triggered detectors participating in the event has been defined, the reconstructed LDF is emulated by fitting an *experimental* LDF, which depends on the detector array under consideration, and is not necessarily the *real* LDF used in the first step to generate the event. The LDF fit requires an estimate of the core position. Such estimate is an important component of the analysis of the event and comes, in practice, from a global reconstruction procedure which implies an energy dependent error in the inferred position of the core. In our algorithm, we simulate this error by shifting the reconstructed core position according to its experimentally determined Gaussian distribution function. For each *shifted* core position an independent LDF fit is performed. We define the optimum distance to the core r_{opt} as the interpolated distance at which the dispersion between the several LDF fits is minimal. We argue that the interpolated signal at this point is the optimum estimator of the energy of a real event and constitutes the operational definition of our parameter r_{opt}. A previous version of our algorithm has been presented in [189]. More details are given below.

We use the following numerical approach to simulate EAS detection in a surface array. The array is a set of equally spaced detectors, located at the vertexes of an infinite grid of triangular elementary cells with variable spacing. The input parameters of an event are its energy, azimuthal and zenith angles and core position. The identity of the primary particle is not taken into account since differences in composition produce only a small effect in the error distribution function of the reconstructed core position [188] which, in turn, when combined with the use of an experimental LDF maps into a negligible variation in both r_0 and r_{opt}.

Whenever we simulate a water Cherenkov detector, we assume that the true lateral distribution of the signal is best represented by a Nishimura-Kamata-Greisen (NKG) function. This functional form was first obtained in an analytical study of the lateral development of electromagnetic showers in [79], and later extended to the hadronic initiated showers because the electromagnetic particles represent around the 90% of the total particles of the shower. The NKG selected is normalized at 1 km in the same way as the reported by Auger in [6]:

$$S(r,E,\theta) = \frac{7.53\ E^{0.95}\ 2^{\beta(\theta)}}{\sqrt{1+11.8[sec(\theta)-1]^2}} \times r^{-\beta(\theta)} \times (1+r)^{-\beta(\theta)} \tag{4.1}$$

where r is the distance to the shower axis expressed in km, E is the energy of the primary in EeV, θ is the zenith angle and $\beta(\theta) = 3.1 - 0.7sec(\theta)$. The signal in Eq. 4.1 is expressed in vertical equivalent muons (VEM), which correspond to the signal deposited by one vertical muon in an Auger water Cherenkov tank.

We use Eq. 4.1 as the real LDF to simulate any given incoming event. The measured signal at each station is obtained with a Poissonian probability distribution function whose mean is given by Eq. 4.1, the "true" LDF. The trigger condition is set to $S(r) = 3.0$ VEM. The saturation value is fixed at S(0.2 km, 1 EeV, 0^o). These values are compatible with the equivalent Auger parameters.

The uncertainty in core determination depends on the array geometry and primary energy and it has been estimated for a variety of cases in [188]. We simulate the reconstruction uncertainty of the core using a Gaussian distribution function centered at the position of the real core, with standard deviation given by [188] as a function of the energy of the shower for the array spacing under consideration.

For any shower, the following procedure is used to obtain the optimum distance r_{opt}. Throughout the procedure, we try to mimic, as far as possible, the actual reconstruction procedure. As explained earlier, several fits to the LDF are performed for any event, each one with its own estimated core position. Since the exact functional form of the LDF function is not crucial [81] we use a generic LDF parametrization to fit the signals of the triggered stations:

$$\log S(r) = a_1 r^{-a_2} + a_3 \tag{4.2}$$

The uncertainty in the core position used for each one of the LDF fits corresponding to a given event, is accounted for by randomly shifting that point 50 times with the same Gaussian probability distribution function referred above centered at the position of the reconstructed core.

For each new core position, the LDF fit is performed using Eq. 4.2. The slope and the normalization constant of each LDF are determined from the fitting procedure. The r_{opt} value is defined as the point at which the dispersion among the interpolated signals over the several LDFs goes through a minimum.

Therefore, the implementation of the algorithm in a real experiment requires a two step process: first, a global fit to the LDF is performed in order to get an estimate of the reconstructed core position and, second, the reconstructed core itself is fluctuated and r_{opt} obtained using the procedure explained in the previous paragraph.

When simulating scintillators as those of the AGASA experiment, we follow exactly the same procedure as before, but we use the following LDF [56] instead of Eq. 4.1:

$$\begin{aligned}\rho(r,E,\theta) &= 49.676 \times 7.55^{\eta(\theta)-1.2} \times f_s(\theta) \times E^{1/1.03} \times \\ & \left(\frac{r}{r_M}\right)^{-1.2} \left(1+\frac{r}{r_M}\right)^{-(\eta(\theta)-1.2)} \left(1+r^2\right)^{-0.6}\end{aligned} \tag{4.3}$$

where ρ is given in m^{-2}, distances are in km, $r_M = 0.0916\ km$ is the Moliere radius at AGASA altitude, $\eta(\theta) = 3.84 - 2.15(\sec(\theta) - 1)$, the energy E is in EeV and $f_s(\theta)$ is the attenuation curve:

$$f_s(\theta) = exp\left[-\frac{X_0}{\Lambda_1}(sec\theta - 1) - \frac{X_0}{\Lambda_2}(sec\,\theta - 1)^2\right] \tag{4.4}$$

where $X_0 = 920\ g/cm^2$, $\Lambda_1 = 500\ g/cm^2$ and $\Lambda_2 = 594\ g/cm^2$ for showers with $\theta \leq 45^o$. The signal is fluctuated, as always, with a Poissonian distribution.

In this case, the trigger condition is selected in such a way that the signal is not dominated by fluctuations. In particular, we use a vale of ρ such that fluctuations account at most by 50% of the signal. The saturation value is $\rho(0.2$ km $, 1$ EeV $, 0^o)$.

In the AGASA case, we use as shower generator Eq. 4.3, and perform the subsequent fitting procedure using an LDF with the functional form “observed” by AGASA:

$$\log\rho(r) = a_1 - a_2\log(r/r_M) - 0.6\log(1 + (r/1000\text{m})^2) \tag{4.5}$$

which is formally equivalent to Eq. 4.3 for $r >> r_M$.

Fig. 4.1 shows some examples of the fitting procedure and the point of the minimum fluctuations. In the figure the case of water Cherenkov tanks in a grid of 1500 m spacing is shown. The last example is a typical event with a saturated detector.

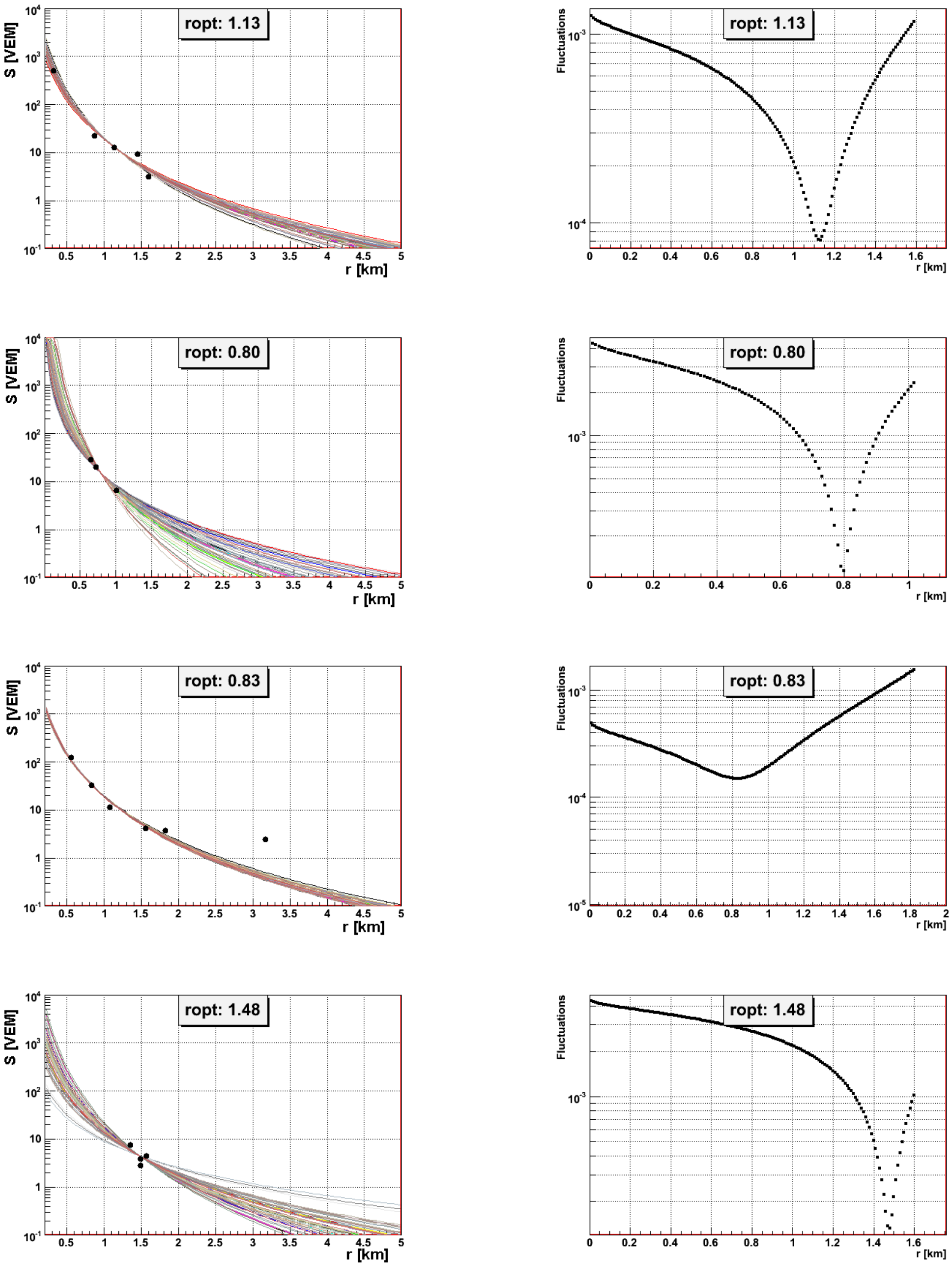

Figure 4.1: Examples of the fitting procedure to find r_{opt} for the 1500 m array and water Cherenkov tanks. Left: 50 LDF fits shifting the core. Right: signal dispersion as a function of distance. r_{opt} is defined as the point at which the signal dispersion reaches the minimum. The last example is a typical event with a saturated detector.

4.3 r_{opt} dependence on the array spacing, the energy and the zenith angle of the event.

We consider in this Section water Cherenkov detectors with separations of 433, 750, 866 and 1500 m, as well as primary energies varying from 10^{17} to $10^{19.5}$ eV. We use Eq. 4.1 to generate the signals and Eq. 4.2 to fit the LDF.

In all cases we consider a uniform distribution in azimuth and zenith angles as explained in each figure. Shower cores are uniformly distributed inside an elementary cell of the array.

Events with and without saturated detectors lead frequently to systematically different behaviors regarding the relationship between r_0 and r_{opt} under discussion and, in principle, should be treated differently during data processing. Thus, in what follows, we will analyze them separately whenever appropriate.

Figure 4.2 shows the dependence of r_{opt} with energy without discriminating whether events have saturated detectors (labeled as *All*). Showers are injected with zenith angles $\theta = 30^o$ and 60^o. It can be seen that r_{opt} is a monotonous increasing function of energy due to the triggering of stations progressively further away from the core as energy increases. In Figure 4.3 the same behavior is shown separately for events with and without saturation.

Since events with and without saturated detectors may, and indeed do, behave in different ways, Figure 4.3 shows r_{opt} results for both separately. All the previous array spacings are considered but only at one zenith angle, $\theta = 30^o$. r_{opt} is an increasing function of energy for both sets of events. It can be seen that r_{opt} is larger for events with saturated detectors at lower energies ($r_{opt}^{sat} > r_{opt}^{non-sat}$) but that, at higher energies, $r_{opt}^{non-sat}$ rapidly grows towards r_{opt}^{sat}. The transition energy region is narrow ($\Delta logE \sim 0.2$) and shifts upwards in energy as the array spacing grows. The symmetry of the triangular array with respect to a shower core located at the center of an elementary triangle (the least likely configuration to saturate at any given energy), manifest itself in ring-like arrangements of triggered stations. The appearance of a third ring of triggered stations is responsible for

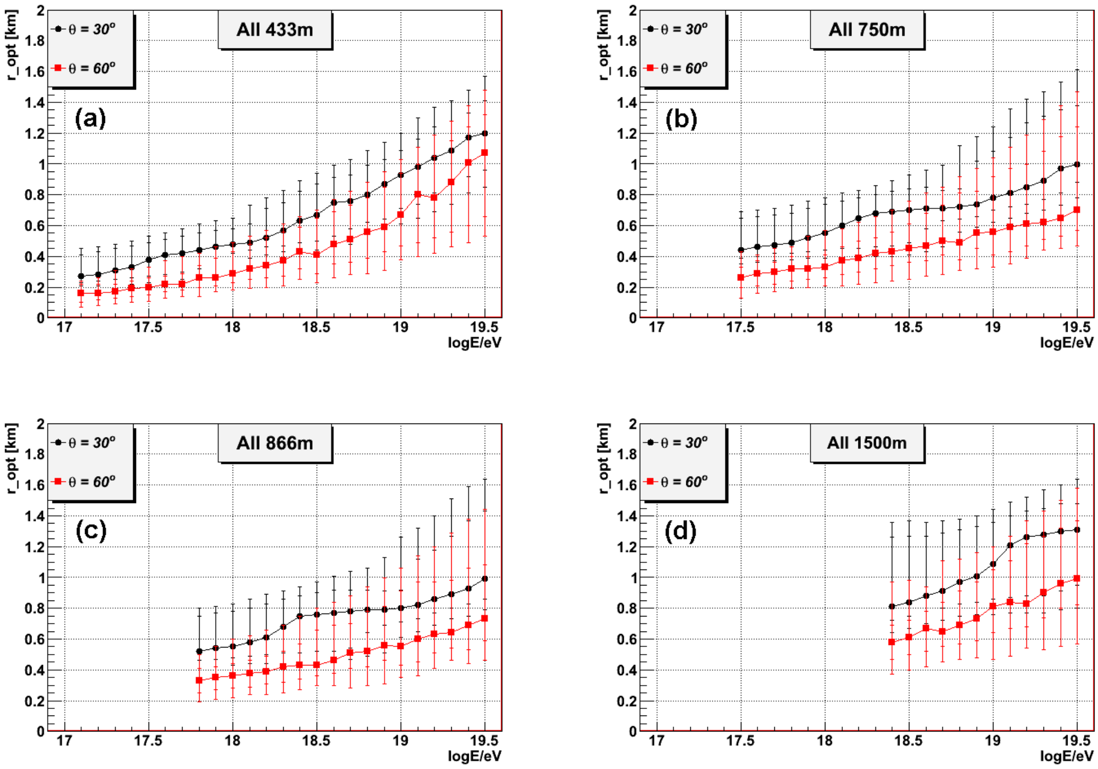

Figure 4.2: r_{opt} vs. energy for different array spacing and zenith angle. The error bars represent the 68% and 95% C.L. The label *All* means that events with and without saturated detectors are both included.

the rapid growth of $r_{opt}^{non-sat}$ over a limited energy interval as shower energy grows. This is shown in Figure 4.4 for the array of 750 m spacing and $\theta = 30^o$, where the distance of the triggered stations to shower axis for non-saturated events the events in the energy bins at $10^{18.4}$ and $10^{18.7}$ eV are shown, which are the energies before and after the jump for this array and zenith angle.

Furthermore, low energy events with saturation have their cores very near the saturated stations. Therefore, the first triggered stations that do not saturate are clustered at the same distance from the core, which is roughly the array separation distance. Therefore, it is at the array separation distance that the dispersion among the several fits to the LDF is minimum (an example could be seen in Fig. 4.1-down). At higher energies, however, the next ring of the array enters into the set of triggered detectors of the event

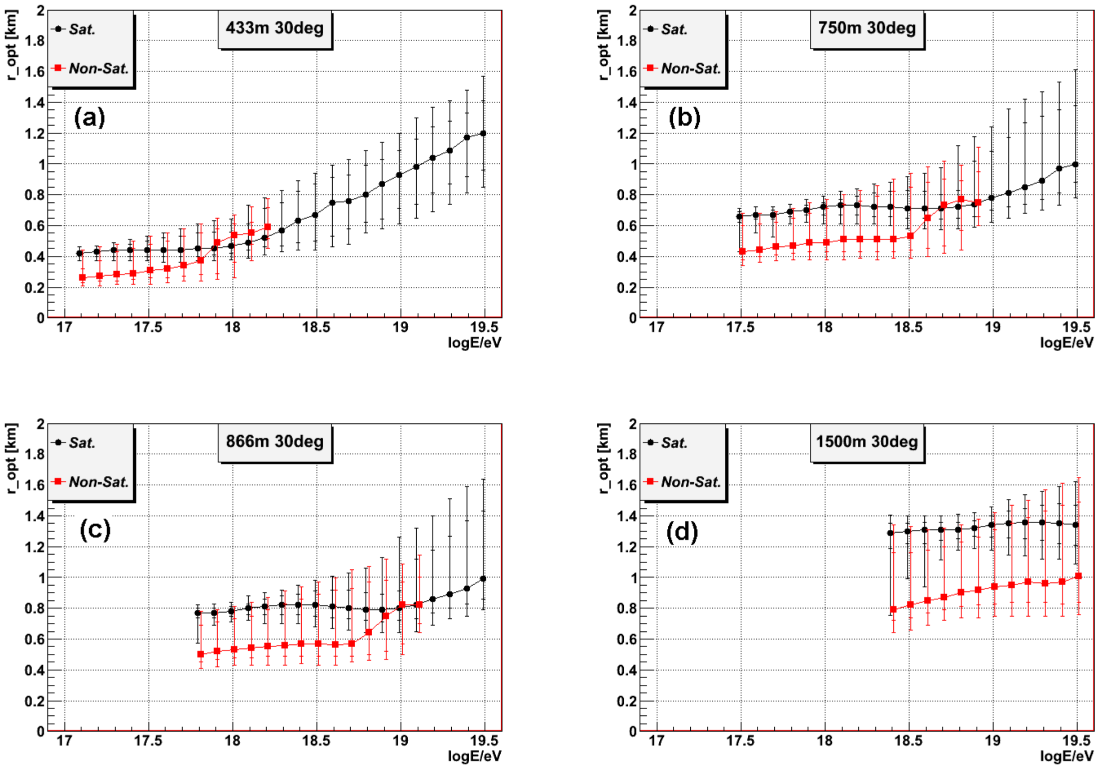

Figure 4.3: r_{opt} vs. energy. Events with and without saturated detectors are shown separately. Zenith angle is $\theta = 30^o$. The error bars represent the 68% and 95% C.L.

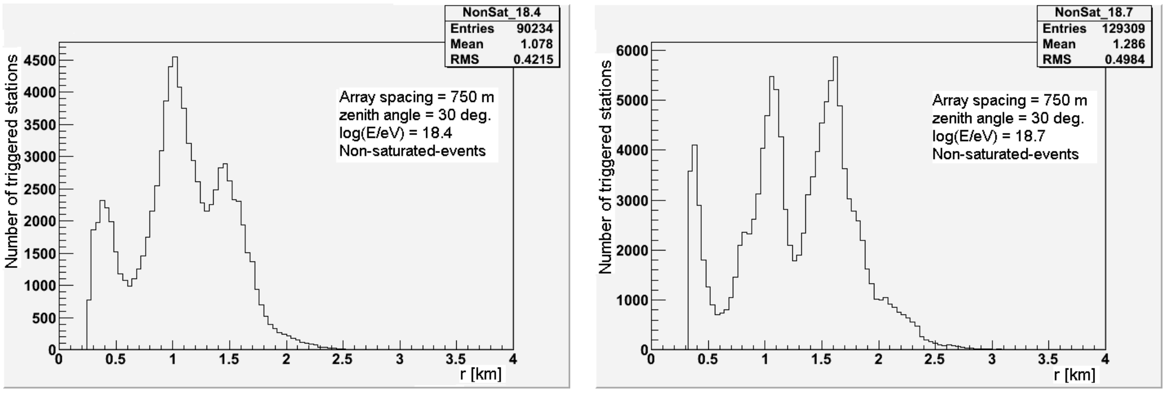

Figure 4.4: Distance of the triggered stations to the shower axis for events without saturation in a 750 m array, for events with $\theta = 30^\circ$ and energies of $10^{18.4}$ (left) and $10^{18.7}$ eV (right). The appearance of a third ring of stations is responsible for the rapid growth of r_{opt} seen in in Figure 4.3.

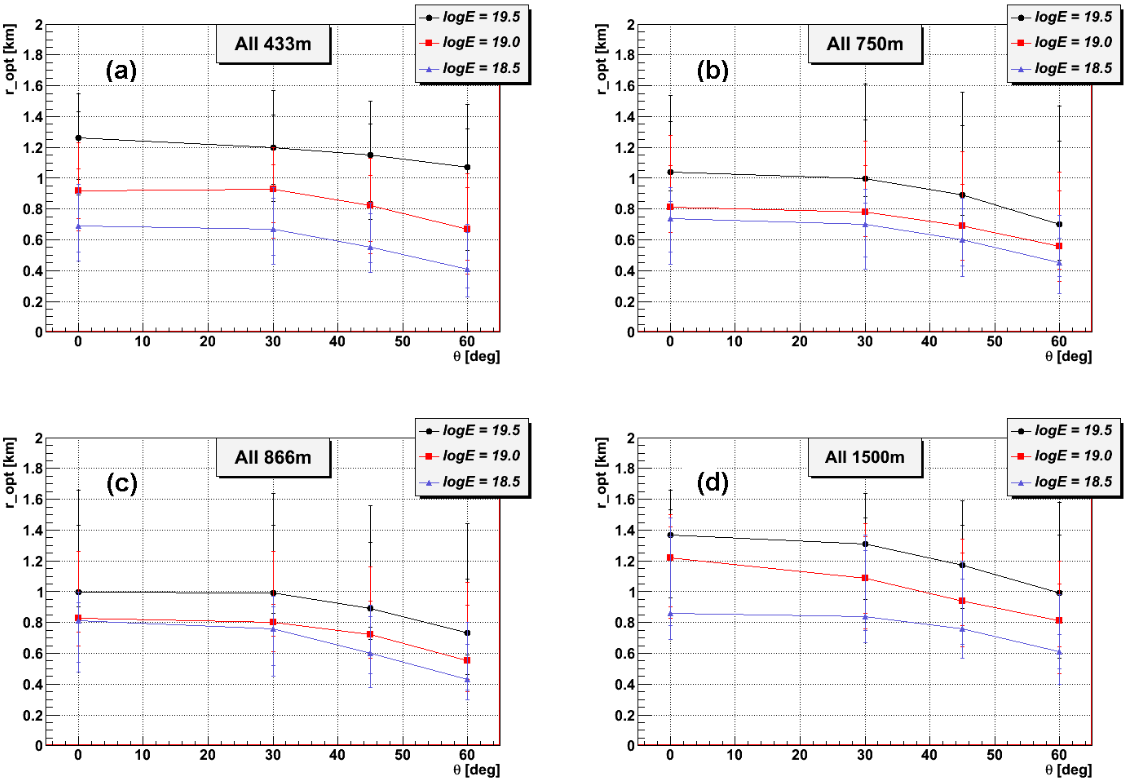

Figure 4.5: r_{opt} vs. zenith angle for different array spacing and energies. The error bars represent the 68% and 95% C.L. Events with and without saturated detectors are both included.

and, naturally, r_{opt} increases. In Figure 4.3 it can be seen that r_{opt} is almost constant and very near to the array separation at the lower energies, and that there is a threshold energy, which depends on the array separation, from which r_{opt} increases steadily with energy.

Figure 4.5 shows the dependence of r_{opt} with zenith angle for the same array spacings and three different input energies: $log(E/eV) = 18.5, 19.0$ and 19.5. Both, events with and without saturated detectors are included. It can be seen that r_{opt} is almost independent of zenith angle for $\theta \lesssim 30^o$ for any array spacing. However, as the zenith angle increases beyond 30^o, r_{opt} decreases with θ, independently of array spacing and energy. The same effect is observed in both sets of events, those with saturation (Figure 4.6.a) and without it (Figure 4.6.b). This result comes from the fact that, for inclined showers, the array

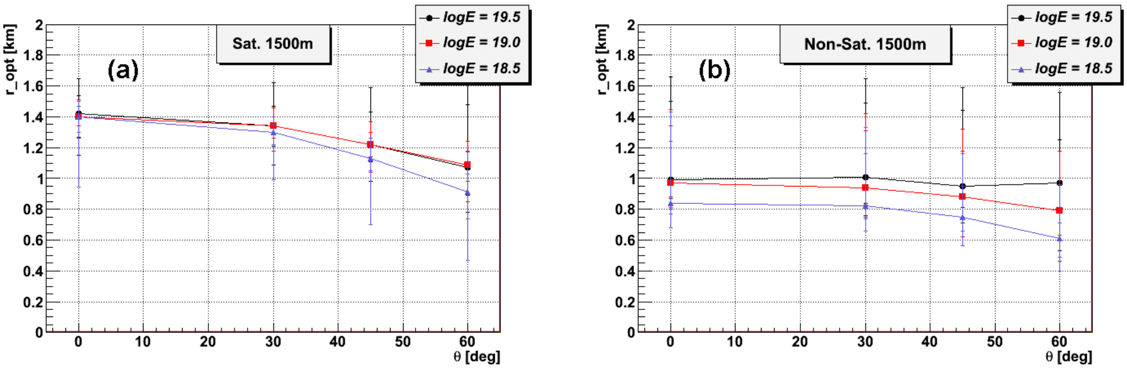

Figure 4.6: r_{opt} vs. zenith angle for 1500 m separation array. The error bars represent the 68% and 95% C.L. (a): Events with saturated detectors. (b): Events without saturated saturated detectors.

spacing projected onto the shower front shrinks with zenith angle and r_{opt} naturally follows this behavior. From the previous results, it is clear that r_{opt} is in general a function of energy and zenith angle for inclined showers. Furthermore, in Figures 4.2, 4.3, 4.5 and 4.6, the error bars indicate the 68% and 95% confidence levels (C.L.) and the central points correspond to the median value of r_{opt}. It can be seen that, in all cases, even if the behavior of the median curves is rather smooth, the C.L. are large and, therefore, considerable fluctuations are expected. Additionally, due to the large relative fluctuations of the signals from detectors located at large distances from the shower axis, the error distributions are skewed in general towards larger values from the median of r_{opt}. These points argues strongly in favor of an r_{opt} determined specifically for each shower since, using a fixed characteristic value, r_0, could compromise the estimation of primary energy. This possibility is analyzed in the following Section.

As it was mentioned in the introduction, although similar in character, the work in reference [187] is rather different in algorithmic approach and scope. Therefore, a comparison between results in both works is not straightforward. Nevertheless, Figures 5 in [187] can be used to some extent to crosscheck our results. Figure 5 bottom-right in [187] shows their r_{opt} as a function of energy. Despite the fact that there are indications of border effects at low energies in their calculation and that different zenith angle events are binned together, the results are similar to those in our Figure 4.3.d. Figure 4.3 shows r_{opt} for events with and without saturated detectors in the energy interval between ~ 10 and 30 EeV, for 433 (a), 750 (b), 866 (c) and 1500 m (d) spacing. It can be seen that, at 433, 750 and 866 m spacing r_{opt} is more or less independent of energy at lower energies but eventually increases steadily above a certain energy. This effect is also expected at a separation of 1500 m for energies beyond those presently plotted in Figure 4.3.d. Reference [187], on the other hand, shows results only for the array of 1500 m separation, where the same trend seems to be suggested for events with saturated detectors (see Figure 5 bottom-right of [187]). Remarkably, although their analysis extends up to 100 EeV, the same trend is not seen for events without saturation. The latter, however, may be due to the fact that in reference [187] showers with all zenith angles are mixed together which, at high energies, implies that their sample must be highly biased to very inclined events (otherwise they would present saturation), masking the effect. In fact, it can be seen from our Figure 4.6.b that, for events without saturation, r_{opt} does decrease at any energy for larger zenith angles.

Again, in Figure 5 bottom-left of [187], and despite the fact that the authors claim only a slight dependence of r_{opt} with zenith angle, we obtain a very similar result for events with saturation in Figure 4.6.a with r_{opt} decreasing markedly with increasing zenith angle. There is no agreement, however, for events without saturation, where they obtain an r_{opt} that increases with zenith angle, while our results (see, Figure 4.6.b) shows an r_{opt} that at low energies decreases as a function of zenith angle, but tends to a constant value as the energy increases. Part of the difference between both results may be due to the fact that in [187] energies randomly selected from a flat spectrum have been binned together. The latter, however, cannot account for their unexpected raise with zenith angle.

4.4 Influence of r_{opt} on the reconstructed energy

In this Section we analyze the effect of adopting a fix characteristic distance, r_0, instead of a shower-specific value, r_{opt}, for the determination of shower energy and energy spectrum. We simulate a detector similar to AGASA (see Section 4.2), i.e., a separation of 1 km between stations and use Eq. 4.3 and Eq. 4.5 in order to generate signals and fit the "observed" LDF respectively. For each event, r_{opt} is estimated using the procedure explained in Section 4.2 while Eq. 4.3 is used in order to estimate the energy for both $r_0 = 600$ m, as AGASA did, and r_{opt}.

Two different input spectra are used. A spectrum with one thousand events per energy bin ($\Delta \log(E) = 0.1$) from $10^{17.8}$ to $10^{20.7}$ eV, is used in order to study the functional form of the energy error distributions and the energy reconstruction bias (Sections 4.1 and 4.2). The energy reconstruction of events with saturated detectors is also analyzed. Second, in Section 4.3, a more structured spectrum extending from $10^{17.7}$ to $10^{20.5}$ eV, which possesses an ankle, a GZK-cut-off, and is exposure-limited at low energy, is used to assess the impact of both techniques in a more realistic situation. For every event, the angular distribution is extracted randomly from an isotropic distribution with a maximum zenith angle of 45^o, as in the case of the AGASA experiment. The azimuthal angles are selected from a uniform distribution between 0^o and 360^o and the core location is randomly located inside an elementary cell.

It must be noted that the results of this Section do not directly apply to the spectrum inferred from surface arrays that relay on the use of hybrid events for the energy calibration. The main reason is that the error in core location for hybrid events is much smaller than for pure surface events. In the case of Auger, for example, the error in hybrid core determination is only around 35 m [190] while for the majority of pure SD events it is estimated to be around 100 m [191]. Therefore, r_{opt} for hybrid events is very much constrained. In addition, hybrid experiments do not directly relate the signal measured at r_{opt} with the primary energy. Instead, they use a calibration with the energy obtained by the fluorescence technique. Finally, the most important uncertainties in the primary energy determination in hybrid experiments come from the fluorescence uncertainties not from the parameter size determination as will be discussed later in detail. In Chapter 6, initial steps to apply the r_{opt} technique to Auger will be explained. Essentially, it requires to find a new calibration formula to relate $S(r_{opt})$ and the energy measured by the fluorescence telescopes, E_{FD}.

4.4.1 Energy error distribution functions

We calculate the distribution functions of the error in the reconstructed energy, i.e. the difference between the reconstructed and the real energy, as a function of the injected energy for both techniques, r_0 and r_{opt}. Figure 4.7 shows, for both r_0 (a) and r_{opt} (b), the 68% and 95% C.L. for the right and left sides with respect to the median of the energy error distribution. It can be clearly seen that the error distribution functions originated by using r_{opt} are much more compact and symmetrical than the corresponding distributions for r_0. The effect is more notable for lower energies where the distribution function for characteristic distance determination is particularly wide and skewed. Although these figures are drawn for the 1000 m separation array, the results apply qualitatively for any of the other spacings considered in previous Section.

Arguably, it is desirable that the errors in energy reconstruction have a Gaussian-like distribution. Gaussian errors, for example, are easier to handle and understand when applying deconvolution techniques in the spectrum determination while assuring that there are no asymmetries or long tails, which further reduces the danger of border effects and biases associated with a rapidly changing spectral index. Again, it can be seen from Figure 4.7.b that the r_{opt} method produce, at any given input energy very approximately normal distributions.

4.4.2 Bias in the reconstructed energies

Figure 4.8 shows the relative reconstruction error as a function of the injected energy for both reconstruction techniques. In the case of events without saturated detectors (Figure

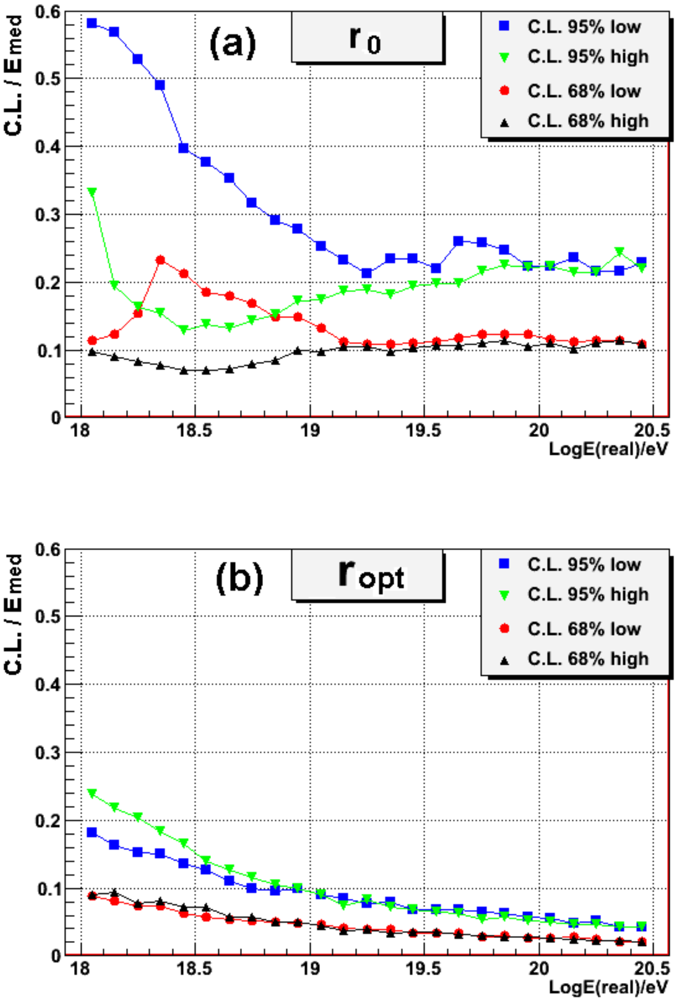

Figure 4.7: 68% and 95% CL over the median, from both its lower and higher energy sides, for the energy error distribution functions determined using either r_0 (a) or r_{opt} (b) methods. See text for more details.

4.8.b), there is no appreciable bias using r_{opt}, while using r_0 there is an energy dependent bias which, at larger energies, can reach $\sim 10\%$. The difference is much more significant in the case of events with saturated detectors (Figure 4.8.a): the r_{opt} approach produces almost negligible bias in the whole energy range while the reconstructed energy is largely underestimated using r_0. The later underlines the fact that r_{opt} is very different for both populations of events. Analogous results are obtained for any array grid size.

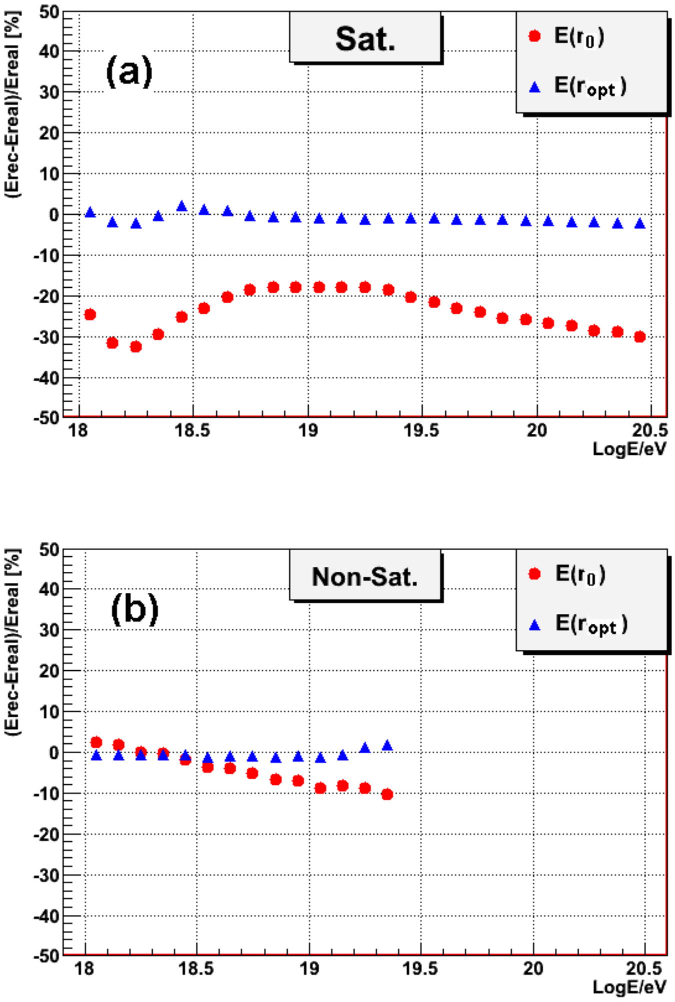

Figure 4.8: Bias in the reconstruction methods for events with (a) and without (b) saturated detectors.

The energy reconstructions of events with and without saturated detectors are shown separately for both techniques in the scatter plots of Figure 4.9. As commented before, better reconstruction is achieved using r_{opt} for both classes of events, while a more significant difference appears for events with saturated detectors.

It is should be noted that the bias in the inferred energy using r_0 is a consequence of the difference between the value assumed (here 600 m as AGASA experiment) and the actual r_{opt} value of the shower, which is the optimum distance to use as energy estimator. In the

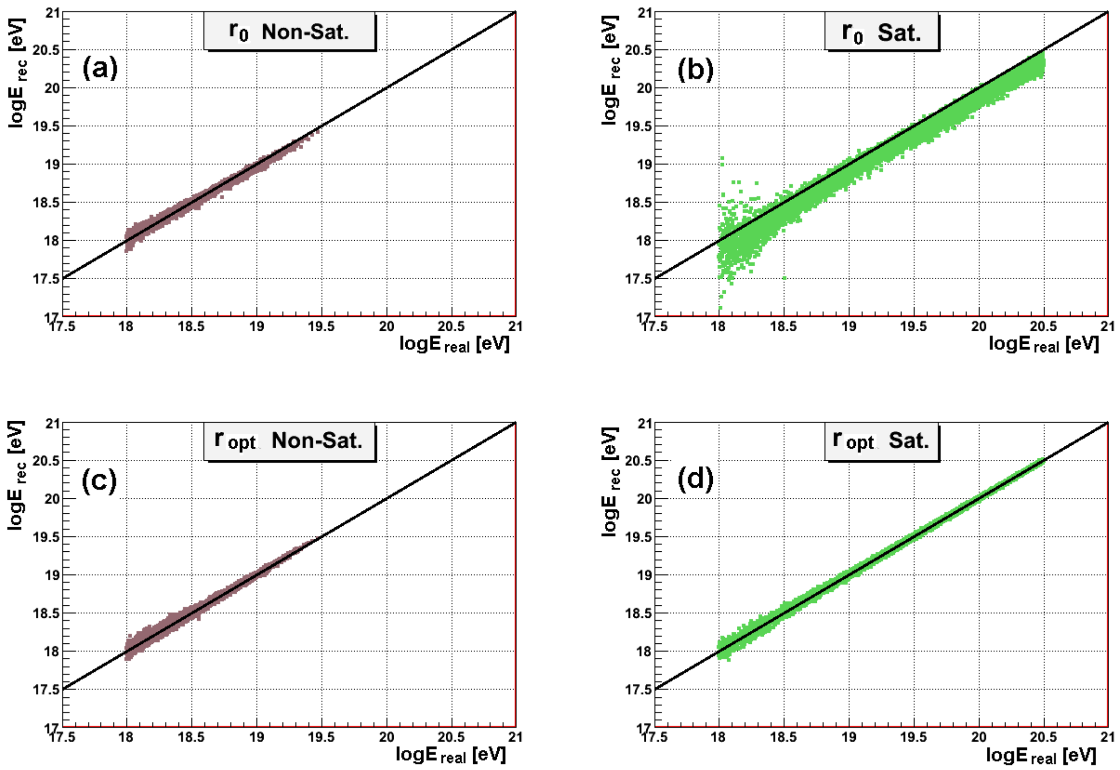

Figure 4.9: Reconstructed energy vs. real energy. Top: using $\rho(r_0)$. Down: using $\rho(r_{opt})$. Left: events without saturation. Right: events with saturation.

Fig. 4.10 the optimum distance as a function of the primary energy is shown in the scatter plot for all the events, and the C.L. of the r_{opt} distributions for events with and without saturation. The selected value of $r_0 = 600$ m is close to r_{opt} for events without saturation, but as energy increases, r_{opt} is greater than r_0 so that the energy is underestimated using the latter (as in Fig. 4.8.b). r_{opt} for events with saturated detectors is close to array separation and increases at higher energies (as in Section 4.3). Therefore, using the signal at 600 m, the inferred energy is largely underestimated (Fig. 4.8.a). If another value for r_0 were used, the bias would change (for example for $r_0 = 1000$ m the bias is shown in Fig. 4.11), but the existence of two different populations with systematically different optimum distances and the dependence of this distance with energy involve that, whatever the value selected, an important bias would exist for a significant fraction of the events.

A main advantage of the proposed method is the treatment of events with saturated

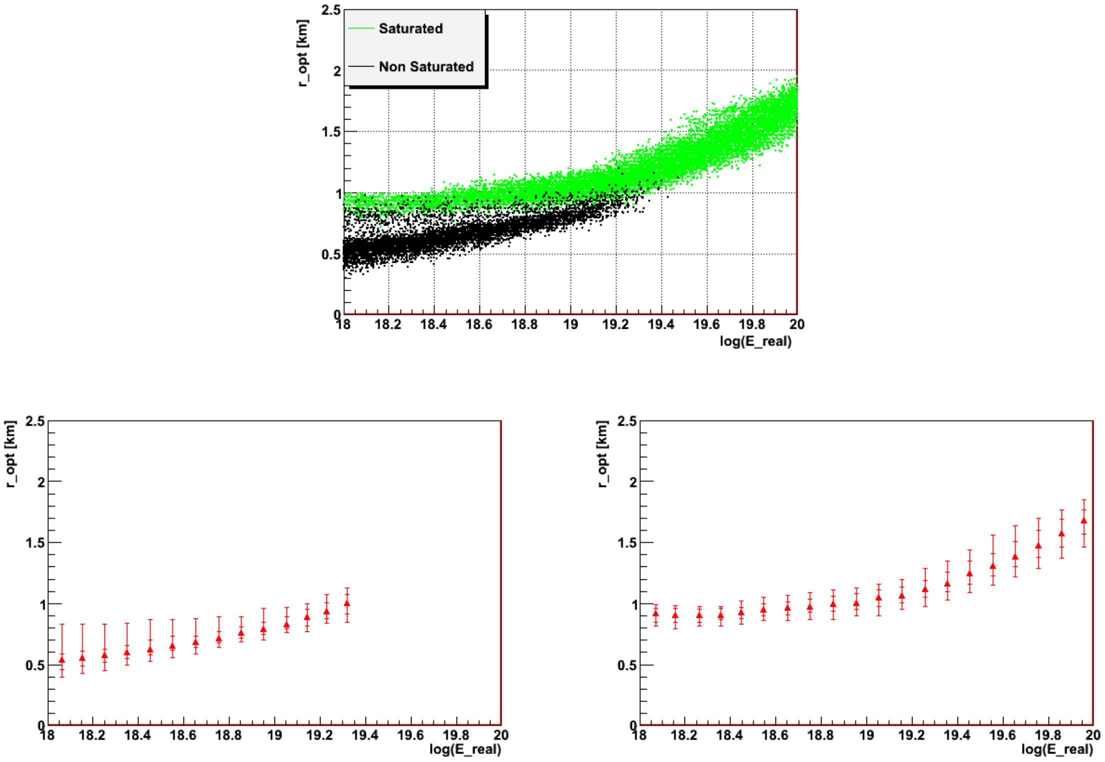

Figure 4.10: Optimum distance as a function of primary energy in a 1 km array. Top: scatter plot, events with and without saturation marked. Left (Right): 68 ad 95% C.L. of the r_{opt} distributions for events without (with) saturated detectors.

detectors, which is shown for r_0 and r_{opt} in Figures 4.9.b and 4.9.d respectively. Using a fix characteristic value r_0, events with saturation are poorly reconstructed, specially at lower energies. Essentially, the main problem is that these events have very few triggered stations and almost at the same distance from core. Consequently, their reconstruction accuracy is quite poor – this is particularly true for the fit to the LDF. In practice, using the r_0 approach, these events probably would not pass the usual quality cuts and would be discarded, or would be reconstructed with a specific procedure. Nevertheless, the r_{opt} approach makes it possible to infer without almost any bias the energy of all the events, with an accuracy comparable to that attained for events without saturation.

The advantage of a homogeneous treatment for all classes of events is further stressed by the fact that events with saturated detectors are in general dominant for most of

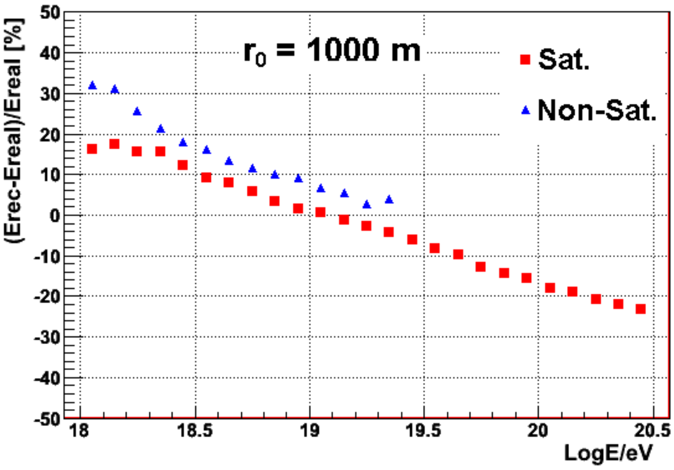

Figure 4.11: Bias for events with a without saturation if $r_0 = 1000$ m were used.

the operational range of an experimental array, regardless of the detector separation (see Figure 4.12). For example, considering the 1 km separation array used in the present Section, the number of triggered detectors in an event varies from 5 to 60 for showers from $10^{17.5}$ to $10^{19.5}$ eV. Considering an incoming event of $E \sim 10^{19.5}$ eV and a zenith angle of $\theta \sim 30$ degrees, any detector located at < 550 m from the shower axis would be saturated. Thus, independently of the position of the core inside the elementary cell, almost 100% of the events have at least one saturated detector at this energy. At still higher energies, even 2 or 3 detectors could be saturated. Furthermore, for the same spacing, 50% of the events will have at least one saturated detector above $E \sim 6$ EeV (see Figure 4.12.a. or 4.14).

4.4.3 Reconstruction of a rapidly changing spectrum

In the previous Section we demonstrated that the energy error distribution functions produced by the r_0 method are wider, more skewed and have more extended tails than those produce by r_{opt}. In principle, depending on the magnitude of these differences, they could affect the determination of spectral features, specially if the spectral index is varying rapidly over a narrow energy interval such as, for example, the ankle region and

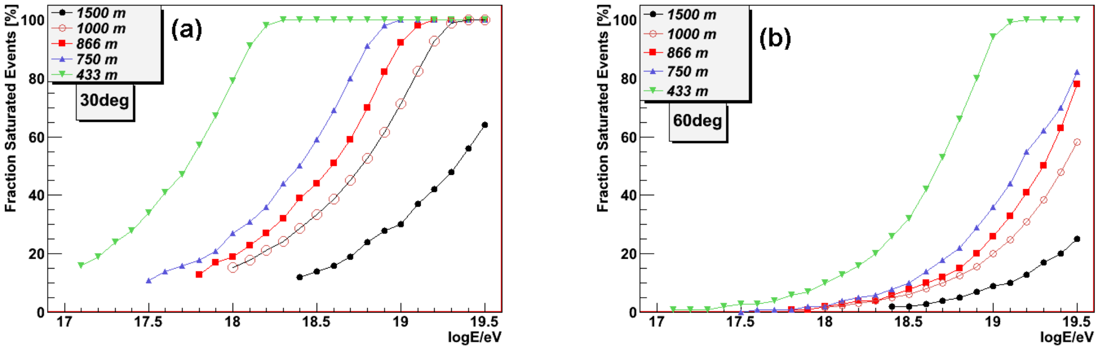

Figure 4.12: Fraction of events with saturated detectors as a function of energy for the different array spacings considered. (a) $\theta = 30^o$. (b) $\theta = 60^o$.

beyond.

In order to assess the potential effects of using either technique for the reconstruction of a structured spectrum with rapid changes as a function of energy, we use the following semi-analytical example. An idealized sectional continuous spectrum is assumed, that resembles the main spectral structures above 10^{18} eV: the ankle, the GZK flux suppression and a smooth low energy cut-off reflecting the discreteness of the surface array. The two latter suppressions in the spectrum are represented by functions of type tanh() of the input energy, while the remaining structures are represented by power laws separated by abrupt discontinuities in the first derivative.

In order to reproduce analytically the energy error distribution functions given in Figure 4.7 as a continuous function of energy, we fit our simulation results with an Asymmetric Generalized Gaussian function (AGG):

$$P_{AGG}(y) = \begin{cases} \left(\frac{c\gamma_a}{\Gamma(1/c)}\right) \exp\{-\gamma_l^c[-(y-\mu)]^c\} & if\ y < \mu \\ \left(\frac{c\gamma_a}{\Gamma(1/c)}\right) \exp\{-\gamma_r^c[(y-\mu)]^c\} & if\ y \geq \mu \end{cases}$$

where,

$$\gamma_a = \frac{1}{\sigma_l + \sigma_r}\left(\frac{\Gamma(3/c)}{\Gamma(1/c)}\right)^{1/2} \quad \gamma_l = \frac{1}{\sigma_l}\left(\frac{\Gamma(3/c)}{\Gamma(1/c)}\right)^{1/2} \quad \gamma_r = \frac{1}{\sigma_r}\left(\frac{\Gamma(3/c)}{\Gamma(1/c)}\right)^{1/2}$$

and σ_l^2 and σ_r^2 are the variances of the left and right sides respectively of the probabil-

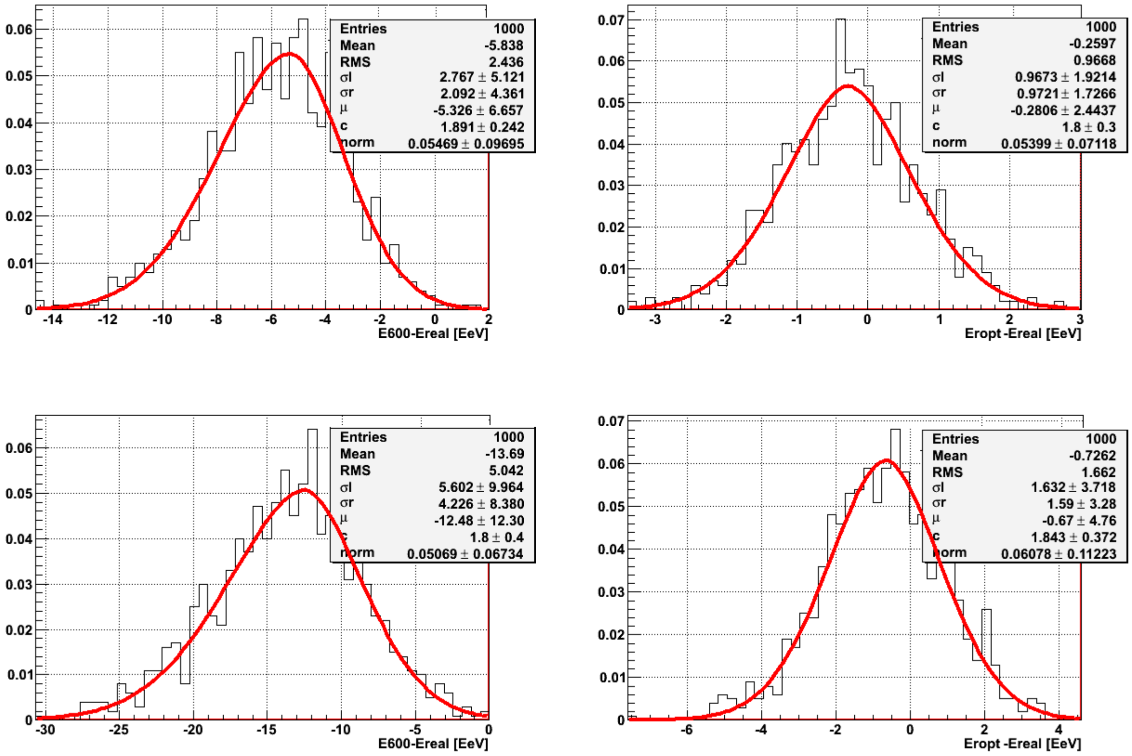

Figure 4.13: Examples of fits of the energy error distributions with the AGG function. Left: two samples using r_0. Right: two samples using r_{opt}.

ity density function and $\Gamma(x)$ is the Gamma function. If $\sigma_l^2 = \sigma_r^2$ AGG is symmetric. Furthermore, if $\sigma_l^2 = \sigma_r^2$ and $c = 2$, AGG reduces to the regular Gaussian distribution function and, for $c = 1$, it represents the Laplacian distribution.

The error functions determined previously in Section 4.4.1 have been fitted using the AGG function for the both techniques: $r_0 = 600$ m and the shower-specific r_{opt}. In the latter case the fit reduces very nearly to a Gaussian distribution function while, for r_0, the best simultaneous fit to the right and left σ_{68} and σ_{95} C.L. shown in Figure 4.7, is obtained for different values of σ_l and σ_r. In this way we can reproduce the asymmetries present on the error distribution functions and analytically map real energies onto reconstructed ones over the whole energy range of the spectrum. Examples of these fits are shown in Figure 4.13.

The results are shown in Figure 4.14. It can be seen that, if both events with and

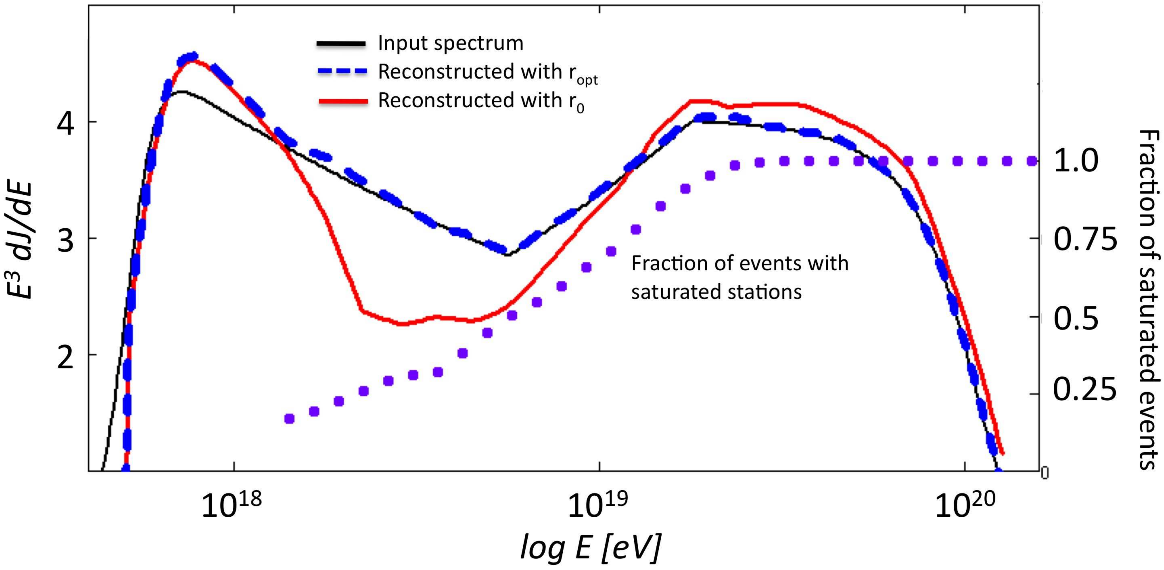

Figure 4.14: Input model spectrum (black/thin line) and the reconstructed spectra using $\rho(r_0)$ (red) and $\rho(r_{opt})$ (blue/dashed line) as energy estimators. The right axis shows the fraction of isotropic events between $\theta = 0^o$ and 45^o with saturation as a function of energy (thick dotted line) for the same array with 1000 m separation.

without saturation are lumped together, the large wings associated with the error distribution functions of the r_0 approach significantly distort the spectral features. In this particular example, the ankle, is widen and shifted, while the bump and GZK suppression are shifted upwards and much more pronounced. The r_{opt} approach, on the other hand, fits very tightly the original spectrum with the exception of very low energies, near the full efficiency edge, due to border effects. The r_0 approach can also give an equivalent fit, although noisier, if only events without saturation are used. However, the decrease in statistics by neglecting events with saturated detectors (cf., the fraction of events with saturation – Figure 4.14, right vertical axis) is so drastic that the reconstructed spectrum is only limited to a much shorter energy interval well before the GZK suppression.

In order to understand the relative magnitude of these effects, one must note that at the AGASA experiment [56], for example, the systematic uncertainty in energy determination is around 18%. Three different sources of uncertainties combine to give this value. The first one is related with the detector, mainly its linearity (±7%) and response (±5%). Second, the uncertainties coming from the lack of knowledge of the LDF (±7%), the attenuation curve used (±5%), the shower front structure and delayed particles (which contribute ±5% each). Finally the relation of $\rho(r_0)$ with energy (due to the hadronic interaction model supposed, simulation codes, chemical composition etc.), introduce an uncertainty of ±12%. In addition, they find an underestimation of 10% in reconstructed energies due to energy calibration with $\rho(r_0)$, which is compensated by the overestimation due to the shower front structure and delayed particles (5% each one). We proposed that the uncertainties related to the LDF and $\rho(r_0)$ determination could be reduced by using an r_{opt} determined on a shower to shower basis. However, while this may be a significant improvement, the other uncertainties would still dominate.

The Auger Observatory, a hybrid detector, reports [163] that, in its case, the largest uncertainties come from the fluorescence yield (±14%), the absolute calibration of FD (±10%) and the FD reconstruction method (±10%). Systematic uncertainties from atmospheric aerosols, the dependence of the fluorescence spectrum on temperature and on humidity are each at the 5% level. These uncertainties added in quadrature give a total uncertainty of 22% in fluorescence energy determination. Therefore, in addition to the fact that the method proposed here does not affect directly hybrid energy reconstruction because of the improved accuracy in the determination of the core position, the total uncertainty in the spectrum determination for hybrid experiments is widely dominated by FD uncertainties.

4.5 Summary and discussion

The primary CR energy is generally estimated in surface arrays by interpolating the lateral distribution function of particles in the shower front at ground level at a fixed distance r_0 from the shower core. This parameter is assumed to be predominantly dependent on the detector separation distance for a given layout geometry and, therefore, is considered as a constant for a given array. In this work we propose an algorithm to evaluate an equivalent, but shower-to-shower optimal distance, which we call r_{opt}. We have performed a thorough analysis of the dependence of r_{opt} on energy and zenith angle, and demonstrate that, contrary to reference [187], these are not negligible factors. In fact, not taking into account an event-specific r_{opt}, produce wider error distribution functions that can even affect the reconstruction of a highly structured, rapidly varying spectrum. The shower-to-shower r_{opt} approach, on the other hand, is an unbiased estimator of the CR primary energy, which produce also narrower, symmetric, almost Gaussian error distribution functions for energy reconstruction. Those properties of r_{opt} can additionally lead to much more reliable spectral reconstruction. The differences emerging from the two procedures, r_0 vs. r_{opt}, when applied to spectral reconstruction may have astrophysical implications, specially in the coming era of improved precision.

An important aspect of the r_{opt} approach is that it has the additional advantage of allowing the same unified treatment for events with and without saturated detectors; something that, in the r_0 approach is generally not possible, requiring either the selection of events through quality cuts, or the separate reconstruction with different techniques of the two types of events. Since the fraction of events presenting saturation is a rapidly increasing function of energy, the later greatly reduces the effective energy range for spectral reconstruction in almost all practical situations.

For practical application to real experiments, a proper calibration curve should be deduced specifically for r_{opt}, which would further optimize it as an energy estimator.

An application of these results in an square surface array as it will be the Pierre Auger North Observatory is in the Appendix A.

Chapter 5

A new surface parameter for composition discrimination of Ultra-High Energy Cosmic Rays

A new family of parameters intended for composition studies is presented. They make exclusive use of surface data combining the information from the total signal at each triggered detector and the array geometry. Here, the study is applied to the Pierre Auger South Observatory, i.e. an array of water Cherenkov detectors located in a hexagonal grid of 1500 m spacing. We perform analytical and numerical studies of these composition estimators in order to assess their reliability, stability and possible optimization. The effects of experimental uncertainties, intrinsic fluctuations and reconstruction errors are discussed in a quantitative way. The stability of the parameter in face of a possible underestimation of the size of the muon component by the shower simulation codes, as it is suggested by experimental evidence, is also analyzed. The potential discrimination power of an optimized realization of these parameters is compared on a simplified, albeit quantitative way, with that expected from other surface and fluorescence estimators in analogous experimental conditions.

5.1 Motivation

Ultra-high energy cosmic rays (UHECR) produce few observables. They are, basically, the arrival direction, the energy and some statistical hint about the identity of the primary particle. Of these three pieces of information, the geometrical one is the most reliable [192]. The energy of the shower can be inferred with an accuracy of around 20% in the case of stereo or hybrid fluorescence reconstruction [193] and this accuracy can be transferred to surface arrays of detectors making use of hybrid events which are observed simultaneously by both techniques [163]. This cross calibration already highlights the existence of yet either unidentified problems with our understanding of the physics involved in shower generation and development or inconsistencies in our implementation of those physical processes into the available shower simulation codes. Most of these problems certainly have their roots in the uncertainties associated with the extrapolation of cross sections, multiplicities and inelasticities from accelerator measurements at much lower energies required to treat the first hadronic interactions suffered by the incoming cosmic ray in the upper layers of the atmosphere. Given the indirect nature of the detection of cosmic rays at the highest energies, those uncertainties permeate, to a larger or lesser extent, all measurements done afterwards. In particular, they have their strongest manifestation in mass composition tracers, since variations in cross section or inelasticity can easily be misinterpreted as changes in baryonic composition.

There are two main observation techniques, fluorescence and surface detection, and they have specific composition indicators (a detailed review is in Section 3.4). The most reliable technique at present for composition studies is fluorescence, where the longitudinal development of the charged component of the atmospheric shower is measured. Differences in composition, manifest themselves through differences in the cross section for interactions with atmospheric nuclei. These, in turn, are mapped as different depth of maximum development of the electromagnetic component in the atmosphere, X_{max}, and as dispersion in the position of this maximum depth ΔX_{max}. If, for example, proton and Iron primaries are compared, the smaller cross section of the former will produce larger

X_{max} and ΔX_{max} than for protons [154, 171]. However, unforeseen changes in cross section as a function of energy can affect these parameters in much the same way as true changes in composition would.

Surface detectors, on the other hand, sample the lateral distribution function of extensive air showers (EAS) at discrete points while they traverse the ground level. Beyond a few tens of meters from the shower axis, the particle content of the shower at ground level is dominated by just two components, electromagnetic (i.e., electrons, positrons and photons) and muonic. These two sets of particles propagate in a different way through the atmosphere: the electromagnetic components propagates diffusively, while the muons do so radially from the last hadronic interaction that produce their parent mesons. Therefore, in a simplified way, the shower front can be thought as the composition of two fronts, a muonic one, temporally thin that arrives first and an electromagnetic disk, more extended in time that follows the muon front. Furthermore, the muon shower front has a much better defined curvature radius, that is larger than the curvature radius of the electromagnetic front. One of the practical effects of these differences is that information about the relative abundance of both components inside the shower is distributed between the slope of the lateral distribution function (LDF) of particles and the radius of curvature and time structure of the shower front. Therefore, information about the identity of the primary particle is also distributed in a non-trivial way throughout these parameters. Thus, several parameters have been proposed to extract composition information from the surface measurements of EAS, e.g., the slope of the LDF, the curvature of the shower front, several indicators of the time structure at a fixed point of the shower like the rise time and fall time of the signal, and the azimuthal asymmetries in the rise time (see Section 3.4).

In general terms, fluorescence composition indicators are regarded as easier to observe and interpret, as well as less prone to systematic errors than surface parameters do. However, fluorescence detectors suffer from a severely constrained duty cycle of approximately 10% of the total time available to surface detectors. This factor alone, which makes the statistics per unit time of surface arrays an order of magnitude larger than that of fluorescence detectors, gives a great attractive to search for reliable surface composition parameters.

In the present work, we propose a new surface parameter which, we argue, for the same integration time can deliver better discrimination power than X_{max}.

The proposed parameter is defined as:

$$S_b = \sum_{i=1}^{N} \left[S_i \times \left(\frac{r_i}{r_0} \right)^b \right] \qquad \text{[VEM]} \tag{5.1}$$

where the sum extends over all the triggered stations N, $r_0 = 1000$ m is a reference distance, S_i is the signal in VEM (Vertical Equivalent Muons, i.e. the signal deposited by one vertical muon in a Auger water Cherenkov tank) measured at the $i-$th station and r_i is the distance of this station to the shower axis in meters. In particular, for water Cherenkov detectors as those used in the Pierre Auger South Observatory, we demonstrate that the primary identity discrimination power goes through a maximum around $b = 3$.

The Chapter is organized as follows: in Section 5.2 an analytical discussion of the properties and stability of the new parameter is presented. Section 5.3 shows in some respects an equivalent numerical study performed with simulations and taking into account the effects of reconstruction. In Section 5.4 we perform a realistic comparative study about the reliability of the inferred proton fraction of selected samples using S_3, X_{max} and the rise time determined at 1000 m from the shower axis. Conclusions are presented in Section 5.5.

5.2 Analytical study

The parameter S_3 for a given event is constructed from the total signal in each triggered Cherenkov detector. Therefore, it depends on the normalization and shape of the lateral distribution function of the total signal. Close to the impact point of the shower, the signal is dominated by the electromagnetic particles (photons, electrons and positrons) whereas at larger distances it is dominated by muons. Fig. 5.1 shows the muon, electromagnetic and total signal in the Cherenkov detectors as a function of the distance to the shower

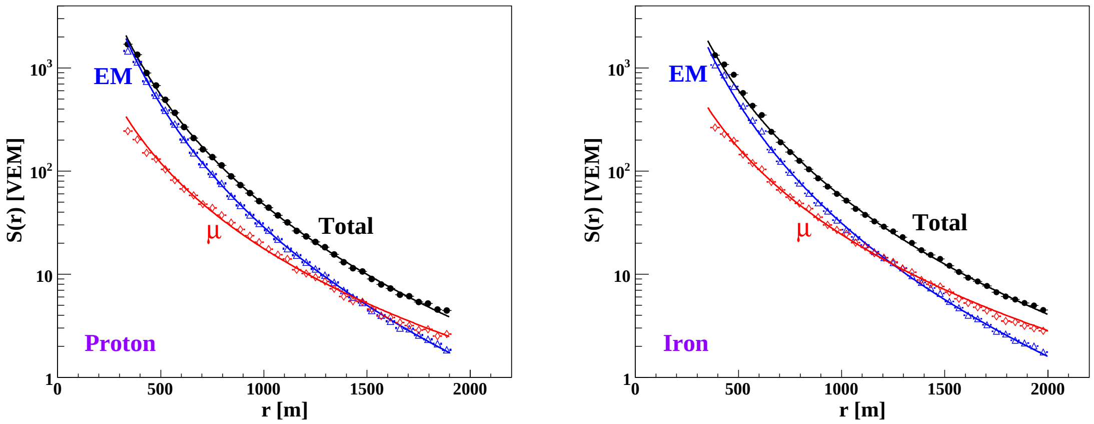

Figure 5.1: Lateral distribution functions of the muon, electromagnetic and total signal in the Cherenkov detectors for simulated protons and iron nuclei of $1 \leq \sec\theta \leq 1.2$ and $19 \leq \log(E/\mathrm{eV}) \leq 19.1$. The hadronic interaction model used to generate the showers is QGSJET-II. The solid lines correspond to the fits with a NKG-like function (see Eq. 5.2).

axis for protons and iron nuclei. The zenith angle of the simulated events considered is such that $1 \leq \sec\theta \leq 1.2$ and the primary energy $19 \leq \log(E/\mathrm{eV}) \leq 19.1$ (see Section 5.3 for details about the simulations). The hadronic model considered is QGSJET-II. Fig. 5.1 also shows the fits of the LDF of each component with a NKG-like function [79, 81]:

$$S(r) = S_0 \left(\frac{r}{r_0}\right)^{\beta} \left(\frac{r+r_s}{r_0+r_s}\right)^{\beta} \tag{5.2}$$

where we fix $r_s = 700$ m and $r_0 = 1000$ m, and S_0 and β are free fit parameters. If we consider proton and iron primaries, the discrimination power of a mass sensitive parameter q, like S_b, can be estimated by using the so-called merit factor

$$\eta = \frac{E[q_{fe}] - E[q_{pr}]}{\sqrt{Var[q_{fe}] + Var[q_{pr}]}} \tag{5.3}$$

where $E[q_A]$ and $Var[q_A]$ are the mean value and the variance, respectively, of the distribution function of parameter q_A with $A = pr, fe$. Note that an alternative definition for the merit factor makes use of the median instead of the mean value and, instead of the variance, $\sigma^2_{68}[q] \equiv [(q_{84}-q_{16})/2]^2$, where q_{84} and q_{16} are the quantiles corresponding to 84% and 16% of probability, respectively. We prefer the second definition because it takes into account possible skewness and asymmetries of the distribution functions. Nevertheless, we use here the definition as it is in Eq. 5.3 to make possible the analytical approach.

5.2.1 Optimization assuming Auger tanks

Assuming that the fluctuations of the total signal in an Auger Cherenkov detector are Gaussian, the distribution function for a given configuration of triggered stations is given by

$$P(s_1,\ldots,s_N;r_1,\ldots,r_N) = \frac{f(r_1,\ldots,r_N)}{(2\pi)^{N/2}\prod_{i=1}^{N}\sigma[S(r_i)]} \exp\left[-\sum_{i=1}^{N}\frac{(s_i - S(r_i))^2}{2\;\sigma^2[S(r_i)]}\right] \tag{5.4}$$

where r_i is the distance to the shower axis of the $i-$th station (the first station, r_1, is the closest one), $S(r_i)$ is the average LDF evaluated at r_i, $\sigma[S(r_i)] = 1.06\ [S(r_i)/\mathrm{VEM}]^{1/2}$ VEM [81] and $f(r_1,\ldots,r_N)$ is the distribution function of the distance of the different stations to the shower axis. Note that just two of the random variables $\{r_1,\ldots,r_N\}$ are independent, for instance, choosing r_1 and r_2 (the first and second closest stations) as the independent ones, we can write $f(r_1,\ldots,r_N) = f_{1,2}(r_1,r_2)\delta(r_3 - r_3(r_1,r_2))\ldots\delta(r_N - r_N(r_1,r_2))$, where $\delta(x)$ is the Dirac delta function.

From Eqs. 5.1 and 5.4 we obtain the expressions for the mean value and the variance of S_b:

$$\begin{aligned} E[S_b] &= \sum_{i=1}^{N} E\left[S(r_i)\left(\frac{r_i}{r_0}\right)^{b}\right], && (5.5)\\ Var[S_b] &= 1.06^2\sum_{i=1}^{N} E\left[S(r_i)\left(\frac{r_i}{r_0}\right)^{2b}\right] + \\ & \sum_{i=1}^{N}\sum_{j=1}^{N} cov\left[S(r_i)\left(\frac{r_i}{r_0}\right)^{b}, S(r_j)\left(\frac{r_j}{r_0}\right)^{b}\right] && (5.6)\end{aligned}$$

where the variables s_i have been already integrated and

$$
\begin{aligned}
E\left[S(r_i)\left(\frac{r_i}{r_0}\right)^x\right] &= \int dr_i\ S(r_i)\left(\frac{r_i}{r_0}\right)^x f_i(r_i) \qquad (5.7)\\
cov\left[S(r_i)\left(\frac{r_i}{r_0}\right)^b, S(r_j)\left(\frac{r_j}{r_0}\right)^b\right] &= \int dr_i\ dr_j\ S(r_i)\left(\frac{r_i}{r_0}\right)^b S(r_j)\left(\frac{r_j}{r_0}\right)^b \times\\
& \quad f_{i,j}(r_i, r_j) \qquad (5.8)
\end{aligned}
$$

Here $f_i(r_i)$ is the distribution function of the distance to the shower axis for the $i-$th station and $f_{i,j}(r_i, r_j)$ is the distribution function of the distance to the shower axis of the $i-$th and $j-$th stations:

$$
f_{i,j}(r_i, r_j) = \int dr_1 \ldots dr_{i-1} dr_{i+1} \ldots dr_{j-1} dr_{j+1} \ldots dr_N\ f(r_1, \ldots, r_N) \qquad (5.9)
$$

In order to simplify the expressions for the mean and variance of S_b, we perform the following approximations:

$$
\begin{aligned}
E[g(r_i)] &\cong g(E[r_i]), \qquad (5.10)\\
cov[g(r_i), g(r_j)] &\cong \left.\frac{dg}{dr}\right|_{E[r_i]} \left.\frac{dg}{dr}\right|_{E[r_j]} cov[r_i, r_j] \qquad (5.11)
\end{aligned}
$$

where $g(r) = S(r)(r/r_0)^b$. Thus, we get:

$$
\begin{aligned}
E[S_b] &= \sum_{i=1}^{N} \left[S(E[r_i])\left(\frac{E[r_i]}{r_0}\right)^b\right] \qquad (5.12)\\
Var[S_b] &= 1.06^2 \sum_{i=1}^{N} \left[S(E[r_i])\left(\frac{E[r_i]}{r_0}\right)^{2b}\right] + \qquad (5.13)\\
& \sum_{i=1}^{N}\sum_{j=1}^{N} \left[\frac{\partial}{\partial r_i}\left(S(r_i)\left(\frac{r_i}{r_0}\right)^b\right)\Bigg|_{E[r_i]} \frac{\partial}{\partial r_j}\left(S(r_j)\left(\frac{r_j}{r_0}\right)^b\right)\Bigg|_{E[r_j]} cov[r_i, r_j]\right] \qquad (5.14)
\end{aligned}
$$

We already have analytical expressions for the average LDFs of proton and iron primaries obtained by fitting the simulated data (Fig. 5.1). The other ingredients needed to calculate the mean value and the variance of S_b are the mean values of the distance to the shower axis for the different stations and the covariance between all pairs of those random variables. We obtain these quantities from a simple Monte Carlo simulation: we

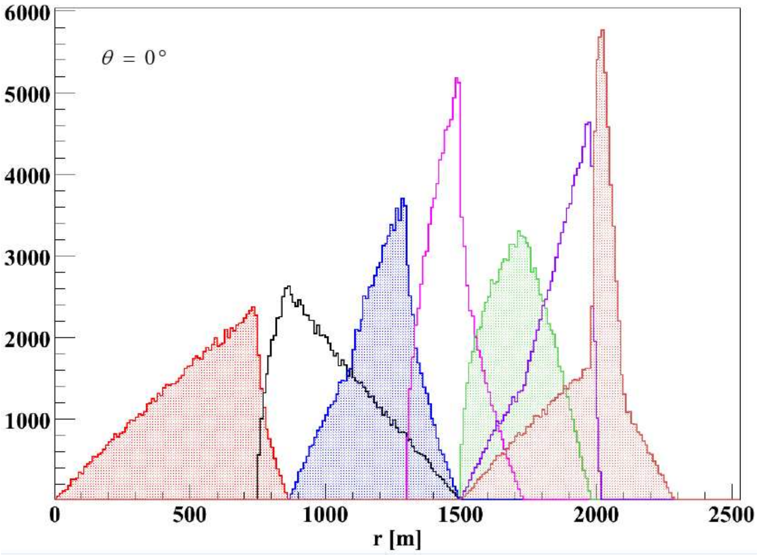

Figure 5.2: Distance of the stations to shower axis for almost vertical showers in a triangular grid of 1.5 km of spacing.

uniformly distribute impact points in a triangular grid of 1.5 km of spacing, like the Auger array, and then, for each event, of zenith angle such that $\sec\theta = 1.1$ and azimuthal angle uniformly distributed in $[0, 2\pi]$, we calculate the distance of each station to the shower axis. The result is shown in Fig. 5.2. From these distributions $E[r_i]$ and $cov[r_i, r_j]$ are easily determined.

Finally, we have all the ingredients needed to calculate the mean and the variance of S_b, and therefore, the merit factor. Thus, Fig. 5.3 shows the discrimination power η as a function of b obtained under the mentioned assumptions and simplifications. We see that η reaches the maximum at $b \cong 3$.

5.2.2 Modifying the slope of the LDF

We also study the discrimination power of S_b when the slope parameter β is modified but keeping constant the integrated signal for distances larger than the Moliere radius (we select $r_M = 80$ m). Thus, the measured energy of the event by surface experiments would not be significantly affected. The modified LDF that fulfills this condition can be written

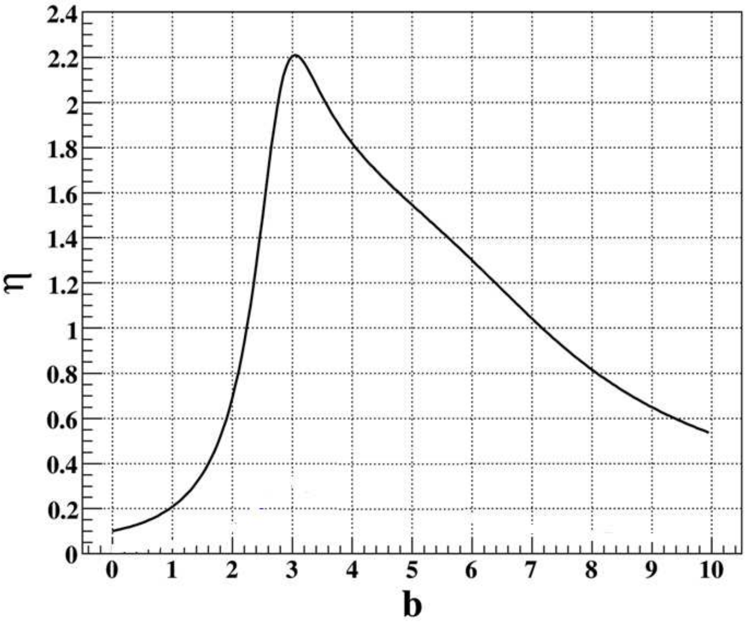

Figure 5.3: η as a function of b for vertical showers ($1 \leq \sec\theta \leq 1.2$) and $19 \leq \log(E/eV) \leq 19.1$. η reaches the maximum at $b \cong 3$.

as

$$S(r,\beta) = \frac{N(r_M, r_0, r_s, \beta_0)}{N(r_M, r_0, r_s, \beta)}\, S_{\beta_0}(r) \tag{5.15}$$

where

$$N(r_M, r_0, r_s, \beta) = \frac{r_s^{2+2\beta}}{r_0^{\beta}(r_s + r_0)^{\beta}}\ \text{Beta}(-r_s/r_M, -2(1+\beta), 1+\beta) \tag{5.16}$$

and

$$\text{Beta}(z,a,b) = \int_0^z dt\ t^{a-1}(1-t)^{b-1} \tag{5.17}$$

and $S_{\beta_0}(r)$ is the LDF of Eq. 5.2 with the parameters S_0 and β_0 originally obtained from the fits in Fig. 5.1.

The slope of the proton LDF is smaller than the corresponding to iron (the absolute value is greater). Then, we modify the slope of both LDFs such that, $\beta_{pr}(\xi) = \beta_{pr}^0 - (\xi - 1)\Delta\beta_0/2$ and $\beta_{fe}(\xi) = \beta_{fe}^0 + (\xi - 1)\Delta\beta_0/2$, where β_{pr}^0 and β_{fe}^0 are the proton and iron slopes, respectively, obtained from the fits of the simulations, $\Delta\beta_0 = \beta_{fe}^0 - \beta_{pr}^0$ and ξ is such that $\Delta\beta(\xi) = \xi\Delta\beta_0$, i.e., $\xi = 1$ corresponds to the non modified case. Note that for $\xi = 0$, $\beta_{pr} = \beta_{fe} = (\beta_{pr} + \beta_{fe})/2$. The mean and the variance of S_b are calculated with

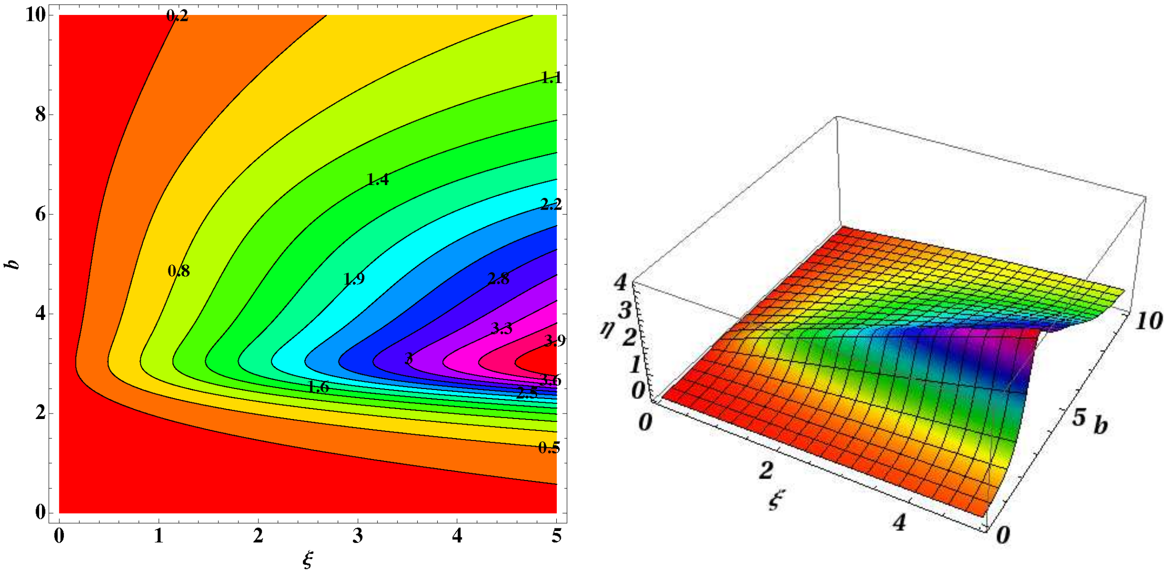

Figure 5.4: Merit factor of S_b as function of b when modifying the slope of the Iron and proton LDFs. See text for details. Left: Contour plot $\eta(\xi, b)/\eta(1, 3)$. Right: 3D plot, η vs. ξ and b.

the same procedure as before, but now $S(E[r_i])$ used before are modified by the factor $N(r_M, r_0, r_s, \beta_0)/N(r_M, r_0, r_s, \beta)$.

Fig. 5.4-left shows a contour plot of $\eta(\xi, b)/\eta(1, 3)$ from where we see that as ξ increases η also increases. We also see that the maximum of η remains close to $b = 3$ almost independent of ξ. In Fig. 5.4-right the 3D plot η vs. (ξ, b) is shown.

5.2.3 Modifying the muon content of the simulated showers

There is experimental evidence of a deficit in the muon content of the simulated showers [149, 150]. It is believed that such deficit is originated in the high energy hadronic interaction models which are extrapolations, over several orders of magnitude, of lower energy accelerator data. As mentioned, the total signal can be decomposed in the muon and the electromagnetic signal. Therefore, in order to study how S_b changes as a function of the muon content of the showers, we modify the total LDFs in the following way, $S(r) = S_{em}(r) + fS_{\mu}(r)$, where f parametrizes the artificial variation in the muon component.

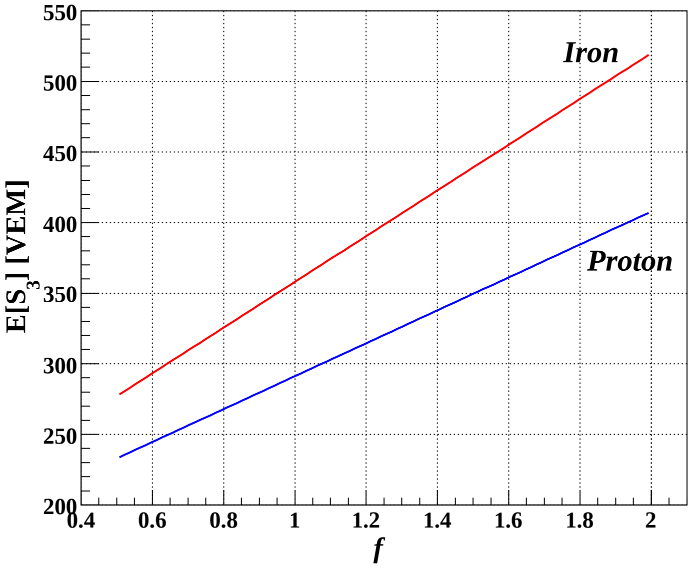

Figure 5.5: Mean value of S_3 for protons and iron nuclei as a function of f, where $f = 1$ corresponds to the muon content predicted by QGSJET-II.

Then, the mean and the variance of S_b are calculated following the same approximations as before. For example, the mean value is given by

$$E[S_b] = \sum_{i=1}^{N} E\left[S(r_i)\left(\frac{r_i}{r_0}\right)^b\right] = \tag{5.18}$$

$$\sum_{i=1}^{N} E\left[(S_{em}(r_i) + fS_\mu(r_i))\left(\frac{r_i}{r_0}\right)^b\right] \simeq \tag{5.19}$$

$$\sum_{i=1}^{N}\left[(S_{em}(E[r_i]) + fS_\mu(E[r_i]))\left(\frac{E[r_i]}{r_0}\right)^b\right] \tag{5.20}$$

and similarly for the variance. The signal of the electromagnetic and the muonic components were also fitted in Fig. 5.1 separately. Fig. 5.5 shows the mean value of S_3 for protons and iron nuclei as a function of f. As expected, they increases with f. We also see that the iron curve increases faster than the proton one, which means that, for larger values of f, the discrimination power of S_3 also increases. This happens because the muon content of the showers is very sensitive to the primary mass. Then, for large values of f the muon component becomes more important increasing the mass sensitivity of S_3.

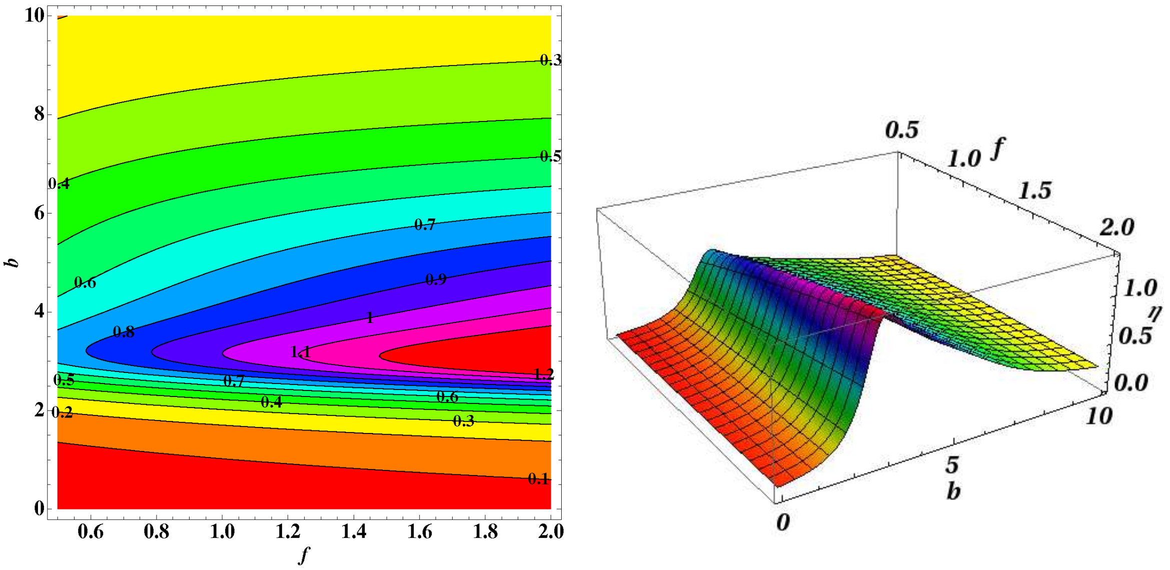

Figure 5.6: Merit factor of S_b as function of b when the number of muons is modified. See text for details. Left: Contour plot of $\eta(f,b)/\eta(1,3)$. Right: 3D plot, η vs. f and b.. $f = 1$ corresponds to the muon content of the showers predicted by QGSJET-II.

Fig. 5.6-left shows a contour plot of $\eta(f,b)/\eta(1,3)$ and Fig. 5.6-right is the 3D plot η vs. (f,b). It can be seen that the discrimination power of S_b increases with the muon content of the showers and that the maximum is reached at $b \cong 3$ almost independently of f.

5.3 Numerical analysis

The simulation of atmospheric showers is performed by using the AIRES Monte Carlo program (version 2.8.4a) [194] with QGSJET-II and Sibyll 2.1 as the hadronic interaction models. Since the number of secondary particles produced in a shower is extremely large (i.e. $\sim 10^{11}$ particles in a proton shower of 10^{20} eV), it is very costly, in processing time and disk space, to follow all of them. Therefore, we used a statistical method called thinning, first introduced by M. Hillas [151, 195], as it is implemented in AIRES. A relative thinning of 10^{-6} and weight factor of 0.2 are used for the generation of the showers. Iron and proton primaries are simulated for both hadronic interaction models, in an energy range from 10^{19} to $10^{19.6}$ eV. The zenith angle distribution of the simulated showers corresponds to an isotropic flux with $0° \leq \theta \leq 60°$. The number of showers for each primary type and hadronic interaction model is comparable to the one corresponding to the Auger spectrum published in [29], where the statistics corresponds to ~ 0.8 years of full operation.

The simulation of the surface detectors response, as well as the shower reconstruction, are performed by using the official Offline reconstruction framework of the Pierre Auger Observatory [196]. The simulation includes a triangular grid of Cherenkov detectors of 1.5 km of spacing. The unthinning method of P. Billoir [152] is used to compensate the thinning in the simulated showers. The GEANT4 package [197] is used to simulate the behavior of particles inside the tanks. The surface detector simulation has been tested and proved to be in good agreement with experimental data [198]. In order to increase the statistics, each shower is recycled 5 times by randomly distributing cores inside the array. (see [153] for a discussion of the statistical effects of recycling air showers in detector simulations).

We consider a Gaussian uncertainty of 18% in the determination of the primary energy, which corresponds to the energy uncertainty obtained from the surface data by the Pierre Auger Observatory [29]. Therefore, the *reconstructed* energy is obtained by fluctuating the real one with this Gaussian function.

The simulations are divided in logarithmic energy bins from $\log(E/eV) = 19$ to $\log(E/eV) = 19.6$ in steps of 0.1. We also consider three different zenith angle bins centered at 30°, 45°, and 55° of 10° wide.

5.3.1 Optimization and comparison with the analytical result

As in the analytical approach, we use the merit factor η to study the discrimination power of the mass sensitive parameter. However, here we prefer to use the median and σ_{68}^2 instead of the mean value and the variance respectively.

Fig. 5.7 shows the merit factor of S_b as a function of b for the hadronic interaction

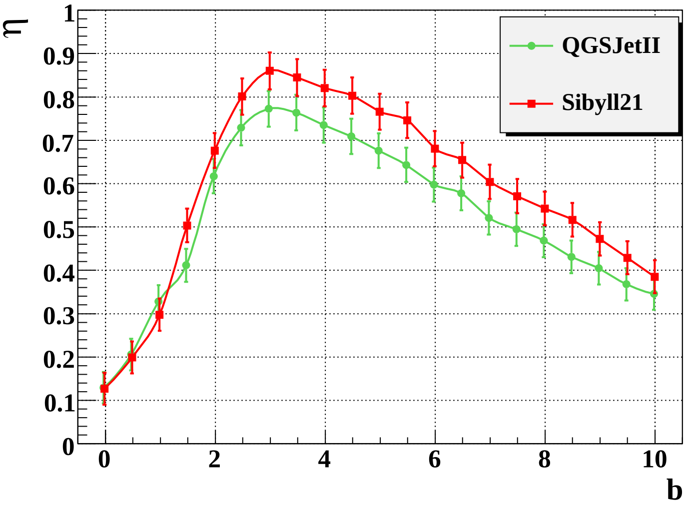

Figure 5.7: Merit factor of S_b as a function of b obtained from simulated data. The proton and iron samples,corresponding to a given hadronic interaction model (QGSJET-II or Sibyll 2.1), used to calculate η include all the events independently on their different energies and zenith angles.

models considered. The maximum is reached at $b \cong 3$ in a very good agreement with the result obtained in the analytical study (see Fig. 5.3). Furthermore, the shape of the curve is quite similar. However, the merit factor is lower and the peak is wider as expected, because in the numerical study, all the fluctuations and the effects introduced by the reconstruction methods are included. Another reason is that, in this calculation, all the simulated events, with different energies and zenith angles, are used to obtain the proton and iron samples of S_b. Hereafter, we consider S_3 (S_b with $b = 3$).

5.3.2 Influence of the detectors far from the shower axis

Several tests have been performed to study if the stations far from the shower axis, whose fluctuations in the signal are more significant, could affect the separation power of S_3. Note that the saturated stations are always rejected. Let us call r_{lim} to the maximum distance of the stations included in the S_3 sum. Five different cuts are studied:

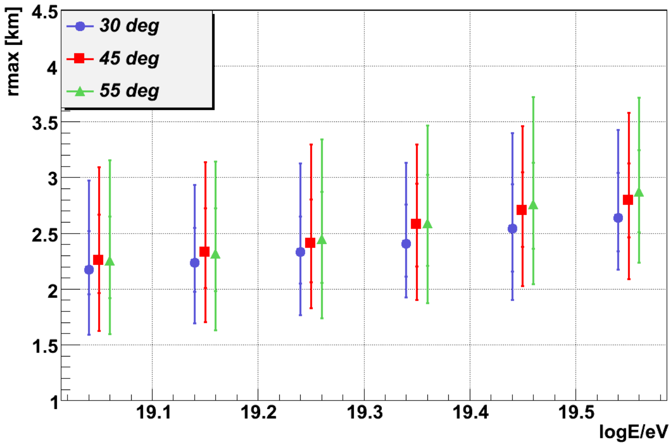

Figure 5.8: Distance of the furthest triggered station to the shower axis for three different zenith angle bins (window of 10^o wide). Both HIMs and type of primaries are included. The error bars corresponds to the regions of 68% and 95% of probability and the points are the median values.

- $r_{lim} = r_{opt}$, which is obtained trying to maximize the merit factor for each hadronic interaction model, zenith angle and energy bin.
- $r_{lim} = 2700$ m, all the stations at a distance from the shower axis larger than 2700 m are excluded.
- $r_{lim} = Med[r_{max}(E, \theta)]$. For a given energy and zenith angle, r_{max} is the distribution of the distance of the furthest station to the shower axis. We choose the median of the r_{max} distribution obtained including the proton and iron events generated with both hadronic models considered (see Fig. 5.8).
- Each term of the S_3 sum, corresponding to a given station, is weighted using the so called Lateral Trigger Probability (LTP), which gives the probability of a shower with certain energy and zenith angle to triggered a detector situated at a given distance from the shower axis. This is an elegant way of switching off smoothly the stations at large distances from the core. The LTP used is calculated from our set of simulations following the procedure described in [188].

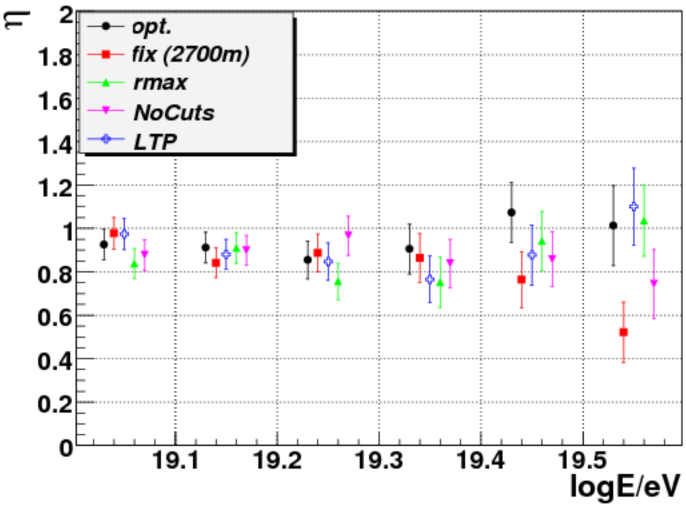

Figure 5.9: Merit factor as a function of energy for several cuts tested (see the text for details) for Sibyll 2.1 and $\theta = 55 \pm 5^{\circ}$. Similar result are obtained for QGSJET-II and other zenith angle bins.

- $r_{lim} \to \infty$: all triggered stations are included. No cut is applied.

Fig. 5.9 shows η as a function of the primary energy obtained by using the different cuts considered. It can be seen that the merit factors are comparable, showing the robustness of the parameter. Therefore, the discrimination power of S_3 is not affected by the stations far from the shower axis where the fluctuations could dominate the signal.

5.3.3 Energy and zenith angle dependence

We also study the dependence of S_3 with primary energy and zenith angle. Fig. 5.10 shows the mean value of S_3 as a function of sec(θ) for protons and Sibyll 2.1, where all energies are considered. It can be seen that there is no significant dependence with the zenith angle. Similar results are obtained for iron primaries and QGSJET-II.

Fig. 5.11 shows the evolution of S_3 with primary energy for proton and iron primaries, $\theta = 30^{\circ}$ and $\theta = 45^{\circ}$ (in a window of 10^{o} wide), and for both HIMs considered. The curves are linear fits to the simulated data. It can be seen that the medians between iron and

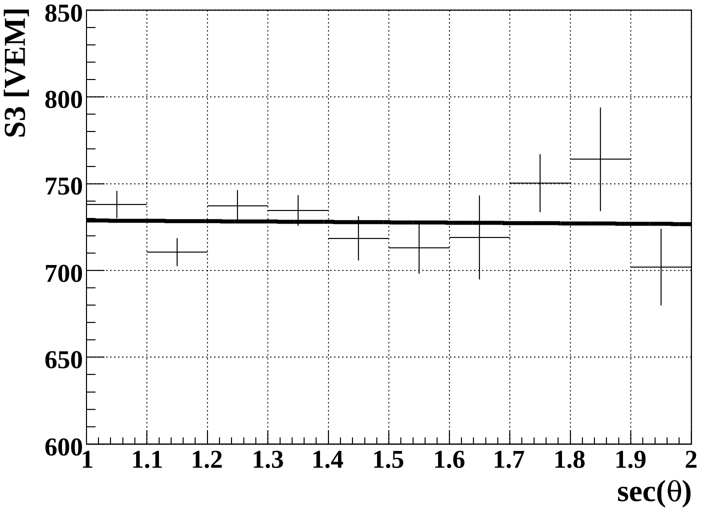

Figure 5.10: S_3 vs. $sec(\theta)$ for proton primaries and using Sibyll 2.1. A linear fit to the points shows a negligible dependence on $sec(\theta)$. Error bars are $RMS/\sqrt{N}$.

proton primaries are closer for QGSJET-II, so that the separation power of S_3 would be lower than for Sibyll 2.1, as will be shown later.

The almost linear dependence of S_3 on energy could hinder its discrimination power if there is an unknown bias in energy reconstruction. One way to circunvect this potential problem is to work with an energy-related direct observational parameter instead of the reconstructed energy. Such parameter could be the signal at a fixed distance from the axis of the shower, $S(r_0)$, or its equivalent value normalized at a certain zenith angle, S_θ, as obtained from a constant intensity cut curve. Fiducial values for the Auger experiment are S_{1000} and S_{38} for $r_0 = 1000$ m and $\theta = 38^o$ respectively [29]. Figure 5.12 is the analog to Fig. 5.11, but drawn as a function of the S_{38} instead of the reconstructed energy. The corresponding S_{38} value for each shower has been obtained from its energy by using the constant intensity cut and energy calibration reported in [163]. It can be seen that the discrimination power of S_3 remains strong for both hadronic interaction models and that, since there is an almost linear relation between S_3 and S_{38}, a valuable astrophysical insight into the evolution of composition as a function of energy can still be gained despite the fact that the exact energy profile is unknown.

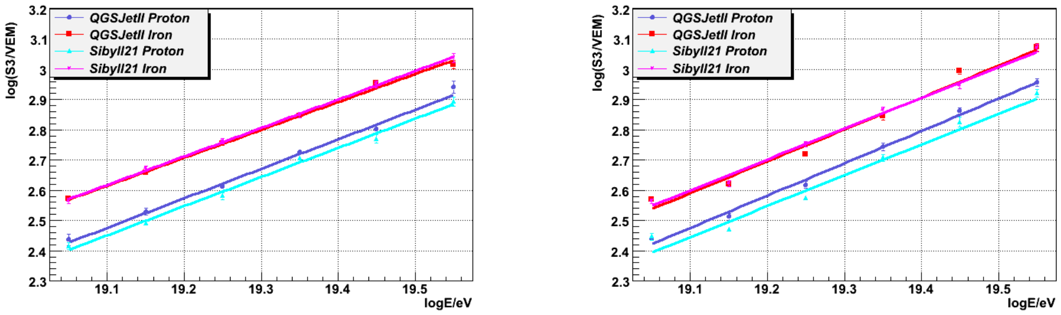

Figure 5.11: $\log(S_3/VEM)$ vs. $\log(E/eV)$ for $\theta = 30°$ (left) and $\theta = 45°$ (right) and for the two hadronic interaction models considered. The error bars are the $RMS/\sqrt{N}$.

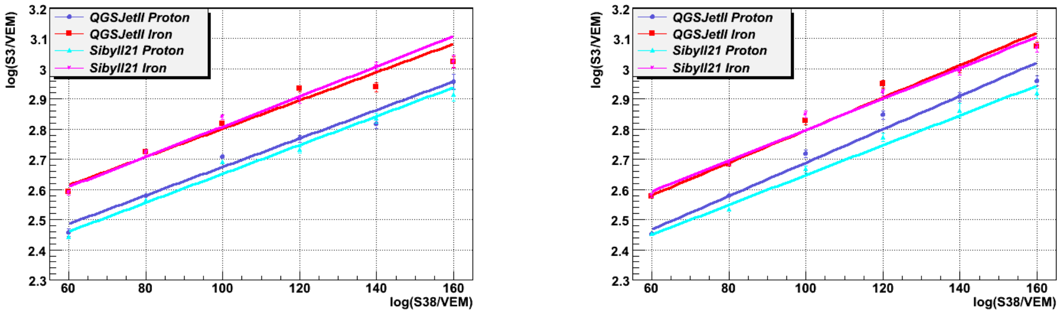

Figure 5.12: As Fig. 5.11 but now $\log(S_3/VEM)$ is plotted vs. $\log(S_{38}/VEM)$

5.4 Application

In this Section we compare the reliability of the composition determination using S_3 and other mass sensitive parameters commonly used in composition analysis. We select the two most useful parameters, one from the surface technique, i.e. the rise time at 1000 m from the shower core, and other from the fluorescence technique, X_{max}, the atmospheric depth at which the maximum development of the cascade is reached. A brief discussion follows on specific details about the determination of these two parameters:

- Rise time at $r_0 = 1000$ m from core, $t_{1/2}(r_0)$ [ns]: The procedure followed by Auger to calculate $t_{1/2}(r_0)$ was explained earlier in Section 3.4.2. First, the rise time of each station is corrected depending on its internal azimuth angle. Second, the rise time at 1000 m is obtained by fitting the corrected rise time of each triggered station using the function $t_{1/2}(r) = (40 + ar + br^2)$ ns. Parameters a and b are free in the fit. Only the stations in the range from 600 to 1500 m from the shower axis and signal greater than 10 VEM, are included in the fit. At least three stations are required. Therefore, in case of showers at large zenith angles is not unusual that there are not enough stations passing the cuts, which reduces significantly the statistics available. Consequently, although the zenith angle distribution of our simulation set is isotropic (peaked at 45°), there are more events whose $t_{1/2}(r_0)$ is available at lower zenith angles.

- X_{max} $[g/cm^2]$: In order to assign a realistic X_{max} value to our simulations, including the response of the detector and the effects of the reconstruction method, we use the value simulated internally in AIRES and fluctuate it with a Gaussian distribution with standard deviation $\sigma[X_{max}] = 20\ g/cm^2$, which is the X_{max} resolution achieved by Auger [154].

We use a maximum likelihood method to compare the reliability of the composition determination using the three parameters. We need samples with large statistics for this method. S_3 and X_{max} are almost independent on the zenith angle, so that it is possible to combine events with different θ in the same sample. Obviously, that is not the case for $t_{1/2}(r_0)$. Thus, a quadratic fit is performed, $t_{1/2}(r_0)$ vs. $sec(\theta)$, for each primary and hadronic model (see Fig. 5.13-left), and using the average values of the fitted parameters, we correct, in a simple way, the zenith angle dependence of $t_{1/2}(r_0)$:

$$t^{corr}_{1/2}(r_0, \sec\theta) = t^{meas}_{1/2}(r_0, \sec\theta) + \left[t^{fit}_{1/2}(r_0, 1.05) - t^{fit}_{1/2}(r_0, \sec\theta)\right] \tag{5.21}$$

The correction does not increase the fluctuations and $t^{corr}_{1/2}(r_0)$ shows a strong reduction on the zenith angle dependence as shown in Fig. 5.13-right. For the subsequent analysis, we consider the lowest energy bin (from 10^{19} to $10^{19.1}$ eV) where we have larger statistics. The sample for each set [primary, HIM, parameter] considered is binned. Let us call $h_p(i)$ and $h_{fe}(i)$ to the number of events in the $i-$th bin for protons and iron nuclei respectively,

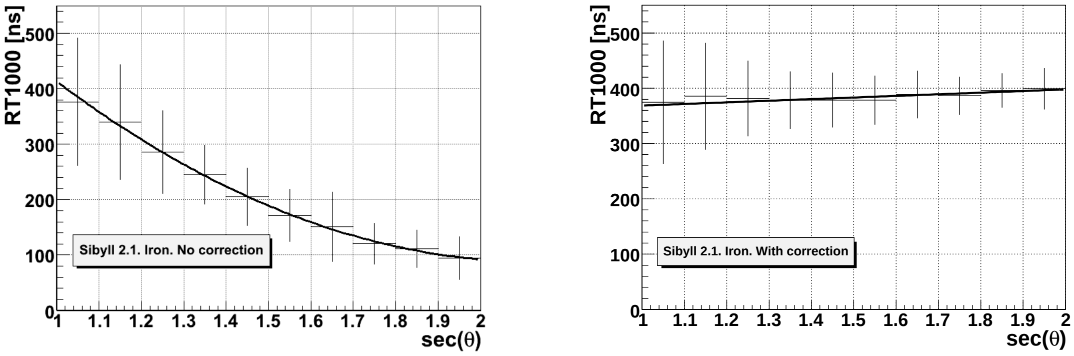

Figure 5.13: $t_{1/2}(1000)$ as function of $sec(\theta)$ for Sibyll 2.1 and iron primaries. Error bars are the RMS. Left panel: $t_{1/2}^{meas}(1000)$ without correction, the data is fitted with a quadratic function. Right panel: $t_{1/2}^{corr}(1000)$ after correction, the data is fitted with a linear function.

normalized to the total number of the events in the sample. The histograms h_p and h_{fe} are assumed to be the distribution of the universe. The proton abundance of a sample is defined as $C_p = N_p/(N_p + N_{fe})$, where N_p and N_{fe} are the number of protons and iron nuclei in the given sample. We create samples of C_p^{true} from 0 to 1 in steps of 0.1. For each value of C_p^{true}, we generate 300 sub-samples of $N_s = 300$ events each by taking them randomly from h_p and h_{fe}. For each sub-sample, we generate a histogram H_s with the same binning used in h_p and h_{fe}. H_s is not normalized so that $\sum_i H_s(i) = N_s$. Thus, assuming Poisson statistics, the probability of the sub-sample H_s is given by

$$P(\{H_s(i)\}_i) = \prod_i \left[\exp\left(-H_t(i)\right) \times \frac{H_t(i)^{H_s(i)}}{H_s(i)!} \right] \tag{5.22}$$

where

$$H_t(i) = N_s[C_p^{inf} h_p(i) + (1 - C_p^{inf}) h_{fe}(i)] \tag{5.23}$$

C_p^{inf} represents the inferred proton abundance and it is obtained by maximizing $\ln P(\{H_s(i)\})$.

Fig. 5.14 (Fig. 5.15) shows the inferred composition as a function of the true one corresponding to QGSJET-II (Sibyll 2.1). Similar results are obtained for both models but smaller error bars (which represent the 68% and 95% C.L.) for the latter, in agreement with the fact that the merit factor for all the parameters is greater for this HIM. The number of events available for each mass sensitive parameter corresponding to a given exposure time, is a key value to compare their discrimination capabilities. Due to the limited duty cycle of the fluorescence telescopes, only 10% of the events are detected, so the statistics for X_{max} are significantly lower than that of surface parameters. From Figs. 5.14 and 5.15, we see that considering the same exposure time, S_3 gives the most accurate results, i.e. its discrimination capability is greater than that of $t_{1/2}(r_0)$ and X_{max}. In order to illustrate the significance of taking into account the limited statistics for X_{max} when doing composition studies, it is also shown the result for X_{max} if the same statistics as the SD parameters were available. The error bars are reduced becoming the smallest ones, but ten times more exposure would be required.

A second study has been performed in order to extend previous results to a larger energy range. Now, a fix true proton fraction $C_p^{True} = 0.5$ is assumed and the inferred proton fraction is calculated in the energy range from $10^{19.0}$ to $10^{19.6}$ eV. In order to improve the small statistics in the higher energy bins, the distributions for each [primary, energy, HIM, parameter] are fitted using the Asymmetric Generalized Gaussian (AGG) function, defined as

$$P_{AGG}(y) = \begin{cases} \frac{c\,\gamma_a}{\Gamma(1/c)}\ \exp[-\gamma_l^c\ (-y+\mu)^c] & if\ y < \mu \\ \frac{c\,\gamma_a}{\Gamma(1/c)}\ \exp[-\gamma_r^c\ (y-\mu)^c] & if\ y \geq \mu \end{cases}$$

where

$$\gamma_a = \frac{1}{\sigma_l + \sigma_r}\left(\frac{\Gamma(3/c)}{\Gamma(1/c)}\right)^{1/2} \quad \gamma_l = \frac{1}{\sigma_l}\left(\frac{\Gamma(3/c)}{\Gamma(1/c)}\right)^{1/2} \quad \gamma_r = \frac{1}{\sigma_r}\left(\frac{\Gamma(3/c)}{\Gamma(1/c)}\right)^{1/2}.$$

which has been already used in Chapter 4.

Fig. 5.16 shows examples of the fits performed for the three parameters considered. It can be seen that it is possible to fit asymmetric distributions with longer tails compared to Gaussian distributions. The fits are very accurate, so that we can extract samples from them and it is feasible to extent the previous study to a larger energy range.

For each set of [primary, energy, HIM, parameter], the samples are generated by randomly sampling the corresponding fitting function. Thus, we generate the histograms

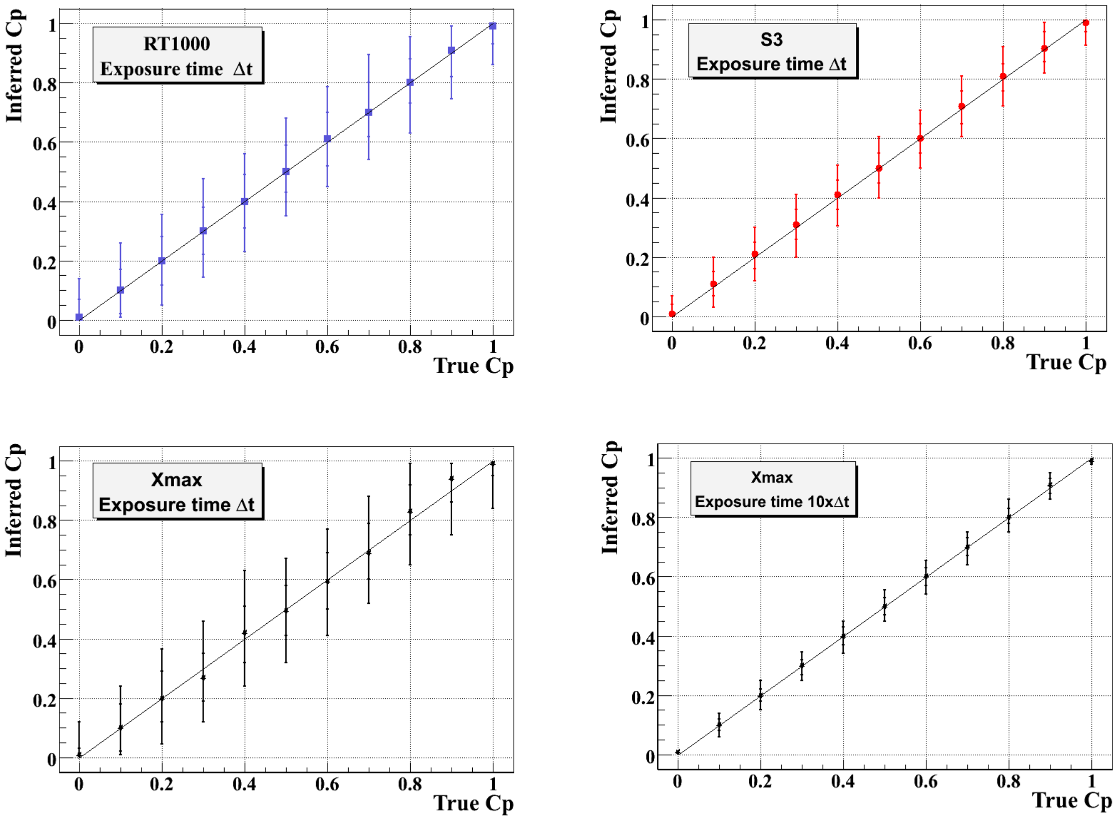

Figure 5.14: Inferred vs. true proton fraction using QGSJET-II for $t_{1/2}(r_0)$ (top-left), S_3 (top-right) and X_{max} (bottom-left) for the same exposure time (the 10% duty cycle of the fluorescence telescopes is taking into account). The bottom-right panel shows the inferred proton fraction obtained using X_{max} and samples with the same statistics as SD parameters, which would required 10 times more exposure time.

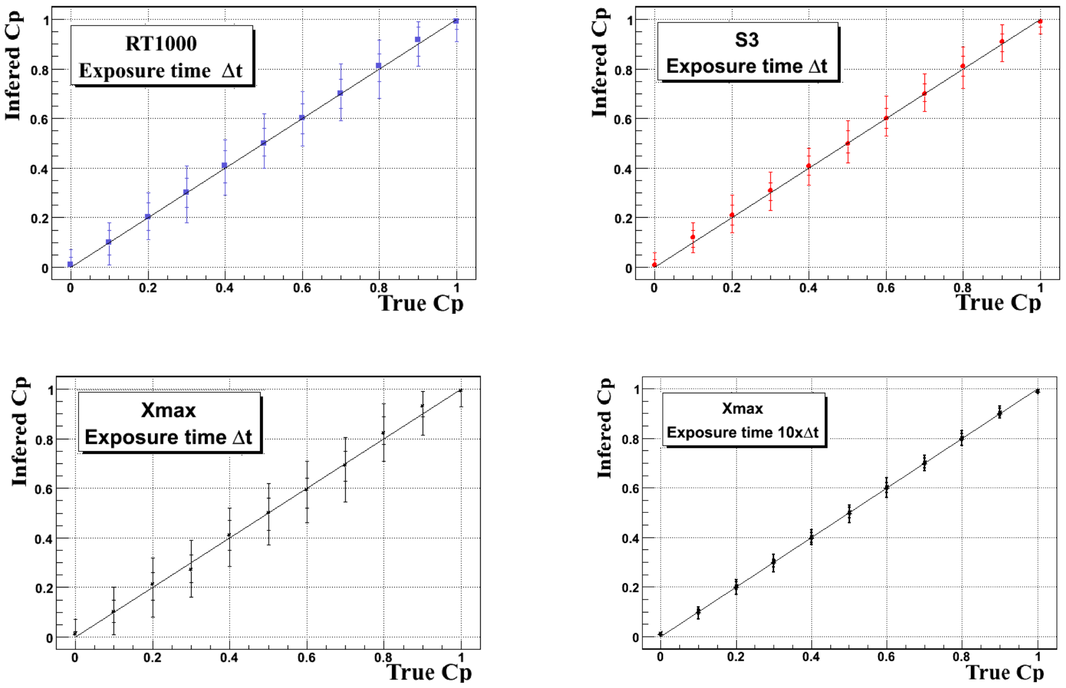

Figure 5.15: As Fig. 5.14 but using Sibyll 2.1.

h_p and h_{fe}, that represent the universe, with 1000 events each. We also generate 200 sub-samples for each case, but the number of events in the sub-samples varies as a function of primary energy because of the steepness of the spectrum. The number of events expected by Auger in 1 and 5 years of full operation are considered (for example, in one year and considering the spectrum reported in [29], around 500 events are expected at $10^{19.0}$ eV and 70 at $10^{19.6}$ eV). To reproduce real conditions, the number of events with available X_{max} is 10% of the total in the sample. The procedure to infer the composition is the same as explained before.

As in the previous case, there is no significant bias in the inferred proton abundance. However, Fig. 5.17 shows the uncertainty (the C.L. at 68%) on the determination of the proton abundance as a function of the primary energy. It can be seen that the best results are obtained by using S_3. As mentioned before, the uncertainties corresponding to Sibyll 2.1 are smaller than that for QGSJET-II because the shower-to-shower fluctuations are in general smaller for Sibyll 2.1.

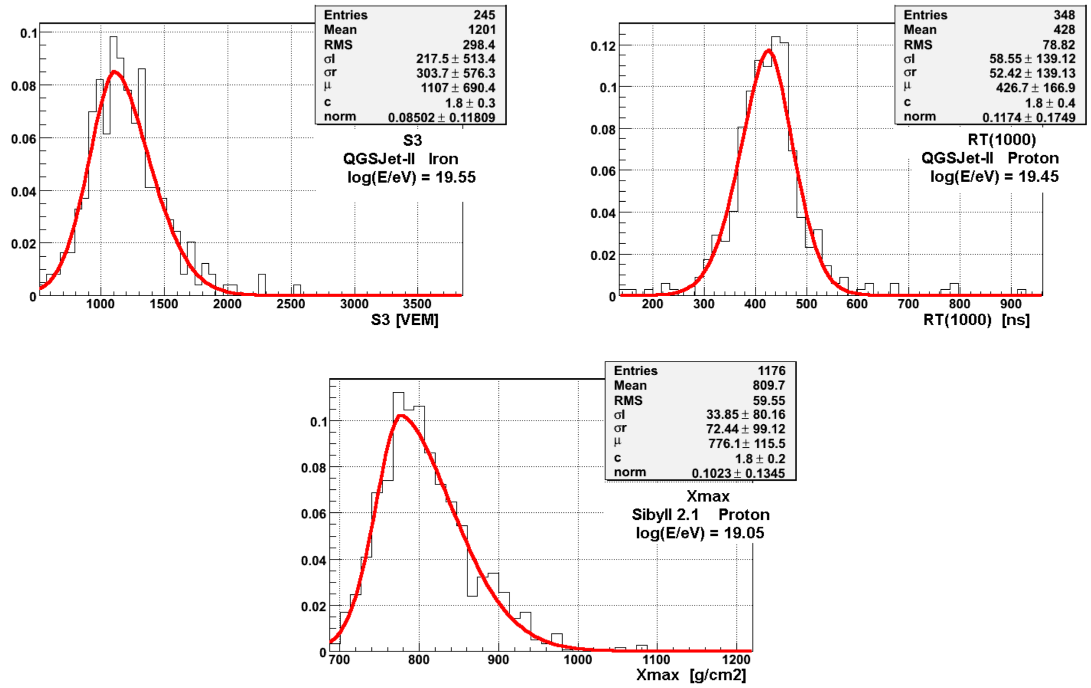

Figure 5.16: Examples of the fits with AGG function for the three parameters considered, different energy bins and hadronic interaction models.

5.5 Summary

We propose a new family of parameters, which we call S_b, for composition analysis in cosmic ray surface detectors. The parameters are evaluated from the total signal and position of each triggered detector, on shower-to-shower basis. In spite of the fact that surface composition parameters are usually more affected by systematic errors than fluorescence parameters do, the former are of great interest because of the larger statistics available for the surface array, while fluorescence telescopes work in a limited duty cycle of around 10%.

We perform an extensive analytical study of the most relevant properties of S_b. In particular, S_b has been optimized to distinguish between Iron and proton primaries assuming Auger water Cherenkov detectors, showing that the discrimination power between both samples reaches a maximum for $b \cong 3$. We have also demonstrated that, in case that the muon size is underestimated by simulation codes, as it is experimentally suggested, S_3 is

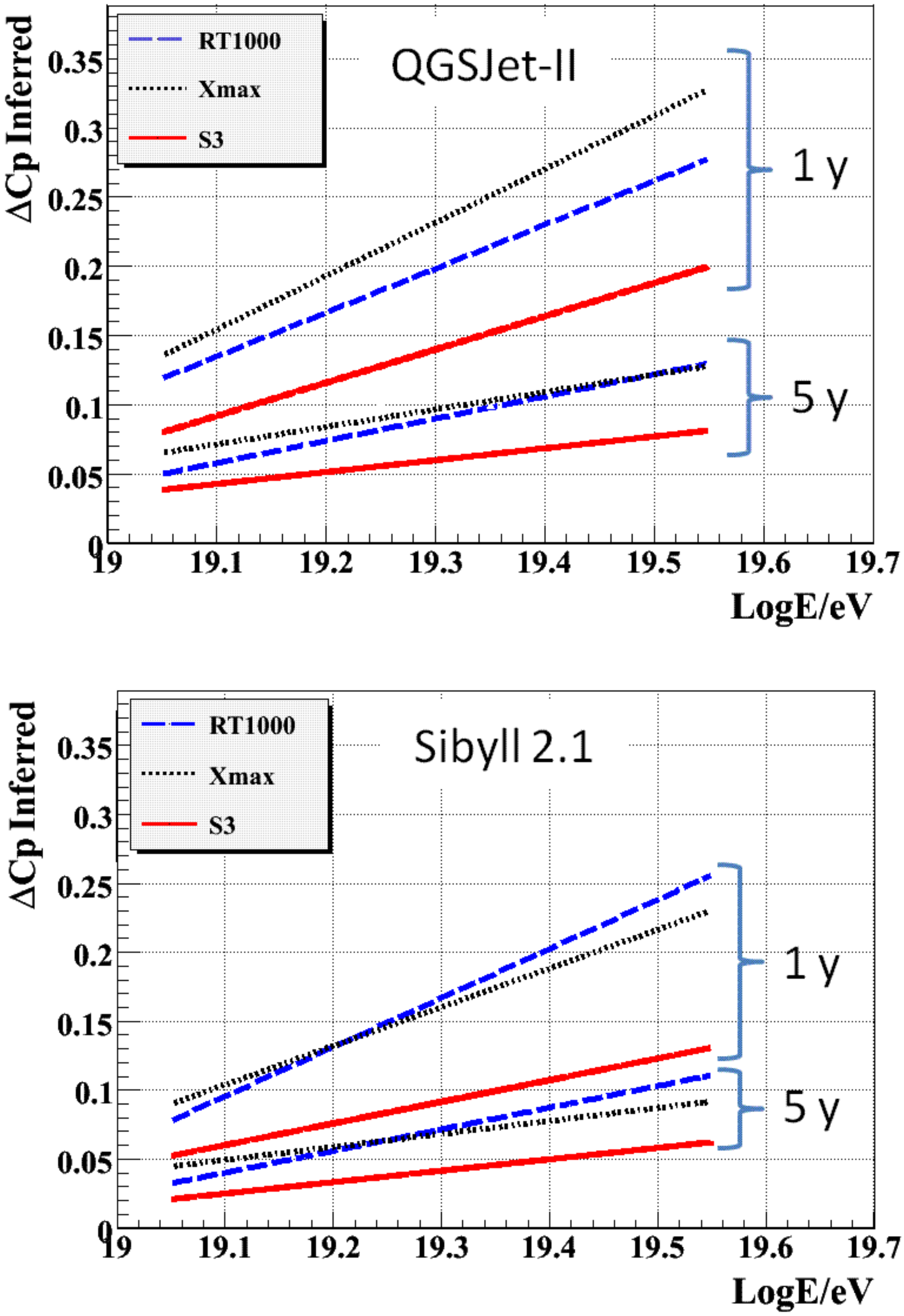

Figure 5.17: Error in the inferred proton abundance determined by using S_3, X_{max} and $t_{1/2}(r_0)$ for 1 and 5 years of Auger exposure and for both hadronic interaction models. For X_{max} only 10% of statistics is used due to the duty cycle of fluorescence telescopes.

not only stable but improves its discrimination power. Therefore, actual uncertainties in the simulation codes goes in the direction of improving S_3 separation power.

A numerical study simulating real experimental conditions and taking into account all the uncertainties involved in the reconstruction procedure is also performed. The numerical result supports that $b \cong 3$ is the value that maximizes the discrimination power of S_b, in agreement with the analytical result. We show that S_3 is almost independent on zenith angle and almost linearly dependent on the primary energy. Our calculations also show that S_3 can be used to infer the composition as a function of S_{38}, instead of energy, in order to avoid uncertainties introduced by possible unknown biases in the reconstructed energy.

A realistic analysis with the expected statistics of Auger in 1 and 5 years of full operation in the energy range from 10^{19} to $10^{19.6}$ eV is done. A likelihood method is used to infer the proton abundances of samples previously generated assuming a fix composition. Three parameters are tested: S_3, the rise time at 1000 m from the shower core $t_{1/2}(1000)$ and X_{max}. For the latter, the limited statistics due to the duty cycle of the fluorescence telescopes is taken into account. The accuracy in the reconstructed proton abundances using S_3 is greater in the whole energy range and for both hadronic interaction models (QGSJET-II and Sibyll 2.1) than that obtained by using $t_{1/2}(1000)$ and X_{max}.

Chapter 6

On-going work and perspectives

In the near future, we plan to continue the work presented in this thesis about the energy and composition determination. The techniques shown in Chapters 4 and 5 have been developed using simulations under conditions as realistic as possible. These promising methods will be applied to real data from the Pierre Auger Observatory. In addition, other related topics will be analyzed and discussed such as the standard Auger fit of the lateral distribution of particles and the reliability of the Monte Carlo air shower simulations.

Energy determination

In [163] the standard calibration formula was obtained from hybrid events by Auger. First, a set of high-quality hybrid events was selected. Data were required that, i) the reduced χ^2 value for the fit of the longitudinal profile were lower than 2.5, ii) the shower maximum depth were within the field of view of the telescopes and, iii) the fraction of the signal attributed to Cherenkov light were less than 50%. As an energy estimator, the signal inferred at a fixed characteristic distance $S(r_0 = 1000m)$ was selected, independently on the shower energy or direction. The Constant Intensity Cut (CIC) method is applied to compensate for the increasing absorption of the atmosphere as the zenith angle of the shower increases, so that $S(1000)$ was transformed into S_{38°, i.e. the $S(1000)$ that the

EAS would have produced if it had arrived at the median zenith angle of 38°. Statistical uncertainties in S_{38° and E_{FD} (i.e. the energy measured by the fluorescence telescopes) were assigned to each event. Averaged over the sample, these uncertainties were 16% and 8% for these magnitudes, respectively. A calibration formula was obtained from the fit, $E = a \cdot S_{38^\circ}^b$. The energy resolution, estimated from the fractional difference between E_{FD} and the derived SD energy E, was also estimated. The bias is 3% and the root-mean-square deviation (RMS) of the distribution is 19%, which is in good agreement with the quadratic sum of the S_{38° and of the E_{FD} statistical uncertainties of 18%.

The alternative method proposed in Chapter 4 to infer the primary energy based on the determination of a optimum shower specific distance, r_{opt}, is currently being applied to Auger data. We have already calculated the optimum distance using the same set of hybrid events as [163] and, as a first approach, we have determined a new energy calibration formula between $S(r_{opt})$ and E_{FD}. In our procedure, no CIC method is, in principle, required since the zenith angle dependence is already taken into account in the determination of r_{opt}. The bias in the relative error of the energy determination by using the new calibration, is 5% and the RMS is around 26%. If the tanks located at distances larger than 1.5 km from the shower axis, whose signal could be dominated by fluctuations, are excluded from the LDF fits to find r_{opt}, the bias is reduced to 1.3% and the RMS to 22%. This is almost compatible with the standard Auger technique. However, at the moment we are working on several methods to improve the new procedure:

- The saturated stations, whose signal could be recovered by different methods as suggested in [106], could be included in the r_{opt} search. It must be studied if this improves the determination of r_{opt}.

- The signal of the stations far from the shower axis could be dominated by fluctuations. As already mentioned, if those stations at distances larger than 1.5 km from the shower axis are not included in the LDF fits to find r_{opt}, the energy resolution is improved. More sophisticated treatments about how to include these stations in the LDF fit must be tested. As an example, the signals could be corrected from their Poissonian fluctuations.

- The first approach to get a calibration curve from r_{opt} has been done by using the hybrid data set selected in [163] to find the standard energy calibration. In order to improve the energy resolution based on the r_{opt} method, it must be studied if more or different selection cuts should be applied.

- The sources and contribution to the statistical uncertainties in $S(r_{opt})$ must be studied. For example, the contributions to the S_{38° uncertainties come from the reconstruction procedure, from the LDF parametrization and from shower to shower fluctuations. Zenith angle uncertainties are neglected (they are related to the CIC method), since their contribution is of only 0.1% [106]. They give a total statistical uncertainty of 16% [163].

In addition, there are other topics to analyze that are related to the new technique:

- As in Section 4.4.1, the distribution function of the relative errors in the energy determination with both techniques, $S(r_0)$ and $S(r_{opt})$, must be compared. We have demonstrated that using $S(r_{opt})$, the distribution functions are expected to be almost Gaussian. Thus, it could be feasible to develop a deconvolution technique for the Auger measured spectrum.

- The comparison between the standard Auger LDF fit and our procedure could be performed. As shown in Appendix B, Auger carries out an iterative process based on a maximum likelihood method where the slope of the LDF is fixed by using a parametrization that depends on S(1000) and on the zenith angle of the incoming event. On the other hand, we fix the core position and leave the slope as a free parameter in the fit. As an example, in Fig. 6.1 the LDF fit performed with both procedures are shown for some events of those selected in [163]. The Auger standard method is optimized to find a reliable S(1000), which is the energy estimator. However, it seems that the closest stations to the shower axis constrain the parameters of the fit, while stations further away from the core are not taken

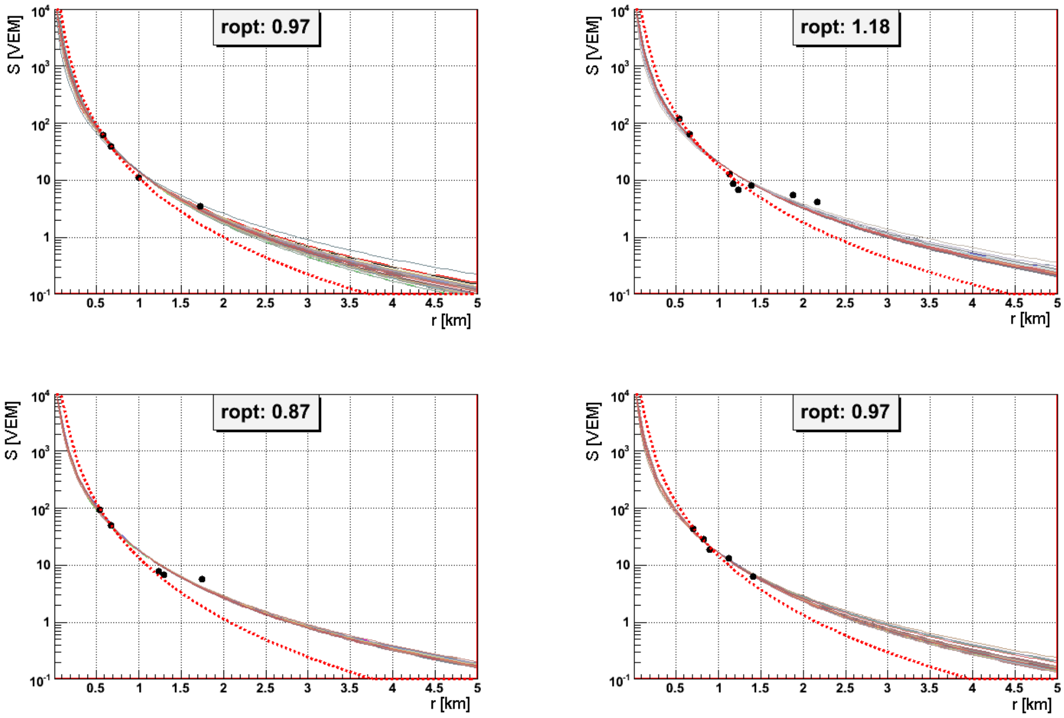

Figure 6.1: Examples of the LDF fit performed by Auger (red dashed line) and the 50 fits used in our technique to find r_{opt}.

into account properly. Although our procedure must be improved to correct the signal from the Poissonian fluctuations, we believe that this procedure allows to perform a more reliable fit of the whole LDF. This is due to the fact that the slope is considered as a free parameter that has to be fitted. In the standard method, the slope is fixed to avoid a bad reconstruction of those events with few triggered stations or with saturation. However, we have shown that our method provides a reliable determination of $S(r_{opt})$ for all the events using an unified treatment, and independently on the number of detectors and whether they are saturated.

- Discussions with the LDF task group during the last years were useful to find some bugs in the standard procedure implemented in Auger Offline Software Analysis. These bugs were fixed in the last released versions. This work should be continued in order to improve the standard LDF fit performed by Auger.

Composition determination

The aim of our composition studies is to determine the average composition of UHECRs from Auger data using surface observables. Among others, these surface observables includes the new parameter S_3 proposed in this work. This allows to compare our results with those obtained by Auger using X_{max}, XAsymMax and the Delta parameter (a review of these parameters was given in Section 3.4.2 and the results were shown in Section 3.5.2). Different aspects must be considered:

- The composition determination on event-by-event basis is unreachable with the present methods. However, a proton fraction, or an average mass, could be determined as a function of energy by several methods such as by defining probabilities for different primaries in a certain parameter space, or by using a maximum likelihood method.

- These techniques are currently being developed. However, better quality cuts are needed in order to get a reliable result. These cuts must be carefully analyzed for each parameter to prevent introducing any bias in the sample of selected events.

- A large amount of simulations are needed to this purpose. We have already performed the end-to-end simulations, including shower and detector simulations and shower reconstruction (using AIRES and Offline programs). The statistics available now are larger than the Auger spectrum presented at the ICRC in 2007, for both Iron and proton primaries and for two hadronic interaction models (QGSJetII-03 and Sibyll 2.1). We are going to continue increasing the statistics available and other nuclei will be used as primaries.

- Sources of systematic errors must be also studied in detail.

- S_3 could be also used to discriminate between photons and hadrons, and compare it with the rise time and the radius of curvature that were used by Auger to find the photon fraction in SD data [28].

In addition, other questions, that may affect composition determination, are being analyzed:

- We are investigating the stability of the array and how this could affect the surface composition estimators. It is crucial to establish this question because a significant fraction of the Auger data were taken during the construction of the array. We are studying how the composition estimators are modified if one station was not working during the event. The T5 trigger (see Appendix B) assures that the first crown around the hottest station was working at the moment of the event, but most energetic events may require different conditions, as considering also the next crowns. In addition, a detector whose signal is closed to the trigger threshold, could be triggered or not, depending on the shower fluctuations. It must be analyzed how these conditions could modify and in what extent, the composition estimators.

- An important handicap when dealing with the composition determination, is the need of simulations. It is known that these simulations are affected by several problems, such as the lack of muons [150] whatever the hadronic interaction model used, artificial fluctuations introduced by the thinning [151] and unthinning [152] algorithms, and additional dispersion in the sensitive composition estimators when the showers are re-used (a common practice in composition studies due to the long CPU time required for shower simulations) [153]. More tests should be performed to check the reliability of the MC simulations to reproduce the spatial and temporal distribution of particles at ground. In this line, we are going to calculate the covariance matrices including several composition observables and other properties such as the energy or zenith angle. If simulations are reliable, there should exist some homogeneity in the covariances between two variables when comparing different hadronic interaction models or energy and zenith angle bins. Moreover, this is important not only for checking the reliability of the simulations, but also for developing multiparametric techniques.

Chapter 7

Conclusions and Outlook

This thesis focuses on cosmic ray detection by a surface detector array and it deals with two questions: i) the technique used to infer the primary energy and ii) how one can determine the chemical composition of ultra-high energy cosmic rays (UHECRs). Both, energy and composition, are linked problems. From the theoretical point of view, this link arises essentially from the acceleration mechanisms (as in the Fermi acceleration, Section 2.2.1) because the source power is directly proportional to the particle charge. Experimentally, the composition sensitive parameters depend on the primary particle energy. As a consequence, the systematic and the statistical uncertainties in energy determination affect the discrimination power of the composition observables.

Regarding the energy determination of UHECRs, big steps forward have been recently given by Auger. For example, the previous controversy between HiRes and AGASA experiments about the existence or not of the GZK suppression has been solved in 2007, when Auger confirmed it with a significance of 6 standard deviations as shown at the 30^{th} International Cosmic Ray Conference in Mérida (Mexico). In addition, the position of the ankle was confirmed, in agreement with HiRes at $10^{18.6}$ eV [166], and in contrast to AGASA, that previously reported it at 10^{19} eV [56]. However, several open questions still remains: Does the second knee in the spectrum exist? What is the transition energy from galactic to extragalactic cosmic rays? What is the reason for the high difference in the measured flux of UHECRs between Auger, HiRes and AGASA (a factor of 2 in flux, 30% in energy or a combination of both)? Is the GZK suppression a consequence of a maximum limit in the acceleration processes at the sources or is it just a consequence of the interaction of cosmic rays with CMB photons? How steep is the fall-off above the suppression? What is the exact shape of the energy spectrum at the highest energies and is it consistent with proton or mixed composition?

In this thesis, a new method to determine primary energy is suggested, which could help to solve these questions. The method and the main conclusions are summarized in the following:

- In surface arrays the primary energy has been traditionally determined from the inferred signal at a fixed *characteristic distance* from the shower axis (r_0). This distance is selected taking into account only the array geometry and detector separation. For example, the AGASA experiment used $r_0 = 600$ m while Auger selects 1000 m due to its larger array spacing. However, we have demonstrated that there exists a shower-specific *optimum distance* (r_{opt}) for energy calibration that depends not only on the array geometry, but also on the energy and incoming direction of the primary. These dependencies are not negligible and, therefore, an optimum distance calculated on shower-to-shower basis is more suitable as an energy estimator.

- A method to find the optimum distance for each individual shower has been developed. Essentially, the measured lateral distribution of particles is fitted assuming a certain functional form. In the fit, the core is fixed and the normalization constant and the slope are kept as free parameters. A set of 50 fits are performed modifying the core position around the estimated one, and taking into account the actual uncertainties in the core position reconstruction. The optimum distance, r_{opt}, is defined as the distance where the dispersion between the interpolated signals from these set of fits reaches a minimum.

- The signal inferred at the optimum distance, $S(r_{opt})$, compared to that at the characteristic distance $S(r_0)$, significantly reduces the bias in the inferred energy. Using $S(r_{opt})$, no bias is introduced in the whole energy range of an experiment for events with and without saturated detectors. On the contrary, the $S(r_0)$ technique gives a significant bias due to the difference between the value selected for r_0 and the actual optimum distance of the shower.

- Using $S(r_{opt})$ as energy estimator improves the behavior of the energy error distribution functions. They are more compact, less skewed and with shorter tails compared with those obtained by using $S(r_0)$. As a consequence, it has been shown that a realistic spectrum can be reliably reconstructed with the $S(r_{opt})$ technique. While the $S(r_0)$ method could change the position and shape of the ankle and shift the GZK suppression, the r_{opt} approach fits very tightly the original spectrum.

- A major advantage of the r_{opt} technique is that it allows the same unified treatment for events with and without saturated detectors. In the r_0 method this is in general not possible for the whole set of events with saturation. It is required to apply selection cuts, to develop new algorithms to recover the signal of these detectors or the separate reconstruction of the two types of events with different techniques. This advantage of the new method is more evident at higher energies since the fraction of events with saturation rapidly increases as a function of energy.

- Previous conclusions also hold for different array geometries and detector types. In fact, the method has been applied to a triangular scintillator array as in the AGASA experiment (Chapter 4), and also to a rectangular array of water Cherenkov detectors as it will be the future Auger North Observatory (Appendix A).

- The application of the r_{opt} method to a real hybrid experiment as Auger, is not straightforward. A relationship for energy calibration must be found to relate $S(r_{opt})$ with the primary energy measured by the fluorescence technique, considering realistic values for the core uncertainties and designing appropriate and efficient quality cuts to select the best hybrid events to achieve this purpose. As it was stated in

> Chapter 6, initial steps have been given in this direction. The main systematic uncertainties in energy determination in a hybrid observatory as Auger, come from the fluorescence technique. This is due to the uncertainties in the fluorescence yield, the FD calibration and reconstruction. Therefore, the improvement would be less significant than in a pure SD array. However, some benefits would be obtained with this technique. First, in principle, a CIC method would be no longer needed since r_{opt} is determined considering the incoming direction of the shower. Second, the energy error distribution functions would be Gaussian-like, which makes easier and more reliable the application of a deconvolution technique in the measured spectrum.

The method summarized above may help to infer more accurately the energy spectrum of UHECRs. For example, in order to distinguish between interpretations about the origin of the suppression, two features must be measured in the spectrum. First, a bump in the flux just at energies where the GZK begins (around $10^{19.6}$ eV as measured by Auger [166]) would be caused by higher energetic particles that interact with the CMB photons and degrade their energy to that point (the cross section of this interaction is significantly reduced below the GZK, see Fig. 2.6). Second, a recovery in the spectrum at even higher energies ($\sim 10^{20.5}$ eV) is expected if the suppression is not a consequence of a maximum limit in the acceleration processes. Such high energies are not reachable by Auger South. However, Auger North could provide enough statistics during several years of operation. More statistics and more accuracy are needed around the GZK energy to distinguish the bump and to interpret correctly the suppression.

In addition, our method could be used to measure more accurately the ankle region and by surface experiments designed to study cosmic rays at lower energies. KASCADE-Grande experiment [55] and the AMIGA extension of Auger [101], analyze energies in the range between 10^{17} - 10^{18} eV, where more accurate and precise measurements could allow to either confirm or to ruled out the existence of the second knee. Moreover, the r_{opt} technique could be applied to AGASA and Auger data in order to understand the origin of the discrepancy in the flux of UHECRs measured by both experiments.

The composition of UHECRs is a more complex issue. In fact, only few methods have been proposed and have been successfully applied. For example, Auger uses the position of maximum development of the shower, X_{max}, and two surface observables related with the rise time (XAsymMax and the Delta parameter), to get the average composition trend as a function of the energy [116, 131]. However, these results only extends up to $10^{19.4}$ eV and some controversy exists about these results and the reported correlation of UHECRs and the position of AGNs [170]. The reason is that the elongation rate at the highest energies goes in the direction of heavy primaries, while the reported correlation between the incoming direction of the most energetic cosmic rays detected by Auger an the position of known AGNs, is valid if primaries are light nuclei, which are slightly deflected by magnetic fields. It is accepted that the main problem regarding these points is the lack of statistics, problem that Auger could solve in the next years of operation. However, new techniques and new sensitive parameters are needed.

The actual techniques only allow to determine an average composition. The composition on event-by-event basis is almost impossible, specially due to the uncertainties in the hadronic interaction models, to intrinsic fluctuations and also due to energy uncertainties (see a discussion in Section 3.4.3).

The knowledge of the composition of UHECRs is also very important to put some constrains on several theoretical models related to their origin. In fact, the actual photon limits published by Auger (Section 3.5.2) disfavor several top-down models about the origin of cosmic rays between 10^{18} and $10^{19.5}$ eV. In addition, composition may allow to distinguish between the models that explain the origin of the ankle in the spectrum and the transition from galactic to extragalactic cosmic rays (full discussion was given in Chapter 2).

In this thesis a new sensitive parameter for composition analysis that make exclusive use of surface data is presented. Some remarks are given in the following:

- Surface composition observables are usually less reliable than fluorescence ones (mainly X_{max} and its fluctuations) because the latter are less prone to systematic errors and easier to interpret. However, the large statistics available from surface array compared to fluorescence telescopes ($\sim$ 10-13% duty cycle), motivates the need to find better surface parameters.

- Here we propose a new family of surface parameters, named S_b. In this work, it is applied to water Cherenkov tank detectors as those found at the Auger Observatory. This parameter could also be useful for scintillator arrays.

- Under these conditions, we have determined that the separation power of S_b, which is obtained by using the merit factor, reaches a maximum for $b \cong 3$.

- We have shown analytically that S_b is sensitive to the different slope of the lateral distribution function between Iron and proton primaries. In addition, current uncertainties in the muonic component of the simulated showers go in the direction of improving the discrimination power of S_b.

- The stability of S_b is stressed by the fact that the value of the exponent b that maximizes the separation power of S_b is always 3.0, independently if the slope of the LDF for proton and Iron primaries or the weight of the muonic component were artificially modified. Additionally, it has been shown that stations further away from the shower axis, whose signal could be dominated by shower fluctuations, do not affect the discrimination power of S_b.

- S_b has been tested under realistic conditions. An end-to-end simulation chain have been performed, including the simulation of the shower development in the atmosphere (using AIRES), the detector response (using Geant4) and the shower reconstruction (using the Auger Offline software). Thus, real experimental conditions, with the uncertainties involved in the reconstruction procedure, have been considered.

- We have shown in this work that S_3 depends almost linearly on S_{38°, the primary energy estimator used at Auger, and the discrimination power as a function of S_{38° is strong. Therefore, a valuable astrophysical insight into the evolution of composition as a function of energy can be gained despite the fact that the exact energy profile is unknown. Regarding the incoming direction of the shower, no dependence of S_3 with the zenith angle is found.

- A detailed and realistic comparison based on a maximum likelihood method has been carried out, comparing surface observables such as the rise time and S_3, and X_{max}, the main fluorescence parameter. The statistics for the latter are reduced due to the limited duty cycle of the fluorescence telescopes ($\sim 10\%$). It is shown that the error in the inferred average mass is significantly lower using S_3 as the composition estimator compared to that obtained by using the rise time and X_{max}. The available statistics when determining composition from data is crucial, as shown by the previous result.

To summarize, the goal of this thesis is to improve energy and composition determination in surface arrays by developing new methods. The reliability of these methods has been demonstrated under the most realistic possible conditions. However, further work needs to be done. Checks, optimizations and application to Auger data are currently under progress (see Chapter 6).

Appendix A

Optimum distance at Auger North array

A.1 Introduction

In this Appendix we apply the study shown in Chapter 4 to a square grid surface array, as it is the design of the Pierre Auger North Observatory. The aim is to study the applicability of the method for a different array geometry, because in Chapter 4 always triangular grids were considered. In addition, in Chapter 4 the study about the energy spectrum was performed assuming an AGASA-like experiment, so that scintillators were used as detectors. Here, water Cherenkov tanks as those at Auger South are used. The preliminary design for the tanks at Auger North is quite similar. Some modifications are being considered, for example, only one PMT is going to be used instead of three and the material will be different [199]. These changes do not affect our study since, in principle, the same lateral distribution function is expected to be valid.

Two different LDFs and energy conversion formulas reported by Auger South are going to be considered. First, an NKG with the $S1000(E,\theta)$ parametrization as reported by Auger at [6] and, second, the current LDF and the energy parametrization obtained from the Constant Intensity Cut method (CIC) and hybrid energy calibration (details below).

We perform a detailed study of r_{opt} as a function of energy and zenith angle. Later, it is investigated how using the signal at a characteristic distance as energy estimator, $S(r_0)$, instead of that at a shower-specific optimum distance, $S(r_{opt})$, could affect the inferred energy spectrum. Special attention is paid to events with saturated detectors.

Details about how the energy calibration from hybrid events is performed by Auger South could be found in [103, 106, 163]. Here it is explained briefly. First, for a given energy the value of S(r_0 = 1000 m) decreases when increasing the zenith angle θ, due to attenuation of the shower particles while traverse the atmosphere and geometrical effect. In order to correct that effect, the CIC method is used assuming an isotropic flux for the whole energy range considered, i.e. the intensity distribution is uniform when binned in $cos^2(\theta)$, so that the shape of the attenuation curve from the data is extracted. The fitted attenuation curve, $CIC(\theta) = 1 + ax + bx^2$, is a quadratic function of $x = cos^2(\theta) - cos^2(38°)$. Since the average angle is $< \theta > \simeq 38°$ we take this angle as reference, and convert S(1000) into $S_{38°}$ by $S_{38°} \equiv S(1000)/CIC(\theta)$. It may be regarded as the signal S(1000) the shower would have produced if it had arrived at $\theta = 38°$. Second, to establish the relation between $S_{38°}$ and the calorimetric energy measurement from the fluorescence detectors, E_{FD}, a set of high quality hybrid events are selected. The data appear to be well described by a linear relation $\log(E_{FD}) = A + B \cdot \log(S_{38°})$, and to avoid possible biases, low energy events are not included in the fit. The uncertainties in $S_{38°}$ and E_{FD} are both considered.

The Appendix follows the structure of Chapter 4. In Section A.2 the algorithm is explained, the r_{opt} dependencies on energy and zenith angle are in Section A.3, and the study of energy spectrum reconstruction is in Section A.4. Discussion about the scope of this work and the conclusions are in Section A.5.

A.2 Algorithm: r_{opt} and energy determination

The algorithm has been already explained in Section 4.2. Here we report only the differences that require the application to the Auger North array: the different array geometry, the *true* LDFs assumed and the functions used to fit the *experimental* LDF.

In order to mimic Auger North Observatory, the surface stations are located at the vertexes of a square grid. There were two proposed configurations of 1 mile and $\sqrt{2}$ miles for the Northern site [200] (2008). Unfortunately, we selected the former for this work, while in the last ICRC that took place in last July, the latter option was finally confirmed [199]. This would modify the values of r_{opt} but the dependencies and the conclusions achieved about the energy spectrum reconstruction will remain the same.

In a square array of 1 mile spacing, full efficiency is assured for energies larger than $10^{19.0}$ eV [201]. Therefore, we analyze primary energies between $10^{19.0}$ and $10^{20.5}$ eV and zenith angles from 0 to 60 degrees. Azimuthal angles have been selected randomly from 0 to 360 degrees and the core positions are chosen randomly inside an square elementary cell.

Given an incoming event with a certain energy E and zenith angle θ, in order to assign the signal at each station located at a distance r from the shower axis, we assume a *true* lateral distribution function. Two options are selected:

- A Nishimura-Kamata-Greisen (NKG) function [79] normalized at 1000 m in the same way as the reported by the Auger in [6]:

$$\begin{aligned} S(r,E,\theta) &= S1000(E,\theta) \times 2^{\beta(\theta)} \times r^{-\beta(\theta)} \times (1+r)^{-\beta(\theta)} \qquad \text{(A.1)} \\ S1000(E,\theta) &= \frac{7.53\ E^{0.95}}{\sqrt{1+11.8[sec(\theta)-1]^2}} \end{aligned}$$

  where r is the distance of the detector to shower axis in km, E is the energy in EeV and $\beta(\theta) = 3.1 - 0.7sec(\theta)$. This is the LDF also used in Chapter 4 for water Cherenkov detectors. We call this option as *OLD* LDF.

- The current LDF determined from data used at Auger South Observatory:

$$S(r,E,\theta) = S1000(E,\theta) \times \left(\frac{r}{1000}\right)^{\beta(\theta,S1000)} \times \left(\frac{r+700}{1000+700}\right)^{\beta(\theta,S1000)} \qquad \text{(A.2)}$$

where the distance to shower axis r is in meters, $\beta(\theta, S1000)$ is given in [202] as

$$\begin{aligned} \beta(\theta, S1000) &= A(\theta) + B(\theta) \times \log_{10}(S1000(E,\theta)) \qquad (A.3)\\ A(\theta) &= a_1 + a_2 sec(\theta) + a_3 sec^2(\theta)\\ B(\theta) &= b_1 + b_2 sec(\theta) + b_3 sec^2(\theta) \end{aligned}$$

where $a_1 = -3.35$, $a_2 = 1.33$, $a_3 = -0.191$, $b_1 = -0.125$, $b_2 = -0.0324$ and $b_3 = -0.00573$, and $S1000(E, \theta)$ is obtained from the primary energy and zenith angle by the CIC curve [203] and the hybrid calibration [204]:

$$\begin{aligned} \log_{10}(E_{FD}) &= 17.117 + 1.105 \cdot \log_{10}(S_{38^\circ}) \qquad (A.4)\\ S1000(E,\theta) &= S_{38^\circ} \times \left[1 + 0.92x - 1.33x^2\right] \qquad (A.5)\\ x &= cos^2(\theta) - cos^2(38^\circ) \end{aligned}$$

We call this option as *NEW* LDF.

The expected signal at each station is then fluctuated with a Poissonian distribution whose mean is given by the *true* LDF selected. We impose as a trigger condition $S(r) \geqslant 3.0$ VEM. Stations with signal $S_i > 1000$ VEM are considered as saturated and are excluded from the LDF fit. In the standard Auger reconstruction algorithm the signal from saturated detectors are used as a lower limit in the fitting process. In addition, several algorithms are being developed and tested to recover the signal of saturated stations [106].

The difference between both LDFs and $S1000(E, \theta)$ parametrizations considered is shown in Fig. A.1. As can be seen, the difference is more important as energy increases, for lower zenith angles and larger distances from core. Therefore, there exists a significant difference in the number of triggered stations (Figure A.2) depending on the LDF selected, that will lead to different values of r_{opt} as will be shown in next Section. The difference is less significant regarding the fraction of saturated events (Figure A.3).

Finally, for each new *shifted* core position, the LDF fit is performed (see Section 4.2 for a description of the algorithm). To that end we use a form of the LDF formally equivalent to Eq.(A.1) and Eq.(A.2) respectively:

$$\log S(r) = a_1 - a_2 \left[\log(r) + \log(1 + r)\right] \qquad (A.6)$$

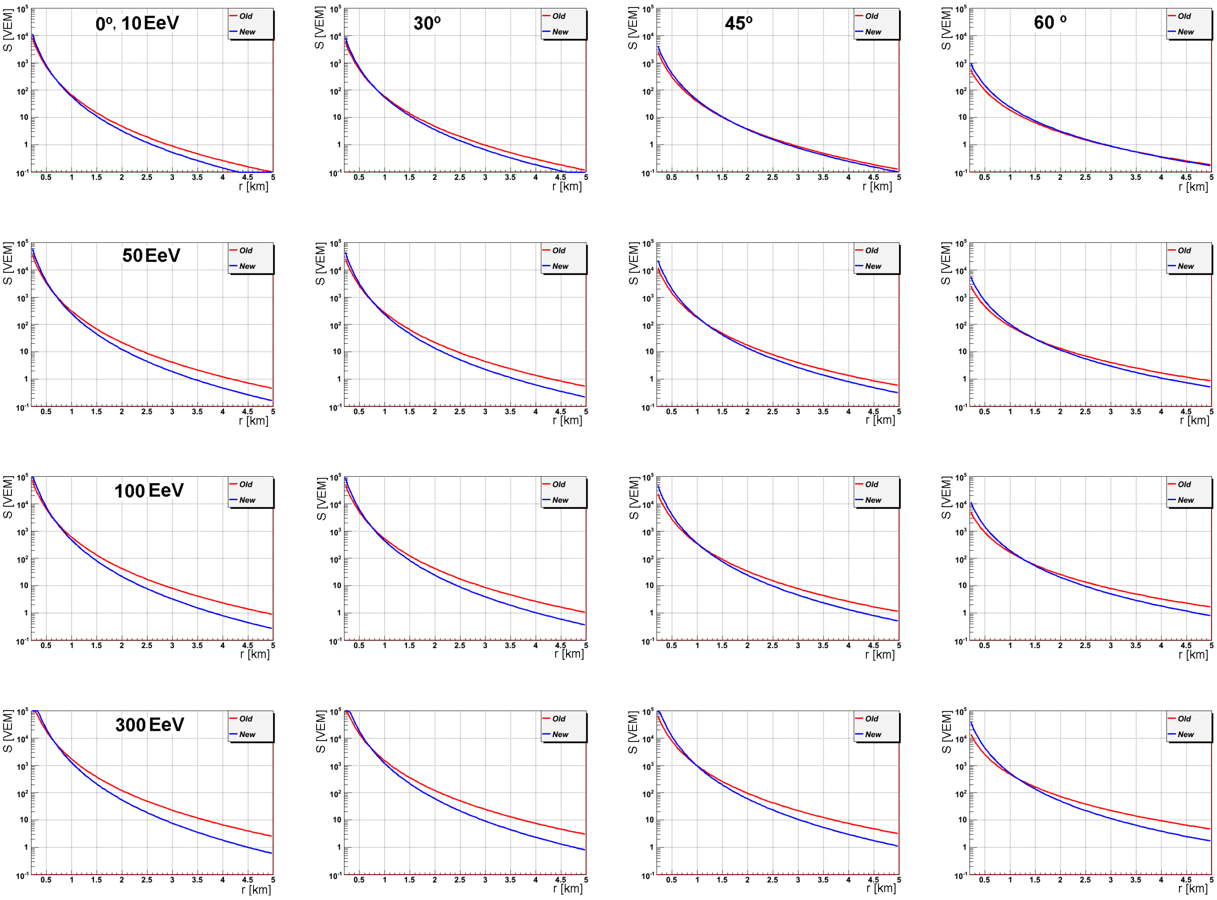

Figure A.1: LDF fits (Signal [VEM] vs Core distance [km]). Both LDFs considered. Old: NKG + $S1000(E,\theta)$ from Auger NIM paper (Red). New: Current Auger LDF + $S1000(E,\theta)$ from CIC and Hybrid calibration (Blue). Rows: 10, 50, 100 and 300 EeV. Columns: zenith angles of 0, 30, 45 and 60 degrees respectively.

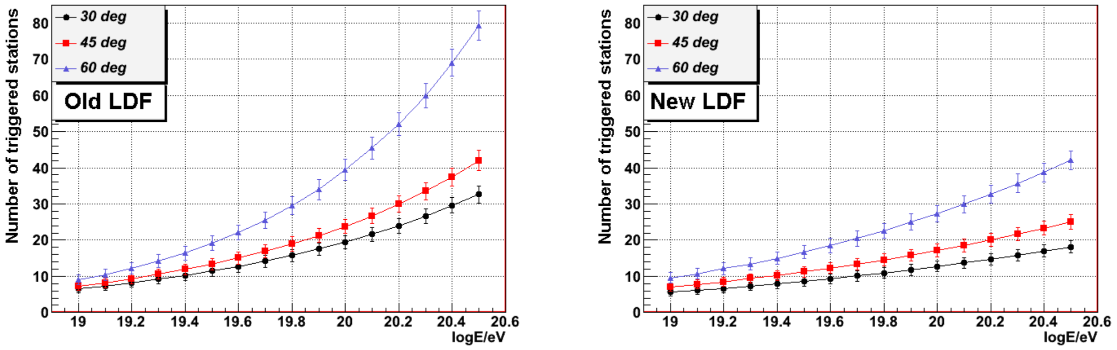

Figure A.2: Number of triggered stations vs. energy for both LDFs and $S1000(E, \theta)$ parametrizations considered and for several zenith angles. Left: Old LDF. Right: New LDF.

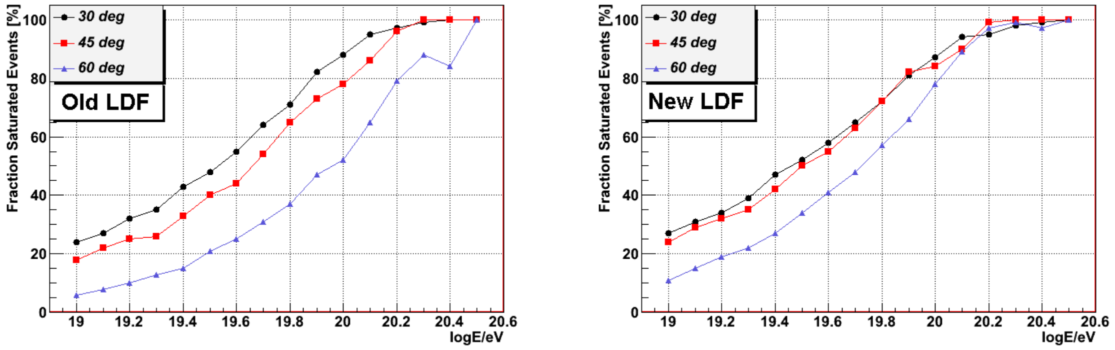

Figure A.3: Fraction of saturated events as a function of energy for several zenith angles. Left: Old LDF. Right: New LDF.

$$\log S(r) = a_1 + a_2 \left[\log \left(\frac{r}{1000} \right) + \log \left(\frac{r + 700}{1000 + 700} \right) \right] \tag{A.7}$$

where the distance to shower axis r is in km in Eq. (A.6) and in m in Eq. (A.7). As in Chapter 4, in each fit the slope of the LDF and the normalization constant are determined while the core position is fixed in the *shifted* core position. The r_{opt} value is defined as the point at which the dispersion among the interpolated signals over the several LDFs goes through a minimum.

To obtained the inferred primary energy the same method as in Chapter 4 is performed for both LDFs and $S1000(E, \theta)$ parametrizations used. Once r_{opt} is calculated, and for the selected value of r_0, the signal at both distances is obtained interpolating from a LDF fit where the core is set in the position of the *reconstructed* one, as it would be done in a real situation. From these signals, $S(r_0)$ and $S(r_{opt})$, and using the corresponding LDF and $S1000(E, \theta)$ parametrization, the reconstructed energy is determined.

A.3 r_{opt} dependence on energy and zenith angle

As in Chapter 4, we use the name *All* in the figures when the events with and without saturated stations are both included, *Sat.* (*Non-Sat.*) when only the former (latter) are considered.

Fig. A.4 shows the dependence of r_{opt} on the primary energy for both LDFs where all the events are included. The relationship between r_{opt} and energy is almost linear due to the triggering of stations progressively further away from the shower core as energy increases. In fact, r_{opt} is a strong function of energy and grows by more than a factor of two in the energy interval from 10^{19} to $10^{20.5}$ eV. Furthermore, the same trend is observed for either saturated or non-saturated events and any zenith angle, regardless of energy, it is always larger for the former (see Fig. A.5). These results are om agreement with Chapter 4 where a triangular grid were assumed.

Naturally, r_{opt} values are greater for the Old LDF compare to New LDF, since the Old LDF triggers stations at larger distances (Figs. A.1 and A.2). Fig. A.6 shows the dependence of r_{opt} on the zenith angle for both LDFs and for all the events, saturated and non-saturated events separately. In general, there is not a significant dependence on the zenith angle. For energies between $10^{19.0}$ to $10^{19.5}$ eV there is a slight decrease of r_{opt} with θ, and at higher energies it slightly increases. It must be noted that this is a property that is not shared by the Auger South and North. In fact, in Chapter 4, it was demonstrated that in a triangular array as the one in Auger South, r_{opt} does not depend on θ up to 30°, but in general it decreases slightly for larger zenith angles.

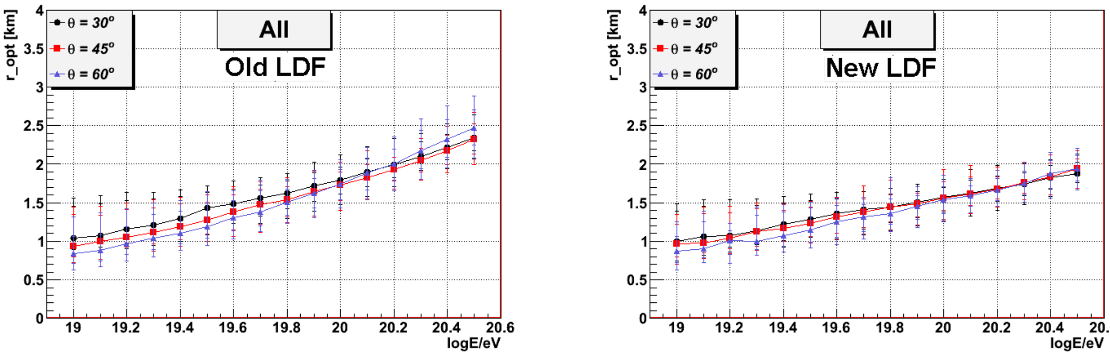

Figure A.4: r_{opt} dependence on the primary energy. Three different zenith angles are shown. Left: Old LDF. Right: New LDF. Error bars are the C.L. at 68% and 95%. The label *All* means that both events with and without saturated stations are included.

The r_{opt} dependence on the primary energy and the larger dispersion shown in the previous figures, show that considerable fluctuations in the optimum distance are expected. Therefore, using the signal at a characteristic fixed distance $S(r_0)$ as energy estimator, or even using a suitable parametrization of $r_{opt}(E, \theta)$ are not properly enough. In the next Section it is analyzed the energy error distribution functions obtained using $S(r_0)$ and $S(r_{opt})$ as energy estimators.

A.4 Energy error distributions

In this Section we analyze the effect of using a characteristic value r_0 instead of a shower-specific optimum distance r_{opt} in the determination of the energy spectrum. We choose $r_0 = 1500$ m as it is suggested in [200] for an square array of 1 mile. This value is also an intermediate value of r_{opt} for both LDFs used (see Fig. A.4), so that it seems a good choice for r_0. We generate the signals at each station with Eq. A.1 (Eq. A.2) and use the Eq. A.6 (Eq. A.7) to fit the "observed" LDF.

A spectrum with one thousand events per energy bin ($\Delta \log(E) = 0.1$) from $10^{19.0}$ to $10^{20.5}$ eV is used as input in our simulation code. The zenith angle is extracted randomly from an isotropic distribution from 0 to 60^o and the azimuth angle is selected randomly

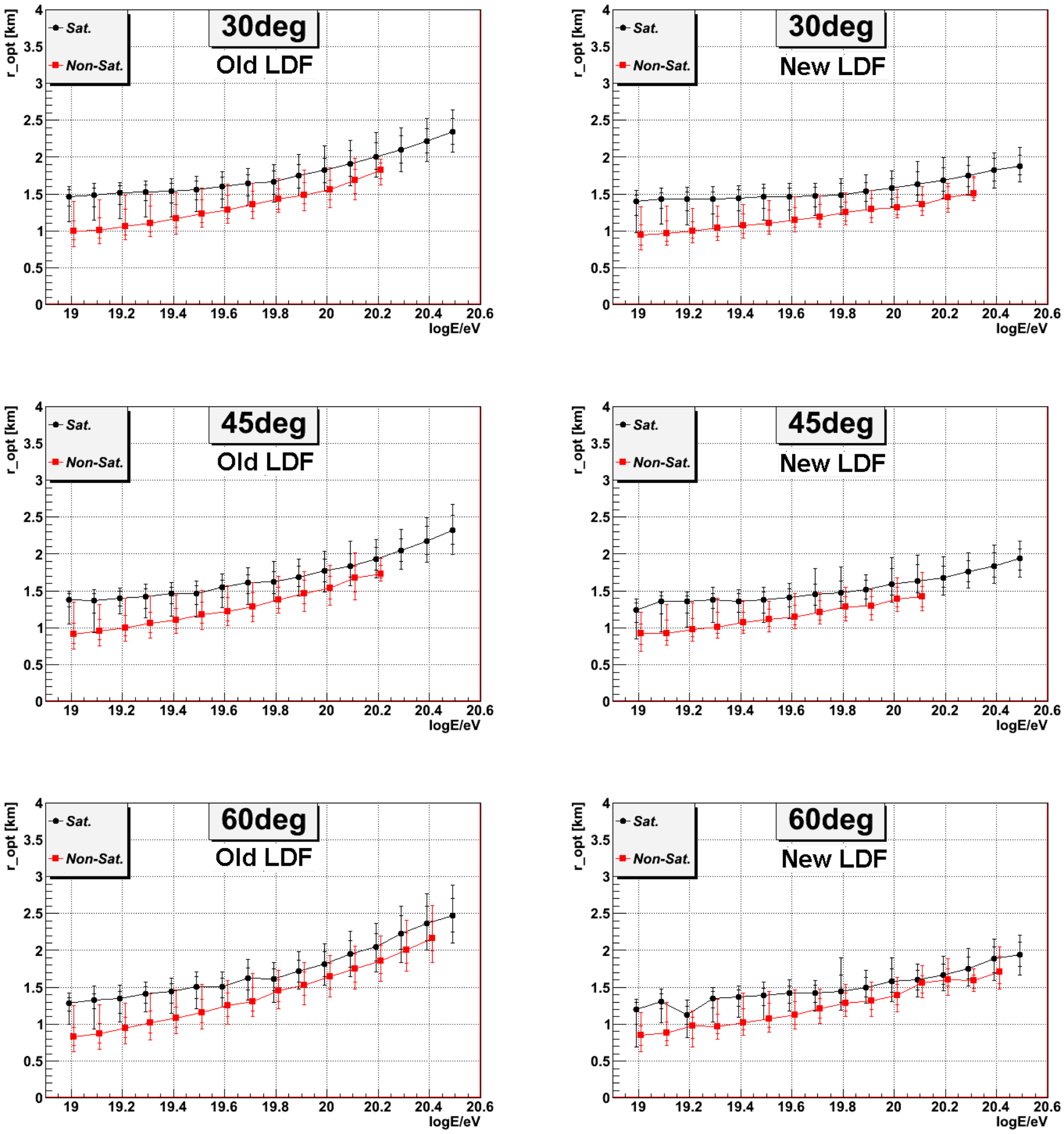

Figure A.5: The optimum distance, r_{opt}, as a function of the primary energy. Events with and without saturated detectors are shown separately for three different zenith angles. Left: Old LDF. Right: New LDF.

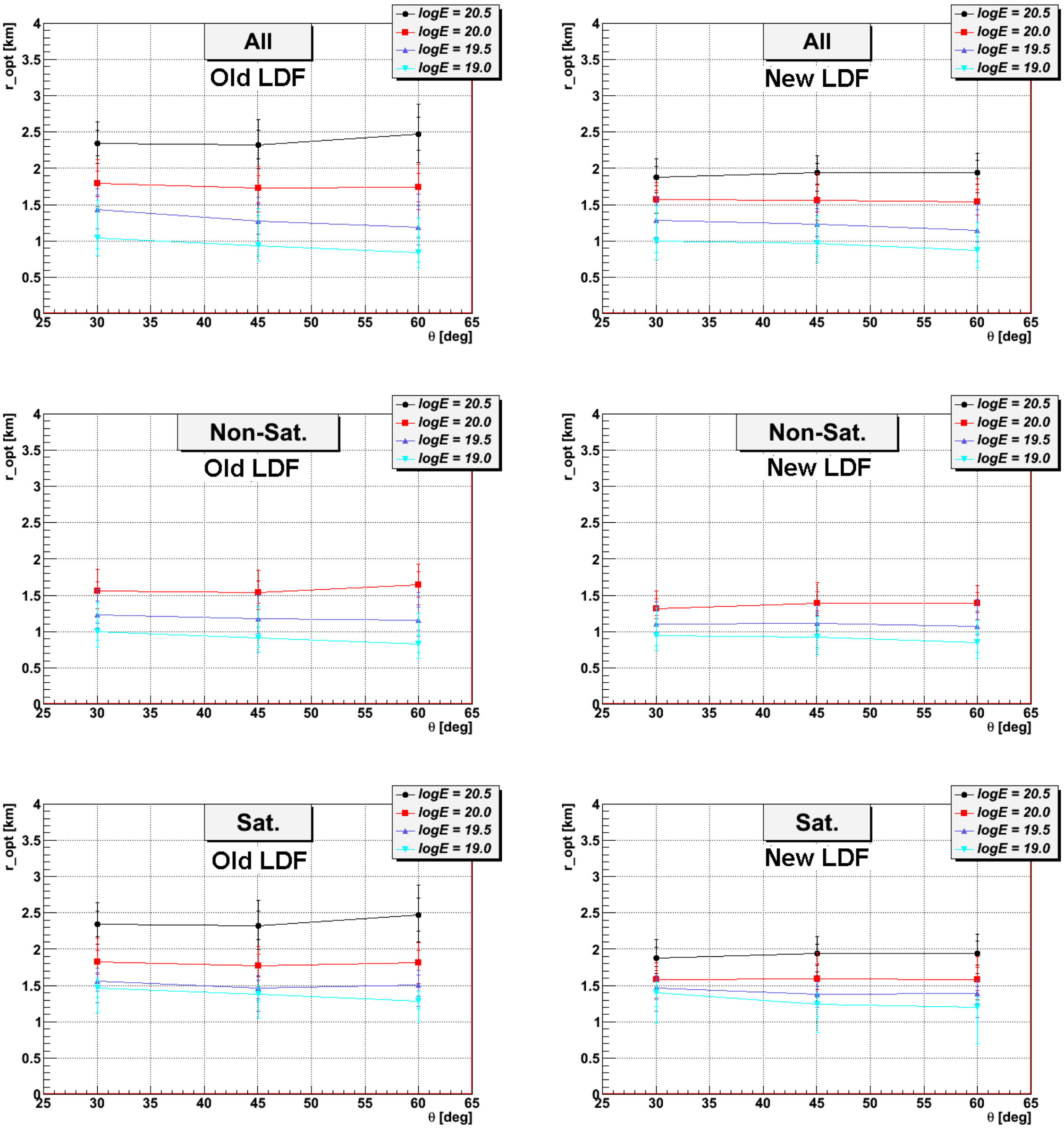

Figure A.6: r_{opt} dependence on zenith angle. Top: All events. Medium: Non-saturated events. Down: Saturated events. Left: Old LDF. Right: New LDF. Error bars are the C.L. at 68% and 95%.

between 0 and 360^o. The core is randomly inside an elementary square cell.

A.4.1 Shape of the energy error distributions

We calculate the distribution function of the errors in the reconstructed energy, i.e. the difference between the reconstructed and the real energy, as a function of the real energy for both techniques, r_0 and r_{opt}. As explained in Section 4.4.1, it is desirable that the errors in energy reconstruction are distributed Gaussianly.

Figure A.7 (Figure A.8) shows for the Old (New) LDF, and for both r_0 and r_{opt}, the 68% and 95% CL for the right and left sides with respect to the median value of the energy error distributions. It can be seen, that the error distribution functions resulting when using $S(r_{opt})$ as energy estimator are, in general, more compact and symmetrical than the corresponding distributions when $S(r_0)$ is used instead. Therefore, the errors in energy reconstruction are lower using $S(r_{opt})$ and the distributions are more Gaussian-like. This behavior appears also if only non-saturated events or only saturated ones are considered. The same result were found in Chapter 4.

A.4.2 Bias in the reconstructed energies

In order to assess any possible bias for both techniques, we compare in Fig. A.9 the relative error in the inferred energy as a function of the injected energy for both LDFs. Again saturated, non-saturated and both types of events together are shown. No bias is introduced when using r_{opt} in any situation. However, the r_0 method introduces a bias due to the difference between $r_0 = 1500$ m and the actual r_{opt} of the event, which is shown as function of energy in Fig. A.10. When $r_0 > r_{opt}$ the energy using r_0 is overestimated and vice versa. In Fig. A.11 the scatter plots of the inferred energy vs the real one are also shown.

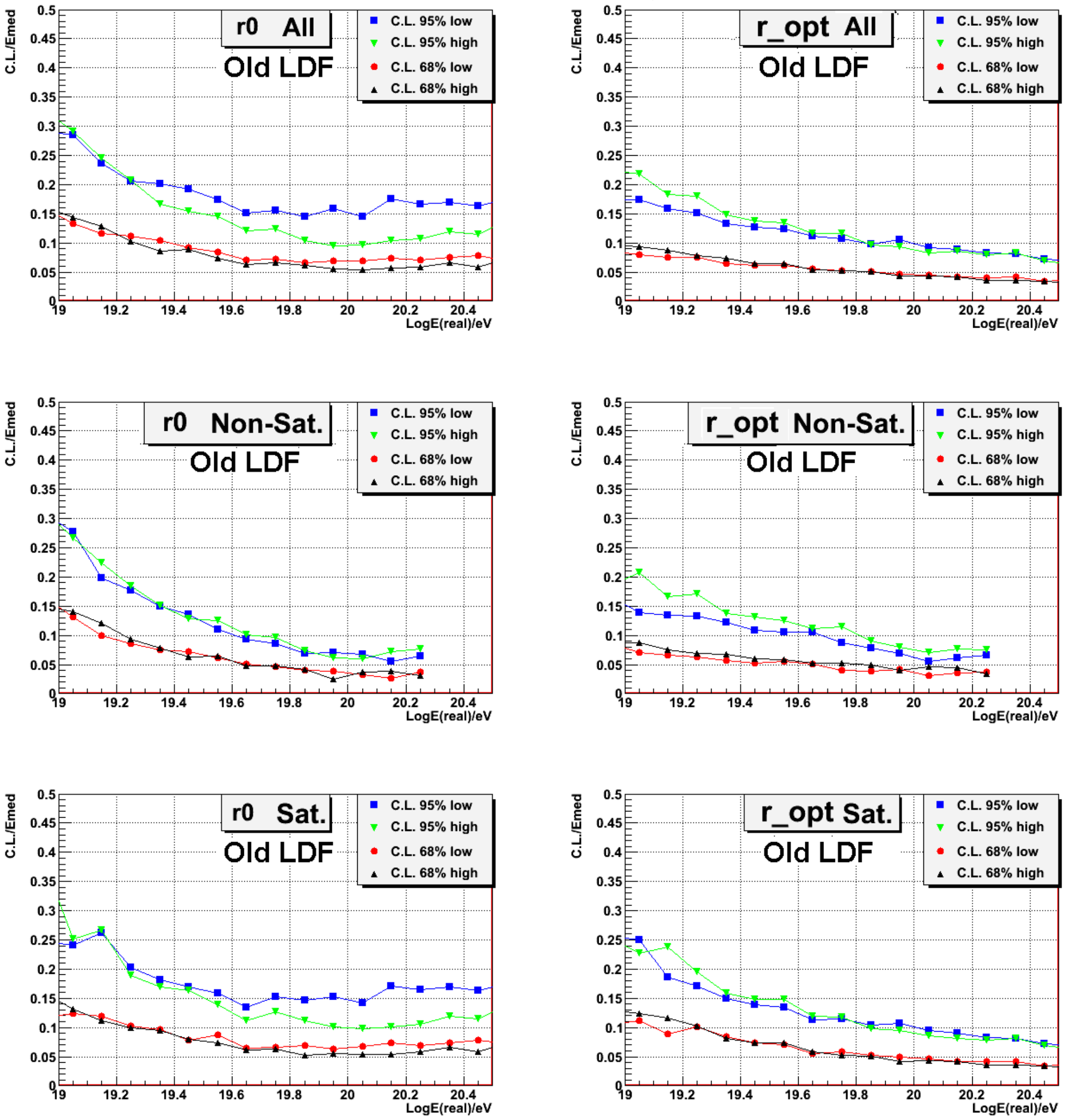

Figure A.7: C.L. at 68% and 95% over the median, from the low and high energy sides, of the energy error distributions obtained when using either $S(r_0)$ (left) or $S(r_{opt})$ (right) as energy estimators. Top: All the events. Med: Non-saturated events. Down: Saturated events. The Old LDF is used.

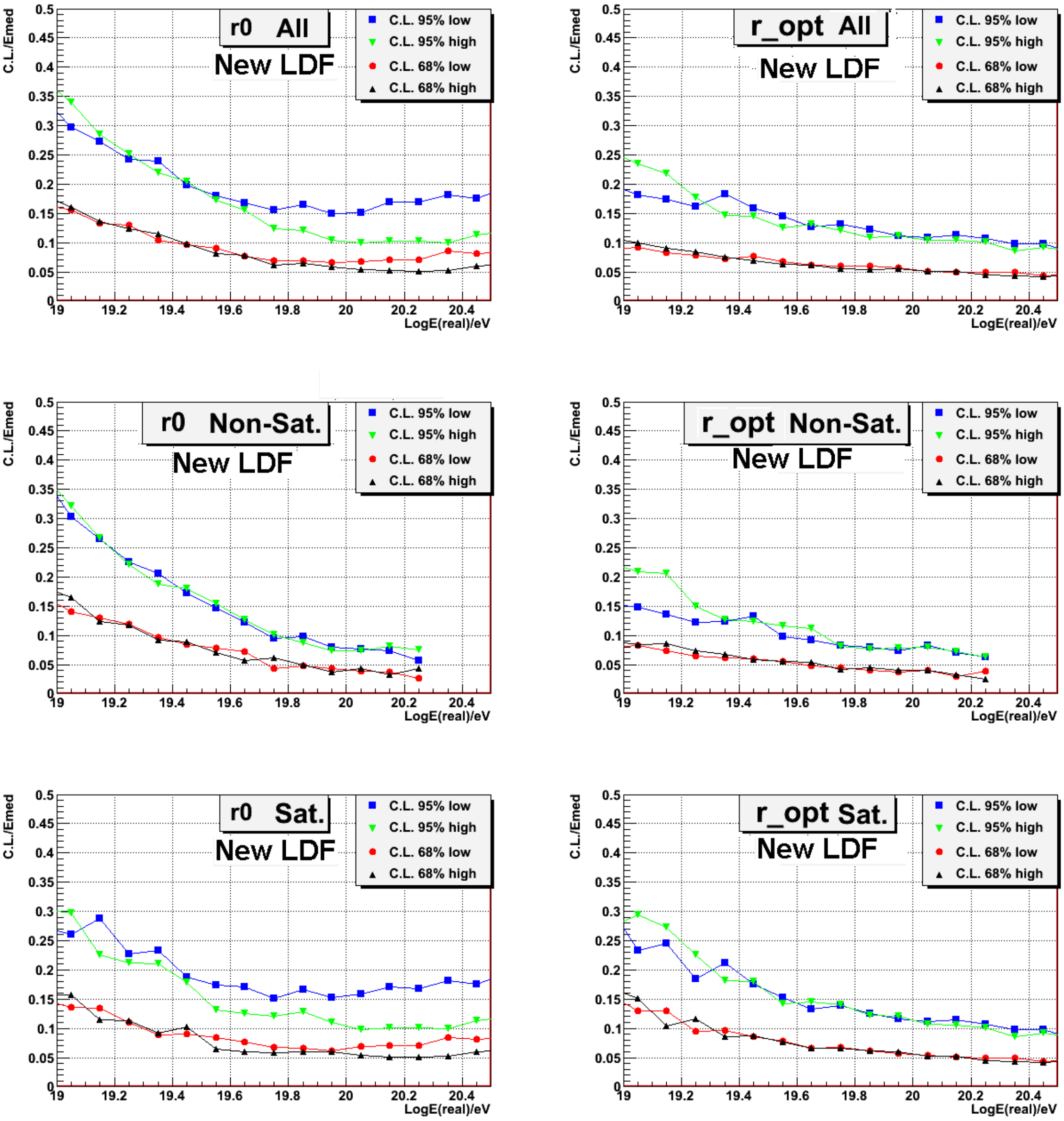

Figure A.8: As Figure A.7 but for the New LDF.

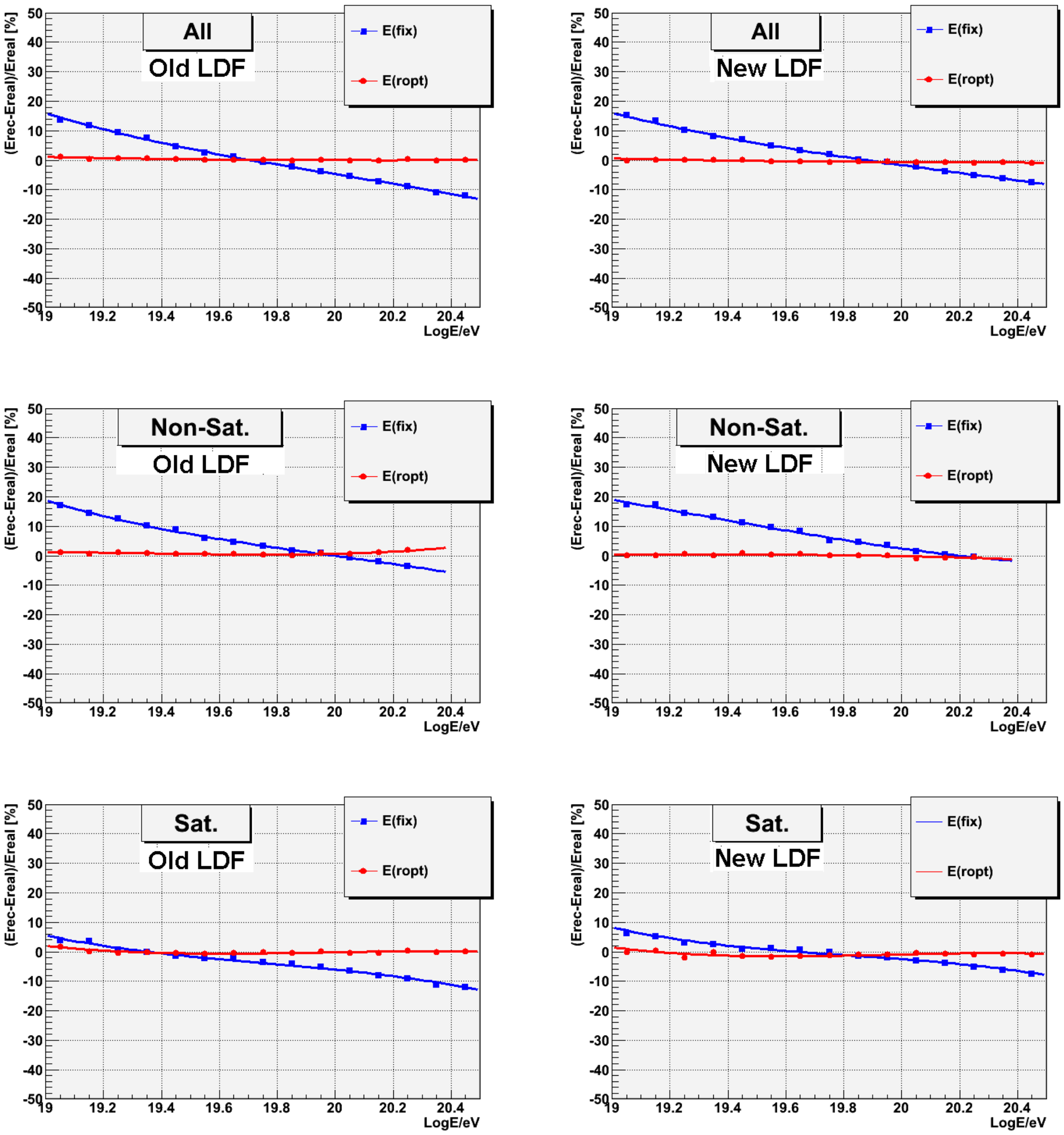

Figure A.9: Bias in the reconstruction methods for all the events (top), non-saturated (medium), saturated (down). Left: Old LDF. Right: New LDF. Comparing with Fig. (A.10), where r_{opt} as a function of energy for saturated and non-saturated events is shown, is clear that the bias is due to the difference between $r_0 = 1500$ m and the actual r_{opt} value of the shower. When $r_0 > r_{opt}$ the energy using r_0 is overestimated and vice versa.

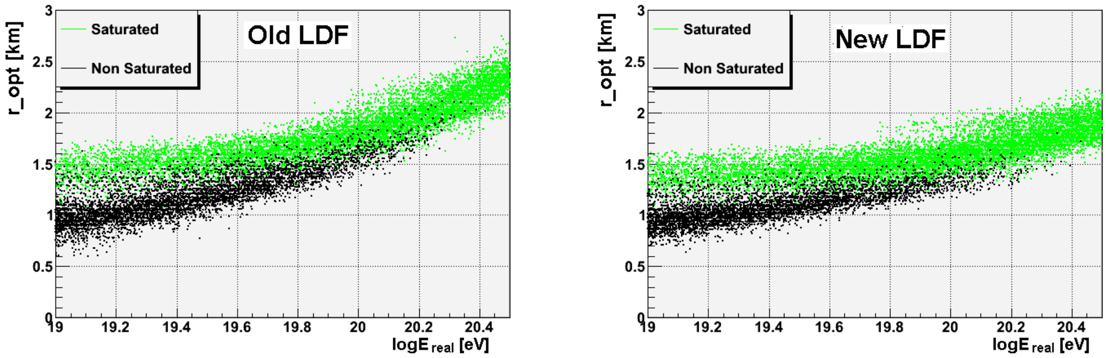

Figure A.10: r_{opt} as a function of energy with saturated and non-saturated events marked. Left: Old LDF. Right: New LDF.

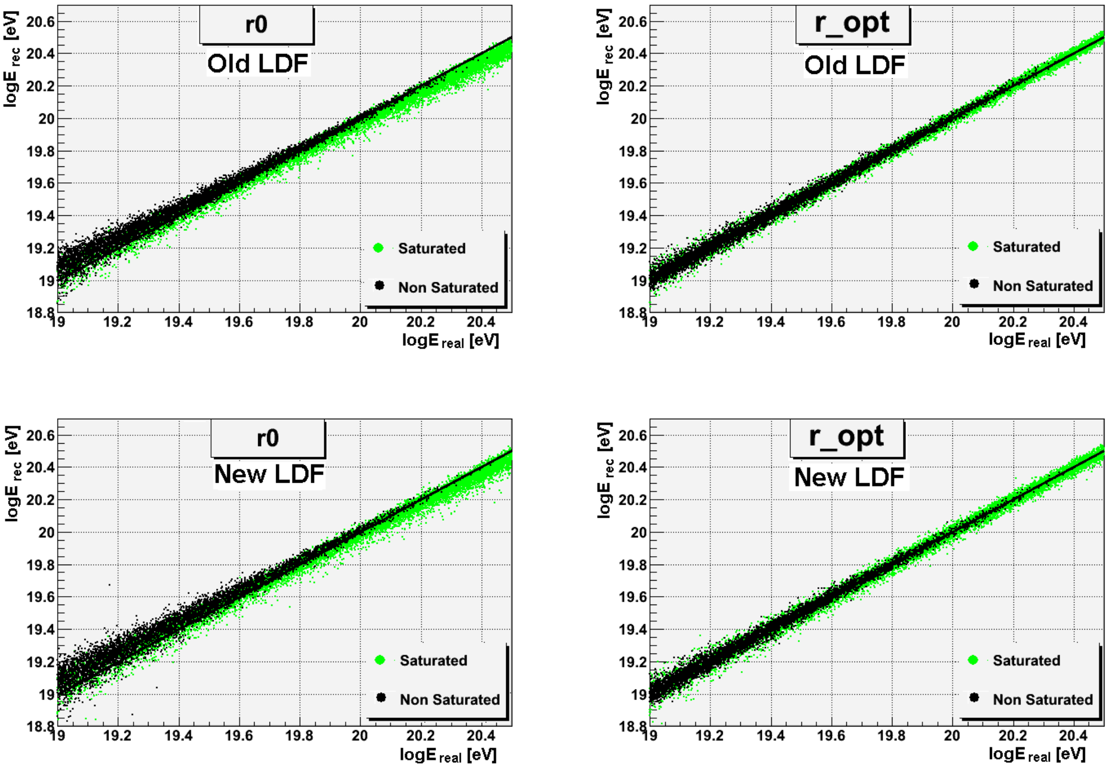

Figure A.11: Inferred energy vs real energy. Left: r_0. Right: using r_{opt}. Top: Old LDF. Down: New LDF. Saturated and non-saturated events marked.

A.5 Summary and discussion

In this Appendix we have applied the study for pure SD arrays presented in Chapter 4 to an square array of 1 mile spacing, that was one of the preliminary designs for the Auger North Observatory. However, in the last ICRC that took place in last July, it was confirmed that the detectors spacing in the final design will be $\sqrt{2}$ miles [199]. This change would modify the values of r_{opt} shown here, but the dependencies and the conclusions achieved about the energy spectrum reconstruction would remain the same. The study is going to updated to a $\sqrt{2}$ mile array soon.

We have calculated with our algorithm, and on shower-to-shower basis, the optimum distance of the LDF, r_{opt}, at which the interpolated signal is the best energy estimator. It is shown the advantages of using the optimum distance for each individual shower instead of a fixed value, as it is done in surface arrays and also at Auger South. Two different LDFs and $S1000(E, \theta)$ parametrizations are considered. First, a NKG form and the parametrization from the Auger NIM paper. Second, the current Auger LDF with the parametrization obtained by using the CIC method and the hybrid calibration. Both are different, which leads to different number of triggered detectors for a given event specially at larger distances from the shower axis. Then, different r_{opt} values are expected for each, while the dependencies on the primary energy and zenith angle are equivalent. The values of r_{opt} obtained with the current Auger LDF and the CIC+hybrid method in Section A.3 are likely more reliable, because this parametrization has been found in a detailed study using a full set of real events [202].

We have shown that r_{opt} increases strongly with energy and depends slightly on the zenith angle. Due to the dependency with energy and the large dispersion of the r_{opt} value at a given energy and zenith angle, using the signal at a fix characteristic value as energy estimator could affect the spectrum determination. We reconstruct an input spectrum with our code, and use the signal at both distances as energy estimators, $S(r_{opt})$ and $S(r_0)$. We demonstrate that the energy error distribution functions obtained with the $S(r_0)$ method, are wider, more skewed and have more extended tails than those produce by using $S(r_{opt})$. Furthermore, the $S(r_0)$ method introduce a significant bias for both saturated nor non-saturated events, while, on the contrary, using $S(r_{opt})$ the bias is negligible in the whole energy range and for any type of event.

Therefore, another important advantage of the r_{opt} method is that it allows the same unified treatment for events with and without saturated detectors. Therefore, specific algorithms designed to recover the signal of saturated detectors are not needed (see for example [106, 205, 206, 207, 208, 209]), which avoids possible uncertainties and/or biases introduced by these procedures.

The LDF fit performed in this work is not the same as the one used in Auger. We reject saturated stations in the fit and those below the triggered threshold. On the other hand, Auger makes a more complex maximum likelihood, fitting the LDF and the geometry of the shower simultaneously. In the Auger standard procedure, the signals of saturated stations are used as a lower limit in the fit and silent stations are also included (see Appendix B for details). A discussion about both methods could be found in Chapter 6.

In addition, the Auger North Observatory is here somehow considered as a pure SD experiment. The two energy calibrations explained before, are used to get the shower size as a function of the primary energy and the zenith angle of the event, as it is done by pure SD experiments. Thus, the procedure used in Chapter 4 could be also performed here. However, the application of the r_{opt} method to a hybrid experiment is not straightforward. A new calibration formula between the signal at the optimum distance, $S(r_{opt})$, and the energy measured by the fluorescence telescopes, E_{FD}, would be needed. Initial steps in this direction were explained in Chapter 6.

The relevance of the conclusions attained in the present study for Auger North will very much depend on the final design of the detector that eventually comes out of the *R&D* phase. In any case, the study helps to fully understand the new SD layout and, specifically, it is useful in order to analyze the r_{opt} dependencies on primary energy and zenith angle in a square grid, to compare them to that in a triangular grid, and to study the applicability of the technique in a SD array of water Cherenkov detectors (in Chapter 4 an AGASA-like experiment was selected, i.e. triangular array of scintillators separated

1 km). Furthermore, the main result obtained here, i.e., that the use of the signal at the optimum distance calculated on a shower-to-shower basis is the best energy estimator, should still be of general applicability.

In conclusion, we propose that Auger, in both Southern and Northern sites, performs an energy calibration using the optimum distance on a shower-to-shower basis. The later would require, of course, a new FD-SD specific calibration for $S(r_{opt})$.

Appendix B

Reconstruction of surface events at Auger

This Appendix is devoted to review the procedure performed by the Pierre Auger Observatory to reconstruct the surface data. It is based on the standard reconstruction as it is implemented in the official software of the Observatory, the Software Offline [196]. Another available software for shower reconstruction is the CDAS [210] that was developed previously. A comparison between both could be found in [106]. This Appendix is based on the official description of the SD reconstruction [211], but more specific details are given in several key points.

The reconstruction procedure uses the time and the signal in the triggered detectors. The objective is to determine the lateral distribution of particles at ground level, the incoming direction of the primary and to find accurately the energy estimator, i.e. the signal at 1000 m from the shower core $S(r = 1000) \equiv S_{1000}$. In addition, several parameters useful for composition studies are determined such as the radius of curvature of the shower front and the slope of the lateral distribution function.

First, the stations belonging to the event are defined by checking the time compatibility between an estimated plane shower front and the start time of the signal in each station. An initial fit assuming a plane shower front gives an estimate of the core position and a pre-value for S_{1000} is determined based on the signal of the station closest to the the distance of 1000 m from shower axis. Thereafter, the fitting procedure is done by using a maximum likelihood in an iterative process, where silent, zero-signal and saturated stations are taken into account. Several approaches are possible as it will be explained below. Later, the curvature of the shower front is reconstructed by fixing the core position. The curvature fit changes the axis of the shower, which modifies the slope of the LDF, so the LDF is fitted again. However, the new LDF fit modifies the core position and, consequently, the shower axis and the curvature, so the process could continue until it is decided to stop. Details are given in the following about each step.

B.1 Station and event selection

Triggered stations could be rejected by several reasons. For example, they could be accidentally triggered and must be identified and discarded for event reconstruction (they are flagged as *accidental*). For example, atmospheric muons could triggered a detector. They are considered as lonely stations and removed if it has no neighbor in 1800 m, or only one in 5000 m, and also based on timing information. In addition, stations with lightning-like signals (oscillations in the FADC traces of all three PMTs), those belonging to the Engineering Array, doublets (i.e. pairs of stations located very close that are used to study signal and timing accuracy) or the infill array (located outside the regular grid of the array which will be used in the AMIGA extension [101]) are also discarded.

Three local triggers are defined for individual stations by coincidences of the three photomultipliers. They are explained in the following.

T1: identifies signals that could be relevant in the event reconstruction procedure. Two different modes are implemented:

- Threshold: The signal level is checked for 3-fold coincidence (all PMTs) above the threshold set at 1.75 VEM.
- Time-over-threshold: The signal is checked for the 2-fold coincidence of the time-

over-threshold trigger (TOT) which requires more than 12 FADC bins with signal 0.2 VEM above baseline in a window of 120 bins (3μs).

T2: selects T1 signals that likely come from air showers. Stations flagged as TOT are accepted without any more requirements, while those that passed the T1 threshold trigger are checked again for a 3-fold coincidence, but now the threshold is set to 3.2 VEM.

T3: this trigger considers several tank configurations that passed T2 level which may have been caused by a single shower. T3 operates in two modes:

- The main T3 trigger condition requires at least three T2 stations in time coincidence that have also passed the TOT condition. In addition a minimum compactness is required, which is fulfilled if 2 stations are within the first ring and 3 within the second one. In this mode, the 90% of the selected events are real showers.
- Another mode is needed to detect horizontal showers that generate fast signals and have wide-spread topological patterns. At least three T2 triggered stations are required around the considered station. At least one of them has to be in the first crown, at least two stations have to be within the second ring and the third one has to be within the fourth ring.

The time compatibility between detectors is checked by performing an initial geometrical reconstruction of the shower arrival. To that end, the three stations that maximize the sum of the signals are used to determine a planar shower front and an initial shower axis (details are given later). The procedure is called *bottom-up* selection. With this axis and supposing that the shower front travels at the speed of light, the time start of the signal at the station located at x_i, is predicted ($t_{sh}(x_i)$) and compared with the measured value (t_i). If the difference $\Delta t_i = t_i - t_{sh}(x_i)$, usually called the time delay of the station, does not satisfy that $-1000ns < \Delta t_i < 2000ns$, the station is not included in the event and it is flagged as *accidental.* The asymmetry in the values is caused by the larger probability for the stations with lower signal to be delayed due to the curved shower front.

The minimum quality of the events useful for reconstruction purposes is set by other three trigger criteria explained below.

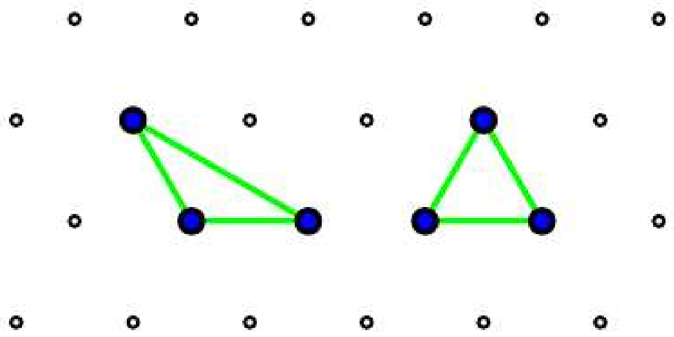

Figure B.1: The two possible 3TOT compact congurations (with addition of all of the symmetry transformations of the triangular grid).

T4: a compact configuration of the selected stations is required to reduce the number of random coincidences. Again two different modes are available:

- 3TOT configuration: it is devised to select physics events with arrival zenith angle of up to 60°. It requires at least 3TOT stations forming a triangle of first neighbors as shown in Fig. B.1. Almost 99% of the events passing this trigger are real showers.
- 4C1 configuration: it is designed to recover the 5% of real events lost by 3TOT configuration and to select events above 60°. It requires 4 stations with any type of T2 trigger where the central station must be surrounded by the other three located in the first crown (see Fig. B.2).

T5: It is a quality trigger that filters out events with a deficient reconstruction caused by the absence of some stations. That mainly happens in those events that fall too close to the edge of the array (that was growing until mid-2008). The station with the largest signal is required to have six nearest (first crown) neighbors present and functioning, though not necessarily triggered, at the time of the shower impact. In this way it is guaranteed that the core of the shower is contained inside the array and a signicant fraction of the shower is sampled to ensure a good reconstruction.

T5 posterior: the T5 trigger cause a significant loss of good quality events when the array was under construction. To recover those events for which no crucial information

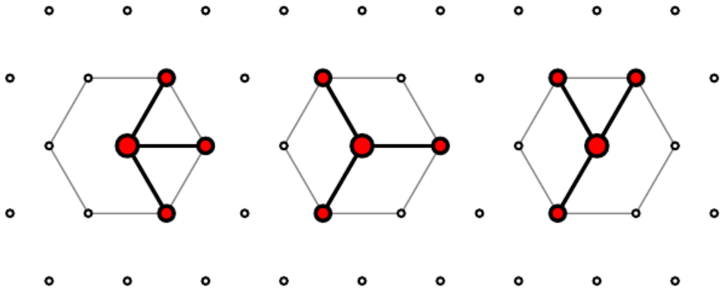

Figure B.2: The three (minimal) 4C1 congurations (with addition of all of the symmetry transformations of the triangular grid).

for reconstruction were lost, a T5 posterior flag is set which relaxes the condition and requires that the shower core position must lie within the equilateral triangle of functioning stations. When doing analysis studies, the user must select if T5 is enough for his study or T5 posterior is preferred.

Regarding saturated detectors, it is possible to recover their signal. Different methods have been tested taking into account the FADC channel overflow as well as the PMT non-linearity. Details could be found in from Ref. [205] to Ref. [209]. The recovered signal is used only when the second derivative of the normalized LDF is smaller than 1 in order to avoid the rapid increase of the LDF approaching the shower core. In the standard reconstruction, the saturation recovery is not set by default, while the user could switch it on.

B.2 Plane fit to the shower front

As a first estimate of the shower axis ($\overrightarrow{a}$), the start time of the signal at each detector and their position is used to determine a planar shower front. The origin of coordinates is the signal-weighted barycenter ($\overrightarrow{b}$) and the weighted bary-time (t_0) of the stations involved in the fit. A shower track (see Fig. B.3) could be visualized as a point $\overrightarrow{x(t)}$ moving with the speed of light c along the straight line with (normalized) axis $\overrightarrow{a}$, and passing the origin at time t_0. Therefore, $-\overrightarrow{a}(\overrightarrow{x(t)} - \overrightarrow{b}) = c(t - t_0)$. Here , the shower barycenter is considered as the core of the shower.

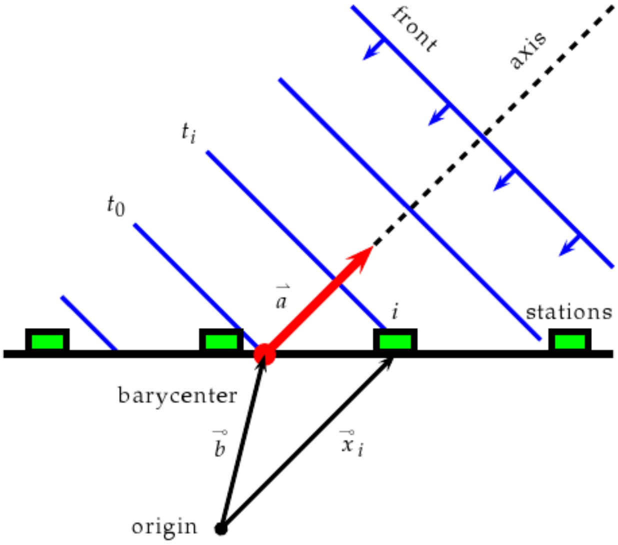

Figure B.3: Sketch of the plane front arrival.

The *shower plane* is a plane perpendicular to the shower axis, moving along with the same speed and containing the shower forehead. To infer the time ($t(\overrightarrow{x})$) when the shower plane pass through a chosen point $\overrightarrow{x}$ on the ground, the point has to be projected to the shower axis:

$$ct(\overrightarrow{x}) = ct_0 - (\overrightarrow{x} - \overrightarrow{b})\overrightarrow{a} \tag{B.1}$$

The only deviations can be due to the time uncertainty σ_t of the signal start because the position of the stations is supposed to be given with absolute precision. Therefore, the function to minimize is the square of the time differences between the measured signal start and the predicted time. If the components of shower axis are $\overrightarrow{a} = (u, v, w)$ and the station coordinates are $\overrightarrow{x}_i = (x_i, y_i, z_i)$ the function is

$$\chi^2 = \sum_i \frac{[ct_i - ct_0 + x_i u + y_i v + z_i w]^2}{c^2 \sigma_{ti}^2} \tag{B.2}$$

with the constraint $u^2 + v^2 + w^2 = 1$ inherited, which makes the problem non-linear. Nevertheless, an approximate solution could be obtained if it is supposed that all stations lay close to a plane, so $z_i << x_i, y_i$ so the z component is neglected (more details in

[211]). This approximation is quite good because the differences in the altitude in Auger layout are lower than 20 meters. In the minimization, the shower axis is determined. This procedure only fails when there is a linear dependence of the z-projected station positions, as for example, when there are three stations in a line. However, for higher station multiplicity this is highly unlikely.

B.3 The lateral distribution function

First, in order to fit the lateral distribution function of particles, the core position is needed. The first estimate was the signal-weighted barycenter of the stations and it was used to estimate the shower axis assuming a plane shower front. Now, the core position is required i) to lie in the plane tangent to the Earth's reference ellipsoid which contains the barycenter of the stations and, ii) to belong to the estimated shower axis determined as explained in previous Section. Thus, the core is determined unambiguously.

The lateral dependence of the signal measured in the tanks is modeled as

$$S(r) = S_{1000}\, f_{LDF}(r) \tag{B.3}$$

where $f_{LDF}(r)$ is a particular shape parametrization normalized such that $f_{LDF}(1000m) = 1$ holds. The uncertainty in the signal is [212]

$$\sigma_S(\theta) = (0.32 + 0.42 \sec\theta)\sqrt{S} \tag{B.4}$$

Several functional forms of the LDF have been investigated [81, 202], and it was found that the best description of data is given by a modified NKG given by

$$f_{LDF}(r) = \left(\frac{r}{1000}\right)^{\beta(\theta)} \left(\frac{r+700}{1000+700}\right)^{\beta(\theta)} \tag{B.5}$$

where r is given in meters. Initial estimate is $\beta(\theta) = 0.9 \sec\theta - 3.3$.

It is an important debate if the slope, β, must be free or fixed in the fit. The actual reconstruction in the Observer [213] which is used by most part of the collaboration and it is based on the official Offline Software [196] with the options set by default, fixes the slope of the LDF according to the following parametrization

$$\begin{aligned}\beta(\theta, S_{1000}) &= -3.35 - 0.125 \log S_{1000} \\ &\quad +(1.33 - 0.0324 \log S_{1000}) \sec\theta \\ &\quad +(-0.191 - 0.00573 \log S_{1000}) \sec^2\theta \qquad \text{(B.6)}\end{aligned}$$

It is also possible to modify the options and to let β free in the fit. However, by this procedure a good reconstruction is achieved only if one of the next conditions, determined by P. Billoir, is fulfilled (otherwise, it must be fixed if a good reconstruction is desired):

- at least 2 candidate stations with r in the interval [500, 1500] m, with maximum difference in r of at least 500 m,
- at least 3 candidate stations with r in the interval [500, 1500] m, with a maximum difference in r of at least 400 m,
- at least 4 candidate stations with r in the interval [500, 1500] m, with a maximum difference in r of at least 300 m.

The official choice is to fix the slope in the standard reconstruction as commented previously. The reason is that it makes possible to reconstruct all the events with the same procedure independently on the number of candidate stations and whether there exist a saturated detector. Then, the reconstruction of the LDF is done with three free parameters: S_{1000} and the core location (x and y components). The uncertainty on S_{1000} from fixing β are obtained doing two additional reconstructions with $\beta \pm 3\%$.

B.4 Maximum Likelihood

A χ^2 minimization could be performed which involve 4 parameters: S_{1000}, the core position (only x and y in the local tangent plane) and the slope parameter β. Details are available at [211]. However the maximum likelihood method explained next is the usual procedure.

In the maximum likelihood the first step is to define an effective particle number, and thus, it is possible to include zero-signal stations, small signals (i.e. small particle densities) by means of Poisson statistics, large signals by a Gaussian approximation and to handle the signal of saturated events. The water Cherenkov tanks provide information about Cherenkov photons, which are released by muons, electrons or converted photons when passing through the tanks. The energy deposit, or equivalently, the number of registered photo-electrons, depends strongly on the particle type, injection point and incident angle. Therefore, it is not straightforward to find a conversion formula from the signal measured to the number of particles.

The total signal measured in a tank has two main contributions, the muonic and the electromagnetic (γ and e^-/e^+). Assuming that a single converted photon and a simple electron equally energetic deposit the same mean signal in the tank, it could be written that $S = S_\mu + S_{e/\gamma}$. A muon is considered to deposit 1 VEM irrespective of incoming angle or distance. On the other hand, the signal $S_{e/\gamma}$ is much smaller than S_μ and the mean conversion factor for electrons and photons to signal is smaller than 1 VEM. The total number of particles that have produces the signal is then estimated as $n = p(r, \theta, E, A)S$, where $p(r, \theta, E, A)$ is called the *Poisson factor*, which, in principle, could depend on the primary energy (E), mass (A), zenith angle (θ) and on the distance from the tank to the shower axis (r). The signal recorded in tanks close to the trigger threshold, specially at larger distances, have a large muon content, so p is taken to be 1. Assuming that the transition when the electromagnetic component deposits half of the signal takes place when the signal exceeds a threshold S^G_{thres} = 15 VEM, the final simplified conversion between signal and effective particle number is

$$n(r, \theta, E, A) = n = \begin{cases} S(r, \theta, E, A) & \text{if } S < S^G_{thres} \\ 2S(r, \theta, E, A) & \text{if } S \geq S^G_{thres} \end{cases} \tag{B.7}$$

Once an effective number of particles has been defined, it is possible to compose the maximum likelihood function to estimate the LDF. The function gathers the information of each tank i at distance r_i and it is given by

$$L = \prod_i f_P(n_i, \mu_i) \prod_i f_G(n_i, \mu_i) \prod_i f_{sat}(n_i, \mu_i) \prod_i f_{zero}(n_i, \mu_i) \tag{B.8}$$

and thus, the log-likelihood function to maximize is

$$l = \sum_i \ln f_P(n_i, \mu_i) + \sum_i \ln f_G(n_i, \mu_i) + \sum_i \ln f_{sat}(n_i, \mu_i) + \sum_i \ln f_{zero}(n_i, \mu_i) \tag{B.9}$$

where n_i is the effective number of particles detected in the tank obtained form Eq. B.7 and μ_i the corresponding theoretical LDF expectation. The different contributions are:

i) Small signals: tanks with signal lower than $S^G_{thres} = 15$ VEM, which corresponds to a lower number of particles, have a Poisson distribution

$$f_P(n_i, \mu_i) = \frac{\mu_i^{n_i} e^{-\mu_i}}{n_i!}, \text{ and} \tag{B.10}$$

$$\ln f_P(n_i, \mu_i) = n_i \ln\mu_i - \mu_i - \sum_{j=1}^{n_j} \ln j \tag{B.11}$$

ii) Large signals ($S_i \geq S^G_{thres} = 15$ VEM): for large number of particles it is possible to apply the Gaussian approximation

$$f_G(n_i, \mu_i) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{(n_i - \mu_i)^2}{2\sigma_i^2}\right), \text{ and} \tag{B.12}$$

$$\ln f_G(n_i, \mu_i) = -\frac{(n_i - \mu_i)^2}{2\sigma_i^2} - \ln\sigma_i - \frac{1}{2}\ln(2\pi) \tag{B.13}$$

Note that the last term is constant so it could be omitted from the minimization procedure.

iii) Saturated signals: The saturated signal, n_i, represents a lower limit on the actual signal. Integrating f_G over all possible values larger than n_i, it is possible to obtain an estimate of the probability of detecting a signal larger than n_i. Therefore,

$$f_{sat}(n_i, \mu_i) = \int_{n_i}^{\text{inf}} f_G(n_i, \mu_i) dn = \frac{1}{2} \operatorname{erfc}\left(\frac{n_i - \mu_i}{\sqrt{2\pi}\sigma_i}\right) \tag{B.14}$$

where $\operatorname{erfc}(x) = 1 - \operatorname{erf}(x)$ is the complementary *error function*. Another possibility is trying to recover the saturated signal as explained in [209], and then, the recovered signal substracted by its uncertainty is used as a lower limit. The recovery method is only used when the second derivative of the normalized LDF is smaller than 1 in order to avoid the rapid increase of the LDF approaching the shower core.

iv) Zero-signal stations: the assumed threshold to trigger a tank is $n_{th} \equiv 3$, i.e. 3 muons hitting the tank. Therefore the contribution of the stations without signal is a sum over all Poisson probabilities with a predicted particle number μ_i and actual particle number $n_i \leq n_{th}$:

$$f_{zero}(n_{th}, \mu_i) = \sum_{n=0}^{n_{th}} f_P(n, \mu_i), \quad \text{and} \tag{B.15}$$

$$\ln f_{zero}(n_{th}, \mu_i) = -\mu_i + \ln \left(\sum_{n=0}^{n_{th}} \frac{\mu_i^n}{n!} \right) \tag{B.16}$$

Maximizing l (Eq. B.9) the lateral distribution function is obtained, and therefore, S_{1000} and the core location.

B.5 Curvature shower front

As a first approximation of the radius of curvature, the shower is approximated as starting at time t_0 from one single point (see Fig. B.4) and propagating towards the stations, so the timing t_i at the station i is $c(t_i - t_0) = |R_c - x_i|$, with R_c the apparent origin of the shower. Therefore, the shower front is described as an expanding sphere. This way, timing information is decoupled from the information about the impact point. With this model and the assumption that $z_i \approx 0$ as in the planar front, it is possible to determine the radius of curvature R_c analytically. The solid angle differences between the plane fit and the curvature fit axis $\overrightarrow{a}$ are of the order of a half degree.

The most realistic model is to extent the method described previously in Section B.2, where a planar shower front were assumed, with a parabolic term which describes the curvature of the shower front near the impact point. Thus, the Eq. B.1 can be modified to get

$$ct(\overrightarrow{x}) = ct_0 - \overrightarrow{a}\,\overrightarrow{x} + \frac{\rho(\overrightarrow{x})^2}{2R_c} \tag{B.17}$$

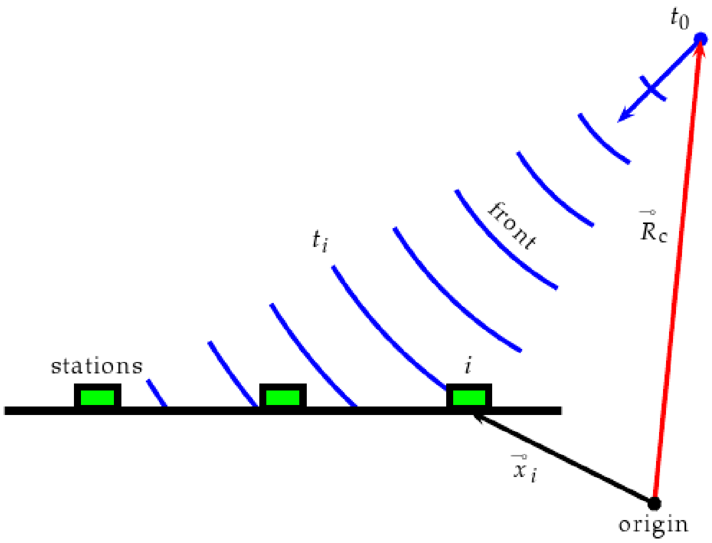

Figure B.4: Sketch of the spherical shower front development.

with perpendicular distance $\rho(\overrightarrow{x})^2 = (\overrightarrow{a} \times \overrightarrow{x}) = x^2 - (\overrightarrow{a}\,\overrightarrow{x})^2$. As in the plane fit, a χ^2 minimization is performed but now taking into account the z coordinate:

$$\chi^2 = \sum_i \frac{[c(t_i - t_0) - |R_c \overrightarrow{a} - \overrightarrow{x_i}|]^2}{c^2 \sigma_{ti}^2} \tag{B.18}$$

The differences to the approximate estimation of R_c (spherical front) and this one are of the order of few tens of meters, while the solid angle differences between the axes is of the order of a few 0.1°.

B.6 Fit stages

First, the shower geometry reconstruction is performed as summing a plane shower front as it was explained above. The core is located at the signal-weighted barycenter of the stations included in the fit. Second, with this core position and an initial estimate for the slope of the LDF ($\beta(\theta) = 0.9 \sec\theta - 3.3$), an S_{1000} pre-factor is estimated based on the station closest to the perpendicular $r = 1000$ m.

In a third stage, using these estimated values of S_{1000} and core location, the fit is performed. The core has only two components, x and y while z is assumed zero. In this minimization process, β could be also determined if it is let to vary, while in the standard reconstruction it is considered fixed using the parametrization given by Eq. B.6 as explained previously.

Later, the curvature of the shower front could be determined. To that end, the core is fixed in the position determined in the previous stage. Unfortunately, the LDF and curvature fits are linked. The curvature determination modifies the axis of the shower, so the LDF fit from the third stage is performed again. This changes the core position and, therefore, the curvature, so the curvature fit is done once again. It is not clear which must be the final step. In the actual standard reconstruction it is the LDF fit while in the last Collaboration Meeting (April 2009) it has been decided that it is better to end with the curvature fit in order to get a better geometrical reconstruction [214]. At present, the LDF task group is working on a global fit, where the slope of the LDF and the curvature are fitted at the same time, which would solve this issue.

B.7 Energy estimation

The energy of surface events is determined by using the hybrid calibration. It is obtained from a selection of high quality *golden* hybrid events, those that could be reconstructed by the surface array and the fluorescence telescopes independently.

A brief description has been given in Section A.1 and all the procedure and systematic and statistical uncertainties are explained in detail in [103, 106, 163]. First, as a result of a CIC analysis, S_{1000} is converted into a reference signal size, S_{38°, by

$$S_{38^\circ} = \frac{S_{1000}}{1 + 0.92x - 1.13x^2}, \tag{B.19}$$

where $x = \cos^2(\theta) - cos^2(38^\circ)$. The CIC parametrization is suitable up to a zenith angle of about 81°.

The effect of atmospheric variations (in pressure, temperature and density) on extensive air showers have been studied in [215], by using about 960.000 events collected by the surface detector of the Pierre Auger Observatory from 1 January 2005 to 31 August 2008. A significant modulation is observed in the rate of events with the atmospheric variables, both on a seasonal scale ($\sim$ 10%) and on a shorter time scale ($\sim$ 2% on average during a day). This modulation can be explained as due to the impact of the density and pressure changes on the shower development, which affects the energy estimator S_{1000}. This affects the trigger probability and the rate of events above a fixed energy. The dominant effect is due to the change with the air density of the Moliere radius near ground.The second effect is due to the pressure changes, which affect, through the variation of the amount of matter traversed, the stage of development of the showers when they reach ground.

The energy estimate is then calculated using the hybrid calibration [163, 164]:

$$E = A \cdot S_{38^\circ}^{B} \tag{B.20}$$

where A = (1.49 ± 0.06 (stat) ± 0.12 (syst)) $\times 10^{17}$ eV and B = 1.08 ± 0.01 (stat) ± 0.04 (syst).

Appendix C

Resumen

C.1 Antecedentes

Los rayos cósmicos han sido desde el siglo pasado una fuente inagotable de información del Universo. Víctor Hess descubrió en 1912 [1] que radiación de origen extraterrestre llega hasta la Tierra, fenómeno al que se denominó *radiación cósmica*. Desde ese momento, físicos de todo el mundo se interesaron en este fenómeno ya que representaba una nueva ventana para estudiar las interacciones entre partículas muy energéticas. Esto condujo al descubrimiento de varias partículas elementales como el positrón [8] o los muones [10]. El interés no ha decrecido en este siglo, sobretodo a las más altas energías, por encima de 10^{18} eV. Esto se debe a que varias de las principales preguntas siguen sin respuesta: ¿de dónde vienen?, ¿cómo alcanzan tan altas energías?, ¿cuál es su composición?, ¿cómo se deben interpretar los cambios que se observan en su espectro de energía?, etc.

A pesar de que estas preguntas aún necesitan una respuesta completa, se ha avanzado mucho en el estudio de los rayos cósmicos. Su espectro se extiende a lo largo de 12 órdenes de magnitud, desde 10^9 hasta más de 10^{20} eV, que son las partículas más energéticas del Universo (incluso 2 órdenes de magnitud superior a las energías que se esperan alcanzar en el Large Hadron Collider del CERN). Su flujo decrece con la energía como $\sim E^{-3.0}$, por lo cual, aunque los rayos cósmicos de energías más bajas bombardean a la tierra a un ritmo de 1000 por segundo y metro cuadrado, los de más alta energía apenas llegan 1 por km^2 y por siglo.

Esta enorme diferencia en el flujo y el amplio rango de energías, hace que sean necesarias técnicas muy diferentes para su estudio. Por un lado, los rayos cósmicos de energías hasta 10^{15} eV pueden ser detectados directamente, mediante globos que se lanzan a las capas más altas de la atmósfera o desde satélites. Los rayos cósmicos de energías mayores llegan con un flujo tan bajo que no se pueden detectar directamente en la práctica. Sin embargo, se pueden estudiar indirectamente mediante la detección de las llamadas cascadas extensas de partículas (EAS, por sus siglas en inglés Extensive Air Showers). Estas cascadas, descubiertas por Pierre Auger en 1938 [13], son originadas por los rayos cósmicos al interaccionar con los núcleos de la atmósfera, lo que produce una serie de interacciones y decaimientos en los que se generan billones de partículas. Del estudio de las propiedades de las cascadas de partículas se pueden deducir las del rayo cósmico primario.

Esta tesis se centra en el estudio de los rayos cósmicos de ultra-alta energía (UHECR, en inglés Ultra-High Energy Cosmic Rays), aquellos con energas mayores de 10^{18} eV. Para el estudio de las cascadas generadas por ellos, existen dos técnicas principales. Por un lado, se colocan detectores al nivel del suelo, habitualmente centelleadores (experimento AGASA [4]) o tanques de agua Cherenkov (experimento Haverah Park [5]), en los cuales las partículas de la cascada depositan su energía, realizando así un muestreo discreto del frente de la cascada. Por otro lado, otra técnica consiste en la detección mediante telescopios de la luz de fluorescencia que emiten las moléculas de Nitrógeno de la atmósfera tras haber sido excitadas por las partículas de la cascada. Esta técnica ha sido utilizada por experimentos como Fly's Eye [2] y HiRes [3].

El Observatorio Pierre Auger [6] constituye un gran paso adelante. Fue inaugurado en el año 2005 y su construcción finalizó en 2008. Se trata del primer observatorio que aúna las dos técnicas, de modo que puede detectar las EAS con los tanques de agua y con los telescopios simultáneamente. Estos eventos se denominan *híbridos*. Esto permite determinar con mucha mayor precisión las propiedades de las cascadas, reducir significativamente los errores sistemáticos y aprovechar las ventajas que cada técnica tiene por separado.

Previamente a los resultados publicados por Auger, existía una gran controversia ya que los dos principales experimentos anteriores, AGASA y HiRes, habían mostrado resultados que daban lugar a interpretaciones muy diferentes sobre el origen, energía y composición de los rayos cósmicos de ultra-alta energía. Principalmente existían dos discrepancias. Por un lado, AGASA había observado que el flujo de rayos cósmicos se extendía más allá de 10^{20} eV [56] mientras que HiRes medía una supresión en el flujo a esas energías [30]. Este resultado tiene importantes implicaciones en astrofísica, cosmología y física de partículas. De hecho, el resultado de AGASA dió lugar a múltiples modelos teóricos, más o menos exóticos, para explicar el origen de rayos cósmicos de tan altas energías, ya que con los mecanismos clásicos no se puede entender fácilmente su origen ni cómo se acelaran a tan altas energías en los objetos astrofísicos. Por otro lado, la existencia de la supresión en el flujo es de esperar según los modelos de propagación de rayos cósmicos que suponen que éstos se originan en fuentes astrofísicas conocidas. Por otro lado, AGASA había detectado un exceso de rayos cósmicos en la dirección del Centro Galáctico, que fue confirmado por SUGAR [88]. Sin embargo, no fue encontrado por HiRes. Además, como ambos experimentos utilizaban técnicas diferentes, la comparación entre ambos es muy compleja.

Desde el comienzo, el Observatorio Pierre Auger ha proporcionado importantes resultados que han arrojado luz sobre estas discrepancias. En el 30^{th} International Cosmic Ray Conference en 2007, Auger confirmó la existencia de la supresión con 6 desviaciones estándar de significación [29]. Y poco antes descartó la existencia de un exceso de rayos cósmicos en la dirección del Centro Galáctico [178].

Auger ha publicado, tal y como se muestra en la Sección 3.5, en estos primeros años de funcionamiento otros importantes resultados. En el futuro se espera que el Observatorio siga resolviendo algunas de las incógnitas que permanecen abiertas. Para ello, se espera también construir un observatorio similar en el Hemisferio Norte [100], especialmente para llegar a energías aún más altas y tener una cobertura completa del cielo. Otra opción, que está en estudio, es el proyecto JEM-EUSO, que pretende colocar un telescopio en la Estación Espacial Internacional [7].

En el Capítulo 2 se presenta un resumen de los aspectos teóricos relacionados con los rayos cósmicos de ultra-alta energía, como los diferentes modelos teóricos para explicar su origen, las posibles fuentes y mecanismos de aceleración, las interacciones y el efecto de los campos magnéticos durante su propagación, su energía y su composición. Además, se explica brevemente la fenomenología de las cascadas extensas de partículas y sus propiedades fundamentales, así como las principales técnicas de detección. El Capítulo 3 se dedica al Observatorio Pierre Auger, con especial atención a los principales parámetros útiles en estudios de composición, que es uno de los temas de esta tesis, y a los principales resultados que ha publicado hasta hoy. Las siguientes secciones de este resumen se centran en los resultados aportados en esta tesis doctoral, que han sido presentados en los Capítulos 4 y 5. Finalmente, se muestran las conclusiones de este trabajo.

C.2 Distancia óptima para la estimación de la energía en los experimentos de superficie

Motivación

En los experimentos de superficie, tanto en los que utilizan centelleadores como tanques de agua, para determinar la energía del primario, se realiza el procedimiento que se detalla a continuación.

En primer lugar, se ajusta la distribución lateral de partículas, es decir, como varía la señal recogida en los detectores con la distancia al eje de la cascada. Para ello, se utiliza la llamada función de distribución lateral (LDF, Lateral Distribution Function), que puede tener diferentes formas funcionales. Este ajuste sufre de incertidumbres relacionadas con las fluctuaciones intrínsecas de la cascada de partículas, con la posición del punto de impacto en el suelo y con la forma funcional de la LDF elegida. La constante de normalización de la LDF es función creciente de la energía, por lo que Hillas [82] propuso utilizar una distancia característica r_0, de modo que la señal interpolada de la LDF en ese punto, $S(r_0)$, se caracteriza porque las incertidumbres mencionadas anteriormente se minimizan. Por tanto, $S(r_0)$ es un buen estimador de la energía del primario. Esta distancia característica depende esencialmente de la geometría de la red de detectores y de la distancia entre ellos, por lo que, hasta hoy, en todos los experimentos se ha utilizado un valor fijo para todos los eventos independientemente de su energía y dirección. Así, AGASA [183], Yakutsk [185] y Haverah Park [186] utilizan $r_0 = 600$ m, mientras que Auger, debido al mayor espaciamiento entre los detectores, elige 1000 m [187]. En segundo lugar, una vez determinada $S(r_0)$, en los experimentos de superficie se relaciona éste con la energía mediante simulaciones Monte Carlo (MC). En el caso de Auger, al ser un detector híbrido, en lugar de utilizar los MC, se obtiene una calibración a partir de los eventos híbridos [163].

En esta tesis, motivados por la idea original de Hillas, se propone un método para calcular la *distancia óptima* de cada lluvia individualmente, r_{opt}, a la cual la señal inferida, $S(r_{opt})$, es el mejor estimador de la energía del primario. Para hallar esa distancia óptima, se propone realizar varios ajustes de la LDF con diferentes posiciones del punto de impacto, el cual se fluctúa de acuerdo a las incertidumbres experimentales en la determinación de su posición. En cada ajuste, se dejan libre la normalización y la pendiente de la LDF y se fija el punto de impacto en el suelo. La distancia óptima se define como la distancia a la cual la dispersión entre las señales interpoladas de todos los ajustes es mínima. En la Fig. 4.1 se muestran varios ejemplos.

En este trabajo se dedica especial atención a aquellos eventos con detectores saturados. El problema de los eventos con saturación es muy común en los experimentos de superficie. Cuando el punto de impacto de la cascada en el suelo es muy cercano a un detector, el número de partículas que atraviesan éste es tan grande, que la luz generada en su interior satura la electrónica y la señal queda incompleta. Algunos experimentos descartan directamente estos detectores mientras que otros usan la señal recogida en ellos como un límite inferior en el ajuste de la LDF (así hace Auger actualmente). Actualmente, Auger está desarrollando métodos para recuperar la señal de estos detectores [106].

Dependencia de r_{opt} con la energía, el ángulo cenital y el espaciamiento de los detectores

Se estudia como varía la distancia óptima con la energía y ángulo cenital del primario y con la geometría y espaciamiento de la red de detectores. Este estudio se realiza para redes triangulares de detectores Cherenkov separados 433, 750, 866 y 1500 metros, obteniéndose idénticos resultados, exceptuando que cuánto mayor es el espaciamiento, mayor es el valor de r_{opt}. La dependencia con la energía es aproximadamente lineal (Fig. 4.2) y, respecto al ángulo cenital, es despreciable para lluvias cuasi-verticales ($\theta \leq 30°$) y disminuye para ángulos mayores (Fig. 4.5). Este efecto es consecuencia de que al aumentar la inclinación de la cascada, la red de detectores se ve más pequeña desde el punto de vista del plano de la lluvia, por lo que r_{opt} decrece, tal y como es de esperar, según lo comentado anteriormente. Los mismos resultados se obtienen para eventos con detectores saturados y los que no los tienen (Fig. 4.3 y Fig. 4.6).

Por tanto, se comprueba que estas dependencias no son despreciables, y además, la dispersión respecto a los valores medios es muy grande, de ahí la importancia de utilizar un valor de la distancia óptima calculado individualmente para cada lluvia y que además tenga en cuenta las fluctuaciones intrínsecas a ésta, en lugar de una distancia fija para estimar la energía.

Además, se demuestra que no es posible definir una única distancia característica r_0 para los eventos con saturación y aquellos que no tienen detectores saturados, debido a que existe una diferencia sistemática entre la distancia óptima entre unos y otros, siendo mayor para los primeros. Sin embargo, el hecho de utilizar la distancia óptima para cada lluvia individualmente r_{opt}, permite utilizar el mismo método para ambos tipos de eventos, lo que representa una ventaja muy importante de este método sin necesidad de recurrir a algoritmos para recuperar la señal de los detectores saturados, los cuales pueden introducir nuevas incertidumbres.

Efecto sobre el espectro de energía

A continuación, se demuestra que la distancia óptima calculada con nuestro algoritmo proporciona un mejor estimador de la energía que considerar una distancia característica fija. Para ello, se elige un espectro de entrada y se reconstruye con ambos métodos, usando la señal inferida del ajuste de la LDF a una distancia fija $S(r_0)$ y a la distancia óptima $S(r_{opt})$. En esta parte, se sigue el procedimiento desarrollado por AGASA y se elige un red de centelleadores separados 1 km y $r_0 = 600$ m. Se comprueba que el error relativo en la determinación de la energía es despreciable utilizando $S(r_{opt})$ en todo el intervalo de estudio (desde 10^{18} hasta $10^{20.5}$ eV), mientras que es muy significativo usando $S(r_0)$, especialmente, en el caso de los eventos con saturación (ver Fig. 4.8). Este hecho es de gran importancia porque los eventos con saturación representan una fracción cada vez mayor del total a medida que crece la energía (Fig. 4.12).

Otra ventaja de utilizar $S(r_{opt})$, es que las funciones de distribución de los errores en la energía reconstruida son mejor comportadas que utilizando $S(r_0)$. De hecho, las funciones de distribución son más simétricas y con colas más cortas en el caso de utilizar la distancia óptima (ver Fig. 4.7). Para $S(r_{opt})$, estas funciones son prácticamente Gaussianas, por lo que son más fáciles de manejar a la hora de aplicar posibles técnicas de deconvolución de los errores en el espectro y asegura que no se produzcan efectos de borde o errores sistemáticos en las regiones del espectro donde el índice espectral pueda cambiar rápidamente.

Por último, para comprobar el efecto de ambas técnicas de reconstrucción en una situación realista, se ha generado un espectro continuo que posee un tobillo (generado con primeras derivadas discontinuas), la supresión GZK y un corte a bajas energías que refleja el umbral inferior de detección de la red de detectores. Para ello, se ha usado la función tangente hiperbólica. Este espectro se ha convoluido analíticamente con las funciones de distribución de los errores en la energía reconstruida mencionadas anteriormente. Para ello, se han ajustado estas distribuciones mediante una función denominada Gaussiana Asimétrica Generalizada. Se obtiene que las largas colas de las funciones de distribución de los errores utilizando $S(r_0)$, distorsionan significativamente el espectro. Por ejemplo, la forma del tobillo se ensancha y su posición se desplaza, mientras que el GZK es también desplazado y más pronunciado. Por otro lado, utilizando $S(r_{opt})$ el espectro reconstruido se ajusta perfectamente al original, exceptuando a las energías más bajas cerca del corte en eficiencia donde sufre de efectos de borde al igual que ocurre con $S(r_0)$ (Fig. 4.14).

Todo ello demuestra los beneficios que esta nueva técnica introduce en la determinación de la energía del rayo cósmico primario en los experimentos de superficie. En el Apéndice A se ha aplicado también al futuro observatorio Pierre Auger Norte, el cual posee una geometría cuadrada y donde se ha utilizado la calibración híbrida de la energía tal y como se determina en Auger Sur. Se ha comprobado que los resultados son compatibles con los obtenidos anteriormente, lo que muestra la viabilidad de la técnica en el caso de detectores Cherenkov ubicados con una geometría y utilizando una calibración en energía diferentes. La aplicación a Auger Sur requiere encontrar la fórmula de calibración entre $S(r_{opt})$ y la energía medida por los telescopios de fluorescencia E_{FD}. Ya se han dado los primeros pasos en esta dirección, tal y como se muestra en el Capítulo 6.

C.3 Un nuevo parámetro para estudios de composición en experimentos de superficie

Motivación

Desde el punto de vista experimental, determinar la composición de los rayos cósmicos de ultra-alta energía es una cuestión muy compleja. Los problemas principales para identificar el primario independientemente del parámetro elegido son dos: las grandes fluctuaciones evento a evento y las incertidumbres en los modelos hadrónicos de interacción a tan altas energías. Las fluctuaciones intrínsecas en la evolución de la cascada se trasladan a cualquier parámetro que se quiera utilizar como discriminador de la masa del primario. En muchos casos, estas fluctuaciones, dado un cierto primario y una cierta energía, pueden ser mayores que las diferencias medias entre primarios de masas tan diferentes como protón y hierro. Al estudiar la composición de los rayos cósmicos, es necesario compararla con simulaciones realizadas para diferentes núcleos. Estas simulaciones utilizan diversos modelos para las interacciones hadrónicas, los cuales han sido afinados utilizando medidas que proceden principalmente de colisionadores de partículas. Sin embargo, los rayos cósmicos de ultra-alta energía están varios órdenes de magnitud por encima, por lo que parámetros como las secciones eficaces, multiplicidades e inelasticidades han de ser extrapolados, lo que introduce incertidumbres muy significativas. De hecho, se ha demostrado que los modelos actuales no son adecuados en varios aspectos, por ejemplo, en todos ellos, al compararlos con los datos reales, faltan muones [149, 150]. Además, variaciones en parámetros como las secciones eficaces o las inelasticidades pueden ser mal interpretados como cambios en la composición. Las dificultades son evidentes si se comparan los resultados publicados por los diferentes experimentos (ver Fig. 3.11), se puede observar que existe una clara discrepancia. A esto se suma otra dificultad, todos los parámetros sensibles a composición dependen de algún modo de la energía del primario. Por ello, las incertidumbres en la determinación de la energía disminuyen la capacidad de discriminación de los parámetros.

Como ya se ha mencionado, existen principalmente dos técnicas de detección de las cascadas de partículas: los telescopios de fluorescencia y la red de detectores de superficie, cada una de las cuales nos proporciona diferentes indicadores de composición. Actualmente, el parámetro más fiable es el punto en el cual el desarrollo longitudinal de la cascada alcanza el máximo, que se designa como X_{max}, y éste se mide con los telescopios de fluorescencia con una precisión de unos 20 g/cm^2 [116]. Diferentes primarios tienen diferentes secciones eficaces de interacción con los núcleos de la atmósfera, lo que se traduce en diferentes profundidades en el máximo del desarrollo de la cascada electromagnética y también en la dispersión de esta posición, ΔX_{max}. Por ejemplo, si se comparan hierro y protón, este último posee una sección eficaz menor lo que implica X_{max} y ΔX_{max} mayores. Desafortunadamente, cambios no esperados en las secciones eficaces como función de la energía pueden afectar a estos parámetros del mismo modo que los cambios reales en composición.

Por otro lado, los detectores de superficie realizan un muestreo del frente de la lluvia, el cual está dominado por dos componentes, la electromagnética (electrones, positrones y fotones) y la muónica. Ambas componentes se propagan de diferente modo. La parte muónica lo hace de forma radial desde el punto en el que se generaron en la atmósfera por lo que llega primero al suelo, mientras que la parte electromagnética lo hace de forma difusiva y llega al suelo más tarde durante un tiempo más largo. Además, la componente muónica tiene un radio de curvatura mejor definido, que es mayor que el de la otra componente. Por todo ello, la información referente a la abundancia relativa de las dos componentes está distribuida a lo largo de la distribución lateral de partículas, el radio de curvatura y la distribución temporal de un modo no trivial. Varios parámetros se han propuesto para extraer información sobre el primario, como la pendiente de la LDF, el radio de curvatura del frente de la cascada, diversos parámetros temporales como el *fall time* y el *rise time*, o las asimetrías azimutales en este último. Una revisión completa de estos parámetros se puede observar en la Sección 3.4.

En general, los parámetros de fluorescencia se consideran menos sensibles a los errores sistemáticos y son más fáciles de medir e interpretar que los de superficie. Sin embargo, tienen el gran inconveniente de la baja estadística, ya que los telescopios de fluorescencia sólo operan las noches claras y sin luna, lo que limita su tiempo de funcionamiento a un 10-13% comparado con los detectores de superficie que funcionan sin interrupción. Por ello, sigue siendo de gran interés encontrar nuevos parámetros de superficie.

En esta línea, esta tesis propone un nuevo parámetro que se define como

$$S_b = \sum_{i=1}^{N} \left[S_i \times \left(\frac{r_i}{r_0} \right)^b \right] \qquad \text{[VEM]} \tag{C.1}$$

donde la suma se extiende sobre todos los detectores activados N, $r_0 = 1000$ m es una distancia de referencia, S_i es la señal de cada detector en VEM y r_i es la distancia del detector al eje de la cascada. En este trabajo se han supuesto como detectores tanques de agua y una red triangular con los detectores separados 1.5 km como en Auger.

Estudio analítico. Optimización y propiedades de S_b

En la primera parte de este trabajo, se ha realizado un estudio analítico del parámetro. Para ello se han calculado el valor medio y la varianza de S_b para dos tipos de primario, hierro y protón, asumiendo que las fluctuaciones en la señal en los tanques Cherenkov es de tipo Gaussiano. Con ello se puede calcular el conocido como *factor de mérito* (ver Eq. 5.3), que representa una medida de la capacidad del parámetro para distinguir entre dos poblaciones, en este caso entre primarios de hierro y protón. En la Fig. 5.3 se representa como varía el factor de mérito con el exponente b (Eq. C.1), y se comprueba que existe un máximo muy destacado para $b \cong 3$. Por tanto, en el caso de un experimento como Auger, la mejor definición del parámetro es con $b = 3$.

Se estudian además dos propiedades importantes de este parámetro. Por un lado, como afecta la diferencia entre la pendiente de la LDF de hierro y protón a su poder de discriminación, y por otro, qué ocurre si se cambia la componente muónica de la cascada. En el primer caso, se ha modificado la pendiente de la LDF de ambos primarios y se ha comprobado que S_b es sensible a esta diferencia, y que a medida que ésta aumenta, el factor de mérito de S_b crece (Fig. 5.4). En el segundo caso, se ha modificado artificialmente el peso de la componente muónica y se ha visto que, si esta componente fuera mayor en las simulaciones, tal y como sugiere la evidencia experimental [149, 150], el factor de mérito crecería (Figs. 5.5y 5.6). Además, en ambas situaciones, bien al modificar la pendiente de las LDFs o bien el peso de la componente muónica, el factor de mérito es máximo para $b \cong 3$, esto muestra la estabilidad del parámetro.

Estudio numérico. Optimización, influencia de las estacines más lejanas y dependencias de S_b con la energía y el ángulo cenital

En la segunda parte del trabajo, se lleva a cabo un estudio numérico del parámetro, para lo cual se realizan simulaciones detalladas de las cascadas en la atmósfera (con el programa AIRES [194]), de la respuesta de los tanques de agua y una completa reconstrucción de estos eventos (para ello se ha usado el programa *Offline* oficial de la colaboración Auger

[196]). Con ello, se tienen en cuenta las incertidumbres experimentales, los errores en la reconstrucción y las fluctuaciones intrínsecas lluvia a lluvia. Además, se han utilizado los dos principales modelos hadrónicos de interacción, QGSJetII-03 [144] y Sibyll2.1 [145], y dos primarios, hierro y protón. El número de simulaciones es equivalente a la estadística de Auger en el espectro presentado en el ICRC de 2007 [29].

Se ha estudiado el rango de energías de 10^{19} a $10^{19.6}$ eV y se ha considerado una incertidumbre gaussiana en la determinación de la energía del 18%, que se corresponde con la precisión que alcanza Auger para los eventos de superficie [29]. En cuanto al ángulo cenital se consideran 3 intervalos de 10° centrados en $\theta = 30$, 45 y 55 grados, respectivamente.

Inicialmente se ha comprobado que el máximo poder separador de S_b corresponde al caso $b \cong 3$ con ambos modelos hadrónicos, de acuerdo con el resultado analítico (Fig. 5.7). Para comprobar que las fluctuaciones en la señal de las estaciones más lejanas (que es donde son más significativas) no afectan a la capacidad del parámetro de discriminar la masa del primario, se han impuesto varios cortes de calidad que esencialmente limitan las estaciones que se incluyen en el sumatorio de S_b. Se ha obtenido que estas estaciones lejanas no afectan (Fig. 5.9) y por tanto, no es necesario eliminarlas ni pesarlas de un modo diferente en el sumatorio. Esto supone una ventaja, ya que el usar cortes de calidad siempre puede introducir errores sistemáticos que hay que estudiar detenidamente.

Respecto a la relación de S_3 con la energía y el ángulo cenital, se ha demostrado que no existe dependencia con θ entre 0 y 60 grados (Fig. 5.10) y que con la energía es aproximadamente lineal (Fig. 5.11). Esto último representa una desventaja frente a X_{max} que depende logarítmicamente de la energía y por tanto, es menos sensible a los errores sistemáticos en su determinación. Esta desventaja es una característica habitual de los parámetros de superficie. A pesar de ello, se ha demostrado que la relación de S_3 con S_{38}, el estimador de la energía que se utiliza en el Observatorio Auger, es también aproximadamente lineal (Fig. 5.12), y que la separación entre hierro y protón en función de S_{38} se mantiene, por lo que los posibles errores sistemáticos procedentes de la calibración S_{38}-Energía no disminuyen significativamente el poder discriminador del parámetro.

Comparación entre S_3, el rise time a 1000 m y X_{max}

Por último, se ha comparado que, en una situación realista, la fiabilidad en la composición inferida del primario usando como parámetro S_3 con otros dos parámetros, X_{max} que es el principal parámetro de fluorescencia y el rise time medido a 1000 metros del punto de impacto, $t_{1/2}(1000)$, que es el parámetro de superficie más utilizado (en el cálculo del rise time se ha tenido en cuenta la corrección por asimétrias, los cortes habituales en las estaciones consideradas y se ha corregido su dependencia con θ para aumentar la estadística disponible).

En primer lugar, dentro del intervalo de energía entre 10^{19} y $10^{19.1}$ eV que es el de mayor estadística disponible, se han escogido muestras mezcla de hierro y protón con una fracción de protón (C_p^{true}) conocida, que varía de 0 a 1 en intervalos de 0.1. Utilizando los tres parámetros, se determina la fracción de protón inferida (C_p^{inf}) mediante el métdo de máxima verosimilitud utilizando muestras aleatorias. Se obtiene que el error en la determinación de C_p^{inf} es menor usando S_3 para cualquier valor de la C_p^{true} elegida (Ver Fig. 5.14 y 5.15). En este cálculo se ha tenido en cuenta que la estadística disponible para X_{max}, dado un cierto tiempo de exposición, es aproximadamente un 10% de los otros dos parámetros de superficie. Para destacar la importancia de este hecho, en las mismas figuras se muestra el resultado en el caso de que se utilizara el 100% de estadística para X_{max}, obteniéndose que en ese caso, el error en la C_p^{inf} es menor para este parámetro. Sin embargo, en la realidad esto requeriría un tiempo 10 veces mayor de exposición.

Para ampliar este estudio a todo el intervalo de energías, desde 10^{19} a $10^{19.6}$ eV, y para tener la estadística necesaria para aplicar el método de máxima verosimilitud, se ajustan las funciones de distribución de cada parámetro para los dos primarios, ambos modelos hadrónicos y para cada intervalo de energía. Para estos ajustes se ha utilizado la función llamada Gaussiana Asimétrica Generalizada (alguno de estos ajustes se muestra en la Fig. 5.16). Con estas funciones de distribución conocidas, es posible tener la estadística suficiente a la hora de extender el análisis a todo el rango de energías. Se ha elegido que el número de eventos utilizado en el cálculo sea el correspondiente a la cantidad que espera detectar Auger en 1 y 5 años de funcionamiento estando completo, y de nuevo, se supone que los telescopios de fluorescencia operan el 10% del tiempo. El resultado se muestra en la Fig. 5.17. El error en el valor inferido de la fracción de protón es mucho menor con S_3 comparado con X_{max} y con $t_{1/2}(1000)$ en todo el intervalo de energía para los dos modelos hadrónicos utilizados.

Todo ello muestra, en una situación lo más realista posible, la fiabilidad y la capacidad de discriminación de la composición del primario del parámetro propuesto, comprobándose que es mejor que otros parámetros de superficie y que, gracias a la estadística disponible con el detector de superficie, es mejor que X_{max}, el parámetro más fiable y utilizado hoy en día.

C.4 Conclusiones

Esta tesis se centra en los experimentos de detección de rayos cósmicos mediante una red de detectores de superficie, y se ocupa de dos cuestiones: i) la técnica utilizada para inferir la energía del rayo cósmico primario, y ii) cómo se puede determinar la composición química de los rayos cósmicos de ultra-alta energía (UHECRs, por sus siglas en inglés). Ambas cuestiones están intrínsecamente unidas. Desde el punto de vista teórico, este vínculo se debe, principalmente, a que en los mecanismos de aceleración (como el de Fermi, Sección 2.2.1) la ganancia en energía es directamente proporcional a la carga de la partícula. Experimentalmente, los parámetros sensibles a la composición del primario dependen de la energía de éste, por lo que los errores sistemáticos y las incertidumbres estadísticas en la determinación de la energía afectan al poder de discriminación de los observables utilizados en estudios de composición.

En los últimos años, el Observatorio Auger ha producido grandes avances en la determinación de la energía de los UHECRs. Por ejemplo, la controversia existente entre experimentos anteriores, como AGASA y HiRes, sobre la existencia o no de la supresión GZK, ha sido resuelta en 2007 (en el 30^{th} International Cosmic Ray Conference en Mérida, México) al confirmar Auger su existencia con una significancia de 6 de desviaciones estándar [29]. Además, la posición del tobillo del espectro se confirmó, en acuerdo con HiRes, en $10^{18.6}$ eV [166], y en contra de AGASA que previamente lo encontró en 10^{19} eV [56]. Sin embargo, varias preguntas siguen abiertas: ¿Existe la segunda rodilla del espectro? ¿Cuál es la energía de transición de los rayos cósmicos galácticos a extragalácticos? ¿Cuál es la razón de la diferencia en el flujo de UHECRs medido por Auger, HiRes y AGASA (un factor de 2 en el flujo, 30% en energía o una combinación de ambos)? ¿Es la supresión GZK consecuencia de que se ha alcanzado el límite máximo de aceleración en los objetos astrofísicos, o es debido a la interacción de los rayos cósmicos con los fotones de la radiación de fondo microondas (CMB)? ¿Cómo de pronunciada es la caída tras la supresión? ¿Cuál es la forma exacta del espectro a las más altas energías y es ésta consistente con una composición predominante de protón o con una composición mixta?

En esta tesis, se propone un nuevo método para la determinación de la energía del primario, lo cual puede ayudar a mejorar la precisión en el espectro medido y con ello, a resolver las cuestiones anteriores. El método y las conclusiones alcanzadas se exponen a continuación:

- En los experimentos de superficie, la energía del primario se determina a partir de la señal interpolada del ajuste de la función de distribución lateral a una *distancia característica* fija del eje de la cascada (r_0). Esta distancia se determina teniendo en cuenta sólo la geometría de la red y la separación entre los detectores. Por ejemplo, el experimento AGASA utilizaba $r_0 = 600$ m, mientras que Auger elige 1000 m debido al mayor espaciamiento de su red. Sin embargo, en este trabajo se ha demostrado que existe una *distancia óptima* para la determinación de la energía, específica de cada lluvia, r_{opt}, que no sólo depende de la geometría de la red, sino también de la energía y dirección del primario. Estas dependencias no son despreciables y, por tanto, una distancia óptima calculada lluvia a lluvia es más adecuada como estimador de la energía que usar una distancia fija.

- Se ha desarrollado un método para encontrar la distancia óptima de cada cascada individual. Esencialmente, la distribución lateral de partículas se ajusta utilizando una cierta forma funcional. En el ajuste, el punto de impacto en el suelo de la cascada es fijo, mientras que la normalización de la función y su pendiente se dejan como parámetros libres. Se realizan 50 ajustes modificando la posición del punto de impacto en torno al punto estimado inicialmente teniendo en cuenta las incertidumbres experimentales en la determinación de esta posición. La distancia óptima se define como la distancia a la cual las fluctuaciones en la señal interpolada de esos ajustes se minimiza.

- Si se utiliza la señal inferida a la distancia óptima ($S(r_{opt})$) como estimador de la energía, se reduce significativamente el error sistemático introducido en la determinación de la energía en comparación con utilizar la señal a la distancia característica ($S(r_0)$). Utilizando $S(r_{opt})$, no se provoca un error sistemático en todo el intervalo de energía, ni para eventos con detectores saturados ni para los que no. Por el contrario, utilizando $S(r_0)$ se introduce un error muy significativo debido a la diferencia entre el valor elegido de r_0 y el valor real de la distancia óptima de cada cascada.
- Utilizando $S(r_{opt})$ como estimador de la energía, se mejora el comportamiento de las distribuciones de los errores en la energía inferida. De hecho, éstas son más compactas, menos asimétricas y con colas más cortas que las obtenidas usando $S(r_0)$. Como consecuencia, se ha mostrado que un espectro realista está mucho mejor determinado utilizando $S(r_{opt})$. Mientras que el uso de $S(r_0)$ puede modificar la posición y forma del tobillo y desplazar la posción del umbral GZK, usando $S(r_{opt})$ el resultado se ajusta con gran precisión al espectro original.
- Una gran ventaja de esta nueva técnica consiste en que permite reconstruir todos los eventos con un mismo método, independientemente de que tengan estaciones saturadas o no. Con $S(r_0)$ esto no es posible en general, sino que se requieren ciertos cortes de selección adicionales o el desarrollo de algoritmos para recuperar la señal del detector saturado. A medida que crece la energía esta ventaja es mayor, ya que la fracción de eventos con saturación crece rápidamente al aumentar ésta.

- Estas conclusiones son válidas para diferentes geometrías de la red y para diversos tipos de detectores. De hecho, el método ha sido aplicado a una red triangular de centelleadores, como el experimento AGASA (Capítulo 4), y también a una red cuadrada de tanques de agua Cherenkov, como el diseño del futuro Observatorio Pierre Auger Norte (Apéndice A).

- La aplicación de éste método a un experimento híbrido como Auger no es directa. Es necesario determinar la curva de calibración que relaciona la señal a la distancia óptima, $S(r_{opt})$, con la energía medida por los telescopios de fluorescencia, E_{FD}. Para ello, es necesario utilizar valores realistas en la indeterminación del punto de impacto y diseñar los cortes de calidad apropiados para la selección de los eventos híbridos que se deben utilizar. Tal y como se comenta el Capítulo 6, se han dado los primeros pasos en esta dirección. Las principales fuentes de error en la energía inferida, en un experimento híbrido como Auger, provienen de la técnica de fluorescencia (calibración, reconstrucción, incertidumbre en el llamado *fluorescence yield*).Por lo tanto, la mejora utlizando la nueva técnica no sería tan importante como en un experimento puro de superficie, aunque sí se conseguirían algunos beneficios. Primero, no sería necesario, en principio, utilizar el método del CIC (Constant Intensity Cut), ya que en la determinación de r_{opt} ya se tiene en cuenta la dirección de llegada del rayo cósmico. Segundo, las funciones de distribución de los errores son más Gaussianas, lo que hace más creible la aplicación de técnicas de deconvolución en el espectro medido.

Este método puede ayudar a determinar con mayor precisión el espectro de energía de los rayos cósmicos ultra energéticos. Por ejemplo, para poder distinguir entre las diferentes interpretaciones sobre el origen de la supresión, dos características del espectro deben ser medidas. Primero, debe existir un aumento en el flujo en torno al umbral GZK si existen partículas a energías mayores. Estas partículas al interaccionar con los fotones del CMB degradan su energía hasta ese umbral, debajo del cual la sección eficaz de esta interacción es mucho menor (ver Fig. 2.6), lo que provoca ese aumento. En segundo lugar, se espera que exista una recuperación en el espectro a energías aún mayores ($\sim 10^{20.5}$ eV) en el caso de que la supresión no sea consecuencia de un límite máximo en los mecanismos de aceleración en los objetos astrofísicos. Estas energías tan altas no son alcanzables por parte de Auger Sur, aunque se espera que Auger Norte pueda proporcionar la estadística suficiente en ese rango tras varios años de funcionamiento. En conclusión, más estadística y más precisión en la determinación de la energía son necesarios en los experimentos actuales para distinguir ese aumento e interpretar correctamente la supresión GZK.

Además, este nuevo método puede utilizarse para medir con mayor precisión la región del tobillo y también por parte de experimentos de superificie que estudian rangos menores de energía. KASCADE-Grande [55] y AMIGA [101] (una extensión de Auger para estudiar energías menores), se centran en el rango de 10^{17} - 10^{18} eV, en el cual métodos más precisos podrían ayudar a descartar o confirmar la existencia de la segunda rodilla del espectro. Para finalizar, esta técnica puede ser aplicada a AGASA y Auger para comprender el origen de la discrepancia el en flujo medido de rayos cósmicos ultra energéticos por parte de ambos experimentos.

El estudio de la composición de los rayos cósmicos de ultra alta energía es un problema más complejo. De hecho, muy pocos han sido los métodos propuestos que han sido aplicados con éxito. Por ejemplo, Auger utiliza la posición del máximo del desarrollo longitudinal de la casacada X_{max}, y dos parámetros de superficie relacionados con el *rise time* (XAsymMax y el parámetro Delta) para determinar la evolución de la composición promedio como función de la energía [116, 131]. Sin embargo, estos resultados sólo se extienden hasta $10^{19.4}$ eV y existe cierta controversia entre estos resultados y la correlación publicada entre la dirección de llegada de los rayos cósmicos de más alta energía detectados por Auger y la posición de AGN cercanos [170]. La razón es que la composición de los rayos cósmicos a altas energías tiende a ser más pesada, mientras que la correlación publicada es válida siempre que los primarios sean elementos ligeros que no son muy desviados por los campos magnéticos. Se acepta que el mayor problema es la fata de estadística, lo cual podrá resolver Auger durante los próximos años de funcionamiento. Sin embargo, es claro que son necesarias nuevas técnicas y nuevos parámetros sensibles a la composición.

Los métodos acutales únicamente permiten determinar la composición media de los rayos cósmicos de ultra alta energía. Actualmente, la composición evento a evento es imposible, debido principalmente a las fluctuaciones intrínsecas de la lluvia y a las incertidumbres en los modelos hadrónicos de interacción y en la determinación de la energía (una discusión más completa puede verse en la Sección 3.4.3).

El conocimiento de la composición es además muy importante para limitar o descartar diversos modelos teóricos acerca del origen de los rayos cósmicos ultra energéticos. Así, los límites en la fracción de fotones publicados por Auger en los últimos años (Sección 3.5.2) desfavorecen varios de los modelos *top-down* como origen de los rayos cósmicos entre 10^{18} y $10^{19.5}$ eV. Además, la composición es la pieza clave para la distinción de los modelos que explican el origen del tobillo del espectro y por tanto, el punto de transición de los rayos cósmicos galácticos a extragalácticos (una discusión completa sobre estos modelos se puede ver en el Capítulo 2).

En esta tesis se propone un nuevo parámetro para estudios de composición, el cual utiliza exclusivamente la información que proviene de la red de detectores de superficie. A continuación se detallan algunos puntos importantes:

- Los parámetros de superficie son habitualmente menos creíbles que los de fluorescencia (principalmente X_{max} y sus fluctuaciones) debido a que son más sensibles a los errores sistemáticos y más difíciles de interpretar. Sin embargo, la gran estadística disponible por parte de los detectores de superficie comparada con los de fluorescencia, cuyo tiempo de funcionamiento es 10-13% respecto a los de superficie, hace que sea de gran interés el buscar nuevos y mejores parámetros de superficie.

- En este trabajo, se propone una nueva familia de parámetros de superficie, llamados S_b. En el Capítulo 5 se aplica al caso de tanques de agua Cherenkov como los de Auger Sur, pero puede ser también utilizado por los experimentos que usan centelleadores.

- Bajo estas condiciones (tanques Cherenkov formando una red de detectores triangular y separados 1.5 km), el poder de discriminación entre hierro y protón de S_b es máximo para $b \cong 3$.

- Se ha demostrado analíticamente que S_b es sensible a la diferencia entre la pendiente de la función de distribución lateral de hierro y protón. Además, las incertidumbres actuales en la componente muónica de los códigos de simulación de lluvias en la atmósfera, van en la dirección de mejorar la capacidad de discriminación del parámetro.

- La estabilidad del parámetro queda clara por el hecho de que el mejor valor del exponente b para maximizar la capacidad de separación de hierro y protón, que es siempre 3.0, independientemente de si la pendiente de la función de distribución lateral de hierro y protón o el peso de la componente muónica sean modificados artificialmente. Además, se ha mostrado que las estaciones muy alejadas al eje de la lluvia, cuya señal puede estar dominada por fluctuaciones, no afectan al poder discriminador de S_b.

- Se ha comprobado la fiabilidad de S_b bajo condiciones realistas. Para ello, se ha realizado la simulación completa de todo el proceso, desde las lluvias en la atmósfera (usando AIRES), la respuesta del detector (basada en GEANT4) y la reconstrucción (usando Offline, el programa oficial de la colaboración Auger). Así, se tienen en cuenta condiciones reales y las incertidumbres propias del proceso de reconstrucción.

- S_3 depende de forma prácticamente lineal con S_{38°, el estimador de la energía utilizado en Auger, y su poder de discriminación como función de S_{38°, permanece estable. Por lo tanto, a pesar de las posibles incertidumbres en la fórmula de calibración S_{38°-energía, se puede tener un valioso resultado sobre la evolución de la composición como función de la energía usando S_3. Respecto a la dirección del rayos cósmico, no se ha encontrado dependencia de S_3 con el ángulo cenital.

- Se ha realizado una estudio detallado y realista, basado en un método de máxima verosimilitud, con el objeto de comparar la capacidad en la determinación de la composición del primario de tres parámetros: el rise time, que es el más utilizado de los de superficie; X_{max}, el más utilizado y creíble actualmente; y S_3. La estadística disponible para X_{max} se ha reducido al 10% debido al limitado tiempo de funcionamiento de los telescopios de fluorescencia. Se ha demostrado que el error en la composición inferida utilizando S_3 es significativamente menor que usando el rise time o X_{max}. Esto muestra la capacidad del nuevo parámetro y la importancia de la estadística disponible a la hora de hacer estudios de composición.

Para concluir, el objetivo de esta tesis ha sido desarrollar nuevos métodos para mejorar la precisión en la determinación de la energía y la composición de los rayos cósmicos de ultra alta energía, en aquellos experimentos que utilizan una red de detectores de superficie. La fiabilidad de los métodos propuestos ha sido demostrada bajo las condiciones más realistas posibles. Sin embargo, este trabajo debe continuar. Nuevas comprobaciones, optimizaciones y especialemente, la aplicación a los datos del Observatorio Auger, están ya en camino (véase el Capítulo 6).

Bibliography

[1] V.F. Hess. Uber Beobachtungen der durchdringenden Strahlung bei sieben Freiballon-fahrten. *Phys. Zeit*, 13:1084–1091, 1912.

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

[3] S. C. Corbato et al. HiRes: A High resolution Fly's Eye detector. *Nucl. Phys. Proc. Suppl.*, 28B:36–39, 1992.

[4] M. Nagano and M. Teshima. Present status of Akeno 100 km^2 air shower array (AGASA). *Nucl. Phys. Proc. Suppl.*, 28B:28–35, 1992.

[5] J. G. Wilson. Haverah park - the first ten years. Proc. Conference On Cosmic Rays (Budapest), vol. 3, 409-414, 1970.

[6] J. Abraham et al. Properties and performance of the prototype instrument for the Pierre Auger Observatory. *Nucl. Instrum. Meth.*, A523:50–95, 2004.

[7] T. Ebisuzaki et al. The JEM-EUSO project: Observing extremely high energy cosmic rays and neutrinos from the International Space Station. *Nucl. Phys. Proc. Suppl.*, 175-176:237–240, 2008.

[8] C. D. Anderson. Early work on the positron and muon. *Am. J. Phys.*, 29:825, 1961.

[9] P. A. M. Dirac. *Proc. Roy. Soc. Lond. A*, 133:60, 1931.

[10] D. Griffiths. *Introduction to Elementary Particles.* John Wiley and Sons, 1987.

[11] C. M. G. Lattes. Process involving charged mesons. *Nature*, 159:694, 1947.

[12] G. D. Rochester and C. C. Butler. Evidence for the existence of new unstable elementary particles. *Nature*, 160:855, 1947.

[13] P. Auger, P. Ehrenfest, R. Maze, J. Daudin, and A. Freon Robley. Extensive cosmic-ray showers. *Rev. Mod. Phys.*, 11:288–291, 1939.

[14] P. Bassi, G. Clark, and B. Rossi. Distribution of arrival times of air shower particles. *Phys. Rev.*, 92:441, 1953.

[15] J. Linsley. Evidence for a primary cosmic-ray particle with energy 10^{20} eV. *Phys. Rev. Lett*, 10:146, 1963.

[16] T. Hara, S. Kawaguchi, S. Mikamo, M. Nagano, K. Suga, G. Tanahashi, K. Uchino, and H. Akiyama. *Acta Phys. Acad. Sci. Hung.*, 29:361, 1970.

[17] H. E. Bergeson et al. Measurement of Light Emission from Remote Cosmic Ray Air Showers. *Phys. Rev. Lett.*, 39:847–849, 1977.

[18] Fred Lawrence Whipple Observatory.
http://www.cfa.harvard.edu/facilities/flwo/directions.html.

[19] J. A. Hinton for the HESS collaboration. The status of the HESS project. *New Astron. Rev.*, 48:331–337, 2004.

[20] Juan Cortina for the MAGIC collaboration. Status and First Results of the MAGIC Telescope. *Astrophys. Space Sci.*, 297:245–255, 2005.

[21] E. Parizot, A. Marcowith, J. Ballet, and Y. A. Gallant. Observational constraints on energetic particle diffusion in young SNRs: Amplified magnetic field and maximum energy. *astro-ph/0603723*, 2006.

[22] S. Lucek and A. Bell. *Mon. Not. R. Astron. Soc.*, 321:433–438, 2001.

[23] M. Berezinsky, V. Kachelrieb and A. Vilenkin. Ultrahigh energy cosmic rays without Greisen-Zatsepin-Kuzmin cut-off. *Phys. Rev. Lett.*, 79, 1997.

[24] P. Bhattacharjee. Origin and propagation of extremely high energy cosmic ray. *Phys. Rep.*, 327:109–247, 2000.

[25] T. J. Weiler. Cosmic-ray neutrino annihilation on relic neutrinos revisited: a mechanism for generating air showers above the Greisen-Zatsepin-Kuzmin cut-off. *Astroparticle Physics*, 11:303–316, 1999.

[26] L. Gonzalez-Mestres. Observing air showers from cosmic superluminal particles. *AIP Conf. Proc.*, 433, 1998.

[27] R. Aloisio, V. Berezinsky, and M. Kachelrieb. *Nuclear Physics B Proceedings Supplements*, 136:319–326, 2004.

[28] Pierre Auger Collaboration. Upper limit on the cosmic-ray photon flux above 10^{19} eV using the surface detector of the Pierre Auger Observatory. *Astroparticle Physics*, 29:243–256, 2008.

[29] The Pierre Auger Collaboration. *Proc. 30th International Cosmic Ray Conference, Mérida, Mexico*, 2007.

[30] HiRes Collaboration. *Proc. 30th International Cosmic Ray Conference, Mérida, Mexico*, 2007.

[31] P. Bhattacharjee and G. Sigl. Origin and propagation of extremely high energy cosmic rays. *Phys. Rept.*, 327:109–247, 2000.

[32] E. Fermi. *Phys. Rev.*, 75:1169, 1949.

[33] R. J. Protheroe. *Acceleration and interaction of ultra high energy cosmic rays. Topics in Cosmic-Ray astrophysics.* M. A. Duvernois, ed., 1999.

[34] A. M. Hillas. *Ann. Rev. Astron. Astrophys.*, 22:425, 1984.

[35] C. De Donato. A phenomenological approach to the study of the cosmic ray galactic-extragalactic transition. *PhD. Thesis*, Universitá degli studi di Milano, 2008.

[36] David Thomas for the Pierre Auger Collaboration. Search for coincidences with astrophysical transients in Pierre Auger Observatory data. *Proceedings of the 31st ICRC, Lodz*, 2009.

[37] R. J. Protheroe and A. P. Szabo. High energy cosmic rays from active galactic nuclei. *Phys. Rev. Lett.*, 69:2885–2888, 1992.

[38] K. Dolag, D. Grasso, V. Springel, and I. Tkachev. Constrained simulations of the magnetic field in the local universe and the propagation of ultrahigh energy cosmic rays. *Journal of Cosmology and Astroparticle Physics*, 1:9, 2005.

[39] J. W. Cronin. *Nucl. Phys. Proc. Suppl.*, 138:465, 2005.

[40] M. Lemoine and G. Sigl (Eds.). *Physics and Astrophysics of Ultra-High Energy Cosmic Rays.* Springer, 2001.

[41] E. Waxman and J. Miralda-Escude. Images of bursting sources of high-energy cosmic rays: Effects of magnetic fields. *Astrophys. J.*, 472, 1996.

[42] S. Lee, A. V. Olinto, and G. Sigl. Extragalactic magnetic field and the highest energy cosmic rays. *Astrophysical Journal*, 455, 1995.

[43] K. Greisen. End to the cosmic-ray spectrum? *Phys. Rev. Lett.*, 16:748, 1966.

[44] Z. T. Zatsepin and V. A. Kuzmin. Upper limit of the spectrum of cosmic rays. *JETP Lett.*, 4:78, 1966.

[45] A. A. Penzias and R. W. Wilson. *Astrophys. J.*, 142:419, 1965.

[46] F. A. Aharonian and J. W. Cronin. *Phys. Rev. D*, 50:1892, 1994.

[47] J. W. Cronin. The highest-energy cosmic rays. *Nucl. Phys. Proc. Suppl.*, 138:465, 2005.

[48] A. D. Erlykin and A. W. Wolfendale. *J. Phys. G*, 23:979, 1997.

[49] J. R. Horandel. *Mod. Phys. Lett. A*, 22:1533, 2007.

[50] A. D. Erlykin and A. W. Wolfendale. *Astropart. Phys.*, 18:151, 2002.

[51] J. Candia, E. Roulet, and L. N. Epele. *JHEP*, 12:033, 2002.

[52] J. R. Hoerandel. Models of the knee in the energy spectrum of cosmic rays. *Astroparticle Physics*, 21:241, 2004.

[53] D. J. Bird et al. *Physical Review Letters*, 71:3401–3404, 1993.

[54] M. Nagano et al. *Journal of Physics G: Nuclear Physics*, 18:423–442, 1992.

[55] G. Navarra et al. (KASCADE-Grande Collaboration). KASCADE-Grande: a large acceptance, high-resolution cosmic-ray detector up to 10^{18} eV. *Nucl. Instr. Meth. A*, 518:207–209, 2004.

[56] Takeda et al. Energy determination in the akeno giant air shower array experiment. *Astroparticle Physics*, 19:447–462, 2003.

[57] A.A. Ivanov et al. *Nucl. Phys. B*, 122:226–230, 2003.

[58] T. Abu-Zayyad et al. *Astrophys. J.*, 557:686, 2001.

[59] M. Ave et al. *Proc. 27th International Cosmic Ray Conference. Hamburg, Germany*, I:381, 2001.

[60] A. Hillas. *Conf. on Cosmology, Galaxy Formation and Astro-Particle Physics on the Pathway to the SKA, Oxford, England*, 2006.

[61] D. De Marco and T. Stanev. *Phys. Rev. D*, 72, 2005.

[62] V. Berezinsky. Transition from galactic to extragalactic cosmic rays. *Proc. 30th International Cosmic Ray Conference, Mérida, Mexico*, 2007.

[63] V. Berezinsky, S. Grigoreva, and B. Hnatyk. *Nucl. Phys. (Proc. Suppl.)*, 151:497, 2006.

[64] D. Allard, E. Parizot, and A. V. Olinto. *Astropart. Phys.*, 27:61, 2007.

[65] C. De Donato and G. A. Medina-Tanco. Experimental constraints on the astrophysical interpretation of the cosmic ray Galactic-extragalactic transition region. *arXiv:0807.4510*, 2008.

[66] V. S. Berezinsky and S. I. Grigor'eva. A Bump in the ultrahigh-energy cosmic ray spectrum. *Astron. Astrophys.*, 199:1–12, 1988.

[67] W. Heitler. Quantum theory of radiation. *Oxford University Press*, 1944.

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

[69] J. Knapp. *Corsika School, Lauterbad*, 2005.

[70] M. Nagano et al. New measurement on photon yields from air and the application to the energy estimation of primary cosmic rays. *Astroparticle Physics*, 22:235–248, 2004.

[71] M. Ave et al. (Airfly Collaboration). Measurement of the pressure dependence of air fluorescence emission induced by electrons. *Astroparticle Physics*, 28:41–57, 2007.

[72] F. Arqueros, J. Blanco, and J. Rosado. Improved model for the analysis of air fluorescence induced by electrons. *arXiv:0712.3536 [astro-ph]*, 2007.

[73] T. Gaisser and A. M. Hillas. *Proc. 15th International Cosmic Ray Conference, Plovdiv*, 8:353, 1977.

[74] Benjamin Rouillé d'Orfeuil for the Pierre Auger Collaboration. Atmospheric effects on extensive air showers observed with the array of surface detectors of the Pierre Auger Observatory. *Proceedings of the 31st ICRC, Lodz*, 2009.

[75] M. S. Longair. *High Energy Astrophysics.* Cambridge University Press, 2nd edition., 1992.

[76] H. M. J. Barbosa, F. Catalani, J. A. Chinellato, and C. Dobrigkeit. *Astropart. Phys.*, 22:159, 2004.

[77] A. M. Hillas. Estimation of the energies of the largest showers. *Astrophysical Aspects of the Most Energetic Cosmic Rays*, 1991.

[78] D. Heck. The influence of hadronic interaction models on simulated air-showers: A phenomenological comparison. *Talk presented at the CORSIKA School, Lauterbad, Germany*, 2005.

[79] K. Greisen. *Progress in Cosmic-Ray Physics, vol. 3.* Interscience Publishers, Inc. New York, 1956.

[80] K. Kamata and J. Nishimura. The lateral and the angular structure functions of electron showers. *Progr Theor. Phys. Suppl.*, 6:93–155, 1958.

[81] D. Barnhill et al. [Pierre Auger Collaboration]. Measurement of the lateral distribution function of UHECR air showers with the Auger Observatory. *Proc. 29th ICRC*, 7:291, 2005.

[82] A. M. Hillas. *Acta Phys. Acad. Sci. Hung.*, 29 Suppl. 3:355, 1970.

[83] W. D. Apel et al. *Astropart. Phys.*, 24:467, 2006.

[84] T. Antoni et al. *Astropart. Phys.*, 14:245, 2001.

[85] P. Auger et al. *Rev. Modern Phys.*, 11:288, 1939.

[86] D. Bergman et al. *arXiv:0704.3721v1 [astro-ph]*, 2007.

[87] H. Tokuno et al. The Telescope Array experiment: Status and prospects. *J. Phys. Conf. Ser.*, 120:062027, 2008.

[88] J. A. Bellido, R. W. Clay, B. R. Dawson, and M. Johnston-Hollitt. *Astropart. Phys.*, 10:303, 1999.

[89] T. Antoni et al. (KASCADE Collaboration). The cosmic-ray experiment KASCADE. *Nucl. Instr. Meth. A*, 513:429, 2003.

[90] R. M. Baltrusaits et al. *NIM A*, 240:410, 1985.

[91] J. V. Jelley et al. Radio pulses from extensive cosmic-ray air showers. *Nature*, 205:327, 1965.

[92] P. R. Barker, W. E. Hazen, and A. Z. Hendel. Radio pulses from cosmic-ray air showers. *Phys. Rev. Lett.*, 18:51–54, 1967.

[93] H. Falcke and P. Gorham. Detecting radio emission from cosmic ray air showers and neutrinos with a digital radio telescope. *Astroparticle Physics*, 19:477–494, 2003.

[94] T. Karg, J. Auffenberg, T. Gaisser, K. Helbing, and A. Karle. A radio air shower detector as an extension for IceCube and IceTop. *Proc. 30th ICRC, Mérida, Mexico*, 2007.

[95] A. M. van den Berg for the Pierre Auger Collaboration. Radio detection of high-energy cosmic rays at the Pierre Auger Observatory. *ArXiv:0708.1709*, 2007.

[96] A. M. van den Berg for the Pierre Auger Collaboration. Radio detection of high-energy cosmic rays at the Pierre Auger Observatory. *Proceedings of the 31st ICRC, Lodz*, 2009.

[97] G. A. Askaryan. *Sov. J. Atom Energy*, 2:921, 1957.

[98] R. Nahnhauer. Alternative detection methods for highest energy neutrinos. *Nuclear Physics B Proc. Suppl.*, 143:387–394, 2005.

[99] N. Hayashida et al. *Astropart. Phys.*, 10:303, 1999.

[100] D. Nitz for the Pierre Auger Collaboration. *arXiv:astro-ph/0706.3940*, 2007.

[101] A. Etchegoyen. AMIGA, Auger Muons and Infill for the Ground Array. *Proc. 30th ICRC. Mérida, Mexico. arXiv:astro-ph/0710.1646*, 2007.

[102] T. Suomijarvi for the Pierre Auger Observatory Collaboration. Surface detector electronics for the Auger Observatory. *Proc. 28th ICRC, Hamburg*, 2001.

[103] Maria Monasor. Energy calibration of the Pierre Auger Observatory. Measurement of the spectrum of Ultra-High Energy Cosmic Rays. *PhD. Thesis at Universidad de Alcalá. Internal GAP-Note 2008-167.*

[104] Nicolas G. Busca. The Ultra High Energy Cosmic Ray flux from the southern Pierre Auger Observatory data. *PhD. Thesis at the University of Chicago. Internal GAP-Note 2006-108.*

[105] A. Tamburro. Neutrino detection with the surface array of the Pierre Auger Observatory. *PhD. Thesis at Karlsruhe Universitat. Internal GAP-Note 2009-002.*

[106] Ioana C. Maris. Measurement of the Ultra High Energy Cosmic Ray flux using data of the Pierre Auger Observatory. *PhD. Thesis at Karlsruhe Universitat. Internal GAP-Note 2008-026.*

[107] H. O. Klages for the Pierre Auger Observatory Collaboration. HEAT: Enhancement telescopes for the Pierre Auger Southern Observatory. *30th ICRC, Mérida, Mexico*, 2007.

[108] H. Gemmeke for the Pierre Auger Observatory Collaboration. The Auger fluorescence detector electronics. *30th ICRC, Mérida, Mexico*, 2007.

[109] M. Ambrosio et al. A new method for the UHECR mass composition studies. *Nucl. Phys. Proc. Suppl.*, 136:301–305, 2004.

[110] Vitor de Souza, Gustavo Medina-Tanco, Jeferson A. Ortiz, and Federico Sánchez. The shower size parameter as estimator of extensive air shower energy in fluorescence telescopes. *Phys. Rev.*, D73:043001, 2006.

[111] J. Linsley. *Proc. 15th ICRC*, 12:89, 1977.

[112] B. R. Dawson for the Pierre Auger collaboration. Hybrid Performance of the Pierre Auger Observatory. *arXiv:0706.1105*, 2007.

[113] D.J. Bird et al. [Flys Eye Collaboration]. *Phys. Rev. Lett.*, 71:3401, 1993.

[114] T. Abu-Zayyad et al. [HiRes-MIA Collaboration]. *Astrophys. J.*, 557:686, 2001.

[115] R.U. Abbasi et al. [HiRes Collaboration]. *Astrophys. J.*, 622:910, 2005.

[116] J. A. Bellido for the Pierre Auger Collaboration. Measurement of the average depth of shower maximum and its fluctuations with the Pierre Auger Observatory. *Proceedings of the 31st ICRC, Lodz*, 2009.

[117] S.Riggi et al. Mass composition analysis from likelihood fit to hybrid X_{max} data. *Internal GAP-Note 2009-006.*

[118] T.K. Gaisser. *Cosmic Rays and Particle Physics.* Cambridge Univ. Press, 1990.

[119] A. Castellina and G. Navarra. The muon/em ratio and the SD energy scale. *Internal GAP-Note 2007-126.*

[120] P. Billoir and B. Gouteraux. Simulation of muon counting in FADC traces of Cherenkov tanks at different frequencies. *Internal GAP-Note 2006-058.*

[121] X. Garrido et al. Measurement of the number of muons in Auger tanks by the FADC jump counting method. *Internal GAP-Note 2007-060.*

[122] A. D. Supanitsky et al. Underground muon counters as a tool for composition analyses. *Astropart. Phys.*, pages 461–470, 2008.

[123] A. Chou. *Proc. 26th ICRC*, 7:319, 2005.

[124] The Pierre Auger Collaboration. Upper limit on the cosmic-ray photon flux above 10^{19} eV using the surface detector of the Pierre Auger Observatory. *Astroparticle Physics*, 29:243, 2008.

[125] M. D. Healy for the Pierre Auger Collaboration. *arXiv:0710.0025*, 2007.

[126] B. Smith. *PhD. Thesis at the University of Leeds. Internal GAP-Note 2008-161.*

[127] C. Wileman. *PhD. Thesis at the University of Leeds. Internal GAP-Note 2008-160.*

[128] A. A. Watson. *Nuclear Physics B,(Proc. Suppl.)*, 151:83, 2006.

[129] B. Smith, C. Wileman, and A. A. Watson. Can the rise time be used to infer the position of shower maximum? *Internal GAP-Note 2007-092.*

[130] B. Smith, C. Wileman, and A. A. Watson. The rise time as an indicator of the mass composition of UHECRs. *GAP-Note 2007-110.*

[131] H. Wahlberg for the Pierre Auger Collaboration. Study of the nuclear mass composition of UHECR with the surface detectors of the Pierre Auger Observatory. *Proceedings of the 31st ICRC, Lodz*, 2009.

[132] F. A. Gómez-Albarracin et al. Study of timing parameters t_y - t_x using SD data. *Internal GAP-Note 2008-017.*

[133] .T. Dova et al. *arXiv:astro-ph/0312463*, 2003.

[134] M. Ave et al. *Astrop. Phys.*, 19:61, 2003.

[135] L. del Peral and M. D. Rodríguez-Frias. Method to infer about the nuclear UHE cosmic ray composition from the EAS lateral development. *Nucl. Instrum. Meth.*, A391:239–248, 1997.

[136] C. D. England. The longitudinal development of extensive air showers at energies above 10^{17} eV. *PhD. Thesis at the University of Leeds*, 1984.

[137] M. T. Dova for the AUGER Collaboration. *Proc. 28th International Cosmic Ray Conference, Tsukuba*, 369, 2003.

[138] M. T. Dova, M. Mancenido, A. G. Mariazzi, H. Wahlberg, F. Arqueros, and D. Garcia-Pinto. UHECR mass determination using asymmetries in time distributions. *Internal GAP-Note 2007-135.*

[139] F. Arqueros, D. Garcia-Pinto, M. Monasor, M. T. Dova, A. G. Mariazzi, and H. P. Wahlberg. Inferring Xmax from asymmetry in risetime. *Internal GAP-Note 2008-179.*

[140] L. Cazon, R. A. Vazquez, A. A. Watson, and E. Zas. *Astropart. Phys.*, 21:71, 2004.

[141] L. Cazon, R. A. Vazquez, and E. Zas. *Astropart. Phys.*, 23:393, 2005.

[142] L. Cazon, M. Roth, R. Ulrich, G. Rodriguez, E. Zas, and R. Vazquez. Reconstruction of the muon production distance distribution. *Internal GAP-Note 2006-029.*

[143] S. Ostapchenko. *arXiv:astro-ph/0412591*, 2004.

[144] S. Ostapchenko. *arXiv:astro-ph/0501093*, 2005.

[145] R. Engel et al. *Proc. 26th Int. Cosmic Ray Conf. Salt Lake City (USA).*, 1:415, 1999.

[146] Klaus Werner and Tanguy Pierog. *arXiv:0707.3330*, 2007.

[147] J. Knapp et al. *Astrop. Phys.*, 19:77, 2003.

[148] A. Haungs et al. *Nuclear Physics B (Proc. Suppl.)*, 151:167, 2006.

[149] Ralph Engel for the Pierre Auger Collaboration. Test of hadronic interaction models with data from the Pierre Auger Observatory. *Proc. 30th International Cosmic Ray Conference. Mérida, Mexico. arXiv:0706.1921*, 2007.

[150] A. Castellina for the Pierre Auger Collaboration. Comparison of data from the Pierre Auger Observatory with predictions from air shower simulations: testing models of hadronic interactions. *Proceedings of the 31st ICRC, Lodz*, 2009.

[151] A. Hillas. *Proc. 19th ICRC*, 1:155, 1985.

[152] P. Billoir. *Astropart. Phys.*, 30:270, 2008.

[153] A. D. Supanitsky and G. Medina-Tanco. *Astropart. Phys.*, 30:264, 2008.

[154] Michael Unger. Study of the cosmic ray composition above 0.4 EeV using the longitudinal profiles of showers observed at the Pierre Auger Observatory. *arXiv:0706.1495*, 2007.

[155] P. Sokolsky and G.B. Thomson. Highest energy cosmic rays and results from the HiRes experiment. *arXiv:0706.1248*, 2007.

[156] N. Hayashida et al. *J. Phys. G: Nucl. Part. Phys.*, 21:1101, 1995.

[157] K. Shinozaki and M. Teshima. *Nuclear Physics B (Proc. Suppl.)*, 136:18, 2004.

[158] M. T. Dova, A. G. Mariazzi, and A. A. Watson. The mass of cosmic rays above 10^{17} eV. 2005.

[159] The Pierre Auger Collaboration. An upper limit to the photon fraction in cosmic rays above 10^{19} eV from the Pierre Auger Observatory. *Astroparticle Physics*, 27:155, 2007.

[160] F. Catalini et al. *arXiv:astro-ph/0703582*, 2007.

[161] The High Resolution Fly's Eye Collaboration. *Astroparticle Physics*, 26:28–40, 2006.

[162] A. K. O. Tiba and G. A. Medina-Tanco. *arXiv:astro-ph/0502255*, 2005.

[163] The Pierre Auger Collaboration. Observation of the suppression of the flux of cosmic rays above 4×10^{19} eV. *Physical Review Letters*, 101:061101, 2008.

[164] Claudio Di Giulio for the Pierre Auger Collaboration. Energy calibration of data recorded with the surface detectors of the Pierre Auger Observatory. *Proceedings of the 31st ICRC, Lodz*, 2009.

[165] Tokonatsu Yamamoto et al. (The Pierre Auger Collaboration). The UHECR spectrum measured at the Pierre Auger Observatory and its astrophysical implications. *Proc. 30th ICRC. Mérida, Mexico*, 2007.

[166] F. Schussler for the Pierre Auger Collaboration. Measurement of the cosmic ray energy spectrum above 10^{18} eV using the Pierre Auger Observatory. *Proceedings of the 31st ICRC, Lodz*, 2009.

[167] R. A. Vzquez for the Pierre Auger Collaboration. The cosmic ray flux observed at zenith angles larger than 60 degrees with the Pierre Auger Observatory. *Proceedings of the 31st ICRC, Lodz*, 2009.

[168] Domenico D'Urso for the Pierre Auger Collaboration. A monte carlo exploration of methods to determine the UHECR composition with the Pierre Auger Observatory. *Proceedings of the 31st ICRC, Lodz*, 2009.

[169] The Pierre Auger Collaboration. Upper limit on the cosmic-ray photon fraction at EeV energies from the Pierre Auger Observatory. *arXiv:0903.1127v2*, 2009.

[170] J. Abraham et al. Correlation of the highest energy cosmic rays with nearby extragalactic objects. *Science*, 318:938–943, 2007.

[171] R. Ulrich et al. *Nucl. Phys. B Proc. Suppl.*, 175:121, 2008.

[172] The Pierre Auger Collaboration. Upper limit on the diffuse flux of ultrahigh energy tau neutrinos from the Pierre Auger Observatory. *Physical Review Letters*, 100:211101, 2008.

[173] The Pierre Auger Collaboration. Limit on the diffuse flux of ultrahigh energy tau neutrinos with the surface detector of the Pierre Auger Observatory. *Physical Review D*, 79:102001, 2009.

[174] J. Tiffenberg for the Pierre Auger Collaboration. Limits on the diffuse flux of ultra high energy neutrinos using the Pierre Auger Observatory. *Proceedings of the 31st ICRC, Lodz*, 2009.

[175] D. Góra for the Pierre Auger Collaboration. UHE neutrino signatures in the surface detector of the Pierre Auger Observatory. *Proceedings of the 31st ICRC, Lodz*, 2009.

[176] I. Valiño. Detection of horizontal air showers and neutrino induced showers with the Pierre Auger Observatory. *Ph.D Thesis at Universidad de Santiago de Compostela. GAP-Note 2008-024.*

[177] David R. Ballantyne, Fulvio Melia, Siming Liu, and Roland M. Crocker. A Possible Link Between the Galactic Center HESS Source and Sgr A*. *Astrophys. J.*, 657:L13–L16, 2007.

[178] M. Aglietta et al. Anisotropy studies around the galactic centre at EeV energies with the Auger observatory. *Astropart. Phys.*, 27:244–253, 2007.

[179] Julien Aublin for the Pierre Auger Collaboration. Discriminating potential astrophysical sources of the highest energy cosmic rays with the Pierre Auger Observatory. *Proceedings of the 31st ICRC, Lodz*, 2009.

[180] J. Abraham et al. Correlation of the highest-energy cosmic rays with the positions of nearby active galactic nuclei. *Astropart. Phys.*, 29:188–204, 2008.

[181] J. D. Hague for The Pierre Auger Collaboration. Correlation of the highest energy cosmic rays with nearby extragalactic objects in Pierre Auger Observatory data. *Proceedings of the 31st ICRC, Lodz*, 2009.

[182] J. R. T. de Mello Neto for the Pierre Auger Collaboration. Intrinsic anisotropy of the UHECRs from the Pierre Auger Observatory. *Proceedings of the 31st ICRC, Lodz*, 2009.

[183] S. Yoshida et al. *J. Phys. G: Nucl. Part. Phys.*, 20:651–664, 1994.

[184] H. Y. Dai, K. Kasahara, Y. Matsubara, Nagano M., and M. Teshima. *J. Phys. G*, 14:793–805, 1988.

[185] N. M. Dyakonov et al. *Proceedings of the 17th International Cosmic Ray Conference, Paris, France*, 6:78, 1981.

[186] D. M. Edge et al. *J. Phys. A*, 6:1612, 1973.

[187] D. Newton, J. Knapp, and A. A. Watson. The optimum distance at which to determine the size of a giant air shower. *Astropart. Phys.*, 26:414–419, 2007.

[188] M. C. Medina et al. Enhancing the Pierre Auger observatory to the 10^{17} eV to $10^{18.5}$ eV range: Capabilities of an infill surface array. *Nucl. Instrum. Meth.*, A566:302–311, 2006.

[189] G. A. Medina-Tanco et al. *Proc. 29th International Cosmic Ray Conference, Pune*, 7:43–46, 2005.

[190] The Pierre Auger Collaboration. *Proc. 30th International Cosmic Ray Conference, Mérida, Mexico. arXiv:astro-ph/0706.1105*, 2007.

[191] The Pierre Auger Collaboration. *Proc. 29th International Cosmic Ray Conference, Pune, India. arXiv:astro-ph/0507029*, 2005.

[192] The Pierre Auger Collaboration. Angular resolution of the Pierre Auger Observatory. *Proceedings of the 29th ICRC, Pune, V7, p. 17*, 2005.

[193] The Pierre Auger Collaboration. Measurement of the UHECR energy spectrum from hybrid data of the Pierre Auger Observatory. *Proc. 30th International Cosmic Ray Conference. Mérida, Mexico.*, 2007.

[194] S. Sciutto. Aires user's manual and reference guide. *http://www.fisica.unlp.edu.ar/auger/aires*, 2002.

[195] A. Hillas. *Nucl. Phys. (Proc. Suppl.) B*, 52:29, 1997.

[196] S. Argiró et al. *Nucl. Instrum. Meth. A*, 580:1485, 2007.

[197] Geant 4. *http://geant4.web.cern.ch/geant4/.*

[198] P. L. Ghia for the Auger Collaboration. *Proc. 30th International Cosmic Ray Conference, Mérida, Mexico*, 2007.

[199] D. Nitz for the Pierre Auger Collaboration. Progress with the northern part of the Pierre Auger Observatory. *Proceedings of the 31st ICRC, Lodz*, 2009.

[200] O. Blanch Bigas, C. Bonifazi, and A. Letessier-Selvon. *Internal GAP-Note 2008-132.*

[201] B. Rouill d'Orfeuil, C. Lachaud, and E. Parizot. *Internal GAP-Note 2008-094.*

[202] T. Schmidt, I. C. Maris, and M. Roth. Fine tuning of the LDF parametrization and the influence on S1000. *Internal GAP-Note 2007-106.*

[203] T. Schmidt, M. Roth, and I. C. Maris. On LDF systematics deducing S38 using the constant intensity cut method. *Internal GAP-Note 2006-070.*

[204] The Pierre Auger Collaboration. Draft on auger energy spectrum.
Internal: http://augerobserver.fzk.de/ roth/doku.php?id=auger:augerspectrum.

[205] P.S. Allison. Testing methods of saturation recovery using per-PMT asymmetry. *Internal GAP-Note 2005-060.*

[206] I.C. Maris, M. Roth, and T. Schmidt. A phenomenological method to recover the signal from saturated stations. *Internal GAP-Note 2006-012.*

[207] I. De Mitri, S. Maglio, and D. Martello. A method for SD traces saturation recovery. *Internal GAP-Note 2006-075.*

[208] M. Aglietta, S. Maldera, G. Navarra, and F. Suárez. SD signals saturation: Summary of the hardware solution studies. *Internal GAP-Note 2007-134.*

[209] M. Aglietta, I. De Mitri, S. Maglio, S. Maldera, I. C. Maris, D. Martello, G. Navarra, and M. Roth. Recovery of saturated signals of the surface detector. *Internal GAP-Note 2008-030.*

[210] CDAS software group. Event reconstruction, release v4r4. *http://www.auger.org.ar/CDAS/.*

[211] D. Veberic and M. for the Offline team Roth. SD reconstruction reference manual. *Internal GAP-Note 2005-035 (Updated in March 2008).*

[212] M. Ave et al. The accuracy of signal measurement with the water Cherenkov detectors of the Pierre Auger Observatory. *Nucl. Instrum. Meth.*, A578:180–184, 2007.

[213] The Auger Observer. *http://observer.fzk.de.*

[214] I. C. Maris in a private communication. June 2009.

[215] Pirre Auger Collaboration. Atmospheric effects on extensive air showers observed with the surface detector of the pierre auger observatory. *Accepted for publication in Astroparticle Physics (June 2009). arXiv:0906.5497*, 2007.