INFERRING THE MASS COMPOSITION OF ULTRA-HIGH-ENERGY COSMIC RAYS THROUGH THE CHARACTERISATION OF THE MUON PRODUCTION PROFILE

Bruno Zamorano García

Universidad de Granada

April 2014

Advisor:
Prof. Antonio Bueno Villar

Co-advisor:
Dr. José Julio Lozano Bahilo

Departamento de Física Teórica y del Cosmos & CAFPE

Editor: Editorial de la Universidad de Granada
Autor: Bruno Zamorano García
D.L.: GR 1904-2014
ISBN: 978-84-9083-082-6

El doctorando Bruno Zamorano García y los directores de la tesis Antonio Bueno Villar y Julio Lozano Bahilo garantizamos, al firmar esta tesis doctoral, que el trabajo ha sido realizado por el doctorando bajo la dirección de los directores de la tesis y hasta donde nuestro conocimiento alcanza, en la realización del trabajo, se han respetado los derechos de otros autores a ser citados, cuando se han utilizado sus resultados o publicaciones.

Granada, 17 de marzo de 2014.

Directores de la tesis | Doctorando

Fdo: Antonio Bueno Villar, Julio Lozano Bahilo | Bruno Zamorano García

D. Antonio Bueno Villar, Catedrático de Universidad,

CERTIFICA: que la presente tesis doctoral, INFERRING THE MASS COMPOSITION OF ULTRA-HIGH-ENERGY COSMIC RAYS THROUGH THE CHARACTERISATION OF THE MUON PRODUCTION PROFILE, ha sido realizada por D. Bruno Zamorano García bajo su dirección en el Dpto. de Física Teórica y del Cosmos de la Universidad de Granada, así como que éste ha realizado una estancia en el extranjero por un periodo superior a tres meses en el *Karlsruher Institut für Technologie*.

Granada, 17 de marzo de 2014

Fdo: Antonio Bueno Villar

Agradecimientos

Grita y despeja de tu alma
lo que el mundo te contagia.
Sacude letargos que te amarran.
Reclama la vivencia de tu ser
sin jugar con tu conciencia,
cercenando ilusiones.
Refréscate con nítida
agua te ofrezco.

Ángel Zamorano Galán

Muchas cosas han tenido que ocurrir para que exista esta tesis. Como en un enorme castillo de naipes, en el que cada carta yace sobre el piso anterior y se sustenta sobre las de su lado, haciendo a su vez de soporte para las que se encuentran sobre ella, una miríada de personas me han traído hasta aquí. Seguramente, la influencia de acontecimientos aleatorios moldean nuestro devenir mucho más de lo que nos gustaría admitir. Si efectivamente esto es así, no puedo quejarme de la suerte que me ha tocado, pues de entre esas personas que se han cruzado en mi camino la mayor parte lo han hecho para bien.

Desde mis profesores del colegio, el instituto y la universidad, pasando por mis compañeros de clase y amigos, y culminando en mi familia, debería extenderme demasiado y poner a prueba mi memoria para agradecer a todos aquellos que han guiado mi camino hasta este punto. Valga una pobre disculpa para cuantos se queden en el tintero.

El protocolo dictamina que el primer agradecimiento debe ser para los directores de la tesis. Nunca he sido muy devoto del protocolo, y sin embargo me faltan palabras para describir todo lo que les debo. Ambos son ejemplos de honestidad, esfuerzo y constancia. Sin esos ejemplos, nunca habría llegado a asumir la importancia que esas cualidades tienen en la investigación.

Antonio es el mejor director que podría haber deseado, y lo es por muchas razones. En primer lugar, porque trata a sus estudiantes con respeto, tanto personal como profesional. Además, sabe sacar partido de las diferentes habilidades que cada estudiante posee, potenciando al mismo tiempo aquellas en las que los sabe menos duchos. Pero sobre todo porque no duda en remangarse y ensuciarse las manos para trabajar codo con codo como si se tratara de un becario más. Ha estado al tanto de todos los pasos de este trabajo desde el principio al fin, abriendo paulatinamente la mano conforme aumentaba mi madurez como investigador. Eso sí, nada de lo anterior impide que además se muestre amistoso y cercano en el trato personal. Los interminables viajes a Malargüe y las largas charlas delante de un bife han forjado una relación que trasciende lo pro-

fesional, y espero que así siga siendo por mucho tiempo. Además, aún me queda un abismo que aprender sobre vinos, así que su tutela no ha hecho más que comenzar.

En cuanto a Julio, a él le debo que me enseñara a dar los primeros pasos en este, por otro lado, árido campo. Ha sido también de gran ayuda en la revisión de esta tesis y de cuantos resultados parciales he ido desarrollando. Aunque sin duda lo que más tengo que agradecerle es que siempre haya estado ahí para portarse como un auténtico amigo, en los buenos y los malos momentos, y siempre con una palabra de aliento. Sólo puedo desearle lo mejor para el futuro, y esperar sinceramente que nuestros caminos vuelvan a cruzarse pronto.

Un despacho de becarios de Física de Partículas puede parecerse a muchas cosas según la ocasión. A veces a un comedor, otras a un foro de debate e incluso a una terapia de grupo. Lo que nunca será es una conserjería, pero sí que ha sido un lugar muy agradable en el que trabajar durante todo este tiempo, y ello es sobre todo gracias a los compañeros con los que lo he compartido.

Jorge, del que siempre contaremos que rompió la silla de tanto trabajar, fue uno de los compañeros más habituales al principio. Mark ya era mi amigo mucho antes de ser compañero de despacho, y aunque la mayoría lo recuerden por tus rabietas yo sé lo crack que es en realidad, igual que sé lo bien que le irá en su nuevo campo de trabajo. Diego no sólo se prestó entonces a enseñarme todo cuanto necesité, sino que a día de hoy sigue siendo colega, maestro y amigo. Roberto es de todos nosotros el más diestro en esto que es vivir, quizá por eso era tan divertido tenerlo en el despacho. Al igual que Adrián, cuyo humor jerezano y mala uva nos hicieron desternillarnos con frecuencia. Luis es un héroe de los de verdad, un triunfador que no sólo logró doctorarse de modo brillante con todo en contra, sino que además ha sido el experimental que mejor futuro se ha labrado, y bien merecido. Alberto ha sido un amigo leal durante todo este tiempo, que no ha dudado en echar un cable siempre que lo he necesitado, además de un gurú de ROOT y de casi todo lo imaginable. Mariano es el único argentino que conozco que detesta el calor, y sin embargo siempre lleva chaqueta; no cabe duda de que es un tipo irrepetible y muy divertido. Laura, *Laumobue* en nuestro pequeño círculo, ha sido un gran fichaje para el grupo, por su capacidad de sacrificio y de esfuerzo y porque se deja liar muy fácilmente si hay una fiesta de por medio. Patricia no le va a la zaga, y resulta divertido verla refunfuñar mientras demuestra día a día que su investigación es la que va más deprisa de cuantas se han sentado en el lado de la ventana. Rafa es un investigador brillante, tanto que tiene tiempo para además estar al tanto de absolutamente todo lo que ocurre en el mundo y para, además, intentar mejorarlo. Alice esconde una friki de aúpa debajo de una capa de sofisticación, y es capaz de bromear en un perfecto andaluz como si fuera su primera lengua. Ben es un caballero inglés cuyos modales le hacen humano, como dirían en su isla.

Fuera del despacho, pero no demasiado lejos, tengo que agradecerle sinceramente a María Dolores su ayuda en la producción de las simulaciones que han hecho posible esta tesis. También le agradezco los ratos de café y los desayunos, que no habrían sido lo mismo sin su gracia cordobesa.

Mis amigos también han puesto su granito de arena para hacer este pasaje más llevadero. Desde mis amigos de toda la vida, los del colegio y el instituto, a los que conocí en la carrera y los que simplemente han aparecido en mi vida, siempre he podido sentir

su confianza en mí, muchas veces más fuerte que la mía propia. Tener amigos que me tienen en tan alta estima es un premio que no merezco. Crecer consiste en abrir nuevos caminos, y algunos de ellos te alejan de gente de la que nunca quisieras separarte. Por eso, amigos como Elia, David, Avo, Migue, Pablo, Nacho, Caniche, Fiz, Rafa y Mark son tan valiosos, porque a pesar de la distancia y el tiempo sin verlos con ellos siempre es como si hubiera pasado solamente un rato desde la última vez. Además, cuando empecé a salir con Gloria recolecté otro grupo de amigos y amigas que empezaron siendo de ella y hoy son también míos. Caro e Ismael, María Cabrera y Florent, María Tejada, Aurora, Valle y los Antonios, tengo mucho que agradeceros por muy buenos ratos durante todo este tiempo y por haberme aguantado cuando no me aguantaba ni yo.

Mi familia paterna sería la envidia de cualquier persona con inseguridades. No sólo por su alegría pertinaz, sino porque su visión de mí y de mis capacidades es tan exagerada que termina siendo contagiosa, y siempre muy refrescante. Mis tíos y primos se deshacen en elogios hacia mí cada vez que me ven, pero lo que no saben es que yo sólo necesito su cariño, que también me dan de forma incondicional. Por si eso fuera poco, perdonan lo olvidadizo que soy y que casi nunca llame ni dé señales de vida. Mi familia materna es pequeña, pero su personalidad la hace grande. Mi tía Amparo fue una de las primeras en darse cuenta cuando dejé de ser un niño, y no ha fallado en ninguna de las grandes ocasiones de mi vida, como tampoco lo ha hecho mi Tata. Mis tíos de Barcelona son de otra pasta, parece que nunca les pudiera el pesimismo, y han sido una suerte de segundos padres para mí. Es natural entonces que mi primo Gabri sea como un hermano, con el que he compartido tantas cosas durante mi infancia y adolescencia que a veces dudo de si yo sería yo de no haberlo tenido a él.

Una de mis abuelas y mis dos abuelos ya no están aquí para alegrarse por mí, pero sé lo orgullosos que se sentirían de poder hacerlo. Al menos mi abuela Desamparados podrá presumir tanto como le apetezca, con esa vitalidad que la ha caracterizado siempre y de la que tanto tenemos que aprender. Si hay algo de lo que me vanaglorio es de lo mucho que de mis abuelos Ángel y Pepe hay en mí. A pesar de ser tan distintos en el modo de ver el mundo, ambos compartían un corazón enorme. Tal vez por eso a los dos les falló tan prematuramente. En cuanto a mi abuela Encarna, ojalá yo tenga el valor que ella demostró en su juventud si la ocasión lo requiere.

Pero sin duda la deuda más grande de todas la tengo con mis padres. Jóvenes, inocentes, seguramente demasiado honestos para el país y el tiempo en que vivimos, la diosa fortuna no les ha repartido las mejores cartas. Y sin embargo, año tras año se han sacrificado para que a mí no me faltara nada. Me han dado estabilidad, apoyo constante y una educación por encima de sus posibilidades reales. Tres décadas han pasado desde que aprendía la lógica de unas luces y un zumbador conectados a unos conmutadores en un improvisado juguete, y sin embargo es esa manera de pensar la que me ha llevado hasta aquí. Vosotros sois los gigantes sobre los que me he aupado para ver un poco más allá.

Mas el tiempo pasa, y el flujo natural de las cosas hace que uno termine fundando su propia familia. Conocí a Gloria en las postrimerías de una cada vez más longeva adolescencia. Este año se cumplirá una década desde que nos conocimos, una vez más por uno de esos azares del destino que en tan buena estima parece tenerme. Parece que fue ayer cuando nos pasábamos las horas muertas en el asiento del Nissan Sun-

ny simplemente conversando de todo cuanto podíamos, como si quisiéramos recorrer rápidamente los años en los que no nos teníamos el uno al otro. Gloria es mucho más que mi pareja, es el centro sobre el que gravita mi vida. Ha sido una compañera fiel y cariñosa todo este tiempo, incluso cuando el estrés me ha hecho mucho más huraño de lo que es natural en mí. Ha perdonado mis días oscuros como si nada, y me ha proporcionado los mejores años de mi vida. Por si eso fuera poco, nuestro amor ha dado el fruto más hermoso que podríamos desear en forma de un hijo del que estamos, si cabe, más enamorados todavía.

Junto con ella también recibí una familia extendida, que se ha portado conmigo durante todo este tiempo de forma inmejorable. El apoyo permanente de mis suegros, su ayuda en absolutamente todos los proyectos en los que nos hemos embarcado estos años y su cariño han sido clave para que no tirase la toalla cuando todo parecía de cara. También la alegría y el cariño de Rocío son revitalizantes, y me gusta mucho la mujer en la que se ha convertido durante estos años. Solamente le puede ir bien en el futuro, es de recibo.

Acknowledgements

Finally, a few words in English to acknowledge my hosts in Karlsruhe. I discovered in Michael Unger an outstanding scientist who is, nevertheless, an unaffected an really cool guy. Markus Roth and Ralph Engel were really warm and friendly to me, and made me feel at home. They even included me in their respective group internal meetings, in which I learnt a lot about the SD and the FD. Ralph Ulrich took me on a fascinating tour around KASCADE and even left me borrow his mugs and steal his coffee. I had a whale of fun sharing our amazement with some of the German habits with Juan. Last but not least, I thank my office mates Detlef, Hans and Alex for very nice moments, including billiard and sports evenings, and for bearing endless videoconferences in Spanish.

A Gloria y Pablo,
mi mejor hallazgo
y mi más sublime contribución.

Contents

Introduction

In the early twentieth century, Victor Hess [1] discovered the extra-terrestrial origin of cosmic rays. For over 100 years, physicists have studied this radiation with increasingly precise techniques. The existence of muons, mesons and even antimatter was discovered by the careful study of cosmic rays. The cosmic ray energy spectrum spans over eleven orders of magnitude, from the GeV cosmic rays of solar origin up the ultra-high-energy cosmic rays (UHECRs) with energies larger than 10^{18} eV. Cosmic rays are continuously bombarding the atmosphere isotropically, but their flux is rapidly decreasing with the energy, so at the highest energies only one particle per square kilometre per century reaches the Earth. This extremely low rate has immediate consequences on their detection, which is generally achieved by deploying detectors covering large areas on ground.

When cosmic rays collide with atmospheric nuclei, they produce secondary particles which are energetic enough to keep the process going, subsequently giving birth to billions of secondary particles that propagate through the atmosphere and eventually reach ground. These secondary particles, known as Extensive Air Showers (EASs), are registered by the detectors, and from them one extracts information regarding the energy, origin and composition of the primary particle. The properties of the mechanism of generation of secondaries are complex, and depend strongly on the first interaction, which takes place at an energy beyond the ones reached at accelerators. As a consequence, the analysis of UHECRs is subject to large uncertainties and hence, many of their properties, in particular their composition, are still unclear.

The Pierre Auger Observatory [2], located in the province of Mendoza, Argentina, was designed to study the properties of UHECRs. It is the largest hybrid detector ever built, combining a fluorescence detector (FD) and a surface detector (SD). The FD collects the ultraviolet light emitted by the de-excitation of nitrogen nuclei in the atmosphere, and can operate only in clear, moonless nights. The SD samples the density of particles at ground level using more than 1600 Water-Cherenkov tanks deployed over an area of about 3000 km^2 and has a nearly 100% duty cycle.

The main goal of this thesis is the measurement of the mass composition of UHECRs using data of the SD. The correct determination of the mass composition is key in order to understand the origin of cosmic radiation and the mechanisms of acceleration that can boost particles up to such enormous energies. These questions remain unanswered after more than one century of cosmic ray Physics.

The Pierre Auger Observatory has published prominent results regarding mass composition of UHECRs using FD measurements [3]. However, it is of much interest to measure the mass composition of UHECRs using only SD data, as they provide an independent measurement that can either support or disfavour FD results. Moreover, the large

duty cycle of the SD yields an event statistics one order of magnitude larger than the FD, which opens the window to a better characterisation of the very largest energies, at which the flux is extremely low.

In this thesis, the analysis of mass composition of UHECRs is performed through the characterisation of the longitudinal development of the muonic component of EASs. Chapter 1 gives an overview of cosmic rays, including some of the most relevant experimental results. In chapter 2, the main features of the Pierre Auger Observatory are described in detail, including the reconstruction of the most relevant variables. Chapter 3 concludes the review of previous results with a thorough description of the shower development and the most relevant variables that have been proposed to infer the mass composition of UHECRs.

Chapters 4 and 5 introduce the main caveats that were found while aiming to obtain a good reconstruction of the variables of interest, together with the proposed solutions. Chapter 6 explores the possibility of measuring shower-to-shower fluctuations with the SD, as an independent hint of mass composition. All of these measurements are interpreted in chapter 7 in terms of mass composition, where some future potentialities of the method are introduced for further work. Finally, a few additional mathematical details are briefly described in the appendixes.

Introducción

A principios del siglo XX, Victor Hess [1] descubrió el origen extraterrestre de los rayos cósmicos. Esta radiación ha sido estudiada durante más de 100 años empleando técnicas cada vez más sofisticadas. La existencia de los muones, los mesones e incluso de la antimateria fueron descubiertas mediante el estudio de los rayos cósmicos. El espectro de energía de los rayos cósmicos abarca once órdenes de magnitud, desde los rayos cósmicos de origen solar, con energías del orden del GeV hasta los rayos cósmicos ultra energéticos (UHECRs), con energías por encima de 10^{18} eV. Los rayos cósmicos bombardean la atmósfera continuamente de forma isótropa, pero su flujo decrece fuertemente con la energía, de modo que a las más altas energías solamente una partícula por kilómetro cuadrado y por siglo llega a la Tierra. Esta tasa tan extremadamente baja impone limitaciones sobre su detección, que solo es posible mediante detectores que abarcan una gran superficie de terreno.

Cuando un rayo cósmico colisiona con un núcleo de la atmósfera se producen partículas secundarias que poseen suficiente energía para repetir el proceso, en una sucesión que da lugar a miles de millones de partículas secundarias, que se propagan a través de la atmósfera y finalmente llegan al suelo. Estas partículas secundarias, conocidas como *cascadas*, o por su acrónimo en inglés (EASs, de *Extensive Air Showers*), se colectan en los detectores, y a través de ellas se extrae información acerca de la energía, origen y composición de la partícula primaria. Las propiedades de este mecanismo de generación de partículas secundarias son complejas, y dependen en gran medida de la primera interacción, que se produce a una energía más alta que la que se puede alcanzar en aceleradores de partículas. Como consecuencia, el análisis de los UHECRs está sujeto a grandes fuentes de incertidumbre y, por ello, muchas de sus propiedades, en particular su composición, siguen sin resolver.

El Observatorio Pierre Auger [2], situado en la provincia de Mendoza, Argentina, se diseñó para estudiar las propiedades de los UHECRs. Es el detector híbrido más grande jamás construido, y combina un detector de fluorescencia (FD) con uno de superficie (SD). El FD recoge la luz ultravioleta emitida durante la desexcitación de los núcleos de nitrógeno de la atmósfera, y puede operar únicamente en noches claras y sin luna. El SD muestrea la densidad de partículas a nivel del suelo mediante más de 1600 tanques Cherenkov de agua espaciados 1.5 km en una red triangular sobre un área de más de 3000 km^2, y su funcionamiento es próximo al 100 % del tiempo.

El objetivo principal de esta tesis es la medida de la composición química de los UHECRs utilizando datos del SD. Una correcta determinación de la composición es crucial para entender el origen de la radiación cósmica, así como los mecanismos de aceleración capaces de impulsar partículas a energías tan enormes. Estas preguntas siguen sin respuesta tras más de un siglo de Física de rayos cósmicos.

El Observatorio Pierre Auger ha publicado importantes resultados acerca de la composición química de los UHECRs utilizando medidas del FD [3]. Sin embargo, existe un gran interés en medir la composición de los UHECRs utilizando para ello datos del SD, ya que esto proporcionaría una medida independiente que podría bien reforzar o poner en cuestión los resultados del FD. Además, el elevado ciclo de trabajo del SD proporciona una estadística de sucesos alrededor de un orden de magnitud mayor que el FD, lo que abre la ventana a una mejor caracterización de las más altas energías, a las cuales el flujo es extremadamente bajo.

En esta tesis, el análisis de la composición química de los UHECRs se lleva a cabo mediante la caracterización del perfil de desarrollo longitudinal de la componente muónica de las EASs. En el capítulo 1 se resumen los principales aspectos de la Física de rayos cósmicos, incluyendo algunos de los resultados experimentales más relevantes. En el capítulo 2 se describen en detalle las principales características del Observatorio Pierre Auger, incluyendo la reconstrucción de las principales variables de interés. El capítulo 3 concluye la revisión de resultados previos con una descripción detallada del desarrollo de las cascadas y las variables más relevantes que se han propuesto para inferir la composición química de los UHECRs.

Los capítulos 4 y 5 introducen las principales dificultades que se encontraron al intentar obtener las variables de interés, junto a las soluciones propuestas al efecto. El capítulo 6 explora la posibilidad de medir las fluctuaciones cascada a cascada con el SD, como una medida independiente de la composición química. Todas estas medidas se interpretan en términos de composición química en el capítulo 7, donde también se presentan algunas posibles vías abiertas de gran interés para el futuro. Finalmente, algunos detalles matemáticos se describen brevemente en los apéndices.

The fact that you are not sure means that it is possible that there is another way someday

Richard P. Feynman

1 Cosmic rays

Cosmic rays of ultra-high energy (UHECRs) are one of the biggest remaining puzzles in Modern Physics [4]. Their energy is far beyond the limits reached in particle accelerators. The centre of mass energy in LHC will reach 14 TeV, while for a proton with a momentum of the order of 10^{20} eV hitting a nitrogen nucleus it is of the order of the PeV. As a consequence, the theoretical models used in the study of cosmic rays are extrapolations of much lower energy data, and thus subject to high uncertainties. Even today, more than 100 years after their discovery [1], we are unable to identify unequivocally their source of production, mechanisms of acceleration or chemical composition.

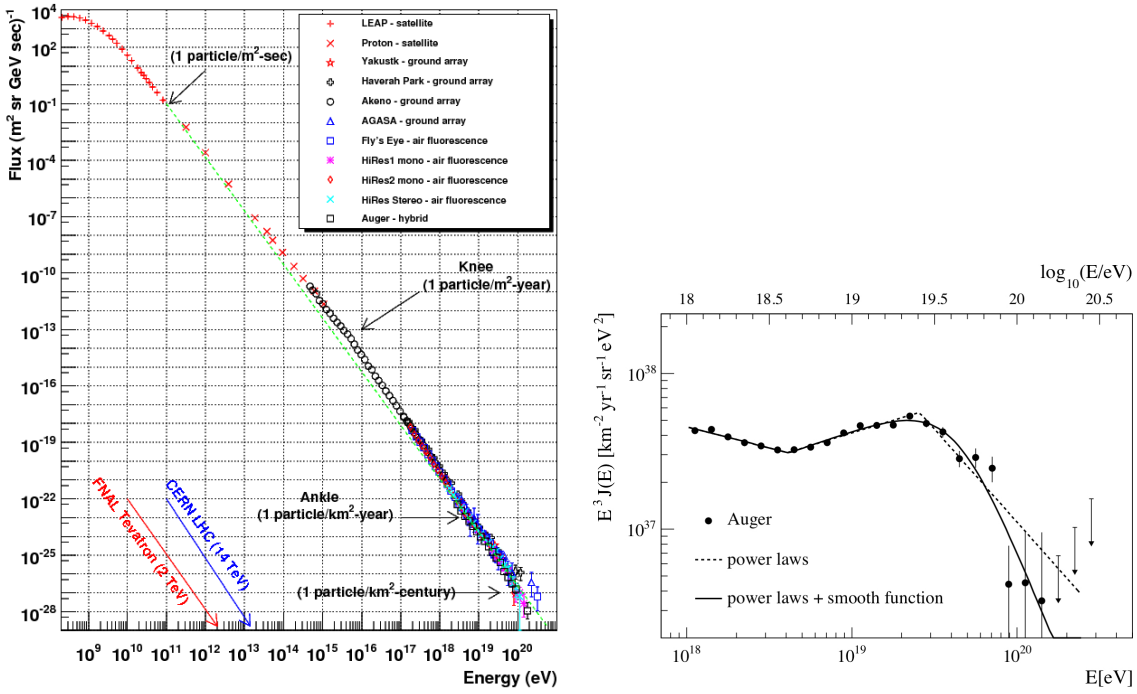
One of the main technical difficulties when dealing with Cosmic Ray Physics is their extremely low event rate (figure 1.1a). In the range of energies in which we are more interested (above 1 EeV[1]), the flux is about one event per km^2 per year, so the only feasible possibility for an experiment to have enough statistics is to arrange detectors covering large areas on the ground. As a result, cosmic rays have to be detected and analysed indirectly, measuring the myriad of secondary particles generated as a shower when a cosmic ray interacts with the atmosphere. This makes the analysis and reconstruction of the primary information a laborious task.

The Pierre Auger Observatory [2] is the largest experiment ever built by humankind, and it has been designed taking these difficulties into account.

1.1 The discovery of cosmic rays

Atmospheric ionisation was discovered by Charles-Augustin de Coulomb ca. 1785 through the spontaneous discharge of electroscopes [5] and confirmed later by Faraday, Kelvin, Crookes and others. With the discovery of natural radioactivity in 1896 by Becquerel [6], the origin of this radiation seemed to be the presence of some unstable

[1] 1 EeV = 10^{18} eV



(a) Cosmic ray flux

(b) Energy spectrum multiplied by E^3

Figure 1.1: Cosmic ray flux as measured by different experiments (left). Energy spectrum obtained by the Pierre Auger Observatory multiplied by E^3 to enhance structures (right).

isotopes in the cortex of the Earth. It was Victor Hess in 1912 [1] who carried scintillation detectors on a balloon in a celebrated experiment (figure 1.2) and finally demonstrated that the amount of ionisation registered by a gas detector increased with altitude. This ruled out the hypothesis of terrestrial origin of this radiation. For his discovery, Hess was granted the Nobel Prize in Physics in 1936.

The term “cosmic rays” nagan

After the primary collision, a cascade of secondary particles is produced, giving birth to the so-called Extensive Air Showers (EASs). They were given this name because they can cover several square kilometres at ground level. The details of this process will be discussed in further sections.

Cosmic rays exhibit a rather unique property. They arrive to the Earth with energies ranging from less than 1 GeV up to 10^{20} eV, that is, more than eleven orders of magnitude (see picture 1.1a). This fact by itself is challenging enough from a theoretical point of view, making very difficult to explain the mechanisms of acceleration capable of accelerating subatomic particles from a few GeV up to macroscopic kinetic energies (well above the Joule).

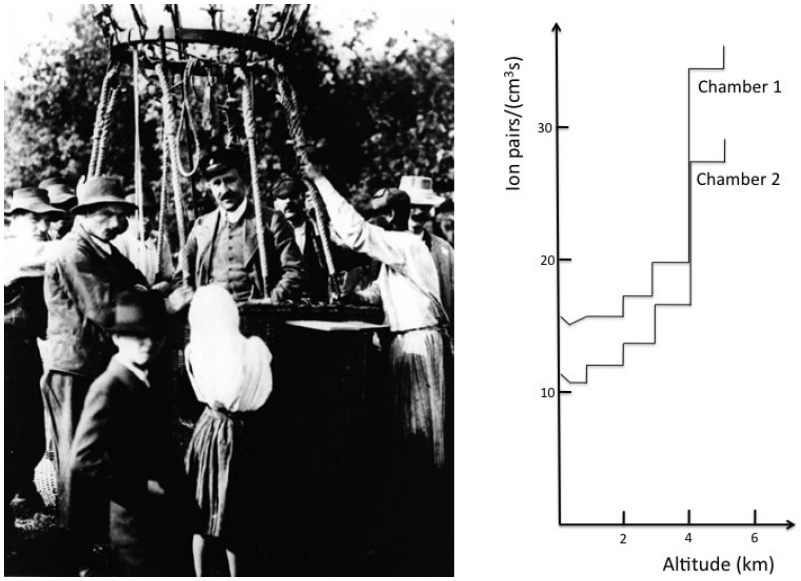


Figure 1.2: Victor Hess during his balloon experiment in 1912 (left). Dependence of the ionisation with the altitude (right).

1.2 Cosmic rays and Particle Physics

The study of cosmic rays led to a series of discoveries, especially during the first half of the twentieth century, that set the pillars for the growth of Particle Physics as a separated field of study. Antimatter, mesons, and a whole set of particle zoology were found in cosmic rays.

The first revolutionary discovery was Skobeltsyn's first observation of the positron in 1929 [7]. He was trying to measure gamma radiation in cosmic rays using a cloud chamber when he detected particles with the same behaviour as electrons, but with opposite direction of curvature when exposed to an external magnetic field. Anderson finally found the positron in 1932 [8] and he was awarded the Nobel Prize in Physics in 1936 for his discovery.

Muons were discovered by Anderson and Neddermeyer in 1936 [9] while studying cosmic rays. They found particles that, when passing through a magnetic field, curved less than protons but more than electrons of the same energy. They concluded that the mass of this new particle would be intermediate, and thus called them *mesons*.

In 1947, Powell discovered the pion [10], which had been proposed by Yukawa in 1935 [11] as the mediator of the strong force. The same year, Rochester and Butler discovered the kaon [12], the first strange particle ever found. This led to the postulation of the strange quark and the quark model by Gell-Mann and Zweig [13, 14].

During the second half of the twentieth century, the field of cosmic rays lost part of its prominent position in Modern Physics, due partly to the experimental difficulties but mostly to the tremendous development of particle accelerators. However, during the last years, there has been a series of new discoveries and open questions that have put cosmic rays in the spotlight again.

1.3 The birth of Astroparticle Physics

Some of the most recent discoveries, such as the neutrino oscillation in the Super-Kamiokande experiment [15], the beginning of gamma-ray Astronomy [16] or the direct searches for dark matter [17, 18, 19], have demonstrated that the connections between Particle Physics, Cosmology and Astrophysics are much closer than ever thought. This has led to the birth of Astroparticle Physics[2] as a new area of expertise with its own entity.

Given the transversal nature of the field, a combination of techniques and methodologies need to be used. This is sometimes known as the *Multi-Messenger* approach. Only by the correct combination of results can astrophysical or cosmological models be ruled out or favoured, and the insight gained by a common effort is becoming increasingly more evident.

Astroparticle Physics addresses almost every unsolved problem in Physics in one way or another, from cosmic inflation to baryon asymmetry or supernovae studies, and opens the window to challenging searches, such as magnetic monopoles, gravitational waves, micro black holes, dark matter and dark energy or supersymmetry. Most of these long-standing questions will probably remain unsolved for many years to come, but both the experimental and theoretical challenges they pose are of much interest.

1.4 Cosmic ray energy spectrum

The cosmic ray spectrum can be described as a rather featureless power law function extending from energies around 10^9 eV up to 10^{20} eV, more than eleven orders of magnitude. The exponent of this power law, referred to as *spectral index*, is almost constant and close to 3

$$J = \frac{dN}{dE} E^{-\gamma}, \qquad \gamma \approx 3 \tag{1.1}$$

Figure 1.1a shows the cosmic ray flux (spectrum per unit of area, solid angle and time) as measured by different experiments.

Three regions of the spectrum exhibit a particularly interesting deviation from the average behaviour, the ‘’Knee”, the ‘’Ankle” and the highest region of the spectrum. The properties of these regions will be explored in the following sections.

In order to visually enhance the properties of these structures, the flux is sometimes multiplied by some power of the energy, e.g. E^3 or $E^{2.7}$, see picture 1.1b. This is somewhat risky when comparing experimental results, as it exaggerates the differences in the energy scale of different experiments.

1.4.1 The Knee

The Knee consists of a steepening in the spectral index from $\gamma \approx 2.7$ to $\gamma \approx 3.1$ that takes place at an energy around 10^{15} eV. Many possible explanations to this feature have been proposed, based mostly on astrophysical reasons, but also on new physical phenomena

[2] Sometimes referred to as Particle Astrophysics, especially in the United States

in the atmosphere leading to energy transfer to particles not yet detected by air shower experiments (see [20] for a very thorough review of models).

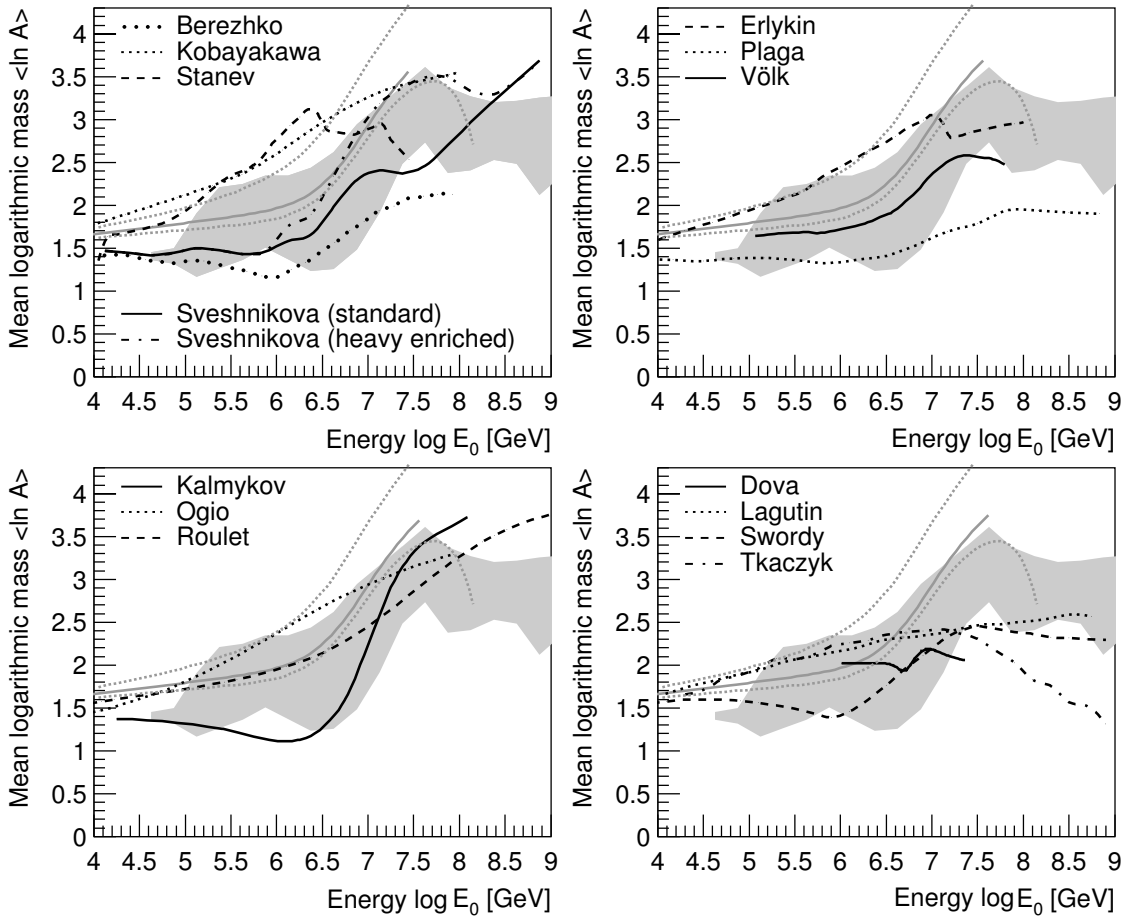


Figure 1.3: Comparison of different theoretical models for the Knee with the mean logarithmic mass derived from many experiments (shaded area) [20].

The existence of the Knee has been confirmed by several experiments, such as Yakutsk [21] and Akeno [22], and studied in depth by KASCADE [23] and its extension KASCADE-Grande [24]. The latter showed recently evidence that the Knee is presumably caused by the decrease in the flux of heavy primaries. However, this result is subject to many uncertainties regarding the high energy models used in the simulations.

The current experimental results are not constraining enough to favour one theoretical model over the rest. Very likely, a combination of different processes regarding acceleration, propagation and interactions of cosmic rays occurs in this particular region of energy, giving rise to the Knee. As it can be seen in picture 1.3, only through a very precise determination of the average mass of cosmic rays, can some models be ruled out or favoured.

1.4.2 The Ankle

According to the measurements of the Pierre Auger Observatory, the Ankle is found at $\log(E/\text{eV}) = 19.62 \pm 0.02$. As pointed out in picture 1.1a, the Ankle consists of a flattening of the spectrum to a spectral index which is again close to $\gamma \approx 2.7$. This flattening has been observed by several other experiments, e.g. HiRes [25].

The origin of the ankle is frequently understood as the transition from galactic to extragalactic cosmic rays. However, the particular characteristics of this transition are different in the various models proposed to understand its features. For instance, the *ankle model* assumes that both the galactic and extragalactic components contribute equally to the total flux, and the extragalactic component is expected to have pure proton composition [26, 27].

On the other hand, the *mixed composition model* [28] is the only model that can accommodate a non pure composition at the highest energies. In this model, the ankle is explained again by the transition from galactic to extragalactic cosmic rays.

Finally, the *dip model* [29, 30, 31], predicts that the extragalactic component, again with a proton-dominated mass composition, is extinguished at much lower energies, and the ankle is explained by the interaction of protons with the cosmic microwave background, where an electron-positron pair is produced and the energy of the primary is reduced. This has a net contribution of suppressing the flux at high energies, increasing it at low ones, and therefore causing the appearance of the ankle.

Disentangling the properties of the ankle to favour on of the models is complicated. However, depending on the mass composition of these two components of the flux, the properties are different, showing again the importance of an accurate measurement of $\langle \ln A \rangle$. For example, in the *mixed composition model* [28], the abundance of heavy elements is of the order of 10% for energies above the Ankle, which is compatible with current observations from the Pierre Auger Observatory, while for the *dip model* this contribution is negligible, and thus is unfavoured by observations.

1.4.3 The most energetic cosmic rays

Shortly after the discovery of cosmic microwave background (CMB) by Penzias and Wilson [32], Greisen [33], and independently Zatsepin y Kuzḿin [34] reached the conclusion that the very existence of it would produce an attenuation in the energy of cosmic rays. This effect is commonly known as the GZK suppression. For nucleons, the reaction taking place is

$$N + \gamma_{\text{CMB}} \longrightarrow \Delta \longrightarrow N + \pi \tag{1.2}$$

This reaction becomes important above 5×10^{19} eV, where the interaction length is about 6 Mpc and the energy loss per interaction of the order of 20%. A direct conclusion of this is that any nucleon above this energy must come from a source closer than ~ 100 Mpc from Earth. Analogously, for photons we have the following process

$$\gamma_{\text{CR}} + \gamma_{\text{CMB}} \longrightarrow e^+ + e^- \tag{1.3}$$

with an attenuation length of the order of 10 Mpc for the energies of interest. Therefore, every stable particle apart from neutrinos suffers a significant attenuation at energies close or above the EeV, meaning that cosmic rays of ultra high energy must have their origin in nearby astrophysical objects.

A key aspect of the GZK effect involves high energy photons and neutrinos, as their observation can provide an indirect hint of the mass composition of ultra-high energy cosmic rays. In fact, in equation (1.2) there are different possibilities for the remnants. The first one would be

$$p + \gamma_{\text{CMB}} \longrightarrow p + \pi^0 \longrightarrow p + \gamma\gamma \tag{1.4}$$

resulting in two cosmogenic photons, and the second one

$$p + \gamma_{\text{CMB}} \longrightarrow n + \pi^+ \longrightarrow n + \mu^+ + \nu_\mu \longrightarrow n + e^+ + \nu_e + \bar{\nu}_\mu + \nu_\mu \tag{1.5}$$

which produces three neutrinos in the final state.

The discovery of cosmogenic neutrinos and photons would favour a light cosmic ray composition at ultra-high energies. This is due to the fact that photo-pion production for heavy primaries has a much higher energy threshold.

Recently, two PeV neutrinos have been detected by the IceCube collaboration [35] with a statistical significance of 2.8σ. These two events may be the first indication of the existence of an astrophysical neutrino flux. However, given the small statistical significance of the result it is just too soon to extract strong physical implications.

However, even if a suppression at the highest energies is observed, it is not guaranteed that it must have originated through the GZK mechanism. In fact, the observation of increasingly heavier composition in the spectrum by the Pierre Auger Observatory, together with a strong spectrum steepening at the highest energies is not fully consistent with the predicted shape of the GZK cutoff [36].

Under the assumption of a rigidity acceleration mechanism, meaning that the maximum acceleration energy for a nucleus of charge number Z is $ZE_{\text{p}}^{\text{max}}$, with a limit around 100-200 EeV, the contribution of lighter elements vanishes with increasing energy and the suppression in the flux is automatically provided by the fact that sources of acceleration cannot reach infinite energies. This model has some rather discouraging consequences, and for that it has been named the *disappointing model*. Some of these consequences are:

- There is no pion photo-production on CMB photons in extragalactic space
- GZK cutoff in the spectrum does not exist
- Cosmogenic neutrinos produced by the interaction with the CMB are absent
- Fluxes of cosmogenic neutrinos produced on infrared-optical background radiation are too low for registration by existing detectors and projects
- Due to nuclei deflection in galactic magnetic fields, the correlation with nearby sources is missing even at the highest energies

This model can be considered as the most conservative scenario in Cosmic Ray Physics. It simplifies the problem of finding sources of acceleration in astrophysical objects, as the maximum energy required is diminished to ∼100 EeV for iron nuclei. On the other hand, any positive result in the consequences mentioned above, would automatically disfavour it.

1.5 Origin of cosmic rays

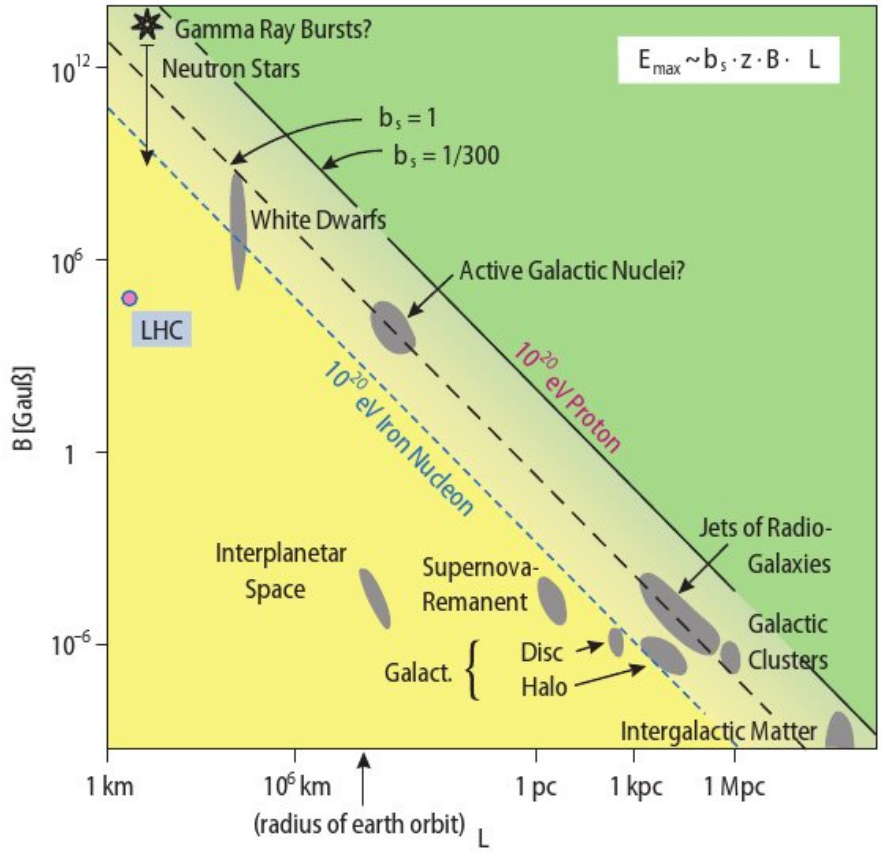


Figure 1.4: The Hillas plot represents the strength of the magnetic field versus the size of possible candidates for ultra-high energy cosmic ray acceleration. Objects below the diagonal can be ruled out.

One of the main puzzles to address when trying to disentangle the origin of cosmic rays is the vast range of energy covered by them (see picture 1.1a for a reminder of the cosmic ray spectrum). It is natural to think that a handful of astrophysical objects may contribute to the different regions of the spectrum, but this gives rise to the question of why this spectrum is, to a great extent, featureless.

Cosmic rays up to some GeV are known to come mostly from the Sun. This is confirmed by direct detection experiments by both their mass composition (compatible with

what we would expect from Solar Physics) and distribution in the sky. However, we know that at higher energies the origin of cosmic rays must be different, as there is not any process involving such a huge energy taking place in the Sun. In order to shed some light to this problem, the different candidates for cosmic ray production are represented in the Hillas plot (picture 1.4).

As we have discussed in the previous section, we know that the most energetic cosmic rays arriving to the Earth must have been originated within a sphere of radius ~ 100 Mpc. But of course, in order for particles to escape whichever the accelerating source, they must have a Larmor radius larger than the typical scale of such object. This constraint can be summarised using the relation between the maximum energy, E_{max}, a particle of charge Z can acquire when accelerated by a magnetic field B in an object of size R:

$$E_{\text{max}} = k\beta Z \left(\frac{B}{\mu\text{G}}\right)\left(\frac{R}{\text{kpc}}\right) \text{ EeV} \tag{1.6}$$

where β is the acceleration velocity in units of c and k the efficiency in the process of acceleration ($k < 1$).

Imposing this condition, several candidates can be ruled out in the Hillas plot, leaving as candidates Gamma Ray Burst (GRBs), Active Galactic Nuclei (AGNs), neutron stars and radio galaxies.

1.5.1 Anisotropy in the TeV scale

Muons reaching underground detectors with GeV energies are mainly produced by primary cosmic rays with TeV energies [37]. Therefore, by studying the arrival direction of muons in underground detectors, possible anisotropies in the TeV scale can be explored.

The largest underground muon detector is IceCube [38]. This experiment has recorded enough events to look for anisotropies of the order of a few per-mile. Their results reveal that the arrival direction distribution is not isotropic. Significant structures are seen at different angular scales. In fact, the sky map of the southern hemisphere shows several regions of significant excess and deficit of cosmic rays. This result is compatible with what was found in the northern hemisphere by the Milagro experiment [39]. The cause for this anisotropy is yet to be explained.

1.5.2 Anisotropies in UHECRs

Disentangling the origin of cosmic rays of ultra-high energy has a lot to do with understanding their mass composition and the large-scale magnetic fields, especially in the neighbourhood of our galaxy.

However, if we consider only the highest part of the spectrum, and make the assumption of a light-dominated mass composition in this region, only a small deviation due to the magnetic fields is expected. Under these conditions, we are capable of performing a point-like analysis in search of a signature of anisotropy.

The Pierre Auger Observatory has an excellent angular resolution of about 1 degree at the highest energies [2], and thus the search for point-like sources is feasible. This sort of searches have produced some evidence of a positive correlation [40, 41, 42] between

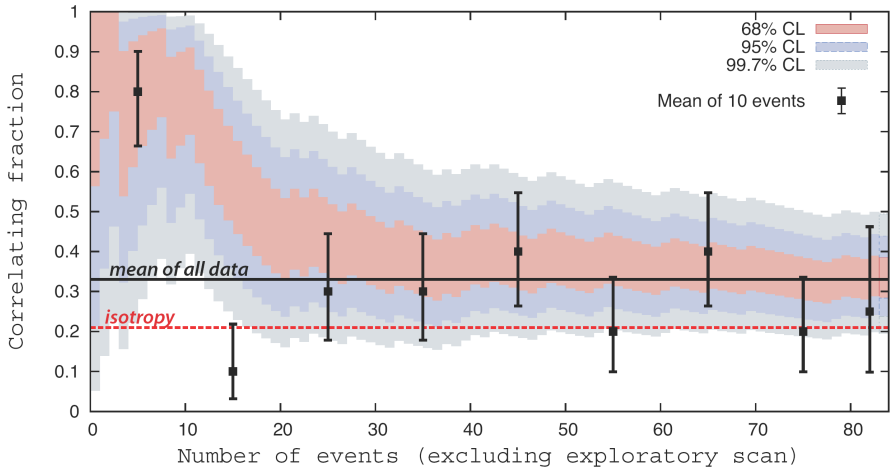


Figure 1.5: Pierre Auger result for the correlation of the incoming direction of ultra-high energy cosmic rays with the location of nearby AGNs. The most likely value of the degree of correlation, p_{data} is plotted as a function of the total number of events in chronological order. The 68%, 95% and 99.7% confidence level intervals are shaded. Isotropy is represented by the horizontal line at $p_{\text{iso}} = 0.21$. Black squares represent the correlation fractions obtained if the events are split in bins of 10 consecutive events.

the arrival direction of ultra-high energy cosmic rays and nearby AGNs from the VCV catalogue [43]. The amount of correlation found for events above 55 EeV is (33 ± 5)%, compared to a 21% expected from isotropy (figure 1.5). This is within the 3σ level and therefore implies only an evidence, yet to be confirmed. Telescope Array has found some degree of correlation as well [44], where 11 out of 25 events above 57 EeV correlate with AGNs for an expected number of 5.9 in the case of isotropy. Both measurements are compatible within uncertainties but none of them yields figures of statistical significance.

1.6 Mass composition of UHECRs

The composition of cosmic rays of energies below the knee can be measured directly by space-based experiments. However, at higher energies, the only possibility is to characterise the properties of the EASs generated by different primaries. However, the large fluctuations that can take place during the shower development, it is extremely difficult to distinguish showers generated by different primaries on an event by event basis. Instead, a statistical approach is generally followed in terms of some observables that are sensitive to the mass composition of the primary particle.

The main observable used to quantify the mass composition of UHECRs is the maximum of the shower development, X_{max}, which can be measured by fluorescence telescopes. Showers generated by protons have an average value of X_{max} about 100g/cm^2 larger than showers produced by iron nuclei of the same energy, so it is possible to measure the energy evolution of this observable. This was achieved at for the first time by the

Fly's Eye collaboration [45] and followed by different experiments since then [46] (see figure 1.6). More details will be given in chapter 3 about this parameter and its experimental measurements.

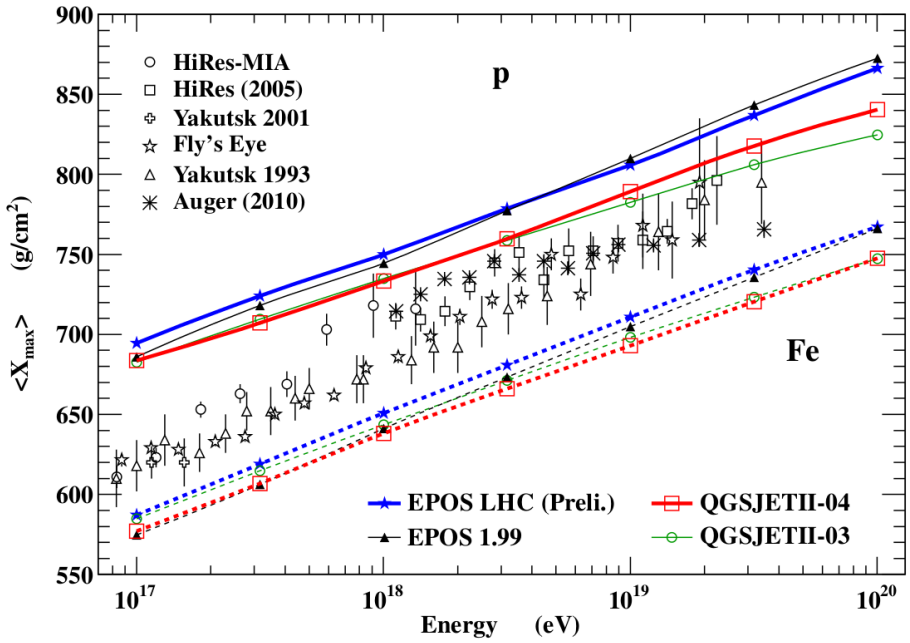


Figure 1.6: The maximum depth of shower as a function of the energy as measured by various experiments before the Pierre Auger Observatory as compared to different Monte Carlo predictions [47].

A complementary approach to determine mass composition consists on the measurement of the relative number of electrons and muons in the shower at ground level. This was proven to contain information related to the mass composition of the primary by Oda and collaborators in 1960 [48] and exploited by many experiments, such as KASCADE-Grande [24, 49] using muon detectors of much greater area.

The different experimental results concerning the measurement of mass composition of UHECRs, as well as their theoretical interpretation, will be discussed in depth in chapter 3.

1.7 Diffuse neutrino flux searches

Neutrinos coming from the Sun, with energies of the order of a few MeV, hit the Earth at an impressively high rate of about 10^8 neutrinos per cm^2 and per second. For that reason, even if the cross section of these elusive particles is very low, their detection has been very successful in the past [50]. Encouraging measurements were performed by detectors observing the evolution of Supernova SN1987A [51]. However, the detection of extraterrestrial neutrinos of higher energy is still an experimental challenge.

Neutrinos of astrophysical origin can be classified in three different categories, de-

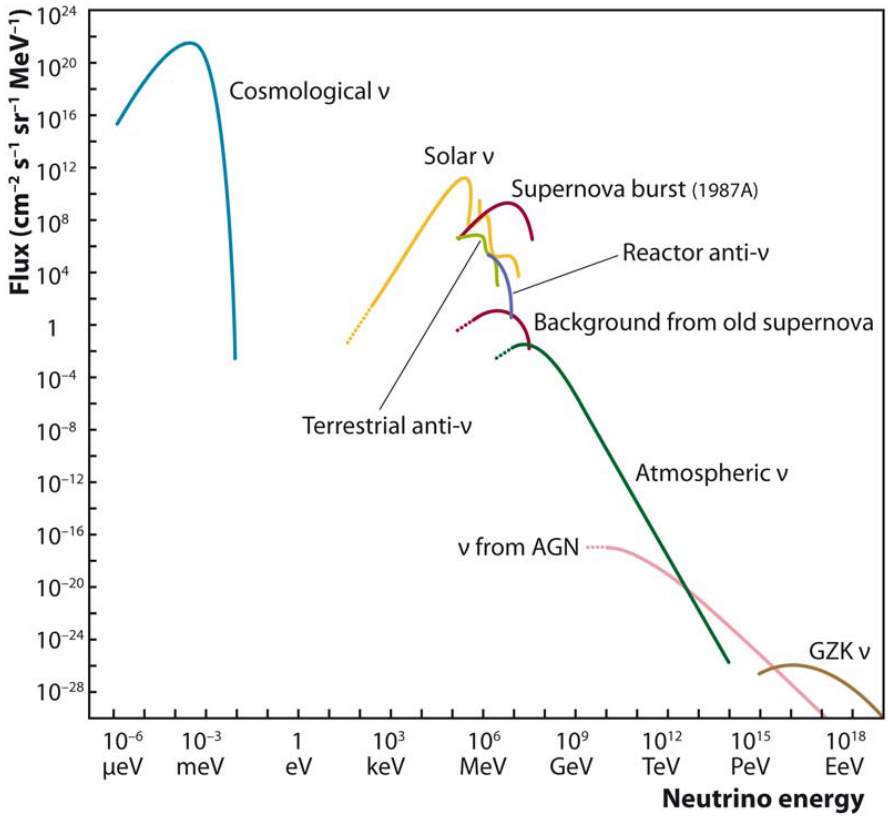


Figure 1.7: The “gran unified” neutrino flux [52].

pending on their production mechanism: astrophysical, cosmological and exotic neutrinos. Figure 1.7 shows the diffuse “grand unified” neutrino flux, from the cosmological relic neutrinos produced $\sim$2 seconds after the Big Bang, with a black body temperature of 1.9 K, to the EeV neutrinos produced through the GZK mechanism. The search for a cumulative flux of high-energy neutrinos is one of the central goals of neutrino telescopes.

In general, no candidates for high-energy neutrinos are found by any experiment, with the remarkable exception of IceCube, as mentioned before [35]. When the amount of candidates is not enough to conclude a discovery, the measured flux of background, in this case high-energy cosmic rays, is used to derive upper bounds on the expected neutrino flux. [53, 54, 55]. In figure 1.8 are the upper limits from different experiments, including observations from the Pierre Auger Observatory [56], Anita-II [57] and IceCube [58].

1.8 Ultra-high energy photon searches

The production of ultra-high energy photons is expected from proton interaction off the cosmological microwave background (CMB). As it was mentioned before, the direct observation of ultra-high energy photons would automatically translate into some restrictions

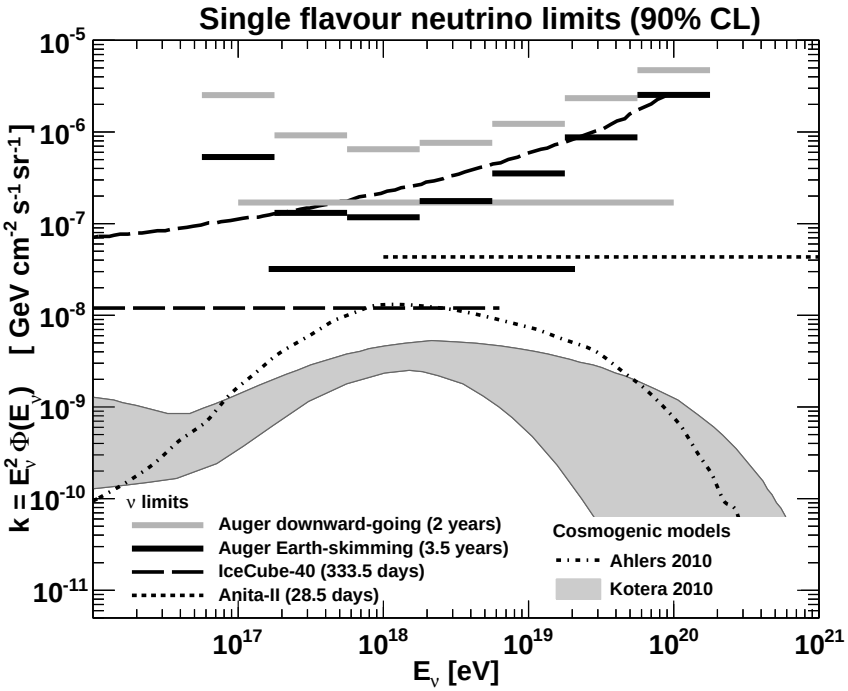


Figure 1.8: Differential and integrated upper limits at 90% C.L. to the diffuse flux of Ultra-high energy neutrinos.

in the mass composition of cosmic rays, favouring a light composition.

On the other hand, photons are expected also as a prediction of the so-called top-down models. This is a generic name that refers to every theoretical model in which ultra-high energy cosmic rays are produced as decay products of some superheavy particle. This theoretical particle can be either metastable or emitted by topological defects. Clearly, such a particle would be a Dark Matter candidate, and thus the canonical interpretation of top-down model is sometimes called SHDM, standing for Super Heavy Dark Matter. However, this kind of models are highly unfavoured by several experimental results, such as the non-observation of ultra-high energy photons or neutrinos, the spectral index in the highest energy region or the signal of correlation with astrophysical objects [59].

No ultra-high energy photon detection has been reported so far. However, upper limits on the fraction of photons in ultra-high energy cosmic rays have been set by various experiments (see figure 1.9 for comparison) [60], including the Pierre Auger Observatory [61, 62], AGASA [63, 64], Yakutsk [65], AGASA-Yakutsk combination [66] and Haverah Park [67, 68]. These limits are already below the predictions of top-down models, favouring models in which ultra-high energy cosmic rays are generated by acceleration of less energetic particles. As a contrast to top-down scenarios, these theories are sometimes called bottom-up models.

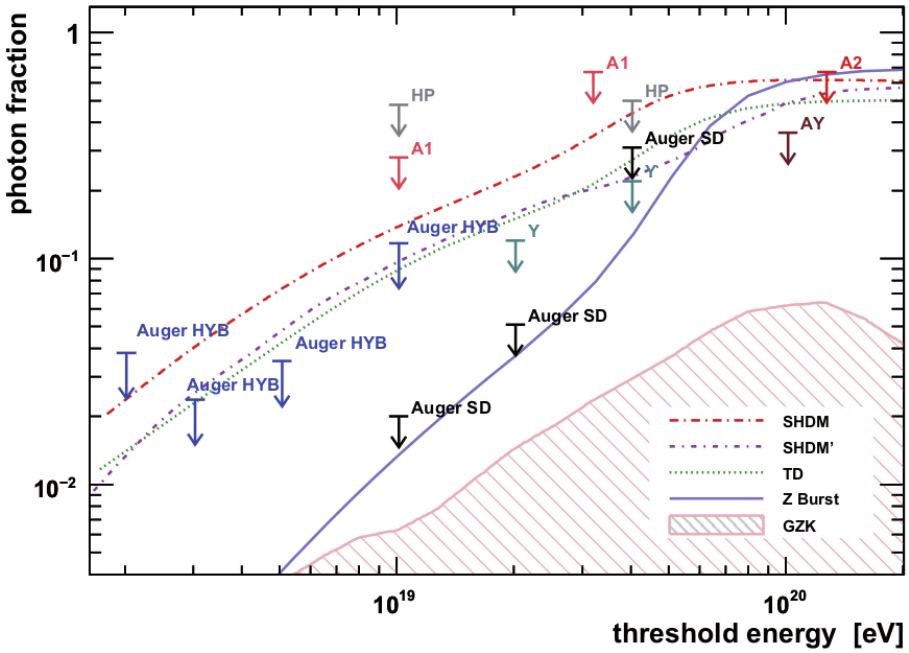


Figure 1.9: Integrated fraction of photons as a function of threshold energy. Experimental limits are shown for Auger hybrid and SD (Auger HYB/SD), AGASA (A1, A2), AGASA-Yakutsk (AY), Yakutsk (Y) and Haverah Park (HP) observations.

1.9 Exotic particle searches

Although cosmic ray experiments lack the precision of colliders, mostly due to uncertainties in high-energy extrapolations of hadronic interactions, they probe energies unexplored in laboratory. This makes cosmic rays the only available window to explore some exotic scenarios.

1.9.1 Magnetic monopoles

Magnetic monopoles are predicted by some grand unification theories (GUT). Their mass is expected to be in the range of 10^8 to 10^{17} GeV. Given their large mass, monopoles are generally thought to be relics of early Universe in GUT models, and therefore they are likely to have encountered enough accelerators during their lifetime to reach energies up to 10^{14} GeV.

A Cherenkov detector can detect magnetic monopoles travelling through it at velocities above the Cherenkov threshold ($\beta > 0.76$). Their signal would be clear, as the radiation they would emit in a Cherenkov detector is proportional to $(gn)^2$, where n is the refraction index of the medium and g is the magnetic charge of the monopole $g = e/2\alpha \approx 68.5e$, meaning they would emit almost 8000 times more light than a single muon.

The more stringent constraint to a magnetic monopole flux so far is given by IceCube [69], and is of the order of 10^{-18} monopoles per square centimetre per second per steradian.

1.9.2 Antinuclei

Antiprotons have been found in cosmic rays of low energy. However, no antiparticles with $|Z| > 2$ have been detected so far. In the standard picture of Cosmology, this asymmetry was caused when the symmetry between matter and antimatter broke, just after the Big Bang. However, the existence of heavier antiparticles in cosmic rays is not theoretically excluded. If ever found, they would be yet another relic from the Big Bang.

The BESS Collaboration has tried to detect antinuclei in cosmic rays over more than 20 years [70]. This detector consists mainly of a magnetic spectrometer to identify the charge of traversing particles, which would tag antimatter unequivocally. The current upper limits indicate that there is less than one antihelium nucleus per 10^7 helium nuclei in the Universe.

1.9.3 Dark matter searches

According to the current standard model of Cosmology, dark matter (DM) accounts for 27% of the total mass-energy content of the Universe [71]. However, despite tremendous efforts carried out during the last years in trying to detect dark matter both directly and indirectly, its nature remains mysterious.

It is normally assumed that the dark matter particles should be weakly interacting massive particles (WIMPs), that were produced in the early Universe. Given their extremely low rate of interaction, they must be either stable or extremely long-living, and of non-baryonic nature [72]. However, through self-annihilation, WIMPs may produce some standard particles that might reach Earth. Thus, it might be possible to connect the detection of very energetic photons to a signature of dark matter annihilation in over-density sites like satellite galaxies, the Galactic centre or galaxy clusters. In these kind of indirect searches of dark matter annihilations, we can consider, to a first approximation, that gamma rays do not interact on their way to Earth, as all relevant sources are relatively close.

The most stringent limits to dark matter annihilation cross section were obtained by MAGIC, through a 30 hours observation of the satellite galaxy Segue 1 [73]. Current limit is

$$\langle \sigma v \rangle \sim 10^{-23} \text{ cm}^3/\text{s, for DM masses } \sim \text{few} \times 100 \text{ GeV} \tag{1.7}$$

where v is the relative velocity of DM particles.

For dark matter mass below a few hundred GeV, the best limits are obtained by Fermi-LAT [74].

$$\langle \sigma v \rangle \sim 10^{-25} \text{ cm}^3/\text{s, for DM masses } \sim 100 \text{ GeV} \tag{1.8}$$

Very likely, the ultimate sensitivity in DM searches using gamma-ray detectors has not been reached yet. It is expected that in the future these efforts will produce a higher sensitivity, especially in the upcoming Cherenkov Telescope Array (CTA) [75].

1.10 Cosmic rays and terrestrial phenomena

The investigation into the possible effects of cosmic rays on living organisms will also offer great interest
Victor F. Hess, Noble Lecture, 12th of December 1936

As we have discussed throughout this chapter, charged particles within a wide range of energies are reaching the Earth permanently. It is clear that these particles should affect a number of physical processes occurring in the Earth system [76]. For example, a large fraction of the energy carried by the primary is deposited in the atmosphere, consequently ionising the surrounding molecules and affecting both their chemistry and the global electric charge.

Also, particles reaching ground contribute to the background of naturally-occurring radiation. We know that life has evolved over the past few billion years under the presence of this radiation, so it might well have contributed to shape the fauna and flora as we know it now. There is growing evidence of an increase in the lightning rate with increasing flux of charged particles, and this might been relevant in the origin of life on Earth [77]. It is suggested that cosmic rays influence short term and long term variation in climate, including cloud formation, cloud coverage and atmospheric temperature [78], although confirming cosmic ray contribution to such effects will need further study.

In a time of universal deceit, telling the truth is a revolutionary act

George Orwell

2 The Pierre Auger Observatory

The Pierre Auger Observatory [2] is devoted to the measurement and analysis of UHECRs. It is located in the Pampa Amarilla, near Malargüe, a small town in the province of Mendoza (Argentina). The location was carefully chosen for a handful of reasons, including low atmospheric pollution, thanks to a relatively small human activity in the area, a remarkably flat landscape, and an altitude such that the maximum of the shower development is close to the ground, 1400 m above sea level, which corresponds to an atmospheric depth of about 875 g/cm^2 for vertical showers. Furthermore, the latitude of the site, 35° south, allows an excellent coverage of the southern hemisphere sky.

The observatory utilises two complementary techniques for the detection of secondary particles in air showers (see figure 2.1). The surface detector (SD) [79] consists of more than 1600 water Cherenkov tanks arranged in a triangular grid of 1.5 km spacing. These stations sample the particles arriving at ground, and hence they provide a description of the lateral distribution of air showers. Not relying on any special ambience light or weather conditions, the SD has a duty cycle of nearly 100%.

The fluorescence detector (FD) [80] consists of four buildings overlooking the SD array, each of them equipped with six telescopes. These telescopes detect the fluorescence light produced by the de-excitation of atmospheric nitrogen after the crossing of charged particles, and they describe the longitudinal development of air showers. Unfortunately, due to the faintness of the fluorescence light, the FD can only work on moonless and clear nights, and thus its duty cycle is reduced to about 13%.

The simultaneous use of these two techniques is one of the features that makes the Pierre Auger Observatory unique as a cosmic ray experiment, being the largest hybrid detector ever built. A particular set of the events are those registered simultaneously by both detectors, the so-called hybrid events. They have an excellent description of both the lateral and longitudinal development, and can provide excellent information. For example, the use of hybrid events can provide a model independent energy calibration of the SD. Hybrid events would be of much interest in the final sections of this thesis.

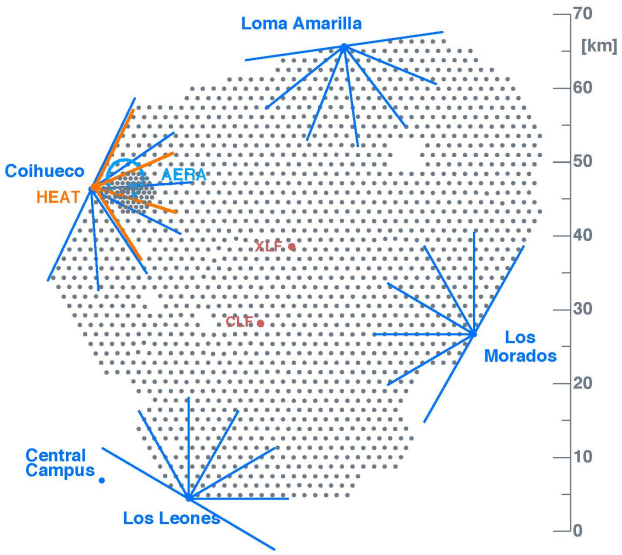


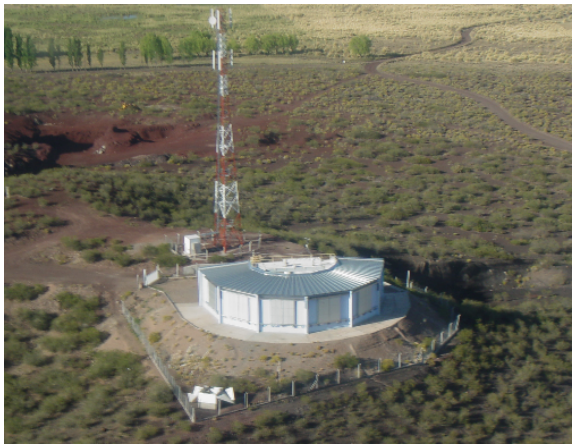
Figure 2.1: Layout of the Pierre Auger Observatory. Dots represent SD stations. Lines show the field of view of the FD telescopes.

In the last few years, a number of new enhancements have been added to the main detectors, mostly to study cosmic rays of lower energy (down to 10^{17} eV). These enhancements include a muon scintillator and water Cherenkov infill array (AMIGA) [81], three high elevation fluorescence telescopes (HEAT) [82], and an engineering radio array (AERA) [83]. These enhancements are in different states of development. For example, HEAT is now considered as a standard part of the Pierre Auger Observatory. However, as we will focus on the highest energy events throughout this thesis, we will fully dedicate the rest of the chapter to an in-depth description of the FD and the SD.

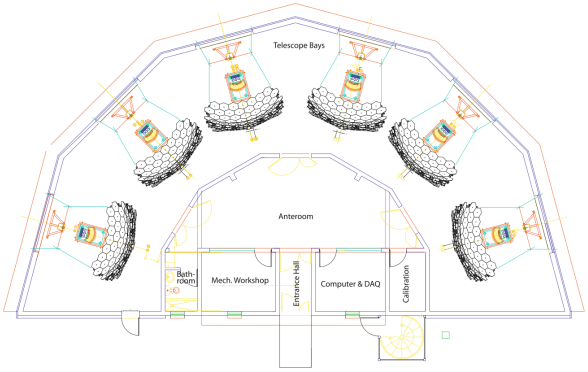
2.1 Fluorescence detector

When charged secondary particles from a cosmic ray shower traverse the atmosphere, they excite the electrons of nitrogen molecules. The de-excitation of these electrons is emitted as high energy photons, with discrete wavelengths between 300 and 400 nm, corresponding to the molecular band structure of nitrogen [84]. This light is detected by the fluorescence telescopes of the FD.

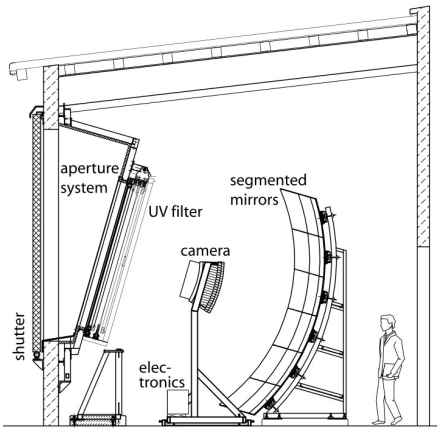
The 24 telescopes are located on four observation sites: Los Leones, Los Morados, Loma Amarilla and Coihueco (see figure 2.1), all of them situated on hills between 40 and 150 metres high surrounding the SD array, which is at an average altitude of 1400 metres above the sea level. Each telescope consists of a 1.7 m diameter diaphragm with a Schmidt optic corrector ring attached to the outer edge. The collected light is reflected by a 12 m^2 spherical mirror with a radius of curvature of 3.4 m and then focused into a camera comprised of 440 hexagonal PMTs in a 22 $\times$ 20 matrix (see figure 2.2). Each



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



(b) Schematic top view of an FD site.



(c) Schematic lateral view of an FD telescope.

Figure 2.2: View and schemes of an SD site of the Pierre Auger Observatory.

PMT has a diameter of 45 mm and a quantum efficiency of about 25%. The light collected by the PMTs is converted into an electrical signal, and finally digitised by fast analogue-to-digital converters (FADCs) into ADC bins.

The detection of the fluorescence light allows not only a description of the longitudinal development of the shower, but also a calorimetric measurement of the energy of the primary cosmic ray. However, in order to be able to make this transformation of ultraviolet light into deposited energy, a very accurate description of the atmosphere, as well as of the physical processes involved in the interaction of charged particles with air is required.

2.1.1 FD calibration

The calibration of the FD is performed following two complementary strategies, a “drum calibration” and a step by step sequence of dedicated calibration runs. The drum cali-

bration consists in the arrangement of a diffuse light source over each of the telescopes. The ratio of the detected to the emitted light is used as an absolute calibration to determine the response of the PMTs as a function of light intensity. This sort of calibration is performed only once every two or three years.

The second kind of calibration is performed at the beginning of every data-taking night and is divided in three steps:

- In the first step, Calibration A, a 375 nm wavelength light source is directed from the mirror centre onto the camera.
- In Calibration B, the same procedure is carried out with light going from the camera edges and onto the mirror.
- Finally, in Calibration C, the light is focused from behind the camera onto reflective targets on the inner shutter edges.

After the three calibrations have taken place, the total charge detected by the camera from each of them is compared to the drum calibration, and in this way a better knowledge of the telescope performance for each night is achieved. At the end of the data taking, Calibration A is repeated for stability checks.

2.1.2 FD Trigger

The fluorescence light produced by the charged particles of a cosmic ray shower, although always faint, can change widely depending on the energy of the primary. Also, the background conditions can change dramatically from one data-taking session to another, and therefore the FD electronics and data acquisition chain are designed so that a high degree of background rejection is achieved, while keeping a large dynamic range. In order to do that, different levels of trigger, both at the hardware and the reconstruction level, are defined.

Hardware triggers: First and Second Level Trigger

The First Level Trigger (FLT), or pixel trigger, is the heart of the digital front-end electronics. Among its tasks, it digitises the signals and measures the pixel trigger rate. The logic of the FLT consists on a running sum over the last n ADC bins, with $5 \leq n \leq 16$, which is compared to an adjustable threshold. A pixel trigger is generated if this sum exceeds the threshold, which is dynamically adjusted to keep the trigger rate close to 100 MHz. When the running sum is smaller than the threshold, the pixel trigger is extended for 5-30 μs, which increases the probability of coincident pixel triggers.

The Second Level Trigger (SLT), is designed to detect patterns of pixels that look similar to straight lines. The algorithm searches for track segments of at least five adjacent triggered pixels in any of the patterns shown in figure 2.3. Patterns obtained starting from these and after a rotation or a mirror reflection are also accepted.

However, in certain situations this criterion can be too restrictive. Sometimes, one of the PMTs may not pass the FLT, because it has collected less light than the threshold or

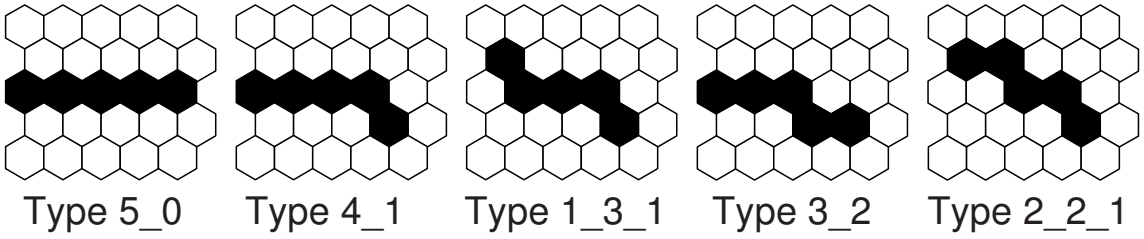


Figure 2.3: Basic patterns of triggered pixels considered by the Second Level Trigger in the FD.

because it is defective at the time of the measurement. This leads to tracks with a hole in the middle. To avoid this situation, the algorithm also accepts events in which only four pixels are triggered, giving rise to 108 different combinations of four-fold patterns from the five-pixel track segments of picture 2.3.

Software triggers: Third Level and Hybrid Trigger

The Third Level Trigger (TLT) is the first software trigger, and its purpose is to discard background events that survive the hardware triggers. These events are mainly caused by lightnings or by muon impacts on the PMTs, as well as by randomly triggered pixels.

In optimal data-taking conditions, only one or two events per minute and per telescope will survive SLT. In contrast, a lightning can trigger hundreds of pixels at the same time, in bursts of several tens of events per second. The TLT can reject lightning events reading the FLT multiplicities and applying cuts based on the time development of the multiplicity and its integral over the whole event. Its rejection power is large, as it removes approximately 99% of all lightning events.

Individual noisy channels far off the light track are most likely generated by fluctuations or atmospheric muon impacts. They can be rejected analysing the space-time correlation of the pixels. In global, the TLT is very efficient at discriminating real showers from background events, as the fraction of true showers that are rejected is below 0.7%.

Finally, once an event survives the TLT, the software merges coincident events from adjacent telescopes and sends a hybrid trigger to the central data acquisition system (CDAS). This trigger, called T3, acts as an external trigger for the SD, allowing to record events below 3×10^{18} eV, where the SD is not fully efficient. At these low energies, only one or two SD stations survive the trigger conditions, but this information together with the FD measurement is enough to ensure a high-quality hybrid reconstruction.

2.1.3 FD reconstruction

Geometry reconstruction

In the FD reconstruction, the firs step consists on the reconstruction of the geometry of the event. When a pixel of FD detector records a light pulse, not only the amount of light is stored, but also the arrival time of the signal. This temporal information is utilised to reconstruct the shower axis through a χ^2 minimisation procedure [80]. The

plane including the location of the detector and the shower axis (the shower-detector plane, SPD) is reconstructed (see fig 2.4). Experimentally, it is the plane passing through the detector that contains most of the pointing directions of the FD pixels centred on the shower axis.

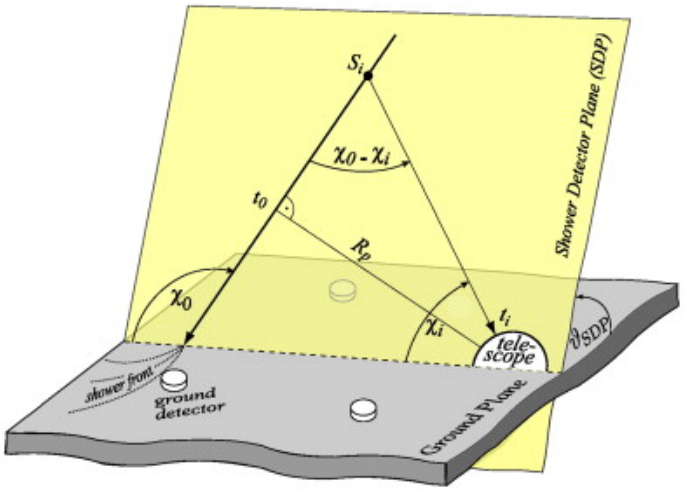


Figure 2.4: Shower-detector plane parameters.

By construction, the shower axis must lay within the SDP, and is defined by two parameters: R_p, the perpendicular distance from the camera to the track, and χ_0, the angle the track forms with the horizontal line in the SDP. The arrival time of the light at the pixel i, t_i is then:

$$t_i = t_0 + \frac{R_p}{c} \tan\left[(\chi_0 - \chi_i)/2\right] \tag{2.1}$$

where χ_i is the angle of the pointing direction of the pixel i with respect to the horizontal line, and t_0 the time at which the shower front on the axis crosses the point at distance R_p. The parameters R_p and χ_0 are obtained by the χ^2 minimisation mentioned above, in which this function is fitted to the data timing information.

Thanks to a fast timing electronics, the reconstruction using only an FD telescope, called monocular, is in general rather accurate. However, when the measured angular speed of the shower does not change much along the observed track-length, the set of parameters (R_p, χ_0) can have a degenerate solution, in which a family of possible geometries can be fitted to the data. In order to avoid this degeneracy, the FD timing information is used simultaneously with the one of SD stations, in what is called the hybrid reconstruction. Figure 2.5 shows how the reconstruction of the shower axis in the monocular mode finds a non optimal solution. Once the information from the SD is added (squares on the top left part of the figure), the hybrid reconstruction is clearly improved.

As the energy of the primary cosmic ray raises, there is an increasing probability for an event to be observed simultaneously by two or more FD sites (stereo, triple and four-fold events). These events are particularly interesting for the evaluation of resolution, as they offer the chance to observe an incoming shower from different directions, performing

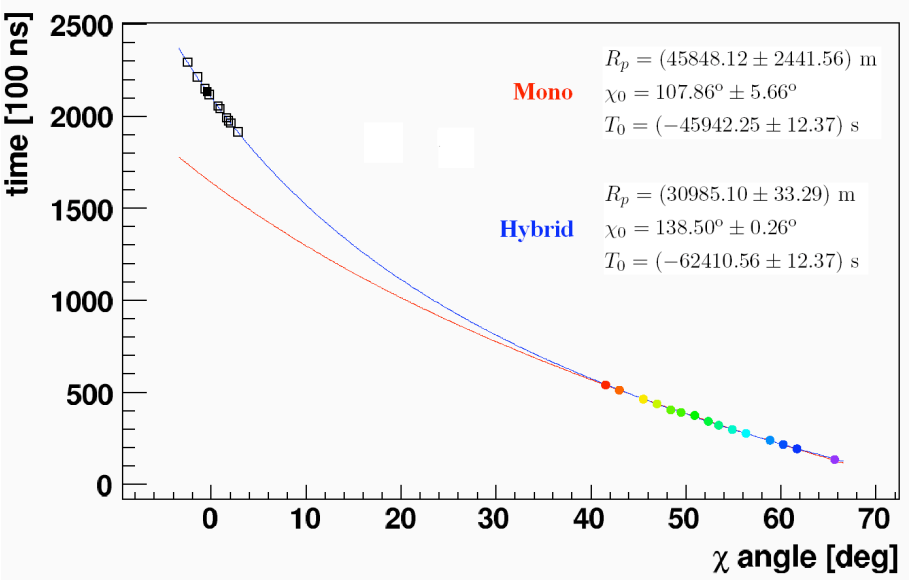


Figure 2.5: Example of a reconstructed shower axis in the monocular and hybrid modes.The addition of the SD stations (squares on the top left) significantly improves the resolution.

an independent hybrid reconstruction for each site. Alternatively, it is also possible to reconstruct the geometry of the shower in a multiple-eye mode by intersecting the different shower detector planes and without information coming from the SD. In the stereo mode, the resolution on the geometry depends mainly on the number of pixels used to find the SDP. A minimum of 6 pixels in each FD station is required to obtain a good reconstruction. The angular resolution of the FD is below 0.6° for events with energy above 3×10^{18} eV (see figure 2.6) [85].

Profile reconstruction

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

$$X = \int_z^\infty \rho\left(z'\right)\, dz' \tag{2.2}$$

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

In order to correctly determine the amount of deposited energy by the shower, the light attenuation from each source has to be estimated, and the different contributions to the produced light must be considered: fluorescence, multiple-scattering and Cherenkov (both direct and scattered). The energy deposited at slant depth X_i is then expressed as

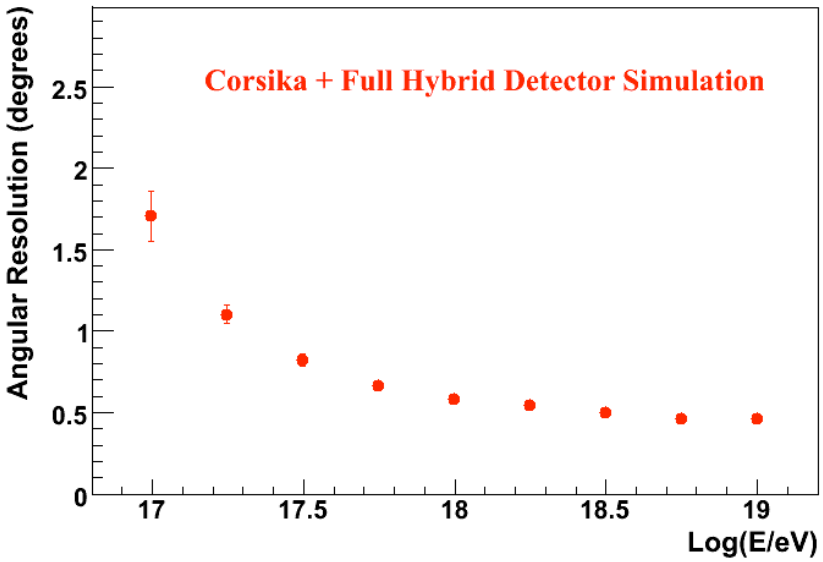


Figure 2.6: Angular resolution of hybrid events.

$$\frac{dE}{dX_i} = N_e\left(X_i\right) \int_0^\infty f_e\left(E, X_i\right) \frac{dE}{dX_e}\left(E, X_i\right)\ dE \tag{2.3}$$

where $f_e\left(E, X_i\right)$ is the normalised electron energy distribution, dE/dX_e the energy loss of a single electron with energy E and $N_e\left(X_i\right)$ represents the number of electrons and positrons above a certain constant energy cut-off [86].

The fluorescence detectors have a limited field of view, and therefore it is highly unlikely to observe the full shower profile. In practice, the shower development in the atmosphere is estimated by the fit of the observed profile to a Gaisser-Hillas function [87]

$$f_{\mathsf{GH}}\left(X\right) = \frac{dE}{dX_{\mathsf{max}}} \times \left(\frac{X - X_0}{X_{\mathsf{max}} - X_0}\right)^{(X_{\mathsf{max}} - X_0)/\lambda} \exp^{(X_{\mathsf{max}} - X)/\lambda} \tag{2.4}$$

where X_{max} represents the depth where the number of particles is maximum and dE/dX_{max} the maximum energy deposit. The shape parameters X_0 and λ do not have a direct physical interpretation, as the Gaisser-Hillas function is a phenomenological function. In fact, the preferred values for X_0 are generally negative. The best set of parameters can be obtained minimising the sum of the squared differences between the reconstructed energy deposit an the fitted value. If a large fraction of the shower profile above and below the maximum are observed, this minimisation converges. Otherwise, the experimental information is not enough to reconstruct all the Gaisser-Hillas parameters [87].

Energy reconstruction

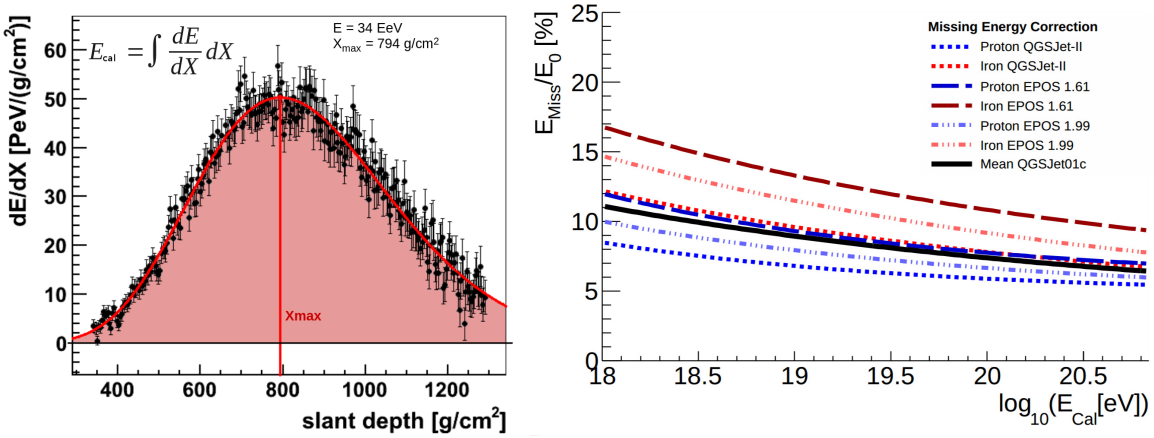
The calorimetric energy measured by the FD is given by the integral over the energy deposit profile (see figure 2.7a)

$$E_{\text{cal}} = \int_0^\infty f_{\text{GH}}(X)\ dX \tag{2.5}$$

However, the electromagnetic component of the shower does not account for the whole energy of the primary particle. The deposition of energy in the atmosphere for muons is negligible, and then they release most of their energy at ground. Also, neutrinos carry an amount of energy that escapes undetected. To take into account this missing energy, called *invisible energy* in this context, the calorimetric energy is multiplied by a correction factor f_{inv}

$$E_{\text{FD}} = E_{\text{cal}} + E_{\text{miss}} = f_{\text{inv}} E_{\text{cal}} \tag{2.6}$$

This correction, shown in figure 2.7b, is obtained from Monte Carlo simulations, and is dependent on energy and primary, mostly due to the energy dependence of mesons decay probability [88]. Above 10^{18} eV, the invisible energy is always below 17% of the total energy, with a decreasing contribution as the total energy grows.



(a) Example of FD profile and Gaisser-Hillas fit.

(b) Missing energy as a function of calorimetric energy.

Figure 2.7: Energy reconstruction in the FD. The calorimetric energy is obtained by the integration of the fit to the profile (left) and then corrected for the missing energy (right).

The resolution in the measurement of the energy achieved by the FD depends on the uncertainties associated to variations in the atmosphere, ranging from 4.5% at 3×10^{18} eV to 6.9% at 10^{20} eV, the invisible energy (1.5%) and the geometry reconstruction, which ranges from 5.2% to 3.3% for the same interval. The resulting overall energy resolution is almost constant with energy in the range $[3, 100]$ EeV, and lays between 7% and 8% [89].

2.1.4 Atmospheric monitoring

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

Atmospheric soundings have been performed periodically, using helium balloons with attached radiosondes. The information obtained by these devices about pressure, temperature and relative humidity were stored into detailed atmospheric models, and taken into account during the FD reconstruction. However, to account for rapidly-changing conditions and non-seasonal variations in the atmospheric parameters, the atmosphere above the detector is continuously monitored. The atmospheric monitoring in the Pierre Auger Observatory [90] is performed by four light detection and ranging (LIDAR) stations adjacent to each FD building, and also by the central laser facility (CLF), located near the centre of the SD array.

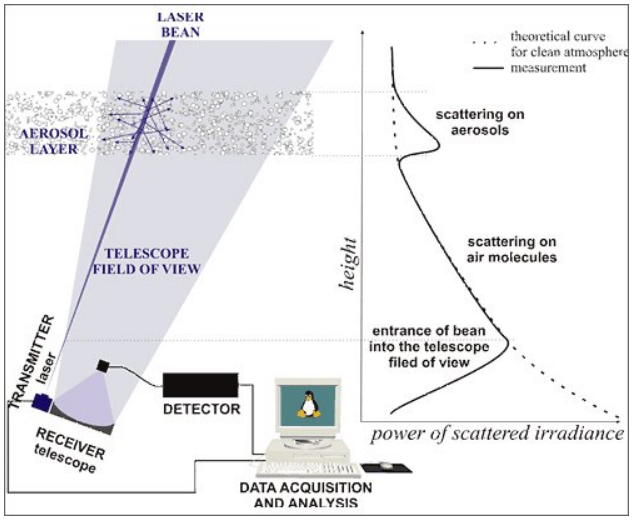
In the four LIDARs, a beam of ultra-violet laser light is directed into the atmosphere at periodic intervals (see figure 2.8a). The intensity and direction of the returning light collected by the LIDAR mirrors is used to measure the optical transmission conditions near the FD telescopes [91]. This allows the estimation of the aerosol content of the atmosphere. Infra-red cloud cameras and meteorological weather stations are also used to measure cloud coverage, humidity and other parameters in the vicinity of each FD site (see picture 2.8b).

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

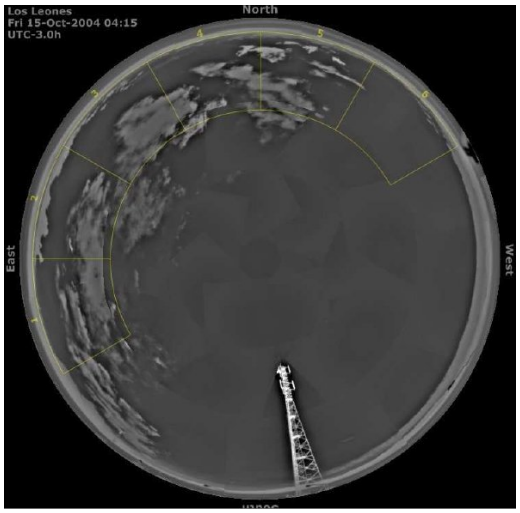
2.2 Surface detector

The vast area covered by the SD, more than 3000 km^2, was chosen to increase the number of detected events at the highest energies, with a grid spacing that, at the same time, provides a sufficiently large number of SD stations in the operating energy range. But of course, as the energy decreases, so does the number of particles reaching the SD, and as a consequence this detector is fully efficient only for events with energy above $10^{18.5}$ eV.

Each of the 1660 water Cherenkov detectors of the SD consists of a cylindrical polyethylene tank of 10 m^2 surface circular area and 1.2 m height, with three 9 inch (22.9 cm) photomultipliers (PMTs) overlooking 12 tonnes of purified water. This water is contained within the tank, inside a Tyvek® liner. The purpose of this liner is double: as a water-tight container to seal out external light, and to reflect Cherenkov light produced within the detector by traversing particles, given its high reflectivity (about 98% depending



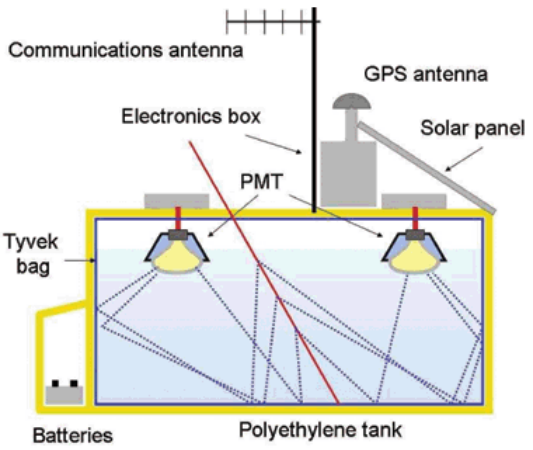
(a) Scheme of the LIDAR operation.



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

Figure 2.8: The LIDAR and cloud cameras are part of the atmospheric monitoring used in the Pierre Auger Observatory.

(a) Photo of an SD station deployed in the field.



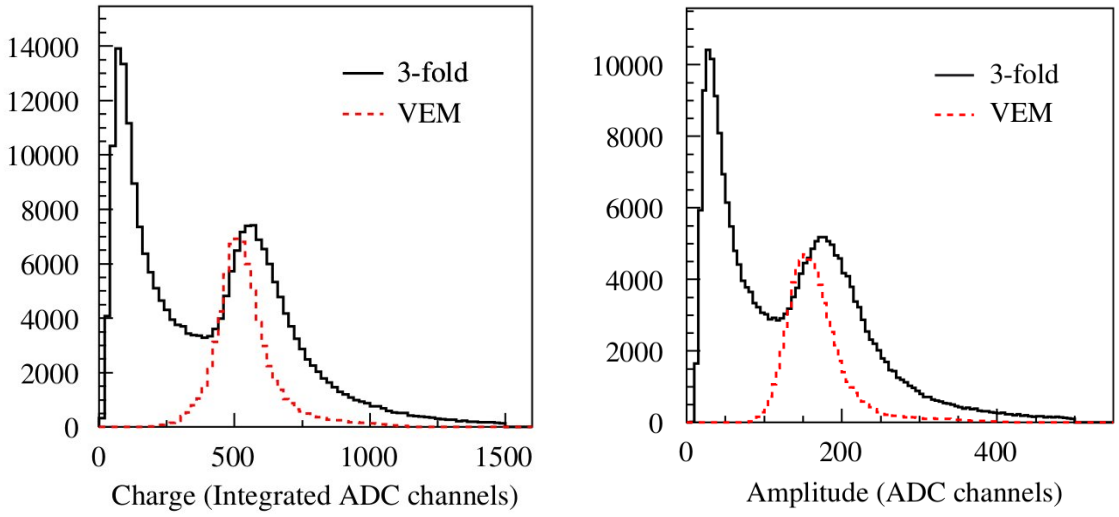
(b) Schematic view of the station components.

Figure 2.9: View and scheme of an SD station of the Pierre Auger Observatory.

on wavelength). The signal registered by the PMTs, both in the anode and last dynode are digitised using FADCs, with a sampling rate of 40 MHz.

Finally, each station is equipped with its own power supply, made up of a 12 V battery that gets charged by two solar panels installed on top of the tank. GPS units establish the timing of the FADC signals with a resolution of 7.2 ns. A radio antenna transmits signals to the closest FD site, where a communication tower sends the data to the CDAS. Signals are previously calibrated locally as discussed in the following section, and then selected by a hierarchical trigger system. A photography and schematic of a surface detector can be seen in picture 2.9.

2.2.1 SD Calibration



(a) VEM Charge.

(b) VEM Peak.

Figure 2.10: Charge and pulse height histograms for an SD station with a 3-fold trigger. The signal is the sum of the three PMTs. In the solid histogram the second peak is produced by vertical through-going atmospheric muons (VEMs), while the first peak is a trigger effect (see text). The dashed histogram is produced by vertical and central muons selected with an external muon telescope.

The main reason for the choice of a local calibration system relies on the available bandwidth for data transmission to the CDAS, 1200 bits per second, making remote calibration impossible. The most important quantity to calibrate an SD station is the average charge, Q_{VEM}, produced in a PMT by a vertical through-going muon traversing the station centrally [92]. This quantity is referred to as vertical-equivalent muon (VEM), and the signal registered by a particular station of the SD is generally expressed in VEM units.

However, in practice, the distribution of charge produced by atmospheric muons produces a peak on both the charge distribution, $Q^{\text{peak}}_{\text{VEM}}$, and the pulse height, $I^{\text{peak}}_{\text{VEM}}$, (see

figure 2.10) which are proportional to the expected values for a vertical muon. The peak $Q_{\text{VEM}}^{\text{peak}}$ is at approximately 1.09 VEM for the sum of the three PMTs, and at 1.03 VEM for a single PMT. This shift is the result of the asymmetry in the distribution of the track length of muons, as well as of defects in collecting the Cherenkov light. In a nutshell, the SD calibration is performed in three steps:

- Adjust the gain of the PMTs so that $I_{\text{VEM}}^{\text{peak}}$ is set in channel 50.
- Calibrate the electronics so that $I_{\text{VEM}}^{\text{peak}}$ is well determined.
- Determine the value of $Q_{\text{VEM}}^{\text{peak}}$, and use it to establish the conversion to VEM units.

The calibration constants $I_{\text{VEM}}^{\text{peak}}$ and $Q_{\text{VEM}}^{\text{peak}}$ are obtained with a 2% resolution every 60 seconds and sent to the CDAS together with every triggering event.

2.2.2 SD Trigger chain

Station triggers: T1, T2 and ToT

The first levels of trigger operate at a station level. The lowest trigger level, T1, works at hardware level and asks either for a coincidence in the three PMTs crossing the threshold of 1.75 VEM (Thr1) or a coincidence of two PMTs with more than 12 FADC bins above 0.2 VEM (above baseline) within a window of 120 time bins.

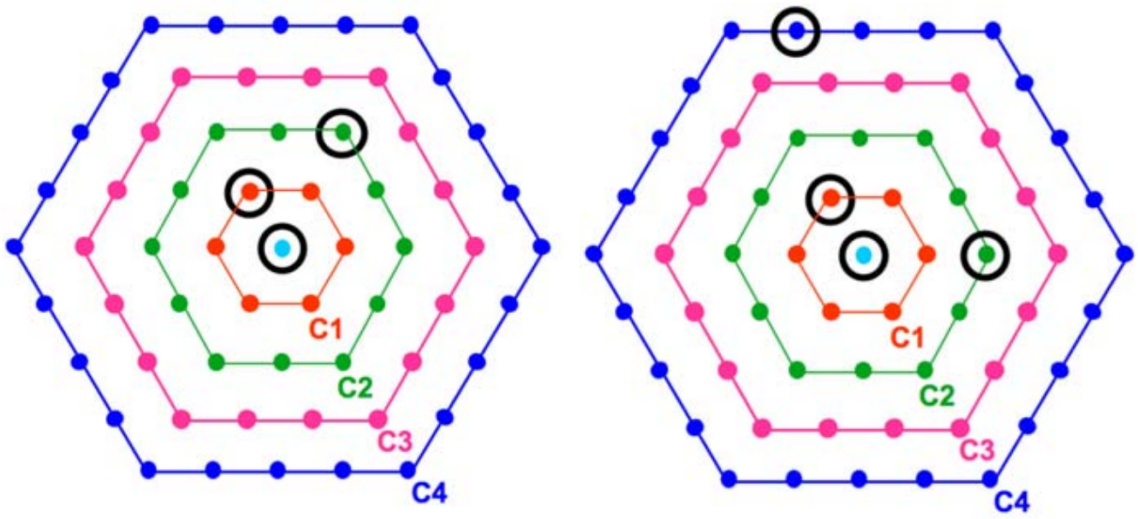
The second level, T2, is a software trigger, and similarly to the T1 level, two possibilities are considered, Thr2 and ToT:

- Threshold trigger Thr2 looks for a coincidence of the three PMTs above 3.2 VEM. The trigger rate is about 20 Hz.
- Time over threshold trigger ToT looks for a coincidence of at least two PMTs with more than 12 bins above 0.2 VEM within a time window of 120 bins. The trigger rate for the ToT is about 5 Hz.

The trigger system is organised hierarchically, and hence a T2 always implies the fulfilling of a T1. In general, ToT trigger performs better at distinguishing signals coming from air showers from background. In fact, for vertical showers, the ToT trigger is the most relevant condition. Signals fulfilling a T2 are sent to the CDAS, where all locally-triggered stations are combined and subject to event trigger conditions.

An improved version of the ToT trigger, the ToTd, was proposed as an optimisation especially suited for low energy events [93, 94]. The underlying idea of this trigger is the deconvolution of the traces using the average response to single peaks. Due to multiple reflections, diffuse light and the decay time of the electronics, single peaks appear in the traces with an exponential tail. As a consequence, a deconvolved background trace would be reduced to a single peak, which would not survive a ToT-like trigger, while a real signal would present several peaks, easily distinguishable from background. This trigger has been recently implemented and is about to finish the testing period.

CDAS trigger T3



(a) T3-3ToT example.

(b) T3-4C1 example.

Figure 2.11: The two possible minimal T3 configurations.

The T3 trigger is the first one at a CDAS level. It looks for time coincidences in the T2 signals, and tries to associate them to a real air shower. Two kind of patterns are taken into account, 3-fold and 4-fold:

- In the 3-fold scenario, T3 requires a coincidence of three neighbour stations with ToT trigger. The allowed time window considers the distance among stations in the following way: two stations have to be in the first two crowns around the first one considered (see figure 2.11a). If this criteria is fulfilled, the pattern is tagged as T3-3ToT.

- For the 4-fold case, four stations with T2 (Thr2 or ToT) have to be in coincidence. In this case, the distance requirement is looser, the fourth station being accepted if it is within four crowns around the reference station (see an example in figure 2.11b). This condition is only relevant for showers of large zenith angle, as nearly all vertical showers fulfil the T3-3ToT condition.

However, T3 trigger does not necessarily guarantee a relevant air shower. It was optimised as a compromise between selection efficiency and purity. The next level of trigger, T4, analyses geometry in a finer way, so that only physically relevant events survive. The trigger logic chain up to T3 level is summarised in figure 2.12.

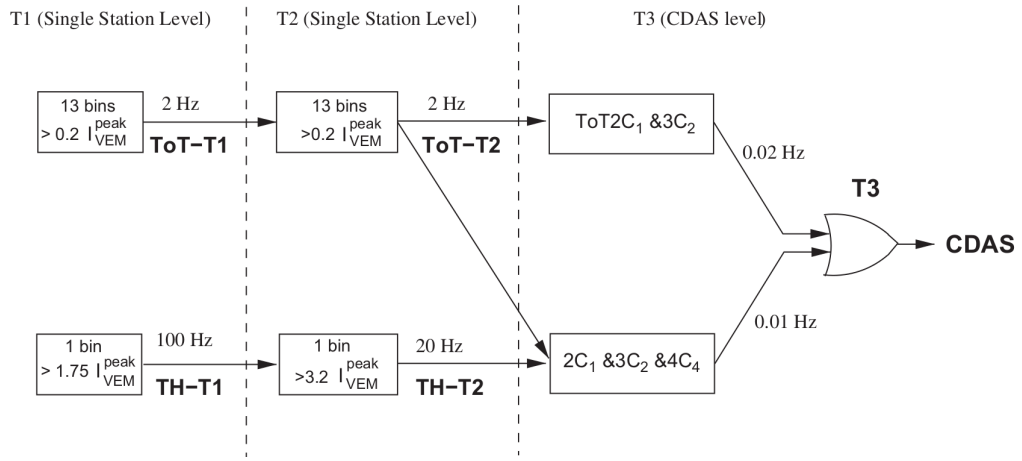
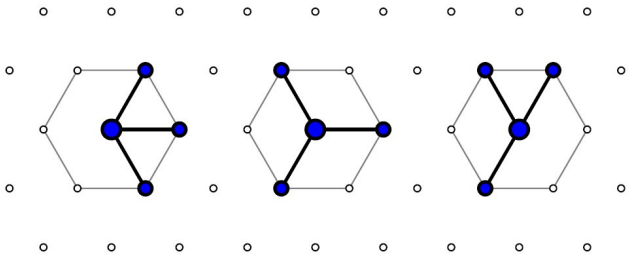
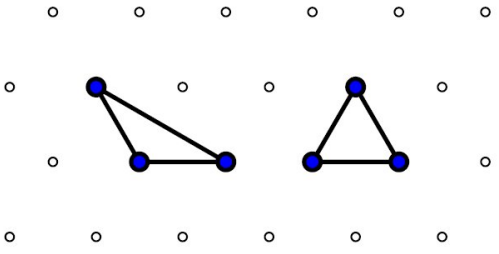


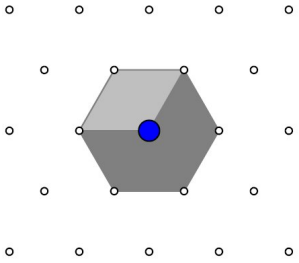
Figure 2.12: Diagram of the SD trigger logic chain in the Pierre Auger Observatory.



(a) Examples of T4-4C1 configurations.



(b) Examples of T4-3ToT configurations.



(c) 6T5 and 5T5 quality configurations.

Figure 2.13: T4 and T5 configurations. 2.13a: The three minimal compact configurations for the T4-4C1 trigger. 2.13b: The two minimal compact configurations for the T4-3ToT configuration. 2.13c: Example of the 6T5 hexagon (shadow) and the 5T5 hexagon (dark shadow).

The “Physics trigger” T4

The selection of physically relevant events is carried out by the T4 trigger, that in a way is a strict and more accurate version of T3. Again, there are two possibilities:

- T4-3ToT is a stricter version of T3-3ToT that asks for a non-aligned compact configuration of at last three neighbour stations with ToT (see figure 2.13b). About 99% of the vertical events are selected with this trigger condition.
- In the T4-4C1, four stations with any T2 trigger must be aligned in such a way that there are four stations in the first four crowns (figure 2.13a). Again, this condition is only important for nearly-horizontal air showers that would not survive the 3ToT.

Finally, a simple time compatibility criterion has to be fulfilled in every T4 event. The difference in their trace start time has to be smaller than d/c, being d the distance between two stations and c the speed of light in vacuum.

Quality selection trigger 6T5

The 6T5 trigger is the highest level of trigger in the Pierre Auger Observatory. Only T4 events survive to this point. The goal of this trigger is to reject events too close to the border of the array. As a general rule, these events have a poor reconstruction and should not be included in Physics analyses.

This trigger requires the station of the event with the largest signal to be surrounded at the time of triggering by six fully functional stations (not necessarily triggered). Another advantage of this trigger is that it keeps the computation of the detector acceptance simple, as it merely consists on the counting of active hexagons. A slightly less restricting version of this trigger is the 5T5, where only five out of the six stations in the hexagon are required to be functioning (see figure 2.13c). In the following we will only be interested in 6T5 events.

Nominal rate of events

The subsequent trigger levels described above aim for a good rejection rate of background events while keeping a high efficiency for interesting events. Table 2.1 summarises the approximate number of events that are detected by the Pierre Auger Observatory in a year. Also, the total number of events registered until December 2012 is presented.

Cut	Annual rate	Up to 2012
Total SD	$\sim$ 500000	3699622
Trigger 6T5	$\sim$ 430000	3217774
$E > 10^{18.5}$ eV	$\sim$ 21000	125133
$E > 10^{19}$ eV	$\sim$ 2300	13269
$E > 10^{19.5}$ eV	$\sim$ 160	1002
$E > 10^{20}$ eV	$\sim$ 1	7

Table 2.1: Approximate number of SD events per year.

The most energetic event registered so far was detected on the 25$^{\text{th}}$ of September 2008, with $\theta = 49.9 \pm 0.1^{\circ}$ and $E = (1.25 \pm 0.03) \times 10^{20}$ eV.

2.2.3 SD reconstruction

Once an event has been selected by the CDAS, its parameters of interest (e.g. energy, zenith and azimuth angles) are reconstructed offline using a dedicated software called $\overline{\text{Off}}\underline{\text{line}}$ [95]. The main feature of this software is its modularity. It was designed in such a fashion that the simulation of the detector and the reconstruction of events can be split in several steps, each of them coded in an individual module. Each module can read the information about the detector status and the event, process this information and write back the result into the event. This strategy is sketched in figure 2.14.

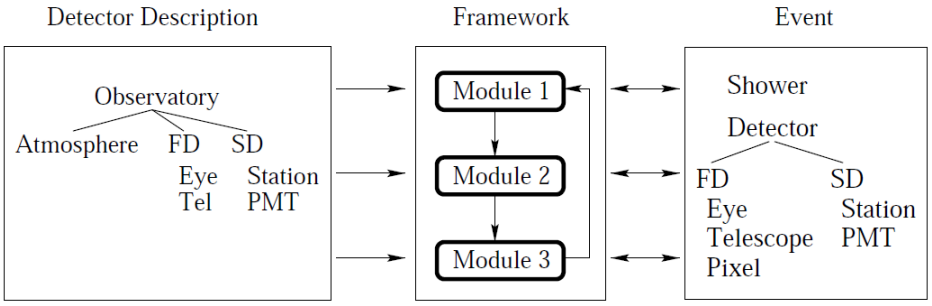


Figure 2.14: Schematic representation of the modularity in $\overline{\text{Off}}\underline{\text{line}}$.

In the following sections we will describe the reconstruction process carried out to obtain some of the parameters of interest in the Pierre Auger Observatory [96].

2.2.4 Event selection

Even if a station has passed all the criteria described in 2.2.2, it is not guaranteed that this station must be considered as part of the event. Accidental muons can trigger a station, and there is a non-zero probability for two events to fall so close in the array that they are seen as one by the CDAS. Moreover, some phenomena such as lightning or malfunctioning stations can distort the event reconstruction, and have to be taken into account. Additionally, 19 stations are located at only 11 m to the standard ones, the so-called *twin* stations. They are useful in resolution studies, but not considered in the final reconstruction of events.

These problems are addressed in the $\overline{\text{Off}}\underline{\text{line}}$ reconstruction, by determining the exact number of *candidate* stations, i.e. those which are used in the event reconstruction. For example, lightning events are seen in a station as a series of oscillations in the FADC traces of all three PMTs. If the total signal in the station is below 1000 FADC counts and crosses zero (after baseline removal) more than three times, the signal is considered as originated by a lightning and the station is removed from the candidate list. Also, for *twin* configurations, the station with the highest ID[1] (i.e. the one that was deployed last) is removed, keeping the standard on-grid station.

[1] ID is a numerical flag identifying every single tank. Represented by a cardinal number, they were assigned chronologically in terms of deployment date.

Finally, to reject *accidental* stations, a criterion based on time compatibility is used. For that matter, a rudimentary geometrical reconstruction is performed with a few stations and then used to reject stations which are incompatible in terms of timing, in an iterative process. In the first step, all stations passing the ToT trigger are considered in the selection of the three stations that will be used in the first approximation to the shower core. If the event is not of the type T4-3ToT, then also stations passing the 4C1 criterion are considered. From this set of stations, the elementary triangle with the largest sum of station signals is selected and used as a *seed* for the event core

$$s_i \in \{s_1, s_2, s_3\} \text{ such that } \sum_i s_i \text{ is maximum} \tag{2.7}$$

The *hottest* station, i.e. the one with the largest signal, is used as the local origin. Then, compatibility with planar shower front propagating at the speed of light is required[2]:

$$c(t_i - t_1) = -\hat{a}\left(\vec{x}_i^{\circ} - \vec{x}_1^{\circ}\right) \tag{2.8}$$

thus obtaining the provisional axis $\hat{a}$. Using this axis and the signal time of the hottest station, t_1, station start times are again checked for compatibility with a planar shower front arrival. The predicted shower time t_{sh} at a point $\vec{x}^{\circ}$ is

$$t_{\text{sh}}\left(\vec{x}^{\circ}\right) = t_1 - \hat{a}\left(\vec{x}^{\circ} - \vec{x}_1^{\circ}\right)/c \tag{2.9}$$

For each station, the time compatibility criterion is applied in terms of the difference of the actual start time and the predicted shower arrival time, i.e. the station delay

$$\Delta t_i = t_i - t_{\text{sh}}\left(\vec{x}_i^{\circ}\right) \tag{2.10}$$

accepting stations only if they satisfy the condition

$$-1000 \text{ ns} < \Delta t_i < 2000 \text{ ns} \tag{2.11}$$

Stations outside this interval are considered as *accidental* and removed from the candidate list. Additionally, stations that have no triggered neighbour stations within 1800 m or only one within 5000 m are considered as isolated and also rejected.

Finally, periods with known software or hardware malfunctioning are either rejected or get a special treatment. They are referred to as “bad periods” and stored in an up-to-date database available to the collaboration [97].

2.2.5 Geometry reconstruction

In the approximation of a plane shower front travelling at the speed of light (see figure 2.15), the projection of any shower track along the shower axis $\hat{a}$ must satisfy

$$-\hat{a}\left(\vec{x}^{\circ}(t) - \vec{b}^{\circ}\right) = c\,(t - t_0) \tag{2.12}$$

[2] Symbol $\vec{x}^{\circ}$ represents point x

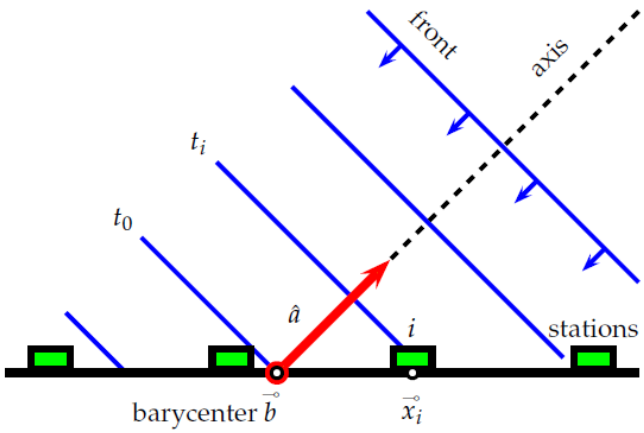


Figure 2.15: Schematic view of a plane shower front.

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

Inferring the time $t\left(\vec{x}\right)$ at which the shower front traverses a certain point on ground $\vec{x}$ consists only of transforming equation (2.12) into

$$ct\left(\vec{x}\right) = ct_0 - \left(\vec{x} - \vec{b}\right)\hat{a} \tag{2.13}$$

Let us assume now that the positions of the stations are known with absolute precision, and that the only deviations from these exact formulae come from the time uncertainty σ_t of the signal start time. Then the process of finding the adequate geometry is achieved by minimising the sum of squares of the deviations from the true values to the model over the uncertainties

$$\chi^2 = \sum_i \frac{\left[t_i - t\left(\vec{x_i}\right)\right]^2}{\sigma_{t_i}^2} = \sum_i \frac{\left[ct_i - ct_0 + \vec{x_i}\hat{a}\right]^2}{c^2\sigma_{t_i}^2} \tag{2.14}$$

with $\vec{x_i} = \vec{x_i} - \vec{b}$ the position[3] and t_i the signal start time of the station i, with variance $\sigma_{t_i}^2$. Writing the axis as $\hat{a} = (u, v, w)$, the station coordinates as $\vec{x_i} = (x_i, y_i, z_i)$ and $c\sigma_{t_i} = \sigma_i$ we are left with

$$\chi^2 = \sum_i \frac{\left[ct_i - ct_0 + x_i u + y_i v + z_i w\right]^2}{\sigma_i} \tag{2.15}$$

subject to the constraint

$$u^2 + v^2 + w^2 = 1 \tag{2.16}$$

[3]Symbol $\vec{x_i}$ represents the vector between two points

It is clear that equation (2.15) is not linear under this constraint, and in fact it can prove to be very difficult to solve. However, an approximate solution can be obtained using a linear model. If all stations lie close to the same plane, then $z_i \ll x_i, y_i$ and a linear approximation can be obtained by neglecting the vertical component.

This approximate solution can serve as a starting point to more elaborate fitting attempts. It is, however, subject to numerical instability if the z-projections of the station positions lay close to a straight line. For events with several stations, the occurrence of such a situation is highly improbable.

A more realistic realisation of the shower front is based on a curved front fit. The underlying idea is to extend the plane fit method with a parabolic term, $\rho \ll R_c$, that describes the curvature of the shower front near the impact point $\vec{c}$. Using $\vec{x} = \vec{x}^{\circ} - \vec{c}$, equation (2.12) is extended in the following way

$$ct\left(\vec{x}\right) = ct_0 - \hat{a}\vec{x} + \frac{\rho\left(\vec{x}\right)^2}{2R_c} \tag{2.17}$$

with perpendicular distance

$$\rho\left(\vec{x}\right)^2 = \left(\hat{a} \times \vec{x}\right)^2 = x^2 - \left(\hat{a}\vec{x}\right)^2 \tag{2.18}$$

Obtaining a first approximation to the radius of curvature R_c is necessary before performing any minimisation following this strategy. In order to do that, a slightly different model is considered. Within this model, the shower development is approximated as starting at a time t_0 from one single point (see figure 2.16) and propagating towards the stations, so the time t_i at the station i is

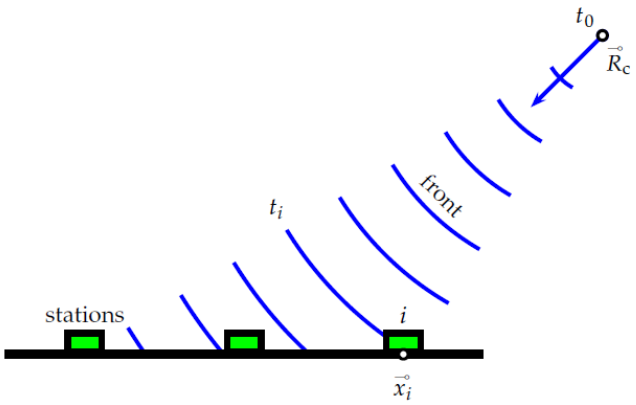


Figure 2.16: Schematic view of a spherical shower front.

$$c\left(t_i - t_0\right) = \left|\vec{R}_c - \vec{x}_i\right| \tag{2.19}$$

with $\vec{R}_c$ the apparent origin of the shower. Time propagation of the shower front is thus described as an expanding sphere. With this model, the timing information is clearly decoupled from any information on the impact point. The only relevant geometrical parameter of the spherical model is the apparent origin of the shower $\vec{R}_c$. The shower axis

is a derived quantity obtained only after the position of the impact point is known. The difference in the solid angle reconstruction between the plane-fit and curvature-fit axis $\hat{a}$ is of the order of half a degree.

The exact curvature fit of the geometry involves a 3D minimisation of the function

$$\chi^2 = \sum_i \frac{\left[c\left(t_i - t_0\right) - \left|R_c\hat{a} - \vec{x}_i\right|\right]^2}{c^2\sigma_{t_i}^2} \tag{2.20}$$

with accurate $z_i \neq 0$ treatment.

The difference between the approximate estimation of R_c and this exact model is of the order of 10 m, while the solid angle difference between the two axes is of the order of a few 0.1°.

2.2.6 LDF reconstruction

The lateral distribution function, LDF, characterises the dependence with the radial distance of the measured signal. In the Pierre Auger Observatory, this dependence is modelled as:

$$S\left(r\right) = S_{1000} f_{\mathsf{LDF}}\left(r\right) \tag{2.21}$$

where S_{1000} is the signal in VEM units estimated at 1000 m from the shower core, and $f_{\mathsf{LDF}}\left(r\right)$ is a particular normalised parametrisation, i.e. $f_{\mathsf{LDF}}\left(1000\text{ m}\right) \equiv 1$. The uncertainty on the signal of the stations, S, depends on the zenith angle θ as [98]:

$$f_S\left(\theta\right) = 0.32 + 0.42/\cos\theta \tag{2.22}$$

$$\sigma_S = f_S\left(\theta\right)\sqrt{S} \tag{2.23}$$

There are currently two parametrisations for the LDF used by the Pierre Auger Observatory collaboration:

The modified power-law LDF

$$f_{\mathsf{LDF}}\left(r\right) = \begin{cases} \left(\frac{r}{r_{1000}}\right)^{\beta+\gamma\ln\left(r_{300}/r_{1000}\right)} & \text{, if } r < r_{300} \\ \left(\frac{r}{r_{1000}}\right)^{\beta+\gamma\ln\left(r/r_{1000}\right)} & \text{, otherwise} \end{cases} \tag{2.24}$$

with r_{300} = 300 m and r_{1000} = 1000 m. The initial estimates for β and γ depend on θ

$$\beta_0\left(\theta\right) = 0.7\arctan\left[6\left(0.65 - \cos\theta\right)\right] - 3 \tag{2.25}$$

$$\gamma_0\left(\theta\right) = 0.05\sin\left[8\left(\cos\theta - 0.6\right)\right] - 0.5 \tag{2.26}$$

If needed, the slopes in higher reconstruction stages (more than three iterations without convergence) are fixed according to the following parametrisation [99]:

$$\begin{aligned}\beta(\theta) &= -4.73 - 0.519\log S_{1000} \\ &\quad + (1.32 + 0.405\log S_{1000})\sec\theta \\ &\quad + (-0.105 - 0.117\log S_{1000})\sec^2\theta\end{aligned} \tag{2.27}$$

The modified NKG LDF

$$f_{\mathsf{LDF}}(r) = \left(\frac{r}{r_{1000}}\right)^{\beta}\left(\frac{r + r_{700}}{r_{1000} + r_{700}}\right)^{\beta+\gamma} \tag{2.28}$$

with r_{700}= 700 m. Initial estimates for β and γ are obtained as [100]:

$$\beta_0 = 0.9\sec\theta - 3.3 \tag{2.29}$$

$$\gamma_0 = 0 \tag{2.30}$$

Again, if needed for convergence, the slope is fixed in higher stages:

$$\begin{aligned}\beta(\theta) &= -3.35 - 0.125\log S_{1000} \\ &\quad + (1.33 + 0.0324\log S_{1000})\sec\theta \\ &\quad + (-0.191 - 0.00573\log S_{1000})\sec^2\theta\end{aligned} \tag{2.31}$$

LDF fitting procedure

Once the list of candidate stations has been established, the final LDF is obtained using a log-likelihood maximisation. In order to achieve that, it is important to correctly estimate the uncertainty in the signal for each station. However, there is a plethora of different situations for an individual station, depending on the amount of collected signal. It can either have a small or large signal or, in more extreme scenarios, it can be saturated if the signal is too large or below trigger if it is too low.

Dealing with these subtleties is easier, from a statistical point of view, once the signal has been rewritten in terms of number of particles detected by the station. Strictly speaking, the energy deposit or equivalently the number of registered photo-electrons depends strongly on the particle type, injection point and incidence angle. Hence, no simple conversion from registered photons to number of detected particles is possible. However, for the purpose of LDF minimisation, a mere approximation is enough, and hence a simplified model is used, so that

$$n = pS \tag{2.32}$$

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

$$p = p(\theta) = \begin{cases} f_S(\theta)^{-2} & \text{, if } f_S(\theta) \geq 1 \\ 1 & \text{, otherwise} \end{cases} \tag{2.33}$$

independent of primary energy or distance to the core.

Within this approximation, the likelihood function to maximise can be split in four parts

$$L = \prod_i f_P(n_i, \mu_i) \prod_i f_G(n_i, \mu_i) \prod_i F_{\text{sat}}(n_i, \mu_i) \prod_i F_{\text{zero}}(n_i, \mu_i) \tag{2.34}$$

and thus the log-likelihood function is

$$\ell = \sum_i \ln f_P(n_i, \mu_i) + \sum_i \ln f_G(n_i, \mu_i) + \sum_i \ln F_{\text{sat}}(n_i, \mu_i) + \sum_i \ln F_{\text{zero}}(n_i, \mu_i) \tag{2.35}$$

with n_i the effective number of particles detected in the tank as obtained using equation (2.32), and μ_i the corresponding LDF expectation. The four members of this equation represent each of the four possibilities for a station status:

- **Small signals:** A Poissonian distribution is assumed

$$\ln f_P(n_i, \mu_i) = n_i \ln \mu_i - \mu_i - \sum_{j=1}^{n_i} \ln j \tag{2.36}$$

- **Large signals:** For signals larger than 15 VEM, the central limit theorem guarantees a Gaussian behaviour

$$\ln f_G(n_i, \mu_i) = -\frac{(n_i - \mu_i)^2}{2\sigma_i^2} - \ln \sigma_i - \frac{1}{2} \ln 2\pi \tag{2.37}$$

- **Saturated signal:** The saturated signal represents a lower limit to the actual signal. We integrate f_G over all possible values above the lower limit, to get an estimate of the probability of detecting a larger signal

$$F_{\text{sat}}(n_i, \mu_i) = \int_{n_i}^{\infty} f_G(n, \mu_i)\, dn = \frac{1}{2} \operatorname{erfc}\left(\frac{n_i - \mu_i}{\sqrt{2}\sigma_i}\right) \tag{2.38}$$

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

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

$$\ln F_{\text{zero}}(n_{\text{th}}, \mu_i) = -\mu_i + \ln \sum_{n=0}^{n_{\text{th}}} \frac{\mu_i^n}{n} \tag{2.39}$$

2.2.7 Fit stages

The previously described algorithms for geometric reconstruction are carried out in a step by step procedure, divided in five stages:

1. **Plane reconstruction:** The shower geometry is reconstructed within the plane shower front approximation .

2. **First estimation:** Using signal-weighted barycentre as the provisional core position and the seed values for β and γ, the S_{1000} normalisation factor of the LDF is estimated based on the signal of the station closest to 1000 m measured along the shower front plane.

3. **Fit for $\mathbf{S_{1000}}$ and core:** This stage performs S_{1000} and core location fitting. The core is parametrised only with the x and y components in the local coordinate system and z is assumed to be zero. This means that the core is allowed to float within the plane tangent to the reference ellipsoid that contains the barycentre. Typical distances to the signal-weighted barycentre amount to a few 100 m.

4. **Fit for β and γ:** Depending on the number of candidate stations, β and γ are gradually included as variable parameters that are subjected to the minimization process.

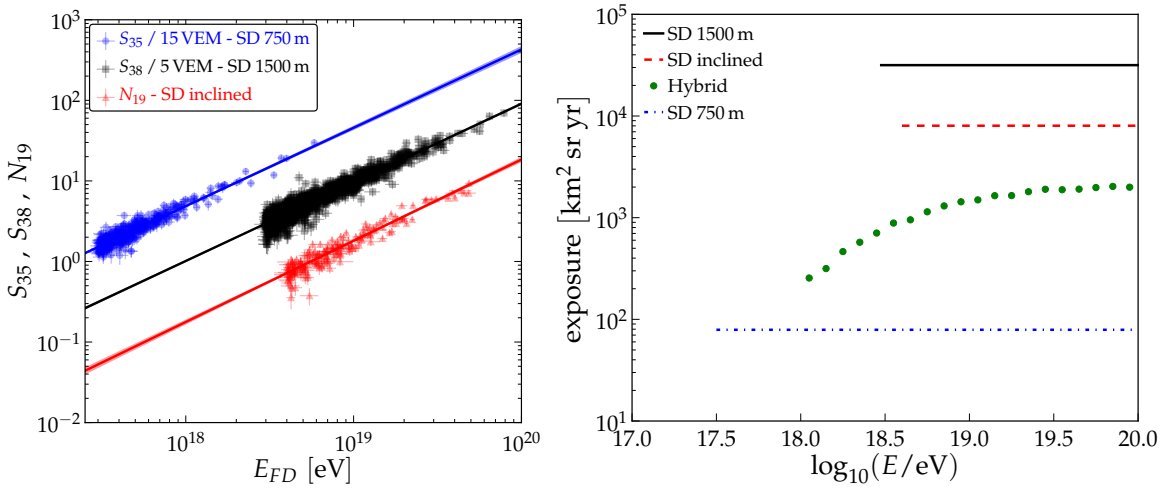
5. **Treatment of zero-signal stations:** Steps from stage 3 are repeated again. This time including zero-signal stations next to stations with signal above the threshold, assuming their respective signal did not exceed the local trigger thresholds. This introduces appropriate terms in the maximum likelihood equation and helps to stabilise the core position for awkward station configurations.

2.2.8 Energy reconstruction

As discussed in the previous section, the calorimetric energy of the FD is much larger than the invisible energy, and therefore the FD can estimate the energy of cosmic rays in an almost model-independent way. In contrast, for the SD there is no direct determination of the energy of the primary particle, E_{SD}, so the reconstruction of the energy must rely on the estimation of some parameter that correlates with the energy.

Different attenuation characteristics of the electromagnetic and muonic shower components lead to different reconstruction methods for different zenith angle ranges. In the following we distinguish between *vertical events* ($\theta < 60^\circ$) and *inclined events* ($60^\circ \leq \theta < 80^\circ$).

The energy reconstruction of vertical events is based on the estimation of the lateral distribution of secondary particles of an air shower reaching ground at an optimal distance to the shower core. The optimal distances are those at which, for a wide range of reasonable lateral distribution functions, the spread in this signal size predicted at that distance is a minimum. For the 1500 m and 750 m arrays the optimal distances, determined empirically, are 1000 m and 450 m respectively [101]. The signals $S(1000)$ and $S(450)$ are corrected for their zenith angle dependence due to air shower attenuation in



(a) S_{38}, S_{35} and N_{19} vs. E_{FD}.

(b) Exposure of the different datasets.

Figure 2.17: Energy calibration in the Pierre Auger Observatory. Correlation between S_{38}, S_{35} and N_{19} with the energy measured by the FD (left). Exposure of the different methods of measuring the energy (right).

the atmosphere with a Constant Intensity Cut (CIC) method [102]. The equivalent signal at median zenith angle of 38° (35°) is used to infer the energy for the 1500 m (750 m) array. Note that for the 750 m array, only events with zenith angle below 55° are accepted.

Inclined air-showers are characterised by the dominance of secondary muons at ground, as the electromagnetic component is largely absorbed in the large atmospheric depth traversed by the shower. The reconstruction of inclined events is based on the estimation of the relative muon content N_{19} with respect to a simulated proton shower with energy 10^{19} eV. Due to the limited exposure of the 750 m array only inclined events from the 1500 m array are included in the present analysis.

Events that have independently triggered the SD array and FD telescopes (called golden hybrid events) are used for the energy calibration of SD data. Only a sub-sample of events that pass strict quality and field of view selection cuts are used [89]. The relations between different energy estimators $\hat{E}$, i.e. S_{38}, S_{35}, N_{19}, and the energies reconstructed by the FD E_{FD} are well described by power-laws $E_{\mathrm{FD}} = A \cdot \hat{E}^B$. The calibration parameters are given in table 2.2 together with the number of golden hybrid events. The correlation between the different energy estimators and E_{FD} is shown in figure 2.17a.

Energy resolution

The resolution of the energy for vertical events is directly related to the uncertainty in the determination of $S(1000)$, and also includes the uncertainty in the lateral distribution functions and the shower-to-shower fluctuations. It starts at 16% at 3×10^{18} eV and improves as the energy grows down to 12% at 10^{20} eV (see figure 2.18) [102]. For the infill array, the resolution goes from 23% at 1.3×10^{17} eV to 12% at 1.4×10^{18} eV

	Auger SD			Auger Hybrid
	1500m vertical	**1500m inclined**	**750m vertical**	
Data taking period	01/2004 - 12/2012	01/2004 - 12/2012	08/2008 - 12/2012	11/2005 - 12/2012
Exposure [km^2 sr yr]	31645±950	8027±240	79±4	See figure 2.17b
Zenith angles [°]	0 - 60	62 - 80	0 - 55	0 - 60
Threshold energy E_{eff} [eV]	3×10^{18}	4×10^{18}	3×10^{17}	10^{18}
No. of events ($E > E_{\text{eff}}$)	82318	11074	29585	11155
No. of events (golden hybrids)	1475	175	414	-
Energy calibration (A) [EeV]	0.190±0.005	5.61±0.1	$(1.21 \pm 0.07) \cdot 10^{-2}$	-
Energy calibration (B)	1.025±0.007	0.985±0.02	1.03±0.02	-

Table 2.2: Summary of the experimental parameters regarding energy measurement in the Pierre Auger Observatory.

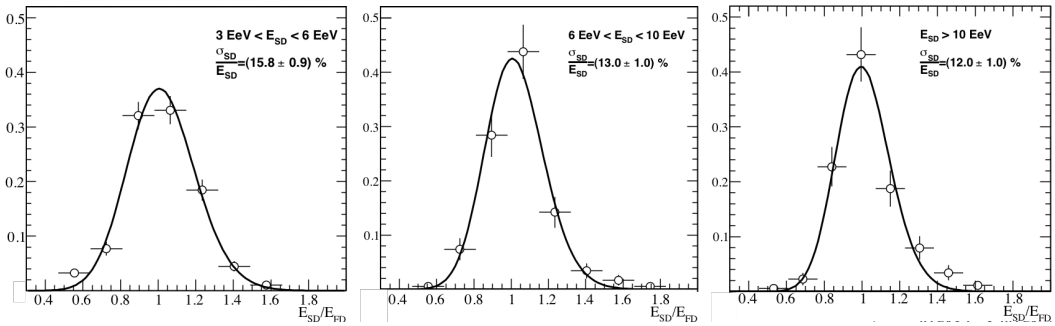


Figure 2.18: Ratio $E_{\text{SD}}/E_{\text{FD}}$ for various ranges of energy

[89]. For inclined events, the resolution mainly depends on the fit of the predicted muon signals at ground to the measured tank signals, and goes from 20% at 4×10^{18} eV to 8% at 6.3×10^{19} eV. [89]. Finally, the systematic uncertainties in the determination of E_{FD} have to be propagated to the determination of E_{SD} in order to obtain the systematic uncertainty of the energy scale (see figure 2.19). The systematics in the determination of the fluorescence yield (3.6%), the aerosol profile (∼5%), the FD calibration (10%), the FD profile reconstruction (∼6%) and the invisible energy (∼3%) contribute to an uncertainty of about 13%. When this value is combined with the uncertainty of the calibration fit (∼1%) and its stability over time (5%), the resulting systematic uncertainty of the energy scale ends up being 14% [89].

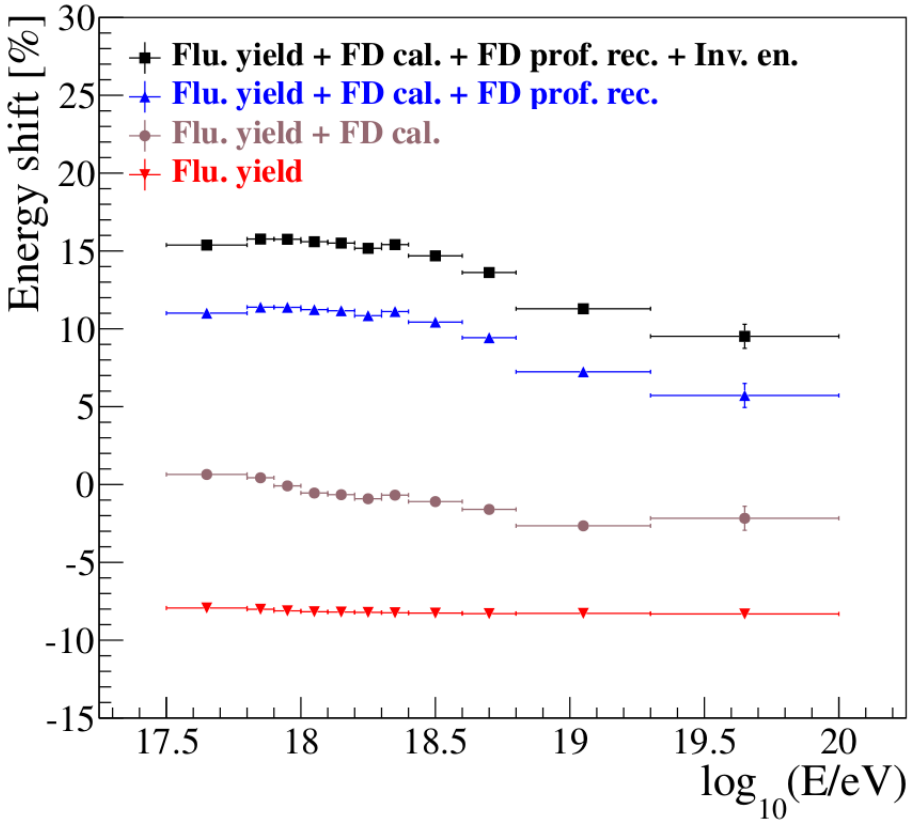


Figure 2.19: Cumulative energy shift as a function of the shower energy when different effects are considered.

La mayor parte de los hechos son inobservables, por lo cual hay que inventar indicadores

Mario Bunge

3

Mass measurements in extensive air showers

The determination of the mass composition of cosmic rays with energies below the TeV scale can be performed directly by satellite experiments (see e.g. [103, 104, 105]). However, at higher energies the cosmic ray flux becomes too scarce, and this strategy is no longer feasible.

Here we will show how, with the careful study of the characteristics of extensive air showers, some insight on the composition of the cosmic ray primary can be obtained through the produced secondary particles reaching ground.

3.1 Extensive air showers

When a cosmic ray hits the atmosphere, the interaction with oxygen or nitrogen nuclei produces a highly energetic inelastic collision. The products of such a collision have enough energy to interact again and generate an even larger set of secondary particles and so on. This process can go on for a few generations, until a vast number of particles (about 10^{10} particles for events of 10^{19} eV) are produced. This mechanism of particle production gives rise to what are called extensive air showers (EASs). These showers of particles can produce a footstep on ground covering a few kilometres, all of them generated by a single primary cosmic ray. The discovery of EASs was made by Pierre Auger and colleagues in 1939 [106], and for that reason the Pierre Auger Observatory was named after him.

Extensive air showers can be described as the superposition of different components (see figure 3.1). The most important components of a shower are the hadronic, muonic and electromagnetic cascades. Other less important components include ultra-violet photons (both from Cherenkov and fluorescence radiation), radio emission and particles contributing to the invisible energy (neutrinos and very low energy particles).

After the first interaction, most of the produced secondary particles are pions, photons

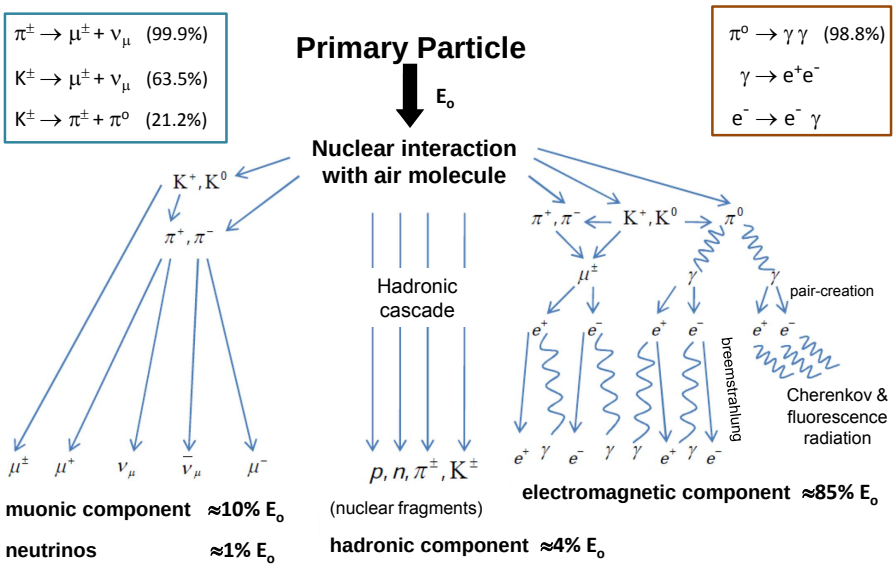


Figure 3.1: Main components of extensive air showers.

and electrons (or positrons). Photons, electrons and positrons initiate the electromagnetic shower. Electromagnetic particles are, by a few orders of magnitude, the most abundant component of the shower, and carry the largest fraction of the total energy. Neutral pions decay into two photons with a branching ratio of 98.8% [107] or to a combination of electrons and photons and contribute to the electromagnetic shower. Charged pions, on the other hand, may interact and produce new secondary particles that will feed the hadronic shower or decay into muons, producing the muonic shower. Hence, the muonic shower inherits some of the characteristics of the hadronic shower.

Although the full description of the evolution of showers can be fairly complicated, it is dominated by electromagnetic processes [108] and can be phenomenologically described by the simplified model developed by Heitler [109]. This model was extended by Matthews [110] to include the description of hadronic showers.

3.1.1 Heitler model of electromagnetic showers

The toy model introduced by Heitler [109], although highly simplified, accounts for some of the main features of the electromagnetic shower development. Hence, it has been very successful for many years as an intuitive introduction to the phenomenology of the development of cosmic rays showers through the atmosphere.

The Heitler model is based on the features of pair production and *bremsstrahlung*. Electrons, positrons and photons interact after travelling an interaction length $d = 2 \ln \lambda_r$, where λ_r is the radiation length of the medium ($\lambda_r = 37$ g/cm^2 in air). After each step, electrons and photons radiate a *bremsstrahlung* photon, and photons produce a pair

e^+e^-. Therefore, after every interaction there are two particles in the final state, each of them carrying one half of the energy. This splitting process goes on until the energy of the resulting particles is below the critical energy $E_c^\gamma \sim 80$ MeV in air, where electrons and positrons lose their energy via ionisation of the nuclei in the atmosphere.

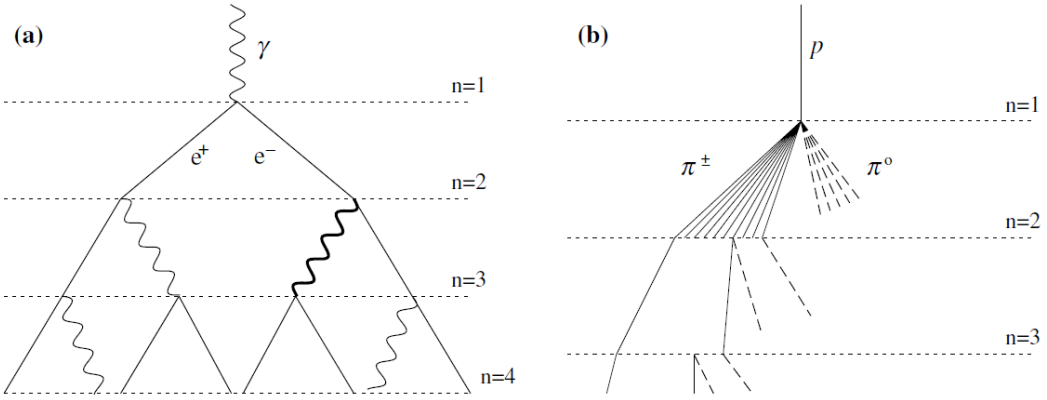


Figure 3.2: Schematic representation of the Heitler model for electromagnetic (a) and hadronic (b) cascade development.

After one interaction length d the shower consists of two particles with half the energy of the primary $E_1 = E_0/2$. After n interaction lengths, there are $N_n = 2^n$ particles in the shower, each of them carrying an energy of $E_n = E_0/N_n$ and the total depth traversed through the atmosphere is just $X = nd$. The whole process is schematically shown in figure 3.2a. Although the Heitler model does not fully describe in detail electromagnetic showers, it correctly predicts some of its main features:

- **The number of particles at the shower maximum is proportional to the energy of the primary particle**. The maximum number of particles, N_{max} is reached right after the energy falls below the critical energy, and so

$$N_{\text{max}} = E_0/E_c^\gamma \tag{3.1}$$

- **The depth at which the number of particles is maximum grows with the logarithm of the energy of the primary particle**. The depth of maximum shower development, X_{max} is determined by the number of interactions needed to reach the critical energy:

$$n_{\text{max}} = \ln(N_{\text{max}})/\ln 2 = \ln(E_0/E_c^\gamma)/\ln 2 \tag{3.2}$$

$$X_{\text{max}} = X_0 + n_{\text{max}} d = X_0 + \lambda_r \ln(E_0/E_c^\gamma) \tag{3.3}$$

  where X_0 is the depth of the first interaction.

- **The elongation rate is proportional to the radiation length**. The elongation rate, D_{10} is defined as the evolution of X_{max} with the decimal logarithm of the primary energy:

$$D_{10} = \frac{dX_{\text{max}}}{d\log E_0} = \lambda_r \ln 10 \approx 85 \text{ g/cm}^2 \tag{3.4}$$

Detailed simulations of electromagnetic cascades confirm these results. However, the Heitler model overestimates the number of particles by a factor 2-3 and the ratio of electrons and positrons to photons by a factor 10-12. These discrepancies are mostly due to the fact that absorption of particles below the critical energy is not considered in the model [110]. Also, more than one photon can be emitted during *bremsstrahlung*. Finally, electrons lose their energy much faster than photons.

3.1.2 Extension of the Heitler model to hadronic showers

The Heitler model provides an intuitive picture of electromagnetic showers that, even if incomplete, describes their main features. It is of great interest, therefore, to extend this model to describe hadronic showers. This was done by Matthews [110] and is thoroughly described in [4].

In this extension, the most relevant parameter is the hadronic interaction length λ_I. The assumption is that after each step of thickness $\lambda_I \ln 2$ hadronic showers produce $2N_\pi$ charged pions and N_π neutral pions. The π^0 immediately decay into two photons [107] and contribute to the electromagnetic cascade, while the $\pi^\pm$ continue to interact deeper in the atmosphere (see figure 3.2b). The hadronic cascade grows in each step, feeding the electromagnetic part via the neutral pions, until the produced charged pions reach their critical energy, at which it is more likely for them to decay than to continue interacting. The remaining charged pions are assumed to decay into muons when they reach the critical energy.

Some of the simplifying assumptions of the model are considering the interaction length and the pion multiplicity ($3N_\pi$ per step) as energy independent, and that energy is equally shared by the secondary pions. In fact, λ_I is not constant, but it does not depend strongly on energy. Between 10 and 1000 GeV, the typical energy of pions in EASs, it can be considered as constant with a value $\lambda_I \approx 120$ g/cm^2 [4, 111]. The critical energy of pions, E_c^π, decreases very slowly with increasing primary energy, ranging from 30 to 10 GeV at primary energies between 10^{14} and 10^{17} eV. A constant value of $E_c^\pi = 20$ GeV is a good approximation for simulations [110]. Finally, the multiplicity increases with laboratory energy, in the range $N_\pi \approx [3, 14]$ between 10 and 10^4 GeV respectively. A constant value of $N_\pi = 10$ is usually assumed [111].

After n interactions, the total number of pions is $(3N_\pi)^n$, each of them carrying an energy of $(1/3)^n E_0$. At the critical energy, then

$$E_c^\pi = \frac{E_0}{(3N_\pi)^{n_c}} \tag{3.5}$$

and therefore

$$n_c = \frac{\ln\left(E_0/E_c^\pi\right)}{\ln\left(3N_\pi\right)} \tag{3.6}$$

To obtain the total number of muons in the shower, we assume that all charged pions decay into muons when they fall below their critical energy

$$N_\mu = \left(\frac{2}{3}N_\pi\right)^{n_c} \tag{3.7}$$

so plugging equation (3.6) into (3.7) we arrive at

$$\ln N_\mu = n_c \ln\left(\frac{2}{3}N_\pi\right) = \ln\left(\frac{E_0}{E_c^\pi}\right)\frac{\ln\left(\frac{2}{3}N_\pi\right)}{\ln N_\pi} = \beta \ln\left(\frac{E_0}{E_c^\pi}\right) \tag{3.8}$$

so finally

$$N_\mu = \left(\frac{E_0}{E_c^\pi}\right)^\beta \tag{3.9}$$

which means that the number of muons does not grow linearly with the energy, as does the number of electrons, but as a power-function. Moreover, the value of β depends on the average pion multiplicity and on the inelasticity of hadronic interactions. Detailed simulations give values of β around 0.9 [112].

Another prediction of the extension of Heitler model to hadronic showers is obtained by comparing equations (3.1) and (3.9)

$$\frac{N_{\mathsf{max}}}{N_\mu} = \left(\frac{E_0}{E_c^\gamma}\right)^{1-\beta}\left(\frac{E_c^\pi}{E_c^\gamma}\right)^\beta \sim O(1000) \text{ for } E_0 \sim 10^{19} \text{ eV} \tag{3.10}$$

which expresses the fact that the number of electrons is much larger that the number of muons present in the shower for UHECRs. As a consequence, the energy of the shower is carried essentially by the electromagnetic cascade. This fact is exploited by fluorescence detectors.

In terms of the elongation rate, the assumption made in [110] is to consider the effect of the hadronic cascade only on the first interaction. Therefore, for proton showers

$$D_{10}^p = D_{10}^\gamma + \frac{dX_0}{d\log E_0} \tag{3.11}$$

where D_{10}^γ is the elongation rate for electromagnetic showers and $X_0 = \lambda_I \ln 2$ the depth of the first interaction. Introducing a realistic parametrisation of the dependence of λ_I as a function of the energy, such as the one given in [112], the elongation rate is $D_{10}^p \approx 64$ g/cm^2. Moreover, since hadronic interaction models predict an approximately logarithmic decrease of λ_I with energy, D_{10}^p is approximately constant.

An important consequence of equation (3.11) is that, regardless of the particular parametrisation of λ_I that is chosen, it will always decrease with increasing energy, and thus the second term in (3.11) is always negative. Therefore, the elongation rate for electromagnetic showers is always bigger than the one for hadronic shower. This fact was stated by Linsley as the *Elongation Rate Theorem* [113].

3.2 The superposition model

The theoretical framework to describe the main differences between air showers initiated by different nuclei is the *superposition model*. In this model, a primary nucleus of mass A and energy E is described as the superposition of A nucleons of energy $E' = E/A$ (see e.g. [114]).

The superposition model is a simplification and cannot fully describe hadronic EASs, as it does not account for nuclear effects such as re-interaction in the target nucleus or nuclear fragmentation. In order to consider all of these processes and others, more realistic transport codes are used, such as CORSIKA [115], AIRES [116] or COSMOS [117], together with hadronic interaction models like EPOS [118], QGSJET-II [119] or SIBYLL [120] (see e.g. [121] for a comprehensive review of air shower simulations).

Even if the description of EASs phenomenology is limited, the superposition model provides a qualitative understanding of the basic features of showers initiated by different nuclei. For instance, some of the most important phenomena that are correctly described are:

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

$$X^A_{\mathsf{max}}(E_0) = X^p_{\mathsf{max}}(E_0/A) = X^p_{\mathsf{max}}(E_0) - \lambda_r A \tag{3.12}$$

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

$$N^A_\mu(E_0) = \sum_i^A N^p_\mu(E_0/A) = N^p_\mu(E_0)\, A^{1-\beta} \tag{3.13}$$

- **The elongation rate is the same regardless of the mass of the primary**

$$D^A_{10} = \frac{dX^A_{\mathsf{max}}}{d\log E_0} = \frac{d\left(X^p_{\mathsf{max}} - \lambda_r A\right)}{d\log E_0} = \frac{dX^p_{\mathsf{max}}}{d\log E_0} = D^p_{10} \tag{3.14}$$

3.3 Mass-sensitive parameters

As mentioned on the first chapter, the determination of the mass composition of cosmic rays is of great value for addressing the long standing conundrum of the origin of cosmic rays and the mechanisms of acceleration they are exposed to. Here we will describe the main observables that have been proposed to measure the composition of cosmic rays, paying special attention to those that will be used in this thesis. In general, all mass-sensitive parameters can be divided into two groups: those reflecting the different longitudinal development of showers and those concerning the larger muonic content of heavier primaries.

3.3.1 Longitudinal development

The features described in the previous section are utilised by EASs experiments to determine the mass composition of UHECRs. For example, if at a given energy the fractions of each nuclei of mass A_i are f_i then the average shower maximum would be

$$\langle X_{\mathsf{max}}\rangle \simeq \sum_i^A f_i \left\langle X^{A_i}_{\mathsf{max}}\right\rangle = \langle X^p_{\mathsf{max}}\rangle - D_p \langle \ln A\rangle \tag{3.15}$$

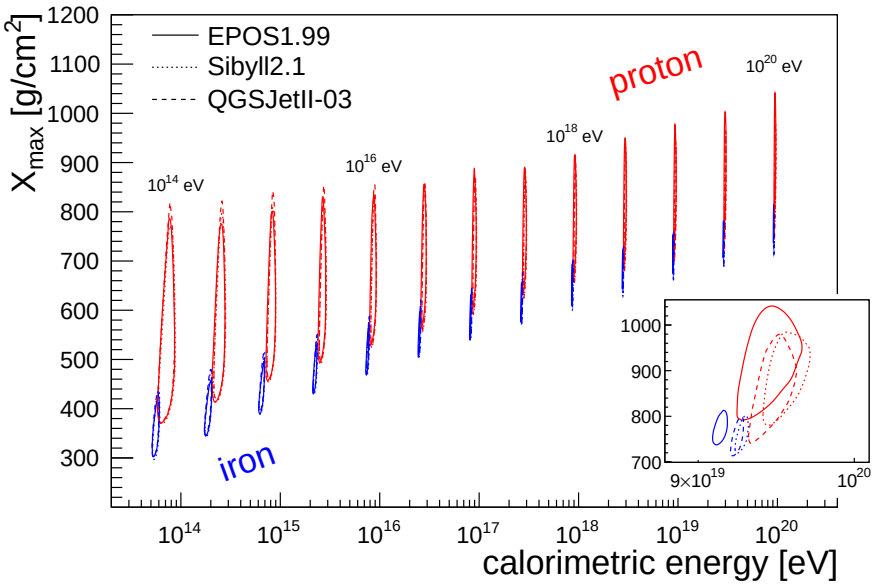


Figure 3.3: Air shower simulations of X_{max} as a function of the energy [114]. Contour lines represent regions containing 90% of the showers.

where $D_p = D_{10}^p / \ln 10$. This equation explicitly demonstrates the relation between the observed average of X_{max} and the average logarithmic mass of the cosmic ray composition

$$\langle \ln A \rangle = \sum_i^A f_i \ln A_i \tag{3.16}$$

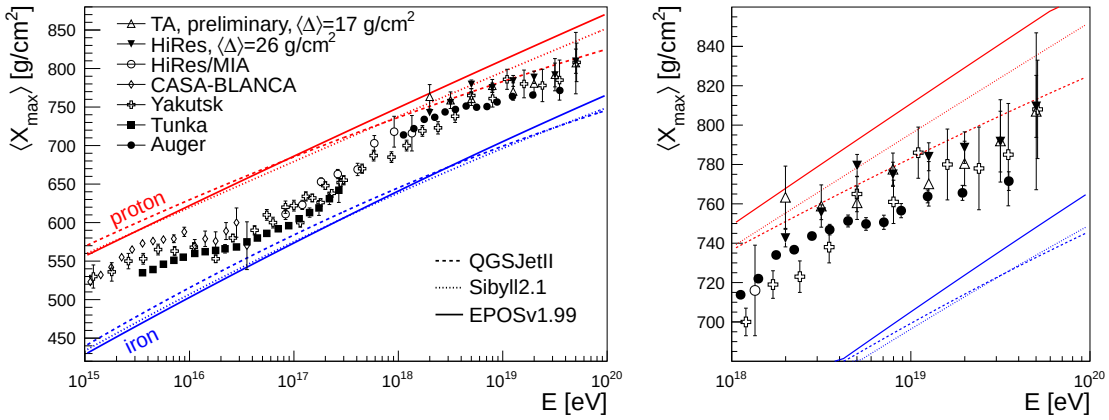


Figure 3.4: Measurements of $\langle X_{\text{max}} \rangle$ for different experiments compared to air shower simulations [114]. HiRes and TA data have been corrected with an anti-bias factor $\langle \Delta \rangle$ to compensate the detector effects. The right panel shows a zoom to the ultra-high energies.

As mentioned before, the numerical value of D_{10}^{p} is of the order of 60 g/cm^2, and therefore proton and iron induced showers are expected to differ by around $D_p\,(\ln 56 - \ln 1) \approx 100$ g/cm^2 (see figure 3.3). In consequence, X_{max} is a *mass-sensitive* parameter, i.e., contains information on the mass composition of UHECRs. Different experimental measurements of the evolution of $\langle X_{\mathrm{max}} \rangle$ with the primary energy are shown in figure 3.4. At the highest energies, all the experiments indicate a gradual increase of the average mass. However, the absolute values of $\langle X_{\mathrm{max}} \rangle$ show some discrepancies, although there is room for compatibility within uncertainties.

Besides the average value of X_{max}, it is interesting to consider the shower-to-shower fluctuations of this observable. The prediction from the superposition model would be that $\sigma_A = \sigma_p/\sqrt{A}$ where σ_A and σ_p are the fluctuations for nuclei of mass A and for protons respectively. Air shower simulations find, in fact, fluctuations that decrease with increasing mass number. However, this decrease is actually milder, going from about 60 g/cm^2 for protons to around 20 g/cm^2 for iron nuclei, reflecting the limited description of the superposition model. For a mixed composition

$$\sigma\,(X_{\mathrm{max}})^2 = \left\langle \sigma_i^2 \right\rangle + \left[\left\langle \langle X_{\mathrm{max}} \rangle_i^2 \right\rangle - \langle X_{\mathrm{max}} \rangle^2 \right] \tag{3.17}$$

where $\langle X_{\mathrm{max}} \rangle = \langle \langle X_{\mathrm{max}} \rangle_i \rangle$ is the mean of the combined distribution. If the composition is assumed to be described only by two components, with f the fraction of the first element, equation (3.17) can be rewritten as

$$\sigma\,(X_{\mathrm{max}})^2 = f\sigma_1^2 + (1-f)\sigma_2^2 + f(1-f)\left(\Delta \langle X_{\mathrm{max}} \rangle\right)^2 \tag{3.18}$$

Hence, depending on the separation between the two mean values, $\Delta \langle X_{\mathrm{max}} \rangle$ and the fraction f, the combined distribution can be actually broader than the individual distributions, and so the interpretation of $\sigma\,(X_{\mathrm{max}})$ in terms of atomic mass A is not as straightforward as in the case of $\langle X_{\mathrm{max}} \rangle$. Expressing equation (3.17) in terms of the average logarithmic mass

$$\sigma^2\,(X_{\mathrm{max}}) = \left\langle \sigma_i^2 \right\rangle + D_p^2 \left(\left\langle \ln^2 A \right\rangle - \langle \ln A \rangle^2 \right) \tag{3.19}$$

and so $\sigma^2\,(X_{\mathrm{max}})$ is proportional to the variance of $\ln A$. Therefore, the spread on X_{max} at a given energy contains information not only on the average composition, but also on the *purity* of the sample of UHECRs. Figure 3.5 shows the results for $\sigma\,(X_{\mathrm{max}})$ obtained by Yakutsk [122] and the Pierre Auger Observatory [123] compared to simulations. There is a clear disagreement between these two measurements, as the data from the Pierre Auger Observatory point to a significant fraction of heavy elements at the highest energies, while Yakutsk is compatible with a light composition at all energies.

3.3.2 Particles at ground

Another way of estimating the mass composition of cosmic rays deals with the measurement of particle densities at ground. One of the drawbacks of this approach is that only one stage on the shower development is sampled, and therefore measurements are more susceptible to shower-to-shower fluctuations. Nevertheless, ground measurements are

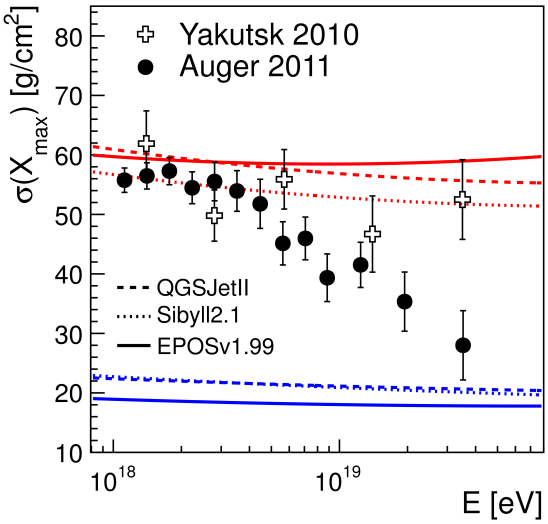


Figure 3.5: Shower-to-shower fluctuations of X_{max} compared to simulations for Yakutsk and Auger [114].

extensively used by cosmic ray experiments, given their large geometrical acceptance and high duty cycle.

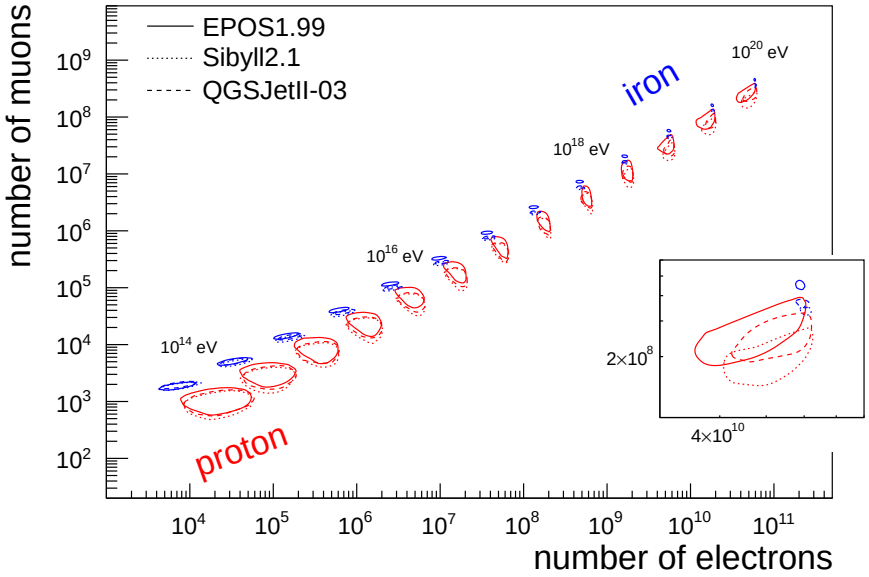


Figure 3.6: Number of muons vs. number of electrons at ground level for vertical showers observed at a $X_{\text{ground}} = 800$ g/cm^2 [114]. Contour lines represent regions containing 90% of the showers.

Equation (3.1) shows how the number of electrons at shower maximum is a good estimate of the primary energy, independently of its mass. On the other hand, from

equation (3.9), the number of muons can be used to infer the mass of the primary particle if the primary energy is known. Moreover, the evolution of the muon number with energy $dN_\mu/d\ln E$ has properties very much like those of the elongation rate. For example, for a constant composition this derivative gives $dN_\mu/d\ln E = \beta$ and deviations from this constant reflect changes on the average mass of the primaries.

However, although the number of electrons at shower maximum can be measured by fluorescence detectors, that is not the case for surface detectors, and the experimental situation is more complex. Indeed, surface detectors do not observe the number of electrons at shower maximum in general, but at a fixed depth $X \simeq X_{\text{ground}}/\cos\theta$ where X_{ground} is the vertical depth at ground level. Then, only an attenuated number of electrons is observed by ground detectors. Since heavy primaries develop higher in the atmosphere, the number of electrons is more attenuated than for light primaries, and as a consequence the number of electrons at ground also depends on composition (see figure 3.6). One way to deal with these attenuation problems is choosing the detector site so that its altitude is near the shower maximum. However, this has implications on the accessible range of energy, as the depth of the maximum grows with $\log E$.

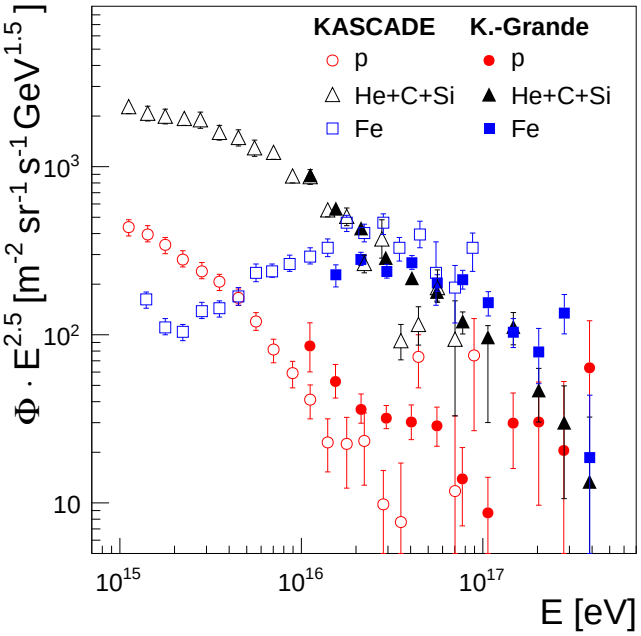


Figure 3.7: Unfolded fluxes from KASCADE and KASCADE-Grande [114].

The effect of the fluctuations on the number of electrons and muons can be properly taken into account using a two-dimensional unfolding technique, first introduced by the KASCADE Collaboration [23, 124]. This technique is based on finding a set of energy spectra of primary mass groups, such that their resulting simulated double differential distribution $d^2/dN_e dN_\mu$ minimises the difference to the real one. Figure 3.7 shows the results obtained for KASCADE and its extension KASCADE-Grande [24].

Steepness of the LDF

The LDF of an event contains information on its relative abundance and distribution of electrons and muons. As a consequence, it contains information about the mass composition of the primary particle. In general, showers initiated by heavy primaries reach their maximum higher in the atmosphere and will therefore have a flatter LDF than light primaries, as the secondary particles are more affected by deflections and scattering. This approach was follow e.g. in the reanalysis [125] of the Volcano Ranch [126] data, yielding an iron fraction of about 75% at 1 EeV for QGSJET-01. However, the sensitivity of the LDF to the primary mass is weaker than the one obtained by measuring the number of electrons and muons.

Rise-time

The rise-time is generally defined as the time it takes for the signal to rise from 10 to 50% of its integral value. This variable is sensitive to the muon-to-electron ratio of the global signal, as muons mostly arrive as a pack near the shower front, while electrons suffer larger attenuation and scattering in the atmosphere. Consequently, the rise-time contains information about the primary mass, as first stated by Bassi *et al.* [127].

Cherenkov detectors in the "early" part of the shower, i.e., those detecting the shower at a earlier stage of development than that of the shower core at ground, have faster rise-time than detectors in the "late" part, due to different attenuations of the electromagnetic component of the shower. This asymmetry, known as the *rise-time asymmetry* [128], disappears for vertical showers, but also for very inclined showers, as they are almost purely muonic. It was also shown that the zenith angle at which the rise-time asymmetry is maximum, θ_{max} is correlated with the shower development [128], and hence with the mass of the primary. The application of this method to real data [123] indicates a transition from light to heavy composition above 3 EeV.

3.3.3 Muon production depth

The muon production depth, MPD, aims for the description of the longitudinal development of the hadronic part of the shower based on measurements made at ground level. This is possible thanks to the model developed in [129, 130] and updated in [131] which basically describes how muons reaching ground retain, to a large extent, information about their production point. Muons are about 200 times heavier than electrons [107], and as a consequence their radiation length is much larger than the whole atmospheric depth, i.e., they behave as minimum ionising particles. Multiple scattering is also negligible for muons, and so it is reasonable to assume that muons travel following almost straight lines close to the shower axis.

If muons travel along straight lines, it is possible to pinpoint their point of origin using elementary geometry (see figure 3.8). The distance travelled by a muon is

$$l = \sqrt{r^2 + (z - \Delta)^2} \tag{3.20}$$

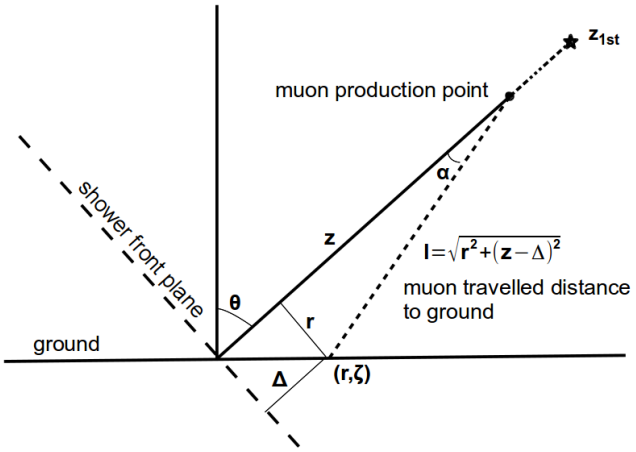


Figure 3.8: Geometry used to obtain the muon travelled distance.

where $\Delta\,(r,\zeta)$ is the distance from the point at ground to the shower front plane, defined as the plane front travelling at the speed of light perpendicularly to the shower axis. For a horizontal ground plane

$$\Delta = r\cos\zeta\tan\theta \tag{3.21}$$

being θ the zenith angle of the shower.

If muons are assumed to travel at the speed of light, it follows that $l = ct_\mu$. The total time it takes for a muon to reach ground is the sum of $t_\mu + t_\pi$, being $ct_\pi = z_{1\text{st}} - z$ the distance travelled by the parent pion. The difference between this time and the time it takes for the shower front to reach (r,ζ), the *geometrical delay* t_g, satisfies:

$$ct_g = ct_\mu + ct_\pi - ct_{\text{front}} = \sqrt{r^2 + (z-\Delta)^2} + (z_{1\text{st}} - z) - (z_{1\text{st}} - \Delta) \tag{3.22}$$

and therefore

$$z = \frac{1}{2}\left(\frac{r^2}{ct_g} - ct_g\right) + \Delta \tag{3.23}$$

However, muons travel at a speed smaller than the speed of light and therefore suffer a *kinematic delay*. Additional sources of delay, such as multiple scattering or the effect of the geomagnetic field are of much lesser importance, as shown in figure 3.9. Current EASs experiments can usually only measure particles at distances larger than $O(\sim 100)$ m, so the geometrical delay is the dominant contribution. For example, for events at $\theta = 60^\circ$ at distances $r > 1000$ m, the kinematic delay, t_ε, acts as a correction to the total delay (typically below 30%) while the other delays are of the order of few percent. On the other hand, a direct measurement of the energy spectrum of muons is generally not accessible, so a complete computation of the kinematic delay is not feasible. In practice, the kinematic delay is approximated by a parametrisation of its average value $t_\varepsilon \to \langle t_\varepsilon \rangle$ [130], and therefore equation (3.23) is rewritten as

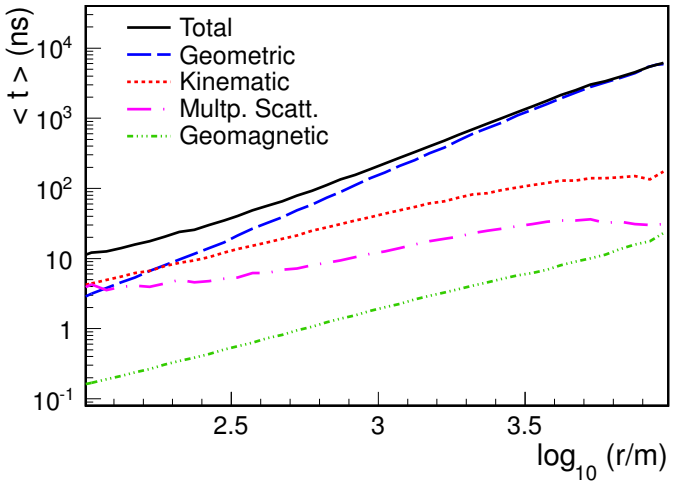


Figure 3.9: Average delays for the different contributions to the total delay for showers at 60° [131].

$$z \approx \frac{1}{2}\left[\frac{r^2}{ct - c\,\langle t_\varepsilon \rangle} - (ct - c\,\langle t_\varepsilon \rangle)\right] + \Delta \tag{3.24}$$

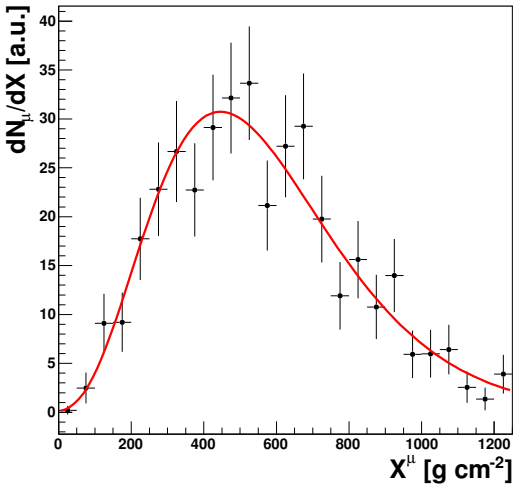
which gives a mapping between the arrival time of muons t and their production distance z. This distance is related to the production depth X^μ using

$$X^\mu = \int_z^\infty \rho\left(z'\right)\ dz' \tag{3.25}$$

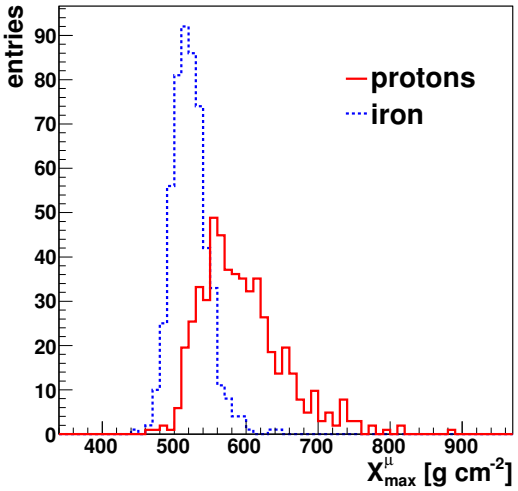
where ρ is the atmospheric density. The MPD is the distribution of produced muons as a function of X^μ (see figure 3.10a). It was first shown in [132] that this method can be applied to the data of the Pierre Auger Observatory, for which the MPD is built from the time information of the Cherenkov detectors. In addition, it was shown that the maximum of the MPD distribution, X^μ_{max} is a mass-sensitive parameter (see figure 3.10b) [123, 89].

The dependence of X^μ_{max} exhibits some properties very much like those of X_{max}. For example, the *muonic elongation rate* $dX^\mu_{\text{max}}/d\log E$ is almost the same regardless of the primary, and the separation between average values for proton and iron primaries is also of the order of 100 g/cm^2 (see figure 3.11b).

The main limitation of this technique at the current stage of development is that it can only be applied to events with energy above $10^{19.3}$ eV and zenith angle between [55°,65°], so the statistics is very limited (see figure 3.11b). Also, the resolution at $E_0 = 10^{19.3}$ is about 90 g/cm^2 (see figure 3.11a), which is almost the total separation between protons and iron nuclei, and goes down to 50 g/cm^2 at 10^{20} eV which is still somehow large. Addressing all of these problems will be the main scope of this thesis, and so they will be discussed thoroughly in the following chapters. More details on the experimental reconstruction of the MPD will also be discussed there.

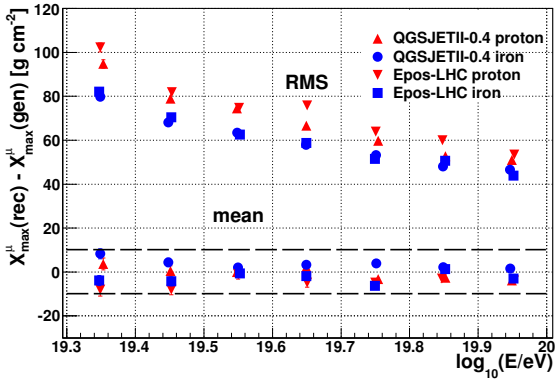


(a) Real reconstructed MPD.

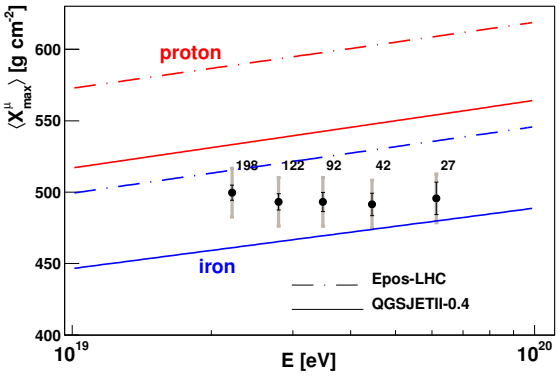


(b) X^{μ}_{max} distributions for proton and iron showers.

Figure 3.10: Reconstructed MPD for an event at $\theta = (59.06 \pm 0.08)^{\circ}$ and $E = (95 \pm 3)$ EeV, with fit to a Gaisser-Hillas function [87] (left). Mass-sensitivity of X^{μ}_{max} shown for simulations at 30 EeV (right) [89].



(a) Bias and resolution of X^{μ}_{max}.



(b) $\langle X^{\mu}_{\mathrm{max}} \rangle$ as a function of energy [89].

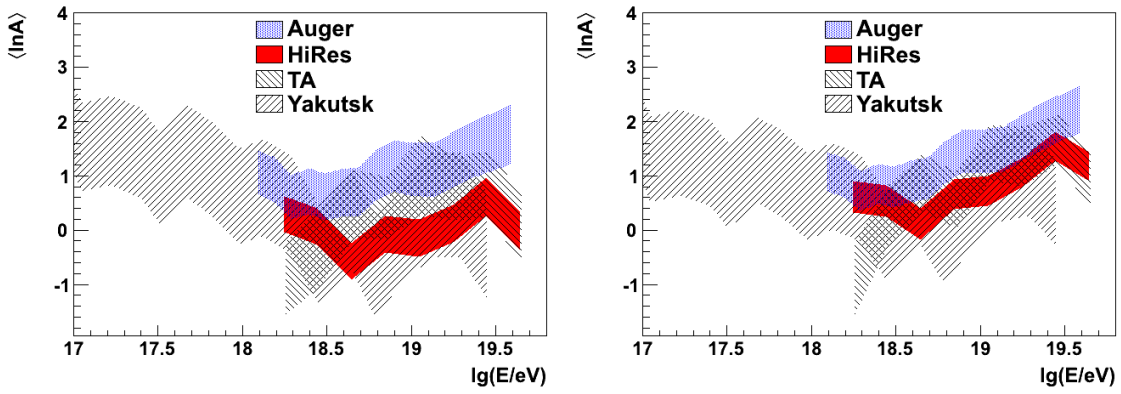
Figure 3.11: Evolution with energy of the mean and standard deviation of the distribution [X^{μ}_{max} (reconstructed) - X^{μ}_{max} (true)] for evens at (55°$\leq \theta \leq$65°) [89] (left). $\langle X^{\mu}_{\mathrm{max}} \rangle$ as a function of energy for the data of the Pierre Auger Observatory (right). The prediction of different hadronic models for proton and iron nuclei are shown. Numbers indicate the amount of selected data in each energy bin and the grey region represents the systematic uncertainty.

3.4 Current status of the measurements of mass composition of UHECRs

The various detectors exploring the ultra-high energy range have very different selection cuts and systematic uncertainties. The most straightforward way to make a comparison between experiments is inferring the average logarithmic mass (3.15), which can be written as:

$$\langle \ln A \rangle = \frac{\langle X_{\text{max}} \rangle_p - \langle X_{\text{max}} \rangle_{\text{data}}}{\langle X_{\text{max}} \rangle_p - \langle X_{\text{max}} \rangle_{\text{Fe}}} \ln 56 \tag{3.26}$$

This way, the different detector effects can be minimised by substituting $\langle X_{\text{max}} \rangle_{p,\text{Fe}}$ by their full simulation values, including detector effects. This is just an approximation, as only an un-biased measurement of $\langle X_{\text{max}} \rangle$ will correlate directly with $\langle \ln A \rangle$ [133], but it will suffice to assess whether all experiments yield compatible results, although of course expression (3.26) is model-dependent. Results are shown in figure 3.12 [134].



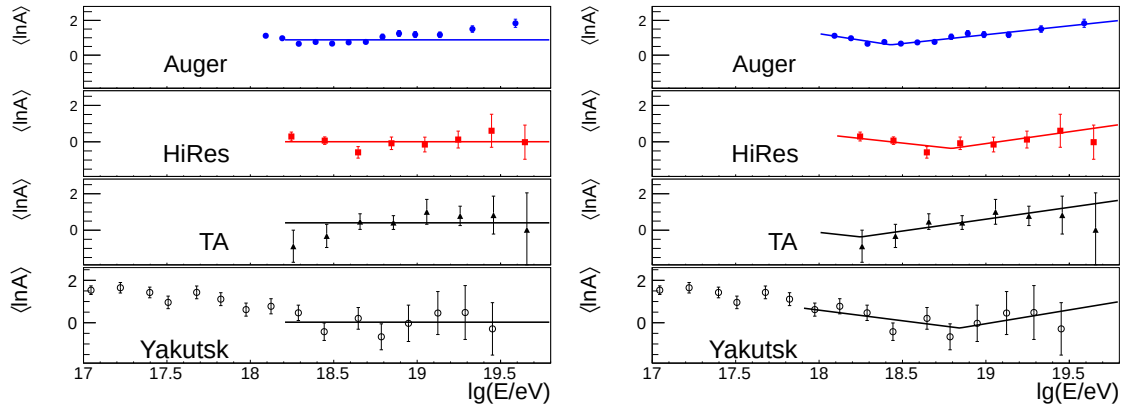
(a) Using QGSJET-II.

(b) Using SIBYLL.

Figure 3.12: Comparison of the average logarithmic mass, $\langle \ln A \rangle$, for different experiments [134]. Shaded regions correspond to the systematic uncertainties.

At ultra-high energies, data from the Pierre Auger Observatory suggest a heavier composition than all other experiments. Nevertheless, results are consistent with TA and Yakutsk within uncertainties. There is some discrepancy between Auger and HiRes at the highest energies when using the model QGSJET-II (figure 3.12a), which almost disappears with SIBYLL (figure 3.12b).

As a way to establish whether different experiments are inferring the same composition *scenario*, i.e., a constant composition regardless of the energy as expected from HiRes results, or a trend to heavier composition as favoured by the Pierre Auger Observatory, a fit to a constant and to a broken line were performed for the different experiments in [134]. Figure 3.13 shows the results for SIBYLL.



(a) Fit to a constant value.

(b) Fit to a broken line.

Figure 3.13: Fit to the logarithmic mass, $\langle \ln A \rangle$, for different experiments [134] using SIBYLL.

All experiments but Auger are compatible with a constant composition. For the Auger data, a fit to a constant results on a large $\chi^2/\text{ndf} = 137/10$, indicating that data from the Pierre Auger Observatory clearly disfavour a constant composition scenario. For the fits to a broken line only the energy and $\langle \ln A \rangle$ at which the lines break were fitted, the slopes fixed to the results of the Auger fit. Although the χ^2/ndf values for these fits are small, the obtained breaking points are incompatible. Further studies and more statistics are necessary to establish the level of compatibility between different experiments.

Apart from X_{max}, some of the forementioned ground level variables can be compared. In particular, Yakutsk uses an array of muon detectors [135] whose data can be compared to the measurements of X^{μ}_{max} [89] and θ_{max} [123] from the Pierre Auger Observatory. The result is shown in figure 3.14 as compared to the fit to $\langle \ln A \rangle$ obtained in the Pierre Auger Observatory using X_{max}.

Despite some differences between measurements from Auger and Yakutsk, all observations suggest a trend to heavier composition above $10^{18.5}$ eV.

The current data coming from different experiments cannot conclude on the mass composition of UHECRs. While many experiments obtain data that are completely consistent with a constant light composition, a changing composition as suggested by the Pierre Auger Observatory cannot be excluded. More statistics, and further constraints on the hadronic models coming from accelerators are necessary in order to clarify the current picture.

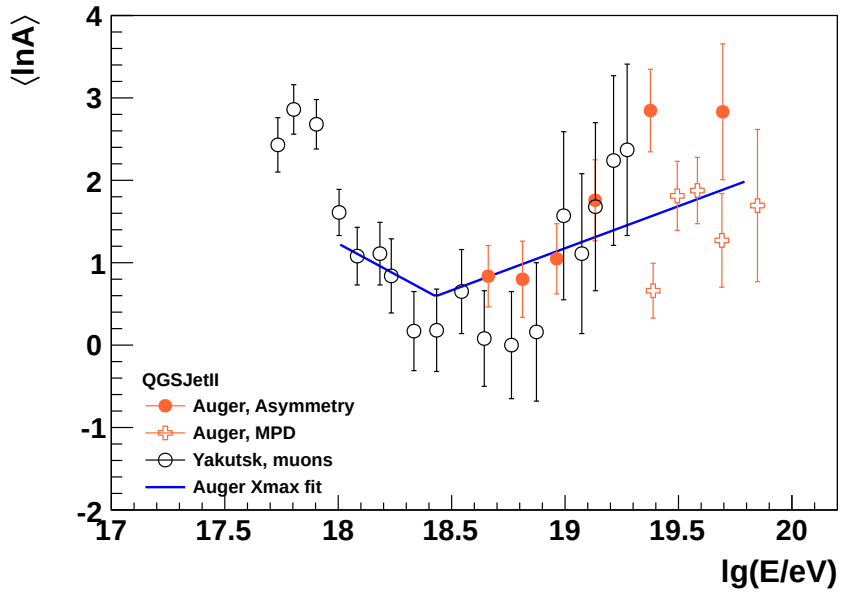


Figure 3.14: Average composition for ground level observables [134] using QGSJET-II.

An idea that is not dangerous is unworthy of being called an idea at all

Oscar Wilde

4

Correction of the radial dependence of MPDs

In this chapter we will address one of the main caveats in the MPD analysis: the dependence of the muon profile on the observation point. It will be shown using CORSIKA [115] simulations that it is possible to correct for this effect. This is the very first step necessary to widen the application range of the MPD technique.

4.1 Dependence of the MPD with the radial distance

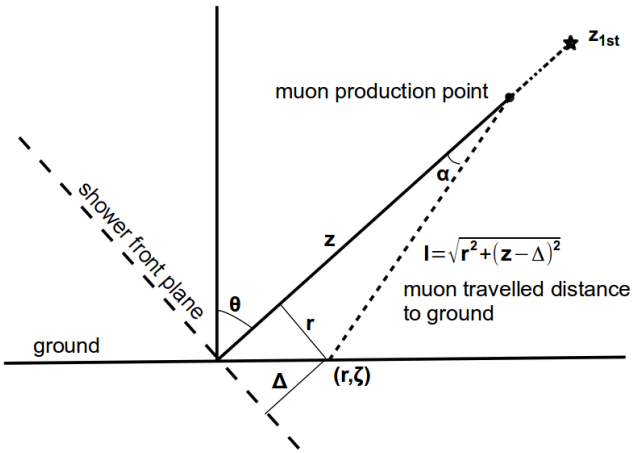


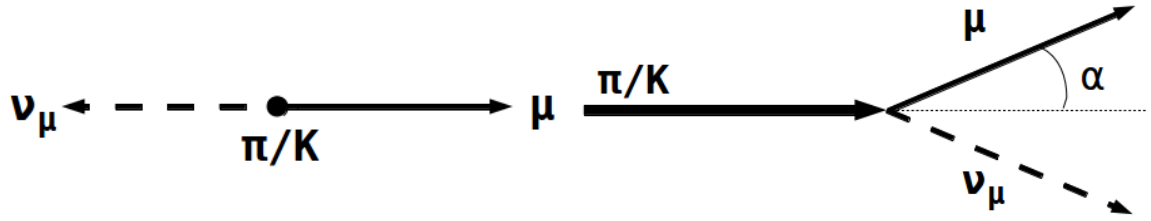
Figure 4.1: Geometry used to obtain the muon travelled distance.

Considering again the geometry that is used to reconstruct the MPD (figure 4.1), it is clear that not all values of z are visible at all distances to the shower axis r. For instance, for a muon produced very deep in the atmosphere it is very unlikely to be detected far from the core, as α is generally small. In the centre of mass of the decaying meson (see

table 4.1 for the different processes in which muons are produced in EASs), the maximum transverse momentum of the outgoing muon is (figure 4.2a):

$$[p_t c]_{\text{max}} = \frac{m_{\pi/K}^2 c^4 - m_\mu^2 c^4}{2 m_{\pi/K} c^2} = \frac{m_{\pi/K} c^2}{2}\left[1 - \left(\frac{m_\mu}{m_{\pi/K}}\right)^2\right] \tag{4.1}$$

which for pions corresponds to 29.8 MeV and for kaons to 235.5 MeV. These values are much smaller than the average energy of muons, which is of the order of some GeV. Therefore, muons are not much deflected and is very unlikely for them to have a large α (figure 4.2b). For example, for a 5 GeV muon coming from a kaon decay $\sin\alpha = p_t/p \approx 0.0471 \rightarrow \alpha \approx 2.7°$.



(a) Centre of mass system (b) Laboratory system

Figure 4.2: Kinematic diagram for the meson decay in the centre of mass and laboratory systems.

$\pi^+ \longrightarrow \mu^+ + \nu_\mu$	$\pi^- \longrightarrow \mu^- + \bar{\nu_\mu}$	99.99%
$K^+ \longrightarrow \mu^+ + \nu_\mu$	$K^- \longrightarrow \mu^- + \bar{\nu_\mu}$	63.51%
$\longrightarrow \pi^+ + \pi^0$	$\longrightarrow \pi^- + \pi^0$	21.16%
$\longrightarrow \pi^+ + \pi^+ + \pi^-$	$\longrightarrow \pi^- + \pi^- + \pi^+$	5.59%
$\longrightarrow \pi^0 + e^+ + \nu_e$	$\longrightarrow \pi^0 + e^- + \bar{\nu_e}$	4.82%
$\longrightarrow \pi^0 + \mu^+ + \nu_\mu$	$\longrightarrow \pi^0 + \mu^- + \bar{\nu_\mu}$	3.18%
$\longrightarrow \pi^+ + \pi^0 + \pi^0$	$\longrightarrow \pi^- + \pi^0 + \pi^0$	1.73%

Table 4.1: Decay modes of charged pions and kaons [107].

As a consequence, the observed distribution of dN^μ/dz for the incoming muons depends on the distance to shower axis, and the average value $\langle z \rangle$ increases with r. This effect is shown in figure 4.3 for four different distances. At large zenith angle, the muon travelled distance becomes very large, and this dependence is minimised. Therefore, at 60°, the distribution of z is approximately *universal*, regardless of the observation distance r. This is the main reason why so far the MPD method has only been applied to showers with zenith angle around 60°.

At every zenith angle, it is guaranteed to exist a certain distance r_0 at which the *observed* z distribution is the same as the integral over $r \in (0, \infty)$ apart from a scaling factor [136]. However, the reconstruction of the MPD at a fixed distance is not feasible,

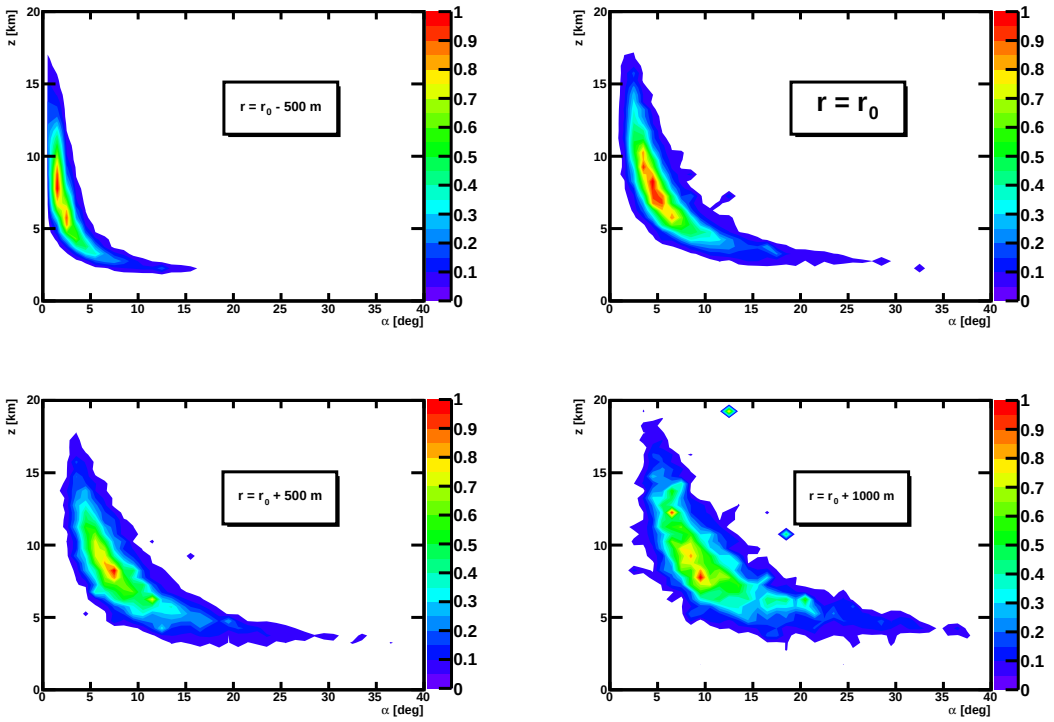


Figure 4.3: Accessible values of z and α at different distances from the shower axis for proton simulations at 50° ($r_0 \approx 800$ m).

due to the small muon flux at ground, and therefore the combination of several SD stations is a must for an event-by-event reconstruction of the MPD. Moreover, r_0 is usually too close to the shower core to be accessible in general, especially at low angle (see figure 4.4).

In order to be able to sum over different distances for events at lower zenith angle, this dependency of z on r must be taken into account and corrected. Otherwise, the combination of different stations would not reconstruct the real muon profile, but a deformed one. The followed approach was to look for a transformation on $z \rightarrow z' = zf\,(r, z)$ so that the distribution dN^{μ}/dz' is approximately universal at all distances and regardless of the zenith angle.

Some of the properties that such transformation $f\,(r, z)$ must satisfy can be summarised:

- It must satisfy $f\,(r_0, z) = 1\ \forall z$, so that no transformation at all is applied at the optimum distance r_0.
- At $r > r_0$ it must satisfy $f\,(r, z) < 1$, at least for large values of z.
- At 60°, the effect of the transformation $f\,(r, z)$ must be negligible or very small.
- It can be argued whether it is necessary to include z in the transformation, or if a

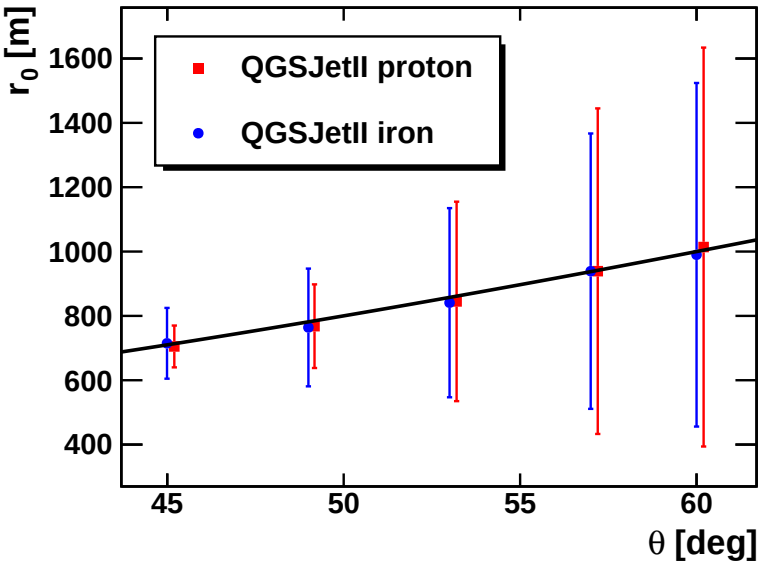


Figure 4.4: Optimum r_0 for simulations. The graph for protons has been slightly displaced to the right for better visibility (by 0.2°). Error bars represent a total bias in X^{μ}_{max} of ± 10 g/cm^2 and hence they grow with θ, showing the approximate universality at large zenith angles. The curve corresponds to a parametrisation $r_0[\text{m}] = 1700 - 1400 \cos\theta$.

form like $\tilde{f}\,(r)$ would suffice. However, we see in figure 4.6a that the deformation on the distribution of dN^{μ}/dz is not just a mere *shift*, but also a deformation on the profile. This deformation can only be taken into account if the transformation is designed as z-dependent.

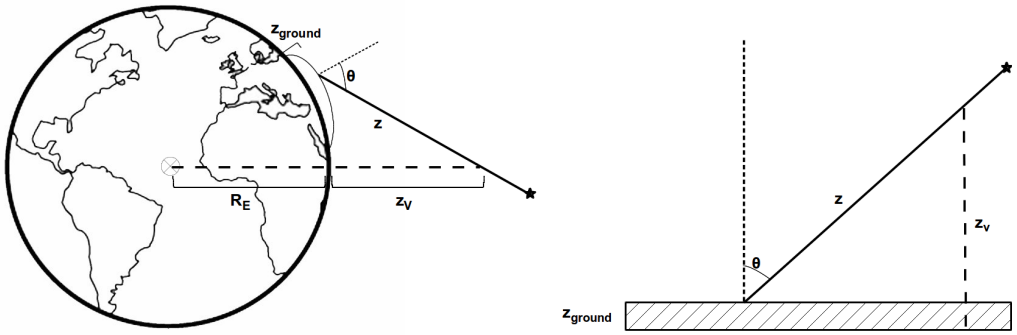
- Numerical values are not expected to be too far from unity, as the effect is in fact *small* in terms of z (but it is amplified by the transformation into X^{μ} due to the approximated exponential density profile of the atmosphere).
- The resulting distribution dN^{μ}/dz' integrated over r must be identical to dN^{μ}/dz.
- As a consequence, the reconstructed value for X^{μ}_{max} must be the same at every distance r.

The coordinate system used by CORSIKA provides z_v, the vertical altitude above sea level (see figure 4.5a). The relationship between this quantity and z can be obtained by simple geometry

$$z_v = \sqrt{z^2 + \left(R_E + z_{\text{ground}}\right)^2 + 2z\left(R_E + z_{\text{ground}}\right)\cos\theta} - R_E \tag{4.2}$$

that can be inverted as

$$z = \sqrt{\left(z_v + R_E\right)^2 + \left(R_E + z_{\text{ground}}\right)^2 \sin^2\theta} - \left(R_E + z_{\text{ground}}\right)\cos\theta \tag{4.3}$$



(a) Curved geometry (b) Flat geometry

Figure 4.5: Coordinate system used by CORSIKA in the curved case (left) and the flat approximation (right).

where R_E is the radius of the Earth and z_{ground} is the altitude above sea level of the observation site. These expressions will be used in the reconstruction of Monte Carlo events and experimental data. However, in simulations it is customary to use the approximation of *flat* geometry, i.e., neglect the curvature of the Earth (see figure 4.5b). In this case, expression 4.2 can be rewritten just as

$$z_v = z\cos\theta + z_{\text{ground}} \tag{4.4}$$

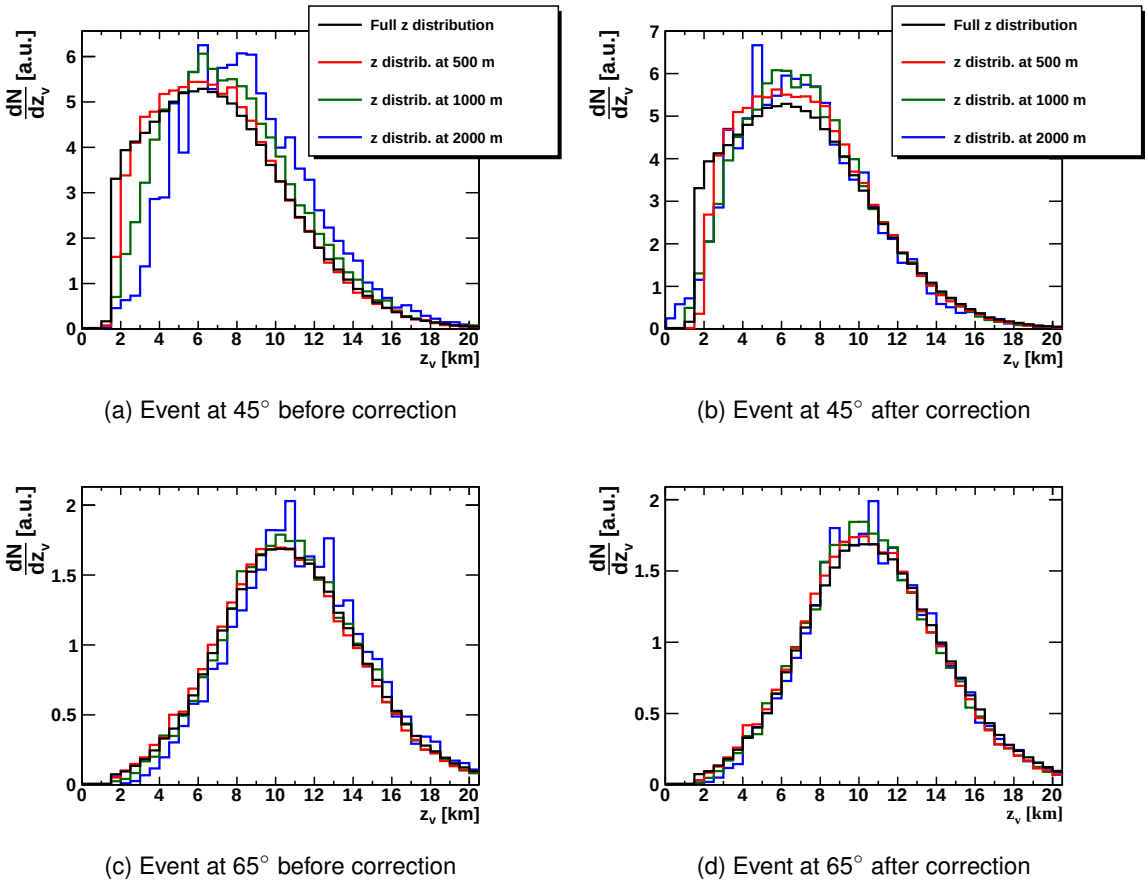
We found empirically a candidate expression for the transformation $f\,(r, z)$ in terms of z_v, which satisfies all the criteria stated above to a good approximation

$$z'_v = z_v \times \frac{z_v \tan^2\theta - r}{z_v \tan^2\theta - r_0} \tag{4.5}$$

This choice for $f\,(r, z_v)$ clearly satisfies $f\,(r_0, z_v) = 1\ \forall z$ by construction. It also holds that $f\,(r > r_0, z_v) < 1$ if $z_v \tan^2\theta > r$. This criterion holds in most of the cases, as z is rarely smaller than ~ 2 km. Finally, the factor $\tan^2\theta$ becomes bigger at larger θ, so that the resulting value for $f\,(r, z_v)$ is closer to unity at large zenith angle. The rest of the properties are less clear to see immediately, but will be shown in the following section.

4.2 Performance of the radial correction factor

The effect of the correction factor (4.5) on the dN^μ/dz distribution is shown in figure 4.6. After applying this correction, the distribution becomes approximately universal regardless of the distance r. This fact can be stressed by showing the evolution with r of the average value of z, $\langle z\rangle$ (see figure 4.7). However, we are interested in conserving not only the average value, but also the *shape* of the whole distribution. In order to quantify to what extent is this true, the standard deviation of the dN^μ/dz distribution, $\sigma\,(z)$, is also shown before and after applying the correction factor in figure 4.7. By analysing these



(a) Event at 45° before correction

(b) Event at 45° after correction

(c) Event at 65° before correction

(d) Event at 65° after correction

Figure 4.6: Example of the performance of the correction factor (4.5) for two events at 45° and 65° respectively.

two figures, it is clear how the correction factor (4.5) makes both statistical parameters *flatter* as a function of r. This figure also shows how both $\langle z \rangle$ and $\sigma\,(z)$ are left unchanged at $r = r_0$.

The same parameters and their evolution after applying (4.5) are shown for an event at $\theta = 65^\circ$ in figure 4.8. In this case the effect of the correction factor is negligible, as expected, since the dependence with the observation distance is much smaller at large zenith angle.

The effect of the correction factor on the radial dependence of z in an event by event basis was found to be consistent regardless of the primary and hadronic model. Also, no significant dependence on the energy was found. The particular value of r_0 for a given zenith angle is slightly different for different primaries or energies, but the differences are of the order of 100 m at most, and do not introduce a significant reconstruction bias in terms of X^{μ}_{max}.

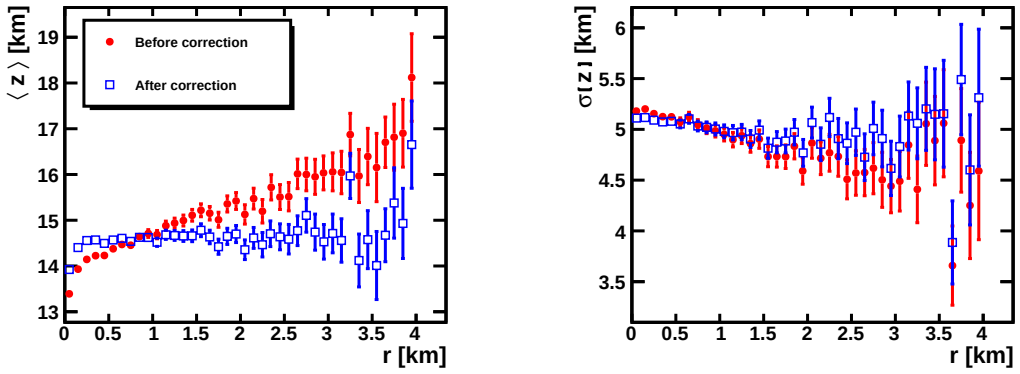


Figure 4.7: Example of the effect of the correction factor (4.5) on $\langle z \rangle$ (left) and $\sigma(z)$ (right) for an event at 45°.

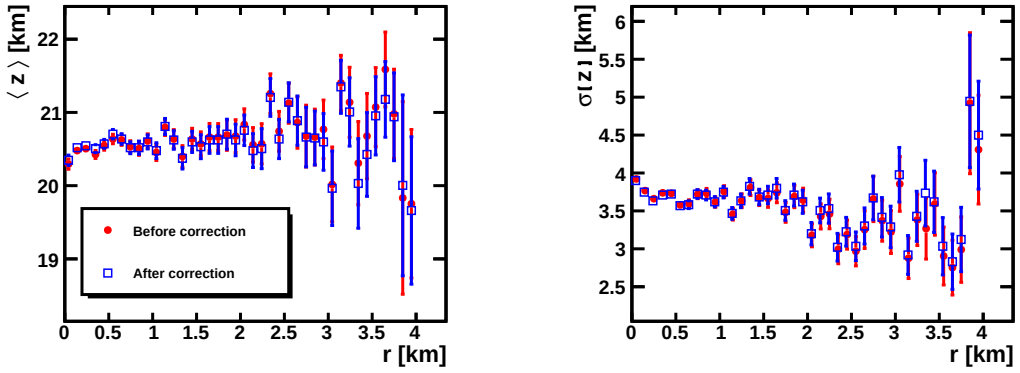


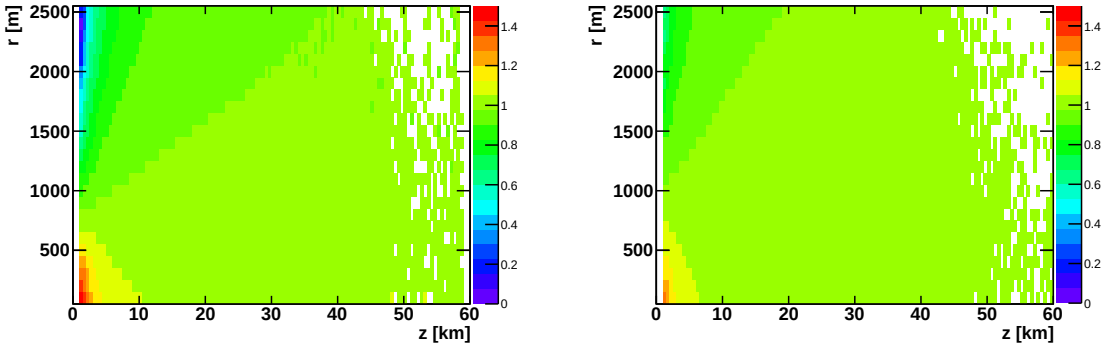
Figure 4.8: Example of the effect of the correction factor (4.5) on $\langle z \rangle$ (left) and $\sigma(z)$ (right) for an event at 65°.

4.2.1 Numerical values of the correction factor

Figure 4.9 shows the average numerical values of $f(r, z)$ for different values of z and r for events at 45° and 65°. It is clear how only for very small values of z and far away from the shower core does this factor deviate notably from unity, correctly accounting for the suppression shown in figure 4.3. Very small values of z near the core are enhanced by the correction factor, which is bigger than one in this region. It will be shown that this correction is not optimum at distances significantly smaller than r_0. However, as that region is not available in current experiments, we will neglect this inaccuracy.

4.2.2 Effect of the correction in terms of atmospheric depth

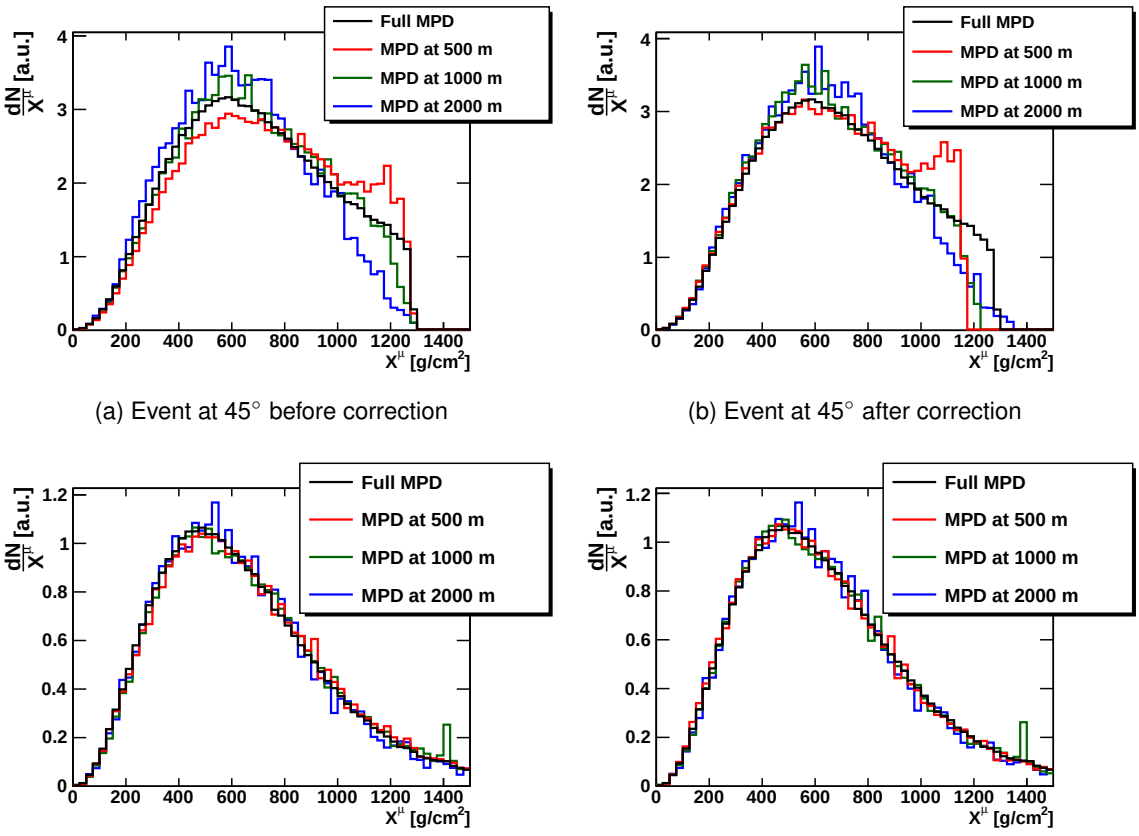
The final validity of equation (4.5) is tested in terms of the distribution dN^{μ}/dX^{μ} as a function of r. This is shown in figure 4.10. Due to the approximately exponential density profile of the atmosphere, the differences in z are enhanced when translated into depth.



(a) $f\,(r,z)$ for events at 45°

(b) $f\,(r,z)$ for events at 65°

Figure 4.9: $f\,(r,z)$ as a function of r and z for events at 45° and 65° respectively.



(a) Event at 45° before correction

(b) Event at 45° after correction

(c) Event at 65° before correction

(d) Event at 65° after correction

Figure 4.10: Performance of the correction factor (4.5) for two events at 45° and 65° respectively in terms of dN^{μ}/dX^{μ} distribution.

Therefore, the good performance of the correction factor shown in figure 4.7 does not guarantee that it should work in terms of X^{μ}. Nevertheless, figure 4.10 exhibits how the correction factor accurately transforms the MPD distribution.

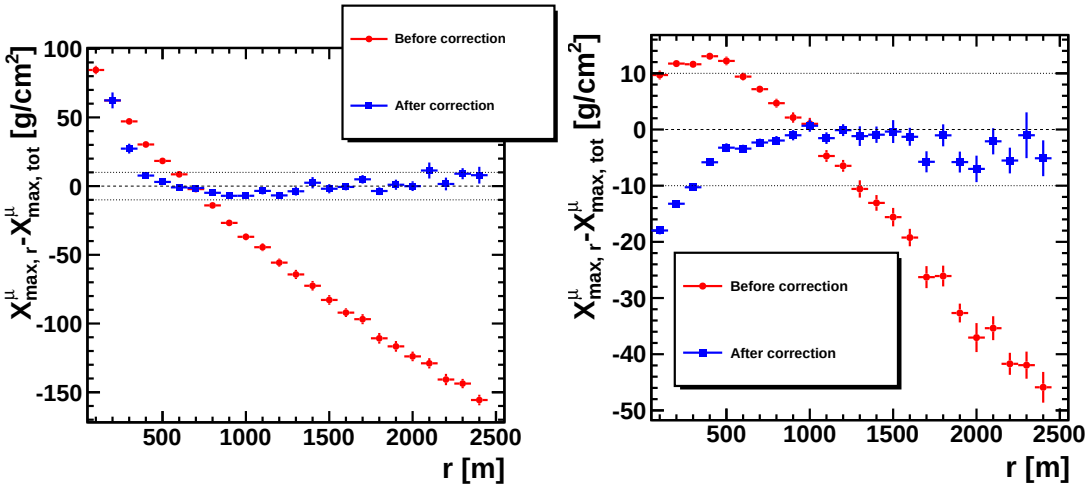


Figure 4.11: Reconstruction of X^{μ}_{max} before and after correction as a function of distance r for events at 45° (left) and at 65° (right).

Finally, in terms of the observable we are more interested in, X^{μ}_{max}, figure 4.11 shows how the same value, within ± 10 g/cm^2, is reconstructed regardless of the distance. As stated before, this is not true for distances to the core below 500 m. However, in practice we will be interested in stations that are at least at 1000 m from the shower core for various reasons. The most important is avoiding a significant contribution of the electromagnetic signal, which becomes increasingly more important as the distance to the core gets smaller.

4.3 Conclusions on the correction factor

It has been shown how equation (4.5) satisfies all the requirements that were introduced in the previous section. However, it is necessary to point out that this formula was obtained numerically and is therefore just an approximation. This means that any expression fulfilling the previous conditions can be used as a correction factor for the radial distance. In particular, we explored the possibility to use an expression of the form $f\,(r,z) = \exp\left((r_0 - r)/\beta\right)$ with β a free zenith-dependent parameter, and its performance was found to be very similar. However, some discrepancies were found between the behaviour for proton and iron when using this formula, the value of β being difficult to accommodate to both primaries, and so in the following chapter the expression (4.5) will be used.

However far modern science and technics have fallen short of their inherent possibilities, they have taught mankind at least one lesson: Nothing is impossible

Lewis Mumford

5

Extension of the MPD analysis to a wider zenith and energy range

We have seen in the previous chapter how the radial dependence of MPDs can be accounted for using an empirical correction factor. However, this is not the only issue that we need to overcome in order to widen the zenith range for the MPD analysis. For example, one of the main disadvantages lays in the fact that at smaller zenith angles the contribution of electromagnetic signal is larger [4]. Hence, we have to develop some algorithm that enhances the signal coming from muons, in order to build a realistic MPD. This and other problems that appear in the reconstruction process will be addressed in this chapter. We will follow the same strategy as in the standard MPD reconstruction, which builds the MPD from the time information of the FADC traces. A different approach, based on the modelling of the propagation of muons and shower development is also feasible. Further details on this approach can be seen in [137].

5.1 Treatment of the electromagnetic contamination

Showers at $\theta = 60^{\circ}$ are basically free of electromagnetic contamination, due to the large amount of matter they traverse before reaching ground. Still, a minor contribution survives and reaches ground detectors, and for that matter a simple approach is followed in the standard reconstruction [132]. This approach is based on the fact that electrons and photons suffer larger deviations and multiple scattering, and hence they produce a signal in the PMTs that is flatter and more spread over time than the one produced by muons. Taking this into account, the electromagnetic contamination rejection in the standard analysis consists of applying a constant threshold of 0.3 VEM to the FADC signals. Only entries above that value are used to reconstruct the MPD. With this approach, above 85% of the signal from muons is kept, according to Monte Carlo simulations [132].

The relative amount of muons, electrons and photons depends on the high energy hadronic model and in lesser extent on the atomic mass of the primary. Moreover, it has been pointed out that all of the simulation codes existing so far have a muon *deficit* as compared to what is observed in data [89]. Therefore, when designing an algorithm to reject electromagnetic signal we chose to always use *relative* quantities, i.e., ratios to the total signal, so the dependence on simulations is minimised as much as possible.

The approach followed in the standard analysis is insufficient for events at smaller zenith angle, and also for stations that are closer to the shower core (in the standard analysis, only stations further than 1700 m are considered [89]). The average number of stations per event decreases at smaller zenith angle, so we are interested in using stations located closer to the core. In addition, the resolution of the technique is expected to improve if closer stations are included in the analysis, as the number of muons falls steeply as a function of the radial distance. However, getting too close to the shower core would introduce a large amount of electromagnetic contamination, so a compromise value of $r_{\text{min}} = 1000$ m was chosen. Another advantage of this moderate value is that stations used in the analysis will be free of saturation.

Considering all of this, we developed a slightly more sophisticated algorithm that cuts the signal both *horizontally* and *vertically*, i.e., in terms of the amount of signal as in the standard approach but also in terms of the arrival time.

Temporal window for FADC traces

As already stated in the previous section, the signal coming from the electromagnetic part of the shower is expected to arrive later for a given position at ground. An example of this fact is shown in figure 5.1. Therefore, the contribution of very late FADC bins is dominated by electromagnetic signal and does not introduce valuable information to the MPD reconstruction. Moreover, the electromagnetic signal that comes together with the muonic front, or the one that is produced by muon decay, the so-called *electromagnetic halo*, is of lesser importance to us, as it does not affect the profile of the MPD, but merely acts as a scaling factor. Considering all of this, a cut in time was applied, in order to keep only early bins, those that are muon-rich.

The strategy followed was to limit the amount of bins that are used in the MPD reconstruction. The maximum bin entering the analysis was obtained so that about 99% of the muonic signal is included in these bins. The position of this limit, i_{max} was parametrised as a function of the total signal, resulting in an approximate $\cos^2\theta$ dependence, reflecting the fact that the spread of signal grows with zenith angle as does the amount of traversed matter. This parametrisation was used to compute the position of this limit for every station, which is chosen so that it satisfies

$$\sum_{i=1}^{i_{\text{max}}} S_i = \xi \; S_{\text{tot}} \tag{5.1}$$

where i runs over the FADC counts, S_i is the signal in the bin i and S_{tot} the total signal of the station. The value of ξ is parametrised as:

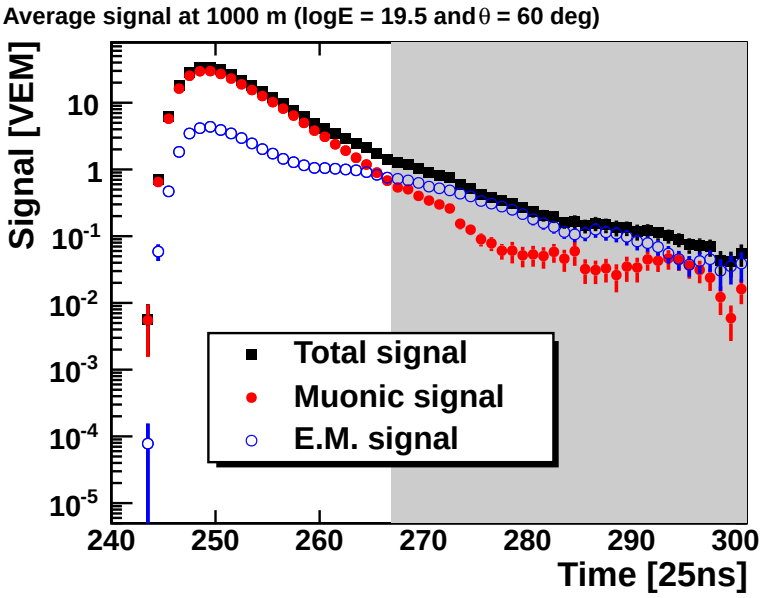


Figure 5.1: Contribution of different components as a function of time. The shaded area represents the region dominated by the electromagnetic signal.

$$\xi = \min\left[1.15 - 0.6\cos^2\theta, 0.99\right] \tag{5.2}$$

which has a minimum of 0.85 at 45° and is fixed at 0.99 above $\theta \approx 58.9°$ (see figure 5.2).

To validate this model, the evolution of the cumulative signal over time as a fraction to the total signal was compared between simulations and experimental data (see figures 5.3 and 5.4 for examples at two different distances to the shower core). The average value for QGSJet-II for proton and iron primaries is used to establish the comparison. The differences between primaries were of the order of one time bin (i.e. 25 ns), while for different hadronic models the differences are negligible.

A natural consequence of the underestimation of the number of muons by Monte Carlo simulations is that the rise-time of the signals is larger in simulations than in data. Therefore, figures 5.3 and 5.4 show how simulations grow systematically slower than data. However, this difference is smaller than 25 ns at all distances and zenith angles for $\xi > 0.85$. On the other hand, the start bin is delayed in simulations with respect to data. Nevertheless, the effect of this difference is minimised by the fact that i_{max} is obtained for every station in terms of the relative total signal, so it will automatically adapt to this discrepancy.

We can therefore trust that at least *qualitatively*, we are actually favouring muonic signal in a consistent way in simulations and in data. Also, the fraction of muonic signal that is kept computed using simulations is not expected to differ much from reality according to figures 5.3 and 5.4.

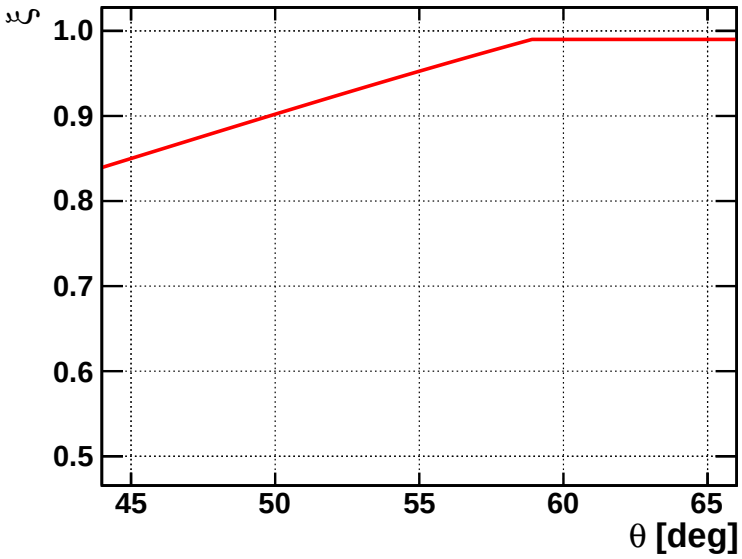


Figure 5.2: Fraction of the total signal ξ kept by the time window as a function of zenith angle.

5.1.1 Adjustable signal threshold

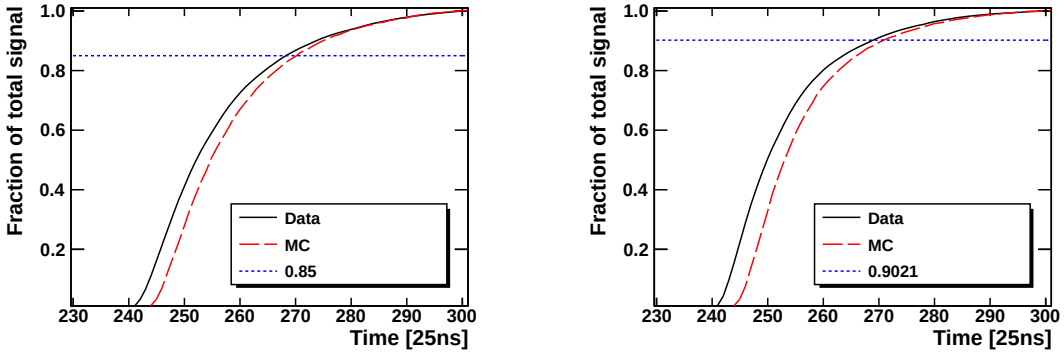
The second part of the algorithm to remove electromagnetic contamination is inspired by the standard reconstruction of MPDs. However, instead of fixing a constant value for the signal threshold, the idea is to relate it to the total signal. More specifically, we will pick a threshold S_{thr} such as

$$\sum_{i}^{i_{\text{max}}} S_i H\left(S_i - S_{\text{thr}}\right) = 0.95\ \xi\ S_{\text{tot}} \tag{5.3}$$

where $H(x)$ represents the Heaviside function. In other words, we will choose the threshold so that 95% of the *remaining* total signal (after applying the time window cut) is kept. In practice, determining S_{thr} to high precision is time-consuming, and hence it will computed in increasing steps of 0.1 VEM until the sum over bins in (5.3) falls below 95%, so the precision of this threshold is ± 0.1 VEM.

Choosing S_{thr} as a function of the total signal has some advantages, including:

- It will automatically adapt if different stations have slightly different calibration constants.
- Seasonal and ageing effects on the stations will also be minimised, as the threshold will also increase or decrease in approximately the same amount as the integrated signal.
- For events at smaller zenith angle, in which the electromagnetic signal is less



(a) $E = 10^{19.15}$ eV and $\theta = 45°$

(b) $E = 10^{19.15}$ eV and $\theta = 50°$

(c) $E = 10^{19.15}$ eV and $\theta = 55°$

(d) $E = 10^{19.15}$ eV and $\theta = 60°$

Figure 5.3: Cumulative signal as a function of time for different zenith angles at $E = 10^{19.15}$ eV and $r = 1000$ m. Horizontal lines represent the value of ξ at that zenith angle.

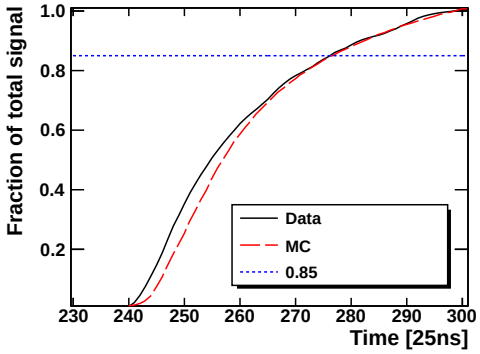
spread over time, the threshold will automatically adapt to approximately the same rejection efficiency as for inclined events.

- Finally, the result is again written in a relative way, so the Monte Carlo dependence is minimised.

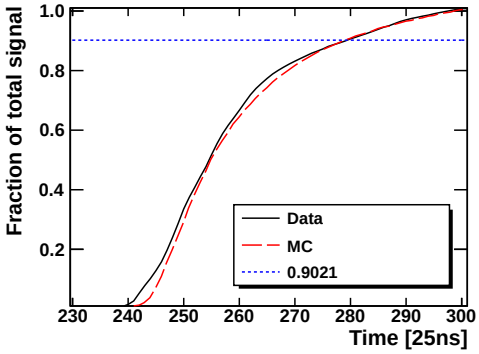
Effect of the threshold on the time shift

No matter how sophisticated the technique to remove electromagnetic contamination is, it will also remove some fraction of the muonic signal. Therefore, it will have some impact on the amount of signal that is collected for single muons. In particular, it will affect the time shift t_0 that needs to be applied to the time reconstructed for every entry.

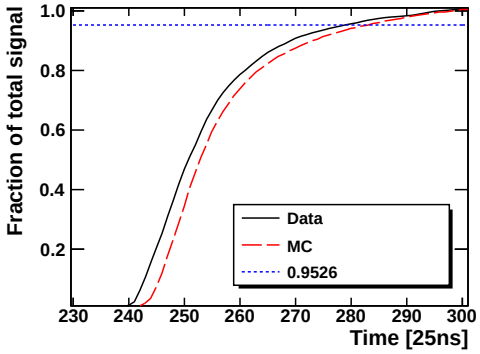
When a muon enters one of the Cherenkov stations, it takes some time for light to travel to the PMTs. In addition to this delay, the PMTs will introduce additional delay, corresponding to the time it takes to produce the avalanche of electrons. Finally, the



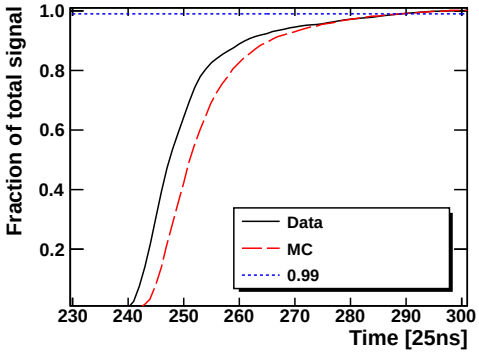
(a) $E = 10^{19.55}$ eV and $\theta = 45°$



(b) $E = 10^{19.55}$ eV and $\theta = 50°$



(c) $E = 10^{19.55}$ eV and $\theta = 55°$

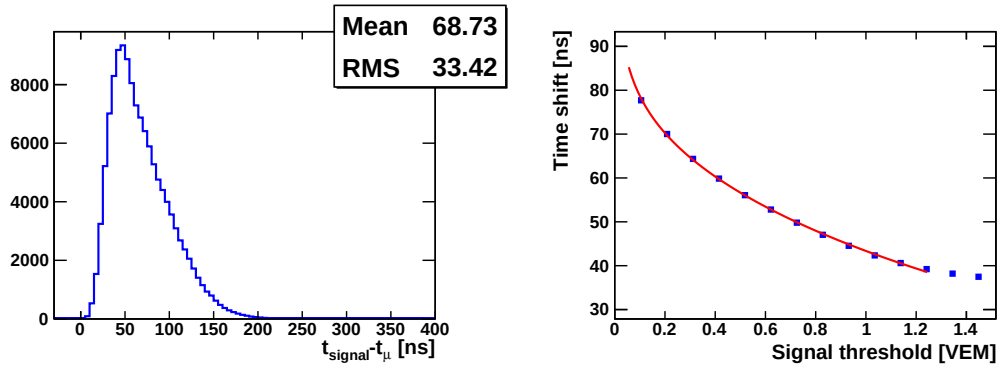


(d) $E = 10^{19.55}$ eV and $\theta = 60°$

Figure 5.4: Cumulative signal as a function of time for different zenith angles at $E = 10^{19.55}$ eV and $r = 1500$ m.

electronics chain will also introduce some delay to build the FADCs. So a muon entering at a given time t_μ will be reconstructed at a delayed time $t_{\text{signal}} = t_\mu + t_s$ where t_s is the total delay introduced by all these effects. In order to account for all of them, a simulation of the station behaviour was carried out. A large amount of muons (50000 muons) were injected randomly distributed within a time bin of 25 ns. The difference between the resulting reconstructed time and the original time of injection provides us a measurement of the *time shift* that must be subtracted to the measured time. The result of this simulation is shown in figure 5.5a when a strategy as the one in (5.3) is used to account for the signal produced by the muon. The effect of the injection direction was found to be negligible, so the case for vertical muons is shown in this figure.

Figure 5.5a yields an average time shift of about 70 ns, that will be taken as a constant for our analysis. This is equivalent to an effective value of $S_{\text{thr}} = 0.22$ VEM, as shown in figure 5.5b. However, this is only the numerical value obtained for single muons. The actual value of S_{thr} for a given station will depend on the electromagnetic background at that particular distance to the shower core and zenith angle. This shift is introduced to



(a) Time shift for method (5.3)

(b) Time shift vs. signal threshold

Figure 5.5: Average time shift for the electromagnetic removal algorithm (left). Parametrisation of the time shift as a function of the signal threshold (right), $t_s \approx 63.68 S_{\text{thr}}^{-0.1087} \exp\left[-0.384 S_{\text{thr}}\right]$ [ns].

account for the fact that muons are not reconstructed as very narrow peaks in the FADC traces, but as relatively wide, asymmetric humps.

5.1.2 Performance of the algorithm to reduce the electromagnetic contamination

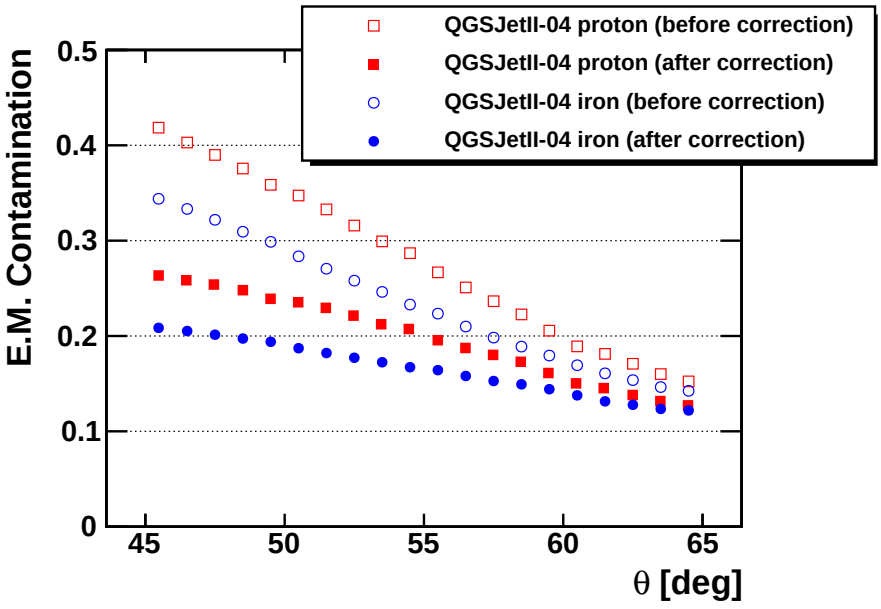
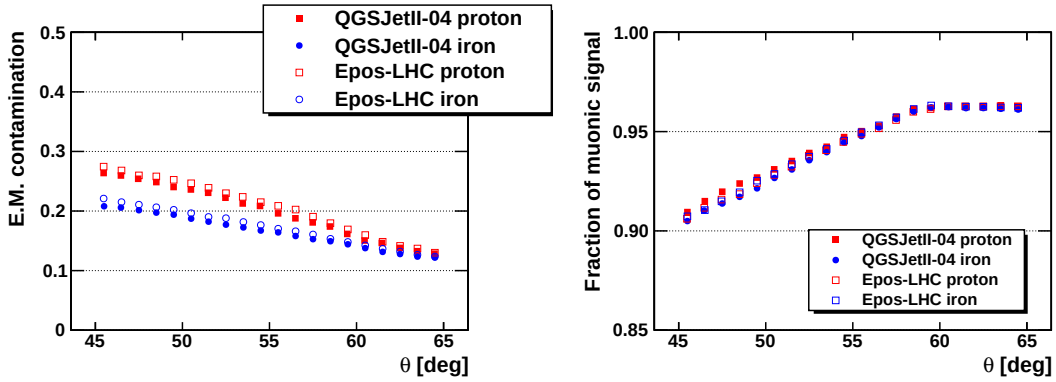


Figure 5.6: Electromagnetic contamination expressed as the fraction of total signal for stations with $r > 1000$ m before and after the rejection algorithm.

The final result on the electromagnetic signal rejection when both parts of the algorithm are applied simultaneously is shown in figure 5.6 for different zenith angles. In the worst possible scenario, which occurs for events at $\theta \approx 45°$, the electromagnetic contamination is below 26%, which means that the muonic signal is about three times larger than the electromagnetic one even at this worst case scenario. This value for the electromagnetic contamination at $\theta \approx 45°$ is very similar to the one obtained in the standard MPD analysis for events at $\theta \approx 60°$.



(a) Electromagnetic contamination

(b) Muonic signal acceptance

Figure 5.7: Final performance of the electromagnetic signal treatment. Evolution of the total electromagnetic contamination for $r > 1000$ m as a function of the zenith angle (left). Fraction of muonic signal that is kept after the electromagnetic signal rejection (right).

This rejection of electromagnetic signal is achieved with only a minor effect on the muonic signal, as shown in figure 5.7b, which shows how more than 90% of the total muonic signal is always preserved for all primaries and hadronic models. Finally, figure 5.7a shows how these conclusions depend very little on the particular choice of hadronic model, and only by a few percent depending on the primary.

5.2 Parametrisation of r_0 in the reconstruction

Although the parameter r_0 was established using simulations in the previous chapter, it is not guaranteed to keep the same value for reconstructed events. In fact, detector effects are expected to change the particular value of r_0. For example, the SD cannot measure the whole muon flux but a much sparser sample. In addition, effects introduced by the electronics such as the ones mentioned in the previous section, or even errors in the calibration of stations can affect the reconstruction of the MPD.

Therefore, the strategy that will be followed is to parametrise again r_0 using simulations. In order to do that, every simulated event will be reconstructed for different values of r_0 running from 0 to 8000 m in steps of 100 m. For a given bin of zenith angle and energy, r_0 will be chosen as the one minimising the absolute difference between the reconstructed value of X^{μ}_{max} and the *true* one, i.e., the one obtained for the Monte Carlo

event before being reconstructed. In fact, in practice we will always refer to the *apparent* MPD, meaning the MPD obtained when only muons between 1000 m and 4000 m from the shower core and that survive until ground level are considered. The upper limit to the distance is due to the fact that for very low muon flux the trigger probability becomes very small, and in practice there are almost no stations further than 4000 m from the core in reconstructed events.

An example of the procedure that was used to obtain r_0 is shown in figure 5.8. For each bin in $(\log E, \theta)$, a linear fit to the profile of the difference between the reconstructed and the real value of X^{μ}_{max} as a function of r_0 is performed. A first seed to r_0 is obtained as the abscissa at which the linear fit equals zero, which represents the distance at which there is no reconstruction bias. Then, this value is used as a seed to a new fit within its neighbourhood, in a band of ± 250 m, from which a new value for r_0 is obtained from the new fit following the same strategy. The process is repeated once again, using the new value of r_0 as the centre of the fit. This process converges very fast, and after two or three iterations the change in r_0 is smaller than 1 m, and then the iteration ends.

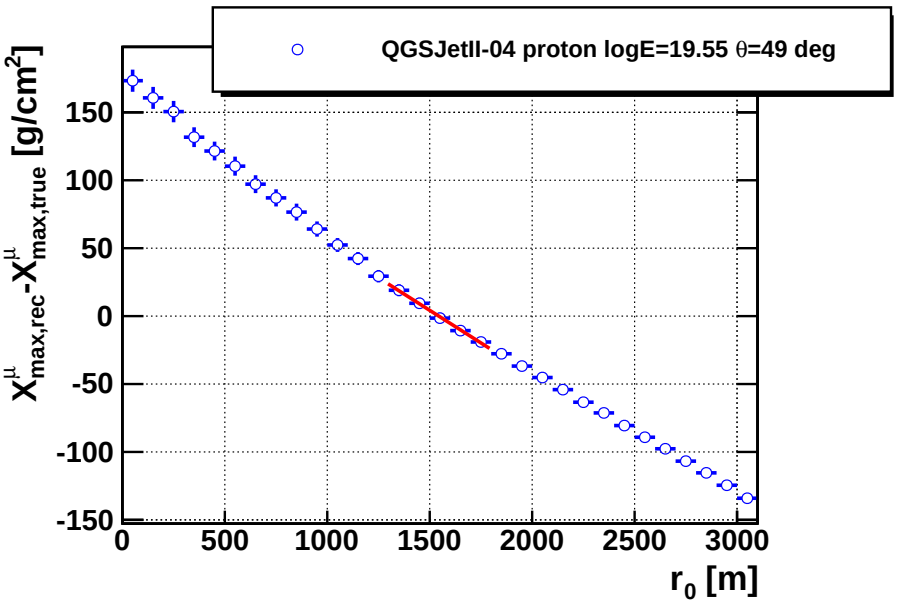


Figure 5.8: Example of the determination of r_0 for a particular bin of energy and zenith angle. The fit is performed between [$r_0 - 250$ m, $r_0 + 250$ m] ($r_0 = 1544$ m in this example).

The values for r_0 are obtained for proton and iron simulations independently in bins of $\log E$ and θ, and then a value averaged over both primaries is finally used. The resulting table of r_0 is shown in figures 5.9 and 5.10 as a function of the primary energy and zenith angle. The bins of $\log(E/\text{eV})$ have width 0.1, running from 19 up to 20. In the case of θ, 28 bins of width 0.0125 rad (approximately 0.7°) between 0.785 (44.98°) and 1.135 (65.03°) were performed.

The dependence of r_0 on the energy of the primary is much smaller than the one on the zenith angle, but not small enough to be neglected, and hence the choice of a bi-dimensional tabulation of the numerical values. The validity of this tabulation will be

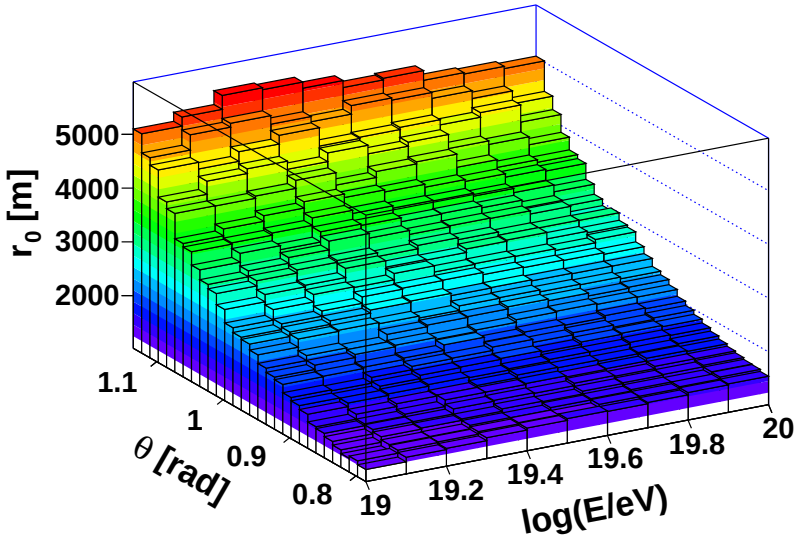


Figure 5.9: Dependence of r_0 with zenith angle and primary energy.

shown in the following sections.

5.3 Effect of the corrections on the average MPD profile

In order to assess the effect of both the correction of radial dependence and the electromagnetic signal removal, we analyse the average MPD profile. This is achieved by the addition of all the MPDs for a given bin of energy and zenith angle. The comparison of the average profile before and after the corrections for CORSIKA simulations and their reconstruction using $\overline{\text{Off}}\underline{\text{line}}$ is shown in figure 5.11 for QGSJETII-04 protons with $\log E \in [19.0, 19.3]$. This comparison is shown for the two zenith windows that will be considered, the one that was already analysed using the standard MPD technique ($55° < \theta < 65°$) and the extension of this work ($45° < \theta < 55°$).

Figure 5.11 shows how the effect of the whole algorithm correctly accounts for the deformation on the profile that is introduced by the reconstruction. There is still some mismatch between the two distributions in the tail. However, we are interested just in the maximum of the distribution, X^{μ}_{max}, so this degree of concordance will suffice for our purposes. The maximum X^{μ}_{max} is obtained through a fit to the MPD histogram, so in order to not introduce a bias in the reconstruction of X^{μ}_{max}, we will fit the histogram up to a maximum depth of 1000 g/cm^2. In the case of QGSJETII-04 protons, only four out of 12000 simulated events have a value of X^{μ}_{max} above 1000 g/cm^2 for the apparent MPD at $r \in [1000\ \text{m}, 4000\ \text{m}]$, so this limit does not introduce any significant loss of statistics.

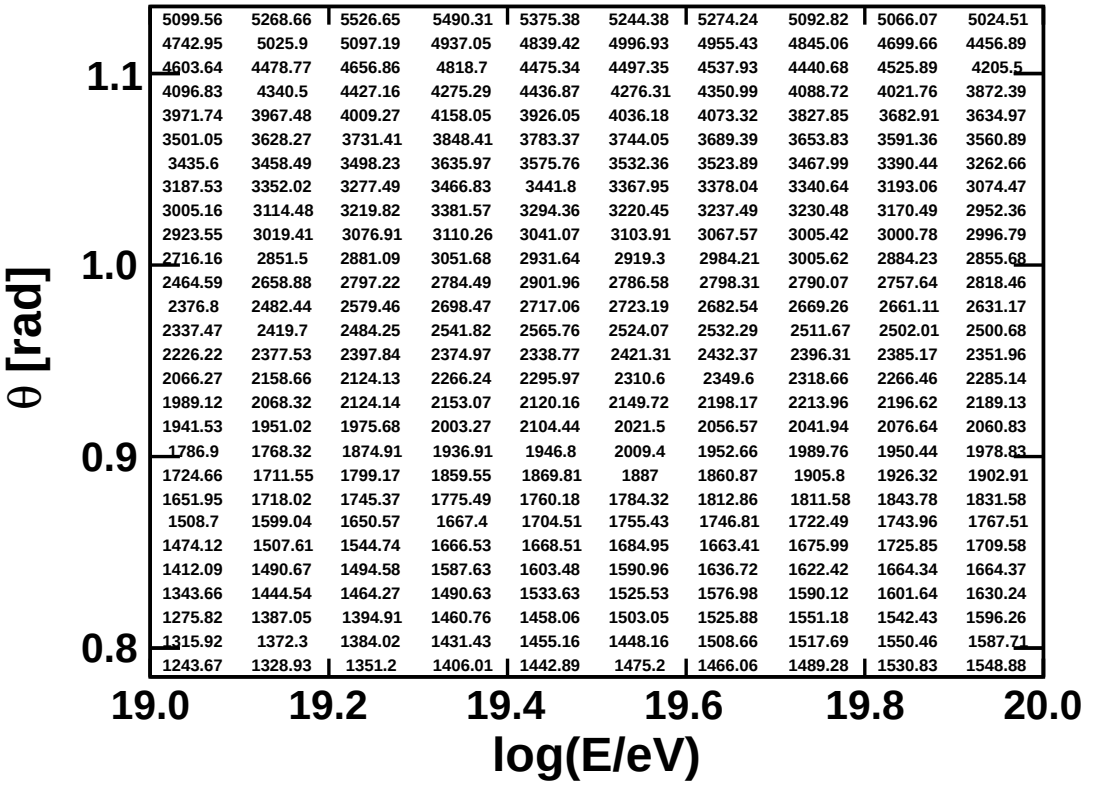


Figure 5.10: Numerical values of r_0 (in metres) used in the reconstruction. A bi-dimensional linear interpolation is used for a particular set of $(\log E, \theta)$.

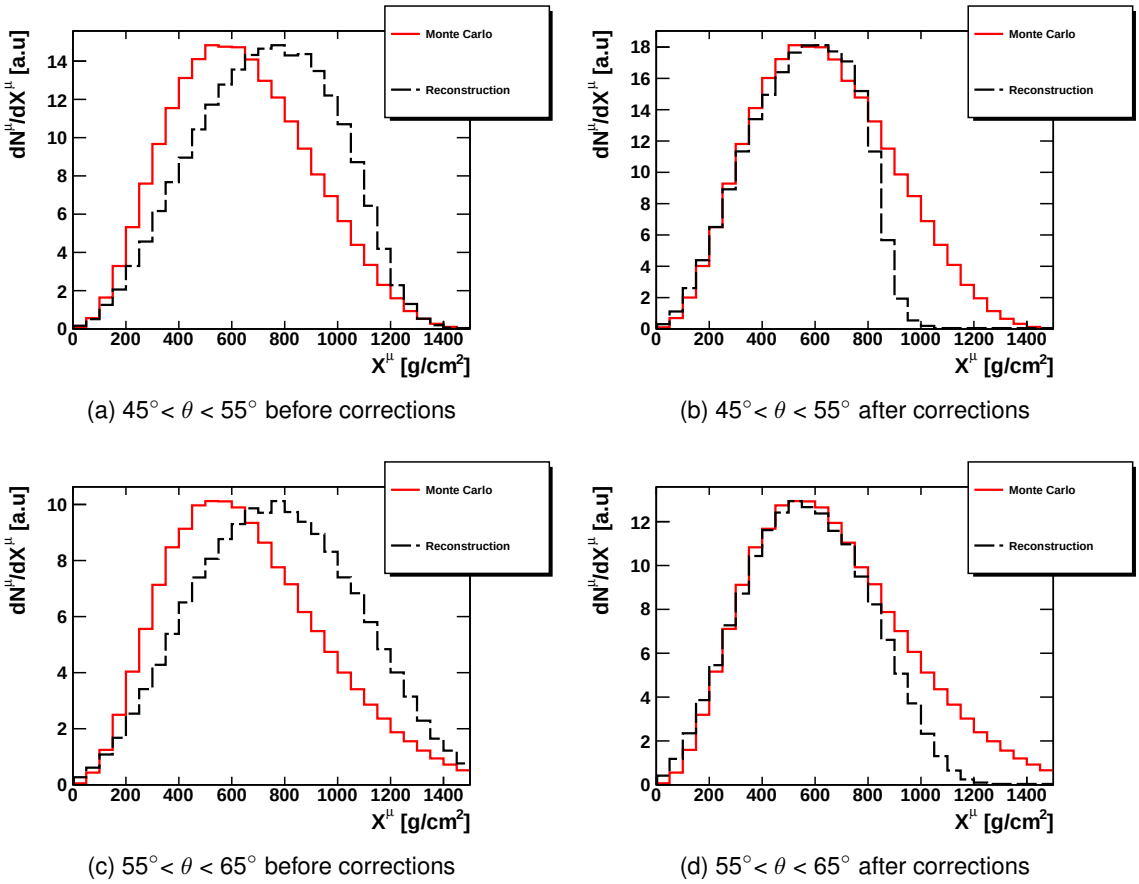
5.4 Final performance of the reconstruction

As mentioned before, the reconstruction of X^{μ}_{max} that will be introduced here is restricted to stations further than 1000 m and closer than 4000 m from the shower core. This represents a compromise between getting as close as possible to the shower core while avoiding as much electromagnetic contamination as possible. Only events above 10 EeV will be considered, as at lower energy the number of stations becomes too small. Finally, the zenith window of analysed events is [45°,65°].

The absolute bias in the reconstruction of X^{μ}_{max}, i.e., the difference between the reconstructed value and the true one, is below 10 g/cm^2 for the standard MPD analysis [132]. We will stick to the same requirement in this new approach. The result for the bias is shown in figure 5.12 divided in the two analysed angular regions. The bias is within ± 10 g/cm^2 for all the primaries and hadronic models in both regions, and so we have an approximately unbiased reconstruction of X^{μ}_{max}.

The resolution, determined as the standard deviation[1] of the distribution $X^{\mu}_{\text{max,rec}} -$

[1]It is customary in the field to refer to the standard deviation of an histogram as the *root mean square*



(a) 45°< θ < 55° before corrections

(b) 45°< θ < 55° after corrections

(c) 55°< θ < 65° before corrections

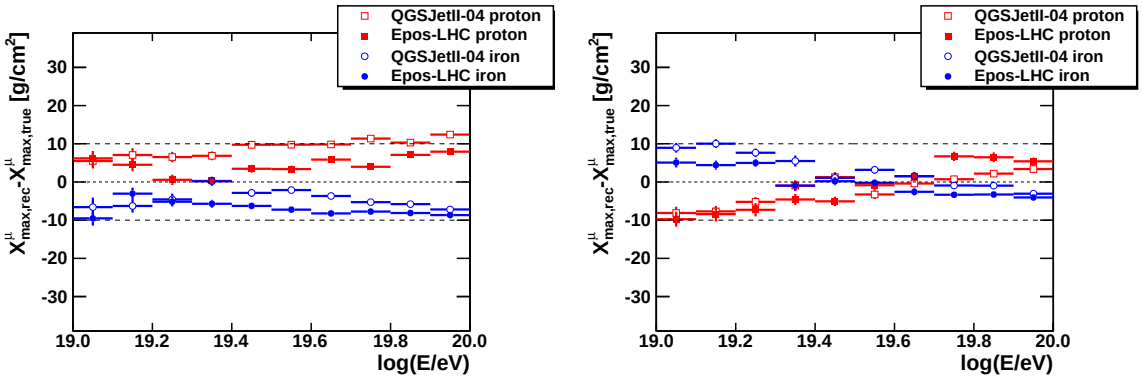
(d) 55°< θ < 65° after corrections

Figure 5.11: Effect of the radial correction and the electromagnetic signal removal on the average MPD profile for the two zenith windows.

$X^{\mu}_{\text{max,true}}$, is shown in figure 5.13. It exhibits a decreasing behaviour as a function of the primary energy. This was an a priori expectation, as the resolution is known to improve when the number of muons increases. For the same reason, the resolution is slightly better for events below 55°, as the number of muons entering the MPD in this window is also slightly larger, due to smaller atmospheric shielding power. Finally, the resolution is also better for iron primaries than for protons, and for EPOS-LHC than for QGSJETII-04, which was again expected due to a larger number of secondary muons.

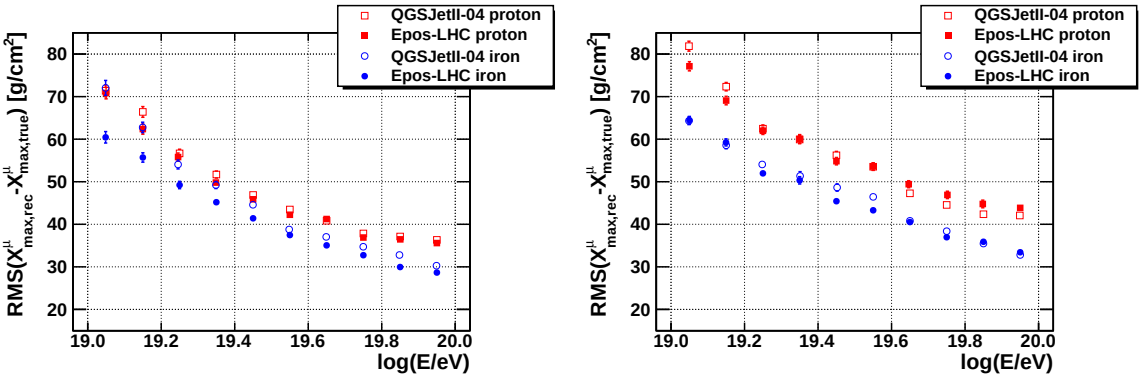
The improvement in the resolution as compared to the standard MPD reconstruction is shown in figure 5.14 for a particular choice of primary and hadronic model. The main reason for this improvement lays on the fact of introducing stations closer to the shower core, which increases the number of muons that are used to reconstruct the MPD. However, this improvement becomes much smaller at the highest energies. This is a consequence

(RMS). This is an abuse of language, as the two values only coincide for variables with null average: $\text{RMS}^2[x] = \langle x \rangle^2 + \sigma_x^2$, as is the approximate case for $x = X^{\mu}_{\text{max,rec}} - X^{\mu}_{\text{max,true}}$.



(a) 45°< θ < 55° (b) 55°< θ < 65°

Figure 5.12: Reconstruction bias for the two zenith windows.



(a) 45°< θ < 55° (b) 55°< θ < 65°

Figure 5.13: Resolution for the two zenith windows.

of the current limitations of the method. At the highest energies, the number of muons becomes larger, and different contributions to the resolution of X^{μ}_{max} get increasingly more relevant. In particular, the resolution in the reconstruction of the shower core poses a limit to the resolution that cannot be solved by a mere addition of muons.

5.4.1 Reconstruction under the standard conditions

In order to validate this approach to the MPD reconstruction, the usual set of parameters was used [89]. This means in practice using stations at 1700 m$< r <$4000 m and analysing events with $E > 10^{19.3}$ eV and $\theta \in$[55°,65°]. The result for the reconstruction bias and the resolution is shown in figure 5.15 for QGSJETII-04 protons. The bias in this context refers to the comparison with the Monte Carlo apparent MPD with

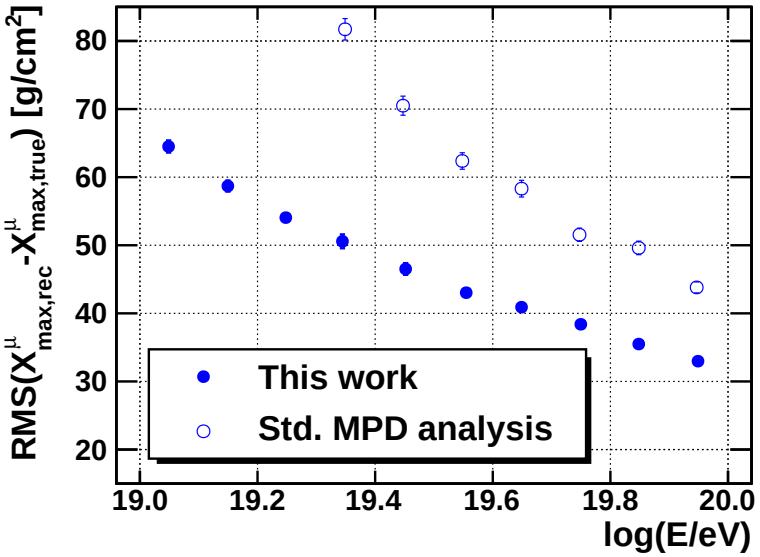


Figure 5.14: Comparison of the resolution for Epos-LHC iron under the standard MPD analysis and this approach.

$r \in [1700\ \text{m}, 4000\ \text{m}]$. The result is in very good agreement with the one obtained in the standard analysis. Different libraries were used to build both graphics, and hence the different error bars.

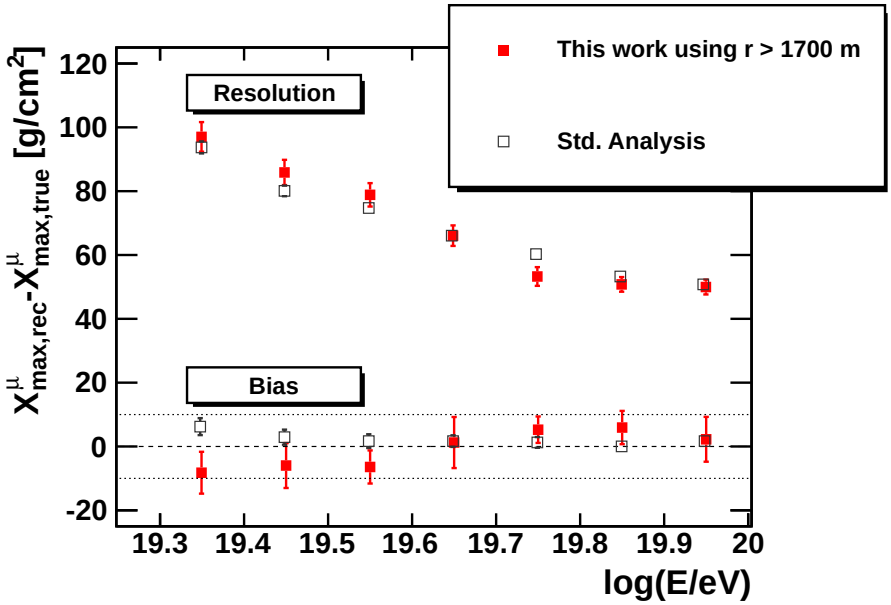


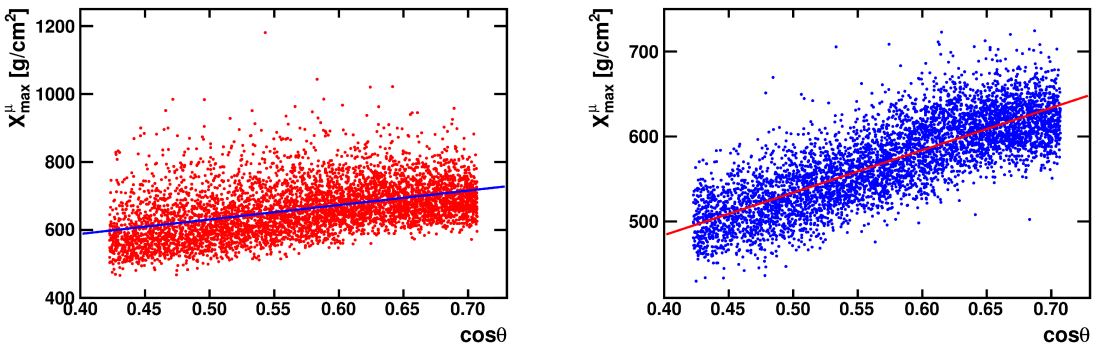
Figure 5.15: Reconstruction bias and resolution for QGSJETII-04 protons under the same conditions as in the standard analysis.

5.5 Unification of datasets: referring to 55°

In order to analyse the whole zenith window as a simple dataset, the dependence of X^{μ}_{max} on the zenith angle needs to be taken into account. It was found for simulations that a simple linear function in $\cos\theta$ correctly describes the dependence of X^{μ}_{max} on the zenith angle (see figure 5.16). However, this dependence is different depending on the atomic mass of the primary. An average value between the slopes obtained for protons and iron nuclei will be used, and the relative difference will be added to the systematics. Therefore, all the values of X^{μ}_{max} can be referred to the equivalent value at 55°, $X^{\mu}_{\text{max,55}}$, using this slope α_X:

$$\alpha_X \equiv \frac{dX^{\mu}_{\text{max}}}{d\cos\theta} \tag{5.4}$$

$$X^{\mu}_{\text{max,55}} \simeq X^{\mu}_{\text{max}}(\theta) - \alpha_X\left(\cos\theta - \cos 55^{\circ}\right) \tag{5.5}$$



(a) QGSJETII-04 protons

(b) QGSJETII-04 iron

Figure 5.16: Zenith dependence of X^{μ}_{max} and linear fit in $\cos\theta$.

The value of α_X for protons is 425 g/cm^2 and for iron 499 g/cm^2, so an average value of 462 g/cm^2 will be used.

5.6 Summary on the reconstruction of the MPD

In this section we will summarise briefly the different steps that are carried out to obtain $X^{\mu}_{\text{max,55}}$ for a given reconstructed event. Most of these steps are common to the standard MPD technique and are therefore described in [132]. We will emphasise those aspect in which this work differs from the standard analysis.

1. All the candidate stations of the event between 1000 m and 4000 m from the shower core are used in the analysis. Stations with signal smaller than 3 VEM are rejected, to minimise the influence of accidental signals.

2. The traces are corrected for the effect of direct light using the algorithm described in [138] to avoid artificial peaks.

3. Compute i_{max} and S_{thr} for the station using equations (5.1) and (5.3).

4. For each of the selected bins, i, obtain the arrival time t as $t = t_0 + (i + 1/2 - i_0) \times 25$ ns $- t_s$ where t_0 is the arrival time of the shower plane, i_0 the start bin of the trace and $t_s = 70$ ns the time shift. The arrival time is referred to the centre of the time bin, and hence the factor $1/2$.

5. Assume that all the time delay comes from the geometrical delay and use (3.23) to obtain an approximate production distance z.

6. Use this value of z to calculate the kinematical delay from the parametrisation given in [136] $t_\varepsilon\,(r, z)$.

7. The time t is corrected by this delay and equation (3.24) is used to calculate z again.

8. This value is corrected for the radial dependence using equation (4.5). As this correction is written in terms of z_v, it is necessary to use equations (4.2) and (4.3) as intermediate steps.

9. X^μ for each time bin is computed by integrating the atmospheric density along the shower axis starting from z (equation (3.25)).

10. The set of X^μ are stored in a histogram, giving the MPD as measured by a single PMT. The weights of this histogram are the signal in VEM of each bin divided by an estimate of the relative average track-length of muons in the tanks for the given zenith angle

$$\frac{1}{L_\mu} = \cos\theta + \frac{2h\sin\theta}{\pi R} \tag{5.6}$$

where $h = 1.2$ m and $R = 1.8$ m are the height and radius of the Cherenkov tanks. This accounts for the different amount of Cherenkov light that is produced by muons depending on the amount of water they traverse. This factor goes from 1.01 at 45° down to 0.81 at 65° and is a mere constant factor that does not affect the reconstruction of X^μ_{max}. It is included, however, to keep the proportionality between the integral of the MPD and the number of muons.

11. The MPD of a given station is obtained as the average of the MPDs obtained by each of its functioning PMTs.

12. Finally, the MPD of the event is built by adding the individual MPDs of all the stations.

13. The final histogram is fitted to a Gaisser-Hillas function (equation (2.4)) in which X_{max} is replaced by X^μ_{max}. To assure convergence, the different parameters are given seeds and variation ranges, which are:

- X_0^μ is fixed to -70 g/cm^2, the average value for proton and iron simulations. The small amount of muons entering the histogram is too small to get a good reconstruction of this parameter, so it cannot be left free.
- The seed to X_{max}^μ is set to 600 g/cm^2. It is left to vary between 100 and 1200 g/cm^2.
- λ is set to 100 g/cm^2 and restricted to values between 10 and 200 g/cm^2.

14. Finally, the value of X_{max}^μ obtained by the previous fit is referred to 55°, using equation (5.5) to get $X_{\text{max,55}}^\mu$.

5.7 Reconstruction of experimental data

The SD data of the Pierre Auger Observatory between the 1$^{\text{st}}$ of January 2004 and the 30$^{\text{th}}$ of April 2013 will be analysed following the prescriptions stated above. This period comprehends the current set of analysable data, starting with the beginning of the data-taking era and ending with the latest available atmospheric databases.

Only events above 10 EeV and with $\theta \in$[45°,65°] are considered. Furthermore, a series of quality cuts were performed. These cuts and their selection efficiency are summarised in table 5.1.

Quality cut	**No. of Events**	**Efficiency**
Energy and zenith ranges	6880	100%
Reject lightning events	6878	99.97%
Reject bad periods	6346	92.24%
Trigger 6T5	4100	59.59%

Table 5.1: Quality cuts applied to data and their relative selection efficiency.

Only events passing the 6T5 trigger will be used, to minimise the influence of the edges of the array. This is the most stringent quality cut. Rejecting lightning events and known bad periods of the detector is required to only analyse physically significant events, and has minimum effect on the resulting statistics.

The evolution of $\left\langle X_{\text{max,55}}^\mu \right\rangle$ as a function of the primary energy is shown in figure 5.17 together with the Monte Carlo predictions for QGSJETII-04 and EPOS-LHC. Error bars represent the statistical uncertainty, and the numbers below the squares are the number of events entering that bin (for a longer discussion on the determination of statistical parameters, see Appendix A). The shaded area represents the systematic uncertainty, that will be discussed in the following section.

Bins in $\log E$ in figure 5.17 have a width of 0.1 in units of $\log E$ up to $\log E = 19.7$. Above this value, there is not enough statistics to keep this binning, so all the events above this energy are represented by the last point. For every energy bin, the average value of $\langle \log E \rangle$ is computed and used as the abscissa for the points. The numerical values of this figure are summarised in table 5.2.

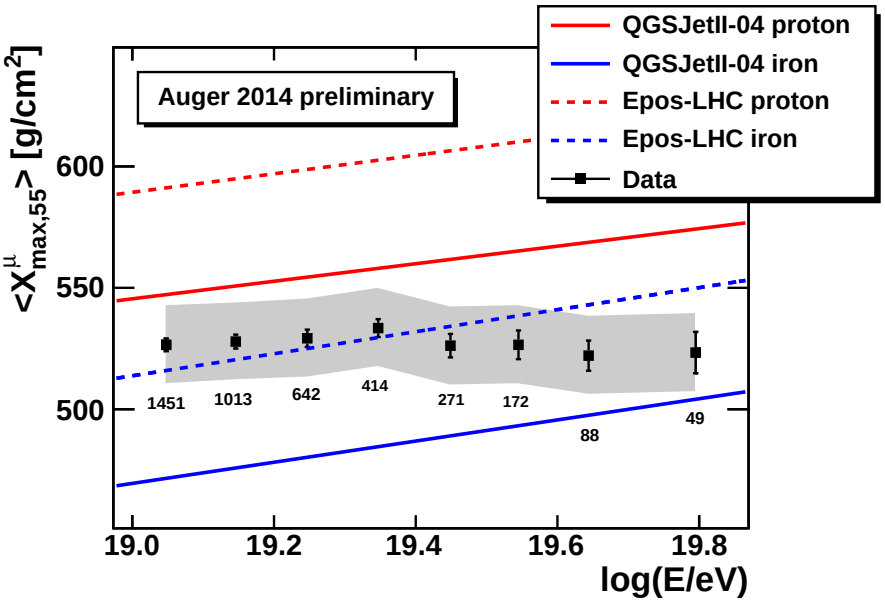


Figure 5.17: Evolution of $\left\langle X^{\mu}_{\text{max,55}} \right\rangle$ as a function of the primary energy for the data of Pierre Auger Observatory compared to Monte Carlo simulations.

$\langle \log(\mathbf{E/eV}) \rangle$	**Entries**	$\left\langle \mathbf{X}^{\mu}_{\text{max,55}} \right\rangle$	$\mathbf{\Delta} \left\langle \mathbf{X}^{\mu}_{\text{max,55}} \right\rangle$
19.048	1451	526.536	2.622
19.146	1013	527.860	2.875
19.247	642	529.223	3.577
19.347	414	533.421	3.692
19.449	271	526.179	4.845
19.545	172	526.533	5.915
19.644	88	522.070	6.213
19.795	49	523.320	8.512

Table 5.2: Numerical values of the evolution of $\left\langle X^{\mu}_{\text{max,55}} \right\rangle$ (in g/cm^2) with $\langle \log E \rangle$.

The interpretation of the data shown in figure 5.17 is very much dependent on the hadronic model. If QGSJETII-04 is chosen, the data exhibit some trend from light-dominated to heavier composition as the energy rises, which is in agreement with the observations of the FD. However, data are difficult to accommodate to the description of EPOS-LHC, as they fall below the average value for protons at the highest energies. This would mean that the average composition is made up of nuclei lighter than protons, which is obviously forbidden. Even though this discrepancy has not much statistical significance with the current systematic uncertainties, this figure indicates how the MPD method can be used to constraint high energy hadronic models.

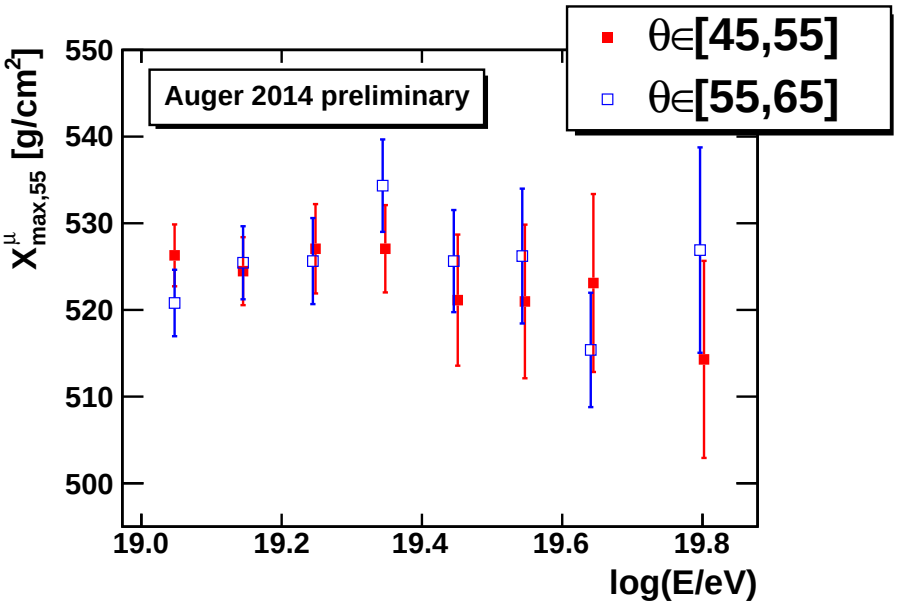


Figure 5.18: $\left\langle X^{\mu}_{\text{max,55}} \right\rangle$ as a function of the primary energy for both zenith windows.

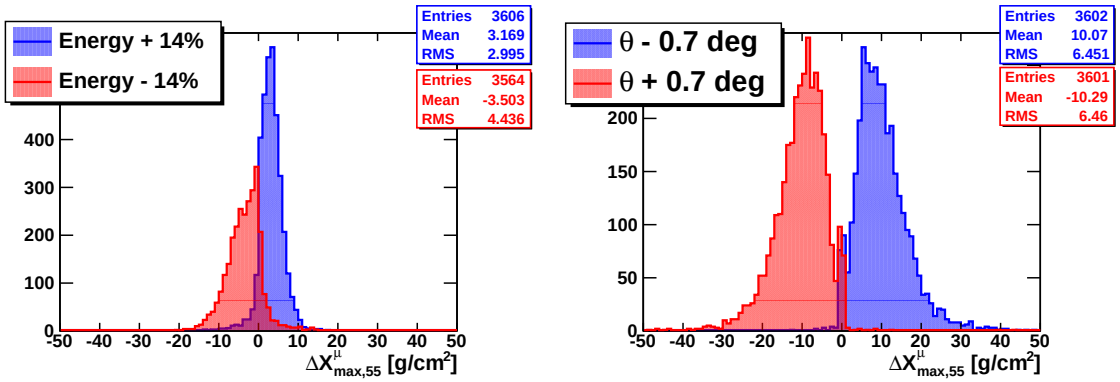
Figure 5.18 shows the result of the evolution of $\left\langle X^{\mu}_{\text{max,55}} \right\rangle$ as a function of the primary energy for the two zenith windows. This demonstrates that the two datasets are equivalent and can be merged safely. Therefore, the interpretation of 5.17 is independent of the extensions to the analysis that we have introduced in this thesis. In chapter 7, a comparison of this result with the standard MPD analysis in terms of mass composition will be shown.

5.7.1 Systematic uncertainties

One of the largest contributions to the systematic uncertainty comes from the difference between proton and iron showers. It has been shown how the reconstruction bias is below 10 g/cm^2 for both primaries regardless of the hadronic model, so the systematic uncertainty coming from the unknown nature of the primaries is quantified as 10 g/cm^2. This value also includes the uncertainties associated to the hadronic interaction models, as the results are consistent for all models than have been considered.

Another contribution comes from the tabulation of r_0 (figure 5.10) in terms of the energy and zenith angle of the shower. The resolution of the energy is 14%, and for the angular region and energies of interest, the resolution of the zenith angle is 0.7°. Figure 5.19 shows the effect of changing r_0 according to these two resolutions. The systematic uncertainty related to the energy resolution is 4 g/cm^2, and the one associated to the zenith resolution is 10 g/cm^2.

Of less importance is the contribution due to seasonal effects. These include changes in the atmosphere as well as in the physical properties of the SD. Figure 5.20a shows that the value of this contribution to the systematic uncertainty is 5 g/cm^2. No significant effect



(a) Energy resolution

(b) Zenith angle resolution

Figure 5.19: Systematic uncertainties associated to the energy (left) and zenith angle resolution (right).

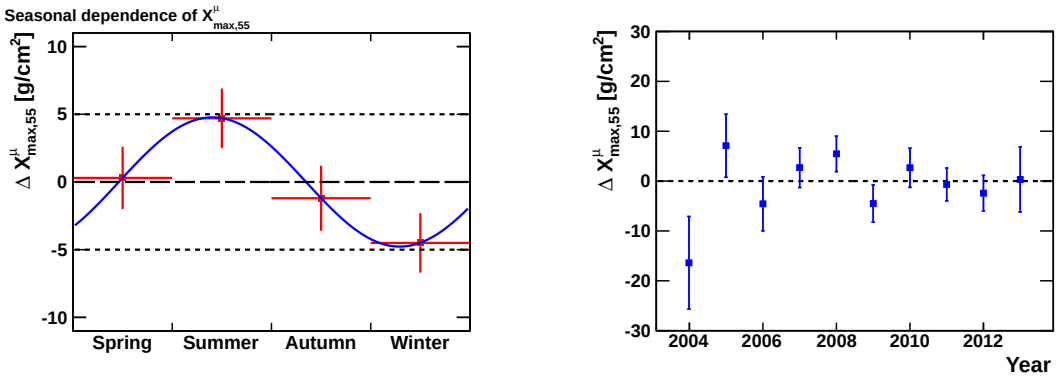
due to the ageing of the detector is observed (figure 5.20b). These two contributions can be checked simultaneously by fitting the value of $X^{\mu}_{\text{max,55}}$ over time with the combination of a linear function and a cosine function with a period of one year. This is shown in figure 5.21. The value of the slope (-0.7 ± 0.6 g/cm^2 per year) is compatible with no ageing effect, while the amplitude of the cosine function (3.3 ± 1.5 g/cm^2) is compatible with the estimate shown in 5.20a, which will be taken as a conservative estimate of the seasonal effect.

The choice of the binning of the histogram, as well as the choice of the seed parameters of the fit have negligible influence, below 1 g/cm^2. The different value of α_X for proton and iron introduces a maximum change in the value of $X^{\mu}_{\text{max,55}}$ when the value of the zenith angle is 45°:

$$\Delta_{\text{max}} X = \frac{\alpha_{X,\text{p}} - \alpha_{X,\text{Fe}}}{2} (\cos 45^\circ - \cos 55^\circ) \approx 5 \text{ g/cm}^2 \tag{5.7}$$

Finally, the effect of accidental signals needs to be taken into account. In real events there is a background of random accidental signals, mostly generated by isolated atmospheric muons. These accidental particles can reach the SD array in any position at any time, uncorrelated with physical showers. Therefore, there is a non-zero probability that one or more of these accidental particles are close enough in time to a physical event to be registered together with the event in the FADC traces.

Random accidental signals can damage the quality of the data. In particular, they can trigger some stations of the array or affect the estimate of the start time of the traces. We analysed the impact of these accidental signals using an unbiased sample of random accidental signals extracted from real events [139] and plugging them inside Monte Carlo simulations. The, the average value of $X^{\mu}_{\text{max,55}}$ was compared between simulations including the effect of systematics and simulation not including them. We found a systematic underestimate for $X^{\mu}_{\text{max,55}}$ of 3 ± 1 g/cm^2, regardless of the primary and hadronic model. We will consider this as a systematic bias of the reconstruction, and correct for it



(a) Seasonal effect (b) Ageing effect

Figure 5.20: Systematic uncertainty associated to seasonal effects (left) and to the ageing of the detector (right). In both cases, the deviation from the average value is depicted as a function of time.

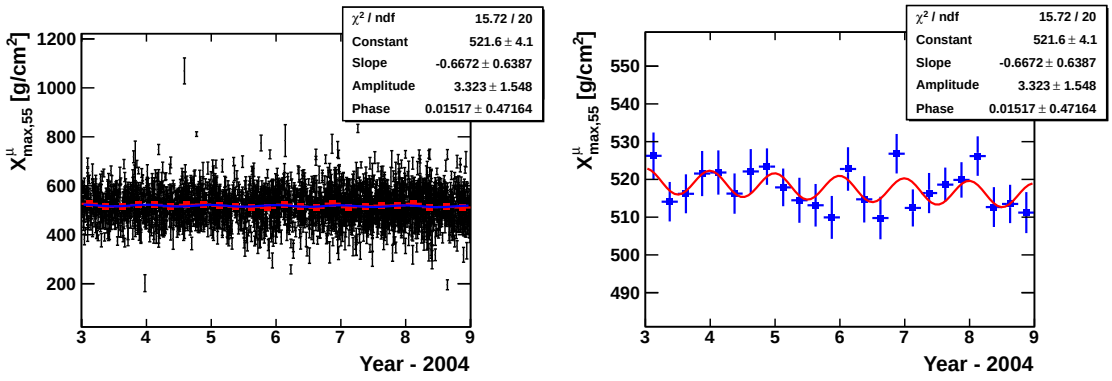


Figure 5.21: Fit of $X^{\mu}_{max,55}$ over time for events detected after the completion of the array (left). Zoom to the profile for better visibility (right).

in the real data. The uncertainty of this correction, obtained by propagation of errors of the difference between two average values, will be added to the systematics.

The full list of contributions to the systematic uncertainty is shown in table 5.3, adding up for a total systematic uncertainty of 16 g/cm^2.

Source	**Error (g/cm^2)**
Hadronic model, primary and reconstruction	10
Zenith angle resolution	10
Energy resolution	4
Seasonal effect	5
Refer to 55°	5
Fitting procedure	1
Accidentals estimate	1
Total:	**16**

Table 5.3: Contributions to the systematic uncertainty of $\left\langle X^{\mu}_{\text{max,55}} \right\rangle$.

Whenever people agree with me I always feel I must be wrong

Oscar Wilde

6

Measuring shower-to-shower fluctuations with the Surface Detector

The aim of this chapter is measuring the evolution with the energy of shower-to-shower fluctuations using only SD data. The main caveat in order to do that is correctly accounting for the detector resolution contributions to the total fluctuations. In fact, for any observable, its fluctuations for a given energy bin satisfy

$$\sigma^2_{\text{obs}} = \sigma^2_{\text{sh}} + \sigma^2_{\text{det}} \tag{6.1}$$

where σ^2_{obs} is the observed variance for the considered observable, σ^2_{sh} the shower-to-shower fluctuations regarding this observable and σ^2_{det} the detector resolution when measuring it. This means that the observed fluctuations will always be *larger* than the real shower-to-shower fluctuations, due to detector effects. For a thorough discussion on this matter check Appendix B.

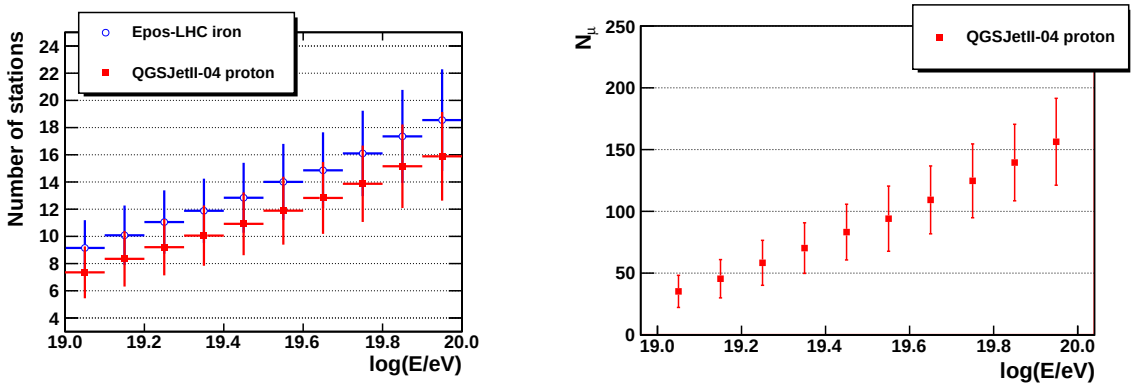
In this work we will concentrate on the measurement of shower-to-shower fluctuations related to X^{μ}_{max}. In particular:

$$\sigma_{\text{obs}} = \sigma\left(X^{\mu}_{\text{max,55}}\right) \tag{6.2}$$

the standard deviation of the distribution of the reconstructed values of X^{μ}_{max}.

Under the conditions of the standard X^{μ}_{max} analysis, this procedure is not possible, as the event statistics and their multiplicity (total number of accepted stations per event) is too small [132]. Therefore, only by taking advantage of the extensions to the method shown in the previous chapter we will be able to correctly determine the resolution of the detector. Figure 6.1a shows the average number of stations used to reconstruct the MPD under our conditions for the two extreme cases: the model and primary giving the largest (Epos-LHC iron) and the smallest (QGSJetII-04 proton) number of muons. Also,

the number of muons as a function of the primary energy is shown in figure 6.1b for QGSJetII-04 protons.



(a) Number of selected stations as a function of energy

(b) Number of muons as a function of energy

Figure 6.1: Number of stations (left) and muons (right) used to reconstruct the MPD as a function of the energy. Error bars represent the standard deviation.

As in the case of the muonic elongation rate, we will refer all values of X^{μ}_{max} to the equivalent value at 55°, $X^{\mu}_{\text{max,55}}$, so we can merge the whole dataset. In the following, when we allude to X^{μ}_{max}, we will always mean the maximum of the *apparent* MPD for $r > 1000$ m and $r < 4000$ m.

6.1 Determination of detector resolution

The general strategy to determine the resolution of a detector relies on having two (or more) simultaneous measurements of the same event. For example, in the FD, this is achieved using stereo hybrid events, for which two independent measurements of X_{max} are available [3]. This approach is clearly not feasible for the SD, as there is only one array for a given event.

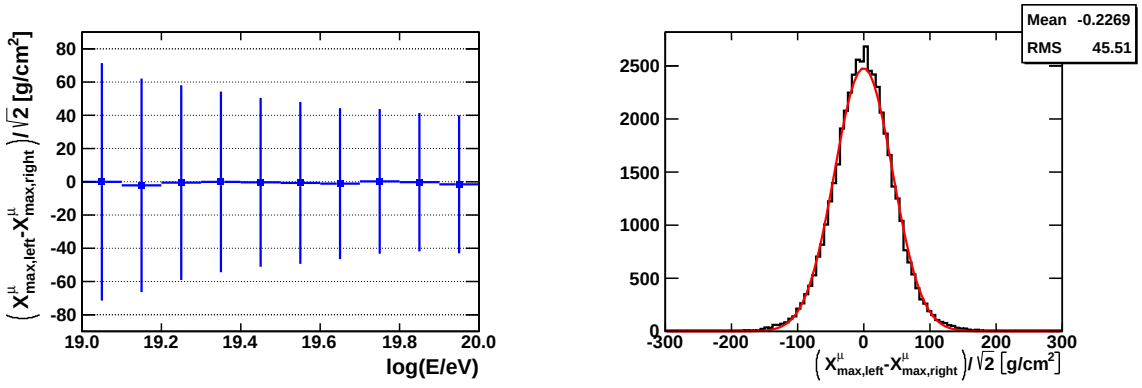
However, for observables that are built-up through the addition of several stations, such as the MPD, events with a large enough multiplicity allow the subdivision of accepted stations in two (or more) sets (see figure 6.1a). For example, taking advantage of azimuthal symmetry, two MPDs can be built for an event depending on the azimuthal angle, ϕ, of the stations in the shower coordinates

$$\text{if } \phi < 0 \longrightarrow \text{MPD “left”}, \qquad \text{if } \phi > 0 \longrightarrow \text{MPD “right”} \tag{6.3}$$

The detector resolution, σ_{det}, is defined as:

$$\sigma_{\text{det}} \equiv \sigma\left(X^{\mu}_{\text{max,rec}} - X^{\mu}_{\text{max,true}}\right) \tag{6.4}$$

the standard deviation of the distribution of the difference between the reconstructed value of X^{μ}_{max} and the real one. As in real data this difference is not accessible (the true value is unknown), the hypothesis will be that the resolution can be estimated through the difference between $X^{\mu}_{\text{max,left}}$ and $X^{\mu}_{\text{max,right}}$, i.e., the values of X^{μ}_{max} obtained for each of the two MPDs. The distribution of these differences is centred at zero by construction (see figure 6.2a) given the azimuthal symmetry and is well described by a gaussian function (see figure 6.2b).



(a) "Left-right" as a function of energy

(b) Distribution at $E = 10^{19.65}$ eV

Figure 6.2: Behaviour of the "left-right" difference for QGSJet-II iron simulations with $\theta \in$[55°,65°]. Error bars on the left-side plot represent the standard deviation.

The main caveat, however, has to do with the fact that the two histograms tagged as MPD "left" and "right" are built with less muons than the global MPD (see figure 6.1b). In fact, the complete MPD for the event is obtained simply by adding both histograms and therefore each of them carries, on average, one half of the total muons. The resolution of the measurement of X^{μ}_{max} is known to decrease as the square root of the number of muons, and hence the statistical factor relating both distributions

$$\sigma_{\text{det}} = \sigma\left(X^{\mu}_{\text{max,rec}} - X^{\mu}_{\text{max,true}}\right) \approx \sigma\left[\left(X^{\mu}_{\text{max,left}} - X^{\mu}_{\text{max,right}}\right)/\sqrt{2}\right] \tag{6.5}$$

This assumption can be tested by analysing different azimuthal divisions of the stations. For instance, instead of just dividing in two halves, it's possible to subdivide in three regions $[-\pi/3, \pi/3]$, $[\pi/3, \pi]$ and $[-\pi, -\pi/3]$. However, in this case the early-late asymmetry of the shower makes the first sector statistically more populated, and therefore only the two late sectors can be considered as equivalent. Analogously, it's possible to subdivide into four sectors, two of them in the "early" part of the shower and the other two in the "late part" (see figure 6.3). Of course, there are a handful of other possibilities, obtained by rotations of these two set-ups, but these will suffice for our purpose.

If a and b are the two MPDs to be compared, obtained by two equivalent sectors of the whole azimuthal distributions, then the statistical factor relating their standard deviation

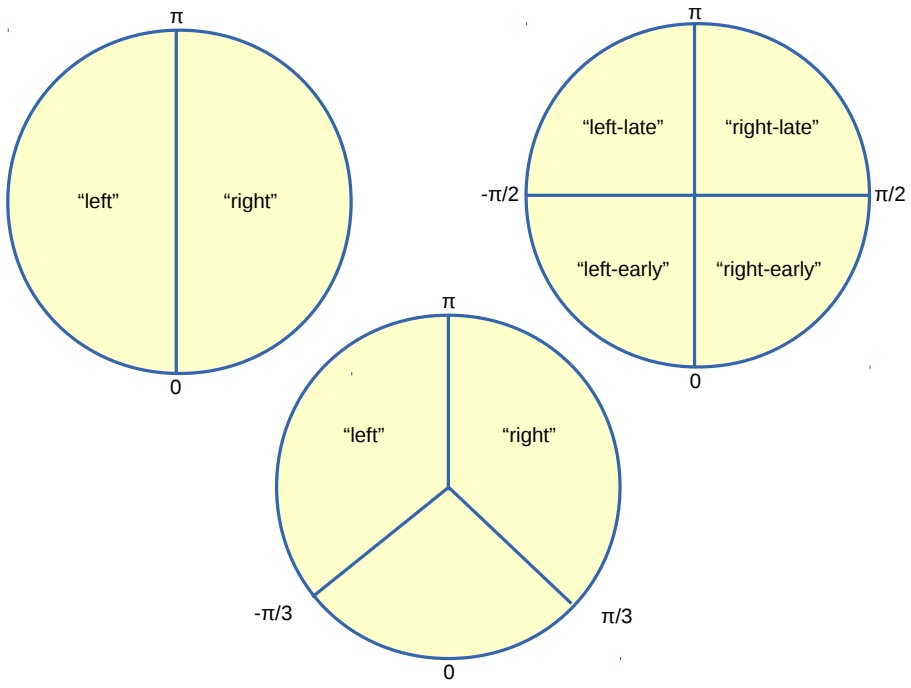


Figure 6.3: Different division schemes in two, three or four sectors of the azimuthal distributions and selected equivalent sectors.

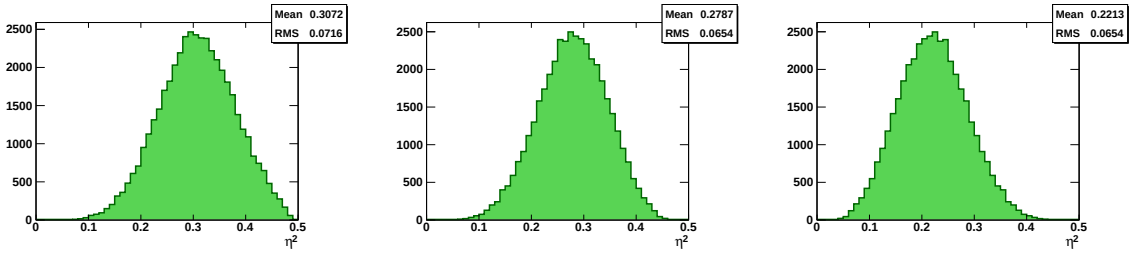
and the resolution, η, is[1]

$$\eta = \sqrt{\frac{\left(N_a^\mu + N_b^\mu\right)/2}{N^\mu}} \tag{6.6}$$

where N_a^μ and N_b^μ are the number of muons entering the histogram of a and b respectively, and N^μ the total number of muons. These quantities are actually unknown, but they are proportional to the integral of the histograms. Trivially, for the case of two halves, $N_a + N_b = N$ and hence $\eta = \sqrt{1/2}$, for the other three scenarios this number is smaller and changes from one event to another (see figure 6.4).

Figure 6.5 shows how the different division schemes lead to very similar results within about ± 5 g/cm^2. However, at the lowest energies the number of stations is too small to perform a good reconstruction when too many divisions are made, leading to just one or two stations per division (see figure 6.1a). All schemes converge at the highest energies, when the number of muons gets larger. Therefore, the final choice, for simplicity and for its better behaviour (minimal division scheme) would be to divide the stations in two halves, as established in expression (6.3). The assumption made on (6.5) was tested, and the results are shown in figure 6.6 for different high-energy hadronic models and primaries. In every case there is a small mismatch between the actual resolution and its estimate, but the error is smaller than 3 g/cm^2, and will just be taken as a systematic. As

[1]This expression was chosen as the average ratio of $N_a^\mu + N_b^\mu$ over N^μ. Choosing e.g. N_a^μ/N^μ works very similarly, the key feature being the dependence on the square root of the number of muons



(a) Three divisions (b) Four divisions (early) (c) Four divisions (late)

Figure 6.4: η^2 for the different division schemes. $\eta^2 = 0.5$ for just one division into two halves.

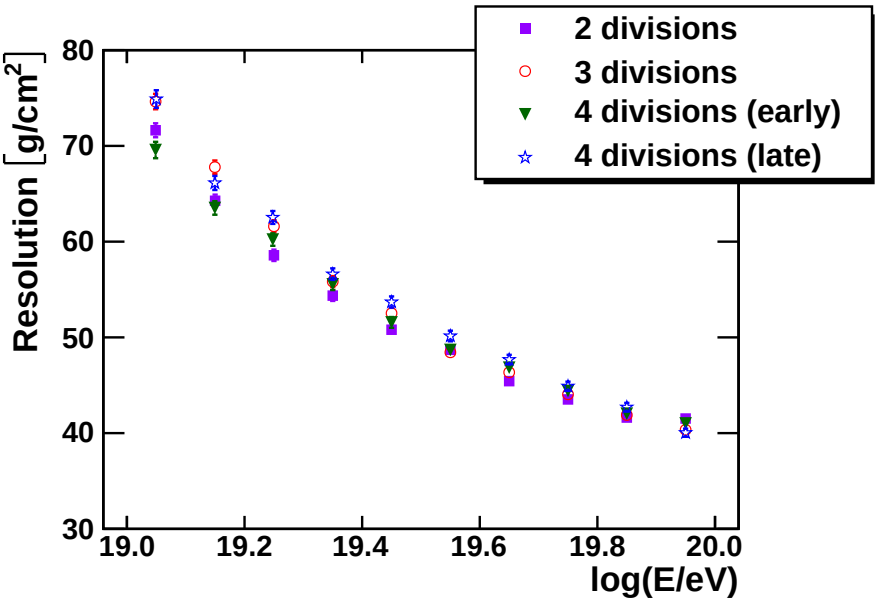
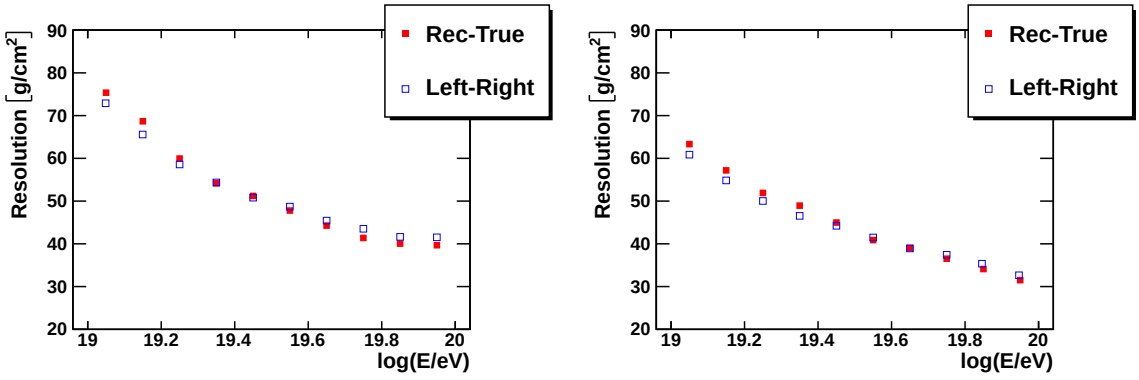


Figure 6.5: Comparison of the resolution obtained using the different division schemes shown in figure 6.3 for QGSJetII-04 protons.

shown in figure 6.7, the detector resolution depends slightly on the primary (the global difference being around 10 g/cm^2), but not on the hadronic model.

6.2 Application to experimental data

The method described above has the advantage of not having to rely on an accurate estimate of the number of muons by simulations. Moreover, by comparing two equivalent measurements of the same observable, many of the systematics disappear and so this technique is well suited to be applied to data (a more thorough discussion on systematics



(a) QGSJetII-04 proton

(b) QGSJetII-04 iron

(c) Epos-LHC proton

(d) Epos-LHC iron

Figure 6.6: Performance of the model to estimate the resolution, equation (6.5).

will be covered on the final part of this chapter).

Our sample covers data collected from January 2004 until April 2013, excluding bad periods. Lightning events are rejected, and only 6T5 events are considered. The quality cuts are the same as in the previous chapter (see table 5.1), so in total we have 4100 events above 10^{19} eV, as seen in figure 5.17.

Figure 6.8 shows the result of the estimate of the resolution for data as compared to proton and iron simulations. The data are well bracketed by the Monte Carlo results, and so we can consider the model to be realistic and well described by simulations. Furthermore, this figure already points to a transition from light to heavy composition as the energy increases. Error bars, however, are too large to make a claim from this figure.

There are in principle two possibilities in order to obtain σ_{sh} from equation (6.1). The fist one would be to use a value for the resolution as the average between proton and iron, and take the difference as a systematic. This is the approach followed by the FD

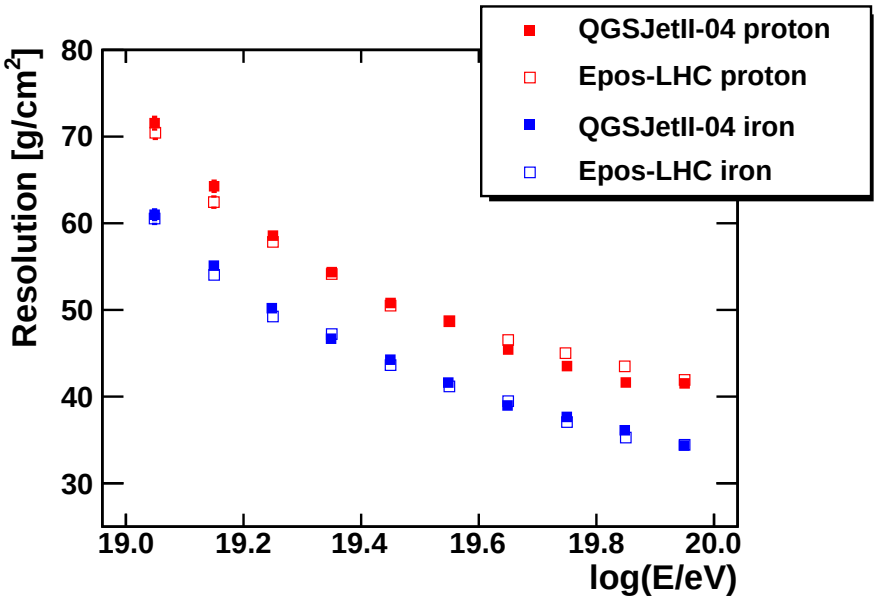


Figure 6.7: Comparison of the resolution obtained for different hadronic models and primaries.

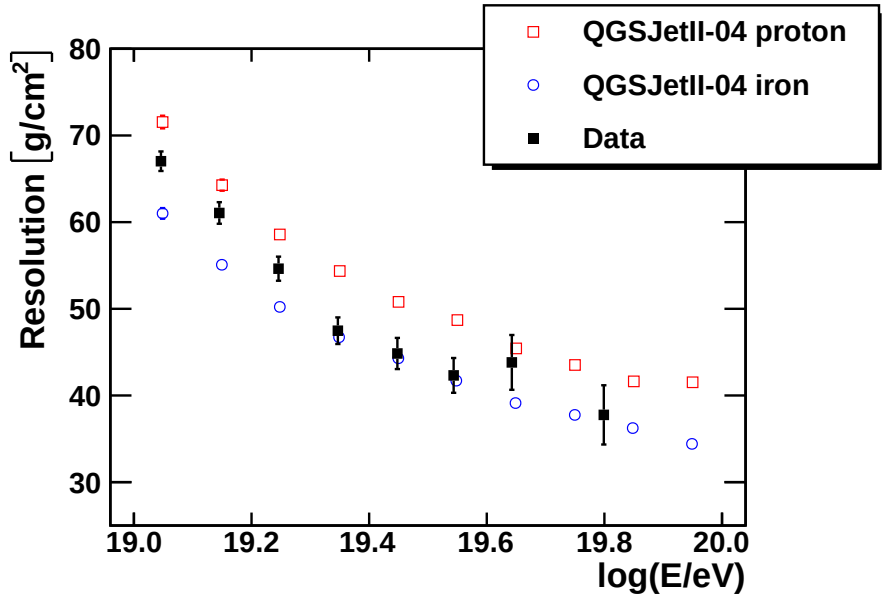


Figure 6.8: Comparison of the resolution obtained for simulations and real data.

[140, 3]. The second one will be our choice, and consists on taking directly the measurement of the resolution (black squares on figure 6.8) as obtained for data. It is clear that the difference between both approaches is smaller than 5 g/cm^2, as the total difference between primaries is of the order of 10 g/cm^2.

The result of σ_{sh}, in this scenario calculated as $\sigma\left[X^{\mu}_{\text{max,55}}\right]$ as compared to simulations is shown in figure 6.9, where the error bars represent the statistical uncertainties and the shaded area the systematic uncertainties (for a thorough discussion on the calculation of statistical parameters see Appendix A). To obtain $\sigma\left[X^{\mu}_{\text{max,55}}\right]$, the detector resolution was obtained using (6.5) and subtracted to the observed values of the standard deviation in quadrature, as established by equation (6.1). The calculation of σ_{sh} and its errors is performed following the same prescriptions as in [140, 3], and hence the binning choice (the number of events on each bin must be larger than 30 to ensure an approximately unbiased reconstruction of the variance). The cumulative bin starts at $\log(E/\text{eV}) > 19.7$ and has 49 events, as shown in figure 5.17.

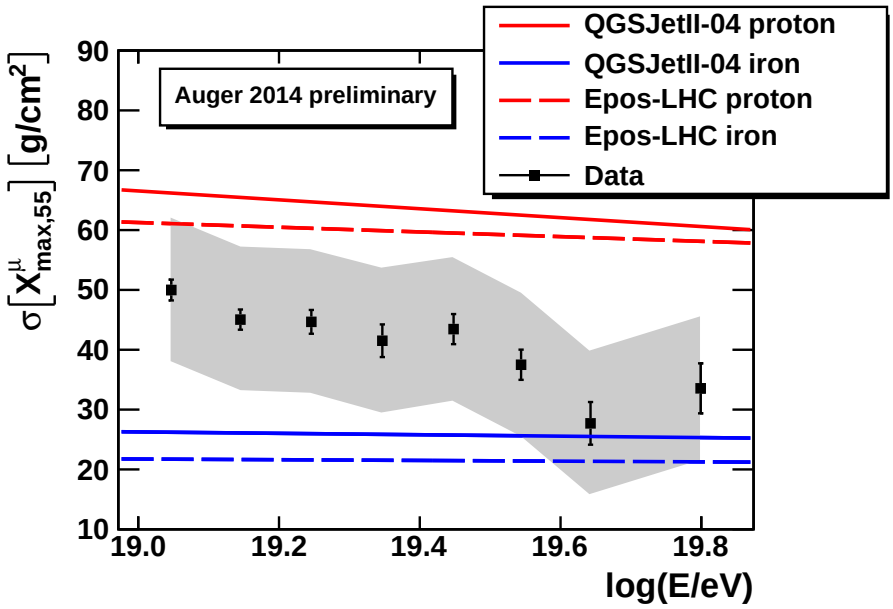


Figure 6.9: $\sigma\left[X^{\mu}_{\text{max,55}}\right]$ as a function of the energy for real data and simulations.

6.2.1 Systematic uncertainties

In order to account for systematic uncertainties, a conservative approach has been followed. For instance, regarding the energy scale, the energy of every event is multiplied (or divided) by 1.14, consistent with the 14% uncertainty of the energy scale, and the change on the spread $\sigma\left[X^{\mu}_{\text{max,55}}\right]$ is evaluated. The approach will be to consider that this change in σ_{obs} is translated directly into σ_{sh}, although we know that σ_{det} will change accordingly, as they are both correlated.

The contribution to the systematic uncertainty due to the energy and zenith resolution is shown in figure 6.10 for the different energy bins. The effect of the energy resolution is negligible, below 1 g/cm^2, while for the zenith angle resolution the systematic uncertainty is of 10 g/cm^2. This can be intuitively understood, as r_0 is nearly independent of the

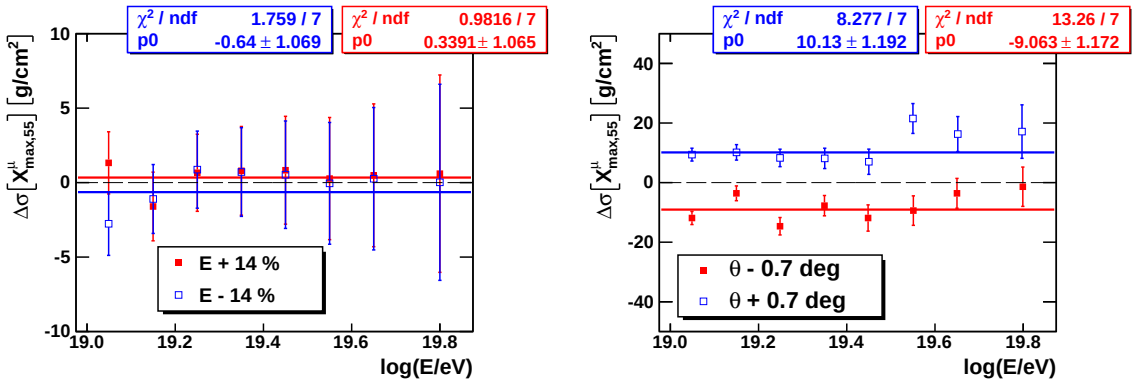
primary energy but depends strongly on the zenith angle, as in the case of $\left\langle X^{\mu}_{\text{max,55}} \right\rangle$.

The systematic uncertainty introduced by referring all X^{μ}_{max} to θ =55° was obtained by calculating the spread of $X^{\mu}_{\text{max,55}}$ using the slope α_X for protons and for iron nuclei and compare them with the one obtained using the average slope. The maximum difference was of 3 g/cm^2.

No significant influence of seasonal effects or of the ageing of the detector were found (see figure 6.11). The list of systematics is summarised on table 6.1, adding up to a total systematic uncertainty of 12 g/cm^2.

Source	**Error (g/cm^2)**
Primary	5
“Left-right” method	3
Energy resolution	1
Zenith resolution	10
Refer to 55°	3
Global	**12**

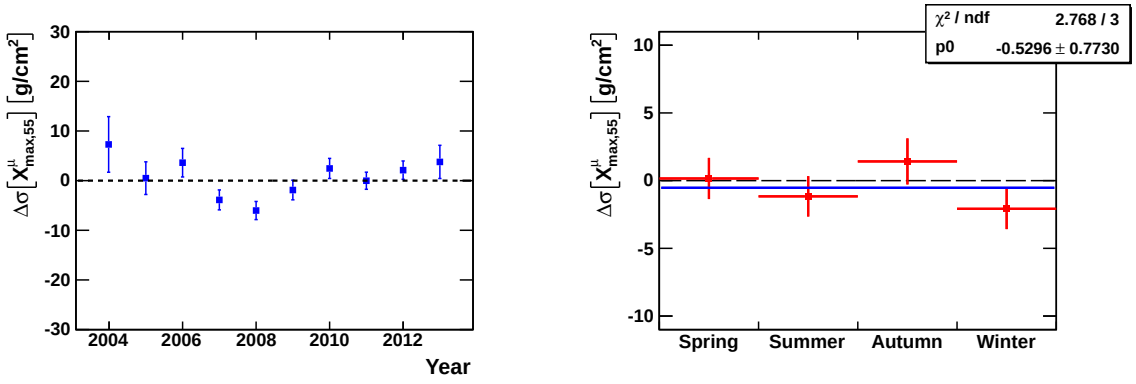
Table 6.1: Contributions to the systematic uncertainty of $\sigma\left[X^{\mu}_{\text{max,55}}\right]$



(a) Energy resolution

(b) Zenith resolution

Figure 6.10: Contributions to the systematic uncertainty in $\sigma\left[X^{\mu}_{\mathrm{max},55}\right]$ relative to the energy (left) and zenith resolution (right).



(a) Ageing

(b) Seasonal effects

Figure 6.11: No contributions to the systematic uncertainty in $\sigma\left[X^{\mu}_{\mathrm{max},55}\right]$ were found for the ageing of the detector (left) nor the seasonal effects (right).

Do not judge me by my successes, judge me by how many times I fell down and got back up again

Nelson Mandela

7 Conclusions and future prospects

In the previous chapters we were able to measure the evolution of $\left\langle X^{\mu}_{\text{max,55}} \right\rangle$ and $\sigma\left[X^{\mu}_{\text{max,55}}\right]$, i.e., the first two moments of the distribution of $X^{\mu}_{\text{max,55}}$, as a function of the primary energy using data of the SD. In this chapter we will try to interpret these measurements in terms of mass composition. This analysis is inspired by the interpretation of the evolution with the energy of $\langle X_{\text{max}} \rangle$ and $\sigma\left[X_{\text{max}}\right]$ shown in [141].

To conclude this thesis, in the final part of this chapter we will introduce some of the future potentialities of this technique regarding the analysis of hybrid events.

7.1 Interpreting the moments of the MPD distributions

Inferring the mass composition of UHECRs from the measurements of $X^{\mu}_{\text{max,55}}$ is only possible by comparing the results to the predictions of hadronic interaction models, and is therefore subject to some level of uncertainty. For instance, in figure 5.17 there is a large discrepancy between experimental data and the predictions of EPOS-LHC which is expected to be reflected in terms of mass composition when this model is used to interpret the data.

It was shown in chapter 3 how the superposition model provides a theoretical framework to interpret the evolution of X_{max} and its fluctuations with the energy. It was also shown that the relation between $\sigma\left(X_{\text{max}}\right)$ and the primary mass is not straightforward, as different combinations of average composition can give rise to the same value of $\sigma\left(X_{\text{max}}\right)$. Only in the case of pure composition is this correspondence unique.

The linear relation between $\langle X_{\text{max}} \rangle$ and $\langle \ln A \rangle$ is expected to hold in the case of the muonic component of the shower, as its longitudinal profile is similar to the electromagnetic one, though with an earlier development [142]. We will assume, then, that the superposition model can be applied to the muonic shower, and hence:

$$\langle X^{\mu}_{\mathsf{max}} \rangle = X^{\mu}_{0} + D_{\mu} \log \left(\frac{E}{E_0 A} \right) \tag{7.1}$$

where X^{μ}_{0} is the mean depth of proton showers at energy E_0 and D_{μ} is the muonic elongation rate. For nuclei of the same mass A this can be rewritten as

$$\langle X^{\mu}_{\mathsf{max}} \rangle = \langle X^{\mu}_{\mathsf{max}} \rangle_{\mathsf{p}} + f_E \ln A \tag{7.2}$$

where $\langle X^{\mu}_{\mathsf{max}} \rangle_{\mathsf{p}}$ is the mean depth at maximum of proton showers and, in the simplistic superposition model, $f_E = -D_{\mu}/\ln 10$.

If the composition is fixed, the dispersion is only due to the shower-to-shower fluctuations

$$\sigma^2 \left(X^{\mu}_{\mathsf{max}} \right) = \sigma^2_{\mathsf{sh}} \left(\ln A \right) \tag{7.3}$$

where $\sigma^2_{\mathsf{sh}} \left(\ln A \right)$ is the X^{μ}_{max} variance for mass number A

$$\sigma^2_{\mathsf{sh}} \left(\ln A \right) = \sigma^2 \left(X^{\mu}_{\mathsf{max}} | \ln A \right) \tag{7.4}$$

If the composition is mixed, there are two contributions to the fluctuations of X^{μ}_{max}: the shower-to-shower fluctuations plus the dispersion arising from the mass distribution. The first term is simply $\left\langle \sigma^2_{\mathsf{sh}} \right\rangle$ weighted by the $\ln A$ distribution, while the second contribution can be written as

$$\sigma^2 \left(X^{\mu}_{\mathsf{max}} \right) = \left\langle \sigma^2_{\mathsf{sh}} \right\rangle + \left(\frac{d \left\langle X^{\mu}_{\mathsf{max}} \right\rangle}{d \ln A} \right)^2 \sigma^2_{\ln A} = \left\langle \sigma^2_{\mathsf{sh}} \right\rangle + f^2_E \sigma^2_{\ln A} \tag{7.5}$$

where $\sigma^2_{\ln A}$ is the variance of the $\ln A$ distribution. To obtain an explicit expression of $\left\langle \sigma^2_{\mathsf{sh}} \right\rangle$ we need a parametrisation for $\sigma^2_{\mathsf{sh}} \left(\ln A \right)$. It was shown in [141] that a quadratic relation correctly describes current models. However, in order to obtain such a parametrisation, it is necessary to arrange for simulations of intermediate mass nuclei. Fortunately, the quadratic term is of the order of 10% of the linear term, so we will approximate to a linear relation using only proton and iron simulations[1], which were the only primaries available for the moment:

$$\sigma^2_{\ln A} \simeq \sigma^2_{\mathsf{p}} + \left(\sigma^2_{\mathsf{Fe}} - \sigma^2_{\mathsf{p}} \right) \frac{\ln A}{\ln 56} = \sigma^2_{\mathsf{p}} \left[1 + a \ln A \right] \tag{7.6}$$

where a is constant. Averaging on $\ln A$

$$\left\langle \sigma^2_{\mathsf{sh}} \right\rangle = \sigma^2_{\mathsf{p}} \left[1 + a \left\langle \ln A \right\rangle \right] = \sigma^2_{\mathsf{sh}} \left(\left\langle \ln A \right\rangle \right) \tag{7.7}$$

Substituting in equation (7.5) and solving for $\sigma^2_{\ln A}$ we obtain finally

$$\sigma^2_{\ln A} = \frac{\sigma^2 \left(X^{\mu}_{\mathsf{max}} \right) - \sigma^2_{\mathsf{sh}} \left(\left\langle \ln A \right\rangle \right)}{f^2_E} \tag{7.8}$$

[1] This approximation can be removed in the future, when simulations of intermediate nuclei become available.

Equations (7.2) and (7.8) are the key tools used to interpret the measurements of X^{μ}_{max} in the Pierre Auger Observatory in terms of mass composition. They will also provide some hints on the validity of hadronic interaction models.

7.1.1 Application to experimental data

Equation (7.2) can be rewritten in a way better suited to be applied to data if we take into account the linear dependence of $\langle X^{\mu}_{\mathsf{max}} \rangle$ with $\ln A$ and we use as reference values the depth of shower maxima for proton and iron [143]. If we assume the composition to be formed by pure iron, equation (7.2) yields

$$\langle X^{\mu}_{\mathsf{max}} \rangle_{\mathsf{Fe}} = \langle X^{\mu}_{\mathsf{max}} \rangle_{\mathsf{p}} + f_E \ln 56 \tag{7.9}$$

and therefore

$$f_E = \frac{\langle X^{\mu}_{\mathsf{max}} \rangle_{\mathsf{Fe}} - \langle X^{\mu}_{\mathsf{max}} \rangle_{\mathsf{p}}}{\ln 56} \tag{7.10}$$

so in order to determine f_E it suffices to determine the average depth of shower maxima for protons and iron nuclei. Hence, at a given energy bin, the average $\ln A$ will be obtained as

$$\langle \ln A \rangle = \ln 56 \, \frac{\langle X^{\mu}_{\mathsf{max}} \rangle_{\mathsf{p}} - \langle X^{\mu}_{\mathsf{max}} \rangle}{\langle X^{\mu}_{\mathsf{max}} \rangle_{\mathsf{p}} - \langle X^{\mu}_{\mathsf{max}} \rangle_{\mathsf{Fe}}} \tag{7.11}$$

and the spread of the $\ln A$ distribution as

$$\sigma^2_{\ln A} = [\ln 56]^2 \, \frac{\sigma^2 \left(X^{\mu}_{\mathsf{max}}\right) - \sigma^2_p \left[1 - a \langle \ln A \rangle\right]}{\left[\langle X^{\mu}_{\mathsf{max}} \rangle_{\mathsf{p}} - \langle X^{\mu}_{\mathsf{max}} \rangle_{\mathsf{Fe}}\right]^2} \tag{7.12}$$

where the value of $\langle \ln A \rangle$ is obtained from equation (7.11) and a is simply

$$a = \left[\frac{\sigma^2_{Fe}}{\sigma^2_p} - 1 \right] / \ln 56 \tag{7.13}$$

so again only simulations of protons and iron nuclei are needed. As shown in chapter 3, the superposition model predicts that the spread of X_{max} (and therefore of X^{μ}_{max}) are smaller for heavier elements, and hence σ^2_{Fe} is always smaller than σ^2_p, so a is negative definite for any hadronic model.

The result of applying equations (7.11) and (7.12) to the data of $X^{\mu}_{\mathsf{max},55}$ is shown in figures 7.1 and 7.2 when QGSJETII-04 and EPOS-LHC are used to interpret the results. These results are compared to the ones obtained by the FD shown in [141]. In the case of the average $\ln A$, they are also compared to the results reported by the standard MPD analysis [89].

The first obvious conclusion that can be extracted by studying figure 7.1 is that this analysis and the standard MPD analysis are in full agreement, regardless of the hadronic model that is used to interpret the data. As a consequence, they both pose the same constraint to EPOS-LHC in terms of $\langle \ln A \rangle$, as the data are difficult to be explained under

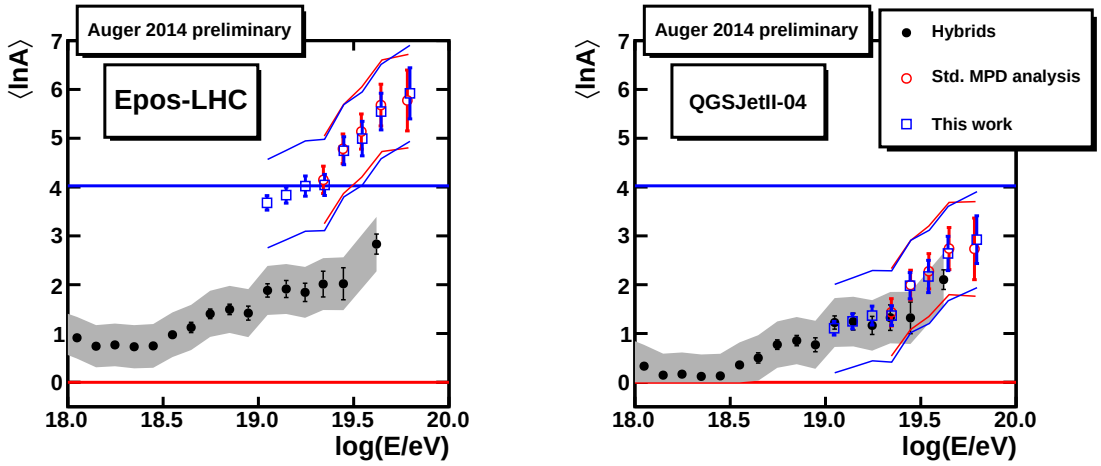


Figure 7.1: $\langle \ln A \rangle$ as a function of $\log E$ for EPOS-LHC and QGSJETII-04. Error bars show the statistical uncertainty. Shaded area is the systematic uncertainty for X_{max} measurements [141], which is shown for X^{μ}_{max} measurements as parallel lines. Horizontal lines show the values for proton ($\ln 1 = 0$) and iron ($\ln 56 \approx 4.03$) respectively.

this model and clearly disagree with the results of the FD. However, if QGSJETII-04 is used, the three datasets agree to a large extent, indicating a trend to heavier composition as the energy rises.

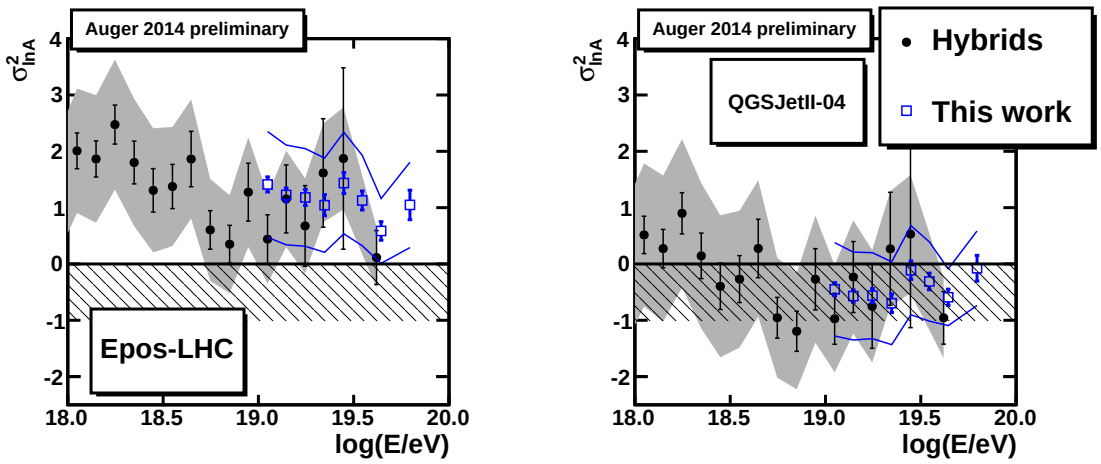


Figure 7.2: $\sigma^2_{\ln A}$ as a function of $\log E$ for EPOS-LHC and QGSJETII-04. Error bars show the statistical uncertainty. Shaded area is the systematic uncertainty for X_{max} measurements [141], which is shown for X^{μ}_{max} measurements as parallel lines. The lower limit of allowed $\sigma^2_{\ln A}$ is shown by the exclusion region, while the upper limit (4.05, corresponding to a 50% mixture of proton and iron nuclei [141]) is just above the maximum of the vertical axis.

In terms of $\sigma^2_{\ln A}$, figure 7.2 shows how this analysis and the results of the FD agree

within uncertainties. However, in the case of QGSJETII-04 they both provide *non-physical* values, as a negative value for the spread of $\ln A$ has no meaning. Even though the statistical significance of this effect is not very large given the systematic uncertainties, we can conclude that neither hadronic model satisfactorily describes simultaneously the average and spread of $\ln A$.

7.1.2 Interpretation of the results

At this stage of development, and under our experimental conditions, the systematic uncertainties are too large to make stringent claims in terms of the mass composition of UHECRs. However, the fact that some of the data lay so far from theoretical predictions can be understood as a mild constraint to hadronic interaction models.

Even under these unfavourable conditions, there is some information that can be extracted from this analysis. For example, regardless of the hadronic model, the energy evolution is common to all models, suggesting an increase of the average mass with decreasing dispersion as the energy rises. This behaviour might imply consequences to the astrophysical scenarios, as there are only a few possible source models that allow a composition with small mass dispersion.

Protons can traverse long distances from sources to the Earth, and the mass dispersion would be null. However, a pure proton composition is disfavoured by the data of the Pierre Auger Observatory, as shown in figure 7.1. On the other hand, nuclei originated nearby (~ 100 Mpc) can be detected with small mass dispersion, as the propagation does not affect much their mass nor their energy spectrum. But, if sources are distributed uniformly, distant sources will produce additional mass dispersion.

The possible end of the injection spectrum (the *disappointing model*) can reduce the light component at high energies, producing a reduction of the mass dispersion at the highest energies. Recent studies [144, 145] based on the assumption of a uniform source distribution show that the composition measured by the Pierre Auger Observatory, together with the energy spectrum measurements, requires hard injection spectra for the different primaries (spectral index below 2), a relatively small energy cut-off that grows linearly with the charge number Z and the possible presence of local sources.

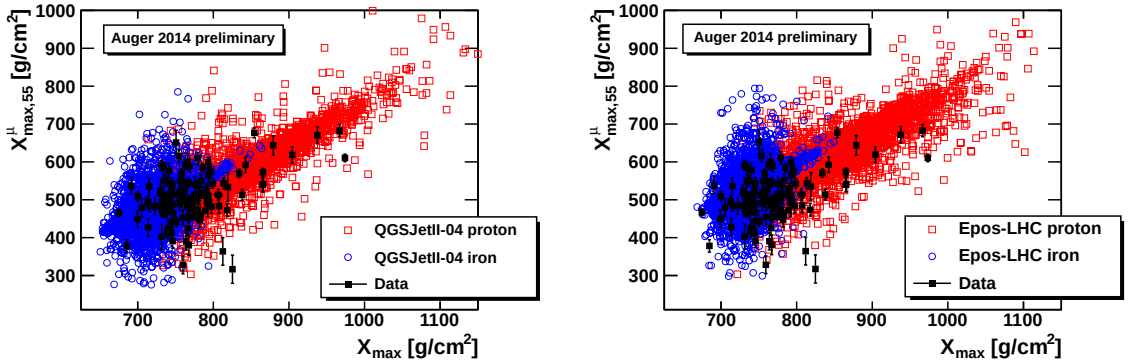
7.2 Correlation between $\mathbf{X_{max}}$ and $\mathbf{X^{\mu}_{max,55}}$

As mentioned before, the shower development is expected to be very similar for the electromagnetic and the hadronic parts of the shower. Therefore, it is expected that the correlation between X_{max} and X^{μ}_{max} is significant. In fact, Monte Carlo simulations provide a very large correlation coefficient, above 0.8, when these two observables are considered together.

However, the simultaneous reconstruction of both parameters is experimentally disfavoured. In the standard MPD analysis, only events of large zenith angle $\theta \approx 60^{\circ}$ and high energy $E > 20$ EeV have a good reconstruction of X^{μ}_{max}. On the other hand, the field of view of the FD drastically falls for large angles. Moreover, the steep energy spectrum

and the reduced duty cycle of the FD provided a set of simultaneous measurements of X_{max} and X^{μ}_{max} of the order of 10-20 events for the standard analysis.

The situation is greatly improved under the approach followed by this work. The widening of the zenith angle window, together with the reduction of the energy threshold, increases the available statistics by a factor of about 10, yielding 121 events that simultaneously survive the X_{max} analysis cuts [89] and have a good $X^{\mu}_{\mathrm{max,55}}$ reconstruction. This is shown in figure 7.3 compared to the predictions by QGSJETII-04 and EPOS-LHC. The correlation coefficient obtained for the experimental data is 0.44. We also observed that most of the outliers have a large value of the relative error $\varepsilon = \Delta X^{\mu}_{\mathrm{max,55}}/X^{\mu}_{\mathrm{max,55}}$. In fact, by removing 7 events with $\varepsilon >$6% the correlation coefficient increases up to 0.54.



(a) QGSJETII-04 (b) EPOS-LHC

Figure 7.3: Correlation between X_{max} and $X^{\mu}_{\mathrm{max,55}}$ for the 121 events passing the FD cuts and with a good reconstruction of $X^{\mu}_{\mathrm{max,55}}$ as compared to the predictions of QGSJETII-04 (left) and EPOS-LHC (right).

In order to fully exploit the potentiality of this analysis, the influence of the detector resolution, especially in the case of $X^{\mu}_{\mathrm{max,55}}$ as it turns out to be considerably large, needs to be carefully taken into account. The simultaneous analysis of these two observables, under the scope of some multi-variate technique, is expected to provide very interesting results in terms of the mass composition of UHECRs in the future. The simultaneous analysis of these two mass-sensitive observables in particular is of great importance, as they rely on very different measurements, have very different systematics and explore complementary components of the shower. In other words, they are orthogonal and independent measurements of the mass composition of the primary and hence are expected to maximise the discrimination power.

Abridgement of results

The main topic addressed by this thesis is the measurement of the mass composition of ultra-high-energy cosmic rays (UHECRs) using data from the surface detector (SD) of the Pierre Auger Observatory. The determination of the mass composition of UHECRs is a long-standing problem of great interest to the scientific community, for it can help understanding the origin and mechanisms of acceleration of these highly energetic particles.

The starting point of this thesis is the reconstruction of the muon production depth profile (MPD), which was introduced some years ago as a proof-of-concept by the same research group. The main goal of this work is to widen the zenith and energy windows that are suitable of being analysed. This poses different problems that were addressed sequentially.

■ Correction of the radial dependence of MPDs

Using Monte Carlo simulations, we introduced an empirical factor that corrects for the dependence of the MPD with the observation distance r. This factor acts as a transformation on the production distance of muons measured vertically, z_v, so that their distribution after the transformation does not depend on r. The analytical expression of this factor is (equation (4.5))

$$z'_v = z_v \times \frac{z_v \tan^2\theta - r}{z_v \tan^2\theta - r_0}$$

where θ is the zenith angle and r_0 a free parameter which represent the optimum observation distance at which the observed MPD is identical to the full MPD apart from a scaling factor.

It was shown that this correction fulfils all the pre-requisites. In particular, it makes for an unbiased reconstruction of the maximum of the MPD X^{μ}_{max} regardless of the observation point.

■ Extension of the MPD analysis to a wider zenith and energy range

In order to apply equation (4.5) to shower reconstructions, parameter r_0 needs to be parametrised in terms of the primary energy and zenith angle. It was shown that this parametrisation is robust under different primaries and hadronic models.

Another caveat of reconstructing MPDs at smaller zenith angle is the increase of the electromagnetic contamination. This problem was solved using a two-fold strategy, that lays on both rejecting very late signals and choosing an appropriate signal threshold,

in such a way that muonic signal is enhanced over electromagnetic signal. Putting everything together, it was shown that it is possible to reconstruct the MPD using stations further than $r > 1000$ m and closer than $r < 4000$ m with a total electromagnetic contamination that is always below 26% and a total efficiency selecting muons above 92%.

Finally, in order to simultaneously analyse a zenith window of 20° width, it becomes necessary to refer all the values of X^{μ}_{max} to a common zenith angle. For that purpose, the slope of the dependence of X^{μ}_{max} versus the cosine of θ was parametrised and used to build $X^{\mu}_{\text{max,55}}$, which represents the equivalent X^{μ}_{max} referred to a 55°.

The full set of improvements with respect to the previous MPD reconstruction technique are summarised in the following table:

	Std. Analysis	**This work**
r_{min}	1700 m	1000 m
Zenith	[55°,65°]	[45°,65°]
$\log(E/\text{eV})$	> 19.3	> 19.0
Statistics	< 500 events	> 4000 events
Resolution	45 - 80 g/cm^2	35 - 75 g/cm^2

The most relevant improvements are a reduction of the resolution of about 20 g/cm^2 at low energy[2] and about 10 g/cm^2 at large energy, and an increase in the available statistics by about a factor eight.

■ Measuring shower-to-shower fluctuations with the SD

Being all of the stated above of much interest, the most relevant result presented in this work has to do with the measurement of the shower-to-shower fluctuations. This thesis demonstrates that the resolution of the detector in the determination of $X^{\mu}_{\text{max,55}}$ can be obtained using a data-driven method. This method is based on the splitting of SD station into two equivalent sets in terms of their azimuth angle, taking advantage of the azimuthal symmetry. It was shown that, apart from a statistical factor, the comparison of the two reconstructed values of $X^{\mu}_{\text{max,55}}$ yields a good estimate of the detector resolution. Furthermore, it was shown that this resolution has negligible dependence with the hadronic model, and that data are well bracketed by Monte Carlo simulations.

The correct determination of the detector resolution allows for a subtraction of the resolution to the observed fluctuations in order to correctly obtain the physical fluctuations. This is an innovative and independent measurement of the mass composition of UHECRs in the context of MPDs. Furthermore, the method developed to determine the resolution of the detector for this particular observable can be used for other observables straightforwardly, and hence constitutes a valuable tool for the future.

[2]The resolution at the same energy $10^{19.3}$ eV for this work is about 60 g/cm^2.

■ Interpretation of the first two moments of the MPD distribution

The measured values of $\left\langle X^{\mu}_{\mathsf{max,55}} \right\rangle$ and $\sigma\left[X^{\mu}_{\mathsf{max,55}}\right]$ can be interpreted within the framework of the superposition model. In particular, the values of $\langle \ln A \rangle$ and $\sigma^2_{\ln A}$, i. e., the first two moments of the $\ln A$ distribution, where A represents the mass number of the primary, can be inferred using (equations (7.11) and (7.12))

$$\langle \ln A \rangle = \ln 56 \, \frac{\langle X^{\mu}_{\mathsf{max}} \rangle_{\mathsf{p}} - \langle X^{\mu}_{\mathsf{max}} \rangle}{\langle X^{\mu}_{\mathsf{max}} \rangle_{\mathsf{p}} - \langle X^{\mu}_{\mathsf{max}} \rangle_{\mathsf{Fe}}}$$

$$\sigma^2_{\ln A} = [\ln 56]^2 \, \frac{\sigma^2 \left(X^{\mu}_{\mathsf{max}}\right) - \sigma^2_p \left[1 - a \langle \ln A \rangle\right]}{\left[\langle X^{\mu}_{\mathsf{max}} \rangle_{\mathsf{p}} - \langle X^{\mu}_{\mathsf{max}} \rangle_{\mathsf{Fe}}\right]^2}$$

The obtained values of $\langle \ln A \rangle$ agree remarkably well with the previous measurements of the standard MPD analysis at 60°. However, the interpretation of $\langle \ln A \rangle$ is difficult when the hadronic model EPOS-LHC is used to interpret data, as they yield a much larger average logarithm mass than the results of the fluorescence detector (FD).

In terms of $\sigma^2_{\ln A}$, the results agree with the FD in the whole energy range. Nevertheless, for the hadronic model QGSJETII-04 some negative values of $\sigma^2_{\ln A}$ are found, which makes no physical sense.

By analysing the trend of these two observables, we can conclude that no hadronic model correctly describes simultaneously the mean and spread of $\ln A$ predicted by the MPD measurement. This technique, hence, can be used to constrain the predictions of hadronic models in the muonic sector.

Concerning mass composition, we conclude that the measurements obtained by the Pierre Auger Observatory both by the SD and the FD clearly disfavour a pure light composition at the highest energies.

■ Correlation between $\mathbf{X}_{\mathbf{max}}$ and $\mathbf{X}^{\mu}_{\mathbf{max,55}}$

The combination of $X^{\mu}_{\mathsf{max,55}}$ with measurements of the FD was also explored. This combination was impossible under the standard MPD analysis, as the statistics was too scarce. It was shown that under the new approach to the MPD reconstruction it is possible to establish the correlation between $X^{\mu}_{\mathsf{max,55}}$ and X_{max}. This analysis has a large potentiality for the future, as the sensitivity of the mass inference significantly improves when adding multiple mass-sensitive variables to the analysis. However, before being able to use these two variables in a multi-dimensional analysis, a full understanding of the influence of the detector resolution, especially in the case of $X^{\mu}_{\mathsf{max,55}}$ given its relatively large value, is required.

Compendio de resultados

El tema de estudio abordado por esta tesis doctoral es la medida de la composición de los rayos cósmicos de ultra alta energía (UHECRs) utilizando para ello datos del detector de superficie (SD) del Observatorio Pierre Auger. La determinación de la composición de los UHECRs es un problema irresoluto desde hace décadas, de gran interés para la comunidad científica dadas las implicaciones que impone acerca del origen y los mecanismos de aceleración de estas partículas tan energéticas.

El punto de partida de esta tesis es la reconstrucción del perfil de densidades de producción de muones (MPD), técnica que fue introducida hace algunos años como una prueba de concepto por el mismo grupo de investigación. El objetivo principal de este trabajo es ampliar la ventana de aplicabilidad de este método a ángulos y energías más bajos. Para ello, se han ido resolviendo secuencialmente las dificultades que se plantean en el proceso.

■ Corrección de la dependencia radial de las MPDs

Mediante el estudio de simulaciones Montecarlo, se introduce en este trabajo un factor de corrección empírico que corrige la dependencia de la MPD con la distancia de observación r. Dicho factor actúa como una transformación sobre la distancia de producción de muones medida verticalmente, z_v, de modo que su distribución tras la transformación es independiente de r. La expresión analítica de este factor de corrección (ecuación (4.5)) es:

$$z'_v = z_v \times \frac{z_v \tan^2\theta - r}{z_v \tan^2\theta - r_0}$$

donde θ es el ángulo cenital y r_0 un parámetro libre que representa la distancia de observación óptima, a la cual la MPD observada es idéntica a la total excepto por un factor de escala.

Se demuestra a lo largo de la tesis que este factor de corrección reúne todos los prerrequisitos, incluyendo una reconstrucción no sesgada del máximo de la MPD X^{μ}_{max}, independientemente del punto de observación.

■ Extensión del análisis de la MPD a energías y ángulos más bajos

Antes de poder aplicar la ecuación (4.5) en la reconstrucción de cascadas, es necesario parametrizar la dependencia de r_0 en función de la energía y ángulo cenital del primario. Se comprobó que dicha dependencia es robusta para distintos primarios y modelos hadrónicos.

Otra dificultad que se presenta al reconstruir MPDs a ánulos más bajos es el aumento de la contaminación electromagnética. Este problema se resolvió utilizando una estrategia doble: en primer lugar se eliminan señales muy tardías, que están dominadas por la componente electromagnética, y en segundo lugar se elige un umbral mínimo para las señales, de tal modo que la señal muónica se realza sobre la señal electromagnética. Al tener todo en cuenta, se demuestra que es posible reconstruir la MPD empleando para ello estaciones a distancias del eje de la cascada mayores de 1000 m y menores de 4000 m con una contaminación electromagnética que es en todo caso menor del 26 % y una eficiencia de selección de señal muónica por encima del 92 %.

Finalmente, para poder analizar una ventana angular de 20°de anchura como un todo, es necesario referir todos los valores de X^{μ}_{max} a un mismo ángulo. Para ello, se obtuvo la pendiente de la dependencia de X^{μ}_{max} con respecto al coseno de θ, y se empleó para construir $X^{\mu}_{\text{max,55}}$, que representa el valor equivalente de X^{μ}_{max} referido a 55°.

El resumen de todas las de mejoras con respecto a la técnica previa de reconstrucción de la MPD se resume en la siguiente tabla:

	Análisis estándar	**Nuevo análisis**
r_{min}	1700 m	1000 m
Ángulo cenital	[55°,65°]	[45°,65°]
$\log(E/\text{eV})$	>19.3	>19.0
Estadística	<500 sucesos	>4000 sucesos
Resolución	45 - 80 g/cm^2	35 - 75 g/cm^2

Las mejoras más significativas son una reducción de la resolución en alrededor de 20 g/cm^2 a baja energía[3] y de 10 g/cm^2 a alta energía, y un aumento de la estadística de sucesos por un factor ocho.

■ Medida de las fluctuaciones cascada a cascada con el SD

Siendo todo lo anterior de gran interés, el resultado más relevante presentado en este trabajo concierne a las fluctuaciones cascada a cascada. En esta tesis se demuestra que la resolución del detector en la medida de $X^{\mu}_{\text{max,55}}$ puede obtenerse empleando para ello únicamente datos reales. Este método se basa en dividir las estaciones del SD en dos conjuntos equivalentes en función de su ángulo acimutal, aprovechando la simetría acimutal del problema. Se demuestra en este trabajo que, excepto por una constante de origen estadístico, la comparación del valor reconstruido de $X^{\mu}_{\text{max,55}}$ de cada una de ambas mitades proporciona una buena estimación de la resolución del detector. Además, se comprueba también que dicha resolución es prácticamente independiente de modelo hadrónico y que el comportamiento de los datos queda dentro de las predicciones de las simulaciones Montecarlo.

Una correcta determinación de la resolución del detector permite sustraerla en cuadraturas a las fluctuaciones observadas para obtener las fluctuaciones de origen físico. Esto representa una medida independiente y novedosa de la composición de los

[3] La resolución para la misma energía $10^{19.3}$ eV con este método es de 60 g/cm^2.

UHECRs haciendo uso de la MPD. Asimismo, el método desarrollado para determinar la resolución para este observable en particular puede ser utilizado para otros observables de forma trivial, y por lo tanto constituye una herramienta valiosa para el futuro.

■ Interpretación de los primeros dos momentos de la MPD

Las mediciones de $\left\langle X^{\mu}_{\mathsf{max},55} \right\rangle$ y $\sigma\left[X^{\mu}_{\mathsf{max},55}\right]$ pueden ser interpretadas en términos de composición empleando el modelo de superposición. En particular, los valores de $\langle \ln A \rangle$ y $\sigma^2_{\ln A}$, es decir, de los dos primeros momentos de la distribución de $\ln A$, donde A es el número másico, pueden deducirse de (7.11) y (7.12):

$$\langle \ln A \rangle = \ln 56 \, \frac{\langle X^{\mu}_{\mathsf{max}} \rangle_{\mathsf{p}} - \langle X^{\mu}_{\mathsf{max}} \rangle}{\langle X^{\mu}_{\mathsf{max}} \rangle_{\mathsf{p}} - \langle X^{\mu}_{\mathsf{max}} \rangle_{\mathsf{Fe}}}$$

$$\sigma^2_{\ln A} = [\ln 56]^2 \, \frac{\sigma^2 \left(X^{\mu}_{\mathsf{max}}\right) - \sigma^2_p \left[1 - a \, \langle \ln A \rangle\right]}{\left[\langle X^{\mu}_{\mathsf{max}} \rangle_{\mathsf{p}} - \langle X^{\mu}_{\mathsf{max}} \rangle_{\mathsf{Fe}}\right]^2}$$

Los valores obtenidos para el promedio $\langle \ln A \rangle$ están en muy buen acuerdo con las medidas previas de la MPD a 60°. Sin embargo, la interpretación de las medidas de $\langle \ln A \rangle$ utilizando el modelo hadrónico EPOS-LHC es difícil, ya que proporciona valores mucho mayores que los medidos por el detector de fluorescencia (FD).

En términos de la varianza $\sigma^2_{\ln A}$, los resultados concuerdan con el FD en todo el rango de energías. No obstante, para el modelo hadrónico QGSJETII-04 se encuentran algunos valores negativos de la varianza, que no tienen sentido físico.

Analizando ambos resultados, podemos concluir que ningún modelo hadrónico reproduce satisfactoriamente las medidas predichas por el análisis de la MPD para $\ln A$ y $\sigma^2_{\ln A}$ simultáneamente. Esta técnica puede, por tanto, ser empleada para constreñir las predicciones de los modelos hadrónicos en el sector muónico.

■ Correlación entre $\mathbf{X}_{\mathsf{max}}$ y $\mathbf{X}^{\mu}_{\mathsf{max,55}}$

La combinación de la medidas de $X^{\mu}_{\mathsf{max},55}$ con resultados del FD también se explora en este trabajo. Dicha combinación era imposible en el análisis previo de la MPD debido a la insuficiente estadística de sucesos. Con esta nuevo método para la reconstrucción de la MPD, se demostró que puede obtenerse la correlación entre X^{μ}_{max} y X_{max}. Esta combinación posee gran potencialidad para el futuro, ya que la sensibilidad en la determinación de la composición mejora significativamente cuando se añaden diversas variables sensibles a composición al análisis. Sin embargo, antes de poder emplear estas dos variables en un análisis bidimensional, es necesario entender en profundidad la influencia que sobre el mismo tiene la resolución del detector, especialmente en el caso de $X^{\mu}_{\mathsf{max},55}$, dado su alto valor.

A

Determination of statistical parameters

The average value of $X^{\mu}_{\text{max,55}}$ for a given energy bin with N_i events is obtained from the arithmetic mean.

$$\left\langle X^{\mu}_{\text{max,55}} \right\rangle_i = \frac{1}{N_i} \sum_j X^{ij} \tag{A.1}$$

where X^{ij} represents the value of $X^{\mu}_{\text{max,55}}$ of entry j, with $j = 1, 2 \ldots N_i$ in the energy bin i. The fluctuations [1], $\text{RMS}'_i = \sqrt{V_i}$ are measured using the unbiased estimator of the sample variance

$$V_i = \frac{1}{N_i - 1} \sum_j \left(X^{ij} - \left\langle X^{\mu}_{\text{max,55}} \right\rangle_i \right)^2 \tag{A.2}$$

The statistical uncertainty of $\left\langle X^{\mu}_{\text{max,55}} \right\rangle_i$ is given by

$$\sigma \left(\left\langle X^{\mu}_{\text{max,55}} \right\rangle_i \right) = \sqrt{\frac{V_i}{N_i}} \tag{A.3}$$

For the uncertainty of the observed fluctuations we use the formula from [107]

$$\sigma \left(V_i \right) = \frac{1}{N_i} \left(m_4 - \frac{N-3}{N-1} V_i^2 \right) \tag{A.4}$$

where m_4 denotes the fourth central moment of the $X^{\mu}_{\text{max,55}}$ distribution. Using this formula, the uncertainty of the observed spread is

$$\sigma \left(\text{RMS}'_i \right) = \frac{\sigma \left(V_i \right)}{\sqrt{4 V_i}} \tag{A.5}$$

[1] We will refer to the fluctuations as *RMS* in this appendix to simplify the notation.

The spread after correcting for the resolution is obtained following equation (6.1)

$$\mathrm{RMS}_i = \sqrt{V_i - \sigma\left(X^{\mu}_{\mathrm{max},55}\right)^2_i} \tag{A.6}$$

and its uncertainty is obtained from error propagation as

$$\sigma\left(\mathrm{RMS}_i\right) = \frac{\sqrt{V_i}}{\mathrm{RMS}_i}\sigma\left(\mathrm{RMS}'_i\right) \tag{A.7}$$

B

Correction for detector resolution

In the following we will demonstrate equation (6.1) following the same approach as the one described in [140].

To understand how to obtain the real shower-to-shower fluctuations starting from the observed distributions of $X^{\mu}_{\text{max,55}}$ it is useful to imagine that there were only two kind of events: a fraction f measured with a good precision d_1 and the rest with a worse precision d_2. If the spread of the physical shower-to-shower fluctuations is s, then the combined measurement of both populations has a variance of

$$\begin{aligned} V\left[X^{\mu}_{\text{max,55}}\right] &= f\left(s^2+d_1^2\right)+(1-f)\left(s^2+d_2^2\right) \\ &= s^2+fd_1^2+(1-d)\,d_2^2 \\ &= s^2+E\left[d^2\right] \end{aligned} \tag{B.1}$$

where $E\left[d^2\right]$ represents the expectation value of the resolution. This means that the effective detector resolution can be used to estimate s^2. Obviously, this calculation can be extended to as many categories as wanted and even to a continuous distribution of the detector resolution.

List of Figures

List of Tables

Bibliography

[1] V. F. Hess, *Über Beobachtungen der durchdringenden Strahlung bei sieben Freiballonfahrten*, Phys. Z. , 1084.

[2] Pierre Auger Collaboration, J. Abraham *et al.*, *Properties and performance of the prototype instrument for the Pierre Auger Observatory*, Nucl.Instrum.Meth. **A523**, 50 (2004).

[3] Pierre Auger Collaboration, J. Abraham *et al.*, *Measurement of the Depth of Maximum of Extensive Air Showers above 10^{18} eV*, Phys. Rev. Lett. **104**, 091101 (2010).

[4] T. Stanev, *High Energy Cosmic Rays* (Springer Berlin Heidelberg, 2010).

[5] C. Coulomb, *Mdm. de l'Acad. des Sciences*, Académie royale des sciences Paris , 612 (1875).

[6] H. Becquerel, *Sur les radiations émises par phosphorescence*, Comptes rendus de l'Acad. des Sciences **420** (1896).

[7] D. Skobelzyn, *Die Intensitätsverteilung in dem Spektrum der γ-Strahlen von Ra C*, Zeitschrift für Physik **43**, 354 (1927).

[8] C. D. Anderson, *The Positive Electron*, Phys. Rev. **43**, 491 (1933).

[9] C. D. Anderson and S. H. Neddermeyer, *Cloud Chamber Observations of Cosmic Rays at 4300 Meters Elevation and Near Sea-Level*, Phys. Rev. **50**, 263 (1936).

[10] C. Lattes, H. Muirhead, G. Occhialini, and C. Powell, *Processes involving charged mesons*, Nature **159**, 694 (1947).

[11] H. Yukawa, *On the Interaction of Elementary Particles. I*, Progress of Theoretical Physics Supplement **1**, 1 (1955).

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

[13] M. Gell-Mann, *A Schematic Model of Baryons and Mesons*, Phys.Lett. **8**, 214 (1964).

[14] G. Zweig, *An SU_3 model for strong interaction symmetry and its breaking; Part II*, p. 80 p (1964).

[15] Super-Kamiokande Collaboration, Y. Fukuda *et al.*, *Evidence for Oscillation of Atmospheric Neutrinos*, Phys. Rev. Lett. **81**, 1562 (1998).

[16] W. Kraushaar *et al.*, *Explorer XI Experiment on Cosmic Gamma Rays.*, Astrophysical Journal **141**, 845 (1965).

[17] A. K. Drukier, K. Freese, and D. N. Spergel, *Detecting cold dark-matter candidates*, Phys. Rev. D **33**, 3495 (1986).

[18] G. Angloher *et al.*, *Results from 730 kg days of the CRESST-II Dark Matter Search*, Eur.Phys.J. **C72**, 1971 (2012).

[19] XENON100 Collaboration, E. Aprile *et al.*, *Dark Matter Results from 225 Live Days of XENON100 Data*, Phys.Rev.Lett. **109**, 181301 (2012).

[20] J. R. Hörandel, *Models of the knee in the energy spectrum of cosmic rays*, Astroparticle Physics **21**, 241 (2004).

[21] A. A. Ivanov, S. P. Knurenko, and I. Y. Sleptsov, *Measuring extensive air showers with Cherenkov light detectors of the Yakutsk array: the energy spectrum of cosmic rays*, New Journal of Physics **11**, 065008 (2009).

[22] M. Nagano *et al.*, *Energy spectrum of primary cosmic rays above 10 17.0 eV determined from extensive air shower experiments at Akeno*, Journal of Physics G: Nuclear and Particle Physics **18**, 423 (1992).

[23] T. Antoni *et al.*, *{KASCADE} measurements of energy spectra for elemental groups of cosmic rays: Results and open problems*, Astroparticle Physics **24**, 1 (2005).

[24] The KASCADE-Grande Collaboration, W. Apel *et al.*, *KASCADE-Grande measurements of energy spectra for elemental groups of cosmic rays*, Astroparticle Physics (2013), arXiv:1306.6283.

[25] D. J. Bird *et al.*, *Evidence for correlated changes in the spectrum and composition of cosmic rays at extremely high energies*, Phys. Rev. Lett. **71**, 3401 (1993).

[26] A. M. Hillas, *Cosmic Rays: Recent Progress and some Current Questions*, (2006), arXiv:astro-ph/0607109.

[27] D. De Marco and T. Stanev, *On the shape of the ultrahigh energy cosmic ray spectrum*, Phys. Rev. D **72**, 081301 (2005).

[28] D. Allard, E. Parizot, and A. Olinto, *On the transition from galactic to extragalactic cosmic-rays: Spectral and composition features from two opposite scenarios*, Astroparticle Physics **27**, 61 (2007).

[29] V. Berezinsky, *Transition from galactic to extragalactic cosmic rays*, (2007), arXiv:0710.2750.

[30] V. Berezinsky, S. Grigorieva, and B. Hnatyk, *Extragalactic {UHE} proton spectrum and prediction for iron-nuclei flux at 108–109 GeV*, Astroparticle Physics **21**, 617 (2004).

[31] V. Berezinsky, *Extragalactic {UHE} proton spectrum and prediction of flux of iron-nuclei at 108–109 GeV*, Nuclear Physics B - Proceedings Supplements **151**, 497 (2006).

[32] A. A. Penzias and R. W. Wilson, *A Measurement of excess antenna temperature at 4080-Mc/s*, Astrophys.J. **142**, 419 (1965).

[33] K. Greisen, *End to the cosmic ray spectrum?*, Phys.Rev.Lett. **16**, 748 (1966).

[34] G. Zatsepin and V. Kuzmin, *Upper limit of the spectrum of cosmic rays*, JETP Lett. **4**, 78 (1966).

[35] IceCube Collaboration, M. Aartsen *et al.*, *First observation of PeV-energy neutrinos with IceCube*, (2013), arXiv:1304.5356.

[36] R. Aloisio, V. Berezinsky, and A. Gazizov, *Ultra high energy cosmic rays: The disappointing model*, Astroparticle Physics **34**, 620 (2011).

[37] J. R. Hoerandel, *Highlights in astroparticle physics: muons, neutrinos, hadronic interactions, exotic particles, and dark matter - Rapporteur Talk HE2 and HE3*, (2012), arXiv:1212.1013.

[38] R. Abbasi *et al.*, *Observation of Anisotropy in the Arrival Directions of Galactic Cosmic Rays at Multiple Angular Scales with IceCube*, The Astrophysical Journal **740**, 16 (2011).

[39] A. A. Abdo *et al.*, *The Large-Scale Cosmic-Ray Anisotropy as Observed with Milagro*, The Astrophysical Journal **698**, 2121 (2009).

[40] Pierre Auger Collaboration, J. Abraham *et al.*, *Correlation of the highest energy cosmic rays with nearby extragalactic objects*, Science **318**, 938 (2007), arXiv:0711.2256.

[41] Pierre Auger Collaboration, J. Abraham *et al.*, *Correlation of the highest-energy cosmic rays with the positions of nearby active galactic nuclei*, Astropart.Phys. **29**, 188 (2008), arXiv:0712.2843.

[42] Pierre Auger Observatory Collaboration, P. Abreu *et al.*, *Update on the correlation of the highest energy cosmic rays with nearby extragalactic matter*, Astropart.Phys. **34**, 314 (2010), arXiv:1009.1855.

[43] M. Veron-Cetty and P. Veron, *A catalogue of quasars and active nuclei: 12th edition*, Astron. Astrophys. **455**, 773 (2006).

[44] T. Abu-Zayyad *et al.*, *Search for Anisotropy of Ultra-High Energy Cosmic Rays with the Telescope Array Experiment*, (2012), arXiv:1205.5984.

[45] T. K. Gaisser *et al.*, *Cosmic-ray composition around 10^{18} eV*, Phys. Rev. D **47**, 1919 (1993).

[46] M. Nagano and A. A. Watson, *Observations and implications of the ultrahigh-energy cosmic rays*, Rev. Mod. Phys. **72**, 689 (2000).

[47] T. Pierog, *LHC results and High Energy Cosmic Ray Interaction Models*, Journal of Physics: Conference Series **409**, 012008 (2013).

[48] S. Fukui *et al.*, *A Study on the Structure of the Extensive Air Shower*, Progress of Theoretical Physics Supplement **16**, 1 (1960), http://ptps.oxfordjournals.org/content/16/1.full.pdf+html.

[49] KASCADE-Grande Collaboration, J. Arteaga-Velazquez *et al.*, *Study of the muon content of very high-energy EAS measured with the KASCADE-Grande observatory*, (2013), arXiv:1308.3202.

[50] Kamiokande-II, K. Hirata *et al.*, *Real time, directional measurement of B-8 solar neutrinos in the Kamiokande-II detector*, Phys. Rev. **D44**, 2241 (1991).

[51] KAMIOKANDE-II, K. Hirata *et al.*, *Observation of a Neutrino Burst from the Supernova SN 1987a*, Phys. Rev. Lett. **58**, 1490 (1987).

[52] Appec/aspera astroparticle physics roadmap phase 1.

[53] E. Waxman and J. Bahcall, *High energy neutrinos from astrophysical sources: An upper bound*, Phys. Rev. D **59**, 023002 (1998).

[54] K. Mannheim, R. J. Protheroe, and J. P. Rachen, *Cosmic ray bound for models of extragalactic neutrino production*, Phys. Rev. D **63**, 023003 (2000).

[55] J. Bahcall and E. Waxman, *High energy astrophysical neutrinos: The upper bound is robust*, Phys. Rev. D **64**, 023002 (2001).

[56] Pierre Auger Observatory Collaboration, P. Abreu *et al.*, *A Search for Ultra-High Energy Neutrinos in Highly Inclined Events at the Pierre Auger Observatory*, Phys.Rev. **D84**, 122005 (2011), arXiv:1202.1493.

[57] ANITA, P. W. Gorham *et al.*, *New Limits on the Ultra-high Energy Cosmic Neutrino Flux from the ANITA Experiment*, Phys. Rev. Lett. **103**, 051103 (2009), arXiv:0812.2715.

[58] IceCube, A. Achterberg *et al.*, *First year performance of the IceCube neutrino telescope*, Astropart. Phys. **26**, 155 (2006), arXiv:astro-ph/0604450.

[59] M. Kachelriess, *The rise and fall of top-down models as main UHECR sources*, (2008), arXiv:0810.3017.

[60] D. Kuempel, GAP Report No. 053, 2011 (unpublished).

[61] J. Abraham *et al.*, *Upper limit on the cosmic-ray photon flux above 10^{19} eV using the surface detector of the Pierre Auger Observatory*, Astroparticle Physics **29**, 243 (2008).

[62] J. Abraham *et al.*, *Upper limit on the cosmic-ray photon fraction at EeV energies from the Pierre Auger Observatory*, Astroparticle Physics **31**, 399 (2009).

[63] K. Shinozaki *et al.*, *Upper Limit on Gamma-Ray Flux above 10^{19} eV Estimated by the Akeno Giant Air Shower Array Experiment*, The Astrophysical Journal Letters **571**, L117 (2002).

[64] M. Risse *et al.*, *Upper Limit on the Photon Fraction in Highest-Energy Cosmic Rays from AGASA Data*, Phys. Rev. Lett. **95**, 171102 (2005).

[65] A. Glushkov *et al.*, *Constraints on the fraction of primary gamma rays at ultra-high energies from the muon data of the Yakutsk EAS array*, JETP Letters **85**, 131 (2007).

[66] G. I. Rubtsov *et al.*, *Upper limit on the ultrahigh-energy photon flux from AGASA and Yakutsk data*, Phys. Rev. D **73**, 063009 (2006).

[67] M. Ave, J. A. Hinton, R. A. Vázquez, A. A. Watson, and E. Zas, *New Constraints from Haverah Park Data on the Photon and Iron Fluxes of Ultrahigh-Energy Cosmic Rays*, Phys. Rev. Lett. **85**, 2244 (2000).

[68] M. Ave, J. A. Hinton, R. A. Vázquez, A. A. Watson, and E. Zas, *Constraints on the ultrahigh-energy photon flux using inclined showers from the Haverah Park array*, Phys. Rev. D **65**, 063007 (2002).

[69] J. Posselt, Search strategies for relativistic magnetic monopoles with the IceCube neutrino telescope, in *International Cosmic Ray Conference*, , International Cosmic Ray Conference Vol. 5, p. 97, 2011.

[70] A. Yamamoto *et al.*, *Search for cosmic-ray antiproton origins and for cosmological antimatter with {BESS}*, Advances in Space Research **51**, 227 (2013), The Origins of Cosmic Rays: Resolving Hess's Century-Old Puzzle.

[71] Planck Collaboration, P. Ade *et al.*, *Planck 2013 results. I. Overview of products and scientific results*, (2013), arXiv:1303.5062.

[72] G. Bertone, D. Hooper, and J. Silk, *Particle dark matter: evidence, candidates and constraints*, Physics Reports **405**, 279 (2005).

[73] J. Aleksić *et al.*, *Searches for dark matter annihilation signatures in the Segue 1 satellite galaxy with the MAGIC-I telescope*, Journal of Cosmology and Astroparticle Physics **2011**, 035 (2011).

[74] A. A. Abdo *et al.*, *Observations of Milky Way Dwarf Spheroidal Galaxies with the Fermi-Large Area Telescope Detector and Constraints on Dark Matter Models*, The Astrophysical Journal **712**, 147 (2010).

[75] M. Wood *et al.*, *Prospects for Indirect Detection of Dark Matter with CTA*, (2013), arXiv:1305.0302.

[76] D. Atri and A. L. Melott, *Cosmic Rays and Terrestrial Life: a Brief Review*, Astropart.Phys. **53**, 186 (2014), arXiv:1211.3962.

[77] D. Siingh and R. P. Singh, *The role of cosmic rays in the Earth's atmospheric processes*, Pramana **74**, 153 (2010), arXiv:0908.4156.

[78] J. Duplissy *et al.*, *Results from the CERN pilot CLOUD experiment*, Atmospheric Chemistry and Physics **10**, 1635 (2010).

[79] I. Allekotte *et al.*, *The surface detector system of the Pierre Auger Observatory*, Nucl. Instrum. Methods A **586**, 409 (2008).

[80] J. Abraham *et al.*, *The fluorescence detector of the Pierre Auger Observatory*, Nucl. Instrum. Methods A **620**, 227 (2010).

[81] F. S. for the Pierre Auger Collaboration, *The AMIGA Detector of the Pierre Auger Observatory: An Overview*, International Cosmic Ray Conference Proceedings, Beijing (2011).

[82] T. M. for the Pierre Auger Collaboration, *The HEAT Telescopes of the Pierre Auger Observatory. Status and First Data.*, International Cosmic Ray Conference Proceedings, Beijing (2011).

[83] J. K. for the Pierre Auger Collaboration, *AERA: The Auger Engineering Radio Array.*, International Cosmic Ray Conference Proceedings, Beijing (2011).

[84] T. Waldenmaier, J. Blümer, and H. Klages, *Spectral resolved measurement of the nitrogen fluorescence emissions in air induced by electrons*, Astroparticle Physics **29**, 205 (2008).

[85] C. Bonifazi, *The angular resolution of the Pierre Auger Observatory*, Nuclear Physics B - Proceedings Supplements **190**, 20 (2009), Proceedings of the Cosmic Ray International Seminars.

[86] M. Unger, GAP Report No. 010, 2006 (unpublished).

[87] T. K. Gaisser and A. M. Hillas, *Reliability of the method of constant intensity cuts for reconstructing the average development of vertical showers*, International Cosmic Ray Conference **8**, 353 (1977).

[88] A. G. Mariazzi and M. Tueros, GAP Report No. 026, 2013 (unpublished).

[89] Pierre Auger Collaboration, A. Aab *et al.*, *The Pierre Auger Observatory: Contributions to the 33rd International Cosmic Ray Conference (ICRC 2013)*, (2013), arXiv:1307.5059.

[90] The Pierre Auger collaboration, P. Abreu *et al.*, *The rapid atmospheric monitoring system of the Pierre Auger Observatory*, Journal of Instrumentation **7**, P09001 (2012).

[91] L. Wiencke and the Pierre Auger Collaboration, *Atmospheric calorimetry above 10 19 ev: Shooting lasers at the Pierre Auger Cosmic-Ray Observatory*, Journal of Physics: Conference Series **160**, 012037 (2009).

[92] X. Bertou *et al.*, *Calibration of the surface array of the Pierre Auger Observatory*, Nucl. Instrum. Methods A **568**, 839 (2006).

[93] P. Billoir, GAP Report No. 179, 2009 (unpublished).

[94] P. Billoir, P. Ghia, D. Nitz, and R. Sato, GAP Report No. 032, 2011 (unpublished).

[95] S. Argirò *et al.*, *The offline software framework of the Pierre Auger Observatory*, Nucl. Instrum. Methods A **580**, 1485 (2007).

[96] D. Veberic and M. Roth, GAP Report No. 035, 2005 (unpublished).

[97] Main page of SD Auger Trigger and Acceptance Group, `http://ipnweb.in2p3.fr/~auger/AugerProtected/AcceptMain.html`.

[98] M. Ave *et al.*, *The accuracy of signal measurement with the water Cherenkov detectors of the Pierre Auger Observatory*, Nucl. Instrum. Methods A **578**, 180 (2007).

[99] Pierre Auger Collaboration, D. Barnhill *et al.*, *Measurement of the lateral distribution function of UHECR air showers with the Pierre Auger Observatory*, (2005), arXiv:astro-ph/0507590.

[100] Pierre Auger Collaboration, M. Roth, *The Lateral distribution function of shower signals in the surface detector of the Pierre Auger Observatory*, p. 333 (2003), arXiv:astro-ph/0308392.

[101] D. Newton, J. Knapp, and A. Watson, *The optimum distance at which to determine the size of a giant air shower*, Astroparticle Physics **26**, 414 (2007).

[102] Pierre Auger Collaboration, P. Abreu *et al.*, *The Pierre Auger Observatory I: The Cosmic Ray Energy Spectrum and Related Measurements*, (2011), arXiv:1107.4809.

[103] E. S. Seo *et al.*, *Measurement of cosmic-ray proton and helium spectra during the 1987 solar minimum*, The Astrophysical Journal **378**, 763 (1991).

[104] M. Aguilar *et al.*, *Relative Composition and Energy Spectra of Light Nuclei in Cosmic Rays: Results from AMS-01*, The Astrophysical Journal **724**, 329 (2010).

[105] M. Aguilar *et al.*, *Isotopic Composition of Light Nuclei in Cosmic Rays: Results from AMS-01*, The Astrophysical Journal **736**, 105 (2011).

[106] P. Auger, P. Ehrenfest, R. Maze, J. Daudin, and R. A. Fréon, *Extensive Cosmic-Ray Showers*, Rev. Mod. Phys. **11**, 288 (1939).

[107] Particle Data Group, J. Beringer *et al.*, *Review of Particle Physics*, Phys.Rev. **D86**, 010001 (2012).

[108] L. Anchordoqui *et al.*, *High energy physics in the atmosphere: phenomenology of cosmic ray air showers*, Annals of Physics **314**, 145 (2004).

[109] W. Heitler, *The Quantum Theory of Radiation.* Dover Books on Physics and Chemistry (Dover Publications, 1954).

[110] J. Matthews, *A Heitler model of extensive air showers*, Astroparticle Physics **22**, 387 (2005).

[111] J. R. Hörandel, *Cosmic rays from the knee to the second knee: 10^{14} to 10^{18} eV*, Modern Physics Letters A **22**, 1533 (2007).

[112] J. Alvarez-Muñiz, R. Engel, T. K. Gaisser, J. A. Ortiz, and T. Stanev, *Hybrid simulations of extensive air showers*, Phys. Rev. D **66**, 033011 (2002).

[113] J. Linsley, *Structure of large air showers at depth 834 g/cm^2. III - Applications*, International Cosmic Ray Conference **12**, 89 (1977).

[114] K.-H. Kampert and M. Unger, *Measurements of the cosmic ray composition with air shower experiments*, Astroparticle Physics **35**, 660 (2012).

[115] D. Heck, G. Schatz, T. Thouw, J. Knapp, and J. Capdevielle, *CORSIKA: A Monte Carlo code to simulate extensive air showers*, (1998).

[116] S. Sciutto, *AIRES: A System for air shower simulations. User's guide and reference manual. Version 2.2.0*, (1999), arXiv:astro-ph/9911331.

[117] S. Roh *et al.*, *Comparison of CORSIKA and COSMOS simulations*, **1367**, 177 (2011), arXiv:1104.1005.

[118] T. Pierog and K. Werner, *Muon Production in Extended Air Shower Simulations*, Phys. Rev. Lett. **101**, 171101 (2008).

[119] S. Ostapchenko, *Monte Carlo treatment of hadronic interactions in enhanced Pomeron scheme: QGSJET-II model*, Phys. Rev. D **83**, 014018 (2011).

[120] E.-J. Ahn, R. Engel, T. K. Gaisser, P. Lipari, and T. Stanev, *Cosmic ray interaction event generator* SIBYLL *2.1*, Phys. Rev. D **80**, 094003 (2009).

[121] J. Knapp, D. Heck, S. Sciutto, M. Dova, and M. Risse, *Extensive air shower simulations at the highest energies*, Astroparticle Physics **19**, 77 (2003).

[122] S. P. Knurenko and A. Sabourov, *The depth of maximum shower development and its fluctuations: cosmic ray mass composition at E_0>10^{17} eV*, Astrophysics and Space Sciences Transactions **7**, 251 (2011).

[123] Pierre Auger Collaboration, K.-H. Kampert, *Highlights from the Pierre Auger Observatory*, (2012), arXiv:1207.4823.

[124] W. Apel *et al.*, *Energy spectra of elemental groups of cosmic rays: Update on the {KASCADE} unfolding analysis*, Astroparticle Physics **31**, 86 (2009).

[125] M. Dova, M. Manceñido, A. Mariazzi, T. McCauley, and A. Watson, *The mass composition of cosmic rays near 1018 eV as deduced from measurements made at Volcano Ranch*, Astroparticle Physics **21**, 597 (2004).

[126] J. Linsley, *Volcano Ranch evidence on cosmic ray composition*, (1973).

[127] P. Bassi, G. Clark, and B. Rossi, *Distribution of Arrival Times of Air Shower Particles*, Phys. Rev. **92**, 441 (1953).

[128] M. Dova *et al.*, *Time asymmetries in extensive air showers: A novel method to identify {UHECR} species*, Astroparticle Physics **31**, 312 (2009).

[129] L. Cazón, R. Vázquez, A. Watson, and E. Zas, *Time structure of muonic showers*, Astroparticle Physics **21**, 71 (2004).

[130] L. Cazón, R. Vázquez, and E. Zas, *Depth development of extensive air showers from muon time distributions*, Astroparticle Physics **23**, 393 (2005).

[131] L. Cazon, R. Conceição, M. Pimenta, and E. Santos, *A model for the transport of muons in extensive air showers*, Astroparticle Physics **36**, 211 (2012).

[132] D. García-Gámez, *Muon Arrival Time distributions and its relationship to the mass composition of Ultra High Energy Cosmic Rays: An application to the Pierre Auger Observatory*, PhD thesis, Universidad de Granada, 2010.

[133] R. Ulrich and L. Cazon, *The non-linearity between and induced by the acceptance of fluorescence telescopes*, Astroparticle Physics **38**, 41 (2012).

[134] Barcikowski, E. *et al.*, *Mass composition working group report*, EPJ Web of Conferences **53**, 01006 (2013).

[135] L. G. Dedenko *et al.*, *The composition of the primary particles at energies 3×10 17 –3×10 19 eV observed at the Yakutsk array*, Journal of Physics G: Nuclear and Particle Physics **39**, 095202 (2012).

[136] L. Cazon, *Modelling the muon time distribution in extensive air showers*, PhD thesis, Universidade de Santiago de Compostela, 2004, ISBN 84-9750-467-4.

[137] E. M. M. dos Santos, *Measurement of the muon production depth with a time-domain fit at the Pierre Auger Observatory*, PhD thesis, Universidade de Lisboa, 2014.

[138] B. Smith, C. Wileman, and A. Watson, GAP Report No. 110, 2007 (unpublished).

[139] P. Billoir, J. L. Navarro, and S. Navas, GAP Report No. 038, 2011 (unpublished).

[140] E. J. Ahn *et al.*, GAP Report No. 078, 2009 (unpublished).

[141] Pierre Auger Collaboration, P. Abreu *et al.*, *Interpretation of the Depths of Maximum of Extensive Air Showers Measured by the Pierre Auger Observatory*, JCAP **1302**, 026 (2013), arXiv:1301.6637.

[142] S. Andringa, L. Cazon, R. Conceicao, and M. Pimenta, *The Muonic longitudinal shower profiles at production*, Astropart.Phys. **35**, 821 (2012), arXiv:1111.1424.

[143] L. Cazon *et al.*, GAP Report No. 068, 2012 (unpublished).

[144] A. M. Taylor, M. Ahlers, and F. A. Aharonian, *The need for a local source of UHE CR nuclei*, Phys.Rev. **D84**, 105007 (2011), arXiv:1107.2055.

[145] D. Allard, *Extragalactic propagation of ultrahigh energy cosmic-rays*, Astropart.Phys. **39-40**, 33 (2012), arXiv:1111.3290.