Departamento de Física Teórica y del Cosmos
Universidad de Granada

HEAVY QUARK SIMULATION AND IDENTIFICATION AT THE PIERRE AUGER OBSERVATORY

Author:
Alberto Gascón Bravo

Directors:
Antonio Bueno Villar
Julio Lozano Bahílo

Doctoral program:
Programa Oficial de Posgrado en Física (P16.56.1)

– December 2013 –

Editor: Editorial de la Universidad de Granada
Autor: Alberto Gascón Bravo
D.L.: GR 1030-2014
ISBN: 978-84-9028-978-5

D. Antonio Bueno Villar y D. Julio Lozano Bahilo

CERTIFICAN: que la presente memoria, HEAVY QUARK SIMULATION AND IDENTIFICATION AT THE PIERRE AUGER OBSERVATORY, ha sido realizada por D. Alberto Gascón Bravo bajo su dirección en el Departamento de Física Teórica y del Cosmos, así como que éste ha disfrutado de estancias en el extranjero, durante un periodo superior a tres meses, en el Laboratório de Instrumentaçao e Física Experimental de Partículas (LIP) en Lisboa, Portugal.

Granada, a 2 de diciembre de 2013

Fdo: Antonio Bueno Villar

Fdo: Julio Lozano Bahilo

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

Granada, a 20 de noviembre de 2013

Directores de la tesis

Fdo: Antonio Bueno Villar

Fdo: Julio Lozano Bahilo

Doctorando

Fdo: Alberto Gascón Bravo

Agredecimientos

It was the best of times, it was the worst of times [...]
we had everything before us, we had nothing before us.
Charles Dickens, A Tale of Two Cities

Si de alguna manera me gustaría empezar este tesis es dando las gracias. 4 años (largos) es un camino que difícilmente se recorre solo, y hay muchas personas que durante este tiempo, conscientemente o sin saberlo, han estado a mi lado y me han ayudado. Son muchos los gigantes a cuyos hombros he ido subido, y si no he conseguido ver más lejos no ha sido por su falta de altura, sino por mi falta de miras.

Los primeros son mis directores, Antonio y Julio. Mis problemas han sido sus problemas, y sin ellos desde luego que esta tesis no habría salido adelante. Entre otras muchas cosas, de ellos he aprendido el valor de la rectitud científica y de un trabajo bien hecho. Muchas gracias por todo.

También mis padres, que a cada paso me han dado su apoyo y me han escuchado. En cada momento de desánimo, de duda o de flaqueza, han sabido ayudarme sin interferir, aconsejarme dejándome elegir. Y a mis hermanas, Pilar e Irene, sin las cuáles no sería quien soy hoy.

Suelo decir que aparte de dos hermanas tengo seis hermanos más. Álvaro, Javi, Jose, Julián, Rafa y Rafa, aunque las circunstancias nos separen, seguís estando ahí. Aunque un poco más tarde, Ari, Laura y Lourdes os habéis unido a nosotros. Sois todos para mí una segunda familia de la que me acuerdo más si cabe cuando no os tengo cerca.

Es mucha gente a la que he podido conocer en el despacho 29. Ahora soy yo el que lo dejo, pero guardaré siempre el grato recuerdo del tiempo que he pasado allí. Por una lado están lo que ya terminaron. Diego, Luis, Roberto y Adrián. Muchas gracias por los momentos que hemos compartido y por vuestra ayuda. Con los que siguen aquí he pasado 4 años compartiendo muchas horas, tanto dentro como fuera del despacho, que dan para conocerse bien. Bruno, empezamos en esto a la vez y terminamos a la vez también. Quien sabe si en algún momento estaremos trabajando juntos de nuevo. Estoy seguro de que Gloria y tú sereis unos padres maravillosos para Pablito. Laura, han sido muchas de horas de despacho aguantando juntos hasta tarde y compartiendo alguna pena de vez en cuando. Ya mismo estarás escribiendo tú un párrafo como éste. Patricia, aunque sueles ver las cosas más negras de la cuenta, casi nos alcanzas. Mucho

ánimo, seguro que todo te saldrá bien. No podría nunca olvidarme de "los teóricos". Rafa, con su buen humor. Alice, con la sonrisa más sincera que he visto. Mariano, con su finísimo sentido del humor. Ben, del que me asombrará siempre lo claro que tiene quién y cómo quiere ser. De todos echaré algo de menos. María Dolores y Ginés, muchas gracias también, ya que parte del trabajo que aquí hay os lo debo a vosotros.

Como fuera del despacho también hay mundo, aunque a veces no lo parezca, hay más gente con la que he coincidido y a la que querría dar las gracias. En los casi 4 meses que estuve de estancia en Lisboa tuve la suerte de vivir con Sonia. Creo que eres la única persona con la que me he "peleado" para poder cocinar. Esos meses no habrían sido lo mismo sin ti. Un beijinho.

Aurora, siempre recordaré los viajes, tanto los que hicimos como los que se quedaron en el papel, y la parte de este camino que hicimos juntos. Muchas gracias a ti también.

En mi primer año en Granada acabamos formando un grupo todos los que no pertenecíamos a ninguno. Azahara, Domi, Marco, Djamil, Matthias, Olga, Gerard, Luis, Timo, Reme, Alex, Christian; un abrazo muy fuerte para todos.

Ana y Tere, si tuviera que hacer una lista de todas las películas que hemos visto (siempre fieles al miércoles) creo que hasta nos asustaríamos. Sara y Laura, me acogísteis como a uno más en vuestra pequeña familia. 6 meses de brasero y de llevarnos el café al parque no van a ser fáciles de olvidar. Rebeca, te conocí demasiado tarde pero que sé que te conoceré durante mucho tiempo. Aunque no lo creas me has ayudado mucho en este último tramo.

No gostaria acabar estes reconhecimientos sem umas palavras para as pessoas que conheci no LIP: Mário, Catarina, Lorenzo, Pedro, Bernardo, Sofia. Muito obrigado pela vosa acolhida e ajuda. Foi um prazer trabalhar con vocês. Fiquei namorado de Lisboa, e sei que vou ter sempre saudades dela.

Contents

Introducción

A comienzos del siglo XX, una serie de avances tecnológicos dieron lugar al descubrimiento de la radiación cósmica, y al de las cascadas de rayos cósmicos pocas décadas después. Durante los últimos 100 años los rayos cósmicos han sido objeto de un profundo estudio para intentar ampliar el conocimiento que tenemos de su naturaleza, sus propiedades y su origen. El espectro de rayos cósmicos abarca 11 órdenes de magnitud, desde los rayos cósmicos solares con energías próximas al GeV, hasta los rayos cósmicos ultraenergéticos (o UHECR, acrónimo en inglés para *Ultra High Energy Cosmic Rays*), que sobrepasan energías de 10^{20} eV. Sin embargo, el flujo de rayos cósmicos disminuye muy rapidamente conforme aumenta su energía. Los rayos cósmicos poco energéticos pueden detectarse con instrumentos instalados en satélites. Por el contrario, de los rayos cósmicos de más energía no esperamos más de una partícula por siglo y kilómetro cuadrado, lo que reduce enormemente la posiblidad de detectarlos directamente. Con un flujo tan reducido es muy difícil detectar grandes cantidades de cascadas lo que hace que tanto el origen, como la composición y los mecanismos de aceleración de los rayos cósmicos de mayor energía sean aún una incógnita. Aún así, al tener estos rayos cósmicos energías inaccesibles a los aceleradores actuales, cualquier avance en el conocimiento de sus propiedades se convierte inmediatamente en un avance en el campo de la física de partículas. Cuando los rayos cósmicos llegan a la tierra interactúan con las moléculas de los gases que forman la atmósfera, produciendo nuevas partículas. Los secundarios de esta interacción se propagan y vuelven a interactuar, produciendo a su vez más partículas. Este proceso se repite generando millones de partículas dando lugar a una Cascada de Rayos Cósmicos (o EAS, de *Extensive Air Shower*, en inglés). Las propiedades de una EAS varían en función de la naturaleza y de la energía del primario que la genera. Normalmente, medimos las características de una EAS, y a través de éstas intentamos inferir las propiedades del primario.

El Observatorio Pierre Auger se diseñó con el objetivo de estudiar las propiedades de los rayos cósmicos con energías a partir de 10^{18} eV hasta [1]. Es el detector de rayos cósmicos más grande del mundo, y emplea un método híbrido de detección, combinando detectores de fluorescencia y de superficie. El detector de superfice (SD, de *Surface Detector*) usa tanques de agua para detectar la radiación Cherenkov de las partículas que los atraviesan y de esa forma medir la densidad de partículas en el suelo. El detector de fluorescencia (FD, de *Fluorescence Detector*) recoge la luz ultravioleta emitida por el nitrógeno atmosférico, excitado por las partículas cargadas de la cascada. La detección de los hadrones que forman la cascada está más allá de las capacidades del detector, pero sus interacciones y sus desintegraciones generan partículas que pasan a formar parte de la cascada electromagnetica, contribuyendo a su desarrollo. Por otro lado, los muones que se producen en las desintegraciones de piones y kaones se detectan en los tanques del detector de superficie. El Observatorio Pierre Auger ha obtenido muchos resultados en el campo de UHECR. Algunos de los más importantes son la medida del límite superior del espectro de rayos cósmicos [2], el análisis de su composición [3], la obtención de límites superiores a los flujos de fotones y neutrinos [4, 5, 6] y la estimación de la sección eficaz protón-aire [7, 8].

Las interacciones y los procesos que ocurren en los aceleradores suceden también, pero a

energías mucho mayores, en las interacciones de rayos cósmicos en la atmósfera. La producción de quarks pesados (*charm* y *bottom*) y la propagación de los hadrones que éstos forman, que normalmente no se tiene en cuenta en la simulación de EAS, es uno de estos procesos. Debido a su corta vida media a energías por debajo de 0.1-1 EeV, se considera que su efecto en la propagación de EAS en la atmósfera es despreciable. Sin embargo, si se producen con una energía suficientemente alta su longitud de desintegración pasa a ser mayor que su longitud de interacción, de forma que pueden interactuar en la atmósfera y alcanzar grandes profundidades en lugar de desintegrarse nada más producirse. Para la elaboración de esta tesis hemos estudiado estas interacciones de hadrones pesados con núcleos de aire, empleando un modelo de interacción específico para éstas [9, 10]. Si además estos hadrones pesados se producen con una fracción suficientemente elevada de la energía de la partícula primaria su propagación puede afectar significativamente al desarrollo longitudinal de la cascada. En 1982 se hicieron las primeras sugerencias de que, en base a estas propiedades de su propagación, sería posible inferir la presencia de hadrones pesados en EAS [11].

La Cromodinámica Cuántica (QCD, de *Quantum Chromodynamics*) explica la producción de quarks pesados a la escala de energía de los aceleradores, y su predicción es que se producen con fracciones pequeñas de la energía del proyectil. Un hadrón pesado formado a partir de estos quarks no tendrá ningún impacto en el desarrollo de EAS. Pese a describir con precisión una gran cantidad de observaciones experimentales, hay algunos procesos medidos en aceleradores que QCD no puede explicar. Dos ejemplos son la producción de hadrones pesados con fracciones muy grandes de la energía del proyectil en colisiones pN y πN [12, 13, 14, 15], y la producción de parejas J/Ψ muy energéticas en colisiones de $\pi^- N$ [16, 17]. Existen modelos para la producción de quark pesados que son capaces de explicar esta fenomenología. Estos modelos predicen la producción de quarks pesados con fracciones mucho mayores de la energía del proyectil. En esta tesis investigamos las implicaciones de la producción de quarks pesados mediante uno de ellos, el mecanismo de Quark Intrínseco (IQ, de *Intrinsic Quark*) [18, 19, 20] en el desarrollo longitudinal de EAS, usando los datos recogidos por el Observatorio Pierre Auger.

De los capítulos 1 al 3 hacemos un repaso sobre los rayos cósmicos, la física de EAS y el Observatorio Pierre Auger. El capítulo 1 está dedicado a aspectos generales de la física de rayos cósmicos, desde las circunstancias históricas en las que se desarrolló este campo, hasta la presentación de algunos de sus últimos resultados. En el capítulo 2 hacemos un repaso al desarrollo de EAS en la atmósfera y a los distintos métodos de detección que existen. Los detectores del Observatorio Pierre Auger, y la reconstrucción de EAS que hace cada uno de ellos, aparecen explicados en el capítulo 3.

Los puntos clave en la historia de la detección de quarks pesados, haciendo hincapié en las observaciones en experimentos de rayos cósmicos, aparecen en el capitulo 4. Se explican en detalle la producción (sección 4.2), interacción (sección 4.3) y propagación (sección 4.4) de hadrones pesados dentro de EAS, ya que a energías ultra altas esperamos que tengan un comportamiento diferente del que tienen en los aceleradores de partículas.

La simulación y reconstrucción de cascadas con producción de quarks pesados, y su identificación usando un discriminante multivariable se encuentrann en el capítulo 5. Investigamos qué variables son mejores para distinguir las cascadas en las que se producen hadrones pesados ultraenergéticos (señal) de aquellas en las no se producen (fondo). Examinamos también la eficiencia de selección para cascadas generadas por otros primarios hadrónicos (núcleos de helio, nitrógeno y hierro) y para fotones, ya que éstos últimos constituyen una posible fuente de eventos de fondo.

En el capítulo 6 usamos la eficiencia de selección obtenida en el capítulo anterior para calcular la exposición híbrida del observatorio, y el número esperado de eventos tanto si hay como si no hay hadrones pesados ultraenergéticos. Comparando el número observado y el esperado de eventos, y conocidas las eficiencias de selección, es posible establecer un límite a la sección eficaz de producción de quarks pesados en el módelo de Quark Intrínseco.

Al final de la tesis se encuentra un estudio del trigger del detector de superficie para cascadas inclinidas (apéndice A). Además hay una descripción detallada de las modificaciones realizadas en el código de Monte-Carlo para poder simular la propagación de hadrones pesados (apéndice

B), así como del modelo teórico detrás de la producción intrínseca de quarks (apéndice C).

Introduction

At the turn of the 20^{th} century, a series of technological developments led to the discovery of cosmic radiation and of extensive air showers few decades later. During the last 100 years cosmic rays have been the subject of a thorough study to try and increase the knowledge we have of their nature, properties and origin. The cosmic ray spectrum spans over 11 orders of magnitude, from the GeV solar cosmic rays up to the ultra-energetic cosmic rays (UHECR) with energies above 10^{20} eV. The flux of cosmic rays decreases rapidly with increasing energy. The lowest energy cosmic rays can be detected directly with satellite detectors but UHECR, on the upper limit of the spectrum, are expected with a rate of about 1 particle per century and square kilometer. Such a meager flux difficults the detection of large samples of showers and consequently the origin, chemical composition and acceleration mechanisms of UHECR are to a great extent unknown. Having energies far above those achieved at accelerators, any increase in the knowledge of UHECR influences directly particle physics. When cosmic rays reach the Earth they collide with atmospheric nuclei, producing new particles. The secondaries of this interaction undergo the same process, eventually producing millions of particles that propagate through the atmosphere and reach the Earth surface. This process is known as an Extensive Air Shower (EAS). The properties of EAS vary depending on the nature and the energy of the primary particle inducing it. Current studies try to extract information of the primaries from the analysis of the showers they generate.

The Pierre Auger Observatory has been conceived to study the properties of cosmic rays with energies above 10^{18} eV [1]. It uses a hybrid detection technique, combining a fluorescence and a surface detector to detect cosmic rays, and it is currently the world largest cosmic ray observatory. The surface detector (SD) samples the densities of particles on ground, using an array of Water-Cherenkov tanks. The fluorescence detector (FD) collects the isotropic ultraviolet fluorescence light emitted by atmospheric nitrogen, excited by the secondary charged particles in the shower. The detection of the hadronic component of the shower exceeds the capabilities of the detector, but the interactions and decays of hadrons produce particles that feed the electromagnetic shower, contributing to the development of the EAS. Moreover, muons produced by the decay of pions and kaons are detected in the Water-Cherenkov tanks of the SD. With the data colleted by the observatory it is possible to perform many different physics analyses. The Pierre Auger Observatory has reported many results in the field of UHECR. The measurements of the upper-end of the energy spectrum [2], the analysis of the mass composition of cosmic rays [3], the derivation of upper limits on the cosmic-ray photon and neutrino fluxes [4, 5, 6], or the estimation of the proton-air cross-section [7, 8] are among the most important scientific outcomes of the collaboration.

The processes that take place at accelerators are replicated, with larger energies, in cosmic rays interactions. One of these processes, normally neglected in EAS simulations, is that of heavy quark production and heavy hadron propagation. The short mean lives of heavy hadrons over a wide range of energies make them unimportant in the propagation of showers in the atmosphere unless they are produced with sufficiently high energy. In that case their decay lengths are longer than their interaction lengths and they will suffer a number of interactions

and reach deeper in the atmosphere, rather than decaying just after their production. In the context of this Ph.D thesis we have studied these interactions of heavy hadrons with air nuclei in the atmosphere, using a dedicated interaction model [9, 10]. If, in addition, these heavy hadrons carry a large fraction of the primary particle energy then they will significantly affect the shower development. The first suggestions of the possibility of looking for heavy hadrons in ground based cosmic ray observatories relying on their propagation properties dates back to 1982 [11]. Quantum Chromodynamics (QCD) explains the production of heavy quarks at accelerator energies, and predicts heavy quarks produced with small fractions of the primary energy. Any hadron formed from these quarks can hardly have any influence in the EAS development. Though very sucessful at describing a broad collection of observations, QCD is unable to explain part of the phenomenology seen at accelerators, namely the presence of leading charmed and bottom hadrons in pN and πN collisions [12, 13, 14, 15] or the production of highly energetic pairs of J/Ψ in $\pi^- N$ interactions [16, 17]. Other models for heavy quark production, able to account for this phenomenology, predict heavy hadrons with much larger fractions of the primary energy. In this Ph.D thesis we explore the implications of heavy quark production by the Intrinsic Quark mechanism [18, 19, 20] in the EAS longitudinal development using the data collected by the Pierre Auger Observatory.

Chapters 1 to 3 give an overview of cosmic rays, EAS and the Pierre Auger Observatory. Chapter 1 is devoted to the general aspects of cosmic rays physics. The historical circumstances under which the research on cosmic rays started are introduced, followed by some of the most recent advances in the field. Chapter 2 describes the basics of EAS development in the atmosphere, and the various detection techniques employed to detect them. The instruments of the Pierre Auger Observatory, and the reconstruction of the showers as detected by the surface and fluorescence detector are discussed in chapter 3.

In chapter 4 we make a historical approach to the cornerstones of heavy quark detection, with special emphasis on their observation (proved or claimed) in cosmic ray experiments. We focus on the problem of treating the production (section 4.2), interactions (section 4.3) and propagation (section 4.4) of heavy hadrons inside EAS at ultra-high energies which we expect to be significantly different to their behavior at accelerators.

The simulation and reconstruction of showers with production of UHE heavy hadrons is described in chapter 5. The identification of these showers in the FD, using a multivariate analysis method (Boosted Decision Trees), is described. We investigate which variables better distinguish showers where heavy hadrons were produced (signal showers) from those with no heavy hadron production (background showers). The selection efficiency of showers induced by other hadronic primaries (helium, nitrogen and iron nuclei), and of photon-induced showers is inspected too. The latter could be a potential source of background showers.

In chapter 6 we use the selection efficiency obtained in the previous chapter to compute the hybrid exposure of the observatory, and the expected rates of events in the presence or the absence of UHE heavy hadrons. Comparing the expected rate of events with the actual number of detected events it is possible to set a limit to the production of heavy quarks in the Intrinsic Quark model.

At the end of the document there is a study of the trigger of the surface detector for inclined showers (appendix A). We also give a detailed description of the modifications made to the Monte-Carlo code to allow for heavy hadron propagation (appendix B), and the theoretical basis of the the heavy quark production model implemented (appendix C).

1

Cosmic Ray Radiation

With a history spanning more than a century, cosmic rays physics is one of the most enduring fields of research in physics. At the beginning of the 20^{th} century, the improvement of the electrometer by Theodor Wulf gave rise to the first experiments able to measure the rate of ionization inside sealed containers. Wulf (in 1910) and Domenico Pacini (in 1911) studied the dependence of air ionization with altitude, concluding that part of the ionization was due to sources other than the Earth. However, their results were not widely accepted. Radioactivity had been discovered in 1896, and the general agreement was that the analysis was compatible with the terrestrial origin of the radiation. Any discrepancies were attributed to residual radioactive components present in the electrometers, or to an incomplete knowledge of the absorption properties of air. The existence of some externally originated radiation became recognized only after 1912, when Victor Hess measured the ionization rate of air as a function of altitude, flying electrometers to an altitude of 5300 m [21]. He found an increase of ionizing radiation with height, and inferred that radiation was penetrating the atmosphere from outer space. He measured radiation levels during a solar eclipse obtaining essentially the same results, hence concluding that cosmic rays were mostly coming from a source other than the sun. Werner Kolhörster confirmed Hess' results during 1913-1914, measuring the ionization rate at a height of 9 km. The name cosmic rays was coined by Robert Millikan, who believed that the primary particles were gamma rays, the most penetrating radiation known at that moment, generated during the nucleosynthesis of carbon and oxygen elements. However, in 1930 J. Clay concluded that cosmic rays consisted of charged particles, based on the dependence of their intensity with magnetic latitude. This dependence was confirmed in three different experiments by Thomas H. Johnson, Arthur Compton and Bruno Rossi [22, 23, 24], who also proved that most primaries are positively charged. The term cosmic rays now refers to charged nuclei of extraterrestrial origin. In 1938, Kolhörster [25] and Pierre Auger [26] reported, in different experiments, coincident signals between separated detectors. They concluded that the particles triggering the detectors were generated in the atmosphere, originating from a single primary cosmic ray, in a process now called Extensive Air Shower (EAS). Using balloons at high altitudes, Marcel Schein showed that primary particles with positive charge were atomic nuclei moving close to the speed of light [27]. The relative abundance of nuclei up to Z$\approx$40 was determined, finding hydrogen and helium to be the most frequent primaries, and elements more massive than iron to be rare. Starting in the mid 1940s, large detector arrays were built to measure EAS. The pioneering ground arrays of Volcano Ranch [28], SUGAR [29], Haverah Park [30], Yakutsk [31] and Akeno [32] made many important contributions to our understanding of the physics of cosmic rays above 10^{17} eV.

In the late 70's the first successful fluorescence light detector, the Fly's Eye [33], was set up. The combination of both surface detector arrays and fluorescence detector has further improved the knowledge of the features of the energy spectrum. The Telescope Array [34] and the Pierre Auger [35] experiments are outstanding examples of the hybrid technique.

Ultra High Energy Cosmic Rays (UHECR) are the most energetic particles observed in

nature, with energies about 10^{20} eV. The observation of these cosmic rays evidences some of the unknowns still present in the field:

- What is the origin of the different features present in the energy spectrum?
- Is there an end to the spectrum of cosmic rays? At what energy does it occur?
- What is the composition of cosmic rays as a function of energy? Is it predominantly composed of heavy elements or is it light-like? Is the composition pure or a mixture of different species?
- What are the sources of the highest energy cosmic rays? Is there anisotropy in the cosmic ray flux that points to sources? What are the acceleration mechanisms capable of generating cosmic rays of energies up to 10^{20} eV?

Measuring the cosmic ray flux, composition and arrival directions at the highest energies is fundamental to separate the different scenarios of origin and propagation of cosmic rays. In the following sections we will review the results from each of these areas.

1.1 Spectrum of Cosmic Rays: from the Sun to the GZK-cutoff

The solar system, and hence the Earth, is continuously exposed to a flux of cosmic rays, their energies ranging from the low energetic MeV cosmic rays up to $\sim 10^{20}$ eV. The flux as a function of the energy, the spectrum, follows a power law ($J(E) \propto E^{-\gamma}$) with almost no features: γ, the spectral index, is almost constant with $\gamma \simeq$ -2.7 above the GeV. Figure 1.1 (left) shows a general picture of the spectrum from the lowest energies keV cosmic rays up to the highest energies. The flux decreases steeply as a function of energy: from about 1000 particles per second and m^2 at energies around the GeV, to less than one particle per century and km^2 above 100 EeV[1]. Such a strong decrease in flux poses an experimental challenge: with increasing energy our knowledge about the nature of the primaries and their origin becomes increasingly limited. A closer inspection of the flux reveals some distinct features. Around 3-5·10^{15} eV the spectral index changes from $\simeq$ -2.7 to $\simeq$ -3.1, steepening the flux [36]. This change is referred to as the *knee*. At about 4·10^{18} eV, the slope flattens again. This change is known as the *ankle* (see figure 1.1, right). In the next sections, we review the origin and characteristics of cosmic rays across the spectrum.

1.1.1 From the lowest energies to the GeV: Solar and Anomalous Cosmic Rays

The lowest energy cosmic rays are produced in the sun, originated in violent events such as solar flares and coronal mass ejections. Their maximum energies are of the order of $\sim$ 0.1-10 GeV [37]. The end of the spectrum of Solar Cosmic Rays overlaps with the so called Anomalous Cosmic Rays. This component arises from the interaction of the neutral interstellar gas with the heliosphere. The abundances of nuclei of this component are very different from the abundances found in Galactic Cosmic Rays, hence their name. The particles forming the gas penetrate to the inner heliosphere, where solar UV radiation ionizes them. Once charged, the solar magnetic field transports them to the outer heliosphere, where they accelerate by a mechanism called diffusive shock acceleration [38]. The ions repeatedly collide with the termination shock wave in the solar wind, gaining energy in the process.

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

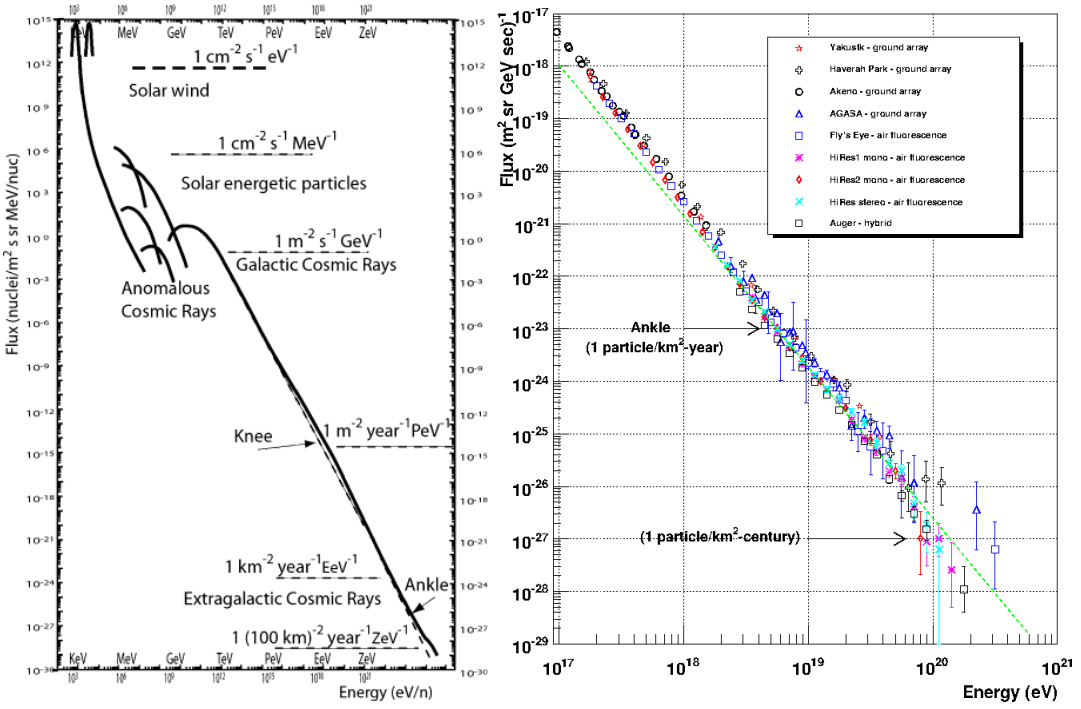


Figure 1.1: Left: All-particle flux, from solar cosmic rays up to the GZK cut-off. The dashed line corresponds to a power law with spectral index $\gamma = 3$ [37]. Right: Zoom of the cosmic rays spectrum on the ultra-high energy regime, from 10^{17} eV up to the highest energies, with data from several experiments.

1.1.2 Galactic Cosmic Rays: acceleration at SNRs and the Knee

At even larger energies the sources of cosmic rays are outside the solar system, but still in the vicinity of the galaxy. Galactic Cosmic Rays are believed to be accelerated at supernova remnants (SNRs). Their large size and long life allow for the process of acceleration to be carried on. A rate of about three supernovae per century in a typical galaxy would account for the energy of Galactic Cosmic Rays if only a 5-10% of the kinetic energy released were transferred to accelerated particles [39].

The mechanism of acceleration at SNRs is similar to the acceleration in the solar wind. The particles travel from the upstream (unshocked) region to the downstream region and back. After each cycle the particle gains an amount of energy $\Delta E \propto E$, but there is a certain probability that it will exit the shock region and will not return. After a time T the maximum energy attained is $E_{max} \sim Ze\beta_s BTV_s$ where Ze is the charge of the particle, $\beta_s = V_s/c$ is the velocity of the shock and B is the intensity of the magnetic field [40]. Using average values from Type II supernovae in a typical interstellar medium this upper limit turns out to be $E_{max} \sim Z \cdot 10^{14}$ eV [41].

The energy range in which the cosmic ray spectrum changes its slope is called the *knee*. Up to 10^6 GeV, the spectrum is a power law with spectral index $\gamma \sim 2.70$-2.75. Above the knee it increases by $\Delta\gamma \sim 0.3$. The KASCADE experiment has shown that the knee in the all-particle spectrum is due to a decrease in the flux of light nuclei [42]. First, around $4{\cdot}10^{15}$ eV a distinct break in the proton flux appears, followed by a break in the helium flux at a slightly larger energy [43]. Figure 1.2 shows the all-particle spectrum obtained with KASCADE-Grande compared to results from other experiments.

Different explanations have been put forward to elucidate the origin of the knee. One group of scenarios assume that the knee results from the interactions of cosmic rays with background particles during their propagation process. This situation would result in the spallation of heavy

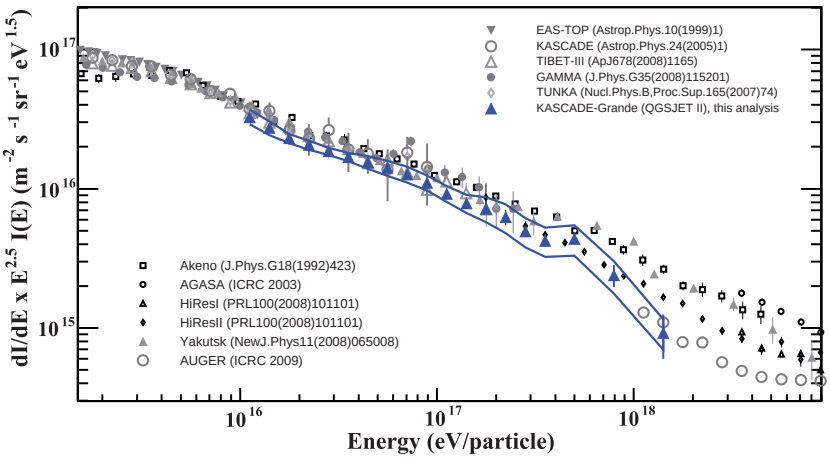


Figure 1.2: All-particle energy spectrum obtained by various experiments. The band corresponds to the systematic uncertainties on the measurements of KASCADE-Grande [36].

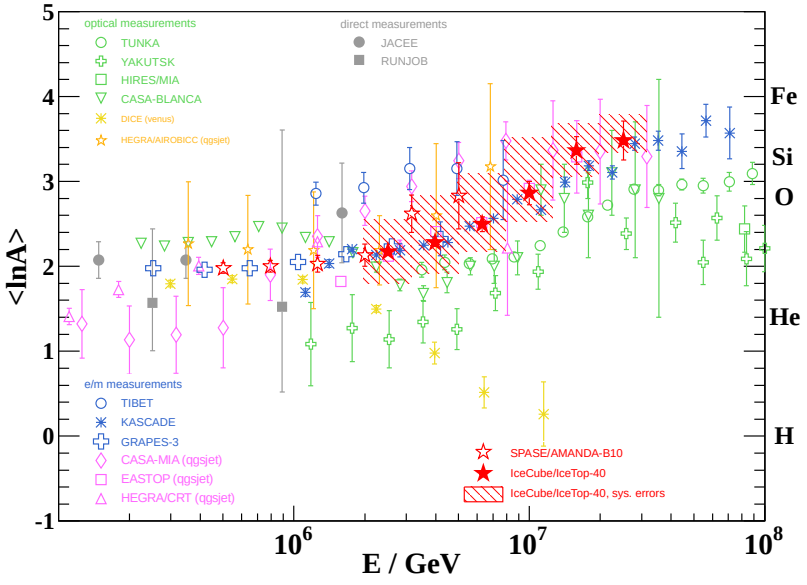


Figure 1.3: Mean logarithmic mass measured by several experiments, showing an increase in primary mass with increasing energy in the range 10^{15} eV to 10^{17} eV [44].

nuclei into lighter particles, that would dominate the spectrum between 10^{16} eV and 10^{17} eV. However, as shown in figure 1.3, data indicate a trend towards heavy composition in this energy range [44].

Other models predict break-offs for individual elements dependent on their charge, Z [45]. In particular, models describing the acceleration of cosmic rays in SNRs predict a maximum energy achievable during the acceleration proportional to Z and the magnetic field at the source. On the other hand, models describing the diffusive propagation of cosmic rays in the galaxy find a maximum rigidity E/Z above which the nuclei of charge Z stop being magnetically bound to the galaxy. In both cases, a knee-like structure in the spectrum of the heavy component is expected in the energy range between $4{\cdot}10^{16}$ eV and $1.2{\cdot}10^{17}$ eV. Recently, the

KASCADE-Grande experiment reported a knee-like feature in the spectrum of heavy particles at $\log_{10}(E/eV) = 16.92 \pm 0.04$ [46], which gives strong support to a rigidity-dependet origin of the knee.

The spectrum that results of assuming any of the two latter models alone shows knees for individual elemental groups which are not compatible with the measured fall-offs [47]. The origin of the knee in the energy spectrum of cosmic rays is most likely a combination of a maximum energy in the sources and the leakage from the galaxy due to the propagation process [48].

1.1.3 Extragalactic Cosmic Rays: the Ankle and acceleration up to 100 EeV

The origin and acceleration of cosmic rays above the ankle are two of the most important open questions. The estimate of the maximum energy released in SNRs is comparable to the position of the knee, so cosmic rays above the knee should have a different origin. The end of the galactic component, and the transition to cosmic rays of extragalactic origin is one of the causes proposed to explain the transition from the knee to the ankle. At some point between these two energies particles would stop being accelerated at galactic astrophysical objects, and cosmic rays of extragalactic origin would make their entry [49]. This transition raises the problem of combining, in a smooth way, two different spectra with different slopes.

A different approach assumes that the same class of sources accelerates all particles up to ankle energies, and that the knee-to-ankle transition results from the propagation of cosmic rays in the galaxy. Once cosmic rays leave their sources, they propagate through the interstellar medium, suffering fragmentation in nuclear reactions, and ionization energy losses due to Coulomb interactions. Both processes modify the cosmic rays composition and energy spectrum. In addition, the galactic and intergalactic magnetic fields deflect their trajectories. Our understanding of the strength and distribution of galactic magnetic fields is very limited, complicating the determination of cosmic rays origin.

The efficient acceleration of cosmic rays (microscopic particles) to energies up to 10^{20} eV (macroscopic energies) is not well understood. *Bottom-up* acceleration scenarios assume that particles at the highest energies originate from low energy particles accelerated at the sources or nearby them. The main bottom-up mechanisms are diffusive shock acceleration, based on the Fermi mechanism [50], and acceleration in very intense electric fields [51].

Considering acceleration through diffusive shocks, in a manner similar to galactic cosmic rays in SNRs, the relation between the maximum energy (E_{max}) attainable by a particle of charge Ze and the magnetic field strength (B) and size (R) of the region of the shock is [52]:

$$E_{max} \simeq \cdot Z \cdot \beta \left(\frac{R}{kpc}\right) \cdot \left(\frac{B}{\mu G}\right) [EeV] \tag{1.1}$$

where β is the shock velocity in units of c. The Hillas plot (figure 1.4) summarizes the candidates to be possible sites of acceleration, though the list is scarce. Among the possible candidates, one finds Active Galactic Nuclei (AGN), Gamma Ray Bursts (GRB), neutron stars and radio galaxies. In principle, all objects above the iron line are suitable sites for acceleration up to the highest energies. However, the energy density and hence the probability of losing energy in the surroundings of the sources is not negligible. Synchrotron radiation, Compton processes or photodisintegration compete with the acceleration process, decreasing the energy of the particle.

Top-down scenarios avoid the need of an accelerator, relying on mechanisms that would produce particles with energies above 100 EeV. For instance, the existence of super-massive unstable relic particles or Topological Defects are invoked to solve many problems related to the acceleration of UHECRs. The decay of these particles produces a cascade of energetic photons, light leptons and neutrinos, along with protons and neutrons and thus no acceleration is required. Independently of the nature of the mechanism invoked, all top-down models share an identifying signature: the high-energy end of the spectrum is dominated by photons and neutrinos. Above a certain energy the fraction of photons should become very large, prediction which is dismissed by the observations of the Pierre Auger Observatory [54, 55].

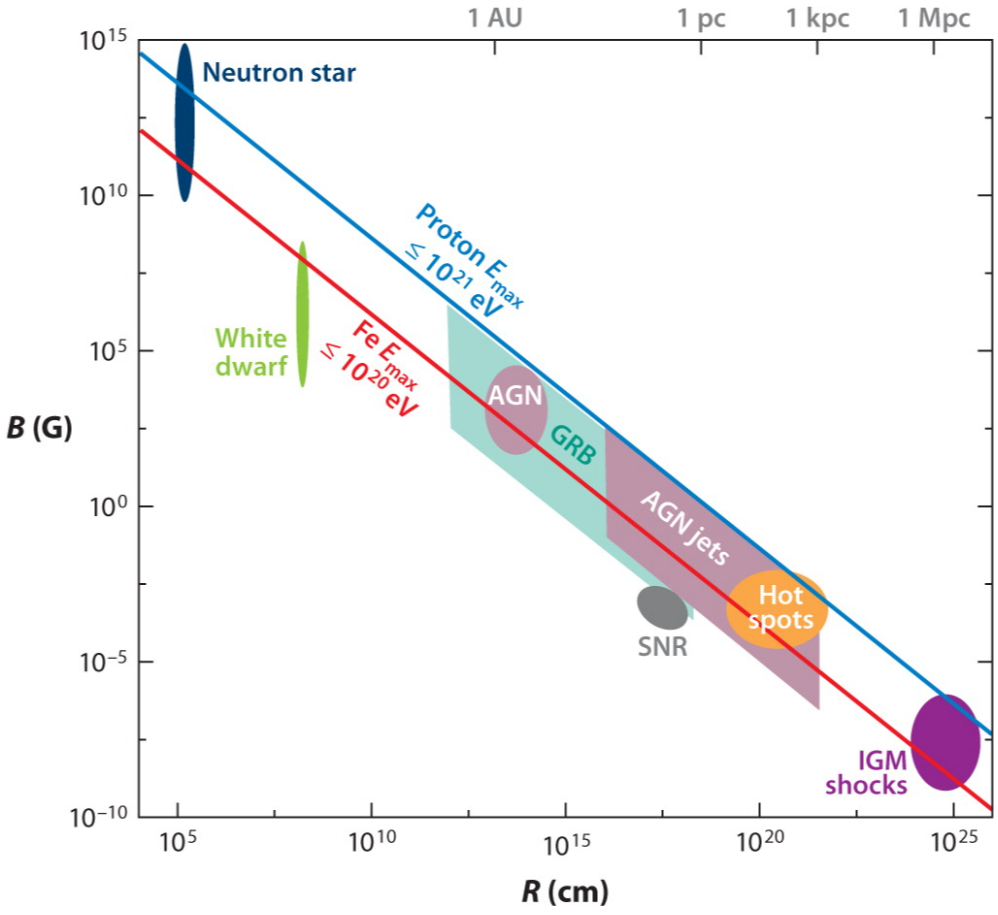


Figure 1.4: Hillas plot, in which astrophysical objects which can be the source of UHE cosmic rays are sorted according to their magnetic field strength and size. Objects below the diagonal line are not suitable sources [53].

1.1.4 The GZK cut-off

A particularly important feature, related to the natural end of the spectrum, is the suppression of the flux at the highest energies. In 1965 Penzias and Wilson discovered the Cosmic Microwave Background radiation (CMB) [56], with which ultra-high energy protons interact via photo-pion production and pair production:

$$p + \gamma_{CMB} \rightarrow \Delta^+(1232) \rightarrow \pi^0 + p \tag{1.2}$$
$$\rightarrow \pi^+ + n$$
$$p \rightarrow p + e^+ e^- \tag{1.3}$$

The energy thresholds for these processes are $\sim 10^{19.6}$ eV and 10^{18} eV, respectively. However, the energy loss per interaction in pair production is only around 0.1%, while for pion production is around 20%, and energy losses are dominated by the latter process. In 1966, a cut-off in the spectrum around $\sim 10^{19.6}$ eV was predicted by Greisen [57], and Zatsepin and Kuzmin [58] (hence termed the GZK cut-off), ascribed to the interactions of protons with the Cosmic Microwave Background (CMB).

UHE nuclei interact with the cosmic radiation backgrounds too, both with the CMB and the infra-red background radiation (IRB), losing energy by photodisintegration and pair production:

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

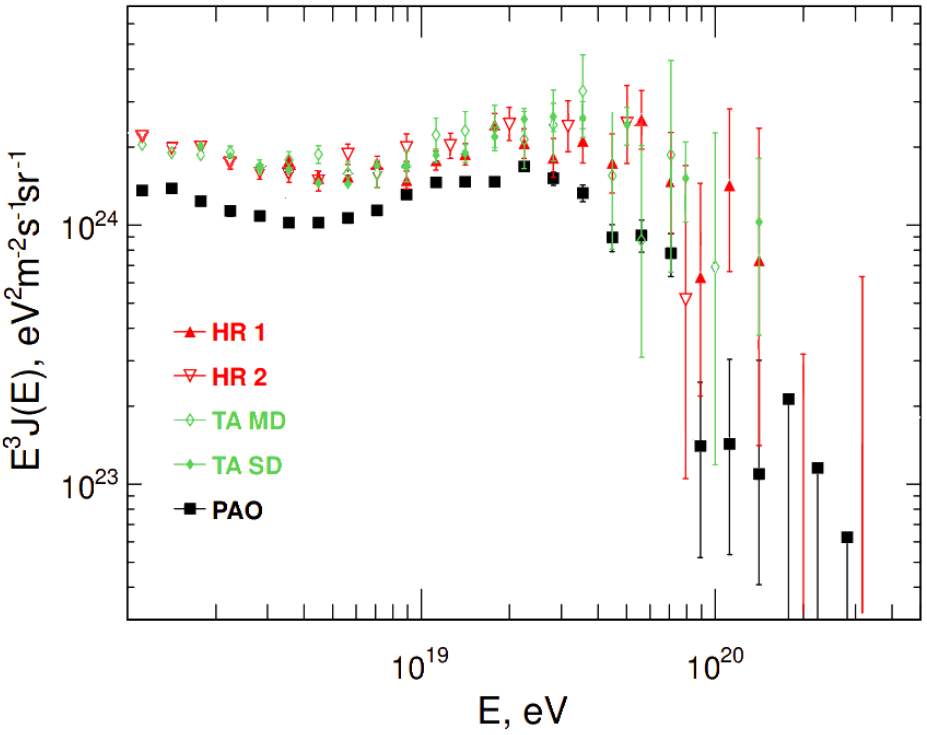


Figure 1.5: Upper end of the spectrum of UHECR as observed by the HiRes, Telescope Array and Pierre Auger experiments [64].

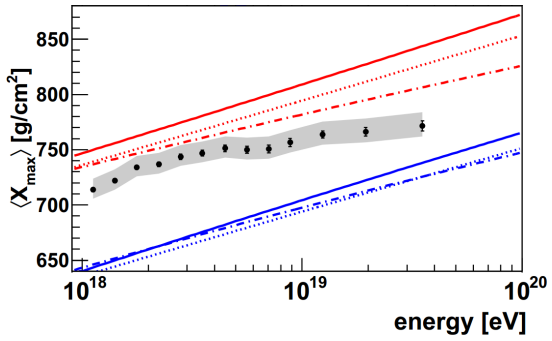
$$\rightarrow \quad A + e^+ + e^- \tag{1.5}$$

where N is a nucleon. The main channel corresponds to n=1, where a proton or a neutron is produced. Double-nucleon emission is possible, but is one order of magnitude less probable than single-nucleon emission [59]. The energy loss due to IR photons is only effective below $5 \cdot 10^{19}$ eV, while energy loss with CMB photons is dominant above $2 \cdot 10^{20}$ eV. Between these two energies, pair production is the most significant source of energy loss in nucleons with background radiation. If UHE cosmic rays are nuclei, a feature similar to the GZK is expected, but the shape and energy threshold of the suppression will be different to that of protons.

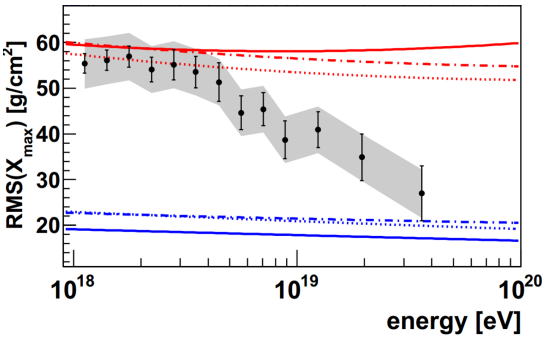
The end of the spectrum for various experiments is shown in figure 1.5. In general all experiments show this flux suppression at the highest energies, which was first established by HiRes [60] and later confirmed by the Pierre Auger Observatory [2]. Recent data from the Telescope Array collaboration also supports this observation [61]. Even though the presence of a flux suppression similar to the GZK cut-off is firmly established, with only AGASA showing no evidence for a flux suppression[62], the possibility that the observed softening is due to the maximum energy attainable at the sources is not easily dismissed [63].

1.2 Mass composition

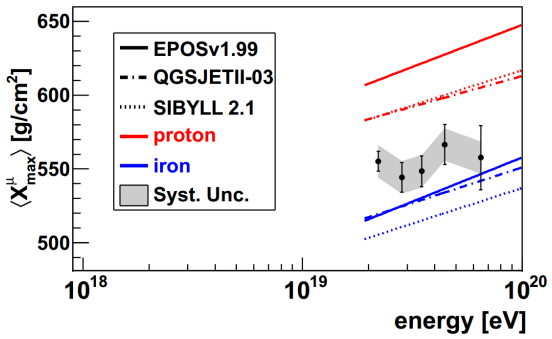
The composition of cosmic rays up to $\sim 10^{14}$ eV can be directly measured with space-based experiments. At higher energies, one has to use the properties of EAS (described in chapter 2) generated by a primary particle to identify it. However, due to the high level of fluctuations in the shower development it is extremely difficult to distinguish showers originating from different hadronic primaries on an event by event basis; it can only be done on a statistically significant set of showers.



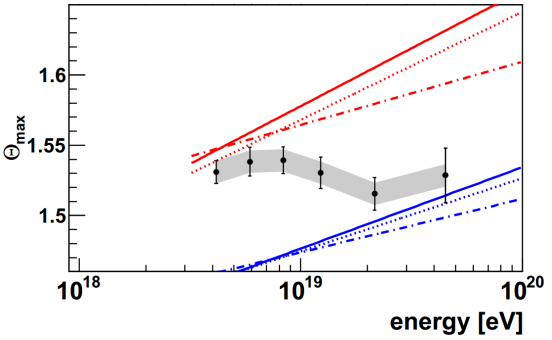
(a) $\langle X_{max} \rangle$ vs energy



(b) $\sigma(X_{max})$ vs energy



(c) $\langle X^{\mu}_{max} \rangle$ vs energy



(d) Θ_{max} vs energy

Figure 1.6: Average X_{max} (a), $\sigma(X_{max})$ (b), average X^{mu}_{max} (c) and Θ_{max} (d) as a function of energy as measured by the Pierre Auger Observatory. Statistical uncertainties are shown as error bars. Systematic uncertainties are represented as a band. Predictions from various hadronic models for proton and iron are presented as different types of lines.

Shower observables that are sensitive to mass composition are needed to discriminate between primary masses. The longitudinal shower development, measured with fluorescence detectors, is different depending on the nature of the primary cosmic ray. This difference is usually quantified observing the depth of the shower maximum X_{max}, the depth of maximum particle production. For identical energies, the average shower maximum $\langle X_{max} \rangle$ of proton-induced showers is around 100 g·cm^{-2} larger than that of showers generated by iron primaries. At the same time, the fluctuations of X_{max}, $\sigma(X_{max})$, become smaller as the mass of the incident primary increases, with values about 40 g·cm^{-2} smaller for iron- than for proton-induced showers. These trends hold regardless of the hadronic interaction model used to describe the shower development [59].

The measurement of these observables from the Pierre Auger Observatory suggests an evolution from light to massive composition starting at 3 EeV when compared to EAS simulations [3]. The results are shown in figures 1.6(a) and 1.6(b).

In spite of the precision achieved by the fluorescence technique, fluorescence detectors have a reduced duty cycle limited to moonless nightly measurements. It is worth investigating surface detector observables sensitive to the shower longitudinal development. Muon signals at ground give an indirect measure of the longitudinal development of the hadronic component of EAS [65, 66]. It is possible to reconstruct the Muon Production Depth distribution (MPD) using the signals of surface detectors far from the shower core. The average depth, along the shower axis, where the number of muons produced reaches a maximum, $\langle X^{\mu}_{max} \rangle$, is another mass sensitive parameter [67]. The azimuthal asymmetry of the rise-time of signals between detectors carries information about the longitudinal development of the showers as well [68]. At a given energy, the zenith angle where the asymmetry amplitude is maximum defines Θ_{max}. Figures 1.6(c) and 1.6(d) show the Pierre Auger Observatory analysis of the evolution of $\langle X^{\mu}_{max} \rangle$ and Θ_{max} with energy, respectively [67].

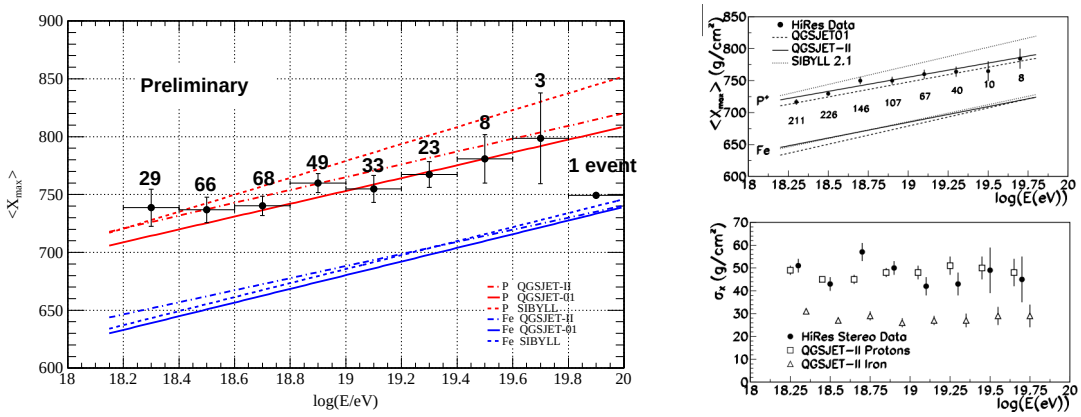


Figure 1.7: Left: $< X_{max} >$ measured by the Telescope Array experiment [74]. Right: $< X_{max} >$ (top) and $\sigma(X_{max})$ (bottom) measured by the HiRes experiment [69]. Lines are predictions from various hadronic models for proton and iron showers.

The analysis of the evolution of $\langle X_{max} \rangle$, $\sigma(X_{max})$, $\langle X^{\mu}_{max} \rangle$ and Θ_{max} with energy suggests a composition becoming consistently heavier up to the highest energy measured though the decrease of $\sigma(X_{max})$ is not consistent with a simple change of the cosmic ray composition from pure proton to pure iron [49]. Data from other experiments claim different results. Telescope Array (figure 1.7, left) and HiRes (figures 1.7, top right and 1.7, bottom right) show a proton dominated composition up to the highest energies [69, 70]. There is a joint effort of the Auger, HiRes, Telescope Array and Yakutsk collaborations to elucidate the origin of this discrepancy. During the 2012 *International Symposium of Future Directions in UHECR* [71] working groups were formed, whose focus was establishing a common view on the experimental status of composition measurements at ultra-high energies. Some of the conclusions they reached were that Auger results are consistent within systematic uncertainties with the results from Telescope Array and Yakutsk, but not fully consistent with HiRes. On their part, Yakutsk, HiRes and TA are consistent within ~ 5 g·cm^{-2}. The compatibility between Auger and HiRes depends on the particular interaction model used to interpret the $\langle X_{max} \rangle$ observations, converting them into $\langle A \rangle$, the logarithmic mean mass. HiRes is compatible with the interpretation of the Auger data only at energies below $10^{18.5}$ eV when using QGSJet-II. When using the SIBYLL model, Auger and HiRes are compatible within a larger energy range [72]. These differences do not seem to stem from issues related to the analysis, such as a time dependence of the results, the range of zenith angles analyzed or the dependence on detector simulations [73]. As such, the chemical evolution of cosmic rays is still an unsettled issue.

1.3 Anisotropies and correlations

One of the keys to understanding the nature of the UHECRs is their distribution over the sky. Measurements of the anisotropies in the distribution of arrival directions of UHECRs, when combined with information on their chemical composition and spectral features can provide valuable information on the nature of the sources and acceleration mechanisms. A precise determination of the arrival direction is fundamental to assess the possible clustering of events from particular directions.

Data collected by the HiRes experiment in the north hemisphere working in stereo mode (detected jointly by its two detectors) was analyzed in search of anisotropies. In this mode the angular resolution in cosmic rays pointing directions is about 0.8°. The analysis from the HiRes data with energy thresholds of 10 EeV and 40 EeV are compatible with an isotropic flux at 95% CL. Data using an energy threshold of 57 EeV were found to be only marginally consistent with an isotropic flux [75].

AGASA has an energy dependent angular resolution, improving from 4° around $10^{18.5}$ eV down to 1.3° at 10^{20} eV. First analyses of AGASA data claimed small scale anisotropies, with 1 triplet and 6 doublets observed using an energy threshold of 4×10^{19} eV within angular windows of 2.5°, whereas only 2 doublets were expected from an isotropy hypothesis [76]. The evidence for clustering in the AGASA data set arising from more recent analyses is weaker than what had been been previously claimed. In fact, it is consistent with the null hypothesis of isotropically distributed arrival directions at the 8% level [77]. The results of statistical tests show that there is no signicant correlation with AGNs, but rather the distribution of AGASA data seems to be isotropic [78].

Most recent experiments, namely the Pierre Auger Observatory and Telescope Array, show different results. The Pierre Auger Observatory is able to reconstruct the direction of cosmic rays with an accuracy of $\sim 1^{\circ}$. Point-like sources searches have resulted in evidence for anisotropy in the distribution of arrival directions of the highest energy events [79, 80, 81]. The arrival directions of the events with energies above 55 EeV show correlation within an angular scale of $\sim 3^{\circ}$ with the positions of nearby (within $\lesssim 75$ Mpc) AGNs from the VCV (Veron-Cetty & Veron) catalog [82], which is above that expected from chance coincidences in an isotropic sky. 28 out of 84 events with E > 55 EeV were found to correlate with AGNs, which corresponds to a fraction of correlating events equal to (33±5)% [83]. For an isotropic distribution of sources, the rate of correlating events would be 21%. The chance probability of observing a 33% correlation from a random distribution is below 1%.

The angular resolution of the Telescope Array SD detector for events with E > 10 EeV is approximately 1.5°. In the full Telescope Array SD data set in the first 40 months of its operation, there are 11 correlating events out of 25 total (44%), while the expected number of random coincidences for this total number of events is 5.9 (23.6%). Such an excess has probability of $\sim 2\%$ to occur by chance given an isotropic distribution of arrival directions [84]. Both the Pierre Auger Observatory and the Telescope Array data are in good agreement yielding a combined chance probability of observing such a correlation at the 10^{-3} level [83]. The weakness of the anisotropy points to a scenario where AGNs are possible point-sources of UHECRs, with a large isotropic background.

1.4 The Multi-messenger information: photons and neutrinos

Even though the vast majority of cosmic ray primaries are of hadronic nature, essentially all models of UHECR production predict, to some extent, fluxes of UHE photons and neutrinos. The interactions of cosmic rays both within their sources and with background radiation fields during their propagation produce charged and neutral pions which decay to neutrinos and photons, respectively [85]. The interest in multi-messenger observations lies in their potential to do astronomy: UHE neutrinos and photons propagate along straight lines, undeflected by magnetic fields, and point to their production sites. They function as direct probes of their source locations and the mechanisms of acceleration.

1.4.1 UHE Photons

The first interaction of photons primaries in the atmosphere is dominated by electron-positron pair production, and EAS induced by photons are almost completely electromagnetic. Photon primaries might in principle initiate hadronic EAS. However, photon hadro-production cross-sections are roughly 3 orders of magnitude smaller than pair production cross-sections ($\sigma_{\gamma \to hadrons}/\sigma_{\gamma \to e^+e^-} \sim 2.8 \cdot 10^{-3}$) [86]. During propagation only a small fraction of the primary energy in photon showers is generally transferred into secondary hadrons and muons.

There are various signatures that characterize photon-induced EAS. Due to the reduced number of secondary particles per interaction, photon showers develop more slowly, and hence have, on average, a larger value of X_{max}. Their slower development results in larger curvature

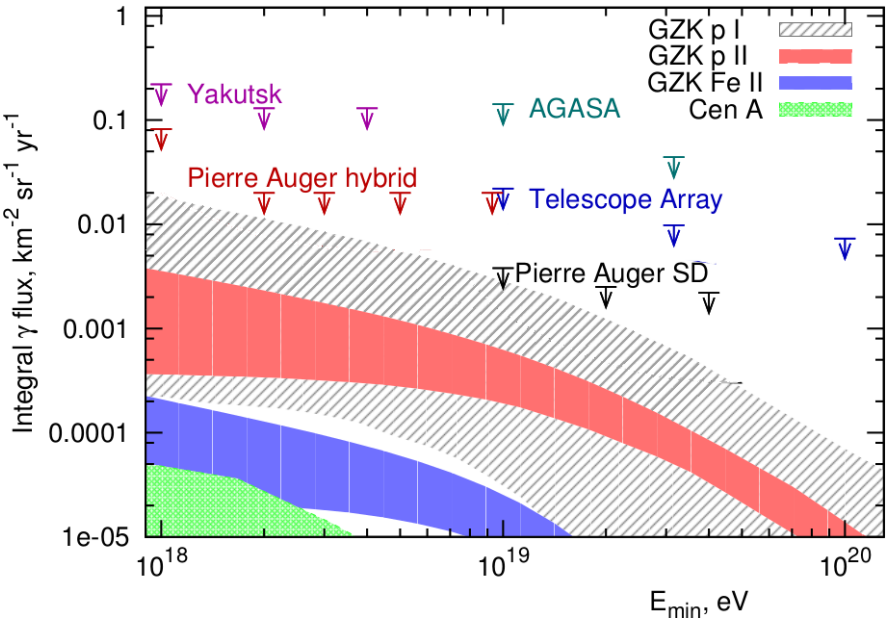


Figure 1.8: Limits to the UHE photon flux derived by the Pierre Auger Observatory, AGASA, Yakutsk and Telescope Array [85]. The predicted GZK photon flux by various models and a Cen A source model are shown in the shaded region [87, 88, 89].

and width of the shower front [49]. They are also characterized by a smaller number of secondary muons, and by a more compact signal distribution on ground [4].

The results of independent experiments are similar: no photon-induced EAS have been detected, and the number of candidates found is consistent with the expectation from nuclei-induced EAS. Current upper limits on the photon flux derived by the Pierre Auger Observatory, AGASA, Yakutsk and Telescope Array [85] are shown in figure 1.8. Upper bounds to the photon fraction derived by the Pierre Auger Observatory correspond to 0.4%, 0.5%, 1.0%, 2.6% and 8.9% for energies above 1, 2, 3, 5 and 10 EeV, respectively [4].

1.4.2 UHE Neutrinos

The extremely small interaction cross-section of neutrinos allow them to cross large amounts of matter without interacting. This is the main idea behind neutrino detection in air showers, ice and water. If neutrinos interact in the atmosphere at all, they would interact very deep in inclined showers[2], only after having traversed a huge amount of matter [49]. This way, they generate showers in an early stage of their development (young showers) close to the ground, whereas nuclei-initiated inclined deep showers are almost completely absorbed by the time they reach ground. Upward-going[3] tau neutrinos can interact within the earth crust, generating a tau lepton that will in turn initiate an EAS.

Detection in ice and water volumes focus on the interactions of neutrinos in dense media. These interactions generate charged particles traveling faster than light in those media, consequently emitting Cherenkov radiation. IceCube [96] and ANTARES [97] use ice and water as targets, respectively, due to their abundance, large density, and transparency to visible wavelengths. Recently, IceCube reported the detection of the two first PeV neutrinos, with energies 1.04 ± 0.16 and 1.14 ± 0.17 PeV [98].

Radio detection of neutrinos is possible as well. Particles traveling faster than light in a dense dielectric produce a shower of charged secondaries which contains a charge anisotropy and emits

[2]Inclined Extensive Air Showers have zenith angles in the range $62^o - 80^o$.
[3]Upward-going Extensive Air Showers have zenith angles in the range $90^o - 95^o$.

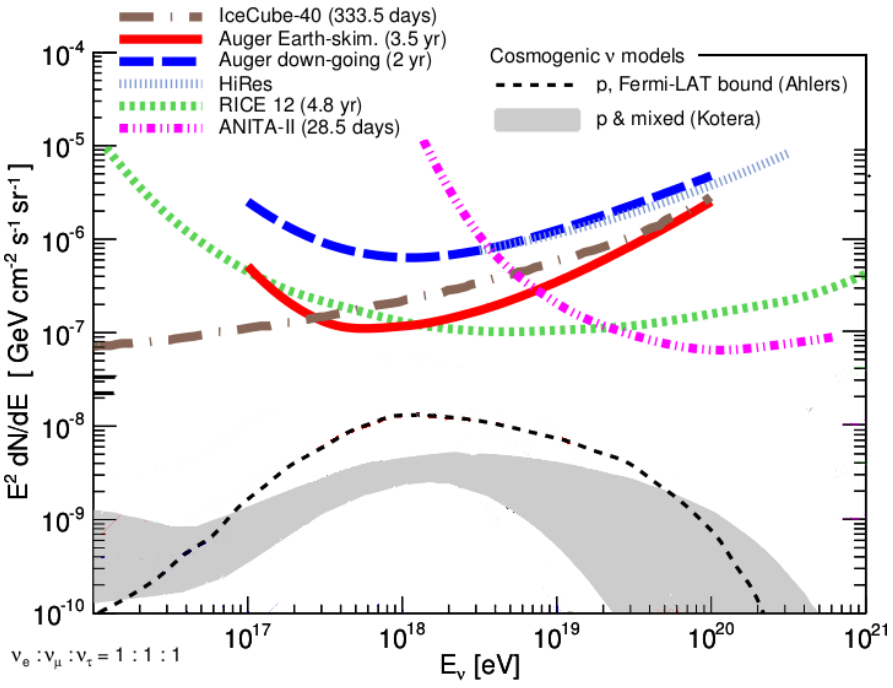


Figure 1.9: Differential flux limits derived by IceCube [90], the Pierre Auger Observatory (for Earth-skimming [5] and down-going [6] neutrinos), the HiRes experiment [91], RICE [92] and ANITA [93]. The gray band represents the flux expected from a set of models with pure proton and mixed compositions at the sources, and different assumptions on the evolution of the sources as well as on the transition from Galactic to extragalactic sources [94]. The dashed line is a cosmogenic model constrained by Fermi-LAT observations of the GeV-TeV diffuse γ-ray background [95].

a cone of coherent radiation in the MHz-GHz range. This phenomenon is called the Askar'yan effect [99]. Experiments such as ANITA [100], RICE [101], ARA [102] or ARIANNA[103] exploit this technique to detect neutrinos in ice. LUNASKA tries to detect radio pulses produced in neutrino-induced cascades in the moon [104].

So far, no UHE neutrinos have been reported, and upper limits to the flux of UHE neutrinos have been set instead. The differential flux limits set by IceCube [90], the Pierre Auger Observatory (for Earth-skimming [5] and down-going [6] neutrinos), the HiRes experiment [91], RICE [92] and ANITA [93] can be seen in figure 1.9.

2

Extensive Air Showers

When cosmic rays arrive to the Earth they interact inelastically with nuclei in the atmosphere, producing a plethora of secondary particles, predominantly electrons, positrons, photons and pions. These also interact with air nuclei, generating new particles. The cascade of particles produced is called an Extensive Air Shower (EAS).

EAS develop in a complex way. Photons, electrons and positrons produced in the first interaction initiate an electromagnetic (EM) shower. EM particles are the most abundant component of the shower, and carry the largest fraction of the shower energy. They develop mainly by 2 processes. In bremsstrahlung, the emitted photon carries, on average, half the electron energy. Pair production generates particles sharing equally the parent's energy at each generation. Particle multiplication and ionization energy losses are competing processes. When particles reach a critical energy, ionization losses start to dominate, and the shower is gradually absorbed.

Most hadrons produced during the first interaction are pions. Neutral pions decay immediately into electromagnetic particles, initiating sub-EM showers. Charged secondaries may interact, generating new secondaries and feeding a hadronic shower, or decay into muons. After a few hadronic interactions most of the energy of the hadronic component is transferred to the electromagnetic part of the shower. Hadronic showers develop a significant muonic component, whereas there are fewer muons in EM showers.

The basic properties of the development of the cascade can be extracted from a simple model due to Heitler, describing the evolution of purely EM cascades [105]. This model was extended by Matthews to describe hadronic showers [106].

2.1 Heitler's model for EM showers

Cascade development and its most important features are easy to understand in the toy model suggested by Heitler [105]. Heitler described the evolution of the EM shower as a perfect binary tree, in which electrons, positrons and photons always interact after traveling an interaction or splitting length, d. This length is given by the radiation length of the medium, $d = 2 \ln \lambda_r$, with $\lambda_r = 37$ g·cm^{-2} in air. At each step, electrons and positrons radiate a photon via bremsstrahlung, and photons split into a e^+e^- pair. As such, two particles arise after the interaction, each carrying one half of the primary energy by construction. The process of particle multiplication continues until the energy of the particles falls below a critical energy, E_c^γ, about 80-85 MeV in air.

After one interaction length d the cascade consists of two particles with half the primary energy, $E_1 = E_0/2$. After $2d$, there are $N_2 = 2^2$ particles with an energy $E_2 = E_0/2^2$ each. After n interaction lengths, the particle number is $N_n = 2^n$, the energy of the individual particles is $E_i = E_0/N_n$ and the depth reached in the atmosphere is $X = nd$. A sketch of the development of the EM shower according to this model is depicted in figure 2.1 (left). This picture does not

capture all the details of EM showers, but is able to account for three important features:

The number of particles at the shower maximum is proportional to the energy of the primary particle. The cascade reaches its maximum size, N_{max}, when all particles fall below the critical energy E_c^γ and the particle production process stops.

$$\begin{aligned} E_0 &= E_c^\gamma \cdot N_{max} \\ N_{max} &= E_0/E_c^\gamma \end{aligned} \tag{2.1}$$

The depth of maximum shower development, X_{max}^γ, is logarithmically proportional to E_0. The depth at which the EM shower reaches its maximum depends on the number of interactions needed for the energy of each particle to be reduced to E_c^γ, n_{max}:

$$\begin{aligned} N_{max} &= 2^{n_{max}} \\ n_{max} &= \ln(E_0/E_c^\gamma)/\ln 2 \\ X_{max} &= X_0 + \lambda_r \ln 2 \cdot n_{max} = X_0 + \lambda_r \cdot \ln(E_0/E_c^\gamma) \end{aligned} \tag{2.2}$$

where X_0 is the depth of the first interaction.

The elongation rate, defined as the evolution of X_{max} with energy

$$D_{10} = dX_{max}/d\log_{10} E_0 = 2.3\lambda_r \simeq 85 \text{ g cm}^{-2} \tag{2.3}$$

is proportional to the radiation length.

The simulations of EM cascades, where a detailed description of their development is included, confirm these properties. However, the predicted number of particles at the shower maximum is overestimated by a factor 2-3 and the ratio of electrons to photons is overestimated by a factor 10-12. These discrepancies appear because the absorption of particles above the critical energy is not accounted for, multiple photons are often radiated during bremsstrahlung, and electrons lose energy much faster than photons do [106].

2.2 Extension to hadronic showers

The model developed by Heitler can be adapted to describe the showers initiated by hadrons [106]. The relevant parameters are now the hadronic interaction length, λ_I, and the pion critical

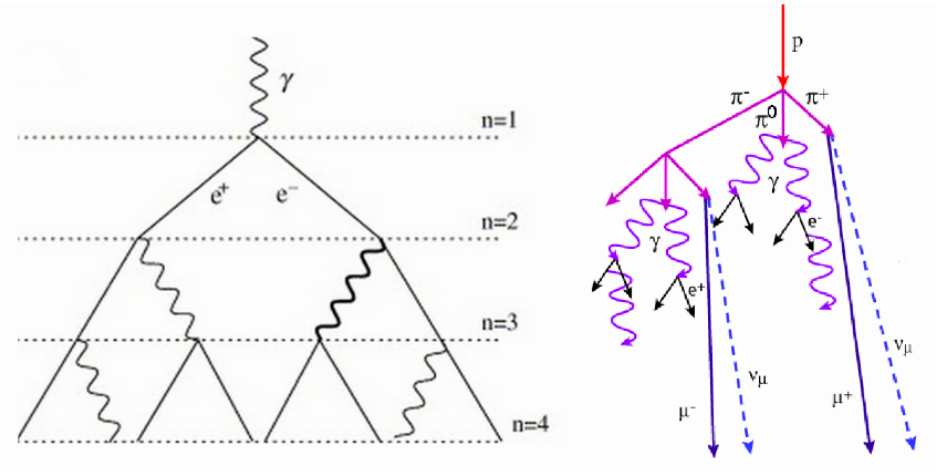


Figure 2.1: Schematic representation of the development of electromagnetic (left) and hadronic (right) showers according to the Heitler's and Matthews' models, respectively.

energy, E_c^π. λ_I is not constant, but it does not depend strongly on energy. Between 10 and 1000 GeV, typical energies of pions in EAS, the inelastic πp cross-section is approximately constant and equal to 20 mb. Scaling this value to collisions with air, a value of $\lambda_I \sim 120$ g·cm^{-2} is a good approximation [86, 107]. The critical energy is the energy at which the decay length of a charged pion becomes smaller than the distance to the next interaction point. It decreases very slowly with increasing primary energy, taking values of 30, 20 and 10 GeV at primary energies of 10^{14}, 10^{15} and 10^{17} eV, respectively. A constant value $E_c^\pi = 20$ GeV for pions in air is a good approximation [106]. When the energy of individual charged pions falls below E_c^π they are assumed to decay, producing muons [106, 107].

After each step of thickness $d = \lambda_I \ln 2$ the hadron interacts, producing N_{ch} charged pions and $\frac{1}{2}N_{ch}$ neutral ones. The multiplicity in πN interactions increases very slowly with laboratory energy. For example, in πN^{14} collisions N_{ch} is approximately 5, 11 and 27 at 10, 10^3 and 10^4 GeV, respectively. A constant value of $N_{ch} = 10$ is usually adopted for energies around 10 GeV [108]. Neutral pions decay to electromagnetic particles on the spot, initiating EM showers. Charged pions interact, producing a new generation of charged and neutral pions. After n interactions, the total number of charged pions is $N_{\pi^\pm} = (N_{ch})^n$. The total energy carried by these pions is $(2/3)^n E_0$, assuming that energy is shared evenly between charged and neutral pions during particle production. Then, the energy per charged pion in the n^{th} interaction layer is $E_\pi = E_0 \cdot (2/(3N_{ch}))^n$. The schematic development of a hadron-initiated shower can be seen in figure 2.1 (right). After a certain number n_c of generations, E_π falls below E_c^π. The number of interactions needed to reach E_c^π is calculated as:

$$\begin{aligned} E_\pi &= \frac{E_0}{\left(\frac{3}{2}N_{ch}\right)^n} \\ E_c^\pi &= \frac{E_0}{\left(\frac{3}{2}N_{ch}\right)^{n_c}} \\ n_c &= \frac{\ln(E_0/E_c^\pi)}{\ln(3N_{ch}/2)} \end{aligned} \tag{2.4}$$

To obtain the number of muons in the shower one assumes that all charged pions decay to muons when they fall below their critical energy:

$$N_\mu = (N_{ch})^{n_c} \tag{2.5}$$

Inseting equation 2.4 into equation 2.5 we obtain:

$$\ln N_\mu = n_c \ln N_{ch} = \ln\left(\frac{E_0}{E_\pi^c}\right) \cdot \frac{\ln N_{ch}}{\ln(3N_{ch}/2)} \tag{2.6}$$

$$= \beta \ln\left(\frac{E_0}{E_\pi^c}\right) \tag{2.7}$$

$$N_\mu = \left(\frac{E_0}{E_\pi^c}\right)^\beta \tag{2.8}$$

It is worth noticing that the number of muons does not grow linearly with energy as the number of electrons does. Moreover, the value of β depends on the average pion multiplicity chosen and the inelasticity of hadronic interactions. Changing any of N_{ch} or the amount of energy transferred to pions would modify the precise value of β.

The definition of X_{max} for showers initiated by hadrons is the same: the depth at which the electrons and photons of the air shower reach their maximum number. The electromagnetic component is generated by photons from the decays of neutral pions. The first interaction diverts $1/3E_0$ into these channels, followed by additional showers from each subsequent interaction point [106]. A simple estimate of the hadronic showers X_{max}, in which only the first generation of electromagnetic showers is used, describes remarkably well the elongation rate, though not the absolute value of X_{max}, agreeing to a high degree with simulated showers [107]. Figure 2.2 shows the elongation rates arising from this model, shifted 100 g·cm^{-2} upwards, compared to

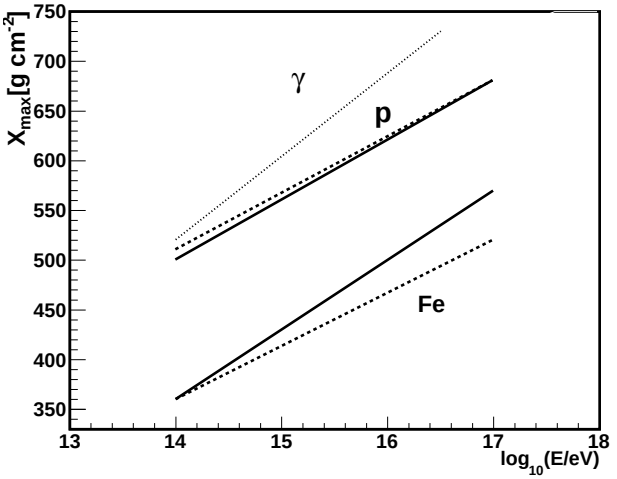


Figure 2.2: X_{max} as a function of the primary energy for photon initiated air showers (dotted line) and proton and iron initiated showers (dashed and solid lines). Dashed lines correspond to the theoretical model uniformly shifted 100 g·cm^{-2} to higher values. Solid lines are from full simulations of p and Fe showers [106].

those obtained from simulations. Following [106], the elongation rate for proton showers can be written as:

$$D_{10}^{p} = D_{10}^{\gamma} + \frac{dX_0}{d\log_{10} E_0} \tag{2.9}$$

where D_{10}^{γ} is the elongation rate of electromagnetic showers, and $X_0 = \lambda_I \ln 2$ is the depth of the first interaction. Using the parameterization from λ_I in [109], the elongation rate yields $D_{10}^{\gamma} = 64$ g·cm^{-2}. Whatever the parameterization used, interaction lengths decrease with rising energy. As such, the elongation rate for electromagnetic showers becomes an upper limit to the elongation rate for hadronic showers. This bound was formulated for the first time by Linsley as the Elongation Rate theorem [110].

2.3 Nuclear primaries

The application of this description to nuclear primaries relies on the superposition model, in which a nucleus with total energy E_0 and nuclear number A is assumed to behave as A nucleons, with energy E_0/A each. Showers are treated as the superposition of A independent showers, all starting at the same point. In this picture, the nucleus is treated as if A nucleons entered the atmosphere at the same time, and not as if the nucleon, upon its first interaction, splitted into A nucleons. The shower observables can be computed substituting the lower primary energy of individual nucleons into the various expressions derived previously for proton showers and summing A such showers where appropriate:

1. as the showers initiated by lower energy primaries have smaller values of X_{max}, nuclei initiated showers will be less deep than proton showers of the same energy:

$$X_{max}^{A}(E_0) = X_{max}^{p}(E_0/A) = X_{max}^{p}(E_0) - \lambda_r A \tag{2.10}$$

2. the number of muons is larger for nuclear primaries than for proton primaries of the same primary energy:

$$N_{\mu}^{A}(E_0) = \sum_{i}^{A} N_{\mu}^{p}(E_0/A) = N_{\mu}^{p}(E_0) A^{1-\beta} \tag{2.11}$$

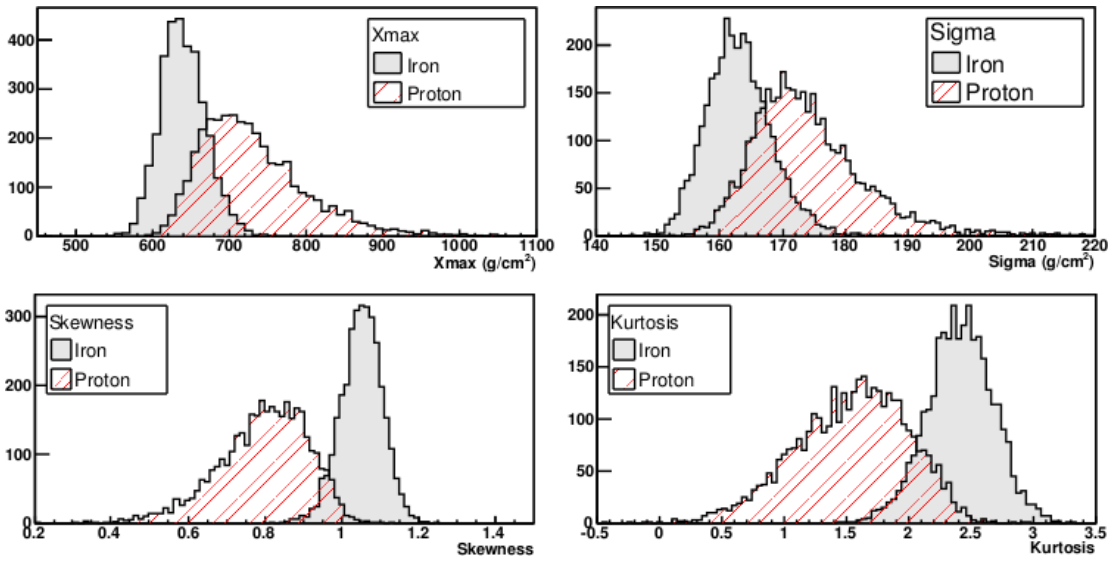


Figure 2.3: Distribution of X_{max}, RMS of X_{max}, skewness and kurtosis for proton- and iron-iniatiated showers with primary energy $E_0 = 10^{18} eV$ [111].

3. the elongation rate is equal for proton- and nuclei- initiated showers:

$$D_{10}^{A} = \frac{dX_{max}^{A}}{d\log_{10}} = \frac{d(X_{max}^{p} - \lambda_r A)}{d\log_{10}} = \frac{dX_{max}^{p}}{d\log_{10}} = D_{10}^{p} \tag{2.12}$$

In a X_{max} vs E_0 plot, different nuclei will appear as parallel lines, with values smaller for larger A nuclei, but with identical slopes (dashed lines in figure 2.2).

Another feature of nuclei-initiated showers is that, as the sum of many equivalent sub-showers, any variable is expected to fluctuate less for nuclei than for protons. In figure 2.3 we depict four different features of the shower profile, namely the value of X_{max}, the RMS of the X_{max}, the kurtosis and the skewness (moments of order 3 and 4 of the distribution, respectively) for proton- and iron-initiated showers of 10^{18} eV [111]. The distributions are always narrower, i.e. show less fluctuations, for showers induced by iron nuclei, statistically stabilized around the expected value for each of its sub-showers, than for showers initiated by protons.

2.4 Detection techniques

At the lowest energies the fluxes of different cosmic rays nuclei are high enough that can be measured in balloon- or satellite-borne detectors. Above 10^{14} eV the flux is too low, and on-orbit detectors, with small detection areas, have little chance of detecting any particle [112]. The Earth's atmosphere converts cosmic rays into an EAS and beyond these energies, cosmic rays are detected on ground using installations with large collection areas.

To unmistakably identify single primaries they would have to be detected before they collided with atmospheric nuclei and initiated an EAS. Once the development of the EAS starts, any information about the primary particle properties (energy, mass and arrival direction) cannot be measured directly. Rather, one has to infer these properties indirectly from the measurements of the EAS. EAS are an imprecise indicator of the primary particle properties because of the numerous interactions that occur between its entry at the top of the atmosphere and the detection at ground level of the shower induced. There are various complications inherent to the indirect detection of EAS [113]:

- The atmosphere becomes part of the detection system, acting as a calorimeter with variable properties.

- Shower measurements are always incomplete, due to the sampling imposed by the detector discreteness and the impossibility to observe the whole shower development.
- Moreover, shower development is a stochastic process: particle densities fluctuate around their expected values. Even a perfectly measured shower would not uniquely determine the nature of the primary particle.
- Expected shower properties have to be inferred from the extrapolation of theoretical models at lower energies.

There are two general classes of air shower detectors: those that sample the flux of secondary particles at ground level (surface detectors) and those that record radiation from the shower front as it traverses the atmosphere (atmospheric radiation detectors). Among surface arrays one finds muon detector arrays (Yakutsk [114]), scintillator arrays (Volcano Ranch, AGASA and the Telescope Array surface detector [115, 116, 117]), and Water-Cherenkov tank arrays (Haverah Park, the Pierre Auger Observatory surface detector and HAWC [30, 118, 119]). Atmospheric detectors measure the longitudinal development of the shower gathering the radiation induced by shower particles at differente frequencies. Within this category there are nitrogen fluorescence detectors (Fly's Eye, HiRes, and the Pierre Auger Observatory and the Telescope Array fluorescence detectors [33, 34, 120, 121]), air Cherenkov detectors (Tunka and CTA [122, 123]), and radio and microwave antenna arrays (LOPES and MIDAS [124, 125]).

2.4.1 Surface arrays

EAS generate a large number of particles, which span a very large area when they reach ground. Surface detector arrays sample the flux of secondary particles at a given observation level, observing only a single slice of the EAS longitudinal development. This means that the determination of primary particle energy and composition require model-dependent extrapolations, which may disagree depending on the model used.

The area of the detector array depends on the cosmic ray flux in the energy region of interest and the rate of events needed. For studies above 10^{19} eV, the expected rate of events is less than 1 per km^2 and year. To record a significant amount of data, areas of thousands of km^2 are required. The distance between individual detectors is optimized to match the area spanned by the EAS (footprint) at the observation level. Showers initiated by low energy primaries have smaller footprints and for a given detector spacing they are detected with less probability than high-energy showers. For example, a triangular grid of detectors spaced 666, 1332 and 1880 m becomes fully efficient at approximately 0.5, 4.5 and 20 EeV, respectively [126]. The shower axis, and hence the direction of the primary particle, is deduced from the relative arrival times of signals at a minimum of three non-collinear detectors. The total shower energy is approximately proportional to the signal in the detectors at a certain distance far from the core. The precise distance depends on the energy range and the detector spacing, and is chosen so that shower to shower fluctuations are minimized [127].

Scintillators and Water-Cherenkov tanks are the most common particle detectors employed on surface arrays. Scintillation detectors are equally sensitive to all charged particles. Their deployment is simple and their use is straightforward but usually restricted to angles below 45°: being flat their effective area falls sharply with rising zenith angle, and so does their aperture. Besides, a large part of the electromagnetic component is absorbed in the atmosphere and does not reach the detector at zenith angles above 45°. If scintillators are installed underground they can work as muon counters, with different threshold energies depending on their depth. Examples of experiments using scintillator arrays are The Volcano Ranch array [115], Yakutsk [114], AGASA [116] and Telescope Array [117].

Water-Cherenkov tanks are sensitive to both the electromagnetic and the muonic components of the shower. Unlike scintillators they are not flat, and at large zenith angles they present a sizeable effective surface. This extends their aperture to nearly horizontal EAS. Experiments that employ Water-Cherenkov tanks include Haverah Park [30], the Pierre Auger Observatory [118] and HAWC [119].

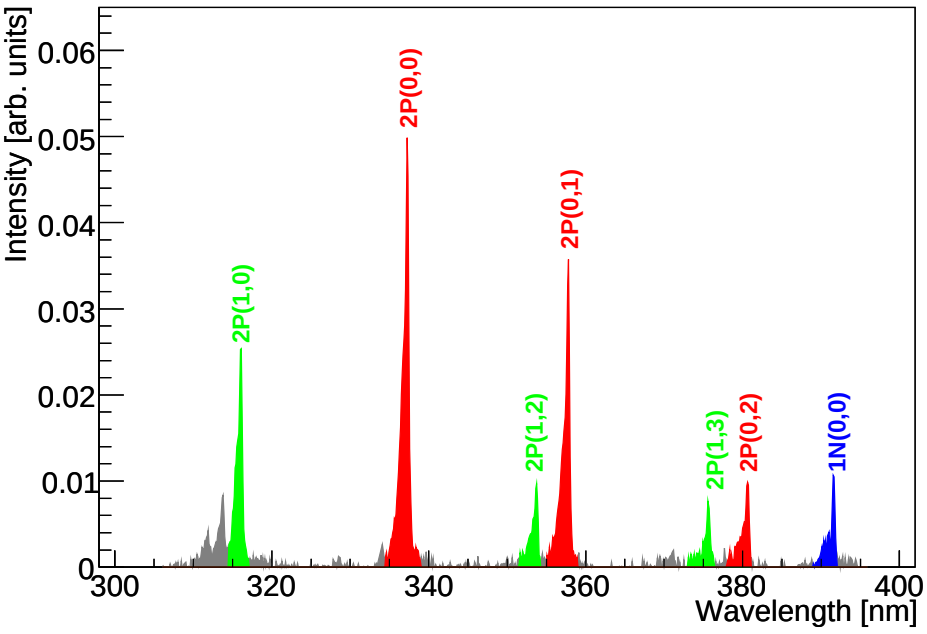


Figure 2.4: Nitrogen fluorescence spectrum between 300 nm and 400 nm in dry air at 1013 hPa [128]

2.4.2 Atmospheric radiation arrays

Atmospheric radiation detectors measure the longitudinal development of the shower in the atmosphere. An air shower almost entirely dissipates its energy through ionization and atmospheric radiation detectors record the emitted radiation. Nitrogen fluorescence emission, even though the most popular technique, is not the only detectable and exploitable radiation. From the low frequency radio emission to UV fluorescence emission, several techniques aim to detect the radiation emitted by EAS.

Fluorescence light

The charged secondary particles of the EAS, mainly electrons and positrons, deposit their energy in the atmosphere by exciting air molecules, which may afterwards decay to a lower energy state emitting fluorescence photons isotropically. Most of these emissions, in several broad spectral bands, come from different states of excited molecular nitrogen. The molecular spectrum is splitted in bands due to the vibrational and rotational movements of the molecular nuclei, which modify the energy states of the electrons. In particular, transitions from the 2P (second positive) band system of N_2 and the 1N (first negative) band system of N_2^+, generate photons in the wavelength range 300-400 nm. The strongest bands are located at 337.1 nm and 357.7 nm, from the 2P transition, and at 391.4 nm from the 1N transition [128]. The fluorescence spectrum of molecular nitrogen and the maxima corresponding to different transitions is shown in figure 2.4.

In general, electrons release their energy undergoing different processes, and only a small fraction of the energy is finally converted into fluorescence photons. The average yield is of 4 photons per electron per meter, with a soft dependence on atmospheric pressure and temperature, rising when any of them decreases [128, 129].

Most of the attenuation suffered by fluorescence light is due to molecular and aerosol scattering. The molecular scattering length for the fluorescence wavelength is ~ 14 km, and aerosol scattering is significant in the first 2 km above ground, but then falls exponentially with increasing altitude. This way, fluorescence light suffers almost no attenuation for distances up to 20 km

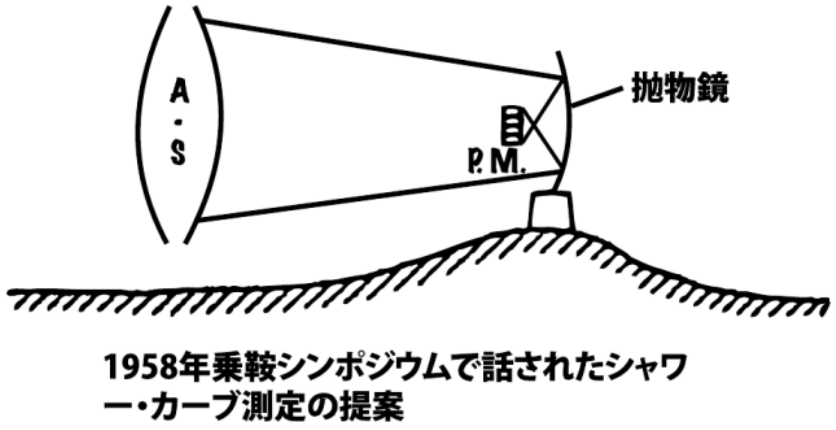


Figure 2.5: Reproduction from Proceedings of Norikura Meeting in Summer 1957. The text translates as *"parabolic mirror"* and *"A proposal for the shower curve measurement in Norikura symposium, 1958"*. Image from [132].

and moderate zenith angles and a large fraction of the fluorescence light reaches the detector in a direct line from the source [113].

These photons can be collected by a light collector system (such as reflector mirrors) and recorded with UV sensitive detectors (like photomultiplier tubes). The UV photons emitted as the shower develops define a moving track through the atmosphere, from which one can reconstruct the longitudinal shower profile, and hence the position of the shower maximum and a calorimetric estimate of the shower energy [130]. Conceptually, the energy determination by a fluorescence detector is straightforward, the amount of emitted fluorescence light being proportional to the energy losses of the charged particles. Thus, measuring the fluorescence emission from the whole shower should yield the total electromagnetic shower energy [131]. However, this simple calorimetric method encounters some difficulties and some corrections have to be made. Fluorescence light suffers scattering while it propagates through the atmosphere, the Cherenkov light emitted by the particles of the shower can reach the detector, either directly or scattered, and the optical properties of the atmosphere are not constant. The atmosphere possesses a variable density and composition, which makes its careful monitoring necessary.

It is not clear who first had the inspiration of using the excitation of atmospheric nitrogen for cosmic ray studies. The first discussions regarding the use of air fluorescence to detect another radiation date back to the late 1940s, during the Manhattan project nuclear bomb tests. At this time, the focus was on using air-fluorescence induced by X-rays from nuclear explosions as a monitoring tool. During the decade of 1950 in Japan a great deal of work was devoted to the investigation of the features of fluorescence light induced by high-energy cosmic rays, and the techniques most suited to its exploitation. Figure 2.5 shows a sketch of the concept of a PMT camera viewing the fluorescence light from an air shower collected with a mirror. It is worth noticing the similarity of this layout compared to the instruments used by HiRes, the Pierre Auger Observatory and Telescope Array. This work led to the first discussion at an international conference of the detection of high-energy cosmic rays with this method. It was presented at the *Fifth Interamerican Seminar on Cosmic Rays*, celebrated in La Paz in 1962 [133]. During the mid-1960s efforts were being made at the University of Cornell to detect fluorescence radiation using Fresnel lenses. However, the first detection of air showers by the fluorescence method was due to a detector installed at Mt. Dodaira, Japan, in 1969 [134]. Following these efforts, a small fluorescence detector was operated in coincidence with the Volcano Ranch scintillation array, showing convinging demonstrations of the method [135]. The first independent fluorescence array was Fly's Eye. The original Fly's Eye (FE I) was completed in 1981 and consisted of sixty seven 1.5 m diameter mirrors covering the entire sky. To improve the shower reconstruction in the absence of a ground array a second eye (FE II) was added in 1986 with 36 mirrors at a

distance of 3.4 km from FE I, pioneering the stereoscopic observations of EAS [132].

Normally, it is the number of electromagnetic particles as a function of atmospheric depth that is measured, parameterized by the 4-parameter Gaisser-Hillas function [136]:

$$N_e(X) = N_{max}\left(\frac{X - X_0}{X_{max} - X_0}\right)^{(X_{max}-X_0)/\lambda} e^{(X_{max}-X_0)/\lambda} \tag{2.13}$$

The integral of this function yields the total number of electromagnetic particles in the shower, which scaled by the average energy loss per particle, 2.2 MeV/g·cm^{-2}, gives a calorimetric calculation of the shower energy. This energy is corrected for the so called *missing energy*: the fraction of the primary energy assumed to be transferred to the hadronic cascade, neutrinos and penetrating muons, and not converted into fluorescence light. This quantity is model-dependent, but it is estimated to be around 15% (10%) for 10^{18} eV (10^{20} eV) iron-induced EAS, and 10% (7%) for proton-induced EAS of the same energy [137].

Cherenkov light

During the shower development charged particles emit Cherenkov radiation [138]. The Cherenkov radiation intensity is proportional to the primary energy, while the slope of the lateral distribution is related to the depth of the maximum shower development. The Cherenkov radiation is strongly beamed along the shower axis and measurements have to be made very close to the core of the shower, which demands very small spacings between detectors. This restriction, in addition to its low duty cycle, limits the use of this technique to detect EAS beyond 10^{17} eV, energy above which large detector areas are needed due to the low flux of cosmic rays.

The HEGRA (**H**igh-**E**nergy **G**amma-**R**ay **A**stronomy) collaboration pioneered the field of TeV gamma ray astronomy applying the stereoscopic observation mode with Cherenkov telescopes [139]. HEGRA took data between 1987 and 2002, when it was dismantled to build MAGIC (**M**ajor **A**tmospheric **G**amma-ray **I**maging **C**herenkov Telescopes), its sucessor. One of the most important achievements of the instrument was the detection of the most energetic photons observed (up to 16 TeV) from an extragalactic object, namely from the Blazar Markarian 501 [140].

MAGIC consists of two 17 m diameter Imaging Atmospheric Cherenkov Telescopes (IACTs). The first telescope started operating in standalone mode in 2004, and the second one was installed in 2009. It explores the very high energy sky, from some tens of GeV up to tens of TeV [141]. Some recent results from MAGIC are the precise determination of the spectrum of the Crab Nebula and the discovery of the gamma ray binary LS I +61° 303 [141].

HESS (**H**igh **E**nergy **S**tereoscopic **S**ystem) is a system of four IACTs designed for the investigation of gamma rays in the sub TeV and TeV energy range. The HESS telescope system was extended by a much larger telescope, HESS II, in 2012. The aim of this enhancement is closing the gap between ground based and space based gamma ray detection. HESS has a dedicated program of observation of the Galactic plane. Through the detection of VHE gamma rays it is able to identify a rich population of SNRs, pulsar wind nebulae and binary systems, among others [142].

The CTA (**C**herenkov **T**elescope **A**rray) instrument, still under development, is designed to detect the Cherenkov radiation emitted by very high energy ($\geq$ 10 GeV) gamma ray-induced electromagnetic showers [123]. However, it is expected to detect EAS induced by charged cosmic rays in the TeV-PeV regime as well [143].

Radio detection

The idea of radio detection from EAS was proposed for the first time by Askar'yan in 1962. According to [99], the electromagnetic component of the shower would present a negative charge excess, giving rise to Cherenkov radiation at radio frequencies whose intensity scales linearly with the shower energy. The first EAS radio detectors date back to 1965 [144].

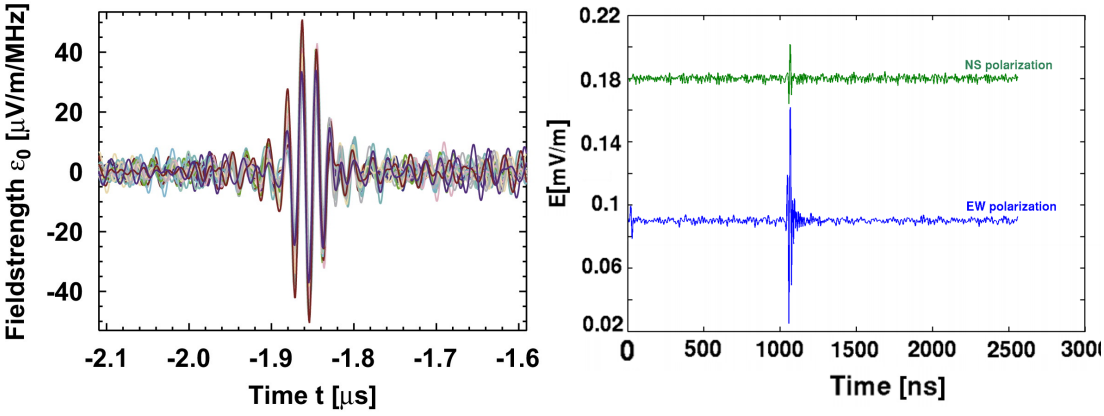


Figure 2.6: Left: Cosmic ray radio pulses seen in individual LOPES antennas [151]. Right: Wave shape for both EW and NS polarizations of the electric induced by an EAS in the 30-80 MHz band as seen by the CODALEMA antennas [152].

Very Low Frequency (VLF) radiation, generally below 0.5-1 MHz, is currently explained by a mechanism called Transition Radiation [145]. Transition Radiation is emitted when a uniformly charged object crosses the boundary between two media having different dielectric properties. There seems to be no unanimity regarding the time duration, generation mechanism, or intensity of these pulses. Radio detection in the very low frequency range has been described in [146] but there have been few claims of radio detection below 500 kHz [147], and no detector is currently working in the detection of this frequency band.

Radio signals in the Very High Frequency (VHF) band (10-100 MHz) have, in contrast, a firm theoretical background, and there are some experimental installations dedicated to their study. The dominant contribution in this band is explained by geosynchrotron emission. Electrons and positrons propagating through the geomagnetic field are deflected in different directions. This effect creates an electric dipole that propagates at the speed of light and emits in radio frequencies through synchrotron radiation [148]. The LOPES collaboration [149], instrumented to detect EAS in the range 40-80 MHz, and the CODALEMA experiment [150], able to detect EAS in the range 1-200 MHz, have detected air showers with energies up to 10^{18} eV [151, 152]. Cosmic ray-induced radio pulses detected by LOPES and CODALEMA are shown in figure 2.6 (left) and figure 2.6 (right), respectively. AERA [153] and EASIER [154] are the radio extensions of the Pierre Auger Observatory, though EASIER is designed to detect microwave radiation induced by EAS as well. Both instruments work in a very similar frequency range. AERA is instrumented with antennas sensitive to radio emission in the frequency range from 30 MHz to 80 MHz, while EASIER works in the 30-70 MHz band. An example of a radio signal from a self-triggered cosmic ray event detected by AERA is shown in figure 2.7 [155]. Radio detection of showers presents two principal advantages. Detectors are built using wire antennas, making them cheap and easy to deploy. In addition, the absorption at radio frequencies in the atmosphere is negligible, and the radio signal travels essentially unaltered from its source to the detector. An array of individual antennas observing the wavefront at different positions with respect to the shower axis allows to reconstruct the properties of the air shower [156]. For instance, the curvature of the electromagnetic radio front, measured through the arrival times of the radio pulses at individual antennas is related to the depth of the shower maximum [157]. Information of the air shower evolution is encoded in the pulse shapes of the radio signals. The technique remains reliable above 40 MHz under all circumstances except for extreme events like thunderstorms [124]. At frequencies between 100 and 200 MHz this emission is radiated into a wide angular range, which increases the acceptance for UHECRs [158].

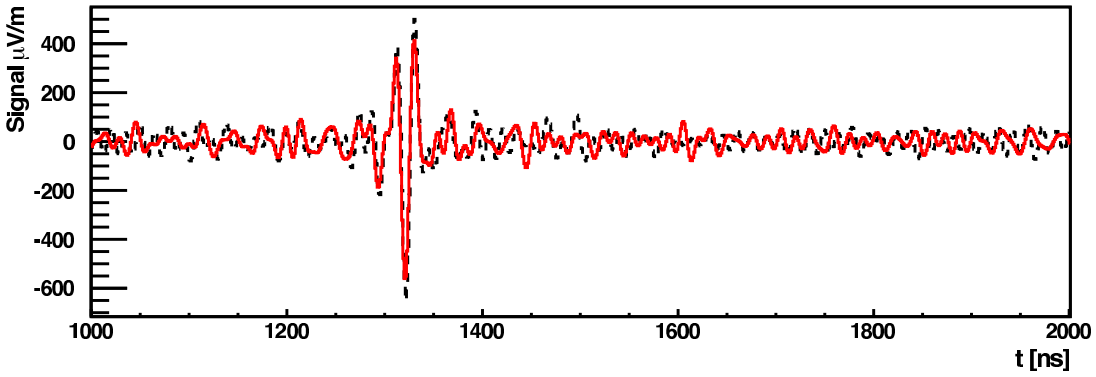


Figure 2.7: Calibrated radio pulse recorded for a 5.7 EeV cosmic ray event by AERA [155].

Microwaves

The first indication of emission at microwave frequencies ($\gtrsim$ 1 GHz) came from accelerator experiments [159]. Frequencies in the 1-10 GHz range are very well suited for shower observation, due to the low natural and man-made backgrounds and interferences.

The origin of signals above the GHz lies in molecular bremsstrahlung (MBR). In MBR, low-energy electrons (E $\leq 10^{10}$ eV) are accelerated in collisions with the fields created by molecules in the ambient medium [159]. Unlike radiation at lower frequencies, MBR radiation is expected to be isotropic and unpolarized allowing to perform shower calorimetry, measuring the MBR intensity along the shower development [160]. In this respect, it is useful to think of MBR emission as analogous to "radio fluorescence". Apart from continuous operation (not being subject to operation in dark conditions) microwave detectors also benefit from an extremely transparent atmosphere, with power attenuation less than 0.05 dB/km [161], which makes atmospheric monitoring unnecessary. Furthermore, MBR intensity is expected to be proportional to the EAS ionization rate, which is known to be itself proportional to the total number of charged particles in the shower [159]. This leads to a direct relationship between MBR intensity and shower energy.

Another possible source of microwave radiation is Cherenkov radiation, emitted in the MHz range, but compressed to GHz frequencies close to the Cherenkov angle [160]. AMBER and EASIER [162, 154] are prototypes installed at the Pierre Auger Observatory, that use the trigger from the Auger detectors to record the microwave emission. MIDAS is a self-triggering system which is foreseen to be installed at the Pierre Auger Observatory as well [125]. The CROME experiment is installed within the KASCADE-Grande array, and measures the GHz signal of EAS triggered by the KASCADE trigger system [163]. Recently, CROME has reported the detection of microwave signals for more than 30 showers with energies above $3{\times}10^{16}$ eV [164]. These constitute the first direct measurements of the basic features of microwave radio emission from EAS. Contrary to the expectations, the measurements are consistent with a mainly forward-beamed, coherent and polarised emission process. These findings, however, do not exclude a sub-leading signal component resulting from an isotropic emission process as expected for MBR. Figure 2.8 shows the time trace of microwave signals measured for an event with reconstructed energy $2.5{\cdot}10^{17}$ eV as a function of the time relative to the KASCADE-Grande trigger.

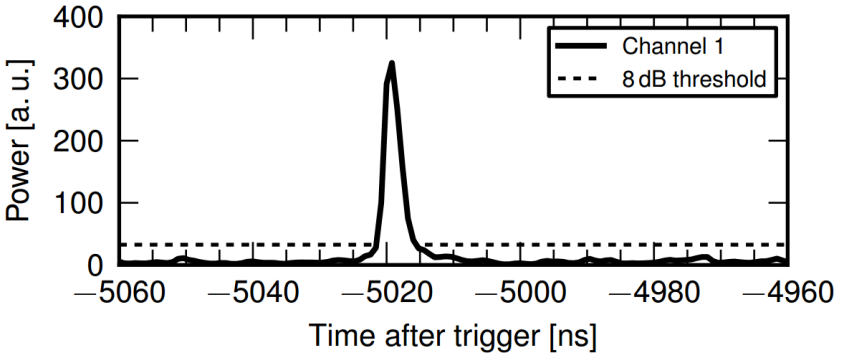


Figure 2.8: Power received by the CROME antenna amplifier as function of the time relative to the KASCADE-Grande trigger [164]. The dashed line represents the 8 dB pulse amplitude threshold used to select events.

3

The Pierre Auger Observatory

The Pierre Auger Observatory [1] is a hybrid air shower experiment that uses two independent, well-established techniques to detect and study high-energy cosmic rays: an array of Surface Water-Cherenkov Detectors (SD) combined with air Fluorescence Detectors (FD).

1660 tanks on a triangular grid, separated 1500 m and covering an area of roughly 3000 km^2 are overlooked by 27 fluorescence telescopes, deployed at four different sites. Together they constitute a powerful instrument for air shower detection. It is the world largest UHECR facility, designed to measure, with high statistical significance, the flux, arrival direction distribution and mass composition of cosmic rays from 0.1 EeV to the highest energies.

The observatory is located at the "Pampa Amarilla", close to the city of Malargüe in the province of Mendoza, Argentina (69°W, 35°S). It has been gathering data since 2004 and it was completed in 2008. The site is relatively flat and near the base of the Andes mountain range, at an altitude of 1400 m above sea level, corresponding to a vertical atmospheric depth of $\sim 880\,\mathrm{g{\cdot}cm^{-2}}$.

The air fluorescence telescopes sample the development of the electromagnetic shower recording the scintillation light emitted by atmospheric nitrogen after its interaction with the charged particles in the shower. The surface detector measures the particles densities at ground, using the Cherenkov light emitted when they propagate through the water inside the detectors. Its hybrid nature allows to make combined estimations of the EAS properties, to run consistency checks and to intercalibrate the detectors, overcoming the limitations of measurements made by any of the detectors alone.

In addition to the SD and the FD, other instruments have been installed, or are foreseen to be installed. These enhancements of the observatory intend to investigate EAS at lower energies, and to detect additional EAS signals. AERA is being used to study the geosynchrotron emission of radio waves from extensive air showers in a frequency range from 30 to 80 MHz [153]. AMBER, EASIER and MIDAS are being developed to measure the microwave emission between 3 and 15 GHz and prototypes are now operated at the Pierre Auger Observatory [165]. The AMIGA project aims at providing full efficiency detection of cosmic rays down to 10^{17} eV and a better mass discrimination through an infill of 61 water Cherenkov tanks separated 750 m with muon counters buried alongside [166]. The low-energy extension of the FD is HEAT, a set of three high elevation telescopes located close to one of the FD stations [167]

Since the atmosphere acts as a calorimeter with variable properties, a series of atmospheric monitoring stations are installed across the observatory. The Central Laser Facility (CLF) and the Extreme Laser Facility (XLF) produce vertical UV laser tracks, recorded by the FD stations and used to estimate the aerosol distribution in the atmosphere at different heights [168]. Four LIDAR stations detect clouds and aerosols analyzing the backscattered light from laser pulses [169]. Finally, infrared cameras installed on the top of each FD building map the cloud coverage over each site [170].

The configuration of the SD and FD appears in figure 3.1. The positions of the surface

detectors (section 3.1) are represented by dots, while labels in the boundaries of the array correspond to the fluorescence detectors (section 3.2).

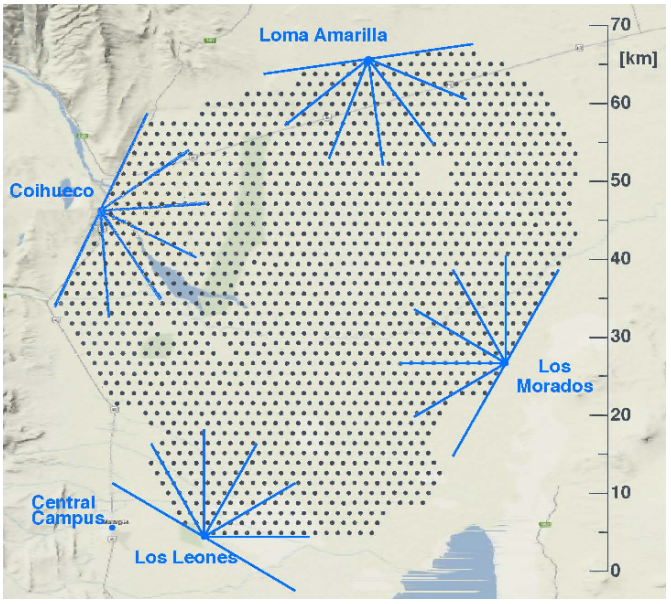


Figure 3.1: Layout of the Pierre Auger Observatory. Surface detector stations are represented as dots, and fluorescence detector stations are labeled on the boundaries of the array.

3.1 The Surface Detector (SD)

The secondary particles produced by a cosmic ray can be detected with stations deployed at the observation level. The Surface Detector (SD) of the Pierre Auger Observatory is composed of 1660 Water-Cherenkov tanks, laid on a triangular grid with 1500 m spacing between tanks. Water Cherenkov tanks were chosen for their robustness and low cost. They exhibit a rather uniform exposure up to large zenith angles and are sensitive to both charged particles and energetic photons. Water-Cherenkov tanks have been used with success in the Haverah Park array [30] and will be employed by the High Altitude Water-Cherenkov observatory (HAWC) [119].

Each detector consists of an opaque polyethylene tank, 1.55 m high and 3.6 m wide, surrounding a liner filled with 12000 l of ultra pure water. The tank structure, shown in figure 3.2, encloses a cylindrical volume of water 1.2 m deep and with a horizontal area of 10 m^2. The liner is an olefin polymer bag (Tyvek®), fulfilling several functions: it works as a seal for the water inside, protecting it from contamination and inhibiting bacteriological activities; it is an efficient reflector of Cherenkov light; and it works as a secondary barrier against external light sources [1]. The top of the tank houses three photomultiplier tubes (9" Photonis XP1805), symmetrically distributed at 1.2 m from the center of the tank, accessing the water volume through polyethylene windows on the top of the liner. Two signals are read from each PMT, an amplified (x32) signal from the last dynode, and a signal from the anode. The two signals provide enough dynamic range to cover with good precision total signals from the highest ($\sim$ 1000 particles μs^{-1} near the shower core) to the lowest ($\sim$ 1 particles μs^{-1} far from the shower core) particle fluxes. The readout of the six signals from each tank is accomplished using front-end electronics having 40 MHz Fast Analog to Digital Converters (FADCs) [1]. The signal recorded by the FADC is referred to in units of ADC counts (a measure of the current from the PMT). Each FADC bin corresponds to 25 ns. Digitized signals are sent to a PLD (programmable logic

device) board, which implements various trigger levels, and then sent to a central data acquisition system (CDAS). The station trigger time is decisive to determine the shower direction. It is measured at each local station using a commercial Motorola GPS board, achieving a time precision of ~ 8 ns. Each station is self-contained and autonomous. Two solar panels provide an average of 10 W and two 12 V batteries power the PMTs and the electronics. Communication with the central station is accomplished through a wireless LAN radio system [1, 118].

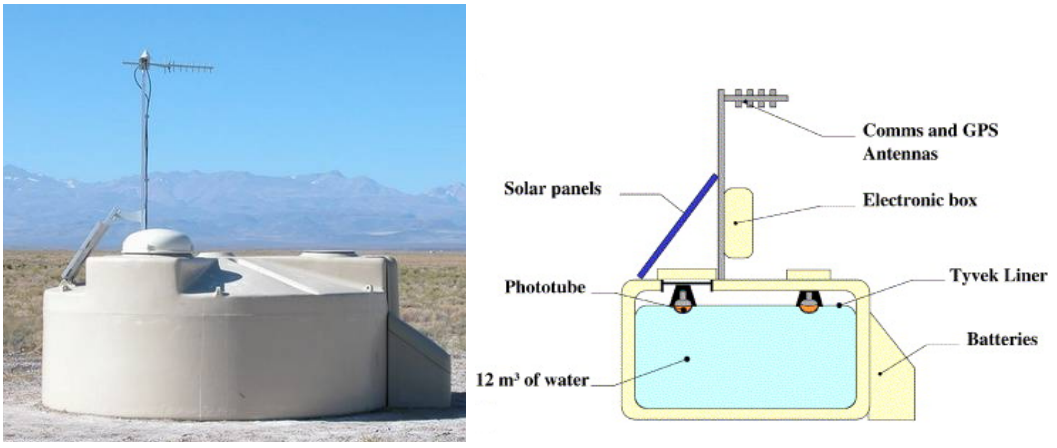


Figure 3.2: Left: Photograph of a Water-Cherenkov tank deployed in the field. Right: Schematic view of the station components.

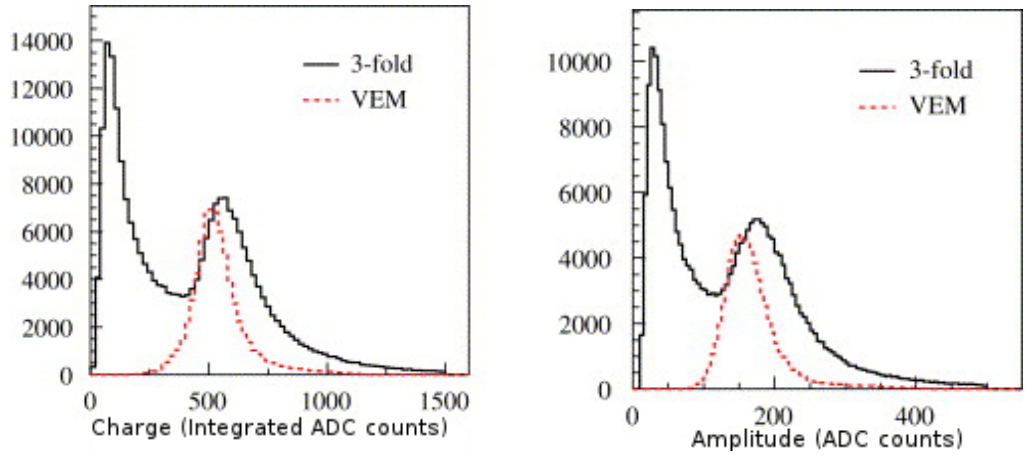
3.1.1 SD Calibration

The reconstruction of an EAS with the surface detector relies upon the values of the signals registered by the tanks and the relative timing between them. A precise cross-calibration of the PMTs is thus mandatory. Given the large number of detectors, and the long distance between them, a self-calibration procedure with the local electronics instead of detector inter-calibration is more feasible.

The Cherenkov light emitted by the particles traversing the tanks is measured in units of the average charge collected when a vertical muon crosses a tank through its center, termed vertical equivalent muon (VEM, or Q_{VEM}). The goal of the calibration is obtaining the value of 1 VEM in electronic units (integrated counts), and setting a common trigger threshold in detector-independent units.

The high rate of atmospheric muons crossing the tanks ($\sim$ 2500 Hz) provides an excellent method for measuring 1 VEM precisely. Even though tanks have no means of selecting only vertical central-crossing muons, to which the VEM is related, atmospheric muons form an extremely well understood and uniform background across the array. The distribution of the integrated pulses from atmospheric muons is shown in figure 3.3(a) (solid line). The first peak is caused by the convolution of the trigger on a steeply falling distribution from low-energy atmospheric muons. The second peak of the distribution (Q^{peak}_{VEM}) can be related to the Q_{VEM}: this peak is generated by vertical through-going atmospheric muons and corresponds to $Q^{peak}_{VEM} = 1.09\ Q_{VEM}$ for the average of the 3 PMTs and $(1.03 \pm 0.02\ Q_{VEM})$ for each PMT, both measured in a reference tank, using a muon telescope to select vertical muons [171]. These values are different because the sum of the PMTs measures the total signal in the tank, whereas individual PMTs measure the portion of the signal deposited closest to them. The geometrical arguments guiding to this relation can be found in [172].

In addition to a reference unit for the integrated pulse the stations must have a reference unit for the measured current too. The first level trigger of the stations relies on the PMTs' signals exceeding a certain threshold, and this trigger is set in electronic units (ADC counts,



(a) Charge histogram

(b) Pulse height histogram

Figure 3.3: Solid histograms correspond to a SD station triggered by a 3-fold coincidence between all 3 PMTs, using atmospheric muons. The trigger level is set at five counts above baseline. The signal from all 3 PMTs is averaged. An external muon telescope selects only vertical and central muons, used to generate the dashed histogram.

or current units). The distribution of pulse heights for atmospheric muons is shown in figure 3.3(b) (solid line). As in the distribution of integrated pulses, there are two peaks, which have an equivalent interpretation. The second peak, I_{VEM}^{peak}, is related to the peak current produced by vertical muons, I_{VEM}, and is used as the common reference for the threshold levels. The conversion from electronic units to I_{VEM}^{peak} needs a continuous update (so a proper trigger level can be maintained) and it should be roughly equivalent between the PMTs, to ensure that the signals recorded by each one are similar. The whole process of calibration to VEM units can be summarized in three steps:

1. Adjust the high-voltage of each PMT to set up its value of I_{VEM}^{peak} to 50 ADC counts.
2. Continuous calibration at station level to determine the value of I_{VEM}^{peak} in counts needed to keep the rate of events at 100 Hz.
3. Determine Q_{VEM}^{peak} precisely with charge histograms and use the known conversion from Q_{VEM}^{peak} to 1 VEM to convert it from the integrated signals of the PMTs to VEM units.

The calibration parameters mentioned above are determined every 60 s and transmitted to the CDAS with each event, stored along with the event data. Precisions of 3% and 6% are achieved in the determination of Q_{VEM}^{peak} and I_{VEM}^{peak}, respectively [172].

3.1.2 SD Trigger

The tanks of the surface detector are continuously being traversed by particles coming from the atmosphere, but not all have their origin in high-energy EAS. A trigger system is needed to separate physical events from background particles, setting a series of constraints on the signals of the tanks. Additionally, the trigger system must comply with the technical requirements imposed by the detector, mainly from those derived from the wireless communication system. The maximum sustainable rate of events per detector is < 1 per hour, whereas the station counting rate is ~ 3 kHz, due to the atmospheric muon flux. The first objective of the trigger system is to reduce the single station counting rate, while keeping the interesting events. It has been designed following a hierarchical structure: with each trigger level, discrimination against background becomes increasingly stricter, and the single station counting rate decreases

accordingly. The total bandwidth available for data transmission from the detectors to the CDAS is 1200 bits·s^{-1}. This rate constraints the first and second level triggers to be performed at station-level, to avoid the saturation of the central trigger system.

Local triggers

The first and second level triggers (called **T1** and **T2**) are formed locally at each detector, upon the analysis of the signals of the PMTs. Two independent and complementary trigger modes are implemented as T1; they are conceived to detect the electromagnetic and muonic components of the EAS.

The T1 Simple Threshold trigger (or **TH**) requires a three-fold coincidence of the PMTs in a single time bin, each above $1.75 I^{peak}_{VEM}$. This trigger is efficient at selecting large and narrow signals, characteristic of the muonic component, and reduces the rate of atmospheric muons from $\sim$ 3 kHz to $\sim$ 100 Hz. The second T1 mode is intended to select sequences of small signals spread in time, dominant in two different scenarios: near-by, low-energy showers, with a strong electromagnetic component; and high-energy showers with cores far away from the tank. It is called Time-over-Threshold (**ToT**), and requires a minimum of 13 bins (325 ns) in 120 FADC bins of a sliding window of 3 µs with signals above a threshold of 0.2 I^{peak}_{VEM} in 2 out of 3 PMTs. The spread of the signals arises from a combination of scattering (relevant for the electromagnetic component) and geometrical effects (relevant for muons). The ToT rate at each detector is < 2 Hz.

The T2 trigger is applied in the station controller to reduce the rate of events per detector to around 20 Hz. T2 triggers are sent to the CDAS for the formation of the trigger at array-level. Whereas ToT-T1 triggers are automatically promoted to the T2 level, TH-T1 triggers are requested to pass another threshold of $3.2 I^{peak}_{VEM}$, again in a single-bin three-fold coincidence, to be considered T2.

Array triggers

The third level trigger (**T3**) initiates the data acquisition from the CDAS. It is based on the spatial and temporal combination of second level triggers. There are two different T3 trigger modes, sensitive to different space configurations of the stations (footprints). The first mode requires at least three detectors passing the ToT trigger condition in a particular spatial configuration: at least one of the detectors must have one of its closest neighbors, and one of its second closest neighbors, with trigger. This configuration is called ToT2C$_1$3C$_2$. C$_\mathrm{n}$ stands for the n^{th} crown of neighboring tanks around a reference one, and mC$_\mathrm{n}$ means at least m tanks satisfying the trigger condition within the first n crowns around the reference tank. A timing criteria is added to this spatial coincidence: each T2 must be within (6+5C$_\mathrm{n}$) µs of the reference one. The rate of this T3 mode is around 1600 events per day, 90% of which are real showers. For the detection of horizontal showers, a different definition of the T3 trigger is more efficient. A four-fold coincidence of any T2 in a configuration 2C$_1$3C$_2$4C$_4$ is required (2 stations within the first crown, 3 within the first two crowns, 4 within the first four crowns). In this thesis we have treated inclined events, and we studied their trigger on the surface detector on detail. Generally, the estimates of the inclined trigger efficiency have been derived from studies of simulated events, generating two-dimensional T2 maps. In appendix A we describe a method for estimating the T3 trigger efficiency directly from the data, using local trigger probability functions. Examples of the two trigger configurations are shown in figure 3.4, along with the structure of crowns around a central tank. This trigger selects about 1200 events per day, with a real shower detection efficiency around 100%.

The next level of trigger concerns the physics selection. Real showers have to be selected among the stored events that fulfill the T3 criteria. Due to the large number of detectors and to the possible combinations between them, a large number of events is expected to arise from chance coincidences. The fourth level trigger, **T4**, is based on the space and time configurations of the detectors and holds two different criteria, with different aims. The first one corresponds to a ToT-3C$_1$ configuration (**3ToT** from now on) and it requires three adjacent T2-ToT stations

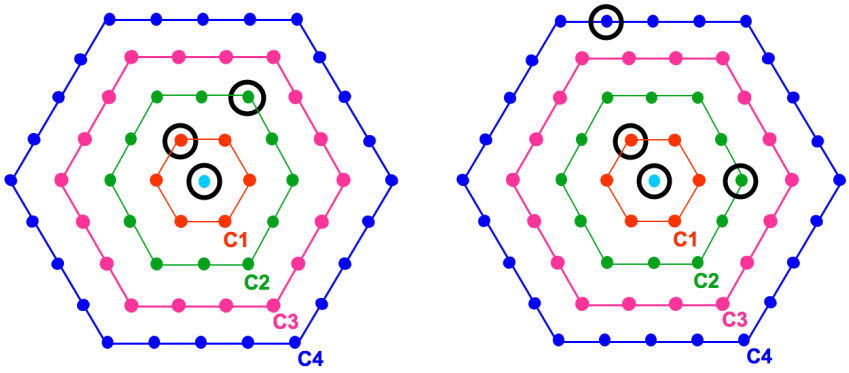


Figure 3.4: Minimal T3 trigger configurations. Left: ToT2C$_1$3C$_2$. Right: ToT2C$_1$3C$_2$4C$_4$

in a triangular pattern, whose possible minimal configurations are shown in figure 3.5 (dashed lines). The 3ToT is an excellent criterion to find real physics events up to 60º, with a selection efficiency above 98%. A second criterion, T2-4C$_1$ (abbreviated as **4C1**), is used to detect showers at larger zenith angles. The 4C1 needs four nearby stations with any T2 trigger. In figure 3.5 (solid lines) we show the three minimal 4C1 configurations. In any of these two triggers the stations times must fit to a plane shower front moving at the speed of light. The 3C$_1$ configuration with the highest total signal where the tanks are not aligned is chosen as a reconstruction seed. The arrival direction of the shower is determined fitting the arrival times of the signals of the detectors of the seed to a plane shower front moving with the speed of light. In any event, there can be stations that have signals seemingly in time with the rest of stations in the event, but are not really part of the shower. To identify these stations the time delays with respect to the fitted shower front are computed. Those stations with delays outside a window [−2 µs,+1 µs] are flagged as *accidentals* and rejected. This is not the only requirement for a station to be flagged as accidental; e.g. detectors with no triggered neighbors within 3 km are flagged as *lonely*, and are always removed. The joint detection efficiency of the 3ToT and 4C1 triggers is ∼100% for showers below 60º and enhances the selection of inclined events.

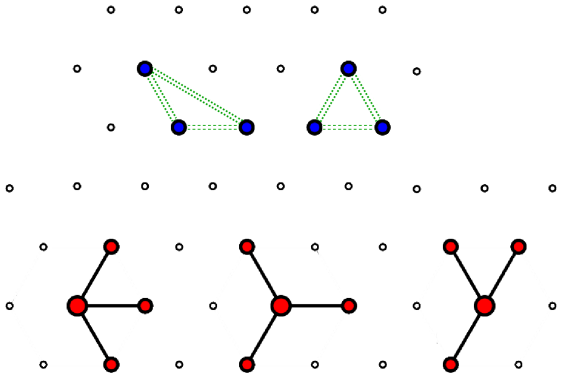
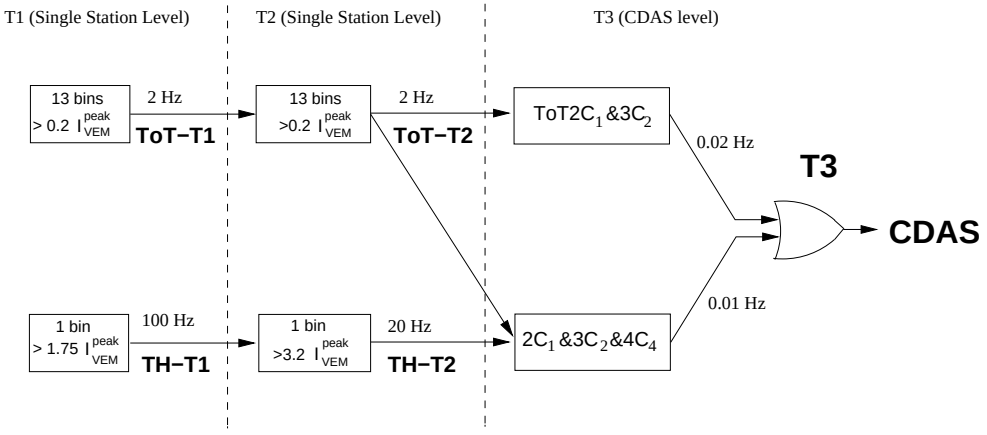


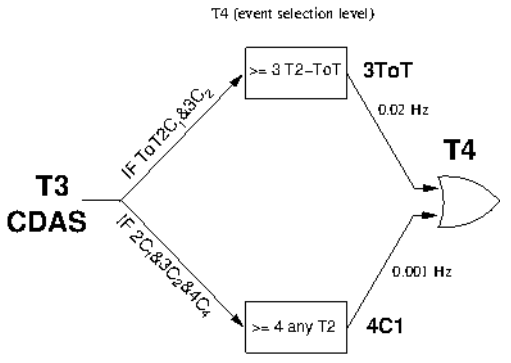
Figure 3.5: Minimal T4 configurations: 3ToT (dashed lines on top) and 4C1 (solid lines, bottom).

The finite size of the array implies that part of the shower will not be detected in events falling close to the border of the array. The missing information from the SD can lead to incorrect core reconstructions and wrong energy assignments. The **T5** trigger (or fiducial trigger) selects events well contained within the array, ensuring a proper core reconstruction. It requires the detector with the highest signal to be surrounded by a hexagon of working stations. This is called the 6T5 trigger. A less restrictive criterion, called 5T5, requires only 5 working stations around

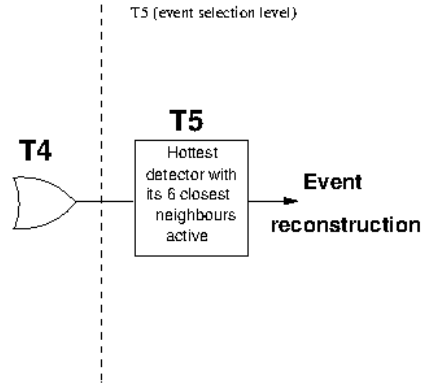
the station with the largest signal. Due to the large number of detectors of the surface array, about 1% of them is expected to malfunction at any time, even with constant maintenance. Thus, the T5 trigger will discard events that, even if contained within the array, fall close to a non-working detector. The full trigger chain, from the station T1 trigger, to the event T5 trigger, is summarized in figures 3.6(a) to 3.6(c).



(a) T3 trigger formation from T1 and T2 level triggers



(b) T4 trigger formation from T3 level trigger



(c) T5 trigger requirements

Figure 3.6: Summary of the trigger system of the Pierre Auger Observatory SD, from the local station (T1) to the event (T5) trigger.

3.1.3 SD Reconstruction

From showers fulfilling the T5 trigger it is possible to extract enough information to allow for the reconstruction of the energy and direction of the primary cosmic ray. Using the relative timing of the surface detectors, the shower direction and front curvature can be extracted, while the shower energy can be derived from the lateral distribution of signals.

However, not every triggered station is used during the reconstruction process. Only stations considered to be *candidates* are used. During the calibration step, stations with random triggers are removed and those that yield no data are flagged as *accidental*. Events with signals due to lightnings are removed as well. The clue to detect lightning-originated signals is a series of oscillations in the FADC traces of all the PMTs of the station: if the signal does not exceed 1000 FADC counts and makes more than three baseline crossings, it is considered to be generated from a lightning. Finally, stations are checked for time compatibility. Given the reconstruction seed (the 3C1 stations with maximal sum of signals) we require the compatibility with a planar

shower front propagating with the speed of light. A sketch of the planar shower front is shown in figure 3.7 (left). The station with the highest signal is used as the local origin of position and time (x_1, t_1):

$$c(t_i - t_1) = -\hat{a}(\vec{x}_i - \vec{x}_1),\ i \in \{1, 2, 3\} \tag{3.1}$$

where c is the speed of light, and $\vec{x}_i$ and t_i are the position and time of the i^{th} station, respectively. This equation defines two projections that determine a linear system from which a provisional axis $\hat{a}$ is obtained. Using this axis and the time of the reference station, t_1, the start time of the rest of stations is checked. The predicted shower time at position $\vec{x}$ is:

$$t_{sh}(\vec{x}) = t_1 - \hat{a}(\vec{x} - \vec{x}_1)/c \tag{3.2}$$

For each station, the difference between the actual and predicted start times, i.e. the station delays, must satisfy the condition:

$$-1000 \text{ ns} < t_i - t_{sh}(\vec{x}_i) < 2000 \text{ ns} \tag{3.3}$$

Otherwise, the station is flagged as *accidental*. Stations with no neighbors within 1800 m or with only one within 5000 m are flagged as *lonely*, and discarded as well.

The signal weighted barycenter, $\vec{b}$, and barytime, t_0, are set as the new origin from which all distances and times are measured at the first stage of the reconstruction. A shower track can be visualized as a point $\vec{x}(t)$ moving with the speed of light along the line defined by the axis, and hitting the ground (passing through $\vec{b}$) at time t_0:

$$\vec{x}(t) - \vec{b} = -c(t - t_0)\hat{a} \tag{3.4}$$

The shower plane is the first approximation to the shower front: it is a plane perpendicular

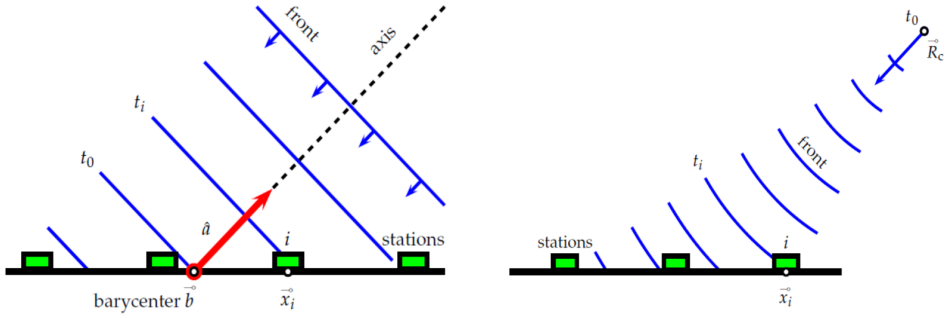


Figure 3.7: Schematic of the plane (left) and spherical (right) shower plane arrival.

to the shower axis, moving along with the same speed and containing the shower most ahead component. The time when the shower plane passes through some point on the ground, $t(\vec{x})$, can be predicted projecting that point onto the shower axis:

$$ct(\vec{x}) = ct_0 - (\vec{x} - \vec{b})\hat{a} \tag{3.5}$$

The shower plane is then obtained minimizing the sum of the squared time differences between the measured signal start times and the predicted times. Assuming that the positions of the stations are given with absolute precision and the only deviations are due to the uncertainty of the signal start time, σ_t, the function to minimize takes the form:

$$\chi^2 = \frac{1}{\sigma_t^2}\sum_i [t_i - t(\vec{x}_i)^2] = \frac{1}{c^2\sigma_t^2}\sum_i [ct_i - ct_0 + \vec{x}_i\hat{a}]^2 \tag{3.6}$$

where $\vec{x}_i = \vec{x}_i^{\circ} - \vec{b}^{\circ}$ and t_i are the position and time of the i^{th} station, respectively. Denoting the axis as $\hat{a} = (u, v, w)$, the station coordinates with $\vec{x}_i = (x_i, y_i, z_i)$ and $\sigma = c\sigma_t$ we can write:

$$\chi^2 = \frac{1}{\sigma^2} \sum_i [ct_i - ct_0 + x_i u + y_i v + z_i w]^2 \tag{3.7}$$

The components of the shower plane have to fulfill the condition

$$\hat{a} \cdot \hat{a} = 1 \rightarrow u^2 + v^2 + w^2 = 1 \tag{3.8}$$

This constraint introduces a dependence in $w = \sqrt{1 - u^2 - v^2}$ in equation 3.7, and we are confronted with a non-linear problem. Still, an approximate solution can be obtained if all stations lay close to some plane, and $z_i \ll x_i, y_i$. The z component is neglected and a linear approximation to equation 3.7 is obtained. This approximate solution can be used as a seed for more complex fitting attempts. In any case, the plane shower front is an approximation that can be refined assuming a curved front fit, such as shown in figure 3.7 (right). Equation 3.5 can be extended to describe the curvature of the shower front near the impact point $\vec{c}^{\circ}$, i.e. $\rho << R_c$, introducing a parabolic term:

$$ct(\vec{x}^{\circ}) = ct_0 - \hat{a}\vec{x} + \frac{\rho(\vec{x})^2}{2R_c} \tag{3.9}$$

where $\rho(\vec{x})^2 = (\hat{a} \times \vec{x})^2 = x^2 - (\hat{a}\vec{x})^2$ is the perpendicular distance and $\vec{x} = \vec{x}^{\circ} - \vec{c}^{\circ}$. The shower development is depicted as starting at time t_0 from a single point, $\vec{R}_c^{\circ}$, and propagating towards the stations. Thus, the start time of the i^{th} station, t_i, is given by:

$$c(t_i - t_0) = |\vec{R}_c^{\circ} - \vec{x}_i^{\circ}| \tag{3.10}$$

The propagation of the shower front is described as an expanding sphere, and the timing information is decoupled from the determination of the impact point. The shower axis becomes a derived quantity obtained only after the position of the impact point is known. The solid angle difference between the axis $\hat{a}$ obtained in the plane- and curvature-fit is of the order of half a degree. The exact curvature fit involves the minimization of the function

$$\chi^2 = \frac{1}{\sigma^2} \sum_i [c(t_i - t_0) - |R_c\hat{a} - \vec{x}_i|]^2 \tag{3.11}$$

without the assumption of small values of z_i.

The angular resolution of the axis is determined from simulations, by computing the angle between the injected shower axis and the reconstructed one, applying the same reconstruction procedure used for real data. The resolution depends on the number of tanks used in the reconstruction. It is around 2.2^{o} for events with energy $E < 4$ EeV, which on average only have 3 candidate tanks, and is better than 1^{o} above $E > 10$ EeV, where events show a large multiplicity [173].

3.2 The Fluorescence Detector (FD)

The Fluorescence Detector (FD) consists of 4 observation sites - Los Leones, Los Morados, Loma Amarilla and Coihueco - each one containing 6 independent telescopes. All fluorescence sites have been completed and are in operation. The last site to be completed and start operation was Loma Amarilla, in 2007. An aerial view of the FD site at Coihueco can be seen in figure 3.8 (left). The aim of the FD is detecting the fluorescence light emitted by the atmospheric nitrogen, excited by charged particles generated during the development of the shower. The detection of ultra-high energy cosmic rays using nitrogen fluorescence is a well established technique, used previously in Fly's Eye [33] and HiRes [120], and used by the recently finished Telescope Array as well [174].

Each telescope is housed in a clean, climate-controlled building. The optical system of the telescope consists of a filter at the entrance window, a circular aperture, a corrector ring, a mirror and a camera with photomultipliers. Figure 3.8 (right) shows an schematic of the components of a telescope. All the elements except the filter constitute a modified Schmidt camera design that partially corrects spherical aberration and eliminates coma aberration [121].

The entrance window is an optical filter designed to absorb visible light while transmitting UV photons in the range 290-410 nm. This range of frequencies includes almost all the nitrogen fluorescence spectrum, shown in figure 2.4 from section 2.4.2. Should the filter window be removed, the fluorescence signals would fade in the noise of visible photons. The size of the aperture is optimized to keep the angular spread of the light around 0.5°. In comparison, the field of view of any single camera pixel is 1.5°. The corrector ring doubles the aperture area of the telescope, while keeping the properties of the Schmidt system. Regarding the mirrors, two different mirror systems are used. A tessellation of 36 rectangular anodized aluminum mirrors of three different sizes is used in the 12 telescopes of Los Leones and Los Morados. Two layers are glued to the aluminum surface, a sheet of $AlMgSiO_5$ alloy to achieve reflectivity, and an aluminum-oxide layer to provide additional protection. In Loma Amarilla and Coihueco, a structure of 60 hexagonal glass mirrors with reflective coatings is installed. The reflective layer is made of one layer of aluminum and another layer of silicon dioxide (SiO_2). In both types of mirrors, the average reflectivity at λ=370 nm exceeds 90%.

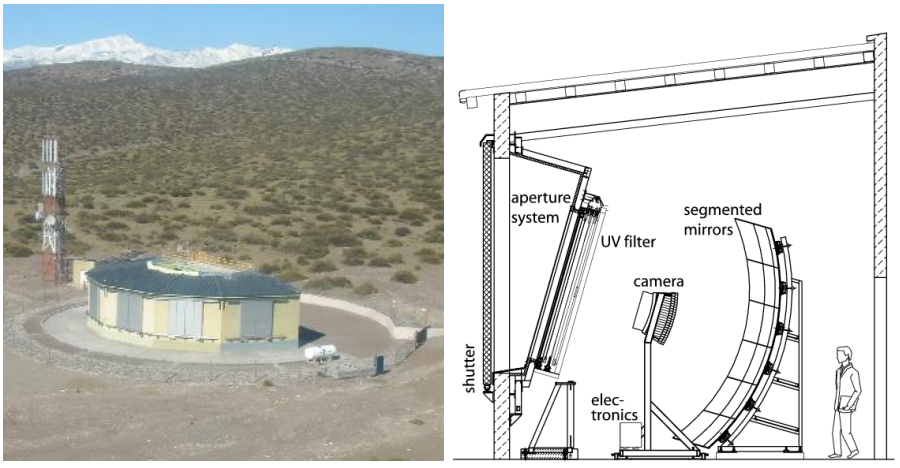


Figure 3.8: Left: Aerial view of the FD building at Coihueco. Right: Scheme of the components of a Fluorescence Detector telescope.

Each camera is a matrix of 440 hexagonal pixels located on the focal surface of the telescope mirror. The camera pixels are arranged in a matrix of 22 rows by 20 columns, with a total field of view of 30° in azimuth and 28.1° in elevation. Pixels are instrumented with an eight-stage PMT tube (model Photonis XP3062), inside a hexagonal window (40 mm side to side) complemented by light collectors that also guarantee a smooth transition between adjacent pixels. High voltage is provided by a CAEN SY527 system to 10 groups of PMTs with similar gains, and a commercial power supply provides the low voltage. Both the high and low voltage are distributed by 10 power control boards at the back of the camera. The PMT signals are received by a set of 20 front-end boards, each serving 22 pixels of a camera column. The signals are continuously digitized by 10 MHz 12 bit ADCs.

The FD is operated during nights with moon fraction below 60%. The observation period lasts 16 days per month, with an average observation time of about 10 h (a maximum of 14 h in June, a minimum of 5 h in December). FD data-taking can only take place under high-quality environmental and atmospheric conditions. Either the presence of the sun or nearly full-moon in the sky, the moon in any phase within 5° of the FOV of a telescope, poor weather or weather conditions dangerous for operation (rain, snow, high wind speed) force the shutters

to automatically close. The operation of the fluorescence detector is usually characterized by the uptime, the fraction of the total time during which the FD is acquiring data, $\sim 11\%$ averaging over all telescopes. The uptime fraction for each telescope in the observation period between January 2004 and December 2012 is shown in figure 3.9.

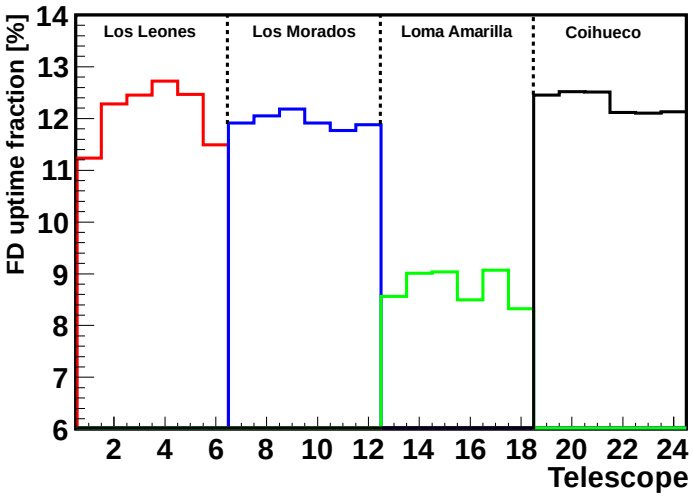


Figure 3.9: Uptime fraction between 2004 and 2012 for all fluorescence telescopes. Telescopes are numbered as follows: Los Leones site 1-6, Los Morados site 7-12, Loma Amarilla site 13-18, Coihueco site 19-24.

3.2.1 FD Calibration

The reconstruction of the longitudinal profile of an EAS and the determination of its total energy rely on the ability to convert pixel ADC counts into a light flux. Thus, the evaluation of the response of each pixel to a certain flux of incident photons is of the utmost importance. A step-by-step calibration of each pixel would need to account for a large number of effects, such as mirror reflectivity, pixel light collection efficiency, cathode quantum efficiency and PMT gains. Therefore, a single end-to-end calibration is performed instead. A calibrated 2.5 m diameter light source, called the *drum*, consisting on a pulsed UV LED, with wavelength 375 ± 12 nm is installed at the telescope aperture. The known characteristics of the light source together with the response of the acquisition system provides a known, uniform response for each pixel in each camera of the FD detector. The average response of the FD is $\sim$5 photons/ADC bin [121]. The drum calibration is performed once per year, but three additional calibration procedures are used as well:

Spectral calibration Since the fluorescence detector is not only sensitive to a single wavelength, but to a range, a calibration of the response at wavelengths of 320, 337, 355, 380 and 405 nm is performed. This calibration defines a spectral response curve, relative to the reference wavelength of 380 nm. A series of filters are used to select each wavelength from a xenon flasher mounted at the back of the drum.

PMT inter-calibration Before and after each night of data taking a relative calibration between different PMTs is performed. Three different positions of each camera are illuminated, monitoring different groups of detector components. This calibration tracks short- and long-term changes in the detector response.

Calibration cross-check Remote vertical laser shots of known intensity and wavelength are used as a cross-check of the drum calibration. A calculable fraction of the photons is

scattered to the aperture of the FD detector, yielding a known number of photons that arrive to the detector for each pixel. The response of each pixel to the arriving photons constitutes another end-to-end calibration method.

3.2.2 FD Trigger

The electronics and data acquisition system of the FD face the challenge of recording the nitrogen fluorescence signals, with a widely varying intensity, over a background of intense and changing light. A large dynamic range is needed, and it should provide a strong background rejection while accepting any physically plausible air shower. To that purpose, various levels of trigger are defined at hardware and software level.

Hardware triggers: First and Second Level Trigger

The 22 front-end boards that record the signal received by the pixel cameras are also used to implement the First Level Trigger (**FLT**), or pixel trigger. A running sum of the last n ADC bins ($5 \leq n \leq 16$) is compared to an adjustable threshold. A pixel trigger is generated whenever the sum exceeds the threshold. The trigger rate of individual pixels is continously measured, and the threshold value is dynamically adjusted to keep the trigger rate as close as possible to 100 MHz. When the running sum drops below the threshold, the pixel trigger is extended for a period of 5-30 µs, increasing the chance of coincident pixel triggers. The multiplicity (number of pixels triggered simultaneously within 100 ns) is calculated for each 22 pixel column and for the full camera. The chronological sequence of multiplicity values carries information about the temporal development of the camera image.

The Second Level Trigger (**SLT**) is designed to detect straight-track patterns in the pixels triggered by the FLT. The algorithm searches for track segments of at least five adjacent triggered pixels in any of the patterns depicted in figure 3.10, as well as those generated by their rotations and mirror reflections. If certain PMTs are not collecting enough light to trigger or are somehow defective it is possible to find untriggered pixels in the middle of a track. To remedy this situation the algorithm requires only four out of five triggered pixels, giving rise to 108 different combinations of four-fold patterns from the five-pixel track segments.

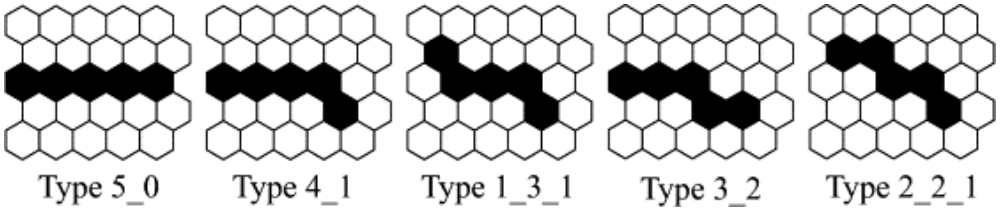


Figure 3.10: Basic patterns of triggered pixels used in the Second Level Trigger of the Fluorescence Detector.

Software triggers: Third Level and Hybrid Trigger

The Third Level Trigger (**TLT**) is a software algorithm intended to remove noise events that survive the first and second level triggers. Lightning events, triggers caused by muon impacts and randomly triggered pixels are the main causes of false triggers. In optimal working conditions, the SLT will detect one or two events per minute per telescope. A lightning can trigger hundreds of pixels at the same time, in bursts of severals tens of events per second. Such a high rate of events can saturate the FLT and SLT systems, increasing the dead-time considerably. The TLT filters lightning events reading the FLT multiplicities. Cuts based on the time development of the multiplicity and its integral over the whole event are able to cut approximately 99% of all lightning events in a 50 µs decision window. Noisy channels far off the light track, with

signals generated by fluctuations or atmospheric muon impacts, can be discarded analyzing the space-time correlations of the pixels. The TLT is very efficient discriminating real showers from background events, keeping the fraction of true showers rejected below 0.7%.

Once an event passes the TLT, the software system merges coincident events from adjacent telescopes and sends a hybrid trigger, called T3, to the CDAS. The T3 acts as an external trigger for the SD, allowing to record events below $3{\cdot}10^{18}$ eV, where the array is not fully efficient and would seldom generate an independent trigger. At these energies no more than one or two SD stations are triggered, but this limited information is enough to ensure a high-quality hybrid reconstruction.

3.2.3 Atmospheric monitoring

The atmosphere is not only responsible for producing light from air showers, but it is also the medium through which it is transmitted to the detector. Since the properties and local composition of the atmosphere are highly variable, any measurement performed with the FD has to be corrected for these changing conditions. To remove the effect of atmospheric fluctuations that would impact FD measurements, an extensive atmospheric monitoring program is carried at the Pierre Auger Observatory. In figure 3.11 the different experimental setups installed at Malargüe to monitor the atmosphere are listed. In the lowest 15 km of the atmosphere where

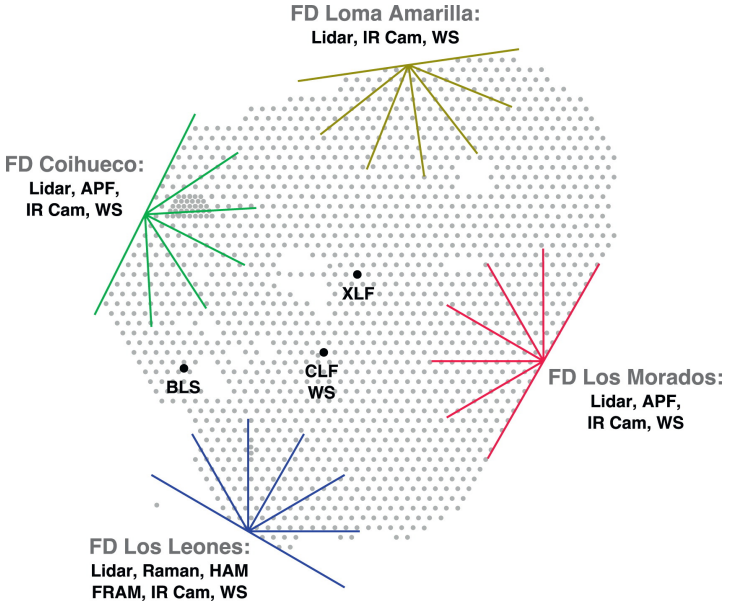


Figure 3.11: Situation of the all the apparatuses installed at the Pierre Auger Observatory to monitor the atmospheric conditions. The balloon, XLF and CLF stations are represented as red dots inside the array. The instruments installed at each FD station are labeled over each one.

EAS development occurs, Mie scattering due to aerosols with sizes from sub-µm to mm play an important role in modifying the light transmission. The vertical density profile of aerosols, as well as their size, shape and composition, vary strongly with location and time, even on an hourly basis. If neglected or not measured properly, such dynamic and variable conditions can bias the shower reconstruction [175]. During the dark periods suitable for FD data-taking hourly measurements of aerosols are made using the FD telescopes. They record data from vertical UV laser tracks produced by the Central Laser Facility (CLF), located near the center of the array, and the Extreme Laser Facility (XLF) [168]. In addition, four LIDAR stations located near each FD building [169] continuously operate outside the FD field of view and detect clouds and aerosols by analyzing the backscatter signal of a 351 nm pulsed laser beam [176]. Two Aerosol

Phase Function Monitors (APFs) are used to determine the aerosol scattering properties of the atmosphere. A Xenon flash lamp at two of the FD sites fires a set of five horizontal shots through the field of view of five out of the six telescopes with a repetition rate of 0.5 Hz once every hour [177]. The resulting signal distribution gives the scattering properties of the aerosols as a function of the scattering angle. Neglecting the presence of aerosols causes an underestimate of the energy that goes from 8% at the lowest energies up to 25% at the highest energies, and a shift in X_{max} between -1 g·cm^{-2} and 10 g·cm^{-2}. Approximately 5% of CLF measurements have optical depths greater than 0.1, i.e. the light transmission factor in the atmosphere is lower than 90%. To avoid making very large corrections to the expected light flux from distant showers, these nights are rejected for air shower reconstruction [175, 178].

Cloud coverage has a major influence on the reconstruction of air showers. Clouds can block the transmission of light from EAS to the detectors, or enhance the observed light flux due to multiple scattering of the intense Cherenkov light flux. The determination of cloud composition is nontrivial, and the estimates of scattering properties are highly unreliable. Since it is difficult to correct for the transmission of light through clouds it is safer to remove cloudy periods from the data taking process. Cloud coverage is measured by infrared Raytheon 2000B cameras located at the roof of each FD building. The cameras were designed to measure infra-red light in the 7-14 µm wavelength band, suitable for distinguishing warm clouds from the cold clear sky [170] These cameras photograph the field of view of each FD station every 5 minutes, producing 5 images that together cover the individual fields of view of each telescope. A cloud index, representing the fraction of clouds in its field of view, is generated for each FD pixel. This index can be used to remove single cloudy pixels from the reconstruction. However, if the cloud fraction is larger than 25%, events are rejected as a whole. Approximately 30% of the events are rejected due to cloudy conditions [179].

3.2.4 Axis, profile and energy reconstruction

Each pixel of the detector system records a pulse of light, from which it is possible to determine its time of origin with a certain uncertainty. Using this temporal information it is possible to generate a trial geometry for the shower axis, eventually resulting in a prediction for the signal arrival times at each detector component. The uncertainty-weighted differences between the predicted and measured times are used to build a χ^2 value. Finally, the hypothesis with the minimum value of χ^2 is chosen as the reconstructed shower axis [121].

In the FD reconstruction, the first step is the determination of the the plane that includes the location of the detector and the line of the shower axis, called the shower-detector plane (SDP) (figure 3.12, left). Experimentally, it is the plane passing through the detector that contains most of the pointing directions of the FD pixels centered on the shower axis. The timing information of the FD pixels is then used to reconstruct the real shower axis within the SDP. The shower axis is defined by two parameters: R_p, the perpendicular distance from the camera to the track, and χ_0, the angle the track makes with the horizontal line in the SDP. Using these parameters, the arrival time of the light at the i^{th} pixel, t_i, can be written as:

$$t_i = t_0 + \frac{R_p}{c} \tan\left[(\chi_0 - \chi_i)/2\right] \tag{3.12}$$

where χ_i is the angle of the pointing direction of the i^{th} pixel with the horizontal line, and t_0 is the time at which the shower front on the axis passes the point at distance R_p. R_p and χ_0 are determined fitting the data information from the pixels to this function. Thanks to the fast timing electronics, even monocular reconstruction can be fairly accurate. However, when the measured angular speed of the shower does not change much over the observed track-length, the solution to the axis determination is degenerated, and there is a family of possible solutions (χ_0, R_p). This degeneracy can be avoided combining the FD timing information with that of the SD stations, in a reconstruction mode called **hybrid**. In figure 3.12 (right) there is an example of a fit to the shower axis in which the monocular reconstruction (dashed line fitting only the points at large χ) fails to find the best solution. However, once the information from the SD stations (squares on the top left of the graph) is added, the hybrid reconstruction significantly

improves the fit (solid line). Multiple-eye events (stereo, triple and four-fold events) offer the

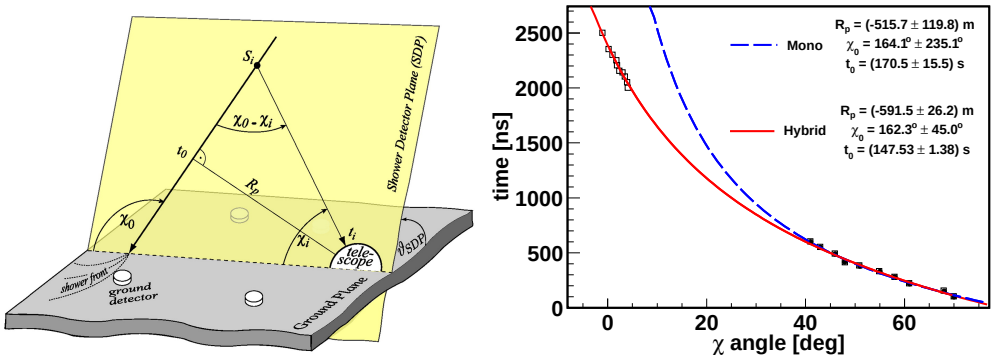


Figure 3.12: Left: Geometrical shower reconstruction and SDP fit parameters. Right: example of a reconstructed shower axis in the monocular (top fit) and hybrid (bottom fit) cases. The addition of the surface stations (squares on the top left of the graph) significantly improves the reconstruction.

chance to observe an incoming shower from different directions, performing an independent hybrid reconstruction for each triggered eye. Alternatively, the geometry of the shower may be reconstructed in a multiple-eye mode by intersecting the shower detector planes of the triggered eyes without using the time information of any surface detector. In the stereo mode, the geometry resolution depends on the SDP determination, mainly on the number of pixels used in the SDP fit. A minimum of 6 pixels in each FD station is required to select well reconstructed events. The angular resolution of the Fluorescence Detector is better than 0.6º at energies above $10^{18.5}$ eV [180].

Once the geometry of the event has been determined, the light flux at the aperture can be converted to energy deposited as a function of the slant depth[1]. To that purpose, the light attenuation from each source to the telescope has to be estimated and the different light sources that constitute the flux have to be considered: fluorescence, direct and scattered Cherenkov, and multiple-scattered light. The energy deposited at slant depth X_i can be expressed as:

$$\frac{dE}{dX_i} = N_e(X_i) \int_0^\infty f_e(E, X_i) \frac{dE}{dX_e}(E, X_i) dE \tag{3.13}$$

where $f_e(E, X_i)$ is the normalized electron energy distribution, $dE/dX_e(E, X_i)$ is the energy loss of a single electron with energy E and $N_e(X_i)$ denotes the number of electrons and positrons above a certain constant energy cutoff [181]. The fluorescence detectors have a limited field of view that makes it very difficult, if not impossible, to observe the full shower profile. Instead, the shower development in the atmosphere is estimated by fitting a Gaisser-Hillas function [136] to the reconstructed energy deposit profile:

$$f_{GH}(X) = dE/dX_{max} \cdot \left(\frac{X - X_0}{X_{max} - X_0} \right)^{(X_{max} - X_0)/\lambda} e^{(X_{max} - X)/\lambda} \tag{3.14}$$

where X_{max} represents the depth of maximum particle production and dE/dX_{max} the maximum energy deposit. X_0 and λ are shape parameters, sometimes wrongly identified with the depth of the shower first interaction and the absorption length (actually, the preferred values of X_0 are negative). For a given reconstructed shower profile, the best set of Gaisser-Hillas parameters can be obtained minimizing the squared differences between the reconstructed energy deposit and the fitted value. As long as a large fraction of the shower above and below the shower maximum

[1] Atmospheric depth measured along the shower axis.

has been sampled, the minimization procedure is successful. Otherwise, the problem is under determined, and the experimental information is not enough to reconstruct all the Gaisser-Hillas parameters. The calorimetric energy is given by the integral over the energy deposit profile:

$$E_{cal} = \int_0^\infty f_{GH}(X)\, dX \tag{3.15}$$

However, the electromagnetic component of an EAS does not account for the whole energy of the shower. Neutrinos scape undetected, and muons need long path lengths to release their energy. This missing energy, called *invisible energy*, is taken into consideration multiplying the calorimetric energy by a correction factor, f_{inv}:

$$E_{FD} = E_{miss} + E_{cal} = f_{inv} \cdot E_{cal} \tag{3.16}$$

This correction, shown in figure 3.13, is obtained from Monte-Carlo simulations, with differences for different primaries and a mild energy dependence, arising from the energy dependence of the mesons decay probability [137].

The resolution of the energy determined by the FD depends on the uncertainties associated to variations in the atmosphere (4.5%-6.9%), the invisible energy (1.5%) and the EAS geometry (5.2%-3.3%). The intervals correspond to the uncertainties at $10^{18.5}$ eV and 10^{20} eV, respectively. The overall energy resolution is almost constant with energy in the range $10^{18.5}$-10^{20} eV, and lies between a 7% and a 8% [182].

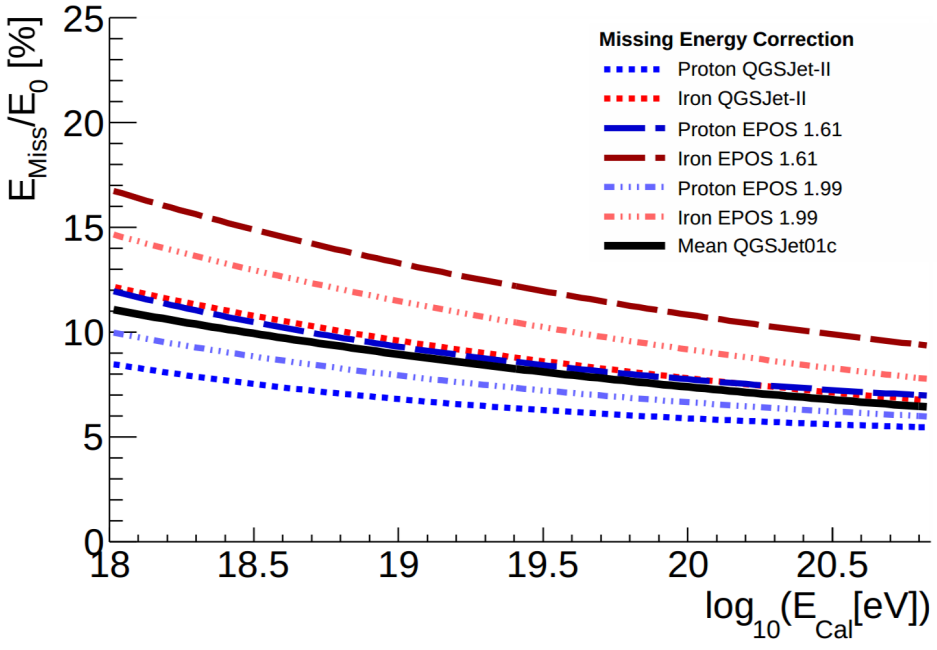


Figure 3.13: Missing energy correction as a function of the calorimetric energy for different hadronic models and primary masses.

3.3 Hybrid energy calibration

From the reconstruction of the events detected by the Fluorescence Detector we obtain an almost model-independent estimate of the cosmic ray energy, E_{FD}. For the Surface Detector there is no direct determination of the energy of the primary particle, E_{SD}, but we can measure estimators of the shower size, S_{SD}, that it correlates to. In vertical events this estimator is the signal at a

fixed distance from the shower axis, $S(d)$, obtained from a fit to a lateral distribution function [183], while in inclined events the muon content relative to simulated proton showers with energy 10^{19} eV, N_{19}, is used instead [35].

The value of d is chosen so that the fluctuations of the signal are minimized. For an array with 1500 m spacing, $d = 1000$ m allows a determination of the signal with an uncertainty smaller than 5%. For a smaller spacing of 750 m, as that of the infill array from the AMIGA enhancement (see section 3.4.3), the optimal value is $d = 450$ m [184]. However, the value of the signal at a fixed distance decreases with increasing zenith angles, due to the attenuation of the shower particles travelling larger distances from their production point to the detector. Values of $S(1000)$ at different zenith angles are converted to a common reference value of 38°, S_{38}, the value of $S(1000)$ the shower would have produced had it arrived with $\theta = 38°$. In the case of the array with 750 m spacing, $\theta = 35°$ is chosen as the reference angle.

In order to estimate the energy of an SD event, the shower size estimator S_{SD} (be it S_{38}, S_{35} or N_{19}) needs to be calibrated, using pairs (E_{FD}; S_{SD}) from a subset of high-quality hybrid events. The current energy calibration for vertical events of the regular array is performed using 1475 events above $10^{18.5}$ eV and $\theta < 60°$, detected between January 2004 and December 2012. Inclined events during the same period, with energies above $10^{18.6}$ eV, amount to 175. The calibration for the infill with 750 m spacing uses 414 events above $10^{17.5}$ eV and $\theta < 55°$, recorded from August 2008 to December 2012 [182]. The relation between S_{SD} and E_{FD} is well described by a power-law function $E_{FD} = AS_{SD}^{B}$. Figure 3.14 (left) shows the correlation between the different shower size estimators and the energy determined by the Fluorescence Detector.

The resolution of the energy assigned to vertical SD events detected with the regular array depends on the statistical error of the fit perfomed to compute $S(1000)$, the uncertainty in the lateral distribution functions and on the shower-to-shower fluctuations. The resolution obtained is a 17% at $10^{18.5}$ eV and improves to a 12% at 10^{20} eV [182]. In the case of the infill array, the resolution goes from a 23% at $1.3{\cdot}10^{17}$ eV to a 12% at $1.4{\cdot}10^{18}$ eV [166]. The resolution of the energy for inclined SD events depends on the fit of the predicted muon signals at ground (so called "muon maps") to the measured tank signals, once the average contribution of the electromagnetic component (typically a 20%) has been removed. The achieved resolution improves from 20% at $10^{18.6}$ eV to 8% at $10^{19.8}$ eV [185]. Finally, the systematic uncertainties in the determination of the energy by the FD have to be propagated to the determination of the energy by the SD. The systematics in the determination of the fluorescence yield (3.6%), the aerosol profiles ($\sim$ 5%), the FD calibration (10%), the FD profile reconstruction ($\sim$ 6%), the invisible energy ($\sim$ 3%) contribute to an uncertainty about 13%. This value, combined with the systematic uncertainties due to the calibration fit ($\sim$ 1%) and its stability over time (5%) amount to a total systematic uncertainty of a 14% in the energy scale.

3.4 Detector upgrades

As can been seen in figure 3.14, above an energy of $10^{18.5}$ eV the surface detector of the Pierre Auger Observatory reaches full shower detection efficiency. This energy threshold can be lowered to E $\sim 10^{18}$ eV in the case of hybrid detection. However, a better discrimination between astrophysical models requires the knowledge of the evolution of the cosmic ray composition along the transition region starting at the second knee, E $\sim 10^{17}$ eV. Following this idea, several improvements over the original design of the detector have been developed. They can be categorized into two different groups. The first group aims at detecting additional EAS signals, focusing on radio and microwave frequencies. The second group tries to investigate EAS signals at lower energies both in the SD and the FD, providing full detection efficiency starting at E $\sim$ 10^{17} eV. The location in the array of the different enhancements is shown in figure 3.15.

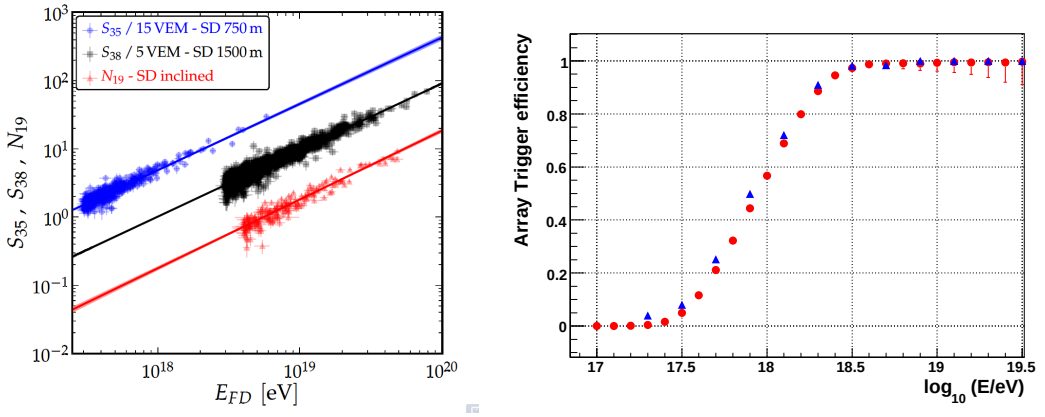


Figure 3.14: Left: Correlation between S_{38}, S_{35} and N_{19} with the energy determined by the FD [186]. Right: Trigger efficiency as a function of energy, derived from events detected only by the Surface Detector (blue triangles) and from hybrid events (red circles).

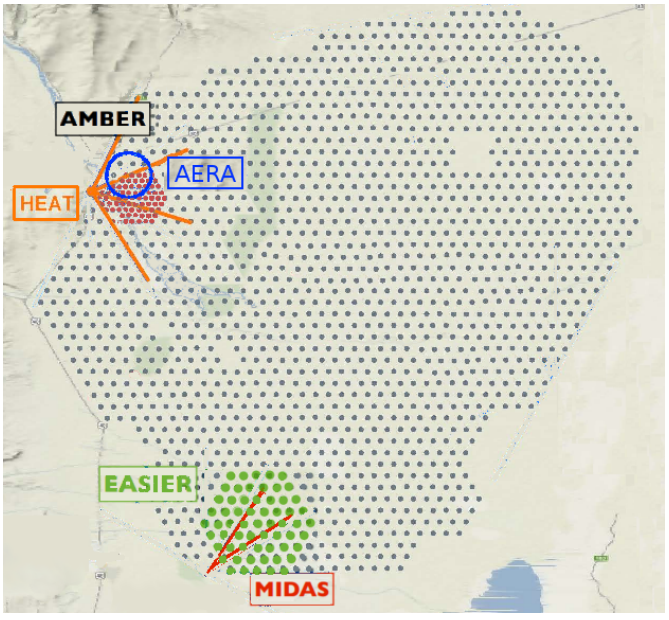


Figure 3.15: Situation of the enhancements of the Pierre Auger Observatory. The field of view of HEAT, installed next to Coihueco (top left of the array), is represented as orange lines. The infill array is represented as a denser group of red dots in front of Coihueco. The radio system AERA, and the microwave detection prototype AMBER are pictured next to the infill array. The surface detectors instrumented with EASIER antennas are represented as green dots to the south of the array. The red lines at the same location represent the field of view of the MIDAS antennas.

3.4.1 Radio and microwave detection

The foundation behind radio and microwave detection is that an array of antennas observing the radio or microwave wavefront at different positions with respect to the shower axis allows a reconstruction of the properties of the shower. In the low frequency range (MHz), there are two radio setups already operating at a small scale (3 antennas in an equilateral triangle) inside the area covered by the SD. These setups are located close to the Balloon Launching Stations (BLS)

and the Central Laser Facility (CLF) [187, 188]. Both detectors have observed radio signals in coincidence with air showers, showing the feasibility of the radio technique at the Pierre Auger Observatory.

- The purpose of AERA (**A**uger **E**ngineering **R**adio **A**rray) is the large scale detection of radio waves in the frequency range from 30 to 80 MHz. Its first phase, AERA24, started operating in April 2011, and consisted of 24 stations with logarithmic periodic dipole antennas (LPDAs) distributed over an area of 0.5 km^2 with a spacing of 125 m. The next phase will deploy 136 additional stations, with a spacing of 250 m [153]. The final detector array will consist of 160 autonomous radio-detector stations, covering at least 10 km^2. Co-located with AMIGA, its position is chosen to maximize the number of showers detected in coincidence with the other detectors. It is expected to record several thousands of cosmic rays showers in the range 10^{17} to 10^{19} eV.

- MIDAS (**MI**crowave **D**etection of **A**ir **S**howers) [125] and AMBER (**A**ir-shower **M**icrowave **B**remsstrahlung **E**xperimental **R**adiometer) [159] are the two first setups intended to observe air showers at GHz frequencies at the Pierre Auger Observatory. Radiation in the GHz regime, due to molecular bremsstrahlung, is expected to be emitted isotropically. This allows to track the shower development along its axis, in a fashion similar to that of the fluorescence detector but without the constraint of measuring in dark conditions. As noted in section 2.4.2, recent results by CROME [164] suggest that the main component in microwave emission is actually forward oriented, in a manner similar to Cherenkov emission. However, their results do not exclude the presence of a sub-leading isotropic component, for whose detection MIDAS and AMBER have been designed. AMBER and MIDAS are imaging telescopes like a FD, instrumenting an array of feed horn antennas at the focus of a parabolic dish. The AMBER and MIDAS prototypes are shown in figures 3.18(a) and 3.18(b) The major difference between both detectors is their trigger philosophy: whereas MIDAS will work with a self-triggering system, AMBER will be subordinate to the trigger of the SD. The MIDAS prototype is currently installed at the University of Chicago. Both setups will be installed in the near future close to Coihueco, joining the variety of techniques employed for shower detection in the north-west part of the array.

- EASIER (**E**xtensive **A**ir **S**hower **I**dentification using **E**lectron **R**adiometers) is a prototype that combines two different detectors aimed at detecting radio signals both in the MHz and the GHz range [154]. The EASIER stations consist of antennae mounted on the tanks of the surface detector (see figure 3.18(c)), and take advantage of their data aquisition system. The GHz design, with a system similar to that of MIDAS, operates in the 3.4-4.2 GHz range and will look for molecular bremsstrahlung emission. The MHz detector operates in a frequency range of 30 MHz to 70 MHz, with antennae designed following the design of those of CODALEMA [150] and is intended to detect geosynchrotron emission. Both detectors depend on the trigger of the SD, but do not interfere with its normal operation. So far 14 SD tanks have been instrumented, half with MHz detectors and half with GHz detectors, in separate hexagons near Los Leones FD building. The first detection of an air shower in the microwave range was performed in June 2011. It was registered in coincidence with an air shower detected by the SD, with energy $E = 13.2$ EeV and zenith angle of 29.7^{o}. Figure 3.16 shows the recorded GHz signal along with the PMT traces of the station closest to the shower core. The maximum of the microwave signal was more than 11 times larger than the noise fluctuations and occured just one time bin (25 ns) before the signal in the Water-Cherenkov detector [165]. When MIDAS is installed, it will overlook the GHz detectors of EASIER.

- FDWave will use modifications of the existing FD hardware. The sixth telescope at Los Leones site has a camera with 176 empty cells and the FDWave prototype intends to make use of these vacant cells, instrumenting them with radio detectors suitable for frequencies above 9 GHz. This detector will not trigger independently, operating in a passive mode dependent on the FD trigger [154].

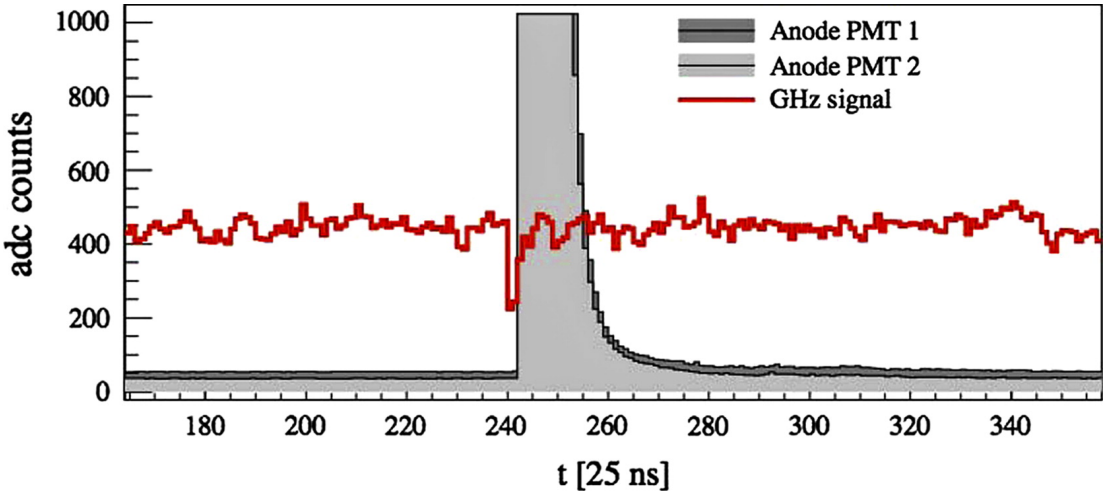


Figure 3.16: Microwave trace (red) of the first event recorded by an EASIER detector. In gray the signals of two low gain PMT channels are shown. The shower core is only 136 m away from the detector, and the traces are saturated.

3.4.2 HEAT

With decreasing primary energies, the shower development becomes faster, and the maximum of the shower is displaced to higher altitudes in the atmosphere. Thus, extending the telescopes field of view to higher depths provides a lower energy threshold for the fluorescence detectors. With this purpose, HEAT (**H**igh **E**levation **A**uger **T**elescope) was installed and has been taking data regularly since 2009. Placed close to the location of the Coihueco site, it comprises three telescopes with the same field of view as the regular telescopes. Its distinctive feature is that it can be tilted up to 30°, allowing it to overlook higher depths in the atmosphere. The HEAT telescopes working in tilted mode are shown in figure 3.18(d). The energy threshold is lowered down to 10^{17} eV for showers seen by HEAT. Working together, Coihueco and HEAT record different and complementary parts of the same shower. In figure 3.17 there is an example of a shower reconstructed with both detectors, in which X_{max} would have fallen outside the field of view if only Coihueco had recorded it.

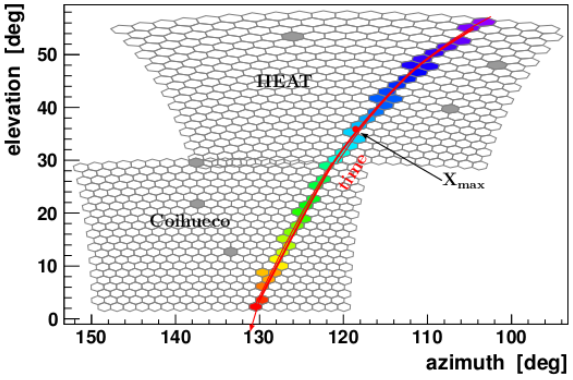


Figure 3.17: Shower detected by the FD site Coihueco and the high elevations extension HEAT, in which X_{max} would fall outside the field of view Coihueco alone.

3.4.3 AMIGA

AMIGA (**A**uger **M**uons and **I**nfill for the **G**round **A**rray) consists of an array of Water-Cherenkov tanks, deployed in the north-west of the array inside the regular 1.5 km grid, with accompanying scintillators buried alongside. It comprises a dense array with 61 tanks in a 750 m grid and 24 stations on a 433 m grid, which are expected to provide full efficiency detection of EAS down to $3{\times}10^{17}$ eV and $\sim 10^{17}$ eV, respectively. As the location of AMIGA is close to that of HEAT it will allow for low energy hybrid measurements. The array of scintillators is to be installed under the infill array, buried next to each of the stations. The shielding of the scintillators strongly suppresses the electromagnetic component allowing them to count muons above 1 GeV. The first seven muon detectors are being deployed in an engineering array, consisting of 30 m^2 counters, to validate the detection technique and the detector design [166].

(a) AMBER dish antenna

(b) MIDAS prototype at the university of Chigago

(c) EASIER antenna mounted on a surface detector.

(d) Tilted HEAT telescope.

Figure 3.18: Photographs of various enhancements of the Pierre Auger Observatory

4

Heavy quark production and propagation in cosmic rays

In the early days of particle physics cosmic rays were the only source of high energy particles. Still today, cosmic rays provide us with particles at energies higher than those available at accelerators, and cosmic ray detectors allow to study particles emitted in the very forward angular region. A significant number of accelerator-based discoveries were first glimpsed in cosmic ray experiments, initially considered anomalies in their interactions. A partial list of such examples might be [189]:

- the rise in the total proton-proton cross-section [190, 191, 192] and the central rapidity plateau [193] with increasing energy;
- the behavior of the mean multiplicity as a function of energy [194, 195];
- the increase in mean transverse momentum of secondaries with increasing particle and energy density [196, 197];
- the observation of scaling at the forward regions [198];
- the observation of jets [199];
- the observation of charmed particles [200].

The last point in particular was only appreciated after the charmed particles were discovered at accelerators. With the development of accelerator technology the importance of cosmic rays as a source of high energy particles greatly diminished, and to a large extent nowadays particle research is done exclusively at accelerators. In section 4.1 we review the first experiments and detections of heavy hadrons, with special emphasis on those using cosmic rays as the source of primary particles.

At accelerator energies heavy hadrons decay right after their production, and their production is of no interest. However, at cosmic rays energies this situation could be very different. In section 4.2 we review the state-of-the-art theoretical models and experimental results (section 4.2.1) of heavy quark production. On both sides there exists a gap from accelerator energies to cosmic rays. It is thus mandatory to study in detail the production, interaction and propagation of heavy hadrons in EAS. In section 4.2.2 we describe models of heavy charm and bottom production at cosmic rays energies. Section 4.3 explains the interaction model we have used. Finally, section 4.4 gives an overview of the implementation of these models in CORSIKA Monte-Carlo code.

4.1 History of charm and bottom detection in cosmic rays

There is a long tradition of looking for new particles in cosmic rays. The success of the quark model [201] in hadronic spectroscopy and scattering phenomenology stimulated the search for free quarks. Cosmic rays experiments were part of this trend as well. Between the 1960s and 1980s there were numerous searches for free quarks in the primary cosmic ray flux or as part of the secondaries produced in the interactions of UHECRs in the upper atmosphere. Comprehensive reviews as in [202, 203] describe in detail the experiments carried out during this period. Some of them reported the observation of free quarks (see for example [204, 205]), but were later disregarded as experimental failures. In addition, in later reviews [206], all detection hints were considered to be explained by fluctuations due to low-energy delayed hadrons in air showers. The lack of success in the free-quark search experiments eventually gave rise to the concept of confinement [203]: color charged particles cannot be isolated singularly, and therefore cannot be directly observed. Since the 1990s there have been only three extensive searches for free quarks in cosmic rays [207], the three of them yielding negative results [208, 209, 210].

One of the possible explanations put forward at the moment to account for the difficulty in observing free quarks was their assumed large mass. As such, free quark searches were closely related to those of massive long-lived particles, either composed of new quarks or of a different nature. The first approach to searches of heavy quarks looked for penetrating long-lived particles, using emulsion chambers as the main study technique. The experiments searching for these particles (called *plutons* at the time) were based upon the following considerations [211]:

- If heavy particles existed, they should be produced in EAS induced by primaries of sufficient energy.
- They should be produced high in the atmosphere (15-25 km), where the flux of high energy particles is larger, and due to their mass they should be delayed with respect to the arrival of the shower disk.

The experimental problem was reduced to searching for high energy particles delayed with respect to the shower front. That kind of experiment would need a detector capable of registering delayed particles penetrating a large amount of matter or releasing a large amount of energy in a single interaction or decay [211]. The delay for such a particle traveling through the atmosphere, Δt, is related to its mass and energy through [212]:

$$\begin{aligned} \Delta t &\sim \left(\frac{h}{\beta c} - \frac{h}{c}\right) \\ &\sim 1667 \left(\frac{h m^2}{E^2}\right) \text{[ns]} \end{aligned} \tag{4.1}$$

where h is the distance in kilometers traveled by the particle in the atmosphere from its production point to the detector. Delays of tens of nanosecond could be explained by particles with mass in the range 1-10 GeV, with small inelasticity, produced during the first interactions of the EAS. There was a long-standing claim of the discovery of heavy long-lived particles, with masses in the range 3-5 GeV and lifetimes $\sim 10^{-6}$-10^{-8} s [213, 214]. Even though these masses are comparable to those of some charmed and bottom hadrons (e.g. $m_{J/\Psi} \simeq 3.096$ GeV, $m_{B^+} \simeq 5.3$ GeV) the latter have lifetimes in the range $\tau \sim 10^{-12}$-10^{-14} s. For, say B mesons, to have any possibility to reach ground they should be so energetic that according to equation 4.1 no delay with respect to the shower front would be observed. As such, they were never candidates to produce this kind of delays in the detection of EAS in the first place. In addition, even at that time the interpretation of the results in terms of heavy long-lived particles was dismissed in favor of instrumental-related explanations [215].

4.1.1 Airborne emulsion chambers

The first real detection of heavy quarks (in the form of heavy hadrons) in cosmic rays came from the hand of an airplane-borne emulsion chamber experiment, in 1971, where a Japanese

group reported the observation of a possible new particle [200]. To investigate ultra high-energy showers in detail this group flew an emulsion chamber to a depth of about 260 g·cm^{-2}, finding a striking event: they observed the pair production and decay of two particles with lifetimes $\sim 10^{-13}$ s and masses in the range 2-3 GeV. One of these particles showed a sudden direction change, accompanied by a π^0 decaying into two γ rays in the forward direction, with an energy one order of magnitude larger than the rest of the γ rays in the event. Combining the tracks of the particles in the emulsion chamber, they reconstructed a 3-dimensional view of the most forward part of the event, which is shown in figure 4.1(a):

- Tracks B' and C' after the kinks were attributed to hadrons, because they crossed seven radiation lengths without any cascading.
- Tracks B, B' and the flight path of the π^0 satisfied a coplanarity condition, and the event was attributed to a 2-body decay of the parent particle B into the charged hadron B' and the π^0 meson, $B^{\pm} \to B'^{\pm}\pi^0$.
- Due to the p_t value of the B' particle, much higher than the possible maximum p_t for strange particles, the parent particle B could not have been a strange particle.
- With an estimated decay time of O(10^{-13}) s it could not have been a resonance undergoing strong decay.

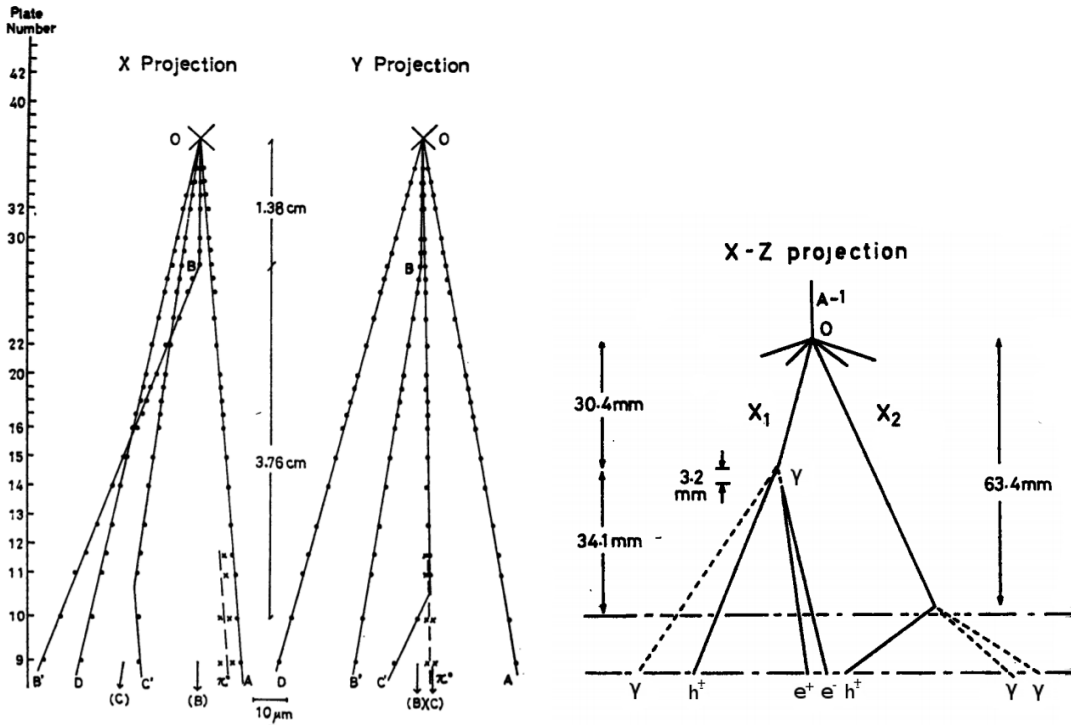
Considering all these facts, the pair production and decay of a new particle (called a X-particle at the time) was reported at the 12^{th} *International Cosmic Ray Conference*, in 1971 [216]. Assuming $B'^{\pm}$ to be a meson, the mass of the X-particle had to be about 1.8 GeV [217]. One can see that their estimates were very close to the values known for the $D^{\pm}$ mesons ($m_{D^{\pm}} = 1.869\,\text{GeV}$, $\tau_{D^{\pm}} = 1.040 \cdot 10^{-12}\,\text{s}$). Actually, $D^{\pm} \to \pi^{\pm}\pi^0$ is one of the possible decay modes of charged D mesons. The presence of very forward and energetic gamma rays, originated in the decay of π^0s, would be used by latter experiments as a smoking gun to look for X-particles. Soon after, there was many a theoretical suggestion that the existence of such new type of hadrons with a long life-time could be possible only by introducing a new degree of freedom into the hadrons, this is, the existence of a new quark [218].

The most complete pair of charm particles produced in cosmic rays was observed by a balloon-borne emulsion chamber in 1974 [219]. The emulsion chamber, designed to study nuclear interactions above 10 TeV was flown to an altitude of 10 g·cm^{-2}. Among the charged secondaries, two were considered to have decayed in flight: one of them into a charged secondary and a η meson, and the other one into a charged particle and a π^0. These decaying particles were very similar to the X-particle found in [200]. In figure 4.1(b) there is a schematic view of the event. The masses of each decaying particle were analyzed as a function of the nature of the unidentified charged hadron (h), assuming it was either a kaon or a pion. The results are shown in table 4.1. Depending on the nature of the unidentified charged hadron, $h^{\pm}$, their combined invariant

h	m_{X_1} [GeV]	m_{X_2} [GeV]
K	1.66 ± 0.42	1.74 ± 0.44
π	1.55 ± 0.38	1.59 ± 0.40

Table 4.1: Possible invariant masses of X-particles in the event depicted in figure 4.1(b), according to the nature of the accompanying hadron h.

masses were estimated between 3.8 and 4.1 GeV, The event could be due to the production of a $c\bar{c}$ resonance at the top of the emulsion chamber, that then decayed into a $D^{\pm}$ or a $D_s^{\pm}$ pair, directly or through the feed-down to less massive $c\bar{c}$ states. Sadly, the emulsion chamber did not allow for a better precision in the prediction of the masses and lifetimes of the X-particles. The predictions are consistent with current values, but one has to take into account that all the values listed in table 4.1 have relative errors of $\sim 25\%$.



(a) Event detected by Niu et al. [200] (b) Event detected by Sugimoto et al. [219]

Figure 4.1: Reconstruction of the event observed by a) Niu [200] and b) Sugimoto [219]

After the discovery of the X-particle, emulsion chambers specially designed to detect similar particles were built and exposed to proton beams of 205 GeV at Fermilab in 1975. One of the first events of a neutral X-particle (neutral charmed hadron) decaying to two charged particles is shown in figure 4.2. Averaging over the possible final states, the mass of the decaying particle was estimated to be 1.74 GeV and its lifetime 0.998×10^{-13} s, close to the mass and lifetime of the D^0 meson ($m_{D^0} = 1.86486 \pm 0.00013$ GeV, $\tau_{D^0} = 4.101 \pm 0.015 \times 10^{-13}$s). Another event, and a detailed explanation of the geometry and tracks of these events can be found in [220].

It was not until 1974 that the J/Ψ was unambiguously detected, simultaneously, in two accelerator experiments. One of them was the Brookhaven fixed target experiment AGS, in proton-Beryllium collisions [221]. The other was SLAC's SPEAR storage ring, in e^+e^- collisions [222]. None of them was actually looking for the J/Ψ state: the former was searching for the possible existence of a heavy photon, while the latter was investigating the energy dependence of e^+e annihilation into hadrons [217]. The detection of other heavy mesons and baryons at accelerators soon followed. The first observation of a single D^0 meson was reported in 1976, from e^+e^- annihilation in the Mark I detector at SLAC [223]. Also in 1976, the bottom quark was discovered in $p\bar{p}$ collisions at Fermilab. A narrow state around 9.5 GeV was observed, the Υ state, which was soon identified as a $b\bar{b}$ quarkonium state [224, 225, 226]. And the first observation of $D^0\bar{D}^0$ pair production in proton interactions was reported in 1978, exposing emulsion chambers to 400 GeV proton beams at Fermilab [227].

The discovery of the J/Ψ and the assumption that it was a composite system of $c\bar{c}$ led to the reinterpretation of many events found in emulsion chambers that had remained unnoticed [228]. Figure 4.3 shows three emulsion chamber events from experiments carried in the 1950s and 1960s that show the characteristic feature of photons with energy one order of magnitude larger than any other particle in the event, considered an indication of the presence of X-particles:

- figure 4.3(a) represents the "T-Star" event seen by the Rochester Group in 1952, during

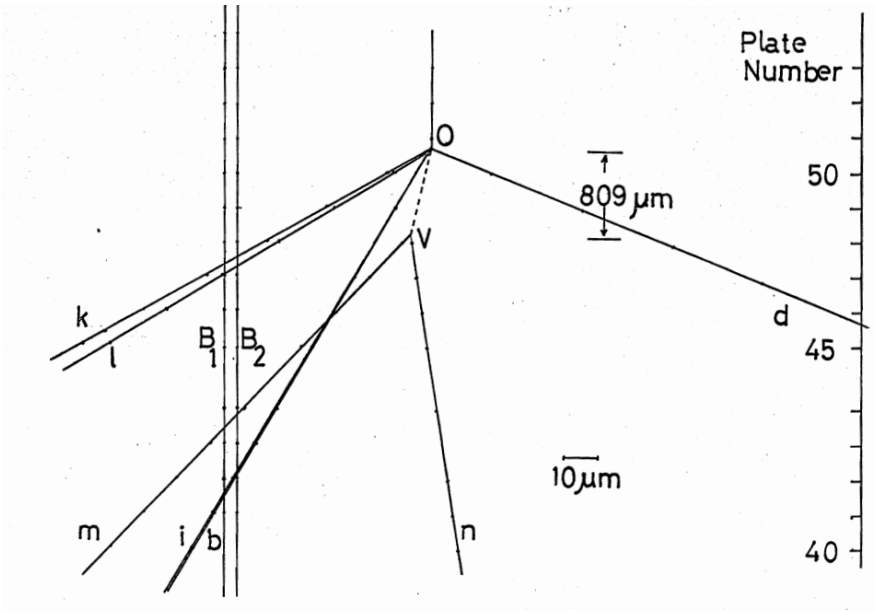


Figure 4.2: Sketch of the first detection of a massive neutral short-lived particle (X-particle) in an emulsion chamber, at Fermilab [220].

the run of their first experiment using emulsion cloud chambers to study cosmic rays [229].

- figure 4.3(b) corresponds to event "ST-2" observed in the first balloon experiment of the Japanese Emulsion Chamber Group in 1956 [230].

- figure 4.3(c) shows event "11c-34", from an experiment at the Tata Institute of Fundamental Research (Bombay) in 1965 [231].

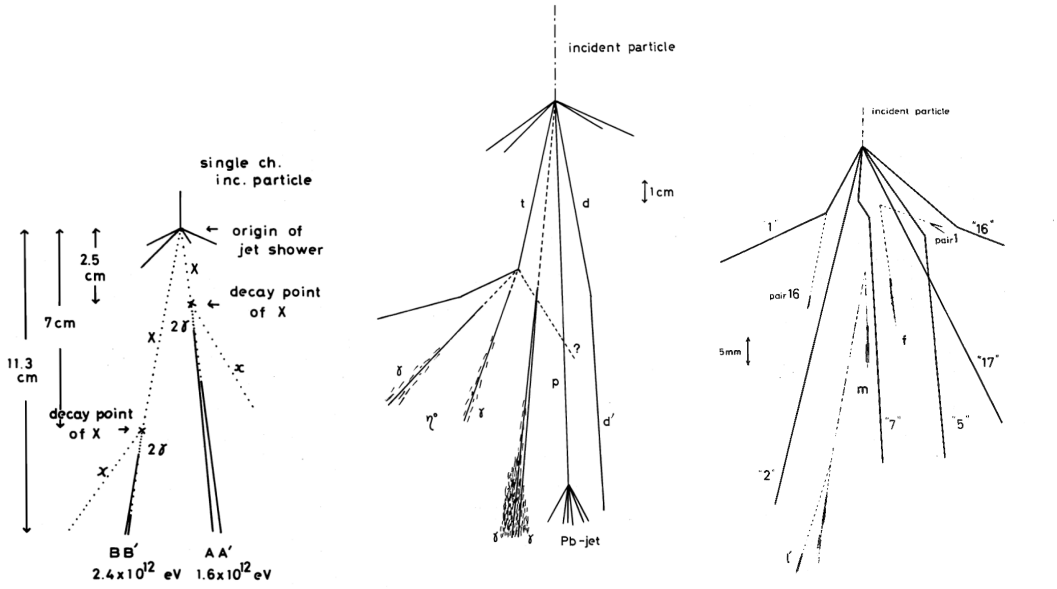
In all cases, these experiments were not planned to measure specifically the production of heavy hadrons, and any explanation in terms of charm production would be difficult without more information than that provided solely by the emulsion chambers.

It is possible that the first observation of the production and decay of open beauty hadrons occurred in emulsion chambers at high altitude. In 1979, Fumuro [232] reported 12 events induced by primaries with energies in the range 10-100 TeV. Among these events they found two which showed the decay of a number of unstable massive particles, coming from the decay of another massive short-lived particle. They were interpreted as the production of charmed hadrons from excited charmed states, or the feed-down from bottom hadrons. However, the limited information of the event made its interpretation difficult, and it joined the rest of complicated phenomena found in high-altitude emulsion chambers.

The first successful observation of beauty meson production, and its subsequent decay into charmed hadrons was reported in 1985, by the WA75 experiment at CERN, using a π beam of 350 GeV [233]. A sketch of the observed process, in which the decays $B^- \to D^0\mu^-$ and $\bar{B}^0 \to D^- h^{+1}$ were identified is shown in figure 4.4.

The first confirmed observation of the decay of a pair of beauty mesons in a cosmic ray experiment was reported by the JACEE collaboration. The JACEE (Japanese-American Collaborative Emulsion Experiment) detectors consisted of series of balloon-borne lead-emulsion chambers designed to directly measure the primary composition and spectra of cosmic rays at energies in the region of 1 TeV - 1 PeV [234]. Among the interactions in the JACEE chambers, 15 events with primary energies above 1 TeV/nucleon and secondary particle multiplicities less

[1] h^+ is a positively charged light hadron.



(a) T-Star (b) ST-2 (c) 11-c 24

Figure 4.3: Reanalysis of the 3 emulsion chambers events. Left: T-Star event, Rochester group, University of Rochester, New York (1952) [229]. Center: ST-2 event, Japanese Emulsion Chamber group, Wakayama University, Japan (1959) [230]. Right: 11-c 24 event, Tata Institute of Fundamental Research, Bombay; and the University of Bristol (1965) [231].

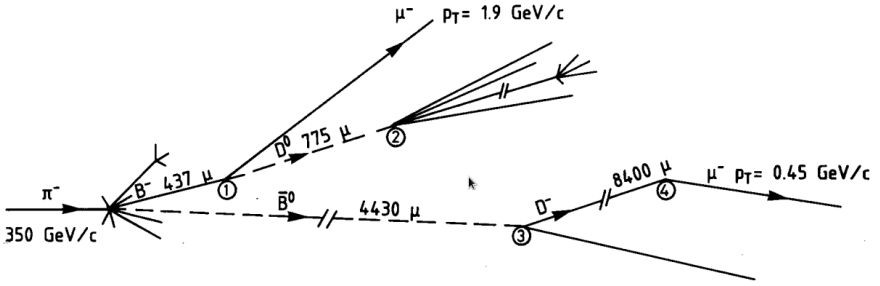


Figure 4.4: Sketch of the first observation of B meson production

than 50 were studied in detail for secondary particle decays. Two events were found that contained decays with almost identical topologies: a singly-charged particle track showing a kink, with four photons converted into e^+-e^- pairs pointing to the kink. Multiple kinks as those presented by the event shown in figure 4.5 (left) are more probably due to particle decay than to multiple interactions. The lower limit for the mass of particle 1 in this event (assuming that no other particles but the 4 photons were emitted) was 3.8 ± 0.5 GeV. With this mass, particle 1 is presumably a bottom hadron. Particle 1.1 is also a short-lived particle, probably a charmed one. It is unlikely that it was a kaon, whose decay probability on the observed path is $2{\cdot}10^{-5}$.

The second event is shown in figure 4.5 (right). The lower limit on mass of particle 1 in this case was 2.5 ± 0.6 GeV. As such, it could be either a charmed or a bottom hadron. The similarity of the decays suggests that particle 1 in the second event might also be a bottom hadron. In both events, photon conversion distances and emission angles were very well determined and there was no doubt that the four photons, in both cases, pointed to the decay vertex of the first particle, and not to the primary interaction vertex. Finally, both events were identified as a

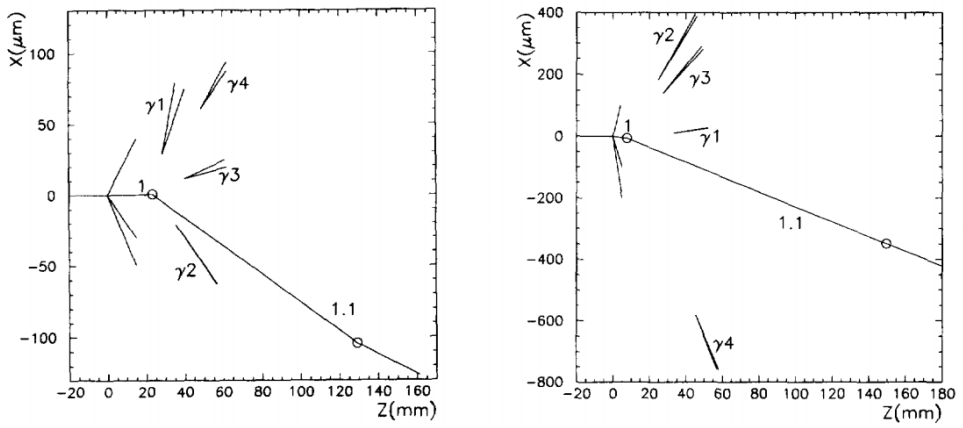


Figure 4.5: Events detected with the JACEE emulsion chambers, identified as the decay $B^- \to D_s^- \eta\eta$.

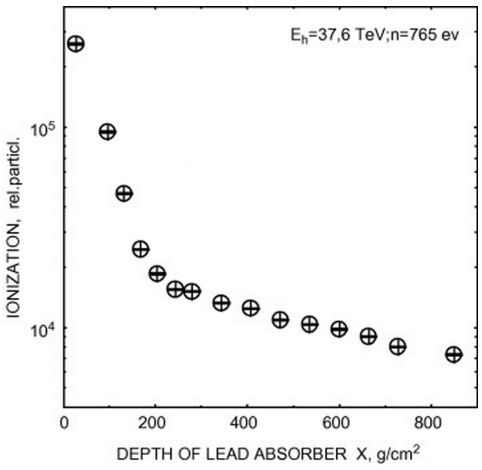
$B^- \to D_s^- \eta\eta$ decay. The importance of this observation lies in being one of the first transitions $b \to u$ detected. Until then, the V_{ub} element of the CKM matrix was assumed to be 0.

4.1.2 Ground-based observatories

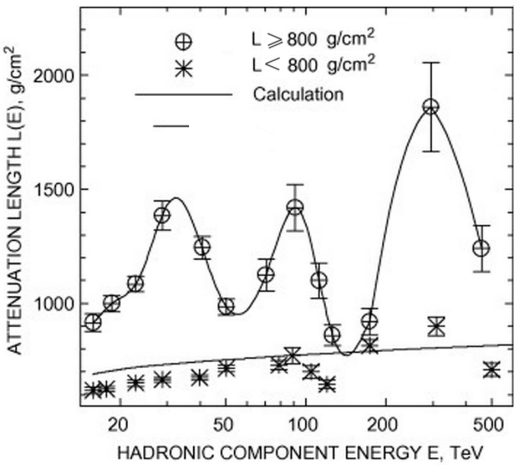
Another long standing claim of the influence of charmed hadrons in cosmic rays experiments is that of the long-flying component (LFC). As a way to study the hadronic cores of EAS and to better understand shower development, starting on the 1960s several hadronic calorimeters were deployed at high altitudes. They intended to trigger on $\sim$ PeV showers initiated right above the detector, and thus study the development of the hadronic component of the shower. Two of these projects were undertaken by the Tien-Shan and the Pamir experiments, located at the mountains of the same name, at 3440 m ($\sim$ 700 g·cm^{-2}) and 4300 m ($\sim$ 600 g·cm^{-2}) above sea level, respectively. One of their studies consisted on measuring the dependence of the shower ionization as a function of the depth traversed in a calorimeter. Unexpectedly, their results showed that the characteristic exponential behavior of the shower ionization became flatter, and thus the attenuation length longer, with rising primary energy [235]. Since the first analysis of the results (circa 1970) an explanation based on a new type of particle, with enhanced elasticity, and lifetimes $\tau \sim 10^{-12} - 10^{-13}$ s was invoked. Once the J/Ψ was discovered, the responsibility of the long-flying component was ascribed to the production and propagation of charmed particles [236].

Hadronic showers develop with a typical attenuation length of 600-700 g·cm^{-2} in lead. Figure 4.6(a) shows an average cascade of energy 37.6 TeV. The electron-photon component of EAS dominates at the depths less than 133 g·cm^{-2}, The energy of the hadronic component can be estimated as the energy released in the calorimeter at depths from 133 to 850 g·cm^{-2}. The average ionization curves are approximated by an exponential function, $\exp(-z/L)$, in the interval 344-850 g·cm^{-2}, where L is called the attenuation length. This quantity is not the real attenuation length of the shower, but rather the apparent attenuation length measured by the calorimeter.

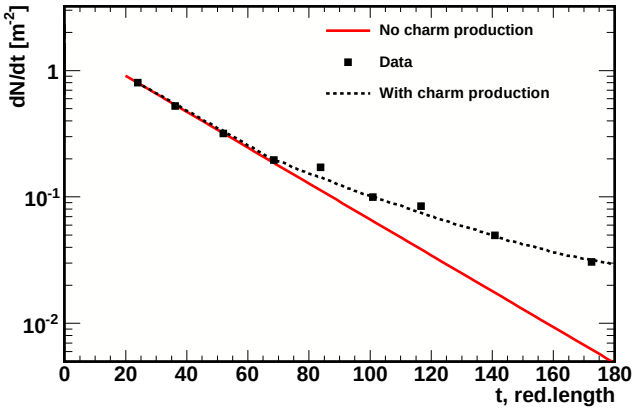
In some events charmed particles might be produced. At low energies their decay lengths are short enough and decay occurs within the calorimeter. However, with rising energy charmed particles are able to penetrate larger depths into the calorimeter before decaying. If they carry a large enough fraction of the shower energy, they delay the shower development, increasing the attenuation length. At even higher energies a charmed hadron decay length will become of the order of the calorimeter size, crossing it without decaying, and the attenuation length will decrease again.



(a) Tien-Shan: average shower profile.



(b) Tien-Shan: energy dependence of the attenuation length.



(c) Pamir: distribution of shower absorption depth.

Figure 4.6: a) Average profile of showers with energy 37.6 TeV measured in the iron calorimeter of the Tien-Shan experiment [238]. b) Energy dependence of the attenuation length for showers with average attenuation length above (circles) and below (stars) 800 g·cm^{-2}, measured by the Tien-Shan experiment [238]. c) Distribution of shower absorption depth measured by the Pamir experiment. The solid (dashed) line corresponds to the prediction without (with) charm production [239].

In particular, in the Tien-Shan experiment, showers were separated into two samples (see figure 4.6(b)): one with average attenuation length (L) larger than 800 g·cm^{-2} and the other with average attenuation length smaller than 800 g·cm^{-2}. The most remarkable difference between the two groups lies in the energy dependence of the attenuation length: those showers with L below the average (stars) show a mild dependence with energy , whereas showers with L above the average (circles) present an irregular behavior with pronounced maxima. The showers were thoroughly analyzed to ensure that this effect was not simply due to showers starting deep into the calorimeter. The interpretation of the results relied on a shower component carrying off the energy deep into the cascade. $D^{\pm}$ and D^0 mesons, and Λ_c baryons have different lifetimes and reach decay lengths larger than the calorimeter size at increasingly larger energies. They were thought to be responsible of the oscillating behavior of the attenuation length with energy of elongated cascades [237, 238].

Regarding the Pamir experiment another unusual but related phenomenon was observed.

The absorption curve for hadrons with energies greater than 6.3 TeV was obtained in an extremely deep uniform lead chamber 120 cm thick. The distribution of showers absorbed within 70 radiation lengths follows an exponential law with index $\lambda = (200 \pm 5)$ g·cm^{-2}. However, at larger depths (> 70 radiation lengths), the absorption length of hadrons in lead is different and λ becomes as high as $\lambda = (340 \pm 80)$ g·cm^{-2}. Figure 4.6(c) shows the experimental distribution of absorption depths (black squares), the prediction assuming no charm production (red solid line) and the prediction allowing for charm production (black dashed line). The interpretation of the measurements was that the excess of hadrons at larger depths could be due to large fractions of charmed Λ_c and D-mesons produced in the upper layers of lead of the emulsion chamber. The decay path for charmed particles with energy ≥ 6.3 TeV is about 1 m, which is comparable with the emulsion chamber depth [239].

The idea of actively looking for the effect of heavy hadrons in ground-based high energy cosmic rays detectors was first formulated by Stodolsky and McLerran in 1982 [11]. Based on their smaller interaction cross-sections and the enhanced elasticity of their interactions, they suggested that charmed hadrons would be more penetrating in passing through matter than pions and protons. As general conditions for their observation they argued that given that charmed particles were produced in the first collision at the top of the atmosphere they may carry away a significant fraction of the shower energy. This energetic and relatively penetrating component may then lead to an elongated hadron shower with, perhaps, a burst at its end when the particle decays. Manifestations of this would be a "long flying" component and an increase in multiplicity and its fluctuations, due to the decay of the unstable particle. Generally speaking, these properties would apply to any hadron with a small interaction cross-section and sufficient lifetime to traverse the atmosphere.

4.1.3 Prompt lepton fluxes

Atmospheric fluxes of muons and neutrinos at low energies have been extensively studied. They arise mainly from decay products of charged pions and kaons. Up to about 1-10 TeV they give rise to the *conventional* atmospheric lepton flux. Shorter lived hadrons are also produced at high energies, which also contribute to the lepton flux. Among the latter, the main contribution comes from the decay of charmed mesons, such as $D \to K + \mu + \nu$ and $\Lambda_c \to \Lambda_0 + \mu + \nu$. Because their decays occur right after their production they are called the *prompt* component of the flux. As their energies increase, the decay lengths of πs, Ks and Ds become longer than their interaction lengths in the atmosphere, suppressing the production of neutrinos and muons. This effect becomes patent for light mesons at much lower energies than for charmed mesons. At energies below ~ 10 TeV, the prompt contribution is hidden by the much more abundant contributions from decay of charged pions and kaons. Above ~ 100 TeV, the semileptonic decay of very-short lived charmed particles, with a harder energy spectrum, becomes the dominant atmospheric leptonic component, despite their low production rate.

The first publications on the subject, by Hinchliffe, Lewellyn Smith, Lederman and Cronin, date back to 1975 [240, 241, 242, 243]. An extensive review of atmospheric muon fluxes (both conventional and prompt), including a review of charm production and its relation to prompt muons can be found in [244]. More recent publications about prompt muon and neutrino fluxes and their role and detection in EAS can be found in [245] and [246]. Experiments such as IceCube are reaching sensitivities close to the theoretical expectations of prompt neutrino fluxes [247, 248].

4.2 Models for Heavy quark production

In the previous sections we have seen that cosmic rays experiments hinted at the existence of massive short-lived particles, even though hindered by small exposures and fluxes. Nevertheless, those experiments were able to flag features difficult to explain if one neglects the presence of heavy hadrons in EAS. In a detector such as the Pierre Auger Observatory, direct detection of heavy hadrons is out of reach. Neither the fluorescence nor the surface detector have access to

the particles produced during the first interaction, but rather to their products many interactions afterwards. Any chance of inferring the presence of heavy hadrons in EAS is tied to the analysis of their effect in the shower development, in a way similar to that pointed by [11].

At accelerator energies the propagation of heavy hadrons is of no interest because they decay much before they can reach the calorimeters. In table 4.2 we can see that charmed and bottom hadrons have mean lives much shorter than those of light hadrons as π, K or Λ. Last column in table 4.2 shows the critical energy above which the decay length of these hadrons becomes larger than their interaction length. While propagating in the atmosphere with energies above the critical one they would interact with air nuclei rather than decay, and a new propagation regime could appear.

Particle	Quark content	Mass [MeV]	$c\tau$ [µm]	ε_{crit} [GeV]
π^+	$u\bar{d}$	139.57018 ± 0.00035	$(7.810 \pm 0.002) \cdot 10^6$	228.9
K^+	$u\bar{s}$	493.677 ± 0.016	$(3.714 \pm 0.006) \cdot 10^6$	1702.32
Λ	uds	1115.683 ± 0.006	$(7.90 \pm 0.06) \cdot 10^4$	1.8 10^5
D^+	$c\bar{d}$	1869.62 ± 0.15	312 ± 2	$7.67{\cdot}10^7$
D^0	$c\bar{u}$	1864.86 ± 0.13	123 ± 1	$1.94{\cdot}10^8$
Λ_c	udc	2286.46 ± 0.14	60 ± 2	$4.88{\cdot}10^8$
D_s^+	$c\bar{s}$	1968.50 ± 0.32	150 ± 2	$1.68{\cdot}10^8$
B^+	$u\bar{b}$	5279.26 ± 0.17	492 ± 2	$1.37{\cdot}10^8$
B^0	$d\bar{b}$	5279.58 ± 0.17	455 ± 2	$1.48{\cdot}10^8$
Λ_b	udb	5619.4 ± 0.6	429 ± 7	$1.67{\cdot}10^8$

Table 4.2: Properties of light baryons, open charm and bottom hadrons. Antihadrons are not listed.

4.2.1 Charm and bottom at accelerator energies: theory and data

The production of heavy quarks at accelerator energies is currently explained by perturbative Quantum Chromodynamics (pQCD) [249]. Perturbative mechanisms rely on the small value of the strong coupling constant, α_s, at high energies and short distance interactions. Examples of models and Monte-Carlo codes based on pQCD are the Dual Parton Model (DPM) [250], QGSJET (quark-gluon-string model with jets) [251, 252] or SIBYLL [253]. At leading order (LO) heavy flavor production in hadronic collisions can only be produced in quark-quark annihilation and gluon fusion [254] (see figure 4.7). Calculations of pQCD processes can be factorized into three different parts:

- The non-perturbative initial conditions (describing the state before the collision), determined by the fractional momenta x the interacting partons of the colliding hadrons carry. These distributions are parameterized as parton distribution functions (PDFs).

- The hard process itself, described at LO by the partonic cross-section:

$$\hat{\sigma}_{ij}^{LO}(\hat{s}, m_Q^2, Q^2) = \frac{\alpha_S^2(Q^2)}{m_Q^2} \cdot f_{ij}^{0,0} \cdot \left(\frac{m_Q^2}{\hat{s}}\right), \; ij = q\bar{q},\, gg \tag{4.2}$$

  It depends on the mass of the heavy quark, m_Q, the strong coupling constant evaluated at scale Q^2, $\alpha_s(Q^2)$, and $\hat{s}$, the squared partonic center of mass energy. $f_{ij}^{0,0}$ is a dimensionless scaling function determining the energy dependence of the heavy flavour production cross-section. Then, the cross-section to produce a heavy quark pair in a proton-proton collision results from the convolution of the perturbative partonic cross-section with the parton

distribution functions of the interacting hadrons:

$$\sigma_{Q\bar{Q}}^{PP} = \sum_{i,j} \int dx_1\, dx_2\, f_i^p(x_1, Q^2)\, f_j^p(x_2, Q^2)\, \hat{\sigma}_{ij}^{LO}(\hat{s}) \tag{4.3}$$

- Hadronization or fragmentation of heavy quarks into heavy-flavor hadrons. Two cases are distinguished:
  - formation of open heavy hadrons, in which the heavy quark and antiquark individually fragment into hadrons. Fragmentation functions are extracted from e^+e^- collisions data [255]. The fragmentation of a heavy quark does not depend on the mechanism by which this quark was produced [256].
  - typically 1-2% of the produced heavy quark-antiquark pairs form a bound quarkonium state instead of a pair of hadrons with open heavy flavor.

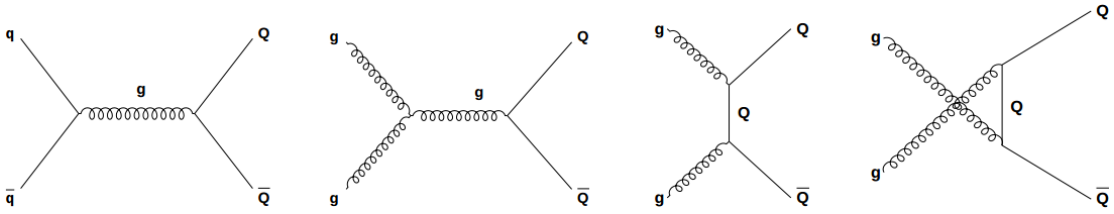


Figure 4.7: Heavy flavor production mechanisms at leading order.

Experiment/Collaboration	Reaction	Beam E(GeV/c)	σ_{D,D_s,Λ_c}(µb)
NA32/ACCMOR col. at CERN SPS	p-Si	200	1.5 ± 0.7
E769/Fermilab	p-(Be,Cu,Al,W)	250	8.8 ± 1.5
WA89/CERN SPS	Σ^--(Cu,C)	340	5.3 ± 0.4
Experiment/Collaboration	Reaction	Beam E(GeV/c)	σ_D(µb)
NA27/LEBC-EHS col.	p-p	400	15.1 ± 1.5
E743/LEBC-MPS col.	p-p	800	24 ± 6
E653/Fermilab	p-emulsion	800	38 ± 10
HERAB	p-(C,Ti,W)	920	48.7 ± 4.7
NA27/LEBC-EHS	π^--p	360	12.6 ± 2.2
NA32/ACCMOR col. at CERN SPS	π^--Si	200	4.1 ± 0.6
NA32/ACCMOR col. at CERN SPS	π^--Cu	230	7.6 ± 1.1
E653/Fermilab	π^--emulsion	600	24.6 ± 4.3
E769/Fermilab	π^--(Be,Cu,Al,W)	210	6.4 ± 0.8
E769/Fermilab	π^--(Be,Cu,Al,W)	250	9.4 ± 0.7
WA92/CERN Ω' spectrometer	π^--(Cu,W)	350	8.8 ± 0.5
E791/Fermilab	π^--(C,Pt)	500	15.4 ± 1.8

Table 4.3: Charm production total cross-sections summary [257].

Charm and bottom total and differential cross-sections have been measured in a wide range of energies and reactions. In table 4.3 there is a summary of the most relevant fixed-target experiments involved in the study of heavy quark production and their results [257]. To select events with charm or bottom, experiments usually employ high-resolution detectors in the target region, to observe primary and secondary vertices, signaling the decay of the heavy flavored hadrons. Bubble chambers, emulsions and silicon tracking telescopes are the detectors commonly used. The latter have a poorer spatial resolution, but can operate at much higher interaction

rates (a must when looking for rare processes [254]). Most of the experiments had particle tracking devices and muon spectrometers. Experiments measuring charmed hadrons in hadronic decay channels used calorimeters and particle identification detectors, to distinguish between pions, kaons and protons.

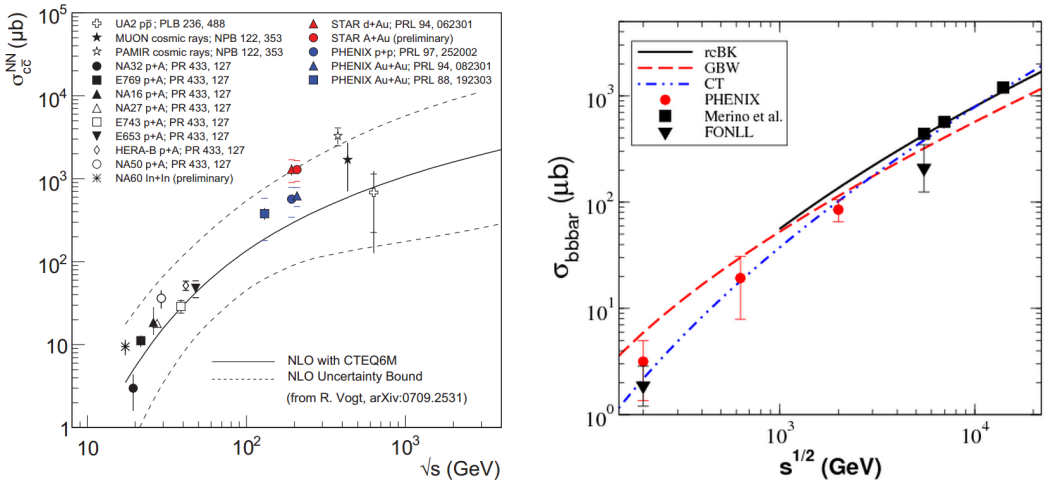


Figure 4.8: Data on charm and bottom production from several experiments. Captions from [258] (left) and [259] (right).

The measurement of heavy-flavor production in pp collisions provides a crucial testing ground for pQCD calculations. The state-of-the-art calculations with pQCD describe well the production cross-sections of open heavy-flavor hadrons measured at RHIC ($\sqrt{s}$ = 0.2 and 0.5 TeV), Tevatron (1.96 TeV) and LHC (2.76 and 7 TeV) [256, 260, 261, 262, 263]. A compilation of data on $c\bar{c}$ and $b\bar{b}$ production from various experiments can be seen in figure 4.8. Extensive reviews of heavy-hadron production in pp, pA and AA collisions, with detailed analyses of individual experiments can be found in [254, 256, 258, 264].

4.2.2 Extrapolation to EAS energies

In cosmic rays collisions we are confronted with the problem of the energy scale, i.e. we study collisions whose center of mass energies are of the order of 100 TeV. The agreement between calculations and data at accelerators make extrapolations to energies somewhat higher than those at accelerators, where data are unavailable, safer. However, accelerator data are several orders of magnitude below the energies of the primary particles in cosmic rays showers, and extrapolation over large ranges of energy, prone to uncertainties as large as ∼40% [265], are not reliable.

To circumvent this difficulty with rising energies, analyses based on effective theories are derived. The solution to this problem is not unique, and different regimes allow for different approaches, which in turn offer qualitatively correct results under certain limits. The Dual Parton Model [250], the Lund Fragmentation Model [266] and the Color Glass Condensate [267] are some of the most well known effective theories.

4.2.3 Color Glass Condensate

The most complete QCD-based effective theory that describes the physics of hadronic interactions at very high energies is the Color Glass Condensate model [267]. At the highest energies, it allows for a unified description of various high-energy phenomena ranging from deep inelastic scattering to heavy-ion, or proton-proton collisions, and to cosmic rays interactions [268]. It

has a firm theoretical basis, supported by its analytic equivalence with the gluon-gluon fusion mechanism of the parton model [269].

In this model, a heavy flavor quark-antiquark pair ($Q\bar{Q}$) is created through the fluctuation of a gluon from the projectile hadron. Upon the interaction with the target, these heavy quarks are released, and charmed and bottom hadrons are formed from hadronization of those heavy quarks with sea quarks. This hadronization mechanism is called Uncorrelated Fragmentation, and all heavy hadrons have the same probability of being formed from the heavy quarks produced. We assume that hadronization occurs without energy loss, and thus the differential production probability for charmed (bottom) hadrons is identical to that of charm (bottom) quarks. Those distributions can be seen in figure 4.9 (left), both scaled to the same integral. The average ($\pm 1\sigma$ deviation) energy carried by charm quarks, as a percentage of the primary energy is $< x_c >= 3.32 \pm 1.87$ %. Regarding bottom quarks, on average they are produced with a fraction of the primary energy equal to $< x_b >= 2.74 \pm 1.74$ %. A more technical discussion of the Color Glass Condensate model can be found in [259, 270, 271]. In figure 4.10 we show a comparison of data with the predicted cross-sections from the Color Glass Condensate model both for charm (solid line) and bottom (dashed line) production. The model shows a reasonable agreement with the set of current data. The Color Glass Condensate describes diverse phenomena, not only those related to heavy quark production. A thorough review of recent results and developments of the model can be found in [272]. A brief summary of experimental data explained by the model could be the following:

Color Glass Condensate

- Fit of HERA data on electron-proton deep inelastic scattering at very small x [273].
- Description of particle multiplicity, and its rapidity and centrality dependence in heavy ion collisions at RHIC [274, 275, 276].
- Nuclear modification factor measured by BRAHMS, PHENIX, PHOBOS and STAR at RHIC in deuteron-Au collisions [277].

4.2.4 Intrinsic Quark production

Not all features present in data are explained by the Color Glass Condensate model. According to the factorization theorem of QCD, quarks hadronize independently of the initial state [278]. As a result, the structure functions of quarks in the nucleon $q(x)$ (probability of finding a quark with a certain fraction of the nucleon energy) should have certain properties:

- $\bar{u}(x)$ and $\bar{d}(x)$ distributions should be identical.
- $s(x)$ and $\bar{s}(x)$ should be the same and fall off fast with rising x.
- c and $\bar{c}$ quarks should be produced with identical energy distributions.

However, measurements of deep inelastic electron and neutrino scattering show that these predictions are wrong. Experiments show that for sea quarks $\bar{u}(x) \neq \bar{d}(x)$ [279]. The HERMES collaboration measured the strange quark momentum distribution in the proton finding a fast falling distribution up to $x \sim 0.1$ followed by a flat component in the range $0.1 < x < 0.5$ [280]. And in $\pi^-(\bar{u}d)$ interactions with hadrons or nuclei, D^- ($\bar{c}d$) carry on average a larger fraction of energy than D^+ ($c\bar{d}$) [18, 281]. These discrepancies between perturbative QCD and the charm hadroproduction data suggest the presence of another $Q\bar{Q}$ production mechanism important at large x and low p_T [282, 283].

To explain these discrepancies theoretical models coincide in invoking a charm or bottom component inside the nucleon. The Meson-Cloud model [284], the Recombination Mechanism [285], or the Intrinsic Quark mechanism [18, 19, 20] are examples of these models, yet the nature

and evolution of the heavy component differs between them. They produce similar results, but we will focus on the last model, which provides a simple mechanism for producing the flavor correlations present in data, and is also able to accommodate some other discrepancies.

Heavy quarks are produced by the processes $q\bar{q} \to Q\bar{Q}$ and $gg \to Q\bar{Q}$. When these heavy quarks arise from fluctuations of the initial state, its wave function can be represented as a superposition of Fock state fluctuations:

$$|h>= c_0|n_v> + c_1|n_v g> + c_2|n_v q\bar{q}> + c_3|n_v Q\bar{Q}> ... \tag{4.4}$$

where $|n_v>$ is the hadron ground state, composed only by its valence quarks. When the projectile scatters in the target the coherence of the Fock components is broken and the fluctuations can hadronize, either with sea quarks or with spectator valence quarks. The latter mechanism is called Coalescence. For instance, the production of Λ_c^+ in p-N collisions comes from the fluctuations of the Fock state of the proton to $|uudc\bar{c}>$. The co-moving heavy and valence quarks have the same rapidity in these states but the larger mass of the heavy quarks implies they carry most of the projectile momentum. Heavy hadrons formed from these states can have a large longitudinal momentum and carry a large fraction of the primary energy [286]. The differential energy fraction distribution for charmed and bottom hadrons, averaged over all possible final states, can be seen in figure 4.9 (right). On average they carry a $34.4 \pm 18.6\%$ of the proton energy when they are produced.

In appendix C we have included a thorough description of the model. Nevertheless, detailed explanations of the model along with theoretical expressions for the differential cross-sections can be found in [18, 282, 287] for intrinsic charm production and [281] for intrinsic bottom production. Unfortunately, so far no experiment has measured the cross-section for Intrinsic Quark production as a function of energy. The region of the phase space most sensitive to this production model is that of large values of rapidity and small transverse momentum, which is not readily accessible at accelerators. There are, however, experimental measurements that current QCD cannot account for, but that can be accommodated assuming an Intrinsic Quark component inside the nucleus. What follows is a selection of some of these measurements:

Intrinsic Quark model

- The anomalous growth of the $p\bar{p} \to \gamma cX$ inclusive cross-section observed by the D0 collaboration is interpreted as the underestimation of the charm distribution at $x > 0.10$ [288, 289].

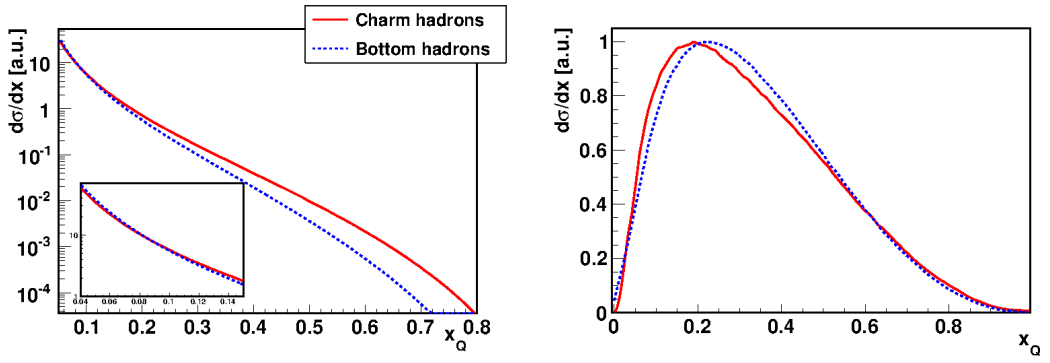
- The EMC collaboration data show an excess of events in the charm quark distribution at $x > 0.3$, at a rate at least one order of magnitude beyond predictions [290]. According to [291, 292] an intrinsic charm component, with probability $\sim$1%, is needed to fit the data.

- Production of pair of J/Ψ mesons or doubly-charmed baryons at large x and low p_t can be interpreted in terms of the presence of an intrinsic $c\bar{c}$ component. [17]:

  - Pairs of J/Ψ at high combined $x = 0.66 \pm 0.15$ have been observed by the NA3 experiment [16].
  - The SELEX collaboration reported the observation of the doubly charmed baryon $\Xi_{cc}^+(3520)$ with average $< x > \sim 0.33$ [293].

- Production of leading charm and bottom hadrons in $pp \to DX$ [12], $pp \to \Lambda_c X$ [13, 14], $pp \to \Lambda_b X$ [15], $\pi^{\pm} N \to D^{\pm} X$ and $pN \to D^{\pm} X$ [287, 19]. In the two last processes, an intrinsic $c\bar{c}$ production cross-section of $\sigma_{ic}(\pi N) \simeq 0.5$ µb and and $\sigma_{ic}(pN) \simeq 0.7$ µb at 200 GeV is sufficient to explain both the magnitude and kinematic dependence of the leading component.

- The branching ratios of the processes $J/\Psi \to \rho\pi$ and $D \to \Phi K$ are much larger than expected. Additional decay channels mediated by the presence of intrinsic $c\bar{c}$ are likely solutions to the discrepancy between predictions and experiments [294, 295].



(a) **Color Glass Condensate** (b) **Intrinsic Quark**

Figure 4.9: a) Differential fraction of primary energy carried by heavy hadrons produced in the Color Glass Condensate model. The mean values for each distribution are ≈3%. The inset zooms the region where the heavy quarks carry a small fraction of the initial energy. b) Distribution of the fraction of primary energy in the Intrinsic Quark production model for charmed and bottom hadrons, averaged over all final states. The mean value of the distributions is ≈34%

- The SELEX collaboration measured the scaling of hadroproduction cross-sections from proton to nuclear targets, $A^{\alpha(x)} \simeq \sigma_{hN}/\sigma_{hA}$. According to the conventional quark fragmentation α should decrease with rising x, but it was found to rise instead. Production of intrinsic $c\bar{c}$ pairs to a level of the 1% is able to account for this dependence [296].

Several studies [296, 297, 298, 299, 300, 301] have estimated the probability to find Intrinsic Charm in the proton. All of them quote values between 0.5% and 3.5%.

As mentioned in section 4.1.3, charm production could play an important role in the spectrum of atmospheric lepton fluxes. Discussions of the influence of an intrinsic charm component in the fluxes of muons and neutrinos measured at ground can be found in [244, 302, 303, 304].

4.2.5 Comparison between models

The implemented models show a conspicuous difference: the fraction of the initial energy heavy quarks carry away. The available phase space is similar; however, as shown in figure 4.9 (left), values smaller than 5% are favored in the Color Glass Condensate model, and the probability of carrying larger fractions falls off very rapidly. For the Intrinsic Quark model (figure 4.9 (right)), fractions of the primary energy ∼30% are the most probable values and larger fractions are not unlikely. This difference is something to be expected given the rather different initial assumptions the two models rely upon.

Heavy hadrons with small fractions of the primary energy will hardly affect the development of the EAS where they are produced for two reasons: at smaller energies, the probability of decaying before interacting rises; in addition, since the EAS size is proportional to energy, the contribution of a heavy component with a small fraction of the primary energy will be comparatively very small in regard to the total size of the shower. We can expect that heavy hadrons produced in the Color Glass Condensate model will have, on average, a modest effect, if any, on the shower development. On the contrary, heavy hadrons produced in the Intrinsic Quark model, given that they carry on average a much larger fraction of the primary energy, will noticeable affect the shower observables related to the longitudinal development through the atmosphere. In this analysis, we will assume that if heavy hadrons are detected in EAS, this can only be due to production through the Intrinsic Quark model. Showers with production of heavy hadrons produced through the Color Glass Condensate model will be indistinguishable from proton-initiated showers where no heavy quarks are produced.

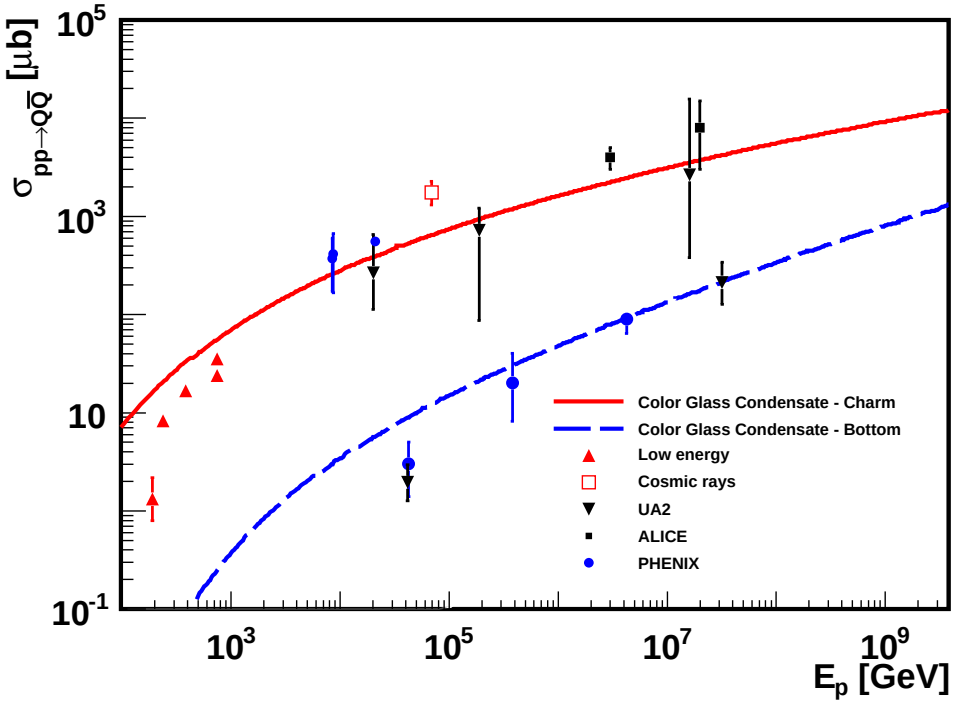


Figure 4.10: Charm and bottom production cross-sections in proton-proton collisions. The solid (dashed) line corresponds to the predictions of the Color Glass Condensate model for charm (bottom) quarks production. Experimental data from [259, 270, 305].

4.3 Interactions of heavy hadrons with air nuclei

The realistic implementation of heavy hadron propagation in EAS needs, apart from the values of cross-sections and interaction lengths, the elasticity distributions of their interactions. In accelerators, heavy *quarks* energy losses are studied while they propagate in dense media[2] before hadronization occurs. However, we deal with a rather different problem: the energy losses of heavy *hadrons* in collisions with atmospheric nuclei, the atmosphere being a much more diluted medium. Since we cannot extrapolate from accelerator conditions to study the interactions of heavy hadrons with air we will apply a detailed treatment of these interactions instead.

We use the modifications of the Monte-Carlo code PYTHIA [306] described in [9, 10, 307] to simulate the collisions of charmed and bottom hadrons with protons and air. PYTHIA distinguishes two types of interactions: diffractive processes where the two hadrons as a whole exchange momentum, and non-diffractive or partonic processes, where gluons are exchanged between the partons in the colliding hadrons. After the collision we look for the energy carried by the leading heavy hadron, and compute the elasticity as the ratio of this to the initial energy of the hadron containing the heavy quark. The emerging heavy hadron needs not to be of the same species as the incoming heavy hadron. Thus, particle type transitions are permitted during the collision.

4.3.1 Diffractive processes

A heavy charm or bottom hadron H_Q can be modeled as a proton or a pion but containing a heavy core of mass $m_c = 1.27$ GeV or $m_b = 4.7$ GeV, respectively. A sketch of a heavy hadron

[2]Such as the Quark Gluon Plasma formed in the collisions of heavy nuclei.

in this picture can be seen in figure 4.11. In a long-distance collision (where the q^2 transferred is small) with an atmospheric nucleon, the heavy core will not take part on the collision, as only interactions with $q^2 \geq m_Q^2$ can resolve it. The momentum exchanged in the collision through non-perturbative processes is assumed to depend only on the light degrees of freedom in H_Q. The light degrees of freedom carry a fraction

$$w = (m_{H_Q} - m_Q)/m_{H_Q} \tag{4.5}$$

of the energy E of H_Q. In a diffractive process H_Q will be seen by the target as a light hadron of energy wE. To estimate the momentum q^μ absorbed by H_Q in the process, a proton (for $H_Q = \Lambda_c$) or a pion (for $H_Q = D^+$) of energy wE is used as a projectile and q^μ is assumed to be the same when the incident particle is a charmed hadron. In the case of bottom hadrons, a proton is simulated for a Λ_b^0, whereas a pion is simulated for a $\bar{B}^0$.

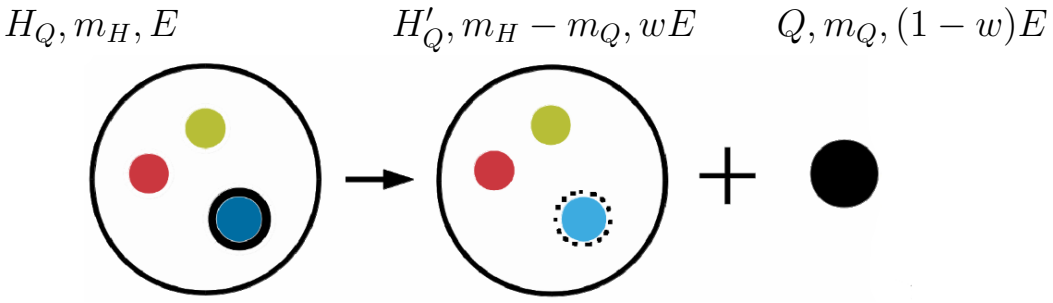


Figure 4.11: Sketch of the behavior of a heavy hadron in a diffractive process. The heavy hadron (left) is effectively seen as a light hadron with energy wE (center). The heavy core (right) does not take part in the interaction.

After the momentum transfer, H_Q becomes a diffractive system with mass M, depending on the value of q^μ. Different values of M generate different processes:

1. If no diffractive system is formed, the projectile does not fragment, behaving as in an elastic interaction.

2. If the diffractive mass M is smaller than the primary mass plus 1 GeV, a two body decay takes place.

3. If the diffractive mass M is greater than the primary mass plus 1 GeV, the system evolves into a string stretching between a quark and a diquark. Quark fragmentation and the decay of baryonic resonances results then into a leading baryon plus a series of hadrons of lower energies.

4. Higher values of the diffractive mass increase the multiplicity of the final state and reduce the elasticity of the leading baryon.

All these processes are treated using PYTHIA, yielding different values of the collision elasticity.

4.3.2 Partonic collisions

Non-diffractive processes dominate the inelastic cross-section in PYTHIA simulations. The heavy hadron H_Q is modeled as a system with the same parton content as the corresponding proton (for H_Q=Λ_c, Λ_b^0) or π (for $H_Q = D^+, \bar{B}^0$), changing a valence quark u for the charm quark c, or a valence quark d for the bottom quark b, respectively. Figure 4.12 shows a schematic of the heavy hadrons modeling for partonic processes. As in section 4.3.1, a heavy hadron with energy E is associated to a light hadron of energy wE. Instead of changing the whole hadron for its charmed (bottom) counterpart, it is the u (d) valence quark that is changed. If it carries

a fraction x of the light hadron momentum, it is replaced with the correspondent heavy quark with a fraction x_Q of the heavy hadron momentum, where

$$x_Q = \frac{m_Q}{m_{H_Q}} + \frac{m_{H_Q} - m_Q}{m_{H_Q}}\, x \tag{4.6}$$

Thus, the excess of energy is carried solely by the heavy quark, and the light partons shared by H_Q and the light hadron carry the same amount of energy.

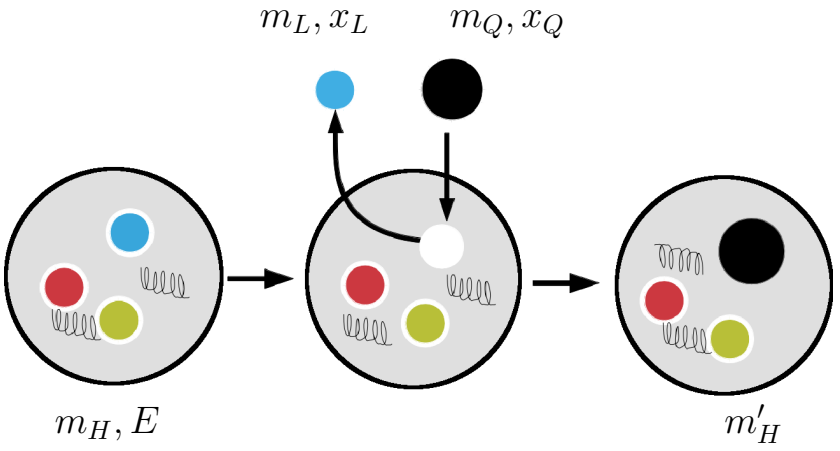


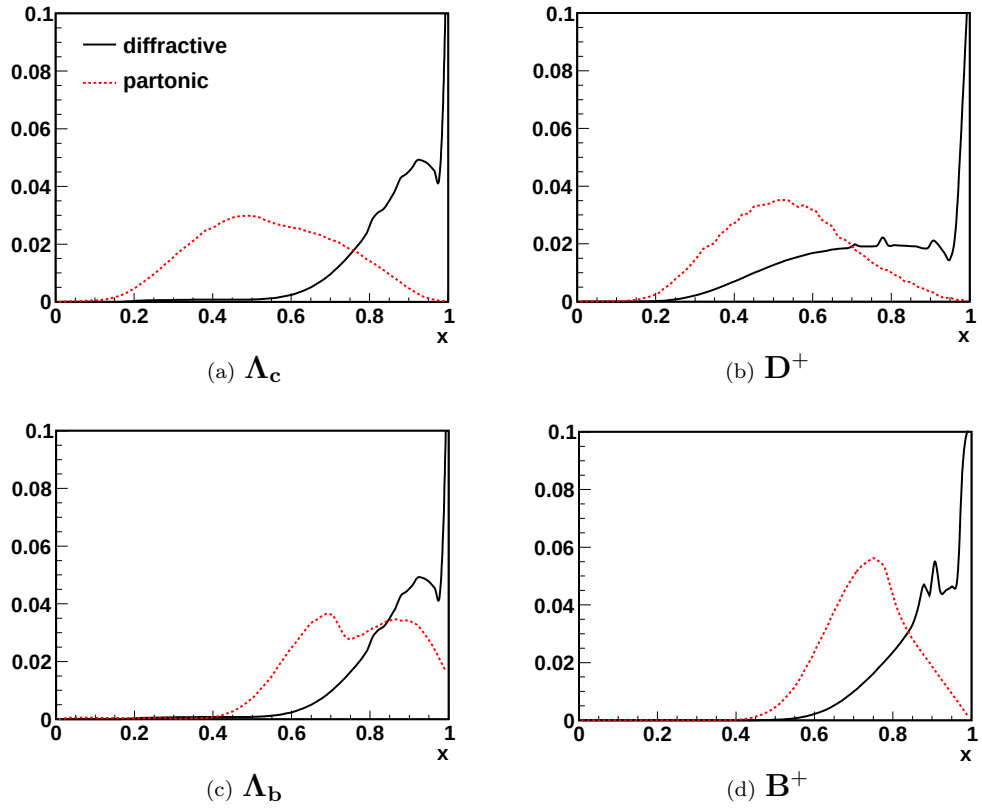
Figure 4.12: Schematic of the modeling of a heavy hadron during partonic processes. A light hadron (left) is used to simulate by a heavy hadron (right), replacing the corresponding light quark by a heavy quark (center).

We use these two models to obtain the elasticitiy distributions for charmed and bottom hadrons in diffractive and partonic collisions with protons and their associated cross-sections. Charmed and bottom hadron collisions with protons are simulated in the energy range $10^{17} - 10^{20}$ eV. Figures 4.13(a) and 4.13(b) show the distributions of the energy fraction taken by the leading hadron in $\Lambda_c p$ and Dp diffractive and partonic collisions, respectively. Figures 4.13(c) and 4.13(d) show the same processes but for $\Lambda_b p$ and Bp collisions. We find that the elasticity distributions are very weakly dependent on the projectile energy, with almost no changes in the energy range simulated. The total inelastic cross-section for each projectile is the sum of the partonic and diffractive cross-sections. The diffractive cross-section is ∼30% of the total cross-section for Λ_c collisions. For D, Λ_b and B this value is 32%, 26% and 29%, respectively. The cross-sections as a function of the projectile energy in collisions with protons can be seen in figure 4.14(a).

4.3.3 Hadron-nucleus collisions

Up to this point we have only considered collisions with protons but, during their propagation in the atmosphere, heavy hadrons will interact with air nuclei. PYTHIA deals with hadron-nucleon collisions based on the Lund string model, but so far there is no agreement on how hadron-nucleus collisions should be treated within this model. There is ample literature dealing with this problem (see for example [308, 309, 310, 311]), but we will use the method described in [312], which is the approach SIBYLL takes.

Considering hadron-nucleon collisions, after the collision we find a leading hadron, carrying the largest fraction of the primary energy, and a series of secondary particles, sharing the rest of the energy. In a simplistic setting we could picture a hadron-nucleus interaction as a series of independent hadron-nucleon collisions with every nucleon composing the nucleus. As a result, the elasticity of the final leading hadron would decrease and the number of low energy secondaries produced would rise with each consecutive collision. Even though this is the behavior we expect,



(a) $\mathbf{\Lambda_c}$ (b) $\mathbf{D^+}$

(c) $\mathbf{\Lambda_b}$ (d) $\mathbf{B^+}$

Figure 4.13: Elasticity distributions of the leading hadron in Λ_c, D, Λ_b and B diffractive and partonic collisions with protons. All distributions are normalized to integral 1.

the underlying assumptions are not correct. First, the time scales of the projectile traversing the nucleus and that of the recombination of partons inside the proton are fairly different, the former being much shorter. Thus, there is no time for the partons resulting from the interaction to recombine into a hadron and suffer a second hadron-nucleon collision before it exits the nucleus. In addition, when a hadron collides with an air nucleus, not every nucleon within will participate in the interaction. To compute the number of participating nucleons, N_{part}, we use the FORTRAN routine NUCOGE [313], where the probability of an inelastic hadron-nucleon hit is determined by the choice of the hadron-nucleon overlap function. The program provides three different options: a gray disk function, a gaussian overlap function, and a eikonal-type overlap function. We will use the last one, defined by the functions G and Ω:

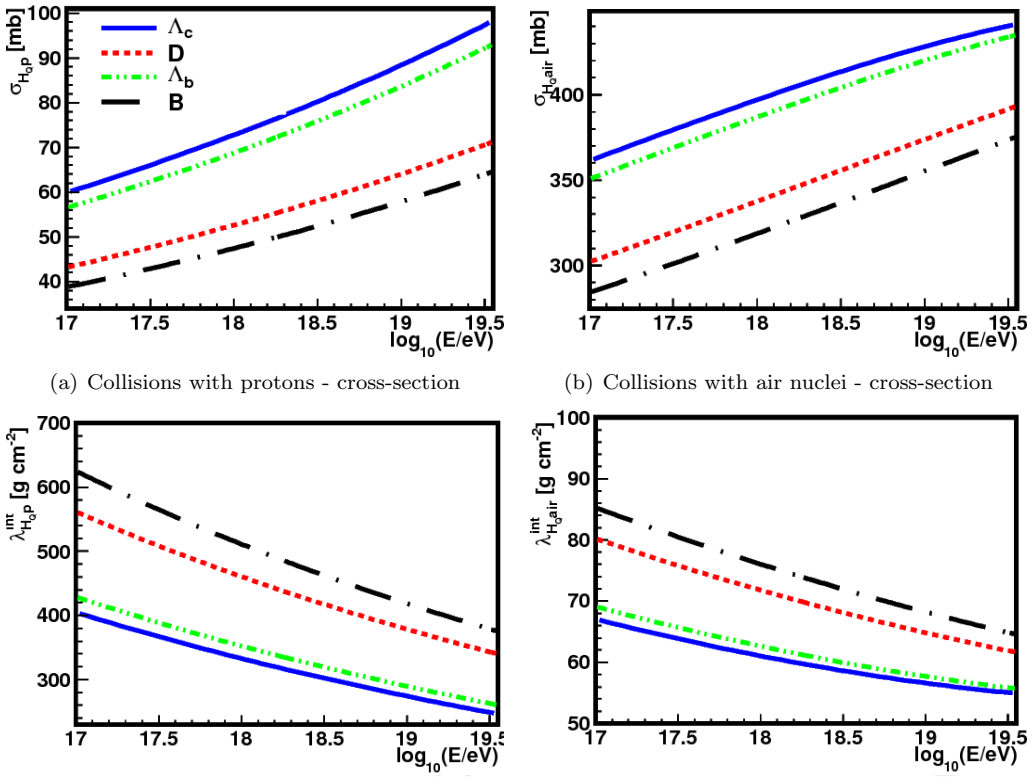
$$\begin{aligned} G(b) &= 1 - \exp(-2\Omega(b)) \\ \Omega(b) &= \Omega_0 \exp(-\beta b^2) \end{aligned} \tag{4.7}$$

where b is the impact parameter of the collision. $G(b)$ and $\Omega(b)$ have two parameters, Ω_0 and β, whose values are determined imposing that the integrals of $G(b)$ and $\Omega(b)$ fulfill:

$$\sigma_{inel} = \iint G(b) d^2b \quad = 2\pi \int b\, G(b) db \tag{4.8}$$

$$\sigma_{tot} = 2 \iint \Omega(b) d^2b = 2\pi \int 2\, b\, \Omega(b) \tag{4.9}$$

where σ_{inel} and σ_{tot} are the inelastic and total projectile-proton cross-sections, respectively. We use the values of σ_{inel} obtained in section 4.3.2 and shown in figure 4.14(a). To obtain the



(a) Collisions with protons - cross-section

(b) Collisions with air nuclei - cross-section

(c) Collisions with protons - interaction length

(d) Collisions with air nuclei - interaction length

Figure 4.14: Interaction cross-sections and interaction lengths of Λ_c, D, Λ_b and B in collisions with protons (a),c)) and air nuclei (b),d))

H_Q	Λ_c	D^+	Λ_b	B^+
$\sigma_{inel}^{H_{QP}}$ [mb]	88.4	59.2	82.1	54.4
$\sigma_{tot}^{H_{QP}}$ [mb]	107.8	64.08	100.117	58.8
Ω_0	1.074	0.346	1.070	0.354
β	0.4911	0.312	0.527	0.347

Table 4.4: σ_{tot} and σ_{inel} used to solve integrals 4.3.3 and values of Ω_0 and β obtained using Λ_c, D^+, Λ_b and B^+ as projectiles.

value of σ_{tot} we assume that the fraction of the total cross-section corresponding to inelastic processes is the same in heavy hadron-proton and light hadron-proton collisions[3].

There is a set of possible values of (Ω_0, β) that solves each integral. For instance, figure 4.15 (left) shows these values for Λ_b collisions with protons. The solid (dashed) line represents the values of (Ω_0, β) that solve integral 4.8 (4.9). The intersection of these curves gives the values of (Ω_0, β) that solve both integrals. Since the values of σ_{inel} and σ_{tot} are energy dependent, the values of (Ω_0, β) that solve the integrals need not to be constant with energy. However, their energy dependence is very mild, and we use the solutions obtained at 10^{19} eV for collisions at all energies above 10^{16} eV. The values of the cross-sections used and the pairs (Ω_0, β) obtained are shown in table 4.4 for four different projectiles.

Once the parameters (Ω_0, β) are determined, for a given target NUCOGE samples the number of participating nucleons, N_{part}. We obtain the distribution of the number of participants in

[3]Proton-proton and π-proton total cross-sections are available at [314].

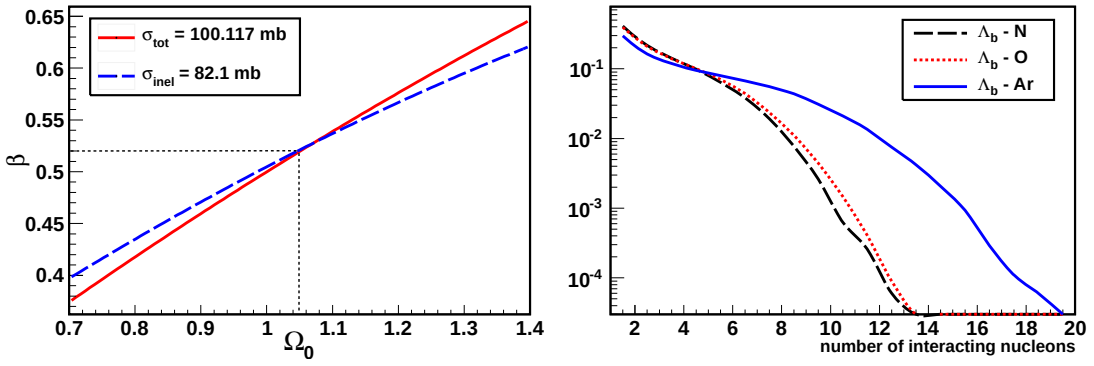


Figure 4.15: Left: Possible values of (Ω_0, β) that solve integrals 4.8 (blue dashed line) and 4.9 (red solid line). The intersection of the two curves gives the values that solve both equations at the same time. Right: Distribution of the number of interactions in Λ_b collisions with nitrogen, oxygen and argon nuclei. The distributions are scaled to integral 1.

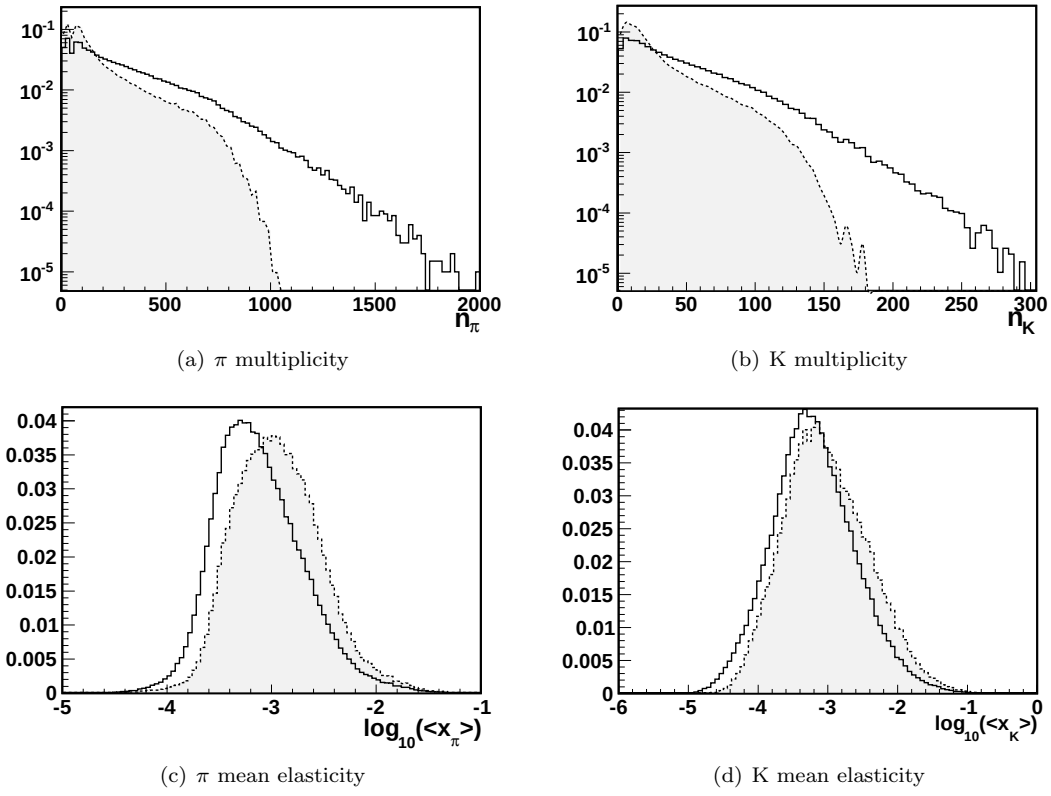
$H_Q A$ collisions, using Λ_c, D^+, Λ_b and B^+ as projectiles and nitrogen, oxygen and argon as targets. As an example, the resulting number of interacting nucleons in collisions of Λ_b with N,O and Ar (all distributions scaled to integral 1) can be seen in figure 4.15 (right).

After determining N_{part}, the nature of the hadron-nucleus interactions should be decided (either diffractive or partonic). Each of the N_{part} nucleons has a probability p_{diff} of suffering a diffractive interaction, and different for each projectile. In the case of charmed hadrons it is 0.30 for Λ_c and 0.32 for D. Turning to bottom hadrons, the values are 0.26 for Λ_b and 0.29 for B. Let N_{diff} be the number of nucleons interacting diffractively. We will consider that the interaction is diffractive as a whole if, and only if, the N_{part} participating nucleons interact diffractively ($N_{diff} = N_{part}$). Else, we consider that $N_{inel} = N_{part} - N_{diff}$ nucleons interact inelastically. To treat the inelastic interaction of a heavy hadron, with energy E_H, off an air nucleus we use the following prescription:

- First, from the N_{inel} participating nucleons, all but one are split in quark-diquark pairs, i.e we have $N_{inel} - 1$ pairs and one unbroken nucleon.
- Then, $N_{inel} - 1$ quark-antiquark pairs are generated in the projectile, with total energy $E_{q\bar{q}}$.
- The partonic interaction occurs between the projectile, with energy $E_H - E_{q\bar{q}}$, and the nucleon in the target that remains unaltered.
- The quark-antiquark pairs are matched with the quark-diquark pairs and hadronize.

$N_{spec} = A - N_{inel}$ nucleons remain as spectators, not participating in the collision. Both the partonic interaction and the hadronization are performed by PYTHIA. As a final state, we find the particles resulting from the hard interaction plus all the particles coming from the hadronization of the $N_{inel} - 1$ pairs formed.

The effects of the transition from hadron-nucleon to hadron-nucleus collisions are increasing the multiplicity of produced particles, and decreasing their mean elasticity. In figure 4.16 we show the effect of the transition from hadron-proton (dashed line) to hadron-air (solid line) collisions in case a Λ_b is used as the probing projectile. All distributions are scaled to the same integral. In figures 4.16(a) and 4.16(b) we observe that the multiplicity of pions and kaons is larger in collisions with air. At the same time (4.16(c) and 4.16(d)) the mean energy transferred to the pion and kaon component (for each collision, the sum of the energy of all pions (kaons) divided by the total number of pions (kaons)) barely rises and thus the mean elasticity per secondary particle is smaller (by a factor 1.5): a larger number of particles is sharing roughly the same amount of energy.



(a) π multiplicity

(b) K multiplicity

(c) π mean elasticity

(d) K mean elasticity

Figure 4.16: Meson multiplicities and mean elasticities for Λ_b-proton collisions (dashed lines) and Λ_b-nucleus scattering (solid lines).

	Λ_c	D^+	Λ_b	B^+
p	0.62	0.65	0.72	0.75
Air	0.56	0.59	0.68	0.72

Table 4.5: Mean elasticity values for the collisions of heavy hadrons with protons and air.

Interacting with air nuclei instead of protons also affects the leading hadron. With rising number of participating nucleons the elasticity decreases. In figure 4.17 we can see the distribution of the number of participants (left) and the mean elasticity as a function of the number of participants (right) for projectiles of any energy. In about 40% of the collisions only one nucleon participates, making these analogous to hadron-proton collisions. However, there is a significant probability of having from two to five participants. We show this behavior in figure 4.18: all elasticity distributions in hadron-nucleus collisions (dashed lines) are shifted to smaller values when compared with hadron-nucleon collisions (solid lines).

From the simulations of partonic and diffractive collisions with protons we obtained the inelastic cross-sections for Λ_c, D, Λ_b and B in collisions with protons at rest. To scale these cross-sections to collisions with air nuclei we apply the following prescription used in CORSIKA[4]. Let σ^{H-p} be the hadron-proton cross-section. Then, the hadron-air cross-section is obtained as:

$$\sigma^{H-air}\ [\text{mb}] = (1 - 4\sigma_{45}^2) \cdot p_0 + \sigma_{45}(2\sigma_{45} - 1) \cdot p_1 + \sigma_{45}(2\sigma_{45} + 1) \cdot p_2 \qquad (4.10)$$

[4]The parameterization is inside the subroutine **BOX2**. The function of this subroutine is explained in section 4.4.2.

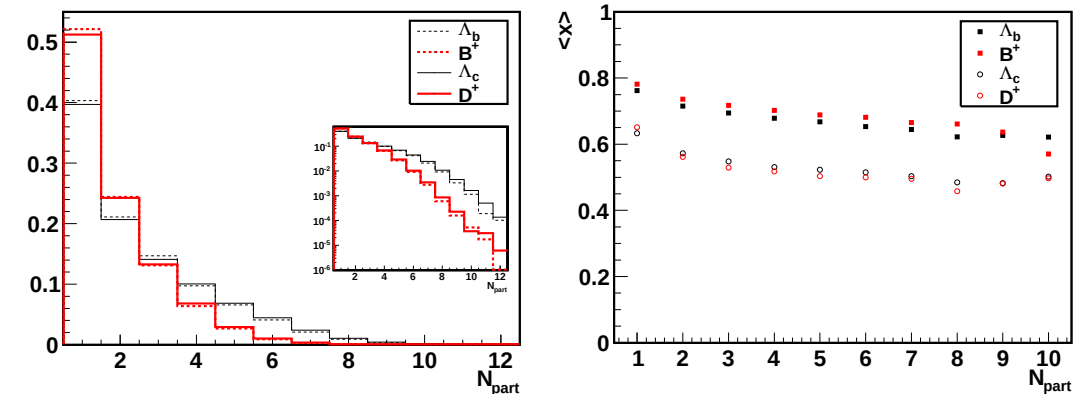
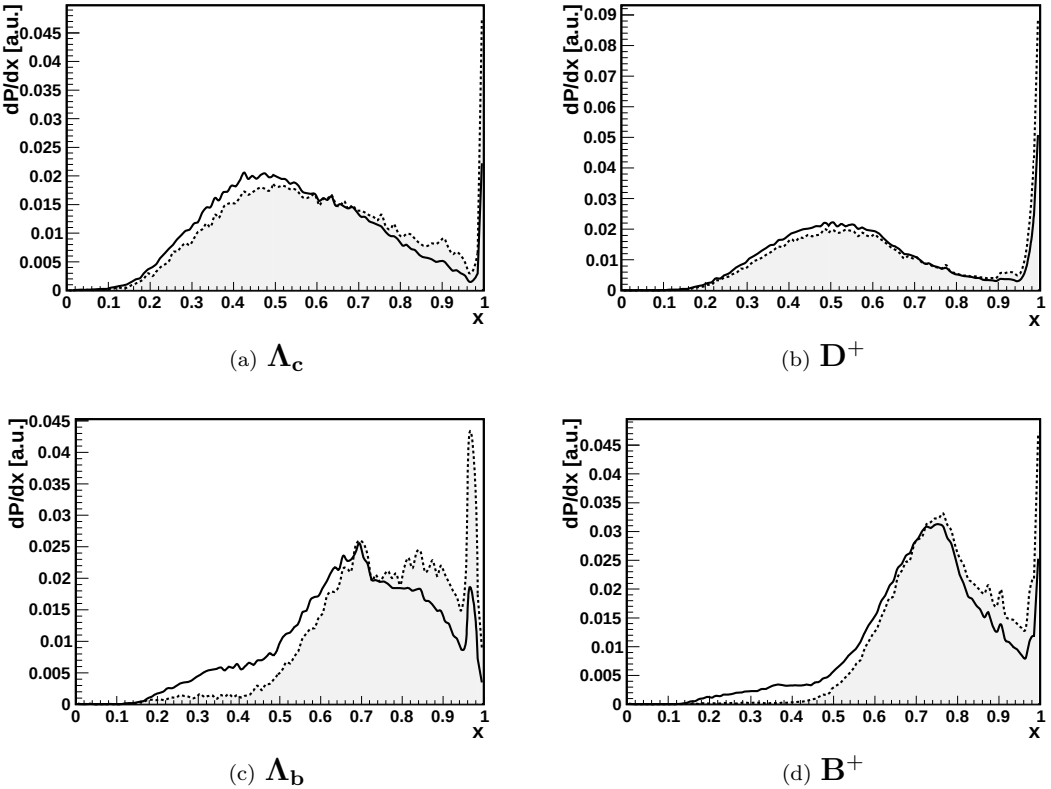


Figure 4.17: Left: Distribution of the number of participants in H_Q-air collisions for different primaries, normalized to 1. Right: Mean elasticity as a function of the number of participants.



(a) $\mathbf{\Lambda_c}$ (b) $\mathbf{D^+}$

(c) $\mathbf{\Lambda_b}$ (d) $\mathbf{B^+}$

Figure 4.18: Elasticity distributions of the leading hadron after collisions off protons (dashed lines) or air (solid lines), for four different projectiles.

where

$$\begin{aligned} \sigma_{45} &= (\sigma^{H-p}\,[\text{mb}] - 45\ \text{mb})/30 \\ p_0 &= 309.4268\,\text{mb} \\ p_1 &= 245.0771\,\text{mb} \\ p_2 &= 361.8057\,\text{mb} \end{aligned}$$

From these cross-sections we can obtain the associated mean interaction length as

$$\langle \lambda_{int} \rangle = \langle m_{air} \rangle / \sigma^{H-air} \tag{4.11}$$

In figure 4.14 we can see the resulting cross-sections (top right) and interaction lengths (bottom right) for heavy hadrons above 10^{16} eV. We use Λ_c and Λ_b cross-sections as representative of all charmed and bottom baryons, respectively. In the same way, D and B cross-sections are used for all charmed and bottom mesons.

4.4 Simulation chain and code implementation

Once we have a model for UHE charm and bottom hadrons production in cosmic rays, and a clear picture of their interactions with air nuclei, the next step is simulating them with a Monte-Carlo code. To simulate the propagation of heavy hadrons in EAS with CORSIKA several parts of the code had to be modified. A partial list of these changes would be:

- Bottom hadrons were not considered by CORSIKA. Their particle codes, masses and lifetimes had to be included and some subroutines had to be modified to recognize them.
- The interaction model described in the previous section was coded as a set of new subroutines.
- Charmed and bottom hadron interaction cross-sections in CORSIKA were changed to those shown in figures 4.14(b).

A complete description of the modifications made to the CORSIKA source code, with a list of the subroutines that were modified or added can be found in appendix B. In the following, we will describe the different steps involved in the Monte-Carlo simulation of heavy hadrons in EAS and reference the subroutines modified or newly written.

The CORSIKA simulation chain consists of several steps. A simplified scheme of the program flow diagram is in figure 4.19. Initially, we simulate the primary particle first interaction, choosing whether charmed hadrons, bottom hadrons or none of them are produced in the collision. The propagation of heavy hadrons across the atmosphere takes place along with the rest of the shower, but according to the interaction model described in section 4.3. The decay of both charmed and bottom hadrons is performed by PYTHIA.

4.4.1 First interaction

Heavy quarks can be produced in any of the collisions taking place along the shower development, provided the interaction is energetic enough. However we restrict our interest only to heavy hadrons produced in the first interaction of the primary particle with an atmospheric nucleus. Charmed and bottom hadrons produced in subsequent interactions are much less energetic and therefore their influence in the longitudinal development of the shower will be small. And, even though they could be produced deeper in the atmosphere, it is the energy at production, and not the production depth, that rules the propagation. To check this, we simulated the production and propagation of charmed (Λ_c, D^0, D^+) and bottom (Λ_b, B^0, B^+) hadrons produced by a proton of energy $10^{19.5}$ eV with a uniform distribution in $\log_{10}(\mathrm{E/eV}) \in [16.5, 19.5]$. Their production depth corresponds to the depth of a $10^{19.5}$ eV proton first interaction, distributed as

$$P(X_0; \lambda_{int}^{p-Air}) = \frac{1}{\lambda_{int}^{p-Air}} \exp(-X_0/\lambda_{int}^{p-Air}) \tag{4.12}$$

In figure 4.20 (left) we plot the mean number of interactions suffered by B^+ and D^0 in the atmosphere before decaying as a function of their initial energies, for 3 different ranges of production depth. In figure 4.20 (right) we can find the mean number of interactions suffered by the B^+, as a function of X_0, for different primary energy ranges. We can see that the number

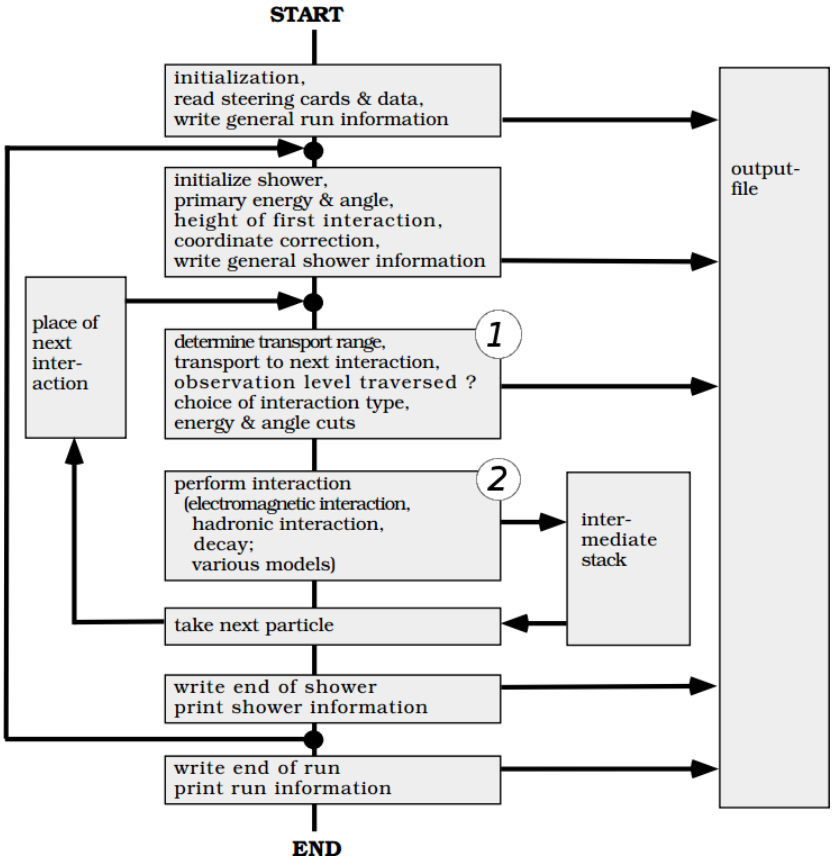


Figure 4.19: Simplified flow diagram of CORSIKA [315]

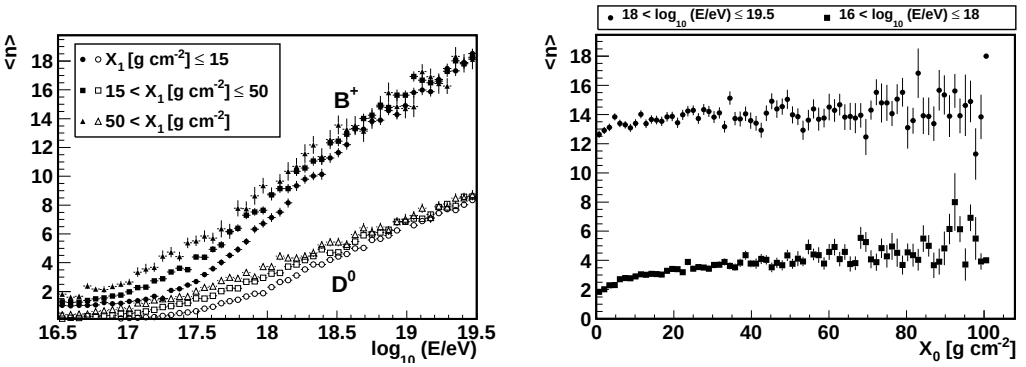


Figure 4.20: Left: Mean number of interactions, <n> as function of energy, suffered by B^+ and D^0, for different production depth bins. Right: Mean number of interactions <n> of a B^+ meson as function of depth, for different energy bins.

of interactions suffered before decay increases rapidly as the meson initial energy grows, almost independently of X_0. This observation is more clearly seen in figure 4.20 (right) where for fixed energies, the number of interactions is roughly constant with growing production depth. Even though only D^0 and B^+ are shown, the results for the other charmed (bottom) hadrons are similar to those of D^0 (B^+).

The already existing CORSIKA subroutine **NUCINT** selects the type of interaction suffered by the particles in the EAS. From there, the new subroutine **COLLIDE** is called during the proton first interaction to generate charmed or bottom hadrons, or to let the high-energy interaction model handle the first interaction.

4.4.2 Propagation

After their production, heavy hadrons generated during the first interaction have to be propagated. As CORSIKA has been modified to recognize particles with bottom quarks, both charmed and bottom hadrons can be propagated using the standard machinery built in CORSIKA. During their propagation, these particles will interact with nuclei in the atmosphere or will decay in flight. Whether any of these happens depends on the values of the interaction and decay lengths. The propagation is performed in CORSIKA in the step marked as **1** in the flow diagram depicted in figure 4.19. The subroutine **BOX2** is in charge of computing the value of the interaction cross-section, the interaction and decay lengths and determining whether the propagated particle interacts or decays. The mean interaction length in units of depth is given by equation 4.11, which yields the values plotted in figure 4.14(d). The mean decay length, i.e the mean distance a particle traverses before it decays, in units of distance is given by:

$$\langle \lambda_{dec} \rangle = \frac{Ec\tau}{m} \tag{4.13}$$

where τ is the particle mean life-time and m its mass. In CORSIKA, the actual values for the interaction and decay lengths are sampled from the following exponential distributions:

$$P(\lambda_{int}; \langle \lambda_{int} \rangle) = \frac{1}{\langle \lambda_{int} \rangle} \quad \exp(-\lambda_{int} / \langle \lambda_{int} \rangle) \tag{4.14}$$

$$P(\lambda_{dec}; \langle \lambda_{dec} \rangle) = \frac{1}{\langle \lambda_{dec} \rangle} \quad \exp(-\lambda_{dec} / \langle \lambda_{dec} \rangle) \tag{4.15}$$

If $\lambda'_{dec} < \lambda_{int}$, where λ'_{dec} is the decay length expressed in depth units, the particle travels a distance λ'_{dec} and decays. Else, the particle travels λ_{int} before interacting with an atmospheric nucleus. Energy losses during the particle time-of-flight are treated by CORSIKA standard routines.

4.4.3 Interaction

As heavy hadrons cross the atmosphere, they will collide with atmospheric nuclei. We treat the collisions according to the model described in section 4.2.2. The new subroutine **HEPARIN** links with the PYTHIA routines that treat the interaction of heavy hadrons with air nuclei, instead of calling the high-energy hadronic model chosen during compilation. The number of interacting atmospheric nucleons is calculated using the function **NNY**, which incorporates the results of the analysis from section 4.3.3. It also assigns whether the interaction is diffractive or partonic. After each collision numerous particles are generated, and usually the particle containing the heavy quark carries away the largest energy fraction. All the collision products are passed back to the CORSIKA stack of particles using the subroutine **PYTSTO** and tracked as any other particle that contributes to the shower development.

4.4.4 Decay

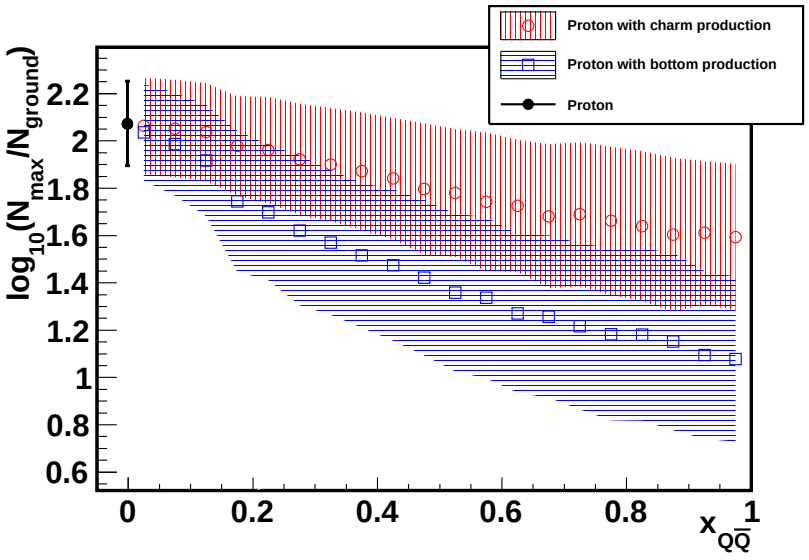
During their propagation in the atmosphere the heavy hadrons will lose energy due to ionization and bremsstrahlung, but specially because of their collisions with nuclei. The decrease in energy modifies the values of both the interaction and decay lengths, rising the former and reducing the latter, and thus increasing the decay probability. At the same time, the particle approaches ground and the atmosphere grows thicker, reducing the distance between interactions. The interplay of these effects will determine where the decay occurs. These processes take place at the stage labeled as **2** in the shower flow diagram.

The decay of both charmed and bottom particles is performed within CORSIKA. **CHRMDC** used to treat only the decay of charmed hadrons. Now it is able to perform the decay of bottom hadrons as well.

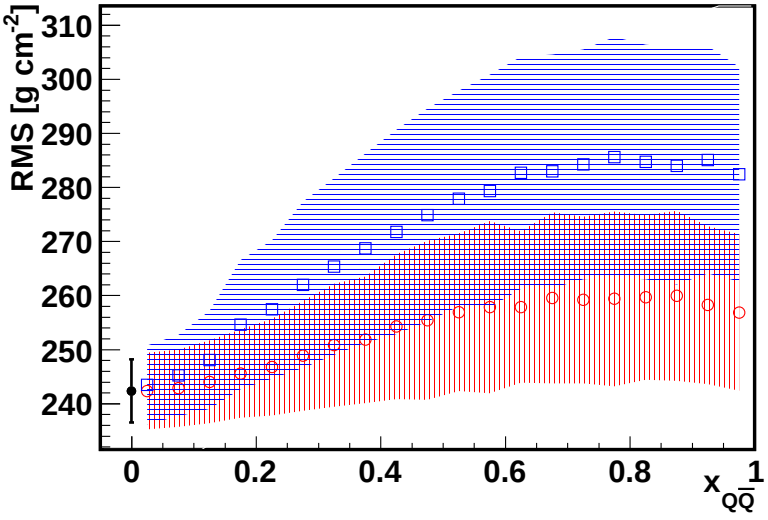
4.5 Effects of heavy quark production on EAS development

In section 4.2.5, based solely on the distributions of energy carried by the produced heavy quarks, we anticipated that only heavy quarks produced through the Intrinsic Quark mechanism could induce modifications on the shower development. Once the production and propagation have been implemented, we can check this hypothesis. According to the suggestion made in [11], the propagation of a relatively penetrating shower component, if energetic enough, could delay the development in the shower. If the energy deposition in the shower is shifted to larger depths, the longitudinal profile of the shower will be affected: the number of particles at maximum will decrease, while at the same time the number of particles that reach ground will increase. Since the interactions of heavy hadrons are more elastic and tend to slow the shower development, we expect shower profiles with a leading heavy hadron to be wider on average, with larger RMS.

Figure 4.21 shows the ratio of the number of particles at ground level to the number of particles at shower maximum (4.21(a)), and the width of the shower profile (4.21(b)) as a function of the proton energy fraction carried by the heavy component during the first interaction, regardless of the production model used. For comparison, the average value $\pm 1\sigma$ deviation for proton showers where heavy quark production has been turned off (black circle) is plotted. We can see that only for values of $x_{Q\bar{Q}}$ above 30-40% the effect starts to be noticeable. Such large fractions of energy in a $H_{Q\bar{Q}}$ pair cannot be achieved by production in the Color Glass Condensate model. As such, all showers containing $H_{Q\bar{Q}}$ produced within that model will be indistinguishable from proton showers without heavy quark production. Only showers with heavy hadrons produced through the Intrinsic Quark model will have any possibility of inducing enough modifications in the shower for it to be distinguished from proton showers without production.



(a) $N(X_{max})/N(X_{ground})$ ratio.



(b) Shower profile width.

Figure 4.21: Evolution of the ratio of particles reaching ground to particles at maximum (a)) and shower profile width (b)) with the fraction of proton energy carried by the $Q\bar{Q}$ pair.

In this chapter we have reviewed the cornerstones of heavy quark detection, with special emphasis on cosmic rays experiments. The effect of heavy hadrons in the shower development largely depends on the fraction of the primary energy they carry. **Large fractions are predicted by the Intrinsic Quark model**, whose effect we consider during the proton first interaction. Heavy hadrons produced during subsequent interactions will not have enough energy to significantly affect the shower.

Due to their large mass, **charmed and bottom hadrons interactions are more elastic than those of lighter hadrons**. In collisions with air nuclei bottom hadrons keep, on average,

a 70% of their energy. This value is of $\sim$ 58% for charmed hadrons.

The production and interaction of these particles has been implemented in CORSIKA. Showers with an energetic heavy hadron component propagating show a **displacement of the number of produced particles towards larger depths**. As a result, the ratio of particles on ground to the maximum number of particles in the shower grows. At the same time the shower development is slower, and the width of the shower increases. Both effects are more prominent with rising fractions of the primary energy carried by the heavy hadrons.

5

Simulation and reconstruction of showers and development of analysis tools

The production and propagation of heavy quark hadrons in EAS are random processes, with an effect on the shower development that can only be assessed or dismissed by means of extensive Monte-Carlo simulations. This is specially true when the expected effects or rate of events are small.

The whole Monte-Carlo simulation has to account for the production of these and other particles during the first interaction, their propagation and decay in the atmosphere, and the response of the detectors to the particles that reach them. The production of showers has three steps. The first step, described in section 5.1, is the simulation of the shower physics and development in the atmosphere. A version of CORSIKA that includes the modifications mentioned in chapter 4.2 is used to generate the showers. For each bin of energy and type of heavy quark produced during the primary interaction (bins are summarized in table 5.1), we simulate 2000 proton showers. The primary flux is not composed only of protons, but we will see that only heavy hadrons generated in proton-induced showers are prone to detection.

Once the particles of the shower have been propagated along the atmosphere and have reached ground, the response of both the fluorescence and the surface detector to these particles has to be simulated (section 5.2), and the properties of the shower reconstructed (section 5.3). Both simulation and reconstruction are performed with the official package $\overline{\text{Off}}\underline{\text{line}}$ of the Pierre Auger Collaboration [316].

The reconstructed showers where heavy hadrons are produced during the first interaction will form our *signal training sample*. The other set of reconstructed showers, those in which heavy hadrons are *not* produced will constitute the *background training sample*. The analysis we have performed using these simulation showers is explained through section 5.4. A series of quality cuts needed to select well reconstructed events is explained in section 5.4.1. We use the events fulfilling the cuts to train a multivariate discriminant, a Boosted Decision Tree, with a series of observables that will try and maximize the separation between the signal and the background samples. Boosted Decision Trees, their implementation and the selection of variables are discussed in section 5.4.2. Finally, in section 5.4.4 we show the resulting selection efficiencies of showers induced by nuclei heavier than protons and photons.

5.1 Atmospheric shower simulation

CORSIKA allows to combine various high- and low-energy models to treat the interactions of particles during the shower development. The showers used in this analysis were simulated using QGSJet01c [251] as the high-energy interaction model and FLUKA [317, 318] as the low energy interaction model. The energy threshold at which the high-energy model is replaced by the low-

energy model is set at 200 GeV. We have used only one high-energy interaction model among the available ones. Since the physical description of the shower is not completely equivalent between them, we have studied the possible systematic uncertainty associated to the choice of the interaction model, and an estimate is given in section 6.5.1.

For energies of the primary particle above 10^{16} eV the computing times, which scale roughly with the primary energy, become excessively long. Reducing computing times is necessary to generate large samples of simulated showers in reasonable periods of time. This situation is remedied introducing the "thin sampling" mechanism [319]. When thinning is active, particles emerging from an interaction are subject to the algorithm. Among all particles below an adjustable fraction of the primary energy (called thinning level $\varepsilon_{th} = E/E_0$) only one is followed, and an appropriate weight to account for the untracked particles is given to it. The rest of the particles below the thinning level are removed from the simulation. Weights too large can result in undesired statistical fluctuations, specially far from the shower core. This situation is solved limiting the maximum weight of particles emerging from an interaction, excluding those above the limit set by the thinning algorithm [320]. A third algorithm to save space on disk reduces the number of particles close to the shower core where detectors will saturate. Particles arriving at the detector level with core distance r within a selected maximum distance r_{max} are selected with a probability $\propto (r/r_{max})^4$. The weight of the selected particles is multiplied by the inverse of this probability, irrespective of exceeding the weight limiting. Further details on the thinning formalism can be found in [319, 321, 322].

We include an example of a CORSIKA input file, showing relevant parameters used in one particular simulation. The parameters of the simulation run are controlled by keywords that choose the primary particle, modify the thinning level or set the random number seeds. We give a brief explanation of those keywords for which we have not used the default values that can be found in the CORSIKA manual [323]:

```
RUNNR   157616
NSHOW   1
PRMPAR  14
COLLDR  1 3
SIGMAQ  0.D0 0.D0 0.D0 0.D0
PROPAQ  1
ERANGE  3.162278e+10 3.162278e+10
THETAP  6.000000e+01 6.000000e+01
PHIP    -180.  180.
SEED 315233 0 0
SEED 315234 0 0
OBSLEV  1.452E+05
MAGNET  19.86 -14.29
LONGI   T  5.  T  T
THIN   1.E-6 3.162278e+04 1.0E+04
USER   agascon
EXIT
```

RUNNR, NSHOW and PRMPAR are an identifying run number of the simulation, the number of showers to be generated in the run, and the type of the primary particle, respectively. `PRMPAR 14` corresponds to proton primaries.

COLLDR, SIGMAQ and PROPAQ modify features of the production and propagation of heavy hadrons in the shower. A detailed explanation of these keywords can be found in appendix B.

ERANGE, THETAP and PHIP select the minimum and maximum energy, zenith angle and azimuth angle, respectively, of each shower generated. Values are generated with a uniform random distribution between the minimum and the maximum.

SEED contains the seeds for the random number generators. At least two seeds should be initialized.

OBSLEV is the altitude at which the properties of the particles are recorded. It has been set to the height of Malargüe above the sea level (∼1452 m)

MAGNET sets the horizontal (pointing north) and vertical (downwards) components of the Earth's magnetic field (in μT). We have used the values of the magnetic field at Malargüe.

LONGI selects whether the longitudinal development of various components of the shower will be sampled, and the depth sampling step, which we set to $5\,\mathrm{g{\cdot}cm^{-2}}$.

THIN defines the thinning level, the weight limit for thinning, and the maximum radius (in cm) for ground radial thinning.

There are three major output files produced by a simulation run [323]. One is the control printout, that allows to monitor the simulation and gives general information of the program settings: interaction models selected, keywords used, physical constants, the atmospheric model and the primary particle. It also contains the number of secondaries reaching the observation levels, interaction statistics for nucleons, pions, kaons, and strange baryons per kinetic energy interval, interaction length statistics for the above particles and decay statistics for muons

The particle output file contains information about all the particles that reach the observation level, in this case defined to be the ground level at Malargüe. The type, momentum, position, timing and weight of each particle are recorded. In addition, a file containing the energy deposit and number of particles at different depths along the shower axis, i.e. the longitudinal shower development, is generated. The energy deposition of particles along the atmosphere is needed to simulate the light detected by the fluorescence telescopes.

$\log_{10}$ (E/eV)	18.5 ı 18.625 ı 18.75 ı 18.875 ı 19.0 19.125 19.25 ı 19.375 ı 19.5 ı 19.625 ı 19.75
θ [deg]	60
First interaction	No heavy quark production Intrinsic charm production Intrinsic bottom production

Table 5.1: Summary of the simulated CORSIKA showers bins. 2000 showers were generated for each $(E, \theta,$ First interaction) bin. Each shower is simulated and reconstructed 5 times, to a total of 10000 reconstructed showers per bin.

5.2 Detector response

Once the CORSIKA simulation has been completed we have a clear picture of the shower development and the type, position and momentum of particles reaching ground. At this point we have to simulate the detector response to the particles of the shower.

The surface detector is discrete, not continuous, and the majority of the particles that reach ground escape detection. Besides, the Water-Cherenkov tanks do not measure the nature, momentum and direction of every single particle traversing them. Instead, they give a time-dependent signal that accounts for all the particles crossing through them. Thus, the sparse nature of the detector, the response to the particles crossing it, and that of the electronic systems within have to be simulated.

With respect to the fluorescence detector, CORSIKA samples the number of different types of particles at different depths in the atmosphere, which is the basis of the FD response simulation. In a real shower, particles traversing the atmosphere excite atmospheric nitrogen (see section 2.4.2) which in turn emits fluorescence light. This radiation might scatter or propagate unaltered, and then reach the telescope. Atmospheric conditions have to be taken into account too. The cloud distribution over the fluorescence detector stations and the presence of aerosols

at different heights affect the propagation of light from its production point to the detector. To obtain the response of the fluorescence detectors to the shower, the nitrogen emission, the Cherenkov production, and their propagation from the production point to the telescopes have to be simulated, along with the response of the electronics and the trigger system.

The Offline package allows to simulate these processes. It is a general purpose framework that supports a variety of different computational tasks necessary to analyze the observatory data. A collection of processing modules can be assembled and sequenced to perform the different steps of the simulation and reconstruction. They can be configured by external files containing the diverse parameters and options [316]. The sequence of modules used for the hybrid simulation is the following:

```
<sequenceFile>

  <enableTiming/>
  <moduleControl>

    <loop numTimes="1" pushEventToStack="yes">

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

      <loop numTimes="5" pushEventToStack="yes">

        <module> EventGeneratorOG </module>

        <!-- Sd simulation -->
        <loop numTimes="unbounded" pushEventToStack="no">
          <module> CachedShowerRegeneratorOG </module>
          <module> G4TankSimulatorOG </module>
        </loop>
        <module> SdSimulationCalibrationFillerOG </module>
        <module> SdPMTSimulatorOG </module>
        <module> SdFilterFADCSimulatorMTU </module>
        <module> SdBaselineSimulatorOG </module>
        <module> TankTriggerSimulatorOG </module>
        <module> TankGPSSimulatorOG </module>

        <!-- Fd simulation -->
        <module> FdSimEventCheckerOG </module>
        <module> ShowerLightSimulatorKG </module>
        <module> LightAtDiaphragmSimulatorKG </module>
        <module> ShowerPhotonGeneratorOG </module>
        <module> TelescopeSimulatorKG </module>
        <module> FdBackgroundSimulatorOG </module>
        <module> FdElectronicsSimulatorOG </module>
        <module> FdTriggerSimulatorOG </module>

        <!-- Build event -->
        <module> CentralTriggerSimulatorXb </module>
        <module> CentralTriggerEventBuilderOG </module>
        <module> EventBuilderOG </module>

        <!-- Output event -->
        <module> EventFileExporterOG </module>
      </loop>
    </loop>
  </moduleControl>
</sequenceFile>
```

The inner loop contains the essential elements for the simulation. The outer loop allows to process all Monte-Carlo showers in a file or collection of files.

- The first module used, *EventFileReaderOG*, is in charge of loading simulated data from a variety of formats, and making the data therein available. Using this information, the *EventGeneratorOG* sets the core position and time for the simulated shower at some location on the array or in front of a telescope. In our simulations, core positions were

generated around stations chosen randomly from the list of stations.

The modules used for the surface detector simulation deal with the whole simulation chain, from the regeneration of the *thinned* particles to the simulation of the local station triggers.

- Module *CachedShowerRegeneratorOG* takes weighted particles from the shower simulation program, regenerates them as a set of particles with unity weight, and injects these particles into stations to simulate the detector response.

- *G4TankSimulatorOG* implements the physics processes of the detector: it simulates the particle's transport and interactions inside the detector where most relevant physics processes are accounted for. The generation of the Cherenkov radiation of the particles in the tank, its reflection in the tank walls and its transport to the PMTs are simulated by this module too.

- Modules *SdSimulationCalibrationFiller*, *SdPMTSimulator*, *SdFilterFADCSimulator* and *SdBaselineSimulator* set the simulated detector calibration constants, produce a charge pulse out of the PMTs and simulate the filter response and the baseline of the signal, respectively.

- Finally, *TankTriggerSimulator* decides whether a signal from a certain station fulfills the local trigger criteria or not.

The FD simulation chain covers all steps starting from the simulation of the produced light to the trigger system of the fluorescence detector:

- Modules *ShowerLightSimulator* and *LightAtDiaphragmSimulator* simulate the Fluorescence and Cherenkov light along the shower axis and the light flux that reaches the FD telescopes, respectively.

- The distribution of injected photons and background light is simulated by modules *ShowerPhotonGenerator* and *FdBackgroundSimulator*, and the response to this light is performed by *TelescopeSimulator*.

- Finally, the various levels of trigger of the detector are simulated by *FdTriggerSimulator*.

The *CentralTriggerSimulator*, *CentralTriggerEventBuilder* and *EventBuilder* inspect the triggers at event level, combining the information of the different detectors. At the end of the simulation chain the *EventFileExporter* module writes events in a format easily readable by the user and apt for the reconstruction chain.

Each CORSIKA shower generated is simulated in the array five times with random core positions, increasing the statistics to 10000 showers per energy and particle type bin. The randomization of the core position allows us to inject each shower more than once, effectively considering it a different shower. However, to avoid spurious correlations in the data sample each shower should not be replicated more than 10 times.

5.3 Shower reconstruction

After detector simulation, the next step deals with the event reconstruction. Combining the signals and times simulated in the Water-Cherenkov tanks and the fluorescence telescopes, a geometry and energy has to be found for the simulated shower. The shower reconstruction is also performed using the $\overline{\text{Off}}\underline{\text{line}}$ package, applying the same modular structure where each step of the reconstruction is in a separate module. The sequence of modules used for the hybrid reconstruction, which is the same for simulated showers and real data collected by the observatory, is the following:

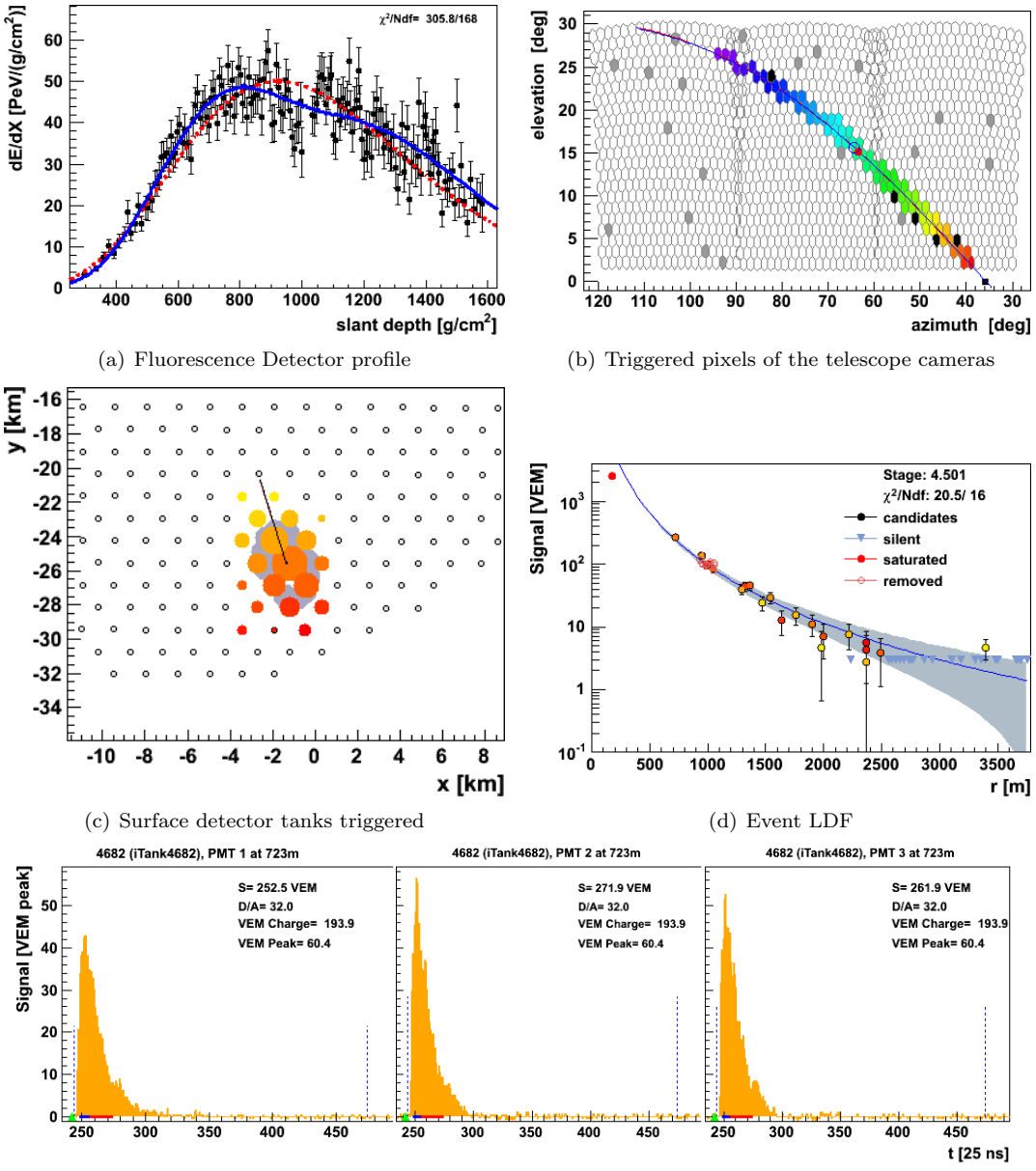
```
<sequenceFile>

  <enableTiming/>

  <moduleControl>

    <!-- <loop numTimes="1" pushEventToStack="yes"> -->
    <loop numTimes="unbounded" pushEventToStack="yes">

      <module> EventFileReaderOG      </module>
      <module> EventCheckerOG         </module>

      <loop numTimes="1" pushEventToStack="yes">

        <try>
          <module> FdCalibratorOG         </module>
          <module> SdCalibratorOG         </module>
          <module> FdPulseFinderOG        </module>
          <module> FdSDPFinderOG          </module>
          <module> FdAxisFinderOG         </module>
          <module> HybridGeometryFinderOG </module>
          <module> FdApertureLightOG      </module>
          <module> FdProfileReconstructorKG </module>
        </try>

        <!-- SD reconstruction -->
        <try>   <!-- limit how far a Continue goes -->
          <module> SdEventSelectorOG </module>
          <module> SdMonte-CarloEventSelectorOG </module>
          <module> SdPlaneFitOG </module>
          <module> LDFFinderKG </module>
          <module> Risetime1000LLL </module>
          <module> SdEventPosteriorSelectorOG </module>
        </try>

        <module> AnalysisGL </module>
        <module> RecDataWriterNG </module>
      </loop>
    </loop>
  </moduleControl>
</sequenceFile>
```

The first inner loop of the sequence is in charge of the fluorescence detector reconstruction:

- First, the fluorescence and the surface detectors have to be calibrated. *FdCalibrator* and *SdCalibrator* transform simulated raw data into physical quantities and *FdPulseFinder* processes the traces recorded by the fluorescence telescopes.
- Next, the plane containing the shower axis and the eye which detected it are determined. Within this plane, modules *FdSDPFinder*, *FdAxisFinder* and *HybridGeometryFinder* perform a complete geometrical fit, taking into account both the timing of the shower image as it traverses the telescope pixels and the timing and impact point on the surface detectors.
- Module *FdApertureLightOG* calculates the light flux reaching the telescope aperture, and inside *FdProfileReconstructor* the shower profile is reconstructed. This process converts the fluorescence light profile recorded by the telescopes into a determination of the energy deposit at a given atmospheric depth along the shower axis.

The next loop of the reconstruction chain is devoted to the surface detector:

- *SdPlaneFit* uses the signal timing of the stations to fit the shower plane, the axis and the impact point on the ground.
- The signals in the stations and their distance to the impact point projected on the shower plane are used by *LDFFinder* to fit the Lateral Distribution Function (LDF), from which a determination of the energy is obtained.



(a) Fluorescence Detector profile

(b) Triggered pixels of the telescope cameras

(c) Surface detector tanks triggered

(d) Event LDF

(e) SD tank PMT traces of one station

Figure 5.1: $10^{19.75}$ eV proton shower simulated in CORSIKA, and then simulated and reconstructed in Offline. a) shows one fluorescence profile, where the solid blue line corresponds to the simulated shower, the black dots correspond to the reconstructed shower profile and the dashed red line is the Gaisser-Hillas fit of the profile. In b) there are the corresponding triggered pixels of the telescopes cameras. c) shows the distribution of the triggered Water-Cherenkov tanks, where the size of the circles is proportional to the tank signal. In d) there is the event Lateral Distribution Function (LDF). e) shows the traces of the station with the second highest signal.

- Module *RecDataWriterNG* stores the relevant information from the simulation and reconstruction in ADST[1] files [324]. A set of basic data is always stored, but the total amount of information parsed to the output file can be changed.

[1] Advanced Data Summary Trees.

After the reconstruction, all the relevant information concerning the detectors and the reconstructed quantities can be retrieved: from top-level variables, such as energy, direction, X_{max} (and their errors) to low-level observables, such as telescope single pixel responses, or the PMT traces. Figure 5.1 shows one of the fluorescence profiles and the Lateral Distribution Function, along with the PMT traces of one Water-Cherenkov tank of a reconstructed $10^{19.75}$ eV shower where a $b\bar{b}$ carrying about a 60% of the proton energy was produced. The simulated shower profile, deformed by the propagation of the heavy hadrons within, is shown as the blue solid line in figure 5.1(top left).

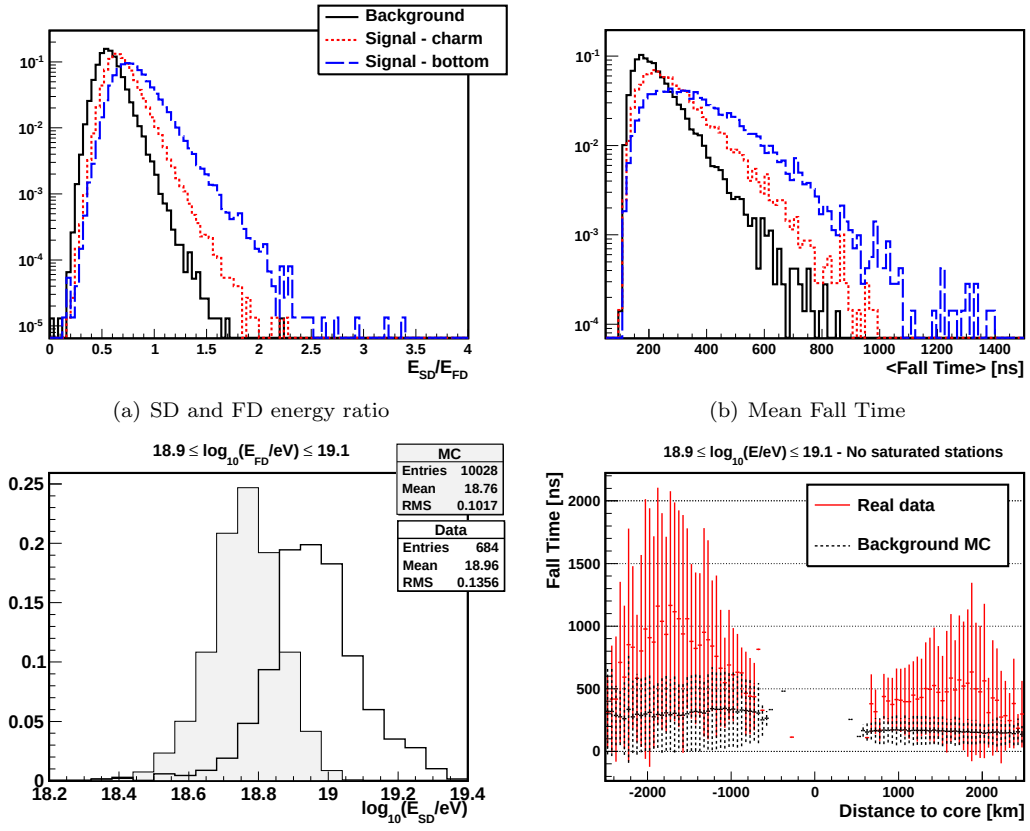
5.4 Shower analysis: tools and selection of variables

The propagation of energetic heavy hadrons in the shower modifies the event as detected both by the Fluorescence and the Surface Detector. Since they are able to reach large depths in the atmosphere, interacting rather than decaying, they displace a fraction of the shower energy, delaying the shower development. The resulting longitudinal profiles are wider, have more particles at ground level and end their development at much larger depths. In addition, showers that transport energy deeper in the atmosphere result in a larger number of EM inside the Water-Cherenkov tanks. In inclined showers this means that tanks that would otherwise only detect muons, since the electromagnetic component is largely absorbed, will have a significant component of EM origin. This fact translates, for instance, in broader signals, with larger average fall-times and rise-times. Besides, very distorted showers which are wrongly fitted by a Gaisser-Hillas function, or showers with a large number of particles on ground, will be reconstructed with different energies in the SD and the FD. As such, the ratio of the energies reconstructed by each detector is different for signal and background showers. The mean fall-time and the ratio of SD to FD reconstructed energy distributions, compared for signal and background, are shown in figures 5.2(a) and 5.2(b), respectively.

However, the usage of these SD variables is not straightforward. High-energy hadronic interaction models have the problem of yielding less muons than those found in real data [325]. This results in simulations of the SD reconstructing energies systematically lower than those of the primary particle, while not affecting the FD reconstruction. Figure 5.2(c) shows the distribution of reconstructed SD energies for a bin of FD energies, for simulations and real data. One can see that in the case of simulations, the energies reconstructed with the surface detector are systematically lower. As such, even if the ratio of energies reconstructed by the surface and fluorescence detector could be a potential observable, this particular feature of the simulation makes it unsuitable. Besides, the tank ADC traces are not 100% compatible to that of real events. In figures 5.2(d) we can see a comparison of the fall-time of individual tanks as a function of distance to the core, at a fixed energy bin and $\theta = 60^{\circ}$, both real events and MC protons. Not only there is a difference in the scale, but also in the shape of the distributions. This divergence in the behavior with time and distance is pervasive for most SD observables extracted from the FADC traces (rise-time, trace length, shape). It is possible to build SD observables compatible between data and simulations, combining variables and restricting them to specific ranges of time and distance to the core. However, there is not a strong foundation for using these variables, other than the matching between data and simulations, and they show a mild separation power. Therefore, we have chosen to perform an analysis using only the variables extracted from the fluorescence detector.

5.4.1 Energy and angular reconstruction of the fluorescence detector: quality cuts

To extract any information from the reconstructed shower profile we have to ensure that it is reconstructed with a minimum of quality. Otherwise, any observed features could potentially be due to a poor reconstruction of the shower profile. The field of view, defined as the difference of the depth of the last and first light collected, is determinant for a successful reconstruction. Depending on the position and length of the field of view some showers might have a poor



(a) SD and FD energy ratio

(b) Mean Fall Time

(c) SD energy reconstruction

(d) Real data vs Monte-Carlo Fall Time distribution

Figure 5.2: a) Comparison of the energies reconstructed with the surface detector for a fixed bin of energy reconstructed with the FD. b) Comparison of the Fall Time distribution of tanks with respect to their distance to the core positions (negative distances indicate tanks with start times before the time of the shower core). Only showers with $\log_{10}(E/eV)$ between 18.9 and 19.1 are selected. Stations with saturated signals are removed.

reconstruction even if they pass some of the trigger levels of the detector. The position of the shower core with respect to the fluorescence stations, and the relative position between the cores reconstructed by the fluorescence and surface detector also help to remove showers with wrong reconstructions. As such, not every reconstructed shower will be used in the analysis. A series of quality cuts on the reconstructed showers are needed to ensure a high-quality energy and direction reconstruction:

- **Field of view:**
  - We select showers whose reconstructed shower profile has a field of view larger than $300\,\mathrm{g{\cdot}cm^{-2}}$.
  - In addition, the field of view should start before $1000\,\mathrm{g{\cdot}cm^{-2}}$. Showers whose development is sampled too close to ground often can be mistaken with deep showers of smaller energy.
  - Usually the reconstructed value of X_{max} is required to lie inside the field of view. This cut is important to obtain a good energy reconstruction, but it is very restrictive as well. In figure 5.4(a) we can see the energy resolution as a function of the distance of the X_{max} to the beginning of the field of view. It is possible to allow for values of X_{max} up to $250\,\mathrm{g{\cdot}cm^{-2}}$ before the beginning of the field of view without impoverishing

the energy reconstruction. This increases the selection efficiency an average of 10% relative to the tighter cut that forces X_{max} to be inside the field of view.

- The hole fraction between two consecutive points of the reconstructed shower profile is defined as

$$hole_{i-1;i} = max(X_i - X_{i-1})/(X_{high} - X_{low}) \tag{5.1}$$

To avoid showers with a large field of view but where a considerable part of the shower inside it was not observed we keep the maximum hole fraction below 40%.

- **Shower core:**
  - The distance between the cores reconstructed by the surface detector and the fluorescence detector has to be smaller than 750 m. Disagreeing calculations of the core position point to wrongly reconstructed showers.

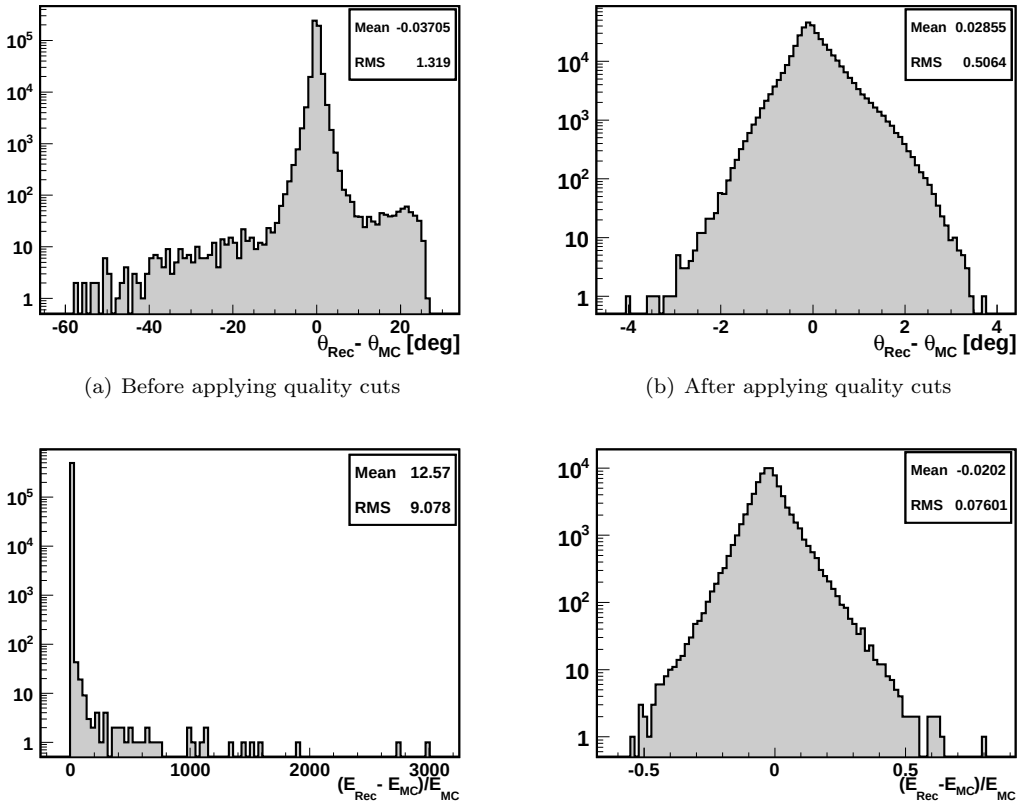
- **Light collection:**
  - The number of FD pixels used in the determination of the shower axis has to be larger than 5.
  - The fraction of Cherenkov light should be smaller than 50%.

This set of cuts eliminates those events that present a poor reconstruction, while trying to maximize the selection efficiency at the same time. Figures 5.3(a) to 5.3(d) show the distributions of reconstructed zenith angle and energy resolution, respectively, before (left) and after (right) applying the cuts. Figure 5.5 shows the quality cut selection efficiency as a function of energy for signal and background showers. We can see that the set of quality cuts accepts equally signal and background showers.

Zenith angle dependence

At the beginning of our study we simulated showers with zenith angle $\theta = 60^{\circ}$ and energies $\log_{10}(\mathrm{E/eV}) \in [18.5, 19.75]$. The reason behind choosing this zenith angle is that any effect due to the production of heavy hadrons will be noticeable deep in the atmosphere. The ground level at Malargüe is at an atmospheric depth of $890\,\mathrm{g{\cdot}cm^{-2}}$. This is roughly the value of X_{max} for vertical showers. Since the early stages of the shower longitudinal development contain no information of the heavy hadron propagation, we have to use showers with larger zenith angles if we want to observe the late longitudinal development. However, in this energy range, considering a zenith angle resolution of $\Delta\theta = 1.5^{\circ}$) there are only 384 events in real data passing the quality cuts in the period between January 2004 and December 2012. Eventually we will apply a discriminant to that sample of data, reducing the statistics even further. Performing a search experiment with poor statistics will inevitably lead to an imprecise result, so increasing the statistics is fundamental to improve the final result. We could either apply looser quality cuts or consider energies and zenith angles outside the simulated bins. The first option is not viable, since relaxing the quality cuts might result in poorly reconstructed showers populating the data sample. Alternatively, we could consider showers with energies below $10^{18.5}$ eV or above $10^{19.75}$ eV, but that would barely increase our statistics. Events with energies below the former have small selection efficiencies, and the flux is drastically reduced for energies above the latter. We are left with the option of considering a wider range of zenith angles.

If we want to analyze real showers at angles different of 60° we should first check the zenith dependence, if any, of the selection cuts. We applied the quality cuts to a set of already simulated proton showers with energies $\log_{10}(\mathrm{E/eV}) = \{18.5, 18.75, 19, 19.25, 19.5, 19.75\}$ and zenith angles θ=$\{41^{\circ}, 45^{\circ}, 49^{\circ}, 53^{\circ}, 57^{\circ}, 60^{\circ}, 63^{\circ}\}$. Since we did not simulate these showers, the energy bins do not exactly match those used in our simulations (see table 5.1). However, they do cover the same range of energies. Given that the quality cuts depend on quantities that vary with θ we expect a dependence on the zenith angle. In figure 5.4(b) we show the selection efficiency of the quality cuts, as a function of $\cos(\theta)$, for each of the simulated energy bins.



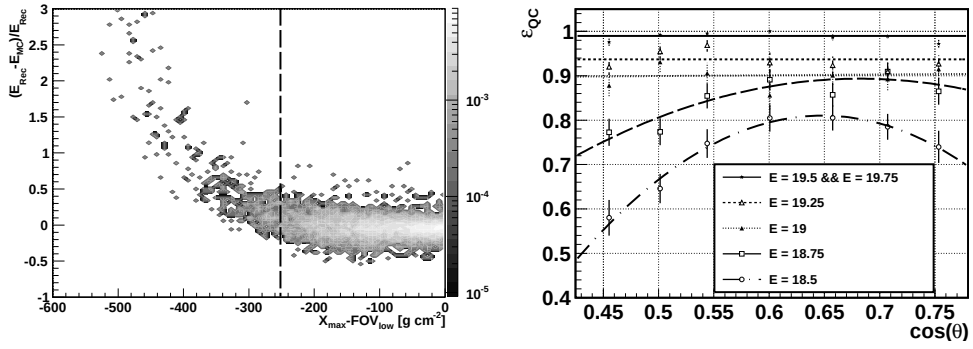
(a) Before applying quality cuts

(b) After applying quality cuts

(c) Before applying quality cuts

(d) After applying quality cuts

Figure 5.3: Reconstructed zenith angle and energy resolution before ((a),(c)) and after ((b),(d)) applying the quality cuts.



(a)

(b)

Figure 5.4: a) Energy resolution as a function of the distance of X_{max} to the lower limit of the field of view. The dashed line marks the maximum distance allowed before the energy reconstruction starts worsening. b) Quality cuts: Selection efficiency as a function of zenith angle for different energies.

At the lower energies there is a clear dependence of the efficiency on the zenith angle of the shower. As the energy of the showers rises this dependence becomes weaker, and at the highest

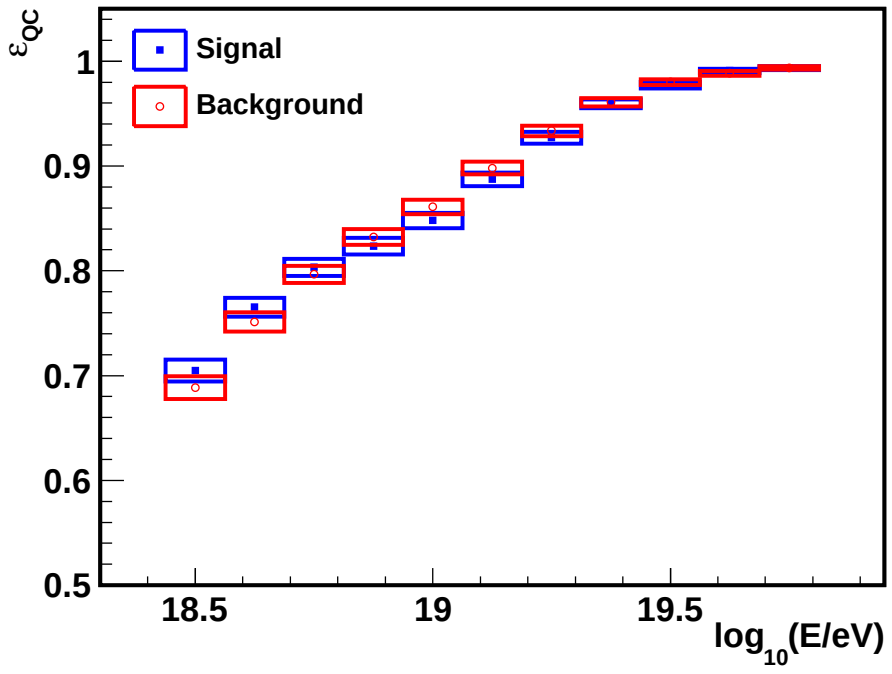


Figure 5.5: Quality cut selection efficiency as a function of energy for signal (full squares) and background (open circles) showers. Error bands correspond to 95% Wilson confidence levels.

energies there is no dependence of the zenith angle at all. To account for this dependence in θ, whenever we have to consider selection efficiencies at zenith angles different to 60° we will scale the efficiency with respect to the value at 60°.

5.4.2 Multivariate discriminant: Boosted Decision Trees

With the purpose of enhancing the discrimination of proton showers containing heavy hadrons produced according to the Intrinsic Quark model we use a multivariate analysis tool: Boosted Decision Trees (**BDTs** from now on). BDTs method is fairly simple, involving only a one-dimensional cut optimization at each training step. Little training is needed to achieve reasonably good results. There are some precedents for the use of this method in accelerator physics [326, 327], neutrino experiments [326] and astrophysics [328, 329].

A decision tree, T, is an example of a multistage decision process (figure 5.6 shows a sketch of a decision tree). The algorithm starts with a main node (full circle) that contains all the events $\{x\}$, both signal and background. For each of the variables defined $\{v\}$ the decision tree scans its distribution to find the cut value that maximizes the separation of signal and background[2]. Then, the variable that provides maximum separation is used to split the event sample, generating two new nodes (hatched circles), i.e. two subsets of events. Each node independently undergoes the same process of finding the best cut, such that each subset might find different best cuts on different variables. Repeated cuts are taken on a single variable at a time, until a stop condition is met, usually when a node reaches a minimum number of events. This way, the phase space is splitted into many regions that are eventually classified as signal or background (full squares in figure 5.6), depending on the majority of events that end up

[2]A variety of separation criteria can be configured, with no significant performance disparity between them. We adopted the criterion of maximizing the *Gini Index*, defined by $p \cdot (1-p)$, where $p = S/(S+B)$, and S and B are the number of signal and background events in the node after the cut, respectively.

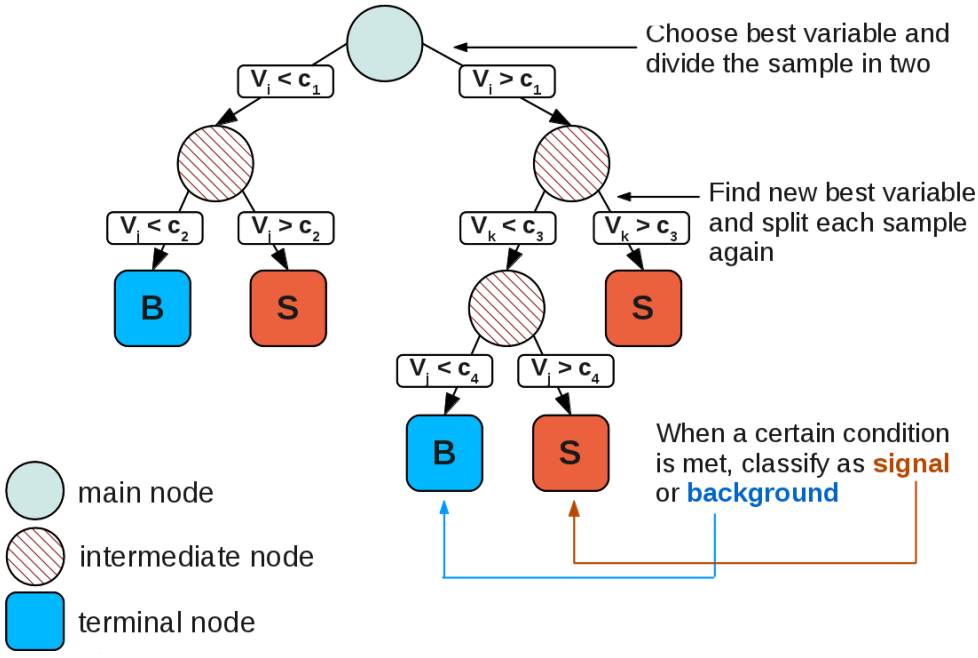


Figure 5.6: Schematic view of a decision tree [331].

in the final node. Each event is assigned a value equal to 1 if it is finally classified as signal ($T(x_i) = 1$), or -1 if it is classified as background ($T(x_i) = -1$). Decision trees have two major advantages: they are capable of modeling complex nonlinear decision boundaries, and they select those variables best suited for the classification as part of the tree construction process [330].

One decision tree is no different from a series of linear cuts, which in principle makes them powerful but unstable, with small changes in the training sample potentially resulting in large changes in the results. This complication is partly solved through a process known as *boosting*. The purpose of boosting is to sequentially apply the classification algorithm to repeatedly modified versions of the data, combining many Decision Trees to achieve a final powerful classifier. The method starts building a decision tree in the usual manner, where each event is assigned an initial weight w_i^0. After each realization of a decision tree, the weights of the events that were misclassified are increased and the Decision Tree is applied again to the weighted observations. As iterations proceed, observations that are difficult to classify receive increasing influence.

The boosting algorithms are varied [332], but in this work we will use the so called AdaBoost (Adaptative Boost). For the m^{th} tree, an error measure e_m is computed as:

$$e_m = \frac{\sum_{i=1}^{n_{ev}} w_i \cdot I(x_i)}{\sum_{i=1}^{N} w_i} \tag{5.2}$$

where $I(x_i)$ is equal to 0 if the event was correctly classified, or to 1 otherwise. i.e. e_m is the sum of the weights of the misclassified events, normalized by the sum of all weights. Then, each misclassified event is reweighed and the weight of all events is renormalized

$$\bar{w}_i^m = w_i^{m-1} \cdot \ln \alpha_m \tag{5.3}$$

$$w_i^m = \bar{w}_i^m / \sum_{i=1}^{n_{ev}} \bar{w}_i^m \tag{5.4}$$

where $\alpha_m = ((1 - e_m)/e_m)^\beta$ and β is the *boost* parameter. The final score of a given event x_i is a value in the range [-1,1] given by the weighted sum of the scores of the event over the

individual trees:

$$T(x_i) = \frac{1}{N_{tree}} \sum_{m=1}^{N_{tree}} \alpha_m T_m(x_i) \tag{5.5}$$

A usual criterion to choose the best variables for the analysis relies on their frequency of use. Variables that are chosen often to split a node are assumed to have a high discriminatory power. Variables that provide little or no separation between signal and background are seldom chosen to split the tree and thus they are effectively ignored [333]. Therefore, decision trees appear insensitive to the inclusion of poorly discriminating input variables. In spite of this, a careful selection of the variables to be used is still needed. At each node all variables are scanned to decide which is the optimum cut, hence the addition of poor variables increases the computing time considerably. Moreover, too many combinations of variables hinder the interpretation of the physics we are trying to describe. As such, it is advisable to drop those variables that are not really contributing. Considering this, we apply a particular ordering scheme to choose our analysis variables.

- Starting with any of the possible observables as the first variable (or seed), we build all possible two-variables BDTs.
- The pair of variables that generated the tree with maximum separation is chosen.
- We build all possible three-variables BDT, adding leftover observables to the selected pair.
- This procedure is repeated to find the ordering of the variables. In our case, after choosing 5-6 variables, the further addition does not improve the discriminant anymore.

Starting from different variables may span different sets of variables and all possible seeds should be investigated. This mechanism has yielded good results previously [334], and is consistent with the idea of classifier-specific variable selection being better than classifier-independent methods.

As explained in section 4.5 there are certain features that are clearly different for showers containing heavy hadrons that carry a significant fraction of the primary energy. They present a deformed shower profile, which can be characterized by a slower development, wider shower profiles and larger relative number of particles on ground. We trained different BDTs with many variables describing these aspects of the shower profile. We applied the ordering scheme described above and obtained the following set of variables:

L: The slower development of the shower can be inspected through the Universal Shower Profile (USP) parameter $L = \sqrt{\lambda \cdot |X_0 - X_{max}|}$ [g·cm^{-2}] [335]. L is used instead of λ from the Gaisser-Hillas fit due to its smaller event-by-event fluctuations, of the order of a few percent. In practice it is a measure of the width of the shower. Showers that displace a significant fraction of its energy to larger depths will be wider.

GH($\mathbf{X_{fixed}}$)/GH($\mathbf{X_{max}}$) and GH($\mathbf{X_{max} + \Delta X}$)/GH($\mathbf{X_{max}}$): Showers with a slower development more slowly will also have a larger number of particles at a fixed depth, compared to that of regular showers. Even though this observable is a good measure of a shower slow development, it is possible that showers with a large value of X_1, the point of first interaction, but with a completely regular development, will have a large number of particles at a fixed depth, just because their development started later. We can overcome this difficulty evaluating the profile at a fixed distance *from* the shower maximum, $GH(X_{max} + \Delta X)/GH(X_{max})$.

The distributions of $GH(X)/GH(X_{max})$ and $GH(X_{max}+\Delta X)/GH(X_{max})$ for signal and background showers were compared for different values of X_{fixed} and ΔX. We found that the best values were $X_{fixed} = 2000\,\mathrm{g{\cdot}cm^{-2}}$ and $\Delta X = 1260\,\mathrm{g{\cdot}cm^{-2}}$.

Shower development: An elongated shower will end its development, defined as the integral up to a certain depth normalized by the total integral, at depths much larger than regular

showers. We define f_{99} as the depth at which the shower reaches 99% of its development. The integral of the shower profile can be rewritten as a gamma function and f_{99} can be computed as

$$f_{99} = P^{-1}((X_{max} - X_0)/\lambda + 1; 0.99) \tag{5.6}$$

where P^{-1} is the inverse of the lower incomplete gamma function[3].

f_{99} presents a similar feature to that of evaluating the profile at a fixed depth: showers that initiated deeper in the atmosphere will have a larger value of f_{99} even if they have a regular development. Knowing that differences in shower development of signal and background showers are expected to appear at large depth and not at beginning of the shower, we define $f_{0.95-0.50}$ as the depth interval between the values at which the shower reaches a 95% and a 50% of its development.

$$f_{95-50} = P^{-1}((X_{max} - X_0)/\lambda + 1; 0.95) - P^{-1}((X_{max} - X_0)/\lambda + 1; 0.50) \tag{5.7}$$

Slope: Showers with a slower development are absorbed at a different rate, which can be seen as a different curvature after the shower maximum. We will use the slope of a linear fit to the right hand side of the profile, between $X_{max} + 100$ and $X_{max} + 400$, as a measurement of the speed of the absorption of the shower in the atmosphere.

Figures 5.7(a) to 5.7(f) show the distributions of the variables for background showers (solid line) and for signal showers both with charm (dotted line) and bottom (dashed line) production. We have chosen the variables according to the bottom-up procedure explained in section 5.4.2. In the figure we can see that not all the variables are equally separating, and some of them are more powerful than the others. However, we have kept the number of variables low, and by construction BDTs are not very sensitive to the presence of weak variables. As such, the presence of the less discriminating variables should not deteriorate the performance of the BDT, and could add some extra information.

Each individual background event is weighted by a factor $w = E_{MC}^{-1} N_{B;E}^{-1}$, which accounts for the different probability of occurrence according to the primary energy, and normalizes to the number of events in each class[4]. Signal events are weighted by $w = E_{MC}^{-1} N_{S;E}^{-1}$ if they correspond to charm production, and $w = 2.7 \cdot 10^{-3} \cdot E_{MC}^{-1} N_{S;E}^{-1}$ if they correspond to bottom production. The factor $w = 2.710^{-3}$ accounts for the relative probability of producing bottom quarks with respect to producing charm quarks (see equation C.1 in appendix C for details).

The analysis variables were chosen so that they would not depend on the zenith angle of the shower. If the variables are θ-independent, the discriminant, which is a function solely of these variables, will be θ-independent as well. In figures 5.8(a) to 5.8(f) we show the distribution of the variables as a function of the reconstructed zenith angle, obtained using the same set of proton showers at various zenith angles used in section 5.4.1. We can see that, as expected, there is no dependence with the zenith angle of the incoming shower.

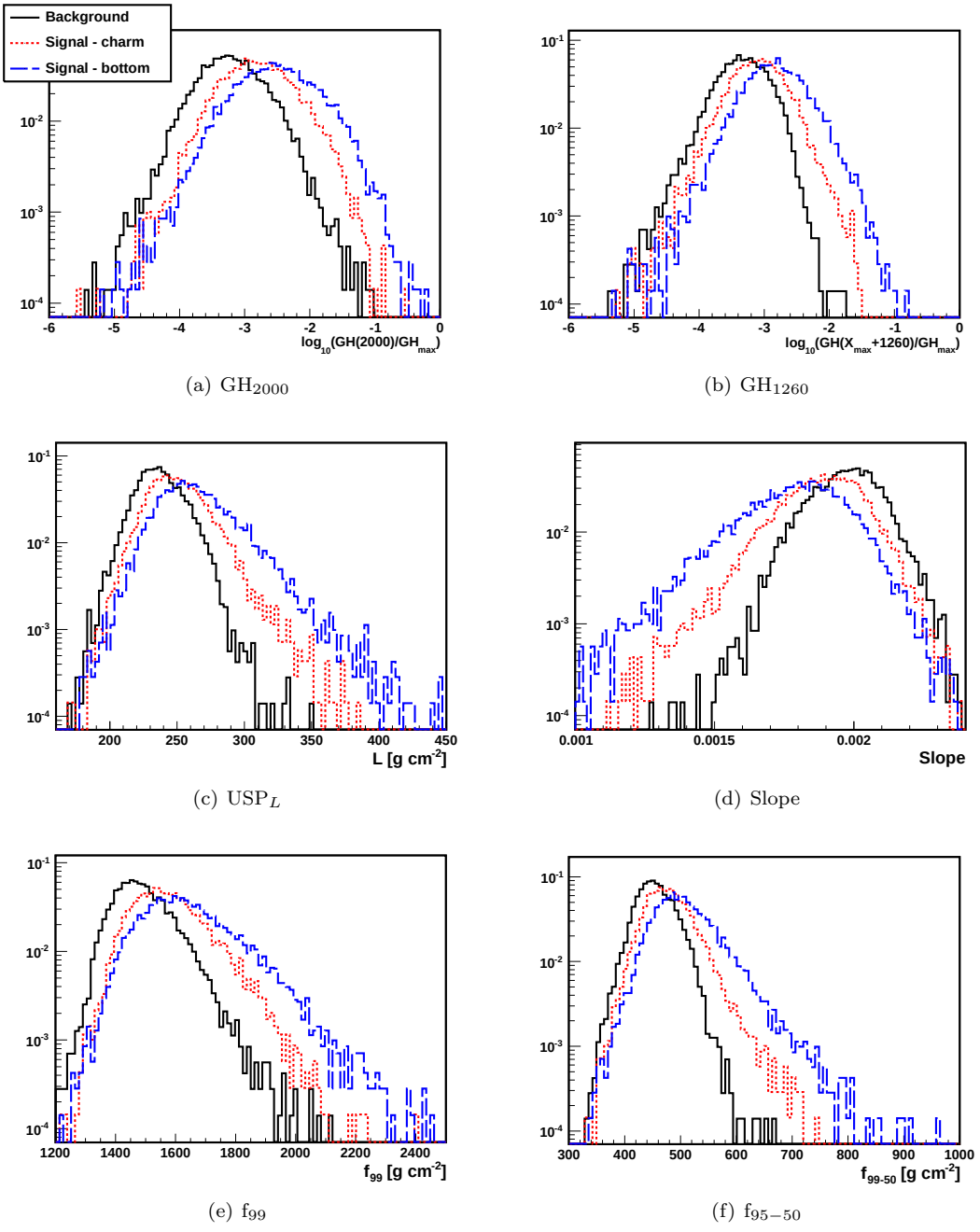
Discriminant optimization

We have used the implementation of Boosted Decision Trees provided by the TMVA package (*Toolkit for Multivariate Analysis with ROOT*) [331]. It is possible to modify some of the BDT characteristics to maximize its discriminatory power. The main configuration options are related to the number of boosts (or trained trees), the size of the tree (maximum number of nodes (n_{max})) and the boost parameter β. Deciding the best set of options for a problem requires a scan of the different available parameters.

The performance of the tree evolves with the number of boosted trees. A small number of boosts leads to under-performance, because misclassified events do not have the chance to be

[3] P^{-1} is implemented in ROOT, within the MATH namespace http://root.cern.ch/root/htmldoc/MATH_Index.html.

[4] We have generated bins with equal statistics but only occasionally the procedure reconstruction fails, and bins are not populated with exactly the same number of events.



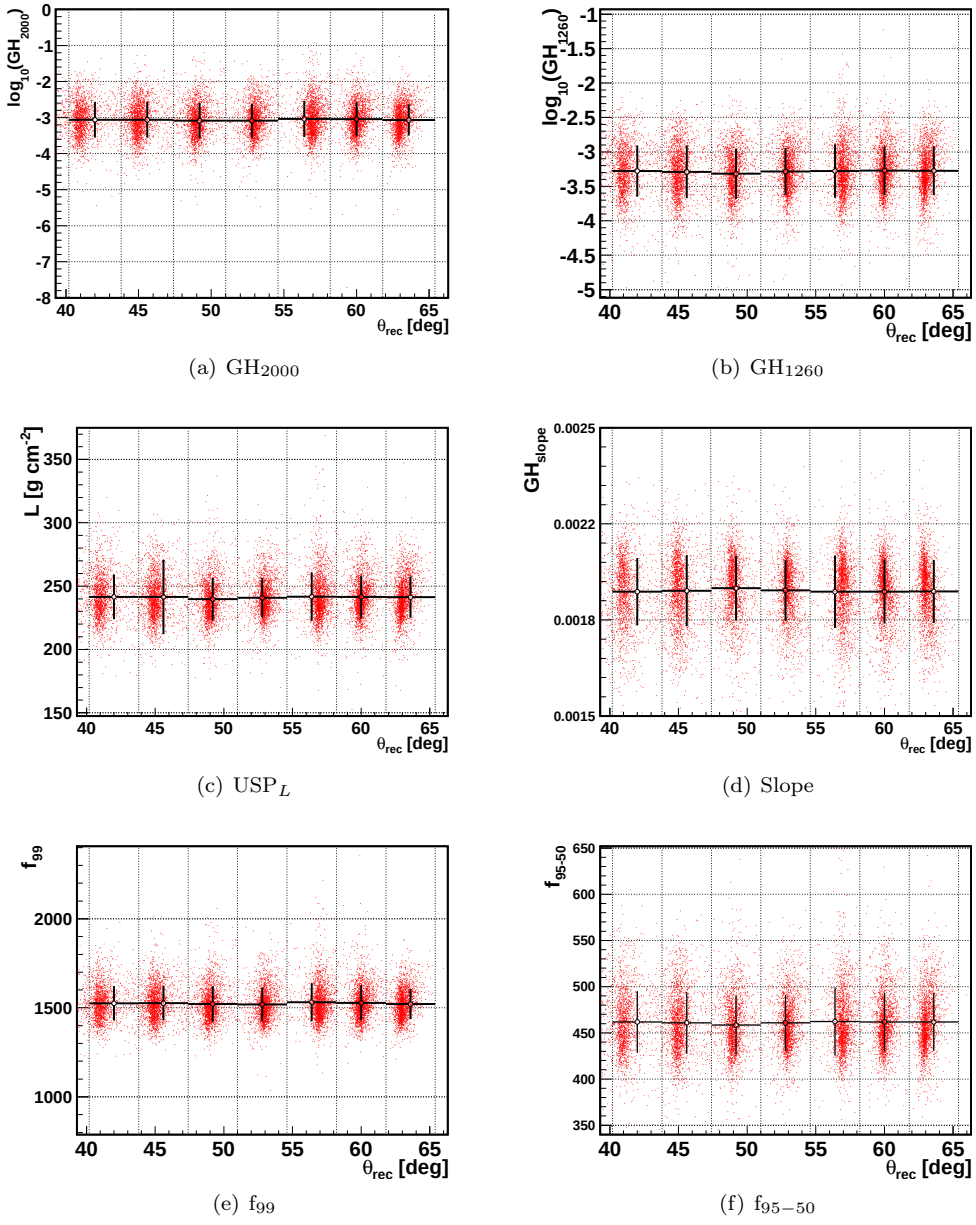
(a) GH_{2000} (b) GH_{1260}

(c) USP_L (d) Slope

(e) f_{99} (f) f_{95-50}

Figure 5.7: Distribution of the observables for background (solid lines) and signal (dotted and dashed lines) 10^{19} eV showers. All distributions normalized to integral 1.

reclassified. With rising number of boosts the performance improves, but too many boosts result in overtraining: the discriminant describes the training sample so accurately that applies poorly to the test sample. β is analogous to a "speed of learning". Large values make the updated weights rapidly closer to the original weights, while small values need a larger number of boosts to match the performance of BDTs with larger β. However, "fast learning" (large values of β and few boosts) and "slow learning" (small β and many boosts) are not equivalent and do not perform equally.



(a) GH_{2000} (b) GH_{1260}

(c) USP_L (d) Slope

(e) f_{99} (f) f_{95-50}

Figure 5.8: Zenith dependence of the variables used in the analysis. Dots indicate values for individual showers and black lines corresponds to the mean value and $\pm 1\sigma$ deviation of the variables in each zenith angle bin.

To find the optimal number of boosts, value of boost parameter β and maximum number of nodes, we train many discriminants scanning the parameter phase space. For each discriminant, we look for the best value of $S/\sqrt{B}$, where S and B are the number of signal and background events, respectively, above a certain value of the BDT response. In a first iteration we train trees with:

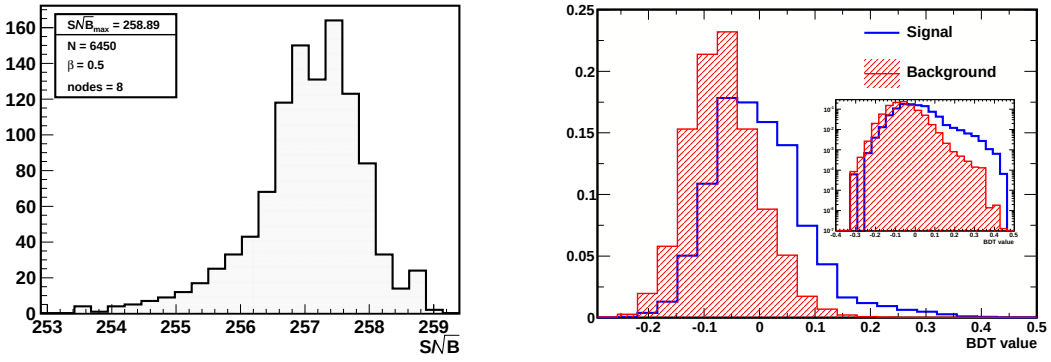
- number of boosts ranging from 2000 to 9000, in steps of 1000;
- β from 0.2 to 0.9 in steps of 0.05; and

- maximum number of nodes equal to 5, 8, 11, 14 and 17.

followed by a second iteration with smaller sampling in the region where the maximum seems to be. For this second iteration we use:

- number of boosts ranging from 6000 to 7000, in steps of 50; and
- β from 0.45 to 0.55 in steps of 0.025

In figure 5.9(a) we show the distribution of $S/\sqrt{B}$ values obtained after these two scans. We find the parameters of the best performing tree $N = 6450$, $\beta = 0.5$ and $n_{max} = 11$. We will adopt these parameters to build the BDT used in our analysis.



(a) Separation distribution

(b) BDT distribution

Figure 5.9: a) Distribution of $S/\sqrt{B}$ for all the simulated trees. b) BDT response distributions for signal (empty histogram) and background (hatched histogram), for the BDT generated with $(N, \beta, nodes)$ that maximizes $S/\sqrt{B}$.

The next step should be deciding where to place the cut on the BDT distribution. Even though signal showers have some distinct features that distinguish them from background showers, both are hadronic-initiated showers. In figure 5.9(b) we can see that the discriminant distribution is shifted to larger values in the case of signal showers, but there is a significant overlap. A cut following the criteria of reducing the background to 0 (such as in neutrino [336] and photon [337] searches) would considerably reduce the selection efficiency. Instead, we look for the value that maximizes $S/\sqrt{B}$ at each energy bin. Figure 5.11(a) shows the distribution of the discriminant values both for signal (contour plot) and background (black dots). In the same figure we show a linear fit to the cut value that maximizes the separation. The selection efficiencies of signal (full squares) and background (full circles) according to this cut are shown in figure 5.11(b). Signal showers have a selection efficiency that rises from ∼35% at $\log_{10}$(E/eV)= 18.5 to ∼60% at $\log_{10}$(E/eV)= 19.75. With rising primary energy, the average energy of the heavy quark component rises as well. Even if its energy as a fraction of the primary proton remains the same, the effects due to its propagation are more noticeable, and hence the selection efficiency improves. In contrast, even with rising energy the development of background proton showers remains regular. As such, the selection efficiency of background showers stays constant around ∼20%.

5.4.3 Event scan

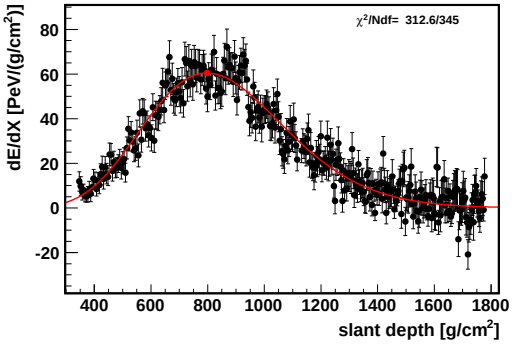
Since we have decided to set the cut on the discriminant maximizing $S/\sqrt{B}$ instead of seeking for 0 background events we find many events in real data passing the selection cuts. That does not mean they are signal, since they are compatible with the background expectation. Given that there are too many, a close inspection of each one of them is not viable. However, it might

be instructive to look at those selected in the high energy region, where there are few events selected. Inspecting the last two energy bins we find 7 events. None of these events shows a wrong reconstruction that could be responsible of their selection. As an example we show the reconstructed profiles of event with id 3167610, and analyze the signal-like features it presents.

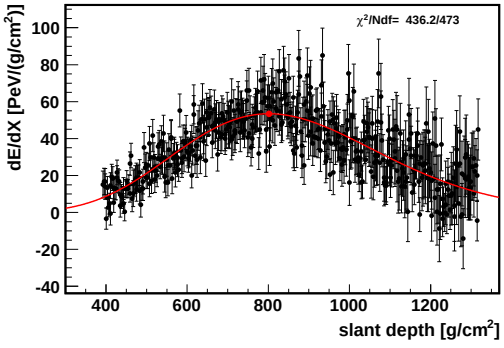
The event is stereo, detected by Coihueco and Los Leones sites. The reconstruction of the zenith angle, the energy and the X_{max} are in good agreement in the two telescopes.

$$\begin{bmatrix} \theta^{CO} = 65.2 \pm 0.3\,^{\circ} & \theta^{LL} = 65.3 \pm 0.2\,^{\circ} \\ E^{CO} = (4.25 \pm 0.28) \cdot 10^{19}\,\mathrm{eV} & E^{LL} = (3.80 \pm 0.27) \cdot 10^{19}\,\mathrm{eV} \\ X^{CO}_{max} = 798 \pm 5\,\mathrm{g\,cm^{-2}} & X^{LL}_{max} = 802 \pm 6\,\mathrm{g\,cm^{-2}} \end{bmatrix}$$

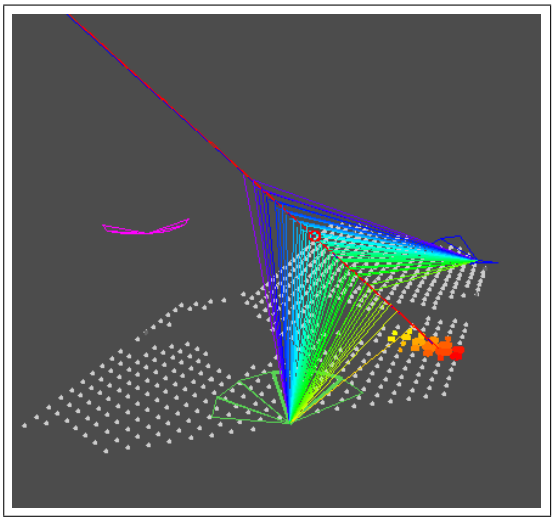
Even though the event does not show any striking feature we can recognize some of the charac-



(a) Coihueco



(b) Los Leones

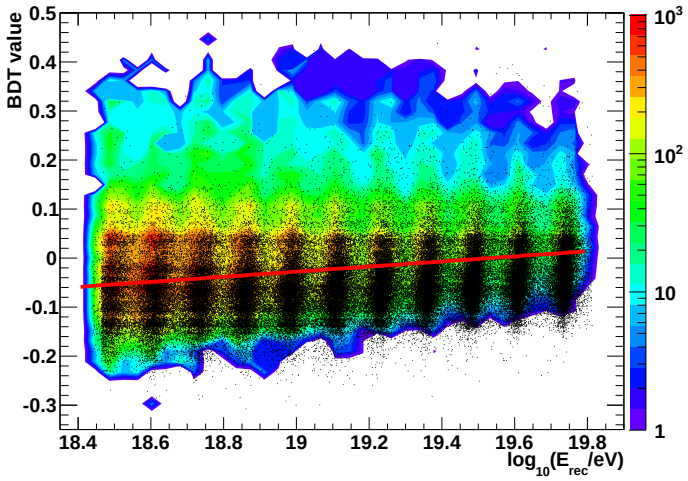


(c) 3d view

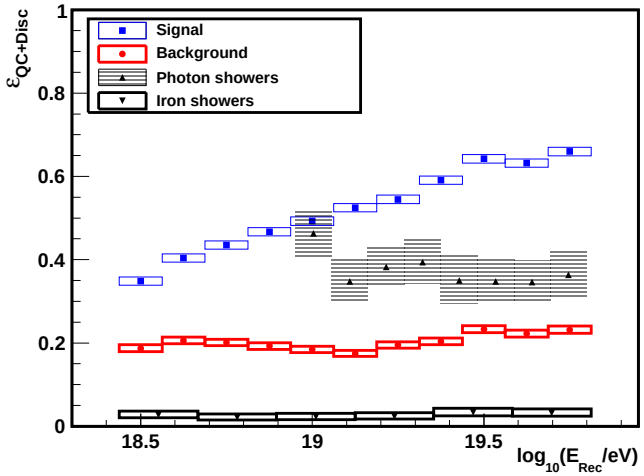
Figure 5.10: a) and b) show Coihueco and Los Leones sites shower profiles from event 3167610, respectively. c) shows the 3D view of the same event.

teristics of signal events. At $\theta = 65^{o}$ the atmosphere has a slant depth of $\sim 2100\,\mathrm{g{\cdot}cm^{-2}}$. This event has an EM component that almost reaches ground, still not absorbed above $1600\,\mathrm{g{\cdot}cm^{-2}}$. As such, this event will show large values of the ratios of particles (GH_{2000} and GH_{1260}) and a slow development (large f_{99} and f_{95-50}).

5.4.4 Sensitivity of the discriminant to nuclei- and photon-induced showers



(a) BDT distribution



(b) Selection efficiency

Figure 5.11: a) BDT distribution for signal (contour plot) and background (black dots). In the same figure we show a linear fit to the cut value that maximizes the separation. b) Overall selection efficiency (quality cuts and discriminant) for signal and background showers (squares and hatched circles, respectively), photon-initiated showers (triangles, hatch) and iron-initiated showers (triangles, no hatch).

The primary flux of cosmic rays is not composed solely of protons. In section 1.2 we showed that the evolution of X_{max} and X^{μ}_{max} with energy shows a composition turning heavier with increasing energy. The analysis in terms of $< \mathrm{Ln}A >$ and $\sigma_{<\mathrm{Ln}A>}$ show the same trend, confirming that the composition is not pure. From the analysis of X_{max}, a fit to the fraction of four different elements is done (p, He, N, Fe), obtaining that the most likely composition at the highest energies is a mixture of helium and nitrogen and barely no traces of iron. The fraction of protons is significant at energies between 10^{18} eV and $10^{18.5}$ eV but decreses with rising

energy. Figure 5.12 shows the mean values and uncertainties of the proton, helium, nitrogen and iron fractions (from top-left to bottom-right, respectively) obtained using the high-energy interaction model Sybill 2.1 and the low-energy hadronic model UrQMD [338]. Since there seem to be noticeable amounts of helium and nitrogen, we analyzed showers induced by these nuclei, finding a selection efficiency of 14.6% for helium-initiated showers and of 12.6% for nitrogen-initiated showers. Even though iron primaries seem to be absent at all energies, we considered iron showers as an extreme case. For these showers, the selection efficiency is reduced to ∼2-3% (down-pointing triangles in figure 5.11(b)).

Heavy quark production by the Intrinsic Quark model can occur in the collisions of any nuclear primary with air. Moreover, with rising mass number, the probability that *any* of the nucleons of the projectile will develop a $Q\bar{Q}$ fluctuation and produce a heavy hadron increases. However, each of the nucleons shares E_0/A of the primary energy. The energy of any heavy hadron produced will be a fraction of E_0/A. Considering helium nuclei, the lightest nuclei after proton and as such the most favorable case, a heavy hadron produced will carry, on average $< x_Q > \cdot E_0/4 = 0.34 \cdot E_0/4 \sim 0.08 \cdot E_0$, a 8% of the primary energy. Taking the case of an iron nuclei, this value decreases to $< E_Q > \cdot E_0/56 = 0.34 \cdot E_0/56 \sim E_0 \cdot 6 \cdot 10^{-3}$. As we presented in figure 4.21, the effects of bottom hadrons propagation on shower development starts being noticeable when the hadrons carry fractions of the primary energy above 30-40% (and even larger fractions in the case of charmed hadrons). Even if the heavy hadron produced would propagate in the atmosphere, no effects would be noticeable. All this considered, the selection efficiency for showers induced by nuclei heavier than proton will be equal for showers producing heavy hadrons and for showers without heavy hadrons.

Regarding photons, in section 1.4.1 we discussed their possible sources and showed that no primary photons have been so far identified in the data collected by the Pierre Auger Observatory [4]. Upper limits to the fraction of photons have been derived instead. We should test whether photons are an important source of background, or if we can safely neglect them. Showers initiated by UHE photons develop differently from showers induced by nuclear primaries. Photon showers are almost purely electromagnetic. Electromagnetic interactions have smaller multiplicities compared to those of hadrons, and showers are expected to develop deeper in the atmosphere, yielding a larger X_{max}. In addition, above 10^{18} eV the Landau-Pomeranchuk-Migdal (LPM) effect further delays the shower development by suppressing Bremsstrahlung and pair-production cross-sections [337]. All this considered, we expect photon showers to be very deep, which could make them to be mistakenly identified as deep proton showers with a heavy hadron component. However, the development of photon showers is very regular compared to that of hadron initiated showers. Particle production is due to pair production and bremsstrahlung. It is unlikely that we will find large elasticity inequalities, or leading components inside the shower. We used a sample of photon showers with $\log_{10}(\mathrm{E/eV})$ between 19 and 19.7, in steps of 0.1 in $\log_{10}(\mathrm{E/eV})$. The selection efficiency for photon showers is ∼35% and approximately constant with energy (up-pointing triangles in figure 5.11(b)). According to [339] the upper bound on the fraction of photons for energies above 3 EeV is 1.0%. Considering this fraction and the efficiency for photon showers we expect less than 10 photon events in a sample of roughly 2000 events. Such a small fraction of photon events will not significantly affect our final result.

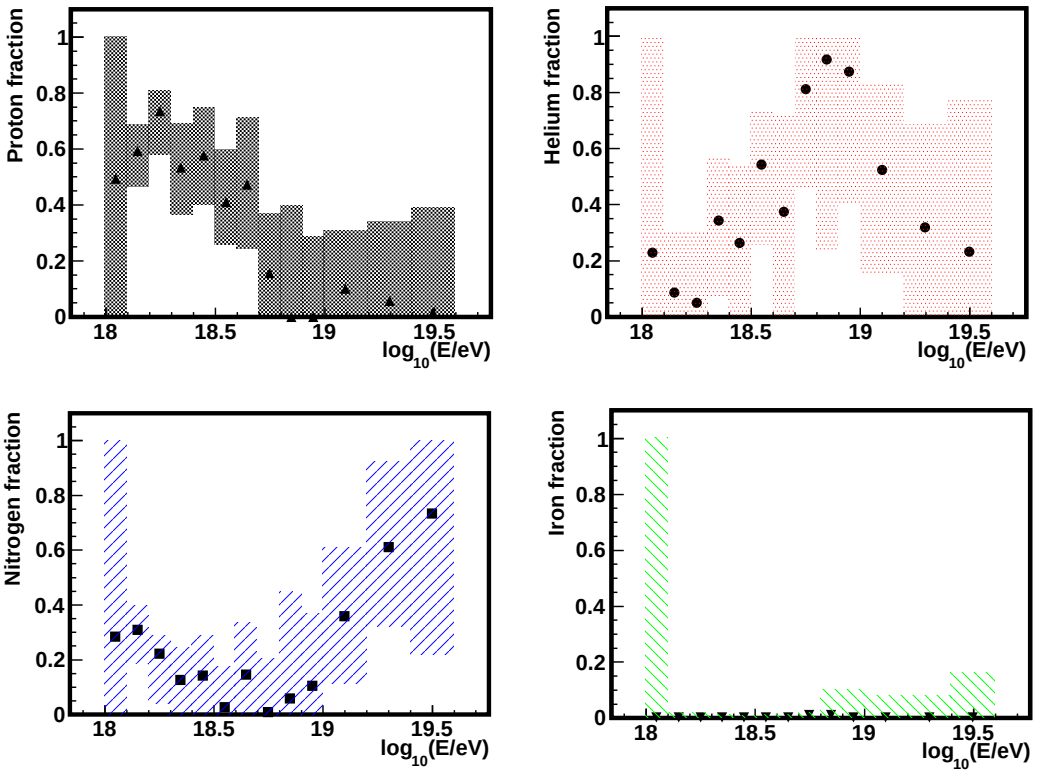


Figure 5.12: Best fit to the fraction of proton, helium, nitrogen and iron nuclei in the primary cosmic ray flux using the hadronic model Sybill 2.1 and the low-energy hadronic model UrQMD. The shaded regions represent systematic uncertainties.

In this chapter we have seen that the simulation of showers with heavy quark production in the Pierre Auger Observatory involves three different steps:

- The generation of the Extensive Air Shower, with the Monte-Carlo code CORSIKA.
- The detector response to the shower, using the $\overline{\text{Off}}\underline{\text{line}}$ framework.
- The reconstruction of the physical parameters of the shower, within the $\overline{\text{Off}}\underline{\text{line}}$ framework as well.

We have used a set of **high-quality reconstructed showers** to identify variables sensitive to the production and propagation of heavy hadrons. We have combined these variables through a multivariate discriminant, namely a **Boosted Decision Tree**, optimized to separate showers with and without heavy quark production. We have obtained a **signal selection efficiency** $\sim 35\%$ at unit[$\log_{10}$](E/eV)=18.5 that rises to $\sim 65\%$ at unit[$\log_{10}$](E/eV)=19.75. The **background selection efficiency** stays at 20% at all energies. We have checked the energy dependence of the selection efficiency for nuclear primaries other than protons and for photon showers. The larger number of nucleons in nuclei compared to protons make them unsuitable

to generate showers with heavy hadrons that could be detected. The low flux of photons turns them into a negligible source of background.

6

Seach for heavy quarks using the Pierre Auger data

In the previous chapter we described Boosted Decision Trees (section 5.4.2), a multivariate discriminant method. Using the variables defined to identify showers containing ultra-high energy heavy hadrons, the discriminant yielded a certain signal and background selection efficiencies (section 5.4.4). These efficiencies are necessary to compute the detector exposure to the signal and background components and eventually the rates of expected events.

In section 6.1 we introduce the concepts of effective area, aperture and exposure of a cosmic ray detector. To compute them, a careful monitoring of the detector status is needed, which is explained in sections 6.2 to 6.4. Finally, in section 6.5 we detail the calculation of the exposure, its systematics, and compute the rate of expected events, from which we obtain a limit to the cross-section of heavy quark production in the Intrinsic Quark model.

6.1 Effective area, aperture, exposure

The effective area of a cosmic ray detector is defined as:

$$A_{eff} = \int_A \varepsilon \cos\theta \, dA \tag{6.1}$$

where ε is the detection efficiency, $\cos\theta \, dA$ is the differential area projected onto the normal direction of the incoming shower, θ is the shower zenith angle and A is the area where shower events hit the ground.

In our analysis, ε is the overall detection efficiency, including detection, reconstruction and event selection, and is a quantity that depends on the energy of the shower, its direction and core position, and the configuration of the array at a certain moment. The aperture can be expressed as the integral of the effective area over the whole detection solid angle:

$$\mathscr{A}(E,t) = \int_\Omega A_{eff} \, d\Omega = \int_\phi \int_\theta \int_A \varepsilon \, \sin\theta \, \cos\theta \, dA \, d\theta \, d\phi \tag{6.2}$$

The aperture of a cosmic ray instrument is per se a figure of merit of its observation capability. The time integrated aperture is commonly referred to as the exposure:

$$\mathscr{E}(E) = \int_T \mathscr{A}(E,t) \, dt = \int_T \int_\phi \int_\theta \int_A \varepsilon \, \sin\theta \, \cos\theta \, dA \, d\theta \, d\phi \, dt \tag{6.3}$$

The exposure is needed to compute the expected number of detected events. Given an incoming flux of primary particles, $\Phi(E)$, and a selection efficiency, ε, the number of detected events per

energy bin dE, solid angle $d\Omega$, time interval dt surface element dA is:

$$\frac{dN}{dE\, d\Omega\, dA\, dt} = \Phi(E)\, \mathscr{E}(E, \Omega, A, t) \tag{6.4}$$

The total number of expected events is obtained integrating over all the variables above:

(6.5)

If we want to predict the number of background and signal events with certain confidence, we have to consider that the flux is a mixture of various components, which have different selection efficiencies (and hence different exposures). We can rewrite the equation above as:

$$N = \sum_A N^A = \sum_A \int_E \Phi^A(E)\, \mathscr{E}^A(E)\, dE \tag{6.6}$$

$$= \sum_A \int_E \Phi(E)\, f_A(E)\, \mathscr{E}^A(E)\, dE \tag{6.7}$$

where $\Phi^A(E) = \Phi(E) f_A(E)$ is the flux of nuclei elements with mass A, f_A is the fraction of nuclei of elements with mass A at energy E, and $\mathscr{E}^A(E)$ is the exposure of the detector for those nuclei.

In ideal conditions, all detectors function perfectly during the whole working time. However, under real conditions, experiments often start taking data before they are completely finished and detectors stop working due to temporary hardware problems and environmental conditions. As such, the exposure of the observatory depends on the detector configuration over time. If the detector configuration changes, the exposure has to be evaluated with the new configuration. If we want to predict accurately the number of expected events in an experiment, we need to take into account the dynamic nature of the observatory.

Our simulations do not reflect the changing status of the array. They were performed always assuming an ideal array, where all stations are deployed and working all the time. However, the status of the SD and the FD are both continuously and carefully monitored and we can use this information to replicate the status of the detector at any time.

6.2 SD monitoring

The surface detector configuration has been continuously changing over the period of data collection. The disposition and number of working stations of the array changed over time because new tanks were deployed since the beginning of the experiment until the completion of the array, making the size of the array dependent on time. Besides, even in a steady configuration, some SD stations are temporarily out of service at any time. The SD status is monitored by updating each second the list of active stations. The response of each tank affects a small region around its position, which in turn reflects directly onto the local trigger efficiency.

A SD hexagon is defined by the surface spanned by one station and the crown of its first neighbors. The so called *Hexalife*[1] files contain a monthly monitoring of the SD hexagons. For each hexagon, the file contains an entry with its position, the gps time when it started being active, and the gps time when it stopped being active; both times are referred to the beginning of the month.

However, the use of the lifetime files alone is not sufficient for the evaluation of a correct exposure. These files indeed give information on the single stations trigger rates but not on the performance of the central trigger (CT) and CDAS. Time periods for which the rate of T5 events is below a certain threshold value are rejected. These periods are termed *Bad Periods*[2].

[1] SD stations and hexagons lifetime files can be found at
http://ipnweb.in2p3.fr/∼auger/AugerProtected/AcceptMain.html

[2] The up to date list of Bad Periods is available at
http://ipnweb.in2p3.fr/∼auger/AugerProtected/AcceptBadPeriods.html.

The main contribution to Bad Periods comes from time intervals when the CDAS was stopped and thus no events were acquired.

6.3 FD monitoring

The efficiency of fluorescence and hybrid data taking are influenced by many effects. These can be external (lightning or storms) or internal (data acquisition system failures). The determination of the hybrid *on-time* and the successful data-taking periods of the Pierre Auger Observatory needs taking into account all these occurrences. Data losses and inefficiencies can occur at different levels, from the photomultiplier units of the FD (pixels) to the joint SD-FD data taking of the observatory.

The FD *Uptime*[3] files contain the details of the status and condition of each telescope and FD station averaged over periods of $\Delta T_{UP} = 10$ minutes. During each period, each of the 6 telescopes in a station may be active or not, and able to send data or not, independently of the status of the others. The FD monitoring database began to be filled during 2007, but part of the monitoring database was not available until 2008, and the uptime is calculated differently before and after this date. Before the end of 2007[4] the fraction of time the *j-th* telescope in the *i-th* station was running and recording reconstruction-worth data (**up**) is calculated as:

$$f_{i,j(\Delta T)} = (U_{i,j} \times \delta_{i,j}) \cdot (U_i \times \delta_i) \cdot (U_{CDAS} \times \delta_{CDAS}) \tag{6.8}$$

where $\mathbf{U_{i,j}}$, $\mathbf{U_i}$ and $\mathbf{U_{CDAS}}$ are the fractions of ΔT_{UP} that the *j-th* telescope in the *i-th* station, the *i-th* station and CDAS were taking data, respectively. $\delta_{\mathbf{i,j}}$, $\delta_{\mathbf{i}}$ and $\delta_{\mathbf{CDAS}}$ are their statuses: a running data acquisition system does not mean that the detector is working properly. If the working conditions of the telescope, the station or CDAS are considered to be inadequate, the corresponding status is set equal to 0.

From 2008 onward, the status is calculated as:

$$f_{i,j}(\Delta T) = U_{i,j} \times U_i \times U_{CDAS} \times V_{FDAS} \times V_{CDAS} \tag{6.9}$$

where $U_{i,j}$, U_i and U_{CDAS} have the same definition as in equation 6.3 and $\mathbf{V_{FDAS}}$ and $\mathbf{V_{CDAS}}$ correspond to vetoes on data acquisition imposed by the FDAS[5] and the CDAS, respectively. The former is applied when the delay between forming a FD-T3 at site level and sending it to CDAS is larger than 10 seconds. The latter is used to prevent the acquisition of nonphysical events when FD-T3 trigger[6] rates are too high (>0.1 Hz) due to lightnings or scattered laser shots. Thus, for each period ΔT_{UP} we calculate $f_{i,j}(\Delta T)$, the fraction of that interval telescope j in station i was working.

In addition to SD Bad Periods, there is a small list of FD Bad Periods in which one or more FD stations, or telescopes within FD stations, were not working properly[7]:

- Between 14/2/2008 and 23/10/2008 there was a failure in the GPS clock from Loma Amarilla. During this period, events from Loma Amarilla should be skipped[8].
- Unstable baselines: Los Leones, telescope 4 had an ill defined baseline. Between 19/3/2009 and 10/10/2009[9], if telescope 4 was taking data, events should be skipped. For Loma Amarilla, telescope 3, events between 31/3/2010 and 11/5/2010[10] should be skipped if telescope 3 was active.

[3]With previous authorization the *Uptime* files can be downloaded from http://augerobserver.fzk.de/doku.php?id=datatree:uptime and http://paomon.physik.uni-wuppertal.de/UpTime.

[4]GPS second 883224013.

[5]Fluorescence Data Acquisition System.

[6]The third level trigger of the FD is explained in section 3.2.2.

[7]The FD Bad Periods can be obtained from modules FDSelection.cc and AugerUpTime.cc of the ADST package [324].

[8]The GPS times corresponding to these dates are 887000000 and 908800000.

[9]GPS seconds 921542451 and 944524815.

[10]GPS seconds 954115215 and 957657614.

- Even if uptime information exists, each station has a first valid date, before which it does not contribute to the exposure. Los Leones and Coihueco first valid date is 1/12/2004. The first date for Los Morados is 2/6/2005. For Loma Amarilla, the last instrumented station, the date is 1/5/2007.

6.4 Atmospheric monitoring

Atmospheric fluctuations have a significant impact on measurements performed with the FD. Besides replicating the changing status of the array due to technical reasons, to compute the exposure we ought to mimic the atmospheric conditions of the observatory as well. A detailed monitoring is carried to record, in short time intervals, a large number of variables describing the status of the atmosphere (see section 3.2.3) Temperature, humidity, aerosol concentration at different heights and cloud coverage (among others) are measured over each FD station. Atmospheric data is structured in MySQL databases. The database *Atm_Aerosol_1_A* contains the values of the vertical aerosol optical depth (VAOD) at different heights, recorded with the CLF, in intervals of 1 hour. Cloud information is stored in the database *Atm_Quality_0_A* [11]. This data is used extensively to grant an accurate reconstruction, correcting the measured data due to the effect of the atmosphere when needed. However, under very unfavorable weather conditions very large corrections are needed and the corresponding periods are removed from the data taking process:

aerosols: the value of the Vertical Aerosol Optical Depth (VAOD) at 4.5 km is the reference normally adopted to decide if the content of aerosols in the atmosphere is too large. Approximately 5% of CLF measurements have VAOD (4.5km) greater than 0.1. To avoid making very large corrections to the expected light flux from distant showers, these periods are not used in the FD analysis [175].

cloud coverage: the presence of clouds has a major influence on the reconstruction of air showers. Clouds can either block the transmission of light from air showers, or enhance the observed light flux due to multiple scattering of the intense Cherenkov light beam. Clouds can reduce the event rate from different parts of the fluorescence detector, affecting the exposure of the detector. Periods with a cloud coverage greater than 25% are rejected [179].

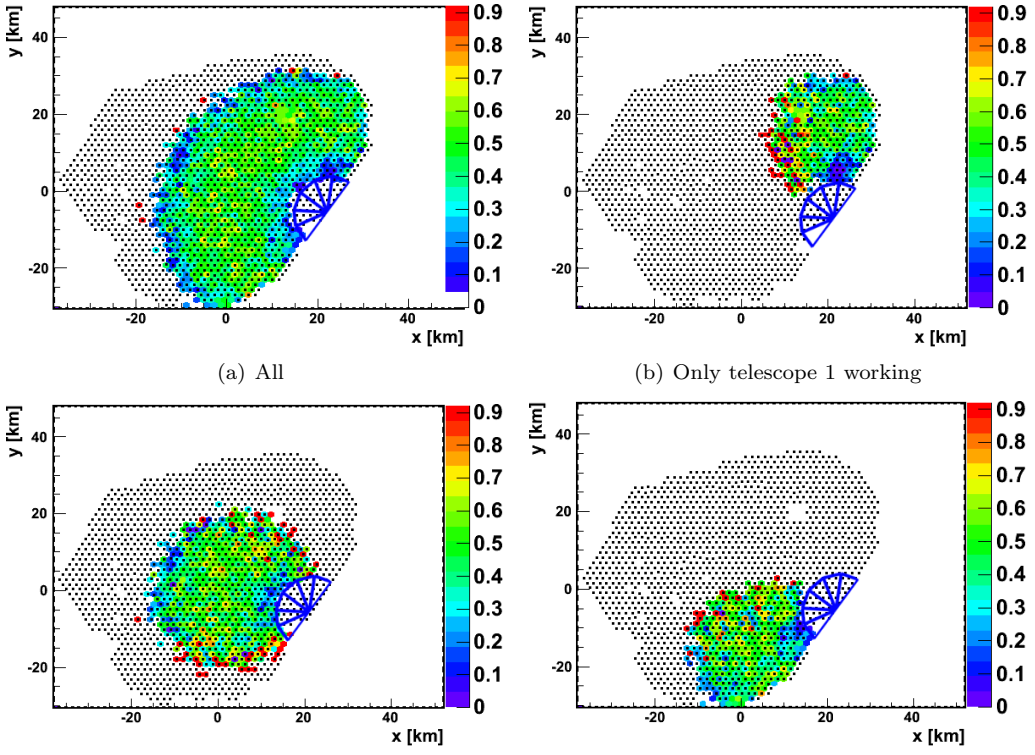
Atmospheric databases contain hourly information of aerosols, coming from CLF and Lidars, and of cloud coverage, measured by the Lidar stations and infrared cloud cameras (IRCC) installed on the top of each FD building. Events detected by a station during poor atmospheric conditions should be discarded, and these stations should be accordingly removed from the exposure calculation.

6.5 Exposure calculation

The calculation of the detector exposure is key to make a prediction of the number of expected events of a certain kind. Given the configuration of the real detector at time t, obtained from the monitoring files, the usual approach to calculate the hybrid aperture consists in replicating the real configuration of the detector at that time and then simulating and reconstructing a large sample of showers. The selection efficiency for a certain shower energy, direction and core position is given by the ratio of showers passing the selection cuts over the total number of simulated showers. The hybrid exposure is obtained repeating this process during the lifetime of the detector, updating the detector configuration over time. This process is extremely time-consuming, given the enormous phase space defined by all the possible combinations of energy, direction and core position for a certain configuration, and the many different configurations over the functioning time of the detector [340].

[11] Atmosphere MySQL databases are available at *offline.ung.si*, with prior authorization

For this analysis we have adopted a different approach. As mentioned in section 5.1 for each of the energy and heavy hadron bins we have simulated 2000 CORSIKA showers with zenith angle θ=60 °. Each shower has been reconstructed five times, to a total of 10000 reconstructed events for each energy and heavy quark type bin. Each of these showers has a random core position distributed over the array and is independently detected by any of the four FD stations. Given a shower with energy E and core position $\vec{r}$, detected by one or more FD stations we can calculate the probability that it will pass the selection cuts: the number of showers with that energy and position that pass the cuts divided by the total number of showers with that same energy and position. We can access the information of which FD stations, and which telescopes within them, detected a shower. As such, we can build efficiency maps for different configurations of working stations and telescopes. As an example, in figure 6.1 we show the quality cut selection efficiency for 10^{19} eV showers of Los Morados station when all telescopes are working (6.1(a)), and when only telescopes 1 (6.1(b)), 3 (6.1(c)) or 6 (6.1(d)) are working.



(a) All

(b) Only telescope 1 working

(c) Only telescope 3 working

(d) Only telescope 6 working

Figure 6.1: Selection efficiency for different configurations of working telescopes.

Now we face the problem of determining the selection efficiency during a particular period of time, where the conditions of the array are extracted from the monitoring information available. The probability that the *i-th* telescope in the *j-th* station will detect a shower initiated by a nucleus with mass A and energy E falling at position $\vec{r}$ is $p^A_{i,j}(E,\Omega,\vec{r})$. However, not all telescopes are continuously working. The detection probability over a period ΔT is obtained weighting the detection probability of each telescope by the fraction of ΔT it has been working, $f_{i,j}(\Delta T)$ (obtained from the uptime files):

$$\varepsilon^A_{i,j}(E,\Omega,\vec{r},\Delta T) = f_{i,j}(\Delta T) \times p^A_{i,j}(E,\Omega,\vec{r}) \tag{6.10}$$

We assume that a shower is detected by a station if any of the telescopes within detects it:

$$\varepsilon^A_{j}(E,\Omega,\vec{r},\Delta T) = \cup_{i=1}^{6}\, \varepsilon^A_{i,j}(E,\Omega,\vec{r},\Delta T) \tag{6.11}$$

where

$$\varepsilon^A_{i,j_1} \cup \varepsilon^A_{i,j_2} = \varepsilon^A_{i,j_1} + \varepsilon^A_{i,j_2} - \varepsilon^A_{i,j_1} \cdot \varepsilon^A_{i,j_2} \tag{6.12}$$

if ε^A_{i,j_1} and ε^A_{i,j_2} are not mutually exclusive (more than one telescope may detect the shower at the same time) and are independent.

$\varepsilon^A_j(E, \Omega, \vec{r}, \Delta T)$ is the probability that a shower with energy E falling on $\vec{r}$ will be detected by station j. If any of the four FD stations detects it, the shower will be detected:

$$\varepsilon^A(E, \Omega, \vec{r}, \Delta T) = \cup^4_{j=1} \varepsilon^A_j(E, \Omega, \vec{r}, \Delta T) \tag{6.13}$$

where the operator $\cup$ has the same meaning as in equation 6.5.

To compute the exposure first we loop over each 10-minutes entry of the FD uptime files. Figure 6.2 shows an schematic of the exposure calculation. For each ΔT_{UP} period:

- First, we establish the status of the FD detector. We compute which fraction of ΔT_{UP} each eye was working during this time. If no eye was working during this period, we skip it and move to the next period.
- Then, we check if the period coincides with a SD Bad Period. If it falls completely inside, we skip it and move to the next period. If there is only an overlap between the current ΔT_{UP} and the Bad Period, we move the limits of the integration window until there is no overlap with the Bad Period.
- We check the atmospheric conditions over each FD station corresponding to the current time period. If the VAOD was larger than 0.1 or the cloud fraction was above 25% (hatched regions in figure 6.2 (top)), the FD station is removed from this period.

Using the criteria above we divide each ΔT_{UP} uptime period in sub-periods with a fixed number of FD stations working ($\Delta T_{UP,j}$ in figure 6.2 (top)).

The *HexaLife* files provide the changing status of one hexagon over each $\Delta T_{UP,j}$ (rows in figure 6.2 (bottom)). Hexagons are either flagged as working (full rectangles) or "dead" (empty rectangles). The effective area and exposure of a single hexagon located at position $\vec{r}_i$ for a shower with energy E and direction Ω is:

$$A^A_{eff}(E, \Omega; \vec{r_i}, \Delta T_{UP,j}) = A_{hex}\, \varepsilon^A(E, \Omega; \vec{r}_i, \Delta T_{UP,j}) \tag{6.14}$$

$$\mathscr{E}^A(E, \Omega; \vec{r_i}, \Delta T_{UP,j}) = A_{hex}\, \varepsilon^A(E, \Omega; \vec{r}_i, \Delta T_{UP,j}) \sum_{t=t^j_0}^{t=t^j_f} \delta_{T2}(\vec{r_i}, t) \tag{6.15}$$

$$= A_{hex}\, \varepsilon^A(E, \Omega; \vec{r}_i, \Delta T_{UP,j})\, t_w(\vec{r_i}; \Delta T_{UP,j}) \tag{6.16}$$

where δ_{T2} is the status of the *i-th* hexagon at time t, t_w is its integrated working time of the *i-th* station during ΔT_j, and A_{hex} is the area of an hexagon. The exposure of the whole array during $\Delta T_{UP,j}$ is equal to the sum of the exposure of all hexagons:

$$\mathscr{E}^A(E, \Omega; \Delta T_{UP,j}) = \sum_{i=0}^{i=n_{hex}} \mathscr{E}^A(E, \Omega; \vec{r_i}, \Delta T_{UP,j}) \tag{6.17}$$

Finally, the exposure of the array for showers with energy E and direction Ω during the whole working time of the array is:

$$\mathscr{E}^A(E, \Omega) = \sum_{j=start}^{j=end} \mathscr{E}^A(E, \Omega; \Delta T_{UP,j}) \tag{6.18}$$

We have applied this procedure to compute the exposure at zenith angles $\theta = 40^o, 45^o, 50^o, 55^o, 60^o, 65^o, 70^o$. At zenith angles different of 60° we scale the selection efficiency from 60° to that angle according to the dependence shown in figure 5.4(b). Then, using the scaled efficiency we compute the exposure at that particular zenith angle. At zenith angles below 40° the atmosphere is not deep enough to observe any effects due to the propagation of heavy hadrons. Above 70° both the selection efficiency and aperture decrease, and we expect few events.

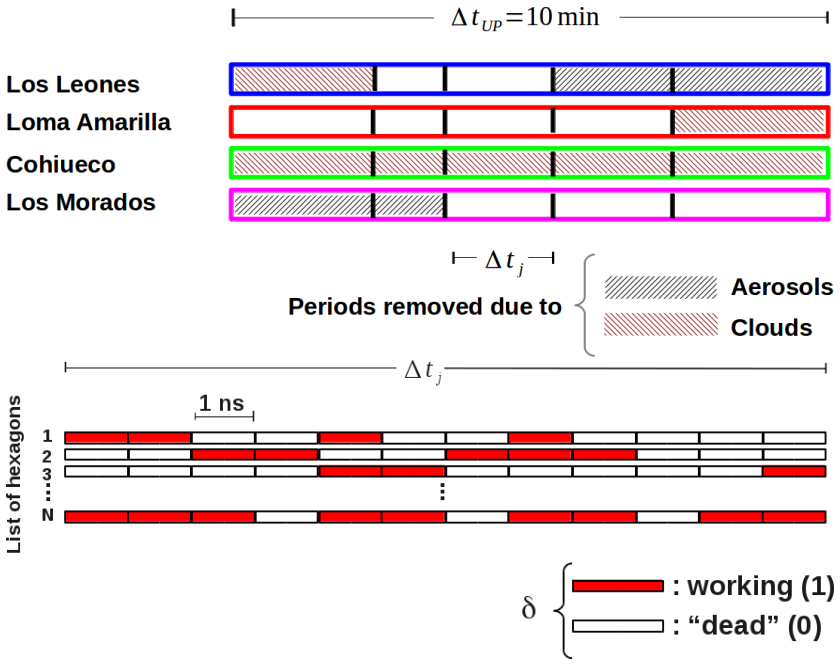


Figure 6.2: Top: Representation of the partition of a 10 minutes uptime period depending on atmospheric conditions. Bottom: Representation of the status monitoring of the surface stations. During each nanosecond the station is either flagged as "working" (full rectangles) or "dead" (empty rectangles).

6.5.1 Systematics

The calculation of the exposure incorporates several quantities, each having its own uncertainty. An exhaustive calculation of the systematic uncertainties of the exposure would entail recalculating the exposure varying each of the parameters according to their uncertainties. Since this is extremely demanding computing- and time-wise, we will only compute the systematic this way whenever viable, looking for different ways of estimating the systematic otherwise.

Energy Scale An incorrect assignment of the energy results in upward or downward displacement of the detection efficiency, and consequently affects the exposure (energy scale). The energy scale contains the total uncertainty deriving from the fluorescence yield, FD calibration, FD profile, and atmosphere. It amounts to a total of a 14% [182] The uncertainty on the energy scale results in a $[+2.4\%, -3\%]$ uncertainty on the exposure.

Bad Periods The exposure depends critically on the correct determination of FD and SD badtimes and uptimes. As a result of the checks with CLF laser shots and between SD data and the Monte Carlo simulations, the exposure should be reduced by 8% to account for lost events [341].

Hadronic models the physical description is not identical between different hadronic models. The extrapolations from accelerator data to cosmic ray energies differ, resulting in different cross-sections, abundances of particles and elasticities. Overall, all hadronic models give a good description of the cosmic ray physics, and we do not expect large differences in the shower development. We compare our simulations, performed using QGSJet01c, with proton simulations using Sybill and QGSJetII, and assume that any systematic difference in the result affects equally signal and background showers. For the energy bins used, $10^{18.75}$ eV and $10^{19.25}$ eV, the results are systematically higher for both hadronic models compared to QGSJet01c, ∼7% for Sybill and ∼3.6% for QGSJetII. Taking the largest

difference, we quote a systematic uncertainty due to the change of the hadronic model of +7%.

Atmosphere The properties of the atmosphere affect the propagation of fluorescence light from the shower to the detector. Since the atmospheric conditions change with the season, there might be an effect associated to assuming an static atmosphere. The effect of the seasonal variation of the atmosphere has been estimated to be ∼4-5% [337].

Mass composition The fractions of proton, helium and nitrogen shown in figure 5.12 have an uncertainty represented by the hatched bands. According to the values and their uncertainties, we can compose the heaviest possible composition assuming, in each bin, the maximum fraction of nitrogen possible, then the maximum fraction of helium given the former nitrogen fraction and let the rest of the composition be protons. Maximizing first the fraction of protons, then that of helium finally nitrogen we can determine the lightest possible mass combination. The result obtained using the lightest (heaviest) composition instead of the central values is 11% (9%) higher (lower). The fraction estimates presented in figure 5.12 were calculated using the high-energy hadronic interaction model Sybill 2.1 and the low-energy hadronic model UrQMD. The estimates computed using other model show slight deviations, but all are compatible within their systematic uncertainties.

Shower profile reconstruction The variables used on the analysis are extracted from the fit to the longitudinal profile. The limited field of view of the fluorescence detector and the propagation of light from the shower to the telescopes result in a degree of uncertainty in the fit to the longitudinal profile. We varied the parameters of the Gaisser-Hillas fit (X_{max}, X_0, λ)[12] according to their uncertainties. We obtained a systematic uncertainty $[+1.8\%, -2\%]$ for the selection efficiency, from which an uncertainty of the same magnitude on the exposure follows directly.

The total systematic uncertainty amounts to $[+13.9\%, -10.5\%]$ for the exposure after considering all the contributions above.

Source		
Energy scale	+2.4%	-3%
Gaisser-Hillas parameters	+1.8%	-2%
Atmosphere (seasonal)	+4%	-4%
Hadronic model	+7%	
Mass composition	+11%	-9%
Total	+13.9%	-10.5%

Table 6.1: Main sources of systematic uncertainty and their influence on the exposure.

6.6 Data analysis and experimental results

Using equation 6.1, the exposure for background showers derived in the previous section and the fractions of each element from figure 5.12, we can compute the expected number of events that should pass all the selection cuts in case there is no heavy quark production through the Intrinsic Quark model. As the selection efficiencies of background showers initiated by primaries heavier than protons we have used those derived in section 5.4.4. From the sample of real data, we analyzed showers between January 2004 and December 2012, in the energy range $E \in [2.7{\cdot}10^{18} - 6.5{\cdot}10^{19}]$ [eV]. In figure 6.3 we can see the predicted number of background events compared to the number of selected events in the sample of real data. Solid bars correspond to

[12]None of the observables used depends on the size of the Gaisser-Hillas function, $(dE/dX)_{max}$. As such, there is not a systematic associated to its uncertainty.

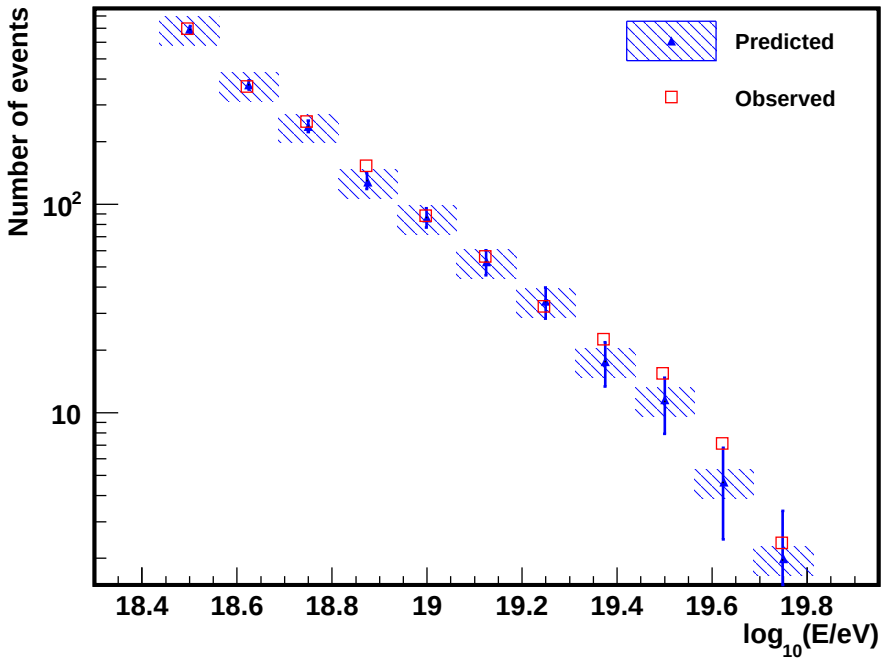


Figure 6.3: Predicted and observed number of events passing the selection cuts in the period between January 2004 and December 2012. The hatched regions corresponds to the systematic uncertainties. Solid bars represent the statistical uncertainties.

$\log_{10}(E/eV)$	N_{obs}	N_{pred}
18.5	691	$691.85^{+138.37}_{-117.61}$
18.625	366	$374.43^{+74.89}_{-63.65}$
18.75	247	$236.24^{+47.25}_{-40.16}$
18.875	150	$128.26^{+25.65}_{-21.80}$
19	86	$85.92^{+17.18}_{-14.60}$
19.125	54	$53.03^{+10.60}_{-9.02}$
19.25	31	$34.29^{+6.57}_{-5.83}$
19.375	22	$17.64^{+3.53}_{-3.00}$
19.5	14	$11.82^{+2.36}_{-2.00}$
19.625	6	$4.80^{+0.96}_{-0.82}$
19.75	1	$2.09^{+0.42}_{-0.36}$
total	1668	$1639.81^{+327.8}_{-278.6}$

Table 6.2: Number of observed and expected events.

the statistical uncertainties, while the shaded region accounts for the systematic uncertainties. The number of predicted and observed events in each bin is shown in table 6.2.

The possible room for a signal corresponds to the difference between the observed and predicted number of events. There are bins where the predicted background is larger than the number of detected events. In this bins, we would find a *negative* signal, which does not make

sense. Instead, we compute the upper and lower limits to the number of signal events corresponding to the measured and predicted background. There is not a unique way of calculating confidence intervals, and different statistical methods result in different values of the upper and lower limit. The classical construction, due to Neyman [342], leads to nonphysical or empty set results when less events than background are found. An alternative, due to Feldman and Cousins [343], solves this issue applying a particular ordering scheme, based on likelihood ratios. Following their prescription, upper limits or confidence belts are derived.

The Feldman and Cousins approach is implemented as a class in the ROOT framework (*TFeldmanCousins*[13]). Using 90% confidence levels, for 1668 events observed and 1639.81 detected, the lower and upper limits to the number of signal events is

$$N_{low}^{FC;90\%CL} = 0 \tag{6.19}$$

$$N_{up}^{FC;90\%CL} = 96.73 \tag{6.20}$$

The Feldman-Cousins treatment is a fully frequentist construction and should be used in case of no (or negligible) uncertainties. However, it is not capable of handling background expectations which are known only with some limited accuracy, and does not account for the possible existence of systematics in the efficiencies of signal and background. The method developed by Rolke, López and Conrad [344] allows for the calculation of limits when the distributions have statistical and systematic errors. Using this approach we can incorporate the systematic errors computed for the exposure in section 6.5.1 in the derivation of the lower and upper limit of the number of signal events. The systematic uncertainty on the exposure to background showers was estimated to be $[+13.9\%, -10.5\%]$. The 14% systematic uncertainty on the energy scale results into a systematic uncertainty of the same magnitude in the flux, the other ingredient along with the exposure in the calculation of the number of events. Since the number of events depends on their product we add both systematic uncertainties in quadrature, finally quoting a $[+19.7\%, -17.5\%]$ systematic in the number of events.

For a number of observed events $n_{obs} = 1668$, a number of background events $n_{exp} = 1639.81$ and a Gaussian distribution of the background systematic uncertainty, the limits to the number of signal events in the Conrad approach is:

$$N_{low}^{Conrad;90\%CL} = 0 \tag{6.21}$$

$$N_{up}^{Conrad;90\%CL} = 96.26 \tag{6.22}$$

Assuming a mixed composition, the number of signal events expected in the data set analyzed can be written as the sum of the signal events generated by each nuclei:

$$N_S = \sum_A \int_E \Phi(E)\, f_A(E)\, \mathscr{E}^A(E)\, \frac{\sigma_{IQ}^{A-air}(E)}{\sigma_{inel}^{A-air}(E)}\, dE \tag{6.23}$$

where $\sigma_{IQ}^{A-air}(E)/\sigma_{inel}^{A-air}(E)$ is the fraction of the inelastic cross-section due to Intrinsic Quark production in nucleus-air interactions and $\mathscr{E}^A(E)$ is the corresponding exposure. From equations C.2 and C.3 in appendix C, we can see that the ratio of the intrinsic quark production cross-section to the total inelastic cross-section, $\sigma_{IQ}^{AN}(E)/\sigma_{inel}^{AN}(E) = w_A$, is energy independent. In addition, the Intrinsic Quark cross-section and the inelastic cross-section both scale from hadron-hadron to nucleus-nucleus collisions in the same way [283], and $w_A = w$ for all values of A. We can then write:

$$N_S = w \sum_A \int_E \Phi(E)\, f_A(E)\, \mathscr{E}^A(E)\, dE \tag{6.24}$$

$$w = \frac{N_S}{\sum_A \int_E \Phi(E)\, f_A(E)\, \mathscr{E}^A(E)\, dE} \tag{6.25}$$

[13] http://root.cern.ch/root/html/TFeldmanCousins.html.

We have considered that only heavy hadrons produced in proton showers will be detectable. For nuclei heavier than protons, we assume that heavy hadrons produce no detectable effect, and the selection efficiency for these showers is the same as in showers not producing heavy hadrons. As in the calculation of the background, we assume a primary composition given by figure 5.12. In this way, $\sum_A \int_E \Phi(E)\, f_A(E)\, \mathscr{E}^A(E)\, dE = 1748.01$ corresponds to the number of signal events in the simulations that pass the quality and the selection cuts. With this number and the upper and lower limits to $N_{S;low}$ obtained above we can compute an upper limits to w.

The integral upper limit to the fraction of the inelastic nucleus-air cross-section producing heavy hadrons in the Intrinsic Quark model is:

$$\left[w_{up;90\%CL} = \left(\frac{\sigma_{IQ}}{\sigma^{inel}_{pAir}} \right)_{up;90\%CL} = \frac{N_{S;up;90\%CL}}{\sum_A \int_E \Phi(E)\, f_A(E)\, \mathscr{E}^A(E)\, dE} \leq 5.5\% \right]$$

in the energy range $E \in [2.7 \cdot 10^{18} - 6.5 \cdot 10^{19}]$ [eV].

In this chapter we have seen how the number of expected events of a certain type depends on the exposure of the observatory. The exposure, defined as the effective area integrated over time, is a function on the changing configuration of the detector and the selection efficiency to the particular process under study.

The surface and fluorescence detectors are continuously monitorized, recording the time evolution of the observatory. Using the selection efficiency derived in the previous chapter and the monitoring information we have computed the exposure and number of expected signal and background showers.

We have applied the Boosted Decision Tree discriminant to the data collected by the Pierre Auger Observatory. The comparison of the number of events passing the selection cuts and the expected number of background events gives a **lower and upper limit to the number of signal events**. We can use these limits to derive a **limit to the heavy quark production cross-section in the Intrinsic Quark model**, written as a fraction of the total inelastic cross-section.

Summary and conclusions

Charm and bottom quark production is ubiquitous in accelerators. Since cosmic rays interactions grant access to an energy range well above that achieved at man-made accelerators, we expect those quarks to be produced copiously in cosmic ray interactions as well. In Monte-Carlo simulations they are usually neglected. Most of the time they carry energies well below their critical ones and they decay on the spot, not having any effect in the shower development. However, if they are produced with large enough energies they become long-lived particles, and interact rather than decay while propagating in the atmosphere. Due to their masses, larger than those of pions and kaons, we expect heavy hadrons interactions to be more elastic, keeping a larger fraction of their energy after each interaction. If in addition they carry a significant fraction of the primary particle energy they can play an important role in the development of the shower.

This Ph.D thesis focuses on the production and possible detection of heavy quarks in EAS. We have analyzed the consequences of an energetic heavy hadronic component produced in the first interaction and propagating in the atmosphere. QCD effective theories at ultra high energies predict charm and bottom quark production with an average $< x_Q >= 3\%$ of the primary proton energy. In contrast, the Intrinsic Quark model predicts the production of heavy hadrons with average fractions of the primary particle much larger, $< x_Q >= 34\%$. The heavy hadrons produced in this way interact with air nuclei and propagate in the atmosphere. **A dedicated interaction model has been used to simulate these interactions, obtaining elasticites $\sim$60% for charmed hadrons and $\sim$70% for bottom hadrons, much larger than those of proton-air and π-air interactions. The interaction routines written have been implemented inside CORSIKA and are part of the official version of the program since version v7.35**.

The propagation of highly energetic heavy hadrons in the shower results in the displacement of a significant component of the shower energy to large depths, leading to a modification of the shower development:

- On average these **showers are wider** and the **ratio of the number of particles that reach ground and the number of particles at the depth of maximum production in the shower is larger**.
- Their **development is slower** compared to that of showers with no heavy quark production and reach the **same shower development stage at larger depths**.
- The deformation of the shower profile and the increase of particles that reach ground introduces a **mismatch between the energetic reconstructions of the surface and the fluorescence detector**.

These features have been used to train a multivariate discrimant, namely a Boosted Decision Tree, to maximize the separation between showers with heavy hadrons and those without them. Decision Trees are highly customizable and we scanned the possible parameters to find the optimal configuration for our analysis. **We built a discriminant with a constant selection efficiency $\sim$20% for background showers, and a selection efficiency growing from $\sim$35% to $\sim$65% with rising energy for signal showers.**

We have studied the selection efficiency dependence on the shower energy, core position, and the configuration of fluorescence telescopes detecting it, generating selection efficiency maps. Using these maps for different detector configurations we have replicated the changing status of the observatory over the period analyzed, from January 2004 to December 2012, and calculated the corresponding hybrid exposure. With this exposure **we have derived an upper limit to the Intrinsic Quark production cross-section, as a fraction of the inelastic proton-air cross-section**:

$$\left[\left(\frac{\sigma_{\mathbf{IQ}}}{\sigma^{\mathbf{inel}}_{\mathbf{pAir}}}\right)_{\mathbf{up;90\%CL}} \leq \mathbf{5.5}\%\right]$$

in the energy range $E \in [2.7 \cdot 10^{18} - 6.5 \cdot 10^{19}]$ [eV]. This limit is compatible with the values of the Intrinsic Quark component derived from accelerator data, which set it at the $\sim$1-3% level [296, 297, 298, 299, 300, 301]. The limit we have derived is not as stringent, which is logical considering the difficulties associated to the observation of particular features in the shower profile of an EAS. Nevertheless, **this is the first estimate of the Intrinsic Quark cross-section made using cosmic ray data at ultra-high energies**.

Resumen y conclusiones

Los quarks charm y bottom se producen abundantemente en los aceleradores. Dado que las interacciones de los rayos cósmicos en la atmósfera ocurren a energías mucho más altas de las que se pueden alcanzar en los aceleradores de partículas, esperamos que estos quarks también se produzcan en gran cantidad en las interacciones de rayos cósmicos, pese a que en los códigos de simulación Monte Carlo normalmente se ignora su producción. Normalmente se producen con energías muy por debajo de su energía crítica, desintegrándose prácticamente en el punto de producción, y por lo tanto no teniendo ningún efecto en el desarrollo de la EAS. Sin embargo, si se producen con suficiente energía se convierten en partículas con una vida larga, y es posible que interactúen antes de desintegrarse. Debido a que su masa es mucho mayor que la los piones y los kaones, esperamos además que las interacciones de los hadrones pesados sean mucho más elásticas, reteniendo una fracción mayor de su energía después de cada interacción. Si además la energía con la que se producen es una fracción considerable de la enegía del rayo cósmico primario, pueden pasar a tener importancia en el desarrollo de la EAS. En 1982, Stodolsky y McLerran estudiaron teoricamente qué efecto tendría la propagación de una componente pesada muy energética en el desarrollo de la cascada [11]. Desde entonces ningún estudio ha abordado el problema de la detección de hadrones con charm y bottom en EAS.

Esta tesis se ha enfocado en la producción y posible detección de quarks pesados en cascadas de rayos cósmicos. Hemos analizado las consecuencias de la propagación en una EAS de una componente pesada y muy energética producida durante la primera interacción. Las teorías efectivas de QCD a energís ultra altas predicen la producción de quark pesados con energías promedio $< x_Q >= 3\%$ de la energía del protón primario. A diferencia de éste, El modelo de Quark Intrínseco predice la producción de quarks pesados con fracciones elevadas de energía del primario, $< x_Q >= 34\%$. Los hadrones pesados que se producen de esta manera interaccionan con los núcleos de aire y se propagan a través de la atmósfera. **Hemos usado un modelo de interacción específico para simular estas interacciones, obteniendo como resultado que su elasticidad es mucho mayor ($\sim$60% para hadrones con charm y $\sim$70% para hadrones con bottom)** que en colisiones de protones y piones con núcleos de aire. Las rutinas de interacción que se han programado forman parte de la versión oficial de CORSIKA desde la versión v7.35.

La propagación en la atmósfera de hadrones pesados muy energéticos da lugar al desplazamiento de fracción importante de la energía de la cascada a profundidades mayores. Como resultado, se produce una modificación del desarrollo de la cascada:

- En promedio **las cascadas son más anchas y la relación entre número de partículas que llegan al suelo y el número de partículas en el máximo de la cascada es mayor**.

- Su **desarrollo es más lento** que el de las EAS que no tienen hadrones pesados, y **alcanzan el mismo desarrollo (medido como % del desarollo total) a profundidades mayores**.
- La deformación del perfil longitudinal de la cascada y el aumento de partículas en el suelo genera una **diferencia entre las energías reconstruídas por el detector de superficie y el de fluorescencia**.

Hemos usado estas características para entrenar un discriminante multivariable, en particular, un árbol de Decisión (o Boosted Decision Tree, en inglés), para intentar maximizar la separación entre cascadas con producción de quarks pesados y aquellas que no la tienen. Los árboles de Decisión permiten un alto grado de personalización, ajustando sus parámetros al estudio de un problema en particular. Hemos escaneado las posibles combinaciones de parámetros hasta encontrar la mejor configuración para nuestro análsis.

El discriminante resultante posee una eficiencia de selección $\sim$20% constante con la energía para cascadas de fondo, y una eficiencia de selección que aumenta desde un $\sim$35% para la energía simulada más baja hasta $\sim$65% en la más alta en el caso de cascadas de señal.

Hemos estudiado la dependencia de la eficiencia de selección en la posición de la enegía y la posición del *core* de la cascada y de la configuración de los telescopios que la observaron, creando mapas de eficiencia de selección. Usando estos mapas hemos replicado el estado del observatorio entre enero de 2004 y diciembre de 2012 y hemos calculado la correspondiente exposición híbrida del detector.

A partir de esta exposición hemos derivado un límite superior a la sección eficaz de producción de quarks pesados en el módelo de Quark Intrínsico, expresado como la fracción de la sección eficaz inelástica protón-aire:

$$\left[\left(\frac{\sigma_{\mathbf{IQ}}}{\sigma^{\mathbf{inel}}_{\mathbf{pAir}}}\right)_{\mathbf{up;90\%CL}} \leq \mathbf{5.5}\%\right]$$

en el rango de energías $E \in [2.7 \cdot 10^{18} - 6.5 \cdot 10^{19}]$ [eV]. Este límite es compatible con los valores de la sección eficaz de producción de Quark Intrínseco obtenidos en aceleradores, que se sitúan en torno al $\sim$1-3% [296, 297, 298, 299, 300, 301]. El límite que hemos conseguido no es tan restrictivo, lo que por otra parte es lógico teniendo en cuenta las dificultades asociadas a la observación de deformaciones del perfil longitudinal de EAS. **Aún así, este es el primer límite a la sección eficaz de producción de Quark Intrínseco obtenido usando rayos cósmicos de ultra alta energía.**

A

Surface Detector Trigger in inclined Extensive Air Showers

A.1 Introduction

The purpose of this appendix is to estimate the fraction of events that cause a central trigger for any given energy and direction. The analysis of this thesis was initially intended only for showers around 60°, zenith angle that marks the transition from vertical to inclined showers. Before the analysis, we made a study of the surface detector trigger for inclined showers. In the case of vertical showers, where there is an approximate cylindrical symmetry around the axis of the shower, it is possible to produce a set of Lateral Trigger Probability distribution functions (LTPs) that can be determined by looking at the data directly [345]. These distributions can then be used to estimate the trigger efficiency [346]. However, in the case of inclined showers, the presence of the geomagnetic field breaks the azimuth symmetry. It is not possible to group events with the same zenith angle since two events with different azimuth angles will produce different footprints. This in turn limits the statistics. To overcome this problem, we assume that the shape of the muon map does not change with energy and that it correctly describes the expected signals in an inclined shower. In this way we can group events with different energies to produce local trigger probability functions and use these to estimate the central trigger efficiency. This method is similar to the method found in [347, 348]. Up to now, the estimates we have of the inclined trigger efficiency come from studies of simulated events. These studies consisted in generating two-dimensional triggering probability maps (T2 maps). These are used in toy Monte Carlo algorithms to estimate the probability of triggering for events of given energy and arrival direction. Only recently there have been efforts to estimate the efficiency from hybrid data as well as infill data [349]. In what follows we describe the method for estimating the central trigger efficiency (T3 efficiency) directly from the data, using local trigger probability functions. The main methods have been implemented using Offline and the current HASOffline package for analysis of inclined events [350].

A.2 Station Trigger Probabilities

The main tool for estimating the central trigger probability will be what we call T2 maps. For a given energy and direction, a T2 map is a two-dimensional function that gives the probability density that a station at position (x,y) will trigger[1]. In order to make a T2 map we first calculate the signal expected in a station, s_m, given the energy and arrival direction of the shower. This is done using the tools provided in HASOffline:

[1]The position is measured in the system at ground

$$s_m =< S_\mu > (1 + f_{EM}(x, y)) \tag{A.1}$$

where f_{EM} is the fraction of the signal due to the electromagnetic component of the cascade and S_μ is the signal expected from the muon component [351]. This in turn is estimated using the tabulated detector response to n muons, $s_\mu(\theta, n)$, assuming Poisson fluctuations around a mean given by the muon map (N_μ):

$$S_\mu = \sum_{n=0}^{\infty} \frac{\exp^{-N_\mu}}{n!} N_\mu^n \cdot s_\mu(\theta, n) \tag{A.2}$$

Using this, one can determine the local station trigger probability, $P_{T2}^{Th}(s_m)$, as a function of the expected signal. This is just the fraction of working stations (N_{ON}) with expected signal s_m that give a second level trigger:

$$P_{T2}^{Th}(s_m) = \frac{N_{T2}(s_m)}{N_{ON}(s_m)} \tag{A.3}$$

This approach gives the probability for a series of discrete bins of zenith angle and signal. However, we want it for every value of tank signal and zenith angle, thus we should look for the appropriate function that parametrizes this dependence. The chosen function reads as follows:

$$P_{T2}(s, \theta) = \frac{1}{1 + \exp a(\theta) \cdot s^{b(\theta)}} \tag{A.4}$$

where s is the signal and the parameters a and b are θ dependent, but not E dependent. In the following sections we analyze the effect of the shower parameters on this function.

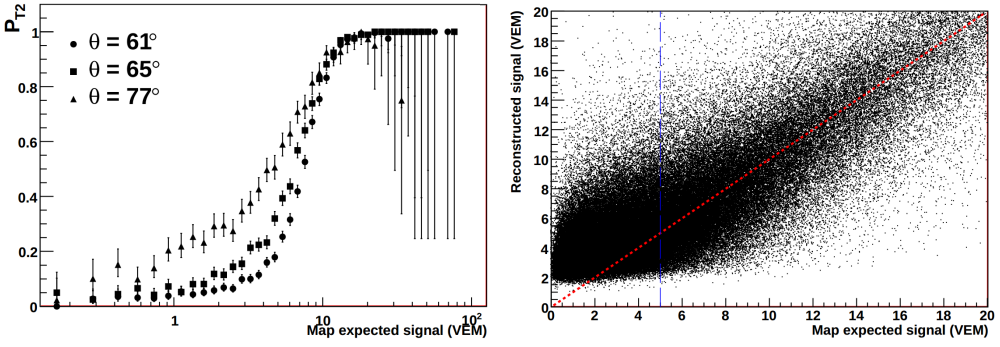


Figure A.1: Left: Local station T2 probability at different zenith angles. No correction for the bias at low expected signals has been applied. Right: Comparison of expected and recorded signal. The vertical line marks $s_{map} = 5$ VEM. The inclined line represents $s_{rec} = s_{map}$.

A.2.1 The data set

We analyzed surface events recorded from January 2004 to May 2008. The reconstruction sequence used is the one described in [350] using the Aachen muon maps and tank response [352]. We work with T5 events checking the following in addition:

- We reject events falling in the so called Bad Periods.
- For every station, we check whether it was alive at the time of the event.

- Stations with saturated signals or flagged as rejected by the reconstruction are not considered.

We calculated the local trigger probability as explained above, separating the events in zenith angle bins centered in $\theta = \{61^o, 63^o, 65^o, 67^o, 69^o, 71^o, 73^o, 75^o, 77^o, 79^o\}$. In this way we select every event in the range $60^o < \theta < 80^o$. A subset of the resulting trigger probabilities is shown in figure A.1 (left).

A.2.2 Signal bias

It is expected that the signal recorded in a station will fluctuate around the expected value. The local trigger probability will then be affected by upward fluctuations and the estimated value will be larger than the actual one. In figure A.1 (right) we plot, for all the selected stations, the reconstructed signals as a function of the expected one, obtained from the muon maps. We see a clear bias for low expected signals. Above 5 VEM, where the effect of the bias is small, the reconstructed signals for a given value of the muon map signal neatly follow a gaussian distribution (with mean close to the expected signal value and a bigger σ as, both, the zenith angle and the mean value of the signal grow). To obtain more accurate results, the parameterization of the reconstructed signal as a gaussian is done for each zenith angle and expected signal bin. We observed that σ grows linearly with the map signal, steeper the higher the zenith angle. For values of the map signal around 5 VEM, σ varies between 1.5 and 2 for all zenith angles. For higher signals, the differences between different zenith angle bins increase.

Parameter	p_0	p_1
$a = p_0 + p_1\theta$	29.0±0.8	0.33 ± 0.01
$b = p_0 + p_1\theta$	13.9 ± 0.5	0.149 ± 0.007

Table A.1: Values of the parameters a and b appearing in equation A.2.

For instance, at a expected signal of 25 VEM, σ grows from 3 for $\theta = 61^o$ to 6 for $\theta = 79^o$. This stems from the smaller statistics available as the signal and angle increase. To correct the bias, we extrapolated this modeling for signals below 5 VEM. Given a certain signal assigned by the muon map, s_m , we get the probability of having a reconstructed signal s_i in station i:

$$P(s_i|s_m) = \frac{1}{\sqrt{2\pi}\sigma} \exp^{-\frac{1}{2}(\frac{s_i - s_m}{\sigma})^2} \tag{A.5}$$

where σ, as explained in the previous paragraphs, is a function of the signal and the zenith angle.

It is clear that the probability computed in equation A.2 is biased. To unbias it, we have to unfold, for each of the stations, the effects caused by the fluctuations. These are taken into account in $P(s_i|s_m)$. The unbiased probability $P_{T2}(s_m, \theta)$ now reads:

$$P_{T2}(s_m, \theta) = \frac{\sum_{i=1}^{N_{ON}} p_{T2}(s_i, \theta)) P(s_i|s_m)}{\sum_{i=1}^{N_{ON}} P(s_i|s_m)} \tag{A.6}$$

where p_{T2} is one for stations with a T2 trigger and zero otherwise and the sum spans over all the active stations that, showing different values for the reconstructed signal s_i , have the same expected signal, s_m. Fitting (7) we obtain the shape of the trigger probability for the case of no bias. To avoid a proliferation of free parameters while unbiasing, we express each of them as a function of θ, ending up with a set of just 4 unknowns. Figure A.2 (left) shows the result of this parameterization. The values of the parameters are compiled in table 1. The result of the unbiased global parameterization is shown in figure A.2 (right). We observe that after the unbiasing, the probabilities vanish as the signal goes to zero. This effect is especially observed for large zenith angles (note the difference with the behavior depicted in figure A.1 (left)).

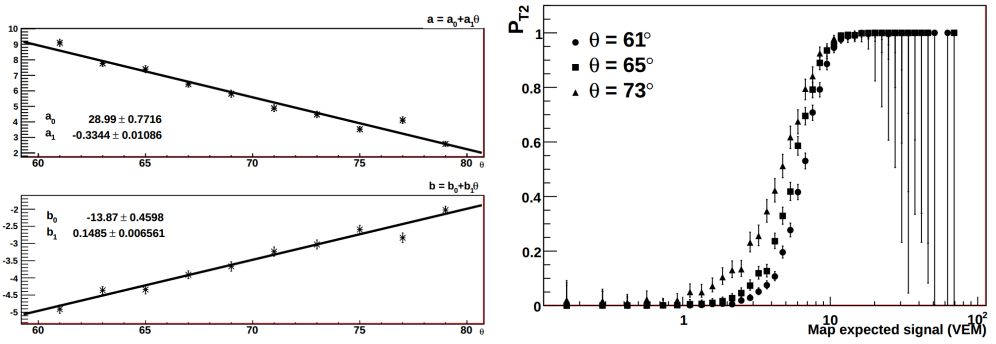


Figure A.2: Left: Result of the parameterization of parameters a and b as function of θ. Functions used and fit values are shown in each panel individually. Right: Local station T2 probability at different zenith angles, after the correction for the bias at low expected signals has been applied.

A.2.3 Early-late dependence

We expect that muons coming from the early and the late part of the shower show an asymmetry in their signal. However, we are always assuming that the muons arrive in the direction of the cascade. The effect of the incident angle in the early and late regions tends to displace the resulting curve in opposite directions. Therefore, the net effect is very small compared to the uncertainties. Properly taking into consideration this effect would unnecessarily complicate the analysis. In figure A.3 one can see the local trigger probability for tanks belonging to late and early part of the shower, for four different zenith angles. There is no significant difference between tanks in the early and late parts of the shower.

A.2.4 Energy and direction dependence

In figure A.5 one can see a comparison of the trigger probability for different values of N_{19} (0.2, 0.6, and 1) and θ (61° , 63° , 65° and 67°). As we can see, there is little dependence on the primary N_{19}. For the sake of simplicity we disregard this dependence hereafter. This assumption will let us increase our statistics, and later parameterize taking only into account the θ dependence. We also checked for a possible dependence on the azimuth angle. In figure A.4 (left) one can see the local trigger probability for different values of ϕ and all possible θ values. There is no significant dependence on the azimuth. However, there is a clear dependence on zenith angle, as shown in figure A.4 (right).

A.3 Central Trigger Efficiency

Using the expected signal and the T2 probability function, a T2 map can be computed for every position (x,y) with respect to the shower core position. Figure A.6 shows examples of T2 maps for different shower configurations. These T2 maps allow to compute the central trigger efficiency (T3 efficiency). This is done using a toy Monte Carlo algorithm that consists in generating mock events and randomly assigning them a core location in an array unit cell. Given the core location, each station will trigger or not according to the local trigger probability given by the T2 map. The resulting configuration is then passed to the central trigger simulator to determine if the configuration gives a central trigger. We repeat this process several times and, finally, the T3 efficiency is simply the fraction of these mock events that result in a central trigger. This was done for the following values of zenith angle (61°, 63°, 65°, 67°, 69°, 71°, 73°, 75°, 77°, 79°). The algorithm is implemented in two Offline modules, one that creates the mock events and another that just keeps the count of the events. Finally, to convert N_{19} to energy

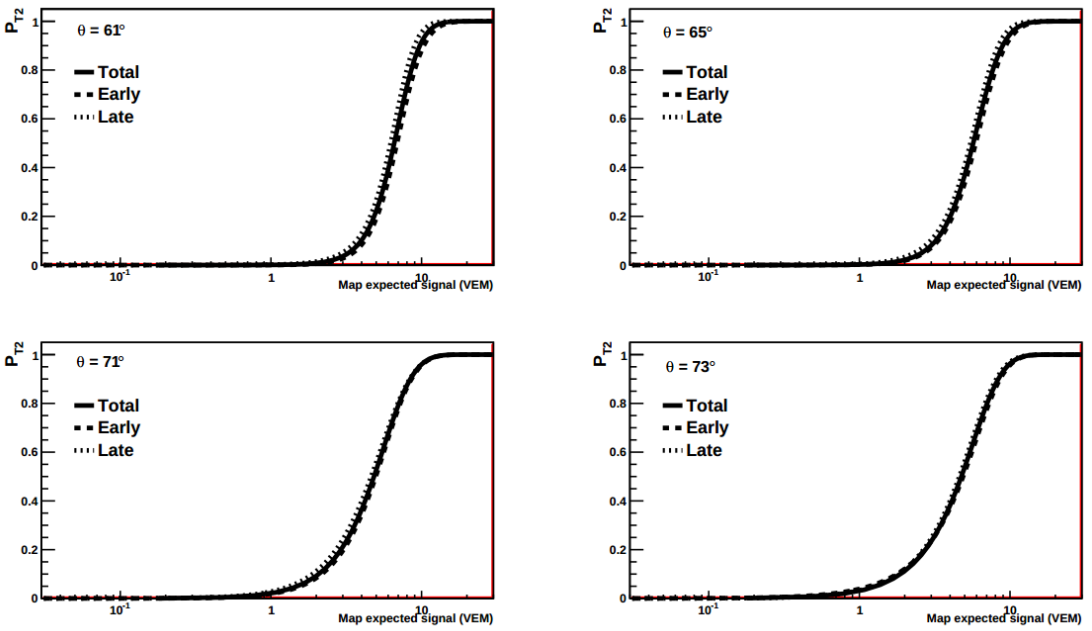


Figure A.3: Local station T2 probability using the stations in the early and late parts of the shower. There is no significant difference among the considered probabilities.

we use the calibration from [353], and the average of the efficiency over a zenith angle range, $P_{T3}(N_{19})$, is given by:

$$P_{T3}(N_{19}) = \frac{\int_{\theta_{min}}^{\theta_{max}} P_{T3}(N_{19},\theta)\cos\theta\sin\theta d\theta}{\int_{\theta_{min}}^{\theta_{max}} \cos\theta\sin\theta d\theta} \tag{A.7}$$

The exact sequence used is the following:

```
<sequenceFile>
  <moduleControl>
    <loop numTimes="5000" pushEventToStack="yes">
      <module> EventGeneratorOG </module>
      <module> MockEvent </module>
      <module> TankGPSSimulatorOG </module>
      <module> CentralTriggerSimulatorXb </module>
      <module> T3Module </module>
    </loop>
  </moduleControl>
</sequenceFile>
```

The resulting T3 efficiency is shown in figure A.7 and A.8 together with the results presented in [349]. The detector reaches full efficiency at $E \sim 10^{18.5}$ eV. Our calculation predicts full efficiency for energies above $10^{18.6}$ eV, which is very close to the expected value. A comparison with a simulation made with the Monte-Carlo simulator AIRES [354] at a fixed zenith angle of 70° is shown in A.8 for comparison. The data gives a smaller efficiency at lower energies, since all zenith angles are considered. This is consistent with the results shown in [348]: the more vertical the zenith angle is, the smaller the trigger probability. The results of this study are in agreement with the Monte-Carlo results of [349] but show a discrepancy with the efficiency they calculated using the infill data. The advantage of our study is that the statistical uncertainty is smaller given the larger data set used.

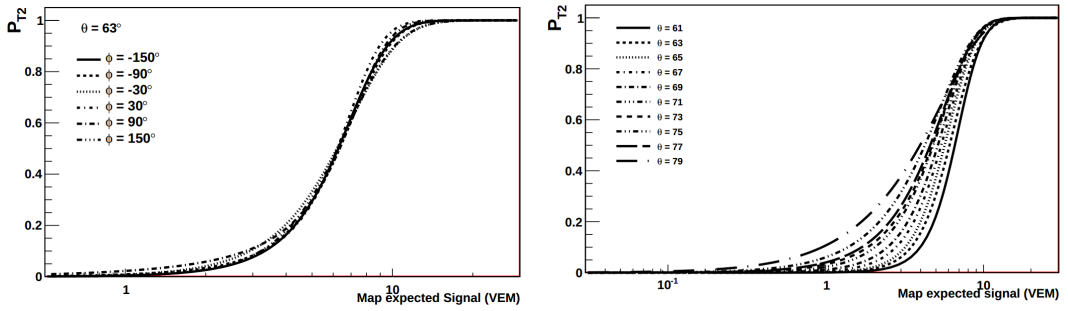


Figure A.4: Local station T2 probability at different azimuth (left) and zenith (right) angles.

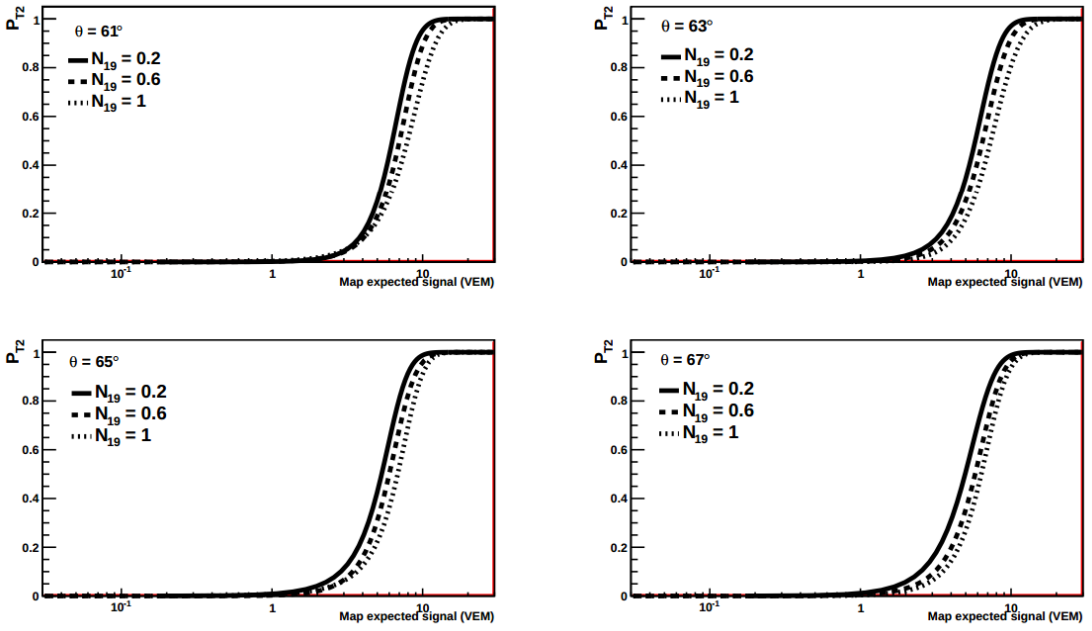


Figure A.5: T2 threshold trigger probability for different values of N_{19} and θ (61°, 63°, 65°, 67°, from top left to bottom right).

We have estimated the probability that stations will give T2 threshold triggers given the core position, energy and arrival direction of a shower in the case of inclined events. The parameterization of this probability in terms of shower energy and arrival direction provides a set of T2 maps that can be used to estimate the T3 efficiency for inclined showers. Our values for the T3 efficiency show a discrepancy with the ones computed using the infill and, as expected, the statistical errors associated with our method are smaller.

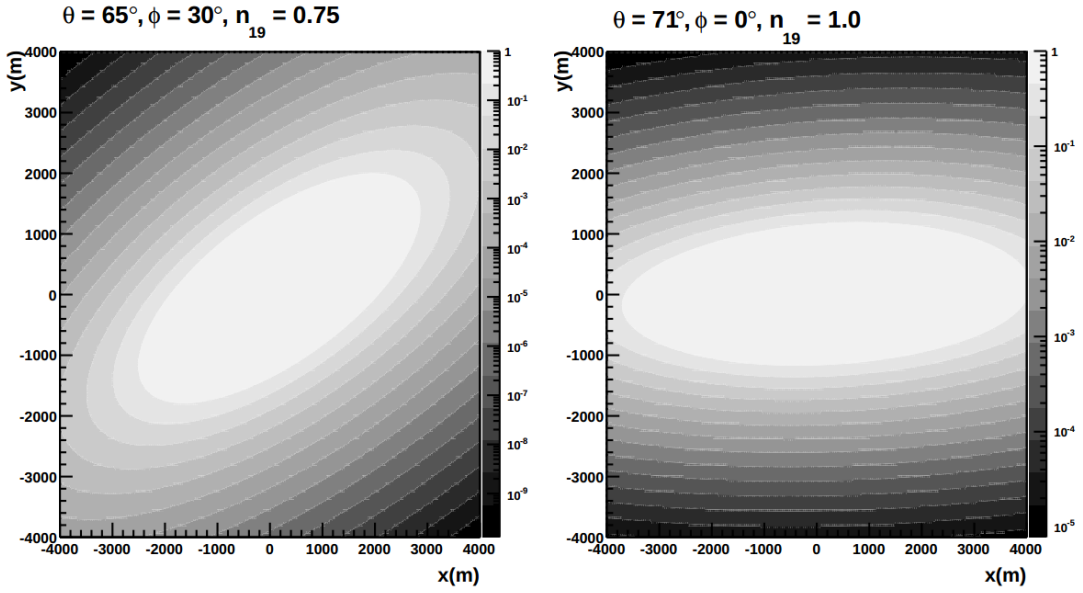


Figure A.6: Example of T2 maps.

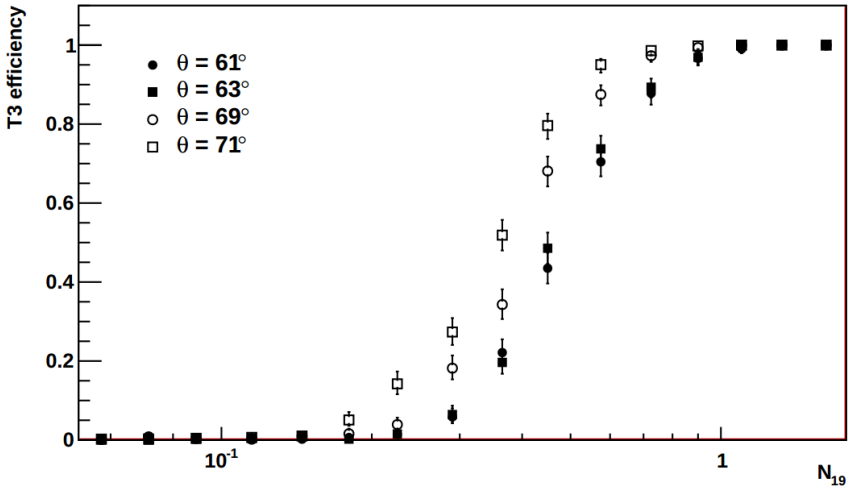


Figure A.7: Left: T3 efficiency as a function of N_{19}.

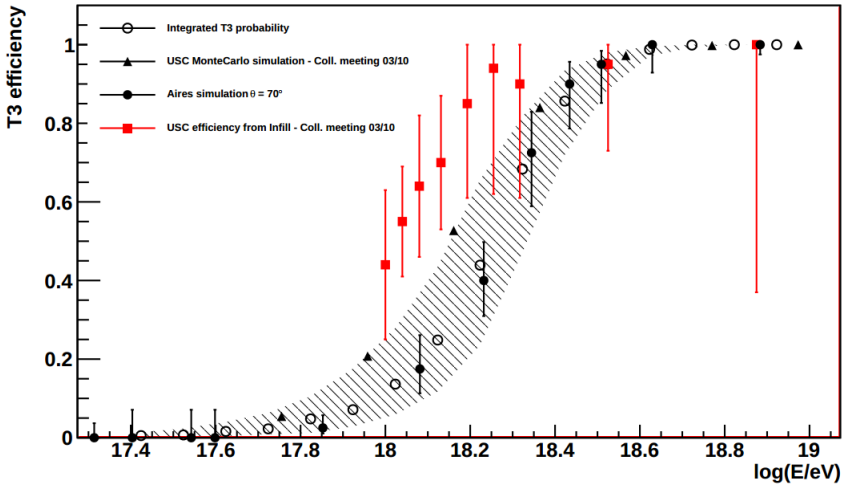


Figure A.8: T3 efficiency as a function of energy. The hatched area corresponds to the error associated to our measurement and includes systematic and statistical errors added in quadrature. The updated results from [5] are also shown for comparison, as well as those from an AIRES simulation done at a fixed zenith angle of 70°.

B

MonteCarlo modifications

The physical models used to produce and propagate heavy hadrons in EAS need to be included in MC generators. As we are using CORSIKA as MC program, we decided to implement the modelled physics as new subroutines, so they became part of the standard CORSIKA framework. For the sake of simplicity we tried to adopt the CORSIKA philosophy, making the transport and production optional through keywords, and allowing a certain flexibility in the values of some physical constants.

In the following, we detail the modifications of the particle lists (section B.1), input file (section B.2) and the source code (B.3) that were needed to adapt CORSIKA for charm and bottom propagation. We made the modifications using version 6.990 of CORSIKA. Those modifications regarding bottom particles properties, the propagation of charmed and bottom hadrons and their decays have been part of the official version of CORSIKA since version 7.35. The modifications listed have been previously reported in [10].

B.1 Particles considered and particle codes

The bottom quark is not considered in CORSIKA because none of the hadronic interaction models it implements produces bottom hadrons.

Basic information about leptons, neutrinos and hadrons is stored in the subroutine **PAMAF**. Six arrays contain the masses in GeV (**MASSES** and **MASSES2**), the electric charge in electron charge units (**CHARGE** and **CHARGE2**) and the mean life-times in s (**DECTIME** and **DECTIME2**). Each couple of arrays can hold up to 200 different particles.

If we want CORSIKA to propagate bottom hadrons, we first have to include them as eligible particles. The last particle included is $\bar{\Sigma}_c^{*0}$, with code 173. We use the empty codes starting from 176 to include the new bottom hadrons. Bottom mesons and their antiparticles are identified by codes from 176 to 183. Λ, Σ, Ξ and Ω baryons and their antiparticles have codes from 184 to 197. Only ground states of the particles have been introduced. Details on the particle codes, masses and lifetimes, obtained from the Particle Data Book [314], are shown in table B.1.

B.2 Input file

CORSIKA reads a series of keywords to select the parameters of the simulations. These keywords have to be provided by the user as an input file. What follows is an example of a simple input file: in addition to the standard keywords, we have underlined those keywords needed to use the subroutines that control the physics of heavy hadron production and propagation:

```
RUNNR 1 number of run
EVTNR 100400 no of first shower event
SEED 100401 0 0 seed for hadronic part
```

```
SEED 100402 0 0 seed for EGS4 part
COLLDR 1 3
SIGMAQ  0.D0 0.D0 0.D0 0.D0
PROPAQ  1
NSHOW 10 no of showers to simulate
PRMPAR 14 primary particle code (proton)
ERANGE 1.00E10 1.00E10 energy range of primary (GeV)
THETAP 60. 60. range zenith angle (deg)
PHIP -180. 180. range azimuth angle (deg)
EXIT
```

- **COLLDR** determines the type of the heavy quarks produced during the first interaction (first argument), and the production mechanism (second argument). The first argument accepts the values 1 for charm production, 2 for bottom production or 0 in case the first interaction is simulated by the chosen hadronic interaction model (be it SIBYLL, QGSJET, ...) without heavy quark production. The second one takes the values 1 for production via the Color Glass Condensate model, or 3, for production using the Intrinsic Quark model.

- **SIGMAQ** takes four arguments, the cross-sections (in mb) for interaction with protons of charmed mesons, charmed baryons, bottom mesons and bottom baryons, respectively. If the values are equal to 0 the parameterization shown in figure 3 is used.

- **PROPAQ** toggles the propagation of heavy hadrons with the new subroutines. If equal to 0, the propagation of heavy hadrons is performed by the high energy interaction model. If equal to 1, the propagation is dealt with using **HEPARIN**.

B.3 Source code modifications

New functions have been written to perform specific parts of the simulation and some others have been modified inside the source code to allow the propagation of the new particles. We list the files that have been modified, those new files added, and overview the changes made to the code.

The directory **corsika-6990/src/** contains the main source files needed to run CORSIKA. Inside the file **corsika.F** we have made several modifications to already present subroutines:

- **DATAC** reads the CORSIKA input file. This subroutine has been modified to accept the new keywords described in section B.2.

- the subroutine **NUCINT** selects the type of interaction process according to the particle energy. Now it includes a call to the new subroutine **COLLIDE**, to simulate the first interaction with production of heavy hadrons. The selection of interaction or decay routines for different particles types is extended to treat bottom hadrons. Both charmed and bottom hadrons interactions are treated in the new subroutine **HEPARIN**.

- **PAMAF** initializes the masses in GeV, the electric charge in electron charge units and the mean life-times in s of the particles defined in CORSIKA. We modify it to hold the bottom hadrons defined in section B.1 as well.

- **BOX2** determines the point of interaction or decay for any particle. It now uses the interaction cross-sections of charmed particles with air shown in figure 4.14 to calculate their interaction lengths and whether they decay or interact. It has been extended to treat bottom hadrons as well.

- **PYTSTO** transports the particles resulting from PYTHIA to the CORSIKA stack. It is modified to accept bottom hadrons too.

We have also added new subroutines:

- **HEPARIN** links with the PYTHIA routines that treat the interaction of heavy hadrons with air nuclei, instead of calling the high-energy model chosen during compilation.
- **NNY** samples the number of interacting nucleons in the collisions of heavy hadrons with air nuclei. The sampled distributions are obtained using a modified version of NUCOGE [313].
- **BTTMDC** is called to perform the decay of bottom hadrons.

Some of these modifications need the definition of new variables. These have been included in the header file **corsika.h**.

The file **qgsjet01c.f** simulates the physics of the model QGSJET01c. We have modified it to suppress the production of heavy quarks during the first interaction. Thus, only **COLLIDE** (see below) produces them at that step of the shower.

The directory **corsika-6990/pythia** contains all the PYTHIA routines called during the simulation of the shower. The source files of some of the new subroutines are here:

- the subroutine **COLLIDE**, in the file **collider.f**, produces the charmed or bottom hadrons at the first proton interaction. We assume that the first interaction $pA \to H_Q H_{\bar{Q}} X$ can be described as the superposition of the shower generated by the heavy hadrons (H_Q and $H_{\bar{Q}}$) and the shower started by a proton of energy $E'_p = E_p - E_{H_Q} - E_{H_{\bar{Q}}}$. We use a proton as a primary and, once the depth of the first interaction (X_0) has been computed, we generate the pair H_Q and $H_{\bar{Q}}$ at depth X_0, sampling the fractions of the proton energy carried away (x_1, x_2) from the corresponding distributions. The energy of the proton is scaled to E'_p and the proton shower starts at X_0. The particles are transferred to the CORSIKA stack using the subroutine **PYTSTO**. The rest of the shower development follows the usual procedure. The type of particle produced (charm or bottom) and the production model (Color Glass Condensate or Intrinsic Quark) are chosen setting new keywords in the datacard (see section B.2).
- subroutines **CHABADIF**, **CHABAPAR**, **CHAMEDIF**, **CHAMEPAR**, **BOBADIF**, **BOBAPAR**, **BOMEDIF** and **BOMEPAR** (defined in the files with the same names and extension **.f**) are called from **HEPARIN** to treat the diffractive and partonic interactions of heavy hadrons. The interactions are simulated according to the model described in section 4.2.2.

The processes included in the subroutines above need the modification of two PYTHIA source files, pypdfu.f and pyspli.f.

Table B.1: CORSIKA particle codes extension. *: Σ_b^0, $\bar{\Sigma_b}^0$ are forced to decay whenever they are produced.

Particle code	Particle name	Particle mass [GeV]	Particle life-time [s]	Particle code	Particle name	Particle mass [GeV]	Particle life-time [s]
176	B^0	5.27958	$1.519{\cdot}10^{-12}$	187	Ξ_b^0	5.788	$1.49{\cdot}10^{-12}$
177	B^+	5.27925	$1.641{\cdot}10^{-12}$	188	Ξ_b^-	5.7911	$1.56{\cdot}10^{-12}$
178	B^-	5.27925	$1.641{\cdot}10^{-12}$	189	Ω_b^-	6.071	$1.1{\cdot}10^{-12}$
179	$\bar{B}^0$	5.27958	$1.519{\cdot}10^{-12}$	190	$\bar{\Lambda_b}^0$	5.6194	$1.425{\cdot}10^{-12}$
180	B_s^0	5.36677	$1.497{\cdot}10^{-12}$	191	$\bar{\Sigma_b}^+$	5.8155	$1.3{\cdot}10^{-22}$
181	$\bar{B_s}^0$	5.36677	$1.497{\cdot}10^{-12}$	192	Σ_b^-	5.8113	$1.68{\cdot}10^{-23}$
182	B_c^+	6.277	$0.483{\cdot}10^{-12}$	193	$\bar{\Xi_b}^0$	5.788	$1.49{\cdot}10^{-12}$
183	B_c^-	6.277	$0.483{\cdot}10^{-12}$	194	$\bar{\Xi_b}^+$	5.7911	$1.56{\cdot}10^{-12}$
184	Λ_b^0	5.6194	$1.425{\cdot}10^{-12}$	195	$\bar{\Omega_b}^+$	6.071	$1.1{\cdot}10^{-12}$
185	Σ_b^-	5.8155	$1.3{\cdot}10^{-22}$	196	Σ_b^0	5.8155	0 *
186	Σ_b^+	5.8113	$6.8{\cdot}10^{-23}$	197	$\bar{\Sigma_b}^0$	5.8155	0 *

C

Intrisic quark model details

In section 4.2.4 we explained that not all features present in data are explained by QCD models. For instance, leading particle asymmetries reported from fixed target experiments that show a strong correlation between the quantum numbers of the projectile and those of the final state hadron [12, 13, 14, 18]. According to the QCD factorization theorem heavy quarks should hadronize independently of the initial state [278], and neither leading-order nor next-to-leading-order corrections can produce flavor correlation such as those observed in leading charm production [355].

C.1 Intrinsic particle production

The wavefunction of a hadron in QCD can be represented as a superposition of Fock state fluctuations.

$$|h>=\alpha_0|n_V>+\alpha_1|n_V g>+\alpha_2|n_V Q\bar{Q}>+... \tag{C.1}$$

The extra gluons and quark pairs in the higher Fock states arise from the QCD interactions. Contributions which are due to a single gluon splitting such as $g \rightarrow c\bar{c}$ are *extrinsic* to the bound state nature of the hadron. In contrast, $c\bar{c}$ pairs multiply connected to the valence quarks cannot be attributed to the gluon substructure, and are *intrinsic* to the hadron's substructure. Extrinsic quarks and gluons are generated on a short time scale associated with large transverse momentum processes. Intrinsic quarks and gluons exist over a time scale which is independent of any probe momentum transfer.

When the projectile scatters in the target, the coherence of the Fock components is broken an the fluctuations can hadronize. Uncorrelated fragmentation (*UC*, from now on) is the process of hadronization with quarks from the sea of the hadron. This is the mechanism that is unable to generate flavor asymmetries. Coalescence, on the contrary, is the process of hadronization with spectator valence quarks [18]. The coalescence mechanism introduces flavor correlations between the projectile and the final-state hadrons. While uncorrelated fragmentation distributions are assumed to be equal for all heavy charm or bottom hadrons, coalescence distributions are specific for the individual states.For instance, the production of Λ_c^+ in p-N collisions comes from the fluctuations of the Fock state of the proton to $|uudc\bar{c}>$. To obtain a Λ_c^- in p-N interactions a fluctuation to $|uudu\bar{u}d\bar{d}c\bar{c}>$ would be required. Thus, since the probability of a five quarks state is large than that of a 9 quarks state, Λ_c^+ production is favoured over Λ_c^- in proton reactions. Intrinsic $Q\bar{Q}$ Fock components are dominated by configurations with equal rapidity constituents so that, unlike sea quarks, the intrinsic heavy quarks carry a large fraction of the parent momentum.

The intrinsic charm and bottom production cross-sections can be related to the inelastic

hadron-nucleon cross-section by

$$\sigma_{ic}(hN) = P_{ic}\,\sigma_{hN}^{inel}\,\frac{\mu^2}{4\hat{m}_c^2}\alpha_s^4(M_{b\bar{b}}) \tag{C.2}$$

$$\sigma_{ib}(hN) = P_{ib}\,\sigma_{hN}^{inel}\,\frac{\mu^2}{4\hat{m}_b^2}\alpha_s^4(M_{b\bar{b}}) \tag{C.3}$$

where P_{iQ} is the probability of finding a $Q\bar{Q}$ fluctuation in the hadron wavefunction, and $P_{ib} = P_{ic}\cdot(\hat{m}_c/\hat{m}_b)^4$. The only energy dependence of σ_{iQ} is through σ^{inel} and hence $\sigma_{iQ}/\sigma^{inel}$ is an energy-independent quantity. The relation between the two cross-sections is:

$$\sigma_{ib}(hN) = \left(\frac{\hat{m}_c}{\hat{m}_b}\right)^4\cdot\left(\frac{\alpha_s(M_{b\bar{b}})}{\alpha_s(M_{c\bar{c}})}\right)^4\sigma_{ic}(hN) \tag{C.4}$$

$$\frac{\sigma_{ib}(hN)}{\sigma_{ib}(hN)} = \left(\frac{\hat{m}_c}{\hat{m}_b}\right)^4\cdot\left(\frac{\alpha_s(M_{b\bar{b}})}{\alpha_s(M_{c\bar{c}})}\right)^4 \sim 2.7\times10^{-3} \tag{C.5}$$

The frame-independent probability distribution of an n-particle $Q\bar{Q}$ Fock state is

$$\frac{dP_{ic}^n}{dx_1\cdots dx_n} \propto \frac{\delta(1-\sum_{i=1}^n x_i)}{(\hat{m}_h^2-\sum_{i=1}^n(\hat{m}_i^2/x_i))^2} \tag{C.6}$$

In uncorrelated fragmentation the $Q\bar{Q}$ pair hadronizes with quarks from the sea of the projectile or the target. i.e the fragmentation of a c quark into a D meson has the following differential probability:

$$\frac{dP_{ic}^{UF}}{dx_D} = \int dz\prod_{i=1}^n dx_i\frac{dP_{ic}^n}{dx_1\cdots dx_n}D_{D/c}(z)\delta(x_D-zx_c) \tag{C.7}$$

This distribution is assumed to be valid for intrinsic charm and bottom production by uncorrelated fragmentation of all possible final states.

If the projectile has the corresponding valence quarks, the heavy quark can also hadronize by coalescence with the valence spectators. The coalescence distribution are specific for the individual heavy hadrons, because not all hadrons containing heavy quarks can be formed from the minimal intrinsic charm or bottom Fock state $|n_V Q\bar{Q}>$. These contributions do not include any binding energy of the produced hadrons or any mass effect. This mechanism introduces flavor correlations between the projectile and the final-state hadrons.

The coalescence contribution to leading D production is:

$$\frac{dP_{ic}^C}{dx_D} = \int\prod_{i=1}^n dx_i\frac{dP_{ic}^n}{dx_1\cdots dx_n}\delta(x_D-x_c-x1-x2) \tag{C.8}$$

where the delta function stands for the coalescence of quarks in the Fock state configuration that make up the valence quarks of the final-state hadron. As an example of the formation of a baryon, the coalescence contribution to leading Λ_c production is:

$$\frac{dP_{ic}^C}{dx_{\Lambda_c}} = \int\prod_{i=1}^n dx_i\frac{dP_{ic}^n}{dx_1\cdots dx_n}\delta(x_{\Lambda_c}-x_c-x1-x2) \tag{C.9}$$

Coalescence may also occur within higher fluctuations of the intrinsic charm Fock state, such as $|n_V c\bar{c}u\bar{u}>$, $|n_V c\bar{c}d\bar{d}>$, $|n_V c\bar{c}s\bar{s}>$ or $|n_V c\bar{c}u\bar{u}d\bar{d}>$. The highest the degree of intrinsic quark fluctuation, the less likely this state will form, and hadrons formed inherit a smaller fraction of the initial momentum since the Fock state momentum is distributed over more partons. Thus, as more partons are included in the Fock state, the coalescence distributions soften and will eventually resemble the fragmentation distributions [287]. We will consider Fock states up to 7 particles only.

The probability distribution is the sum of all contributions from the $|n_V Q\bar{Q}>$ and $|n_V Q\bar{Q}q\bar{q}>$ configurations with $q = u, d, s$. Terms due to uncorrelated fragmentation are always included. Those due to coalescence are only present when the quark configuration allows for the formation of the state. We will assume that all the lowest lying charm and bottom hadrons produced by uncorrelated fragmentation have the same probability of being produced. There are 10 charm/bottom hadrons, and the same number of anticharm/antibottom hadrons (excluding excited states). As for coalescence, only a fraction of the final state hadrons can be produced by coalescence. A simple counting scheme is used. For instance, in the $|uudb\bar{b}>$ fluctuation the possible final states containing the b quark are 2 $\Lambda_b^0(\Sigma_b^0)$, Σ_b^+ and Υ, whereas the states containing the $\bar{b}$ quark are $2B^+$, B^0 and Υ. For $|uudb\bar{b}q\bar{q}>$ configurations, the final states are depending on the nature of q. Table C.1 contains all the hadrons that can be formed for each different fluctuation. For the sake of clarity, we discuss in detail the case of B^+ x_F distribution. The distribution due to uncorrelated fragmentation can be directly written as:

$$\frac{dP^F_{B^+}}{dx_F} = \left(\frac{1}{10}\frac{dP^{5F}_{ib}}{dx_F}\right) + \left(\frac{1}{10}\frac{dP^{7F}_{ibu}}{dx_F}\right) + \left(\frac{1}{10}\frac{dP^{7F}_{ibd}}{dx_F}\right) + \left(\frac{1}{10}\frac{dP^{7F}_{ibs}}{dx_F}\right) \tag{C.10}$$

A final state B^+ can be produced by coalescence from the $|uudb\bar{b}>$ fluctuation in 2 out of 4 combinations ($2B^+$, B^0, Υ). The $|uudb\bar{b}u\bar{u}>$, $|uudb\bar{b}d\bar{d}>$ and $|uudb\bar{b}s\bar{s}>$ configurations yield B^+ in 3 out of 5, 2 out of 5 and 2 out of 5 cases, respectively. The distribution due to coalescence is then:

$$\frac{dP^C_{B^+}}{dx_F} = \left(\frac{1}{2}\frac{dP^{5C}_{ib}}{dx_F}\right) + \left(\frac{3}{5}\frac{dP^{7C}_{ibu}}{dx_F}\right) + \left(\frac{2}{5}\frac{dP^{7C}_{ibd}}{dx_F}\right) + \left(\frac{2}{5}\frac{dP^{7C}_{ibs}}{dx_F}\right) \tag{C.11}$$

When a particular final state can be produced both by coalescence and uncorrelated fragmentation, the sum of the probabilities is multiplied by 0.5 to keep the total probability fixed. As such, the x_F distribution for B^+ hadrons is finally written as:

$$\begin{aligned}\frac{dP_{B^+}}{dx_F} &= \frac{1}{2}\frac{dP^F_{B^+}}{dx_F} + \frac{1}{2}\frac{dP^C_{B^+}}{dx_F} \\ &= \frac{1}{2}\left(\frac{1}{10}\frac{dP^{5F}_{ib}}{dx_F} + \frac{1}{2}\frac{dP^{5C}_{ib}}{dx_F}\right) + \frac{1}{2}\left(\frac{1}{10}\frac{dP^{7F}_{ibu}}{dx_F} + \frac{3}{5}\frac{dP^{7C}_{ibu}}{dx_F}\right) \\ &+ \frac{1}{2}\left(\frac{1}{10}\frac{dP^{7F}_{ibd}}{dx_F} + \frac{2}{5}\frac{dP^{7C}_{ibd}}{dx_F}\right) + \frac{1}{2}\left(\frac{1}{10}\frac{dP^{7F}_{ibs}}{dx_F} + \frac{2}{5}\frac{dP^{7C}_{ibd}}{dx_F}\right)\end{aligned} \tag{C.12}$$

Each of the heavy quarks produced will hadronize by either fragmentation or coalescence. In case both quarks undergo fragmentation, any bottom hadron (antihadron) can be formed from the corresponding quark (antiquark). If both quarks hadronize by coalescence, the possible final states are limited, since not all combinations can be formed with the available quarks.

Configuration	Possible final states with c (b)	Possible final states with $\bar{c}$ ($\bar{b}$)
$\|uudc\bar{c}>$	2 $\Lambda_c^+(\Sigma_c^+)$, Σ_c^{++},J/Ψ	2 $\bar{D}^0$, D^-, J/Ψ
$\|uudc\bar{c}u\bar{u}>$	$3\Lambda_c^+(\Sigma_c^+)$, $3\Sigma_c^{++}$, D^0, J/Ψ	D^-,$3\bar{D}^0$,J/Ψ
$\|uudc\bar{c}d\bar{d}>$	4 Λ_c^+ (Σ_c^+), Σ_c^{++}, Σ_c^0, D^+, J/Ψ	$2D^-$, $2\bar{D}^0$, J/Ψ
$\|uudc\bar{c}s\bar{s}>$	$2\Xi_c^+$, Ξ_c^0, 2 Λ_c^+ (Σ_c^+), Σ_c^{++}, D_s^+, J/Ψ	2 $\bar{D}^0$, D^-, D_s^-, J/Ψ
$\|uudb\bar{b}>$	2 $\Lambda_b^0(\Sigma_b^0)$, Σ_b^+,Υ	2 B^+, B^0, Υ
$\|uudb\bar{b}u\bar{u}>$	$3\Lambda_b^0(\Sigma_b^0)$, $3\Sigma_b^+$, B^-, Υ	B^0,$3B^+$,Υ
$\|uudb\bar{b}d\bar{d}>$	4 Λ_b^0 (Σ_b^0), Σ_b^-, Σ_b^+, $\bar{B}^0$, Υ	$2B^0$, $2B^+$, Υ
$\|uudb\bar{b}s\bar{s}>$	$2\Xi_b^0$, Ξ_b^-, 2 Λ_b^0 (Σ_b^0), Σ_b^+, $\bar{B}_s^0$, Υ	2 B^+, B^0, B_s^0, Υ

Table C.1: Possible final charm and bottom states for each fluctuation considered

The differential probability distributions for charm hadrons, following the combinations from

$c\bar{c}$ fluctuations in table C.1 are:

$$\frac{dP_{\Lambda_c^+}}{dx_F} = \frac{1}{2}\left(\frac{1}{10}\frac{dP_{ic}^{5F}}{dx_F} + \frac{1}{2}\frac{dP_{ic}^{5C}}{dx_F}\right) + \frac{1}{2}\left(\frac{1}{10}\frac{dP_{icu}^{7F}}{dx_F} + \frac{3}{8}\frac{dP_{icu}^{7C}}{dx_F}\right) + \frac{1}{2}\left(\frac{1}{10}\frac{dP_{icd}^{7F}}{dx_F} + \frac{1}{2}\frac{dP_{icd}^{7C}}{dx_F}\right) + \frac{1}{2}\left(\frac{1}{10}\frac{dP_{ics}^{7F}}{dx_F} + \frac{1}{4}\frac{dP_{icd}^{7C}}{dx_F}\right) \tag{C.13}$$

$$\frac{dP_{\bar{D}^0}}{dx_F} = \frac{1}{2}\left(\frac{1}{10}\frac{dP_{ic}^{5F}}{dx_F} + \frac{1}{2}\frac{dP_{ic}^{5C}}{dx_F}\right) + \frac{1}{2}\left(\frac{1}{10}\frac{dP_{icu}^{7F}}{dx_F} + \frac{3}{5}\frac{dP_{icu}^{7C}}{dx_F}\right) + \frac{1}{2}\left(\frac{1}{10}\frac{dP_{icd}^{7F}}{dx_F} + \frac{2}{5}\frac{dP_{icd}^{7C}}{dx_F}\right) + \frac{1}{2}\left(\frac{1}{10}\frac{dP_{ics}^{7F}}{dx_F} + \frac{1}{5}\frac{dP_{icd}^{7C}}{dx_F}\right) \tag{C.14}$$

$$\frac{dP_{D^-}}{dx_F} = \frac{1}{2}\left(\frac{1}{10}\frac{dP_{ic}^{5F}}{dx_F} + \frac{1}{4}\frac{dP_{ic}^{5C}}{dx_F}\right) + \frac{1}{2}\left(\frac{1}{10}\frac{dP_{icu}^{7F}}{dx_F} + \frac{1}{5}\frac{dP_{icu}^{7C}}{dx_F}\right) + \frac{1}{2}\left(\frac{1}{10}\frac{dP_{icd}^{7F}}{dx_F} + \frac{2}{5}\frac{dP_{icd}^{7C}}{dx_F}\right) + \frac{1}{2}\left(\frac{1}{10}\frac{dP_{ics}^{7F}}{dx_F} + \frac{1}{5}\frac{dP_{icd}^{7C}}{dx_F}\right) \tag{C.15}$$

$$\frac{dP_{\Sigma_c^{++}}}{dx_F} = \frac{1}{2}\left(\frac{1}{10}\frac{dP_{ic}^{5F}}{dx_F} + \frac{1}{4}\frac{dP_{ic}^{5C}}{dx_F}\right) + \frac{1}{2}\left(\frac{1}{10}\frac{dP_{icu}^{7F}}{dx_F} + \frac{3}{8}\frac{dP_{icu}^{7C}}{dx_F}\right) + \frac{1}{2}\left(\frac{1}{10}\frac{dP_{icd}^{7F}}{dx_F} + \frac{1}{8}\frac{dP_{icd}^{7C}}{dx_F}\right) + \frac{1}{2}\left(\frac{1}{10}\frac{dP_{ics}^{7F}}{dx_F} + \frac{1}{8}\frac{dP_{icd}^{7C}}{dx_F}\right) \tag{C.16}$$

$$\frac{dP_{\Sigma_c^+}}{dx_F} = \frac{dP_{\Lambda_c^+}}{dx_F} \tag{C.17}$$

$$\frac{dP_{\Sigma_c^0}}{dx_F} = \frac{1}{10}\frac{dP_{ic}^{5F}}{dx_F} + \frac{1}{10}\frac{dP_{icu}^{7F}}{dx_F} + \frac{1}{2}\left(\frac{1}{10}\frac{dP_{icd}^{7F}}{dx_F} + \frac{1}{8}\frac{dP_{icd}^{7C}}{dx_F}\right) + \frac{1}{10}\frac{dP_{ics}^{7F}}{dx_F} \tag{C.18}$$

$$\frac{dP_{D^0}}{dx_F} = \frac{1}{10}\frac{dP_{ic}^{5F}}{dx_F} + \frac{1}{2}\left(\frac{1}{10}\frac{dP_{icu}^{7F}}{dx_F} + \frac{1}{8}\frac{dP_{icu}^{7C}}{dx_F}\right) + \frac{1}{10}\frac{dP_{icd}^{7F}}{dx_F} + \frac{1}{10}\frac{dP_{ics}^{7F}}{dx_F} \tag{C.19}$$

$$\frac{dP_{D^+}}{dx_F} = \frac{1}{10}\frac{dP_{ic}^{5F}}{dx_F} + \frac{1}{10}\frac{dP_{icu}^{7F}}{dx_F} + \frac{1}{2}\left(\frac{1}{10}\frac{dP_{icd}^{7F}}{dx_F} + \frac{1}{8}\frac{dP_{icd}^{7C}}{dx_F}\right) + \frac{1}{10}\frac{dP_{ics}^{7F}}{dx_F} \tag{C.20}$$

$$\frac{dP_{D_s^-}}{dx_F} = \frac{1}{10}\frac{dP_{ic}^{5F}}{dx_F} + \frac{1}{10}\frac{dP_{icu}^{7F}}{dx_F} + \frac{1}{10}\frac{dP_{icd}^{7F}}{dx_F} + \frac{1}{2}\left(\frac{1}{10}\frac{dP_{ics}^{7F}}{dx_F} + \frac{1}{5}\frac{dP_{ics}^{7C}}{dx_F}\right) \tag{C.21}$$

$$\frac{dP_{D_s^+}}{dx_F} = \frac{1}{10}\frac{dP_{ic}^{5F}}{dx_F} + \frac{1}{10}\frac{dP_{icu}^{7F}}{dx_F} + \frac{1}{10}\frac{dP_{icd}^{7F}}{dx_F} + \frac{1}{2}\left(\frac{1}{10}\frac{dP_{ics}^{7F}}{dx_F} + \frac{1}{8}\frac{dP_{ics}^{7C}}{dx_F}\right) \tag{C.22}$$

$$\frac{dP_{\Xi_c^0}}{dx_F} = \frac{1}{10}\frac{dP_{ic}^{5F}}{dx_F} + \frac{1}{10}\frac{dP_{icu}^{7F}}{dx_F} + \frac{1}{10}\frac{dP_{icd}^{7F}}{dx_F} + \frac{1}{2}\left(\frac{1}{10}\frac{dP_{ics}^{7F}}{dx_F} + \frac{1}{8}\frac{dP_{ics}^{7C}}{dx_F}\right) \tag{C.23}$$

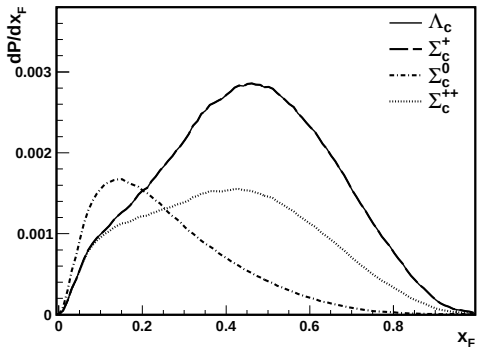
$$\frac{dP_{\Xi_c^+}}{dx_F} = \frac{1}{10}\frac{dP_{ic}^{5F}}{dx_F} + \frac{1}{10}\frac{dP_{icu}^{7F}}{dx_F} + \frac{1}{10}\frac{dP_{icd}^{7F}}{dx_F} + \frac{1}{2}\left(\frac{1}{10}\frac{dP_{ics}^{7F}}{dx_F} + \frac{1}{4}\frac{dP_{ics}^{7C}}{dx_F}\right) \tag{C.24}$$

$$\frac{dP_{Other}}{dx_F} = \frac{1}{10}\frac{dP_{ic}^{5F}}{dx_F} + \frac{1}{10}\frac{dP_{icu}^{7F}}{dx_F} + \frac{1}{10}\frac{dP_{icd}^{7F}}{dx_F} + \frac{1}{10}\frac{dP_{ics}^{7F}}{dx_F} \tag{C.25}$$

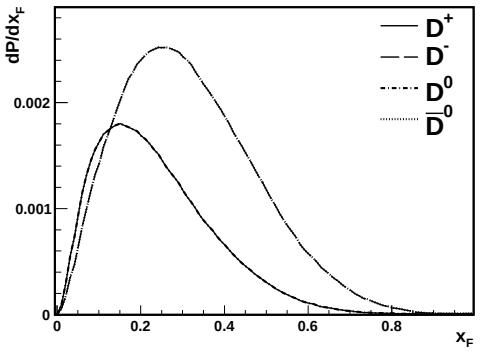
Figures C.1(a) to C.1(c) and C.2(a) to C.2(c) show the x_F distributions for charm and bottom hadrons, respectively, obtained from the equations above. The corresponding distributions for bottom hadrons formed from $b\bar{b}$ fluctuations, following table C.1, are:

$$\frac{dP_{\Lambda_b^0}}{dx_F} = \frac{1}{2}\left(\frac{1}{10}\frac{dP_{ib}^{5F}}{dx_F} + \frac{1}{2}\frac{dP_{ib}^{5C}}{dx_F}\right) + \frac{1}{2}\left(\frac{1}{10}\frac{dP_{ibu}^{7F}}{dx_F} + \frac{3}{8}\frac{dP_{ibu}^{7C}}{dx_F}\right) + \frac{1}{2}\left(\frac{1}{10}\frac{dP_{ibd}^{7F}}{dx_F} + \frac{1}{2}\frac{dP_{ibd}^{7C}}{dx_F}\right) + \frac{1}{2}\left(\frac{1}{10}\frac{dP_{ibs}^{7F}}{dx_F} + \frac{1}{4}\frac{dP_{ibd}^{7C}}{dx_F}\right) \tag{C.26}$$

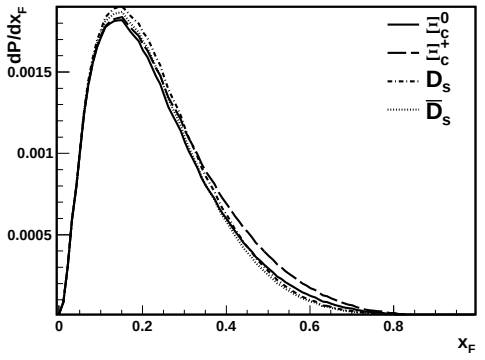
$$\frac{dP_{B^+}}{dx_F} = \frac{1}{2}\left(\frac{1}{10}\frac{dP_{ib}^{5F}}{dx_F} + \frac{1}{2}\frac{dP_{ib}^{5C}}{dx_F}\right) + \frac{1}{2}\left(\frac{1}{10}\frac{dP_{ibu}^{7F}}{dx_F} + \frac{3}{5}\frac{dP_{ibu}^{7C}}{dx_F}\right)$$



(a) Λ_c, Σ_c^0, Σ_c^+ and Σ_c^{++}



(b) $D^{\pm}$, D^0 and $\bar{D}^0$



(c) Ξ_c^0, Ξ_c^+, D_s and $\bar{D}_s$

Figure C.1: x_F distributions for bottom hadrons produced in the Intrinsic Quark model. Λ_c and Σ_c^0 distributions are coincident. The same is true for D^+, D^- distributions, and for D^0, $\bar{D}^0$ distributions.

$$+\frac{1}{2}\left(\frac{1}{10}\frac{dP^{7F}_{ibd}}{dx_F}+\frac{2}{5}\frac{dP^{7C}_{ibd}}{dx_F}\right)+\frac{1}{2}\left(\frac{1}{10}\frac{dP^{7F}_{ibs}}{dx_F}+\frac{2}{5}\frac{dP^{7C}_{ibd}}{dx_F}\right) \tag{C.27}$$

$$\frac{dP_{B^0}}{dx_F}=\frac{1}{2}\left(\frac{1}{10}\frac{dP^{5F}_{ib}}{dx_F}+\frac{1}{4}\frac{dP^{5C}_{ib}}{dx_F}\right)+\frac{1}{2}\left(\frac{1}{10}\frac{dP^{7F}_{ibu}}{dx_F}+\frac{1}{5}\frac{dP^{7C}_{ibu}}{dx_F}\right)$$
$$+\frac{1}{2}\left(\frac{1}{10}\frac{dP^{7F}_{ibd}}{dx_F}+\frac{2}{5}\frac{dP^{7C}_{ibd}}{dx_F}\right)+\frac{1}{2}\left(\frac{1}{10}\frac{dP^{7F}_{ibs}}{dx_F}+\frac{1}{5}\frac{dP^{7C}_{ibd}}{dx_F}\right) \tag{C.28}$$

$$\frac{dP_{\Sigma_b^+}}{dx_F}=\frac{1}{2}\left(\frac{1}{10}\frac{dP^{5F}_{ib}}{dx_F}+\frac{1}{4}\frac{dP^{5C}_{ib}}{dx_F}\right)+\frac{1}{2}\left(\frac{1}{10}\frac{dP^{7F}_{ibu}}{dx_F}+\frac{3}{8}\frac{dP^{7C}_{ibu}}{dx_F}\right)$$
$$+\frac{1}{2}\left(\frac{1}{10}\frac{dP^{7F}_{ibd}}{dx_F}+\frac{1}{8}\frac{dP^{7C}_{ibd}}{dx_F}\right)+\frac{1}{2}\left(\frac{1}{10}\frac{dP^{7F}_{ibs}}{dx_F}+\frac{1}{8}\frac{dP^{7C}_{ibd}}{dx_F}\right) \tag{C.29}$$

$$\frac{dP_{\Sigma_b^0}}{dx_F}=\frac{dP_{\Lambda_b^0}}{dx_F} \tag{C.30}$$

$$\frac{dP_{\Sigma_b^-}}{dx_F}=\frac{1}{10}\frac{dP^{5F}_{ib}}{dx_F}+\frac{1}{10}\frac{dP^{7F}_{ibu}}{dx_F}+\frac{1}{2}\left(\frac{1}{10}\frac{dP^{7F}_{ibd}}{dx_F}+\frac{1}{8}\frac{dP^{7C}_{ibd}}{dx_F}\right)+\frac{1}{10}\frac{dP^{7F}_{ibs}}{dx_F} \tag{C.31}$$

$$\frac{dP_{B^-}}{dx_F}=\frac{1}{10}\frac{dP^{5F}_{ib}}{dx_F}+\frac{1}{2}\left(\frac{1}{10}\frac{dP^{7F}_{ibu}}{dx_F}+\frac{1}{8}\frac{dP^{7C}_{ibu}}{dx_F}\right)+\frac{1}{10}\frac{dP^{7F}_{ibd}}{dx_F}+\frac{1}{10}\frac{dP^{7F}_{ibs}}{dx_F} \tag{C.32}$$

$$\frac{dP_{\bar{B}^0}}{dx_F}=\frac{1}{10}\frac{dP^{5F}_{ib}}{dx_F}+\frac{1}{10}\frac{dP^{7F}_{ibu}}{dx_F}+\frac{1}{2}\left(\frac{1}{10}\frac{dP^{7F}_{ibd}}{dx_F}+\frac{1}{8}\frac{dP^{7C}_{ibd}}{dx_F}\right)+\frac{1}{10}\frac{dP^{7F}_{ibs}}{dx_F} \tag{C.33}$$

$$\frac{dP_{B_s^0}}{dx_F}=\frac{1}{10}\frac{dP^{5F}_{ib}}{dx_F}+\frac{1}{10}\frac{dP^{7F}_{ibu}}{dx_F}+\frac{1}{10}\frac{dP^{7F}_{ibd}}{dx_F}+\frac{1}{2}\left(\frac{1}{10}\frac{dP^{7F}_{ibs}}{dx_F}+\frac{1}{5}\frac{dP^{7C}_{ibs}}{dx_F}\right) \tag{C.34}$$

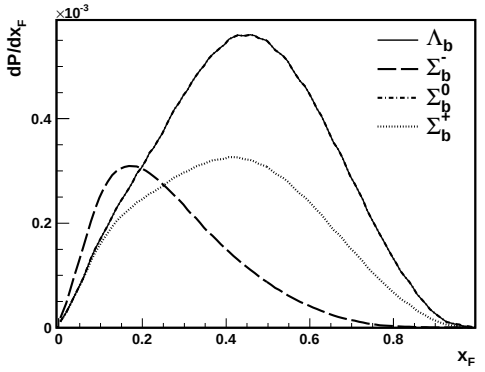
$$\frac{dP_{\bar{B}_s^0}}{dx_F}=\frac{1}{10}\frac{dP^{5F}_{ib}}{dx_F}+\frac{1}{10}\frac{dP^{7F}_{ibu}}{dx_F}+\frac{1}{10}\frac{dP^{7F}_{ibd}}{dx_F}+\frac{1}{2}\left(\frac{1}{10}\frac{dP^{7F}_{ibs}}{dx_F}+\frac{1}{8}\frac{dP^{7C}_{ibs}}{dx_F}\right) \tag{C.35}$$

$$\frac{dP_{\Xi_b^-}}{dx_F}=\frac{1}{10}\frac{dP^{5F}_{ib}}{dx_F}+\frac{1}{10}\frac{dP^{7F}_{ibu}}{dx_F}+\frac{1}{10}\frac{dP^{7F}_{ibd}}{dx_F}+\frac{1}{2}\left(\frac{1}{10}\frac{dP^{7F}_{ibs}}{dx_F}+\frac{1}{8}\frac{dP^{7C}_{ibs}}{dx_F}\right) \tag{C.36}$$

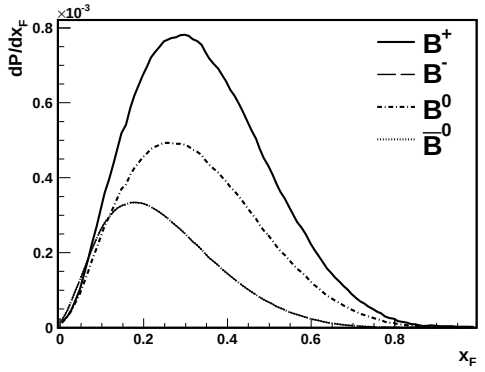
$$\frac{dP_{\Xi_b^0}}{dx_F}=\frac{1}{10}\frac{dP^{5F}_{ib}}{dx_F}+\frac{1}{10}\frac{dP^{7F}_{ibu}}{dx_F}+\frac{1}{10}\frac{dP^{7F}_{ibd}}{dx_F}+\frac{1}{2}\left(\frac{1}{10}\frac{dP^{7F}_{ibs}}{dx_F}+\frac{1}{4}\frac{dP^{7C}_{ibs}}{dx_F}\right) \tag{C.37}$$

$$\frac{dP_{Other}}{dx_F}=\frac{1}{10}\frac{dP^{5F}_{ib}}{dx_F}+\frac{1}{10}\frac{dP^{7F}_{ibu}}{dx_F}+\frac{1}{10}\frac{dP^{7F}_{ibd}}{dx_F}+\frac{1}{10}\frac{dP^{7F}_{ibs}}{dx_F} \tag{C.38}$$

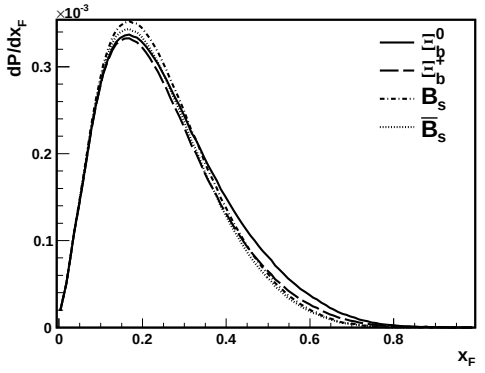
dP_{Other}/dx_F is the differential probability for any charmed or bottom hadron that cannot be generated by coalescence from 5- or 7-particle Fock states. i.e, to generate a $\Omega_c^0(\Omega_b^-)$ from coalescence, a fluctuation to a $|uudc\bar{c}s\bar{s}s\bar{s}>$ ($|uudb\bar{b}s\bar{s}s\bar{s}>$) 9-particle state would be needed. Considering only states up to 7 particles, $\Omega_c^0(\Omega_b^-)$ can only be formed by fragmentation.



(a) Λ_b, Σ_b^-, Σ_b^0 and Σ_b^+



(b) $B^{\pm}$, B^0 and $\bar{B}^0$



(c) Ξ_b^0, Ξ_b^+, B_s and $\bar{B}_s$

Figure C.2: x_F distributions for bottom hadrons produced in the Intrinsic Quark model. Λ_b and Σ_b^0 distributions are coincident. The same is true for B^-, B^0 distributions, and for B_s, $\bar{B}_s$ distributions.

List of figures

List of tables

Bibliography

[1] J. Abraham et al. Properties and performance of the prototype instrument for the Pierre Auger Observatory. *Nucl. Instrum. Meth. A*, 523(1-2):50–95, 2004.

[2] J. Abraham et al. Observation of the Suppression of the Flux of Cosmic Rays above $4 \cdot 10^{19}$ eV. *Phys. Rev. Lett.*, 101:061101, 2008.

[3] A. Creusot. Latest results of the Pierre Auger Observatory. *Nucl. Instrum. Meth. A*, 662(0):S106–S112, 2012.

[4] P. Abreu et al. The Pierre Auger Observatory III: Other Astrophysical Observations. *arXiv:1107.4805 [astro-ph.HE]*, 2011.

[5] P. Abreu et al. Search for point-like sources of ultra-high energy neutrinos at the Pierre Auger Observatory and improved limit on the diffuse flux of tau neutrinos. *Astrophys. J.*, 755:L4, 2012.

[6] P. Abreu et al. Search for ultrahigh energy neutrinos in highly inclined events at the Pierre Auger Observatory. *Phys. Rev. D*, 84:122005, 2011.

[7] Ralf Ulrich. Measurement of the proton-air cross-section with the Pierre Auger Observatory. *EPJ Web Conf.*, 53:07005, 2013.

[8] Jan Ebr. Measurement of the proton-air cross-section at $\sqrt{s} = 57$ TeV with the Pierre Auger Observatory. pages 341–347, 2012.

[9] A. Bueno, A. Gascón, J.I. Illana, and M. Masip. Propagation of B mesons in the atmosphere. *JCAP*, (02):028, 2012.

[10] A. Bueno and A. Gascón. CORSIKA implementation of heavy quark production and propagation in extensive air showers. *Comp. Phys. Comm.*, 185(2):638 – 650, 2014.

[11] L. Stodolsky and L. McLerran. Penetrating flavors. *Phys. Lett. B*, 109:485–489, 1982.

[12] E.M. Aitala et al. Differential cross-sections, charge production asymmetry, and spin density matrix elements for $D^{*\pm}(2010)$ produced in 500-GeV/c π-nucleon interactions. *Phys. Lett. B*, 539:218–226, 2002.

[13] P. Chauvat et al. Production of Λ_c with large x_f at the ISR. *Phys. Lett. B*, 199:304, 1987.

[14] E.M. Aitala et al. Asymmetries in the production of Λ_c^+ and Λ_c^- baryons in 500-GeV/c π^-–nucleon interactions. *Phys. Lett. B*, 495:42–48, 2000.

[15] G. Bari, M. Basile, G. Bruni, G. Cara Romeo, R. Casaccia, et al. The Λ_b^0 beauty baryon production in proton proton interactions at $\sqrt{s} = 62$ GeV: A Second observation. *Nuovo Cim. A*, 104:1787–1800, 1991.

[16] J. Badier et al. Evidence for $\psi\psi$ production in π^- interactions at 150 GeV/c and 280 GeV/c. *Phys. Lett. B*, 114:457, 1982.

[17] S.J. Brodsky, B. Kopeliovich, I. Schmidt, and J. Soffer. Diffractive Higgs production from intrinsic heavy flavors in the proton. *Phys. Rev. D*, 73:113005, 2006.

[18] T. Gutiérrez and R. Vogt. Leading charm in hadron nucleus interactions in the intrinsic charm model. *Nucl. Phys. B*, 539:189–214, 1999.

[19] R. Vogt and S.J. Brodsky. QCD and intrinsic heavy quark predictions for leading charm and beauty hadroproduction. *Nucl. Phys. B*, 438(1-2):261–277, 1995.

[20] E. Norrbin and R. Vogt. Bottom Production Asymmetries at the LHC. *arXiv:0003056 [hep-ph]*, 2000.

[21] V.F. Hess. Über Beobachtungen der durchdringenden Strahlung bei sieben Freiballonfahrten (Observation of Penetrating Radiation in Seven Balloon Flights). *Phys. Zeitschr.*, 13:1084, 1912.

[22] T.H. Johnson. The Azimuthal Asymmetry of the Cosmic Radiation. *Phys. Rev.*, 43:834–835, 1933.

[23] L. Álvarez and A.H. Compton. A Positively Charged Component of Cosmic Rays. *Phys. Rev.*, 43:835–836, 1933.

[24] B. Rossi. On the Magnetic Deflection of Cosmic Rays. *Phys. Rev.*, 36:606–606, 1930.

[25] W. Kolhörster, I. Matthes, and E. Weber. Gekoppelte Höhenstrahlen. *Naturwissenschaften*, 26:576–576, 1938.

[26] P. Auger, R. Maze, and T. Grivet-Mayer. Extensive cosmic showers in the atmosphere containing ultra-penetrating particles. *C. R. Acad. Sci.*, 206:1721–1722, 1938.

[27] M. Schein, W.P. Jesse, and E.O. Wollan. The Nature of the Primary Cosmic Radiation and the Origin of the Mesotron. *Phys. Rev.*, 59:615–615, 1941.

[28] J. Linsley. Evidence for a Primary Cosmic-Ray Particle with Energy 10^{20} eV. *Phys. Rev. Lett.*, 10:146–148, 1963.

[29] M.M. Winn, J. Ulrichs, L.S. Peak, C.B.A. McCusker, and L. Horton. The Cosmic Ray Energy Spectrum Above 10^{17} eV. *J. Phys. G*, 12:653–674, 1986.

[30] M.A. Lawrence, R.J.O. Reid, and A.A. Watson. The Cosmic ray energy spectrum above $4 \cdot 10^{17}$ eV as measured by the Haverah Park array. *J. Phys. G*, 17:733–757, 1991.

[31] N.N. Efimov, T.A. Egorov, A.V. Glushkov, M.I. Pravdin, and I.E. Sleptsov. The energy spectrum and anisotropy of primary cosmic rays at energy $E > 10^{17}$ eV observed in Yakutsk. 1990. Published in Kofu 1990, Proceedings, Astrophysical aspects of the most energetic cosmic rays, pages 20-33.

[32] M. Nagano, M. Teshima, Y. Matsubara, H.Y. Dai, T. Hara, N. Hayashida, M. Honda, H. Ohoka, and S. Yoshida. Energy spectrum of primary cosmic rays above 10^{17} eV determined from extensive air shower experiments at Akeno. *J. Phys. G*, 18(2):423, 1992.

[33] R.M. Baltrusaitis, R. Cady, G.L. Cassiday, R. Cooper, J.W. Elbert, et al. The Utah Fly's Eye Detector. *Nucl. Instrum. Meth. A*, 240:410–428, 1985.

[34] H. Tokuno, Y. Tameda, M. Takeda, K. Kadota, D. Ikeda, et al. New air fluorescence detectors employed in the Telescope Array experiment. *Nucl. Instrum. Meth. A*, 676:54–65, 2012.

[35] P. Abreu et al. The Pierre Auger Observatory I: The Cosmic Ray Energy Spectrum and Related Measurements. *arXiv:1107.4809v1 [astro-ph.HE]*, 2011.

[36] W.D. Apel et al. The spectrum of high-energy cosmic rays measured with KASCADE-Grande. *arXiv:1206.3834 [astro-ph.HE]*, 2012.

[37] R. Battiston, P. Blasi, M.T. Brunetti, D. de Marco, P. Lipari, B. Sacco, and A. Santangelo. Cosmic ray physics from low to extreme energies: Status and perspectives. *Adv. Space Res.*, 37(10):1834–1840, 2006.

[38] B. Klecker. Anomalous Cosmic Rays: Our present understanding and open questions. *Adv. Space Res.*, 23(3):521–530, 1999.

[39] V.I.L. Ginzburg and S.I. Syrovatskii. *The Origin of Cosmic Rays.* A Pergamon Press Book. Macmillan, 1964.

[40] J.R. Hörandel. Cosmic-ray composition and its relation to shock acceleration by supernova remnants. *Adv. Space Res.*, 41(3):442 – 463, 2008.

[41] P.O. Lagage and C.J. Cesarsky. The maximum energy of cosmic rays accelerated by supernova shocks. *Astron. Astrophys.*, 125:249–257, 1983.

[42] W.D. Apel, J.C. Arteaga-Velázquez, K. Bekk, M. Bertaina, J. Blümer, et al. Ankle-like Feature in the Energy Spectrum of Light Elements of Cosmic Rays Observed with KASCADE-Grande. *Phys. Rev. D*, 87:081101, 2013.

[43] T. Antoni et al. KASCADE measurements of energy spectra for elemental groups of cosmic rays: Results and open problems. *Astropart. Phys.*, 24:1–25, 2005.

[44] A. Tamburro. Cosmic-ray Spectrum and Composition with the IceCube Observatory. *arXiv:1210.7526 [astro-ph.HE].*

[45] J.R. Hörandel. Models of the knee in the energy spectrum of cosmic rays. *Astropart. Phys.*, 21:241–265, 2004.

[46] W.D. Apel et al. Kneelike structure in the spectrum of the heavy component of cosmic rays observed with KASCADE-Grande. *Phys. Rev. Lett.*, 107:171104, 2011.

[47] J.R. Hörandel. The Composition of Cosmic Rays at the Knee. *AIP Conf. Proc.*, 1516:185–194, 2012.

[48] J.R. Hörandel, N.N. Kalmykov, and A.V. Timokhin. Propagation of super high-energy cosmic rays in the Galaxy. *Astropart. Phys.*, 27:119–126, 2007.

[49] A. Letessier-Selvon and T. Stanev. Ultrahigh energy cosmic rays. *Rev. Mod. Phys.*, 83:907–942, 2011.

[50] E. Fermi. On the Origin of the Cosmic Radiation. *Phys. Rev.*, 75:1169–1174, 1949.

[51] W.I. Axford. The Origins of high-energy cosmic rays. *Astrophys. J., Suppl. Ser.*, 90:937–944, 1994.

[52] A.M. Hillas. The Origin of Ultra-High Energy Cosmic Rays. *Ann. Rev. Astron. Astrophys.*, 22(1):425–444, 1984.

[53] K. Kotera and A.V. Olinto. The Astrophysics of Ultrahigh Energy Cosmic Rays. *Ann. Rev. Astron. Astrophys.*, 49:119–153, 2011.

[54] J. Abraham et al. An upper limit to the photon fraction in cosmic rays above 10^{19} eV from the Pierre Auger Observatory. *Astropart. Phys.*, 27(2-3):155–168, 2007.

[55] J. Abraham et al. Upper limit on the cosmic-ray photon fraction at EeV energies from the Pierre Auger Observatory. *Astropart. Phys.*, 31(6):399–406, 2009.

[56] A.A. Penzias and R.W. Wilson. A Measurement of Excess Antenna Temperature at 4080 Mc/s. *Astrophys. J.*, 142:419–421, 1965.

[57] K. Greisen. End to the Cosmic-Ray Spectrum? *Phys. Rev. Lett.*, 16:748–750, 1966.

[58] G.T. Zatsepin and V.A. Kuzmin. Upper limit of the spectrum of cosmic rays. *JETP Lett.*, 4:78–80, 1966.

[59] M. Nagano and A.A. Watson. Observations and Implications of the Ultra-high Energy Cosmic Rays. *Rev. Mod. Phys.*, 72:689–732, 2000.

[60] R.U. Abbasi, T. Abu-Zayyad, M. Al-Seady, M. Allen, J.F. Amann, et al. Measurement of the Flux of Ultra-High Energy Cosmic Rays by the Stereo Technique. *Astropart. Phys.*, 32:53–60, 2009.

[61] T. Abu-Zayyad, R. Aida, M. Allen, R. Anderson, R. Azuma, et al. The Cosmic Ray Energy Spectrum Observed with the Surface Detector of the Telescope Array Experiment. *Astrophys. J.*, 768:L1, 2013.

[62] K. Shinozaki and M. Teshima. AGASA Results. *Nucl. Phys. B (Proc. Suppl.)*, 136(0):18–27, 2004.

[63] A.V. Olinto. Cosmic Rays at the Highest Energies. *Journal of Physics: Conference Series*, 375(5):052001, 2012.

[64] R. Aloisio. Propagation and Energy Spectrum of Ultra High Energy Cosmic Rays. *Nucl. Phys. B (Proc. Suppl.)*, 239 240:169 – 175, 2013.

[65] L. Cazón, R.A. Vázquez, and E. Zas. Depth development of extensive air showers from muon time distributions. *Astropart. Phys.*, 23(4):393–409, 2005.

[66] L. Cazón. Extensive Air Showers: from the muonic smoking guns to the hadronic backbone. *arXiv:1301.3340 [astro-ph.HE]*, 2013.

[67] P. Abreu et al. The Pierre Auger Observatory II: Studies of Cosmic Ray Composition and Hadronic Interaction models. *arXiv:1107.4804 [astro-ph.HE]*, 2011.

[68] M.T. Dova, M.E. Manceñido, A.G. Mariazzi, H. Wahlberg, F. Arqueros, and D. García-Pinto. Time asymmetries in extensive air showers: A novel method to identify UHECR species. *Astropart. Phys.*, 31(4):312–319, 2009.

[69] P. Sokolsky. Final Results from the High Resolution Fly's Eye (HiRes) Experiment. *Nucl. Phys. B (Proc. Suppl.)*, 212-213(0):74–78, 2011.

[70] R.U. Abbasi et al. Indications of Proton-Dominated Cosmic-Ray Composition above 1.6 EeV. *Phys. Rev. Lett.*, 104:161101, 2010.

[71] *Proceedings of International Symposium on Future Directions in UHECR Physics (UHECR2012)*, volume 53, 2013.

[72] E. Barcikowski et al. Mass Composition Working Group Report at UHECR-2012. *EPJ Web Conf.*, 53:01006, 2013.

[73] J. Bellido. Review of UHECR composition data Working Group. *Proceedings of International Symposium on Future Directions in UHECR Physics (UHECR2012)*, 2012.

[74] Y. Tsunesada. Highlights from Telescope Array. *arXiv:1111.2507 [astro-ph.HE]*, 2011.

[75] R.U. Abbasi et al. Analysis of Large-scale Anisotropy of Ultra-high Energy Cosmic Rays in HiRes Data. *Astrophys. J. Lett.*, 713(1):L64, 2010.

[76] M. Teshima et al. The arrival direction distribution of extremely high energy cosmic rays observed by AGASA. *28th ICRC, Tsukuba, Japan*, pages 437–440, 2003.

[77] H.B. Kim and J. Kim. Update of Correlation Analysis between Active Galactic Nuclei and Ultra-High Energy Cosmic Rays. *Int. J. Mod. Phys. D*, 22:1350045, 2013.

[78] H.B. Kim and J. Kim. Correlation of Ultra-High Energy Cosmic Rays with Active Galactic Nuclei. *AIP Conf. Proc.*, 1367:193–196, 2011.

[79] J. Abraham et al. Correlation of the Highest-Energy Cosmic Rays with Nearby Extragalactic Objects. *Science*, 318(5852):938–943, 2007.

[80] J. Abraham et al. Correlation of the highest-energy cosmic rays with the positions of nearby active galactic nuclei. *Astropart. Phys.*, 29(3):188–204, 2008.

[81] P. Abreu et al. Update on the correlation of the highest energy cosmic rays with nearby extragalactic matter. *Astropart. Phys.*, 34(5):314–326, 2010.

[82] M.P. Véron-Cetty and P. Véron. A catalogue of quasars and active nuclei: 12th edition. *Astropart. and Astroph.*, 455(2):773–777, 2006.

[83] K.H. Kampert. Highlights from the Pierre Auger Observatory. *arXiv:1207.4823 [astro-ph.HE]*, 2012.

[84] T. Abu-Zayyad, R. Aida, M. Allen, R. Anderson, R. Azuma, et al. Search for Anisotropy of Ultra-High Energy Cosmic Rays with the Telescope Array Experiment. *Astrophys. J.*, 757:26, 2012.

[85] J. Álvarez Muñiz, M. Risse, G.I. Rubtsov, and B.T. Stokes. Review of the Multimessenger Working Group at UHECR-2012. *Proceedings of International Symposium on Future Directions in UHECR Physics (UHECR2012)*, 2012.

[86] T.K. Gaisser. *Cosmic Rays and Particle Physics*. Cambridge University Press, 1990.

[87] G. Gelmini, O.E. Kalashev, and D.V. Semikoz. GZK photons as ultra high energy cosmic rays. *J. Exp. Theor. Phys.*, 106, 2008.

[88] D. Hooper, A.M. Taylor, and S. Sarkar. Cosmogenic photons as a test of ultra-high energy cosmic ray composition. *Astropart. Phys.*, 34, 2011.

[89] M. Kachelriess, S. Ostapchenko, and R. Tomas. TeV gamma-rays from UHECR interactions in AGN cores: Lessons from Centaurus A. *Publ. Astron. Soc. Austral.*, 27, 2010.

[90] R. Abbasi et al. Constraints on the extremely-high energy cosmic neutrino flux with the IceCube 2008-2009 data. *Phys. Rev. D*, 83:092003, 2011.

[91] R.U. Abbasi et al. An Upper Limit on the Electron-Neutrino Flux from the HiRes Detector. *Astrophys. J.*, 684(2):790, 2008.

[92] I. Kravchenko, S. Hussain, D. Seckel, D. Besson, E. Fensholt, J. Ralston, J. Taylor, K. Ratzlaff, and R. Young. Updated results from the RICE experiment and future prospects for ultra-high energy neutrino detection at the south pole. *Phys. Rev. D*, 85:062004, 2012.

[93] P.W. Gorham et al. Observational constraints on the ultrahigh energy cosmic neutrino flux from the second flight of the ANITA experiment. *Phys. Rev. D*, 82:022004, 2010.

[94] K. Kotera, D. Allard, and A.V. Olinto. Cosmogenic Neutrinos: parameter space and detectabilty from PeV to ZeV. *JCAP*, 1010:013, 2010.

[95] M. Ahlers, L.A. Anchordoqui, M.C. González-García, F. Halzen, and S. Sarkar. GZK neutrinos after the Fermi-LAT diffuse photon flux measurement. *Astropart. Phys.*, 34(2):106 – 115, 2010.

[96] J.A. Aguilar. Neutrino searches with the IceCube telescope. *arXiv:1301.6504 [astro-ph.HE]*, 2013.

[97] H. Costantini. Status and recent results of the ANTARES deep-sea neutrino telescope. *Nucl. Instrum. Meth. A*, 692:41–45, 2012.

[98] M.G. Aartsen et al. First observation of PeV-energy neutrinos with IceCube. *arXiv:1304.5356 [astro-ph.HE]*, 2013.

[99] G.A. Askaryan. Excess Negative Charge of an electron-photon Shower and its coherent radio emission. *Sov. Phys. JETP-USSR*, 14(2):441–443, 1962.

[100] R.J. Nichol. Radio detection of high-energy particles with the ANITA experiment. *Nucl. Instrum. Meth. A*, 626-627:S30–S35, 2011.

[101] I. Kravchenko. Status of RICE and preparations for the next generation radio neutrino experiment in Antarctica. *Nucl. Instrum. Meth. A*, 692:233–235, 2012.

[102] K. Hanson. The Askar'yan Radio Array. *J. Phys. Conf. Ser.*, 375:052037, 2012.

[103] S.R. Klein. ARIANNA: A radio detector array for cosmic neutrinos on the Ross Ice Shelf. *IEEE Trans. Nucl. Sci.*, 60:637–643, 2013.

[104] J.D. Bray, R.D. Ekers, R.J. Protheroe, C.W. James, C.J. Phillips, et al. LUNASKA neutrino search with the Parkes and ATCA telescopes. *arXiv:1301.6490 [astro-ph.IM]*, 2013.

[105] W. Heitler. *The Quantum Theory of Radiation*. Dover Books on Physics Series. New York – Dover Publications, Incorporated, 1954.

[106] J. Matthews. A Heitler Model of Extensive Air Showers. *Astropart. Phys.*, 22(5-6):387–397, 2005.

[107] J.R. Hörandel. Cosmic Rays from the Knee to the Second Knee: 10^{14} to 10^{18} eV. *Modern Phys. Lett. A*, 22(21):1533–1551, 2007.

[108] J.R. Hörandel. On total inelastic cross sections and the average depth of the maximum of extensive air showers. *J. Phys. G*, 29(11):2439, 2003.

[109] J. Álvarez-Muñiz, R. Engel, T.K. Gaisser, J.A. Ortiz, and T. Stanev. Hybrid Simulations of Extensive Air Showers. *Phys. Rev. D*, 66:033011, 2002.

[110] J. Linsley. Structure of large air showers at depth 834 g/cm^2 III - Applications. In *International Cosmic Ray Conference*, volume 12 of *International Cosmic Ray Conference*, pages 89–96, 1977.

[111] F. Catalani, J.A. Chinellato, V. de Souza, J. Takahashi, and G.M.S. Vasconcelos. Statistical methods applied to composition studies of ultrahigh energy cosmic rays. *Astropart. Phys.*, 28:357–365, 2007.

[112] A.M. Hillas. Cosmic Rays: Recent Progress and some Current Questions. *arXiv:0607109v2 [astro-ph]*, 2006.

[113] P. Sokolsky, P. Sommers, and B.R. Dawson. Extremely high-energy cosmic rays. *Phys. Rep.*, 217:225–277, 1992.

[114] S.P. Knurenko and A. Sabourov. Study of cosmic rays at the Yakutsk EAS array: Energy spectrum and mass composition. *Nucl. Phys. Proc. Suppl.*, 212-213:241–251, 2011.

[115] J. Linselly. New Detectors at Volcano Range for Investigating Primary Composition of Highest-Energy Cosmic Rays. In *International Cosmic Ray Conference*, volume 11 of *International Cosmic Ray Conference*, page 155, 1979.

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

[117] M. Sasaki. The Telescope Array project. ICRR-REPORT-481-2001-11. 2001.

[118] I. Allekotte et al. The Surface Detector System of the Pierre Auger Observatory. *Nucl. Instrum. Meth. A*, 586:409–420, 2008.

[119] J.A. Goodman. HAWC: The high altitude water Cherenkov observatory. *AIP Conf. Proc.*, 1516:265–268, 2012.

[120] R.W. Springer. Recent results from the HiRes air fluorescence experiment. *Nucl. Phys. B (Proc. Suppl.)*, 138(0):307–309, 2005.

[121] J. Abraham et al. The fluorescence detector of the Pierre Auger Observatory. *Nucl. Instrum. Meth. A*, 620(2-3):227–251, 2010.

[122] S.F. Berezhnev, D. Besson, N.M. Budnev, A. Chiavassa, O.A. Chvalaev, et al. The Tunka-133 EAS Cherenkov light array: status of 2011. *Nucl. Instrum. Meth. A*, 692:98–105, 2012.

[123] M. Persic. CTA: the future of ground-based gamma-ray astrophysics. *Nucl. Phys. B (Proc. Suppl.)*, 239-240:210–215, 2013.

[124] H. Flacke for the LOPES Collaboration. Radio Detection of Ultra-High Energy Cosmic Rays, 30th ICRC, Mérida, México. *arXiv:0804.0548 [astro-ph]*, 2007.

[125] P. Facal San Luis, I. Alekotte, J. Álvarez, et al. Microwave detection of air showers with MIDAS. *Nucl. Instrum. Meth. A*, 662(0):S118–S123, 2012.

[126] K. de Souza, M. Gillman, S. Hart, J. Lloyd-Evans, J.E. McMillan, and A.A. Watson. Design studies for a very large detector array to study cosmic rays at the highest energies. *Nucl. Phys. B (Proc. Suppl.)*, 28(2):135–142, 1992.

[127] D. Newton, J. Knapp, and A.A. Watson. The Optimum Distance at which to Determine the Size of a Giant Air Shower. *Astropart. Phys.*, 26:414–419, 2007.

[128] T. Waldenmaier, J. Blümer, and H. Klages. Spectral resolved measurement of the nitrogen fluorescence emissions in air induced by electrons. *Astropart. Phys.*, 29(3):205–222, 2008.

[129] J.W. Belz et al. Measurement of pressure dependent fluorescence yield of air: Calibration factor for UHECR detectors. *Astropart. Phys.*, 25:129–139, 2006.

[130] S. Yoshida. Energy determination of trans-EeV cosmic rays. *CR. Phys.*, 5(4):483–493, 2004.

[131] P. Sommers. Extensive air showers and measurement techniques. *CR. Phys.*, 5(4):463–472, 2004.

[132] B. Falkenburg and W. Rhode. *From Ultra Rays to Astroparticles: A Historical Introduction to Astropart. Phys.* SpringerLink : Bücher. 2012.

[133] Unesco. Science Cooperation Office for Latin America. *Proceedings of the Fifth Interamerican Seminar on Cosmic Rays: La Paz (Bolivia) 17-27 July 1962*. Number v.1. Laboratorio de Física Cósmica de la Universidad Mayor de San Andrés, 1962.

[134] T. Hara, F. Ishikawa, S. Kawaguchi, Y. Miura, M. Nagano, K. Suga, and G. Tanahashi. Detection of the atmospheric scintillation light from air showers. In *International Cosmic Ray Conference*, volume 3 of *International Cosmic Ray Conference*, page 369, 1970.

[135] H.E. Bergeson, G.L. Cassiday, T.W. Chiu, et al. Measurement of Light Emission from Remote Cosmic-Ray Air Showers. *Phys. Rev. Lett.*, 39:847–849, 1977.

[136] T. K. Gaisser and A. M. Hillas. Reliability of the method of constant intensity cuts for reconstructing the average development of vertical showers. In *International Cosmic Ray Conference*, volume 8 of *International Cosmic Ray Conference*, pages 353–357, 1977.

[137] A.G. Mariazzi and M.J. Tueros. Estimating the missing energy of Extensive Air Showers at the Pierre Auger Observatory (Internal note). GAP 2013-026, 2013.

[138] P.A. Cherenkov. Radiation of Ultralight Speed Charged Particles and its Utilization in the Physics of High-Energy Particles - Report FTD-TT-612-846. 1962.

[139] N. Gotting. Recent results from HEGRA: Gamma-ray observations with the HEGRA stereoscopic system of 5 Cherenkov telescopes. *Eur. Phys. J. C*, 33:S932–S934, 2004.

[140] S.K. Basu. *Encyclopaedic Dictionary of Astrophysics.* Global Vision Publishing House, 2007.

[141] R. Zanin. Overview of galactic results obtained by MAGIC. *Nucl. Phys. B (Proc. Suppl.)*, 239-240:55–60, 2013.

[142] C. Stegmann. Highlights from H.E.S.S. *AIP Conf. Proc.*, 1505:194–201, 2012.

[143] J. Vandenbroucke, B. Humensky, J. Holder, and R.A. Ong. Fundamental Physics with Charged Particle Measurements at the Cherenkov Telescope Array. *arXiv:1305.0022 [astro-ph.HE]*, 2013.

[144] J.V. Jelley, J.H. Fruin, N.A. Porter, T.C. Weekes, F.G. Smith, et al. Radio Pulses from Extensive Cosmic-Ray Air Showers. *Nature*, 205:327–328, 1965.

[145] J. Nichimura. Radio wave emitted by an extensive air showers in 10 kHz to 1 MHz region. In F.C. Jones, editor, *International Cosmic Ray Conference*, volume 7 of *International Cosmic Ray Conference*, pages 308–311, 1985.

[146] J.H. Hough. Calculations on the radio emission resulting from geomagnetic charge separation in an extensive air shower. *J. Phys. A*, 6:892–900, 1973.

[147] K. Boruah, N.M. Saikia, M. Rahman, and P.K. Boruah. Low frequency radio-emission associated with UHE cosmic rays. *Nucl. Phys. B (Proc. Suppl.)*, 212-213(0):317–322, 2011.

[148] T. Huege and H. Falcke. Radio emission from cosmic ray air showers: Coherent geosynchrotron radiation. *Astron. Astrophys.*, 412:19–34, 2003.

[149] W.D. Apel et al. Progress in air shower radio measurements: Detection of distant events. *Astropart. Phys.*, 26(4-5):332–340, 2006.

[150] D. Ardouin, A. Bellétoile, D. Charrier, et al. Radio-detection signature of high-energy cosmic rays by the CODALEMA experiment. *Nucl. Instrum. Meth. A*, 555(1-2):148–163, 2005.

[151] T. Huege et al. The LOPES experiment – Recent results, status and perspectives. *Nucl. Instrum. Meth. A*, 662(0):S72 – S79, 2012.

[152] D.T. Machado and the CODALEMA Collaboration. Latest results of the CODALEMA experiment: cosmic rays radio detection in a self trigger mode. *Journal of Physics: Conference Series*, 409(1):012074, 2013.

[153] F.G. Schröder for the Pierre Auger Collaboration. Radio detection of air showers with the Auger Engineering Radio Array. *arXiv:1307.5059 [astro-ph.HE]*, 2013.

[154] C. Williams. Microwave Detection of Cosmic Ray Air Showers at the Pierre Auger Observatory, an R&D Effort. *Phys. Proc.*, 37(0):1341–1348, 2012.

[155] J.L. Kelley. AERA: the Auger Engineering Radio Array, 31st ICRC, Beijing, China. 2011.

[156] P. Abreu et al. Antennas for the Detection of Radio Emission Pulses from Cosmic-Ray. *JINST*, 7:P10011, 2012.

[157] J. Petrovic et al. Radio emission of highly inclined cosmic ray air showers measured with LOPES: Possibility for neutrino detection. *J. Phys. Conf. Ser.*, 39:471–474, 2006.

[158] O. Scholten, J. Bacelar, R. Braun, A.G. de Bruyn, H. Falcke, et al. Optimal radio window for the detection of ultra-high energy cosmic rays and neutrinos off the moon. *Astropart. Phys.*, 26:219–229, 2006.

[159] P.W. Gorham, N.G. Lehtinen, G.S. Varner, J.J. Beatty, et al. Observations of Microwave Continuum Emission from Air Shower Plasmas. *Phys. Rev. D*, 78:032007, 2008.

[160] S. Baur, M. Bertaina, J. Blümer, et al. On the Microwave Signal of Air Showers Observed with CROME (Internal note). GAP 2012-061, 2012.

[161] J.J.P. Burrows, P. Borrell, and U. Platt. *The Remote Sensing of Tropospheric Composition from Space*. Physics of Earth and Space Environments. Springer Berlin Heidelberg, 2011.

[162] P. Abreu et al. The Pierre Auger Observatory V: Enhancements. *J. Phys. Conf. Ser.*, 375:052006, 2012.

[163] R. Smida et al. First results of the CROME experiment. *arXiv:1108.0588 [hep-ex]*, 2011.

[164] R. Smida, F. Werner, R. Engel, J.C. Arteaga-Velázquez, K. Bekk, et al. Observation of Polarised Microwave Emission from Cosmic Ray Air Showers. *arXiv:1306.6738 [astro-ph.IM]*, 2013.

[165] R. Gaïor for the Pierre Auger Collaboration. Detection of cosmic rays using microwave radiation at the Pierre Auger Observatory. *arXiv:1307.5059 [astro-ph.HE]*, 2013.

[166] F. Suárez for the Pierre Auger Collaboration. The AMIGA muon detectors of the Pierre Auger Observatory: overview and status. *arXiv:1307.5059 [astro-ph.HE]*, 2013.

[167] C. Meurer and N. Scharf. HEAT - a low energy enhancement of the Pierre Auger Observatory. *Astrophys. Space Sci. Trans.*, 7:183–186, 2011.

[168] B. Fick, M. Malek, J.A.J. Matthews, J. Matthews, R. Meyhandan, M. Mostafá, M. Roberts, P. Sommers, and L. Wiencke. The Central Laser Facility at the Pierre Auger Observatory. *J. Instrum.*, 1(11):P11003, 2006.

[169] S.Y. BenZvi, R. Cester, M. Chiosso, B.M. Connolly, A. Filipcic, et al. The Lidar System of the Pierre Auger Observatory. *Nucl. Instrum. Meth. A*, 574:171–184, 2007.

[170] J. Chirinos for the Pierre Auger Collaboration. Cloud Monitoring at the Pierre Auger Observatory. *arXiv:1307.5059 [astro-ph.HE]*, 2013.

[171] M. Aglietta et al. Calibration of the surface array of the Pierre Auger Observatory – FERMILAB-CONF-05-281-E-TD. 2005.

[172] A. Etchegoyen, P. Bauleo, X. Bertou, et al. Muon-track studies in a water Cherenkov detector. *Nucl. Instrum. Meth. A*, 545(3):602–612, 2005.

[173] C. Bonifazi. The angular resolution of the Pierre Auger Observatory. *Nucl. Phys. B (Proc. Suppl.)*, 190:20–25, 2009.

[174] H. Tokuno et al. New air fluorescence detectors employed in the Telescope Array experiment. *Nucl. Instrum. Meth. A*, 676(0):54–65, 2012.

[175] J. Abraham et al. A study of the effect of molecular and aerosol conditions in the atmosphere on air fluorescence measurements at the Pierre Auger Observatory. *Astropart. Phys.*, 33(2):108–129, 2010.

[176] P. Abreu et al. Techniques for Measuring Aerosol Attenuation using the Central Laser Facility at the Pierre Auger Observatory. *JINST*, 8:P04009, 2013.

[177] S.Y. BenZvi, B.M. Connolly, J.A.J. Matthews, M. Prouza, E.F. Visbal, et al. Measurement of the Aerosol Phase Function at the Pierre Auger Observatory. *Astropart. Phys.*, 28:312–320, 2007.

[178] K. Loudec and M. Urban. Where do aerosols come from? Study done with HYSPLIT (Internal note). GAP 2011-058, 2011.

[179] K. Loudec for the Pierre Auger Collaboration. Atmospheric Monitoring at the Pierre Auger Observatory - Status and Update. *arXiv:1107.4806 [astro-ph.IM]*.

[180] C. Bonifazi. Angular resolution of the Pierre Auger Observatory – FERMILAB-CONF-05-301-E-TD. 2005.

[181] M. Unger. Shower Profile Reconstruction from Fluorescence and Cherenkov Light (Internal note). GAP 2006-010, 2006.

[182] V. Verzi for the Pierre Auger Collaboration. The energy scale of the Pierre Auger Observatory. *arXiv:1307.5059 [astro-ph.HE]*, 2013.

[183] D. Barnhill et al. Measurement of the lateral distribution function of UHECR air showers with the Pierre Auger Observatory. 2005.

[184] D Newton, J. Knapp, and A.A. Watson. The Optimum Distance at which to Determine the Size of a Giant Air Shower. *Astropart. Phys.*, 26:414–419, 2007.

[185] I. Valiño for the Pierre Auger Collaboration. A measurement of the muon number in showers using inclined events recorded at the Pierre Auger Observatory. *arXiv:1307.5059 [astro-ph.HE]*, 2013.

[186] A. Schulz for the Pierre Auger Collaboration. The measurement of the energy spectrum of cosmic rays above $3{\cdot}10^{17}$ eV with the Pierre Auger Observatory. *arXiv:1307.5059 [astro-ph.HE]*, 2013.

[187] B. Revenu. Radio detection of cosmic ray air showers by the RAuger experiment, a fully autonomous and self-triggered system installed at the Pierre Auger Observatory. *Nucl. Instrum. Meth. A*, 662(0):S130–S133, 2012.

[188] H. Schoorlemmer. Results from polarization studies of radio signals induced by cosmic rays at the Pierre Auger Observatory. *Nucl. Instrum. Meth. A*, 662:S134–S137, 2012.

[189] J. Rich, D. Lloyd Owen, and M. Spiro. Experimental Particle Physics Without Accelerators. *Phys. Rep.*, 151:239–364, 1987.

[190] V.V. Akimov, N.L. Grigorov, N.A. Mamontova, V.E. Nesterov, V.L. Prokhin, et al. Measurements of the inelastic proton-proton and proton-carbon cross-sections at energies 10^{10} to 10^{12} eV on board the satellites proton 1, 2 and 3. *Conf. Proc.*, C690825:211–214, 1969.

[191] G.B. Yodh, Y. Pal, and J.S. Trefil. Evidence for Rapidly Rising p–p Total Cross Section from Cosmic-Ray Data. *Phys. Rev. Lett.*, 28:1005–1008, 1972.

[192] T.K. Gaisser, U.P. Sukhatme, and G.B. Yodh. Hadron cross sections at ultrahigh energies and unitarity bounds on diffraction. *Phys. Rev. D*, 36:1350–1357, 1987.

[193] T. H. Burnett et al. Extremely High Multiplicities in High-Energy Nucleus-Nucleus Collisions. *Phys. Rev. Lett.*, 50:2062–2065, 1983.

[194] L.W. Jones et al. Charged-Particle Multiplicities of Proton-Proton Interactions Between 90 and 800 GeV. *Phys. Rev. Lett.*, 25:1679–1683, 1970.

[195] P.K. MacKeown and A.W. Wolfendale. The interpretation of the charge ratio of cosmic-ray muons. *P. Phys. Soc.*, 89(3):553, 1966.

[196] C.M.G. Lattes, Y. Fujimoto, and S. Hasegawa. Hadronic interactions of high energy cosmic-ray observed by emulsion chambers. *Phys. Rep.*, 65:151–229, 1980.

[197] T.H. Burnett et al. Average transverse momentum and energy density in high-energy nucleus-nucleus collisions. *Phys. Rev. Lett.*, 57:3249–3252, 1986.

[198] E.L. Feinberg. Multiple production of hadrons at cosmic ray energies (experimental results and theoretical concepts). *Phys. Rep.*, 5:237–350, 1972.

[199] K. Rybicki. Meson production in central collisions of heavy primaries of the cosmic radiation. *Nuovo Cim. Ser. 10*, 28(6):1437–1454, 1963.

[200] Y. Maeda, E. Mikumo, and K. Niu. A Possible decay in flight of a new type particle. *Prog. Theor. Phys.*, 46:1644, 1971.

[201] M. Gell-Mann. A Schematic Model of Baryons and Mesons. *Phys. Lett.*, 8:214–215, 1964.

[202] L.W. Jones. A Review of Quark Search Experiments. *Rev. Mod. Phys.*, 49:717, 1977.

[203] L. Lyons. Current status of quark search experiments. *Prog. Part. Nucl. Phys.*, 7:157–167, 1981.

[204] V.D. Volovic and A.D. Ivanov. On a possibility to search quarks in wide atmospheric showers of cosmic rays. *Izv. Vuz. Fiz.*, 9:121–123, 1971.

[205] J. Bjørneboe and Z. Koba. Space-Time Distribution of Heavy Penetrating Particles (Quarks) in Extensive Air Showers. *Prog. Theor. Phys. Supp.*, 37-38:192–210, 1966.

[206] A. Mincer, H. Freudenreich, J.A. Goodman, et al. Search for heavy long-lived particles in high-energy cosmic rays. *Phys. Rev. D*, 32:541–546, 1985.

[207] M.L. Perl, E.R. Lee, and D. Loomba. A Brief review of the search for isolatable fractional charge elementary particles. *Mod. Phys. Lett. A*, 19:2595–2610, 2004.

[208] M. Aglietta, P. Antonioli, G. Badino, C. Castagnoli, A. Castellina, et al. Search for fractionally charged particles in the Mont Blanc LSD scintillation detector. *Astropart. Phys.*, 2:29–34, 1994.

[209] M. Mori et al. Search for fractionally charged particles in Kamiokande-II. *Phys. Rev. D*, 43:2843–2846, 1991.

[210] M. Ambrosio et al. Final search for lightly ionizing particles with the MACRO detector. *arXiv:0402006 [hep-ex]*, 2004.

[211] G. Damgaard, P. Grieder, K.H. Hansen, C. Iversen, E. Lohse, B. Peters, and T. Rengarajan. A new method for investigating the production of heavy particles in the cosmic radiation. *Phys. Lett.*, 17(2):152–154, 1965.

[212] P.N. Bhat, S.K. Gupta, P.V. Ramana Murthy, B.V. Sreekantan, S.C. Tonwar, et al. Delayed Energetic Events in Extensive Air Showers. *Phys. Rev. D*, 25:2820–2836, 1982.

[213] P.C.M. Yock. Heavy particles in the cosmic radiation? *Phys. Rev. D*, 22:61–64, 1980.

[214] P.C.M. Yock. Further Evidence for Heavy Particles in the Cosmic Radiation. *Phys. Rev. D*, 23:1207, 1981.

[215] C.D. Orth. Alternative Explanations for the Observations of Unknown Heavy-Mass Particles in Cosmic-Ray Showers. In *International Cosmic Ray Conference*, volume 3 of *International Cosmic Ray Conference*, page 2112, 1973.

[216] K. Niu, I. Ohta, and E. Mikumo. Observation of High Energy Interactions at Aeroplane Altitude. In *International Cosmic Ray Conference*, volume 3 of *International Cosmic Ray Conference*, page 1262, 1971.

[217] S. Bianco, F.L. Fabbri, D. Benson, and I. Bigi. A Cicerone for the physics of charm. *Riv. Nuovo Cim.*, 26N7:1–200, 2003.

[218] T. Hayashi, E. Kawai, M. Matsuda, S. Ogawa, and S. Shige-Eda. A possible interpretation of the new event in the cosmic ray experiment. *Prog. Theor. Phys.*, 47:1998–2014, 1972.

[219] H. Sugimoto, Y. Sato, and T. Saito. A New Particle Observed in High Energy Cosmic Ray Interactions. *Prog. Theor. Phys*, 53(5):1541, 1975.

[220] K. Hoshino, S Kuramata, Y. Maeda, E. Mikumo, K. Niu, K. Niwa, S. Tasaka, and N. Ushida. X-Particle Production in 205 GeV/c Proton Interactions. *Prog. Theor. Phys.*, 53(6):1859–1862, 1975.

[221] U. Becker et al. Experimental Observation of a Heavy Particle J. *Phys. Rev. Lett.*, 33:1404–1406, 1974.

[222] J.E. Augustin et al. Discovery of a Narrow Resonance in e^+e^- Annihilation. *Phys. Rev. Lett.*, 33:1406–1408, 1974.

[223] G. Goldhaber et al. Observation in e^+e^- Annihilation of a Narrow State at 1865 MeV/c^2 decaying to $K\pi$ and $K\pi\pi\pi$. *Phys. Rev. Lett.*, 37:255–259, 1976.

[224] D.C. Hom, L.M. Lederman, H.P. Paar, et al. Observation of High-Mass Dilepton Pairs in Hadron Collisions at 400 GeV. *Phys. Rev. Lett.*, 36:1236–1239, 1976.

[225] S.W. Herb, D.C. Hom, L.M. Lederman, et al. Observation of a Dimuon Resonance at 9.5 GeV in 400-GeV Proton-Nucleus Collisions. *Phys. Rev. Lett.*, 39:252–255, 1977.

[226] W.R. Innes, J.A. Appel, B.C. Brown, et al. Observation of Structure in the Υ Region. *Phys. Rev. Lett.*, 39:1240–1242, 1977.

[227] N. Ushida, H. Fuchi, K. Hoshino, S. Kuramata, K. Niu, et al. Observation of a pair decay of short-lived neutral particles produced in 400-GeV/c proton interactions. *Lett. Nuovo Cim.*, 23:577–585, 1978.

[228] K. Niu. Discovery of naked charm particles and lifetime differences among charm species using nuclear emulsion techniques innovated in Japan. *Proc. Jpn. Acad., Ser. B*, 84, 2008.

[229] M. Kaplon, B. Peters, and D.M. Ritson. Emulsion Cloud-Chamber Study of a High Energy Interaction in the Cosmic Radiation. *Phys. Rev.*, 85:900–903, 1952.

[230] K. Nishikawa. A Jet Shower with Energy of $5\cdot 10^{13}$ eV found in Emulsion Chamber. *J. Phys. Soc. Jpn.*, 14:880, 1959.

[231] P.K. Malhotra et al. A study of the production of γ-rays in high-energy nuclear interactions. *Nuovo Cim. A*, 40(2):404–423, 1965.

[232] F. Fumuro, R. Ihara, J. Iwai, T. Ogata, I. Ohta, et al. Some 10-100 TeV Events. *AIP Conference Proceedings*, 49(1):133–144, 1978.

[233] J.P. Albanese, V. Alpe, S. Aoki, R. Arnold, G. Baroni, et al. Direct Observation of the Decay of Beauty Particles Into Charm Particles. *Phys. Lett. B*, 158:186, 1985.

[234] H. Wilczynski. JACEE results on very high energy interactions. *Nucl. Phys. B (Proc. Suppl.)*, 52:81–91, 1997.

[235] J.L. Pinfold. Links between astroparticle physics and the LHC. *J. Phys. G*, 31:R1–R74, 2005.

[236] V.S. Aseikin, G.Y. Goryacheva, S.I. Nikolsky, and V.I. Yakovlev. Long-Flying Component in EAS Cores. In *International Cosmic Ray Conference*, volume 7 of *International Cosmic Ray Conference*, page 2462, 1975.

[237] V.I. Yakovlev. Long flying component produced by protons at energies above 12 TeV in the ionization calorimeter. *Nucl. Phys. B*, 122:417–421, 2003.

[238] V.I. Yakovlev and I.M. Dremin. Charm in cosmic rays (The long-flying component of EAS cores). *Astropart. Phys.*, 26:1–9, 2006.

[239] S.A. Slavatinsky. Results of emulsion chamber experiments with very high-energy cosmic rays which are difficult to explain in the framework of the standard model. *Nucl. Phys. B (Proc. Suppl.)*, 122(0):3–11, 2003.

[240] J.W. Cronin. Review of Direct Lepton Production in Nucleon-Nucleon Collisions. *Erice Subnucl.Phys 1975:0929.* Lecture Presented at Int. School of Subnuclear Physics, Erice, Italy, Jul 11-31, 1975.

[241] L.M. Lederman. Lepton Production in Hadron Collisions. *Lepton-Photon Symp. 1975:265.* Proc. Intern. Symposium on Lepton and photon interactions, Stanford.

[242] C.H. Lewellyn-Smith and I. Hincliffe. Can charm account for prompt lepton production? *Phys. Lett. B*, 61:472, 1976.

[243] C.H. Lewellyn-Smith and I. Hincliffe. Charm as a possible source of prompt leptons. *Nucl. Phys. B*, 114:45–60, 1976.

[244] V.A. Naumov, T.S. Sinegovskaya, S.I. Sinegovsky, N. Takahashi, A. Misaki, and E.V. Bugaev. Atmospheric muon flux at sea level, underground, and underwater. *Phys. Rev. D*, 58:054001.

[245] D. Nosek, P. Travnicek, P. Necesal, and J. Ridky. Prompt muons in extended air showers, 30th ICRC, Mérida, México. *arXiv:0706.2145 [astro-ph]*, 2007.

[246] O. Saavedra and L.V. Volkova. Prompt muon production in cosmic rays. *Astropart. Phys.*, 32:136–139, 2009.

[247] A. Schukraft. A view of prompt atmospheric neutrinos with IceCube. *arxiv:1302.0127 [astro-ph.HE]*, 2013.

[248] R. Laha, J.F. Beacom, B. Dasgupta, S. Horiuchi, and K. Murase. Demystifying the PeV Cascades in IceCube: Less (Energy) is More (Events). *Phys. Rev. D*, 88:043009, 2013.

[249] R.K. Ellis, W.J. Stirling, and B.R. Webber. *QCD and Collider Physics.* Cambridge University Press, 1996.

[250] A. Capella, U. Sukhatme, C.I. Tan, and J. Tran Thanh Van. Dual parton model. *Phys. Rep.*, 236(4–5):225–329, 1994.

[251] N.N. Kalmykov, S.S. Ostapchenko, and A.I. Pavlov. Quark-gluon-string model and EAS simulation problems at ultra-high energies. *Nucl. Phys. B (Proc. Suppl.)*, 52(3):17–28, 1997.

[252] S. Ostapchenko. QGSJET-II: results for extensive air showers. *Nucl. Phys. B (Proc. Suppl.)*, 151:147–150, 2006.

[253] E.J. Ahn et al. Cosmic ray interaction event generator SIBYLL 2.1. *Phys. Rev. D*, 80(9):094003, 2009.

[254] C. Lourenço and H.K. Wöhri. Heavy-flavour hadro-production from fixed-target to collider energies. *Phys. Rep.*, 433(3):127–180, 2006.

[255] M. Cacciari and P. Nason. Is there a significant excess in bottom hadroproduction at the Tevatron? *Phys. Rev. Lett.*, 89:122003, 2002.

[256] R. Averbeck. Heavy-flavor production in heavy-ion collisions and implications for the properties of hot QCD matter. *Prog. Part. Nucl. Phys.*, 70:159–209, 2013.

[257] E.A. Blanco-Covarrubias. *Measurement of the cross section of charmed hadrons and the nuclear dependence α*. PhD thesis, San Luis Potosi University, 2009.

[258] A.D. Frawley, T. Ullrich, and R. Vogt. Heavy flavor in heavy-ion collisions at RHIC and RHIC II. *Phys. Rep.*, 462:125–175, 2008.

[259] E.R. Cazaroto, V.P. Gonçalves, and F.S. Navarra. Heavy quark production at LHC in the color dipole formalism. *Nucl. Phys. A*, 872(1):196–209, 2011.

[260] A. Dainese. Heavy-quark production in heavy-ion collisions. *J. Phys. Conf. Ser.*, 446:012034, 2013.

[261] M. Cacciari, S Frixione, N. Houdeau, M.L. Mangano, P. Nason, et al. Theoretical predictions for charm and bottom production at the LHC. *JHEP*, 1210:137, 2012.

[262] M. Cacciari. Rise and fall of the bottom quark production excess. *arXiv:0407187 [hep-ph]*, 2004.

[263] M.L. Mangano. The Saga of bottom production in $p\bar{p}$ collisions. *AIP Conf. Proc.*, 753:247–260, 2005.

[264] J.A. Appel. Review of heavy quark production at fixed target experiments. *Frascati Phys. Ser.*, 20:155–172, 2001.

[265] M. Cacciari, S. Frixione, M.L. Mangano, P. Nason, and G. Ridolfi. QCD analysis of first b cross section data at 1.96 TeV. *JHEP*, (07):033, 2004.

[266] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjöstrand. Parton fragmentation and string dynamics. *Phys. Rep.*, 97(2-3):31–145, 1983.

[267] V.P. Gonçalves and M.V.T. Machado. Parton Saturation approach in heavy quark production at high energies. *Mod. Phys. Lett. A*, 19(34):2525–2539, 2004.

[268] E. Iancu. Color Glass Condensate and its relation to HERA physics. *Nucl. Phys. B (Proc. Suppl.)*, 191:281–294, 2009.

[269] J. Raufeisen and J. Peng. Relating the parton model and color dipole formulation of heavy quark hadroproduction. *Phys. Rev. D*, 67:054008, 2003.

[270] V.P. Gonçalves and M.V.T. Machado. Saturation physics in ultra high energy cosmic rays: heavy quark production. *JHEP*, (04):028, 2007.

[271] L. McLerran. What is the Color Glass Condensate? *Nucl. Phys. A*, 699:73–81, 2002.

[272] A. Dumitru. Review of Recent Developments in the CGC. *Nucl. Phys. A*, 904-905:51c–58c, 2013.

[273] E. Iancu, K. Itakura, and S. Munier. Saturation and BFKL dynamics in the HERA data at small x. *Phys. Lett. B*, 590:199–208, 2004.

[274] D. Kharzeev and E. Levin. Manifestations of high density QCD in the first RHIC data. *Phys. Lett. B*, 523:79–87, 2001.

[275] T. Hirano and Y. Nara. Hydrodynamic afterburner for the color glass condensate and the parton energy loss. *Nucl. Phys. A*, 743:305–328, 2004.

[276] I.G. Bearden et al. Pseudorapidity distributions of charged particles from Au+Au collisions at the maximum RHIC energy. *Phys. Rev. Lett.*, 88:202301, 2002.

[277] J. Jalilian-Marian and Y.V. Kovchegov. Saturation physics and deuteron-gold collisions at RHIC. *Prog. Part. Nucl. Phys.*, 56:104–231, 2006.

[278] J. Qiu and G. Sterman. Power corrections in hadronic scattering (II). Factorization. *Nucl. Phys. B*, 353(1):137–164, 1991.

[279] S. Brodsky, G. de Teramond, and M. Karliner. Puzzles in Hadronic Physics and Novel Quantum Chromodynamics Phenomenology. *Ann. Rev. Nucl. Part. Sci.*, 62:1–35, 2012.

[280] W. Chang and J. Peng. Extraction of Various Five-Quark Components of the Nucleons. *Phys. Lett. B*, 704:197–200, 2011.

[281] E. Norrbin and R. Vogt. Bottom Production Asymmetries at the LHC. *Report number: LU TP 00-10, LBNL-45275.*

[282] R. Vogt. Charm Production in Hadronic Collisions. *Nucl. Phys. A*, 553:791–798, 1993.

[283] R. Vogt, S.J. Brodsky, and P. Hoyer. Systematics of charm production in hadronic collisions. *Nucl. Phys. B*, 383:643–684, 1992.

[284] F.S. Navarra, M. Nielsen, C.A.A. Nunes, and M. Teixeira. Intrinsic charm component of the nucleon. *Phys. Rev. D*, 54:842–846, 1996.

[285] J.C. Anjosa, J. Magnin, and G. Herrera. On the intrinsic charm and the recombination mechanism in charm hadron production. *Phys. Lett. B*, (523):29–34, 2001.

[286] N. Sakai, P. Hoyer, C. Peterson, and S.J. Brodsky. The intrinsic charm of the proton. *Phys. Lett. B*, 93:451–455, 1980.

[287] R. Vogt and S.J. Brodsky. Charmed hadron asymmetries in the intrinsic charm coalescence model. *Nucl. Phys. B*, 478:311–332, 1996.

[288] S.J. Brodsky. Novel QCD Phenomenology at the LHeC. *arXiv:1106.5820 [hep-ph]*, 2011.

[289] V.M. Abazov et al. Measurement of $\gamma + b + X$ and $\gamma + c + X$ production cross sections in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV. *Phys. Rev. Lett.*, 102:192002, 2009.

[290] S.J. Brodsky and S. Gardner. Evading the CKM hierarchy: Intrinsic charm in B decays. *Phys. Rev. D*, 65:054016, 2002.

[291] B.W. Harris, J. Smith, and R. Vogt. Reanalysis of the EMC charm production data with extrinsic and intrinsic charm at NLO. *Nucl. Phys. B*, 461:181–196, 1996.

[292] F.M. Steffens, W. Melnitchouk, and A. W. Thomas. Charm in the nucleon. *Eur. Phys. J. C*, 11:673–683, 1999.

[293] A. Ocherashvili et al. Confirmation of the double charm baryon $\Xi_{cc}^{+}(3520)$ via its decay to $pD^{+}K^{-}$. *Phys. Lett. B*, 628:18–24, 2005.

[294] J.R. Ellis, Y. Frishman, A. Hanany, and M. Karliner. Quark solitons as constituents of hadrons. *Nucl. Phys. B*, 382:189–212, 1992.

[295] S.J. Brodsky and M. Karliner. Intrinsic charm of vector mesons: A Possible solution of the '$\rho\pi$ puzzle'. *Phys. Rev. Lett.*, 78:4682–4685, 1997.

[296] A. Blanco-Covarrubias et al. Nuclear Dependence of Charm Production. *Eur. Phys. J. C*, 64:637–644, 2009.

[297] J. Pumplin. Light-cone models for intrinsic charm and bottom. *Phys. Rev. D*, 73:114015, 2006.

[298] J. Pumplin, H.L. Lai, and W.K. Tung. The Charm Parton Content of the Nucleon. *Phys. Rev. D*, 75:054029, 2007.

[299] P.M. Nadolsky, H. Lai, Q. Cao, J Huston, J. Pumplin, et al. Implications of CTEQ global analysis for collider observables. *Phys. Rev. D*, 78:013004, 2008.

[300] J. Peng and W. Chang. Nucleon sea and the five-quark components. *PoS*, QNP2012:012, 2012.

[301] J. Alwall. Quark asymmetries and intrinsic charm in nucleons. *arXiv:0508126 [hep-ph]*, 2005.

[302] P. Desiati and T.K. Gaisser. Seasonal Variation of Atmospheric Leptons as a Probe of Charm. *Phys. Rev. Lett.*, 105:121102, 2010.

[303] T.K. Gaisser. Atmospheric leptons, the search for a prompt component. *arXiv:1303.1431 [hep-ph]*, 2013.

[304] C.G.S Costa. The prompt lepton cookbook. *Astropart. Phys.*, 16:193–204, 2001.

[305] B. Abelev et al. Measurement of charm production at central rapidity in proton-proton collisions at $\sqrt{s} = 2.76$ TeV. *JHEP*, 1207:191, 2012.

[306] http://home.thep.lu.se/ torbjorn/Pythia.html.

[307] R. Barceló, M. Jenkins, J.I. Illana, and M. Masip. Propagation in the atmosphere of ultrahigh-energy charmed hadrons. *arXiv:1010.0894v1 [hep-ph]*, 2010.

[308] X. Wang and M. Gyulassy. Systematic study of particle production in $p+p$ ($\overline{p}$) collisions via the HIJING model. *Phys. Rev. D*, 45:844–856, 1992.

[309] X. Wang and M. Gyulassy. HIJING: A Monte–Carlo model for multiple jet production in pp, pA, and AA collisions. *Phys. Rev. D*, 44:3501–3516, 1991.

[310] K. Werner. Multiparticle production in p-p, p-Air, and p-Xe collisions at 200 GeV by the multistring model Venus. *Phys. Rev. D*, 39:780–791, 1989.

[311] K. Werner. Strings, pomerons and the VENUS model of hadronic interactions at ultra-relativistic energies. *Phys. Rep.*, 232(25):87–299, 1993.

[312] P. Lipari R.S. Fletcher, T.K. Gaisser and T. Stanev. SIBYLL:An event generator for simulation of high energy cosmic ray cascades. *Phys. Rev. D*, 50(9):5710–5726, 1994.

[313] L. Ding and E. Stenlund. A Monte–Carlo program for nuclear collision geometry. *Comp. Phys. Comm.*, 59(2):313 – 318, 1990.

[314] J. Beringer et al. 2012 Review of Particle Physics. *Phys. Rev. D*, 86(010001), 2012.

[315] D. Heck. *Report FZK 6019*. http://www-ik.fzk.de/physics-description/corsika-phys.html, 1998.

[316] S. Argirò, S.L.C. Barroso, J. González, L. Nellen, et al. The offline software framework of the Pierre Auger Observatory. *Nucl. Instrum. Meth. A*, 580(3):1485–1496, 2007.

[317] A. Fassò, A. Ferrari, S. Roesler, et al. The physics models of FLUKA: status and recent development. *Computing in High Energy and Nuclear Physics (CHEP03), La Jolla, Ca, USA, 2003*.

[318] A. Fassò, A. Ferrari, J. Ranft, and P.R. Sala. *Report CERN-2005-10 (2005)*.

[319] D. Heck and J. Knapp. Report FZKA 6097. Technical report, Forschungszentrum Karlsruhe, 1998. http://www-ik.fzk.de/ heck/publications/.

[320] M. Kobal. A thinning method using weight limitation for air-shower simulations. *Astropart. Phys.*, 15:259–273, 2001.

[321] W.R. Nelson, H. Hirayama, and D.W.O. Rogers. The Egs4 Code System - SLAC report 265. 1985.

[322] A.M. Hillas. Shower simulation: Lessons from MOCCA. *Nucl. Phys. B (Proc. Suppl.)*, 52:29–42, 1997.

[323] http://www ik.fzk.de/corsika/usersguide/corsika_tech.html.

[324] I.C. Mariş, F. Schussler, R. Ulrich, and M. Unger. Data Summary Trees and Shower Visualization for Reconstructed Auger Events (Internal note). GAP 2006-081, 2006.

[325] A. Yushkov. Measurements of the muon shower content at the Pierre Auger Observatory. *EPJ Web Conf.*, 53:07002, 2013.

[326] H. Yang, B.P. Roe, and J. Zhu. Studies of boosted decision trees for MiniBooNE particle identification. *Nucl. Instrum. Meth. A*, 555(1-2):370–385, 2005.

[327] V.V. Gligorov. Performance and upgrade plans of the LHCb trigger system. *Nucl. Instrum. Meth. A*, 718, 2012.

[328] Y. Becherini, A. Djannati-Atai, V. Marandon, M. Punch, and S. Pita. A new analysis strategy for detection of faint gamma-ray sources with Imaging Atmospheric Cherenkov Telescopes. *Astropart. Phys.*, 34:858–870, 2011.

[329] A. Fiasson, F. Dubois, G. Lamanna, J. Masbou, and S. Rosier-Lees. Optimization of multivariate analysis for IACT stereoscopic systems. *Astropart. Phys.*, 34:25–32, 2010.

[330] A.R. Webb and K.D. Copsey. *Statistical Pattern Recognition*. Hoboken/Wiley, 2001.

[331] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss. TMVA: Toolkit for Multivariate Data Analysis. *PoS*, ACAT:040, 2007.

[332] T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer New York, 2009.

[333] G. Cowan. Topics in statistical data analysis for high-energy physics. *arXiv:1012.3589 [physics.data-an]*, 2010.

[334] Y. Liu and I. Stancu. Cascade training technique for particle identification. *Nucl. Instrum. Meth. A*, 578(1):315–321, 2007.

[335] S. Andringa, R. Conceição, F. Diogo, and M. Pimenta. Sensitivity to primary composition and hadronic models from average shape of high energy cosmic ray shower profiles. *arXiv:1209.6011 [hep-ph]*, 2012.

[336] J.L. Navarro Quirante. *Search for Ultra-High Energy Neutrinos at the Pierre Auger Observatory*. PhD thesis, Dpto. Física Teórica y del Cosmos, Universidad de Granada, 2012.

[337] M. Settimo. *Hybrid detection of Ultra High Energy Cosmic Rays with the Pierre Auger Observatory*. PhD thesis, Facolt di Scienze Matematiche, Fisiche e Naturali. Universit del Salento, 2010.

[338] E.J. Ahn. Composition fit to hybrid Xmax data, part II (Internal note). GAP 2012-028.

[339] V. Scherini. Search for ultra-high energy photons at the Pierre Auger Observatory. *EPJ Web Conf.*, 53:05002, 2013.

[340] P. Abreu et al. Measurement of the Cosmic Ray Energy Spectrum Using Hybrid Events of the Pierre Auger Observatory. *Eur. Phys. J. Plus*, 127:87, 2012.

[341] P. Abreu et al. The exposure of the hybrid detector of the Pierre Auger Observatory. *Astropart. Phys.*, 34:368–381, 2011.

[342] J. Neyman. *A Selection of Early Statistical Papers of J. Neyman*. The Selected Papers of Jerzy Neyman and E. S. Pearson. University of California Press, 1967.

[343] G.J. Feldman and R.D. Cousins. Unified approach to the classical statistical analysis of small signals. *Phys. Rev. D*, 57:3873–3889, 1998.

[344] W.A. Rolke, Á.M. López, and J. Conrad. Limits and confidence intervals in the presence of nuisance parameters. *Nucl. Instrum. Meth. A*, 551(2–3):493–503, 2005.

[345] J. Aublin, D. Allard, I. L'Henry-Yvon, and E. Parizot. The lateral trigger probability (LTP) and the lateral distribution of over threshold (LADIBOT) as powerful tools to test MC simulations (Internal note). GAP 2007-013, 2007.

[346] Proc. of the 29th ICRC 2005. *Aperture calculation of the Pierre Auger Observatory Surface Detector*, 2005.

[347] P. Facal San Luis, V.M. Olmos-Gilbaja, G. Parente, G. Rodríguez-Fernández, I. Valiño, R.A. Vázquez, E. Zas, and L. Cazón. Cosmic ray spectrum with inclined showers: November 2006 update (Internal note). GAP 2007-027, 2007.

[348] I. Allekote, H. Asorey, and M. Gómez Berisso. Improving the determination of the Auger Surface Detector trigger probability from real data (Internal note). GAP 2009-019, 2007.

[349] A. Parra, V.M. Olmos-Gilbaja, J. Álvarez-Muñiz, R. Pelayo, R.A. Vázquez, and E. Zas. The trigger efficiency of the surface array of the Pierre Auger Observatory for inclined showers (Internal note). GAP 2009-135, 2009.

[350] H. Dembinski, J. González, M. Roth, V. Olmos-Giljaba, T. Schmidt, and I. Valiño. Reconstruction of horizontal events within offline (Internal note). GAP 2009-135, 2009.

[351] Proc. of the 31th ICRC 2009. *The electromagnetic component of inclined air showers at the Pierre Auger Observatory*, 2009.

[352] H. Dembinski, T. Hebbeker, and M. Leuthold. A comparison of Monte–Carlo generated muon maps with near horizontal SD showers (Internal note). GAP 2007-124, 2007.

[353] H. Dembinski. Update of the SD-FD energy calibration at $\theta > 60^{\circ}$. Pierre Auger Collaboration meeting, 2010.

[354] S.J. Sciutto. AIRES: A system for air shower simulations (Version 2.2.0). *arXiv:astro-ph/9911331v1*.

[355] S. Frixione, M.L. Mangano, P. Nason, and G. Ridolfi. Charm and bottom production: Theoretical results versus experimental data. *Nucl. Phys. B*, 431:453–483, 1994.