Core-corona hadronization model and its impact on muon content of extensive air showers

Zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

von der KIT-Fakultät für Physik des
Karlsruher Instituts für Technologie (KIT)
und der
Universidad Nacional de San Martín (UNSAM)

angenommene

Dissertation

von

Lic. Matias Perlin

aus Buenos Aires

Tag der mündlichen Prüfung: 19. February 2021

Referent: Dr. F. A. Sánchez

Korreferent: Prof. Dr. J. Blümer

Betreuer: Dr. T. Pierog

Core-corona hadronization model and its impact on muon content of extensive air showers

Tesis presentada para optar por el título de

Doctor en Astrofísica

del Instituto de Tecnología "Prof. Jorge A. Sábato" de la
Universidad Nacional de San Martín (UNSAM)
y del
Karlsruher Institut für Technologie (KIT)

por

Lic. Matias Perlin

Fecha de la defensa oral: 19 de Febrero de 2021

Director: Dr. F. A. Sánchez

Co-director: Prof. Dr. J. Blümer

Colaborador: Dr. T. Pierog

Abstract

Cosmic rays were discovered at the beginning of the 20th century and their study has contributed extensively to science areas such as Astrophysics and Particle Physics. However, there are still big open questions. The sources and the acceleration mechanisms of ultra-high energy cosmic rays yet remain elusive. To improve both descriptions, an accurate measurement of mass composition is required. Furthermore, mass composition is closely related to the so-called *muon puzzle*, which states that the current understanding of hadronic interactions is not enough to explain the results of several cosmic-ray experiments. The main problem is that the muon content in extensive air shower simulations is significantly less than the one measured. Several exotic mechanisms have been proposed but improvements in measurements at colliders make them unlikely to be viable processes. In hadron collisions the fraction of energy going into electromagnetic particles has a large impact on the number of muons produced in air shower cascades. Therefore, any proposed model needs to find a way to transfer energy from the electromagnetic cascade to the hadronic one to increase the muon production.

We developed a new technique to quickly test on air showers the effect of any kind of modification of the hadronic interaction properties. In this novel approach the secondary particle spectra in the CONEX framework are modified. There are many different ways of implementing these modifications, each representing different interesting scenarios.

As an application, we chose to test the core-corona model of heavy ion interactions on air shower development. The basic assumption of the core-corona model is a mixture of different underlying particle production mechanisms such as a collective statistical hadronization in large density regions (core) in addition to the expected string fragmentation in low density regions (corona). Recent measurements at the LHC confirm features that can be linked to this model.

Since the two mechanisms present in the core-corona model imply different electromagnetic energy fractions, each mechanism impacts the muon production in extensive air showers in a different way. Our tool allows, for the first time, to compare the energy evolution of muon-based observables accessible to experiments and to find the most adequate parameters for the core-corona model.

The simplified core-corona model based results were compatible with a significant increase of the number of muons in simulations. Hence, we have included direct muon measurements into our analysis and compared them with full end-to-end simulations in the context of the Pierre Auger Observatory.

As part of the ongoing upgrade of the Observatory, dubbed AugerPrime, the Underground Muon Detector (UMD) offers a unique and straight-forward opportunity to directly

measure high-energy muons of extensive air showers. Thus, the first two and a half years of data acquired with the array of the UMD have been reconstructed and analyzed. In particular, we have built data driven average muon lateral distributions (MLDF) in bins of energy and zenith angle and we have compared them with full detector simulations in a core-corona scenario.

Thanks to a new technique providing fast and realistic air shower simulations with modified hadronic interaction properties, we showed that a model partially based on hadronization properties of a QGP applied gradually from low energy to the highest energy significantly reduces the muon deficit. Further detailed investigation is needed to determine if it can be integrated in a future hadronic interaction model.

Zusammenfassung

Die kosmische Strahlung wurde zu Beginn des 20. Jahrhunderts entdeckt und ihre Erforschung hat zu erheblichen Fortschritten in wissenschaftlichen Bereichen wie Astrophysik und Teilchenphysik beigetragen. Es gibt jedoch noch tiefgreifende offene Fragen. Die Quellen und Beschleunigungsmechanismen der kosmischen Strahlung ultrahoher Energie sind noch nicht vollkommen erfasst. Um beide Phänomene besser su verstehen, ist eine genaue Messung der Massenzusammensetzung der kosmischen Strahlung erforderlich. Darüber hinaus ist die Massenzusammensetzung eng mit dem sogenannten Myonen-Puzzle verbunden, welches besagt, dass das aktuelle Verständnis von hadronischen Wechselwirkungen nicht ausreicht, um die Ergebnisse mehrerer Experimente, die kosmische Strahlung messen, zu erklären. Das Hauptproblem besteht darin, dass der Myonengehalt in Luftschauersimulationen deutlich geringer ist als der gemessene. Es wurden mehrere exotische Mechanismen vorgeschlagen, aber Fortschritte in den Messungen in Teilchenbeschleunigern ergeben, dass diese unwahrscheinlich sind. Der Energieanteil, der in Hadronenkollisionen in elektromagnetische Partikel gelangt, hat einen erheblichen Einfluss auf die Anzahl der in Luftschauern produzierten Myonen. Deswegen muss jedes neue Modell, die Möglichkeit einer Energieübertragung von der elektromagnetischen Komponente zur hadronischen aufweisen.

Wir haben ein neues Verfahren entwickelt, welches ermöglicht, effizient den Effekt jeder Art von Modifikation der hadronischen Wechselwirkungseigenschaften an Luftschauern zu testen. Mit dieser neuen Methode werden die Sekundärteilchenspektren im Rahmen des CONEX-Programms modifiziert. Es gibt viele verschiedene Möglichkeiten, diese Änderungen umzusetzen, und jede repräsentiert ein anderes interessantes Szenarium.

Als Anwendung haben wir den Einfluss des Core-Corona-Modells für Wechselwirkungen schwerer Ionen auf die Entwicklung von Luftschauern analysiert. Die Basis des Core-Corona-Modells ist eine Mischung zweier verschiedener Partikelproduktionsmechanismen: kollektive statistische Hadronisierung in Regionen großer Dichte (Core) und Stringfragmentierung in Bereichen geringer Dichte (Corona). Neueste Messungen am LHC bestätigen, dass bestimmte Beobachtungen mit diesem Modell verknüpft werden können.

Da beide im Core-Corona-Modell anwesenden Mechanismen unterschiedliche elektromagnetische Energiefraktionen bedeuten, wirkt sich jeder Mechanismus auf unterschiedliche Weise auf die Myonenproduktion in Luftschauern aus. Mit unserem Modell, ist es zum ersten Mal möglich, die Energieabhängigkeit von auf Myonen basierenden messbaren Werten, die in Experimenten zugänglich sind, zu analysieren und die geeignetsten Parameter für das Core-Corona-Modell zu finden.

Die auf dem vereinfachten Core-Corona-Modell basierenden Ergebnisse sind mit einer erheblichen Erhöhung der Myonenanzahl in Simulationen kompatibel. Aus diesem Grund

haben wir direkte Myonenmessungen in unsere Analyse aufgenommen und diese mit vollständigen Simulationen im Kontext des Pierre Auger Observatoriums verglichen.

Im Rahmen des gängigen Upgrades des Observatoriums, AugerPrime genannt, bietet der sogenannte Underground Muon Detector (UMD) eine einzigartige Gelegenheit hochenergetische Myonen von Luftschauern direkt zu messen. Dafür wurden die in den ersten zweieinhalb Jahren mit dem UMD-Array erfassten Daten rekonstruiert und analysiert. Insbesondere haben wir auf Daten basierende durchschnittliche laterale Myonenverteilungen (MLDF) in bins von Energie und Zenitwinkel erstellt und diese mit vollständigen Detektorsimulationen in einem Core-Corona-Szenario verglichen.

Dank einer neuen Methode, die schnelle und realistische Luftschauersimulationen mit modifizierten hadronischen Wechselwirkungseigenschaften bietet, haben wir gezeigt, dass ein Modell, welches teilweise auf den Hadronisierungseigenschaften eines Quark-Gluon-Plasmas basiert und für niedrige bis zu den höchsten Energien angewendet wird, das Myonendefizit erheblich reduziert. Weitere detaillierte Nachforschungen sind erforderlich, um festzustellen, ob dieses Modell in einem zukünftigen Modell hadronischer Interaktionen integriert werden kann.

Resumen

Los rayos cósmicos fueron descubiertos a principios del siglo XX y su estudio ha contribuido a significativos avances en campos científicos como la astrofísica y la física de partículas. Sin embargo, todavía quedan importantes preguntas abiertas. Las fuentes y los mecanismos de aceleración de los rayos cósmicos de ultra alta energía aún no se comprenden completamente. Para entender mejor ambos fenómenos, es necesaria una medición exacta de la composición de masa de los rayos cósmicos. Además, la composición de masa está estrechamente relacionada con el llamado *rompecabezas de muones*, que establece que la comprensión actual de las interacciones hadrónicas es insuficiente para explicar los resultados de diversos experimentos de rayos cósmicos. El principal problema es que el contenido de muones en las simulaciones de lluvias extensas de rayos cósmicos es significativamente menor al medido. Se han propuesto varios mecanismos exóticos, pero los avances en las mediciones en aceleradores de partículas sugieren que estos son poco probables. La fracción de energía destinada a las partículas electromagnéticas en las colisiones de hadrones tiene una influencia significativa en la cantidad de muones producidos en las lluvias extensas de rayos cósmicos. Por lo tanto, todo modelo propuesto debe implementar una transferencia de energía desde el componente electromagnético al componente hadrónico para incrementar la producción de muones.

Hemos desarrollado un nuevo método, que permite probar de manera eficiente el efecto de cualquier tipo de modificación de las propiedades de interacción hadrónica en lluvias extensas. En este enfoque, se modifican los espectros de partículas secundarias utilizados por el programa de simulacion CONEX. Hay muchas formas diferentes de implementar los cambios y cada una representa un escenario interesante.

Como aplicación, analizamos el impacto de utilizar el modelo core-corona de las interacciones de iones pesados en el desarrollo de las lluvias extensas. El modelo core-corona se basa en la combinación de dos mecanismos de producción de partículas diferentes: hadronización estadística en regiones de alta densidad (core) y fragmentación de cuerdas en áreas de baja densidad (corona). Recientes mediciones del LHC confirman que ciertas observaciones pueden entenderse con este modelo.

Dado que ambos mecanismos presentes en el modelo core-corona implican diferentes fracciones de energía electromagnética, cada mecanismo afecta la producción de muones en las lluvias extensas de distinta manera. Con nuestro método, es posible por primera vez, comparar la evolución en energía de observables basados en muones accesibles en distintos experimentos y encontrar los parámetros más adecuados para el modelo core-corona.

Los resultados basados en un modelo core-corona simplificado mostraron ser compatibles con un aumento significativo en el número de muones en las simulaciones. Por esta

razón, incluimos mediciones directas de muones en nuestro análisis y las comparamos con simulaciones completas en el contexto del Observatorio Pierre Auger.

Como parte de la actualización del Observatorio, denominada AugerPrime, el Detector de Muones Subterráneo (UMD) ofrece una oportunidad única para medir directamente los muones de alta energía de las lluvias extensas. Para ello, se reconstruyeron y analizaron los datos adquiridos con el arreglo UMD en los primeros dos años y medio. En particular, se estudiaron las distribuciones laterales medias de muones (MLDF) en bines de energía y ángulo cenital y las comparamos con simulaciones detalladas de los detectores en un escenario de core-corona.

Gracias al nuevo método que ofrece simulaciones de lluvias extensas rápidas y realistas con propiedades de interacción hadrónica modificadas, hemos demostrado que un modelo parcialmente basado en las propiedades de hadronización del plasma de quark-gluón (QGP) y aplicado gradualmente desde energías bajas a más altas redujo considerablemente el déficit de muones. Sería necesaria una investigación más detallada para determinar si este modelo se puede incorporar a un futuro modelo de interacciones hadrónicas.

Contents

Acknowledgments

This Ph.D. would not have been possible without the help and support from many people over the last years.

In first place, I would like to thank Professors Alberto Etchegoyen and Johannes Blümer for giving me the opportunity to take part in the Double Doctorate Degree in Astrophysics and for laying the foundations of this joint German and Argentine program.

To my advisor from the Argentine side, Federico Sánchez, for his guidance from my master thesis. To my advisor from the German side, Tanguy Pierog, for his continuous support and great help in a research field new for me.

To all my colleagues at ITeDA and KIT, with whom I spent very good moments. I want to specially thank to Ana Botti, Belén Andrada, Isabel Goos and Nicolas Gonzalez, whom I consider great friends and were very helpful at different times for different reasons. To Max Reininghaus, Alex Streich, Alvaro Taboada for making my stay in Germany very pleasant. To Diego Melo, Diego Ravignani, Markus Roth, Ralf Ulrich, Brian Wundheiler for the helpful discussions. I also have to mention Christian, Ronco, all the engineers, Mariela, Adrian, Ana Laura...

To Marie-Christine Mundt my sincerest gratitude for taking care of any issues I had, even the smallest one. Also to Sabine Bucher, who also help me in my stay in Germany and for taking care of submitting this thesis. I would also like to mention Isabel and Belén again who greatly contributed with corrections to this writing. To the Pierre Auger Collaboration, in which fortunately I found warm and interesting people on a personal and professional level.

I owe the deepest gratitude to my family and friends, who care about me and help me in countless ways. To my life partner, Jesica, for sharing our days since the first times of the university and taking care of each other, for supporting me when I decided to spend a year abroad, for understand or at least trying to when I spent long nights and weekends working. To my mother and brother, Mercedes and Exequiel, whose unconditional support and love allowed me to get here. And to my grandma Estela, whose memories accompany me every day.

CHAPTER 1

Cosmic Rays and their detection

Cosmic rays (CRs) are particles, mainly ionized nuclei, that arrive at Earth from the outer space. These particles cover a wide range of energies from less than a GeV up to energies around 10^{20} eV. Their sources may be either galactic or extra-galactic depending on their energy. When CRs enter the atmosphere, they may interact with the air producing a shower of secondary particles which propagate towards the ground.

At the beginning of the 20th century, the ionization of atmospheric molecules was believed to be caused only by radioactive elements of the Earth. Numerous experiments were carried out to discover the origin of this radiation. In 1909, Theodor Wulf measured on the top of the Eiffel Tower more radiation than at its base. Two years later, Domenico Pacini also pointed to the atmosphere as the source of this radiation. In 1912 Victor Hess performed a series of measurements on a hot-air balloon and observed that the ionizing radiation increased in intensity as a function of the balloon height [1]. The radiation level at 5000 m from the surface was almost 4 times greater than that detected on Earth. Later observations also showed that this effect was not due to solar radiation since it remains the same during solar eclipses.

Robert Millikan introduced in 1925 the concept of *Cosmic Rays*. He believed that they were γ radiation [2], which was the most penetrating particle type known at the time. This implied cosmic rays could not interact with the geomagnetic field. However, Arthur Compton in 1932 [3] showed that the intensity of cosmic radiation depended on geomagnetic latitude. Further analysis confirmed that cosmic rays are predominantly composed by charged particles, and the study of secondary particles allowed the discovery of new particles such as the muon [4].

The discovery of extensive air showers (EAS) is commonly attributed to Pierre Auger. In 1938, he initiated the study of extended atmospheric cascades [5] and observed events in coincidence using surface detectors separated by distances up to 75 m. The temporal correlation indicated the same physical event which was suggested to be a high-energy cosmic ray that interacted with the particles of the atmosphere. These measurements allowed to extend the estimated range of cosmic-ray energies up to 10^{15} eV. In the 1950s, larger and more complex arrays of detectors were built. The first measurements of cosmic rays energy above 10^{15} eV were achieved in 1954 by the Cosmic-Ray Group of the University of Massachusetts lead by Bruno Rossi. A decade later, the energy spectrum of the primary cosmic rays was extended beyond 10^{20} eV [6].

After the discovery of the cosmic microwave background (CMB) radiation in 1964 [7], Greisen, Zatsepin and Kuzmin [8,9] formulated the GZK cut-off in the energy spectrum of protons. Protons with energies above 5×10^{19} eV interact with CMB photons and in this way, they loss continuously their energy until it falls below the energy threshold of the interaction. This produces a decrease in the proton flux and consequently a spectrum suppression at distances larger than 100 Mpc. However, it does not set a constraint on the energy that protons may have at their source.

The inconclusive experimental situation about the existence of the GZK cut-off [6,10–13] was a main reason for the design of new ultra-high-energy cosmic ray (UHECR) experiments such as the Pierre Auger Observatory [14] and Telescope Array [15]. Nowadays, the spectrum suppression has been established but remains a subject of study because the acceleration mechanisms and propagation of cosmic rays have not been fully explained. Moreover, the mass composition of cosmic rays at different energies is still unclear.

1.1 Energy spectrum

The energy spectrum of the cosmic rays observed at Earth covers more than 12 orders of magnitude in energy, from 10^9 eV (mainly of solar origin) to more than 10^{20} eV (of extra galactic origin) and more than 30 orders of magnitude in flux. The flux follows a power-law form like $E^{-\gamma}$ as Fig. 1.1 shows, where the spectral index is $\gamma \approx 3$ above $E = 10^{14}$ eV. It means that the number of cosmic rays decreases by a factor $\sim 10^3$ for 1each decade increase in energy. For example, the flux is around one particle per m^2 per year at 10^{15} eV and around one particle per km^2 per century at 10^{20} eV.

The spectrum at high energies shown in Fig. 1.1 is characterized by some inflection points. The *knee* around $10^{15.5}$ eV, where the flux of light mass primaries decreases changing the composition from light to heavy elements [16]. The *second knee* around 10^{17} eV, which is assumed to be related with the rapid fall-off of the galactic heavy primaries [17]. The *ankle* around 5×10^{18} eV , it is explained by the transition from galactic to extra-galactic sources [18]. And the spectrum suppression around $10^{19.5}$ eV. However, the explanation of these features in not closed since it is strongly dependent on the astrophysical model used to describe the mass composition and sources of cosmic rays [19].

The abrupt suppression of the flux around $10^{19.5}$ eV can be explained in terms of the GZK cut-off as mentioned above. In the case of protons, photo-pion production through the Δ resonance decreases its energy by mean of these processes:

$$\begin{aligned} p + \gamma_{CMB} &\rightarrow \Delta^+ \rightarrow n + \pi^+ \\ p + \gamma_{CMB} &\rightarrow \Delta^+ \rightarrow p + \pi^0. \end{aligned} \tag{1.1}$$

Fig. 1.2 shows how protons with an average energy up to 10^{22} eV cannot have energies greater than 10^{20} eV after travel $\sim$ 100 Mpc due to these interactions. In consequence, the flux of extragalactic cosmic rays above this energy is suppressed producing the well-known GZK cut-off feature in the cosmic rays spectrum.

The energy of heavy nuclei is reduced by an analogous mechanism, where the main processes are the one-nucleon and two-nucleon photodisintegration. In the case of gammas, the pair production via $\gamma + \gamma_{CMB} \rightarrow e^+ + e^-$ can happen at gamma energies from 10^{14} eV up to 10^{19} eV.

The energy where the spectrum is suppressed was calculated by the Pierre Auger Observatory as the energy at which the integral spectrum drops by a factor of two below

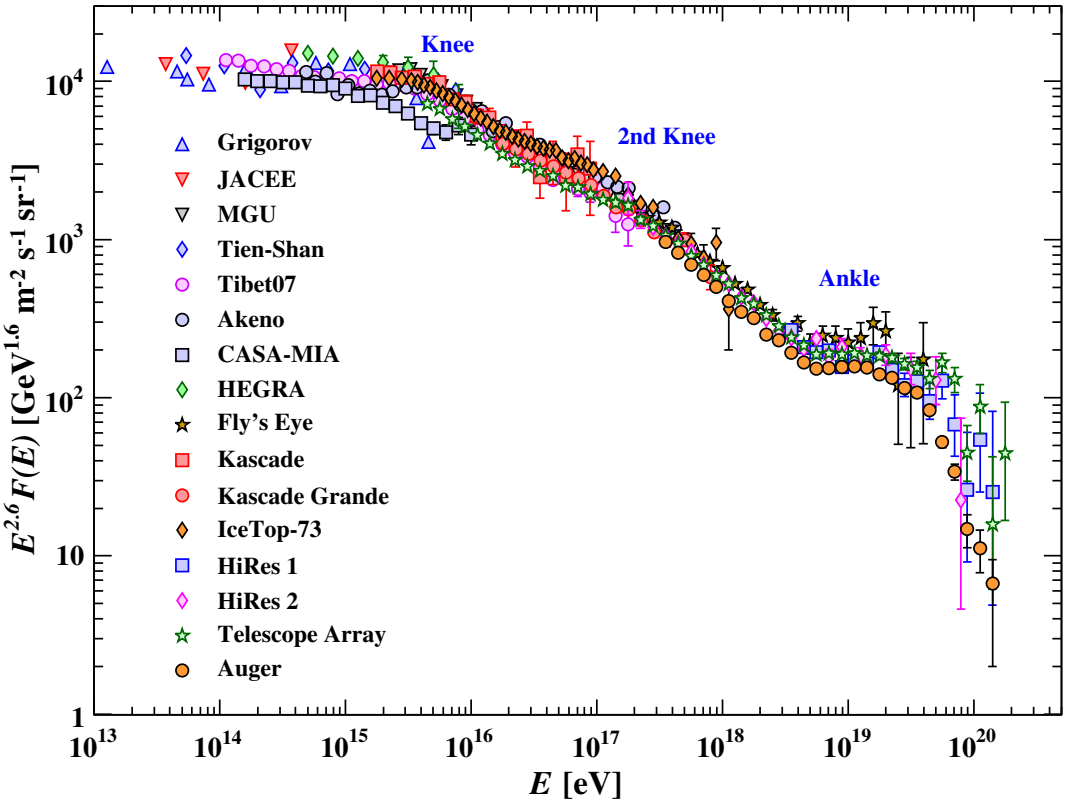


Figure 1.1: Energy spectrum of high-energy cosmic rays measured by several experiments [20].

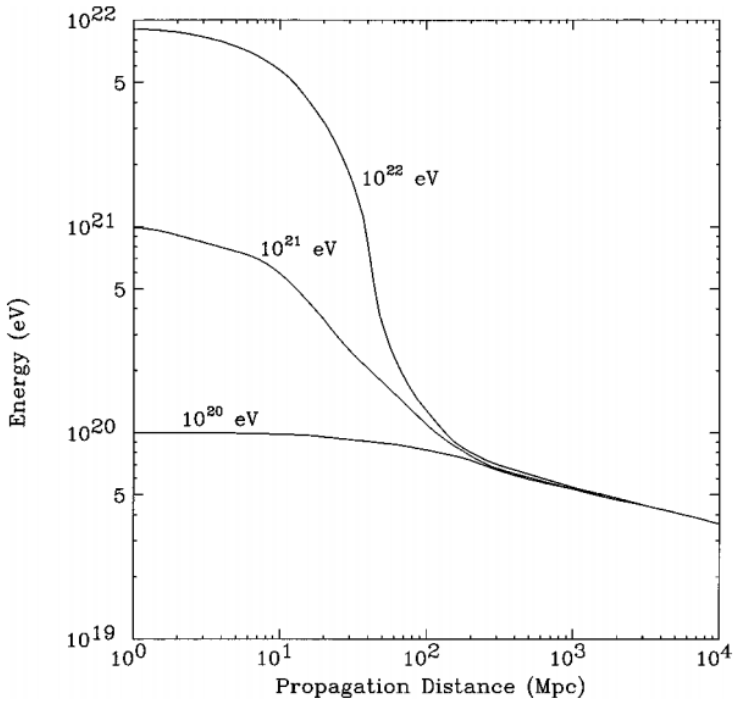


Figure 1.2: Average proton energy as a function of the distance traveled through the CMB for different initial energy [21].

what would be expected with no cutoff, and its value is $E_{1/2}$ = (2.26 $\pm$ 0.08 (stat.) $\pm$ 0.4 (syst.)) $\times 10^{19}$ eV [22]. This energy is considerably close to the prediction by the GZK model $E_{gzk} = 5.3 \times 10^{19}$ eV. However, this prediction depends on the cosmic ray mass composition. On the other hand, there has been some discrepancies between Telescope Array and Auger Observatory about the spectrum suppression [23] but both collaboration are working even together to understand them [24].

1.2 Extensive Air Showers

The cosmic rays flux for energies from 10^{14} eV becomes so low that the only way to measure them is indirectly. When a high energy cosmic ray reaches the Earth, it interacts with an atom in the atmosphere producing a large cascade of secondary particles which emit Cherenkov and fluorescence light. These cascade of particles are called *extensive air showers* (EAS) or simply *air showers*. The shower consists of a core of high-energy hadrons that continuously feeds the electromagnetic part of the shower, mainly by photons coming from neutral pion decays. Lower-energy charged pions and kaons decay to feed the muonic component. The electromagnetic component also produces low energy muons by photoproduction or muon pair-production. The observation of the longitudinal development of the air shower and the particles that reach the ground allows to infer properties of the primary cosmic rays.

A simple but effective model to describe the main features of the air shower development was presented by Heitler in 1950s [25]. The Heitler model was originally developed for electromagnetic particles as primary cosmic ray. However, it can be extended to air showers induced by hadron particles [26]. This model assumes that photons create electron-positron pair and these particles in turn are affected by radiative losses producing photons and ionization losses. These processes take place after the particles travel a fixed distance $d = \lambda_r \ln 2$ where λ_r is the radiation length in the medium. In fact, d is the atmospheric depth where an electron loses half its energy by radiation. The picture of the model is that each particle after travel a distance d is split into two particles with the energy equally divided between them. After n splitting lengths or generations, the atmospheric depth along the shower axis is $X = n\,d$ and the total number of particles is:

$$\mathrm{N}(\mathrm{X}) = 2^{\mathrm{n}} = e^{\mathrm{X}/\lambda_r}. \tag{1.2}$$

Assuming that the initial particle has an energy E_0, the energy of each particle at a given atmospheric depth X is:

$$\mathrm{E}(\mathrm{X}) = \frac{\mathrm{E}_0}{\mathrm{N}(\mathrm{X})}. \tag{1.3}$$

The number of particles cannot increase indefinitely. The particle-multiplication process continues until particles reach the *critical energy* E_c, where ionization-energy losses equals radiative losses. In the case of air $E_c = 86$ MeV. When the particles reach the critical energy, the number of particles in the shower reaches the maximum $\mathrm{N}_{\mathrm{max}}$ at the generation n_c.

$$\mathrm{N}_{\mathrm{max}} = \frac{\mathrm{E}_0}{\mathrm{E}_\mathrm{c}} = 2^{n_c} = 2^{\mathrm{X}_{\mathrm{max}}/\lambda \ln 2}. \tag{1.4}$$

From this expression, $\mathrm{X}_{\mathrm{max}}$ is given by

$$\mathrm{X}_{\mathrm{max}} = \lambda \ln\left(\frac{\mathrm{E}_0}{\mathrm{E}_\mathrm{c}}\right). \tag{1.5}$$

Here, two important features of electromagnetic air showers are obtained. The maximum number of particles is proportional to the primary cosmic ray energy E_0 and the atmospheric depth along the shower axis where it happens is proportional to the logarithm of E_0:

$$\begin{aligned} N_{max} &\propto E_0 \\ X_{max} &\propto \ln E_0. \end{aligned} \tag{1.6}$$

The generalization of the Heitler model to air showers initiated by hadrons is called *Heitler-Matthews model* [26]. The atmospheric depth of each generation is $\lambda_I \ln 2$ where λ_I is the interaction length. For pions in air, $\lambda_I \sim 120\,\mathrm{g\,cm^{-2}}$. This model assumes that each hadron produces N_{ch} charged pions and $N_{ch}/2$ neutral pions per generation. In the next generation, all pions decay into photons ($\pi^0 \to 2\gamma$) initiating electromagnetic showers and charged pions interact again. The hadron cascade stops when the average energy of hadrons is below the critical energy E_{dec} where all charged pions are assumed to decay ($\pi^- \to \mu^- \bar{\nu}_\mu$). Therefore, the total number of muons produced by a hadron shower with energy E_0 is:

$$N_\mu = N_{ch} = (N_{ch})^{n_c} = \left(\frac{E_0}{E_{dec}}\right)^\beta \tag{1.7}$$

where $\beta = \frac{\ln n_{ch}}{\ln n_{tot}} \approx 0.82 \ldots 0.94$ is calculated by Monte Carlo simulations and is model dependent.

In the case of an air shower initiated by a nucleus of A nucleons with energy E_0, the superposition model considers the nucleus as A independent nucleons with energy E_0/A and the $N_{EM,max}$, N_μ and X_{max} of the generated air shower are the following [27]:

$$\begin{aligned} {N^A}_{EM,max}(E_0) &= A{N^p}_{EM,max}(E_0/A) \approx {N^p}_{EM,max}(E_0) \\ {X^A}_{max}(E_0) &= {X^p}_{max}(E_0/A), \\ {N^A}_{\mu}(E_0) &= A\left(\frac{E_0/A}{E_{dec}}\right)^\beta = A^{1-\beta}{N^p}_{\mu}(E_0). \end{aligned} \tag{1.8}$$

These expressions show the sensitivity of X_{max} and N_μ to the composition of the primary cosmic ray.

In Sect. 1.1 was discussed how the chemical or mass composition of cosmic rays is very important in the understanding of the energy spectrum. Changes in mass composition produce changes in the spectral index γ. A better knowledge of the mass composition and energy spectrum could help to distinguish or reject different models of production and propagation. Since high energy cosmic rays can be only studied by means of EAS, it is necessary to use EAS observables sensitive to energy and mass composition of the primary cosmic ray. As we have already seen, the depth of the maximum development of the EAS (X_{max}) and the muon content are some of these observables [28]. The X_{max} is a useful observable since EAS produced by heavy nuclei are expected to develop earlier than proton ones. In addition, X_{max} fluctuations are lower for heavy nuclei than protons primaries, since heavy nuclei showers can be thought as an average lower energy proton shower using the superposition model.

1.3 The Pierre Auger Observatory

The Pierre Auger Observatory [29] is located close to the city of Malargüe, Argentina, around 1400 m above sea level. Its design was discussed in the early 1990s to give answer to the main

question about Ultra High Energy Cosmic Rays (UHECR) which have energies above 10^{18} eV. Among others, these questions point to definitely know the mass composition of primaries, measure the suppression of the energy spectrum, individualize sources and test hadronic interaction models at these energies. The construction of the Auger Observatory began in 2001, three years later entered into the production phase and in 2008 the construction of the base design was concluded [30] .

UHECRs are indirectly studied through the detection of EASs, explained in Sec. 1.2. These cascades of particles can be measured while they travel through the atmosphere and when particles reach the ground. In the following sections the different detection techniques implemented in Auger are presented.

1.3.1 Hybrid detection

The Auger Observatory implements a hybrid detector to observe the air shower development through the atmosphere and particles at ground simultaneously. On one hand, particles at ground are measured by the surface detector (SD) [31], which is composed of 1660 water-Cherenkov detectors (WCDs). The standard array (SD-1500) is arranged in an equilateral triangular grid of 1500 m covering a total area of 3000 km^2. On the other hand, the fluorescence light produced by the interaction of the electromagnetic component of the shower with nitrogen molecules in their travel through the atmosphere is measured by the fluorescence detector (FD), which consists of 24 fluorescence telescopes distributed in 4 buildings. The combination of these two detectors results in a hybrid detector [29] which reduce systematic uncertainties. The hybrid detector is shown in Fig. 1.3 together with other sites where are located atmospheric monitors as the Central Laser Facility (CLF) [32], the eXtreme Laser Facility (XLF) and the Balloon Launching Station (BLS). In addition, the figure shows some enhancements described in Sect. 1.3.5.

A hybrid event is an EAS that is simultaneously detected by the FD and the SD. The picture of the reconstruction of a hybrid event is shown in Fig. 1.4. The detection principles of both SD and FD are presented in the following sections.

1.3.2 The Surface Detector

The surface detector consists of an array of more than 1600 stations separated by a distance of 1.5 km and arranged in a triangular grid, as mentioned before. This spacial distribution is the needed to obtained enough statistics for UHECRs. Each SD station is formed by one WCD detector which consists of a cylindrical tank with 3.6 m diameter and 1.5 m height with a sealed liner with a reflective inner surface filled with 12000 liters of ultra pure water. Three photomultiplier tubes (PMTs) are symmetrically distributed on the surface of the liner looking downward into the water through windows of clear polyethylene. When a charged ultra-relativistic particle passes through the water, Cherenkov radiation is produced which is then detected by the PMTs. In this way, the lateral particle distribution of air showers is measured and the energy of the primary cosmic ray can be inferred.

Moreover, each SD station is equipped by a GPS in order to determine the time of detection needed to build the hybrid events and to reconstruct the shower geometry [34, 35] and solar panels provide the energy needed by the PMTs and the electronics. A picture of an SD station is shown in Fig. 1.5.

1.3.3 The SD triggers

The SD has implemented triggers for different kinds of purposes. The main trigger results from the SD traces and has two levels of selection [36]. The first level called T1 has two

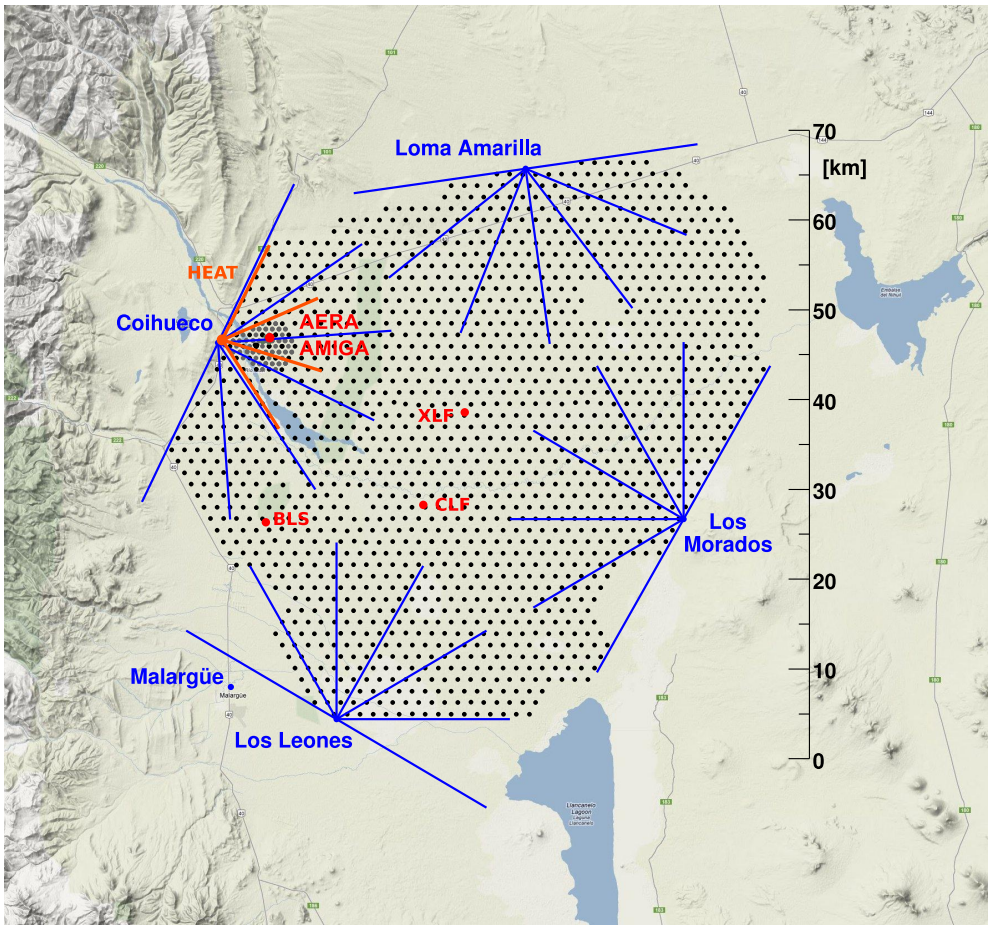


Figure 1.3: Map of the Auger Observatory and its detectors locations. Each black dot corresponds to one of the 1660 surface detector stations (see Sect. 1.3.2). The fluorescence light detectors are distributed in four buildings surrounding the surface detector array labeled in blue and their field of view are also shown by blue straight lines (see Sect. 1.3.4). Low-energy enhancements located close to the Coihueco site are also shown (see Sect. 1.3.5).

independent trigger modes to detect the electromagnetic and muonic components in a complementary way. The first T1 mode is a simple threshold trigger (TH) requiring the coincidence signal of all three PMTs to be above a certain threshold. The second T1 mode is a Time-over-Threshold trigger (ToT) which requires two out of the three PMTs to be above a lower threshold for a given minimum time in coincidence. The TH trigger is effective for the detection of very inclined showers where the muonic component dominates while the ToT trigger is efficient for the detection of vertical events.

The second level trigger is called T2. T1-ToT triggers are automatically promoted to T2-ToT, but only T1-TH triggers passing a single threshold in coincidence for the three PMTs become T2-TH triggers. All T2 triggers send their time-stamp to the Central Data Acquisition System (CDAS) for the global level trigger T3 determination. The T3 trigger is based on the spatial and temporal combination of T2 to select the possible physical events. The first T3 mode, called ToT2C1&3C2, requires the coincidence of at least three stations that have passed the ToT condition, approve a timing criteria and they have to be placed similar to the picture of Fig. 1.6-left. The second T3 mode, called "2C1&3C2&4C4", is more permissive since it requires a four-fold coincidence of any T2 with moderate compactness as Fig. 1.6-right shows. This trigger is useful for the detection of horizontal showers.

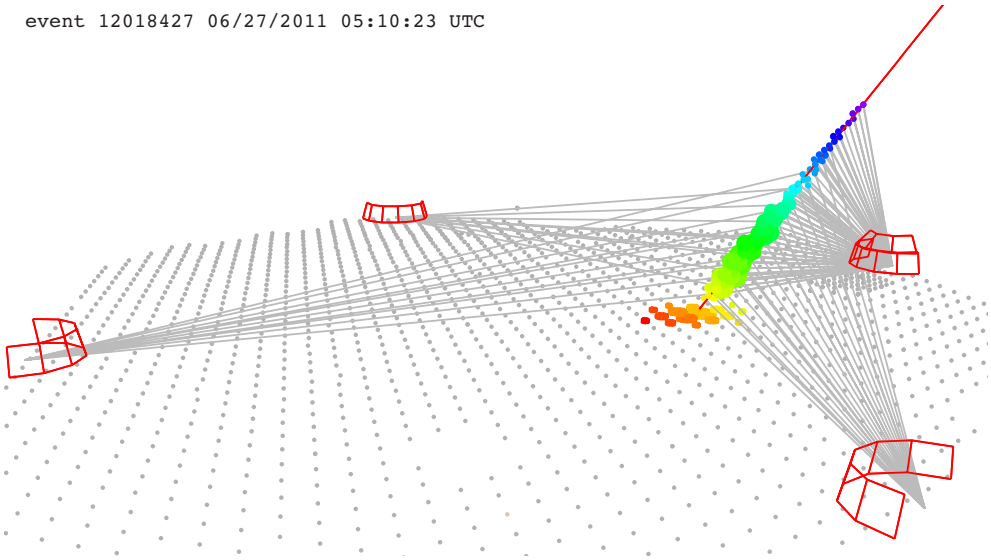


Figure 1.4: Reconstructed event by the Auger Observatory measured by both the SD and the FD [33]. Shower time along its axis and at ground is shown by colors where the size of the markers correspond to the magnitude of measured signals In this example, the reconstructed energy was E = $(4.7 \pm 0.1) \times 10^{19}$ eV and X_{max} = (730 ± 3) g/cm^2.

Figure 1.5: Picture of a Water-Cherenkov detector at the Pierre Auger Observatory.

Two additional T1 triggers were implemented in 2013, applying more sophisticated analysis to the WCD traces. These are the time-over-threshold-deconvolved (ToTd) and the

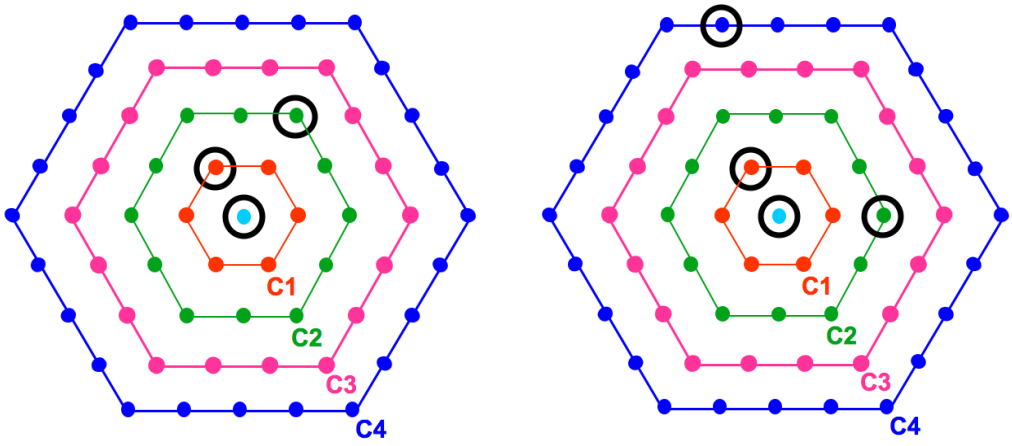


Figure 1.6: Example of T3 configurations: the 3-fold T3 mode ToT2C1&3C2 is shown on the left and the 4-fold mode 2C1&3C2&4C4 on the right where Cn indicates the nth set of neighbors [37].

Figure 1.7: Picture of the FD building and its communication tower at Los Leones [14]. In this case, the shutters were open because of maintenance.

multiplicity-of-positive-steps (MoPS) trigger, which improve the sensitivity for lower energy events and the sensitivity to photon and neutrino primaries thanks to a less influence of muons in the trigger.

1.3.4 The Fluorescence Detector

The FD is constituted by 24 fluorescence telescopes placed at four sites surrounding the SD array. These sites are Los Morados, Loma Amarilla, Los Leones and Cohiueco. Each one has a building where six telescopes with a field of view of $30^\circ \times 30^\circ$ in azimuth overlook the SD as Fig. 1.3 shows. The building at *Los Leones* is shown in Fig. 1.7.

The FD measures the fluorescence light produced in the development of EAS through the atmosphere. The electromagnetic particles in the EAS excite nitrogen molecules of the atmosphere which emit fluorescence light during the de-excitation. The fluorescence light

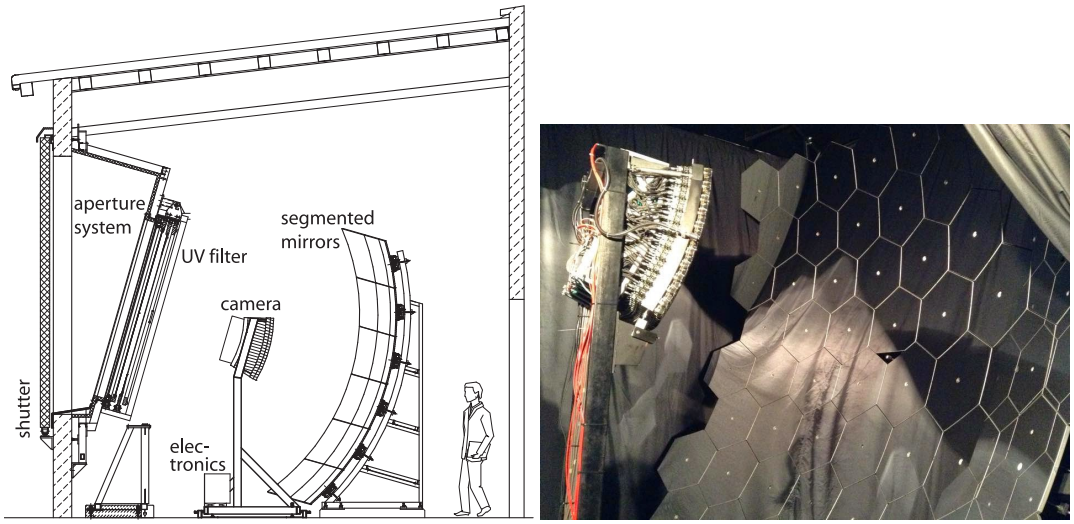


Figure 1.8: Schematics of a fluorescence telescope is shown on the left [40] and a picture of the telescope mirror and camera on the right.

that arrive at each building enter through a circular diaphragm with a diameter of 1.1 m and a UV filter allows only photons with wavelength between 300 and 410 nm to pass through. Then the light is focused into a camera with an array of 20 x 22 PMT as shown Fig. 1.8. The FD is only operated at moonless nights with clear skies. For this reason, their duty cycle is limited up to $\sim$19% of the time, which is normally reduced to $\sim$15% due to weather conditions as storms and high wind [14]. The FD operation and the analysis of the measured signals require the knowledge of several atmospheric parameters [38]. For example, clouds and aerosols may scatter and attenuate the produced fluorescence light. For that reason, aerosol monitoring is performed using four elastic scattering LIDAR stations and the Central and eXtreme Laser Facilities (CLF/XLF), among other facilities [39].

The intensity of the produced fluorescence light depends on the number of particles in the EAS, which is related to the energy of the primary particle (see Sec. 1.2). Therefore, the fluorescence light measured with the FD telescopes is used to estimate the calorimetric energy of the primary particle. In addition, other parameters can be estimated as the X_{max} by analyzing the intensity profile or the arrival direction of the primary cosmic ray by analyzing the arrival time of photons.

1.3.5 Enhancements of the Observatory

The Auger Observatory has implemented several enhancements in order to, among other things, improve its sensitivity to different observables and to extend its detection energy range. The main low-energy enhancements are the Auger Engineering Radio Array (AERA), the Auger Muon and Infill for the Ground Array (AMIGA) and the High Elevation Auger Telescopes (HEAT). Fig. 1.9 indicates where they are deployed.

AERA is built by 150 radio stations and enables the study of EAS's radio emission. The radio emission is mainly generated by the geomagnetic deflection of electrons and positrons during EAS development. Therefore, radio detectors are sensitive to the electromagnetic component. HEAT is a low energy extension of the FD, where three fluorescence telescopes covering the elevation range from 30° to 58° were placed at the Coihueco site. As the data analysis of the AMIGA detectors is a central part of this thesis, the AMIGA extension is described in the next subsection and its muon detectors are described in detail in Ch. 5.

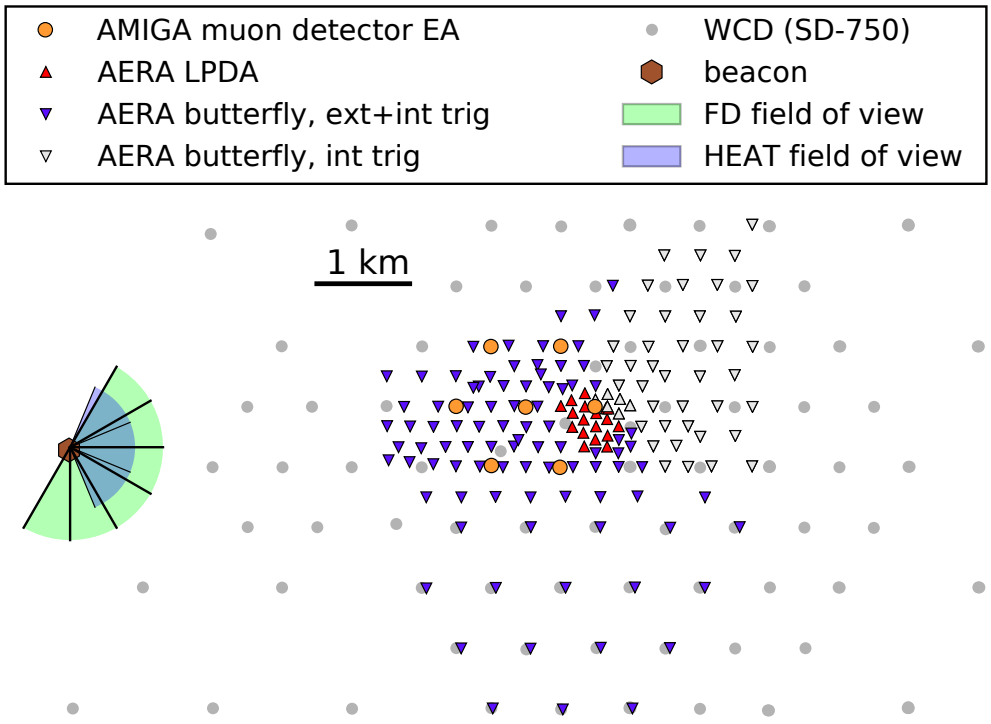


Figure 1.9: Map of the low-energy enhancements area [37], where the AERA antennas, the AMIGA UMDs, the SD-750 stations and HEAT field of view are shown.

The Pierre Auger Observatory is planned to be operated until the end of 2025. AugerPrime is a major upgrade of the current design of the Observatory [41]. This upgrade consists on the addition of new detectors and the enhancement of existing ones and aims to provide additional composition information that will improve the reconstruction of the properties of the primary particles at the highest energies.

Auger Muon and Infill For the Ground Array (AMIGA)

The AMIGA enhancement comprised the deployment of denser SD arrays and the deployment of a dedicated Underground Muon Detector (UMD). On one hand, a denser array with a 750 m spacing (SD-750) extending over 23.5 km^2 was deployed to lower the energy detection range of the SD [42]. Furthermore, an even denser array of 1.9 km^2 with a 433 m spacing (SD-433) is being deployed to extend the SD energy threshold down to $10^{16.5}$ eV [43]. The location of the AMIGA site is illustrated in Fig. 1.3 and 1.9.

On the other hand, the muonic component of air showers can be directly measured thanks to the AMIGA underground muon detector UMD. The muon content of EASs is an observable sensitive to the mass composition of the primary cosmic ray (see Sec. 1.2 and Ref. [44]) These UMDs are deployed close to each SD-750/SD-433 station and are buried at 2.3 m to shield them from the electromagnetic component.

Each UMD consists of three 10 m^2 modules segmented into 64 plastic-scintillator strips with wavelength-shifting (WLS) optical fibers and photodetectors located at the center of the module [45]. An engineering array was deployed during the prototype phase for physics validation and optimization of the detector design. Two kind of photo-detectors were analyzed, a 64-pixel multi-anode photomultiplier tube (PMT) and an array of 64 silicon photomultipliers (SiPMs), being the SiPM chosen for the production phase of the AMIGA UMDs. A picture of a UMD is shown in Fig. 1.10.

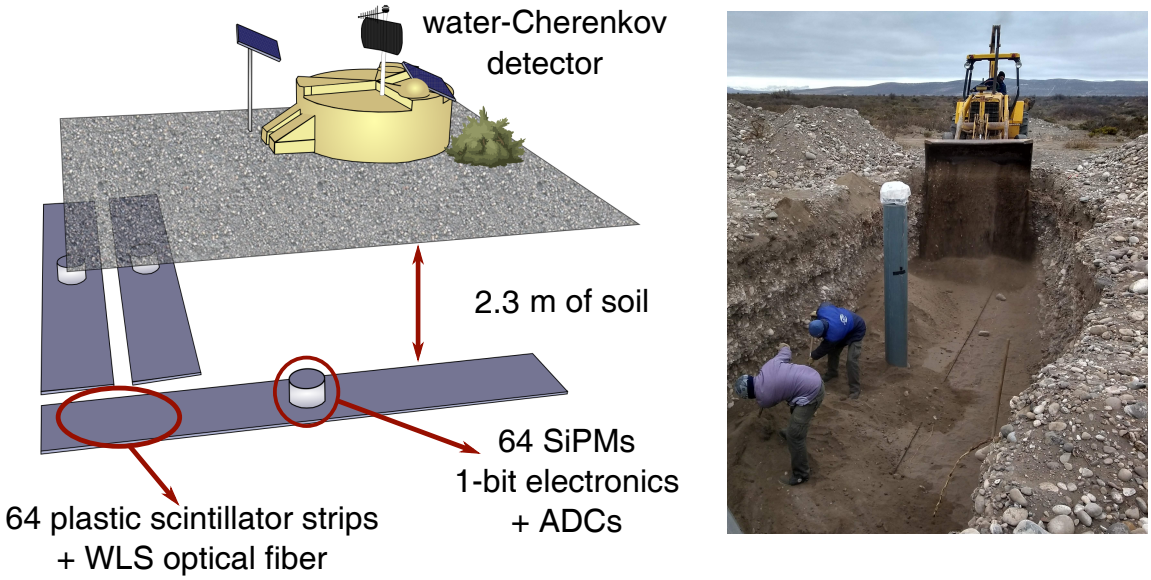


Figure 1.10: Layout of the AMIGA production-phase design where each WCD of the SD-750/433 array is coupled to three $10\,\mathrm{m}^2$ buried scintillator modules on the left and a picture of the a detector deployment in the field on the right [37].

The muon detector works in tandem with each SD station. When a T1 trigger condition is found by the WCD, it sends a trigger signal to the UMD to keep its buffered data. And once a T3 condition is met, the UMD sends the event information to the CDAS. Further details on the AMIGA UMDs are presented in Chapter 5.

1.3.6 Published results highlights

The Auger Observatory has published a large and important number of results in the field of UHECR physics. Among others achievements, Auger confirmed the suppression in the energy spectrum above 5×10^{19} eV [46], discovered a dipole in the arrival directions of cosmic rays with energies greater than 8×10^{18} eV [47], and set leading photon [48] and neutrino [49] limits. These results are very important to understand the origin and propagation of cosmic rays. In addition, Auger measured proton-air cross-sections at $\sqrt{s} \approx 57$ TeV [50].

One important result, that motivates this thesis, is the muon deficit in EAS simulations using the LHC-tuned hadronic interaction models when compared with measurements [51], giving the possibility to perform studies related to particle physic. In Sect. 1.4 the results of the muon deficit from different experiments are analyzed.

1.4 The muon puzzle

Since high energy cosmic rays can only be indirectly measured by extensive air showers (EAS), a deep knowledge of particle interactions in the air shower development is fundamental to infer their mass and energy. The tracking of the longitudinal shower development by the fluorescence technique reduces the model-dependence for the energy measurement. On the other hand, the main observables used to infer the mean logarithmic mass of the primary, $\langle \ln A \rangle$, are the depth of the shower maximum in the atmosphere X_{max} and the number of muons in the shower N_μ (see Sect. 1.2) . Each observable depends mainly on a different component of the air showers. The electromagnetic component contributes mainly to X_{max}

while the hadronic cascade is directly related to N_μ. The uncertainty in the interaction models used to simulate air showers has a large impact on the inferred value of $\langle \ln A \rangle$. Moreover, $X_{\max}$ and N_μ measurements in terms of mass interpretation are not consistent according to the simulations done with the leading hadronic interaction models. This implies that the description of the hadronic physics in these models needs to be improved. This situation can be seen as an opportunity to test models at phase-space regions beyond the reach of the LHC.

Most measurements of the electromagnetic component of air showers show acceptable agreement with simulations using the post-LHC generation of hadronic interaction models. The most direct measurement of a hadronic interaction property using air showers is the proton-air cross section measured by the Auger Observatory and the Telescope Array. It is based on the slope of the tail of the X_{max} - distribution [52, 53]. These measurements constrain hadronic interaction models since the mass composition dependence is reduced. The average longitudinal shapes of air showers, obtained by the Auger Observatory [54, 55], and the signal attenuation with zenith angle measured by the Telescope Array [56] are both compatible with simulations.

On the other hand, most measurements of the muonic component of air showers show disagreement with simulations. There are eight EAS experiments which have reported the muon content. In the year 2000, the HiRes/MIA collaboration reported a discrepancy of the number of muons in simulated and measured air showers between 10^{17} to 10^{18} eV [57]. A decade later, NEVOD-DECOR reported [58, 59] an increase of the muon density relative to simulations from 10^{15} to 10^{18} eV. The SUGAR array also observed an excess in data [60]. EAS-MSU [61] as well as KASCADE-Grande [62] reported no muon number discrepancy in this energy range. Auger Observatory [51, 63] and the Telescope Array [64] also observed a muon deficit in simulations at 10^{19} eV with the latest models.

In addition, other measurements show tension when data is compared to simulations. The muon production height measured by Auger [65] and KASCADE-Grande [62] are inconsistent with for QGSJet-II.04 and EPOS-LHC models at EeV energies. Measurements of the slope of the lateral density profile at the Icecube surface array and muon bundles below a kilometer of ice [66], which both are sensitive to cosmic-rays mass composition, show inconsistent results for SIBYLL-2.1 and EPOS-LHC models. The attenuation of the muon lateral profile as a function of the zenith angle measured by KASCADE-Grande [58] does not agree with the leading hadronic interactions models. The atmospheric flux [67, 68] and the lateral separation [69] of TeV muons measured by IceCube and the rise-time of the shower front measured by Auger [70, 71] are other examples with disagreement between data and simulations.

1.4.1 Measurements of the muon lateral density

The WHISP group [72, 73] analyzed the lateral muon density measured from the following eight cosmic-ray experiments: EAS-MSU [61], IceCube Neutrino Observatory [74], KASCADE-Grande [58], NEVOD-DECOR [58, 59], Pierre Auger Observatory [51, 63, 75], SUGAR [60], Telescope Array [64] and Yakutsk based on preliminary unpublished data [76].

A direct comparison between these lateral muon density measurements is not possible because they are performed under very different conditions and using different techniques. The muon density at ground depends on many parameters which differ from experiment to experiment. These are the cosmic ray energy and zenith angle, the energy threshold of the muons detectors, the lateral distance from the shower axis and the shower age (i.e. the stage of the air shower development). The parameter space covered by each experiment is shown in Fig. 1.11. The approach used by the WHISP group is that two measurements are

only comparable, if simulations with the same hadronic interaction model are available for both. In that case, the ratio of data and simulation is comparable between experiments.

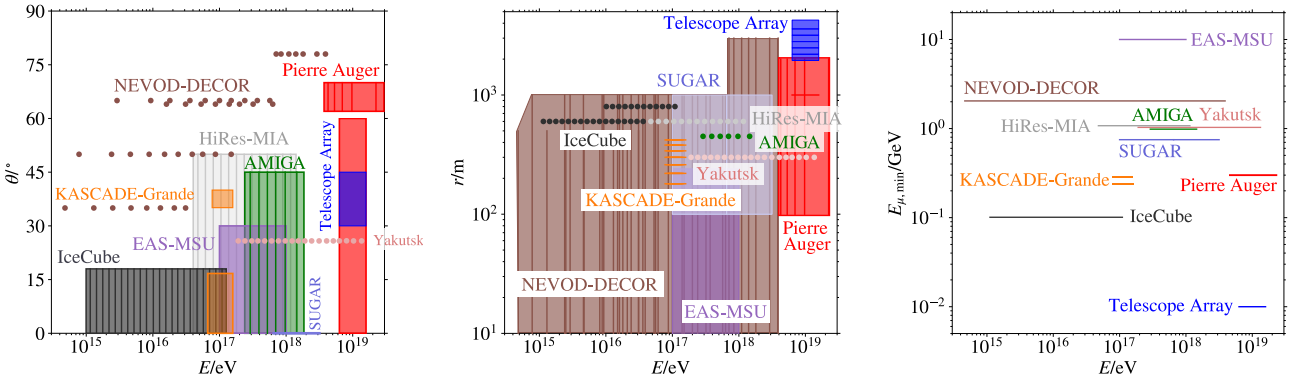


Figure 1.11: The muon density at ground has been measured by different experiments under various conditions. These plots show the parameter space measured as a function of the air shower energy for: the zenith angle of air showers (left), the lateral distance of the muon density measurement (middle) and the energy threshold for the muons [73].

Different kinds of parameters need to be set in simulations. Most of them are easily extracted from data but the cosmic ray energy is not one of them and this has a large impact. According to the Heitler-Matthews model of air showers [26] the muon number depends on the energy E and the mass A of the cosmic ray according to (from eq. 1.7 and 1.8)

$$N_\mu = A^{1-\beta}(E/C)^\beta \tag{1.9}$$

where $\beta \sim 0.9$ and C is a constant. So, the muon number scales almost linearly with the cosmic-ray energy. Taking into account that air shower experiments usually have independently calibrated energy scales with systematic uncertainties of 10 % to 20 %, the ratio of data and simulations will have a similar offset because measurements are compared to air showers simulated at different energies. The way to remove these offsets is making a cross-calibration of the energy scale. This is done taking the cosmic ray flux as a universal reference and considering that any deviations in the measured fluxes are due to the energy-scale offsets. Making fluxes of all experiments overlap, the relative energy-scale ratio $E_{\text{data}}/E_{\text{ref}}$ is calculated for each experiment. E_{ref} was set, by the Spectrum Working Group formed by the Pierre Auger and Telescope Array collaborations, to be placed between the two experiments.

The cross-calibration factors are given in Table 1.11. The resulting overall energy-scale is assumed to have an uncertainty of at least 10 %.

The Auger Observatory, Telescope Array, Yakutsk and IceCube Neutrino Observatory can calculate the cosmic ray energy E and lateral muon density. The first three experiments calculate the calorimetric energy of showers, E_{cal}, by the measurement of the Cherenkov or fluorescence light emitted (see Sect. 1.3.4). E_{cal} is converted to E with low model-dependence. This is done almost independently of the muon density at ground. These experiments can compute the data/MC ratio for showers with the same energy. Instead, IceCube can calculate E directly with surface detectors, because its altitude is close to the average depth of the shower maximum [77].

KASCADE-Grande and EAS-MSU measure the density of electrons and muons separately. These experiments compute the data/MC ratio for showers in the same electron-density interval, which can contain different primary energies. The electron density is correlated to the cosmic-ray energy, but also to the muon density [78]. So, these ratios cannot be directly compared with those of former experiments.

Experiment	$E_{\text{data}}/E_{\text{ref}}$	sec θ	$E_{\mu\,\text{prod}}$/GeV
EAS-MSU	-	1.1	11.9
IceCube Neutrino Observatory	1.19	1.0	0.7
KASCADE-Grande	-	1.0 , 1.3	1.5 , 2.1
NEVOD-DECOR	1.08	2.3 , 4.8	8.4 , 18.6
Pierre Auger Observatory	0.948	1.3 , 2.4	1.8 , 4.0
AMIGA	0.948	1.2	2.4
SUGAR	0.948	1.0	1.9
Telescope Array	1.052	1.3	1.4
Yakutsk EAS Array	1.24	1.1	2.6

Figure 1.12: For each experiment, the table shows in the second column the energy-scale adjustment factors obtained from cross-calibration, the median sec θ of several data sets, and the minimum energy of the muons at production (due to energy loss in the atmosphere and detector shielding) [73].

NEVOD-DECOR and SUGAR are pure muon detectors, without a separate energy estimator. The flux of showers is measured in intervals of muon density, which is then compared with a simulated flux. An average cosmic-ray flux from several experiments is used by NEVOD-DECOR, while the flux measured by the Pierre Auger Observatory is used by SUGAR.

Taking the logarithm on both sides of Eq. 1.9 and then the mean, a linear equation is obtained

$$\langle \ln N_\mu \rangle = (1-\beta)\langle \ln A \rangle + \beta \langle \ln(E/C) \rangle \tag{1.10}$$

Based on this relation the z-factor is defined in order to compare all the lateral muon density measurements:

$$z = \frac{\langle \ln N_\mu \rangle - \langle \ln N_\mu \rangle_{\text{p}}}{\langle \ln N_\mu \rangle_{\text{Fe}} - \langle \ln N_\mu \rangle_{\text{p}}} \tag{1.11}$$

Here N_μ is the muon density estimate as seen by the detector, while N_μ^{p} and N_μ^{Fe} are the simulated muon density estimates for proton and iron showers obtained by a full detector simulation. The z-factor is 0 for proton showers and 1 for iron showers. This definition has the feature that any bias of the form $\ln N_\mu = A + B \ln N_\mu^{\text{true}}$ cancels in z.

The measurements of all considered experiments converted to z as a function of the cosmic ray energy are shown in Fig. 1.13. For some models it was not possible to calculate N_μ^{p} or N_μ^{Fe}, so z is not available for them. The data suggest an energy-dependent trend, but with a large scatter. Applying the cross-calibration, the scatter is drastically reduced as is shown in Fig. 1.14 for EPOS-LHC and QGSJet-II.04 models. The cross-calibration causes a shift in the simulated values N_μ^{p} and N_μ^{Fe}, which were computed for the energy E_{data}, but are needed for E_{ref}.

The points move horizontally by the relative amount $(E_{\text{data}}/E_{\text{ref}})^{-1}$, which has a low impact, and vertically by

$$z_{\text{ref}} = z_{\text{data}} + \frac{\beta \ln(E_{\text{data}}/E_{\text{ref}})}{\langle \ln N_\mu \rangle_{\text{Fe}} - \langle \ln N_\mu \rangle_{\text{p}}} \tag{1.12}$$

with $\beta = 1 - (\langle \ln N_\mu \rangle_{\text{Fe}} - \langle \ln N_\mu \rangle_{\text{p}})/\ln 56$. It is worth noting that the z-factor in all plots can collectively vary by about ± 0.25, since the reference energy-scale after cross- calibration has a remaining uncertainty of at least 10 %.

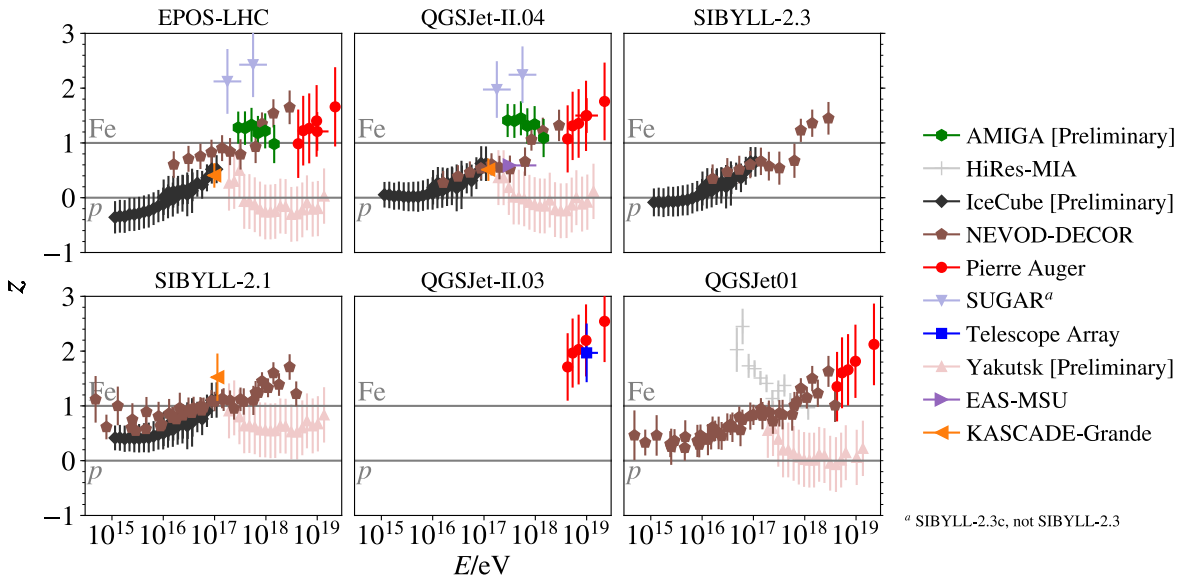


Figure 1.13: Muon density measurements converted to the z-scale for each hadronic interaction model [72]. For some experiments the corresponding simulations are missing.

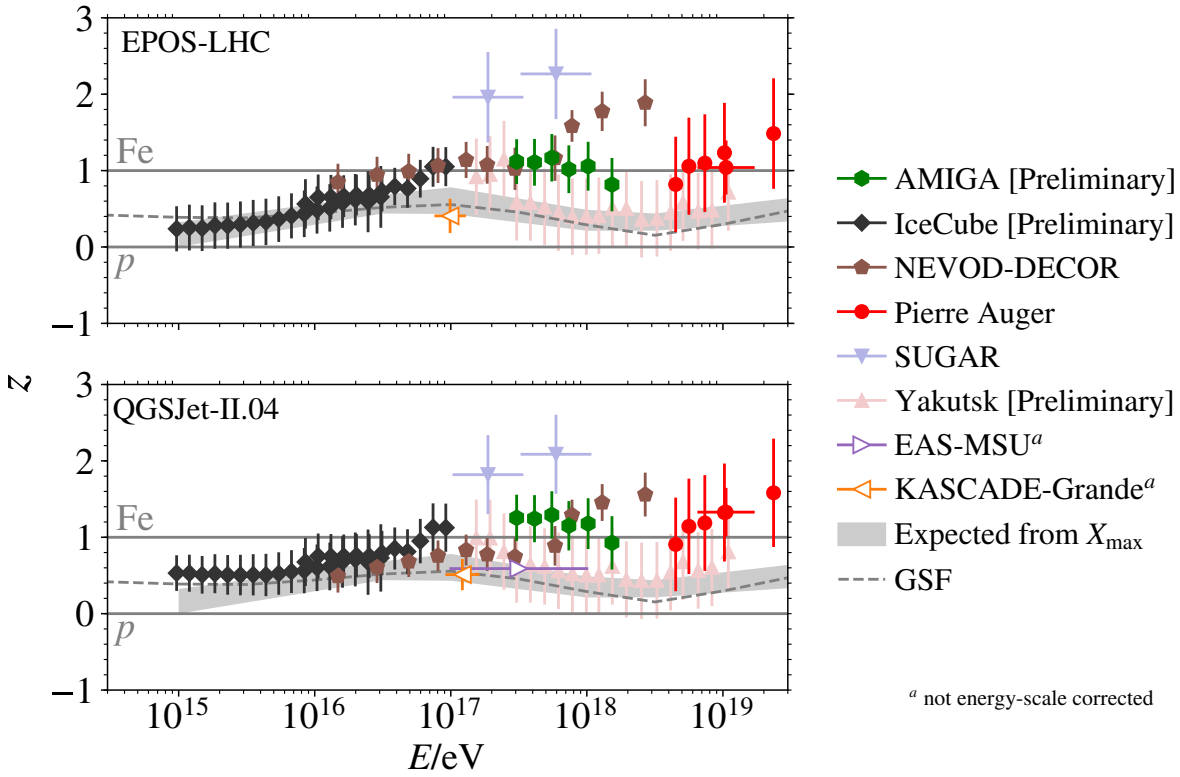


Figure 1.14: Data from Fig. 1.13 after applying energy-scale cross-calibration for EPOS-LHC and QSGJet-II.04 [72]. The points for KASCADE-Grande and EAS-MSU cannot be cross-calibrated and are only included for comparison.

The effect of an energy-dependent mass composition was also analyzed. From Eq. 1.10 and 1.11, the expected value for a given mean-logarithmic-mass $\langle \ln A \rangle$ is

$$z_{\text{mass}} = \frac{\langle \ln A \rangle}{\ln 56}, \tag{1.13}$$

The optical measurements of X_{max} from several experiments, were converted to $\langle \ln A \rangle$ based on air shower simulations with EPOS-LHC. This result is shown as a band in Fig. 1.14 and it is independent of the muon measurements and can therefore be used as a reference. In the same figure, the Global spline fit (GSF) model was used to calculate the z_{mass} value (see the details in [79]), and its line mostly falls inside the envelope.

If the measured z values follow z_{mass}, the model describes the muon density at the ground consistently. This is overall not the case. EPOS-LHC, QGSJet-II.04 and SIBYLL-2.3 give a reasonable description of data up to a few 10^{16} eV. At higher shower energies, a muon deficit in simulations is observed ($z > z_{\mathrm{mass}}$) in all models.

In order to remove the effect of the changing mass composition the difference $\Delta z = z - z_{\mathrm{mass}}$ is shown in Fig. 1.15 where an energy-dependent trend in Δz remains. The positive slope in the trend implies a continuous increase with energy which starts at a relatively low energy. It means that the muon puzzle is not due to a dramatic change of hadronic interactions at very high energy. In this way, any approach that seeks to explain the current situation has to do it at low energies, too. In Sec. 3.3, we present a new versatile tool to test the change of hadronic interactions in the whole range of interest.

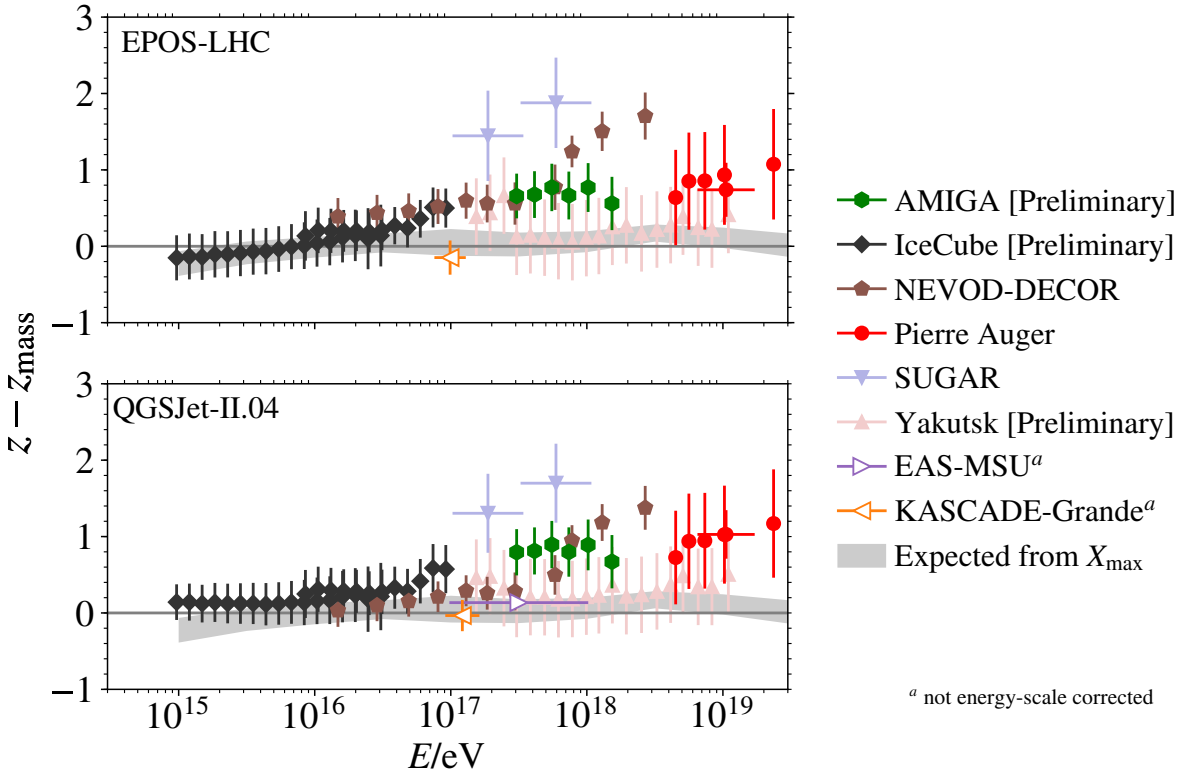


Figure 1.15: Data from Fig. 1.14 after subtracting z_{mass} [72].

CHAPTER 2

Hadronic interactions

The muon puzzle presented in Sec. 1.4 reveals that the current description of the hadronic interactions is still incomplete. Moreover, the discrepancy between data and simulations cover several orders of magnitude in energy. Before trying to change hadronic interactions, it is important to understand what is currently used in the models. In this chapter the basic theory about hadronic interactions is presented.

2.1 Introduction

Describing the hadronic structure is a complex challenge. Usually the preference of a given model depends on the scale of interest. On small scales hadrons can be described by Quantumchromodynamics (QCD), a quantum field theory where they are formed by quarks and gluons. Before the validation of the quark model in 1969, Feynman proposed the parton model for hadron description, where hadrons were considered to be formed by point-like objects called *partons*. Indeed, partons are quarks and gluons and the parton model was successfully applied to electron-proton deep inelastic scattering. On larger scales, a hadron description based on parton collections is not enough and it is necessary to consider collective interactions on parton groups.

Hadron collisions are mainly characterized by the total energy in the center-of-mass frame (cm), E_{cm}, which is usually expressed in terms of the Lorentz-invariant variable $\sqrt{s}$:

$$s = (p^{\mu}_{\text{beam}} + p^{\mu}_{\text{target}})^2 = (E_{\text{beam}} + E_{\text{target}})^2 = E^2_{\text{cm}} \tag{2.1}$$

where p^{μ} is the particle four-momentum of hadrons. The z-axis is commonly taken along the direction of the momentum, where the hadron with momentum $+p_z$ is called the beam and the one with $-p_z$ the target.

In the cm frame the hadrons have a null transversal momentum $p_{\perp} = \sqrt{p_x^2 + p_y^2}$. However, each individual parton has a typical transversal momentum $p_{\perp} \sim \frac{1}{R} \sim 0.4\,\text{GeV}$ from the uncertainty relation knowing that the hadron size is $R \sim 1\text{fm}$.

Over the momentum direction the Lorentz factors $\gamma = E/m$ for protons at the LHC is up to 10^4. So, hadron collisions at high energy have to be pictured as a collision of a narrow disk of partons. Using again the uncertainty relation, the typical parton parallel momentum is $\Delta p_{\parallel} \sim \frac{1}{R}\frac{E_p}{m_p} \sim \frac{E_p}{5}$.

It is worth noting that the term *hadronic interaction* typically only refers to strong interactions, though hadrons can interact via any of the interactions in the Standard Model.

Interactions of hadrons can be separated into two main categories, *elastic* and *inelastic* interactions. The former happen when the final state particles are the same as the initial ones. Since momentum and energy must be conserved, only the direction of the momentum vector can be changed. From the parton picture, elastic interaction requires the coherent interaction of all partons of both hadrons. So, this becomes more unlikely with increasing energy. On the other hand, *inelastic* interactions produce new particles in the final state, where hadrons can exchange energy, momentum or any quantum number. In this way the total cross sections, which give the probability of an interaction, can be written as

$$\sigma_{\text{tot}} = \sigma_{\text{ela}} + \sigma_{\text{inel}}. \tag{2.2}$$

Inelastic interactions can be further separated into *diffractive* and *non-diffractive* interactions. In diffractive interactions only energy and momentum are exchanged between the hadrons, but one or two hadrons transition into an excited state given a single or double diffractive interaction. The excited states eventually decay yielding new particles, typically pions, and therefore the interaction is counted as an inelastic process. This kind of interactions also requires the coherent interaction of the hadrons, so the fraction of diffractive events decreases with energy. The rest of the interactions, where any number of gluons or quarks can be exchanged, are called non-diffractive interactions. In this case hadrons are connected by multiple color fields that fragment into new hadrons.

Experimentally, non-diffractive and diffractive interactions can only be distinguished to a certain degree. Especially if the acceptance of the detectors does not cover the entire phase space.

2.2 Kinematic variables and phase space

To describe the final state of hadronic interactions different variables are used, depending on the process or phase space of interest. Transverse phase space is useful to study the structure of hadrons, the fundamental interactions or to find new particles and interactions. Instead, longitudinal phase space is much more important for the study of air showers and hadronic interactions in astrophysical scenarios, since this is where particles carry most of the energy. In these cases the typical interaction is looked for rather than the rare exception.

The transverse phase space is usually described directly with transverse momentum $p_\perp = \sqrt{p_x^2 + p_y^2}$. Since transverse momentum in the initial hadron beam is zero, the partons can only acquire transverse momentum in the scattering process. From the uncertainty relation in transverse space ($\Delta b \Delta p_\perp \sim 1$), large $p_\perp$ imply small length scales. So the study of parton interactions or parton substructure, are all carried out by looking at collisions with high $p_\perp$ particles (>10GeV). The large scale structure of hadrons on the other hand determines the low $p_\perp$ region.

The typical variable to describe longitudinal phase space is the Feynman-x [80]. For inclusive production of some particle A, i.e. $pp \to A + \ldots$, the Feynman-x is defined as

$$x_F = \left. \frac{p_\parallel^A}{p_{\parallel_{\text{max}}}^A} \right|_{\text{cm}} \tag{2.3}$$

where $p_\parallel^A$ is the longitudinal momentum of particle A in the cm frame and $p_{\parallel_{max}}^A$ is the maximum momentum that A could ever have in this frame, based on the energy of the collision and the masses of the particles. It is clear from the definition that x_F can vary

between -1 and 1. The original idea was that data for A production at different p+p beam energies might show a common behavior plotted against x_F, and this would factor out the beam-energy dependence of the shape of the spectrum. Data which follow this behavior are said to show x_F *scaling* which is why x_F is referred to as a scaling variable.

A similar variable x_L can be defined for air showers, where the reference frame is the fixed-target frame (lab frame) and hence the total momentum comes from the beam. The x_L variable is the longitudinal momentum of a particle A relative to the beam particle

$$x_L = \left.\frac{p_\parallel^A}{p_{\text{beam}}}\right|_{\text{lab}}. \tag{2.4}$$

The relation between x_F and x_L can be found applying a Lorentz boost transformation to x_F and expanding the ratio of the transverse mass $m_T = \sqrt{p_\perp^2 + m^2}$ to particle energy E

$$x_F = x_L - \frac{m_T^2}{2mE} + O[(m_T/E)^2]. \tag{2.5}$$

Note that $x_F \to x_L$ for small $p_\perp$ and large x_F since $m_T/E \to 0$.

The transverse and longitudinal phase spaces are combined in the *rapidity* variable.

$$y = \frac{1}{2}\ln\left(\frac{E+p_\parallel}{E-p_\parallel}\right) = \ln\left(\frac{E+p_\parallel}{m_T}\right). \tag{2.6}$$

The main feature of the rapidity is its simple transformation rule under Lorentz boosts, which is helpful to calculate multiple interactions:

$$y' = y + \ln\sqrt{\frac{1-\beta}{1+\beta}} = y - \tanh^{-1}\beta \tag{2.7}$$

where β is the relative velocity between the reference frames. The rapidity reaches its maximum value when $p_\perp \to 0$. For the initial particles the maximum rapidity is given by

$$y_{\text{max}}^{\text{beam}} = \ln\left(\frac{E+p_\parallel}{m_{\text{beam}}}\right) \approx \ln\left(\frac{\sqrt{s}}{m_{\text{beam}}}\right). \tag{2.8}$$

This means, all heavier particles cover the rapidity range $(-y_{\text{max}}^{\text{beam}}, y_{\text{max}}^{\text{beam}})$. Lighter particles, e.g. pions, can take larger rapidity values

$$y_{\text{max}}^{\text{pion}} = y_{\text{max}}^{\text{beam}} - \ln\left(\frac{m_{\text{pion}}}{m_{\text{beam}}}\right). \tag{2.9}$$

From an experimental point of view, it is not always possible to determine the identity of the particle when its momentum is measured. In those cases the rapidity cannot be calculated, since the total energy requires the knowledge of particle masses. To avoid this problem the *pseudorapidity* is commonly used. Here, the energy is replaced by the momentum in the rapidity definition. Moreover, the pseudorapidity can be written as a function of the scattering angle θ. So, in fact it is not necessary to measure the momentum, it is sufficient to track the particle:

$$\eta = \frac{1}{2}\ln\left(\frac{p+p_z}{p-p_z}\right) = \ln\left(\frac{p+p_z}{p_\perp}\right) = -\ln\left(\tan\frac{\theta}{2}\right) \tag{2.10}$$

The pseudorapidity range goes from $-\infty$ to ∞, taking these values in the beam direction. Particles created with momentum perpendicular to the beam axis have $|\eta| = 0$. This quantity

is useful to describe the phase space cover by different kinds of detectors used at colliders. Fig. 2.1 shows the typical pseudorapidity ranges covered at the LHC. Unlike charged particle detectors, neutral particle detectors can cover the forward phase space very close to the beam direction. The CMS and ATLAS detectors only cover $|\eta| \lesssim 5$ since their goal is to look for new physics, i.e. high $p_\perp$, and most particles are produced precisely at mid-rapidity as Fig. 2.1 on the bottom shows. However, the same figure shows that the energy flow $\mathrm{dE}/\mathrm{d}\eta$ goes in the forward region carried by few particles, which lies outside of the charged detector range at LHC. The accurate description of these particles at large rapidity is fundamental for the understanding of air shower development.

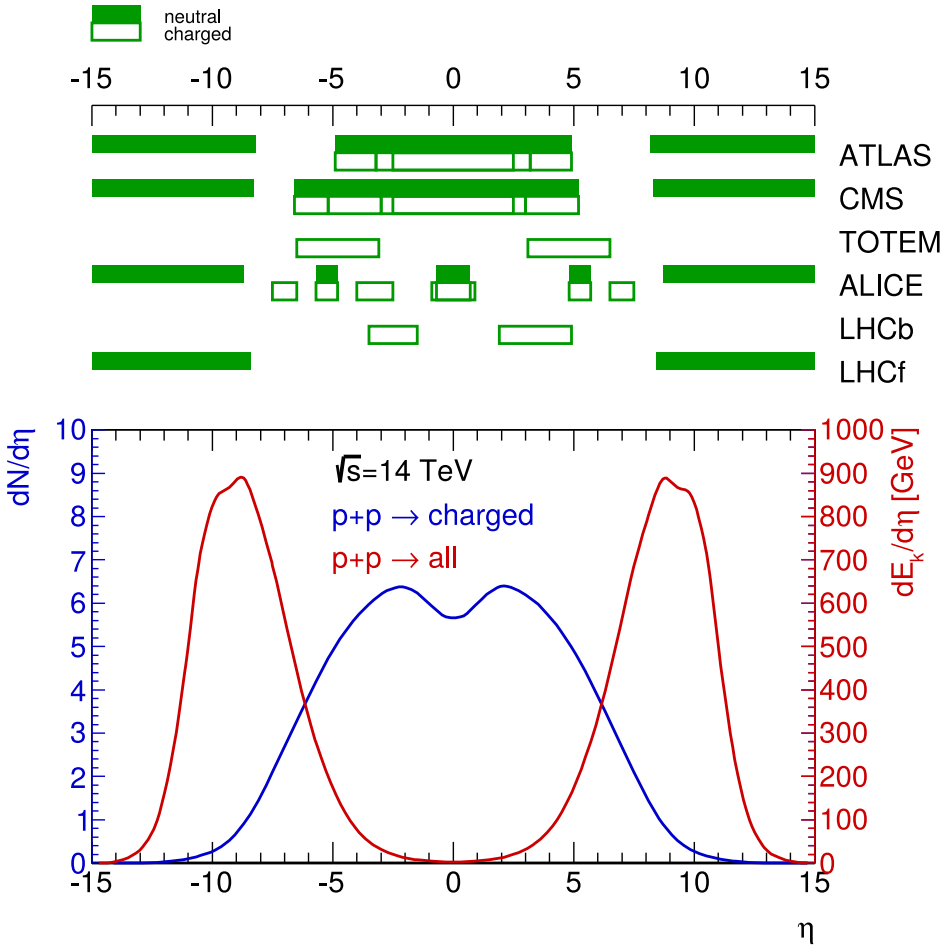


Figure 2.1: Energy and particle flow at $\sqrt{s} = 14\,\mathrm{TeV}$ as a function of pseudorapidity at the LHC [81]. The green bands represent the coverage of different detector components.

2.3 Regge phenomenology

The quark model assumes that hadrons are bound states of quarks, called *constituent* or *valence* quarks. In this picture, baryons are formed by three valence quarks qqq and mesons by one quark and one antiquark $q\bar{q}$. Most hadron properties and quantum numbers can be explained by this model. However, hadrons are formed by the valence quarks and a large number of gluons and quarks with low momentum. These gluons and quarks, known as *sea quarks*, are short-lived fluctuations of $q\bar{q}$ pairs produced in the color field of the bound hadron and contribute significantly to the overall momentum and spin of a hadron.

The masses of the quarks in the Standard Model, called *current quark masses*, do not give the whole mass of the hadrons. The masses of hadrons are mainly generated dynamically by interaction with low momentum exchanges. So, it is not possible to use the perturbative QCD framework. Effective potentials for quark interactions or lattice gauge theory can be used to model hadron masses.

An important empirical observation is that hadrons with the same quantum numbers only differ in the orbital momentum of the quarks, hence their spins J are different and their masses can be described by

$$m^2(J) \approx a \cdot J + m_0^2 \tag{2.11}$$

where a and m_0 are constants related to the quantum numbers of the hadron group. Plots of this relation are called *Chew–Frautschi plot* [82]. As an example the masses of the excited states of the $\rho(770)$ are shown in Fig. 2.2.

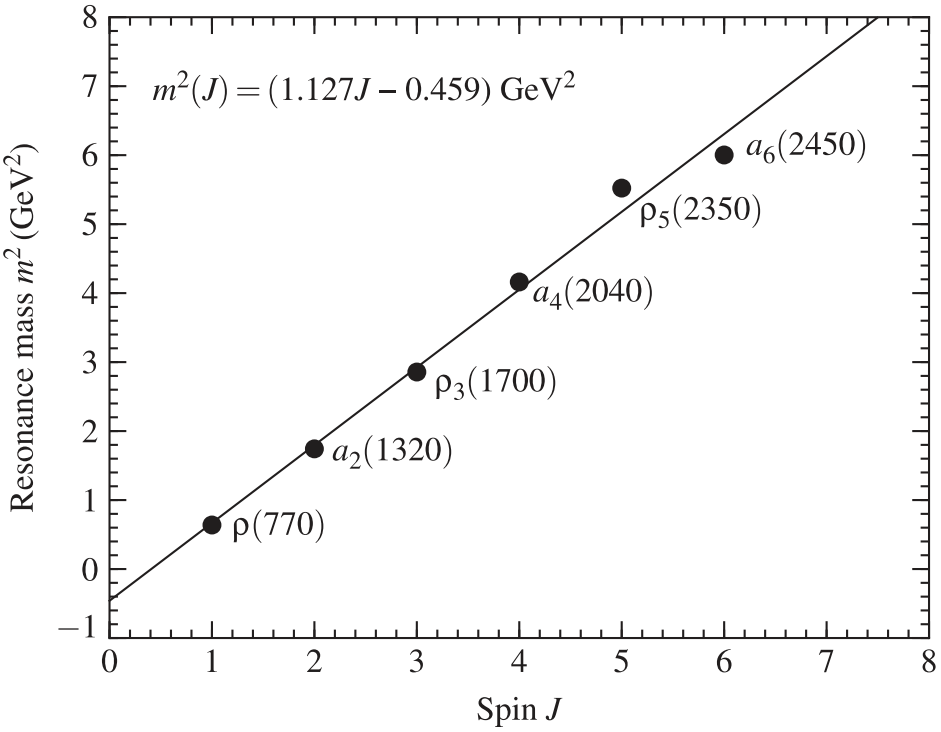


Figure 2.2: Chew–Frautschi plot for the resonances belonging to the $\rho(770)$ trajectory (isospin I=1, natural parity) [83].

From eq. 2.11, J is easily solved to be $J = \alpha(m^2) = \frac{m^2 - m_0^2}{a}$, where $\alpha(m^2)$ is called the *Regge trajectory* of the corresponding group of hadrons. Regge trajectories do not have to be linear functions of m^2 though in most cases this is a good phenomenological approximation, so it is usually parameterized as

$$\alpha(t) = \alpha(0) + \alpha'(0)\, t \tag{2.12}$$

where $t = m^2$ and $\alpha(0)$ is called *Regge intercept*. From these trajectories and based on general assumptions such as unitarity and maximum analyticity of scattering amplitude, the framework of Regge theory of hadronic scattering is formulated and an efficient parameterization of hadronic cross sections at intermediate energies is obtained. The elastic scattering

between particles a and b by the exchange of particles of the Regge trajectory α_k is given by the Regge amplitude

$$A_k(s,t) = \left(-\frac{1+\tau e^{-i\pi\alpha_k(t)}}{\sin(\pi\alpha_k(t))} \right) \beta_{a,k}(t)\beta_{b,k}(t) \left(\frac{s}{s_0}\right)^{\alpha_k(t)} \tag{2.13}$$

where τ is the parity of the Regge trajectory and the first bracket is called the *signature factor* of the Regge trajectory k, which determines whether the contribution to the total cross section of the Regge term is positive or negative. A_k describes the contributions of all exchanged particles of a Regge trajectory, which is equivalent to the exchange of a quasi-particle of non-integer spin, called *reggeon*, with the angular momentum α. The functions $\beta_{a,k}(t)$ and $\beta_{b,k}(t)$ can be interpreted as coupling constants of the reggeon k to the incoming particles a and b. Fig. 2.3 shows diagrams of Reggeon exchange.

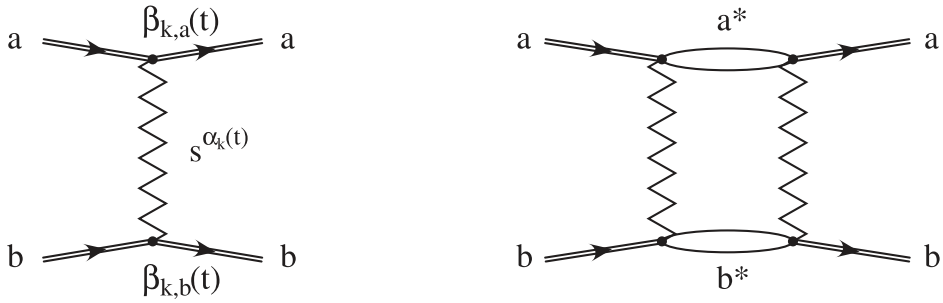


Figure 2.3: Reggeon exchange diagrams [83]. The Regge amplitude is interpreted as a coupling constant together with a reggeon propagator (left) or multiple reggeon exchanges (right).

The differential elastic cross section can be written as

$$\frac{d\sigma}{dt} = \frac{1}{16\pi} \left| \sum_k A_k(s,t) \frac{s_0}{s} \right|^2 . \tag{2.14}$$

The energy dependence of the total cross section at high energy is determined by the largest Regge intercept $\alpha_k(0)$

$$\sigma_{tot} = \frac{1}{s} \lim_{t\to 0} \mathrm{Im} \sum_k A_k(s,t) \propto \beta_{a,k}(t)\ \beta_{b,k}(t)\ s^{\alpha_k(0)-1}. \tag{2.15}$$

The intercepts of all known Regge trajectories are $\alpha_k(0) \ll 1$. The existence of another Regge trajectory with intercept $\alpha(0) \approx 1$ was postulated by Pomeranchuk. The corresponding quasi-particle is called *pomeron*. The pomeron has vacuum quantum numbers, which means that it couples to all hadrons in a similar way. It is assumed that *glueballs* are the bound states of the pomeron trajectory but experimental searches for glueballs have been inconclusive until now. The parameters of the pomeron trajectory are estimated from cross section data at high energy [83]:

$$\alpha_P(t) \approx \alpha_P(0) + \alpha'_P(0)t \approx 1.08 + 0.25\ \mathrm{GeV}^{-2}\ t \tag{2.16}$$

The Regge theory allows a simple parameterization of a large variety of hadronic cross sections at intermediate and high energy [84]:

$$\sigma_{\mathrm{tot}}(s) = Xs^{\epsilon} + Ys^{-\eta} \tag{2.17}$$

with exponents $\epsilon = 0.0808$ and $\eta = 0.4525$, corresponding to pomerons and reggeons respectively, and particle-dependent constants X and Y. The total elastic cross section is then given by

$$\sigma_{\text{ela}} \approx (1+\rho^2)\frac{\sigma_{\text{tot}}^2}{16\pi B_{\text{ela}}} \tag{2.18}$$

where ρ is the ratio of the real part to the imaginary part of the signature factor in eq. 2.13 for $t \to 0$. Assuming a Gaussian profile for the impact parameter amplitude:

$$\text{Im } a(b=0,s) = \frac{\sigma_{\text{tot}}}{8\pi B_{\text{ela}}} = \frac{2}{1+\rho^2}\frac{\sigma_{\text{ela}}}{\sigma_{\text{tot}}}. \tag{2.19}$$

In the 1970s, motivated by the observation of an approximately constant ratio $\sigma_{\text{ela}}/\sigma_{\text{tot}}$ at intermediate energies ($10 \lesssim \sqrt{s} \lesssim 100$ GeV), it was thought that the scattering amplitude satisfies geometric scaling [85]. Consequently, one gets $\sigma_{\text{ine}} \propto \sigma_{\text{tot}}$. It is now known that the opacity of the scattering amplitude is increasing and latest measurements show that the black disk limit has been reached for central collisions at LHC energies.

The simple parameterizations of cross sections given by eq. 2.17 work well at intermediate and high energies, but cannot be extrapolated to very high energies. This is due to the unitarity constraint on individual partial waves in the model of the Gaussian impact parameter amplitude. So, it is expected that corrections become important well before this bound is saturated. In terms of Regge theory this means that multi-pomeron exchange becomes an important factor.

An intuitive understanding of the expected high-energy behavior of hadronic cross sections can be developed with a generic model for the impact parameter amplitude that accounts for the short-range character of hadronic interactions. An upper bound estimation of the cross section at high energy can be calculated considering that any elementary interaction amplitude cannot grow with energy faster than a power law. A rigorous calculation in scattering theory gives

$$\sigma_{\text{tot}} \leq \frac{\pi}{m_\pi^2}\ln^2(s/s_0), \tag{2.20}$$

which is known as the *Froissart* or *Froissart–Martin bound* [86,87]. Here $s_0 = m_\pi^2\sqrt{2}/(17\pi^{3/2})$ is an energy scale [88]. The total p-p and p-$\bar{p}$ cross sections measured rise proportionally to $\ln^2 s$, but they are numerically lower than the Froissart bound [89].

2.4 Particle production mechanism

Particle production mechanisms are different at low, intermediate and high interaction energies. Interactions at very low energy, in the called *resonance region*, are dominated by the formation and subsequent decay of hadronic resonances. The energy range goes from the particle production threshold to about $\sqrt{s} \lesssim 2$ GeV.

The cross sections and distributions of secondary particles can be described by isobar models [90], which are based on conservation of isospin and angular momentum. An advantage of isobar models is the detailed description of the hadronic final state as a superposition of different decay distributions. However, the number of resonances that should be included in the calculations grows very fast with energy, so that an extension to higher energies is not feasible.

At higher energies, follows the so-called *region of scaling* and extends to $\sqrt{s} \lesssim 100$ GeV. The scaling law proposed by Feynman says that the differential cross sections in hadron collisions at high energies satisfy the so-called *Feynman scaling*:

$$E\frac{d^3\sigma}{d^3\bar{p}} = f(x_F, p_\perp) \tag{2.21}$$

This scaling law is violated at $x_F \simeq 0$ but it is expected to be valid at $x_F \simeq 1$ in the so-called *fragmentation region*. One important characteristic of interactions at these energies is the *leading particle effect*. Almost half of the total momentum is carried by the most energetic secondary particle. This leading particle also has the same or almost the same quantum numbers as the interacting particle.

The Regge theory can be used to calculate cross sections for a planar interaction topology, which corresponds to the exchange of a meson. However, a pomeron exchange implies a cylinder topology since the pomeron interacts with all hadrons with a similar strength. Fig. 2.4 (left) shows these topologies. The optical theorem relates the color flow topology of the elastic scattering amplitude to that of inelastic final states. Taking the imaginary part of the amplitude corresponds to putting particle propagators on mass shell. This is typically visualized as a *unitarity cut*, see Fig. 2.4 (right). Unitarity cuts of reggeon and pomeron amplitudes lead to one and two chains of hadrons in the final state, respectively. On the basis of the phenomenology described here, the very successful *dual parton model* [91] (DPM) and the *quark gluon string model* [92] (QGS) were constructed.

These hadron chains are taken to be the fragmentation products of color strings in Monte Carlo event generators. The only unknown quantity is then the flavor and the momentum distribution of the quarks and diquarks at the string ends, for which different assumptions are made in different models.

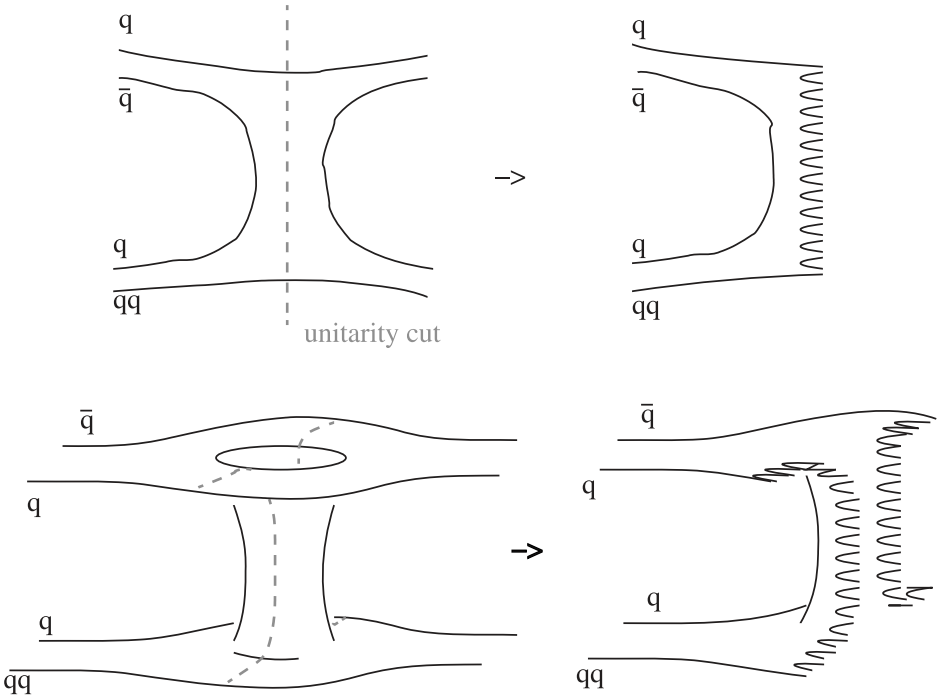


Figure 2.4: Color flow topologies of elastic scattering due to reggeon (top left) and pomeron (bottom left) exchange and the expected particle distributions (right) for inelastic interactions [83].

At energies higher than $\sqrt{s} \gtrsim 100$ GeV, hadronic interactions are most efficiently described by perturbative QCD in terms of asymptotically free partons and their interactions.

The production of partonic jets of a few GeV in transverse momentum, called *minijets*, becomes the dominating process. In the perturbative picture, sea partons in a hadron are quantum fluctuations with an average lifetime larger than the typical hadronic interaction time. So, an interacting hadron can be considered as a frozen-in configuration of independently acting partons as long as the parton virtuality is not too high.

In this context, it is necessary to extend the definition of the cross section to particle distributions for the beam and the target hadrons. The impact parameter amplitude can be written using the *eikonal function* $\chi(b)$:

$$a(b,s) = \frac{i}{2}(1 - e^{-\chi(b)}) \tag{2.22}$$

In the Monte Carlo models EPOS [93,94] and QGSJet [95,96], the expansion of the impact parameter amplitude of eq. 2.22 is used to identify the different terms explicitly with multi-pomeron exchange amplitudes that include highly virtual states that correspond to hard scattering after unitarity cuts are applied. Multi-pomeron amplitudes can be calculated within Gribov's reggeon field theory [97], leading to eikonal-like expressions [98].

Another approach is followed by the Monte Carlo event generator Sibyll [99] where different interaction processes have to be added up linearly in the eikonal function

$$\chi(b,s) = \frac{1}{2}(\sigma_{inc,1}A_1(b) + \sigma_{inc,2}A_2(b) + \ldots) \tag{2.23}$$

2.5 Phase diagram for QCD matter

Quantum Chromodynamics (QCD) is a solid quantum field theory of the strong interactions and essential in the understanding of a large variety of phenomena at hadron colliders as the LHC. The description of proton-proton interactions at LHC energies is widely accepted, however it is not clear how nucleus-nucleus interactions should be described taking into account the knowledge of small systems. For example, hadronization and jet formation do not necessarily have to be same. In fact, the physics of heavy-ion collisions at relativistic energies explore regimes of QCD under extreme conditions where the matter is highly compressed and becomes a thermal bath. In this situation it can still be ordinary nuclear matter or phase transition into new states of matter.

QCD describes the strong interactions in an $SU(3)_c$ gauge theory framework, where its strength α_s runs with the energy scale according to a wide range of experiments as in Fig. 2.5. It is formulated in terms of elementary fields, quarks and gluons, whose interactions obey the principles of a relativistic quantum field theory (QFT) with non-abelian gauge invariance. The color potential between quarks is particular because it increases linearly with distance. Consequently, hadrons combine to zero net color charge hadrons given the so-called *confinement of quarks*. This confinement can be seen as a direct consequence of the gluon self-interaction.

The strength of the strong interactions is characterized by the coupling constant $\alpha_s = g_s^2/4\pi$. Its running can be shown by experiments, asking that all relevant observables $\mathcal{O}(Q^2, \alpha_s)$ constructed with QCD should be renormalization scale independent. This is satisfied by the Renormalization Group Equation:

$$\left(\mu^2 \frac{\partial}{\partial \mu^2} + \beta(\alpha_s)\right) \mathcal{O}(\alpha_s, Q^2/\mu^2) = 0 \tag{2.24}$$

where the beta function is defined as

$$\beta(\alpha_s) = \frac{d\,\alpha_s}{d\,log\,Q^2/\mu^2}. \tag{2.25}$$

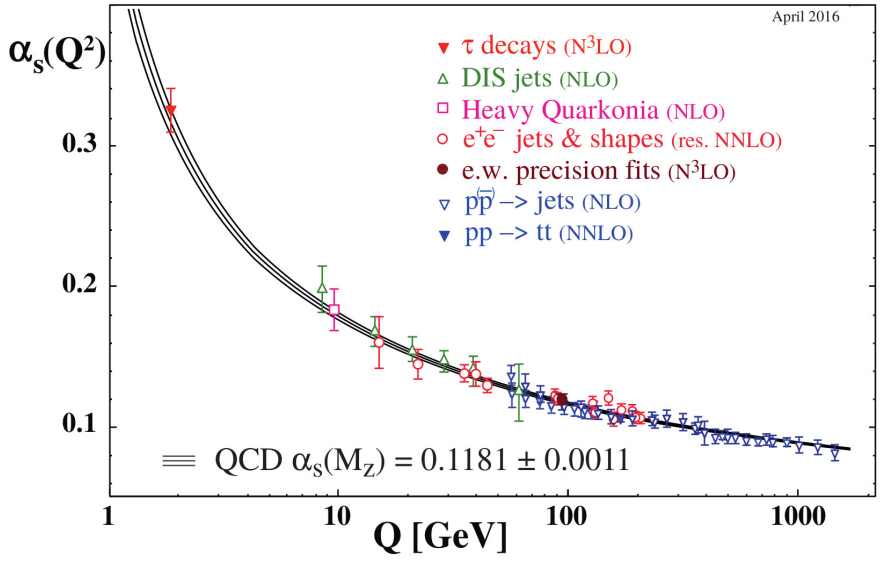


Figure 2.5: Measurements of α_s as a function of the energy scale Q [100]. The levels of QCD perturbation theory used to calculate α_s by each experiment is indicated in brackets.

The beta function can be obtained perturbatively in the high energy regime, where coupling values are small. At lowest order,

$$\beta(\alpha_s) = -b_0\alpha_s^2 \tag{2.26}$$

with $b_0 = \frac{1}{2\pi}\frac{11}{6}N_C - \frac{1}{3}N_f$, $N_c = 3$ and N_f are the number of color and flavour degrees of freedom of the theory, respectively. Eqs. 2.25 and 2.26 can be solved, since $b_0 > 0$, giving the famous *running* of α_s:

$$\alpha_s(\mu^2) = \frac{1}{b_0 log(\mu^2/\Lambda^2)} \tag{2.27}$$

The parameter Λ describes the boundary condition of the first order differential equation and corresponds to the scale at which the coupling becomes infinity. Since α_s is not an observable, it can contain all the terms that are μ dependent, in order to achieve a μ-independent observable $\mathcal{O}$ that has a power-series representation in terms of α_s .

If Λ_{QCD} is defined as a small enough transferred momentum such that the coupling goes to infinity, then

$$\Lambda_{QCD}^2 = \mu^2 e^{-\frac{1}{b_0\alpha_s(\mu^2)}} \tag{2.28}$$

where Λ_{QCD} is the intrinsic scale in QCD and a renormalization scheme dependent quantity. In the $\overline{MS}$ scheme and for three active flavors, its value is of order $\Lambda_{QCD} \sim 200$ - 300 MeV.

In heavy-ion collisions at relativistic energies the confinement of quarks is broken becoming a medium of thermally equilibrated hadronic matter where now the quarks and gluons can move freely. The main parameters that describe this new behavior are the temperature T and the baryon number density n_B or its conjugate variable, the baryon chemical potential μ_b. The partons-to-hadrons phase transition is expected around $T \simeq \Lambda_{QCD} \sim 200$ MeV and $n_B \sim \Lambda_{QCD}^3 \sim 1\text{fm}^{-3}$. The phase diagram for QCD matter is shown in Fig. 2.6-left, where the Hadron Gas phase and the Quark-Gluon plasma phase are separated by a phase transition line that ends at the *critical end point* (CEP). Going down to smaller μ_b values the transition disappears and it becomes a continuous crossover around Λ_{QCD}.

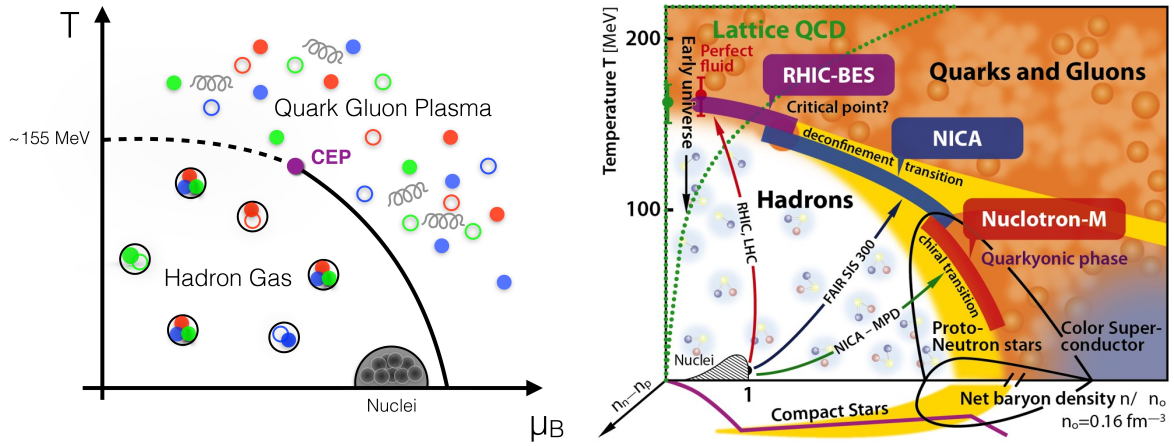


Figure 2.6: Phase diagrams for QCD matter. Left: in the temperature (T) and baryon chemical potential (μ_B) plane [101]. Right: in the temperature (T) and baryon density (n) normalized to the cold nuclei baryon density n_0 plane [102].

Phase transition and CEP location studies are not possible at RHIC and LHC since the QGP created at high energies contains almost equal amounts of matter and antimatter. So the corresponding experimental scan stays close to the vertical axis: low μ_b or n_B and high T. For that reason, colliders as FAIR and NICA are designed to create a QGP with excess of matter over antimatter to explore the bulk of the phase diagram. Taking the last heavy-ion collision experiments and results from effective models and lattice QCD calculations into account, the phase diagram becomes more complex as is shown in Fig. 2.6(right), where there is a new phase called *quarkyonic phase*. These three phases are separated by a chiral phase transition, a deconfinement transition and a pseudocritical crossover line.

2.6 Relativistic heavy-ion collisions

Heavy-ion collisions at relativistic energies can be thought of as the collision of two thinning disks due to the Lorentz contraction. Since the volume of nuclei is proportional to the mass number, the nuclei length over the motion direction at high energies is of the order of $A^{1/3}/\gamma$ Fm. Typical sizes of Pb and Au nuclei are 14 fm and the γ factor is of the order of 100 at RHIC and 2500 at LHC. So, a nuclei collision is in fact the collision of two disks of size 0.14 Fm at RHIC and 0.006 Fm at LHC. When the Lorentz contracted heavy-ions collision occurs, a region of high temperature and energy density is created. Then the system expands and cools down, eventually fragmenting into hadrons. Fig. 2.7 shows the stages of a heavy-ion collision. The interacting heavy ions are thin discs of color fields and have on average more quarks than antiquarks with color charges, so that they become sources of colored gluons. When the nuclei pass through each other, long color fields fill the space between the receding two Lorentz contracted ions, which makes them loose energy and then they gradually decay into q-$\bar{q}$ pairs and gluons.

According to lattice QCD, QCD matter in thermal equilibrium at $T \approx 300$ MeV has an energy density $\epsilon \approx 12\,T^4 = 12.7\,\mathrm{GeV/fm^3}$. On the other hand, in heavy-ion collisions $\langle\epsilon\rangle \approx 12\,\mathrm{GeV/fm} \approx 20\epsilon_{hadron}$ at the LHC with $\sqrt{(s_{NN})} = 2.76$ TeV. This implies that the medium formed after the collisions cannot be described just as a collection of distinct individual hadrons, rather it has to be made out of a high density of quarks and gluons. Another feature is the quick increase of entropy after the collision, since final states can

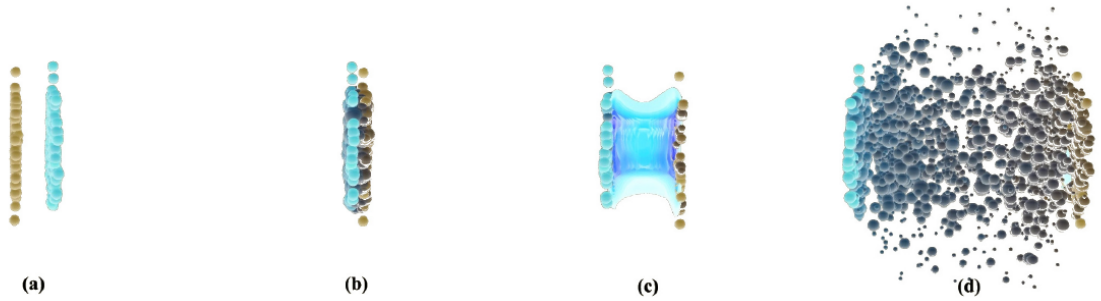


Figure 2.7: Stages of a heavy-ion collision [103]: (a) the Lorentz contracted heavy ions move towards each other; (b) they travel across each other; (c) a volume of high temperature and energy density is generated; (d) the system expands and cools down, eventually fragmenting into hadrons

be formed with more than 10^4 particles while initial particles have almost null entropy. Furthermore, the ions may collide head-on or may only partially overlap at the collision stage as in shown Fig. 2.8, so in the overlap region there are conditions that facilitate QGP formation.

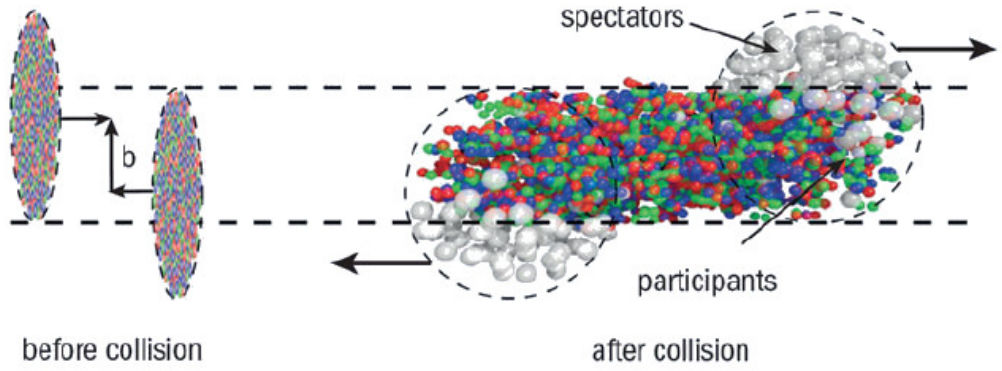


Figure 2.8: A schematic representation of the collision centrality [103].

2.7 Quark-Gluon-Plasma (QGP)

The QGP is made of strongly coupled quarks and gluons. They form a collective medium that expands and flows as a relativistic hydrodynamic fluid, with a low viscosity to entropy density ratio $\eta/s \geq 1/4\pi$. In heavy-ion collisions (see Sec. 2.6), a QGP fluid is formed and it flows hydrodynamically. In the last decade, the nuclear physics community has converged towards an initial consensus on a model with marked epochs in the time evolution of heavy-ion collisions. The main features of this model of the stages in the evolution of heavy-ion collisions are as follows [103, 104]

1. The Lorentz-contracted nuclei collide in a short time $\Delta t \ll 1$ fm/c. The energy deposited into the medium mainly through gluon field interactions, creates an inhomogeneous initial condition in the transverse plane.

2. Matter expands at almost the speed of light trying to reach the equilibrium.

3. When matter is almost in equilibrium, it behaves like a fluid and shows collective modes. Viscous relativistic hydrodynamics and an equation of state from lattice QCD is commonly used to describe the QGP. The small deviations from equilibrium happen because of the shear/bulk viscous medium.

4. The fluid temperature decreases and when the cross-over temperature of T $\sim$ 170 MeV is reached, the fluid breaks up into hadrons.

5. Finally, the hadrons scatter inelastically until no more decays or secondary production is possible, reaching a chemical freeze-out. Particles continue to scatter elastically until their momentum distribution is set, reaching the kinetic freeze- out. Now particles can be considered in a final-state with momenta as measured experimentally.

This model of particle production in heavy-ion collisions with an intermediate epoch during which a hydrodynamic fluid forms and expands, is quite different from the ones used for particle production in elementary hadron collisions in which only a few new particles are created. In fact, the question of whether QGP is created in p + p collisions has caused a big excitement and the possibility to have collective effects in relativistic collisions of small systems is now being pursued with both theoretical and experimental approaches. See Ch. 4 for more details about it.

2.8 QCD Topological expansion and String Fragmentation Models

Soft hadronic interactions are characterized by momentum transfers $Q \lesssim \Lambda_{QCD}$. In these cases, perturbative calculations in powers of α_s cannot be used since the coupling constant is of the order of unity. This led to the development of different approaches to understand non-perturbative phenomena in QCD.

In 1970s, 't Hooft analyzed the importance of different diagrams in QCD for processes with low momentum transfer but in the limit $N_c \to \infty$ and $g_s^2 N_c$=const with N_c the number of colors and g_s the coupling constant [105]. The extension of this approach where $N_c \to \infty$ and N_c/N_f=const, with N_f the number of flavors, is called *topological expansion of QCD*. In this limit, quarks are represented by a color line and gluons by two color lines, and diagrams can be drawn without crossing lines. The topological importance of these diagrams is related to the genus of the surface. Diagrams with the lowest genus are the leading terms in the expansion and each increase of the genus is suppressed by a factor of $1/N_c^2$. This classification of diagrams is expected to be approximately valid for $N_c = 3$.

In the topological expansion of QCD, the effective degrees of freedom in interactions are described by *strings*. An open string connects two valence quarks and represents a meson. The movement of these strings in space-time produces two-dimensional surfaces because planar diagrams are dominating. Fig. 2.9 shows some examples. The description of baryons is not straightforward [106]. However, a simple approach is the *diquark model* where the valence partons of a baryon are taken to be a quark and a diquark.

Interaction and propagation of quarks and gluons are understood for large momentum transfers within perturbative QCD. However, their transition to hadronic particles in the final state of a process is a non-perturbative process called *hadronization*. Some approaches to describe hadronization for particular cases use the asumption of local parton–hadron duality [107] or universal fragmentation functions. Monte Carlo methods can be used within phenomenological fragmentation models giving a better description of final states. A very successful model of this kind is the *Lund string fragmentation model* [108], which is implemented in Monte Carlo event generators for hadronic interactions of cosmic-ray air showers. The basic principle of the string fragmentation and hadronization is shown in Fig. 2.10.

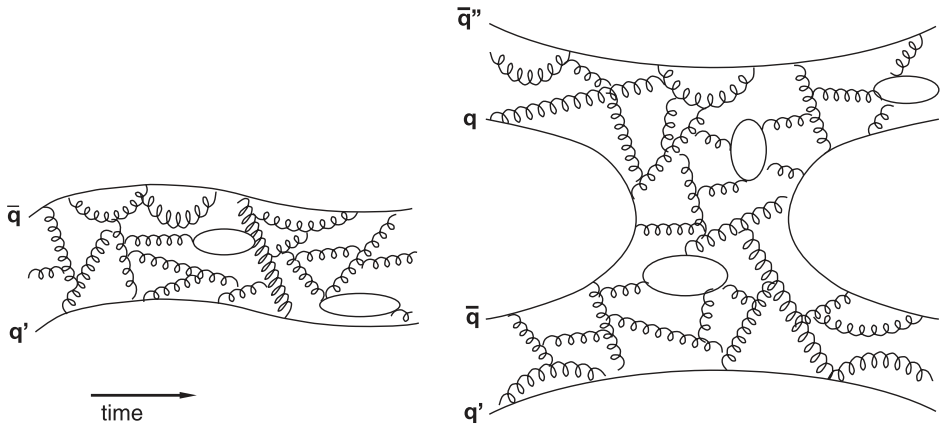


Figure 2.9: Examples of planar diagrams [83]. Left: meson propagation in time. Right: meson–meson interaction.

The kinetic energy of a high energy quark–antiquark pair is converted to energy stored in the color string stretched between the two particles, which form a color-neutral system. Quantum fluctuations, as quark–antiquark pairs, lead to a break-up of the string once the stored energy density is high enough. Baryon–antibaryon pairs are also produced in string break-ups involving diquarks. The process of subsequent string break-ups continues until the energy stored in the strings is too small to materialize further quantum fluctuations. The remaining color-neutral objects are then identified as hadrons and hadronic resonances according to their flavor content and invariant mass.

A different approach is chosen in *statistical fragmentation or hadronization models* [109,110] (see Sect. 2.9). In this model, final state hadrons are assumed to originate from hot fireballs of a hadron gas at thermal and partial chemical equilibrium. The hadron abundances and other interactions can be described using canonical statistics with a small number of universal parameters, one of which is the critical transition temperature $T \approx 170\,\text{MeV}$, which depends on the energy density produced in the collisions.

2.9 Statistical Hadronization Model

The study of hadron production by the Statistical Hadronization Model (SHM) is motivated by its application to heavy ion collisions. The picture of a high energy collision in this model is that of a non-perturbative QCD process, eventually giving rise to the formation of massive colorless extended objects called *clusters* or *fireballs*. These clusters are formed at some critical energy density or another parameter, to decay later coherently into multihadronic states [111]. The decay rate into any state compatible with the cluster quantum numbers is only given by its phase space, in the so-called *phase space dominance*. So, the clusters decay statistically into hadrons, since all accessible states are equally likely.

The statistical equilibrium would be the effect of equal quantum transition probabilities from a cluster to all accessible final states instead of a collisional thermalization process between formed hadrons. A distinctive feature of the statistical model in comparison with other cluster models is that clusters have a finite spatial extension. Probably, the best known model with relativistic extended massive objects is the bag model [112] and indeed the SHM can be considered as a model for the strong decays of bags.

The statistical ensemble to use in a single cluster hadronization should be the *microcanonical* one. This ensemble is defined as the set of states with fixed energy-momenta, angular

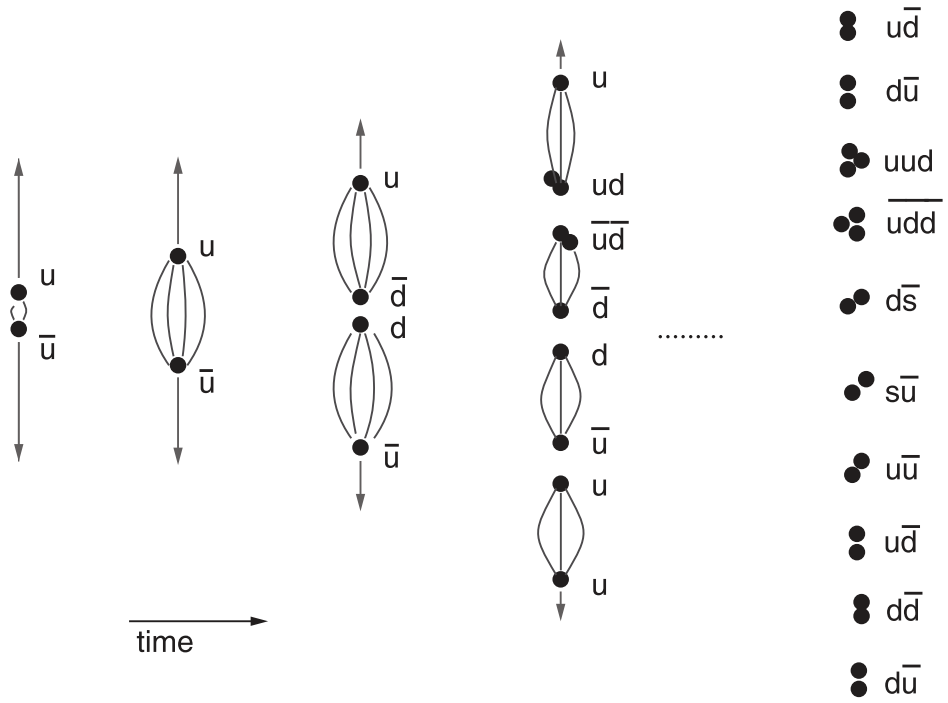


Figure 2.10: Illustration of string fragmentation of a quark–antiquark pair [83].

momenta, parity and internal charges. Although, the set of states with only fixed energy-momentum and internal charges is commonly called with the same name.

However, almost all data analysis using the statistical model has been done in the framework of the *canonical* or the *grand-canonical* ensemble. In these ensembles the conservation of energy-momentum and angular momentum are taken only on average, and the source is described in terms of a temperature. The reason for their use is that in high energy collisions many clusters are produced and their fluctuations of masses and volumes reduce the importance of exact conservation of energy and momentum. The canonical ensemble is a better approximation of the microcanonical one for larger values of mass and volume of the cluster. But their values are taken just to be in agreement with the data. On the other hand, an explicit calculation of the microcanonical ensemble is needed for energies close to the resonance region (see Sect. 2.4), where conservation laws have a larger impact.

In the high energy physics community, there are several discussions about the meaning of the statistical model. The main reasons (given in [113]) are related to the use of thermodynamical quantities, as the temperature, even in systems with small volumes and masses, or the fact that statistical equilibrium is not derivable from QCD, indeed when the hadronization temperature is really close to the estimated critical temperature of QCD. One of the fundamental assumptions in the SHM is that the statistical equilibrium must be an inherent property of hadronization itself instead of a thermal process after the hadronization through inelastic collisions. The results of the statistical model can be obtained from other models but using supplementary assumptions or the use of many free parameters.

Moreover, the use of the *statistical* and *thermal* words as synonymous is under discussion. In [113] the author argues that the word *statistical* means equal probability in phase space, which is appropriately measured with $d^3x d^3p$ for any particle and a volume is involved, then *statistical* and *thermal* should be taken as synonymous. But other authors [114] make a clear distinction between the temperature determined in the SHM by fitting particle abundances and a *proper* temperature which would be achieved through inelastic reinteractions of formed particles.

2.9.1 SHM and Data

Among all heavy-ion collision hadronization models, SHM is a unique phenomenological approach linking the production of hadrons to the QCD phase diagram [115]. The model is based on simple assumptions, but is very powerful considering its small set of parameters, i.e. temperature T, baryochemical potential μ_B and cluster volume V. In order to make calculations, all hadron states should be taken into account. All currently known hadrons are considered, but missing states could in principle still have an effect on T.

In the commonly used grand canonical approach, chemical potentials μ_B ensure conservation of the average of additive quantum numbers such as baryon number, isospin and strangeness, which are fixed by initial conditions. The standard implementation of the model was successfully used to describe hadron production in heavy-ion collisions over a wide range of collision energies (for example [116,117]). Several versions with an extended set of parameters have been proposed [118, 119]. In particular, the possible departure of equilibrium of hadrons containing strange quarks [117] and the flavor-dependent freeze-out [120–122] were studied. Also the possible repulsive interactions among hadrons are modeled in a hard-sphere excluded-volume approach. However, the measurements do not seem to require any of these extensions [123]. For small systems and/or low energies, a canonical treatment is needed [124], usually implemented only for strangeness. Such recent studies in p–Nb and Ar–KCl collisions [125] and in pp collisions [126–128] lead to values of T comparable to (or even larger than) those in (central) Au–Au or Pb–Pb collisions. The studies performed by ALICE in p–Pb [129] and $p-p$ [130] collisions revealed that in high-multiplicity events hadron production in these small systems resembles that in Pb–Pb collisions.

In [123] a model based on the statistical operator for the hadron resonance gas is used and it leads to an equation of state in good agreement with lattice QCD calculations. This model describes the LHC data in central (0-10%) Pb–Pb collisions at s = 2.76 TeV as Fig. 2.11 shows. The best fit values and their uncertainties are: T = 156.5 $\pm$ 1.5 MeV, μ_B = 0.7 $\pm$ 3.8 MeV, $V_{\Delta y=1}$ = 5280 $\pm$ 410 fm^3.

The value of the (pseudo-)critical temperature, T_c , at vanishing baryochemical potential (μ_B) is currently calculated in lattice QCD [131,132] to be 155$\pm$9 MeV. Remarkably, at low μ_B chemical freeze-out coincides with T_c, indicating hadron formation from deconfined matter. The agreement between the results from several independent analyses [133–136] is very good.

2.10 The core-corona effect in heavy ion interactions

Experimental data at RHIC for $\sqrt{s} = 130$ GeV and 200GeV demonstrated the formation of a thermalized system in Au + Au collisions, since it was found that thermal spectra with collective transverse flow component are well described by ideal fluid dynamics. The particle production per participant in heavy-ion collisions at RHIC energies was found significantly larger than in p+p or d+Au collisions [137–140], and it deviates from a simple superposition of independent production from all participant nucleons. That was no surprise, since for a high enough density of elementary collisions a phase transition with a complete chemical equilibrium to the quark-gluon plasma was expected, which would behave as an ideal fluid if the plasma was strongly interacting near the critical temperature. In the boundaries of the interaction region the density of participants and therefore the density of the deposited energy is not high enough to have a fast thermalization. These results of heavy-ion collisions at RHIC energies drove to the picture of a thermalized source in the region with the largest density, the so-called *core* and outer *mantle* or *corona*.

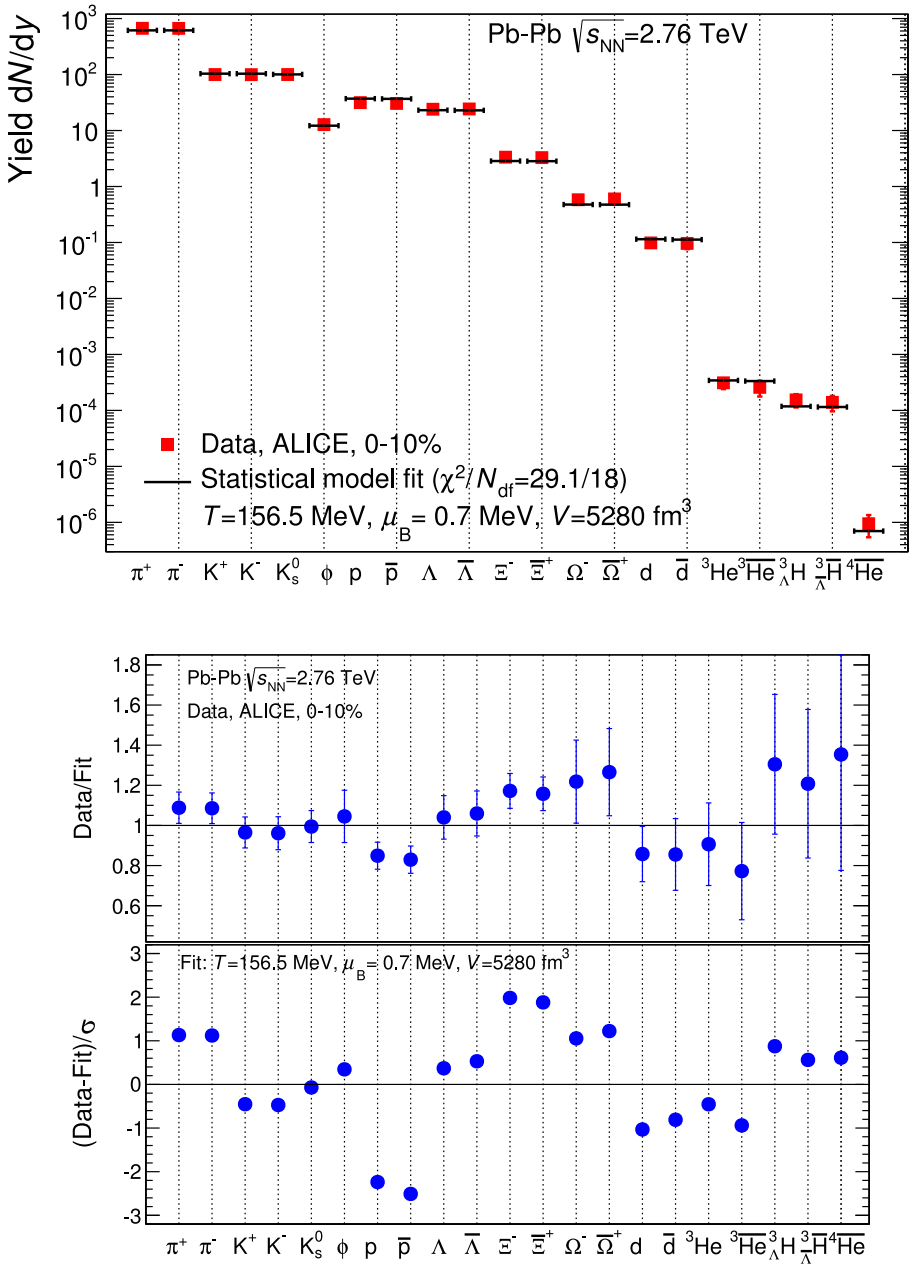


Figure 2.11: Hadron multiplicities in central (0-10%) Pb–Pb collisions at the LHC and best fit [123]. The lower panels show the ratio of data values to the fit and the difference between data and the model fit in units of the experimental uncertainty.

The first study with this approach was done by Bożek [141] who analyzed the core dependence on the centrality of collisions. Later, Werner [142] parameterized the core-corona contributions with global parameters so as to reproduce the density of charged particles, based on string density before hadronization within the EPOS model. The effects of the core-corona separation were also estimated using the UrQMD hybrid model [143]. From these initial works, further studies using the core-corona approach were done: The strangeness production at SPS and RHIC and their centrality dependence were well described in a core-corona model framework [144–147]. Also the centrality dependence of the elliptic flow was quantitatively described for SPS to RHIC energies in a very simple core-corona model [148–150]. In addition, another recent core-corona approach was implemented in the *dynamical core–corona*

initialization (DCCI) model [151], where the fluids that are in local thermal and chemical equilibrium form the core, while the corona is the system of nonequilibrated partons traversing the fluids or the vacuum.

The core-corona model was fully implemented in the event generator EPOS 3 [152], where small and large systems are described by a hydrodynamical evolution. In that work the simulated transverse momentum p_t distribution of several kinds of hadrons are mostly in good agreement with collider measurements as ALICE, ATLAS and CMS. In a further work [153], resonance production was studied for different collision systems. As an example, the contribution of the core-corona to the particle ratio as a function of multiplicity is shown in Fig. 2.12. The use of the core-corona model improves the agreement with ALICE data considerably. It is worth noting that the core yield and corona yield do not depend on multiplicity. What changes is the fraction of particles produced by the core and the corona leading to a smooth transition from corona yield to core yield with multiplicity. This feature is taken into account in the core-corona approach presented in Sect. 4.3 using the tool presented in Sect. 3.3.

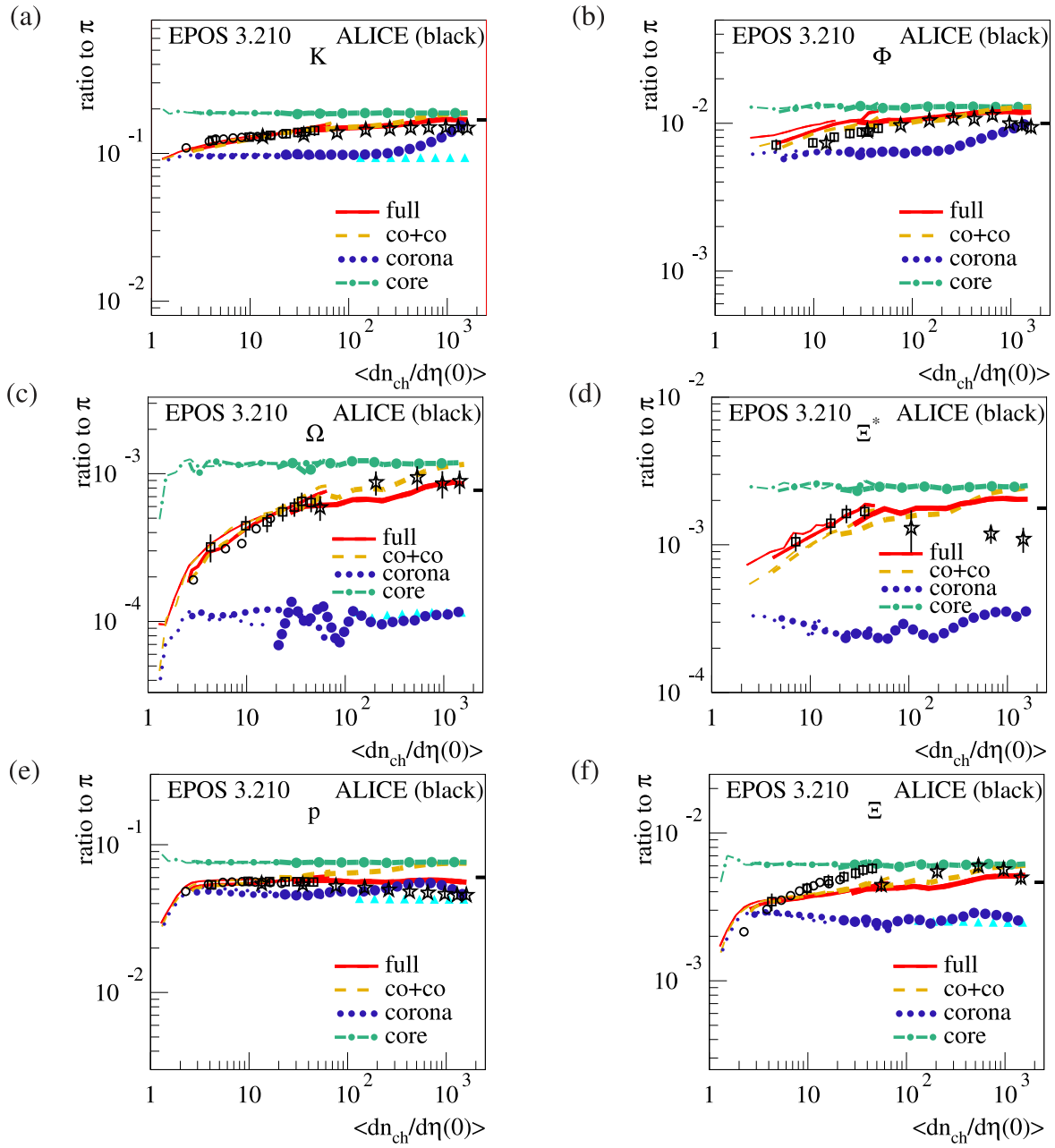


Figure 2.12: Particle to pion ratio for different particle species versus multiplicity, for different contributions for EPOS 3.210 simulations and for the systems pp, pA, AA (slim, medium and thin lines) [153]. ALICE data is also plotted for the same systems pp, pA, AA (open circles, open squares, open stars). The contribution *co-co* is for core-corona and *full* is the core-corona with hadron-hadron rescatterings.

CHAPTER 3

Cascade Equations and CONEX

Nowadays, Monte Carlo (MC) simulation is the most common method used to describe in detail the extensive air shower (EAS) development. However, full MC simulations of ultra-high energy EAS require very large computing time. Some interesting tools have been developed to deal with this problem. One of them is the so-called *thinning* algorithm [154] which follows a small number of particles and assigns each particle a weight factor. Maximum weights must be constrained to avoid significant artificial fluctuations of EAS observables, so this method at high energies still needs a considerable computing time.

A different approach is to describe the EAS development numerically, based on the solutions of the corresponding cascade equations (CE) [155–157]. Shower simulation with an explicit MC treatment for high energy particles and a numerical development for low energy particles gives accurate results for average showers and their fluctuations. The air shower simulation program *CONEX* [158] implements this hybrid approach. The MC treatment for particles with energies above a given threshold is carried out in the standard way similar to the implementation in CORSIKA [159]. The numerical description of lower energy sub-cascades is based on the solution of hadronic and electomagnetic CE. This type of air shower calculation is ideal to test modifications of hadronic properties to study the muon puzzle because it is fast and modifications can be done at all energies and not only in the first interaction.

There are two kinds of CE necessary to describe EAS: one of them is hadronic and the other one is electromagnetic (EM). The particles considered in the hadronic CE by CONEX, i.e. their interactions and propagation, are protons, neutrons, charged pions, charged and neutral kaons. In this context, these particles are called *projectile particles*. The only species of particles produced in a projectile-air interaction are protons, neutrons, charged pions, charged and neutral kaons, photons, muons and electrons. These produced particles are called *secondary particles*. Other kinds of hadrons produced by decay or through interaction are assumed to decay immediately into secondary particles. Muons are treated as hadrons, but all possible high energy interactions are neglected in cascade equations. When photons and electrons are produced, they are transferred directly to the EM CE.

Air showers are simulated by CONEX as follows: The hadronic and EM cascades are simulated by an explicit MC, until all produced secondaries have an energy below the threshold E_{thr}. All sub-threshold particles form the *source terms* of the CE, giving the initial conditions for the numerical analysis. The hadronic and EM cascades at energies below E_{thr} are calculated by solving their corresponding CE for each depth level in a sequential order. First, the hadronic cascade is solved starting at the first depth level with the initial conditions given

by the source terms. The result of this is a discretized energy spectrum of all hadrons for the next depth level. EM particles are also produced, which are added to the corresponding EM source term and then the EM CE are solved. All hadrons produced in the EM cascade by photonuclear interaction and all pair-produced muons are added to the hadronic source term at its next depth level. This procedure is repeated for the following depth levels, each time using the hadronic and EM source terms of the previous level.

3.1 Hadronic cascade equations

The hadronic cascade equations take account of all possible processes in the following integro-differential equation:

$$\begin{aligned}\frac{\partial h_a(E,X)}{\partial X} = &- \frac{h_a(E,X)}{\lambda_a(E)} - \frac{h_a(E,X)}{\tau_a(E)\,\rho_{air}(X)} + \frac{\partial}{\partial E}\,\beta_a^{ion} h_a(E,X) \\ &+ \sum_d \int_E^{E_{max}} dE'\; h_d(E',X)\;\left(\frac{W_{d\to a}(E',E)}{\lambda_d(E')} + \frac{D_{d\to a}(E',E)}{\tau_d(E')\,\rho_{air}(X)}\right) \\ &+ S^{had}(E,X) \end{aligned} \tag{3.1}$$

where $h_a(E,X)$ is the differential energy spectrum of the hadron of type a, with energy E at depth position X along a given straight line trajectory, $\beta_a^{ion} = -dE_a/dX$ is the ionization energy loss per depth unit, $\lambda_a = m_{air}/\sigma_{inel}^{a-air}$ is the mean free path and τ_a is the life time in the laboratory system (related to the proper life time by $\tau_a = \tau^0 E/m$). $W_{d\to a}$ and $D_{d\to a}$ are the inclusive secondary spectra for interactions and decays respectively. Muons are treated as hadrons but without interaction term. The five terms in Eq. 3.1 describe the variation in the hadron number due to: interactions with air nuclei, particle decays, ionization loss, hadron production from higher energy parents and source terms, respectively. The particle decay term comes from the decay rate $\mathrm{d}h_a = -h_a \mathrm{d}L/\tau_a$ and $\mathrm{d}L/\mathrm{d}X = \rho_{air}^{-1}(X)$.

The source term $S_a^{had}(E,X)$ defines the initial conditions and is determined during the MC simulation of above-threshold particle cascading. It consists of contributions of all sub-threshold hadrons produced at that stage

$$S_a^{MC\to had}(E,X) = \sum_i \delta_{d_i}^a \delta(E-E_i)\delta(X-X_i) \tag{3.2}$$

with d_i, E_i, X_i being type, energy, and depth position of the source particle i. The photoproduction of hadrons is taken into account via a source term, where the particle production distributions are approximated by those of π^0-air interaction

$$S_a^{em\to had}(E,X) = \int_E^{E_{max}} dE'\; l_\gamma(E',X)\; W_{\pi^0\to a}(E',E)\; \tilde{\sigma}_\gamma^{\text{photonuclear}}(E') \tag{3.3}$$

where l_γ is the energy spectrum of photons. Moreover, the photoproduction of muon pairs gives another contribution to the hadronic source term:

$$S_a^{em\to\mu}(E,X) = \int_E^{E_{max}} dE'\; l_\gamma(E',X)\; W_{\gamma\to\mu}(E',E)\; \tilde{\sigma}_\gamma^{\mu-\text{pair}}(E'). \tag{3.4}$$

Therefore the total source term is given by

$$S_a^{had}(E,X) = S_a^{MC\to had}(E,X) + S_a^{em\to had}(E,X) + S_a^{em\to\mu}(E,X). \tag{3.5}$$

3.2 Electromagnetic cascade equations

The EM cascade development can be described by the following system of integro-differential equations

$$\begin{aligned}
\frac{\partial l_{e^-}(E,X)}{\partial X} &= -\tilde{\sigma}_{e^-}(E)\, l_{e^-}(E,X) + \frac{\partial}{\partial E}\,\beta^{ion}_{e^-} l_{e^-}(E,X) \\
&\quad + \int_E^{E_{max}} dE' \,\left[l_{e^-}(E',X)\, W_{e^-\to e^-}(E',E) + l_{e^+}(E',X)\, W_{e^+\to e^-}(E',E) \right. \\
&\quad \left. + l_\gamma(E',X)\, W_{\gamma\to e^-}(E',E)\right] + S^{em}_{e^-}(E,X) \qquad (3.6)\\
\frac{\partial l_{e^+}(E,X)}{\partial X} &= -\tilde{\sigma}_{e^+}(E)\, l_{e^+}(E,X) + \frac{\partial}{\partial E}\,\beta^{ion}_{e^+} l_{e^+}(E,X) \\
&\quad + \int_E^{E_{max}} dE' \,\left[l_{e^+}(E',X)\, W_{e^+\to e^+}(E',E) + l_\gamma(E',X)\, W_{\gamma\to e^+}(E',E)\right] \\
&\quad + S^{em}_{e^-}(E,X) \qquad (3.7)\\
\frac{\partial l_\gamma(E,X)}{\partial X} &= -\tilde{\sigma}_\gamma(E)\, l_\gamma(E,X) + \int_E^{E_{max}} dE' \,\left[l_{e^-}(E',X)\, W_{e^-\to\gamma}(E',E) \right. \\
&\quad \left. + l_{e^+}(E',X)\, W_{e^+\to\gamma}(E',E) + l_\gamma(E',X)\, W_{\gamma\to\gamma}(E',E)\right] + S^{em}_{\gamma}(E,X) \qquad (3.8)
\end{aligned}$$

where l_{e^-}, l_{e^+} and l_γ are the energy spectra of electrons, positrons and photons. The cross section in units of area/mass $\tilde{\sigma}_i = \sigma_{i-air}/m_{air}$ takes into account the particle interaction processes Bremsstrahlung, Bhabha, Moeller, annihilation, electron and muon pair production and photonuclear production and $W_{a\to b}$ are the inclusive production cross sections of secondary particles in units of area/(mass energy).

3.3 Tool development: Modification of the CONEX secondary particle spectra

In this section a technique that we developed to implement any type of modification of the hadronic interactions in the CONEX framework is presented. The goal is to study new solutions for the muon puzzle.

CONEX implements the latest updated high energy hadronic interaction models: EPOS LHC, QGSJetII.04 and SIBYLL 2.3d. For the numerical analysis, the same models are used to pre-calculate the spectra of secondary particles. This means that for each hadronic interaction model there is a spectrum for each secondary particle, for each projectile particle and for each projectile energy.

As already explained, CONEX considers in the cascade equation analysis five different projectile particles (p, n, $\pi^\pm$, K_l, and $K^\pm$) and eleven secondary particles (p, n, $\pi^\pm$, π^0, K_l, $K^\pm$, γ, μ^+, μ^-, e^+ and e^-). The projectile energy is discretized in 20 logarithmic energy bins per decade from 1 GeV to 10^{10} GeV. This means thus CONEX has 11000 secondary particle spectra for each high energy hadronic model. For example, Fig. 3.1(left) shows all the energy spectra of secondary particles for an interaction between a proton at 10^{19} eV and air. Fig. 3.1(right) shows the charged pion energy spectra for all proton energies in proton-air interactions for the EPOS-LHC model where each line corresponds to different proton energy.

Any change in these spectra has an impact on the air shower simulation means of the CE. So it is possible to implement modifications in particle interactions through changes in the secondary particle spectra. In order to do that, we developed a tool which allows to make

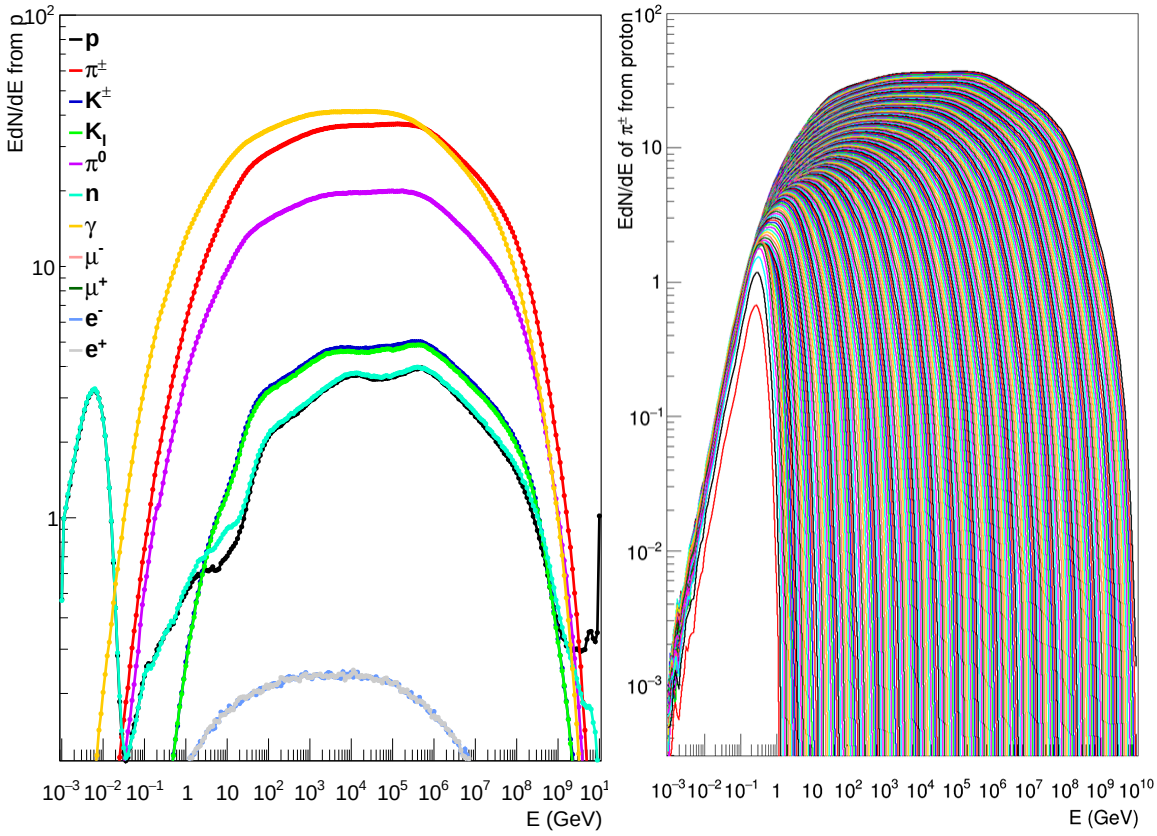


Figure 3.1: Energy spectra in CONEX for EPOS LHC. Left: Spectra of secondary particles in p+air interactions at 10^{19} eV. Right: $\pi^{\pm}$ spectra in p+air interactions for all proton energies.

several kinds of modifications over the particle spectra ensuring energy conservation. These secondary particle spectra are discretized in CONEX. Here, the function $ppj(j, i, n_2, n_1)$ is defined to manage all spectra in an easy way. This function gives the spectrum value, where n_1 is the projectile particle identification number, n_2 is the secondary particle identification number, i is the energy bin of the projectile particle and j is the energy bin of the secondary particle.

Each projectile type has different secondary particle spectra depending on its energy. The first step is to select an energy range of the projectile particle, where the corresponding secondary particle spectra will be changed. This is done following these steps:

1. Set the secondary particle type over which the changes will be implemented. This particle is called the *reference particle*.

2. Set the energy range of the reference particle spectra where the changes will be implemented.

3. Set the amount of energy E_S that will be transferred from the selected range to another particle spectrum.

4. Set how this energy will be taken from the selected energy range through a given predefined *transfer function*.

5. Set how this energy will be shared among other secondary particles, i.e. the energy range of the chosen secondary particles and the transfer function.

In figure 3.2 a schematic picture of this procedure is shown. Consider two energy spectra of the secondary particles B and C, which are the result of the interaction of the projectile particle A and air. The addition of their integrated spectra gives the total energy E of the interaction (Fig. 3.2-a). In Fig. 3.2-b (left) the spectrum B is shown where the hatched region corresponds to the region from where energy will be taken. The energy E_S is taken using a triangular function (Fig. 3.2-b,center). Afterwards, E_S is transferred to the spectrum C into its hatched range (Fig. 3.2-c, left) using another function (Fig. 3.2-c, center). At the end of these steps, the total energy of both spectra is the same as at the beginning, E, which means that the energy is conserved.

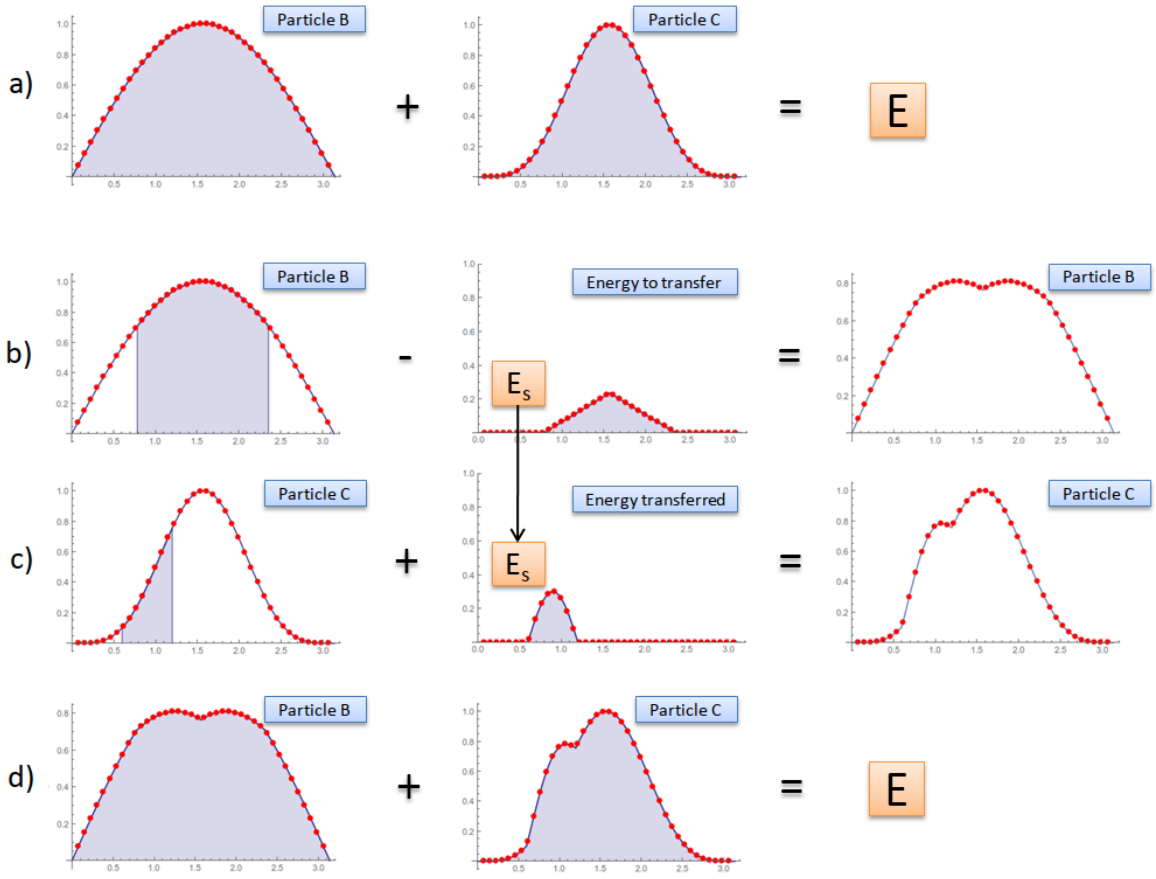


Figure 3.2: Scheme of the procedure used to modify the secondary energy spectra of CONEX.

In order to modify the spectra consistently, this tool needs an input file like the following:

```
Input      InputName
Output     OutputName
Projectile n1  E0  E1  Efunction TypeTransfer
Secondary  n2  -E  RangeType Ei Ef functionName coef1 coef2
Secondary  n2   E  RangeType Ei Ef functionName coef1 coef2
Secondary  n2   E  RangeType Ei Ef functionName coef1 coef2
...
```

The parameter `InputName` is the input filename which contains all the spectra. Usually it is the default file created by CONEX. The `OutputName` is the output filename where the modified spectra will be saved.

The third line sets which spectra will be modified. n1 is the projectile type. E0 and E1 are the lower and higher limits of the energy range of n1, where the corresponding secondary particle spectra will be changed.

`Efunction` sets how the projectile energy dependence of the changes is. `Efunction` can be set as *constant* applying the same changes at all spectra of the projectile energies ranging from E0 to E1. Otherwise it can be *logE* applying no change at E0 and increasing linearly with the logarithm of the energy reaching the maximum change at E1.

The fourth line sets how the spectrum of the reference particle will be modified. n2 is the secondary particle type. Ei and Ef are the lower and higher limits of the energy range where the changes will be applied. This limit can be expressed in absolute energy values or relative to the projectile energy setting `RangeType` as *abs* or *rel*, respectively. A negative energy *-E* indicates the reference particle, where *E* is the percentage of the energy contained in the energy range that will be transferred. The parameters `functionName,coef1` and `coef2` give the kind of function that will take this energy. The total energy E_s taken from the reference particle in terms of the function ppj is defined as

$$E_s = E \sum_{j=i_0}^{i_f} E_j \; ppj(j, i, n_2, n_1) \tag{3.9}$$

The fifth and following lines set how E_s is shared between the other particles.

The parameter `TypeTransfer`has to be set as *in* (or *out*) if the energy E_s is taken from (or added to) the reference particle.

Constant Energy Distribution

The simplest way to transfer energy is using a constant energy distribution. In a spectrum the way to take an energy E_s from the energy bin E_{i_0} to E_{i_f} using a constant function is given by

$$E_s = \sum_{j=i_0}^{i_f} a \; E_j \Rightarrow \; a = \frac{E_s}{\sum_{j=i_0}^{i_f} E_j}$$

So the spectrum is modified by

$$ppj(j, i, n_2, n_1) \quad = \quad ppj(j, i, n_2, n_1) \; \pm \; a$$

Fig. 3.3 shows how this constant energy distribution in a small range (3.3-a) and in the whole range (3.3-b) is implemented. In the small-range case the spectrum is not continuous, for that reason its implementation alone has no physical motivation. On the other hand the whole-range case can be understood as an increment in the total number of particles with no change in the shape of the energy spectrum.

In the whole range case this option has a problem with energy conservation due to the first and last bins which have small values. The spectrum which transfers energy could have negative values in the first and last bins. The first bin does not imply a considerable problem because the impact on the energy conservation is negligible but the last bins have a big impact and should be taken into account. To avoid this problem the trapezoid function is useful.

Fig. 3.4 shows how the proton and charged pion spectra change after a trivial modification using the constant function. These changes are implemented with the following input file:

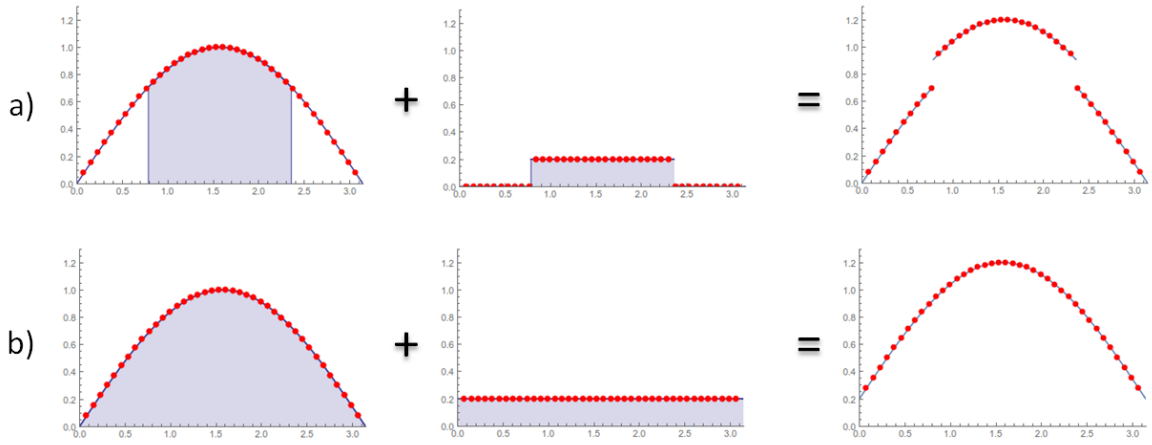


Figure 3.3: Schematic picture of spectra change using a constant function.

```
Projectile 1  1e5  1e10  constant out
Secondary  2  -40   rel  .4  .7  constant 0  0
Secondary  1  100   rel  .7  .8  constant 0  0
```

This file has to be read as follow: The particle spectra of the projectile id=1 (proton) with energy from 10^5 GeV to 10^{10} GeV have to be modified with a constant energy evolution (line 1). 40% of the energy contained in the relative energy range $(0.4; 0.7)$ have to be taken in a constant way from the spectra of the secondary particle id=2 (pions) (line 2). All this energy has to be added to the spectra of secondary particle id=1 (proton) in the relative energy range $(0.7; 0.8)$ in a constant way (line 3).

Scale factor Energy Distribution

Another simple way to take energy from a spectrum is using a scale factor.

$$E_s = \sum_{j=i_0}^{i_f} a\ E_j\ ppj(j,i,\hat{n}_2,n_1) \Rightarrow\ a = \frac{E_s}{\sum_{j=i_0}^{i_f} E_j\ ppj(j,i,\hat{n}_2,n_1)}.$$

So the spectrum is modified by a:

$$ppj(j,i,\hat{n}_2,n_1) \quad = \quad ppj(j,i,\hat{n}_2,n_1)\ (1 \pm a).$$

Trapezoidal Energy Distribution

The problem of the constant and scale factor functions is the non-continuity when the range is not the whole spectrum. In order to ensure continuity the trapezoidal function must be used, see Fig. 3.5:

$$E_s = \sum_{j=i_0+1}^{i_1} a_1\ (i - i_0)\ E_i + \sum_{j=i_1+1}^{i_2} a_2\ E_i + \sum_{j=i_2+1}^{i_3} a_3\ (i - i_3)\ E_i.$$

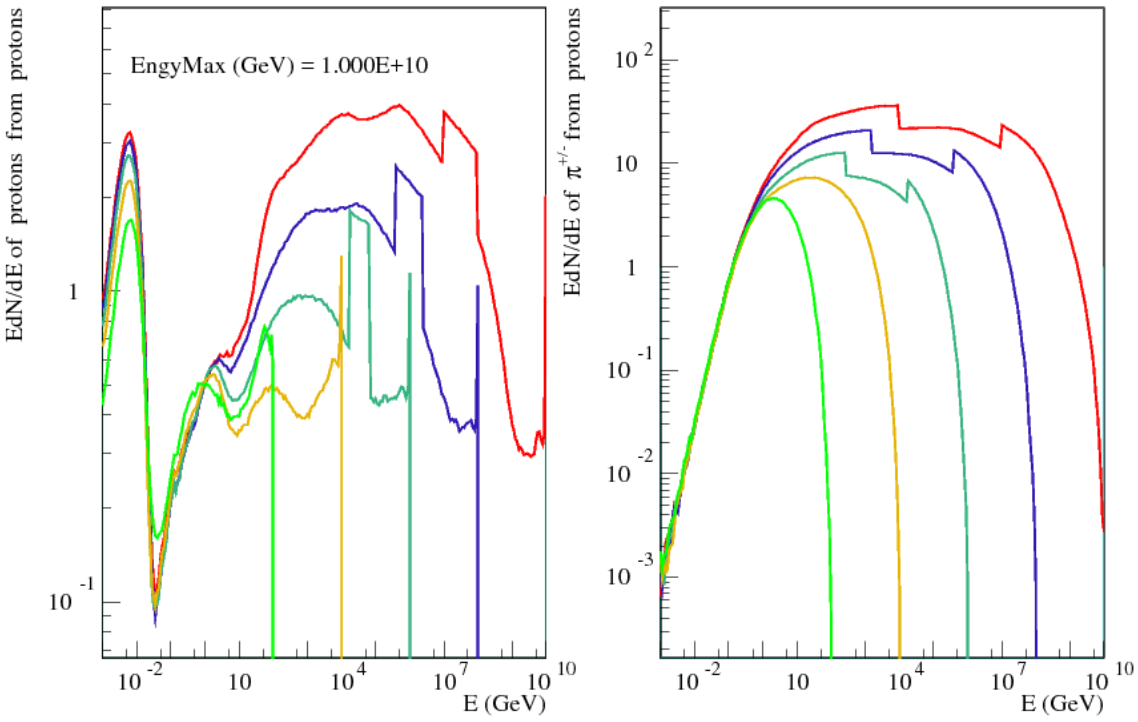


Figure 3.4: Trivial spectra modification by the constant function. The energy taken from the spectra on the right is added to the spectra on the left. Five spectra at different energies are shown. Low energy spectra are not modified.

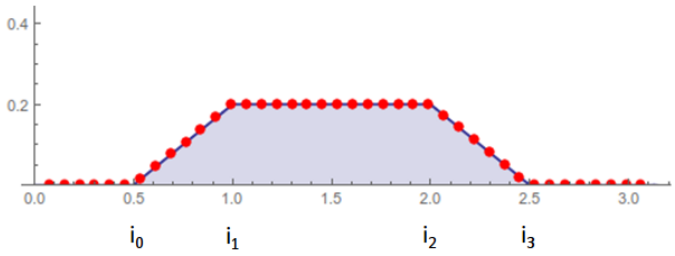


Figure 3.5: Schematic picture of the trapezoidal function.

Requiring continuity

$$
\begin{aligned}
a_1 &= E_s \Big(\sum_{j=j_0+1}^{j_1} (j - j_0)\, E_j + \sum_{j=j_1+1}^{j_2} (j_1 - j_0)\, E_j + \sum_{j=j_2+1}^{j_3} \frac{j_1 - j_0}{j_2 - j_3}\, (j - j_2)\, E_j \Big)^{-1} \\
a_2 &= a_1 \frac{j_1 - j_0}{j_2 - j_3} \\
a_3 &= a_1 \, (j_1 - j_0)
\end{aligned}
$$

Photon spectra from π^0 spectra

In the electromagnetic development, CONEX uses the photons spectra which are obtained from the neutral pion spectra. Hence, if neutral pion spectra are modified, it is necessary to calculate the new photon spectra.

Consider a π^0 which propagates along the +z axis and that the angle of the photon to the z axis in the rest frame is θ^*: The decay $\pi^0 \rightarrow \gamma\gamma$ is isotropic in the rest frame, so the distribution is flat as a function of $\cos\theta^*$

$$\frac{dN}{d\cos\theta^*} = \frac{1}{2}. \tag{3.10}$$

The distribution of photon energies is

$$\frac{dN}{dE_\gamma} = \frac{dN}{d\cos\theta^*}\frac{d\cos\theta^*}{dE_\gamma} = \frac{1}{2}\frac{d\cos\theta^*}{dE_\gamma}. \tag{3.11}$$

The energy distribution of photons from a π^0 with momentum p_π is obtained from the transformation between rest and laboratory frame.

$$\frac{dN}{dE_\gamma} = \frac{2}{p_\pi} \tag{3.12}$$

The number of photons with energy E_γ is then:

$$N_\gamma(E_\gamma) = \int_{E_\gamma}^{\infty} \frac{2}{p_\pi} dp_\pi \tag{3.13}$$

The spectra in CONEX are discrete and have a log energy spacing, so the number of photons in the energy range (E_i,E_i+ ΔE_i) is

$$N_\gamma(E_i) \approx \sum_{j=i}^{j_{max}} \frac{2}{E_j} \Delta E_i \; N_\pi(E_j). \tag{3.14}$$

The last step is to apply a scale factor to the new photons spectra in order to have energy conservation. The photons spectra obtained with this procedure are compared with the original spectra in Fig. 3.6. The differences at the highest energy bins between both spectra produce a little change in X_{max} by ± 2 g cm^{-2}, which is corrected in all shower simulations.

Particle ratio modification

One interesting application of this tool is to modify the secondary particle ratios. The particle ratios are given by the hadronization model implemented in the hadronic interaction model. Consequently it is possible to modify the hadronization in an effective way just changing the particle ratios of secondary energy spectra. The hadronic models consider two types of hadronization in hadronic interactions, one for the central part of the collision and another one for the remnants. The standard string fragmentation is commonly used by the hadronic models to describe the central part, see Sec. 2.8. This kind of hadronization produces particles in different parts of the energy spectrum. The string fragmentation produces most particles at mid-rapidity which corresponds to the broad peak in the center of the spectra in Fig. 3.1(left). The remnant hadronization has its main contribution when the secondary particle is the same as the projectile particle. This is the already mentioned *leading particle effect*, see Sec. 2.4. When the projectile and secondary particles are the same, their spectrum has a *diffractive peak* at $x_F \simeq 1$, i. e. at the maximum energy. Fig. 3.7 shows the spectra with the leading particle contribution. If one wants to modify a given hadronization model the other one must remain untouched.

In this tool, it is possible to modify the particle ratios by the following procedure: the spectra of all particles, except the spectra with leading particle contribution (i.e. protons in

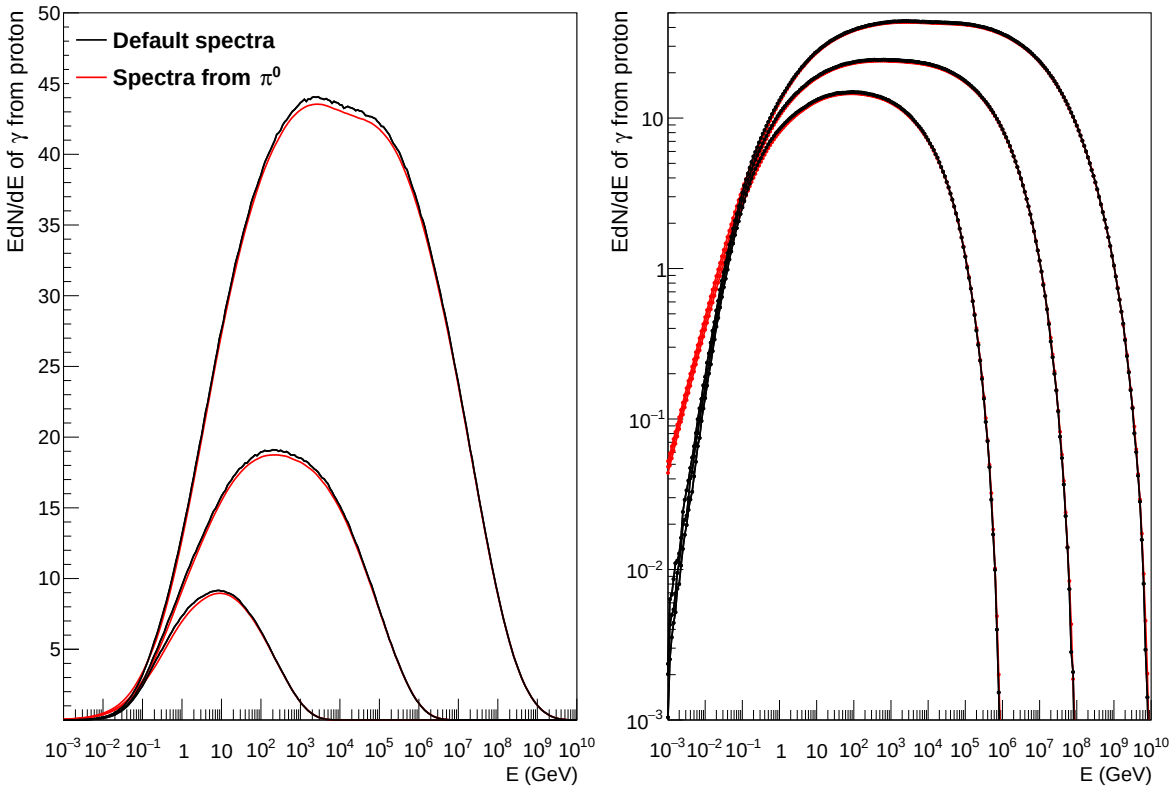
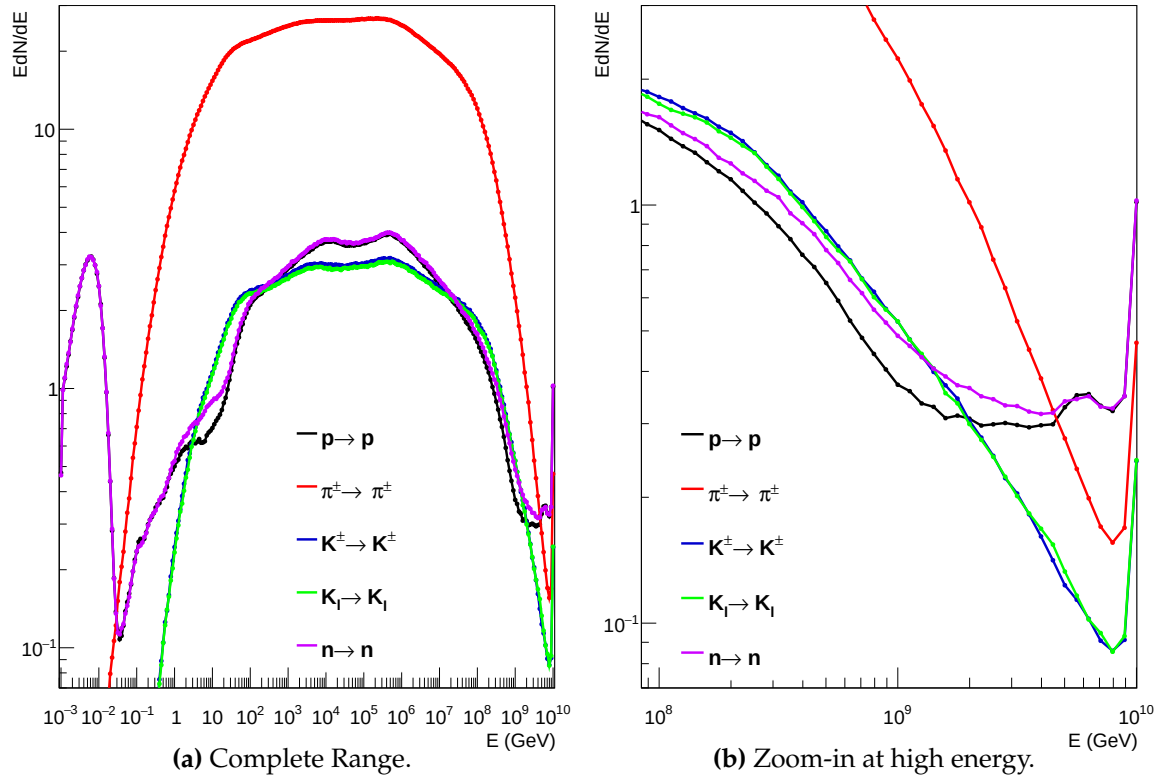


Figure 3.6: Photon spectra calculated from the π^0 spectra in linear scale (left) and log scale (right).

proton-air, kaons in kaon-air interactions, and so on), are modified by a scale factor. And in the last step, the spectra with the leading particle contribution change in order to maintain energy conservation. In Fig. 3.8 an example is shown where the spectra are modified linearly with logE at three different projectile energies and where the neutral pion spectrum is modified by a scale factor as well as for the other particles. At the end, the proton spectra are modified in order to preserve energy conservation. The leading particle contribution is not modified.



(a) Complete Range. **(b)** Zoom-in at high energy.

Figure 3.7: Spectra with leading particle contribution.

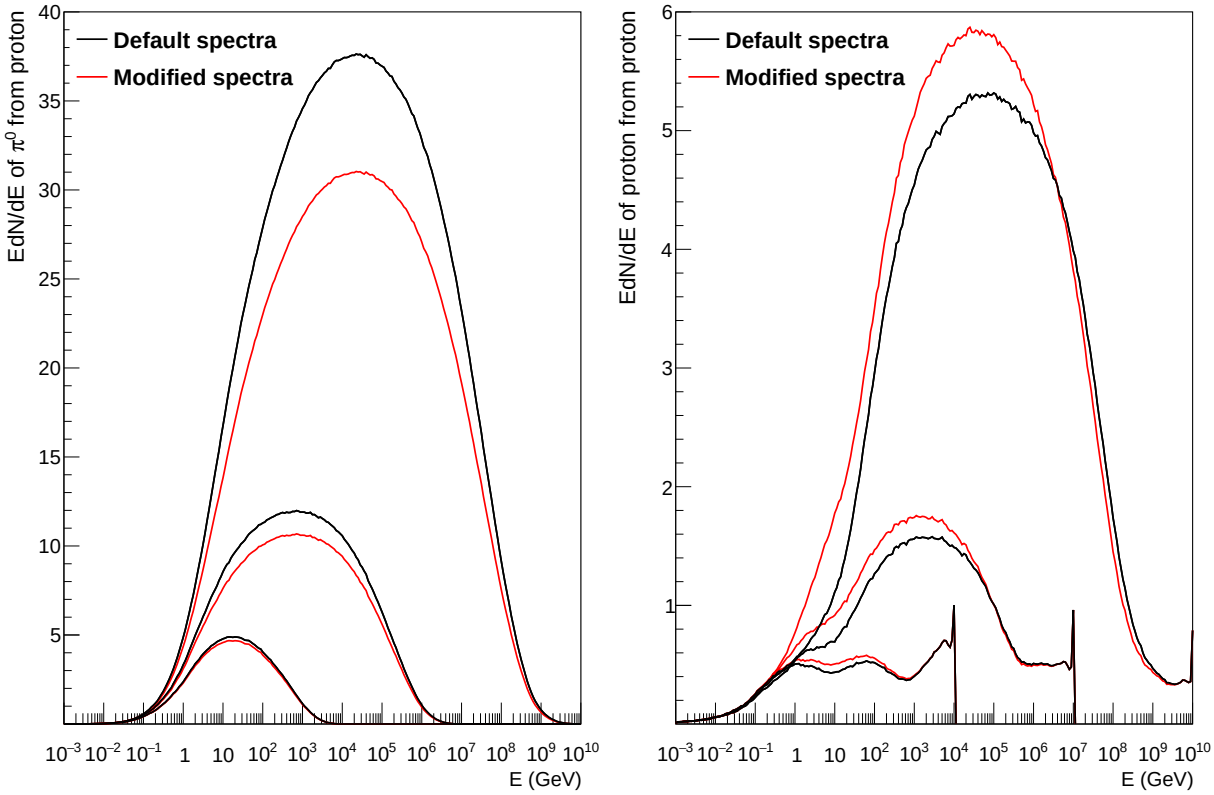


Figure 3.8: Default and modified spectra of secondary π^0 (left) and proton (right) at three different proton-projectile energies for QGSJETII.04.

CHAPTER 4

Collective hadronization and muon production in air showers

The muon production in air showers depends strongly on the amount of energy ending up in the electromagnetic channel in hadronic collisions. Therefore, the parameter $R = \frac{E_{\text{em}}}{E_{\text{had}}}$ is useful to describe the muon content in air showers [160–162]. This parameter, in turn, is closely related to excited partonic system hadronization.

In this chapter, the newly developed tool presented in Sect. 3.3 is used in order to implement the core-corona model in the CONEX framework where R is expected to change significantly.

4.1 Introduction

In the hadronic interaction models used to simulate air showers, the hadronization is mainly done using a string fragmentation model (see Sect. 2.8), which was successfully developed to describe the hadron production in e^+-e^- collisions and low energy proton-proton collisions. In systems with higher energy densities, such as heavy ion collisions, a statistical hadronization (see Sect. 2.6 and 2.9) of a fluid is expected, where the production of heavy particles is favored, reducing the fraction of π^0 compared to other types of particles. In the early 2000s, *collective effects* have been observed in heavy ion collisions, often referred to as *large* systems, at RHIC [163–166]. Similar effects have been predicted [167–172] for proton-proton collisions (aka *small* systems) and were eventually discovered at the LHC [173] (see Refs. [174, 175] for detailed reviews).

While a fluid-like behavior (referred to as collective effects in the following) is confirmed in both large and small systems, their origin is still unclear. In large systems the existence of a quark-gluon-plasma (QGP) is commonly assumed as a phase of parton matter (see Sect. 2.7). This QGP will evolve according to the laws of hydrodynamics and eventually decay statistically. There are various expected consequences of such a scenario, such as long-range two-particle correlations, the so-called "ridge" phenomenon [173, 176], jet quenching [177, 178], or enhanced production of strange hadrons [130]. It was initially a surprise when such effects were also discovered in small systems. While it was argued that also in central collisions of small systems the energy densities may be high enough to allow for the formation of a QGP [167], other recent studies have shown that collective effects can be achieved by alternative mechanisms such as microscopic effects in string fragmentation [179] or QCD

interference [180]. The possibility of collective effects in smaller systems opens the door to study the impact of a different hadronization scheme in high energy interactions also within air showers. Air shower cascades are driven by collisions of hadrons and light nuclei at ultra-high energies.

The following sections shows that statistical hadronization in collisions of hadrons and nuclei can play a so far underestimated importance in the understanding of muon production in air showers [181, 182].

The underlying mechanism responsible for the production of these effects is expected to produce characteristic observables in the final state of hadron collisions. In particular, the statistical hadronization affects the energy fraction contained in electromagnetic versus hadronic particles, R, which has important possible implications for the muon production in cosmic ray air showers.

4.2 The muon problem and the R observable

The dominant mechanism for the production of muons in air showers is via the decay of light charged mesons. The vast majority of mesons are produced at the end of the hadron cascade after typically five to ten generations of hadronic interactions (depending on the energy and zenith angle of the primary cosmic ray). The energy carried by neutral pions, however, is directly fed to the electromagnetic shower component and is not available for further production of more mesons and subsequently muons. The energy carried by hadrons that are not neutral pions is, on the other hand, able to produce more hadrons and ultimately muons in following interactions and decays. Using a simple Heitler type toy-model [183] based on [26], the *neutral pion fraction* $c = N_{\pi^0}/N_{\text{mult}}$, defined as the number of neutral pions N_{π^0} divided by the total number of final-state particles N_{mult} in a collision, was found to have a strong impact on the muon number and in particular on the slope of the energy dependence of the muon production. Indeed in this model

$$N_\mu = \left(\frac{E_0}{E_{\text{dec}}}\right)^\beta \quad \text{with} \quad \beta = 1 + \frac{\ln(1-c)}{\ln N_{\text{mult}}}, \tag{4.1}$$

where E_0 is the energy of the primary cosmic ray particle and E_{dec} is the typical energy at which mesons decay in the cascade (see Sect. 1.2). So the muon number N_μ increases strongly with decreasing c, which is understandable since more hadrons are available to produce muons. A second quantity with a strong impact on the muon number is the hadron multiplicity N_{mult}.

The value of c is very important for the muon production. Unfortunately, it is difficult to measure both N_{π^0} and N_{mult} experimentally (for example at the LHC) since neutral particles cannot be counted easily individually. In general, secondary particle identification is unavailable at large pseudorapidities η where the energy flow is large enough to become relevant for the air shower development.

The ratio of the electromagnetic to the hadronic energy density has been proposed in [184] as a new observable sensitive to properties of the hadronization and can be directly related to c. This observable is given by

$$R(\eta) = \frac{\langle \mathrm{d}E_{\text{em}}/\mathrm{d}\eta \rangle}{\langle \mathrm{d}E_{\text{had}}/\mathrm{d}\eta \rangle}\,. \tag{4.2}$$

where the energy densities $\langle \mathrm{d}E/\mathrm{d}\eta \rangle$ are obtained by summing the energy of all final-state particles except for neutrinos in bins of η and averaging over a large number of collisions.

The neutral pion fraction c can be easily related to the energy ratio R, since both are very similar kinematic aspects of final state distributions. If all particles have the same energy such as in the generalized Heitler model, then $R = c/(1-c)$. But R is experimentally much easier to measure, since, using a calorimeter, the signals deposited by electromagnetic particles and by hadrons are characteristically different.

The influence of various effective parameters in interaction models (like R, c, or N_{mult}) on the main air shower observables was investigated in a previous study [160], in which the behavior of hadronic interaction models in air shower simulations was modified in an energy-dependent way during full air shower cascade simulations within CONEX [158]. The correlated impact of R and N_{mult} on X_{max} and $\ln N_{\mu}$ were analized in [184] in full air shower simulations, where the change of N_{mult} is not enough to close the gap to the data, but modifications of R mainly affect the muon number and leave X_{max} unchanged. So it is possible to make the simulations compatible with the air shower data at 10^{19} eV.

Furthermore, in Sect. 1.4 it was established that the muon discrepancy in simulations increases smoothly with energy. Thus, the slope of the energy dependence introduced in eq. (4.1) is also affected, pointing to a too small value of β. This may be related to a too large π^0 production. This energy dependence is explored in more detail in the next section.

4.3 Core-corona Implementation

The discussion in the previous section suggests that a change of R (or c, equivalently) is a potential way to reduce the discrepancy between measurements and air shower simulations. Nevertheless, R is quite well constrained by theory as well as laboratory measurements and, thus, cannot be changed entirely arbitrarily. In a naive model like Ref. [26] where only pions are considered as secondary particles, $R = 0.5$. In a more realistic approach based on string fragmentation $R \approx 0.41$. But as shown in Ref. [130], particle ratios such as K/π, p/π or Λ/π change with increasing secondary particle density, saturating to the value given by a thermal/statistical model with a freezeout temperature of 156.5 MeV [123] yielding $R \approx 0.34$. Such a behavior can be explained in terms of a core-corona picture presented in Sect. 2.10. This approach has been used in the framework of realistic simulations [142, 153], but also in simple model calculations [144, 146, 149, 185]. The basic idea is that some fraction of the volume of an event (or even a fraction of events) behaves as a quark gluon plasma and decays according to statistical hadronization (core), whereas the other part produces particles via string fragmentation (corona). The particle yield N_i for particle species i is then a sum of two contributions

$$N_i = \omega_{\mathrm{core}}\, N_i^{\mathrm{core}} + (1-\omega_{\mathrm{core}})\, N_i^{\mathrm{corona}}, \tag{4.3}$$

where N_i^{core} represents statistical (grand canonical) particle production, and N_i^{corona} is the yield from string decay. Crucial is the core weight ω_{core}. In order to explain LHC data [130] the weight ω_{core} needs to increase monotonically with the multiplicity, starting from zero for low multiplicity pp scattering, up to 0.5 or more for very high multiplicity pp, reaching unity for central heavy ion collisions (PbPb).

In the following, a straightforward core-corona approach is employed, based on eq. (4.3), for the different hadronic interaction models in CONEX air shower simulations. The particle yield from the chosen interaction model is by definition considered to be the corona yield, whereas the standard statistical hadronization is used (also referred to as resonance gas) for the core part. So $\omega_{\mathrm{core}} = 0$ would be the "normal" simulation with the default interaction model. Choosing $\omega_{\mathrm{core}} > 0$ amounts to mixing the yields from the interaction model according to the core-corona superposition shown in eq. (4.3). The core will certainly help

concerning the “muon problem", because statistical hadronization produces more heavy particles and less pions compared to string fragmentation, and therefore R is smaller [181,182].

Technically, we directly modify individual particle ratios of the secondary particle spectra $\mathrm{d}N_i/\mathrm{d}E_j$, for particle species i and energy bins $\mathrm{d}E_j$, of hadronic interactions with air nuclei used by CONEX for numerical air shower simulations based on cascade equations. More details about CONEX and its secondary particle spectra are given in Sect. 3.3. So, knowing the initial ratios $\pi^0/\pi^\pm, p/\pi^\pm, K^\pm/\pi^\pm, p/n, K^0/K^\pm$ (taking into account strange baryon decays) from a corona type model and the value of the same ratios from the core model, we compute new spectra in which the particle yields include both, core and corona according to ω_{core} and the preservation of the leading particle spectra as explained in Sect. 3.3. The particle ratios from the core are taken from Fig. 2.11. But as CONEX decays all strange baryons and only consider nucleons, pions and kaons, the particle ratios become $p/\pi^\pm = 0.0951$, $K^\pm/\pi^\pm = 0.167$, $p/n = 1.17$, $K^0/K^\pm = 0.98$, $\pi^0/\pi^\pm = 0.504$ where $p = p + \bar{p}$, $n = n + \bar{n}$ and $K^0 = K^0 + \bar{K}^0$.

The core weight ω_{core} is expected to increase with energy in a logarithmic way, so the energy-dependent factor is defined as:

$$F(E_{\mathrm{lab}}; E_{\mathrm{th}}, E_{\mathrm{scale}}) = \frac{\log_{10}(E_{\mathrm{lab}}/E_{\mathrm{th}})}{\log_{10}(E_{\mathrm{scale}}/E_{\mathrm{th}})} \text{ for } E_{\mathrm{lab}} > E_{\mathrm{th}}, \tag{4.4}$$

representing the assumption that models are well constraint by accelerator data at lower energies (below E_{th}), where $F(E_{\mathrm{lab}}) = 0$, while they become logarithmically unconstrained going to higher energies. Typical threshold values are $E_{\mathrm{th}} \simeq s_{\mathrm{Tevatron}}/(2m_{\mathrm{p}}) \approx 1\,\mathrm{PeV}$, using the center-of-mass energy of the Tevatron accelerator. However, in particular particle production, in the important forward phase space, may be largely unconstrained by both Tevatron and LHC data, allowing much lower values of E_{th} to be explored. In order to do that, the energy threshold is fixed by $E_{\mathrm{th}} = 100\,\mathrm{GeV}$ for fixed-target energies. The parameter E_{scale} is the *reference* energy scale. Different energy dependencies are explored by changing E_{scale} from 100 GeV (corresponding to a step function), to 10^6 GeV, and 10^{10} GeV. In addition $F(E_{\mathrm{lab}}; E_{\mathrm{th}}, E_{\mathrm{scale}}) = 1$ is required for all $E_{\mathrm{lab}} \geq E_{\mathrm{scale}}$. It is a key point of the application of eq. (4.4) inside CONEX that a significant fraction of the air shower cascade is consistently modified during the simulations. Thus, to model the energy dependency

$$\omega_{\mathrm{core}}(E_{\mathrm{lab}}) = f_\omega\, F(E_{\mathrm{lab}}; E_{\mathrm{th}}, E_{\mathrm{scale}}) \tag{4.5}$$

is used, where the f_ω scale is varied by 0.25, 0.5, 0.75 and 1.0 as depicted in Fig. 4.1. Each $\omega_{\mathrm{core}}(E_{\mathrm{lab}})$ gives a different $R = R(E_{\mathrm{lab}})$. As an example, the default $R = R(E_{\mathrm{lab}})$ in Fig. 4.2 is compared with the case $f_\omega = 1$ at mid-rapidity and forward.

4.4 Assumptions and limitations

The core-corona implementation presented in the previous section has some theoretical limitations. One of them is that in this CONEX approach there is no information about the η distribution of the core and the corona. This could be a problem because the core could be limited to low η values. For example, Fig. 4.3 shows the different core and corona contributions at different rapidity values for EPOS LHC which has a very conservative core-corona model properly implemented. In the forward region the density of particles is low and the core is not expected to form there.

However, the particle density at which the core is formed is unknown (underestimated in EPOS LHC for instance) and other processes like the saturation of the remnants could lead to core hadronization forward. As a consequence the rapidity at which the core contribution

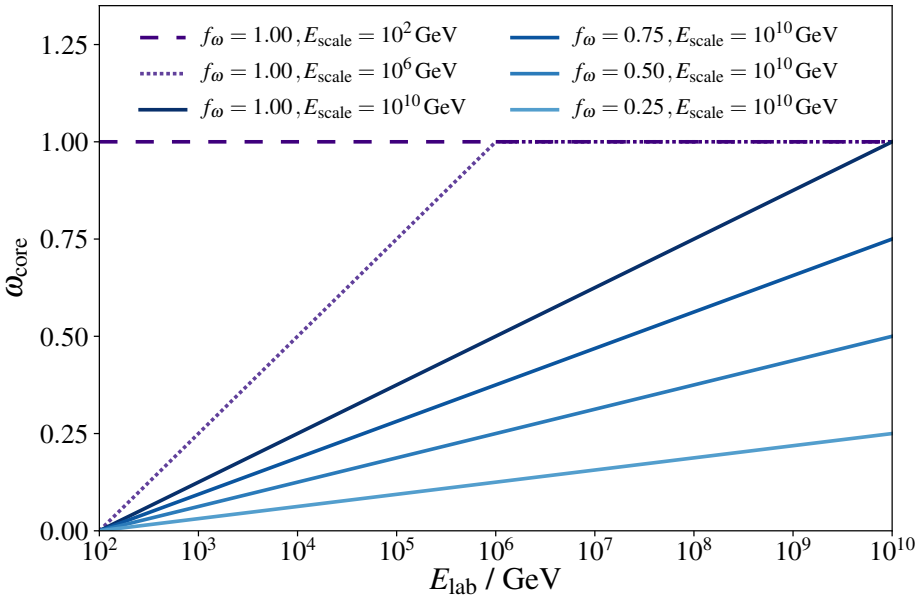


Figure 4.1: Different energy evolutions probed for ω_{core}. The solid lines represent changes in the scale f_ω of the effect, while the dashed lines also indicate the effect of changing E_{scale}.

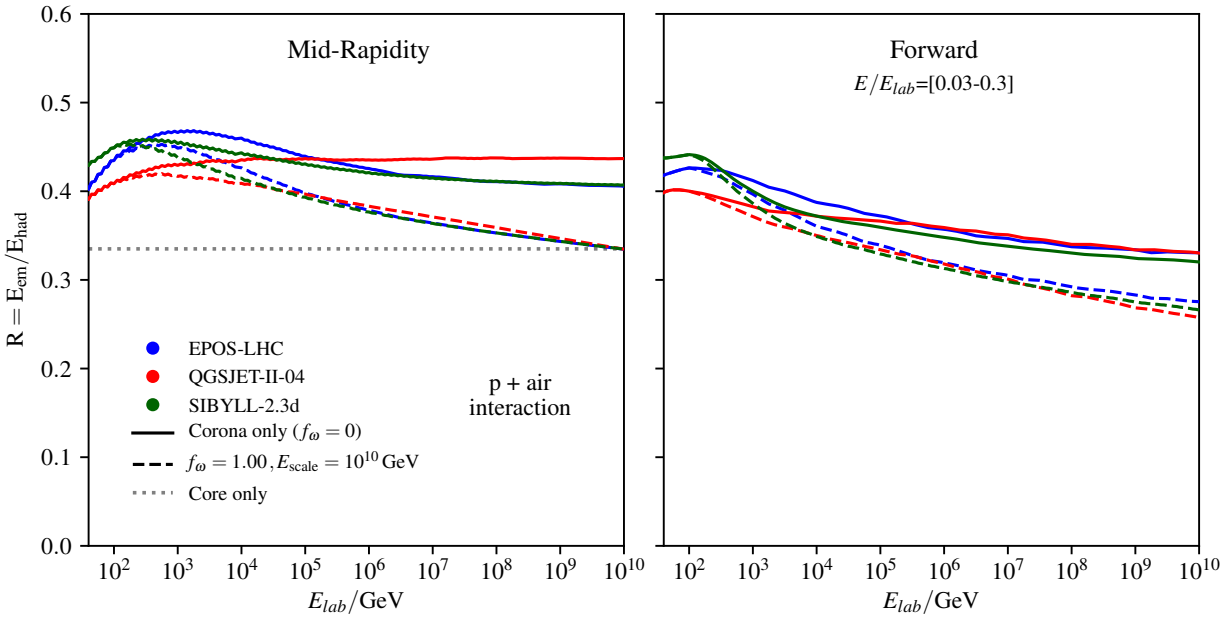


Figure 4.2: R values as a function of the projectile energy E_{lab} at mid-rapidity (left) and forward $0.03 \leq x_{\text{F}} \leq 0.3$ (right) for the default models (solid line) and the modified one with $f_\omega = 1.00$ and $E_{\text{scale}} = 10^{10}$ GeV (dashed lines) where R reaches the value given by the statistical model at the maximum energy.

becomes negligible is unknown from the theoretical point of view and there is no data available to give further information. Thus for simplicity, we assume that the core and the corona have the same rapidity distribution in order to take this approach as an upper limit of the core-corona model that could be later implemented in a Monte-Carlo model.

Furthermore, as presented in Sect. 4.3 the relative fraction of core and corona contributions in a general core-corona model depend on the interaction multiplicity event-by-event. Fig. 4.4 shows how the core contribution increases with multiplicity at a given energy in

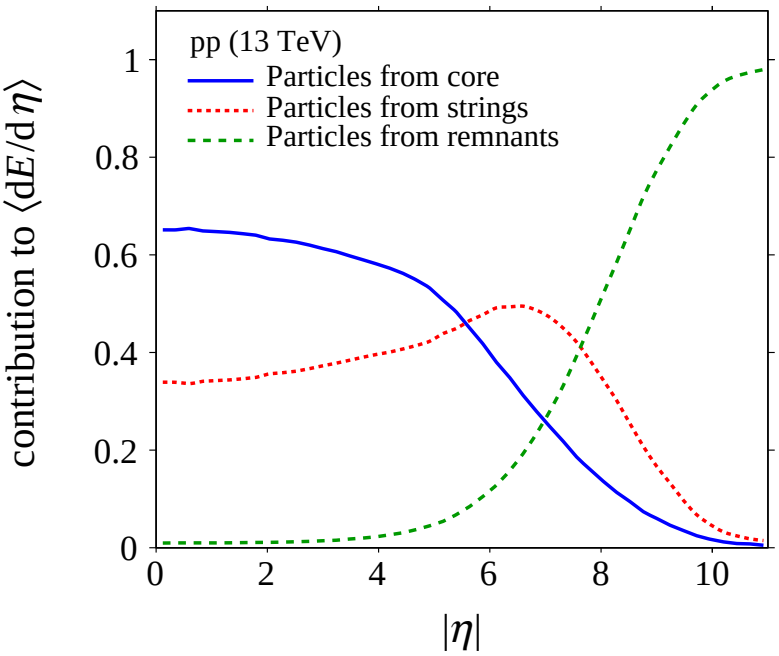


Figure 4.3: Fractional contribution of particles originating from different production mechanisms to the total energy density $dE/d\eta$ as predicted by EPOS LHC as function of $|\eta|$ [184].

EPOS LHC. CONEX spectra as used to solve cascade equations are averages of thousand of events, so the event-by-event multiplicity cannot be used as a parameter to determine the core fraction in our approach. But since the average multiplicity increases with the interaction energy as shown in Fig. 4.5, it is reasonable to assume that the average fraction of the core contribution also increases with the energy. As a consequence, the scenarii chosen to test different core-corona fraction evolution with energy are arbitrary but realistic enough to set an upper limit of what can be expected from such a model.

Last but not least, the minimum energy at which the modification is gradually applied is as low as 100 GeV in the lab system. This is given by the minimum energy in the tables for the high energy models. Again the minimum value is chosen to get the maximum impact to set an upper-limit. On the one hand, if simulations with the hadronization parameters from a QGP are not able to reproduce the observed number of muons in such favorable conditions, it could be excluded as a possible solution to solve the muon puzzle. On the other hand, if simulations are close to observed air shower data, it would justify further development in hadronic interaction models to implement such scenario in a realistic way.

4.5 Results

All the scenarios of Fig. 4.1 have been used to simulate full air showers with CONEX, using cascade equations from the first interaction to the ground, for proton and iron primary particles at $E_0 = 10^{19}$ eV. In Fig. 4.6 the results are shown in the $X_{\max}$-lnN_μ plane for the models EPOS LHC, QGSJETII.04 and SIBYLL 2.3D. Lines in this figure show all possible resulting mean values of $X_{\max}$ and lnN_μ for any mass composition of cosmic rays between pure proton (bottom right end of lines) and pure iron (top left end of lines). The resulting values of $X_{\max}$ and lnN_μ are located on a straight line because the mean values for both are linear functions of the mean-logarithmic mass of cosmic rays [187, 188] given a fixed air shower energy. The line-shape is universal, but its location, and to a lesser degree the slope and length, depend on the hadronic interaction model. Current hadronic interaction models

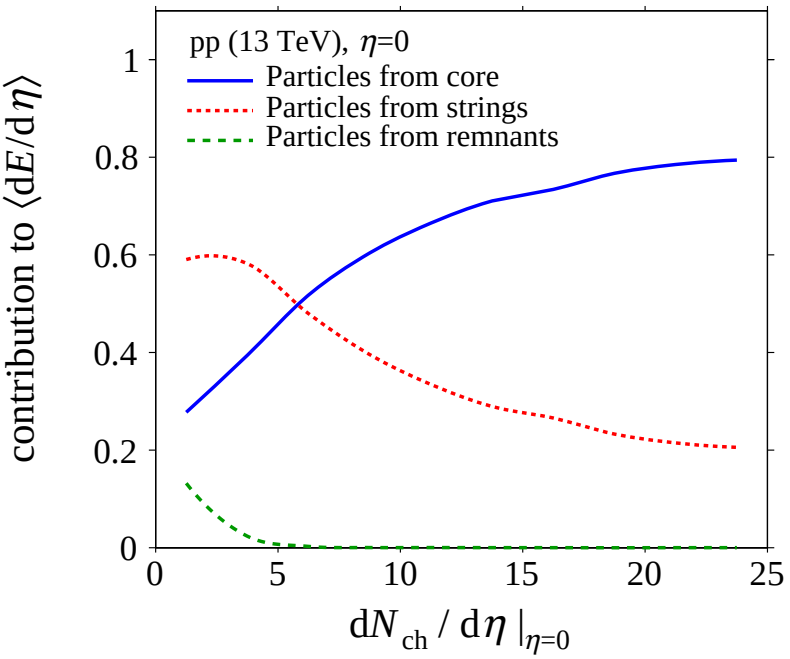


Figure 4.4: Fractional contribution of particles originating from different production mechanisms to the total energy density $\mathrm{d}E/\mathrm{d}\eta$ as predicted by EPOS LHC as a function of the charged particle density at $\eta = 0$ [184].

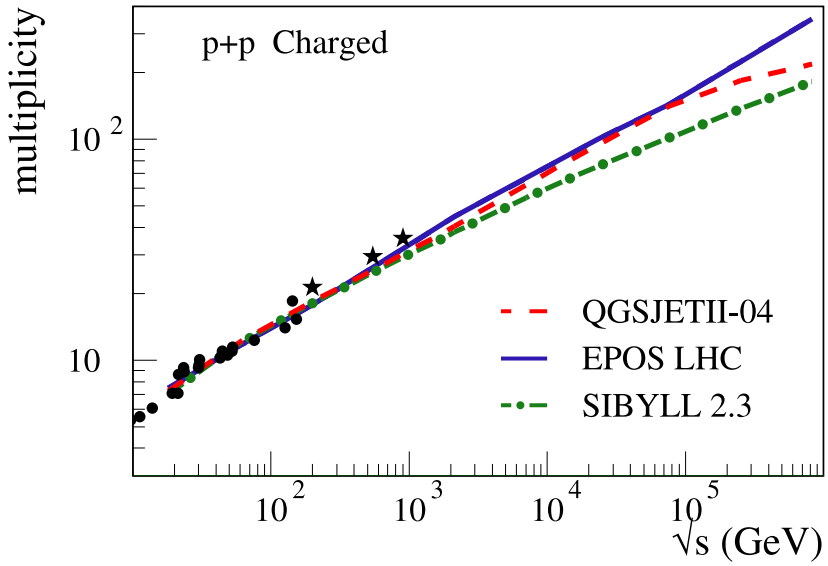


Figure 4.5: Averaged multiplicity as a function of center of mass energy. Lines are simulations and points are data [186].

predict lines which are too low compared to experimental data from air showers, as indicated by the vertical gap between the representative data point from the Auger [51] and the model line. This discrepancy is the expression of the muon problem outlined above. These examples illustrate that it is well possible with modified hadronization in air shower cascades to describe the data of the Auger Observatory. As expected, more core-like contributions are needed compared to what is currently provided by the models. This means QGP-like effects also in light colliding systems and starting in central collisions at much lower center-of-mass energies may play a decisive role.

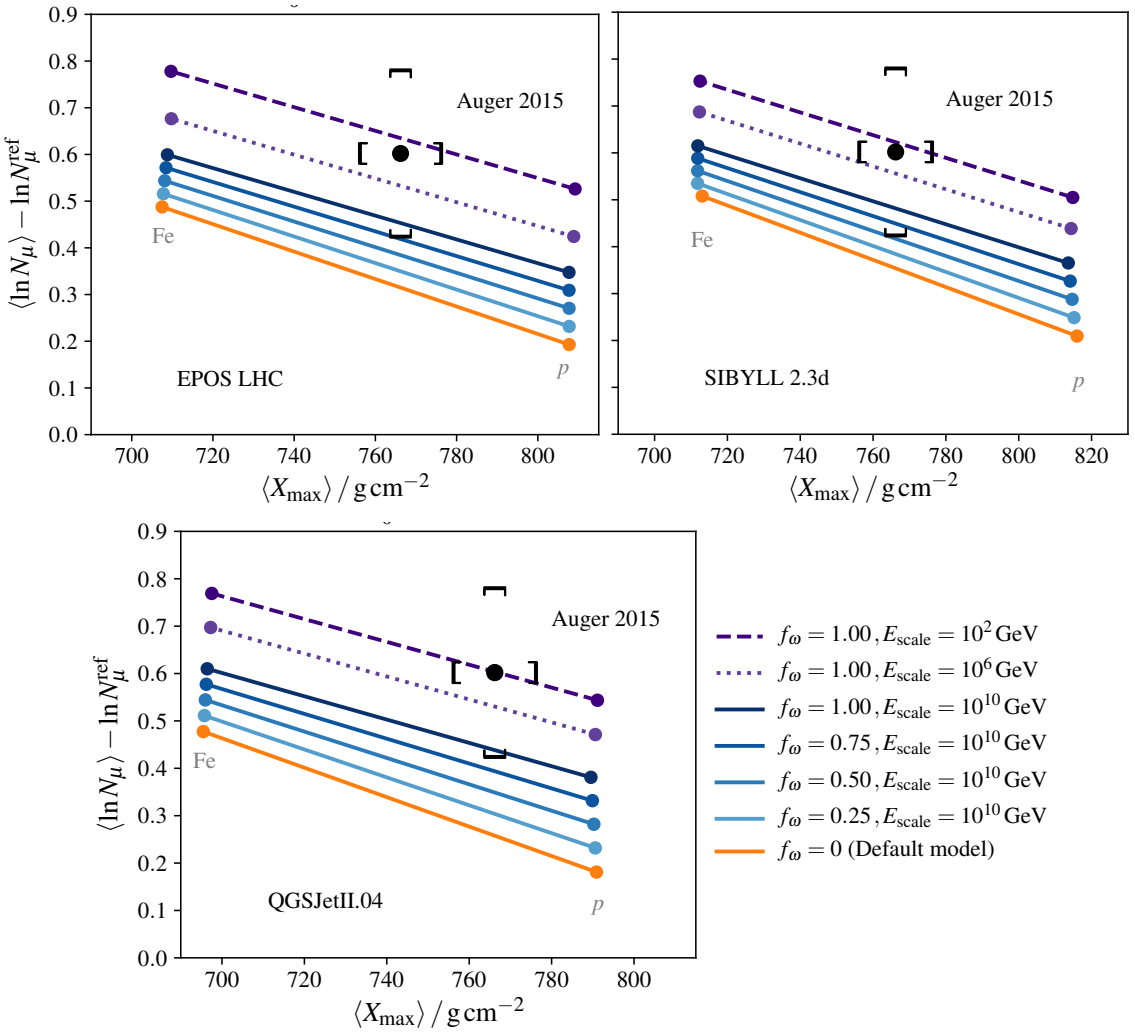


Figure 4.6: Comparison of different core-corona mixing scenarios, as described in the text, on air shower simulations at 10^{19} eV using EPOS LHC (top-left), SIBYLL 2.3D (top-right) and QGSJETII.04 (bottom) in the $X_{\rm max}$-lnN_μ plane. The solid lines represent changes in the scale f_ω, while the dashed lines also indicate the effect of changing $E_{\rm scale}$. The *default* model corresponds to the corona-only simulations. The datum is from the Auger Observatory [51]. Each model line represents all values that can be obtained for any mixture of cosmic nuclei from proton (bottom right) to iron (top left).

In order to study the effect of the core-corona model on the muon production as a function of the shower energy, the different scenarios with the compilation of experimental data presented in Sect. 1.4 can be compared using the z-scale defined in Eq. 1.11. In this way, z-scale allows a direct comparison between different muon observables. For instance, Auger SD+FD data at 10^{19} eV and SD+MD data at $10^{17.5}$ eV are compared in the $X_{\rm max}$-z plane in Fig. 4.7, where MD data seems to show that more core-like hadronization is needed to have compatible simulations.

Considering the energy dependence of z, there is an implicit dependence on the cosmic-ray mass A given by Eq. 1.13. As pointed out in Sect. 1.4, $\Delta z = z - z_{\rm mass}$ is zero for the case of full consistency between all experimental observables and the simulations based on a valid reference model. This means, plotting Δz for experimental data, $\Delta z = 0$ implies the reference model is perfect, whereas $\Delta z > 0$ implies a muon deficit in the simulations.

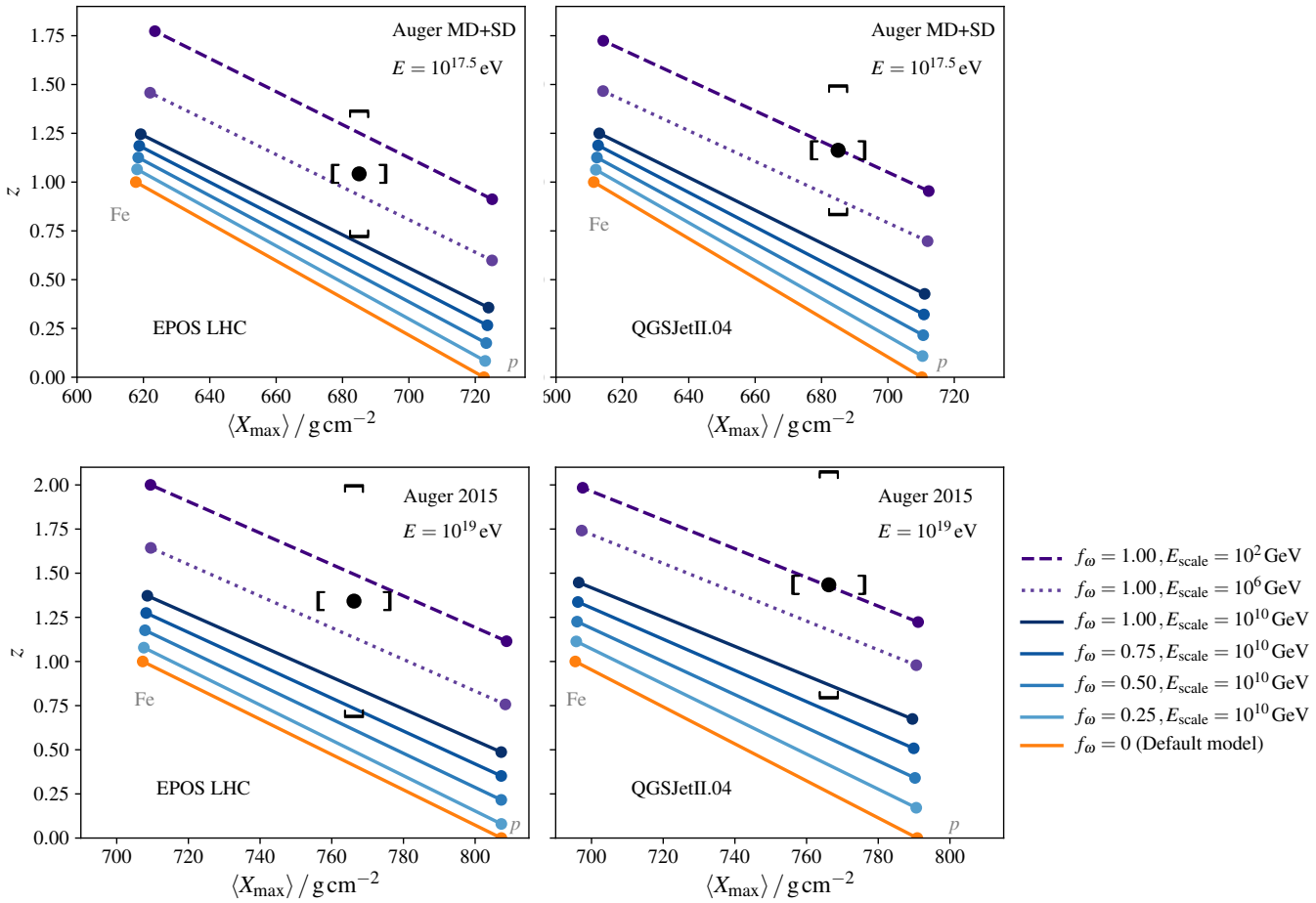


Figure 4.7: Comparison of different core-corona mixing scenarios on air shower simulations at $10^{17.5}$ eV (top) and 10^{19} eV (bottom) using EPOS LHC (left) and QGSJETII.04 (right) in the X_{max}-z plane. The solid lines represent changes in the scale f_ω, while the dashed lines also indicate the effect of changing E_{scale}. The *default* model corresponds to the corona-only simulations. The z-scale value of data are taken from [189]. Each model line represents all values that can be obtained for any mixture of cosmic nuclei from proton (bottom right) to iron (top left).

Fig. 1.15 shows that for all models the data have a positive Δz showing a significant logarithmic increase with the primary energy, indicating an increasing muon deficit in the simulations. The effect of the different energy evolution of ω_{core} for EPOS LHC and QGSJETII.04 on Δz are shown in Fig. 4.8. Here, the new simulations are treated like data and the z-scale is calculated using the original (quoted) models as a reference such that the new Δz can be compared to the data points directly. The positive Δz of the lines indicate a larger muon production when ω_{core} increases and the positive slopes mean that the slope of the muon production as a function of the primary energy is larger when ω_{core} increases. By including a consistent core-like hadronization, it is possible to reproduce the energy evolution as found in the data. This is even possible for values $\omega_{core} < 1$.

The possibility to see the effect of a core hadronization (QGP or similar more exotic phenomena) on air shower physics have already been studied in the literature [190–193]. Changes in the muon production because of a change of R under either extreme or exotic assumptions (which were not yet observed at the LHC) are usually assumed. Furthermore, it was shown that the production of a core only in very central, high-density, collisions is not sufficient to significantly change the muon numbers in air shower simulations [194].

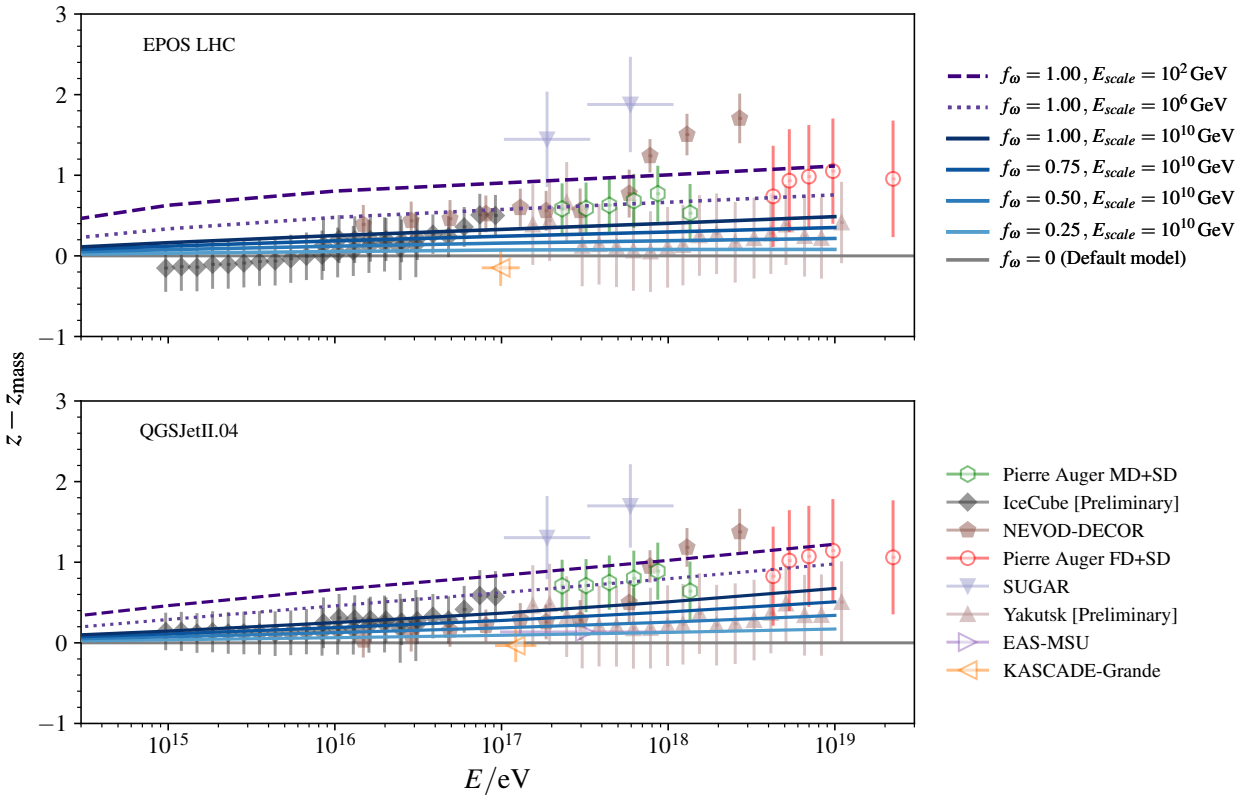


Figure 4.8: Evolution of the mass corrected z-scale, $\Delta z = z - z_{\mathrm{mass}}$, as a function of the primary energy. z values are taken from [72] and the updated Auger data from [189]. Overlayed are predictions obtained from changing the scale f_ω (solid lines) and E_{scale} (dashed and dotted lines) obtained with EPOS LHC (top) and QGSJETII.04 (bottom) air shower simulations.

In contrast to the new results presented here, in previous studies the core-like production does not cover sufficient phase space to change the muon production in air showers significantly. We demonstrate that core-like effects potentially starting at much smaller colliding systems, and at much lower center-of-mass energies as studied here, have an important impact on muon production in air showers. There are various indications at the LHC in pp and pA collisions that such a scenario is compatible with current data [130, 173], or even suggested, at energy densities as those reached by cosmic rays interacting with the atmosphere [182]. Studying LHC data at mid-rapidity it is found that for events with $\langle \mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta \rangle_{|\eta|<0.5} \sim 10$ (corresponding to typical proton-air interactions) ω_{core} is already $\approx$ 50–75%. Since this study is based on the simple assumption that the full phase space has a modified π^0 ratio, it remains crucial for cosmic ray physics to conduct further dedicated measurements at the LHC to better understand π^0 production relative to other particles. The phase space for the formation of core-like effects is potentially significantly larger than previously studied, and in particular may extend towards larger rapidities.

CHAPTER 5

Auger Muon and Infill For the Ground Array (AMIGA)

In Ch. 4 the muon content of EAS was analyzed and compared with the Core-Corona model predictions. That was done by means of the z factor which allows to compare several experiments at the same time. However, almost all of these experiments obtain the muon content of EAS in an indirect way from the measurement of the lateral distribution. In contrast, the Underground Muon Detector (UMD) of the Pierre Auger Observatory presents a great opportunity to directly measure the muon lateral distribution function (MLDF). The UMD working principle is presented in detail in this chapter, and the corresponding data analysis including the Core-Corona predictions is presented in Ch. 6.

5.1 Underground Muon Detector

The layout of the UMD stations was presented in Sect. 1.3.5. Each UMD is located close to a WCD and consists of three $10\,\mathrm{m}^2$ modules segmented into 64 plastic-scintillator strips with wavelength-shifting (WLS) optical fibers to collect the light produced by particles. The optical fibers transport the light toward an array of 64 SiPMs. Modules are buried 2.3 m underground to shield them from the electromagnetic component of air showers, so the muonic component can be measured in isolation. A module under construction is shown in Fig. 5.1. Finished modules are cover with a PVC container to protect them from the environment.

5.1.1 Scintillators

Organic scintillators emit photons by means of the fluorescence mechanism. When particles pass through the scintillator, they transfer energy to the medium generating electronic transitions from the ground state S_0 to vibrational levels of the singlet S_1. As the characteristic times of the vibrational states are smaller than the radioactive transitions ones, molecules reach their thermal equilibrium before the backward transition (S_1 to S_0). A diagram of the characteristic energy levels of a plastic scintillator is shown in Fig. 5.2. The absorption and emission spectra is not the same, because the electrons lose part of the energy in the backward transition. The shift between both spectra is known as *Stoke's Shift* [195]. This feature is necessary to avoid the re-absorption of fluorescence photons. Once the scintillator is excited,

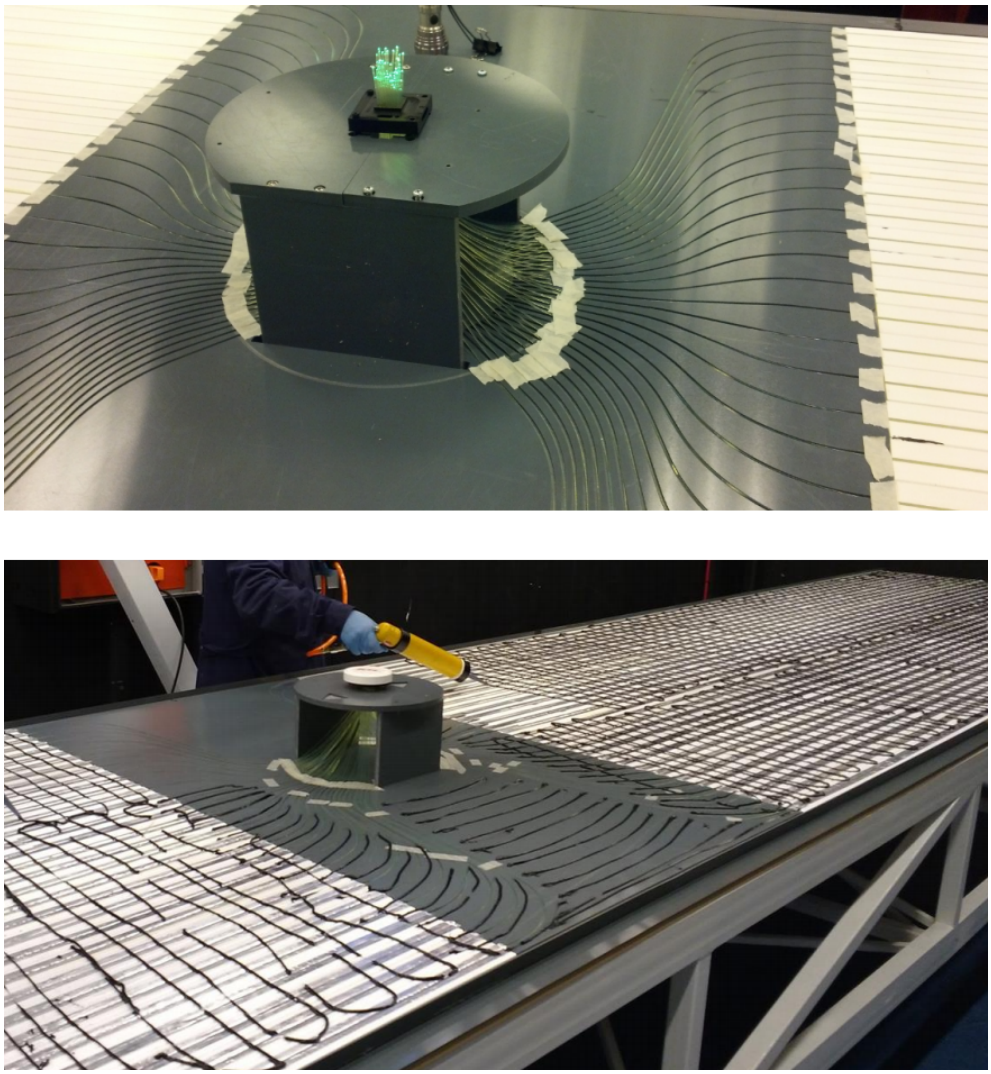


Figure 5.1: Pictures of a UMD module under construction. The optical connector with WLS optical fibers is shown at the top and the module before closing the PVC casing at the bottom [37].

it emits the maximum of light at a given time. Then, it continues to emit photons for longer times with a decreasing probability that decays exponentially in time. This is important to understand the measured signals.

The scintillation bars used in the UMD modules are extruded plastic scintillators produced at Fermi National Accelerator Laboratory [196]. Each scintillator strip is 4 m long, 4 cm width and 1 cm height, built with polystyrene doped with fluor and it has a 0.25 mm polystyrene reflective layer with 12% TiO_2 which prevents the light from leaving the scintillator. The core of the scintillation strip consists of a compound of Dow Styron 663W polystyrene as the base material doped with PPO (at 1% by weight) and with POPOP (0.03% by weight). The polystyrene base absorbs the incident radiation and emits photons in the ultraviolet spectrum. The first dopant re-emits ultraviolet photons with a larger attenuation length. And the second one absorbs the photons and re-emits 420 nm photons which are transparent for the scintillator. Therefore, most of these photons are not absorbed and they are transmitted towards the optic detection system. The composition of the scintillator used in the AMIGA UMDs is presented in Fig. 5.3. In addition, these strips have a middle groove to glue the WLS optical fiber (see Sect. 5.1.2).

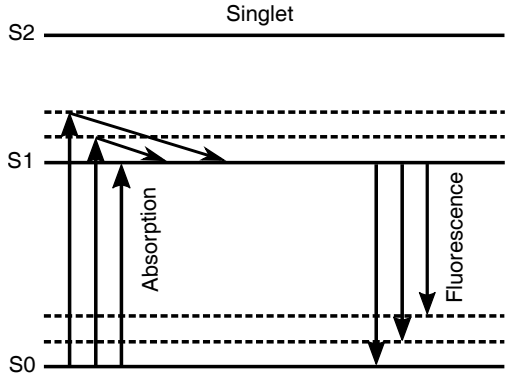


Figure 5.2: Schematics of the fluorescence mechanism in organic scintillators [37].

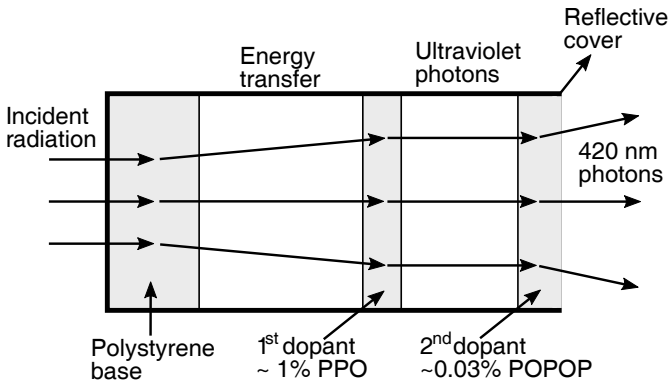


Figure 5.3: Diagram of mechanisms inside a plastic scintillator [37]. Two dopants are mixed homogeneously in a plastic base.

5.1.2 Optical fiber

While the mean attenuation length of photons produced in the scintillator is between 5 and 25 cm, the scintillator strip length is 4 m. To solve this problem a WLS optical fiber is used to collect the scintillation light, re-emit photons with longer wavelengths, and conduct it to a photodetector. These optical fibers have a radius of 1.2 mm and are coupled to the scintillator strip with optical cement.

Fig. 5.4 shows the schematics of the scintillator and optical-fiber system. Blue photons of 420 nm produced in the scintillator are reflected by the TiO_2 layer and eventually are absorbed by the optical fiber and re-emitted as green photons with $\sim$ 500 nm in typical decay times between 3 - 10 ns. In this way, light can propagate several meters through the optical fiber until it is collected by a photodetector. However, optical fibers produce attenuation, so the amount of light that arrives at the photodetector depends on how far away from the photodetector the particle enters the fiber.

The optical fiber used in UMDs is the Saint-Gobain BCF-99-29AMC multi-clad WLS fiber [197]. It consists of a polystyrene core doped with a fluorophore, and cladded with a thin transparent film (2-5 μm) of PMMA. Photons can travel several meters through this fiber thanks to the total internal reflection at the interface between the core and the cladding. The efficiency to capture photons is $\sim$ 6% since the refractive index of the core and first cladding is 1.60 and 1.49 respectively. This value is improved up to $\sim$ 10% using a second EMA cladding with a refractive index of 1.42. In Fig. 5.5 a diagram of the WLS fiber is presented.

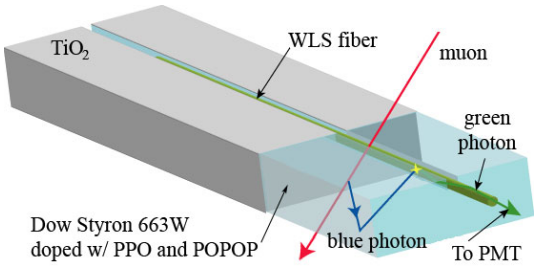


Figure 5.4: Schematics of the UMD scintillator with embedded optical fiber. Blue photons emmited by the scintillators are absorbed with the WLS optical fiber and re-emitted into green photons that are propagated towards a photo-detector [198].

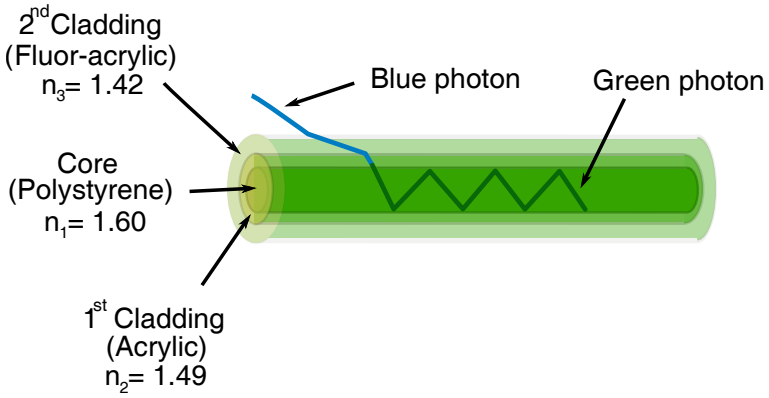


Figure 5.5: WLS optical fiber diagram [37].

5.1.3 From photomultiplier tubes to silicon photomultipliers

In the engineering phase, UMDs were operated with multi-anode PMTs Hamamatsu H8804-200MOD. The first AMIGA physics results were published [199, 200] using more than a whole year of calibrated PMT data. After several considerations, it was decided to replace the PMTs by arrays of SiPMs. The main reasons were that SiPMs consume ten times less power and are up to four times cheaper than PMTs. The SiPM array has no crosstalk between pixels and its photodetection efficiency ($\sim$ 40%) [201] is higher than PMT's quantum efficiency ($\sim$ 20%) [202] at the emission peak ($\sim$ 492 nm). However, the use of SiPM has other kind of problems to consider. For example, SiPMs have a high thermal noise which can be filtered with the read-out electronics and the data analysis. Also, the SiPM gain strongly depends on on the SiPM temperature, thus a self-calibration and temperature compensation are needed.

The first station equipped with SiPM was deployed in December 2016 and the deployment of SiPMs in the whole AMIGA engineering array was completed in January 2018

5.1.4 Silicon photomultipliers

SiPMs are arrays of avalanche photodiodes (APDs) operated in Geiger mode (G-APDs). They are based on the working principle of p-n junctions, where the n-type dopant is rich in free charge carriers (e^-) and the p-type dopant is rich in holes. A typical APD consists on two p-type regions (p and a lightly doped π) between a heavy doped p-type (p^+) and a heavy doped n-type (n^+) as shows Fig. 5.6-top. At the p-n interface, the n-type becomes positively charged while the p-type becomes negatively charged. This produces an accumulation of charge and a strong electric field, forming a depletion layer. If a reverse bias voltage

is applied at the side of the junctions, very little current flows until it reaches a critical level where the diode breaks down. Once the electric field intensity increases beyond a critical level, corresponding to the breakdown voltage V_{br}, the p–n junction depletion zone breaks down and current begins to flow, usually by an avalanche breakdown process. The electric field in an APD after applying a reverse bias is shown in Fig. 5.6-bottom. There are two different regions: the absorption region R_{abs} produces electron-hole pairs, which are accelerated by the electric field. And the multiplication region R_{mult}, also called avalanche region, generate a macroscopic current due to the exponential generation of electron-hole pairs.

APDs can be operated in different modes depending on the applied reverse bias V_{bias} [203] as is shown in Fig. 5.7. If $V_{bias} < V_{APD}$ the device behaves as a photodiode and the gain is 1. It means that a single electron is collected for each electron-hole pair created. This mode is useful to measure high-levels of light. If $V_{APD} < V_{bias} < V_{br}$, the detector runs in a linear mode, commonly called APD mode. For every electron-hole pair produced in the depletion region, a finite multiplication is generated. A higher V_{bias} results in a higher gain. This mode can be used for detecting low level of light but not for photon counting due to the large gain fluctuation. When $V_{br} < V_{bias}$, the APD is on *Geiger-mode*. A charge carrier generated in the depletion region can trigger a diverging avalanche multiplication of carriers. This avalanche does not stop by itself and must be quenched by the voltage drop across a high-ohmic serial resistor or by an active quenching circuit. Geiger mode provides a uniform output, which mainly depends on the applied V_{bias} and the equivalent quenching resistor.

SiPMs are arrays of G-APDs, which can be triggered once within its recovery time. Therefore, the SiPM output depends on the number of triggered cells and not on the number of electron-hole pairs produced by the incident radiation. The number of triggered cells is commonly refer as *photon-equivalents* (PEs). SiPMs used in the UMDs have a total of 1584 cells [201], so each SiPM can detect a maximum of 1584 simultaneous photons. However, the probability of photons hitting more than once on the same cell increases with the number of incident photons and the SiPM saturation can be produced by photons that pile up in the same cell.

Other relevant characteristics of SiPM are the following: The probability for a SiPM to detect an incident photon is called *photodetection efficiency* (PDE) and depends on the direction of the incident light and voltaje applied. Optical *crosstalk* between cells may occur when electrons are recombined during the avalanche emitting photons, which hit a neighbor cell producing a simultaneous avalanche giving two PEs instead of one. *Afterpulsing* by electrons may be produced when electrons in an avalanche are trapped into silicon impurities and get released after the avalanche was quenched giving a second avalanche. A dark count can also originate an afterpulsing event or crosstalk. Once a cell breaks down, the time that takes for detection efficiency to recover is known as *recovery time*.

5.1.5 Read-out electronics

The readout electronics of the UMD modules implements both *binary* and *integrator* modes. In this thesis only the binary mode is considered because it is used for data analysis in Ch. 6.

The *binary* mode independently handles the 64 SiPM signals through a pre-amplifier, fast-shaper and a discriminator, built within each channel of two 32-channel Application-Specific Integrated Circuits (ASICs), in this case CITIROCs [205] furnished by WEEROC. The discriminator signal is sampled at 320 MHz (3.125 ns sample time) with a Field-Programmable Gate Array (FPGA) into 64 2048-bit traces, giving a trace length of 6.4μs. In this way, SiPMs pulses are converted into binary traces. In each trace, a "1"-bit is set if the signal is above the discriminator threshold and a "0"-bit otherwise. The discriminator threshold is set at 2.5 PE

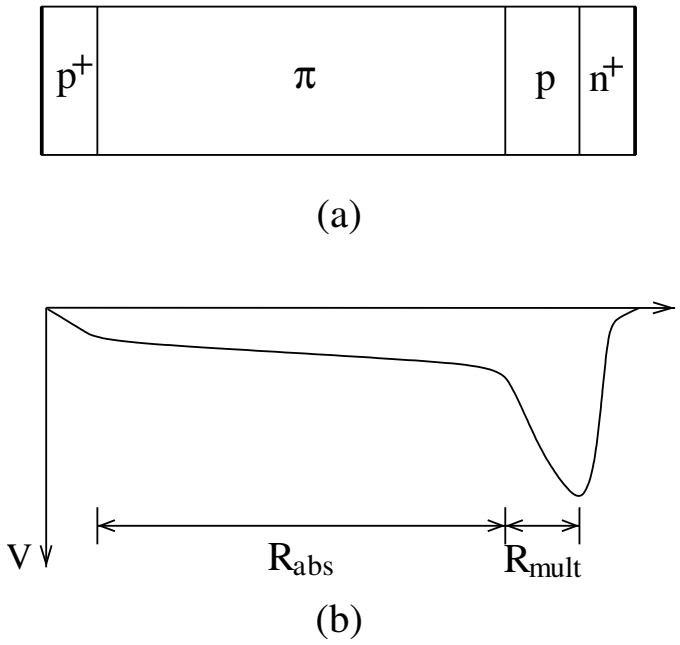


Figure 5.6: Top: schematics of APD design. π represents a lightly doped p-material while the superscript (+) refers to heavy doping. Bottom: profile of the electric field of the APD after applying a reverse bias voltage [204].

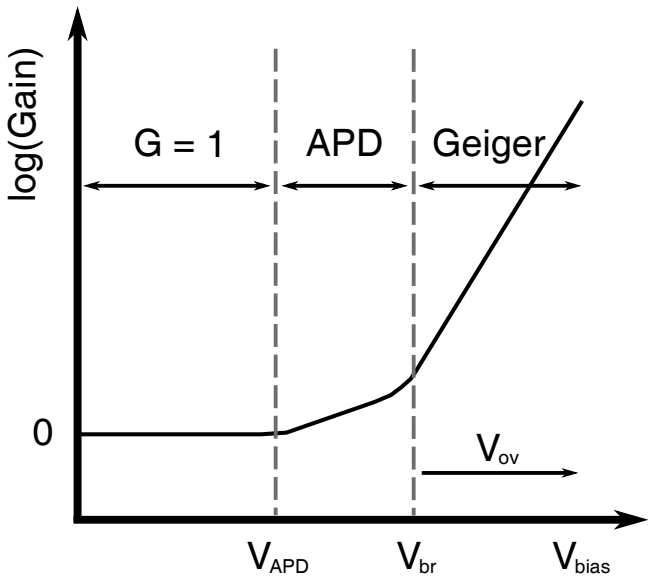


Figure 5.7: APD operation modes as a function of the applied voltaje [37].

to reject most of the SiPM dark rate [206]. Muons can be identified as sequences of "1"s in the binary traces. A simulated single-muon trace is shown in Fig. 5.8.

It is worth noting that the *binary* mode is limited by the detector segmentation since two muons arriving at the same strip simultaneously will be counted as a single particle. This effect limits the number of muons that can be detected at the same time.

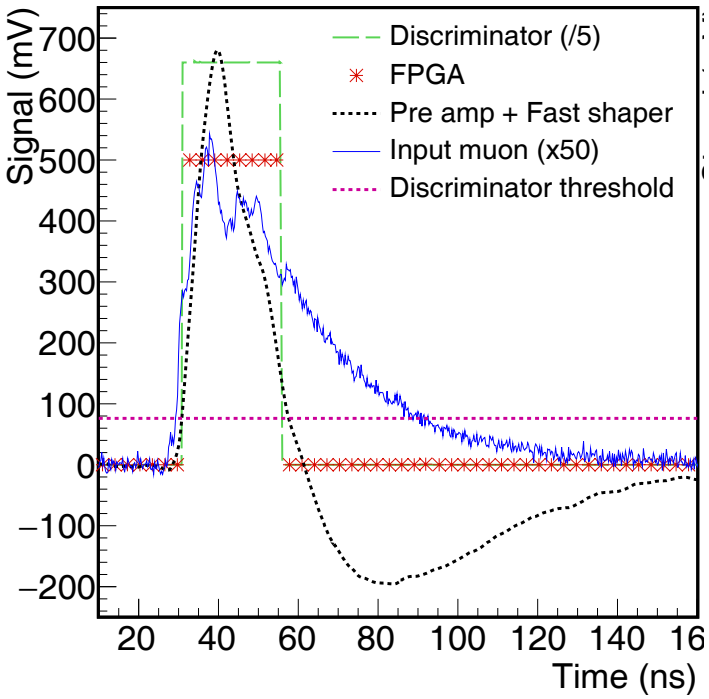


Figure 5.8: Example of a simulated single-muon trace in the *binary* mode [37]. The FPGA outputs a "1"-bit in the binary trace when the fast-shaper output is above the discriminator threshold.

5.2 Calibration

The calibration of the binary channel [206] consists of determining each SiPM V_{bias} and discriminator threshold to obtain an equalized response in the whole SiPM array. By measuring the photo-equivalent peak (PE_{peak}) as a function of V_{bias}, the V_{br} of each SiPM is obtained. The final V_{bias} is set to be 3.5 V larger than the the V_{br} and the discriminator threshold is set to 2.5 PE. During operation, the V_{bias} is re-adjusted to compensate for the temperature fluctuations.

Dark-rate pulses can be used to calibrate SiPM thanks to their well-defined PE spectrum. The PE amplitude spectrum can be obtained by measuring the dark rate as a function of an amplitude threshold. The presence of a plateau corresponds to the transition between PE peaks as shown in Fig. 5.9-left. On the other hand, the transition between plateaus corresponds to the PE amplitude that can be obtained by means of the derivative curve. In Fig. 5.9-right the measured dark-rate curve (blue) and its derivative (green) with a high-voltage of 55.54 V is shown.

5.3 Signal characterization and counting strategy

In Sect. 5.1.3 was mentioned that the UMD engineering array has been operating with SiPMs since January 2018. To analyze this newly acquired data, it is necessary to find proper counting strategies to maximize the signal-to-noise ratio. By employing laboratory measurements with a setup similar to a UMD module and using a muon telescope to trigger the signal, the generated data and noise were analyzed [206]. An example of an event output in this setup is shown in Fig. 5.10. Once the T1 trigger is received, the muon signal is outputted as a sequence of positive samples.

SiPM muon signals are expected to have no structure and be compact, meaning they have no 0s in between its positive samples [206]. For this reason, the muon width is defined as

the distance (in samples) between the first and last positive sample. The number of positive samples ("1"s) in the binary traces as a function of the signal width along with the signal width histograms are shown in Fig. 5.11. The signal from the strips is shown in black, while the signals from the passive channels (SiPM noise) and the ones from the strips outside the trigger scope are shown in red and blue, respectively. Most of the strip signals are distributed along an identity function. However, data points away from the identity are also found and correspond to muon signals with noise at the beginning or end of the trace. The signal mean widths (μ) and standard deviations (σ) are also shown. The muons signal widths is $\mu = 7.8$ samples (24.4 ns) and $\sigma = 1.5$ samples (4.7 ns). An inhibition window of 12 samples (37.5 ns) contains more than 99% of the signals. On the other hand, the signal width at the passive channels is $\mu = 3.0$ samples (9.4 ns) and $\sigma = 0.7$ samples (2 ns). Therefore, identifying a muon as a signal with more than three positive samples will reject more than 95% of the SiPM noise. In addition, with these counting patterns, wide signals such as those produced at the beginning of the strips do not introduce over-counting as they need 16 positive samples to be counted twice.

5.4 Binary mode optimization

There are two types of bias sources in the *binary* mode: one is related to the timing of particles in the shower and the other one related to the shower geometry.

Pile-up

In order to reconstruct the number of muons, it is necessary to define an inhibition window to avoid counting the same muon more than once. However, when two particles reach the same scintillator strip within the inhibition window, the second particle will not be detected and the number of muons will be under-estimated. This effect is known as muon *pile-up*.

Given 64 strips hit by k muons, the probability of missing the k+1 muon within the same inhibition window is $k/64$. Therefore, the probability of under-counting depends on the number of impinging muons. This can be considered in the reconstruction with a maximum-

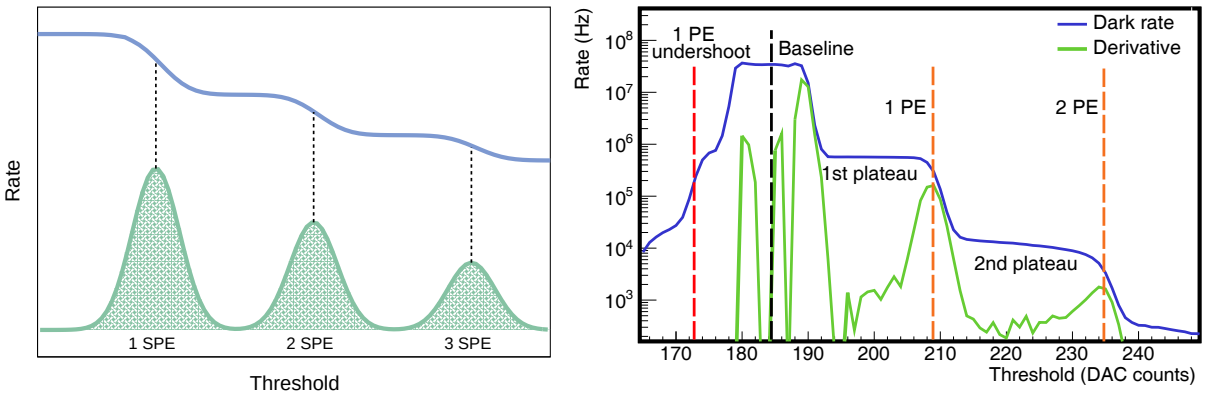


Figure 5.9: Left: Schematic of the dark-rate curve as a function of an amplitude threshold (blue); its derivative (green) shows the PE amplitude spectrum. Right: Dark-rate curve measured as a function of the discriminator threshold (blue). The derivative of this rate curve (green) readily identifies the PE amplitude spectrum. [37]

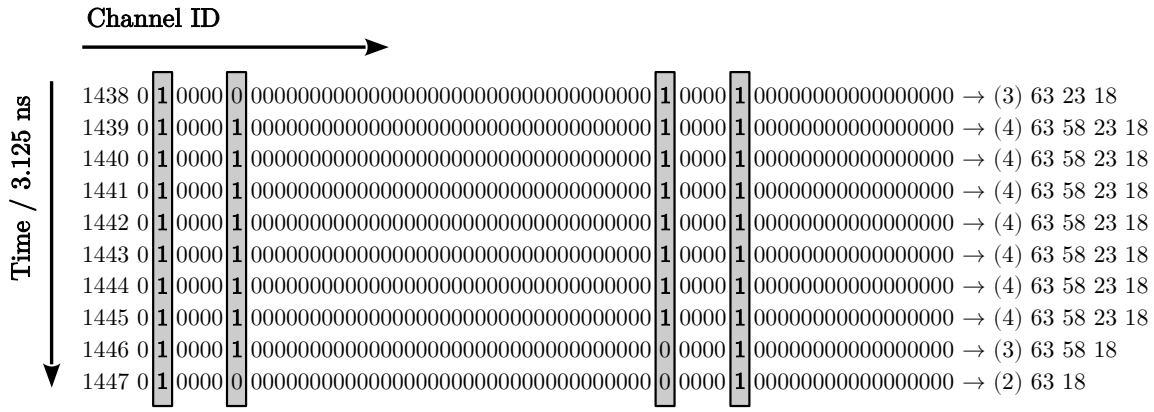


Figure 5.10: An example of a typical event, composed by binary traces from available channels. The muon signal is visible as sequences of positive samples in the four active channels [37].

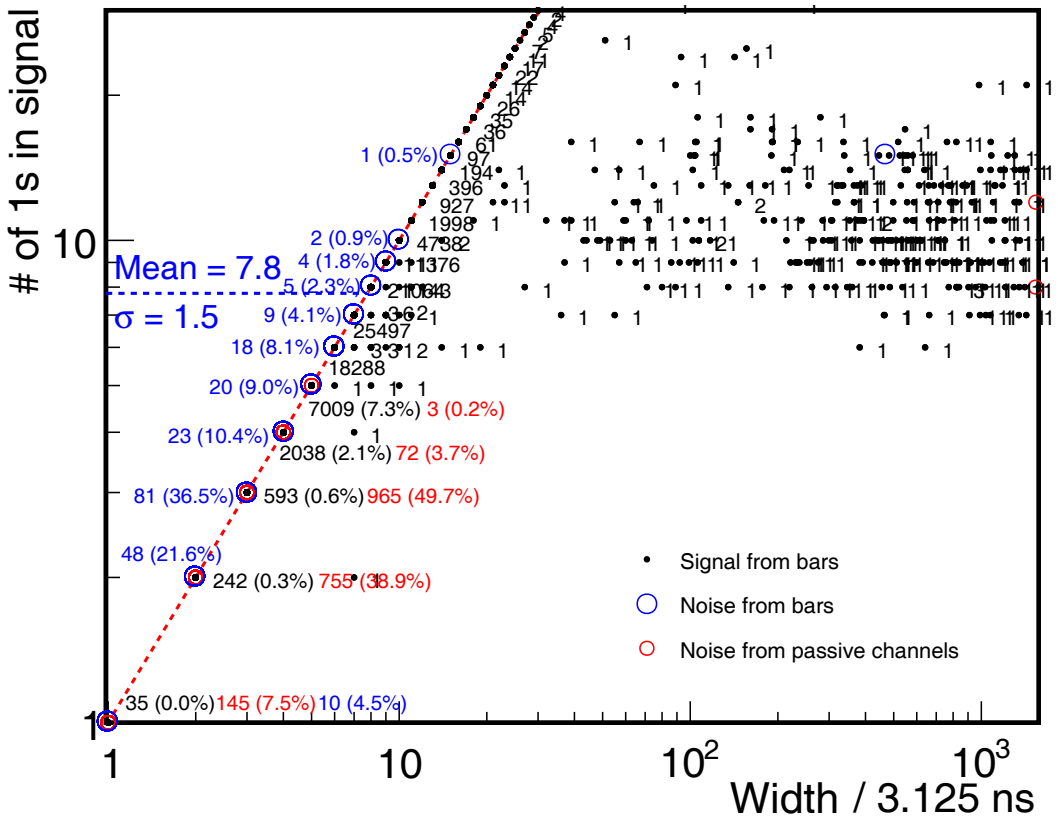


Figure 5.11: Number of positive samples ("1"s) as a function of the signal width. In black, the total signal of the channels with strips is shown. In red, the signal from the passive channels, corresponding to SiPM noise. In blue, the signal from the strips outside the trigger scope. Most of the signals are located along an identity function, indicated with the red dashed line [37].

likelihood approach assuming a binomial distribution [207]. In a 6.4 μs trace and assuming an inhibition window of 37.5 ns (12 samples) (see Sect. 5.3), there are $\sim$ 170 windows.

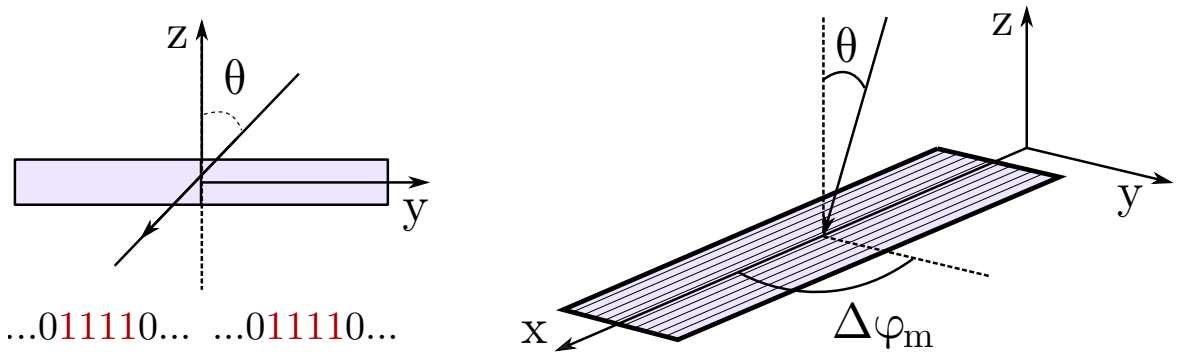


Figure 5.12: Left: Schematic of frontal view of two scintillator strips and an impinging muon depositing energy in both. An example of the binary output for this event is also shown. Right: Schematic of the geometric dependency of corner clipping muons, θ and $\Delta\varphi_m$ [37].

The total number of muons is estimated by summing over all time windows and it is given by:

$$\hat{\mu} = -64 \sum_{i=1}^{170} \ln\left(1 - \frac{k_i}{64}\right). \tag{5.1}$$

where k_i is the number of muons counted for the ith window.

Corner clipping muons

Corner clipping muons are muons that deposit energy in more than one scintillator strip. These muons may produce under- or over-counting and therefore might be interpreted as 0 (not enough energy in either channel), 1 (enough energy in a single channel) or 2 (enough energy in both channels) muons. The probability of corner clipping depends on the direction of the shower axis as Fig. 5.12 shows.

The reconstruction bias is determined by two parameters. The first one is the muon inclination angle, which is relative to the zenith angle of the shower axis. The second one is the difference in azimuth between the direction of the shower and the orientation of the module in the ground plane ($\Delta\varphi_m = \varphi - \varphi_m$) . A muon cannot pass through two neighboring strips in the only case when both θ and $\Delta\varphi_m$ are zero. Maximal corner clipping occurs when the muon momentum is perpendicular to the orientation of the scintillator strips ($\Delta\phi_m = 90^\circ$) and increases with θ.

Reconstruction analysis

In order to verify the UMD reconstruction, a comparison of the mean muon lateral distribution function (MLDF) of CORSIKA showers at the UMD depth, before and after UMD reconstruction is done. The inputs are 120 full Monte Carlo proton showers at $10^{17.5}$ eV and $\theta = 0^\circ$ from the KIT server. As CORSIKA showers give particles at ground, each muon was propagated underground $d = 2.3\,\mathrm{m}$ through soil with an average density of $\rho = 2.38\,\mathrm{g\,cm^{-3}}$ and considering only ionization losses by $\frac{dE}{dX} = -1.815\,\mathrm{MeV\,cm^2\,g^{-1}}$. The underground energy E_{ug} for a muon with energy at ground E_g and an angle ϕ between its momentum and the vertical is given by

$$E_{\mathrm{ug}} = E_{\mathrm{g}} + \frac{dE}{dX}\ \rho\ \frac{d}{\cos(\phi)} \tag{5.2}$$

On the other hand, the same showers are passed through the Auger detectors, the SD and UMD arrays, where the detector reconstruction was done by Offline reconstruction

software and each shower was used ten times changing the core position inside the array. Fig. 5.13 shows the underground MLDF obtained directly from the Corsika showers and the reconstructed MLDF by the UMD array. There are significant differences at short distances and over 1000 m where statistical fluctuation increases, so these distances are not used in the further analysis presented in Ch. 6.

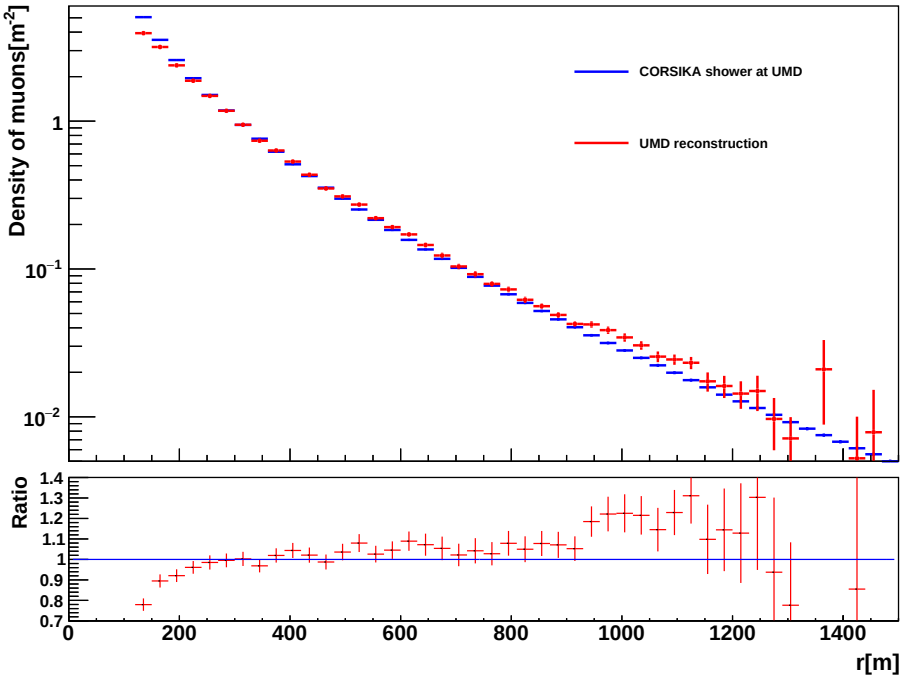


Figure 5.13: MLDF of full MC showers at UMD depth (blue line) and after UMD reconstruction (red line).

CHAPTER 6

Muon Lateral Distribution Analysis

In this chapter, data from the UMD is analyzed in Sect. 6.1 and the corresponding muon lateral distribution function (MLDF) is obtained for several primary energy and zenith angle bins. Then, in Sect. 6.2, the constant intensity cut (CIC) method is applied to data. The comparison of these MLDFs with simulations and Core-Corona cases is done in Sect. 6.3 and 6.4.

6.1 UMD data analysis

The Underground Muon Detector (UMD) is part of the AugerPrime upgrade of the Pierre Auger Observatory and its main feature is the direct measurement of the muon component of EAS. Each station is composed of modules of 10 m^2 with 64 scintillation plastic bars buried beneath about 2.3 m of soil. The muon measurements were performed by the operational UMD stations equipped with SiPMs. The muon counting strategy, by which the muon density was obtained from the binary traces, was based on the recognition of the pattern "1111" [208] with a veto window of 56.25 ns (18 time bins). This veto window has been observed to provide a more accurate estimation of the number of muons when considering the time distribution of muons in the shower front. Lastly, muon measurements performed with saturated UMD stations are not considered in the analysis. A detailed explanation of the UMD is presented in Ch. 5.

The event reconstruction was performed with the official Auger software $\overline{\text{Off}}\underline{\text{line}}$, in particular the trunk revision 33890, using the low energy threshold TOTd-MOPS (see Sect. 1.3.3). The input data set was comprised of reconstructed events with at least one triggered water-Cherenkov detector (WCD) paired with an operational UMD station from January 1st 2018 to August 31st 2020. Events were selected according to the primary energies between 10^{16} eV and 10^{20} eV, with a zenith angle below 45 degrees and without any low-gain saturated SD station. The events are required to have at least a T4 trigger. After these cuts more than 67000 events were accepted.

The number of events over time is shown in Fig 6.1. This figure shows almost a constant number of events over 2018 since at that time the engineering array (EA) was deployed, which is a hexagon of seven UMDs. The data acquisition over 2019 was affected by a major problem in the communication system of the stations, which was completely solved in the beginning of 2020. The deployment of new UMDs over 2020 produced a notable increase in the number of events being its first eight months comparable to the previous two years.

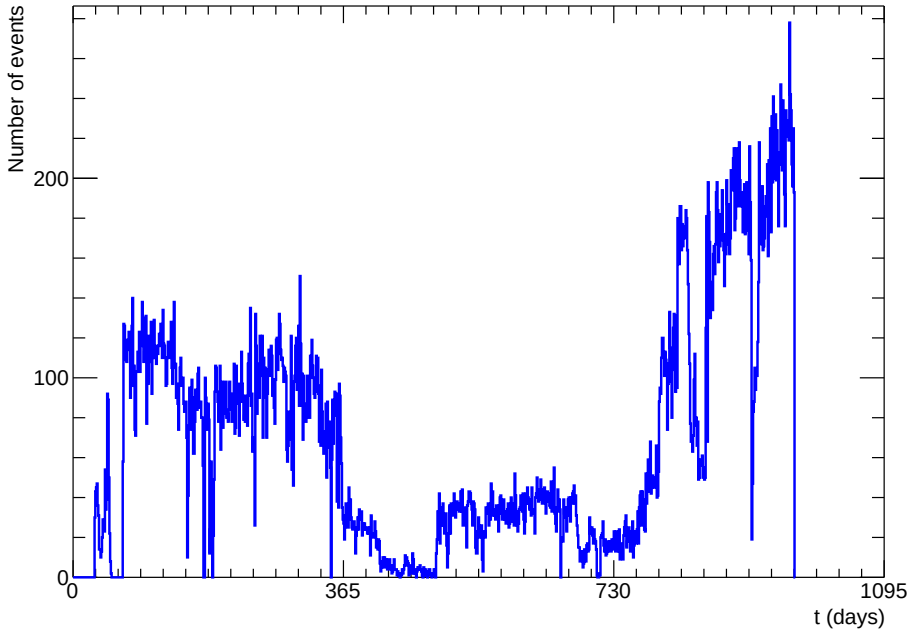


Figure 6.1: Number of events with UMD measurements as a function of days from January 1 2018 to August 31 2020.

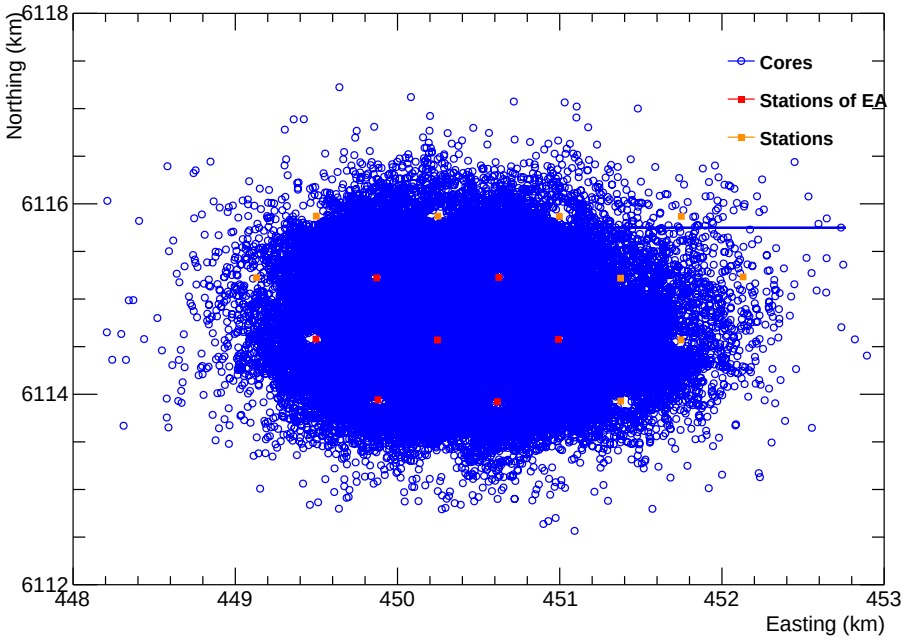


Figure 6.2: Current UMDs deployed and core's positions. Red squares are the UMDs of the EA.

This increase is partially explained because of the larger exposure but the main reason is that the number of hexagons of UMDs increased from one in 2019 to five on August 2020. That is because the number of hexagons is closely related to the number of T4 triggered events. The UMDs and the event core's position of selected events are shown in Fig. 6.2, where the red squares are the UMDs of the EA while the orange squares are new UMDs.

The number of events as a function of primary energy is shown in Fig. 6.3. The maximum number of events is found at log(E/eV)=17.0. However, from Fig. 6.5 the full efficiency of the trigger is only for events with a zenith angle $\theta < 30^{\circ}$. In order to get a full efficiency

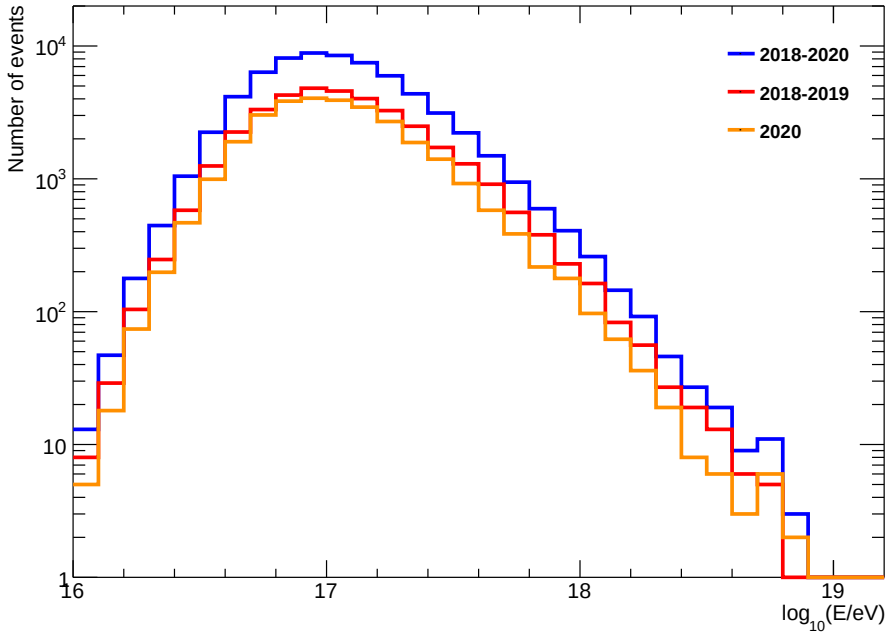


Figure 6.3: Number of UMD events as a function of primary energy. Subset of events of the years 2018-2019 and 2020 are also plotted.

trigger for all analyzed zenith angles, i.e. $\theta < 45^\circ$, only events with energy larger than $\log(E/\text{eV})$=17.2 are considered in the analysis. Fig. 6.3 also shows that for $\log(E/\text{eV})\approx$18.5 the number of events is around ten, so the analysis at these energies is dominated by large fluctuations.

The impact of the full trigger efficiency can be seen in Fig. 6.4 where the number of events as a function of zenith angle is shown. In the case where all energies are considered, the number of events increases with θ due to the larger exposure area up to $\theta < 30^\circ$. For larger θ values the number of events decreases due to lower trigger efficiency. Taking the subset of events with an energy such that $\log(E/\text{eV})$>17.2 the number of events increases up to $\theta = 45^\circ$. The subset of events with energies such that $\log(E/\text{eV})$>17.5 is also shown. That energy was the lowest full efficiency energy used by the old triggers. Both cases show the same profile, as is expected.

The reconstructed events are divided into primary energy bins of 0.15 in logarithmic scale and in ten equal-exposure zenith bins. The muon density measured by each UMD and its distance to the shower axis r give the muon density $\rho_\mu(r)$. The muon density $\rho_\mu(r)$ is averaged over each bin (E, θ), where r is discretized in bins of 30 m. These bin widths are roughly in accordance with the reconstruction resolution of the SD-750 [210]. The average profile $\rho_\mu(r|E,\theta)$ is fitted through a χ^2 minimization of the function

$$\rho_\mu^{model}(r|E,\theta) = \rho_0 \left(\frac{r}{r^*}\right)^{-\alpha} \left(1+\frac{r}{r^*}\right)^{-\beta} \left(1+\left(\frac{r}{10r^*}\right)^2\right)^{-\gamma} \tag{6.1}$$

which is known to suitably represent the MLDF measured by the UMD [211]. To reduce the fit uncertainty, fixed values were assigned to some of the parameters, namely $r^* = 320$ m, $\alpha = 0.75$ and $\gamma = 3.0$, thus leaving ρ_0 and β as free parameters. A set of requirements is imposed to select only bins with a valid MLDF: the fitting process must converge with a non-null number of degrees of freedom (ndf), with positive values for the two free parameters and with a p-value larger than 0.01. If the MLDF does not satisfy any of these conditions, the fitting process is repeated changing the fitting range $(r_{\text{min}}; r_{\text{max}})$. The minimization is

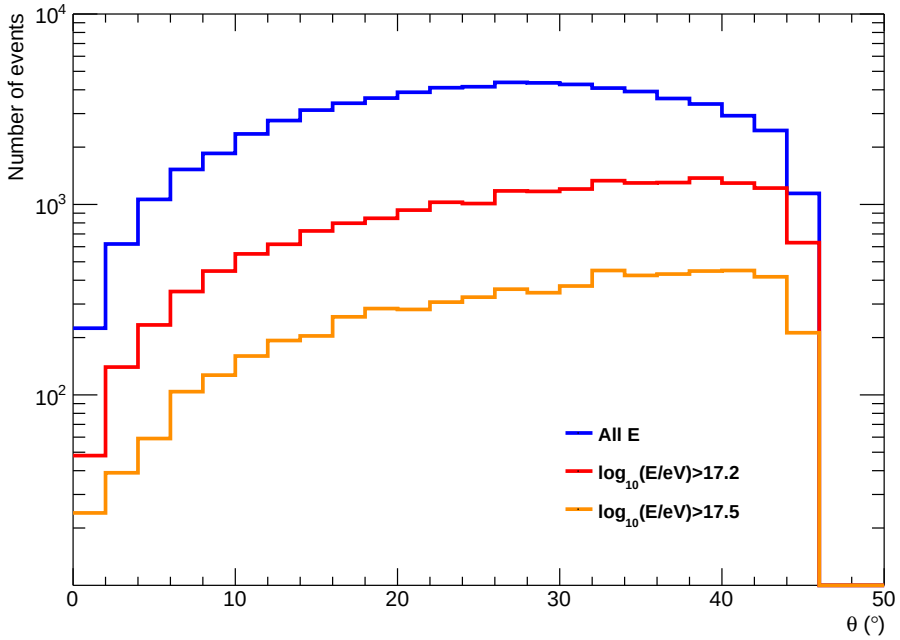


Figure 6.4: Number of UMD events as a function of zenith angle for all events and for two energy cuts.

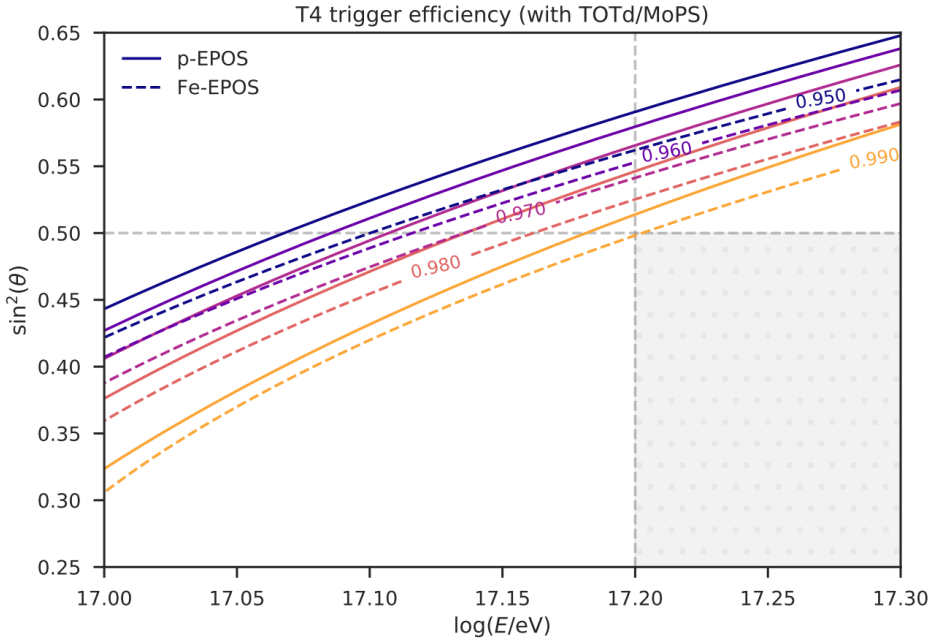


Figure 6.5: T4 trigger efficiency for proton and iron showers for EPOS LHC in a zenith angle vs energy plot. Plot by F. Gesualdi from [209].

performed extracting the first bin with entries to avoid fluctuation due to saturated stations, and successively the last bins to avoid large fluctuations due to low entries. Each accepted MLDF is fully expressed by ρ_0 and β with the associated uncertainties rising from the fitting procedure. As an example, one accepted MLDF is displayed in Fig. 6.6.

The fitting results show that r_{min} increases with shower energy from 120 m for the energy bin $(10^{17.2};10^{17.35})$ eV to 240 m for the bin $(10^{18.1};10^{18.25})$ eV. And r_{max} is 600 m for the energy bin $(10^{17.2};10^{17.35})$ eV and 990 m for the largest bins where the muon density is around the

UMD resolution, i. e. one muon detected by the $10\,m^2$ detector gives 0.1 muon m^{-2}. Taking these ranges, all zenith bins are accepted from $10^{17.2}$ eV to $10^{18.1}$ eV as Fig. 6.7 shows, while several larger energy bins are rejected due to the lack of statistics. The reduced χ^2 value of almost all the accepted MLDFs is between 0.5 to 2 as Fig. 6.8 shows. As an example, some of the accepted MLDFs for a given zenith angle bin are shown in Fig. 6.9 and some of the accepted MLDFs for a given energy bin are shown in Fig. 6.10.

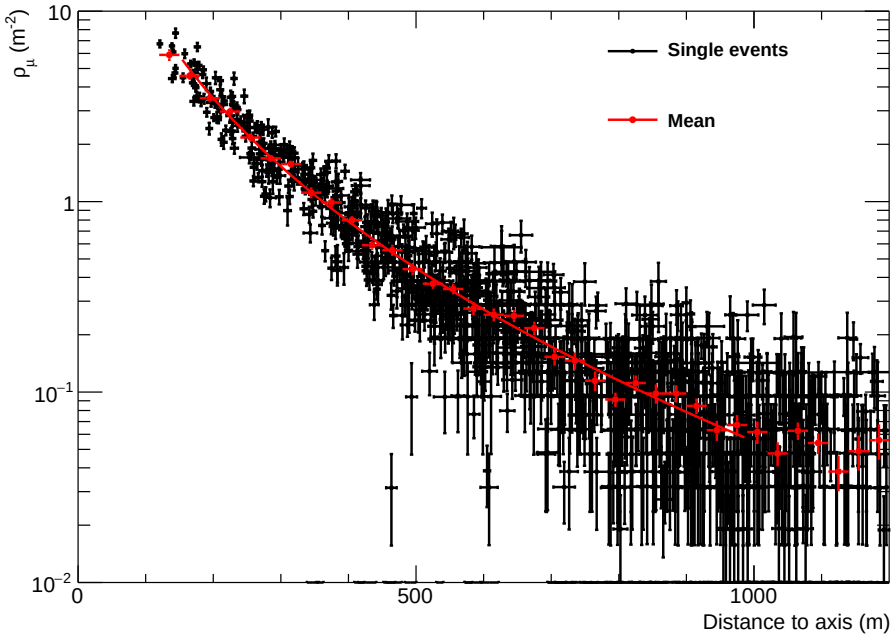


Figure 6.6: One of the accepted MLDFs, corresponding to E = ($10^{17.5}$ eV - $10^{17.65}$ eV) and $\theta = (0^\circ - 13^\circ)$. Black (red) dots represent single events (averaged over distance bins), while the red line corresponds to the fitted model of Eq. 6.1.

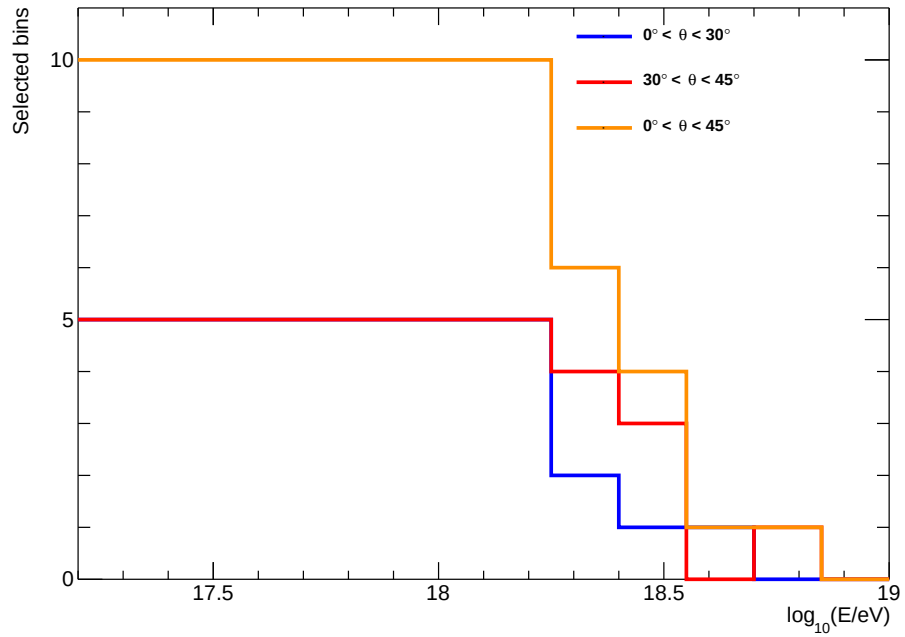


Figure 6.7: The number of selected bins with valid MLDFs in terms of the primary energy bin.

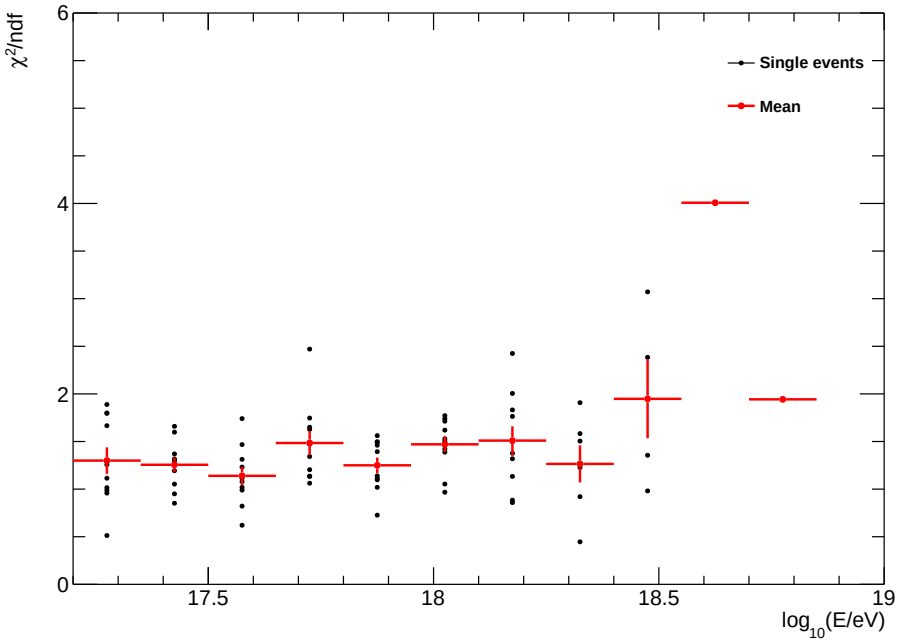


Figure 6.8: The reduced χ^2 value of the selected MLDFs in terms of the primary energy.

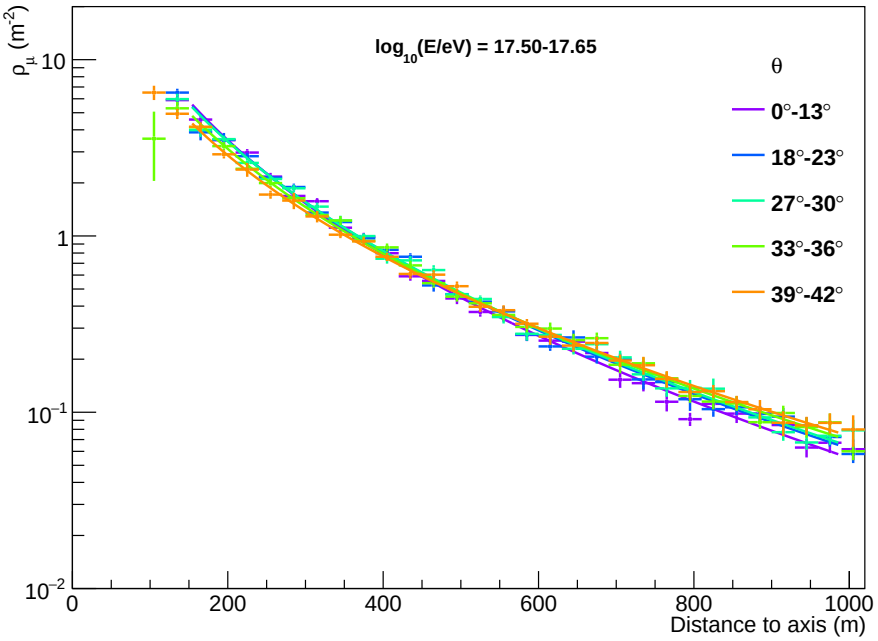


Figure 6.9: Accepted MLDFs of the energy bin $E = (10^{17.5} - 10^{17.65})$ eV.

Fig. 6.9 shows how the MLDFs depend strongly on the scale parameter ρ_0 while β changes slightly. This behavior is confirmed in Fig. 6.11 and 6.12, which show that ρ_0 increases linearly with the logarithm of the energy while β is almost constant for a given angular band. The zenith angle dependence of ρ_0 reflects a well-known attenuation feature of the atmosphere and the soil shielding of the muonic EAS component.

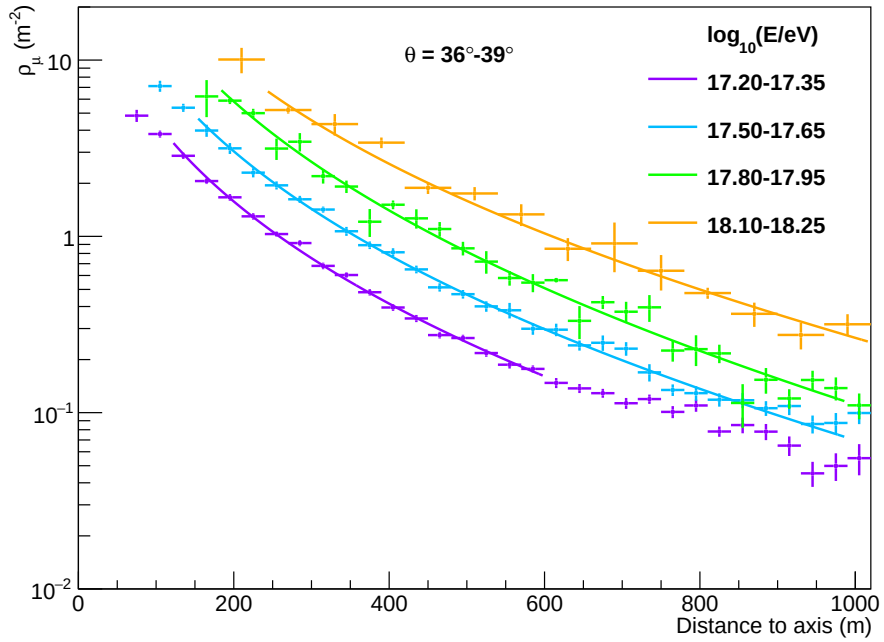


Figure 6.10: Some of the accepted MLDFs of the zenith angle bin $\theta = (36^\circ - 39^\circ)$.

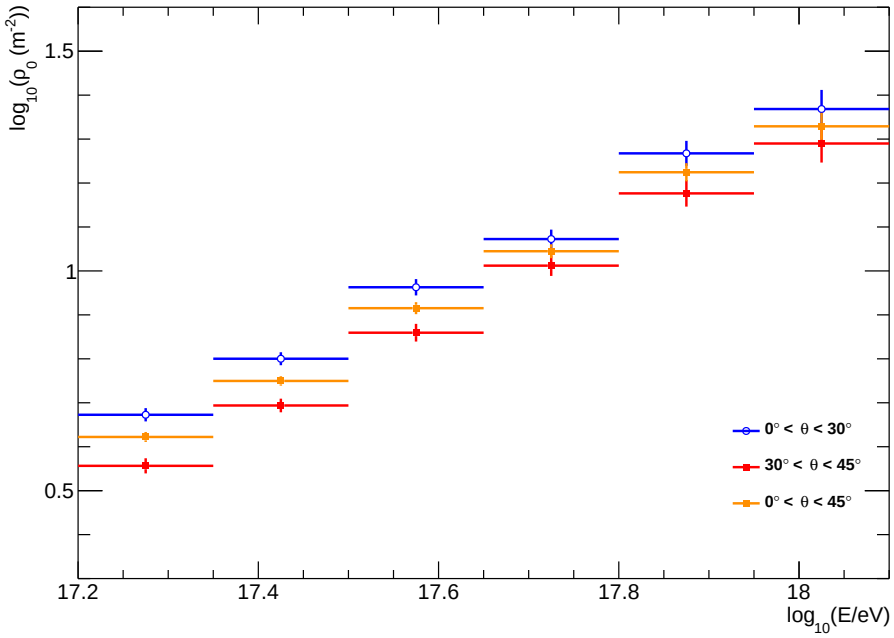


Figure 6.11: The normalization ρ_0 of the MLDF in terms of primary energy in different angular bands.

In fact, from Fig. 6.12, β has no dependence with energy at first order considering the fluctuations. This assumption allows to parameterize the β dependence with θ following the Linsley's elongation rate theorem

$$\beta(\theta) = b_0 + b_1 sec(\theta). \tag{6.2}$$

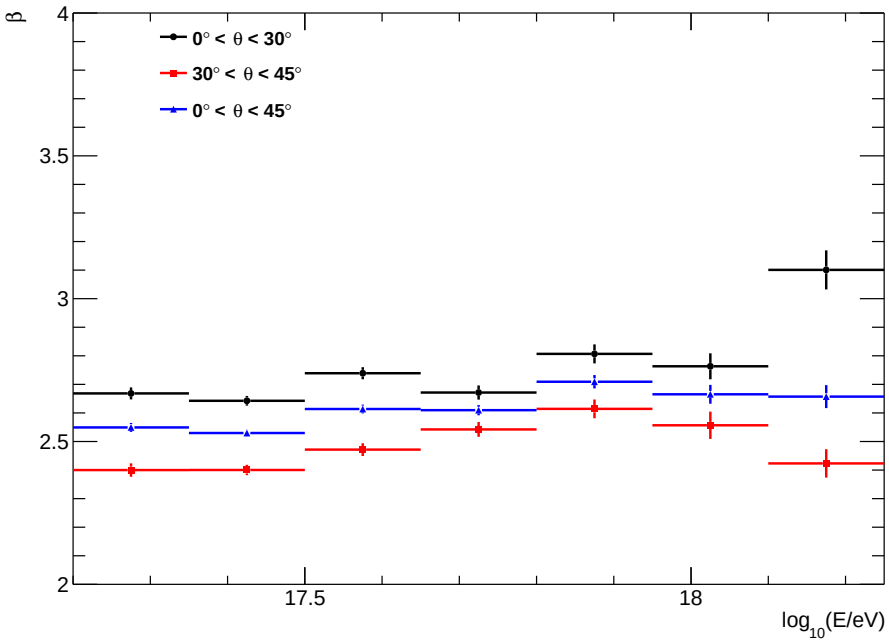


Figure 6.12: The evolution of the weighted average MLDF slope β as a function of primary energy considering all zenith angles together and discriminating by angular bands.

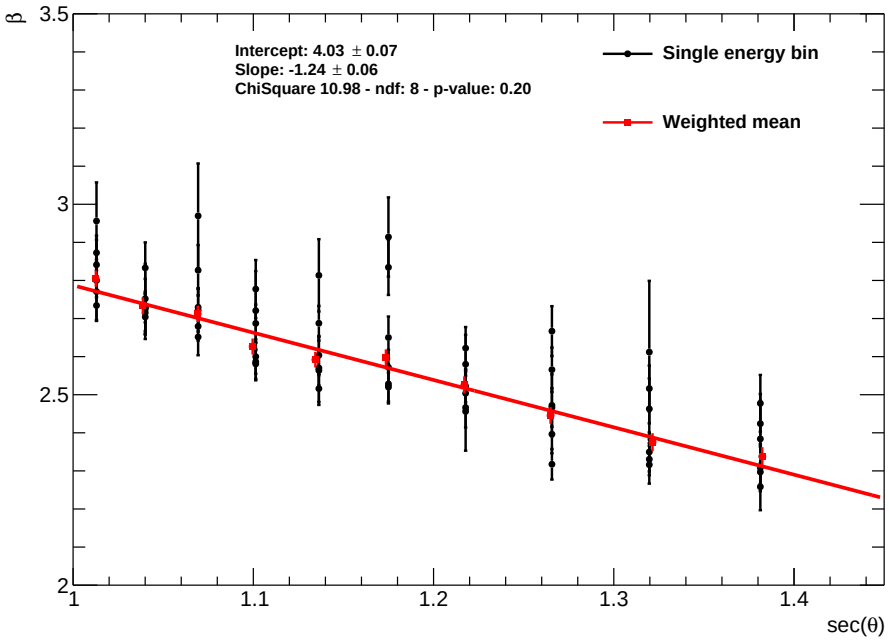


Figure 6.13: The evolution of the weighted average MLDF slope β in terms of the secant of the zenith angle considering the selected bins with energies between $10^{17.2}$ eV and $10^{18.1}$ eV. The red line represents the fitted model of Eq. 6.2.

The parameters are obtained by fitting the β values of all the accepted MLDFs for energy bins between $10^{17.5}$ eV and $10^{18.1}$ eV . From Fig. 6.13 the parameters are $b_0 = 4.03 \pm 0.07$ and $b_1 = -1.24 \pm 0.06$

On the other hand, Fig. 6.10 shows that the MLDFs of different zenith angle bins for a given energy bin have a similar muon density around $r = 450\,\mathrm{m}$ while the β value changes. For that reason the average MLDF is commonly parameterized as

$$\rho_{\mu}^{\mathrm{model}}(r, E, \theta) = \rho_{450}(E, \theta)\, f(r, \theta) \tag{6.3}$$

where the normalization parameter is taken at a radial distance of 450 m since the fluctuations are reduced at this value and the average structure function $f(r, \theta)$ is normalized at $r = 450\,\mathrm{m}$

$$f(r, \theta) = \left(\frac{r}{r^*}\right)^{-\alpha} \left(1 + \frac{r}{r^*}\right)^{-\beta} \left(1 + \left(\frac{r}{10r^*}\right)^2\right)^{-\gamma}. \tag{6.4}$$

6.2 Constant intensity cut method

The value of the average MLDF at 450 m from the shower axis, $\rho_{450}(E, \theta)$, is linked to the primary energy and composition. Since muons are partially absorbed in the atmosphere and in the soil shielding the detector, there is an intrinsic modulation dependent on the zenith angle acting on any muon estimator. Specifically, ρ_{450} can be factorized into an energy-dependent estimator $\rho_{\mathrm{ref}}(E)$ and the mentioned angular dependence $f_{\mathrm{att}}(\theta)$ as

$$\rho_{450}(E, \theta) = \rho_{\mathrm{ref}}(E) f_{\mathrm{att}}(\theta). \tag{6.5}$$

The magnitude of the angular component can be found with the well-known constant intensity cut method (CIC) [212] which is based on the assumption that the CR flux is isotropic. This requirement is vastly verified considering that no significant anisotropies in the CR arrival directions have been identified below 10^{18} eV [213]. The CIC method is applied to correct each value of $\rho_{450}(E, \theta)$ as if the EAS were coming with a reference zenith angle of $\theta_{\mathrm{ref}} = 35^{\circ}$ as for the SD-750 [210].

For the CIC correction all events are reconstructed a second time by means of the MLDF defined in Eq. 6.3 using the β parameterization from Eq. 6.2. The values of ρ_{450} for each event are estimated in the standard Offline reconstruction by minimizing a likelihood function where all UMDs are taken into account, including those saturated and non triggered.

Then, the events are divided into the same ten zenith bins of equal exposure used before. If these values for all MLDFs in each zenith bin are sorted in decreasing order, a fixed numbered position (i.e. constant intensity) should represent the same average muon density or, likewise, the muonic component of EAS generated by the integral CR flux, if no zenith dependence were present.

The number of events with a ρ_{450} value greater than a certain threshold $\rho_{450}^{\mathrm{cut}}$ is defined as *intensity* n and is shown in Fig. 6.15. A larger number of events is observed at low values of $\rho_{450}^{\mathrm{cut}}$ for smaller zenith angle bins. This is due to the attenuation of low-energy and non-vertical events. As the cut value is increased, all profiles show a consistent drop with similar rates. A horizontal cut at a constant intensity n_{cut} corresponds to different $\rho_{450}^{\mathrm{cut}}$ values in each zenith bin since the profiles do not lie on top of each other. If no attenuation effect were in play, $\rho_{450}^{\mathrm{cut}}$ would be independent of the zenith angle at a fixed n_{cut}. From Fig. 6.15, the $\rho_{450}^{\mathrm{cut}}$ value for each n_{cut} and zenith angle bin is calculated and shown in Fig. 6.16.

The correction factor $f_{\mathrm{att}}(\theta)$ can be modeled with a quadratic function of $x = \cos^2(\theta) - \cos^2(35^{\circ})$:

$$f_{\mathrm{att}}(x) = 1 + f_1 x + f_2 x^2. \tag{6.6}$$

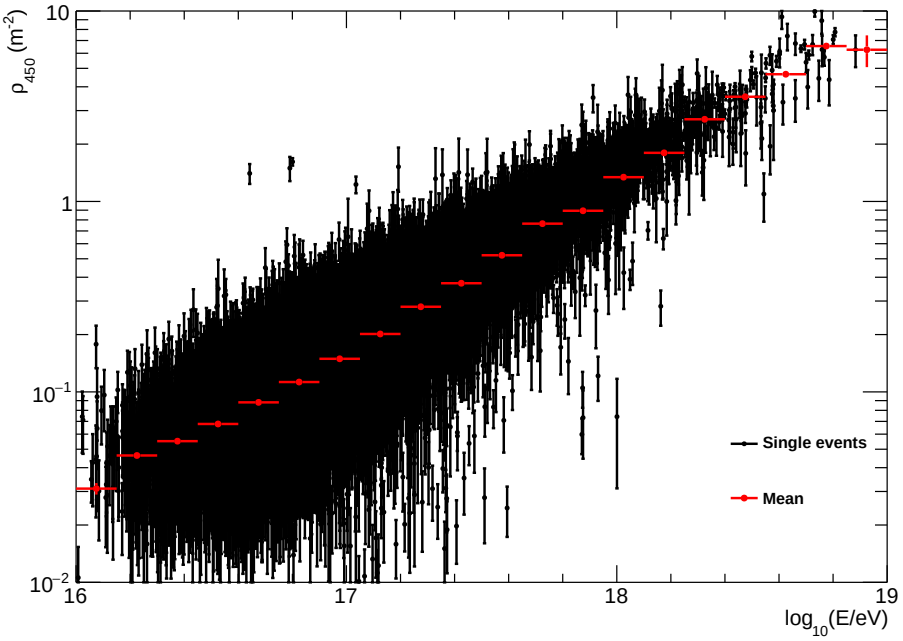


Figure 6.14: The muon density at 450 m from the shower axis obtained from the reconstructed event-wise MLDF using the slope parameterization from Eq. 6.2.

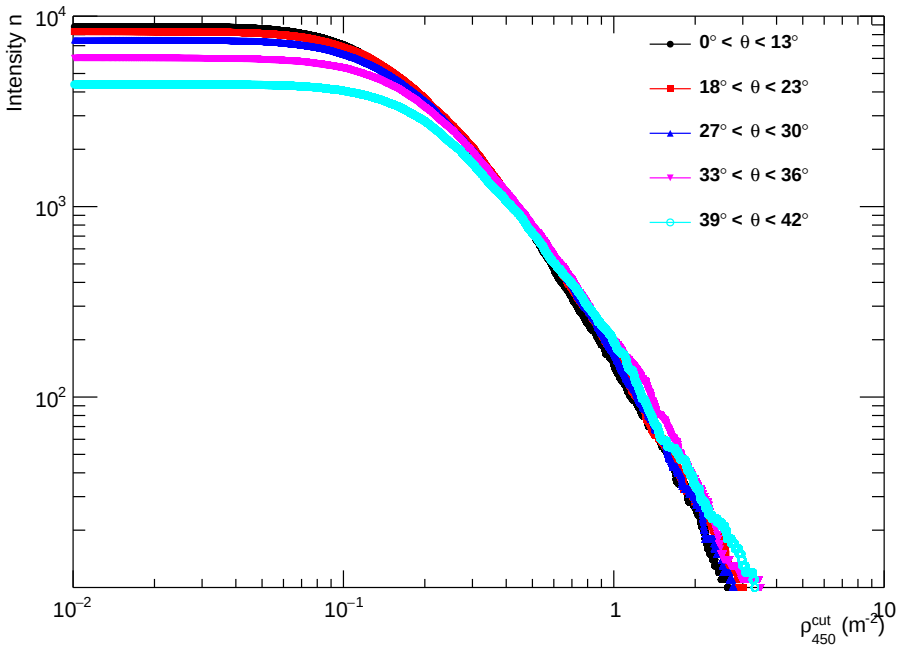


Figure 6.15: The intensity n, i. e. the number of events with values of ρ_{450} above a given threshold ρ_{450}^{cut}, for the ten equal-exposure zenith angle bins.

Fig. 6.17 shows the threshold ρ_{450}^{cut} fitted with a function proportional to $f_{\text{att}}(\theta)$ for different values of n_{cut}. To obtain the attenuation function, the fitted function is normalized by its value at $\theta = 35^{\circ}$.

The parameter values as functions of n_{cut} are shown in Fig. 6.18 where the first parameter monotonically increases while the second one has a range of almost constant values in the interval $n_{\text{cut}} = (800; 1200)$. The p-values corresponding to the fits are presented in Fig. 6.19.

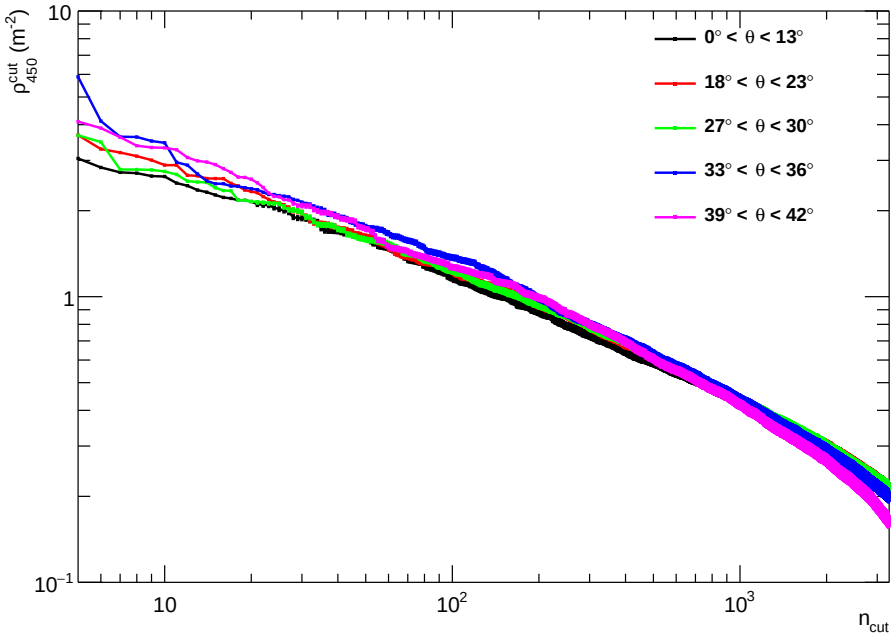


Figure 6.16: The threshold ρ_{450}^{cut} as a function of the constant intensity cut n_{cut}.

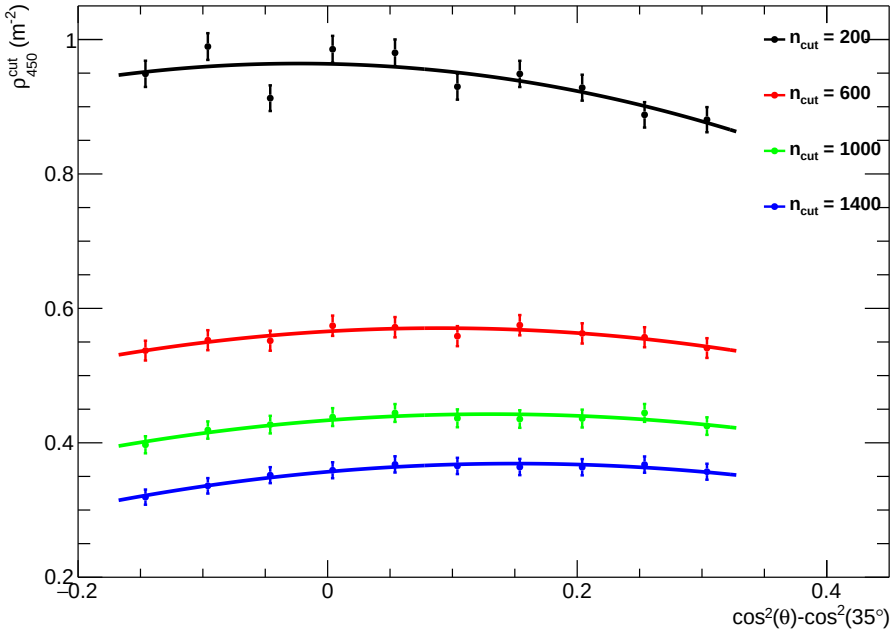


Figure 6.17: The threshold ρ_{450}^{cut} at different constant intensity cuts n_{cut} as a function of $\cos^2\theta - \cos^2 35^\circ$.

They are practically one for $n_{\text{cut}} > 500$. Hence, a weighted average of the parameters in $n_{\text{cut}} = (800; 1200)$ is performed. From Fig. 6.16, these n_{cut} values correspond to $\rho_{450}^{\text{cut}} \sim 0.4\,\text{m}^{-2}$, and this threshold corresponds to shower energies $E \sim 10^{17.4}$ eV. So the attenuation function is optimized to this energy. The values obtained for the parameters are $f_1 = 0.322 \pm 0.005$ and $f_2 = -1.25 \pm 0.03$. The comparison of these values to those found in the EA phase, $f_1 = 0.54 \pm 0.1$ and $f_2 = -1.02 \pm 0.69$, is easier to understand by plotting the attenuation curve in Fig. 6.20. This new data analysis presents a smaller attenuation than the previous

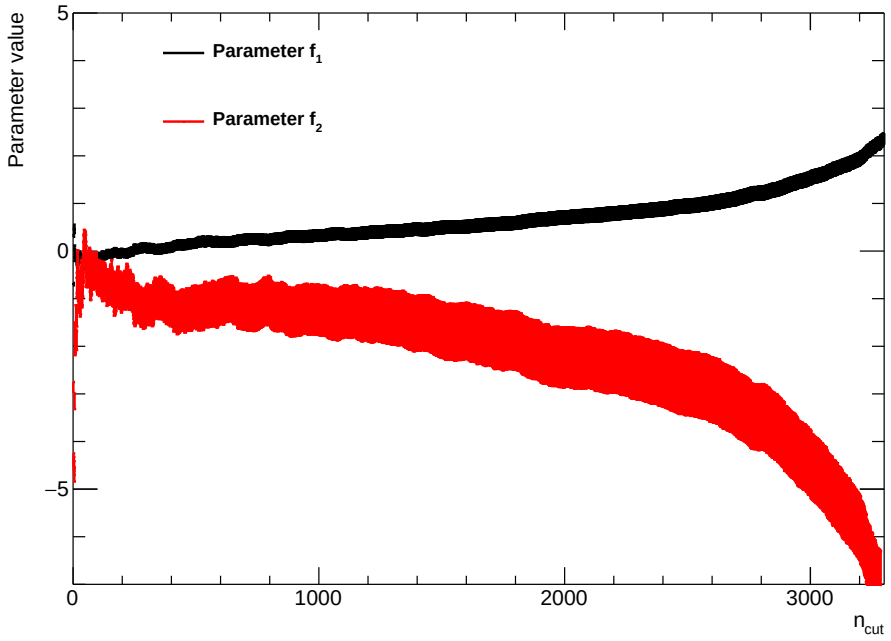


Figure 6.18: The coefficients of the attenuation function f_{att} versus n_{cut}.

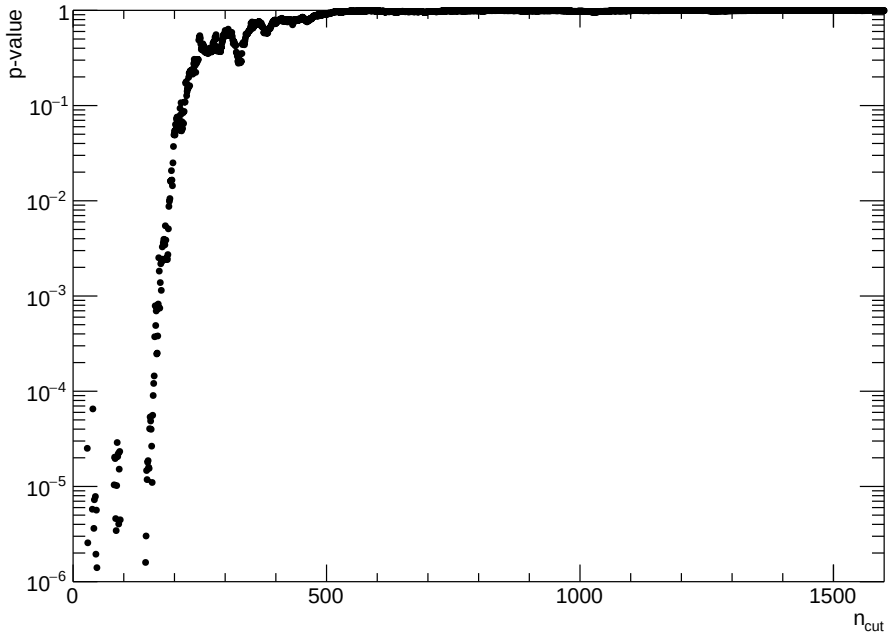


Figure 6.19: The p-value of the quadratic fits to ρ_{450}^{cut} with $\cos^2\theta - \cos^2 35^\circ$ profiles.

study and it has a larger impact on large zenith angles. The attenuation curve is between 0.97 to 1.02 for all zenith angles below $\sim 39^\circ$.

Once the attenuation function is obtained, it is possible to calculate the reference muon density $\rho_{35}(E)$ for all the accepted MLDFs $\rho_\mu(r; E, \theta)$ of Fig. 6.7 by:

$$\rho_{35}(E) = \frac{\rho_\mu(r = 450\,\text{m}; E, \theta)}{f_{\text{att}}(\theta)} = \frac{\rho_{450}(E, \theta)}{f_{\text{att}}(\theta)} \tag{6.7}$$

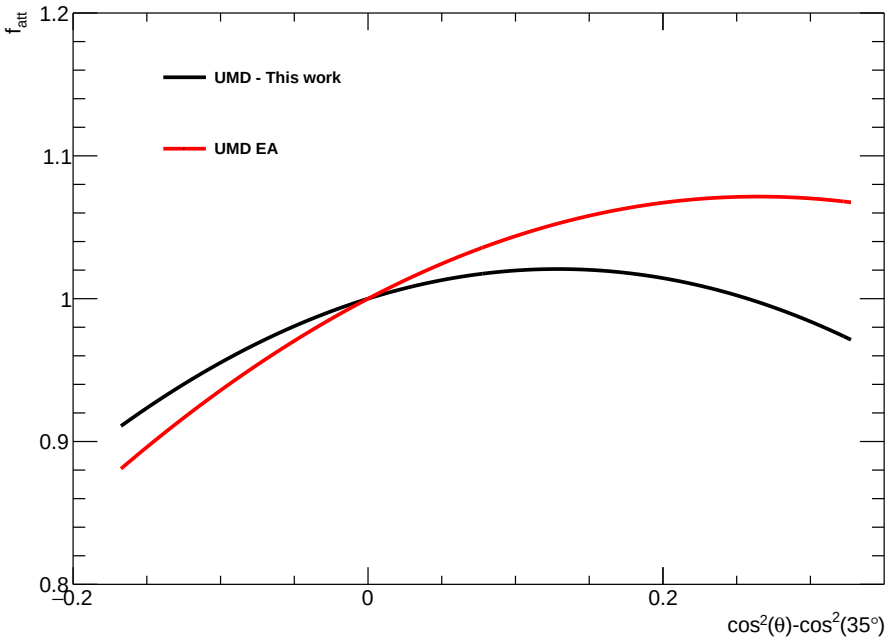


Figure 6.20: Comparison of the attenuation curve obtained in this study and in the EA phase.

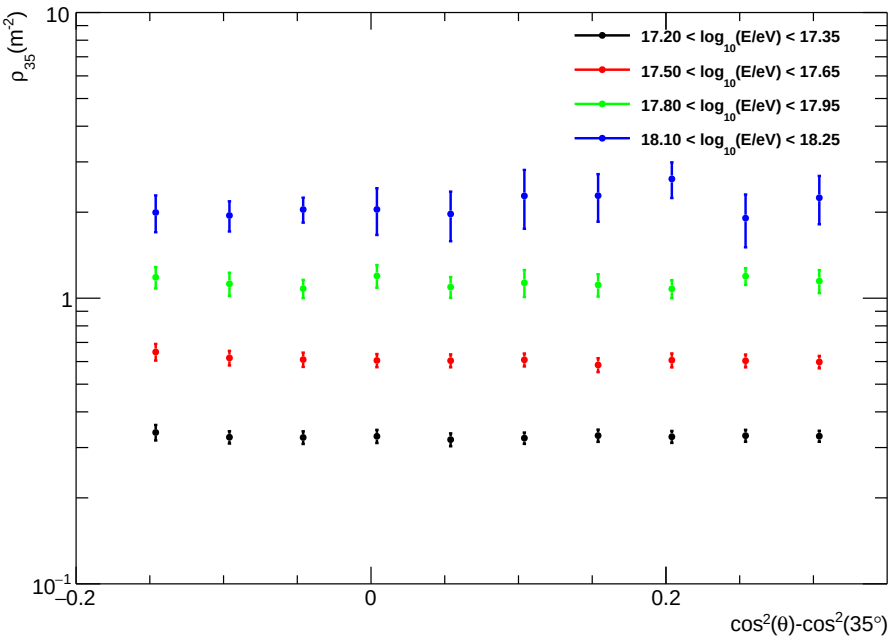


Figure 6.21: The reference ρ_{35} value as a function of the mean zenith angle bin for different primary energy bins.

Fig. 6.21 shows no zenith angle dependence of ρ_{35} which is expected. Only fluctuations are observed for larger energy bins. On the other hand, the ρ_{35} values as a function of energy are shown in Fig. 6.22. A power law dependence with energy is expected from Monte-Carlo simulations if mass composition was constant. So in the same figure ρ_{35} is averaged and fitted through a χ^2 minimization of the function:

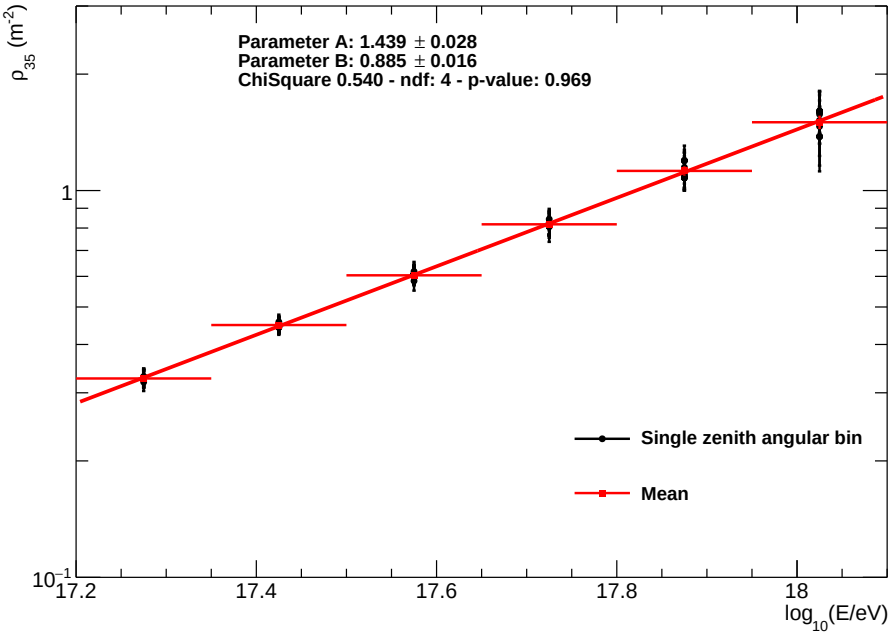


Figure 6.22: Muon densities ρ_{35} as a function of energy fitted by Eq. 6.8.

$$\rho_{35}(E; A, B) = A \left(\frac{E}{10^{18}\text{eV}} \right)^{B}. \tag{6.8}$$

The best-fit parameter values are $A = 1.44 \pm 0.03\,\text{m}^{-2}$ and $B = 0.89 \pm 0.002$. These values can be compared with the previous study [189] in the engineering array phase where the UMDs were equipped with photo-multiplier tubes (PMTs) instead of the current silicon photo-multipliers (SiPMs). In that study the reported values were $A = (1.75 \pm 0.05(\text{stat.}) \pm 0.05(\text{sys.}))\,\text{m}^{-2}$ and $B = 0.89 \pm 0.04(\text{stat.}) \pm 0.04(\text{sys.})$. The values of the logarithmic gain B are compatible. But the values of A, the average muon density at $E = 10^{18}\,\text{eV}$, have a significant difference. In this study the muon content is 18% lower than in previous analysis. This relatively large difference can be explained partly by the differences in the detector itself, but most likely by the technique to fit the LDF. In our analysis we focus on the intermediate distance range to avoid the problems due to saturated stations while for the EA, the full range was used. In that case a larger normalization as observed is expected. The comparison with EAS simulation is done in the next section.

6.3 Shower simulations

The muon content measured by the UMDs can be analyzed in terms of mass composition or, if the mass is fixed by another technique (like fluorescence), can be used to test the hadronic interaction models. In that case it is necessary to compare it with shower simulations. A set of showers at some of the energy and angle bins used in the last section are simulated using the CORSIKA event generator software. Proton and iron showers in energy ranges $(10^{17.2}, 10^{17.35})$ eV, $(10^{17.5}, 10^{17.65})$ eV and $(10^{17.8}, 10^{17.95})$ eV and zenith angle ranges $(0^\circ, 13^\circ)$, $(27^\circ, 30^\circ)$ and $(39^\circ, 42^\circ)$ are used. For every setting, 400 showers are available resulting from 200 simulated showers with EPOS LHC that were tossed two times on the SD-750 array each. This number of showers ensures low fluctuations due to detector simulation. The

showers are simulated with an energy spectrum index $\gamma = -3.27$ reported by Auger in the energy range $(10^{17} - 10^{18})$ eV.

While the Monte Carlo method gives the best description of the shower development, the numerical development by means of the cascade equations, already explained in Ch. 3, is used to simulate the showers, because it allows to implement the Core-Corona scenarios presented in Sect. 6.4 and reduce the simulation time considerably. An average shower, corresponding to a very large number of individual showers, can be obtained in only five minutes at $10^{17.5}$ eV. CORSIKA has to be compiled with the CONEX option and the most important flags used for the simulation are CONEX 111 to describe the shower using cascade equations from the first interaction and CASCADE TFF to have an accurate description of the muon and EM shower components. Fig. 6.23 compares full MC and CONEX simulations after detector reconstruction. There are some statistical fluctuations but the ratio shows a difference lower than 10%.

Simulated showers are reconstructed by $\overline{\text{Off}}\underline{\text{line}}$ repeating the same procedure as in the previous section. Fig. 6.24 shows the flat energy evolution of the weighted average MLDF slope β and the parameterization of β in Fig. 6.25 is in agreement with that found in Fig. 6.13. The attenuation function is calculated for simulations simply by fitting the ρ_{450} by Eq. 6.6. The comparison of the attenuation curves for proton and iron showers with data is shown in Fig. 6.26. The data curve is between the proton and iron curves as expected. The ρ_{35} values as a function of energy are fitted by Eq. 6.8 for proton and iron showers in Fig. 6.27 and 6.28. The slopes are similar to the slope found for data.

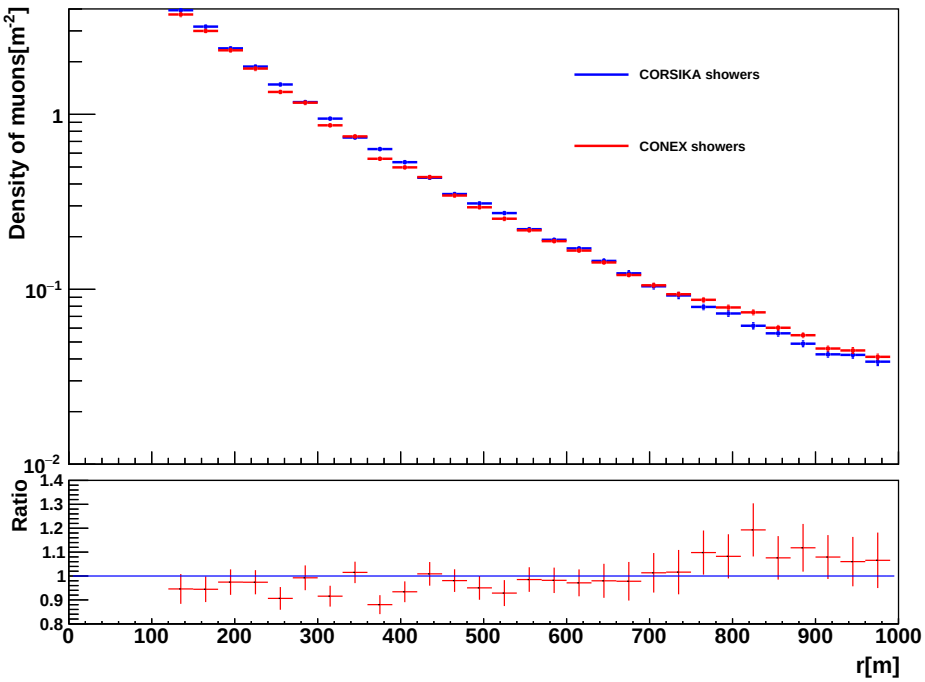


Figure 6.23: MLDF of full MC (blue line) and CONEX (red line) showers, both at UMD depth.

6.4 UMD data comparison with simulations and Core-Corona scenarios

The UMD data analyzed in Sect. 6.1 can be compared with shower simulations presented in Sect. 6.3. Fig. 6.29 shows the MLDF of UMD data and for simulated proton and iron showers

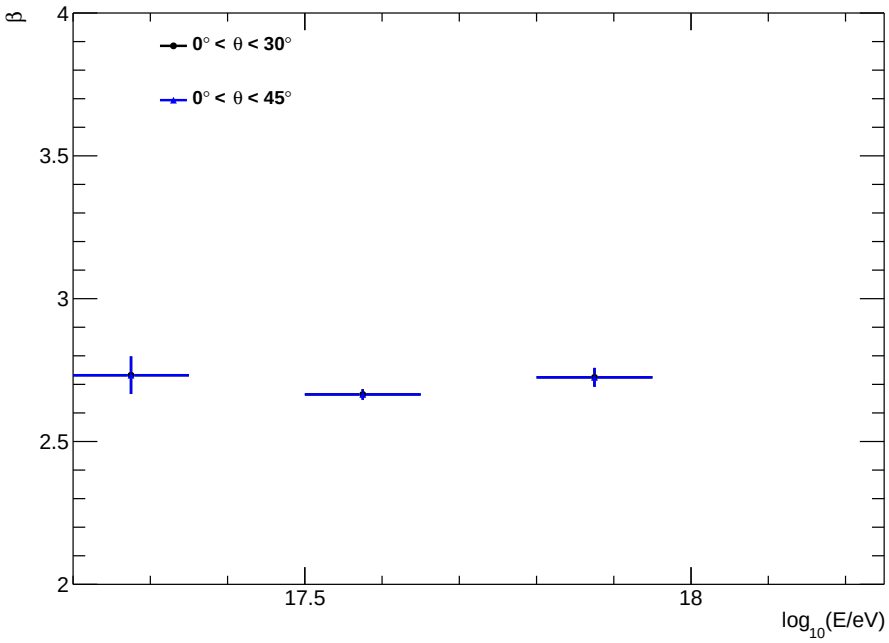


Figure 6.24: The evolution of the weighted average MLDF slope β as a function of primary energy considering all zenith angles together and discriminating by angular bands for simulated proton showers.

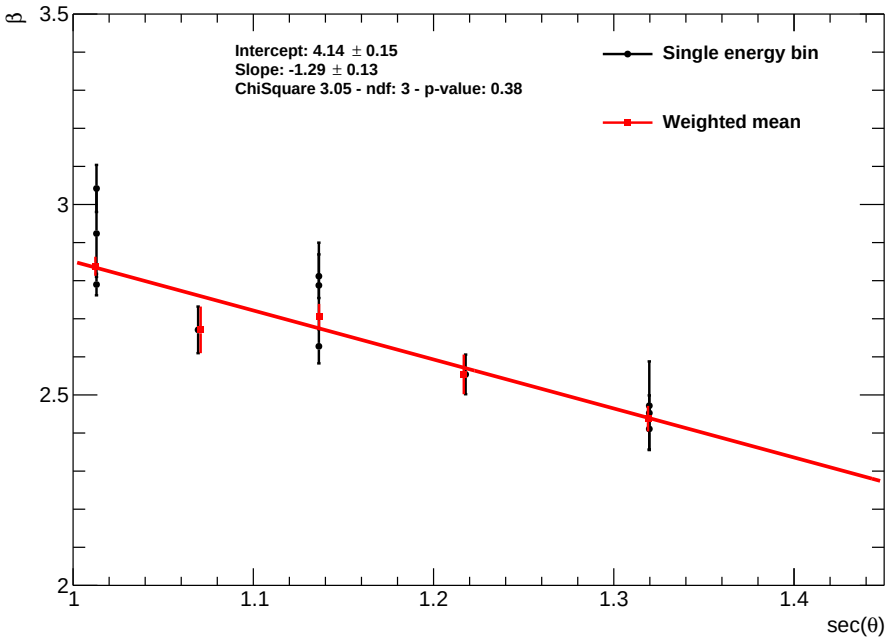


Figure 6.25: The evolution of the weighted average MLDF slope β in terms of the secant of the zenith angle. The red line represents the fitted model of Eq. 6.2 for simulated proton showers.

for the energy range $(10^{17.5} - 10^{17.65})$ eV and zenith angle range $0^\circ - 13^\circ$. These MLDFs are parameterized by Eq. 6.1 where the β values are very similar while the normalization values are different. The curve of UMD data falls between the proton and iron curves, which is the first test of consistency between simulation and data. Looking at the ratio plot, UMD data is around 37% larger than the result for proton showers. On the other hand, shower simulations of the core-corona case with $f_\omega = 1$ and $E_{\text{scale}} = 1$ of Fig. 4.1 show in Fig. 6.30

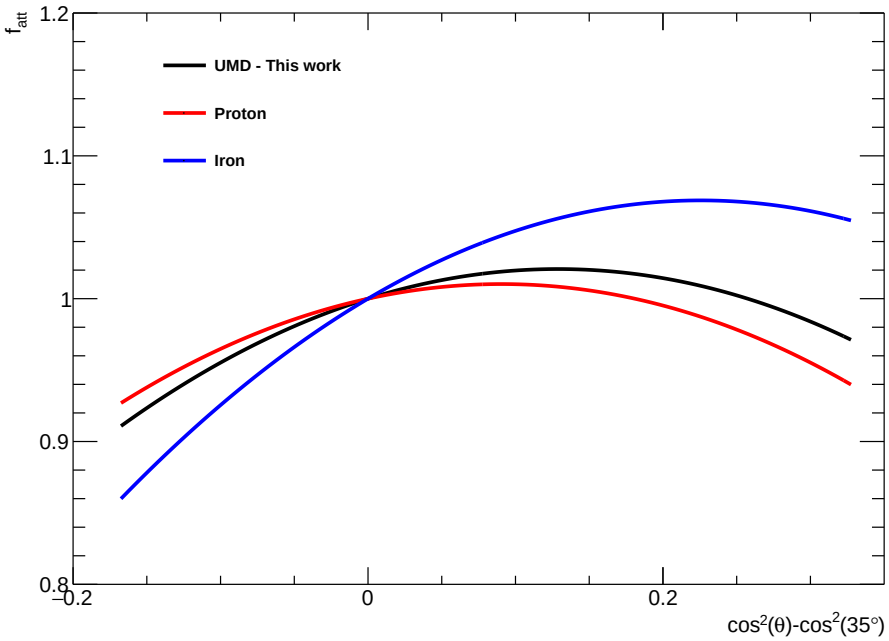


Figure 6.26: Comparison of the attenuation curve obtained in the UMD data analysis and those corresponding for proton and iron simulations.

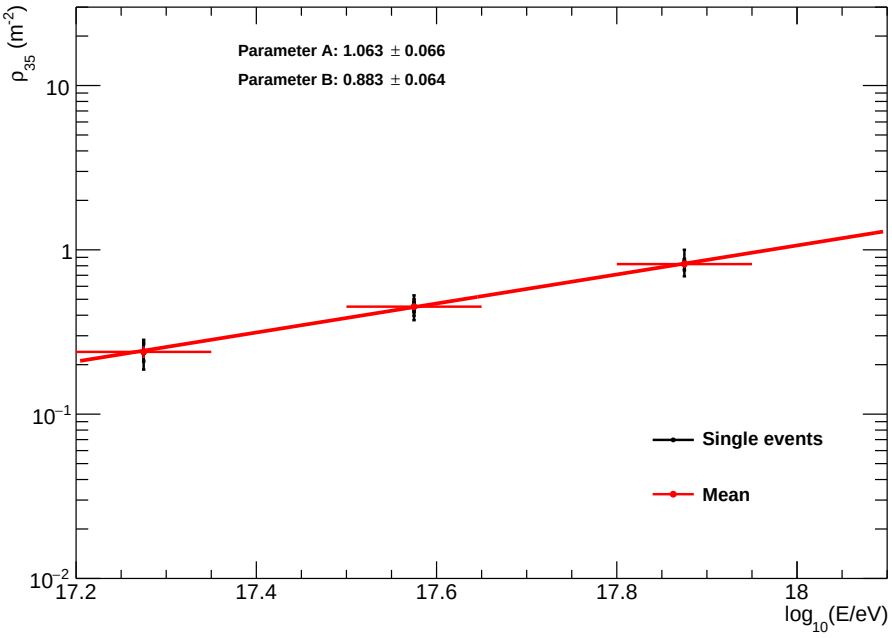


Figure 6.27: Muon density ρ_{35} as a function of energy for proton showers.

that UMD data is around 23 % larger. The MLDF of proton and iron simulation have a larger normalization parameter than the default case. In particular, the proton MLDF increases more than the iron one because iron showers are, by the superposition model, 56 proton showers with lower energy so that the core-corona mix is less evident for these showers. The β values are more similar than in the default case but it could be due to fluctuations.

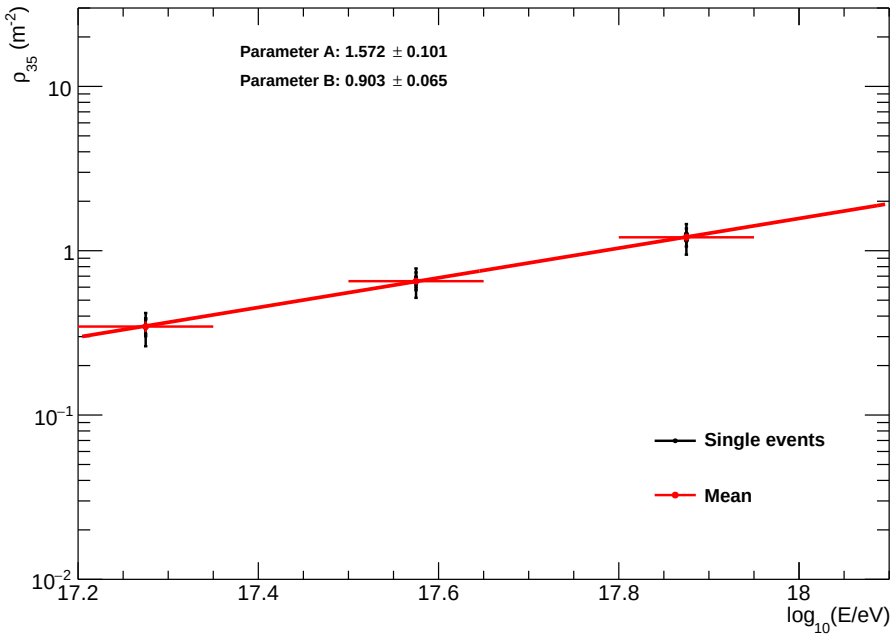


Figure 6.28: Muon density ρ_{35} as a function of energy for iron showers.

In order to compare these results with observables sensitive to mass composition measurements, the expected ratio of UMD data $\tilde{\rho}_{450}$ and proton simulations ρ^{p}_{450} can be calculated from Eq. 1.11 as follows:

$$\frac{\tilde{\rho}_{450}}{\rho^{p}_{450}} = \left(\frac{\rho^{Fe}_{450}}{\rho^{p}_{450}}\right)^{\tilde{z}} \tag{6.9}$$

where $\tilde{z}$ is the measured z-scale for a given observable. Measurements of the mean depth of shower maximum X_{max} by the FD give $\tilde{z} \sim 0.35$ at energies around $10^{17.5}$ eV. The expected ratio obtained by Eq. 6.9 is shown in the lower plots of Fig. 6.29 and 6.30 for all distances as a reference. In the former case where proton and iron showers are simulated with the default EPOS LHC model the ratio of data and protons is ~ 1.37 while the expected value is ~ 1.21. It means that data has $\sim 13\%$ more muons than expected from simulations. On the other hand, when simulations are done with EPOS LHC with the conservative Core-Corona model, the ratio of data and protons is ~ 1.23 while the expected value is ~ 1.18. In this case the data muon content is just $\sim 4\%$ larger than the one expected from simulations. Hence, the muon density by the Core-Corona model with the parameters $f_{\omega} = 1$ and $E_{\mathrm{scale}} = 1$ shows to be in agreement with the measurements in terms of mass composition. It is interesting to notice that in Fig. 6.30, the core-corona MLDF seems to be more parallel to the data compared to the default model. It could be due to fluctuations in the fit, but with more statistic it could be an interesting point to be studied in further analysis.

The superposition of Fig. 6.22, 6.27 and 6.28 in Fig. 6.31 make it possible to compare the muon content of data and simulation in terms of ρ_{35} as a function of energy. Data and simulations have the same slope, which means no change in mass composition over these energies. Data curves are close and below the iron curve. The blue line represents, as previous plots, the expected data ratio taking into account mass composition measurements. The ratio of data and protons is ~ 1.35 while the expected value is ~ 1.14. It means that data has $\sim 18\%$ more muons than expected from simulations. Fig. 6.32 is the analogous plot for

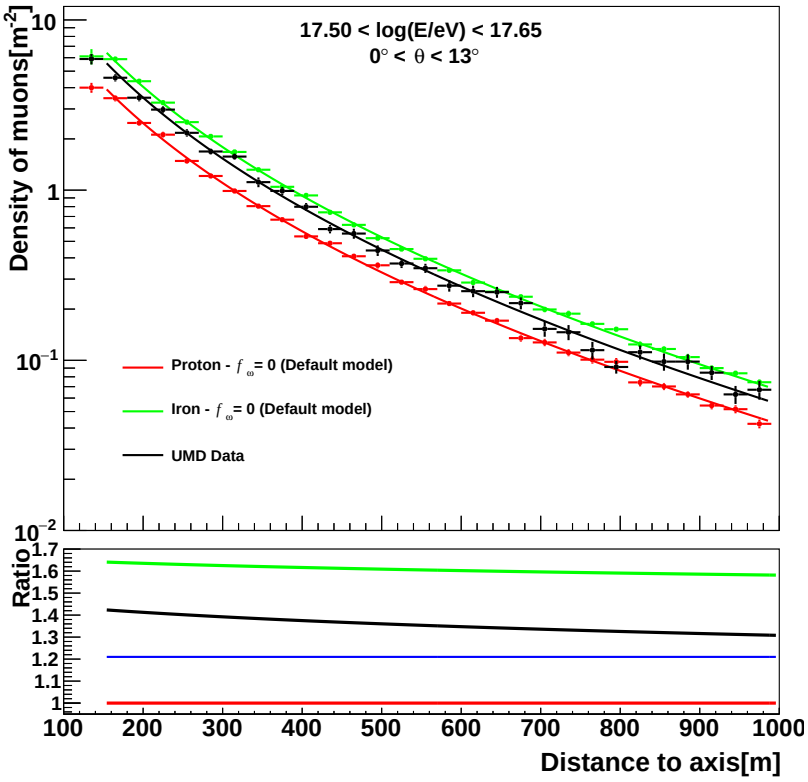


Figure 6.29: MLDF of proton and iron showers for default EPOS LHC model, i.e. $f_\omega = 0$, with UMD reconstruction and UMD data. Blue line in the bottom plot is the expected data ratio from mass composition measurements.

the core-corona case, where the data has $\sim 6\%$ more muons than simulations. These results confirm the previous result obtained in the vertical zenith bin at $10^{17.5}$ eV.

The result at $10^{17.5}$ eV can also be compared with the $X_{\max}$ measurement by Auger. In order to do that the z-scale of Eq. 1.11 is used together with the muon-density estimator ρ_{35}. The z-scale found at $E = 10^{17.5}$ eV is $z = 0.82 \pm 0.37$ where a systematic error of 10% is taken for ρ_{35} from the EA study. This error includes the uncertainty due to the soil density, MLDF and the calibration of SiPM and is applied for both data and simulations to be conservative since it is related to the detector simulation. Fig. 6.33 is Fig. 4.7 including the new result of this analysis. The previous UMD result in the EA phase was $z = 1.04 \pm 0.3$, so both results are in agreement, but the mean of the new result is lower. The conservative core-corona case is now very close to the data point.

At this point it is necessary to emphasize that we cannot conclude that the core-corona model actually solves the muon puzzle or not. The core-corona implementation by means of CE and their energy spectra modification gives the flexibility to be used with any hadronic interaction model in an easy way. However, the method is not dependent on the pseudo rapidity or multiplicity distribution of particles from the core or the corona. Hence, the results presented in the last chapter should be taken as an indication to investigate further this model in a full Monte-Carlo analysis.

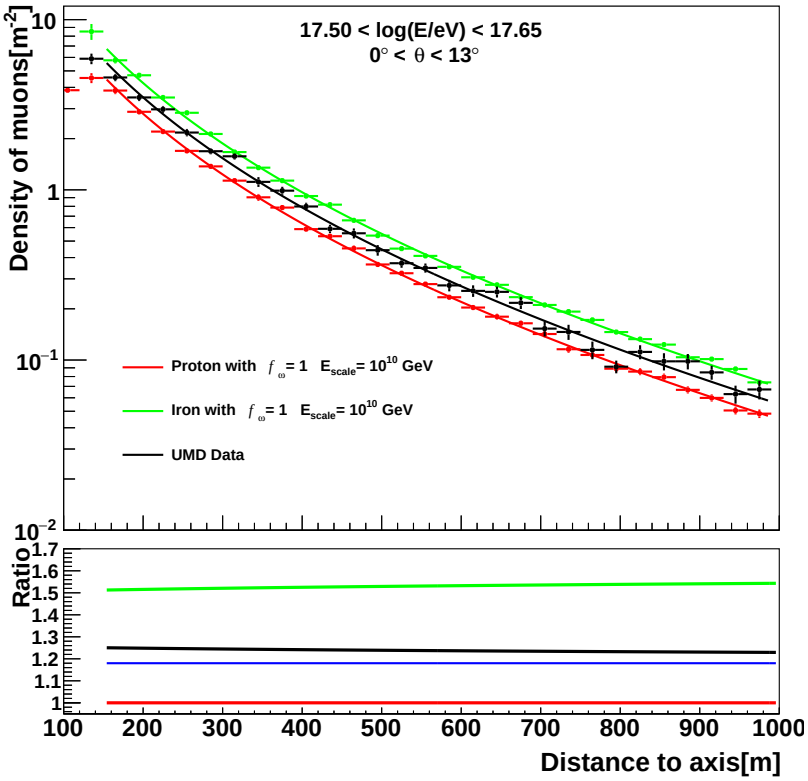


Figure 6.30: Plot analogous to Fig. 6.29 where proton and iron showers are simulated with the conservative core-corona scenario $f_\omega = 1$ and $E_{\text{scale}} = 10^{10}$ GeV.

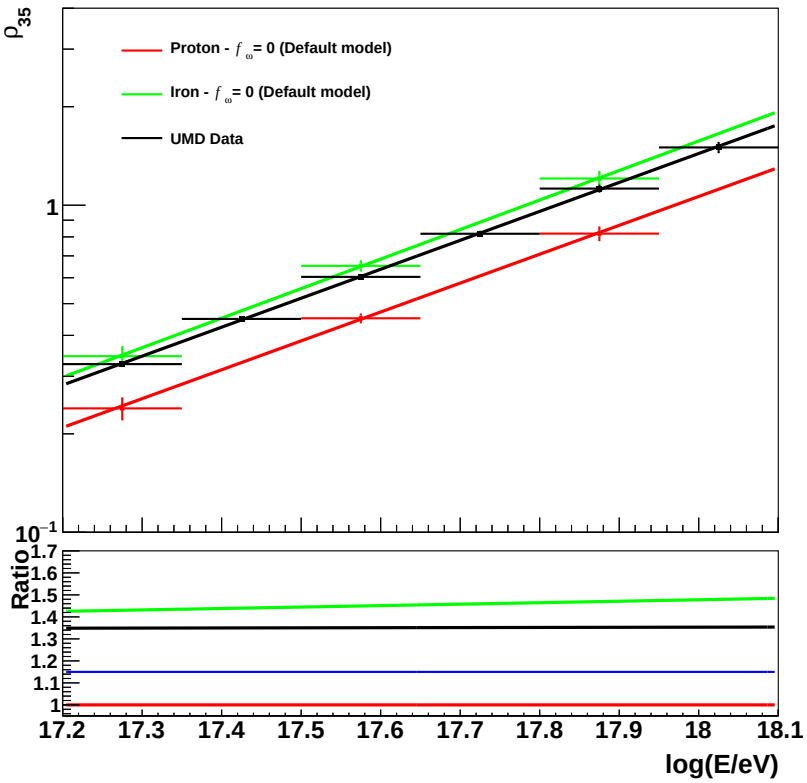


Figure 6.31: Mean ρ_{35} as a function of energy for UMD data and shower simulations of proton and iron with the default EPOS LHC model. the blue line in the bottom plot is the expected data ratio from mass composition measurements.

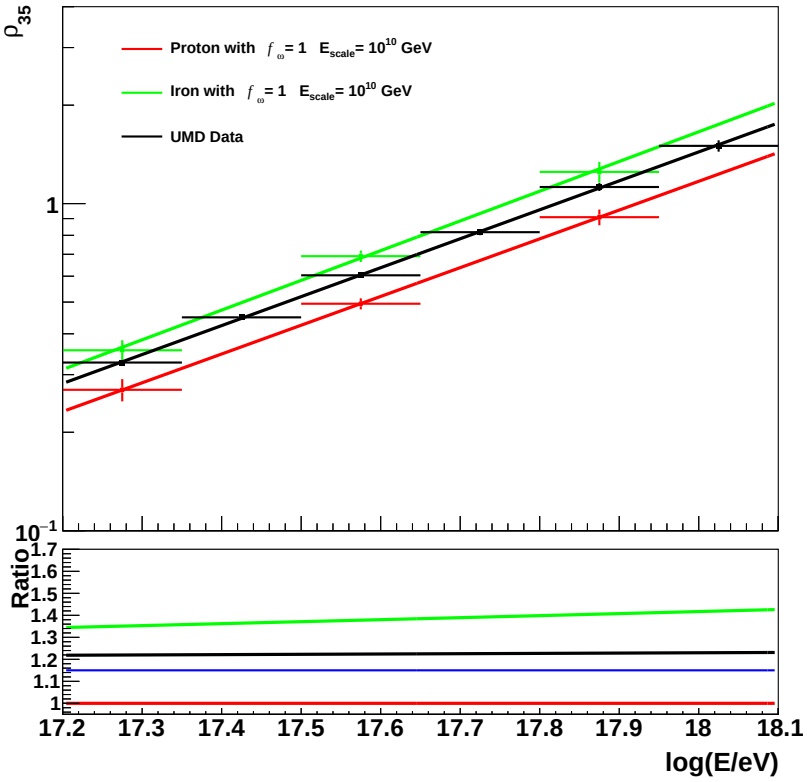


Figure 6.32: Mean ρ_{35} as a function of energy for UMD data and shower simulations of proton and iron with EPOS LHC in a conservative core-corona mixing scenario. The blue line in the bottom plot is the expected data ratio from mass composition measurements.

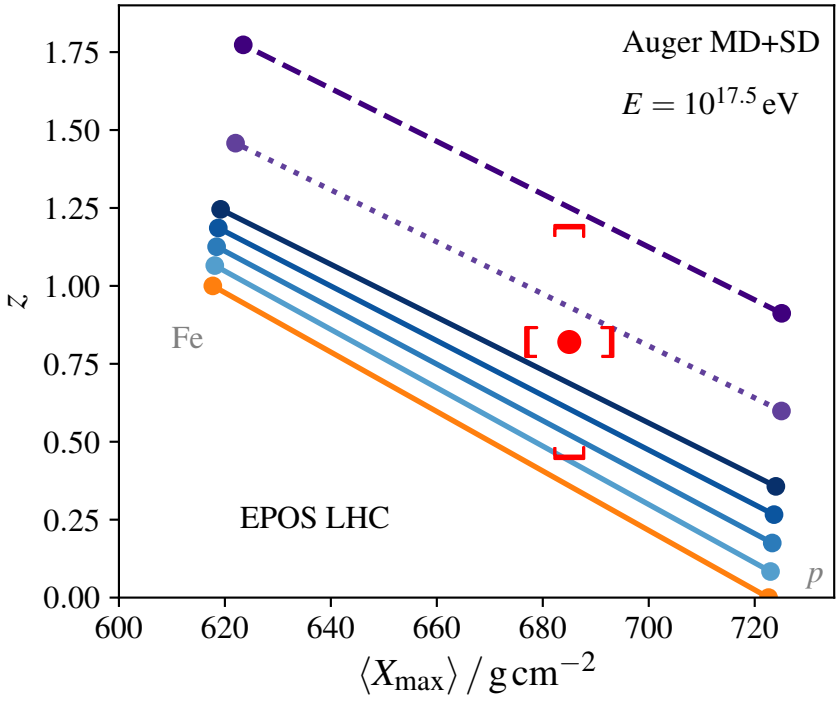


Figure 6.33: Fig. 4.7 including the UMD data (red point) as calculated in this study.

CHAPTER 7

Conclusions

The understanding of the muon puzzle is one of the most challenging problems in high energy physics. Several experiments measured the muon content of extensive air showers (EAS). Each has a different muon energy threshold and the stage of the shower development is not the same for all. The muon content obtained in simulations is significantly lower than the one measured in most of these experiments. This fact implies that the hadronic component of EAS has to retain in some way more energy than in current hadronic interaction models. Exotic mechanisms proposed to explain the muon deficit in simulations have been disfavoured by measurements at colliders. However, recent measurements at the LHC show collective behaviour even in small systems. These collective effects are commonly related to the existence of a QGP in high density regions where the hadronization mechanism is purely statistical.

In this thesis, a novel approach was used to modify the hadronization mechanism simulated by the hadronic interactions models. This is achieved by modifying the part where EAS are described numerically by CE in the framework of the CONEX event generator. This was done by means of a tool developed to apply changes in the effective energy spectra of secondary particles of hadron interactions. These spectra are closely related to the underlying particle production mechanisms, thus any change in the spectra implies a modified hadronization mechanism in the EAS development. One feature of this tool is its versatility since it can be applied to any hadronic interaction model. In addition, there is the possibility to reproduce exotic models in an easy way.

In order to study the effects of collective hadronization in EAS muon production, this tool was used to implement the Core-Corona model. In the Core-Corona model the particles in large density regions behave as a quark gluon plasma and decay according to statistical hadronization (core), whereas those in low density regions produce particles via string fragmentation (corona). Both mechanisms produce different particle ratios, so the secondary particle spectra of the hadronic interaction models were shifted from the default ratios in the corona to the ratios expected in the core. The modified particle ratios correspond to the particles produced at mid-rapidity while the leading particle effect in the forward region is not modified. These changes increase with the interacting energy in a logarithmic way and a transition from full Corona to full Core was analyzed. The number of produced muons increases when there are more particles in the core because the ratio of the electromagnetic to the hadronic energy density R is lower in core hadronization.

It was found that the muon deficit observed by the Pierre Auger Observatory is lower when more Core hadronization is taken into account. In adittion, the comparison of sev-

eral cosmic-ray based experiments by means of the z-scale is presented, which allows a direct comparison between different muon observables. This shows that more Core-like contributions are needed compared with what is currently provided by the models. Even conservative Core-Corona scenarios seem to be enough to reproduce the data. This means that QGP-like effects also in light colliding systems and starting in central collisions at much lower center-of-mass energies than what was considered until now may play a decisive role in muon production. The most important results of this analysis were published in [184].

In the second part of this thesis, direct muon measurements were included into the analysis and compared with full end-to-end simulations in the context of the Pierre Auger Observatory. The Underground Muon Detector (UMD), as part of the upcoming upgrade of the Observatory, gives a unique and straight-forward opportunity to directly measure high-energy muons of extensive air showers. Thus, the first two and a half years of data acquired with the production array of the UMD have been reconstructed and analyzed. In particular, a data driven analysis was built on average muon lateral distributions (MLDF) in bins of energy and zenith angle. The constant intensity cut method was applied to compensate the shower attenuation at different zenith angles. The observable ρ_{35} was used to calculated the muon content of EAS. It was found 18% less than the previous result in the engineering array (EA) phase. However, the muon deficit is still observed when data is compared with mass composition measurements. Data results were compared with full detector simulations where the muon deficit in simulation for EPOS LHC model was found $\sim$ 18%. Applying the conservative Core-Corona to the same hadronic interaction model the muon deficit decreased to $\sim$ 6%. UMD data was compared with full detector simulations and the obtained z-factor is 0.82 ± 0.37 which is compatible with the result observed in the EA phase. Also this result shows to be compatible with a conservative Core-Corona scenario.

Despite of these encouraging results, this implementation of the Core-Corona model is mostly to test the hypotheses that the ratio of electromagnetic and hadronic particles R for QGP is low enough to increase the muon content of EAS as it is needed from data measurements. It is worth emphasizing that the changes done in the hadronization mechanism are implemented in the numerical description of the EAS and not in the Monte-Carlo one. Hence, as an outlook, detailed studies on complete MC simulations of EAS need to be done to confirm these results. However, the analyses presented in this thesis seem to go in the right direction to understand the muon deficit helping to reconcile the present observations with the predictions of post-LHC hadronic interaction models.

Bibliography

[1] V. Hess. Penetrating radiation in seven free ballon flights. *Phys. Z.*, 13:1804, 1912.

[2] R. A. Millikan. High frequency rays of cosmic origin. *Science*, 62(1612):445–448, 1925.

[3] A. H. Compton. A geographic study of cosmic rays. *Phys. Rev.*, 43:387–403, Mar 1933.

[4] S. H. Neddermeyer and C. D. Anderson. Note on the nature of cosmic-ray particles. *Phys. Rev.*, 51, 01 1937.

[5] P. Auger, *et al.* Extension et pouvoir pénétrant des grandes gerbes de rayons cosmiques. *Comptes rendus Physique*, 208:1641, 1939.

[6] J. Linsley. Evidence for a primary cosmic-ray particle with energy 10^{20} eV. *Phys. Rev. Lett.*, 10:146–148, Feb 1963.

[7] Arno A. Penzias and Robert Woodrow Wilson. A measurement of excess antenna temperature at 4080-Mc/s. *Astrophys. J.*, 142:419–421, 1965.

[8] K. Greisen. End to the cosmic ray spectrum? *Phys. Rev. Lett.*, 16:748–750, 1966.

[9] G. T. Zatsepin and V. A. Kuzmin. Upper limit of the spectrum of cosmic rays. *JETP Lett.*, 4:78–80, 1966. [Pisma Zh. Eksp. Teor. Fiz.4,114(1966)].

[10] D. J. Bird et al. Detection of a cosmic ray with measured energy well beyond the expected spectral cutoff due to cosmic microwave radiation. *Astrophys. J.*, 441:144–150, 1995.

[11] M. Takeda, N. Sakaki, and K. Honda *et al.* Energy determination in the akeno giant air shower array experiment. *Astroparticle Physics*, 19(4):447 – 462, 2003.

[12] M A Lawrence, R J O Reid, and A A Watson. The cosmic ray energy spectrum above 4×10^{17}eV as measured by the Haverah Park array. *Journal of Physics G: Nuclear and Particle Physics*, 17(5):733–757, may 1991.

[13] G. A. Medina-Tanco. The energy spectrum observed by the AGASA experiment and the spatial distribution of the sources of ultrahigh-energy cosmic rays. *Astrophys. J.*, 510:L91, 1999.

[14] A. Aab *et al.* (Pierre Auger Collab.). The Pierre Auger Cosmic Ray Observatory. *Nucl. Instrum. Meth.*, A798:172–213, 2015.

[15] W. F. Hanlon. Recent results from the Telescope Array Project. *Nuclear and Particle Physics Proceedings*, 279-281:15 – 22, 2016. Proceedings of the 9th Cosmic Ray International Seminar.

[16] K. H. Kampert *et al.* Cosmic rays in the 'knee'-region - recent results from KASCADE. *Acta Phys. Polon.*, B35:1799–1812, 2004.

[17] 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.

[18] G. Giacinti, M. Kachelriess, D. V. Semikoz, and G. Sigl. Cosmic Ray Anisotropy as Signature for the Transition from Galactic to Extragalactic Cosmic Rays. *JCAP*, 1207:031, 2012.

[19] M. Unger. Cosmic rays above the knee. In *21st European Cosmic Ray Symposium*, pages 48–55, 2009.

[20] Particle Data Group. Review of particle physics. *Phys. Rev. D*, 98:030001, 2018.

[21] J. W. Cronin. The highest-energy cosmic rays. *Nuclear Physics B - Proceedings Supplements*, 138:465 – 491, 2005. Proceedings of the Eighth International Workshop on Topics in Astroparticle and Undeground Physics.

[22] M. Unger for the Pierre Auger Collaboration. Highlights from the Pierre Auger Observatory. *PoS*, ICRC2017:1102, 2018. [35,1102(2017)].

[23] D. Ivanov. Report of the Telescope Array - Pierre Auger Observatory Working Group on Energy Spectrum. *PoS*, ICRC2017:498, 2018.

[24] O. Deligny for the Pierre Auger Collaboration. The energy spectrum of ultra-high energy cosmic rays measured at the pierre auger observatory and at the telescope array. *PoS*, ICRC2019:234, 2019.

[25] W. Heitler. *The quantum theory of radiation*, volume 5 of *International Series of Monographs on Physics*. Oxford University Press, 1936.

[26] J. Matthews. A Heitler model of extensive air showers. *Astropart. Phys.*, 22:387–397, 2005.

[27] R. Engel, D. Heck, and T. Pierog. Extensive air showers and hadronic interactions at high energy. *Annual Review of Nuclear and Particle Science*, 61(1):467–489, 2011.

[28] A Watson. The mass composition of cosmic rays above 10^{17} eV. *Nuclear Physics B - Proceedings Supplements*, 136:290–300, 10 2004.

[29] J. Abraham *et al.* (Pierre Auger Collab.). Properties and performance of the prototype instrument for the Pierre Auger Observatory. *Nucl. Instrum. Meth.*, A523(1-2):50 – 95, 2004.

[30] P. Privitera. The Pierre Auger Observatory. *Nuclear Physics B - Proceedings Supplements*, 110:487 – 490, 2002.

[31] I. Allekotte *et al.* (Pierre Auger Collab.). The Surface Detector System of the Pierre Auger Observatory. *Nucl. Instrum. Meth.*, A586:409–420, 2007.

[32] F. *et al.* Arqueros. The Central Laser Facility at the Pierre Auger Observatory. In *29th International Cosmic Ray Conference (ICRC29)*, volume 8 of *International Cosmic Ray Conference*, page 335, January 2005.

[33] M. Unger. Mass sensitive observables of the Pierre Auger Observatory. *EPJ Web of Conferences*, 53:04009–, 06 2013.

[34] C. Bonifazi. The angular resolution of the Pierre Auger Observatory. *Nuclear Physics B - Proceedings Supplements*, 190:20 – 25, 2009. Proceedings of the Cosmic Ray International Seminars.

[35] C. Bonifazi, A. Letessier-Selvon, and E.M. Santos. A model for the time uncertainty measurements in the Auger surface detector array. *Astroparticle Physics*, 28(6):523 – 528, 2008.

[36] J. Abraham *et al.* (Pierre Auger Collab.). Trigger and aperture of the surface detector array of the Pierre Auger Observatory. *Nucl. Instrum. Meth.*, A613:29–39, 2010.

[37] A. M. Botti . *Determination of the chemical composition of cosmic rays in the energy region of 5 EeV with the AMIGA upgrade of the Pierre Auger Observatory*. PhD thesis, 2019.

[38] 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.*, 33:108–129, 2010.

[39] P. Abreu *et al.* The Rapid Atmospheric Monitoring System of the Pierre Auger Observatory. *JINST*, 7:P09001, 2012.

[40] J. Abraham *et al.* (Pierre Auger Collab.). The fluorescence detector of the Pierre Auger Observatory. *Nucl. Instrum. Meth.*, A620:227 – 251, 2010.

[41] A. Castellina. AugerPrime: the Pierre Auger Observatory Upgrade. *EPJ Web Conf.*, 210:06002, 2019.

[42] C. Medina *et al.* . Enhancing the Pierre Auger Observatory to the $10^{17} - 10^{18.5}$eV range: Capabilities of an infill surface array. pages 302–311, 10 2006.

[43] N. Gonzalez *et al.* . Reconstruction performance of simulated events with the SD-433: a WCD+UMD analysis. *Pierre Auger internal note*, page GAP2019_036, 2019.

[44] A. D. Supanitsky, A. Etchegoyen, G. Medina-Tanco, I. Allekotte, M. Gomez Berisso, and Maria Clementina Medina. Underground Muon Counters as a Tool for Composition Analyses. *Astropart. Phys.*, 29:461–470, 2008.

[45] A. Aab *et al.* Prototype muon detectors for the AMIGA component of the Pierre Auger Observatory. *JINST*, 11(02):P02012, 2016.

[46] J. Abraham *et al.* (Pierre Auger Collab.). Measurement of the energy spectrum of cosmic rays above 10^{18} eV using the Pierre Auger Observatory. *Phys. Lett. B*, 685(4-5):239 – 246, 2010.

[47] A. Aab *et al.* Observation of a Large-scale Anisotropy in the Arrival Directions of Cosmic Rays above 8×10^{18} eV. *Science*, 357(6537):1266–1270, 2017.

[48] M. Settimo (Pierre Auger Collab.). An update on a search for ultra-high energy photons using the Pierre Auger Observatory. In *Proc. 32nd Int. Cosmic Ray Conf.*, Beijing, China, 2011.

[49] P. Abreu *et al.* (Pierre Auger Collab.). Ultrahigh Energy Neutrinos at the Pierre Auger Observatory. *Adv. High Energy Phys.*, 2013:708680, 2013.

[50] P. Abreu *et al.* Measurement of the proton-air cross-section at $\sqrt{s} = 57$ TeV with the Pierre Auger Observatory. *Phys. Rev. Lett.*, 109:062002, 2012.

[51] A. Aab *et al.* Muons in Air Showers at the Pierre Auger Observatory: Mean Number in Highly Inclined Events. *Phys. Rev. D*, 91(3):032003, 2015. [Erratum: Phys.Rev.D 91, 059901 (2015)].

[52] R. Ulrich. Extension of the measurement of the proton-air cross section with the Pierre Auger Observatory. *PoS*, ICRC2015:401, 2016.

[53] R.U. Abbasi *et al.* Measurement of the proton-air cross section with Telescope Array's Middle Drum detector and surface array in hybrid mode. *Phys. Rev. D*, 92(3):032007, 2015.

[54] A. Aab *et al.* Measurement of the average shape of longitudinal profiles of cosmic-ray air showers at the Pierre Auger Observatory. *JCAP*, 03:018, 2019.

[55] S. Andringa *et al.* Average shape of longitudinal shower profiles measured at the Pierre Auger Observatory. *EPJ Web Conf.*, 210:02015, 2019.

[56] D. Ivanov. Telescope Array Experiment (2018). *PoS*, TeVPA 2018, 2018.

[57] T. Abu-Zayyad *et al.* Evidence for changing of cosmic ray cmposition between 10^{17}eV and 10^{18}eV from multicomponent measurements. *Phys. Rev. Lett.*, 84:4276–4279, 2000.

[58] A. G. Bogdanov, D. M. Gromushkin, R. P. Kokoulin, G. Mannocchi, A. A. Petrukhin, O. Saavedra, G. Trinchero, D. V. Chernov, V. V. Shutenko, and I. I. Yashin. Investigation of the properties of the flux and interaction of ultrahigh-energy cosmic rays by the method of local-muon-density spectra. *Phys. Atom. Nucl.*, 73:1852–1869, 2010. [Yad. Fiz.73,1904(2010)].

[59] A. G. Bogdanov, R. P. Kokoulin, G. Mannocchi, A. A. Petrukhin, O. Saavedra, V. V. Shutenko, G. Trinchero, and I. I. Yashin. Investigation of very high energy cosmic rays by means of inclined muon bundles. *Astropart. Phys.*, 98:13–20, 2018.

[60] J. A. Bellido, R. W. Clay, N. N. Kalmykov, I. S. Karpikov, G. I. Rubtsov, S. V. Troitsky, and J. Ulrichs. Muon content of extensive air showers: comparison of the energy spectra obtained by the Sydney University Giant Air-shower Recorder and by the Pierre Auger Observatory. *Phys. Rev.*, D98(2):023014, 2018.

[61] Yu. A. Fomin, N. N. Kalmykov, I. S. Karpikov, G. V. Kulikov, M. Yu. Kuznetsov, G. I. Rubtsov, V. P. Sulakov, and S. V. Troitsky. No muon excess in extensive air showers at 100–500 PeV primary energy: EAS–MSU results. *Astropart. Phys.*, 92:1–6, 2017.

[62] Apel, W.D. *et al.* Muon production height studies with the air shower experiment KASCADE-Grande. *Astropart. Phys.*, 34:476–485, 2011.

[63] A. Aab *et al.* Testing Hadronic Interactions at Ultrahigh Energies with Air Showers Measured by the Pierre Auger Observatory. *Phys. Rev. Lett.*, 117(19):192001, 2016.

[64] R.U. Abbasi *et al.* Study of muons from ultrahigh energy cosmic ray air showers measured with the Telescope Array experiment. *Phys. Rev. D*, 98(2):022002, 2018.

[65] A. Aab *et al.* Muons in Air Showers at the Pierre Auger Observatory: Measurement of Atmospheric Production Depth. *Phys. Rev. D*, 90(1):012012, 2014. [Addendum: Phys.Rev.D 90, 039904 (2014), Erratum: Phys.Rev.D 92, 019903 (2015)].

[66] S. De Ridder, E. Dvorak, and T. K. Gaisser. Sensitivity of IceCube Cosmic-Ray measurements to the hadronic interaction models. *PoS*, ICRC2017:319, 2018.

[67] M.G. Aartsen *et al.* Characterization of the Atmospheric Muon Flux in IceCube. *Astropart. Phys.*, 78:1–27, 2016.

[68] T. Fuchs. Development of a Machine Learning Based Analysis Chain for the Measurement of Atmospheric Muon Spectra with IceCube. In *25th European Cosmic Ray Symposium*, 1 2017.

[69] D. Soldin. Atmospheric Muons Measured with IceCube. *EPJ Web Conf.*, 208:08007, 2019.

[70] A. Aab *et al.* Inferences on mass composition and tests of hadronic interactions from 0.3 to 100 EeV using the water-Cherenkov detectors of the Pierre Auger Observatory. *Phys. Rev. D*, 96(12):122003, 2017.

[71] A. Aab *et al.* Azimuthal Asymmetry in the Risetime of the Surface Detector Signals of the Pierre Auger Observatory. *Phys. Rev. D*, 93(7):072006, 2016.

[72] H.P. Dembinski et al. Report on Tests and Measurements of Hadronic Interaction Properties with Air Showers. *EPJ Web Conf.*, 210:02004, 2019.

[73] L. Cazon. Working Group Report on the Combined Analysis of Muon Density Measurements from Eight Air Shower Experiments. *PoS*, ICRC2019:214, 2020.

[74] J. G. Gonzalez. Measuring the Muon Content of Air Showers with IceTop. *EPJ Web Conf.*, 99:06002, 2015.

[75] S. Mueller *et al. EPS Web Conf.*, 210:02013, 2019.

[76] A.V. Glushkov, M.I. Pravdin, and A. Sabourov for the Yakutsk collaboration, private comm. (2018).

[77] M.G. Aartsen et al. Measurement of the cosmic ray energy spectrum with IceTop-73. *Phys. Rev. D*, 88(4):042004, 2013.

[78] Karl-Heinz Kampert and Michael Unger. Measurements of the Cosmic Ray Composition with Air Shower Experiments. *Astropart. Phys.*, 35:660–678, 2012.

[79] H. P. Dembinski, R. Engel, A. Fedynitch, T. Gaisser, F. Riehn, and T. Stanev. Data-driven model of the cosmic-ray flux and mass composition from 10 GeV to 10^{11} GeV. *PoS*, ICRC2017:533, 2018.

[80] R. P. Feynman. Very high-energy collisions of hadrons. *Phys. Rev. Lett.*, 23:1415–1417, 1969.

[81] F. Riehn. *Hadronic multiparticle production with Sibyll*. PhD thesis, KIT, Karlsruhe, 2015.

[82] G.F. Chew and S. C. Frautschi. Regge Trajectories and the Principle of Maximum Strength for Strong Interactions. *Phys. Rev. Lett.*, 8:41–44, 1962.

[83] T. K. Gaisser, R. Engel, and E. Resconi. *Cosmic Rays and Particle Physics: 2nd Edition*. Cambridge University Press, 6 2016.

[84] A. Donnachie and P.V. Landshoff. Total cross-sections. *Phys. Lett. B*, 296:227–232, 1992.

[85] A.J. Buras and J. Dias de Deus. Scaling law for the elastic differential cross-section in p p scattering from geometric scaling. *Nucl. Phys. B*, 71:481–492, 1974.

[86] Marcel Froissart. Asymptotic behavior and subtractions in the Mandelstam representation. *Phys. Rev.*, 123:1053–1057, 1961.

[87] A. Martin. An absolute upper bound on the pion-pion scattering amplitude. *Report SITP-134*, 1964.

[88] A. Martin and S.M. Roy. Froissart Bound on Inelastic Cross Section Without Unknown Constants. *Phys. Rev. D*, 91(7):076006, 2015.

[89] A. Dymarsky. Can Froissart Bound Explain Hadron Cross-Sections at High Energies? *JHEP*, 15:106, 2020.

[90] S.J. Lindenbaum and R.M. Sternheimer. Isobaric nucleon model for pion production in nucleon-nucleon collisions. *Phys. Rev.*, 105:1874–1879, 1957.

[91] A. Capella, U. Sukhatme, C-I Tan, and J. Tran Thanh Van. Dual parton model. *Phys. Rept.*, 236:225–329, 1994.

[92] A.B. Kaidalov. High-energy hadronic interactions (20 years of the quark gluon strings model). *Phys. Atom. Nucl.*, 66:1994–2016, 2003.

[93] K. Werner, F.-M. Liu, and T. Pierog. Parton ladder splitting and the rapidity dependence of transverse momentum spectra in deuteron-gold collisions at RHIC. *Phys. Rev. C*, 74:044902, 2006.

[94] T. Pierog and K. Werner. EPOS Model and Ultra High Energy Cosmic Rays. *Nucl. Phys. B Proc. Suppl.*, 196:102–105, 2009.

[95] S. Ostapchenko. Nonlinear screening effects in high energy hadronic interactions. *Phys. Rev. D*, 74(1):014026, 2006.

[96] S. Ostapchenko. Monte Carlo treatment of hadronic interactions in enhanced Pomeron scheme: I. QGSJET-II model. *Phys. Rev. D*, 83:014018, 2011.

[97] V.N. Gribov. A reggeon diagram technique. *Sov. Phys. JETP*, 26:414–422, 1968.

[98] J.H. Weis. Regge Theory and High-Energy Hadron-Nucleus Scattering. *Acta Phys. Polon. B*, 7:851, 1976.

[99] E.-J. Ahn, R. Engel, T. K. Gaisser, P. Lipari, and T. Stanev. Cosmic ray interaction event generator SIBYLL 2.1. *Phys. Rev. D*, 80:094003, 2009.

[100] M. Tanabashi et al. Review of Particle Physics. *Phys. Rev. D*, 98(3):030001, 2018.

[101] H.-T. Ding, F. Karsch, and S. Mukherjee. Thermodynamics of strong-interaction matter from Lattice QCD. *Int. J. Mod. Phys. E*, 24(10):1530007, 2015.

[102] NICA physics at JINR (Official Web-Page). nica.jinr.ru/physics.php.

[103] M. E. Tejeda-Yeomans. Heavy-ion physics: freedom to do hot, dense, exciting QCD. In *10th CERN–Latin-American School of High-Energy Physics*, 4 2020.

[104] J. L. Nagle and W. A. Zajc. Small System Collectivity in Relativistic Hadronic and Nuclear Collisions. *Ann. Rev. Nucl. Part. Sci.*, 68:211–235, 2018.

[105] G. 't Hooft. A Planar Diagram Theory for Strong Interactions. *Nucl. Phys. B*, 72:461, 1974.

[106] E. Witten. Baryons in the 1n expansion. *Nuclear Physics B*, 160(1):57 – 115, 1979.

[107] V. A. Khoze and W. Ochs. Perturbative QCD approach to multiparticle production. *Int. J. Mod. Phys. A*, 12:2949–3120, 1997.

[108] Bo Andersson, G. Gustafson, G. Ingelman, and T. Sjöstrand. Parton Fragmentation and String Dynamics. *Phys. Rept.*, 97:31–145, 1983.

[109] R. Hagedorn. Statistical thermodynamics of strong interactions at high-energies. *Nuovo Cim. Suppl.*, 3:147–186, 1965.

[110] F. Becattini and U. W. Heinz. Thermal hadron production in p p and p anti-p collisions. *Z. Phys. C*, 76:269–286, 1997. [Erratum: Z.Phys.C 76, 578 (1997)].

[111] R. Stock. Hadron production in relativistic nuclear collisions. *Proc. Int. Sch. Phys. Fermi*, 153:447–472, 2003.

[112] A. Chodos, R. L. Jaffe, K. Johnson, Charles B. Thorn, and V. F. Weisskopf. A New Extended Model of Hadrons. *Phys. Rev. D*, 9:3471–3495, 1974.

[113] F. Becattini. What is the meaning of the statistical hadronization model? *J. Phys. Conf. Ser.*, 5:175–188, 2005.

[114] U. W. Heinz. Concepts of heavy ion physics. In *2nd CERN-CLAF School of High Energy Physics*, pages 165–238, 7 2004.

[115] P. Braun-Munzinger and J. Wambach. The Phase Diagram of Strongly-Interacting Matter. *Rev. Mod. Phys.*, 81:1031–1050, 2009.

[116] A. Andronic, P. Braun-Munzinger, and J. Stachel. Hadron production in central nucleus-nucleus collisions at chemical freeze-out. *Nucl. Phys. A*, 772:167–199, 2006.

[117] F. Becattini, J. Manninen, and M. Gazdzicki. Energy and system size dependence of chemical freeze-out in relativistic nuclear collisions. *Phys. Rev. C*, 73:044905, 2006.

[118] J. Letessier and J. Rafelski. Hadron production and phase changes in relativistic heavy ion collisions. *Eur. Phys. J. A*, 35:221–242, 2008.

[119] M. Petran and J. Rafelski. Universal hadronization condition in heavy ion collisions at $\sqrt{s_{\rm NN}} = 62$ GeV and at $\sqrt{s_{\rm NN}} = 2.76$ TeV. *Phys. Rev. C*, 88(2):021901, 2013.

[120] A. Bazavov et al. Strangeness at high temperatures: from hadrons to quarks. *Phys. Rev. Lett.*, 111:082301, 2013.

[121] R. Bellwied, S. Borsanyi, Z. Fodor, S. D Katz, and C. Ratti. Is there a flavor hierarchy in the deconfinement transition of QCD? *Phys. Rev. Lett.*, 111:202302, 2013.

[122] S. Chatterjee, R. M. Godbole, and Sourendu Gupta. Strange freezeout. *Phys. Lett. B*, 727:554–557, 2013.

[123] A. Andronic, P. Braun-Munzinger, K. Redlich, and J. Stachel. Hadron yields, the chemical freeze-out and the QCD phase diagram. *J. Phys. Conf. Ser.*, 779(1):012012, 2017.

[124] J. Cleymans, H. Oeschler, and K. Redlich. Influence of impact parameter on thermal description of relativistic heavy ion collisions at (1-2) A-GeV. *Phys. Rev. C*, 59:1663, 1999.

[125] G. Agakishiev *et al.* Statistical hadronization model analysis of hadron yields in p + Nb and Ar + KCl at SIS18 energies. *Eur. Phys. J. A*, 52(6):178, 2016.

[126] V. Vovchenko, V. V. Begun, and M. I. Gorenstein. Hadron multiplicities and chemical freeze-out conditions in proton-proton and nucleus-nucleus collisions. *Phys. Rev. C*, 93(6):064906, 2016.

[127] S. Das, D. Mishra, S. Chatterjee, and B. Mohanty. Freeze-out conditions in proton-proton collisions at the highest energies available at the BNL Relativistic Heavy Ion Collider and the CERN Large Hadron Collider. *Phys. Rev. C*, 95(1):014912, 2017.

[128] J. Cleymans, B. Hippolyte, H. Oeschler, K. Redlich, and N. Sharma. Thermal Model Description of Collisions of Small Nuclei. 3 2016.

[129] J. Adam *et al.* Multi-strange baryon production in p-Pb collisions at $\sqrt{s_{\mathbf{NN}}} = 5.02$ TeV. *Phys. Lett. B*, 758:389–401, 2016.

[130] J. Adam *et al.* Enhanced production of multi-strange hadrons in high-multiplicity proton-proton collisions. *Nature Phys.*, 13:535–539, 2017.

[131] S. Borsanyi *et al.* Is there still any T_c mystery in lattice QCD? Results with physical masses in the continuum limit III. *JHEP*, 09:073, 2010.

[132] A. Bazavov *et al.* Equation of state in (2+1)-flavor QCD. *Phys. Rev. D*, 90:094503, 2014.

[133] A. Andronic, P. Braun-Munzinger, and J. Stachel. Thermal hadron production in relativistic nuclear collisions: The Hadron mass spectrum, the horn, and the QCD phase transition. *Phys. Lett. B*, 673:142–145, 2009. [Erratum: Phys.Lett.B 678, 516 (2009)].

[134] J. Manninen and F. Becattini. Chemical freeze-out in ultrarelativistic heavy ion collisions at $\sqrt{s_{\mathrm{NN}}} = 130$ and 200 gev. *Phys. Rev. C*, 78:054901, Nov 2008.

[135] B. I. Abelev *et al.* Identified particle production, azimuthal anisotropy, and interferometry measurements in Au+Au collisions at $\sqrt{s_{NN}}$ = 9.2- GeV. *Phys. Rev. C*, 81:024911, 2010.

[136] M. M. Aggarwal *et al.* Scaling properties at freeze-out in relativistic heavy ion collisions. *Phys. Rev. C*, 83:034910, 2011.

[137] B. B. Back *et al.* Collision geometry scaling of Au+Au pseudorapidity density from $\sqrt{s_{NN}}$ = 19.6-GeV to 200-GeV. *Phys. Rev. C*, 70:021902, 2004.

[138] B. B. Back *et al.* Scaling of charged particle production in d + Au collisions at $\sqrt{s_{NN}}$ = 200-GeV. *Phys. Rev. C*, 72:031901, 2005.

[139] B. B. Back *et al.* Centrality dependence of the charged particle multiplicity near mid-rapidity in Au + Au collisions at $\sqrt{s_{NN}}$ = 130 GeV and 200 GeV. *Phys. Rev. C*, 65:061901, 2002.

[140] S. S. Adler *et al.* Systematic studies of the centrality and $\sqrt{s_{NN}}$ dependence of the d E(T) / d eta and d (N(ch) / d eta in heavy ion collisions at mid-rapidity. *Phys. Rev. C*, 71:034908, 2005. [Erratum: Phys.Rev.C 71, 049901 (2005)].

[141] P. Bozek. Size of the thermal source in relativistic heavy-ion collisions. *Acta Phys. Polon. B*, 36:3071–3082, 2005.

[142] Klaus Werner. Core-Corona separation in ultra-relativistic heavy ion collisions. *Phys. Rev. Lett.*, 98:152301, 2007.

[143] J. Steinheimer and M. Bleicher. Core-corona separation in the UrQMD hybrid model. *Phys. Rev. C*, 84:024905, 2011.

[144] F. Becattini and J. Manninen. Strangeness production from SPS to LHC. *J. Phys. G*, 35:104013, 2008.

[145] F. Becattini and J. Manninen. Centrality dependence of strangeness production in heavy-ion collisions as a geometrical effect of core-corona superposition. *Phys. Lett. B*, 673:19–23, 2009.

[146] J. Aichelin and K. Werner. Centrality Dependence of Strangeness Enhancement in Ultrarelativistic Heavy Ion Collisions: A Core-Corona Effect. *Phys. Rev. C*, 79:064907, 2009. [Erratum: Phys. Rev.C81,029902(2010)].

[147] M. Gemard and J. Aichelin. Core - Corona Model analysis of the Low Energy Beam Scan at RHIC (Relativistic Heavy Ion Collider) in Brookhaven (USA). 2 2014.

[148] J. Aichelin and K. Werner. Core - Corona Model describes the Centrality Dependence of v_2/ϵ. *J. Phys. G*, 37:094006, 2010.

[149] J. Aichelin and K. Werner. Is the centrality dependence of the elliptic flow v_2 and of the average $< p_T >$ more than a Core-Corona Effect? *Phys. Rev. C*, 82:034906, 2010.

[150] P. Bozek. Interplay of the emission from thermal and direct sources in relativistic heavy ion collisions. *Phys. Rev. C*, 79:054901, 2009.

[151] Y. Kanakubo, Y. Tachibana, and T. Hirano. Unified description of hadron yield ratios from dynamical core-corona initialization. *Phys. Rev. C*, 101(2):024912, 2020.

[152] K. Werner, B. Guiot, Iu. Karpenko, and T. Pierog. Analysing radial flow features in p-Pb and p-p collisions at several TeV by studying identified particle production in EPOS3. *Phys. Rev. C*, 89(6):064903, 2014.

[153] K. Werner, A. G. Knospe, C. Markert, B. Guiot, Iu. Karpenko, T. Pierog, G. Sophys, M. Stefaniak, M. Bleicher, and J. Steinheimer. Resonance production in high energy collisions from small to big systems. *EPJ Web Conf.*, 17:10900, 2018.

[154] A. M. Hillas. Two Interesting Techniques for Monte-Carlo Simulation of Very High Energy Hadron Cascades. In *International Cosmic Ray Conference*, volume 8 of *International Cosmic Ray Conference*, page 193, January 1981.

[155] L. G. Dedenko. A new method of solving the nuclear cascade equation. In *International Cosmic Ray Conference*, volume 1 of *International Cosmic Ray Conference*, page 662, January 1965.

[156] A. M. Hillas. Calculations on the propagation of mesons in extensive air showers. In *International Cosmic Ray Conference*, volume 1 of *International Cosmic Ray Conference*, page 758, January 1965.

[157] G. Bossard, H. J. Drescher, N. N. Kalmykov, S. Ostapchenko, A. I. Pavlov, T. Pierog, E. A. Vishnevskaya, and K. Werner. Cosmic ray air shower characteristics in the framework of the parton based Gribov-Regge model NEXUS. *Phys. Rev. D*, 63:054030, 2001.

[158] Till Bergmann, R. Engel, D. Heck, N. N. Kalmykov, Sergey Ostapchenko, T. Pierog, T. Thouw, and K. Werner. One-dimensional Hybrid Approach to Extensive Air Shower Simulation. *Astropart. Phys.*, 26:420–432, 2007.

[159] D. Heck, J. Knapp, J. N. Capdevielle, G. Schatz, and T. Thouw. CORSIKA: A Monte Carlo code to simulate extensive air showers. 1998.

[160] R. Ulrich, R. Engel, and M. Unger. Hadronic Multiparticle Production at Ultra-High Energies and Extensive Air Showers. *Phys. Rev. D*, 83:054026, 2011.

[161] L. Cazon, R. Conceição, and F. Riehn. Probing the energy spectrum of hadrons in proton air interactions at ultrahigh energies through the fluctuations of the muon content of extensive air showers. *Phys. Lett. B*, 784:68–76, 2018.

[162] L. Cazon, R. Conceição, M. A. Martins, and F. Riehn. Probing the π^0 spectrum at high-x in proton-Air interactions at ultra-high energies. 2018.

[163] J. Adams *et al.* Experimental and theoretical challenges in the search for the quark gluon plasma: The STAR Collaboration's critical assessment of the evidence from RHIC collisions. *Nucl. Phys. A*, 757:102–183, 2005.

[164] K. Adcox *et al.* Formation of dense partonic matter in relativistic nucleus-nucleus collisions at RHIC: Experimental evaluation by the PHENIX collaboration. *Nucl. Phys. A*, 757:184–283, 2005.

[165] I. Arsene *et al.* Quark gluon plasma and color glass condensate at RHIC? The Perspective from the BRAHMS experiment. *Nucl. Phys. A*, 757:1–27, 2005.

[166] B. B. Back *et al.* The PHOBOS perspective on discoveries at RHIC. *Nucl. Phys. A*, 757:28–101, 2005.

[167] K. Werner, Iu. Karpenko, and T. Pierog. The 'Ridge' in Proton-Proton Scattering at 7 TeV. *Phys. Rev. Lett.*, 106:122004, 2011.

[168] P. Bozek. Elliptic flow in proton-proton collisions at $\sqrt{s} = 7$ TeV. *Eur. Phys. J. C*, 71:1530, 2011.

[169] D. d'Enterria, G. Kh. Eyyubova, V. L. Korotkikh, I. P. Lokhtin, S. V. Petrushanko, L. I. Sarycheva, and A. M. Snigirev. Estimates of hadron azimuthal anisotropy from multiparton interactions in proton-proton collisions at $\sqrt{s} = 14$ TeV. *Eur. Phys. J. C*, 66:173–185, 2010.

[170] S. K. Prasad, Victor Roy, S. Chattopadhyay, and A. K. Chaudhuri. Elliptic flow (v_2) in pp collisions at energies available at the CERN Large Hadron Collider: A hydrodynamical approach. *Phys. Rev. C*, 82:024909, 2010.

[171] G. Ortona, G. S. Denicol, Ph. Mota, and T. Kodama. Elliptic flow in high multiplicity proton-proton collisions at $\sqrt{s} = 14$ TeV as a signature of deconfinement and quantum energy density fluctuations. 2009.

[172] L. Cunqueiro, J. Dias de Deus, and C. Pajares. Nuclear like effects in proton-proton collisions at high energy. *Eur. Phys. J. C*, 65:423–426, 2010.

[173] V. Khachatryan *et al.* Observation of Long-Range Near-Side Angular Correlations in Proton-Proton Collisions at the LHC. *JHEP*, 09:091, 2010.

[174] K. Dusling, W. Li, and B. Schenke. Novel collective phenomena in high-energy proton-proton and proton-nucleus collisions. *Int. J. Mod. Phys. E*, 25(01):1630002, 2016.

[175] C. Loizides. Experimental overview on small collision systems at the LHC. *Nucl. Phys. A*, 956:200, 2016.

[176] B. I. Abelev *et al.* Long range rapidity correlations and jet production in high energy nuclear collisions. *Phys. Rev. C*, 80:064912, 2009.

[177] G. Aad *et al.* Observation of a Centrality-Dependent Dijet Asymmetry in Lead-Lead Collisions at $\sqrt{s_{NN}} = 2.77$ TeV with the ATLAS Detector at the LHC. *Phys. Rev. Lett.*, 105:252303, 2010.

[178] S. Chatrchyan *et al.* Observation and studies of jet quenching in PbPb collisions at $\sqrt{s_{NN}} = 2.76$ TeV. *Phys. Rev. C*, 84:024906, 2011.

[179] C. Bierlich, G. Gustafson, and L. Lönnblad. Collectivity without plasma in hadronic collisions. *Phys. Lett. B*, 779:58, 2018.

[180] B. Blok, C. D. Jäkel, M. Strikman, and U. A. Wiedemann. Collectivity from interference. *jhep*, 2017(12):074, 2017.

[181] T. Pierog, B. Guiot, I. Karpenko, G. Sophys, M. Stefaniak, and K. Werner. EPOS 3 and Air Showers. *EPJ Web Conf.*, 210:02008, 2019.

[182] L. A. Anchordoqui, C. Garcia Canal, S. J. Sciutto, and J. F. Soriano. Through the Looking-Glass with ALICE into the Quark-Gluon Plasma: A New Test for Hadronic Interaction Models Used in Air Shower Simulations. *arXiv:1907.09816*, 2019.

[183] T. Pierog and K. Werner. Muon Production in Extended Air Shower Simulations. *Phys. Rev. Lett.*, 101:171101, 2008.

[184] S. Baur, H. Dembinski, M. Perlin, T: Pierog, R. Ulrich, and K. Werner. Core-corona effect in hadron collisions and muon production in air showers. *arXiv e-prints*, page 1902.09265, 2019.

[185] J. Manninen and F. Becattini. Chemical freeze-out in ultra-relativistic heavy ion collisions at $\sqrt{s_{NN}}$ = 130 and 200-GeV. *Phys. Rev. C*, 78:054901, 2008.

[186] Tanguy Pierog. Open issues in hadronic interactions for air showers. *EPJ Web Conf.*, 145:18002, 2017.

[187] Pedro Abreu et al. Interpretation of the Depths of Maximum of Extensive Air Showers Measured by the Pierre Auger Observatory. *JCAP*, 02:026, 2013.

[188] H. P. Dembinski. Computing mean logarithmic mass from muon counts in air shower experiments. *Astropart. Phys.*, 102:89, 2018.

[189] The Pierre Auger Collaboration. Direct measurement of the muonic content of extensive air showers between $2 \cdot 10^{17}$ and $2 \cdot 10^{18}$ eV at the Pierre Auger Observatory. *Accepted by Eur. Phys. J. C*, 2020.

[190] J. F. Soriano, L. A. Anchordoqui, T. C. Paul, and T. J. Weiler. Probing QCD approach to thermal equilibrium with ultrahigh energy cosmic rays. *PoS*, ICRC2017:342, 2018.

[191] J. Alvarez-Muniz, L. Cazon, R. Conceição, J. Dias de Deus, C. Pajares, and M. Pimenta. Muon production and string percolation effects in cosmic rays at the highest energies. 2012.

[192] Glennys R. Farrar and Jeffrey D. Allen. A new physical phenomenon in ultra-high energy collisions. *EPJ Web Conf.*, 53:07007, 2013.

[193] L. A. Anchordoqui, H. Goldberg, and T. J. Weiler. Strange fireball as an explanation of the muon excess in Auger data. *Phys. Rev. D*, 95(6):063005, 2017.

[194] D. LaHurd and C. E. Covault. Exploring Potential Signatures of QGP in UHECR Ground Profiles. *jcap*, 2018(11):007, 2018.

[195] J. Lakowicz. Principles of fluorescence spectroscopy. *Principles of Fluorescent Spectroscopy, 3rd Edn*, 2006.

[196] A. Pla-Dalmau, A. D. Bross, and V. V. Rykalin. Extruding plastic scintillator at fermilab. In *2003 IEEE Nuclear Science Symposium. Conference Record (IEEE Cat. No.03CH37515)*, volume 1, pages 102–104 Vol.1, Oct 2003.

[197] Saint-Gobain. Scintillating Optical Fibers BCF-99-29AMC. 2016.

[198] M. Platino *et al.* . AMIGA at the Auger Observatory: the scintillator module testing system. *Journal of Instrumentation*, 6(06):P06006–P06006, jun 2011.

[199] F. Sánchez for the Pierre Auger Collaboration. The muon component of extensive air showers above $10^{17.5}$ eV measured with the Pierre Auger Observatory. *PoS*, ICRC2019:411, 2019.

[200] A. M. Botti for the Pierre Auger Collaboration. The AMIGA underground muon detector of the Pierre Auger Observatory - performance and event reconstruction. *PoS*, ICRC2019:202, 2019.

[201] Hamamatsu Photonics K K. DataSheet MPPC (Multi-Pixel Photon Counter) arrays. S13361-2050 series, 2018.

[202] Hamamatsu Photonics K K. Datasheet photomultiplier tube assembly. H8804, 2012.

[203] S. Dolinsky. Novel approach for calibration breakdown voltage of large area SiPM. *PoS*, PhotoDet2012:005, 2012.

[204] S. N. Ahmed. 6 - scintillation detectors and photodetectors. In *Physics and Engineering of Radiation Detection (Second Edition)*, pages 331 – 433. Elsevier, second edition edition, 2015.

[205] Weeroc. Datasheet - citiroc 1a. v2.5, 2019.

[206] The Pierre Auger Collaboration. Muon counting using silicon photomultipliers in the AMIGA detector of the Pierre Auger observatory. *JINST*, 22:P03002, 2017.

[207] D. Ravignani, A. D. Supanitsky, and D. Melo. Reconstruction of air shower muon densities using segmented counters with time resolution. *Astropart. Phys.*, 82:108–116, 2016.

[208] A. M. Botti, F. Sánchez, A. Fuster, and M. Rolf Hampel *et al.* . Characterization of AMIGA binary signals with SiPMs. *Pierre Auger internal note*, pages GAP2018–049, 2018.

[209] A. Coleman. Measurement of the Cosmic Ray Flux Above 100 PeV at the Pierre Auger Observatory. *Pierre Auger internal note*, pages GAP2018–049, 2018.

[210] A. Coleman. Measurement of the cosmic ray flux near the second knee with the Pierre Auger Observatory. *PoS*, ICRC2019:225, 2020.

[211] A. D. Supanitsky, A. Etchegoyen, G. Medina-Tanco, I. Allekotte, M. Gomez Berisso, and Maria Clementina Medina. Underground Muon Counters as a Tool for Composition Analyses. *Astropart. Phys.*, 29:461–470, 2008.

[212] M. Nagano, T. Hara, Y. Hatano, N. Hayashida, S. Kawaguchi, K. Kamata, T. Kifune, and Y. Mizumoto. Energy Spectrum of Primary Cosmic Rays Between $10^{14.5}$ eV and 10^{18} eV. *J. Phys. G*, 10:1295, 1984.

[213] I. Al Samarai. Arrival direction distribution of cosmic rays from $\simeq$100 PeV to the highest energies detected at the Pierre Auger Observatory. *PoS*, FFP14:042, 2016.

Acronyms

This is a list of acronyms used sorted alphabetically.

AERA	Auger Engineering Radio Array
AMIGA	Auger Muon Detectors for the Infill Ground Array
APD	avalanche photo diode
BLS	Balloon Launching Station
CDAS	Central Data Acquisition System
CE	cascade equation
CEP	critical end point
CMB	Cosmic Microwave Background
CR	cosmic ray
EA	engineering array
EAS	extensive air shower
EMA	ethylene methacrylate
FD	Fluorescence Detector
G-APD	geiger mode avalanche photo diode
GPS	Global Positioning System
GZK	Greisen–Zatsepin–Kuzmin
HEAT	High Elevation Auger Telescopes
MLDF	Muon Lateral Distribtuion Function
PE	photon-equivalent
PMMA	poly(methyl methacrylate)
PMT	photomultiplier tube
PVC	polyvinyl chloride
QGP	Quark Gluon Plasma
SD	Surface Detector
SHM	Statistical Hadronization Model
SiPM	Silicon photo-multipler
UHECR	ultra-high-energy cosmic ray
UMD	Underground Muon Detector
WCD	water-Cherenkov detector
WLS	wavelength shifter